Uncovering the organization of neural circuits with generalized phase locking analysis

Shervin Safavi[1,2], **Theofanis Panagiotaropoulos**[1,3], **Vishal Kapoor**[1], **Juan F. Ramirez-Villegas**[1,4], **Nikos K. Logothetis**[1,6,7], **Michel Besserve**[1,5*]

***For correspondence:**
michel.besserve@tuebingen.mpg.de

[1]Department of Physiology of Cognitive Processes, Max Planck Institute for Biological Cybernetics, Tübingen, Germany; [2]IMPRS for Cognitive and Systems Neurosurgeon, University of Tübingen, Tübingen, Germany; [3]Cognitive Neuroimaging Unit, INSERM, CEA, Université Paris-Saclay, NeuroSpin center, Gif/Yvette, France; [4]Institute of Science and Technology Austria (IST Austria), Klosterneuburg, Austria; [5]Department of Empirical Inference, Max Planck Institute for Intelligent Systems and MPI-ETH center for Learning Systems, Tübingen, Germany; [6]International Center for Primate Brain Research, Songjiang, Shanghai, China; [7]Centre for Imaging Sciences, Biomedical Imaging Institute, The University of Manchester, Manchester, UK

Abstract Spike-field coupling characterizes the relationship between neurophysiological activities observed at two different scales: on the one hand, the action potential produced by a neuron, on the other hand a mesoscopic "field" signal, reflecting subthreshold activities. This provides insights about the role of a specific unit in network dynamics. However, assessing the overall organization of neural circuits based on multivariate data requires going beyond pairwise approaches, and remains largely unaddressed. We develop *Generalized Phase Locking Analysis* (GPLA) as an multichannel extension of univariate spike-field coupling. GPLA estimates the dominant spatio-temporal distributions of field activity and neural ensembles, and the strength of the coupling between them. We demonstrate the statistical benefits and interpretability of this approach in various biophysical neuronal network models and Utah array recordings. In particular, we show that GPLA, combined with neural field modeling, help untangle the contribution of recurrent interactions to the spatio-temporal dynamics observed in multi-channel recordings.

Keywords
Spike-field coupling, microcircuits, prefrontal cortex, neural field models, neural mass models, hippocampus, phase-locking.

Introduction

Understanding brain function requires uncovering the relationships between neural mechanisms at different scales (Einevoll et al., 2019): from single neurons to microcircuits (Rasch et al., 2008, 2009), from microcircuits to a single brain area (Li et al., 2009) and from a single area to the whole brain (Schwalm et al., 2017; Zerbi et al., 2019). Therefore, it is crucial to develop data analysis tools to investigate the cooperative mechanisms that connect one level of organization to the next.

A well-studied example of such a collective organization phenomenon is oscillatory neuronal activity. Neural oscillations are hypothesized to support computations in the brain (Peterson and

Voytek, 2018) and various cognitive functions, such as perceptual binding (Engel et al., 1999), visual awareness (Dwarakanath et al., 2020), attention (Niebur et al., 1993) and memory (Buzsaki, 2006). These oscillations manifest themselves in Local Field Potentials (LFP), a mesoscopic extracellular signal (Liljenstroem, 2012) resulting from ionic currents flowing across the cellular membranes surrounding the electrode. LFP oscillatory activity partly reflects a number of subthreshold processes shared by units belonging to underlying neuronal ensembles and responsible for the coordination of their activity (Buzsaki et al., 2012; Einevoll et al., 2013; Herreras, 2016). As a consequence, a broad range of empirical investigations support the importance of analyzing oscillatory dynamics observed in LFPs (for reviews see Buzsaki et al. (2012, 2013); Einevoll et al. (2013); Herreras (2016); Pesaran et al. (2018)).

In particular, the relationship between spiking activity and LFP has broad implications for mesoscale mechanisms of network coordination. For instance, spike-field coupling relates to synaptic plasticity, triggering changes in the spike sequences generated by neural ensembles (Grosmark et al., 2012; Grosmark and Buzsáki, 2016). Moreover, cognitive functions, such as attention, are hypothesized to rely on interactions between various neural populations coordinated by network oscillations, which modulate the excitability of target populations so that they spike during time windows facilitating their communication (Fries, 2005, 2015; Womelsdorf et al., 2007).

Common approaches for investigating the spike-LFP oscillatory coupling are typically restricted to pairwise spike-LFP interactions (Zeitler et al., 2006; Ashida et al., 2010; Vinck et al., 2010, 2012; Jiang et al., 2015; Li et al., 2016; Zarei et al., 2018) which are suboptimal for modern datasets. Indeed, state-of-the-art multichannel electrophysiology systems (Dickey et al., 2009; Jun et al., 2017; Juavinett et al., 2019) allow simultaneous recording of hundreds or even thousands of sites (Pesaran et al., 2018; Jun et al., 2017; Buzsáki, 2004; Fukushima et al., 2015). This growth in dimensionality and complexity requires the parallel development of appropriate methodologies and models (Stevenson and Kording, 2011; O'Leary et al., 2015; Gao and Ganguli, 2015; Williamson et al., 2019). In particular, recent technological advances offer an unprecedented opportunity to study the relationship of large scale collective organization binding the activity of individual units (e. g. as spiking activity) with spatio-temporal field potential dynamics (e. g. wave patterns (Ermentrout and Pinto, 2007; Ermentrout and Kleinfeld, 2001)). Alongside this experimental progress, there is a growing need for conceptual and methodological frameworks to investigate this relationship. In particular, computing an interpretable summary of the coupling between neurophysiological quantities is of paramount importance in this high dimensional setting.

We develop a "Generalized Phase Locking Analysis" (GPLA) to address this growing need. GPLA characterizes and assesses statistically at once the coupling between the spiking activity of large populations of units and large-scale spatio-temporal patterns of LFP. After demonstrating GPLA's interpretability on toy simulated datasets and statistical benefits with respect to pairwise coupling approaches, we use neural field models to illustrate how this method characterizes key aspects of the underlying neural circuits. In particular, we show, on simulated data with a larger degree of realism, how GPLA can untangle the contribution of recurrent interactions to the observed spatio-temporal dynamics. Finally, GPLA's interpretability is exploited in the analysis of Utah array recordings from the macaque prefrontal cortex.

Results

Generalizing spike-oscillation coupling analysis to the multivariate setting

Our analysis relies first on characterizing the coupling between signals originating from a pair of recording channels. On the one hand, we consider the time-varying spike rate $\lambda(t)$ of a given unit; and on the other hand oscillatory activity $L_f(t)$ is derived from the LFP by band-pass filtering in a band of center frequency f. We assume $L_f(t)$ is the complex analytic signal representation of this oscillation, computed using the Hilbert transform (Chavez et al., 2006), such that $L_f(t) = a_f(t)e^{i\phi_f(t)}$, where $a_f(t)$ and $\phi_f(t)$ are the instantaneous amplitude and phase of the oscillation, respectively.

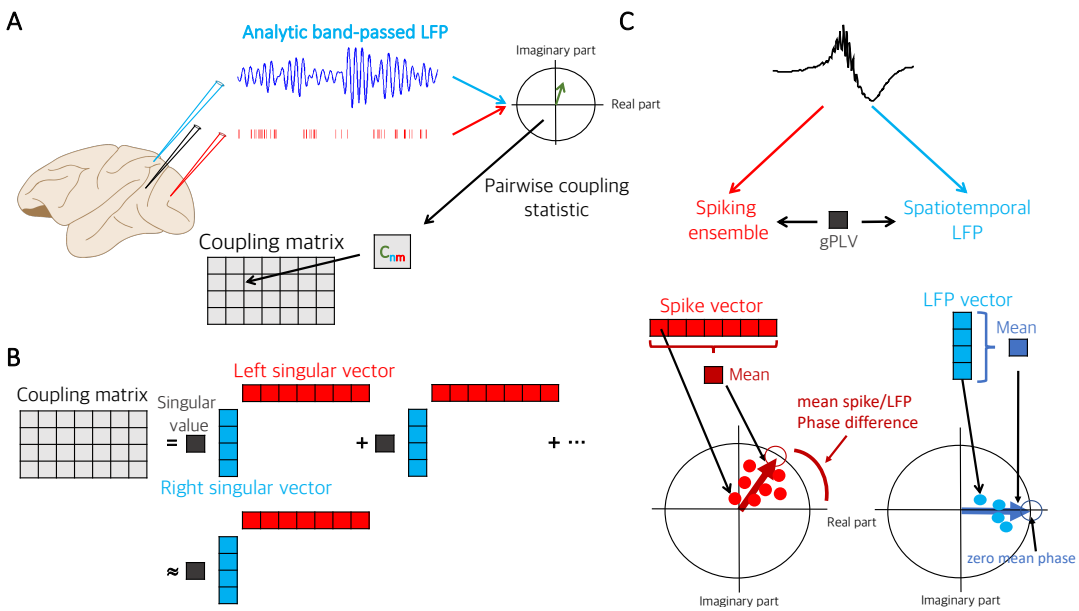

Figure 1. Schematic illustration of Generalized Phase Locking Analysis (GPLA).
(A) The coupling matrix is built from electrophysiology data by gathering pairwise complex phase locking estimates (based on Equation 2) of all spike-LFP pairs in a rectangular matrix. Coefficients (C_{nm}) contain information similar to complex-valued PLV up to a scaling factor: the magnitude indicates the strength of coupling and the angle reflects the average timing of the spike occurrence within the period of the corresponding LFP oscillation. **(B)** The coupling matrix can be approximated using its largest singular value and the corresponding singular vectors. **(C)** Singular vectors represent the dominant LFP (blue array) and spiking patterns (red array) and the singular value (d_1), called generalized Phase Locking Value (gPLV), characterizes the spike-field coupling strength for the phenomenon under study and the chosen frequency. The magnitude of each vector entry indicates a relative coupling of the corresponding unit/channel and the phase indicates the relative timing with respect to other units/channels. By convention, the phase of the LFP vector coefficients' average is set to zero, such that the phase of the spike vector average reflects the overall phase shift of the spike pattern with respect ot the LFP pattern.

The coupling between these signals can be characterized by the covariance

$$c(f) = \left\langle \lambda(t) L_f(t) \right\rangle = \left\langle \lambda(t) a_f(t) e^{i\phi_f(t)} \right\rangle = |c| e^{i\Phi_c} = |c| \left(\cos(\Phi_c) + i \sin(\Phi_c) \right) , \tag{1}$$

where the $\langle \cdot \rangle$ indicates averaging across time and experimental trials. The complex number c then reflects the strength of coupling through its modulus $|c|$, and the dominant LFP phase of spiking through its argument Φ_c (see Figure 1A). This coupling measure is a modification of the Phase-Locking Value (PLV) (Ashida et al., 2010), and differs from the latter mainly through the incorporation of the amplitude of the oscillation in the averaging, and the absence of normalization by the spike rate. Although $c(f)$ is straightforward to estimate based on observed spike times, leading to the quantity noted $\hat{c}(f)$ (see Safavi et al. (2020)), interpreting its value in terms of the underlying neural mechanisms is challenging, due to the largely unknown specific properties of the considered cell and recording site. Alternatively, synthesizing the information provided by a large number of couplings may provide a more reliable picture of the functioning of the underlying circuits.

Generalized Phase Locking Analysis (GPLA) is introduced as a multivariate statistical analysis technique to estimate the key properties of a matrix $\boldsymbol{C}(f)$ consisting of the pairwise coupling between a large number of units and LFP channels. Given the spike times $\{t_k^m\}$ for unit m and the analytic signal $L_f^n(t)$ that is filtered around frequency f for LFP channel n, the (n, m) coordinate of the coupling matrix $\boldsymbol{C}(f)$ is estimated by summing the values taken by the analytic signal at all spike

times (see Figure 1A),

$$\widehat{\boldsymbol{C}}(f)_{n,m} = \sum_k L_f^n(t_k^m)\,. \tag{2}$$

As schematized in Figure 1B, to derive a compact and interpretable representation from this typically large matrix, we compute the Singular Value Decomposition (SVD) of the coupling matrix estimate and approximate it with the term of largest singular value d_1 leading to the approximation

$$\widehat{\boldsymbol{C}} \quad = U\,D\,V^H \quad = \sum_{k=1}^{p} d_k \boldsymbol{u}_k \boldsymbol{v}_k^H \approx d_1 \boldsymbol{u}_1 \boldsymbol{v}_1^H\,, \tag{3}$$

where $\boldsymbol{v}^H$ indicate the transpose conjugate of vector $\boldsymbol{v}$. In this expression, the singular value d_1 is a positive scalar, that we will call generalized Phase Locking Value ($gPLV$), and quantifies the magnitude of the coupling. The associated complex valued singular vectors in this factorization will be respectively called the *LFP vector* $\boldsymbol{u} = \boldsymbol{u}_1$ and the *spike vector* $\boldsymbol{v} = \boldsymbol{v}_1$. As illustrated in Figure 1C, the spike vector indicates the pattern of coordinated spiking activity most coupled to LFP oscillations, while the LFP vector reflects the dominant spatio-temporal pattern of LFP involved in this coupling. Importantly, the relationship between $\boldsymbol{u}$ and $\boldsymbol{v}$ reflects the phase lag between spiking and LFP activities. Multiplication of both singular vectors by the same unit complex number leads to the exact same approximation as Equation 3, reflecting that GPLA only measures the relative phase between LFP and spikes. To resolve this ambiguity in our analyses, we adopt the convention of setting the phase of the average of LFP vector coefficients $\langle \boldsymbol{u} \rangle = \frac{1}{n_c} \sum_k u_k$ to zero, as illustrated in Figure 1C. As a consequence, the phase of the mean of the spike vector coefficients $\langle \boldsymbol{v} \rangle = \frac{1}{n_u} \sum_k v_k$ reflects the difference of mean phases between spiking and LFP activities. See section Detailed GPLA methodology for electrophysiology data in STAR Methods for more details.

Illustration of GPLA on toy examples

To illustrate how GPLA can provide an intuitive summary of the coupling between the population of spiking units and LFPs, we use toy simulations in which transient LFP oscillations (Figure 2A) modulate the firing probability in 18 spike trains (attributed to neuron-like units). The firing patterns of these neurons depend on the coupling strength and the delay with respect to the global oscillation. We introduce four simple models in Figure 2C-F, and demonstrate how the resulting spike vector summarizes the internal structure of the model. For each model, we schematically represent the coupling strength by the line thickness (absence of a line indicates no coupling) and the delay with respect to the global oscillation by the line color (Figure 2C-F, first column). In model 1-3 (Figure 2C-E) the spikes are coupled to the LFP (see STAR Methods, section Simulation of phase-locked spike trains for details) and in model 4 (Figure 2F) there is no coupling and the spikes were generated with a homogeneous Poisson process (presence or absence of coupling is also reflected in the coupling strength assessed by gPLV, as demonstrated in Figure 2B). Exemplary spike trains for each model are displayed in the second column of Figure 2C-F overlaid on the magnified version of the oscillation. Model 1 instantiates a global oscillation driving a synchronous population of neurons and this structure is also reflected in the resulting spike vector of the model as all the coefficients collapse to a single point (Figure 2C, third column). Model 2 corresponds to wave-like discharges of neurons (similar to the case of "delayed excitations from single oscillator" described by Ermentrout and Kleinfeld (2001)) which is also reflected in the evenly distributed phases of the spike vector coefficients over a 180 degrees interval (Figure 2D). Model 3 accounts for the activity of different cell populations that fire together predominately at three distinct phase value of the LFP and this synchrony is also reflected in the spike vector as the coefficients of the spike vector are clustered together in a complex plane (Figure 2E, third column) i. e. the group to which each unit belongs to is reflected in the phase of the associated spike vector coefficient. In model 4, there is no coupling between spikes and LFP, and the coefficients of the spike vector of this model are isotropically distributed in the complex plane (Figure 2F, third column). In all the above cases, the spike vector resulting from GPLA summarizes the coupling structure in an intuitive and compact way. Because

this setting has a single LFP channel, this analysis straightforwardly combines univariate coupling measures of each unit. However, statistical analysis of GPLA is different from the univariate case, as we show next.

Advantages of GPLA over univariate spike-field coupling

Beyond the above benefits for interpreting multivariate data, we investigated whether GPLA can be advantageous over its univariate counterpart from a statistical perspective by simulating transient oscillations (Figure 3A) and phase-locked spikes. The neural populations consisted of 3 groups, each of them locked to a different phase (0, 120, and 240 degrees) of the same oscillation, corrupted with additive noise. An illustrative simulation in the case of low noise and large amount of observed spike is shown in Figure 3B, together with the corresponding spike vector in Figure 3C, providing results similar to Fig. 2E.

For quantitative analysis, we first consider the setting of a single LFP channel and a handful of neurons are the focus of the analysis. Such recordings are still common and valuable in human electrophysiology experiments for understanding cognition (Mukamel and Fried, 2011; Fried et al., 2014). The main difference between this simulation and the one presented in Figure 2E is the weaker coupling of individual neurons leading to values at the edge of significance (assessed with the surrogate-based test, see section Significance assessment of gPLV in STAR Methods). While pooling the spikes from all units into a single spike train to get a *pooled Phase-Locking-Value* (pPLV) would result in a higher statistical power, it requires the distribution of the locking phase to be homogeneous across units (e.g. in the case of Figure 2C, but not for Figure 2D and E). In contrast, GPLA may be able to exploit the spike times from multiple neurons to achieve better statistical power in assessing the global coupling between spikes and LFPs. We ran 5000 simulations with only 3 units and compared the coupling assessment based on PLV, pPLV, and gPLV. Figure 3D represents the estimated PLVs, with averages matching the couplings obtain with a larger number of spikes Figure 3C. Performance of each measure is assessed based on its detection rate, which is defined as the percentage of simulations for which significant coupling is detected, as assessed using spike-jittered surrogate data (see STAR Methods section Significance assessment of gPLV) and with a significance threshold of 5%. As it is demonstrated in Figure 3E, gPLV detection outperforms the competing approaches (PLV and pPLV).

Beyond better detection of significant coupling, GPLA-based estimation of pairwise couplings based on the approximation of Equation 3 may be more accurate than individual estimates when the data is very noisy and multivariate, benefiting from the SVD procedure to disentangle noise from the ground truth coupling. To demonstrate this, we performed simulations similar to the above, but using 50 LFP channels containing oscillations driving spike-LFP coupling, contaminated by different levels of noise (i. e. adding Gaussian noise with different variances to the transient oscillation, see STAR Methods for details), and modulating the firing rates of the units, lower firing rates leading to a larger amount of martingale noise for the PLV estimates (Safavi et al., 2020). An example LFP trace with (black) and without (blue) noise is exemplified in Figure 3F and an example coupling matrix in the presence of noise is also illustrated in Figure 3G. In this case, the coupling matrix has rank one, as all the units are locked to a single frequency. We ran the simulations with different amounts of LFP noise (indicated on the x-axis of Figure 3H-I) and computed the coupling coefficients (similar to Figure 3G) and compare it to ground truth (based on Equation 32). Signal-to-Noise Ratio (SNR) was defined as the ratio of estimation error (the difference between estimated PLV and the ground truth) to coupling strength (PLV) and was used to compare the quality of GPLA-based and univariate estimation (indicated in the y-axis of Figure 3H-I). As this simulation demonstrates, the estimation error of the coupling coefficients is larger for the univariate estimation than for the GPLA-based approach for a broad range of noise levels (Figure 3H-I).

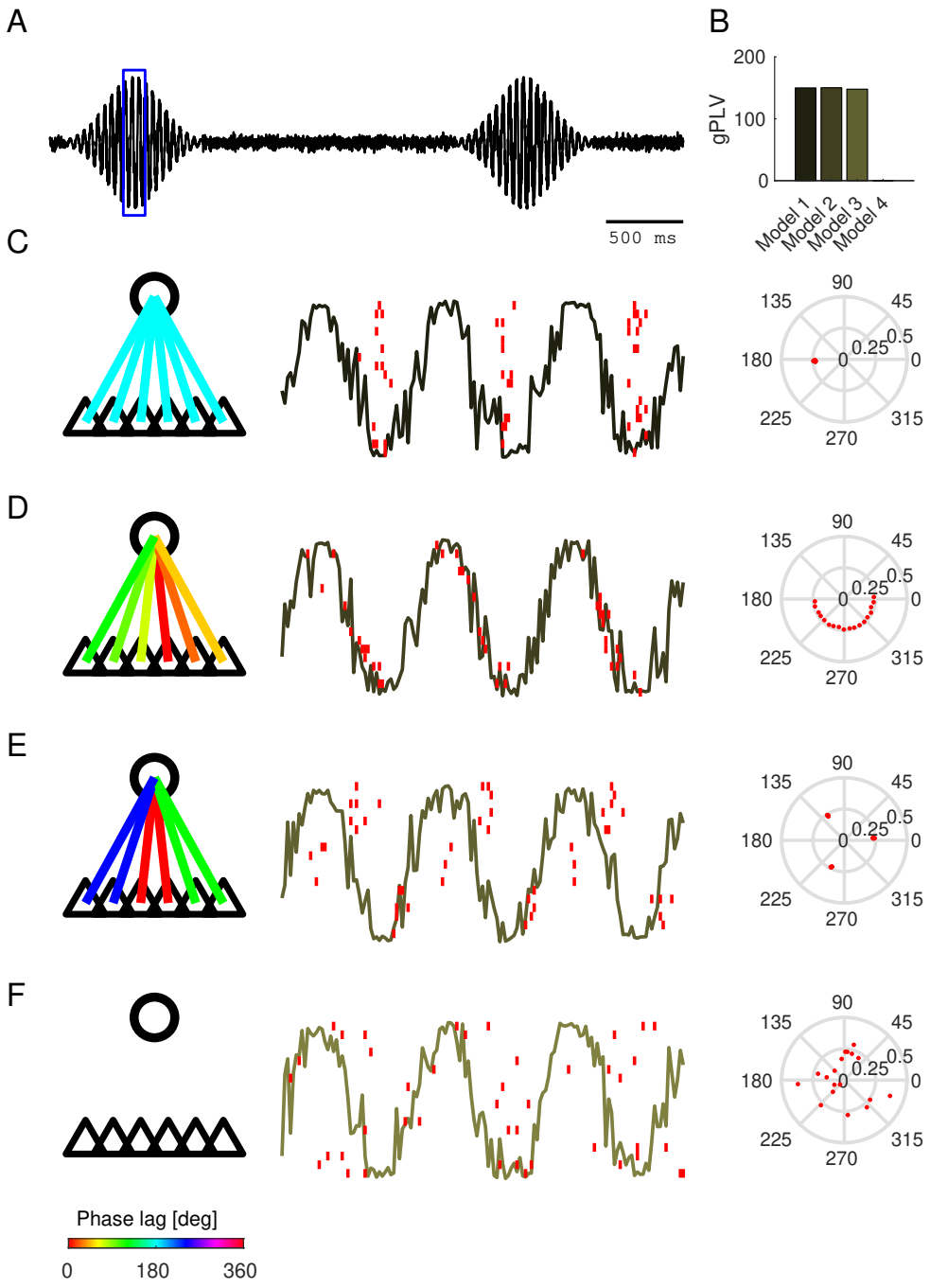

Figure 2. Illustration of GPLA on simple simulations.
(A) Normalized amplitude of LFP-like oscillatory signals. **(B)** gPLVs for different models demonstrated in C-F **(C-F)** Various scenarios of spike-LFP coupling. Left: schematic representation of the modulating LFP oscillation (circle), and 6 representative neuron-like-units (indicated by the triangles). The color of each connecting line indicates the locking phase (see bottom colorbar for color code). Center: LFP-like signals within the window specified by the blue box in A and spikes are represented by overlaid red vertical lines. Right: resulting spike vector is represented in the third column. **(C)** Spiking activity globally synchronized to the trough of the LFP oscillation. **(D)** Sequential discharge of spikes coupled to the LFP. **(E)** Three clusters of neurons discharge at different phases of the LFP oscillation (a similar model was also used in Figure 3). **(F)** Spiking activity uncoupled to LFP oscillation (independent homogeneous spike trains).

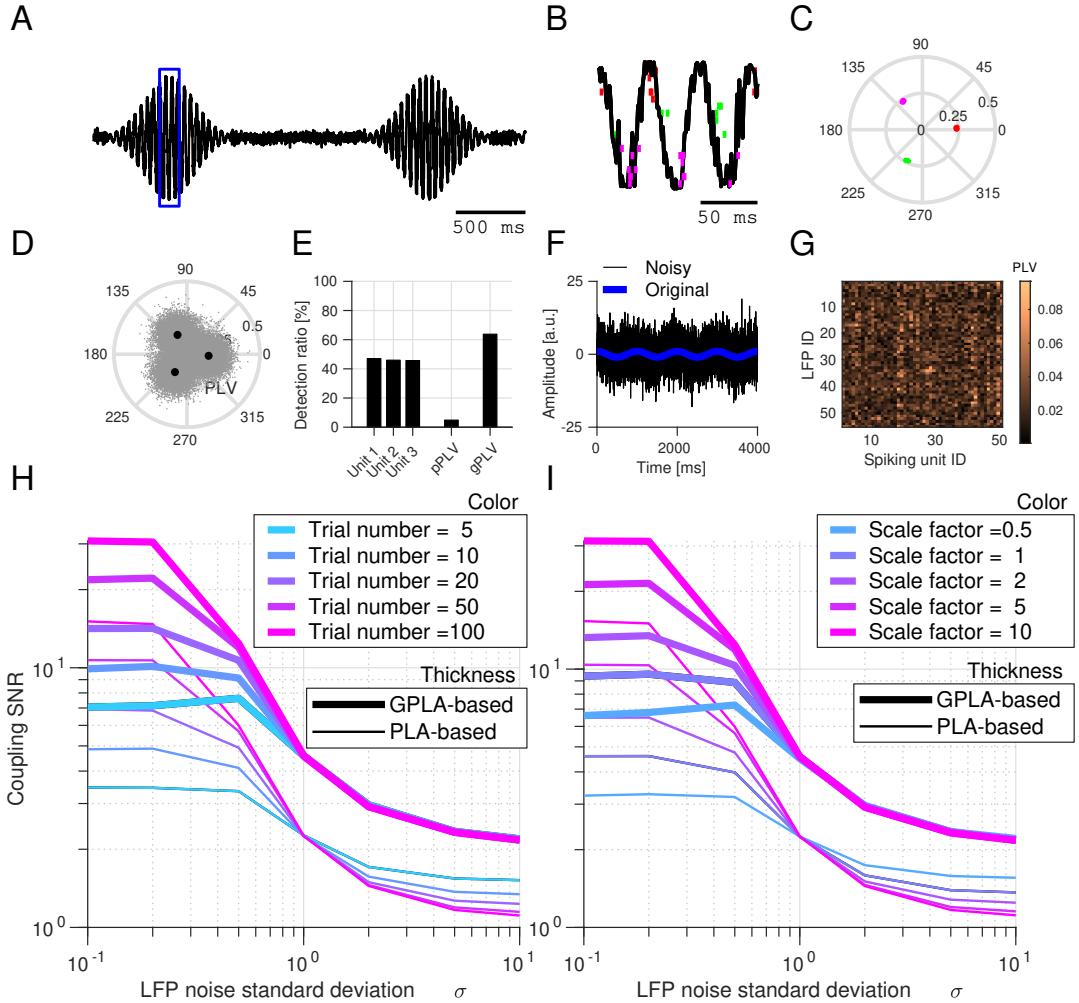

Figure 3. Comparison of GPLA and uni-variate spike-field coupling
(A) Normalized amplitude of LFP-like transient oscillatory signal with additive Gaussian white noise. **(B)** LFP-like signal and overlaid spike raster (colored vertical lines – colors indicate populations with the same locking phase) within the window specified by the blue box in (A). **(C)** Spike vector coefficients in the complex plane (colors correspond to B). Each dot represents one coefficient of the spike vector corresponding to a single neuron (note that within each cluster dots are overlapping as they are similarly coupled). **(D)** Complex PLVs represented in the complex plane. Angles indicate the locking phase and the radius the PLV. The gray point clouds indicate the PLV of multiple simulations and larger black dots indicate the average values. **(E)** Performance comparison (in percentage of simulations with significant coupling) of PLV, pooled PLV (pPLV) and gPLV, for three individual neurons. **(F)** Example oscillation, original (blue trace) and noisy (black trace). **(G)** Example coupling matrix related to simulation with a large amount of noise ($\sigma = 5$) **(H-I)** Comparison of GPLA-based and PLA-based estimation of PLVs for (H) different number of trials and (I) different levels of firing rate. Signal-to-Noise Ratio (SNR) is defined as the estimation error (the difference between estimated PLV and the ground truth) to coupling strength (PLV).

Statistical properties of GPLA

While in the previous section, GPLA's significance was assessed using surrogate data, this approach is computationally expensive and provides limited insights into the statistical properties of GPLA estimates. We now investigate this question using theory, and exploit it to assess more efficiently the significance of multivariate coupling. Singular values and vectors estimated by GPLA have an intrinsic variability due to the stochasticity of spiking activity, which can be investigated through stochastic integration and random matrix theory (Anderson et al., 2010; Capitaine and Donati-Martin, 2016). In the absence of coupling between spikes and LFP, appropriate preprocessing allows deriving analytically the asymptotic distribution of univariate and multivariate coupling measures (Safavi et al., 2020), including the convergence of the squared singular values to the classical Marchenko-Pastur (MP) law (Marchenko and Pastur, 1967). Based on the MP law, we can define an upper bound on the largest singular values of the coupling matrix, such that exceeding this bound indicates the significance of the coupling (for more details see STAR Methods section Analytical test and Safavi et al. (2020)), leading to a *analytical test*.

We assessed the performance of this test on simulated spikes and LFPs with or without coupling as follows. Briefly, we synthesized multivariate LFP activity by linearly superimposing several oscillations ($O_k(t)$ demonstrated in Figure 4A) with different multiplicative weights applied for each LFP channel and generated the spike trains of each unit with Poisson statistics. As for the coupling between spikes and LFPs, $2/5^{\text{th}}$ of the units were coupled to the LFP oscillations (exemplified in Figure 4B), while the remaining units had homogeneous Poisson spike trains (for details see the STAR Methods, section Simulation of phase-locked spike trains). The estimated coupling matrix computed based on Equation 2 for a simulation with 100 spike trains and 100 LFPs is exemplified in Figure 4C, where we have two coupled populations, one coupled to the lowest-frequency and one coupled to the highest-frequency oscillatory component of LFP (respectively top-right and bottom-left relatively bright blocks in Figure 4C and sample spike trains and LFP in Figure 4B).

Computing the SVD of the coupling matrix after application of the preprocessing explained in STAR Methods, section LFP pre-processing, results in a spectral distribution for the squared singular values, which matches the prediction of the theory (Figure 4D-E). In the absence of coupling between spikes and LFP signals (Figure 4D), the distribution of the eigenvalues closely follows the MP law and in the presence of coupling, the largest eigenvalue exceeds the significance bound predicted by Random Matrix Theory (RMT) (see STAR Methods section Analytical test for more details).

We further quantified the type I and II error of this analytical test. For type I error, we ran the simulations with non-zero coupling between spikes and LFP signals. GPLA was able to indicate a significant coupling between spike and LFP even when the coupling strength is as small as 0.05 (no coupling case corresponds to 0 strength and the perfect coupling corresponds to 1). One benefit of our analytical test is that the statistical power does not degrade with the increasing dimension of the data through the number of recording channels (Figure 4F). Notably, this is in contrast with univariate methods, assessing independently the significance of pairwise couplings resulting from each pair of spike train and LFP. In this case, it is necessary to correct for multiple comparisons before defining the significance threshold, Therefore, the statistical power will typically degrade with the increase in the number of units/LFPs. This is particularly relevant for weaker couplings as they may lose significance after correction for multiple comparisons . Lastly, in order to quantify the type II error of the significance test, we ran the simulation with no coupling between spikes and LFP and quantified the number of false positives. Our results show that our analytical test has a negligible ($< 5\%$) false positive error (Figure 4G).

We also quantified the performance of the method for estimating the number of populations coupled to different rhythms. Similar to the simulation explained earlier (Figure 4A-C), we simulate multiple (1-10) non-overlapping cell assemblies synchronized to different LFP rhythms (with different frequencies within a narrow range of 11-15.5 Hz). When the coupling is larger than a minimal

strength (coupling strength of 0.5), the method was able to capture the number of populations with very low error, MSE < 0.015 (Figure 4H).

A neural mass model of spike-LFP dynamics

While the above results have addressed the meaning of GPLA's outcomes in toy models, their neuroscientific interpretation requires modeling of the underlying neural network. We study this question in the context of a two-population *neural field model*: a spatially distributed rate model of the activity of two interacting homogeneous populations: excitatory pyramidal cells (E population) and inhibitory interneurons ("I" population) (Nunez and Srinivasan, 2006; Wilson and Cowan, 1973). The model is governed by three basic input-output relations (see STAR Methods, section Analytical neural field modeling of spike-field coupling) and depicted in Figure 5A: (1) the dynamics of the average somatic membrane potentials V_E and V_I of each population is governed by afferent post-synaptic currents η originating from other cortical or subcortical structures as well as recurrent excitatory and inhibitory post-synaptic currents (EPSC and IPSC) s_E and s_I; (2) the population spike rates λ_E and λ_I are a function of their respective membrane potentials; and (3) EPSC and IPSC are each controlled by the spike rate of their afferent population (E and I respectively). As a consequence, the dynamics of a two population neural mass model can be described based on six state variables: the excitatory variables (V_E, λ_E, s_E) and the inhibitory variables (V_I, λ_I, s_I), together with the additional PSC η reflecting exogenous input to the structure. In the context of large-scale recordings, the neural population can be distributed across one or several spatial directions, possibly following the spatial spread of neurons' soma, axons, and dendrites. Following a classical approximation depicted in Figure 5B, inhibitory connections are considered to be local (Sik et al., 1995; Schomburg et al., 2012; Taxidis et al., 2012), such that global coupling between cells surrounding distinct recording sites is assumed to happen exclusively through excitatory axons ($s_E(x)$ may depend on λ_E at other spatial locations than x) and the possibly space-dependent exogenous input current $\eta(x)$.

For the LFP $L(t)$, resulting from the conduction of trans-membrane currents in the extracellular space, we assume the contribution of currents flowing through the membrane of interneurons is negligible, based on the weakness of the anisotropy induced by their dendritic geometry across the population (Nó, 1947; Lindén et al., 2011; Mazzoni et al., 2015). The LFP thus results exclusively from pyramidal cell's membrane currents. Which currents (IPSC, EPSC, leak current, exogenous current) affect the most the recorded LFP at a given spatial location depends on multiple factors: the geometry of the cells, the distribution of synapses (inhibitory, excitatory, exogenous) onto them, and the geometry of the electrodes (Buzsaki et al., 2012, 2013; Einevoll et al., 2013). Figure 5C provides a schematic of how the differentiated location of synaptic boutons over the dendritic tree may result in variable algebraic contributions of each type of current to each recording channel.

Low rank linear response theory and frequency analysis

Neural mass models as described above are typically assumed non-linear, and use a static sigmoidal transformation to convert membrane potentials into spike rates (see section Analytical neural field modeling of spike-field coupling in STAR Methods and Jirsa and Haken (1996)). However, assuming small-amplitude perturbations in the neighborhood of an operating point, important aspects of their structure can be captured by linearization of such a model (see Methods and e.g. Moran et al. (2007); Ledoux and Brunel (2011); Pinotsis et al. (2012)) This leads to a linear time-invariant model, whose behavior is fully characterized by its amplitude and phase response to oscillatory inputs at each frequency. This allows studying how the underlying microcircuit parameters influence the GPLA properties. We provide qualitative illustrations of this idea in simplistic scenarios. First, as illustrated in Figure 5D, consider the microcircuit receives exogenous input exclusively onto the pyramidal cells' dendrites (no feedforward inhibition) and I cells receive local excitatory input, but do not synapse back onto E cells (no feedback inhibition). We assume additionally that subthreshold activity is dominated by the exogenous input currents and proportional to the measured LFP. The

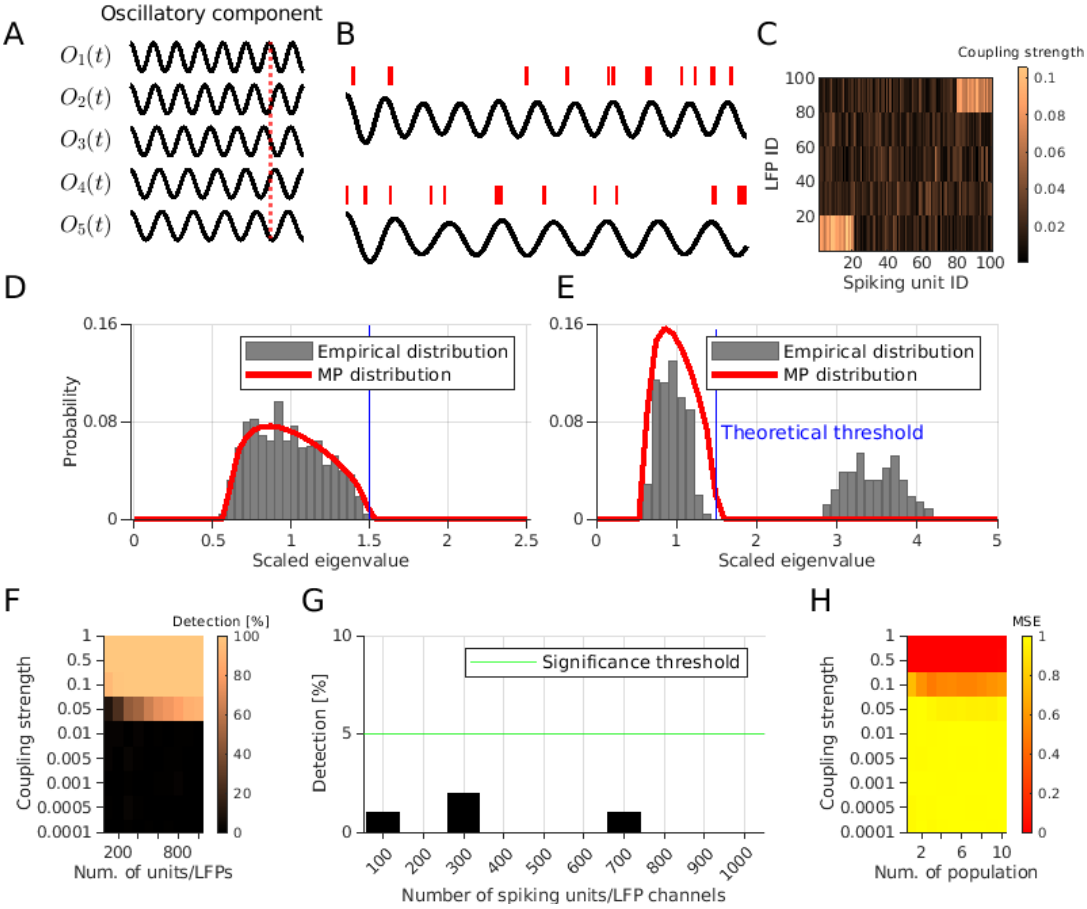

Figure 4. Quantifying statistical properties of GPLA with theoretical significance test.
(A) LFPs are synthesized by mixing several oscillatory components ($O_k(t)$). The vertical red line evidences the phase shift between them. **(B)** Two exemplary spike trains (each from one of the coupled populations) and the corresponding LFPs. In the LFP trace on the top, the oscillatory component with the highest frequency is dominated and for the bottom one, the lowest frequency component is the dominated oscillation. **(C)** An exemplary coupling matrix for a simulation with two coupled populations. **(D-E)** Theoretical Marchenko-Pastur distribution (red lines) and empirical distribution (gray bars) for (D) simulation without coupling and (E) with coupling between multivariate spikes and LFP **(F)** Performance of GPLA for the detection of coupling between spike trains and LFPs for different strength of coupling (y-axis) and different number of spiking units/LFP channels. **(G)** Type I error for different numbers of spiking units/LFP channels (x-axis), quantified as the percentage of simulations wherein a significant coupling between spike trains and LFPs is detected in absence of ground truth coupling. The horizontal green line indicates the %5 threshold. **(H)** Mean-squared-error of GPLA-based estimation of the number of populations coupled to LFP for varying coupling strengths (y-axis) and numbers of coupled populations (x-axis).

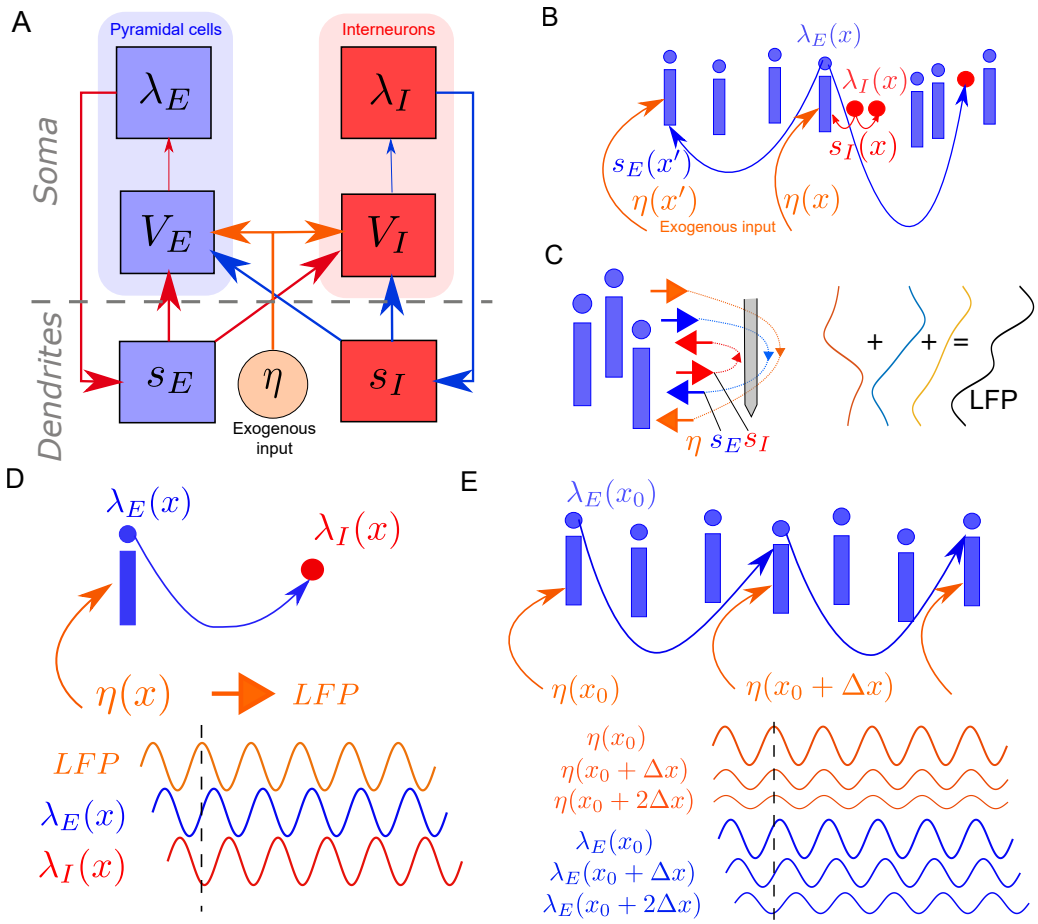

Figure 5. Generative model of spike-LFP coupling.
(A) A two-population neural field model of neural dynamics. V_k, λ_k and s_k indicate respectively somatic membrane potential, firing rate and post-synaptic current for Excitatory (k=E) and Inhibitory (k=I) populations. η indicates the exogenous input to the circuit. Arrows indicates the causal dependence between variables of the model. **(B)** Schematic representation of the model's connectivity: local inhibition and long range excitation, together with the driving by exogenous synaptic currents. **(C)** Schematic representation of the contribution of postsynaptic currents to the electric field, affected by the spatial distribution of synapses over the dendritic tree and the geometry of pyramidal cells. From left to right: Schematic representation of pyramidal neurons, electric field, electrode (gray bar), contribution of each current (EPSC, IPSC and exogenous current, leak current is also contributing to LFP but is not shown) to the LFP profile along the electrode's axis **(D)** Simple microcircuit structure leading to a temporal ordering of the local activities of different kinds $LFP \rightarrow excitation \rightarrow inhibition$ **(E)** Simple microcircuit structure leading to a temporal ordering of activities of the same kind across space: the location receiving stronger exogenous input leads other locations, such that amplitude gradient leads to phase gradients.

lag induced by the membrane potential dynamics results in a positive (frequency-dependent) delay of excitatory activity with respect to the LFP (reflecting the input), while inhibitory activity is itself delayed with respect to excitation. For an exogenous input oscillating at a given frequency, this implies a phase lag configuration between the (oscillatory) response of these variables, irrespective of the considered frequency.

Going beyond the relationship between different variables observed at the same location, we study how the activity of the same variable varies across space. For such a scenario, we extend spatially the previous model (no feedforward and feedback inhibition), by adding horizontal E-E connectivity with strength monotonously decreasing with distance, and study the activity resulting from a spatially inhomogeneous oscillatory input, with larger input amplitude at a given side (on the left in Figure 5E). Again due to the delay induced by membrane dynamics, this results in the propagation of the activity from one side to the opposite. This leads to an interesting relationship between the phase and amplitude of oscillatory activity: the location of largest amplitude is ahead of time with respect to the neighboring locations with smaller amplitudes. Note in particular that these propagation-like patterns are induced by the network's horizontal connectivity, while the input to the structure does not have phase lags at different locations (Ermentrout and Kleinfeld, 2001). These two examples indicate that phase and amplitude of oscillatory activities are informative about the underlying microcircuit structure and dynamics. More realistic scenarios should however take into account recurrent interactions between cell populations, and these above phases-amplitude relationships become frequency-dependent, as we will see in the next sections.

In order to analyze these more complex scenarios, a systematic way to evaluate the coupling between network variables at a given frequency is required. Consider the coupling between firing rate ($\lambda_E(x_1, t)$) and LFP ($L(x_2, t)$) at two different locations x_1 and x_2.

The linearity of the system implies the existence of transfer functions (denoted H_{λ_E} and H_L respectively), linking their activity to the time domain Fourier transform of the exogenous input $\hat{\eta}(x, f)$. Crucially, the first singular value approximation of GPLA can be obtained by assuming the space and time dependence of the input are separable ($\hat{\eta}(x', t) = n(x')\epsilon(t)$ implying $\hat{\eta}(x', f) = n(x')\hat{\epsilon}(f)$). This leads to both rate and LFP being proportional to the exogenous input at a given frequency, with the following expressions based on the transfer functions:

$$\lambda_E(x, f) = \int H_{\lambda_E}(x, x', f) n(x') \hat{\epsilon}(f) dx' = \psi_E(x, f)\hat{\epsilon}(f) \quad \text{and} \quad L(x, f) = \psi_L(x, f)\hat{\epsilon}(f)\,. \tag{4}$$

such that the coupling between spikes and LFP at locations x_1 and x_2 writes

$$C_{x_1,x_2}(f) = \psi_L(x_2, f)\psi_E(x_1, f)^* E(f)\,, \tag{5}$$

where $E(f)$ is the input power spectral density at frequency f and z^* denotes the complex conjugate of z. This shows that the coupling between L at x_2 and λ_E at x_1 is separable in the spatial variables (x_1, x_2), characterized by two functions of space: one for the field, ψ_L, and one for the excitatory spiking, ψ_E. Three quantities, ψ_L, ψ_E, and $E(f)$, are thus enough to reconstruct any pairwise coupling up to a scaling factor. In particular, as $E(f)$ is a positive number, the phase of C reflects a property of the circuit, independent of the exogenous input. Importantly, this analysis also applies to the coupling between the same variables at different locations, phase gradient thus informing us about the underlying spatial connectivity, and can account for excitatory and inhibitory activity at the same time.

In practice, we have access to a finite number of spatial measurement points of $C_{x_1,x_2}(f)$ corresponding to electrode channels where L and λ are recorded. This corresponds to the rectangular matrix $\boldsymbol{C}(f)$ estimated by GPLA, combining excitatory and inhibitory units. The above separability then implies that $\boldsymbol{C}(f)$ is a rank-one matrix, such that we can in principle apply Singular Value Decomposition (SVD) to estimate the coupling properties using the largest singular value and the associated left and right singular vectors, which are spatially discretized approximations of the above functions ψ_L and ψ_E (up to a multiplicative constant). Overall, the spatial distribution of the

phase and magnitude of each singular vector is influenced by the underlying recurrent dynamics and its propagation across horizontal connections (shaping the transfer functions such as H_{λ_E} and H_L), as well as the type of currents that dominate the LFP. As we will demonstrate in the next sections, the analysis of these features across frequencies is a rich source of information to validate assumptions about local network organization based on experimental multivariate data.

Application to spike-field dynamics during sharp wave-ripples

The phenomenon of hippocampal Sharp Wave-Ripples (SWR) is one of the most striking examples of neural activity entangling spike and LFP dynamics in multiple frequency bands, attributable to specific mechanisms of the underlying microcircuit (Buzsáki et al., 1992). Specifically, SWRs are brief episodes observed in hippocampal LFP traces combining a high-frequency oscillation (the ripple) to a low-frequency biphasic deflection (the sharp-wave). Moreover, these LFP activities are well known to be synchronized with spiking activity, with different cell-types firing at specific phases of the ripple oscillation (Buzsáki et al., 1992), although spike-field coupling at lower frequencies has also been demonstrated (Ramirez-Villegas et al., 2015).

We use simulations of in-vivo SWR described in Ramirez-Villegas et al. (2018) in order to demonstrate what insights GPLA can provide about the underlying hippocampal network mechanisms. The model generates realistic spiking and LFP activity in subfields CA1 and CA3, based on populations of two-compartment Hodgkin-Huxley neurons distributed along two one dimensional grids representing the strata of each subfield. In this model, the connectivity of CA3 is characterized by strong recurrent excitatory auto-associational $E-E$ connections, together with $E \to I$ connections and short-range $I \to E$ and $I \to E$ synapses (see Figure 6A for a schematic representation). In contrast, local $E-E$ connections are absent in CA1, but both E and I cells receive feedforward excitation from CA3. LFPs were generated from the total trans-membrane currents using line current density approximation, and measured by two laminar multi-shank electrodes (see STAR Methods, section Simulation of hippocampal sharp wave-ripples for more details).

We first assay the insights provided by GPLA on SWRs by analyzing a single hippocampal subfield, CA1, as various studies suggest SWRs emerge from it in response to afferent CA2- and CA3-ensemble synchronous discharges (Csicsvari et al., 2000; Oliva et al., 2016). In this simulation, LFP and unit recordings are distributed along two orthogonal spatial directions (laminar for LFPs and horizontal for units). We use a total of 157 peri-ripple traces of simulated LFPs and spikes of both populations (inhibitory and excitatory) of duration approximately 1 sec. Exemplary traces of simulated LFP and population firing rate of the CA1 population (pyramidal cells and inhibitory interneurons belonging to CA1) are shown in Figure 6B.

GPLA results for representative frequency bands are provided in Figure 6C-E and for all bands covering the 1-180Hz interval in supplemental Figure 10. The overall coupling magnitude (gPLV) was significant is all frequencies (Figure 6C), according to both surrogate (based on spike jittering, $p<0.05$) and analytical (based on random matrix theory) tests. In particular, the strongest coupling is detected in the ripple band (80-180 Hz), which is compatible with results based on classical univariate techniques (Buzsáki et al., 1992).

The LFP vectors exhibit a biphasic electric potential profile along the spatial axis along which the electrode contacts are distributed, typical of laminar recordings (Figure 6D), and corresponding to the field generated by the dipolar geometric arrangement of sources and sinks in the parallel two-compartment models of pyramidal neurons used for this simulation. Notably, the two-compartment approximation has been shown to lead to results comparable to simulations based on passive membrane currents of realistic morphologies (see Figure 6D (right) and Ramirez-Villegas et al. (2018) for more details). To check the quantitative agreement between the LFP vector and the original model of LFP generation in this simulation, we computed analytically the total LFP generated passively by all pyramidal cells using the original LFP simulation code of Ramirez-Villegas et al. (2018), and assuming all cells have identical trans-membrane currents flowing through their somatic and dendritic compartments (see STAR Methods section Computation of the laminar LFP profile).

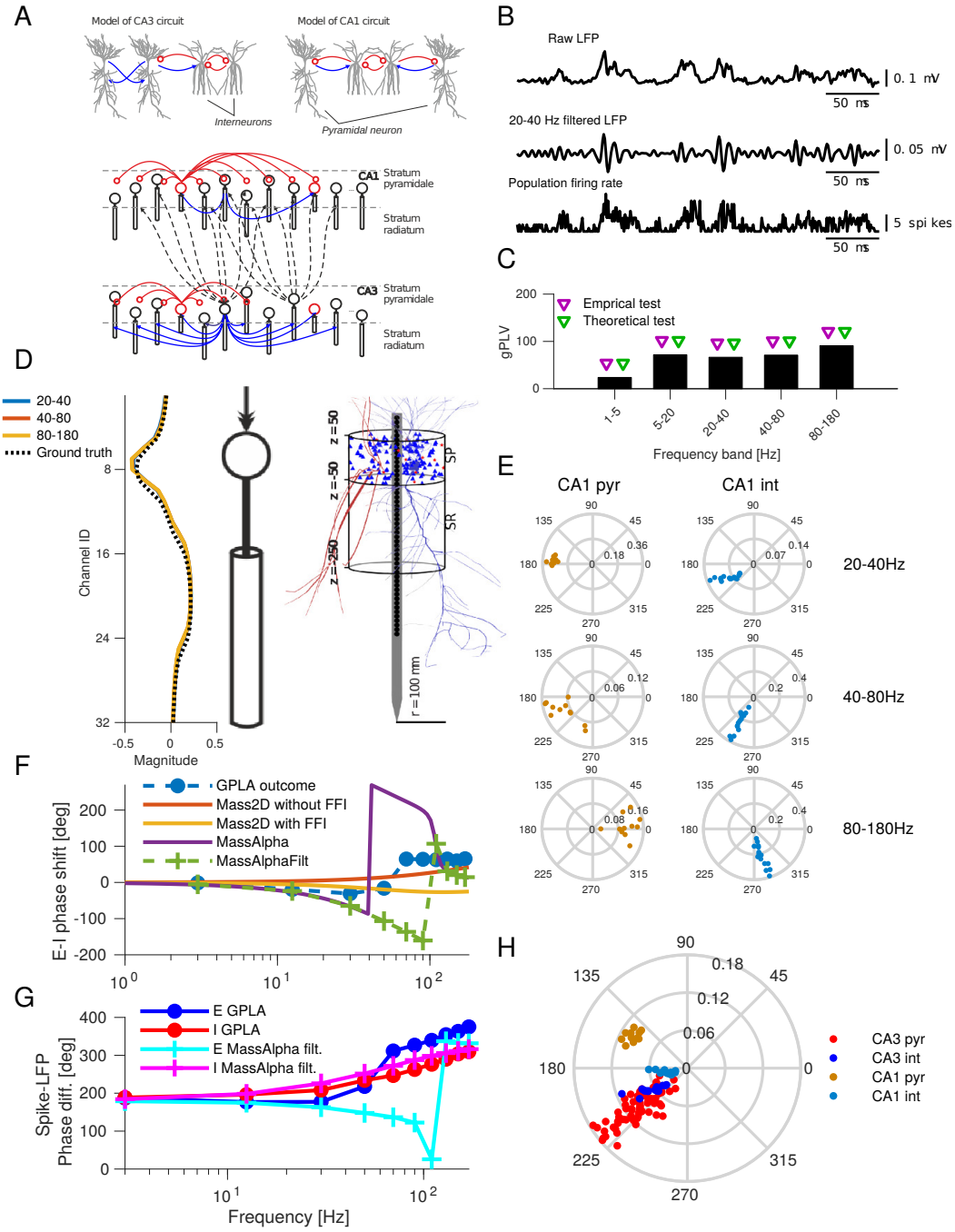

Figure 6. Analysis of hippocampal SWRs generated by a biophysical model

(A) Hippocampal multi-compartment model. Top: Canonical circuits of CA1 and CA3. Bottom: Schematic of the whole model (blue, excitatory connections; red, inhibitory, modified from Ramirez-Villegas et al. (2018). **(B)** From top to bottom: Example broad band CA1 LFP trace, band-pass filtered trace of the CA1 LFP in ripple band (80-180 Hz), and population firing rate of CA1 neurons. **(C)** CA1 gPLVs. Triangles indicate the significance assessed based on empirical (blue triangles, $p<0.05$) and theoretical (red triangles) tests. **(D)** Left: LFP vectors for GPLA of CA1, with a schematic of the pyramidal two-compartment model. Right: Location of the somata and expample morphologies of each cell type (blue, pyramidal cells; red, interneurons) in an equivalent model with many compartments (right hand-side schematic modified from Ramirez-Villegas et al. (2018)). **(E)** Spike vector coefficients for CA1 in several frequency bands (left: pyramidal cells, right: interneurons). **(F)** Average phase lag between LFP and spike vectors across frequencies for: outcome GPLA on hippocampal SWRs, theoretical analysis of *Mass2D* (without and with feedforward inhibition) and *MassAlpha* neural mass models. Dashed green line indicate *MassAlpha* filtered over the frequency bands used for GPLA. **(G)** Difference between phases of E and I populations based on GPLA the *MassAlpha* neural mass model filtered in the same bands (IPSP was used as LFP proxy). **(H)** Spike vector resulting GPLA jointly applied to CA1 and CA3 in the gamma band (20-40 Hz).

While the dendritic current reflects the post-synaptic input of the cell, somatic currents are taken opposite to preserve the charge neutrality of each cell. The resulting theoretical LFP profile of the pyramidal population are highly similar to the LFP vector (cosine similarity > 0.97 for LFP vector of all three frequencies in Figure 6D). Note that the sign of the LFP vector, only reflects our convention of setting the phase of its mean to zero: as the LFP vector coefficients are divided into two groups of opposite sign, a positive sign is attributed to the set of coefficients that weight the most in the overall sum. In the context of laminar recordings, one could as well adopt a different convention ascribing a fixed sign to peri-somatic LFP, reflecting the sign obtained through classical analyses, e. g. triggered averaging based on spikes or oscillatory peak (Sullivan et al., 2011). This suggest that the LFP vector can be exploited to further study the current sources and sinks causing the LFP, e. g. through current source density analysis (Wójcik, 2013).

The spike vectors components distribution in the complex plane (Figure 6E) support that both E and I cells are synchronized to CA1 LFP in the ripple band (80-180 Hz), but at a different phase, in line with experimental and simulation results (Buzsáki et al., 1992; Ramirez-Villegas et al., 2018). This extends our previous observation of Figure 5D to a more realistic case of recurrent E-I interactions. Notably, we observe that while GPLA does not exploit neuron type information, the spike vector can clearly differentiate pyramidal cells from interneurons based on their respective phase in the spike vector, showing that interneurons lead pyramidal cells in lower frequency bands, barring a drastic change in the distribution of phases in high frequencies (see also Figure 6F). This can be used not only for inferring cell types from experimental data Varga et al. (2014), but also to infer the organization and dynamics of neural circuits based on analytical or computational models.

In order to illustrate this second approach, we use the linear response framework (see above section Low rank linear response theory and frequency analysis) to analyze the phase of the populations in neural mass models of the microcircuit activity with different levels of complexity. In line with (Ledoux and Brunel, 2011), we first took into account only membrane time constants in the *Mass2D* model (resulting from membrane capacitance and leak currents), and neglected completely synaptic dynamics (see STAR Methods, section Analysis and simulation of two population neural mass models). As a result, *Mass2D* has a 2 dimensional state space, allowing for resonance due to interactions between pyramidal cells and interneurons (a mechanism exploited in the *PING* model of gamma oscillations (Traub et al., 1999)). However, as shown in Figure 6F, the predicted phase shift across frequencies could neither account for the driving by interneurons in CA1, nor for phase changes in high frequencies (>30Hz). Notably, incorporating strong feedforward inhibition (FFI) did not improve the qualitative match between the analytical predictions and GPLA's outcome. The inappropriateness of *Mass2D* is in line with the current understanding of SWR emergence in CA1 through the pacing of pyramidal activity by delayed I-I interactions (Stark et al., 2014), as the model does not account for such interactions.

The emergence of oscillations through I-I interactions is supported by mathematical modeling showing that sufficiently strong delayed recurrent inhibition gives rise to resonance or sustained oscillations (Brunel and Wang, 2003). The presence of these oscillations can be accounted for in an extended version of our neural mass model by including an additional synaptic delay and/or a synaptic time constant for I-I synapses (Ledoux and Brunel, 2011). We build the *MassAlpha* model that introduces a lag in the post-synaptic dynamics through the use of the so-called alpha synapses (see Methods for details). Interestingly, the resulting E-I phase shift of this model is in qualitative agreement with GPLA estimation (Figure 6F), exhibiting a reversal in the lead-lag relation between populations as frequency grows. This supports the ability of GPLA, combined with appropriate simulations, to capture key characteristics of the underlying circuitry (in this simulation having oscillation dynamics and synaptic delays).

Another interesting property of the network is the phase shift between each individual population and the LFP, which is simply reflected in the phases of the spike vector coefficients averaged across each population (E and I). Given that the LFP is a linear combination of all post-synaptic currents of the network, we can use modeling to evaluate which of these currents is the most repre-

sentative of the observed spike-LFP phase relation. As shown in Figure 6G (also see Supplementary Figure 9 for EPSP proxy), the choice of the IPSP as an LFP proxy in the MassAlpha model accounts qualitatively, as frequency increases, for (1) monotonous phase increase of the I population, (2) the phase slope reversal of the E population. Overall, the simulation results are also in line with the experimentally observed synchronizations of these populations. In contrast, using EPSP as an LFP proxy still fails to reproduce these two aspects (see Supplementary Figure 9), illustrating how GPLA, beyond microcircuit dynamics, may also help address the cellular underpinnings of experimentally observed LFP (Teleńczuk et al., 2017). This overall suggests that GPLA combined with neural mass modeling of a structure can provide insights into the microcircuit dynamics underlying phenomena as complex as sharp-wave ripples, despite neglecting many biophysical details.

Moreover, GPLA can also provide further insights when recordings from multiple regions are available. As an illustration, Figure 6H depicts coefficients of the spike vector for the joint analysis of CA1 and CA3 (LFP and spikes of both areas is used for GPLA) showing CA1 and CA3 neurons are all coupled to the field activity with cell-type-specific phases in the gamma band (20-40 Hz) (see supplemental Figure 11) that are consistent with the GPLAs obtained from individual structures (see supplemental Figure 10). This suggests that the gamma rhythm has a dominant coherent component spanning the two structures, and thus may support communication between subfields during SWR and consequently memory trace replay, consistently with current hypotheses (Carr et al., 2012).

Application to spatio-temporal patterns of neural field models

In order to assess the interpretability of GPLA in the context of electrode arrays covering a piece of the cortical surface, we simulated a 2D neural field model as described in Figure 5. We used an exponentially decaying horizontal excitatory connectivity with a spatial scale constant $r_0 = 440\mu m$, in line with recent analyses of cortical recordings Teleńczuk et al. (2017) (see STAR Methods section Analysis and simulations of neural field models for details). The spatio-temporal dynamics was down-sampled spatially on a grid with a step size $\Delta x = 800\mu m$, representing the inter-electrode distance of a putative electrode array of $1.2cm$ size. We compared the spatio-temporal dynamics for two choices of connectivity for which the input-free network has a stable equilibrium. First, we consider the *weak inhibition* case (Figure 7A), for which inhibitory (I) cells have weak feedback inhibition ($I \rightarrow E$), relative to the self-excitation caused by $E - E$ horizontal connections. The resulting activity is akin to stochastic fluctuations, due to the exogenous input, around a *stable node* equilibrium. Second, in the *strong inhibition* case (Figure 7B), the larger excitability of inhibitory neurons strengthens their influence on excitation and leads to activity fluctuating around a *stable spiral equilibrium*, reflecting a tendency of perturbations to oscillate around this point (Figure 7B) (Onslow et al., 2014). In both cases, the computed excitatory population rate is used to simulate the spike train of one excitatory unit per spatial electrode on this grid, in line with the observation that excitatory units are more easily detected experimentally due to their open field configuration (Logothetis and Panzeri, 2014). GPLA is then computed between this excitatory spiking activity and different LFP proxies. The results in Figure 7C-H are computed using the total EPSP resulting from horizontal E-E connections as LFP proxy (i. e. excluding exogenous excitation). We observe key differences between the GPLA of the two systems, predicted by linear response theory (see STAR Methods section Analysis and simulations of neural field models)

First, as reflected in the gPLV values (Figure 7C), spike-field coupling appears stronger in the lower frequency bands in the case of weak recurrent inhibition, while in the case of strong recurrent inhibition we observe a stronger coupling at intermediate frequencies. Notably, the peak of spike-field coupling in intermediate frequencies for strong inhibition is in line with models of the prefrontal cortex with the same enhanced feedback inhibition (Sherfey et al., 2018), exhibiting a resonance in the beta range (25Hz).

Second, as demonstrated in the previous neural mass model simulation, the global spike-LFP phase shift may also be informative about the underlying neural circuits. We can compute the

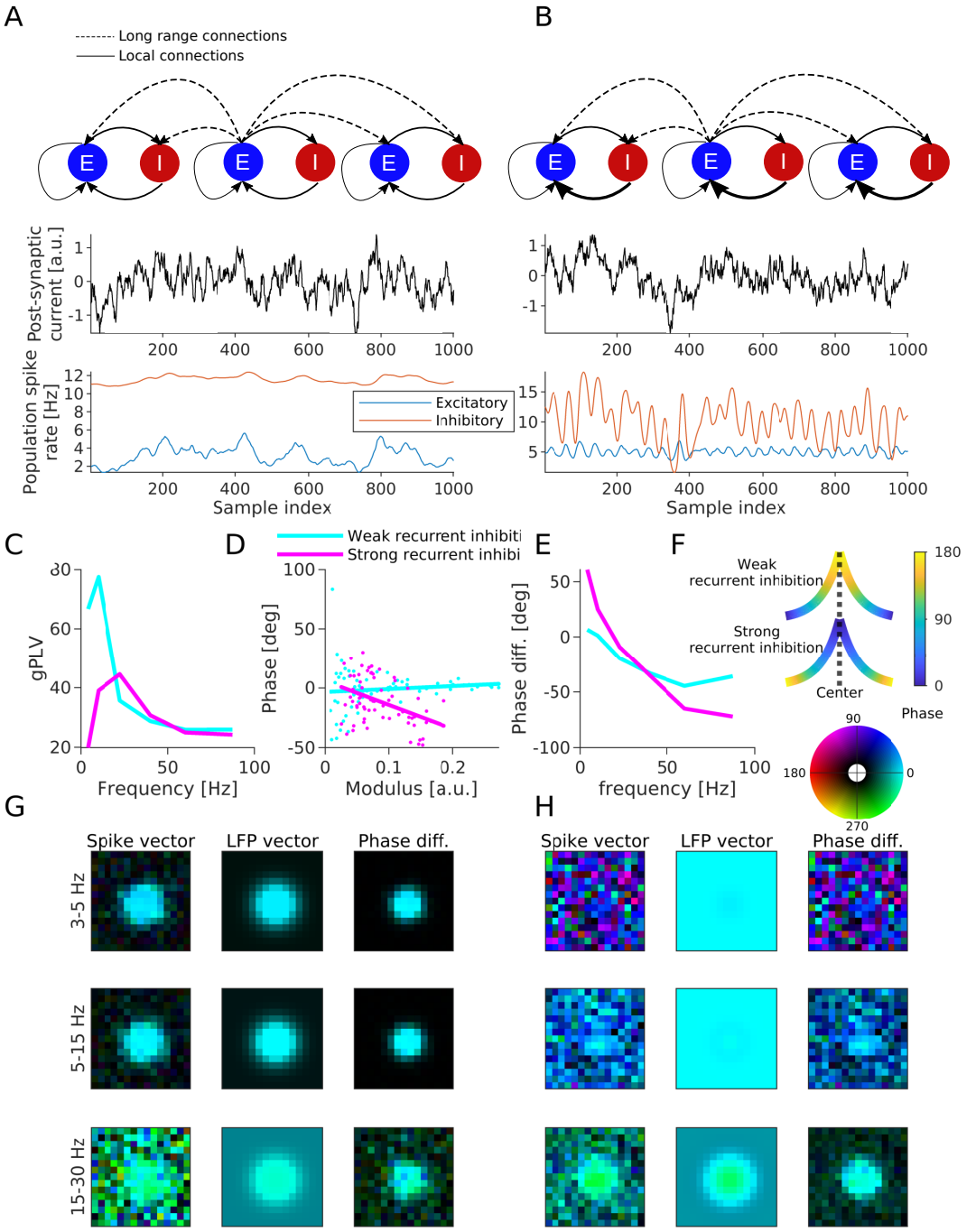

Figure 7. Neural field simulation. (A) Simulation of a neural field with weak recurrent inhibition. Example time course of (from top to bottom) exogenous input, E- and I- populations rates at the center of the neural field (in green, blue and red). **(B)** Same as A for a neural field strong recurrent inhibition, which shows oscillatory dynamics due to increase in recurrent activity. **(C)** gPLV as a function of frequency for both models. **(D)** Phase of spike vector coefficients as a function of its modulus for the frequency band associated with maximum gPLV for both models (each dot one coefficient, and the continuous lines are plotted based on linear regression). **(E)** Shift between the averaged phase of spike vector and averaged phase of LFP vector, as a function of frequency. **(F)** Schematic of the spike vector's phase gradient in the two models according to Equation 6. X-axis is the distance from center and y-axis is the connectivity strength. Line color indicates the phase according to the colorbar on the right. **(G)** Resulting GPLA in 3 frequency bands (indicated on the left) for neural field with weak recurrent inhibition (model schematized in A). **(H)** Same as G for neural field with strong recurrent inhibition (model schematized in B). In both G and H, color of pixel code the values of spike/LFP vector coefficients, with colorbar on top of H. Colors are represented in HSV mode, in which a complex number ($re^{i\phi}$) is represented by hue and brightness of a pixel. Hue of a pixel indicates the phase (ϕ) and the brightness of a pixel indicates the magnitude (r).

average phase shift between the spike and LFP vectors as a function of the frequency band to see a clear difference between the two models. Strong recurrent inhibition leads to phase advance of the spiking activity in the low frequency, in contrast with the weak recurrent inhibition case showing a consistent lag of excitatory spiking across frequencies (Figure 7E).

Third, the relationship between the spatial variations of modulus and phase of the spike vector is different across these two networks. In the simulation with strong recurrent inhibition, the phase of spike vector coefficients as a function of their modulus for the frequency band associated with maximum gPLV for each model indicates that the phase of the spike vector coefficients decreases (i. e. the oscillation lags further relative to the LFP) for larger modulus ($p < 10^{-4}$, linear regression), whereas, in the simulation with weak recurrent inhibition, phase is not correlated with modulus ($p > 0.3$, linear regression) (Figure 7D).

This last difference between the two cases can be directly interpreted based on the spatial maps of spike vector coefficients across the array. Indeed, models exhibit a different radial phase map in both situations (Figure 7E-H), reflecting how phase changes as magnitude decreases when going away from the center (the location with largest input). This gradient can be predicted by theoretical analysis of a one dimensional neural field as we show in detail in STAR Methods section Spatio-temporal phase analysis in 1D. Briefly, to obtain the spike vector, the input spatial pattern at a given temporal frequency f is approximately convolved by a spatial convolution kernel of the form

$$k(x) = e^{-|x|a(f)} = e^{-|x|\mathrm{Re}[a(f)]} e^{-i|x|\mathrm{Im}[a(f)]} \,. \tag{6}$$

The first term of this kernel has a negative real number multiplied by distance in the exponential that makes the activity decrease away from the locations where exogenous input is the highest, as intuitively expected from the horizontal connectivity of the circuit. In contrast, the imaginary number in the argument of the second term enforces a spatial phase gradient in response to the input, which depends on the sign of the imaginary part of a. If this sign is positive, responses at the location of highest input will be ahead of time with respect to their surrounding in the considered band, as reflected by their larger spike vector phase in the top illustration of Figure 7F. In case $\mathrm{Im}[a]$ is negative, locations with highest input are lagging behind (bottom illustration of Figure 7F). The frequency-dependent complex number $a(f)$ that controls this behavior reveals the key microcircuit properties that determine our observation through the approximation relation (valid a low frequency)

$$a^2 \approx \frac{1}{r_0^2} \left(1 + \nu_{E\leftarrow I}\nu_{I\leftarrow E} - \nu_{E\leftarrow E} - \boldsymbol{i} 2\pi\tau f (2\nu_{E\leftarrow I}\nu_{I\leftarrow E} - \nu_{E\leftarrow E})\right) \,. \tag{7}$$

It can be deduced from this expression (taking the square root of this number) that the imaginary part of a will depend on the relative strength of recurrent inhibition onto pyramidal cells, controlled by $\nu_{E\leftarrow I}\nu_{I\leftarrow E}$, with respect to recurrent excitation controlled by $\nu_{E\leftarrow E}$. Intuitively, $\nu_{E\leftarrow I}\nu_{I\leftarrow E} >> \nu_{E\leftarrow E}$ will tend to suppress the input that created the response, generating a wave converging to the points were the input was highest. The theory also predicts that large values of $\nu_{E\leftarrow I}\nu_{I\leftarrow E}$, as used in the strongly recurrent simulation, can generate strong phase gradients. In contrast, linear stability constrains the values of $\nu_{E\leftarrow E}$ to remain small, reflecting our choice for the simulations, and resulting in a comparatively moderate slope for the weakly recurrent case.

Overall, contrasting these two cases of neural field connectivity shows that changes in strength of feedback inhibition are reflected not only in the dominant frequency of spike-LFP synchronization (Figure 7C), but also in the spike-LFP shifts of the GPLA results (Figure 7E), and in the relationship between modulus and phase of spike vector coefficients (Figure 7D). Notably, these observations are being made in absence of specific oscillatory activity nor spatial phase gradient of the exogenous input (which influences the activity synchronously across the array). Therefore, it supports that the observation of complex coordinated activity involving phase spatial gradients, such as oscillations and traveling waves-like phase gradients, in array recordings may emerge from local

recurrent interactions in the recorded regions, instead of resulting from the passive driving by spatio-temporally coordinated activity originating from other brain regions.

Analysis of Utah array data in the prefrontal cortex

After testing the capabilities of GPLA on simulations, we investigated its performance on a large-scale electrophysiological dataset. We applied the method on Utah array (10×10 electrodes, inter-electrode distance $400\mu m$) recordings from the ventrolateral prefrontal cortex of one anaesthetized rhesus monkey (see Figure 8A). LFP signals were preprocessed as described in STAR Methods, section Animal preparation and intracortical recordings, and multi-unit activity with a minimum of 5 Hz firing rate was used. Recorded signals are exemplified in Figure 8B-F. Exemplary LFP traces are illustrated in Figure 8B. Each trace is recorded from the location specified in Figure 8C. Spike trains are also displayed in Figure 8D (for the same epoch used in Figure 8B). As the analysis is performed in band-limited frequency ranges we also exemplified band-passed LFP signals (together with spikes) in Figure 8E and F. The dataset consisted of 200 trials of visual stimulation (10 sec) and intertrials (10 sec) each 20 sec.

Computing GPLA in different frequency bands revealed that the strongest coupling was in the alpha range (5-15Hz) (Figure 8G). Furthermore, we assessed the significance of coupling with both surrogate and analytical tests (see STAR Methods, section Significance assessment of gPLV). GPLA above 50 or 60 Hz should be considered with caution, as in high frequencies the spike-LFP relationships can be largely effected by the local coupling between spikes and LFP recorded from the same electrode. Furthemore, contamination of high frequency LFP bands by spike waveforms (Zanos et al., 2011; Ray and Maunsell, 2011b) may bias spike-LFP coupling towards interactions within the same electrode contact instead of capturing coordination at the mesoscopic scale.

Figure 8I-K further shows the spike and LFP vectors for the three frequencies with largest coupling according to their gPLVs (for other frequencies see Supplementary Figure 13). Representing spike and LFP vectors in the complex plane (Figure 8I-J first column), suggests that the relative phases of spike and LFP vectors are different across these three frequencies. To demonstrate the difference more clearly, we fit von Mises distributions to the pooled phase of all coefficients of the vectors (Figure 8I-J second column). The sign of the spike-LFP phase differences changes across frequencies, with spikes ahead of time with respect to LFP in low frequency, while lagging at higher frequencies. This behavior is similar to the above analysis of strongly recurrent neural field model (Figure 7G), when EPSP is taken as an LFP proxy.

The spatial mappings of the LFP and spike vectors on the Utah array (Figure 8I-J, third and fourth column) also demonstrate a spatial structure in the modulus and phase of the LFP and spike vectors, revealing localized regions with stronger participation in the locking, in particular in the beta range 15-30 Hz (green pixels at the middle-top and -bottom in Figure 8K, fourth column). We hypothesize this is due to a higher activation of spatially localized populations, as supported by anatomical studies of the PFC (Elston, 2000, 2003) and electrophysiological (Safavi et al., 2018) studies.

Furthermore, in the lower frequencies, in particular the alpha band (5-15 Hz) exhibiting the strongest coupling between spike and LFP, the spike vector coefficients' moduli are significantly negatively correlated with their phase (Figure 8H ($p < 10^{-6}$, linear regression). Interestingly, we observe again a similar behavior in the above neural field simulation with strong recurrent inhibition, but not in the simulation with weak recurrent inhibition (Figure 7F). Overall, these results suggest a neural field with excitatory horizontal connections and strong local recurrent inhibition as a plausible model for the recorded prefrontal circuits, in line with what has been suggested by previous modeling work (Sherfey et al., 2018, 2020).

Discussion

In spite of the strong relation between LFP activity and the neural mechanisms involved in the coordination of the underlying networks (Buzsaki et al., 2012, 2013; Einevoll et al., 2013; Herreras, 2016; Pesaran et al., 2018), spike-LFP relationship is still not systematically exploited in extensive

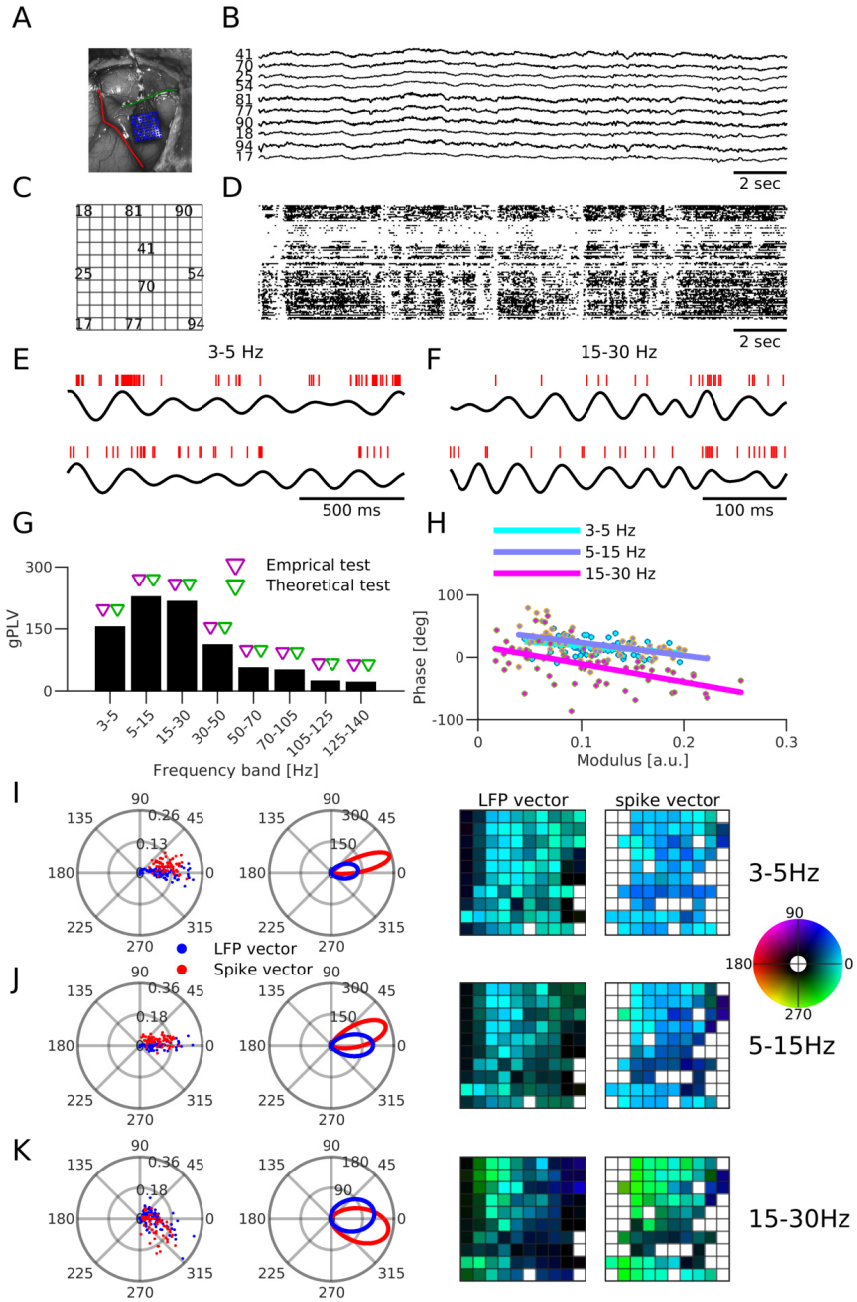

Figure 8. Application to electrophysiological recordings in non-human primate PFC.
(A) Location of the Utah array, anterior to the arcuate sulcus (red line) and inferior to the principal sulcus (green line). **(B)** Broadband trace of the recorded LFP (from the recording channels indicated in C). **(C)** Utah array spatial map identifying channel IDs shown in B. **(D)** Spike rasters for all recorded neurons. **(E-F)** Example spike trains (red bars) and filtered LFP (black traces) in the frequency ranges (E) 3-5 Hz and (F) 15-30 Hz. **(G)** gPLV values. Triangles indicate the significance assessed based on surrogate (blue triangles) and analytical test (red triangles) tests. **(H)** Phase of spike vector coefficients as a function of its modulus for the frequencies indicated in the legend (one dot per coefficient, continuous lines indicate linear regression). **(I-K)** LFP and spike vectors for frequency (I) 3-5 Hz and (F) 15-30 Hz. First column depicts the LFP (blue dots) and spike (red dots) in the complex plane. Second column depicts the fitted von Mises distribution to phase of LFP and spike vectors. Third and forth columns respectively represent the spatial distribution of phase of LFP and spike vectors values on the array (see C). White pixels in the third column (LFP vector) indicate the recording channels that were not used in the recording and in the forth column (spike vector), white pixels indicate the recording channels with insufficient number of spikes (multiunit activity with a minimum of 5 Hz firing). In the last two columns, colors are represented in HSV mode, in which a complex number ($re^{i\phi}$) is represented by hue and brightness of a pixel. Hue of a pixel indicates the phase (ϕ) and the brightness of a pixel indicates the modulus (r). The colorbar is depicted on the right.

quantitative analysis of highly multivariate neural recordings. One potential reason could be the lack of multivariate methodologies for investigating spike-LFP coupling beyond a single pair of spiking unit and a LFP channel.

In this study, we develop the Generalized Phase Locking Analysis (GPLA), to the best of our knowledge, as the first biophysically interpretable *multivariate* method for investigating the coupling between spatially distributed spiking and LFP activity. GPLA summarizes the coupling between multiple LFP spatio-temporal patterns and multiple spiking units in a compact way. At a given frequency, the spike and LFP vectors represent the dominant LFP and spiking spatio-temporal distribution and the generalized Phase Locking Value (gPLV), characterizes the strength of the coupling between LFP and spike patterns.

We demonstrate that by using outcomes of GPLA such as the overall spike-LFP phase shift, the phase shift between different cell types (excitatory and inhibitory), and the spatial phase gradients, we can extract information about the overall organization of the recorded structure. In particular, we demonstrated that GPLA features can reflect properties pertaining to the organization of recurrent interactions in the recorded region that are not directly accessible by simple measurements. First, application to realistic simulations of hippocampal SWR revealed various characteristics of hippocampal circuitry with minimal prior knowledge. Second, in order to better understand the interpretation of spike and LFP vectors' spatial structure, we also simulated spatially extended neural field models and demonstrate that phase gradients of spike and LFP vectors in these neural field models reflect the underlying microcircuit connectivity (such as the strength of recurrent interactions). Finally, the application of GPLA to experimental recordings suggests a global coupling between spiking activity and LFP traveling wave in vlPFC in line with our simulations of a neural field endowed with strong recurrent inhibition.

Statistical properties of the gPLV were investigated to develop an empirical and theoretical framework for the assessment of the significance of the coupling. The theoretical statistical test makes the method applicable to high dimensional electrophysiology data with low run-time complexity, which is important for modern probes such as Neuropixel with 960 recording sites (Jun et al., 2017). Notably, the conventional statistical testing procedures based on the generation of surrogate data are computationally very expensive as their computation time is scaled by the number of recorded neurons and number of surrogate datasets generated for construction of a null distribution.

In summary, GPLA is a powerful tool to quantify, statistically assess and interpret the interactions between spiking activity and large-scale dynamics of current recording techniques.

Comparison to existing approaches

To the best of our knowledge, there are very few studies that include the information of multiple LFP channels *and* multiple spiking units for investigating spike-LFP interactions. Among those few studies which employ a multi-variate approach, none fully exploited the multivariate nature of spiking activity recorded from multiple sites, but instead, restrict the multivariate aspect only to LFP channels.

Spike-Triggered Average (STA) of LFP is one of the common multivariate technique for characterizing spike-LFP relationship (Jin et al., 2008; Nauhaus et al., 2009). Spike-triggered average of LFP has been also interpreted as a measure of functional connectivity (Nauhaus et al., 2009), therefore it is also interpretable (but also see Ray and Maunsell (2011a) and Nauhaus et al. (2012)). Although in LFP spike-triggered averaging the multivariate aspect of the LFP signals is exploited, ultimately only individual spiking units are being used as each spike-triggered average is only computed based on a single spiking unit. However, considering the information of all spiking units simultaneously can be potentially informative. For instance, the phase difference between different types of neurons (e. g. excitatory and inhibitory neurons) (Klausberger et al., 2003) can inform us about the organization of the neural circuit (as we illustrate with one of our simulations) Varga et al. (2014). Notably, even more sophisticated extensions of spike-triggered averaging of LFP (Teleńczuk et al.,

2017) are still using information of the individual spiking units in their methodology. In a similar vein, Canolty et al. (2010) by exploiting maximum entropy models, showed that the probability of spiking can be modulated not only by the phase of the local LFP, but also by the LFP phase in multiple distant regions. This study was also limited to individual spiking units.

Limitations and potential extensions

One limitation of GPLA is that it considers the underlying networks dynamics to be fixed for the analyzed data. Identifying transient activities that potentially show differentiated signatures of cell assemblies activation is of paramount importance (Harris, 2005; Buzsaki, 2010). As mentioned earlier, LFPs result from the superposition of electric potentials from multiple sources and can capture various *coordinated or cooperative* phenomena. LFP decomposition techniques can temporally isolate these epochs of coordinated activity and application of GPLA to these epochs can characterize how each neuron is participating in the collective activity and/or to what degree, it is coupled to the larger-scale dynamics.

Furthermore, GPLA can also be improved by exploiting a better univariate estimation method. Various novel methodologies for assessing the coupling between a pair of spike train and LFP time course have been developed in recent years (Grasse and Moxon, 2010; Vinck et al., 2010; Lepage et al., 2011; Vinck et al., 2012; Jiang et al., 2015; Zarei et al., 2018; Li et al., 2016) each providing some improvements over classical measures such as PLV and SFC. Replacing coefficients of coupling matrix (Eq 15) with better estimators of the coupling between spike and LFPs can bring those benefits to GPLA as well. For instance, Zarei et al. (2018) proposed a bias-free estimation of spike-LFP coupling in low firing regime which can also improve GPLA if we replace the coefficients of the coupling matrix (Equation 15) with the new estimator. Nevertheless, the current settings i. e. using a normalized version of the conventional PLV (Equation 13) is suitable for the theoretical statistical test that we introduce in Analytical test. Alternative univariate coupling estimates should be adapted to the requirement of the Random Matrix Theory to preserve the statistical benefits of our approach. This typically requires knowing the asymptotic distribution of the coupling statistics and devising and appropriate normalization thereafter. In case the new coupling measures are not adaptable to theoretical statistical testing, the Surrogate-based test that we discussed earlier remains applicable at the expense of heavier computational costs.

Neuroscientific interpretation of GPLA

Due to the complexity of the structure and dynamics of spatially extended neural networks, even a statistically sound approach such as GPLA (as demonstrated in simulations), might be challenging to interpret in terms of biological mechanisms. In order to ease interpretability, we related it to linear response analysis of neural field models. Thanks to this approach, we could link several features of GPLA to a mechanistic interpretation. First, we chose a biophysically realistic model of hippocampal ripples in order to use a system for which the underlying mechanism are (1) well understood, (2) more complex than the neural mass models used to interpret GPLA results. Despite the discrepancy between models, this showed that increasing the complexity of neural mass models using properties that are qualitatively in line with the key ground truth underlying mechanisms (e. g. inhibitory synaptic delays), allowed reproducing qualitatively GPLA results of these simulations, making the approach more interpretable. This allowed in particular (1) to show the LFP vector reflects the laminar distribution of field potential generated by current dipoles, (2) to link the phases of the spike vector to cell types and recurrent I-I dynamics,

Next, we used neural field simulations in order to find interpretations of GPLA characteristics that can be exploited in the context of electrode array recording in the cortex. This is an important step as the mechanisms underlying spatio-temporal phenomena observed in *in vivo* recordings remain largely elusive. While keeping the complexity of these models minimal (using exponentially decaying horizontal excitation and local inhibition), we could already observe the microcircuit structure produced non-trivial qualitative changes in the GPLA outcome, in particular regarding

the phase gradients of spike and LFP vectors across the array. Finally, our analysis of Utah array recordings suggests the key GPLA features exhibited in simulation can also be estimated in real data and provide insights in the underlying organization of the circuit.

Overall, the simple rate models we investigated have the benefit of lending themselves to approximate analytical treatment, providing direct insights into the role played by network parameters in GPLA characteristics. However, refinements of these first analyses are possible by improving the biological realism of population models and will allow checking the robustness of our simplifying assumptions. In particular, neural data recorded with multi-electrode arrays may be investigated in light of the knowledge about the horizontal connectivity of the structure, which may not be monotonous (for example see recent findings on non-monotonous correlation structure in V1 (Rosenbaum et al., 2017) and PFC (Safavi et al., 2018)).

More generally, a mechanistic interpretation of GPLA results in a given structure strongly relies on the accuracy of the assumptions made to perform analytical and/or computational modeling. We have shown neural mass models can approximate the population behavior of rather complex dynamics, such as sharp-wave ripples, and further sophistication of mass models will certainly help generalize the interpretability of GPLA analysis to include more biophysically realistic results (Jirsa and Haken, 1996; Qubbaj and Jirsa, 2007; Costa et al., 2016).

Ultimately, we believe GPLA provides insights for understanding the distributed information processing in higher-tier cortical areas such as PFC and hippocampus where investigating distributed spike-LFP interactions has been shown to be insightful for elucidating the mechanisms of cognitive functions such as working memory, memory consolidation and spatial navigation. Indeed, previous studies exploited spike-LFP relationship to shed light on the coordination mechanisms involved in working memory (Liebe et al., 2012; Markowitz et al., 2015), memory consolidation and spatial navigation (Taxidis et al., 2015; Fernández-Ruiz et al., 2017). Certainly, exploiting multivariate methods for investigating spike-LFP relationships can provide further insights about the coordination mechanisms involved in such cognitive functions.

Acknowledgments

We thank Britni Crocker for help with preprocessing of the data and spike sorting; Joachim Werner and Michael Schnabel for their excellent IT support. Andreas Tolias for help with the initial implantation's of the Utah arrays; scidraw.io for providing a free repository of high-quality scientific drawings (in particular Macauley Smith Breault for providing her brain drawing in this repository). This work was supported by the Max Planck Society.

Author contribution

Conceptualization, S.S., T.I.P., M.B.; Methodology, S.S., J.F.R.-V. and M.B.; Software, S.S. and M.B.; Formal Analysis, S.S. and M.B.; Investigation, S.S., T.I.P., V.K. and M.B.; Resources, N.K.L.; Data Curation, S.S., T.I.P., V.K., and M.B.; Writing - Original Draft, S.S. and M.B.; Writing - Review & Editing: S.S., T.I.P., V.K., J.F.R.-V., N.K.L. and M.B.; Visualization, S.S. and M.B.; Supervision and Project administration, T.I.P. and M.B.; Funding acquisition, N.K.L.

References

Aalen, O. O., Borgan, Ø., and Gjessing, H. K. (2008). *Survival and event history analysis: a process point of view*. Statistics for Biology and Health. Springer, New York, NY. OCLC: 254319944.

Abramowitz, M., Stegun, I. A., et al. (1972). Handbook of mathematical functions with formulas, graphs, and mathematical tables.

Anderson, G. W., Guionnet, A., and Zeitouni, O. (2010). *An Introduction to Random Matrices*. Cambridge University Press, Cambridge; New York.

Ashida, G., Wagner, H., and Carr, C. E. (2010). Processing of Phase-Locked Spikes and Periodic Signals. In *Analysis of Parallel Spike Trains*, Springer Series in Computational Neuroscience, pages 59–74. Springer, Boston, MA.

Belitski, A., Gretton, A., Magri, C., Murayama, Y., Montemurro, M. A., Logothetis, N. K., and Panzeri, S. (2008). Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. *J Neurosci*, 28(22):5696–709.

Brunel, N. and Wang, X.-J. (2003). What determines the frequency of fast network oscillations with irregular neural discharges? i. synaptic dynamics and excitation-inhibition balance. *Journal of Neurophysiology*, 90(1):415–430.

Buzsáki, G. (2004). Large-scale recording of neuronal ensembles. *Nature Neuroscience*, 7(5):446–451.

Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford University Press.

Buzsaki, G. (2010). Neural syntax: Cell assemblies, synapsembles, and readers. *Neuron*, 68(3):362–85.

Buzsaki, G., Anastassiou, C. A., and Koch, C. (2012). The origin of extracellular fields and currents–eeg, ecog, lfp and spikes. *Nat Rev Neurosci*, 13(6):407–20.

Buzsáki, G., Horváth, Z., Urioste, R., Hetke, J., and Wise, K. (1992). High-frequency network oscillation in the hippocampus. *Science*, 256(5059):1025–1027.

Buzsaki, G., Logothetis, N., and Singer, W. (2013). Scaling brain size, keeping timing: evolutionary preservation of brain rhythms. *Neuron*, 80(3):751–64.

Canolty, R. T., Ganguly, K., Kennerley, S. W., Cadieu, C. F., Koepsell, K., Wallis, J. D., and Carmena, J. M. (2010). Oscillatory phase coupling coordinates anatomically dispersed functional cell assemblies. *Proc Natl Acad Sci U S A*, 107:17356–61.

Capitaine, M. and Donati-Martin, C. (2016). Spectrum of deformed random matrices and free probability. *arXiv preprint arXiv:1607.05560*.

Carnevale, N. T. and Hines, M. L. (2006). *The NEURON Book*. Cambridge University Press, Cambridge, UK ; New York, illustrated edition edition.

Carr, M. F., Karlsson, M. P., and Frank, L. M. (2012). Transient slow gamma synchrony underlies hippocampal memory replay. *Neuron*, 75(4):700–713.

Chavez, M., Besserve, M., Adam, C., and Martinerie, J. (2006). Towards a proper estimation of phase synchronization from time series. *J Neurosci Methods*, 154(1-2):149–60.

Costa, M. S., Weigenand, A., Ngo, H.-V. V., Marshall, L., Born, J., Martinetz, T., and Claussen, J. C. (2016). A Thalamocortical Neural Mass Model of the EEG during NREM Sleep and Its Response to Auditory Stimulation. *PLOS Computational Biology*, 12(9):e1005022.

Csicsvari, J., Hirase, H., Mamiya, A., and Buzsáki, G. (2000). Ensemble Patterns of Hippocampal CA3-CA1 Neurons during Sharp Wave–Associated Population Events. *Neuron*, 28(2):585–594.

Datta, B. N. (2010). *Numerical Linear Algebra and Applications*.

Denker, M., Roux, S., Timme, M., Riehle, A., and Grun, S. (2007). Phase synchronization between LFP and spiking activity in motor cortex during movement preparation. *Neurocomputing*, 70(10-12):2096–2101.

Dickey, A. S., Suminski, A., Amit, Y., and Hatsopoulos, N. G. (2009). Single-Unit Stability Using Chronically Implanted Multielectrode Arrays. *J. Neurophysiol.*, 102(2):1331–1339.

Dwarakanath, A., Kapoor, V., Werner, J., Safavi, S., Fedorov, L. A., Logothetis, N. K., and Panagiotaropoulos, T. I. (2020). Prefrontal state fluctuations control access to consciousness. *bioRxiv*, page 2020.01.29.924928.

Einevoll, G. T., Destexhe, A., Diesmann, M., Grün, S., Jirsa, V., de Kamps, M., Migliore, M., Ness, T. V., Plesser, H. E., and Schürmann, F. (2019). The Scientific Case for Brain Simulations. *Neuron*, 102(4):735–744.

Einevoll, G. T., Kayser, C., Logothetis, N. K., and Panzeri, S. (2013). Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nat Rev Neurosci*, 14(11):770–85.

Elston, G. N. (2000). Pyramidal cells of the frontal lobe: All the more spinous to think with. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 20:RC95.

Elston, G. N. (2003). Cortex, cognition and the cell: new insights into the pyramidal neuron and prefrontal function. *Cerebral Cortex*, 13(11):1124–1138.

Engel, A. K., Fries, P., Konig, P., Brecht, M., and Singer, W. (1999). Temporal binding, binocular rivalry, and consciousness. *Conscious Cogn*, 8(2):128–51.

Ermentrout, B. and Pinto, D. (2007). Neurophysiology and waves.

Ermentrout, G. B. and Kleinfeld, D. (2001). Traveling electrical waves in cortex: insights from phase dynamics and speculation on a computational role. *Neuron*, 29(1):33–44.

Fernández-Ruiz, A., Oliva, A., Nagy, G. A., Maurer, A. P., Berényi, A., and Buzsáki, G. (2017). Entorhinal-CA3 Dual-Input Control of Spike Timing in the Hippocampus by Theta-Gamma Coupling. *Neuron*, 93(5):1213–1226.e5.

Fisher, N. I. (1995). *Statistical Analysis of Circular Data*. Univ. Press, Cambridge, repr., 1. paperback ed edition.

Fletcher, C. (1991). Computational techniques for fluid dynamics; vol 1.

Fried, I., Rutishauser, U., Cerf, M., and Kreiman, G., editors (2014). *Single Neuron Studies of the Human Brain: Probing Cognition*. The MIT Press, Cambridge, Massachusetts.

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends Cogn Sci*, 9(10):474–80.

Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron*, 88:220–35.

Fukushima, M., Chao, Z. C., and Fujii, N. (2015). Studying brain functions with mesoscopic measurements: Advances in electrocorticography for non-human primates. *Current Opinion in Neurobiology*, 32:124–131.

Gao, P. and Ganguli, S. (2015). On simplicity and complexity in the brave new world of large-scale neuroscience. *Current Opinion in Neurobiology*, 32:148–155.

Grasse, D. W. and Moxon, K. A. (2010). Correcting the bias of spike field coherence estimators due to a finite number of spikes. *Journal of neurophysiology*, 104(1):548–58.

Grosmark, A. D. and Buzsáki, G. (2016). Diversity in neural firing dynamics supports both rigid and learned hippocampal sequences. *Science*, 351(6280):1440–1443.

Grosmark, A. D., Mizuseki, K., Pastalkova, E., Diba, K., and Buzsáki, G. (2012). REM Sleep Reorganizes Hippocampal Excitability. *Neuron*, 75(6):1001–1007.

Grün, S. (2009). Data-Driven Significance Estimation for Precise Spike Correlation. *Journal of Neurophysiology*, 101(3):1126–1140.

Harris, K. D. (2005). Neural signatures of cell assembly organization. *Nature reviews. Neuroscience*, 6(5):399–407.

Herreras, O. (2016). Local Field Potentials: Myths and Misunderstandings. *Front Neural Circuit*, 10:101.

Jiang, H., Bahramisharif, A., van Gerven, M. A. J., and Jensen, O. (2015). Measuring directionality between neuronal oscillations of different frequencies. *NeuroImage*, 118:359–367.

Jin, J. Z., Weng, C., Yeh, C.-I., Gordon, J. A., Ruthazer, E. S., Stryker, M. P., Swadlow, H. A., and Alonso, J.-M. (2008). On and off domains of geniculate afferents in cat primary visual cortex. *Nat. Neurosci.*, 11(1):88–94.

Jirsa, V. K. and Haken, H. (1996). Field theory of electromagnetic brain activity. *Physical Review Letters*, 77(5):960–963.

Juavinett, A. L., Bekheet, G., and Churchland, A. K. (2019). Chronically implanted Neuropixels probes enable high-yield recordings in freely moving mice. *eLife*, 8:e47188.

Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., Lee, A. K., Anastassiou, C. A., Andrei, A., Aydin, C., Barbic, M., Blanche, T. J., Bonin, V., Couto, J., Dutta, B., Gratiy, S. L., Gutnisky, D. A., Hausser, M., Karsh, B., Ledochowitsch, P., Lopez, C. M., Mitelut, C., Musa, S., Okun, M., Pachitariu, M., Putzeys, J., Rich, P. D., Rossant, C., Sun, W. L., Svoboda, K., Carandini, M., Harris, K. D., Koch, C., O'Keefe, J., and Harris, T. D. (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature*, 551:232–236.

Klausberger, T., Magill, P. J., Marton, L. F., Roberts, J. D., Cobden, P. M., Buzsaki, G., and Somogyi, P. (2003). Brain-state- and cell-type-specific firing of hippocampal interneurons in vivo. *Nature*, 421(6925):844–8.

Ledoux, E. and Brunel, N. (2011). Dynamics of networks of excitatory and inhibitory neurons in response to time-dependent inputs. *Frontiers in Computational Neuroscience*, 5:25.

Lepage, K. Q., Kramer, M. A., and Eden, U. T. (2011). The dependence of spike field coherence on expected intensity. *Neural computation*, 23(9):2209–41.

Li, C. Y., Poo, M. M., and Dan, Y. (2009). Burst spiking of a single cortical neuron modifies global brain state. *Science*, 324(5927):643–6.

Li, Z., Cui, D., and Li, X. (2016). Unbiased and robust quantification of synchronization between spikes and local field potential. *Journal of neuroscience methods*, 269:33–8.

Liebe, S., Hoerzer, G. M., Logothetis, N. K., and Rainer, G. (2012). Theta coupling between V4 and prefrontal cortex predicts visual short-term memory performance. *Nat. Neurosci.*, 15(3):456–462, S1–2.

Liljenstroem, H. (2012). Mesoscopic brain dynamics. *Scholarpedia*, 7(9):4601.

Lindén, H., Tetzlaff, T., Potjans, T. C., Pettersen, K. H., Grün, S., Diesmann, M., and Einevoll, G. T. (2011). Modeling the Spatial Reach of the LFP. *Neuron*, 72(5):859–872.

Logothetis, N., Merkle, H., Augath, M., Trinath, T., and Ugurbil, K. (2002). Ultra high-resolution fmri in monkeys with implanted rf coils. *Neuron*, 35(2):227–42.

Logothetis, N. K., Guggenberger, H., Peled, S., and Pauls, J. (1999). Functional imaging of the monkey brain. *Nat Neurosci*, 2(6):555–62.

Logothetis, N. K. and Panzeri, S. (2014). Local Field Potential, Relationship to BOLD Signal. In Dieter Jaeger, R. J., editor, *Encyclopedia of Computational Neuroscience*, pages 1–11. Springer New York.

Marchenko, V. A. and Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Matematicheskii Sbornik*, 114(4):507–536.

Markowitz, D. A., Curtis, C. E., and Pesaran, B. (2015). Multiple component networks support working memory in prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*.

Maynard, E. M., Nordhausen, C. T., and Normann, R. A. (1997). The utah intracortical electrode array: a recording structure for potential brain-computer interfaces. *Electroencephalography and clinical neurophysiology*, 102(3):228–239.

Mazzoni, A., Linden, H., Cuntz, H., Lansner, A., Panzeri, S., and Einevoll, G. T. (2015). Computing the Local Field Potential (LFP) from Integrate-and-Fire Network Models. *PLoS computational biology*, 11(12):e1004584.

Moran, R. J., Kiebel, S. J., Stephan, K. E., Reilly, R., Daunizeau, J., and Friston, K. J. (2007). A neural mass model of spectral responses in electrophysiology. *NeuroImage*, 37(3):706–720.

Mukamel, R. and Fried, I. (2011). Human Intracranial Recordings and Cognitive Neuroscience. *Annu. Rev. Psychol.*, 63(1):511–537.

Nauhaus, I., Busse, L., Carandini, M., and Ringach, D. L. (2009). Stimulus contrast modulates functional connectivity in visual cortex. *Nature neuroscience*, 12(1):70–6.

Nauhaus, I., Busse, L., Ringach, D. L., and Carandini, M. (2012). Robustness of traveling waves in ongoing activity of visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 32(9):3088–94.

Niebur, E., Koch, C., and Rosin, C. (1993). An oscillation-based model for the neuronal basis of attention. *Vision Res*, 33(18):2789–802.

Nó, R. L. D. (1947). Action potential of the motoneurons of the hypoglossus nucleus. *J. Cell. Comp. Physiol.*, 29(3):207–287.

Nunez, P. L. and Srinivasan, R. (2006). *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA.

O'Leary, T., Sutton, A. C., and Marder, E. (2015). Computational models in the age of large datasets. *Current Opinion in Neurobiology*, 32:87–94.

Oliva, A., Fernandez-Ruiz, A., Buzsaki, G., and Berenyi, A. (2016). Role of Hippocampal CA2 Region in Triggering Sharp-Wave Ripples. *Neuron*, 91:1342–55.

Onslow, A. C., Jones, M. W., and Bogacz, R. (2014). A canonical circuit for generating phase-amplitude coupling. *PloS one*, 9(8):e102591.

Pesaran, B., Vinck, M., Einevoll, G. T., Sirota, A., Fries, P., Siegel, M., Truccolo, W., Schroeder, C. E., and Srinivasan, R. (2018). Investigating large-scale brain dynamics using field potential recordings: Analysis and interpretation. *Nat. Neurosci.*, page 1.

Peterson, E. J. and Voytek, B. (2018). Healthy oscillatory coordination is bounded by single-unit computation. *bioRxiv*, page 309427.

Pinotsis, D. A., Moran, R. J., and Friston, K. J. (2012). Dynamic causal modeling with neural fields. *Neuroimage*, 59(2):1261–1274.

Pinsky, P. F. and Rinzel, J. (1994). Intrinsic and network rhythmogenesis in a reduced traub model for CA3 neurons. *J Comput Neurosci*, 1(1):39–60.

Platkiewicz, J., Stark, E., and Amarasingham, A. (2017). Spike-Centered Jitter Can Mistake Temporal Structure. *Neural Computation*, 29(3):783–803.

Qubbaj, M. R. and Jirsa, V. K. (2007). Neural field dynamics with heterogeneous connection topology. *Physical review letters*, 98:238102.

Quiroga, R. (2007). Spike sorting. *Scholarpedia*, 2(12):3583.

Rabiner, L. R., Mcclellan, J. H., and Parks, T. W. (1975). Fir digital-filter design techniques using weighted chebyshev approximation. *Proceedings of the IEEE*, 63(4):595–610.

Ramirez-Villegas, J. F., Logothetis, N. K., and Besserve, M. (2015). Diversity of sharp-wave-ripple LFP signatures reveals differentiated brain-wide dynamical events. *Proceedings of the National Academy of Sciences of the United States of America*, 112:E6379–87.

Ramirez-Villegas, J. F., Willeke, K. F., Logothetis, N. K., and Besserve, M. (2018). Dissecting the synapse- and frequency-dependent network mechanisms of in vivo hippocampal sharp wave-ripples. *Neuron*, 100(5):1224–1240.e13.

Rasch, M., Logothetis, N. K., and Kreiman, G. (2009). From neurons to circuits: Linear estimation of local field potentials. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29:13785–96.

Rasch, M. J., Gretton, A., Murayama, Y., Maass, W., and Logothetis, N. K. (2008). Inferring spike trains from local field potentials. *J Neurophysiol*, 99(3):1461–76.

Ray, S. and Maunsell, J. H. (2011a). Network rhythms influence the relationship between spike-triggered local field potential and functional connectivity. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 31(35):12674–82.

Ray, S. and Maunsell, J. H. R. (2011b). Different Origins of Gamma Rhythm and High-Gamma Activity in Macaque Visual Cortex. *PLOS Biology*, 9(4):e1000610.

Rosenbaum, R., Smith, M. A., Kohn, A., Rubin, J. E., and Doiron, B. (2017). The spatial structure of correlated neuronal variability. *Nature neuroscience*, 20:107–114.

Safavi, S., Dwarakanath, A., Kapoor, V., Werner, J., Hatsopoulos, N. G., Logothetis, N. K., and Panagiotaropoulos, T. I. (2018). Nonmonotonic spatial structure of interneuronal correlations in prefrontal microcircuits. *PNAS*, page 201802356.

Safavi, S., Logothetis, N. K., and Besserve, M. (2020). From univariate to multivariate coupling between continuous signals and point processes: A mathematical framework. *ArXiv200504034 Q-Bio Stat*.

Schomburg, E. W., Anastassiou, C. A., Buzsáki, G., and Koch, C. (2012). The Spiking Component of Oscillatory Extracellular Potentials in the Rat Hippocampus. *J. Neurosci.*, 32(34):11798–11811.

Schwalm, M., Schmid, F., Wachsmuth, L., Backhaus, H., Kronfeld, A., Aedo Jury, F., Prouvot, P. H., Fois, C., Albers, F., van Alst, T., Faber, C., and Stroh, A. (2017). Cortex-wide BOLD fMRI activity reflects locally-recorded slow oscillation-associated calcium waves. In *eLife*, volume 6.

Sherfey, J., Ardid, S., Miller, E. K., Hasselmo, M. E., and Kopell, N. J. (2020). Prefrontal oscillations modulate the propagation of neuronal activity required for working memory. *Neurobiology of Learning and Memory*, 173:107228.

Sherfey, J. S., Ardid, S., Hass, J., Hasselmo, M. E., and Kopell, N. J. (2018). Flexible resonance in prefrontal networks with strong feedback inhibition. *PLOS Computational Biology*, 14(8):e1006357.

Sik, A., Penttonen, M., Ylinen, A., and Buzsaki, G. (1995). Hippocampal CA1 interneurons: An in vivo intracellular labeling study. *J. Neurosci.*, 15(10):6651–6665.

Somers, D. C., Nelson, S. B., and Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *J. Neurosci.*, 15(8):5448–5465.

Stark, E., Roux, L., Eichler, R., Senzai, Y., Royer, S., and Buzsáki, G. (2014). Pyramidal cell-interneuron interactions underlie hippocampal ripple oscillations. *Neuron*, 83(2):467–480.

Stevenson, I. H. and Kording, K. P. (2011). How advances in neural recording affect data analysis. *Nature Neuroscience*, 14(2):139–142.

Sullivan, D., Csicsvari, J., Mizuseki, K., Montgomery, S., Diba, K., and Buzsáki, G. (2011). Relationships between hippocampal sharp waves, ripples, and fast gamma oscillation: influence of dentate and entorhinal cortical activity. *Journal of Neuroscience*, 31(23):8605–8616.

Taxidis, J., Anastassiou, C. A., Diba, K., and Koch, C. (2015). Local Field Potentials Encode Place Cell Ensemble Activation during Hippocampal Sharp Wave Ripples. *Neuron*, 87(3):590–604.

Taxidis, J., Coombes, S., Mason, R., and Owen, M. R. (2012). Modeling sharp wave-ripple complexes through a CA3-CA1 network model with chemical synapses. *Hippocampus*, 22(5):995–1017.

Teleńczuk, B., Dehghani, N., Le Van Quyen, M., Cash, S. S., Halgren, E., Hatsopoulos, N. G., and Destexhe, A. (2017). Local field potentials primarily reflect inhibitory neuron activity in human and monkey cortex. *Sci. Rep.*, 7:40211.

Traub, R. D. and Miles, R. (1995). Pyramidal cell-to-inhibitory cell spike transduction explicable by active dendritic conductances in inhibitory cell. *Journal of computational neuroscience*, 2(4):291–298.

Traub, R. D., Whittington, M. A., Buhl, E. H., Jefferys, J. G., and Faulkner, H. J. (1999). On the mechanism of the gamma –> beta frequency shift in neuronal oscillations induced in rat hippocampal slices by tetanic stimulation. *The Journal of Neuroscience*, 19(3):1088–1105.

Varga, C., Oijala, M., Lish, J., Szabo, G. G., Bezaire, M., Marchionni, I., Golshani, P., and Soltesz, I. (2014). Functional fission of parvalbumin interneuron classes during fast network events. *Elife*, 3.

Vinck, M., Battaglia, F. P., Womelsdorf, T., and Pennartz, C. (2012). Improved measures of phase-coupling between spikes and the Local Field Potential. *J Comput Neurosci*, 33(1):53–75.

Vinck, M., van Wingerden, M., Womelsdorf, T., Fries, P., and Pennartz, C. M. (2010). The pairwise phase consistency: A bias-free measure of rhythmic neuronal synchronization. *Neuroimage*, 51(1):112–22.

Williamson, R. C., Doiron, B., Smith, M. A., and Yu, B. M. (2019). Bridging large-scale neuronal recordings and large-scale network models using dimensionality reduction. *Current Opinion in Neurobiology*, 55:40–47.

Wilson, H. R. and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13(2):55–80.

Wójcik, D. K. (2013). Current Source Density (CSD) Analysis. In Jaeger, D. and Jung, R., editors, *Encyclopedia of Computational Neuroscience*, pages 1–10. Springer, New York, NY.

Womelsdorf, T., Schoffelen, J. M., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., and Fries, P. (2007). Modulation of neuronal interactions through neuronal synchronization. *Science*, 316:1609–12.

Zanos, T. P., Mineault, P. J., and Pack, C. C. (2011). Removal of spurious correlations between spikes and local field potentials. *Journal of neurophysiology*, 105(1):474–86.

Zarei, M., Jahed, M., and Daliri, M. R. (2018). Introducing a comprehensive framework to measure spike-lfp coupling. *Frontiers in Computational Neuroscience*, 12.

Zeitler, M., Fries, P., and Gielen, S. (2006). Assessing neuronal coherence with single-unit, multi-unit, and local field potentials. *Neural computation*, 18:2256–81.

Zerbi, V., Floriou-Servou, A., Markicevic, M., Vermeiren, Y., Sturman, O., Privitera, M., von Ziegler, L., Ferrari, K. D., Weber, B., De Deyn, P. P., Wenderoth, N., and Bohacek, J. (2019). Rapid Reconfiguration of the Functional Connectome after Chemogenetic Locus Coeruleus Activation. *Neuron*, 103(4):702–718.e5.

STAR Methods

Contact for reagent and resource sharing

Further information and requests for reagents and resources may be directed to and will be fulfilled by the Lead Contact, Dr. Michel Besserve (michel.besserve@tuebingen.mpg.de).

Experimental model and subject details

The neural data used in this study were recorded from the ventrolateral prefrontal cortex (vlPFC) of one anaesthetised adult, male rhesus monkey (macaca mulatta) by using Utah microelectrode arrays [Blackrock Microsystems (Maynard et al., 1997)] (more details on these experiments are provided in a previous study exploiting this data by Safavi et al. (2018)). All experiments were approved by the local authorities (Regierungspräsidium) and were in full compliance with the guidelines of the European Community (EUVD 86/609/EEC) for the care and use of laboratory animals.

Method details

Detailed GPLA methodology for electrophysiology data

GPLA proceeds in several steps: preprocessing of multi-channel LFP signals, construction of the coupling matrix, and its low-rank approximation. Finally, parameters of this low-rank approximation are standardized following specific normalization conventions allowing their easy interpretation and comparison. These steps are described in the following subsections.

LFP pre-processing

Prior to computing couplings, the LFP signal is pre-processed, first by filtering in the frequency band of interest. The choice of the filter bandwidth for the purpose of extracting the instantaneous phase or analytic signal in a particular band is subjected to a trade-off. On one hand, the signal requires a narrow enough band-pass filtering to provide us a proper estimate of the phases (Chavez et al., 2006). On the other hand, the filtered signal should preserve the temporal dynamics in the frequency of interest. The second step is extracting the analytical signal using the Hilbert transform, resulting in a complex-valued signal containing both the amplitude and phase of LFP. In the optional third step (see section Necessity of whitening and post-processing), we whiten the LFPs. We need to decorrelate LFP signal recorded in the different channels by applying a whitening operator. It is necessary to be able to use tools from Random Matrix Theory (Anderson et al., 2010) (the rationale for the inclusion of the whitening step is elaborated in section Analytical test and Safavi et al. (2020)).

We consider LFPs and spiking units are recorded repeatedly over K trials, and each trial has length T (number of time-points). We represent LFPs of trial k by $L^{(k)}$, which is a $(n_c \times T)$ matrix, where n_c is the number of LFP recording channels. To simplify the notations, by $L^{(k)}$ we refer to analytical signals, i. e. band-passed in a particular frequency range and Hilbert transformed signals. We denote the collection of $N_m^{(k)}$ spike times of unit m at trial k by $\{t_j^{m,(k)}\}_{j=1\ldots N_m^{(k)}}$ ($\{t_j^{m,(k)}\}$ contains the time-point indices of the LFP data for which spikes occur).

We introduce a *reduced-ranked* whitening operator which is a modified version of the conventional whitening that decorrelate the data, in this case, LFP signals. We customized this procedure in order to accommodate GPLA's needs, i. e. (1) avoid over-amplification of noise components of LFP (which are reflected in smaller eigenvalues of LFP covariance matrix) in the whitening operator, and (2) eliminate factors of variability that are not consistent across trials.

In our *reduced-ranked* whitening, we first reduce the rank of the LFP covariance matrix, by truncating the eigenvalue decomposition of LFP covariance matrix. We choose the number of components such that 99% of variance is explained with the reduced rank covariance matrix. In order to find the number of components that 99% of variance of LFP covariance matrix, we

concatenate LFPs of all trials into a larger $n_c \times KT$ matrix, denoted by L and compute the eigenvalue decomposition of the covariance matrix,

$$\mathrm{Cov}\,(L) = \frac{1}{T} L L^H \,, \tag{8}$$

where $.^H$ indicates the transpose complex conjugate (should be noted that, analytical signal L, is a complex-valued matrix). We denote the number of components needed to explain 99% of variance of LFP covariance matrix by n_c^{eff}. We find the reduced number of components, n_c^{eff}, based on all trials, and we use n_c^{eff} to define the whitening operator of individual trials. The reduced rank single-trial LFP covariance matrix is denoted by $\mathrm{Cov}^{\mathrm{red}}\left(\mathrm{L}^{(\mathrm{k})}\right)$, and computed as follows

$$\mathrm{Cov}^{red}\left(L^{(k)}\right) = \sum_{p=1}^{n_c^{eff}} \lambda_p^{(k)} x_p^{(k)} (x_p^{(k)})^H \,, \tag{9}$$

where $\lambda_k^{(k)}$ and $x_k^{(k)}$ respectively denote the eigenvalue and eigenvectors of the LFP covariance matrix of trial k. We denote the whitened LFP of trial k by $L_w^{(k)}$, and compute it as follows,

$$L_w^{(k)} = (\Lambda^{(k)})^{\frac{-1}{2}} (X^{(k)})^H L^{(k)} \,, \tag{10}$$

where $\Lambda^{(k)}$ is a $n_c^{eff} \times n_c^{eff}$ diagonal matrix containing the eigenvalues of the above single-trial reduced rank LFP covariance matrix, and $X^{(k)}$ is a is a $n_c \times n_c^{eff}$ matrix containing the eigenvectors $x_k^{(k)}$.

Coupling matrix

Given the spike times of a single spike train $\{t_j^{(k)}\}_{j=1\ldots N^{(k)}}$ and $L_w^{(k)}$ a single channel pre-processed LFP analytic signal (as explained in section LFP pre-processing) and its phase ϕ ($= \angle L$), the conventional measure of spike-LFP coupling, Phase Locking Value (PLV), defined as follows:

$$PLV = \frac{1}{N^{tot}} \sum_{k=1}^{K} \sum_{j=1}^{N^{(k)}} \exp\left(\boldsymbol{i}\phi_{t_j^{(k)}}^{(k)}\right) \,, \tag{11}$$

where, $\boldsymbol{i}$ is the imaginary unit ($\boldsymbol{i}^2 = -1$), and $N^{(k)}$ is the number of spikes occurring during the trial k, N_{tot} is the total number of spikes occurred across all trials, i. e.

$$N^{tot} = \sum_{k=1}^{K} N^{(k)} \,. \tag{12}$$

In addition to PLV, we introduce a similar coupling statistics, denoted by c,

$$c = \frac{1}{\sqrt{N^{tot}}} \sum_{k=1}^{K} \sum_{j=1}^{N^{(k)}} L_{t_j^{(k)}}^{(k)} \,, \tag{13}$$

to be used when the theoretical significance test is intended to be used (see section Analytical test). The coupling statistics c is different from PLV in two ways, First, in PLV only the phase information from the continuous signal is used, while for c, we use both the phase and amplitude of the LFP signal. This is motivated by evidence that inclusion of the amplitude can improve the coupling measure (Denker et al., 2007) by weighting the contribution of spikes in the coupling measure by the LFP amplitude at the correspond spike time, as well as by theoretical considerations (see STAR Methods section Analytical test for more details). The second difference is, for c we have normalization by square root of the number of spikes rather the number of spikes (division by $\sqrt{N^{tot}}$ in Equation 13 versus N^{tot} in Equation 11). Basically, a scaling by $\sqrt{N^{tot}}$ is needed to normalize the variance of entires of the coupling matrix to 1, in order to be able to use tools from Random Matrix Theory (Anderson et al., 2010) (see Safavi et al. (2020) for more details).

A multivariate generalization of the coupling statistics, could be achieved by collecting the coupling statistics between all spiking units and LFP signals. Given spike times $\{t_j^{m,(k)}\}_{j=1\ldots N_m^{(k)}}$, $\phi_w^{(k)}$ LFP

phase, and $L_w^{(k)}$ the analytical LFP, we can define the coupling matrix $\boldsymbol{C}$, based on PLV (Equation 11) as follows,

$$(\boldsymbol{C})_{n,m} = \frac{1}{N_m^{tot}} \sum_{k=1}^{K} \sum_{j=1}^{N_m^{(k)}} \exp\left(\boldsymbol{i}(\phi^{(k)})_{n,t_j^{(k)}}\right) , \tag{14}$$

or based on c (Equation 13),

$$(\boldsymbol{C})_{n,m} = \frac{1}{\sqrt{N_m^{tot}}} \sum_{k=1}^{K} \sum_{j=1}^{N_m^{(k)}} (L^{(k)})_{n,t_j^{(k)}} , \tag{15}$$

where m, j and n respectively indicate the index of spiking unit, index of spike time and index of LFP channel and N_m refers to number of spikes recorded in spiking unit m [1].

Let n_c and n_s be the number of LFP channels and number of spiking units, respectively, $\boldsymbol{C}$ is thus a $n_c \times n_s$ complex-valued matrix (or $n_c^{eff} \times n_s$ if whitening is applied). As n_c (or n_c^{eff}) and n_s are not necessarily equal in electrophysiological datasets, the coupling matrix is not square in general.

Our coupling matrix is thus designed as a multivariate generalization of univariate coupling measures in order to capture the overall synchronization between the spiking activity and the phase of a global oscillatory dynamics in a given frequency band.

Low rank decomposition

Each column of the coupling matrix $\boldsymbol{C}$ has a common spiking unit whose locking is computed with respect to different LFP channels (called LFP vectors). Conversely, each row collects the phase locking values of all spiking channels to a common LFP reference channel. In order to achieve a compact and interpretable representation of this high dimensional object, we compute the Singular Value Decomposition (SVD) of the coupling matrix of the form

$$\boldsymbol{C} \quad = U D V^H \quad = \sum_{k=1}^{p} d_k u_k v_k^H , \tag{16}$$

where (d_k) is a tuple of positive scalars, the singular values (SV), listed in decreasing order. The complex valued vectors u_k and v_k are, respectively, the n_c/n_c^{eff}- and n_s-dimensional singular vectors associated to a given SV d_k. One important property of SVD is that keeping only the first term in Equation (16), with SV d_1, achieves the best rank-one approximation of the matrix, $\boldsymbol{C} \approx d_1 u_1 v_1^H$, in the least square sense (Datta, 2010, Theorem 7.29).

Post-processing

In order to make the outputs of GPLA interpretable, we introduce a few post-processing steps. An unwhitening and rescaling procedure is introduced to reverse some normalization discussed in previous sections LFP pre-processing, Coupling matrix, and Low rank decomposition, and a rotational transformation is introduced in order to represent the singular vectors in a more interpretable fashion.

Representation of singular vectors: Following the conventional mathematical representation of SVD in Equation 3, U and V are unitary matrices i. e. $U^H U = I$ and $V^H V = I$.[2] This implies that all singular vectors are unit norm, and all the information regarding the strength of coupling is absorbed in the singular values on the diagonal matrix D. As explained in main text (see sections Low rank linear response theory and frequency analysis and Generalizing spike-oscillation coupling analysis to the multivariate setting), the relative magnitude and phase of singular vectors coefficients can be used to interpret the *relative* contribution of individual LFP channel and individual spiking unit to the coordinated pattern captured by the largest singular value.

We can summarize the coupling matrix with three quantities:

$$\boldsymbol{C} \sim (gPLV).v_{LFP} v_{spike})^H . \tag{17}$$

[1] Reader can also refer to Safavi et al. (2020, Section 4) for a different formulation of computing the coupling matrix.

[2] For spike vector, V, should be noted that, V is a unitary after appropriate normalization that was discussed in the previous paragraph

However the coefficient of both singular vectors can be rotated of the same arbitrary angle in the complex plane, as the rotation transformation in the complex plane does not change the SVD factorization, i. e.

$$udv^H = udv^H e^{-i\theta_0} e^{i\theta_0} = e^{-i\theta_0} ud \left(e^{-i\theta_0} v\right)^H . \tag{18}$$

We exploit this free parameter to make the GPLA more neuroscientifically interpretable by rotating both spike and LFP vectors with $\overline{-\phi_{LFP}}$, where $\overline{\phi_{LFP}}$ and $\overline{\phi_{spike}}$ are the average spike and LFP phases, defined as,

$$\overline{\phi_{LFP}} = \angle \sum_{i=1}^{n_c} (v_{LFP})_i \,, \tag{19}$$

$$\overline{\phi_{spike}} = \angle \sum_{i=1}^{n_u} (v_{spike})_i \,. \tag{20}$$

The rationale behind it is to center the coefficient of the rotated LFP vector ($\widetilde{v_{LFP}} = v_{LFP} e^{-i\overline{\phi_{LFP}}}$) around zero phase in the complex plane and the rotated spike vector, $\widetilde{v_{spike}} = v_{spike} e^{-i\overline{\phi_{LFP}}}$, preserves the angular difference of Φ_d of the spikes with respect to the LFP, defined as

$$\Phi_d = \overline{\phi_{LFP}} - \overline{\phi_{spike}} \,. \tag{21}$$

With this chosen convention, we obtain the final GPLA factorization

$$C \sim (gPLV).\widetilde{v_{LFP}} \widetilde{v_{spike}}^H \,. \tag{22}$$

We can also apply the phase difference between average LFP and spike vectors (Φ_d) to gPLV as it can summarize the overall phase shift between LFP and spikes. Given that gPLV is always a real positive value, by this convention, we add an extra information to gPLV.

We thus define a *complex gPLV* ($\widetilde{gPLV} = gPLV\, e^{-i\Phi_d}$) which its magnitude indicate the coupling strength between spikes and LFPs as in phase locking value (PLV) and its angle indicates the overall phase difference between spiking activity and LFP which is similar to locking phase in classical univariate phase locking analysis. This is an arbitrary choice to some degree, neverthelss it allows to interpret the GPLA output very similar to classical univariate phase locking analysis.

Needless to mention, when the magnitude of gPLV is small, this overall phase difference is not meaningful (similar to the case where PLV is small, the locking phase is not meaningful).

Unwhitening: As discussed in section LFP pre-processing, due to theoretical considerations, and in particular for applicability of our analytical significance test (see Significance assessment of gPLV), we whiten the LFPs prior to any other processing. In order to retrieve the original structure of the LFP i. e. retrieve all the correlations that were present in the original LFP signals but was diminished by the whitening, we need to revert the whitening i. e. unwhitening the LFP vector resulting from GPLA. This can be achieved by computing the unwhiten operator and apply it to the LFP vector,

$$v_{LFP}^{unwhiten} = W^{-1} v_{LFP} \,, \tag{23}$$

where W^{-1} is the unwhitening operator. In order to find this operator, we first we concatenate whitened LFPs of all trials (resulting from Equation 10) into a larger matrix L_w ($n_c^{eff} \times KT$). Then we estimate W^{-1} by using a linear regression with unwhitned and whitened LFPs (W^{-1} is the $n_c \times n_c^{eff}$ matrix of coefficient for regression).

Rescaling: As introduced in Equation 13, coefficient of coupling matrix are normalized by the square root of number of spikes. This choice of normalization, is different from the one use in conventional PLV (Equation 11). This will lead to inhomogeneous weighting of spiking units according to their variability of their firing rate. To avoid this potentially misleading weighting, we divide the spike vector by the square root of number of spikes,

$$v_{Spike}^{rescaled} = v_{Spike} \oslash \vec{N} \,, \tag{24}$$

where $\oslash$ is (entrywise) Hadamard division and $\vec{N} = \{N_m^{tot}\}_{m=1,\ldots,n_s}$, which is a vector consist of total spike counts (similar to Equation 12) of all the neurons (indexted by m) used in GPLA. Furthermore,

to preserve the original norm of the spike vector (unit magnitude), we also need to normalize the spike vector by its norm,

$$v_{Spike}^{final} = \frac{v_{Spike} \oslash \vec{N}}{\left\| v_{Spike} \oslash \vec{N} \right\|} . \tag{25}$$

Necessity of whitening and post-processing

The whitening (and the subsequent post-processing) is necessary to have the advantage of applicability of the analytical significance test. LFPs are typically very correlated signals, and such correlation will be reflected in the magnitude of the singular values (and consequently gPLV), therefore, by whitening such correlations will be removed prior to investigating spike-LFP coupling. Nevertheless, multivariate analysis of spike-LFP coupling with GPLA even without whitening can still be beneficial for investigating spike-LFP relation sip (see Figure 3). If statistical testing based on surrogate data is intended, it is possible to skip the whitening step and proceed directly with the constructing the coupling matrix and low rank estimation. In that case, entries of the coupling can be filled by conventional PLVs (see Equation 14), or other choices of spike-LFP coupling measures (Grasse and Moxon, 2010; Vinck et al., 2010; Lepage et al., 2011; Vinck et al., 2012; Jiang et al., 2015; Zarei et al., 2018; Li et al., 2016) (also see the section Limitations and potential extensions for further elaboration). In this case, whitening of the LFP can be skipped and subsequent "Unwhitening and rescaling" discussed in section Post-processing is not necessary anymore.

Optional normalization for gPLV

As gPLV is a singular value of a matrix, it grows with the dimensions of the coupling matrix. This makes the comparison of gPLV resulting from different datasets difficult. For instance, assume the hypothetical situation of having two datasets recorded from two homogeneous populations of neurons, if the strength of coupling is the same in two populations, the populations with a larger amount of recorded neurons (therefore larger dimension of the coupling matrix) will have larger gPLV. Certainly, this can be misleading for investigating the spike-LFP coupling with GPLA when datasets with variable number of spiking units and/or LFP channels. To overcome this issue we suggest normalizing the gPLV to become independent of the size of the neural population (dimension of the coupling matrix) and the number of channels. When we consider the the entries of coupling matrix, C, to be PLV (LFPs are not whitened and Equation 14 is used for constructing the coupling matrix), pairwise coupling static is bounded ($|PLV| \leq 1$). When all the entities of the coupling matrix C attain their maximum value, gPLV will also gain the maximum possible value. Therefore, we can exploit it to normalize the gPLV. For a coupling matrix having maximum coupling for all pairs ($(C)_{n,m} = 1$ and C, a $n_c \times n_s$ matrix), then $gPLV_{max} = \sqrt{n_c n_s}$. Therefore, if we normalize the original gPLV by the maximum value it can achieve ($gPLV_{max} = \sqrt{n_c n_s}$, calculated is based on the dimensionality of matrix C), then the gPLV will be bounded by 1 as well. Moreover, with this normalization, gPLV is also comparable to PLV (if we have a homogeneous population of neurons, otherwise these quantities are not comparable).

Significance assessment of gPLV

In order to statistically assess the significance of coupling between spikes and LFP based on gPLV, we develop a surrogate- and a Random Matrix Theory (RMT)-based statistical testing framework exposed in Safavi et al. (2020). Hypothesis testing based on the generation of surrogate data is a common method for significant assessment in neuroscience. Nevertheless, not only generating appropriate surrogate data can be challenging (for a review see Grün (2009)), but also computationally expensive. This motivates the development of an analytical test exploiting minimal computational resources.

Surrogate-based test

In contrast to uni-variate methods for which the distribution under a null hypothesis is more likely to be (possibly approximately) derived based on theoretical analysis (e.g. Rayleigh test for PLV (Fisher, 1995, Chapter 4)), such approaches are usually unavailable in multi-variate settings (nevertheless, we have developed one for gPLV, see section Analytical test). Following a common alternative approach, we build the null distribution by generating many surrogate datasets Grün (2009). The resulting gPLVs values forms an empirical H_0 distribution that can be used to compute the p-value for statistical assessment of the significance gPLV in the data. Importantly, the choice of the appropriate surrogate data according to characteristics of neural data is critical. For instance, generating surrogate data by shuffling inter-spike-intervals (ISI) is not an appropriate method when we have non-stationarity in firing rates (Grün, 2009).

In this work, we used an *interval*-jittering rather than a *spike-centered*-jittering[3], as the former was reported to be more reliable for detecting temporal structures in spike data (Platkiewicz et al., 2017). We devised the two following spike-jittering-based methods for GPLA. We also verify the appropriateness of our jittering approaches with various simulations (see the Results).

Simple interval jitter: Each surrogate dataset is generated by jittering all the spikes (from all neurons) with a particular jittering window (or dither width). In the interval jittering, per each spike, a new spike time is drawn within the jittering window around the spike. The timing of jittered spikes should be drawn from a uniform distribution. The size of the jittering window can be specified by the frequency wherein the spike-LFP coupling is being investigated. The smallest jittering window (or dither width) that can be used in order to destroy the temporal structure potentially exists in the range of frequency-of-interest. In the phase-locking analysis of electrophysiological data we usually extract the analytic signal or instantaneous phase of LFP by applying Hilbert transform on band-limited LFP signals (Chavez et al., 2006). The central frequency of the band-limited filter can be used for specifying the jittering window (or dither width), i. e. jittering window is the inverse of this central frequency.

Group preserved jitter: Similar to "simple interval jitter" we generate each surrogate dataset by relocating all the spikes within a window. For each surrogate data, we first divide the spike trains into equally-sized windows. Then we circularly shift the spike sequence within each window for all neurons together using a uniformly distributed time shift. Notably, we use a single random value for circular shifting of all neuron's spiking within the window. This size of this window should be chosen similar to the previous method ("simple interval jitter") i. e. based on the central frequency of the band-limited filter. The rationale behind this method of generation surrogate data is *relative* timing of the spikes could be associated to a large degree to the ansamble activity irrespective of the coupling to the LFP. Therefore, the relative timing of the spikes might not be impaired in the absence of coupling to global dynamics of the LFP. With "group preserved jittering" the relative timing is preserved and the coupling to the LFP is destroyed.

Analytical test

Challenges in generation of surrogate data (Grün, 2009) and considerable increase in the dimensionality of datasets (Pesaran et al., 2018; Jun et al., 2017; Buzsáki, 2004; Fukushima et al., 2015), suggest that deriving mathematically (asymptotic) properties of GPLA under the null hypotheses, as is done for univariate testing (e. g. Rayleigh test for PLV (Fisher, 1995, Chapter 4)) is an interesting alternative.

In a companion work (Safavi et al., 2020), by using martingale theory (Aalen et al., 2008) we derive an asymptotic distribution for the entries of the coupling matrix in fairly general settings. Furthermore, by exploiting RMT (Anderson et al., 2010) we can find a good approximation of the distribution of eigenvalues (or singular values) of the coupling matrix in absence of coupling between spikes and LFPs. This provides a null hypothesis for the statistical testing of the largest eigenvalues (or singular values) of the coupling matrix, which corresponds to gPLV in our setting.

[3] Interval- and spike-centered-jittering are also known as hard and soft dithering respectively.

Table 1

Figure num.	Osc. type	Num. of oscillatory component	Equations
Figure 2	Transient	1	14,31
Figure 3 A-C	Transient	1	14,31
Figure 3 D-I	Sustained	1	31,32
Figure 4 A-G	Sustained	5	15,31
Figure 4 H	Sustained	1-10	15, 31

As mathematical details are described in Safavi et al. (2020), we restrict ourselves to a brief explanation. When the LFP signal is whitened, and under a null hypothesis reflecting an *absence of coupling*, the coupling matrix which is constructed based on Eq 13, asymptotically converges to a matrix with i.i.d. complex standard normal coefficients (Safavi et al., 2020, Theorem 2), and the Marchenko-Pastur (MP) law then provides an approximation of the distribution of its squared singular values (Safavi et al., 2020, Theorem 3).

This law (Marchenko and Pastur, 1967) has the density

$$\frac{d\mu_{MP}}{dx}(x) = \begin{cases} \frac{1}{2\pi\alpha x}\sqrt{(b-x)(x-a)}\,, & a \le x \le b, \\ 0\,, & \text{otherwise}\,, \end{cases} \tag{26}$$

with $a = (1-\sqrt{\alpha})^2$ and $b = (1+\sqrt{\alpha})^2$ which are the upper and low bounds of the support of the distribution. Based on on the these bounds we can define a significance threshold, θ_{DET}, for the largest eigenvalue of hermitian matrix, $\boldsymbol{S} = \frac{K}{n_u}\boldsymbol{C}\boldsymbol{C}^H$:

$$\theta_{DET} = (1+\sqrt{\alpha})^2\,. \tag{27}$$

The null hypothesis can be rejected if, the largest eigenvalue of $\boldsymbol{S}$ (denoted by ℓ_1) is superior to the significance threshold:

$$\ell_1\left(\boldsymbol{S}_n\right) > \theta_{DET}\,. \tag{28}$$

Therefore, there is a significant coupling between the multi-channel spikes and LFPs, if

$$gPLV > \sqrt{n_u\theta_{DET}}\,. \tag{29}$$

As mentioned above to be able to use the result of Safavi et al. (2020, Theorem 3), we need to whiten the LFP signal first, as described in LFP pre-processing. Furthermore, satisfying this theorem requires the coupling matrix to be normalized appropriately based on the spike rate of each unit (as defined in Equation 15).

For computing α on neural data, the *reduced ranked* $n_c^{eff} < n_c$ entailed by the whitening procedure (see LFP pre-processing more details), the *effective* dimensionality of the coupling matrix changes from $n_c \times n_u$ to $n_c^{eff} \times n_u$ (which depends on the spectral content of the LFP). This leads to a modification of Equation 27 as follows:

$$\theta_{DET} = (1+\sqrt{\alpha_{eff}})^2\,, \tag{30}$$

where $\alpha_{eff} = n_c^{eff}/n_u$.

Simulation of phase-locked spike trains

We use simulated phase-locked spike trains and noisy oscillations as a toy model to demonstrate the potential applications of GPLA. The core principles of simulations used in both Figure 2, 4 and 3 are explained in the following paragraphs and the specializations used for individual figures are provided at the end (also summerized in table 1).

For generating phase-locked spike trains, we adopt the method introduced in (Ashida et al., 2010). As the model has already been described elsewhere we restrict ourselves to a brief explanation. We sample the spike times from an inhomogeneous Poisson process with rate $\lambda(t)$,

$$\lambda(t) = \lambda_0 \exp\left(\kappa \cos(2\pi f t - \varphi_0)\right) , \tag{31}$$

where φ_0 is the locking phase of the spikes with respect to the oscillation, κ is the concentration parameter of the spikes around the locking phase, that specify the strength of coupling between spikes and the oscillation, λ_0 is propositional to the average firing rate over time ($\lambda_0 I_0(\kappa)$ is the average firing rate), and f is the frequency of oscillatory modulation of the spike trains.

Furthermore, we can also derive an analytical expression for the complex-valued PLV to be used as ground truth PLV (used in Figure 3),

$$PLV^* = e^{\mathbf{i}\varphi_0} \frac{\int_0^\pi \cos(\theta)\exp(\kappa\cos(\theta))d\theta}{\int_0^\pi \exp(\kappa\cos(\theta))d\theta} = e^{\mathbf{i}\varphi_0} \frac{I_1(\kappa)}{I_0(\kappa)} , \tag{32}$$

where PLV^* indicate the ground truth value, and the I_k's denoting the modified Bessel functions of the first kind for k integer (see e. g. Abramowitz et al. (1972, p. 376)):

$$I_k(\kappa) = \frac{1}{\pi}\int_0^\pi \cos(k\theta)\exp(\kappa\cos(\theta))d\theta . \tag{33}$$

For the simulation used in Figure 4, we construct the LFP by superimposing $N_{osc} \leq 10$ oscillatory components $O_j(t) = e^{2\pi \mathbf{i} f_j t}$, $j \in \{1, \dots, N_{osc}\}$ that the frequency of oscillations are limited in range of $[f_{min}, f_{max}]$. Each LFP signal is a weighted sum of these oscillatory components. We can represent these weights in a $(n_c \times N_{osc})$-variate matrix (we call it *mixing matrix*, and denot it by W), where each row of mixing matrix indicate the weights for the corresponding LFP channel. Thus, the synthetized multichannel LFP ($\mathbf{\Psi}(t) = \{\psi_l(t)\}_{l=1,\dots,n_c}$) can be written as the product of the mixing matrix (W) and the oscillatory basis ($\boldsymbol{O}(t) = \{O_j(t)\}_{j=1,\dots,N_{osc}}$),

$$\mathbf{\Psi}(t) = W\boldsymbol{O}(t) + \boldsymbol{\eta}(t) , \tag{34}$$

where $\boldsymbol{\eta}(t)$ is additive white noise on both real and imaginary parts.

In this simulation, the frequency of the oscillatory components range from 11Hz to 15 Hz, and the mixing is the following,

$$W = \begin{bmatrix} \boldsymbol{w}_d & \boldsymbol{w}_0 & \cdots & \cdots & \boldsymbol{w}_0 \\ \boldsymbol{w}_0 & \boldsymbol{w}_d & \boldsymbol{w}_0 & \cdots & \boldsymbol{w}_0 \\ \boldsymbol{w}_0 & \boldsymbol{w}_0 & \ddots & \ddots & \boldsymbol{w}_0 \\ \boldsymbol{w}_0 & \ddots & \ddots & \ddots & \boldsymbol{w}_0 \\ \boldsymbol{w}_0 & \cdots & \cdots & \boldsymbol{w}_0 & \boldsymbol{w}_d \end{bmatrix} . \tag{35}$$

where $\boldsymbol{w}_d = w_d \mathbf{1}_{N_g}$ and $\boldsymbol{w}_0 = w_d \mathbf{0}_{N_g}$, $w_d = 1$ and $w_0 = 0.1$ ($\mathbf{1}_{N_g}$, is a $(N_g \times 1)$-variate all-one column vector). This simple structure of the mixing matrix (which is close to a block diagonal matrix) implies that each LFP channel contains one dominant oscillatory component with a specific frequency (as in each row, there is only one oscillatory component with large coefficient w_d and a specific frequency).

For the simulation used in Figure 3A-C, oscillations originate from a single oscillatory source, but in order to make transient rather sustained oscillations, they were multiplied by a Gaussian window (with the size of 20 cycles of oscillation) around random events. The timing of transitory events was governed by a homogeneous Poisson process. Moreover, the spiking activities are phase-locked to the phase of the oscillations as the spike times were drawn from an inhomogeneous Poisson process with the rate specified in Equation 31. For the rest of the simulations in Figure 3 a single sustained oscillation has been used.

Simulation of hippocampal sharp wave-ripples

The model was introduced and described in Ramirez-Villegas et al. (2018). We thus restrict ourselves to a brief explanation of the characteristics that are the most relevant to GPLA analysis.

Network architecture

We use a model of part of the hippocampal formation, accounting for the dynamics of CA1 and CA3 subfields during non-Rapid Eye Movement (non-REM) sleep. Cells of each subfield consist of 150 units (135 pyramidal neurons and 15 interneurons), arranged on a one dimensional array, along the x-axis. The connectivity of CA3 is characterized by strong recurrent excitatory auto-associational pyramidal-pyramidal connections, together with pyramidal-interneuron connections and short-range and interneuron-pyramidal synapses. In contrast, CA1 connectivity is implemented as a "feedback and reciprocal inhibition" circuit, including only pyramidal-interneuron, interneuron-pyramidal, and interneuron-interneuron synapses, all located in their peri-somatic region (see Figure 6A for the schema of the model)

Cell dynamics

Each neuron is modeled with two compartments: dendritic and axosomatic, the dynamics of each follows a Hodgkin-Huxley type (conductance-based) equation (Pinsky and Rinzel, 1994; Traub and Miles, 1995). Notably, they include a non-linear slow dendritic calcium channel responsible for the bursting activity.

Computation of the laminar LFP profile

The procedure for computing laminar LFP profiles was also described in Ramirez-Villegas et al. (2018). Briefly, the trans-membrane current of each compartment of each cell is modeled as a line source (Schomburg et al., 2012) that were placed with the equal distance across the horizontally in a Stratum Pyramidale (SP) of 100μm thickness, with an axosomatic compartments height of 80μm for both pyramidal neurons and interneurons (Traub and Miles, 1995). Total dendritic arbor height of pyramidal cells was 200μm, corresponding to the CA1 Stratum Radiatum (SR). LFPs were captured through two multi-channel electrodes (mimicking laminar probes), each with 16 recording sites disposed along the vertical axis (denoted by z), 20 μm apart covering the simulated axosomatic and apical dendritic fields of CA1 and CA3. Each electrode crosses the corresponding linear cell arrangement (perpendicularly) in its middle.

The extracellular medium is modeled as a uniform and isotropic ohmic conductor with resistivity $\rho = 333\Omega\mathrm{cm}$. The potential in the extracellular medium is governed by the Poisson equation $\nabla^2\phi = \frac{1}{\sigma}\frac{d\xi}{dt} = -\frac{I_t}{\sigma}$, where $\sigma = \frac{1}{\rho}$ is the conductivity of the extracellular space $\left[\frac{\mathrm{S}}{\mathrm{m}}\right]$. With these assumptions, the extracellular potential $\phi\left(y_0, r, t\right)$ at the algebraic depth z_0 and a radial distance r, measured over the compartment's length limits (z_1 and z_2, respectively indicate the algebraic depth and the bottom of the top of the cylindrical compartment with length $L = z_2 - z_1$) can be computed by

$$\phi\left(z_0, z_1, z_2, r, t\right) = \frac{1}{4\pi\sigma L}\int_{z_1}^{z_2}\frac{I\left(t\right)}{\sqrt{\left(z - z_0\right)^2 + r^2}}\mathrm{dz} = \frac{1}{4\pi\sigma}\frac{I\left(t\right)}{L}\ln\left[\frac{\sqrt{\left(z_1 - z_0\right)^2 + r^2} - \left(z_1 - z_0\right)}{\sqrt{\left(z_2 - z_0\right)^2 + r^2} - \left(z_2 - z_0\right)}\right]. \quad (36)$$

after solving the integral with standard procedures. Accounting for the contribution of all compartments and cells, the total extracellular potential $\phi_{\mathrm{tot}}\left(z_0, t\right)$ at a given depth z_0 is

$$\phi_{\mathrm{tot}}\left(z_0, t\right) = \sum_i\sum_j \phi_{i,j}\left(z_0, z_{1,i,j}, z_{2,i,j}, r_{i,j}, t\right)\ , \quad (37)$$

where $\phi_{i,j}$ is the potential generated by the total transmembrane current of the jth compartment of the ith cell, located at radial distance $r_{i,j}$ from the electrode.

Note that since the neuron models considered in this work are two-compartmental, charge conservation within the cell implies that the total absolute somatic transmembrane currents equal the absolute of the total dendritic transmembrane currents (which also follows the charge conservation principle), leading to a dipolar distribution of the LFP contribution for each cell.

Equations 36-37 describe the way LFPs are simulated (as the low-frequency parts of the extracellular potential) in the original biologically realistic model that generated the LFP data we use for

GPLA analysis. We also exploit it to provide an approximative LFP laminar profile for the population of pyramidal cells based on this equation, by injecting the same constant current to all cells and compartment of the linear arrangement, but having opposite signs for the axosomatic and dendritic compartment to respect charge neutrality and thus the dipolar structure (Figure 6D (broken line)).

Neuron exclusion criterion for GPLA

To reduce the small sample bias caused by a low number of spike events, we only use neurons that had a minimum average firing rate of 3 Hz firing. Nevertheless, using all neurons did not change the results significantly. For instance, in Figure 6G (in contrast to Figure 6A-F and H where excluded neurons based on their firing), to be compatible with the neural mass model we did not exclude any neuron.

Analytical neural field modeling of spike-field coupling

In order to justify and interpret our approach, we use a rate-based neural field model. Units are grouped in populations according to their cell-type on spatial localization. Spiking activity of a specific population p at possibly multidimensional location x is represented by its average spike rate $\lambda_p(x,t)$. Simultaneously, the LFP $L(X,t)$ is recorded at locations reflected by possibly different coordinates X.

Rate model of circuit dynamics

We follow classical neural field models, stating that the rate of each population evolves as a monotonous function of the membrane potential, itself controlled by post-synaptic currents (PSCs). Dynamics of the membrane potential V_p of each population p, is assumed to be governed by the following differential equation:

$$\frac{dV_p}{dt}(x,t) + \tau_p V_p(x,t) = \alpha_p \eta(x,t) + \sum_k \nu_{p\leftarrow k} s_k(x,t)\,, \tag{38}$$

where η represents the post-synaptic current generated by the external input (for which no spiking activity is available), s_k the normalized[4] post-synaptic current from the afferent population k, whose effect on target population p is scaled by synaptic strength $\nu_{p\leftarrow k}$.

The relationship between the normalized post-synaptic current at location x and spiking activity of the afferent population activity is modeled by spatio-temporal integration (Wilson and Cowan, 1973; Somers et al., 1995; Jirsa and Haken, 1996), possibly taking into account the propagation speed v_0 along the axons

$$s_k(x) = \int c_k(x,X)\lambda_k(x, t - |x-X|/v_0)dX, \tag{39}$$

where the connectivity kernel $c_p(x,X)$ models the density of synapses of neurons whose soma is located at target location x, with afferent neurons having their somas at location X. The integral covers the spatial domain where units' somas can be found, and may thus be 1-, 2- or 3-dimensional depending on the model. The kernel c_k reflects the spatial spread of axonal arborizations and as such can be approximated based on anatomical studies.

The elements finalizing the description of the state of the system are the relations between each population's membrane potential and rate, modeled by

$$\lambda_p = a_p(V_p)\,, \tag{40}$$

where a_p is a typically sigmoidal activation function (these are the only non-linearities considered in our neural mass equations), leading to the overall dynamical system represented in Figure 5A.

[4] By "normalized", we mean that s_k is a numerical quantity independent of the target population, possible target-specific differences in the PSCs being taken into account in the connectivity parameters $\nu_{p\leftarrow k}$, without loss of generality.

From synaptic currents to LFPs

The local field potential is the lower frequency (<150Hz) part of the electrical potential recorded in the extracellular space, generated by the transmembrane currents (Einevoll et al., 2013). Considering that active currents mostly reflect spiking activity, whose dynamics lies mostly above the typical LFP frequency range, we approximate the LFP as resulting from the linear superposition of passive membrane currents triggered by post-synaptic input currents (Mazzoni et al., 2015), leading to the equation

$$L(y,t) = \sum_p \int f_{p,e}(y,x)\eta(x,t)dx + \sum_{p,k} \int f_{p,k}(y,x)s_k(x,t)dx\,,$$

where $f_{p,k}(.,x)$ represents the electrical field spatial distribution generated by trans-membrane currents of the p cells with soma located at x, resulting from exciting them with post-synaptic unit currents of population k. Note due to charge neutrality of the cells, trans-membrane currents across the membrane of individual cells sum to zero, such that input currents at the levels of synapses are compensated by opposite trans-membrane currents away from them, typically leading to dipolar current distributions. Along the same lines, $f_{p,e}$ is the distribution associated with post-synaptic current resulting from exogenous inputs to p cells. Differences between these fields according to the afferent populations are due to the respective distribution of their synaptic button over the efferent cell, preferentially targeting either the peri-somatic or distant dendritic sites, as illustrated in Figure 5C. These field distributions are assumed dominated by currents originating from pyramidal cells, due to their individual and collective geometric arrangement (Nó, 1947; Lindén et al., 2011; Mazzoni et al., 2015), such that we can simplify the above equation to obtain

$$L(y,t) = \int f_{E,e}(y,x)\eta(x,t)dx + \sum_k \int f_{E,k}(y,x)s_k(x,t)dx\,. \tag{41}$$

Spike-LFP relation

Analysis of the frequency response of neural network models is a useful approach to understand their characteristics and underlying mechanisms (Ledoux and Brunel, 2011; Sherfey et al., 2018). In our case, this can be performed analytically by linearizing Equation 40 around an operating point, the neural field model becomes a linear time-invariant system controlled by the exogenous input $\eta(t)$. We can thus compute transfer functions for each variable of the system that will determine their response to a sinusoidal exogenous input at frequency f, based on the computation of temporal Fourier transforms of the signals. For a given signal, $s(t)$ the temporal Fourier transform at frequency f is given by

$$S(f) = \mathcal{F}_t[s](f) = \int_{\mathbb{R}} s(t)e^{-i2\pi ft}dt\,.$$

By applying the Fourier transforms on the left- and right-hand-side of dynamical equations, we can derive transfer functions linking the responses of each network variable to the exogenous input at a given frequency. We provide in Table 2 a list of time-domain variables and their corresponding notations for their time-domain Fourier transform, used in analytical developments.

Specifically, for a general exogenous input signal with time-domain Fourier transform $\mathcal{E}(X,f)$, where f denotes the temporal frequency, we obtain the input-output relation for population rates and LFP activity L

$$\lambda_p(x,f) = \int H_{\lambda_p}(x,X,f)\mathcal{E}(X,f)dX \quad \text{and} \quad L(y,f) = \int H_L(y,X,f)\mathcal{E}(X,f)dX\,.$$

A major simplification of this expression occurs when the exogenous input is separable in time and space,

$$\eta(X,t) = n(X)e(t))\,, \tag{42}$$

leading to $\mathcal{E}(X,f) = n(X)E(f)$ after temporal Fourier transform. This simplifying assumption models a number of typical inputs to the structure, including sinusoidal standing waves ($\eta(x,t) = n(x)e^{i2\pi ft}$) and traveling plane waves ($\eta(x,t) = e^{i(2\pi ft-kx)}$). This results in a simple expression for the covariance

Table 2. List of neural field and mass model variables (k indicates neuron population (afferent in case of synapse property))

Description	Symbol	Temporal Fourier transform	Equation
Exogenous input (spatio-temporal)	$\eta(x,t)$	$\mathcal{E}(x,f)$	38
Exogenous input (temporal)	$e(t)$	$\mathcal{E}(f)$	42
Spike rate	$\lambda_k(x,t)$	$\Lambda_k(x,f)$	43
Membrane potential	$v_k(x,t)$	$V_k(x,f)$	43
Synaptic activity	$s_k(x,t)$	$S_k(x,f)$	38, 39,41
Activation function	$a_k(x,t)$	$A_k(x,f)$	40

Table 3. List of neural mass model parameters

Parameter name	Symbol	Value (*Mass2D*)	Value (*MassAlpha*)
E membrane time constant	τ	$20ms$	$25ms$
I membrane time constant	δ	$5ms$	$25ms$
$E \leftarrow I$ synaptic strength	$\nu_{E\leftarrow I}$	0.01	0.85
$I \leftarrow E$ synaptic strength	$\nu_{I\leftarrow E}$	0.5	0.3
Alpha synapse time constant	σ	n.a.	$15ms$
$I \leftarrow I$ synaptic strength	$\nu_{I\leftarrow I}$	n.a. (accounted for in δ)	12.0

estimated across experimental trials between $rate$ and LFP at two possibly different spatial points (x, y)

$$\left\langle L(y,f)\lambda_p(x,f)\right\rangle = \left(\int H_L(y,X,f)n(X)dX\right)\left(\int H_{\lambda_p}(x,X,f)n(X)dX\right)|e(f)|^2 .$$

in which the input intervenes only as a multiplicative positive constant, and which is separable in both space variables x and y. As a consequence, the rank one approximation of the covariance between spiking units and LFP channels activity estimated by GPLA is informative about the microcircuit properties, as we explained in the STAR methods sections describing the analysis of the neural mass and neural field models.

Analysis and simulation of two population neural mass models

General description

The generic dynamic model of Equation 38 is exploited to describe network activity at a single location (i.e. we neglect the spatial extent of the considered structure) containing two cell types: pyramidal (E) and inhibitory (I), leading to the linear equations:

$$V_E + \tau_E \frac{dV_E}{dt} = \nu_{E\leftarrow E}s_E - \nu_{E\leftarrow I}s_I + \eta \tag{43}$$

$$V_I + \tau_I \frac{dV_I}{dt} = \nu_{I\leftarrow E}s_E - \nu_{I\leftarrow I}s_I + \alpha\eta \tag{44}$$

where $\boldsymbol{\nu}$ is a matrix gathering the non-negative synaptic strengths between populations, η the exogenous input to the network, with $\alpha \geq 0$ controlling the ratio between feed-forward excitation and inhibition. The term $\nu_{kj}s_j$ is the population averaged post-synaptic potential from population j to population k. In order to study quantitatively the effect of connectivity changes in the microcircuit, in this expression of the post-synaptic current, we isolate the synaptic strength coefficient $\nu_{k\leftarrow j}$, from a perisynaptic activity, that summarizes the dynamical processes occurring pre- and post-synaptically (synaptic delay, time constant induced by the post-synaptic channel conductance, ...). In the simplest case we assume peri-synaptic activity s_j can be approximated by the spike rate of population j, λ_j (up to a multiplicative constant that is absorbed by ν_{kj}). Alternatively, we model synaptic dynamics by a linear differential equation controlled by this rate (see model *MassAlpha* below).

The neural mass models will be analyzed with linear response theory, such that the a_k's of Equation 40 will be linearized around an equilibrium point of the dynamical system (that can be computed for vanishing input $\eta = 0$), and the resulting multiplicative constants will be themselves absorbed in the connectivity matrix $\boldsymbol{\nu}$, leading to replacing Equation 40 by

$$\lambda_k = V_k \,. \tag{45}$$

Next we describe the two linearized neural mass models exploited to interpret GPLA results of hippocampal simulations (see Figure 6). Parameter values for both models are reported in Table 3

Mass2D: E-I interactions without synaptic dynamics

Starting from Equation 43 and using Equation 45, the linearized system can be trivially reduced to the two-dimensional dynamical system (up to rescaling of the connectivity matrix coefficients)

$$\lambda_E + \tau \frac{d\lambda_E}{dt} = -\nu_{E\leftarrow I}\lambda_I + \eta \,, \tag{46}$$

$$\lambda_I + \delta \frac{d\lambda_I}{dt} = \nu_{I\leftarrow E}\lambda_E + \alpha\eta \,. \tag{47}$$

where τ and δ are time constants derived from membrane time constant τ_k and recurrent synaptic connection ν_{kk}. Linear response analysis then relies on the Laplace transform (with Laplace variable p) of these equation

$$\Lambda_E(p) + \tau p \Lambda_E = -\nu_{E\leftarrow I}\Lambda_I + N(p) \,, \tag{48}$$

$$\Lambda_I(p) + \delta p \Lambda_I = \nu_{I\leftarrow E}\Lambda_E + \alpha N(p) \,. \tag{49}$$

For the case of *no feedforward inhibition* ($\alpha = 0$), this leads to the ratio of excitatory to inhibitory activity in the Laplace domain

$$\frac{\Lambda_E}{\Lambda_I}(p) = \frac{\delta p + 1}{\nu_{I\leftarrow E}} \tag{50}$$

resulting in excitatory activity being in advance of $\tan^{-1} 2\pi f \tau$ with respect to inhibitory activity at frequency f.

For the case of strong feedforward inhibition ($\alpha = 1$), this leads to

$$\frac{\Lambda_E}{\Lambda_I}(p) = \frac{\delta p + 1 - \nu_{E\leftarrow I}}{\tau p + 1 + \nu_{I\leftarrow E}} \tag{51}$$

such that the phase shift between population is of constant sign across frequencies, but may be positive or negative depending on the exact parameters' values governing the E-I dynamics. Plots summarizing these situations are provided in Figure 6D.

MassAlpha: E-I interactions with alpha type synaptic impulse response

Together with Equation 43, we include in addition a non-trivial synaptic dynamics in the form of the differential equation (assuming same dynamic for both AMPA and GABA synapses),

$$\sigma^2 \frac{d^2 s_k}{dt^2} + 2\sigma \frac{d s_k}{dt} + s_k = \lambda_k \,, \tag{52}$$

This corresponds the classical *alpha synapse* used in computational models (e.g. implemented in the *NEURON* software (Carnevale and Hines, 2006)), modeling the response to a single spike with the alpha function

$$s(t) = \frac{1}{\sigma^2} t e^{-t/\sigma} \,. \tag{53}$$

By combining these equations with linear activations (Equation 45), the dynamics of the circuit is summarized by a 6-dimensional state-space model that can be studied analytically with linear response theory.

Analysis and simulations of neural field models

When taking into account the spatial extension of the network, neural mass models can be extended to neural field models, where the variables described above possibly depend on space. Consider one or two spatial dimensions tangential to the layers of the network (assuming a layered organization like the hippocampus or cortex), the key phenomenon that should be additionally modeled is then the coupling between activity in different locations of the network entailed by horizontal connections. In line with the literature and to simplify the analysis, we will consider only the excitatory connections are spatially extended. With respect to the above generic neural field model Equations 38-40, equations pertaining to synaptic activity and rate need only to be specified. We use the model introduced by Jirsa and Haken (1996, Equation 15) with a spatial diffusion term with characteristic distance r_0 and axonal propagation speed $v_0 = r_0\gamma$, which takes the form of a damped wave equation:

$$\frac{1}{\gamma^2}\frac{\partial^2 s_E}{\partial t^2} + \frac{2}{\gamma}\frac{\partial s_E}{d\partial t} + s_E - r_0^2 \Delta s_E = \lambda_E + \frac{1}{\gamma}\frac{\partial \lambda_E}{\partial t}\,, \tag{54}$$

where Δ is the Laplacian operator, while $s_I = V_I$ to encode purely local inhibition, eliminating redundant multiplicative factors.

To specify Equation 40, we use sigmoid activation functions for both AMPA and GABA synapses,

$$\lambda_E = \frac{Q_E}{1+\exp(-\chi_E.(V_E - V_{th,E}))}\,, \tag{55}$$

$$\lambda_I = \frac{Q_I}{1+\exp(-\chi_I.(V_I - V_{th,I}))}\,, \tag{56}$$

whose parameters (maximum rate Q_k, spiking threshold $V_{th,k}$ and excitability χ_k for population k) are adjusted in order to obtain different types of dynamics, either evolving around a stable equilibrium point (model *FieldStable*) or with a clear oscillatory activity (model *FieldOsc*).

Spatio-temporal phase analysis in 1D

Before simulating 2D neural field models with an explicit method, we investigate analytically properties of a simplified 1D model. In this case, the partial differential Equation 54 corresponds to an exponentially decaying connectivity, with axonal propagation speed $v_0 = r_0\gamma$, such that the resulting post-synaptic current takes the integral form (see Jirsa and Haken (1996, Equation (14)))

$$s_E(x,t) = \frac{1}{2r_0}\int \exp(-|x-X|/r_0)\lambda_E(X, t-|x-X|/v_0)dX\,. \tag{57}$$

If we take the neural field equation in the context of horizontal connection along non-myelinated axons ($v_0 \sim 1m/s$, $r_0 < 1mm$), the typical value of γ is beyond 1000 such that if we focus on frequencies below $200Hz$, we may neglect for the temporal derivatives of the partial differential equation. This leads to the following approximation for the dynamics of excitatory post-synaptic current.

$$s_E - r_0^2\frac{\partial^2 s_E}{dx^2} = \lambda_E\,. \tag{58}$$

In an unbounded 1D medium, assuming that activities vanish at large distances, we can use the spatial Fourier transform $\widehat{f}(t,z) = \mathcal{F}_x\left[f(t,x)\right](z) = \int_{\mathbb{R}} f(t,x)e^{-2i\pi zx}dx$ to derive the expression of s_E as a function of λ_E

$$\widehat{s_E}(t,z) = \frac{1}{1+(2\pi z)^2 r_0^2}\widehat{\lambda_E}(t,z)\,. \tag{59}$$

In order to get back to the original spatial position domain, we use a general Fourier transform relation for an arbitrary complex parameter a such that $\mathrm{Re}[a] \geq 0$:

$$\mathcal{F}_x\left[\frac{1}{2a}e^{-|x|a/r_0}\right](z) = \frac{r_0}{(2\pi z r_0)^2 + a^2}\,, \tag{60}$$

This formula can be inverted, considering an arbitrary complex number b, and defining $\sqrt{b}$ to be the unique complex number such that $\sqrt{b}^2 = b$ and $\mathrm{Re}\sqrt{b} \geq 0$, we get

$$\mathcal{F}_z^{-1}\left[\frac{r_0}{(2\pi z)^2 r_0^2 + b}\right](x) = \frac{1}{2\sqrt{b}} e^{-|x|\sqrt{b}/r_0}\,, \tag{61}$$

For the particular case $b = 1$, we get $\sqrt{b} = 1$, which, in the spatial position domain, leads to

$$s_E(t,x) = \int_{\mathbb{R}} \lambda_E(t,y) h_{r_0}(x-y)dy = h_{r_0} * \lambda_E(t,x)\,, \tag{62}$$

where $*$ denotes spatial convolution and $h_{r_0}(x) = \frac{1}{2r_0} e^{-|x|/r_0}$. This reflects that horizontal connectivity generates EPSCs corresponding to a spatial smoothing of the excitation rate spatial distribution.

After linearizing around the operating point of the network (absorbing again the resulting multiplicative constant in the connectivity matrix), we obtain the equation of the dynamics by modifying Equation 38 (assuming neither long range nor feedforward inhibition)

$$\lambda_E(t,x) + \tau \frac{d\lambda_E}{dt} = \nu_{E\leftarrow E} s_E(t,x) - \nu_{E\leftarrow I}\lambda_I(t,x) + \eta(t,x)\,, \tag{63}$$

$$\lambda_I(t,x) + \delta \frac{d\lambda_I}{dt} = \nu_{I\leftarrow E} s_E(t,x)\,, \tag{64}$$

where the synaptic strength values incorporate multiplicative constants resulting from the linearization of Equations 55-56. By computing the temporal (with frequency variable f) and spatial Fourier transform of each equation, we get (using $p = i2\pi f$)

$$\widehat{\Lambda_E} + \tau p \widehat{\Lambda_E} = \nu_{E\leftarrow E}\widehat{S_E} - \nu_{E\leftarrow I}\widehat{\Lambda_I} + \widehat{H}\,, \tag{65}$$

$$\Lambda_I(t,x) + \delta p \widehat{\Lambda_I} = \nu_{I\leftarrow E}\widehat{S_E}\,. \tag{66}$$

Eliminating $\widehat{\Lambda_I}$ we get

$$(1+\tau p)\widehat{\Lambda_E} = \left(\nu_{E\leftarrow E} - \nu_{E\leftarrow I}\frac{\nu_{I\leftarrow E}}{1+\delta p}\right)\widehat{S_E} + \widehat{H}\,. \tag{67}$$

Combined with the spatially Fourier transformed horizontal connectivity Equation 58 (using $k = i2\pi z$)

$$(1 - r_0^2 k^2)\widehat{S}_E = \widehat{\Lambda_E}\,, \tag{68}$$

this leads to

$$\widehat{S_E} = \frac{1}{1+\tau p}\frac{\widehat{E}}{-r_0^2k^2 + 1 + \frac{1}{1+\tau p}\left(\frac{\nu_f}{1+\delta p} - \nu_{E\leftarrow E}\right)}\,, \tag{69}$$

where we define the *feedback inhibition gain* $\nu_f = \nu_{E\leftarrow I}\nu_{I\leftarrow E}$.

Introducing our time-space separability assumption on the exogenous input

$$\eta(x,\,t) = n(x)\epsilon(t) \tag{70}$$

leads to

$$\widehat{S_E}(z,\,f) = \frac{1}{1+i2\pi\tau f}\frac{E(f)\widehat{n}(z)}{-r_0^2(2\pi z)^2 + 1 + \frac{1}{1+i2\pi\tau f}\left(\frac{\nu_f}{1+i2\pi\delta f} - \nu_{E\leftarrow E}\right)}\,, \tag{71}$$

By defining

$$b = 1 + \frac{1}{1+i2\pi\tau f}\left(\frac{\nu_f}{1+i2\pi\delta f} - \nu_{E\leftarrow E}\right)\,, \tag{72}$$

and using the inverse spatial Fourier transform of Equation 61, we get

$$S_E(x,\,f) = \frac{1}{2r_0\sqrt{b}}\frac{E(f)}{1+i2\pi\tau f} n(x) * e^{-|x|\sqrt{b}/r_0}\,, \tag{73}$$

Assuming the exogenous input does not impose a spatial phase gradient to the structure (i.e. $n(x)$ is positive real for all locations up to a multiplicative constant), the phase gradient at a given frequency

will be controlled by the imaginary part of $\sqrt{b}$. Specifically, to investigate qualitatively the phase gradient around a peak of activity of the exogenous input, we assume that $n(x)$ is a dirac at $x = 0$. Then the spatial variation of the phase around $x = 0$ take the form

$$\phi(x) = -\frac{|x|}{r_0}\mathrm{Re}\left[\sqrt{b}\right] . \tag{74}$$

This dirac approximation, does not match well our simulations (using a Gaussian shape spatial input distribution). However, computing spiking activity form S_E based on Equation 58, to obtain the spatial distribution of the spike vector, will have a deblurring effect compensating the convolution by $n(x)$ in Equation 75, making in closer to a Dirac. As a consequence, we will interpret the data based on the following approximation

$$\lambda_E(x,\ f) \approx C\frac{1}{2r_0\sqrt{b}}\frac{E(f)}{1+i2\pi\tau f}e^{-|x|\sqrt{b}/r_0} , \tag{75}$$

up to a multiplicative constant C.

In order to investigate the qualitative effect of the microcircuit connectivity on this spatial gradient, we assume $\tau = \delta$ and use a low (temporal) frequency assumption of the form $f << 1/\tau$, such that we can exploit a first order expansion for the fractions containing the term $\tau p << 1$. This leads to the approximation

$$b \approx 1 + (1 - i2\pi\tau f)\left(\nu_f(1 - i2\pi\tau f) - \nu_{E\leftarrow E}\right) \approx 1 + \nu_f - \nu_{E\leftarrow E} - i2\pi\tau f(2\nu_f - \nu_{E\leftarrow E}) . \tag{76}$$

Simple geometric considerations show that the sign of the imaginary part of b is the same as the sign of its square root, such that under our simplifying assumption, the sign of the gradient taken algebraically from center ($x = 0$) to surround ($|x| > 0$) is the sign of

$$2\nu_f - \nu_{E\leftarrow E} , \tag{77}$$

showing that strong feedback inhibition will tend to put the populations surrounding $x = 0$ in advance with respect to this center point, while weak feedback inhibition (with respect to feedback excitation), will to generate a phase lag of the surround with respect to the center.

Neural field simulation in 2D

While the above analysis is much easier to perform in 1D, in most structures (and in particular cortex), the domain spanned by horizontal connectivity is better approximated by a 2D domain, which can also be sampled by modern electrode arrays. We thus simulate the dynamics of such 2D system to get insight into the characteristics revealed by GPLA analysis in this context. We use simplified notations for the 2D (in space) time-varying scalar fields $V(t,x,y) = s_E((x,y),t)$ and $I(t,x,y) = \lambda_E((x,y),t)$. Let Δx and Δt be the spatial and temporal grid spacings, and $V^n_{j,l} = V(n\Delta t, j\Delta x, l\Delta x)$ the discretized field. We use a Forward Time Centered Space (FTCS) finite difference scheme to simulate the above neural field model (Fletcher, 1991). FTCS relies on making the approximations

$$\frac{\partial V}{\partial t}(t,x,y) \approx \frac{1}{\Delta t}\left(V^{n+1}_{j,l} - V^n_{j,l}\right) , \quad \frac{\partial^2 V}{\partial t^2}(t,x,y) \approx \frac{1}{(\Delta t)^2}\left(V^{n+1}_{j,l} + V^{n-1}_{j,l} - 2V^n_{j,l}\right) , \tag{78}$$

$$\text{and} \quad \frac{\partial^2 V}{\partial x^2}(t,x,y) \approx \frac{1}{(\Delta x)^2}\left(V^n_{j+1,l} + V^n_{j-1,l} - 2V^n_{j,l}\right) . \tag{79}$$

Applying these approximations to Equation 54, leads to an explicit scheme for the field values at time $(n+1)\Delta t$ based on all values at time $n\Delta t$ and $(n-1)\Delta t$.

$$V^{n+1}_{j,l} = \frac{1}{G+1}\left((G-1)V^{n-1}_{j,l} + 2V^n_{j,l} + G^2\left(R^2\left(K_\ell *_d \boldsymbol{V}^n\right)_{j,l} + I^n_{k,l} + \frac{1}{G}\left(I^n_{k,l} - I^{n-1}_{k,l}\right) - V^n_{j,l}\right)\right) , \tag{80}$$

where $G = \gamma\Delta t$, $R = r_0/\Delta x$ and $K_\ell *_d$ denotes the discrete 2D spatial convolution with the discrete Laplace operator

$$K_\ell = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} . \tag{81}$$

The parameters chosen for both models presented in main text are reported in Table 4.

Table 4. List of neural field model parameters

Parameter name	Symbol	Value (*weak Rec.*)	Value (*strong Rec.*)
E membrane time constant	τ_E	$20ms$	$20ms$
I membrane time constant	τ_I	$20ms$	$20ms$
E-E synaptic strength	$\nu_{E\leftarrow E}$	0.2	0.2
I-I synaptic strength	$\nu_{I\leftarrow I}$	0	0
$E \rightarrow I$ synaptic strength	$\nu_{E\leftarrow I}$	0.2	0.2
$I \rightarrow E$ synaptic strength	$\nu_{I\leftarrow E}$	1	1
E excitability	χ_E	1	1
I excitability	χ_I	0.1	3.33
E sigmoid threshold	$V_{th,E}$	0	0
I sigmoid threshold	$V_{th,I}$	0	0
E maximum rate	Q_E	$20Hz$	$20Hz$
I maximum rate	Q_E	$20Hz$	$20Hz$

Quantification and statistical analysis

Parameter estimation of von Mises distribution

The von Mises distribution (VM), which is also known as "circular normal" distribution is the counterpart of the Gaussian distribution for circular data (Fisher, 1995, Chapter 3). We used it for various purposes in this work (e. g. to model the spiking probability to synthesize phase-locked spike trains).

The VM distribution takes the form,

$$p(\phi|\varphi_0, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\left(\kappa \cos(\phi - \varphi_0)\right), \tag{82}$$

where $I_0(\kappa)$ is the modified Bessel function of order zero (Equation 33). In Figure 8 and 13, we fit a VM distribution to the pooled phases of spike and LFP vectors coefficients. We use a maximum likelihood (ML) method for estimating the two parameters of the VM distribution, φ_0 and κ (Fisher, 1995). The ML estimation of φ_0 is simply the sample mean direction, denoted by $\bar{R}$ (for spike-LFP data is the locking phase). Maximum likelihood estimation of, $\hat{\kappa}$, is the solution of following equation:

$$A_1(\hat{\kappa}) = \bar{R}, \tag{83}$$

and A_1 is a ratio of two modified Bessel functions:

$$A_1(x) = \frac{I_1(x)}{I_0(x)}. \tag{84}$$

Approximate solutions are available for $\hat{\kappa}$ (Fisher, 1995, sec. 4.5.5)

$$\hat{\kappa} = \begin{cases} 2\overline{R} + \overline{R}^3 + 5\overline{R}^5/6 & \overline{R} < 0.53 \\ -0.4 + 1.39\overline{R} + 0.43/(1-\overline{R}) & 0.53 \leq \overline{R} < 0.85 \\ 1/\left(\overline{R}^3 - 4\overline{R}^2 + 3\vec{R}\right) & \overline{R} \geq 0.85 \end{cases} \tag{85}$$

where $\overline{R}$ is the resultant length of the phases.

Animal preparation and intracortical recordings

The methods for surgical preparation, anesthesia, and presentation of visual stimuli for the Utah array recordings have been described in previous studies (see Logothetis et al. (1999, 2002); Belitski et al. (2008); Safavi et al. (2018)).

Data collection

Neural signals were recorded with a NeuroPort Cortical Microelectrode Array (Blackrock Microsystems, Salt Lake City, Utah USA). An array was implanted in the inferior convexity of the prefrontal cortex (see Safavi et al. (2018) for more details). The arrays are 4mm × 4mm with a 10 by 10 electrode configuration. Neural signals recorded from 96 of the available 100 electrodes. Neural activity was recorded in 200 trials. Each trial consisted of a 10s period of movie presentation, followed by 10s of a blank screen (inter-trial).

LFP extraction

The raw signals were low-pass filtered using an 8th order Chebyshev Type 1 filter with a cut-off frequency of 200Hz and a pass-band ripple less than 0.05dB. Forward and backward filtering was used to minimize phase distortions caused by the filtering. Next, the filtered signal was decimated to a sampling frequency of approximately 500Hz.

Spike detection

For detecting multi-unit spikes, the raw signal was band-pass filtered using a minimum–order finite impulse response (FIR) filter (Rabiner et al., 1975) with pass-band cut-off frequencies of 600Hz to 5800Hz and stop-band cut-off frequencies of 400Hz and 6000Hz, with at least 65dB attenuation in the stop-bands and less than 0.002dB ripple within the pass-band. The amplitude threshold for spike detection was set to 5 standard deviations above the average of the filtered signal (Quiroga, 2007). To spare computational costs, the standard deviation of the signal for each channel was estimated using a smaller, randomly chosen section of the filtered signal. Spike times with inter-spike intervals less than the refractory period of 0.5ms were eliminated.