Gamma-band synchronization in the neocortex: novel analysis methods and their application to sensory and motivational systems

Vinck, M.A.

Citation for published version (APA):
CHAPTER 2

Improved measures of phase-coupling between spikes and the Local Field Potential

This article was originally published as


and is a follow-up of the earlier work

Abstract

An important tool to study rhythmic neuronal synchronization is provided by relating spiking activity to the Local Field Potential (LFP). Two types of interdependent spike-LFP measures exist. The first approach is to directly quantify the consistency of single spike-LFP phases across spikes, referred to here as point-field phase synchronization measures. We show that conventional point-field phase synchronization measures are sensitive not only to the consistency of spike-LFP phases, but are also affected by statistical dependencies between spike-LFP phases, caused by e.g. non-Poissonian history-effects within spike trains such as bursting and refractoriness. To solve this problem, we develop a new pairwise measure that is not biased by the number of spikes and not affected by statistical dependencies between spike-LFP phases. The second approach is to quantify, similar to EEG-EEG coherence, the consistency of the relative phase between spike train and LFP signals across trials instead of across spikes, referred to here as spike train to field phase synchronization measures. We demonstrate an analytical relationship between point-field and spike train to field phase synchronization measures. Based on this relationship, we prove that the spike train to field pairwise phase consistency (PPC), a quantity closely related to the squared spike-field coherence, is a monotonically increasing function of the number of spikes per trial. This derived relationship is exact and analytic, and takes a linear form for weak phase-coupling. To solve this problem, we introduce a corrected version of the spike train to field PPC that is independent of the number of spikes per trial. Finally, we address the problem that dependencies between spike-LFP phase and the number of spikes per trial can cause spike-LFP phase synchronization measures to be biased by the number of trials. We show how to modify the developed point-field and spike train to field phase synchronization measures in order to make them unbiased by the number of trials.
2.0 Introduction

Neuronal spiking outputs are often temporally coordinated across neuronal groups. Temporal coordination often manifests itself as rhythmic or phase synchronization, i.e. phase-coupling of oscillatory activity, which has been proposed as a candidate mechanism for the flexible routing of neuronal signals (Abeles, 1982; Fries, 2005; Varela et al., 2001; Burchell et al., 1998; Womelsdorf et al., 2010b; Gielen et al., 2010; Börgers & Kopell, 2008; Salinas & Sejnowski, 2001; Buschman & Miller, 2007; Benchenane et al., 2011; Gregoriou et al., 2009), assembly formation (Singer, 1999; Buzsáki, 2010; Kopell et al., 2011; Canolty et al., 2010; Battaglia et al., 2011), and coding (König et al., 1995; Fries et al., 2007; Panzeri et al., 2010; Vinck et al., 2010a; O’Keefe & Recce, 1993; Azouz & Gray, 2003; Nadasdy, 2010; Reimer & Hatsopoulos, 2010; Siegel et al., 2009; VanRullen et al., 2005; Havenith et al., 2011; Koepsell et al., 2010).

To quantify rhythmic synchronization, we need to relate signals generated by separate sources. An increasingly popular choice of signals is the combination of spiking activity and the LFP (Local Field Potential) (e.g. see Fries et al. (2001b, 1997); Gregoriou et al. (2009); Pesaran et al. (2002, 2008); Ray & Maunsell (2010); Colgin et al. (2009); Csicsvari et al. (2003); Vinck et al. (2010a); van Wingerden et al. (2010b); Chalk et al. (2010); O’Keefe & Recce (1993); Eeckman & Freeman (1990); Womelsdorf et al. (2010a); Brown et al. (2004); Jarvis & Mitra (2001); Zeitler et al. (2006); Lepage et al. (2011); Siapas et al. (2005); Jacobs et al. (2007); Donoghue et al. (1998); Gray et al. (1989); Wehr & Laurent (1996); Cardin et al. (2009); Sohal et al. (2009); Okun et al. (2010); Bichot et al. (2005); Buffalo et al. (2011); Canolty et al. (2010); Montemurro et al. (2008); Kayser et al. (2009); Mitchell et al. (2009); Ray et al. (2008); Saalmann et al. (2007); van der Meer & Redish (2009); Denker et al. (2011); Lakatos et al. (2007); Besserve et al. (2010); Zanos et al. (2011); Destexhe et al. (1999)).

Spike-LFP measures are thought to be more sensitive to phase synchronization than spike-spike measures, because the LFP measures the compound synaptic activity generated by a large pool of neurons (Mitzdorf, 1985; Logothetis, 2003). This sensitivity is especially advantageous for the study of long-range interactions: Many studies have revealed functionally relevant changes in cross-areal spike-LFP phase synchronization (e.g. Pesaran et al. (2008); Womelsdorf et al. (2007); Siapas et al. (2005); Gregoriou et al. (2009); Benchenane et al. (2010); Colgin et al. (2009); Sigurdsson et al. (2010)). In comparison to LFP-LFP measures, spike-LFP measures have the advantage that they provide an indirect measure of the relationship between spiking outputs and synaptic inputs (Okun et al., 2010; Harvey et al., 2009), and that they can be related to neuronal firing rate selectivity (e.g. Gregoriou et al. (2009); van Wingerden et al. (2010b)) and cell type (Benchenane et al., 2010; Tukker et al., 2007; Klausberger et al., 2003; Csicsvari et al., 2003; Hasenstaub et al., 2005; van Wingerden et al., 2010a; Wulff et al., 2009; Sohal et al., 2009; Cardin et al., 2009; Bragin et al., 1995; Gulyas et al., 2010), which can provide useful insights into the mechanisms of rhythmogenesis (for an overview, see Wang (2010)). At the same time, spike-LFP measures avoid the problem that volume-conduction and common referencing can create spurious correlations between LFP signals (e.g. Vinck et al. (2011); Nolte et al. (2004); Sirotta et al. (2008); Stam et al. (2007); Nunez & Srinivasan (2006)).

A wide variety of spike-LFP measures exist, and they can be categorized into two classes. First, some measures are based on a quantification of the consistency of individual spike-
LFP phases at a certain frequency (e.g. Vinck et al. (2010b); Csicsvari et al. (2003); Fries et al. (1997); Grasse & Moxon (2010); Canolty et al. (2010); Fries et al. (2001b); Tukker et al. (2007); Hansen & Dragoi (2011); Siapas et al. (2005); Sirota et al. (2008); Colgin et al. (2009); Denker et al. (2011)). In this paper, we will refer to these measures as **point-field phase synchronization measures** (Fig 2.1), because they relate single points, i.e. spikes, to the ongoing LFP oscillations, i.e. fields. A major problem bound to these point-field phase synchronization measures is that they can be strongly biased by the number of spikes. Previously, we addressed this problem by introducing the pairwise phase consistency (PPC) (Vinck et al., 2010a).

However, there are two further problems related to point-field phase synchronization measures that were not solved by the PPC measure. First, the problem that by the direct quantification of the consistency of individual spike-LFP phases through circular statistics, it is implicitly assumed that individual spike-LFP phases can be modeled as outcomes of statistically independent random variables. This assumption is typically violated because neurons display non-Poissonian history effects, such as refractoriness and bursting (Section 2.1 and 2.2). For example, given the observation of a certain spike-LFP phase for a bursting neuron, we may be more likely to observe a similar spike-LFP phase again, independent of the modulation of spiking probability with LFP phase. To solve the latter problem, we develop a new point-field phase synchronization measure that is not affected by a dependence between spike-LFP phases that were obtained from the same trial (Section 2.2). We achieve this by modifying the PPC such that its computation is now based only on pairs of spike-LFP phases that were obtained from separate trials (see Eq. 2.10 below). A second problem that we address here is that dependencies between spike count and spike-LFP phase, e.g. as consequence of theta or gamma phase precession/shifting (O’Keefe & Recce, 1993; Vinck et al., 2010a; Tiesinga & Sejnowski, 2010; Harris et al., 2002; Mehta et al., 2002; Schmidt et al., 2009; Siegel et al., 2009; Huxter et al., 2008), can cause point-field phase synchronization measures to become biased by the number of trials. We show how the point-field PPC measure can be modified in order to render it unbiased as a function of the number of trials (see Eq. 2.15 below).

The second class of spike-field measures is inspired by the classical coherence approach for EEG and MEG data (e.g., see Nunez & Srinivasan (2006); Mitra & Pesaran (1999); Walter (1963); Adey et al. (1961); Lachaux et al. (1999)) and is characterized by quantifying, for a certain frequency, the consistency of the relative phases between spike train and concurrent LFP signals across trials, as opposed to quantifying it across spikes (Fig 2.2) (e.g. Fries et al. (2008); Gregoriou et al. (2009); Jarvis & Mitra (2001); Pesaran et al. (2008); Chalk et al. (2010); Pesaran et al. (2002); Lepage et al. (2011); Mitra & Bokil (2008)). In this paper, we will refer to this second type of measures as **spike train to field phase synchronization measures**. With this approach, the spike trains are not broken down into single spikes that are then related to the LFPs around them, as opposed to the point-field phase synchronization measures. Instead, the spike train to field phase synchronization measures are based on the computed cross-spectra between the spike train signals of a certain duration and the concurrent LFPs.

Although both point-field and spike train to field phase synchronization measures are widely applied, it is unknown how they compare. In this paper, we derive an exact, analytical relationship (see Eq. 2.28 below) between the point-field PPC measures and the spike train to field PPC, which is a quantity closely related to the squared spike-field coherence (e.g.
see Jarvis & Mitra (2001)), with the only difference that the latter spike train to field index is a concurrent measure of amplitude co-variabilities as well, while the former spike train to field PPC measure only considers the consistency of relative phases. From this relationship, it follows that the spike train to field PPC is strictly positively dependent on the number of spikes per trial if the consistency of individual spike-LFP phases is independent of the number of spikes per trial. These findings are consistent with the previous result that the spike train to field coherence is positively dependent on the number of spikes (Zeitler et al., 2006; Lepage et al., 2011). Our analytic derivation is exact and directly leads to two corrected versions of the spike train to field PPC measures that are fully independent of the number of spikes per trial (Section 2.4).

The paper is organized as follows. We first introduce conventional point-field phase synchronization measures (Section 2.1). In Section 2.2, we study the effect of violations of the assumption of independency between spike-LFP phases (for different spikes) on these conventional point-field phase synchronization measures, and introduce a point-field measure that does not require this assumption. In Section 2.3, we study the consequences of a dependence between spike-LFP phase and spike count, and introduce a measure that is not affected by this dependence. In Section 2.4, we demonstrate an exact, analytical relationship between point-field and spike train to field phase synchronization measures and study the dependence of spike train to field phase synchronization measures on the number of spikes per trial. We end with a comparison of the advantages and disadvantages of the developed measures.

2.1 Definitions of conventional indices

2.1.1 A mathematical framework for point-field phase synchronization measures

We first develop a mathematical framework for point-field phase synchronization measures. Our starting point is the observed data, i.e. spike counts and spike-LFP phases, and we then proceed with modeling the observed data as outcomes of random variables. Suppose we perform simultaneous recordings of an LFP and the spiking activity of a neuronal source. We model the trial-wise spike counts by defining the vector of identically and independently distributed random variables $N \equiv (N_1, \ldots, N_M) \in \mathbb{N}^M$, where $M$ is the number of trials, whose observed value is defined as $n \equiv (n_1, \ldots, n_M) \in \mathbb{N}^M$, i.e. the experimentally observed counts of spikes emitted in each trial. Define $N \equiv \sum_{m=1}^{M} N_m$, and its observed value $n \equiv \sum_{m=1}^{M} n_m$, the observed total number of spikes. The set of trials that contained at least one spike is defined as $M \equiv \{m \in \{1, \ldots, M\} : N_m > 0\}$, such that $|M|$ is the number of trials with at least one spike. For every spike, we take the LFP segment that is centered around it, as illustrated in Figure 2.1, and spectrally decompose it, to obtain the spike-LFP phase at each frequency $f$. The spectral decomposition of the LFP may, without loss of generality, be based on e.g. the Discrete Fourier Transform or the Hilbert Transform on the filtered signal. Thus for all trials $m \in M$ we observe, at a particular frequency $f$, a data vector $\theta_m(f) \equiv (\theta_{1,m}(f), \ldots, \theta_{n_m,m}(f))$ containing $n_m$ spike-LFP phase values, which are depicted in Figure 2.1. The dependence of $\theta_m$ and related statistics on frequency is omitted and implicitly assumed in what follows.

In order to make statements about the statistics of the population from which these sample spike-LFP phase data were drawn, we model the observed spike-LFP phases as outcomes
of random variables. A straightforward approach may seem to define $n$ independent and identically distributed random variables ($\Theta_1, \ldots, \Theta_n$) (as in Vinck et al. (2010b)). However, the number of spikes itself is a random variable, such that we cannot restrict ourselves to defining $n$ random variables. Furthermore, for a given trial, the spike-LFP phase may be statistically dependent on the number of spikes, as it is for example the case when theta phase precession (O’Keefe & Recce, 1993) is present. To account for these two intricacies, we proceed as follows. For all $m \in \{1, \ldots, M\}$, let $\Theta_m = (\Theta_{1,m}, \Theta_{2,m}, \ldots)$ be a vector of random variables. For all $(k, m) \in \{(k, m) : k \in \{1, \ldots, N_m\}, m \in M\}$, we let the observed value of $\Theta_{k,m}$ equal a randomly chosen - without replacement - element of $\theta_m$, and obtain a vector representation of the spike-LFP phases by defining the spike-LFP phase vector $U_{k,m} \equiv (\cos(\Theta_{k,m}), \sin(\Theta_{k,m}))$.

![Diagram](image)

Figure 2.1: Illustration of point-field phase synchronization measures. Spikes (vertical, solid lines) are related to LFPs by determining the instantaneous LFP phases at the times of spiking by centering a window (horizontal, solid lines) around every spike. Thus, the phase for each spike is determined from the LFP snippet that is centered around it. The instantaneous LFP phase at the time of spiking can be directly gauged by using the dotted lines that connect the spikes to the LFP. The frequency resolution, which is inversely proportional to time resolution, is proportional to the window length, which needs to be sufficiently large. However, if the windows are too long, then the loss of time-resolution may lead to a decrease in spike-LFP phase consistency. The instantaneous spike-LFP phases are depicted by the unit arrows that originate from the spikes. The LFP was modeled as band-stop filtered Gaussian white noise, as in Zeitler et al. (2006). Here, we show only two ($m = 1$ and $m = 2$) out of $M$ available trials. The measures of point-field phase consistency are computed over the $n$ relative spike-LFP phases. By vector addition of the $n$ spike-LFP phase vectors, a resultant vector is obtained, whose length is conventionally taken as a measure of point-field phase consistency.

Thus, for all $m \in M$, we model all $\Theta_{k,m}$ elements of $\Theta_m$ for which $1 \leq k \leq N_m$ as iden-
Figure 2.2: Illustration of spike train to field phase synchronization measures. Spikes (vertical, solid lines) are related to LFPs by computing, using the Fast Fourier Transform, the respective phase offsets of the spike train and the LFP at a particular frequency (grey arrows) and subsequently determining the relative phase between these phase offsets (black arrows). Thus, the relative phases are computed for a time-period, not for an individual spike. The LFP was generated using the same model as in Figure 2.1. Shown here are two trials ($m = 1$ and $m = 2$). The measures of spike train to field phase consistency are computed over the $M$ relative phases. By vector addition of the $M$ relative phase vectors, a resultant vector is obtained, whose length is conventionally taken as a measure of the consistency of the relative phases between LFPs and spike trains.

...tically distributed (i.e., across time). In practice, it can occur that spike-LFP phases from the same trial are not identically distributed, because the mean spike-LFP phase can be a function of a time-varying variable, which occurs for example during theta phase precession (O’Keefe & Recce, 1993) and gamma phase-shifting (Vinck et al., 2010a). Mixing distributions with different mean spike-LFP phases causes an average decrease in point-field phase synchronization measures. If the mean spike-LFP phase is indeed non-stationary, then this reduction in point-field phase synchronization measures may be avoided by using one of the following two strategies. First, by computing our point-field phase synchronization measures over smaller trial segments for which the mean spike-LFP phase is approximately stationary, at the cost of an increase in estimator variance. Second, by explicitly modeling the mean spike-LFP phase as a function of a time-varying random variable, and subsequently computing our point-field phase synchronization measures over the residual model error terms, as in Vinck et al. (2010a).

Further, for all $k \in \{1, 2, \ldots\}$, we model all elements of the vector ($\Theta_{k,1}, \ldots, \Theta_{k,M}$) as identically distributed (i.e., across trials). In practice, it can occur that spike-LFP phases are not identically distributed across trials, e.g. as a consequence of pooling trials across different experimental conditions or as a consequence of learning processes. Again, the consistency
of spike-LFP phases is, on average, decreased by mixing distributions with varying mean spike-LFP phases.

Thus far, we did not make statements about the statistical dependence or independence between the different $\Theta_{k,m}$’s (the spike-LFP phases). While the assumption of independence between the spike-LFP phases from the same trial is convenient from a statistical perspective (Section 2.1.2), it may be often violated in practice (Section 2.2). Henceforth, we assume that a dependence between spike-LFP phases can only occur for spike-LFP phases from the same trial, i.e. we assume that random variables are statistically independent if they represent spike-LFP phases from separate trials (Assumption 1).

Assumption 1: Spike-LFP phase vectors from separate trials are conditionally independent given observation of the spike count. This assumption implies that for all $(k,m,j,l)$ for which $l \neq m$, the equation

$$E[U_{k,m} \cdot U_{j,l} | N_m, N_l] = E[U_{k,m} | N_m] \cdot E[U_{j,l} | N_l]$$

(2.1)

holds, where $\cdot$ is the dot product operation. From the independence of $N_m$ and $N_l$, it follows that the equation $E[U_{k,m} U_{j,l}] = E[U_{k,m}] \cdot E[U_{j,l}]$ holds as well. The justification of Assumption 1 is that dependencies will mostly act on short timescales, and that spike train statistics will typically be reset by an inter-trial interval. Assumption 1 does not require that the LFPs are not phase-synchronized across trials. If the LFP attains the same phase offset relative to a trial event in every trial, then the spike-LFP phase vectors from different trials can still be independent of each other, e.g. if the spike trains are generated by a homogeneous Poisson process, or if the inter-trial interval duration is sufficiently long. Nevertheless, phase-synchronization of the LFPs across trials may well cause Assumption 1 to be violated if the inter-trial interval is short relative to the oscillatory cycle duration. In that case, non-Poissonian history effects such as autorhythmicity can cause dependencies between spike-LFP phases across trials. The violation of Assumption 1 has the same consequences as a violation of Assumption 2 (which will be introduced in Section 2.1.2), namely a systematic increase or decrease in the PPC (Section 2.2).

2.1.2 Direct application of circular statistics to spike-LFP phase data

It is common practice in neurophysiology to directly compute circular statistics over the concatenated data vector of spike-LFP phases. Whether this leads to valid conclusions about the relationship between the LFP and spiking activity remains to be demonstrated. Suppose that the following two additional assumptions hold however:

Assumption 2: Spike-LFP phase vectors from the same trial are conditionally independent given observation of the spike count. This assumption implies that for all $(k,j,m)$ for which $k \neq j$ the equality

$$E[U_{k,m} \cdot U_{j,m} | N_m] = E[U_{k,m} | N_m] \cdot E[U_{j,m} | N_m]$$

(2.2)

holds. In other words, this assumption holds that, given observation of the spike count $N_m$, the spike-LFP phase vectors from the $m$th trial are independent, i.e. the observation of one spike-LFP phase vector does not change the expected value of another spike-LFP phase vector from the same trial. The conditional independence (given $N_m$) is important
here: If spike-LFP phase and spike count are statistically dependent, then the unconditional independence between the spike-LFP phase vectors from the same trial does not necessarily hold given Assumption 2. The reason for this is that the observation of a certain spike-LFP phase may change the expected value of $N_m$, which could in turn change the expected value of the other spike-LFP phase vectors. Unconditional independence will only be attained by making the additional assumption of independence between spike count and spike-LFP phase (Assumption 3, see below). Example cases where Assumption 2 is violated, e.g. as a consequence of non-Poissonian history effects, will be presented in Section 2.2.

Assumption 3: Independence between spike count and spike-LFP phase. This assumption holds that the expected value of $U_{k,m}$ is independent of $N_m$, for $k \leq N_m$, i.e. that the expected value of the spike-LFP phase vector $U_{k,m}$ is independent of the number of spikes that were observed in the $m$th trial. Example cases where this assumption is violated will be presented in Section 2.3.

Based on the $\Theta_{k,m}$’s (spike-LFP phases) and $N_m$’s (spike counts), we can now define several statistics that estimate the consistency of spike-LFP phases, e.g. the resultant vector length (Eq. 2.5). In order to define the population parameters corresponding to these estimators, we will now define random variables that are identically distributed to the $U_{k,m}$’s. Let $N^{(1)}$ and $N^{(2)}$ be two random variables independently and identically distributed, such that for any $n_m \in \mathbb{N}$, $\Pr(N^{(1)} = n_m) = \Pr(N^{(2)} = n_m) = \Pr(N_m = n_m \mid N_m > 0)$. Let $\Theta$ be a random variable and $U^{(1)}$ and $U^{(2)}$ two statistically independent random vectors identically distributed to the vector $(\cos(\Theta), \sin(\Theta))$, such that the equation $\mathbb{E}\{U^{(1)} \mid N^{(1)} = n_m\} = \mathbb{E}\{U^{(2)} \mid N^{(2)} = n_m\} = \mathbb{E}\{U_{k,m} \mid N_m = n_m\}$ holds for $1 \leq k \leq n_m$. Given Assumptions 1, 2 and 3, it follows that for any $(N^{(1)}, N^{(2)}) \in \mathbb{N}^2$, the equation

$$\mathbb{E}\{U^{(1)} \cdot U^{(2)} \mid N^{(1)}, N^{(2)}\} = \mathbb{E}\{U^{(1)}\} \cdot \mathbb{E}\{U^{(2)}\} \quad (2.3)$$

holds.

We now proceed with defining conventional statistics quantifying the consistency of spike-LFP phases, relying on Assumptions 1, 2 and 3. A natural statistic for the consistency of the spike-LFP phases is provided by the circular resultant length, which is defined, on the population level, as

$$R \equiv \left| \mathbb{E}\{\exp(i\Theta)\} \right| . \quad (2.4)$$

Its direct estimator is defined as

$$\hat{R} \equiv \left| \frac{1}{N} \sum_{m \in M} \sum_{k=1}^{N_m} \exp(i\Theta_{k,m}) \right| . \quad (2.5)$$

This statistic is sometimes also referred to as the (point-field) PLV (‘phase locking value’). Given our Assumptions 1, 2 and 3, it follows that the estimator $\hat{R}$ is strongly biased by sample size, i.e. approaches $R$ from above. This is a problem when comparing neurons or experimental conditions with a small, varying and uncontrollable number of spikes (Vinck et al., 2010b). Therefore, the PPC estimator (Vinck et al., 2010b), which is graphically illustrated in Figure 2.3A, is our preferred statistic at the descriptive level, and for performing simple parametric inferential statistics at the group level, and is defined as
\[ P_0 \equiv \frac{\left( \sum_{m \in M} \sum_{l \in M} \sum_{k=1}^{N_m} \sum_{j=1}^{N_l} U_{k,m} \cdot U_{j,l} \right) - N}{N(N-1)}. \] (2.6)

The corresponding population parameter of the PPC is defined as
\[ P \equiv \mathbb{E}\{ U^{(1)} \cdot U^{(2)} \}. \] (2.7)

The quadruple sum in Eq. 2.6 computes the dot product for all combinations of two spike-LFP phase vectors \((U_{k,m}, U_{j,l})\) for a given pair of trials \((m \text{ and } l)\), and repeats this for all pairs of trials \((m \in M \text{ and } l \in M)\), such that the dot product is computed for all \(N(N-1)\) combinations of two spike-LFP phase vectors. The numerator term \(-N\) accounts for the fact that the quadruple sum in Eq. 2.6 contains \(N\) dot products of phase vectors for which \((k, m) = (j, l)\). The dot product of a spike-LFP phase vector with itself (the diagonal entries in Figure 2.3A) is always one, and does not contribute to the \(P_0\) statistic.

The PPC can be interpreted as quantifying the average coincidence between any two spikes in the LFP phase domain. Two spikes are maximally coincident in the LFP phase domain if their relative spike-LFP phase equals zero, and they are maximally non-coincident if their relative spike-LFP phase equals \(\pm \pi\). The dot product measures to what extent two spikes (with observed phases \(\theta_{k,m}\) and \(\theta_{j,l}\)) are coincident in the LFP phase domain, by outputting \(\cos(\theta_{k,m} - \theta_{j,l})\). For example, the dot product takes on the value \(-1\) if two spikes are maximally non-coincident, and \(+1\) if two spikes are maximally coincident. The choice of the dot product function instead of the absolute angular distance (i.e., relative phase) as a measure of coincidence is motivated by the relationship \(P = R^2\) between the PPC and the resultant vector length (Vinck et al., 2010b).

The expected value of \(P_0\) is obtained by taking its expected value conditional on \(N\), which is required because the \(N_m\)'s are random variables themselves, and subsequently taking the expected value over all possible realizations of \(N\). Assume that \(\Pr\{\ N > 1 \} = 1\), because the PPC is not defined for \(N < 2\), because we need at least one pair of spike-LFP phase vectors to compute the PPC. Using Assumptions 1 and 2, it follows that
\[
\mathbb{E}\{ \hat{P}_0 | N = n \} = \frac{\sum_{m \in M} \sum_{l \in M} n_m n_l \mathbb{E}\{ U_{k,m} \cdot U_{j,l} | N_m, N_l \} - n}{n(n-1)} = \frac{\sum_{m \in M} \sum_{l \in M} n_m n_l \mathbb{E}\{ U^{(1)} \cdot U^{(2)} | N^{(1)} = n_m, N^{(2)} = n_l \} - n}{n(n-1)}. \] (2.8)

Assumptions 1 and 2 allow us to replace the \(U_{k,m}\) and \(U_{j,l}\) terms by the independent and identically distributed random variables \(U^{(1)}\) and \(U^{(2)}\), such that the term \(n_m n_l\) appears in the numerator. By using Assumption 3 and Eq. 2.3, Eq. 2.8 simplifies to
\[
\mathbb{E}\{ \hat{P}_0 | N = n \} = \mathbb{E}\{ U^{(1)} \cdot U^{(2)} \} \equiv P. \] (2.9)
Eq. 2.9 shows that $E\{\hat{P}_0|N = n\}$ does not depend on $n$. Hence, it follows that $E\{\hat{P}_0\} = E[E\{\hat{P}_0|N = n\}] = P$, showing that, given Assumptions 1, 2 and 3, $\hat{P}_0$ is an unbiased estimator of $P$. Note that while $0 \leq P \leq 1$, we have $-1 \leq \hat{P}_0 \leq 1$.

If we make Assumptions 1, 2 and 3, then $P > 0$ indicates that a neuron has a preferred LFP phase of firing, i.e. that spiking probability is modulated with LFP phase. To test whether the inequality $P > 0$ holds, one could for example use the Rayleigh test, or test whether the distribution of $\hat{P}_0$'s across neurons differs from zero. As will be shown in Section 2.2 and 2.3, the validity of these statistical inferences is based on Assumptions 1, 2 and 3 as laid out above.

2.2 The problem of statistical dependence between spike-LFP phases

The assumption that observed spike-LFP phases can be modeled as outcomes of independent random variables (Assumption 2) is typically violated. One cause of this violation is that neurons typically do not behave as homogenous or inhomogeneous Poisson sources (Shinomoto et al., 2009), because neurons have several physiological properties, e.g. refractoriness, autorhythmicity and burstiness, that cause non-Poissonian history effects (Section 2.2.2). Another reason for dependencies between spike-LFP phases to occur is that the LFP may not cover all phases uniformly in a given trial (Section 2.2.2). If we ignore that Assumption 2 is violated and directly apply circular statistics to the observed vector of spike-LFP phases as in Eqs. 2.5 and 2.6, then we may draw invalid conclusions about the relationship between spiking activity and the LFP, especially if the number of trials is small.

The problem posed by dependencies between spike-LFP phases is that the dot product term $U_{k,m} \cdot U_{j,l}$ in Eq. 2.6 may be biased towards producing higher or lower values if $l = m$, i.e. if spike-LFP phases fall in the same trial, in comparison to the case $l \neq m$, i.e. if spike-LFP phases fall in separate trials.

If we make Assumption 1, but not Assumption 2, then we can conclude that the statistical dependence between spike-LFP phases may especially have a large impact on point-field phase synchronization measures if the number of trials is relatively small, because the number of independent spike-LFP phases scales with $N(N-1)/2$, whereas the number of possibly dependent spike-LFP phases scales only with approximately $M^{\binom{N}{2}}$. For example, suppose that all spike-LFP phases within the same trial are statistically dependent, and that the number of spikes per trial equals $N/M = 40$. For one, two, and ten trials, the fraction of dependent spike-LFP phases equals 1, 0.5 and 0.1 respectively, and it tends to zero as $M \to \infty$. The regime in which $M$ is small can be of great importance to neurophysiology, e.g. when studying the involvement of rhythmic neuronal synchronization in memory and learning processes (e.g., van Wingerden et al. (2010a); Hansen & Dragoi (2011); Tort et al. (2009); Popescu et al. (2009)).

2.2.1 Solution to the dependence problem

The goal of this section is to provide a solution to the problem that spike-LFP phases from the same trial can be statistically dependent. We solve the problem of statistical dependence be-
between spike-LFP phases as follows. Instead of computing the dot product of spike-LFP phase vectors across all combinations of two spikes (Fig 2.3A), we compute it only for combinations of spikes that were recorded from separate trials (Fig 2.3B). Let \( \mathcal{L} \equiv \{ l : l \in \mathcal{M}, l \neq m \} \). We define the point-field PPC \( \hat{P}_1 \) estimator of \( P \) as
\[ \hat{P}_1 = \frac{\sum_{m \in M} \sum_{l \in L} \sum_{k=1}^{N_m} \sum_{j=1}^{N_l} U_{k,m} \cdot U_{j,l}}{\sum_{m \in M} \sum_{l \in L} N_m N_l}. \] (2.10)

Note the similarity between Eq. 2.6 and Eq. 2.10, whereas they differ in that the quadruple sum now runs only over all pairs of separate trials \((m \in M\) and \(l \in L\)). If we make Assumption 3 and the assumption that Pr\(|M| < 2\) = 0, then the conditional expected value of this estimator equals

\[ E\{\hat{P}_1|N = n\} = \frac{\sum_{m \in M} \sum_{l \in L} \sum_{k=1}^{n_m} \sum_{j=1}^{n_l} E\{U_{k,m} \cdot U_{j,l} | N_m = n_m, N_l = n_l\}}{\sum_{m \in M} \sum_{l \in L} n_m n_l} \]

\[ = \frac{\sum_{m \in M} \sum_{l \in L} n_m n_l E\{U^{(1)} \cdot U^{(2)} | N^{(1)} = n_m, N^{(2)} = n_l\}}{\sum_{m \in M} \sum_{l \in L} n_m n_l}. \] (2.11)

Assumption 1 allows to replace the \(U_{k,m}\) and \(U_{j,l}\) terms by the independent and identically distributed random variables \(U^{(1)}\) and \(U^{(2)}\), such that the term \(n_m n_l\) appears in the numerator. The product \(n_m n_l\) acts as a weighting factor, assigning heavier weight to expected values of spike-LFP phase vectors for trials with many spikes. However, if Assumption 3 holds, then this weighing factor falls away by division, and equation

\[ E\{\hat{P}_1|N = n\} = \frac{\sum_{m \in M} \sum_{l \in L} n_m n_l E\{U^{(1)} \cdot U^{(2)}\}}{\sum_{m \in M} \sum_{l \in L} n_m n_l} \]

\[ = E\{U^{(1)} \cdot U^{(2)}\} \equiv P \] (2.12)

holds. Thus, the \(\hat{P}_1\) statistic is not affected by a dependence between the spike-LFP phases (note that the derivation of Eqs. 2.10 and 2.11 does not make use of Assumption 2). This is our first main result, improving the point-field PPC \(\hat{P}_0\) measure. We will show in Section 2.4 that the PPC \(\hat{P}_1\) measure allows to derive an analytical relationship between point-field (Figure 2.1) and spike train to field phase-synchronization measures (Figure 2.2).
It should be noted, however, that the $\hat{P}_1$ estimate cannot be obtained in case only a single trial is available. This problem may to some extent be circumvented by removing pairs of spike-LFP phase vectors on the basis of a time-difference threshold, i.e. by rejecting combinations of spikes if they lie too closely together in time. The same approach, i.e. rejection on the basis of a time-threshold, is possible for data-sets that are not divided into trial and inter-trial interval periods. Finally, it should be noted that the $\hat{P}_1$ estimator has a higher variance than the $\hat{P}_0$ estimator if Assumption 2 does hold, especially if the number of trials is small, because the computation of the $\hat{P}_0$ statistic is based on more combinations of two spike-LFP phase vectors than the computation of the $\hat{P}_1$ statistic.

### 2.2.2 Simulations of the effect of a statistical dependence between spike-LFP phases on the $\hat{P}_0$ and $\hat{P}_1$ estimators

**Effect of refractoriness and burstiness**

To study the effect of a statistical dependence between spike-LFP phases on point-field phase synchronization statistics, we performed simulations with the following model. Initially we assumed that neuronal firing was described as a Poisson process, with the instantaneous rate parameter defined by $\lambda(\phi(t)) = \beta/2\pi$, where the LFP-phase was defined as a linear function of time, i.e. $\phi(t) = 2\pi ft$, such that there was no modulation of spiking probability with LFP phase. The probability of observing a spike in the interval $[t, t + \delta]$ was equal to the product $\delta \lambda(\phi(t))$ for small $\delta$. Thus, spiking probability was a linear function of the parameter $\beta$, which equaled the expected number of spikes per oscillatory cycle. We assumed that the trial duration was of length $k/f$ for integer $k$. In this model, spike-LFP phases can be considered as outcomes of independent random variables: The observation of a spike at a particular time-point does not change the probability density of firing at other time-points. Correspondingly, the observation of a spike at a particular LFP phase does not alter the probability density of firing at other LFP phases, and the point-field PPC $\hat{P}_0$ would be an unbiased estimator.

To demonstrate the effect of a refractory period with simulations, we set $\lambda(\phi(t)) = 0$ for the duration of the refractory period after a spike was fired. We chose $f = 20$ Hz, a time-window that comprised 500 time-steps of 0.0001 s (such that $k = 1$) and $\beta = 100$. Without a refractory period, the number of spikes equaled on average $\beta \cdot 500 \cdot 0.0001 = 5$ per trial. We then examined two lengths of the refractory period: 8 (Figure 2.4A) and 40 ms (Figure 2.4B). Figure 2.4A-B show the expected $\hat{P}_0$ value (solid line) as a function of the number of trials. Figure 2.4A shows that in the presence of a refractory period of 8 ms, the expected $\hat{P}_0$ value attained negative values and approached its asymptotic value from below. This result can be explained as follows: If the refractory period is short in comparison to the oscillatory cycle period, then the observation of a particular spike-LFP phase makes the observation of similar spike-LFP phases less likely, causing the expected value of the dot product of spike-LFP phase vectors from the same trial to decrease. However, when the refractory period was relatively long in comparison to the oscillatory cycle length, then $\hat{P}_0$ approached its asymptotic value from above (Figure 2.4B). The latter result can be explained as follows: If the duration of the refractory period approaches the duration of the oscillatory cycle period, then the observation of a spike-LFP phase makes the observation of similar spike-LFP phases on average more likely, causing the expected value of the dot product of spike-LFP phase vectors from the same trial to increase. In contrast to the point-field $\hat{P}_0$ measure, we did not...
observe any dependence of the $\hat{P}_1$ measure on the number of trials. Thus, the $\hat{P}_1$ measure correctly indicated, on expectation, that there was no modulation of spiking probability with LFP phase.

To demonstrate the effect of burstiness we performed the following simulation. Instead of adding a refractory period to the Poisson model, we now added for every spike fired an additional spike with the same phase (i.e., duplicating all spike-LFP phases). Figure 2.4C shows that a strong negative dependence of the expected $\hat{P}_0$ value on the number of trials emerged, such that $\hat{P}_0$ approached its asymptotic population value from above. Even though the probability of firing was equal for all LFP phases, the expected $\hat{P}_0$ value still attained positive values for a finite number of trials. Again, in contrast to the point-field $\hat{P}_0$ measure, we did not observe any dependence of the $\hat{P}_1$ measure on the number of trials. Thus, the $\hat{P}_1$ measure correctly indicated, on expectation, that there was no modulation of spiking probability with LFP phase. The effect of bursting on the PPC may be explained as follows. The PPC quantifies coincidences of spikes in the LFP phase domain. Spikes from the same neuron are expected to be coincident in the LFP phase domain if the probability of spiking is modulated by LFP phase. Bursting can cause spikes from the same trial to have a certain amount of coincidence in the time domain. Because the LFP phase is an approximately linear function of time (e.g., see SI Figure 1 in Canolty et al. (2010)), this time-domain coincidence can cause spikes from the same trial to become coincident in the spike-LFP phase domain as well, even though spiking probability is not modulated with LFP phase.

**Effect of non-uniform sampling of the LFP phases**

Dependencies between spike-LFP phases can also occur because of statistical dependencies between instantaneous LFP phases within the same trial, i.e. can be a consequence of properties of the LFP signal instead of the spike train signal. For example, a dependence between instantaneous LFP phases occurs if the window length that is used to compute a point-field phase synchronization measure does not equal an integer multiple of $1/f$, causing the instantaneous LFP phases to be non-uniformly covered across time. Consequently, given the observation of a particular instantaneous LFP phase at a certain time-point, regardless of whether a spike was fired there, we will be more likely to observe similar instantaneous LFP phases at other time-points within the same trial. This is exemplified by the simulation that was performed for Figure 2.4D. We set the trial duration to $k/f = 0.01$ s (in 100 steps of 0.0001 s) with $k = 0.2$, $f = 20$, $\beta = 100$, as above, but without further non-Poissonian history effects. The $\hat{P}_0$ statistic was negatively biased by the number of trials (Figure 2.4D), despite the fact that spike trains were generated according to a homogeneous Poisson process. In contrast to the point-field $\hat{P}_0$ measure, we did not observe any dependence of the $\hat{P}_1$ measure on the number of trials, and the $\hat{P}_1$ measure correctly indicated that there was no modulation of spiking probability with LFP phase.

Taken together, the results from Section 2.2.2 show that dependencies between spike-LFP phases within the same trial can arise because of many factors. There are various consequences: (i) The equation $E\{\hat{P}_0\} \geq 0$ does not hold in general. (ii) It is not recommendable to use the Rayleigh test to assess whether spiking probability is modulated with LFP phase because the test assumes that phases are outcomes of independent random variables. (iii) The equation $E\{\hat{P}_0\} > 0$ can hold even in the absence of a modulation of spiking probability with LFP phase. (iv) The estimator $\hat{P}_0$ is unbiased by the number of spikes only if the spike-LFP
phases can be treated as outcomes of independent random variables. However, in practice, they usually cannot, and as a result $\hat{P}_0$ can be both positively and negatively biased by the number of trials. Because the number of spikes per trial determines the ratio of the number of possibly dependent over the number of independent spike-LFP phases, it affects the $\hat{P}_0$ estimator as well. In opposition to our previous claim (Vinck et al., 2010b), we conclude that the $\hat{P}_0$ statistic is not an unbiased estimator if Assumption 2 is violated.

2.3 Effect of dependence between spike count and phase

We dealt with the violation of Assumption 2 by developing the point-field PPC $\hat{P}_1$ measure. In order for this measure to be unbiased as a function of the number of trials, we need to make Assumption 3. However, in practice this assumption may be violated, which can introduce a positive or negative bias as a function of the number of trials.

There are two possible relationships between the spike count and the distribution of spike-LFP phases. First, the consistency of spike-LFP phases and the spike count can be statistically dependent. For example, the strength of V1 gamma-band synchronization has been found to be positively dependent on stimulus size, while the V1 firing rate has been found to be negatively dependent on stimulus size (Gieselmann & Thiele, 2008). To give an example of the effect of a negative dependence between spike-LFP phase consistency and firing rate on the point-field PPC $\hat{P}_1$, we studied the model

$$\lambda(\phi_m(t)) = \beta_m g(\phi_m(t); \mu, \kappa),$$

where $g$ is the von Mises density, $\kappa = 0.9$, $\mu = 0$, and $\beta_m$ was a uniformly distributed random variable on the interval [200, 600]. Spiking probability was a linear function of the parameter $\beta_m$, which equaled the expected number of spikes per oscillatory cycle. The LFP phase was defined as

$$\phi_m(t) = 2\pi f t + 2\pi \epsilon(t) ((\beta_m - 200) / 400)^2,$$

where $\epsilon(t)$ was a random variable uniformly distributed on the interval [0, 1]. For large $\beta_m$, a large error term was added to the LFP phase. Thus, small amplitude noise was added to spike-LFP phases when a trial contained a small number of spikes (i.e., $\beta_m$ small), and large amplitude noise was added when a trial contained a large number of spikes (i.e., $\beta_m$ large), thereby implementing a negative correlation between the firing rate and phase locking as in the V1 data (Gieselmann & Thiele, 2008) mentioned above. Figure 2.5 shows that in this model, the $\hat{P}_1$ measure was positively biased by the number of trials.

Second, it can occur that the mean spike-LFP phase co-varies with the spike count. For example, in case of pooling trials across experimental conditions (e.g. orientations), it can occur that trials with higher firing rates correspond to an advanced spike gamma phase (Vinck et al., 2010a), which can bias the $\hat{P}_1$ measure as a function of the number of trials.

2.3.1 A point-field measure unaffected by rate-phase dependencies

We will now define a statistic that is not affected by phase-spikecount dependencies, and hence not biased by the number of trials. We call the defined statistic the point-field PPC $\hat{P}_2$. 
Figure 2.4: Influence of dependence between spike-LFP phases on point-field phase synchronization measures. (A) Estimated, expected point-field PPC $\hat{P}_0$ (solid), and point-field PPC $\hat{P}_1$ values (dotted) (y-axis) as a function of the number of trials (x-axis), in the presence of an absolute refractory period of 8 ms. The grey portion of the solid $\hat{P}_0$ curve highlights that no $\hat{P}_1$ estimate exists when only one trial is available. (B) Similar to (A), but now in the presence of an absolute refractory period of 40 ms. (C) Influence of burstiness on point-field phase synchronization measures. Estimated expected point-field PPC $\hat{P}_0$ (solid), and point-field PPC $\hat{P}_1$ value (dotted) (y-axis) as a function of the number of trials (x-axis). For every spike fired, we added one extra spike with the same phase. (D) Influence of incomplete coverage of LFP phases on estimated, expected point-field PPC $\hat{P}_0$ (solid), and point-field PPC $\hat{P}_1$ values (dotted) (y-axis) as a function of number of trials (x-axis), with the neuron firing only during 20% of the cycle duration.

measure, and define it as

\[
\hat{P}_2 \equiv \frac{1}{|M|(|M| - 1)} \sum_{m \in M} \sum_{l \in L} \left( \frac{\sum_{k=1}^{N_m} \sum_{j=1}^{N_l} U_{k,m} \cdot U_{j,l}}{N_m N_l} \right). \tag{2.15}
\]
The key difference with the $\hat{P}_1$ measure is that in Eq. 2.15, the sum of dot products across combinations of two spike-LFP phase vectors for a pair of separate trials $(m, l)$ is first averaged, i.e. normalized by the number of combinations $N_m N_l$. The average dot product for a pair of separate trials is then averaged across all pairs of separate trials ($m \in M$ and $l \in L$). Thus, in the computation of $\hat{P}_2$, a trial with few spikes will be given the same weight as a trial with many spikes. Consequently, spike-LFP phases from trials with few spikes are given a stronger weight than spike-LFP phases from trials with many spikes. The rationale underlying the $\hat{P}_2$ measure is the following: Because the number of combinations of two spike-LFP phase vectors equals $N_m N_l$ for a pair of separate trials, both trials are given equal weight in the computation of the average dot product for that pair of trials. As a trivial example, suppose that the first trial contains 20 spikes and the second trial contains 100 spikes, then there are $100 \cdot 20 = 2000$ combinations of spike-LFP phases containing spikes from the first trial, but also 2000 combinations of spike-LFP phases containing spikes from the second trial. For a larger number of trials, the trials with a large number of spikes will contribute to more combinations of spike-LFP phase vectors for the $\hat{P}_1$ statistic. For example if we add a third trial that contains 50 spikes, then the first trial will contribute to $20 \cdot 100 + 20 \cdot 50 = 3000$ combinations, but the second trial will contribute to $100 \cdot 20 + 100 \cdot 50 = 7000$ combinations of two spike-LFP phases. This occurs because the normalization term in Eq. 2.10 ($\sum_{m \in M} \sum_{l \in L} n_m n_l$) is a function of the $n_m$’s itself, such that it is less influenced by a single $n_m$ if the number of trials is large. This explains why in Figure 2.5, the $\hat{P}_1$ measure is negatively dependent on the number of trials.

From Assumption 1, it follows that the expected value of $\hat{P}_2$ conditional on $N$, assuming that $\Pr(|M| < 2) = 0$, equals
\[
E(\hat{P}_2|N = n) = \frac{1}{|\mathcal{M}|(|\mathcal{M}| - 1)} \sum_{m \in \mathcal{M}} \sum_{l \in \mathcal{L}} E \left( \sum_{k=1}^{N_m} \sum_{j=1}^{N_l} U_{km} \cdot U_{jl} \mid n_m, n_l \right) \\
= \sum_{m \in \mathcal{M}} \sum_{l \in \mathcal{L}} E\{U^{(1)} \cdot U^{(2)} \mid N^{(1)} = n_m, N^{(2)} = n_l\} \\
= \sum_{m \in \mathcal{M}} \sum_{l \in \mathcal{L}} \frac{E\{U^{(1)} \cdot U^{(2)} \mid N^{(1)} = n_m, N^{(2)} = n_l\}}{|\mathcal{M}|(|\mathcal{M}| - 1)}.
\]

(2.16)

Note that, in comparison to Eq. 2.10, we removed the weighing factor \(n_m n_l\).

The expected value over all possible outcomes of the spikecounts for which \(|\mathcal{M}| > 1\) is then given as

\[
E[E(\hat{P}_2|N = n)] = \sum_{n_1 = 1}^{\infty} \cdots \sum_{n_n = 1}^{\infty} \prod_{m \in \mathcal{M}} \Pr\{N^{(1)} = n_m\} \sum_{n_1 = 1}^{\infty} \sum_{n_1 = 1}^{\infty} E\{U^{(1)} \cdot U^{(2)} \mid N^{(1)} = n_1, N^{(2)} = n_m\} \Pr\{N^{(1)} = n_m\} \Pr\{N^{(2)} = n_1\} \\
= E\{U^{(1)} \cdot U^{(2)}\} \\
= |E\{U^{(1)}\}|^2 \\
= \hat{P}.
\]

(2.17)

The results of the simulation that are presented in Figure 2.5 show that, indeed, phase-spikecount dependence does not lead to a bias of \(\hat{P}_2\) on the number of trials.

Note that

\[
E[U^{(1)}] = \sum_{n_m = 1}^{\infty} E[U^{(1)} \mid N^{(1)} = n_m] \Pr\{N^{(1)} = n_m\},
\]

i.e. \(E[U^{(1)}]\) corresponds to the expected spike-LFP phase vector, and is not affected by a potential dependence between spike count and spike-LFP phase. If Assumption 2 holds, then \(E(\hat{P}_1) = E(\hat{P}_2)\). Also, for \(|\mathcal{M}| = 2\), the equality \(E(\hat{P}_1) = E(\hat{P}_2)\) holds. However, as \(|\mathcal{M}| \to \infty\), \(\hat{P} \to |E[Z]|^2\) where the resultant vector \(E[Z]\) equals

\[
E[Z] = \sum_{n_m = 1}^{\infty} \Pr\{N^{(1)} = n_m\} n_m E[U^{(1)} \mid N^{(1)} = n_m] \\
= \frac{\sum_{n_m = 1}^{\infty} \Pr\{N^{(1)} = n_m\} n_m \sum_{n_m = 1}^{\infty} \Pr\{N^{(1)} = n_m\} n_m}{\sum_{n_m = 1}^{\infty} \Pr\{N^{(1)} = n_m\} n_m \Pr\{N^{(1)} = n_m\}}.
\]

(2.19)
2.4 Relationship between the point-field PPC $\hat{P}_1$ and the spike train to field PPC

2.4.1 Definition of the spike train to field locking indices

We will now demonstrate a relationship between on the one hand the point-field $\hat{P}_1$ and $\hat{P}_2$ measures, and spike train-LFP phase synchronization measures on the other hand. Suppose we observe $M$ trials of LFP and spike train signals. Again, let $\Theta_{k,m}$ be the relative phase between a concurrent LFP segment and an individual spike at a particular frequency, with the spike-LFP phase now defined relative to the concurrent LFP segment (Figure 2.2), i.e., not in a centered fashion as in Section 2.1.

For all $m \in M$, the cross-spectrum between spike train and LFP in a given trial is then defined as

$$C_m \equiv A_m \sum_{k=1}^{N_m} \exp(i\Theta_{k,m}), \quad (2.20)$$

where $A_m$ is defined as the magnitude of the Fourier spectrum of the $m$th LFP signal at frequency $f$ and the spikes are assigned unit energy at frequency $f$. The single-trial resultant length over the individual spike-LFP phases is defined as

$$R_m \equiv \frac{\left|\sum_{k=1}^{N_m} X_{k,m}\right|}{N_m}, \quad (2.21)$$

where $X_{k,m} \equiv \exp(i\Theta_{k,m})$. Note that $0 \leq R_m \leq 1$. The power of the spike train (at a given frequency) equals $(N_m R_m)^2$. If at a given frequency all the spikes are in phase, then the power of the spike train equals $N_m^2$. The single-trial circular mean spike-LFP phase is defined as

$$\bar{\Theta}_m \equiv \arg \left(\sum_{k=1}^{N_m} X_{k,m}\right), \quad (2.22)$$

and the variables $N_m, R_m$ and $\bar{\Theta}_m$ are related by the equation

$$\sum_{k=1}^{N_m} X_{k,m} \equiv R_m N_m \bar{X}_m, \quad (2.23)$$

where $\bar{X}_m \equiv \exp(i\bar{\Theta}_m)$. Finally, we define the phase vector $V_m \equiv (\Re\{\bar{X}_m\}, \Im\{\bar{X}_m\})$, which is depicted in Figure 2.2. Note that $\bar{\Theta}_m$ is the relative phase between spike train and LFP for the $m$th trial. We define the spike train to field PLV as the resultant length across spike train to field phases, i.e. $\left|M\right|^{-1} \sum_{m \in M} \bar{X}_m$. The spike train to field PLV is a quantity closely related to the ‘spike-field coherence’ (e.g., see Jarvis & Mitra (2001)), except for the fact that the former ignores amplitude co-variations and assigns an equal weight to every trial that contains at least one spike, while the latter weighs the trials according to the product of LFP and spike train spectral magnitude, i.e. $A_m N_m R_m$ (at a given frequency), and is a concurrent measure of amplitude co-variations. We may quantify the consistency of relative
phases across trials by directly using the PPC, which removes the bias of the spike train to field PLV, by defining

\[ \hat{S}_2 = \frac{1}{|M|(|M|-1)} \sum_{m \in M} \sum_{l \in L} V_m \cdot V_l, \]  

(2.24)

and we refer to this measure as the uncorrected spike train to field PPC \( \hat{S}_2 \). By normalizing by \(|M|(|M|-1)\), the measure takes into account that no predictions about the LFP phase can be made if a trial does not contain spikes. Let \( \Phi \) be a random variable identically distributed to \( \bar{\Theta}_m \) for all \( m \in M \). The statistic \( \hat{S}_2 \) is an unbiased estimator of its corresponding population value, which is defined as \( S_2 = |E\{\exp(i\Phi)\}|^2 \) (Vinck et al., 2010b).

It may occur that neuronal output is highly predictive of LFP phase, but that a neuron fires only in a small number of trials, such that the predictive value of the spike train across all trials is much lower than indicated by the \( \hat{S}_2 \) measure. To accommodate this, and to assess to what extent spike trains predict the LFP phase across all trials, we define the measure

\[ \hat{S}_{2}^{(\star)} = \frac{1}{M(M-1)} \sum_{m \in M} \sum_{l \in L} V_m \cdot V_l. \]  

(2.25)

Note that because \( |M| \leq M \), the inequality \( \hat{S}_{2}^{(\star)} \leq \hat{S}_2 \) holds.

In Eq. 2.24, all \(|M|\) relative phases are assigned the same weight. However, the pairwise statistic may be turned into a weighted pairwise statistic, similar to the weighted phase-lag index in Vinck et al. (2011), defining

\[ \hat{S}^{(w)} = \frac{\sum_{m \in M} \sum_{l \in L} W_m W_l (V_m \cdot V_l)}{\sum_{m \in M} \sum_{l \in L} W_m W_l}. \]  

(2.26)

If we define \( W_m = R_m N_m A_m \), then the expected value of the numerator equals the numerator of the squared spike-field coherence (e.g. see Jarvis & Mitra (2001)). A special case of weighing is obtained when we let \( W_m = R_m N_m \), i.e., the spectral magnitude of the spike train at a particular frequency, defining

\[ \hat{S}_1 = \frac{\sum_{m \in M} \sum_{l \in L} R_m R_l N_m N_l (V_m \cdot V_l)}{\sum_{m \in M} \sum_{l \in L} R_m R_l N_m N_l}. \]  

(2.27)

We refer to this measure as the uncorrected spike train to field PPC \( \hat{S}_1 \). The rationale underlying this weighted measure is that every individual spike-LFP phase is given an equal weight by weighing by \( N_m N_l \), as will be shown below. Further, we will show that the weighting factor \( R_m R_l \) can be used to correct for the number of spikes per trial.
2.4.2 Relationship between point-field PPC and spike train to field PPC

We now derive a relationship between the point-field $\hat{P}_1$ measure and the $\hat{S}_1$ measure. We observe that the equations

\[
\hat{P}_1 \equiv \frac{\sum_{m \in M} \sum_{l \in L} \sum_{k=1}^{N_m} \sum_{j=1}^{N_l} U_{k,m} \cdot U_{j,l}}{\sum_{m \in M} \sum_{l \in L} N_m N_l}
\]

\[
= \frac{\sum_{m \in M} \sum_{l \in L} \left( \sum_{k=1}^{N_m} \sin(\Theta_{k,m}) \sum_{j=1}^{N_l} \sin(\Theta_{j,l}) + \sum_{k=1}^{N_m} \cos(\Theta_{k,m}) \sum_{j=1}^{N_l} \cos(\Theta_{j,l}) \right) N_m N_l}{\sum_{m \in M} \sum_{l \in L} N_m N_l}
\]

\[
= \frac{\sum_{m \in M} \sum_{l \in L} R_m R_l N_m N_l V_m \cdot V_l}{\sum_{m \in M} \sum_{l \in L} N_m N_l}
\]

(2.28)

hold.

The final expression in Eq. 2.28 is nearly identical to the right-hand expression in Eq. 2.27. We define the corrected (using the abbreviation ‘corr.’) spike train to field PPC $\hat{S}_{1 \text{corr}}$ as

\[
\hat{S}_{1 \text{corr}} \equiv \frac{\sum_{m \in M} \sum_{l \in L} N_m R_m N_l R_l}{\sum_{m \in M} \sum_{l \in L} N_m N_l}.
\]

(2.29)

Note that this measure is, similar to the $\hat{P}_1$ statistic, not unbiased as a function of the number of trials, unless we make the assumption that the single spike-LFP phases are independent of the spike count (Assumption 3). Similarly, we define the corrected spike train to field PPC $\hat{S}_{2 \text{corr}}$ as

\[
\hat{S}_{2 \text{corr}} = \frac{1}{|M| (|M| - 1)} \sum_{m \in M} \sum_{l \in L} R_m R_l V_m \cdot V_l.
\]

(2.30)

Similar to the $\hat{P}_2$ measure, this measure is not biased by the number of trials, even if Assumptions 2 and 3 are violated. The main difference between the $\hat{S}_{1 \text{corr}}$ and the $\hat{S}_{2 \text{corr}}$ measure is that the former assigns an equal weight to every spike-LFP phase (i.e., $\Theta_{k,m}$), by weighing the influence of a trial by $N_m N_l$, whereas the latter assigns an equal weight to every phase obtained for a trial (i.e., $\Theta_m$), thereby weighing those spikes heavier that occurred in trials with few spikes.

If we define the individual spike-LFP phases in an identical way for $\hat{P}_1$ and $\hat{S}_{1 \text{corr}}$, then the equality $\hat{P}_1 = \hat{S}_{1 \text{corr}}$ holds and the measures have mathematically equivalent definitions.
Thus, we conclude that there exists a simple mathematical relationship between point-field PPC $\hat{P}_1$ and the uncorrected spike train to field PPC $\hat{S}_1$. Because the point-field $\hat{P}_1$ and $\hat{P}_2$ measures are unaffected by the number of spikes per trial, we conclude that the same property holds for the measures $\hat{S}_1^{\text{corr}}$ and $\hat{S}_2^{\text{corr}}$ respectively, solving the problem that their respective uncorrected counterparts can strongly depend on the number of spikes. We may thus regard $\hat{S}_1^{\text{corr}}$ and $\hat{S}_2^{\text{corr}}$ as measures of both the consistency of individual spike-LFP phases and of the relative phases between spike trains and LFP signals across trials, with a correction factor for the number of spikes per trial.

The definitions of $\hat{P}_1$ and $\hat{P}_2$ are equivalent to the definitions of their respective spike train to field counterparts $\hat{S}_1^{\text{corr}}$ and $\hat{S}_2^{\text{corr}}$, expect for the fact that the individual spike-LFP phases are not computed in the same way. We predict that, as a consequence of this difference in phase estimation, the inequalities $\hat{P}_1 > \hat{S}_1^{\text{corr}}$ and $\hat{P}_2 > \hat{S}_2^{\text{corr}}$ typically hold if $\Pr\{N_m > 1\}$ is small. This prediction has the following rationale: For spike train to field phase-synchronization measures, spikes can fall at arbitrary positions relative to the concurrent LFP segment (Figure 2.2), such that they sometimes fall close to the borders of the concurrent LFP segment. On the other hand, for point-field phase-synchronization measures, they fall per definition at the center of the LFP segment that is used to compute the spike-LFP phase (Figure 2.1). The instantaneous LFP phase is not a strict linear function of time, but decorrelates over time, because neuronal (e.g. gamma or theta) oscillations contain energy over a relatively large range of frequencies. We assume that spiking probability at a given time-point is especially modulated by the local fluctuations in the LFP around that time-point, but not by distant past or future fluctuations in the LFP. The decorrelation of LFP phase across time will then cause a decrease in the magnitude of the side-lobes of the spike-triggered LFP average as a function of time (e.g. see Figure 1H-I in Fries et al. (2001b)). If a spike falls at the border of an LFP segment, then its spike-LFP phase is determined predominantly from either past or future LFP fluctuations, and will be computed based on a larger portion of the LFP signal to which the spikes are not locked. For the point-field measure, the spike-LFP phase will be computed based on a larger portion of the LFP signal to which the spikes are locked. To counter the decorrelation in LFP phases across time, the LFP segment could be decreased in duration. However, this would decrease the frequency resolution, and thereby increase the amount of spectral leakage.

2.4.3 Dependence of spike train to field PPC on the mean and variance of the spike count

From the properties of the resultant length $R_m$, it follows that $\hat{S}_1$ and $\hat{S}_2$ are statistically dependent on $N$. The equation $0 \leq R_m \leq 1$ holds, and $R_m$ is positively biased by $N_m$, i.e. approaches its asymptotic value from above (e.g. see (Vinck et al., 2010b)). Thus, given a fixed distribution of individual spike-LFP phases, it follows that the expected values of $\hat{S}_1$ and $\hat{S}_2$ are positively dependent on the expected value of $N_m$, consistent with previous results on the spike-field coherence (Zeitler et al., 2006; Lepage et al., 2011). There are two cases in which $\hat{S}_1$ and $\hat{S}_1^{\text{corr}}$ are approximately equal. First, if for all $m \in \{1, \ldots, M\}$, $\Pr\{N_m > 1\} = 0$, then $\hat{S}_1 = \hat{S}_1^{\text{corr}}$ and $\hat{S}_2 = \hat{S}_2^{\text{corr}}$. Second, if the consistency of the individual spike-LFP phases is very high, then $R_m$ is close to 1, and $\hat{S}_1 \approx \hat{S}_1^{\text{corr}}$. However, if the consistency of the phases is very low and the number of spikes per trial is large, then $\hat{S}_1$ will be small, yet much larger
Figure 2.6: Comparison of uncorrected and corrected spike train to field PPC $\hat{S}_1$. (A) $Y$-axis: corrected spike train to field PPC $\hat{S}_1^{\text{corr}}$ (dotted) and uncorrected spike train to field PPC $\hat{S}_1$ (solid). $X$-axis: number of spikes per trial. Different lines represent different coupling strength, with the von Mises dispersion parameter $\kappa \in \{0.1, 0.5, 1, 20\}$. (B) Empirically observed and analytically predicted correction factor for spike train to field PPC $\hat{S}_1$ ($y$-axis) as a function of the number of spikes per trial ($x$-axis), again for $\kappa \in \{0.1, 0.5, 1, 20\}$

than $\hat{S}_1^{\text{corr}}$ (Figure 2.6).

These relationships can be intuitively explained as follows. The statistics $\hat{S}_1$ and $\hat{S}_2$ measure the consistency of relative phases across trials. For a given trial, the computation of a relative phase is based on averaging across multiple single spike-LFP phases. The average of multiple single spike-LFP phases is more consistent than the single spike-LFP phases themselves, because averaging will remove variance across the single spike-LFP phases. If the single spike-LFP phases are only weakly consistent, then the average over many single
Figure 2.7: Influence of firing rate variance on spike train to field phase synchronization measures. Dashed: spike train to field PPC $\hat{S}_1$ for Poisson distribution minus spike train to field PPC $\hat{S}_1$ for zero-variance distribution, as a function of the number of spikes per trial. Positive values indicate a larger $\hat{S}_1$ value for Poisson distribution. Solid: as for dashed, but now for spike train to field PPC $\hat{S}_2$. Dashed-dotted: as solid, but now for spike train to field PPC $\hat{S}^*_2$.

spike-LFP phases will be much more consistent than the single spike-LFP phases themselves. However, the expression $R_mR_l$ is a function not only of $N_m$ and $N_l$, but also of the statistical dependence between spike-LFP phases. If spike-LFP phases become, ceteris paribus, more positively (negatively) dependent, then the expected value of $R_m$ will increase (decrease), and consequently the expected values of $\hat{S}_1$ and $\hat{S}_2$ will decrease (increase). Thus, the expected values of $\hat{S}_1$ and $\hat{S}_2$ depend on three main factors, namely the number of spike-LFP phases per trial, the statistical dependence of the individual spike-LFP phases, and the consistency of the individual spike-LFP phases. The $\hat{S}^\text{corr}_1$ and $\hat{S}^\text{corr}_2$ measures remove the dependence on the first two factors and offer a purer measure of the last factor.

To demonstrate the relationship between the expected value of $\hat{S}_1$ and the number of spikes by simulations, we drew samples of spike-LFP phases from a von Mises distribution with dispersion parameter $k \in \{0.1, 0.5, 1, 20\}$. The number of spikes per trial $N_m$ was fixed, such that for all $m$, $N_m \in \{1, \ldots, 100\}$. The number of trials $M$ equaled 100. Figure 2.6A shows the estimated expected values of $\hat{S}^\text{corr}_1$ (dotted) and $\hat{S}_1$ (solid) as a function of the number of spikes per trial. The $\hat{S}^\text{corr}_1$ measure attained values independently of $N_m$. For $N_m = 1$, the equality $\hat{S}^\text{corr}_1 = \hat{S}_1$ held. As $N_m$ increased, $\hat{S}_1$ strongly increased. The ratio $\hat{S}^\text{corr}_1/\hat{S}_1$ indeed had the form of the correction factor in Eq. 2.29 (Figure 2.6B). Note that for large expected values of $\hat{S}^\text{corr}_1$, $\hat{S}_1$ rapidly approached one, while it rose linearly with the number of spikes per trial for a small expected value of $\hat{S}^\text{corr}_1$. This linear relationship implies that an $x\%$ increase in firing rate, without a concurrent change in the distribution of single spike-LFP phases, entails an $x\%$ increase in uncorrected spike train to field PPC $\hat{S}_1$ for low expected values of $\hat{S}_1$ and a relatively small number of spikes per trial.

If our aim is to compare $\hat{S}_1$ or $\hat{S}_2$ between two experimental conditions that have a different distribution of spike counts, then it is insufficient that the means of the spike counts
are equal. As we will now show, other parameters of the distribution of spike counts, especially the variance, should be equal as well, and these parameters may very well differ between conditions. For example, the variance of spike counts can be a function of attention (e.g. Mitchell et al. (2007)). To demonstrate the influence of the spike count variance (across trials), we compared two conditions with an equal average spike count, but with a different variance of the spike counts. Phase samples were drawn from a von Mises distribution with dispersion parameter $\kappa = 1$, and we set the number of trials $M = 100$. In the first condition, the variance of the counts equaled zero (zero-variance condition). In the second condition, the spike count was distributed according to a Poisson distribution, i.e. the spike count variance equaled the mean spike count (Poisson condition). Figure 2.7 shows the difference in $\hat{S}_1$, $\hat{S}_2$, and $\hat{S}_{2}^{\star}$ between the Poisson and the zero-variance condition, as a function of the expected spike count per trial, $E\{N_m\}$. The $\hat{S}_1$ measure attained higher values for the Poisson distribution than for the zero-variance condition. The cause of this difference is that the weights $N_m$ in Eq. 2.27 give a stronger weight to trials with more spikes, and that the average spike-LFP phase obtained from these trials is more reliable. The $\hat{S}_2$ measure on the other hand showed two opposite trends. Because trials with zero spikes (i.e., $N_m = 0$) are left out in Eq. 2.24, the effective firing rate was higher for the Poisson condition, which increased the value of $\hat{S}_2$ for the Poisson condition for small $E\{N_m\}$. However, the results of Fig 2.6A imply that the consistency of $\Theta_m$ does not depend linearly on $N_m$, but is a concave function of $N_m$. This concavity led to a relative decrease in the value of $\hat{S}_2$ for the Poisson condition, an effect which became dominant as $E\{N_m\} \geq 5$. For the $\hat{S}_{2}^{\star}$ measure, only the latter effect was present, causing the $\hat{S}_{2}^{\star}$ measure to be higher for the zero-variance condition than for the Poisson condition.

### 2.5 Discussion

#### 2.5.1 Relationship between firing rate and uncorrected spike train to field PPC measures

We demonstrated that the spike train to field PPC measures can be corrected for the positive dependence on the number of spikes per trial. However, correcting for this positive dependence does not preclude the possibility that the consistency of the single spike-LFP phases (i.e., point-field phase-synchronization) is correlated with the firing rate. Still, the latter correlation would be the result of a physiological relationship, and a change in the firing rate could be the consequence of a change in spike-LFP phase-coupling as well. On the other hand, the positive effect of the firing rate on the uncorrected spike train to field PPC measures is a result of a statistical relationship, i.e. higher (lower) spike train to field phase-synchronization appears because a neuron has more (less) spikes per trial.

No a priori statements can be made about the correlation between the consistency of single spike-LFP phases and the firing rate. Independent variation can be caused by subsampling spike trains, or by a failure of one’s measurement or spike-sorting system to detect a spike. A negative correlation between the firing rate and the consistency of spike-LFP phases may be a result of an increase in the amplitude of inputs that are not coherent with the LFP (Zeitler et al., 2006), or it may be a consequence of the inhibitory nature of the rhythm generators (Gieselmann & Thiele, 2008; van Wingerden et al., 2010a; Donoghue et al., 1998; Jensen
& Mazaheri, 2010). In case of such a negative correlation, it can occur that the point-field PPC measures decrease while the increase in firing rate causes a simultaneous increase in the uncorrected spike train to field PPC measures. A positive correlation between point-field phase synchronization and firing rate can be found in visual cortex, where visual stimulation and high contrast stimuli can increase both firing rates and the precision of gamma-band synchronization (Gray et al., 1989; Henrie & Shapley, 2005). Selective visual attention can increase both gamma-band synchronization and firing rate (Fries et al., 2001b, 2008; Gregoriou et al., 2009) in V4 and FEF (but not in V1 (Chalk et al., 2010)). If the firing rate is positively correlated with the consistency of individual spike-LFP phases, then both the point-field PPC measures and the uncorrected spike train to field PPC measures increase, although the proportional increase in the value of the latter will typically be larger.

### 2.5.2 Comparison between single-unit field and multi-unit field phase synchronization measures

Spike-field phase synchronization measures can be based either on single unit activity (SUA), or multi unit activity (MUA). Because MUA contains more spikes per second than SUA, and because spiking activity of different neurons is partially driven by uncorrelated inputs which renders their respective spike-LFP phases more independent, MUA-field measures will in general attain higher uncorrected spike train to field PPC values, as shown by Zeitler et al. (2006).

This relative gain of the MUA is not present for the point-field PPC and corrected spike train to field PPC measures however. In fact, we predict that the point-field PPC measures typically attain lower values for MUA-field pairs than SUA-field pairs: Different neuronal cell types, e.g. fast spiking basket cells and pyramidal cells, fire preferentially at different LFP phases (Klausberger et al., 2003; Tukker et al., 2007; Csicsvari et al., 2003; Hasenstaub et al., 2005; Okun & Lampl, 2008). Furthermore, the preferred LFP-phase of spiking of neurons in the hippocampus and visual cortex is input-dependent (Vinck et al., 2010a; O’Keefe & Recce, 1993). This divergence of preferred LFP-phases, a factor that was ignored in Zeitler et al. (2006), tends to broaden the distribution of individual spike-LFP phases for the MUA. The effect of linearly mixing spiking activity from different cells is similar to linearly mixing trials across experimental conditions with different mean spike-LFP phases (Section 2.2). Thus, MUA-field point-field synchronization will typically be weaker than SUA-field point-field synchronization.

### 2.5.3 Comparison of point-field PPC $\hat{P}_0$ and PPC $\hat{P}_1$

In Section 2.5.3, 2.5.4 and 2.5.5, we will discuss the advantages and disadvantages of the developed measures, whose properties are summarized in Table 2.1. The $\hat{P}_1$ estimator (Eq. 2.10) solves the problem that the $\hat{P}_0$ measure can be affected by dependencies between spike-LFP phases from the same trial. We believe that the effect of dependencies between spike-LFP phases on the $\hat{P}_0$ measure should be considered undesirable, and that the $\hat{P}_1$ measure is therefore to be preferred. It should be emphasized however that e.g. refractoriness or burstiness can cause a true change in the consistency of spike-LFP phases, which should be detected by a point-field phase synchronization measure. However, besides changing the consistency
of spike-LFP phases, refractoriness and burstiness also introduce a statistical dependence between spike-LFP phases from the same trial, an effect which should be distinguished from the change in the distribution of spike-LFP phases. For a large number of trials, \( \hat{P}_0 \) will be exclusively determined by the modulation of spiking probability with LFP phase. For a small number of trials, \( \hat{P}_0 \) can also determined by the dependencies between spike-LFP phases. For example, spiking probability may not be modulated with LFP phase, however burstiness can cause \( \mathbb{E}[\hat{P}_0] \) to exceed zero. The question that we are interested in is whether neurons tend to fire consistently at particular phases of the LFP oscillation, i.e., whether the ongoing phase of the LFP can be predicted from the observation of a spike. However, a dependence between spike-LFP phases will affect the \( \hat{P}_0 \) measure even if we replace the LFP by a white noise signal that bears no relationship to neuronal firing. The availability of a phase consistency measure that alleviates the dependency from firing rate and history-effects goes in the same direction as the derivation of statistically independent measure of single cell spiking intensity and cell-cell interaction in the information geometry paradigm (Nakahara et al., 2006).

<table>
<thead>
<tr>
<th>Point-field PLV (( \hat{R} ))</th>
<th>Dependent on number of spikes?</th>
<th>Effect of dependence between spike-LFP phases?</th>
<th>Effect of dependence between spikecount and phase?</th>
<th>Biased by number trials?</th>
<th>Relationship with other statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>As ( M \rightarrow \infty, ) ( \hat{R}^2 \rightarrow \hat{P}_0 )</td>
<td></td>
</tr>
<tr>
<td>Point-field PPC ( \hat{P}_0 )</td>
<td>0</td>
<td>±</td>
<td>±</td>
<td>As ( M \rightarrow \infty, ) ( \hat{P}_0 \rightarrow \hat{P}_1 )</td>
<td></td>
</tr>
<tr>
<td>Point-field PPC ( \hat{P}_1 )</td>
<td>0</td>
<td>0</td>
<td>±</td>
<td>If spike count and spike-LFP phase are independent, then ( \hat{P}_1 = \hat{P}_2 ) and ( \text{Var}[\hat{P}_2] \geq \text{Var}[\hat{P}_1] )</td>
<td></td>
</tr>
<tr>
<td>Point-field PPC ( \hat{S}_1 )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Only difference to ( \hat{P}_1 ) lies in phase-estimation</td>
<td></td>
</tr>
<tr>
<td>Point-field PPC ( \hat{S}_2 )</td>
<td>0</td>
<td>0</td>
<td>±</td>
<td>Only difference to ( \hat{P}_2 ) lies in phase-estimation</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1: Summary of performance different phase synchronization indices. ‘±’: means that effect can be positive or negative. ‘0’: means that statistic is not affected.

### 2.5.4 Comparison of point-field PPC \( \hat{P}_1 \) and PPC \( \hat{P}_2 \)

The \( \hat{P}_2 \) (Eq. 2.15) estimator differs from the \( \hat{P}_1 \) measure in that the former cannot be affected by dependencies between the spike-LFP phase and the spike count. Whether or not \( \hat{P}_2 \) is a
better measure than $\hat{P}_1$ depends on the properties of the data, and our measurement goals. If there is no relationship between the spike count and the spike-LFP phase distribution, then the $\hat{P}_1$ estimator is to be preferred above the $\hat{P}_2$ estimator, because the former has a smaller estimator variance. Whether spike count and spike-LFP phase are indeed dependent may be tested by using the data itself. On the other hand, if there is a relationship between the spike count and the spike-LFP phase distribution, then the $\hat{P}_1$ and $\hat{P}_2$ measures estimate a different population parameter. The question is then, what do $\hat{P}_1$ and $\hat{P}_2$ measure?

The $\hat{P}_1$ measure gives an answer to the following question: To what extent is the probability of observing a spike modulated with LFP phase? This follows from the fact that it simply counts how many spikes fell at a particular phase, and characterizes that distribution. The probability of observing a spike in a small interval $[\phi(t), \phi(t) + \delta]$ is approximately equal to

$$\Pr(s = 1 \mid \phi(t)) \approx \sum_{n_m=1}^{\infty} \Pr(N_m = n_m) f(\phi(t) \mid N = n_m) \left( n_m \frac{\delta}{T} \right), \quad (2.31)$$

where $s$ is a binary variable indicating the occurrence of a spike, $f(\phi)$ is a probability density function modeling the modulation of spiking probability with LFP phase $\phi(t) \in [-\pi, \pi]$. The modulation of spiking probability with LFP phase can then be quantified by the resultant vector length

$$R = \frac{\left| \int_{\phi(t)=-\pi}^{\phi(t)=\pi} \Pr(s = 1 \mid \phi(t)) \exp(i \phi(t)) \, d\phi(t) \right|}{\int_{\phi(t)=-\pi}^{\phi(t)=\pi} \Pr(s = 1 \mid \phi(t)) \, d\phi(t)}$$

$$= \frac{\int_{\phi(t)=-\pi}^{\phi(t)=\pi} \sum_{n_m=1}^{\infty} \Pr(N_m = n_m) f(\phi(t) \mid n_m) n_m \exp(i \phi(t)) \, d\phi(t)}{\int_{\phi(t)=-\pi}^{\phi(t)=\pi} \sum_{n_m=1}^{\infty} \Pr(N_m = n_m) f(\phi(t) \mid n_m) n_m \, d\phi(t)}$$

$$= \frac{\sum_{n_m=1}^{\infty} \Pr(N_m = n_m) n_m \int_{\phi(t)=-\pi}^{\phi(t)=\pi} \exp(i \phi(t)) f(\phi(t) \mid n_m) \, d\phi(t)}{\sum_{n_m=1}^{\infty} \Pr(N_m = n_m) n_m} \quad (2.32)$$

If for all $n_m \geq 1$ we let $E[\exp(i \Theta) | N^{(1)} = n_m] = \int_{\phi(t)=-\pi}^{\phi(t)=\pi} \exp(i \phi(t)) f(\phi(t) \mid N_m = n_m) \, d\phi(t)$, then this equation equals Eq. 2.19, i.e. the asymptotic value that the $\hat{P}_1$ measure is estimating.

The question that the $\hat{P}_2$ measure answers is the following: If we know that a spike occurred at a certain time-point, then to what extent can the LFP-phase at the time of spiking be predicted, without further knowledge of the spike count? In this case, the spike has already occurred, and the estimated phase vector is given by Eq. 2.18. If we would have used information about the spike count then the expected spike-LFP phase would have changed if spike count and spike-LFP phase were statistically dependent on each other. Thus, the $\hat{P}_2$ measure tells us to what extent we can predict the instantaneous LFP-phase from the occurrence of a spike, if we ignore the information in the spike count, i.e. if we assume that all spike counts are equally likely.

There are two cautionary remarks about the use of the $\hat{P}_1$ measure however. First, when we mix experimental conditions it should be realized that the average will then be dominated
by the experimental conditions with the most spikes. Second, while the $\hat{P}_1$ measure is not biased by the number of spikes or dependencies between spike-LFP phases, it can still be biased as a function of the number of trials. The fact that the $\hat{P}_1$ measure may in that case have some statistical bias can be dealt with by using e.g. randomization statistics (e.g. Maris et al. (2007)). On the other hand, a statistical dependence between spike count and spike-LFP phase cannot ‘create’ an expected $\hat{P}_1$ value that exceeds zero: If for all trials, firing probability is not modulated with LFP phase, then $E(\hat{P}_1) = 0$.

The $\hat{P}_2$ measure may sometimes detect a relationship between spiking probability and the LFP phase, while the $\hat{P}_1$ measure does not, because the LFP phase preference of a neuron can sometimes become visible only over trials. For example, if a neuron has a firing rate non-stationarity then the modulation of spiking probability with LFP phase may only become apparent over trials.

2.5.5 Comparison of point-field and corrected spike train to field phase synchronization measures

The corrected spike train to field phase synchronization measures solved the problem that the uncorrected spike train to field phase synchronization measures are strongly dependent on the spike count distribution. The only difference between the point-field and corrected spike train to field phase synchronization measures is that the latter define the phase in a non-centered way, whereas the former define spike-LFP phases in a centered way. In fact, both measures are mathematically equivalent if the individual spike-LFP phases are defined in the same manner. In practice, this means that both measures are roughly equivalent if the windows that are used to determine spike-LFP phases are equally long, provided that the LFP phase does not decorrelate on a too rapid time-scale. An advantage of the spike train to field approach is that it is computationally more efficient, because in case of point-field PPC measures, the instantaneous LFP phase needs to be determined for every spike separately.

If the spike-LFP phases are defined in a centered way, then the point-field PPC measures to what extent the occurrence of a spike can be predicted from the instantaneous, ongoing LFP phase, or, vice versa, to what extent the instantaneous LFP phase can be predicted from the observation of a spike. Spike train to field phase-synchronization measures, with the phase defined in a non-centered way, indicate to what extent the phase of an LFP signal at arbitrary moments in time, including those moments in time where no spike was observed, can be predicted from the observation of the concurrent spike train. Because spike train to field phase-synchronization measures are inspired on the classical EEG-EEG coherence approach (by taking the complete spike train signal) we expect them to be more closely related to EEG-EEG measures or the coherence of the membrane potential of a neuron to the LFP.

The measures and estimation procedures described in this paper are made available in the open source FieldTrip toolbox (http://www.ru.nl/neuroimaging/fieldtrip).