Journal of Neuroscience Methods 207 (2012) 113–124

Contents lists available at SciVerse ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Computational Neuroscience

Quantifying circular–linear associations: Hippocampal phase precession

Richard Kempter[a,b,*], Christian Leibold[c,d], György Buzsáki[e], Kamran Diba[e,1], Robert Schmidt[a,b,2]

[a] Institute for Theoretical Biology, Department of Biology, Humboldt-Universität zu Berlin, Invalidenstr. 43, 10115 Berlin, Germany
[b] Bernstein Center for Computational Neuroscience Berlin, Phillipstr. 13, 10115 Berlin, Germany
[c] Division of Neurobiology, LMU Munich, Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany
[d] Bernstein Center for Computational Neuroscience Munich, Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany
[e] Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ 07102, USA

ARTICLE INFO

Article history:
Received 20 December 2010
Received in revised form 9 March 2012
Accepted 20 March 2012

Keywords:
Hippocampus
Place field
Phase precession
CA1
Theta oscillation
Circular–linear regression

ABSTRACT

When a rat crosses the place field of a hippocampal pyramidal cell, this cell typically fires a series of spikes. Spike phases, measured with respect to theta oscillations of the local field potential, on average decrease as a function of the spatial distance traveled. This relation between phase and position of spikes might be a neural basis for encoding and is called phase precession. The degree of association between the circular phase variable and the linear spatial variable is commonly quantified through, however, a linear–linear correlation coefficient where the circular variable is converted to a linear variable by restricting the phase to an arbitrarily chosen range, which may bias the estimated correlation. Here we introduce a new measure to quantify circular–linear associations. This measure leads to a robust estimate of the slope and phase offset of the regression line, and it provides a correlation coefficient for circular–linear data that is a natural analog of Pearson's product-moment correlation coefficient for linear–linear data. Using surrogate data, we show that the new method outperforms the standard linear–linear approach with respect to estimates of the regression line and the correlation, and that the new method is less dependent on noise and sample size. We confirm these findings in a large data set of experimental recordings from hippocampal place cells and theta oscillations, and we discuss remaining problems that are relevant for the analysis and interpretation of phase precession. In summary, we provide a new method for the quantification of circular–linear associations.

1. Introduction

Phase precession is a relational code that is thought to be important for animals to learn a sequence of places, and, in general, phase precession might be a basis for episodic(-like) memory (Skaggs et al., 1996). In the hippocampus, the position of an animal in its environment is encoded through activity of so-called place cells. Their action potentials exhibit a precession of firing phases relative to theta oscillations in the local field potential (4–12 Hz; Buzsáki, 2002), and in successive oscillation cycles, the theta phase of action potentials of a pyramidal cell progressively decreases toward earlier phases (O'Keefe and Recce, 1993). Besides the hippocampus, phase precession also exists in the entorhinal cortex (Hafting et al., 2008; Mizuseki et al., 2009; Reifenstein et al., 2012). The mechanisms underlying the generation of phase precession are unknown despite considerable experimental and theoretical work to unravel its origin (e.g., Skaggs et al., 1996; Jensen and Lisman, 1996; Tsodyks et al., 1996; Wallenstein and Hasselmo, 1997; Kamondi et al., 1998; Ekstrom et al., 2001; Harris et al., 2002; Mehta et al., 2002; Lengyel et al., 2003; Huxter et al., 2003; Hasselmo and Eichenbaum, 2005; Lisman et al., 2005; Zugaro et al., 2005; Dragoi and Buzsáki, 2006; Cheng and Frank, 2008; Thurley et al., 2008; Leibold et al., 2008; Harvey et al., 2009; Geisler et al., 2010; Losonczy et al., 2010).

Basic properties of phase precession are the slope and the offset of a regression line as well as the correlation coefficient between the theta phases and the spatial locations at which spikes occur. These characteristic properties of phase precession allow for a comparison across trials, cells, animals, and species, and also to computational models. A fair comparison, however, requires appropriate methods to measure these properties.

Phase precession describes the relation between a linear variable (usually animal position) and a circular variable (spike theta phase). Nevertheless, this relation is commonly quantified through a linear–linear regression (slope and offset) and a linear–linear

Abbreviations: CA, Cornu Ammonis.

* Corresponding author at: Institute for Theoretical Biology, Department of Biology, Humboldt-Universität zu Berlin, Invalidenstr. 43, D-10115 Berlin, Germany. Tel.: +49 30 2093 8925; fax: +49 30 2093 8801.
E-mail address: r.kempter@biologie.hu-berlin.de (R. Kempter).
[1] Present address: Neural Circuits and Memory Lab, University of Wisconsin-Milwaukee, Department of Psychology, PO Box 413, Milwaukee, WI 53201, USA.
[2] Present address: Department of Psychology, University of Michigan, Ann Arbor, MI 48109, USA.

0165-0270/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2012.03.007

correlation coefficient (e.g., O'Keefe and Recce, 1993; Mehta et al., 2002): the circular variable is simply converted to a linear one by restricting the phase to a predefined range. However, such a restriction is artificial. As we will show, for such linear–linear correlations, the sample estimates of basic properties of phase precession are biased and depend on the sample size.

Phase precession is a specific example of circular–linear data. The goal of linear regression in such a case is predicting the circular variable given the linear variable. A procedure for circular–linear regression and a correlation measure with well-defined properties are not available. What is available, on the one hand, is the classical Pearson product-moment correlation coefficient for linear–linear data, and, on the other hand, correlation coefficients for circular–circular data (Fisher, 1995; Zar, 1999; Jammalamadaka and SenGupta, 2001).

Here we develop a new method to quantify circular–linear associations. We show that this measure is robust with respect to noise and small sample sizes, and only weakly depends on prior knowledge on the data. We derive a correlation coefficient for circular–linear data that is a natural analog of the product-moment correlation coefficient for two linear variables; that is, the two correlation measures share as many features as possible and are identical in limiting cases. The new measure can be used to quantify phase precession as well as other circular–linear data.

2. Materials and methods

Our method is demonstrated on experimental data that has been used in several previous studies (Diba and Buzsáki, 2007, 2008; Schmidt et al., 2009) in which experimental procedures have been described in detail. Briefly, three male Sprague-Dawley rats were trained to run back and forth on a linear track to retrieve water rewards at both ends. All protocols were approved by the Institutional Animal Care and Use Committee of Rutgers University. After learning the task, the rats were implanted with silicon probes in the left dorsal hippocampus under isoflurane anesthesia. The silicon probes were lowered to CA1 and CA3 pyramidal cell layers. For this study, all units and the local field potential were taken from CA1 recording sites. In contrast to the previous studies, here we did not exclude spikes that occurred near reward sites located at platforms at the two ends of the linear track, and we did not exclude cells without phase precession. However, spikes were excluded if the instantaneous running speed was smaller than 10 cm/s. In addition, spikes from single crossings of the place field in which the average running speed was smaller than 10 cm/s were excluded. Using only a running speed criterion to select spikes, place fields that extended into the reward sites entered the analysis while spikes that occurred during rest periods with non-theta states were nevertheless excluded.

3. Results

To quantify phase precession, we first motivate the main idea behind our new procedure and state the essential results in Eqs. (1)–(4). The justification of this approach as well as the derivation of the equations and further details are provided in Appendix A.

To outline the basic recipe for applying the new quantification, let us assume that there are n action potentials and that each action potential is characterized by a phase angle ϕ_j and a linear spatial position x_j for $j = 1, \ldots, n$. We further assume that the association between phase and position can be described by a linear model of the form $\tilde{\phi}_j = 2\pi\, a\, x_j + \phi_0$, where a is the slope (in units of cycles per length scale) and ϕ_0 is the phase-offset. The angle $\tilde{\phi}_j$ can then be considered as the prediction of ϕ_j from the measurement x_j.

The parameters a and ϕ_0 of the regression line that represents the data best are found by minimizing the error between the measured angles $\phi_1, \ldots, \phi_n$ and the predicted angles $\tilde{\phi}_1, \ldots, \tilde{\phi}_n$. To account for the circular nature of phase, we use a circular error measure. The circular distance d between two angles ϕ_j and $\tilde{\phi}_j$ is defined as (Lund, 1999)

$$d(\phi_j, \tilde{\phi}_j) = 2[1 - \cos(\phi_j - \tilde{\phi}_j)].$$

The slope parameter a of the regression line can be found by minimizing the mean circular distance. In Appendix A.1 we prove that this minimization is equivalent to maximizing the mean resultant length R (or vector strength) of the residual distribution (FitzGerald et al., 2001),

$$R = \sqrt{\left[\frac{1}{n}\sum_{j=1}^{n}\cos(\phi_j - 2\pi\, a\, x_j)\right]^2 + \left[\frac{1}{n}\sum_{j=1}^{n}\sin(\phi_j - 2\pi\, a\, x_j)\right]^2}. \tag{1}$$

One should note that R is independent of the phase offset ϕ_0. The maximization of R with respect to the slope a, which usually demands numerical methods, leads us to the estimate of the slope, $\hat{a} = \arg\max_a R$ on a cylinder is never unique, a restriction of the range of possible slopes is necessary and can be implemented through constraining $\hat{a}$ to some interval during the numerical optimization procedure.

To estimate the phase offset $\hat{\phi}_0$, we plug $\hat{a}$ into

$$\hat{\phi}_0 = \arctan^*\frac{\sum_j \sin(\phi_j - 2\pi\,\hat{a}\,x_j)}{\sum_j \cos(\phi_j - 2\pi\,\hat{a}\,x_j)} \tag{2}$$

where the function arctan* is the quadrant-specific inverse of the tangent; see Eq. (A.6).

Finally, we calculate the circular–linear correlation coefficient (Jammalamadaka and SenGupta, 2001)

$$\hat{\rho}_c = \frac{\sum_{j=1}^{n}\sin(\phi_j - \overline{\phi})\,\sin(\theta_j - \overline{\theta})}{\sqrt{\sum_{i=1}^{n}[\sin(\phi_i - \overline{\phi})]^2 \sum_{j=1}^{n}[\sin(\theta_j - \overline{\theta})]^2}} \tag{3}$$

where the important new point is to convert the linear variable x_j into a circular variable $\theta_j = 2\pi\,|\hat{a}|\,x_j \,(\mathrm{mod}\, 2\pi)$ that is scaled by the estimated slope $\hat{a}$; we note that $\overline{\phi}$ and $\overline{\theta}$ are circular sample mean values given by

$$\overline{\phi} = \arctan^*\frac{\sum_j \sin(\phi_j)}{\sum_j \cos(\phi_j)} \quad \text{and} \quad \overline{\theta} = \arctan^*\frac{\sum_j \sin(\theta_j)}{\sum_j \cos(\theta_j)}. \tag{4}$$

For large sample sizes, p-values can be derived, as outlined in Jammalamadaka and SenGupta (2001) and Appendix A.2.

The properties of this new approach to estimate the slope $\hat{a}$, the offset $\hat{\phi}_0$, and the degree of association $\hat{\rho}_c$ between a linear variable X and a circular variable Φ, are illustrated in Fig. 1 using surrogate data. There, first samples from a zero-mean bivariate Gaussian distribution in X and Y are generated at some correlation ρ. Then Y is wrapped around the circle to create a circular variate, that is, $\Phi = Y\,(\mathrm{mod}\, 2\pi)$. As a result, X and Φ can be thought of as data lying on a cylinder. Such surrogate data provide a simple phenomenological model of phase precession.

3.1. Comparison of circular–linear and linear–linear regression

We now quantify the surrogate data in Fig. 1 using the new circular–linear as well as common linear–linear approaches. Note that in the figures angles are plotted in units of degrees (0°, …, 360°, …, 720°) to allow for a better comparability to the literature.

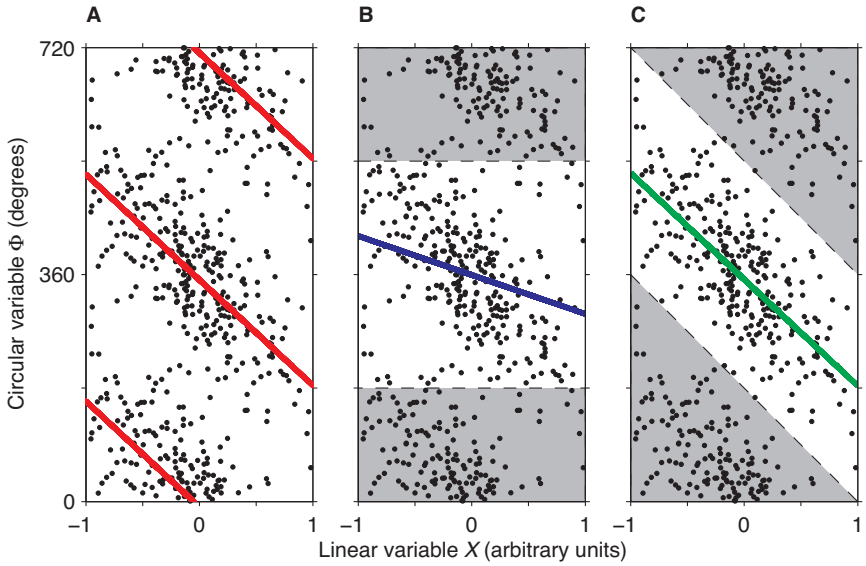

Fig. 1. Three approaches to quantify circular–linear associations. Circular–linear surrogate data (black dots) are zero-mean bivariate Gaussian in X and Y, and the variate Y is wrapped around the circle to generate the circular variate $\Phi = Y \pmod{360^\circ}$. Two cycles of Φ in units of degrees are shown. Colored lines indicate linear regression lines. (A) The circular–linear fit leads to a slope $\hat{a} = -0.46$ (in cycles per unit on the X axis) and an offset (i.e. the phase of the regression line at $X = 0$) $\hat{\phi}_0 = -8.3^\circ$. The interval of possible slopes for the fit was $[-50, 50]$. (B) Linear–linear fit in the selected white region: the data cylinder was cut horizontally at $\phi_{\text{cut}} = 180^\circ \pmod{360^\circ}$ (dashed lines). We find $\hat{a} = -0.17$ and $\hat{\phi}_0 = -0.6^\circ$ through standard linear–linear regression. (C) Linear–linear fit in the selected white region: the data cylinder was cut diagonally at phase $\phi_{\text{cut}} = 180^\circ \pmod{360^\circ}$ and slope $a_{\text{cut}} = -0.50$ (dashed lines). We find $\hat{a} = -0.47$ and $\hat{\phi}_0 = -7.7^\circ$. In (A), (B), and (C), we show the same data sample with $n = 300$ data points in one cycle; the parameters of the wrapped bivariate Gaussian distribution (see Appendix A) are $\rho = -0.8$, $\sigma_X = 0.5$, $\sigma_Y = 2$ rad, $\mu_X = \mu_Y = 0$. To be able to further evaluate the quality of the estimates, we note that the expected slope is -0.51 (from Eq. A.15) and the expected offset is 0°.

3.1.1. Circular–linear approach

In Fig. 1A, the regression line is the result of minimizing the sum of circular distances between data and a linear prediction, as outlined in Eqs. (1) and (2). The regression line matches the data well. To derive these results, no prior knowledge on the data is necessary (parameter-free approach), apart from restricting the possible slopes to a range of reasonable values.

3.1.2. Linear–linear approaches

Two common other approaches to the same problem of circular–linear regression are illustrated in Fig. 1B and C. There, the data cylinder is cut at some line and unwrapped: the circular variable Φ is treated as if it were a linear one. Accordingly, linear–linear least squares regression is performed on data points in the selected white regions.

In Fig. 1B, the data cylinder is cut horizontally at phase ϕ_{cut}, and the cylinder can be reconstructed by joining the two dashed lines. Here, we have chosen $\phi_{\text{cut}} = 180^\circ$, which is the best choice given the underlying zero-mean Gaussian distribution of the surrogate data. However, even under this condition, the regression line in Fig. 1B does not represent the data satisfactorily. The slope is underestimated because points in the lower left and upper right corner of the white region provide a strong bias.

Underestimation of the slope is a general problem with a horizontal phase cut-off, although there is some freedom in choosing ϕ_{cut}: Zugaro et al. (2005) defined it as the circular mean phase of spikes shifted by 180°; in other studies, ϕ_{cut} was defined through maximization of the explained variance of a linear model (Mehta et al., 2002; Foster and Wilson, 2007; Brun et al., 2008; Cheng and Frank, 2008; Hafting et al., 2008; Mizuseki et al., 2009, 2012). This maximization is mostly done using discrete phase steps, for example in 1° steps (Brun et al., 2008; Hafting et al., 2008; Mizuseki et al., 2009, 2012) or in 10° steps (Foster and Wilson, 2007).

Some of the problems created by a horizontal cut of the cylinder might be resolved by a diagonal cut (O'Keefe and Recce, 1993; Huxter et al., 2003, 2008; Dragoi and Buzsáki, 2006; Ego-Stengel and Wilson, 2007; Lenck-Santini and Holmes, 2008), as in Fig. 1C. There, the regression line fits the data well. However, this approach requires fixing two free parameters of the model, namely the phase offset ϕ_{cut} and the slope a_{cut} of the line cutting the cylinder, which were chosen arbitrarily in Fig. 1C. In some studies, the cut parameters were again found by maximization of the explained variance; for example, Lenck-Santini and Holmes (2008) and Huxter et al. (2008) used an iterative approach with a limited number of steps for the unwrapping process.

3.2. Comparison of correlation measures

To verify the strengths and weaknesses of the three approaches outlined in Fig. 1, we next estimate correlations and evaluate the circular–linear and the linear–linear approaches in terms of robustness to noise, dependence on sample size, computational cost, and requirement of prior knowledge on the data.

3.2.1. Circular–linear approach

In our circular–linear approach, a range of slope values a is tested. The best-fitting slope $\hat{a}$ is the one that maximizes the mean resultant length R as defined in Eq. (1). Fig. 2A shows R as a function of a for the data presented in Fig. 1. There is a unique and distinct absolute maximum for a wide range of slopes. Finding the maximum is possible with standard numerical methods because R is continuous and differentiable.

Fig. 2B indicates how the circular–linear correlation coefficient $\hat{\rho}_c$ depends on noise, which is parameterized by the correlation coefficient ρ of the underlying Gaussian model of the surrogate data. The estimates $\hat{\rho}_c$ (dots) for a finite amount of data (here: $n = 300$ for each dot) are represented well by the expected values ρ_c (black solid line). The expected value ρ_c is equal to the estimate $\hat{\rho}_c$ for an infinite number of data points ($n \to \infty$); see also Eq. (A.20) in Appendix A.

How does the *expected* correlation ρ_c depend on the parameter ρ of the underlying Gaussian model? The functions $\rho_c(\rho)$ in Fig. 2B

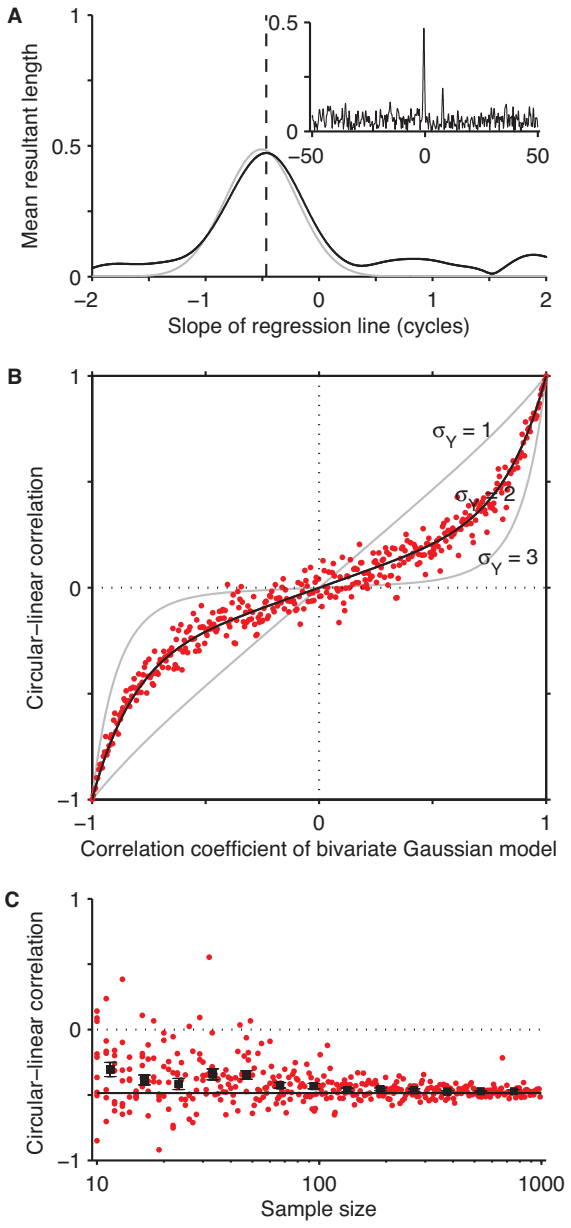

Fig. 2. Circular–linear regression and correlation measure. (A) Mean resultant length R (black line) according to Eq. (1) as a function of the slope a of the regression line for the same surrogate data as in Fig. 1. The vertical dashed line at the maximum of R marks the best slope $\hat{a} = -0.46$ with $R(\hat{a}) = 0.47$. The gray line denotes the expected $\mathcal{R}$ from Eq. (A.14), matching the sample estimate well. The inset indicates sample estimates of R for a wider range of slopes, illustrating that the interval of possible slopes is not critical here. (B) Relation between the actual correlation ρ of the underlying bivariate Gaussian model and the circular–linear correlation. Red dots are sample estimates $\hat{\rho}_c$ for $\sigma_Y = 2$ and $n = 300$, and the solid black line denotes the expected values ρ_c from Eq. (A.20). The two gray lines are expected values ρ_c for $\sigma_Y = 1$, which is close to the identity, and $\sigma_Y = 3$, which is highly nonlinear. (C) Dependence of $\hat{\rho}_c$ (red dots) on the sample size n for $\rho = -0.8$, $\sigma_X = 0.5$, and $\sigma_Y = 2$. The horizontal solid line indicates the expected value $\rho_c = -0.49$. Black dots are mean values of surrounding 30 data points, and bars denote the standard error of the mean. In (B) and (C), dotted lines mark zero correlations. Here, and throughout the rest of the paper, the range of possible slopes was restricted to the interval $[-2, 2]$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(one black and two gray solid lines) are monotonous so that there is a unique and invertible, though nonlinear, mapping between ρ and ρ_c. The degree of nonlinearity of $\rho_c(\rho)$ depends on the variance σ_Y^2 of the Gaussian model. For $\sigma_Y = 1$ (in units of rad), we are close to the identity $\rho = \rho_c$, which becomes exact in the limiting case $\sigma_Y \to 0$. This behavior is one reason why ρ_c is a natural analog of the Pearson product-moment correlation coefficient, although, in general we have $|\rho| > |\rho_c|$. The function $\rho_c(\rho)$ markedly deviates from the identity if the width of the Gaussian exceeds one cycle, which is the case for $\sigma_Y \gtrsim 1.5$.

Finally, in Fig. 2C, the efficiency of $\hat{\rho}_c$ as a function of the sample size n points to a well-behaved estimator. For small sample size, however, there is a systematic underestimation of the expected absolute correlation, and the variability is large.

3.2.2. Linear–linear approaches

Fig. 1B exhibits a common linear–linear approach to quantifying circular–linear data. There, the data cylinder is cut horizontally at some phase ϕ_{cut}. The specific value of the parameter ϕ_{cut} determines the estimated linear correlation $\hat{\rho}$, as indicated in Fig. 3A. Therefore, the free parameter ϕ_{cut} needs to be fixed. Here and in many other applications (e.g., Mehta et al., 2002; Foster and Wilson, 2007; Brun et al., 2008; Hafting et al., 2008; Mizuseki et al., 2009), ϕ_{cut} is defined as the phase that maximizes $|\hat{\rho}|$.

There are several weaknesses associated with this approach. First, finding extrema of $\hat{\rho}$ is numerically inconvenient because $\hat{\rho}(\phi_{\text{cut}})$ is a piecewise constant and discontinuous function, and efficient standard algorithms for finding extrema, which rely on gradients, fail. Note that the function $\hat{\rho}(\phi_{\text{cut}})$ is discontinuous at the measured phases ϕ_j for $j = 1, \ldots, n$, producing up to n discontinuities, resulting in a rough graph as in Fig. 3A. In particular, adding or removing data points might lead to large fluctuations of the maximum of $|\hat{\rho}|$ and the resulting ϕ_{cut}. Moreover, the estimated correlation $\hat{\rho}(\phi_{\text{cut}})$ in Fig. 3A not only has a minimum but also a maximum (see also Hafting et al., 2008; their Supplemental Fig. 10a), adding another difficulty for robustly identifying the maximal modulus $|\hat{\rho}|$.

Another weakness of the linear–linear approach is illustrated in Fig. 3B, which indicates how the estimated correlation coefficient $\hat{\rho}$ depends on the correlation parameter ρ of the underlying Gaussian model of the surrogate data. The estimates $\hat{\rho}(\rho)$ are mostly far from the identity. This behavior is most striking for the case $\rho = \pm 1$, where $\hat{\rho}$ underestimates ρ; and this behavior is also obvious for the case $\rho = 0$, where $\hat{\rho}$ overestimates ρ: there are no values of $\hat{\rho}$ around zero, even if ρ is small, because we have maximized $|\hat{\rho}|$ with respect to ϕ_{cut}. Moreover, the sign of $\hat{\rho}$ can be estimated wrongly, even if $|\rho|$ is large. This is because for data similar in structure to the one shown in Fig. 1, there is typically a distinct minimum and also a distinct maximum of $\hat{\rho}(\phi_{\text{cut}})$ with similar absolute values; see also Fig. 3A. This phenomenon was also described for phase precession in entorhinal grid cells (Hafting et al., 2008; Reifenstein et al., 2012).

Finally, Fig. 3C depicts the strong dependence of $\hat{\rho}$ on the sample size n. This behavior of $\hat{\rho}$ illustrates that the resulting estimator is problematic. In Fig. 3C we, again, recognize a bimodal distribution with positive values of the estimate $\hat{\rho}$ even though $\rho = -0.8$ is negative.

The various problems identified in the approach outlined in Figs. 1B and 3 with a horizontal cut of the data cylinder might be circumvented by a diagonal cut, as indicated in Fig. 1C and evaluated in Fig. 4; this approach might be superior because the unwrapped cylinder represents the presented data better. However, as we now show, diagonally cutting the cylinder to estimate a linear–linear correlation coefficient is even more questionable than cutting horizontally. The main problem arises when we attempt to fix the two free parameters of the cutting line, that is, its slope a_{cut} and its offset ϕ_{cut}, which are similar in nature to the result of a circular–linear regression itself.

Let us illustrate the problem of fixing a_{cut} and ϕ_{cut} for the example from Fig. 1C. We set $\phi_{\text{cut}} = 180^\circ$, which is the best choice given the underlying zero-mean Gaussian model, and then try to find

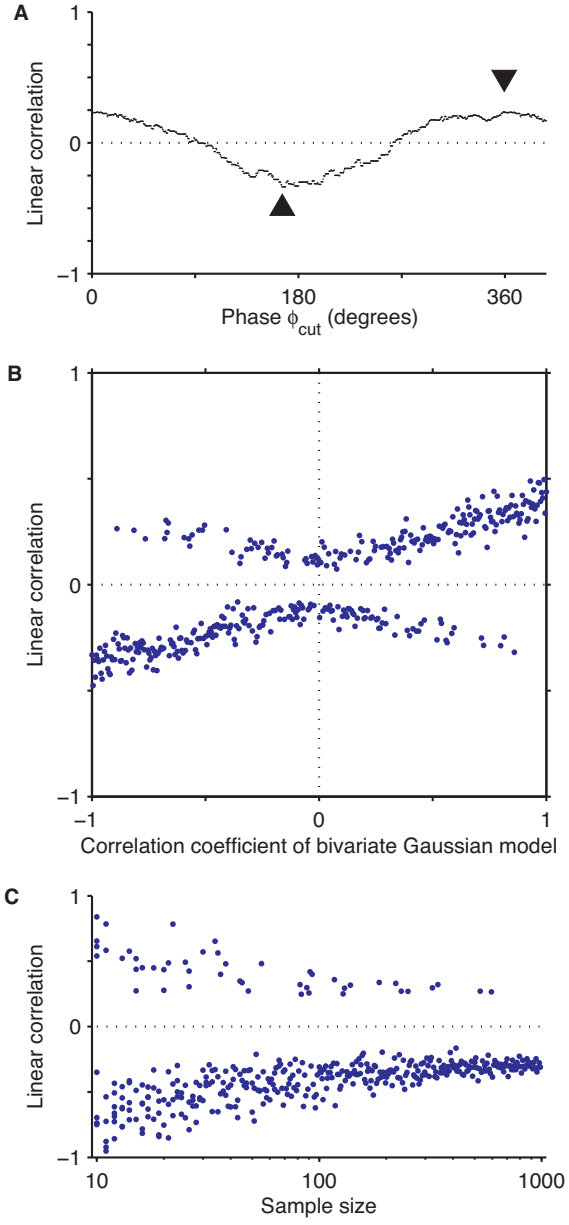

Fig. 3. Linear–linear regression after horizontally cutting the data cylinder, as in Fig. 1B. (A) Linear correlation $\hat{\rho}$ as a function of the phase ϕ_{cut} at which the cylinder is cut, for the same surrogate data as in Fig. 1B ($n = 300$ and $\rho = -0.8$). The upward triangle indicates the minimum $\hat{\rho}_c = -0.34$, which is also the maximum of $|\hat{\rho}|$. The downward triangle indicates the maximum of $\hat{\rho}$; interestingly, the maximum is positive. (B) Dependence of the linear correlation coefficient $\hat{\rho}$ (blue dots) on the parameter ρ of the underlying bivariate Gaussian model of the surrogate data ($n = 300$, $\sigma_X = 0.5$, $\sigma_Y = 2$). There are no values of $\hat{\rho}$ around zero, even if ρ is small. Note that the estimates often have the wrong sign, e.g., the blue dots with $\rho > 0.5$ but $\hat{\rho} < 0$. (C) Dependence of $\hat{\rho}$ (blue dots) on the sample size n for $\rho = -0.8$, $\sigma_X = 0.5$, and $\sigma_Y = 2$. In (A), (B), and (C), dotted lines mark zero correlations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

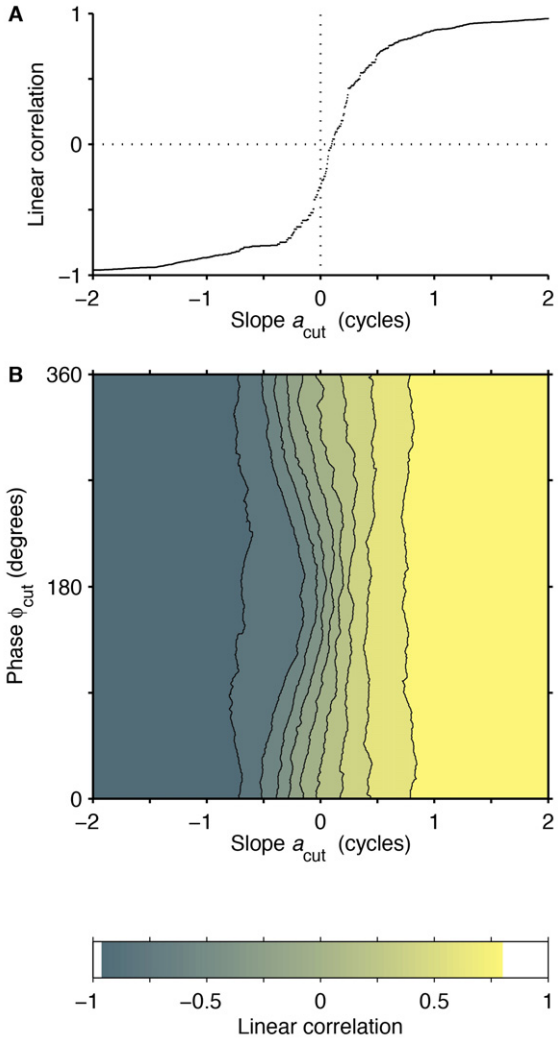

Fig. 4. Linear–linear regression after diagonally cutting the data cylinder as indicated in Fig. 1C. (A) Estimated linear correlation $\hat{\rho}$ as a function of the slope a_{cut}, for the same fixed offset $\phi_{cut} = 180^\circ$ as in Fig. 1C. Dotted lines mark zero values. (B) Sample correlation $\hat{\rho}$ (color coded) as a function of a_{cut} and ϕ_{cut}. There is no distinct extremum. For (A) and (B), the surrogate data shown in Fig. 1 were used. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

an a_{cut} that maximizes $|\hat{\rho}|$. The function $\hat{\rho}(a_{cut})$ is, again, piecewise constant and discontinuous (Fig. 4A). Beyond that, however, $\hat{\rho}(a_{cut})$ does not exhibit a clear extremum, not even near the expected value of -0.51 (from Eq. A.15) of the slope of a regression line. Conversely, extrema are reached in the limiting cases $a_{cut} \to \pm\infty$ with values $\hat{\rho} \to \pm 1$.

To further outline why the approach of diagonally cutting the cylinder is ill-posed, we vary ϕ_{cut} and a_{cut} simultaneously. Fig. 4B demonstrates that in the plane spanned by a_{cut} and ϕ_{cut} there is no clear extremum of $\hat{\rho}$. This example does not rule out that there are other data that might give rise to a clear maximum of $|\hat{\rho}|$ as a function of a_{cut} and ϕ_{cut}. However, without profound prior knowledge on the data we cannot rule out that the approach is non-well behaved, as in our example. We therefore will not elaborate on this approach in more detail and summarize that the choice of ϕ_{cut} and a_{cut} heavily biases the estimated linear correlation coefficient.

Next we also want to exclude the feasibility of combinations of circular–linear regression (defining a line at which the data cylinder is cut) and a subsequent estimate via a product-moment correlation coefficient. Let us show the weakness of such a combined approach for a specific example in which circular and linear variables are independent and the circular variable is uniformly distributed over the whole cycle. We then expect a correlation coefficient at zero, but the circular–linear regression line can have arbitrary slopes. Cutting the data cylinder at a slope that is different from zero, however, gives rise to nonzero linear–linear correlation coefficients. The larger the slope, the larger is the correlation, which is in opposition to the expected zero correlation. A zero correlation is, however, retained when we use the circular–linear correlation ρ_c because

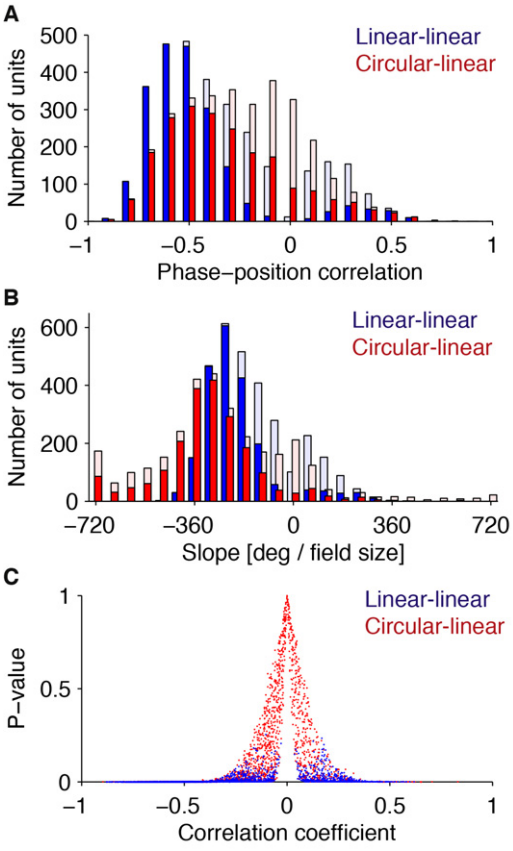

Fig. 5. Hippocampal phase precession: comparison of linear–linear (blue color) and circular–linear (red color) correlation and slope estimates for $n = 3097$ hippocampal units. (A) Histogram of phase-position correlations. For each CA1 unit we calculated the correlation of animal position and theta phase of all spikes. The linear–linear approach yields overall stronger correlations, almost no correlation coefficients around zero, and a distinct peak at positive correlations. The circular–linear approach yields a skewed unimodal distribution with a peak around −0.5. Bars filled with pale colors indicate non-significant correlation coefficients ($p > 0.0001$; Pearson's test is used for linear–linear correlation, the test for circular–linear correlation is described in Appendix A.2.1). (B) Histogram of slopes of the regression line of phase-position data. Pale bars denote slopes that did not yield a significant correlation coefficient, which often have a slope around zero. The circular–linear approach yields steeper (more negative) slopes than the linear–linear approach, in line with results obtained from the surrogate data. (C) Significance (p value) as a function of correlation coefficient. Note the absence of high p values for the linear–linear approach. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

this measure is insensitive to the slope. We thus conclude that the strategy with diagonally cutting the cylinder is, in general, inadequate.

3.3. Quantifying hippocampal phase precession

Hippocampal phase precession describes the association between a linear spatial variable and a circular phase variable. For the quantification of the strength of the association, so far, mostly variations of linear–linear approaches were used (Fig. 1B and C). Here we compare the relation between the linear–linear approach with a horizontal cut (Fig. 1B) to our new circular–linear approach (Fig. 1A) for phase precession in $n = 3097$ hippocampal CA1 units.

The experimental data shown in Fig. 5 exhibits many features seen in the surrogate data. For example, the linear–linear approach yields stronger correlation coefficients because of the maximization of the coefficient of variation (Fig. 5A). Furthermore, the linear–linear approach results in shallower slopes than the circular–linear approach (Fig. 5B). Most importantly, however, the circular–linear correlation coefficient much better indicates significance than the linear–linear coefficient (Fig. 5C); in fact, according to the linear–linear method, only few trials show insignificant correlation. These results indicate that the estimates of the standard linear–linear approach can be biased.

A more detailed analysis of the relation between the two approaches is provided in Fig. 6; in particular, we show scatter plots of phase-position correlations (top left panel) and slopes (top right panel). The first bottom row gives phase-position data from eight representative example units with strong phase precession pooled over multiple trials.

Apart from cells with strong phase precession, the experimental data also includes cells without phase precession. Furthermore, in some units phase precession is visible but not correctly identified by one or both quantification methods. Example units for these cases are shown in the bottom row of Fig. 6.

Overall, the two approaches (circular–linear vs. linear–linear) clearly lead to different results. For example, Fig. 5A and the top left panel of Fig. 6 indicate that there are few linear–linear correlation coefficients near zero, which is expected for a maximization of the square of this coefficient. Moreover, the top left panel of Fig. 6 shows that for a given linear–linear correlation coefficient there is a wide range of circular–linear ones, even with the opposite sign (examples c, e, f, g, and h). The absolute values of the circular–linear coefficients tend to be smaller than those of the linear–linear coefficients, which is reflected in the triangular distribution of dots. This effect could also be explained by the maximization of the square of the linear–linear coefficient (Figs. 2C and 3C).

Also for the estimated slopes in Fig. 5B and in the top right panel of Fig. 6, we observe a gap for small values of the linear–linear slopes. For a given value of the linear–linear slope, Fig. 6 shows that there is also a large range of circular–linear slopes, but the absolute values of the former tend to be smaller (see examples in the top row of Fig. 6 and also c, g, and h in the bottom row). This effect could be explained by the underestimation of slopes by the linear–linear method as indicated in Fig. 1B.

An inspection of the phase-position plots of the example neurons in Fig. 6(a–h) reveals that the circular–linear slopes generally fit the data clouds better than the linear–linear slopes. However, for few units, the sign of the circular–linear correlation coefficient is different from that of the corresponding slope (examples b, c, d, g, and h). This requires a subtle distinction between slope and coefficient regarding their interpretation. Whereas the coefficient is, strictly speaking, only a measure of significance, the slope indicates the direction of the phase drift, independently of significance. For most examples, the sign of correlation coefficient nevertheless indicates this direction correctly.

To conclude, the linear–linear approach tends to overestimate the quality of phase precession and underestimate its slope. The circular–linear approach is better suited for estimating the slope and is more conservative in estimating the correlation. In both approaches, problems in the quantification can arise for nonlinear or bimodal phase precession and for multiple fields.

4. Discussion

We developed a new measure to quantify circular–linear associations. This approach required three parts: first, we fit a linear regression model to circular–linear data by minimizing the circular error between measured and predicted angles. Second, the resulting slope of the regression line was used to scale the linear variable and to transform it into an appropriate circular one. Third, this transformation enabled us to utilize a measure for circular–circular correlations.

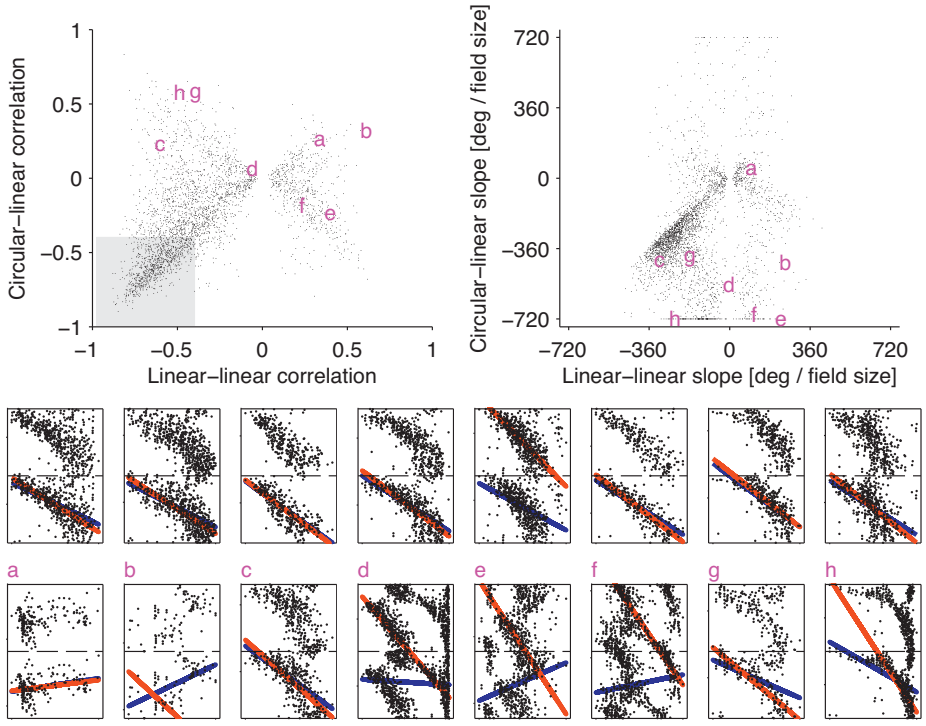

Fig. 6. Quantification of hippocampal phase precession. *Top left*, scatter plot of phase-position correlation coefficients (circular–linear vs. linear–linear with a horizontal cut) for $n = 3097$ CA1 units. The labels a–h refer to examples in the bottom row of the figure. The gray area marks units with strong phase precession, which are pyramidal neurons that are place cells; examples from that area are shown in the first row below. *Top right*, scatter plot of slopes (circular–linear vs. linear–linear). *Bottom*, example units. In every bottom panel, a black dot corresponds to a spike at a specific theta phase (ordinate, two cycles) and a specific spatial position within the place field (abscissa, normalized to field width). Regression lines indicate the circular–linear method (red color) as well as the linear–linear method (blue color) with an optimized horizontal cut of the cylinder (dashed line), where optimization means maximization of the coefficient of determination. (a–h) Eight specific example units; two of them are without phase precession (i.e. *re*cession in examples a and b), and six with phase precession that is not correctly identified by at least one quantification method (c–h). The latter six include cells that suggest phase precession (negative slope) by visual inspection but yield a positive circular–linear correlation (c, g, and h) or positive linear–linear correlation (e, f). This contradiction of the two methods seems to be due to strong bimodal phase distributions (c and g), nonlinear phase precession (g and h), apparent multiple phase precession (d, e, and f), and spike phases distributed across the whole cycle (c–g). The latter case corresponds to a large σ_Y in the bivariate Gaussian model, and similar examples can be observed in Figs. 2C and 3C. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We note that this new approach does not require prior knowledge on the data apart from weakly restricting the range of slopes to a reasonable interval. Altogether, the proposed approach is numerically cheap. Simulations for Gaussian surrogate data indicated that mean values of sample estimates only weakly depended on sample size, and that the estimates matched expected values also in the limiting cases of low and high noise, which points to a well-behaved estimator.

Other common approaches to quantify circular–linear associations rely on cutting the data cylinder at some line, unwrapping the cylinder to create another linear variable, and finally using ordinary least squares regression and a linear–linear correlation coefficient. Results obtained in this way heavily depend on how the cylinder is cut. It is therefore necessary to uniquely define the cutting line. A prominent criterion is to maximize the coefficient of determination, i.e. the square of the linear–linear correlation coefficient. For discrete data, this maximization is numerically inconvenient because the coefficient of determination is a piecewise constant and discontinuous function of the parameters that define the cut.

Numerical simulations based on Gaussian surrogate data where the data cylinder was cut horizontally (at some phase) revealed three further problems: (i) the slope of the regression line was typically underestimated; (ii) the obtained correlation coefficient did not represent the data well, with an overestimation for high noise (low correlations) because of the maximization, and an underestimation for low noise (high correlations) because of the cut; (iii) the distribution of sample estimates of the correlation coefficient was bimodal for low sample sizes, and the most likely values of sample estimates strongly depended on the sample size. Thus, estimating correlation coefficients in this way is not efficient. Alternatively, cutting the data cylinder in other ways, for example a linear diagonal cut, cannot fundamentally solve these problems; conversely, additional problems arise that are related to fixing the parameters of the cut. We conclude that linear–linear approaches applied to circular–linear data are, in general, not robust.

There may, nevertheless, be circular–linear data for which a linear–linear approach can lead to acceptable results. Such cases typically include scenarios in which the circular variable is concentrated, that is, circular measurements are restricted to a small fraction of a cycle. In the limiting case of a highly concentrated circular variable, the described linear–linear approach with an appropriate horizontal cut and our new circular–linear approach are equivalent. Our new circular–linear correlation coefficient is therefore a natural analog of Pearson's product-moment correlation coefficient for linear–linear data. On the other hand, if the circular variable is not concentrated, that is, circular measurements are spread over the whole cycle, linear–linear approaches that maximize the coefficient of determination fail whereas our circular–linear approach is nevertheless valid.

As a specific application, we considered hippocampal phase precession, which describes the relation between a linear spatial variable and a circular phase variable. In most previous approaches, phase precession was quantified through ordinary least squares regression and Pearson's product-moment correlation coefficient (O'Keefe and Recce, 1993; Mehta et al., 2002; Huxter et al., 2003; Zugaro et al., 2005; Dragoi and Buzsáki, 2006; Ego-Stengel and

Wilson, 2007; Foster and Wilson, 2007; Brun et al., 2008; Cheng and Frank, 2008; Hafting et al., 2008; Lenck-Santini and Holmes, 2008; Mizuseki et al., 2009, 2012). As shown in this paper, the linear–linear approach suffers from the basic problems of biasing the estimates of slope, offset, and correlation, with a strong dependence of estimates on sample size and noise. Therefore, even though in different studies the same linear–linear method was used, it is difficult to compare results, in particular if the number of spikes is low and variable, as for example in single trials (Schmidt et al., 2009; Reifenstein et al., 2010, 2012). The circular–linear approach does not suffer from these basic problems, and therefore it provides more robust estimates.

Our comparison of the circular–linear and the linear–linear approaches to quantify hippocampal phase precession for a large sample of place fields showed that the obtained correlation coefficients and slopes are different. The differences match the ones seen in the surrogate data: linear–linear approach yielded larger correlation and smaller slope estimates than the circular–linear approach. However, these differences are difficult to interpret because phase precession is often nonlinear and bimodal. Both approaches cannot account for this.

Another linear–linear approach was proposed by Yamaguchi et al. (2002), who estimated parameters of phase precession by fitting multiple Gaussian probability density functions to the data. A quite different approach to the quantification of phase precession was described by Maurer et al. (2006): for interneurons with overlapping place fields, they identified the slope "through a process of parametrically rotating the smoothed position by phase density matrix", which correctly accounts for the circular nature of phase. Alternatively, to derive the slope, Mizuseki et al. (2009) estimated the frequency difference between spiking of a neuron and the local field potential in theta band.

A circular–linear approach to estimate the slope and offset of phase precession, as outlined in this article, was employed by Schmidt et al. (2009) and Reifenstein et al. (2012). In contrast, Huxter et al. (2008) estimated the slope by means of a diagonal cut of the cylinder, and the cut was defined through maximization of the coefficient of determination. Similar to our approach, they transformed the linear variable into a circular one with a phase range of 180° and applied a circular–circular correlation measure (ρ_T: "T-linear association"; Fisher and Lee, 1983).

The *circular–linear* approach where the circular variable Φ is predicted given the linear variable X is fundamentally different from a *linear–circular* approach where X is predicted given Φ. For the particular class of linear–circular associations in which the data follow a sinusoidal model $X = x_0 + b_0 \cos(\Phi - \phi_0)$ with parameters x_0, b_0, and ϕ_0, Fisher (1995, p. 145) proposed a positive correlation coefficient called "C-linear association", which was used by Lenck-Santini and Holmes (2008) and van der Meer and Redish (2011) to evaluate significances and correlations of distance/phase relationships (see also p. 651 of Zar, 1999; Berens, 2009). However, this sinusoidal model of the data might not always be appropriate to describe phase precession.

To conclude, we introduced and motivated a new method to quantify circular–linear associations. The new method can be applied to all kinds of circular–linear data, for example in the auditory system (FitzGerald et al., 2001); in this context an estimate of the circular–linear slope was introduced. In the auditory system, circular (phase) variables naturally occur due to the band-pass frequency filtering of the basilar membrane, and the linear variable "best frequency" reflects the cochleotopic organization of the auditory pathway (e.g., Yin and Kuwada, 1983). We have applied this method to hippocampal phase precession. A robust quantification of phase precession is essential to identify its functional role in memory (Robbe and Buzsáki, 2009; Lenck-Santini and Holmes, 2008). In particular, a systematic overestimation of the phase-position correlation with linear–linear approaches can lead to a wrong interpretation of the relation between behavioral performance and phase precession in memory tasks. We demonstrated that the circular–linear approach is better suited to quantify phase precession than linear–linear ones.

Acknowledgments

The authors thank Urs Bergmann, Nikolay Chenkov, Andreas V. M. Herz, Jorge Jaramillo, Eric T. Reifenstein, Kay Thurley, John Huxter, and Jozsef Csicsvari for comments. This research was supported by the Bundesministerium für Bildung und Forschung (BMBF) under Grant numbers 01GQ0901 (Bernstein Focus "Neuronal Basis of Learning"), 01GQ0410 + 01GQ1001A and 01GQ440 (Bernstein Centers for Computational Neuroscience, Berlin and Munich) and the Deutsche Forschungsgemeinschaft (DFG) through the SFB 618 "Theoretical Biology" and under Grant numbers Ke 788/1-4 and Le 2250/2-1.

Appendix A.

Here we provide a detailed motivation of the formalism used for circular–linear regression and correlation, which was briefly sketched in Section 3.

A.1. Circular–linear regression

For the regression of a circular dependent variable Φ on a linear independent variable X, an appropriate measure for the distance between two angles is necessary. The squared difference between two angles, as used in ordinary least squares regression, is inappropriate: for example, the two angles 5° and 355° have a squared difference of $(350^\circ)^2$ although they are only 10° apart. To avoid such an overestimation of small circular distances, we utilize a more natural measure.

A.1.1. Circular distance

We define the distance d between two angles ψ_1 and ψ_2 through

$$d(\psi_1, \psi_2) = 2[1 - \cos(\psi_1 - \psi_2)], \tag{A.1}$$

(Lund, 1999; see also Eq. (1)) which is symmetric with respect to the interchange of the two angles. This distance measure is nonnegative and has an upper bound of four, i.e. $0 \le d \le 4$. The value $d = 0$ implies $\psi_1 = \psi_2 \pmod{2\pi}$, and $d = 4$ implies $|\psi_1 - \psi_2| = \pi \pmod{2\pi}$. For small circular differences, $|\psi_1 - \psi_2| \ll 1$, a Taylor series expansion of the cosine function leads us to $d(\psi_1, \psi_2) \approx (\psi_1 - \psi_2)^2$, which is the usual Euclidean measure for distances (Lund, 1999; Jammalamadaka and SenGupta, 2001). To achieve this equivalence for small differences, the factor 2 appears in Eq. (A.1).

A.1.2. Circular–linear model

We use the circular distance d to fit a circular–linear model to the data, in analogy to using the Euclidean distance in ordinary least squares regression. Let us suppose a random sample of observations $(\phi_1, x_1), \ldots, (\phi_n, x_n)$, that is, data on the surface of a cylinder where ϕ_j is a circular and x_j is a linear measurement for $j = 1, \ldots, n$. The data vector can then be denoted as $(\vec{\phi}, \vec{x})$. We assume that the data obeys a linear model of the form

$$\tilde{\Phi} = [2\pi\, a X + \phi_0] \pmod{2\pi}. \tag{A.2}$$

The two parameters of this linear model are the slope a and the phase offset ϕ_0 of the regression line. The model allows to predict the mean angle $\tilde{\Phi}$ for a given $X = x_j$.

We fit the linear model in Eq. (A.2) to the measurements $(\vec{\phi}, \vec{x})$ by minimizing the mean circular distance, or error, D between observed and predicted angles,

$$D(\phi_0, a) = \frac{1}{n}\sum_{j=1}^{n} d(\phi_j, 2\pi\, a\, x_j + \phi_0). \tag{A.3}$$

Using the definition of circular distance in Eq. (A.1), we find

$$D(\phi_0, a) = 2\left[1 - \frac{1}{n}\sum_{j=1}^{n}\cos(\phi_j - 2\pi\, a\, x_j - \phi_0)\right], \tag{A.4}$$

which is a measure of circular variance of the residual distribution.

A.1.2.1. Estimating the phase offset ϕ_0. To obtain an estimate $\hat{\phi}_0$ of the phase offset ϕ_0, we first minimize $D(\phi_0, a)$ in Eq. (A.4) with respect to ϕ_0 for constant a. From $\partial D(\phi_0, a)/\partial\phi_0|_{\hat{\phi}_0} = 0$ we find $\sum_j \sin(\phi_j - 2\pi\, a\, x_j - \hat{\phi}_0) = 0$, which can be rewritten using the trigonometric addition theorem $\sin(\alpha - \beta) = \sin\alpha\cos\beta - \cos\alpha\sin\beta$. With $\alpha = \phi_j - 2\pi a x_j$ and $\beta = \hat{\phi}_0$ we find (Gould, 1969; Fisher, 1995)

$$\hat{\phi}_0(a) = \arctan^*\frac{\sum_j \sin(\phi_j - 2\pi\, a\, x_j)}{\sum_j \cos(\phi_j - 2\pi\, a\, x_j)} = \arg\left[\sum_{j=1}^{n}\exp\{i\,(\phi_j - 2\pi\, a\, x_j)\}\right]. \tag{A.5}$$

Here, arctan* is the quadrant-specific inverse of the tangent, also known as atan2 or four-quadrant arctangent (see also Jammalamadaka and SenGupta, 2001),

$$\arctan^*\left(\frac{S}{C}\right) = \begin{cases} \arctan\left(\frac{S}{C}\right) & \text{if } C > 0, \quad S \geq 0, \\ \frac{\pi}{2} & \text{if } C = 0, \quad S > 0, \\ \arctan\left(\frac{S}{C}\right) + \pi & \text{if } C < 0, \\ \arctan\left(\frac{S}{C}\right) + 2\pi & \text{if } C \geq 0, \quad S < 0, \\ \text{undefined} & \text{if } C = 0, \quad S = 0 \end{cases} \tag{A.6}$$

for $S = \sum_j \sin(\phi_j - 2\pi\, a\, x_j)$ and $C = \sum_j \cos(\phi_j - 2\pi\, a\, x_j)$.

A.1.2.2. Estimating the slope a. Having found $\hat{\phi}_0$ as a function of the slope a in Eq. (A.5), we finally need to obtain the estimate $\hat{a}$. We find $\hat{a}$ at the minimum (with respect to a) of $D(\phi_0, a)$ in Eq. (A.4) with $\phi_0 = \hat{\phi}_0(a)$,

$$D(\hat{\phi}_0(a), a) = 2\left[1 - \frac{1}{n}\sum_{j=1}^{n}\cos(\phi_j - 2\pi\, a\, x_j - \hat{\phi}_0(a))\right]. \tag{A.7}$$

To summarize, fitting a linear model $\Phi = 2\pi\, a\, X + \phi_0 \,(\mathrm{mod}\, 2\pi)$ to circular–linear data $(\vec{\phi}, \vec{x})$ requires a one-dimensional (with respect to the slope a) minimization of the mean circular distance D in Eq. (A.7). The estimate $\hat{a}$ of the slope is

$$\hat{a} = \arg\min_a D(\hat{\phi}_0(a), a), \tag{A.8}$$

which normally demands numerical methods. The estimate $\hat{\phi}_0$ for the phase offset then follows from Eq. (A.5) with $a = \hat{a}$.

A.1.3. Mean resultant length R

To put our ansatz for circular–linear regression in a wider context of circular statistics and to see important properties of the mean circular distance in Eq. (A.7) that considerably simplify its calculation, we show how the mean circular distance is related to the so-called "mean resultant length" or "vector strength". For a set of circular observations $\vec{\psi} = (\psi_1, \ldots, \psi_n)$, each angle ψ_j can be represented by a complex unit vector $\exp(i\,\psi_j)$. The normalized sum $n^{-1}\sum_{j=1}^{n}\exp(i\,\psi_j)$ is the circular mean vector. It has the real component $C = n^{-1}\sum_{j=1}^{n}\cos\psi_j$ and the imaginary component $S = n^{-1}\sum_{j=1}^{n}\sin\psi_j$. The length of the circular mean vector is called the mean resultant length

$$R(\vec{\psi}) = \sqrt{C^2 + S^2}, \tag{A.9}$$

and the angle of the circular mean vector, called circular mean angle, is defined as

$$\overline{\psi} = \arctan^*\left(\frac{S}{C}\right). \tag{A.10}$$

Eq. (A.10) is analogous to Eq. (A.5), that is, for $\vec{\psi} = \vec{\phi} - 2\pi a\vec{x}$ the circular mean angle is $\overline{\psi} = \hat{\phi}_0(a)$. Knowing this relation, we can transform the sum in Eq. (A.7) as follows (Jammalamadaka and SenGupta, 2001):

$$\frac{1}{n}\sum_{j=1}^{n}\cos(\psi_j - \overline{\psi}) = \frac{1}{n}\sum_{j=1}^{n}(\cos\psi_j\,\cos\overline{\psi} + \sin\psi_j\,\sin\overline{\psi}) = C\,\cos\overline{\psi} + S\,\sin\overline{\psi} = R\cos^2\overline{\psi} + R\sin^2\overline{\psi} = R. \tag{A.11}$$

In the first line of Eq. (A.11) we have used a trigonometric addition theorem, and in the second line we have utilized the identities $C = R\cos\overline{\psi}$ and $S = R\sin\overline{\psi}$, which follow from Eqs. (A.9) and (A.10). The mean circular distance in Eq. (A.7) can then be rewritten as

$$D(\hat{\phi}_0(a), a) = 2[1 - R(\vec{\phi} - 2\pi\, a\,\vec{x})]. \tag{A.12}$$

Therefore, minimizing $D(\hat{\phi}_0(a), a)$ for circular–linear regression is equivalent to maximizing the mean resultant length $R(\vec{\phi} - 2\pi\, a\,\vec{x})$, which is defined in Eq. (1); see also Fig. 7. We note that $R(\vec{\phi} - 2\pi\, a\,\vec{x}) = R(\vec{\phi} - 2\pi\, a\,\vec{x} - \phi_0)$ for arbitrary phase shifts ϕ_0 because the length of a circular mean vector does not depend on its direction. Therefore the slope in Eq. (A.8) can be estimated without knowledge about the phase offset.

It is important to note that the estimate $\hat{a}$ according to Eq. (A.8) is not unique for a given data set. In fact, there can be infinitely many solutions. Such solutions correspond to "barber's pole" regression functions that spiral many times around the surface of a cylinder (Gould, 1969; Fisher and Lee, 1992; Fisher, 1995; Downs and Mardia, 2002). As a specific example, we consider the case of

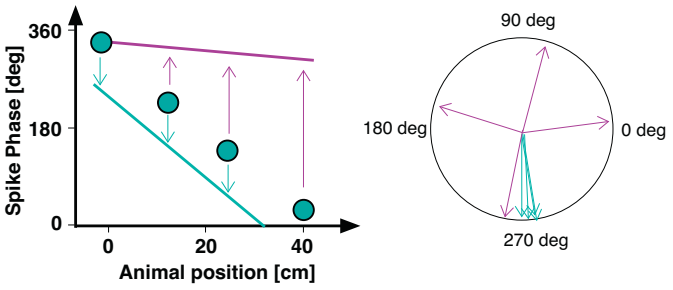

Fig. 7. Illustration of the method to estimate the slope. The green line on the left indicates a slope that matches the data (green dots) well. In this case, in the polar plot on the right, the resulting residual vectors point in similar directions, and the mean resultant length is large. In contrast, the magenta line's slope deviates from the slope of the data points, and the residual vectors point in different directions. In that case, the mean resultant length of the residual vectors is small. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

equidistant linear measurements, that is, $x_j = j\Delta$ with $\Delta > 0$. From Eq. (A.12) we then find that R is periodic in a with period $1/\Delta$. To obtain an unique slope $\hat{a}$, one needs to restrict the range of possible slopes, for example to an interval $[a_{\min}, a_{\max}]$, through some prior knowledge on the data.

A.1.4. Further remarks

In this section we present two remarks and one example to discuss the results obtained so far.

(i) Fitting a linear model to circular–linear data through minimizing the mean circular distance is equivalent to a maximum-likelihood approach when we assume that the data $(\vec{\phi}, \vec{x})$ is a set of independent observations with angles ϕ_j drawn from a von Mises distribution with a mean direction given by the linear model $2\pi\, a\, x_j + \phi_0$ (Gould, 1969); see also Fisher (1995) and Jammalamadaka and SenGupta (2001).

(ii) Having found an approach to circular–linear regression, how can we quantify the degree of association between the circular variable Φ and the linear variable X? One possibility is to use the maximum of the mean resultant length $\hat{R} = \max_a R(\vec{\phi} - 2\pi\, a\, \vec{x})$ from Eq. (A.12), which has some convenient properties. For example, $\hat{R}$ is invariant with respect to shifts in Φ or X because the length of a circular mean vector is independent of its direction. $\hat{R}$ is also invariant to scaling of X because the scale factor a is optimized. Values of $\hat{R}$ are restricted to the interval [0, 1], and the larger the value, the higher is the correlation between Φ and X.

Although $\hat{R}$ has convenient properties, there is, however, a strong argument against using $\hat{R}$ as a measure for the degree of association. Importantly, Φ and X being independent does not imply $\hat{R} = 0$, even in expectation. To see this property, we define the corresponding measure $\mathcal{R}$ for random variates Φ and X:

$$\mathcal{R}(\Phi - 2\pi\, aX) = \sqrt{E[\cos(\Phi - 2\pi aX)]^2 + E[\sin(\Phi - 2\pi aX)]^2} \tag{A.13}$$

with $E[\cdot]$ denoting the expectation value. Let us calculate $\mathcal{R}$ for a specific example.

A.1.4.1. Example: wrapped bivariate Gaussian distribution. As a specific example, we utilize the bivariate Gaussian distribution of two linear variates X and Y, where Y is wrapped around the circle so that we obtain a circular variate $\Phi = Y \pmod{2\pi}$. The bivariate Gaussian distribution is characterized by the mean vectors μ_X and μ_Y, the variances $\sigma_X^2 > 0$ and $\sigma_Y^2 > 0$, and the correlation ρ; the expected mean resultant length $\mathcal{R}$ as defined in Eq. (A.13) is, following Fisher and Lee (1983),

$$\mathcal{R}(\Phi - 2\pi aX) = \exp\left[-\frac{(2\pi a\sigma_X)^2}{2} + \rho(2\pi a\sigma_X)\sigma_Y - \frac{\sigma_Y^2}{2}\right] \tag{A.14}$$

which assumes its maximum at

$$a = \frac{\rho\,\sigma_Y}{2\pi\,\sigma_X} \tag{A.15}$$

with a maximum value $\max_a \mathcal{R} = \exp[-\sigma_Y^2(1-\rho^2)/2]$. Because $\max_a \mathcal{R} > 0$ even for $\rho = 0$ and because $\max_a \mathcal{R}$ increases with decreasing σ_Y, the mean resultant length is not suited for quantifying circular–linear associations. In the next section, we therefore develop a more appropriate measure, which is nevertheless based on circular–linear regression.

A.2. Circular–linear correlation

The goal of this section is to motivate and introduce a natural measure for the association between a circular variable Φ and a linear variable X. This measure should ideally share as many properties as possible with the Pearson product-moment correlation coefficient for two linear variables, and become identical to it in some limiting cases. Such a measure for quantifying a circular–linear association is not yet available. What is available, on the other hand, are measures for associations between two circular variables Φ and Θ (Fisher and Lee, 1983; Jammalamadaka and Sarma, 1988; Fisher, 1995; Jammalamadaka and SenGupta, 2001). Let us now introduce a particular measure for circular–circular associations and then show how it can be adapted to quantify circular–linear associations.

A.2.1. Circular–circular correlation measure ρ_c

For two circular random variates Φ and Θ with a joint distribution on the surface of a torus, Jammalamadaka and Sarma (1988) define as a measure for a circular correlation coefficient

$$\rho_c(\Phi, \Theta) = \frac{E[\sin(\Phi - \overline{\Phi})\,\sin(\Theta - \overline{\Theta})]}{\sqrt{\mathrm{Var}[\sin(\Phi - \overline{\Phi})]\,\mathrm{Var}[\sin(\Theta - \overline{\Theta})]}} \tag{A.16}$$

with the expectation value $E[\cdot]$, the expected variance $\mathrm{Var}[\cdot]$, and the expected circular mean directions $\overline{\Phi}$ and $\overline{\Theta}$ of the variates. The measure ρ_c has the following six properties (Jammalamadaka and SenGupta, 2001):

(i) $-1 \le \rho_c \le 1$
(ii) $\rho_c = \pm 1$ if and only if $\Phi = \pm\Theta + \theta_0$ with a constant angle θ_0
(iii) ρ_c is invariant under choice of origin for Φ and Θ
(iv) $\rho_c(\Phi, \Theta) = \rho_c(\Theta, \Phi)$
(v) $\rho_c = 0$ if Φ and Θ are independent; the converse need not be true!
(vi) If the distributions of Φ and Θ are concentrated in a small neighborhood of their respective mean directions, we have $\rho_c(\Phi, \Theta) \approx \rho(\Phi, \Theta)$, where ρ is the ordinary correlation coefficient for two linear variables.

In summary, ρ_c is a natural analog of the Pearson product-moment correlation coefficient ρ. However in contrast to Pearson's ρ, the circular–circular correlation ρ_c is in general not invariant with respect to scaling of Φ or Θ.

For random circular–circular data $(\vec{\phi}, \vec{\theta}) = [(\phi_1, \theta_1), \ldots, (\phi_n, \theta_n)]$, an estimate of ρ_c is given by the sample correlation coefficient

$$\hat{\rho}_c(\vec{\phi}, \vec{\theta}) = \frac{\sum_{i=1}^n \sin(\phi_i - \overline{\phi})\,\sin(\theta_i - \overline{\theta})}{\sqrt{\sum_{i=1}^n \sin^2(\phi_i - \overline{\phi}) \sum_{j=1}^n \sin^2(\theta_j - \overline{\theta})}} \tag{A.17}$$

where $\overline{\phi}$ and $\overline{\theta}$ are sample mean values; see also Eq. (3).

For testing hypotheses about ρ_c and to calculate p-values when n is sufficiently large, we note that under the null hypothesis $\rho_c = 0$ the quantity $z := \hat{\rho}_c \sqrt{n\,\hat{\lambda}_{20}\,\hat{\lambda}_{02}/\hat{\lambda}_{22}}$ is normally distributed, where $\hat{\lambda}_{ij} = n^{-1}\sum_{k=1}^n \sin^i(\phi_k - \overline{\phi}) \sin^j(\theta_k - \overline{\theta})$ (Jammalamadaka and SenGupta, 2001). The significance value p can therefore be obtained from the cumulative normal distribution: $p = 1 - \mathrm{erf}(|z|/\sqrt{2})$]. To numerically calculate circular–circular correlations and significances, we used the implementation by Berens (2009).

If the sample mean values $\overline{\phi}$ and $\overline{\theta}$ are not well defined because Φ or Θ have uniform distributions, it is more robust to replace the numerator of Eq. (A.17) by the expression $R(\vec{\phi} - \overline{\theta}) - R(\vec{\phi} + \overline{\theta})$ with R as defined in Eq. (A.9); for details we refer to Jammalamadaka and SenGupta (2001). We note that for small sample sizes ($n \lesssim 10$) Eq.

(A.17) in rare cases yields values slightly outside the range from −1 to +1.

A.2.1.1. Example: wrapped bivariate Gaussian distribution. To illustrate properties of ρ_c for an example case, we again use the bivariate Gaussian distribution, which is characterized by the variances $\sigma_X^2 > 0$ and $\sigma_Y^2 > 0$ as well as the correlation ρ of two linear variates X and Y. We assume that X and Y are wrapped around the circle to obtain two circular variates $\Theta = X \pmod{2\pi}$ and $\Phi = Y \pmod{2\pi}$. Then Θ and Φ have a wrapped bivariate Gaussian distribution, and we find from Eq. (A.16) (Johnson and Wehrly, 1977; Jammalamadaka and SenGupta, 2001)

$$\rho_c(\Phi, \Theta) = \frac{\sinh(\rho\,\sigma_X\sigma_Y)}{\sqrt{\sinh(\sigma_X^2)\,\sinh(\sigma_Y^2)}}. \tag{A.18}$$

For $\sigma_X \ll 1$ and $\sigma_Y \ll 1$, the correlation ρ_c in Eq. (A.18) is well approximated by the linear correlation ρ; see also property (vi) above. For arbitrary σ_X and σ_Y, note that even if X and Y are perfectly correlated in the linear sense, that is, $\rho = \pm 1$, it does not follow $\rho_c = \pm 1$ unless $\sigma_X = \sigma_Y$; see also property (ii) above.

To develop an appropriate measure for circular–linear correlations, we aim at retaining the important property that $\rho = \pm 1$ implies $\rho_c = \pm 1$. We therefore need to scale X before wrapping. This approach of appropriately scaling X before wrapping is used in what follows.

A.2.2. Transforming the linear variable into a circular one

Measures for circular–circular associations typically require that the data follow a model of the form $\Phi = \pm\Theta + \phi_0$, that is, the slope is restricted to values of +1 or −1 but the phase offset ϕ_0 can be arbitrary; see property (ii) above. To generate appropriate circular–circular variables Φ and Θ from circular–linear variables Φ and X, we utilize results from the section on circular–linear regression, which provide us with the slope a of the linear regression line of Φ on X. Scaling the linear variable X by the factor $2\pi|a|$ and taking the result modulo 2π, we obtain the circular variable $\Theta = 2\pi|a|X \pmod{2\pi}$. The two circular variables Φ and Θ are then distributed according to the model $\Phi = \pm\Theta + \phi_0$ with slope +1 for $a > 0$ and slope −1 for $a < 0$, as needed for using a measure for quantifying circular–circular associations. For $a = 0$ we take $\rho_c = 0$. The degree of association between Φ and X can therefore be quantified through the degree of association between Φ and $\Theta = 2\pi|a|X \pmod{2\pi}$.

A.2.2.1. Example: wrapped bivariate Gaussian distribution. For a wrapped, and scaled, bivariate Gaussian distribution with circular variates $\Theta = 2\pi|a|X \pmod{2\pi}$ with a slope parameter a and $\Phi = Y \pmod{2\pi}$ we find from Eq. (A.18)

$$\rho_c(\Phi, \Theta) = \frac{\sinh(\rho\,2\pi|a|\,\sigma_X\sigma_Y)}{\sqrt{\sinh([2\pi|a|\,\sigma_X]^2)\,\sinh(\sigma_Y^2)}} \tag{A.19}$$

for $a \neq 0$ and $\rho_c(\Phi, \Theta) = 0$ for $a = 0$. Again, for small σ_X and σ_Y, ρ_c in Eq. (A.19) is well approximated by the linear correlation ρ – independent of the value of a. If X and Y are perfectly correlated in the linear sense, that is, $\rho = \pm 1$, we find $\rho_c = \pm 1$ only if $2\pi|a|\,\sigma_X = \sigma_Y$. This fact indicates, again, that we must know the slope a to obtain, from the measure ρ_c, a reasonable value for the association between Φ and X.

A special value for a can be found by maximizing the expected mean resultant length $\mathcal{R}$ in Eq. (A.13). For the wrapped bivariate Gaussian distribution, $\mathcal{R}$ is given in Eq. (A.14), and $a = \rho\,\sigma_Y/(2\pi\,\sigma_X)$ from Eq. (A.15) was shown to maximize $\mathcal{R}$. Plugging this slope in the correlation ρ_c in Eq. (A.19) we obtain the expected correlation

$$\rho_c = \mathrm{sgn}(\rho)\sqrt{\frac{\sinh(\rho^2\,\sigma_Y^2)}{\sinh(\sigma_Y^2)}} \tag{A.20}$$

which is plotted in Fig. 2B for $\sigma_Y = 1$, 2, and 3.

We have argued that the correlation coefficient ρ_c for two circular random variates Φ and Θ shares many basic properties with the usual product-moment correlation coefficient, but ρ_c is not invariant to scaling of Φ or Θ. However, for circular–linear variates Φ and X, and for $\Theta = 2\pi|a|X \pmod{2\pi}$ where a is obtained from circular–linear regression, ρ_c is independent of the scale at which X is measured because a normalizes X. Accordingly, Eq. (A.20) is independent of σ_X. Still, ρ_c is *not* invariant to scaling of the phase Φ.

References

Berens P. CircStat: a MATLAB toolbox for circular statistics. J Stat Softw 2009;31:1–21.

Brun VH, Solstad T, Kjelstrup KB, Fyhn M, Witter MP, Moser EI, et al. Progressive increase in grid scale from dorsal to ventral medial entorhinal cortex. Hippocampus 2008;18:1200–12.

Buzsáki G. Theta oscillations in the hippocampus. Neuron 2002;33:325–40.

Cheng S, Frank LM. New experiences enhance coordinated neural activity in the hippocampus. Neuron 2008;57:303–13.

Diba K, Buzsáki G. Forward and reverse hippocampal place-cell sequences during ripples. Nat Neurosci 2007;10:1241–2.

Diba K, Buzsáki G. Hippocampal network dynamics constrain the time lag between pyramidal cells across modified environments. J Neurosci 2008;28:13448–56.

Downs TD, Mardia KV. Circular regression. Biometrika 2002;89:683–97.

Dragoi G, Buzsáki G. Temporal encoding of place sequences by hippocampal cell assemblies. Neuron 2006;50:145–57.

Ego-Stengel V, Wilson MA. Spatial selectivity and theta phase precession in CA1 interneurons. Hippocampus 2007;17:161–74.

Ekstrom AD, Meltzer J, McNaughton BL, Barnes CA. NMDA receptor antagonism blocks experience-dependent expansion of hippocampal 'place fields'. Neuron 2001;31:631–8.

Fisher NI. Statistical analysis of circular data. Cambridge, UK: Cambridge University Press; 1995.

Fisher NI, Lee AJ. A correlation coefficient for circular data. Biometrika 1983;70:327–32.

Fisher NI, Lee AJ. Regression models for an angular response. Biometrics 1992;48:665–77.

FitzGerald JV, Burkitt AN, Clark GM, Paolini AG. Delay analysis in the auditory brainstem of the rat: comparison with click latency. Hear Res 2001;159:85–100.

Foster DJ, Wilson MA. Hippocampal theta sequences. Hippocampus 2007;17:1093–9.

Geisler C, Diba K, Pastalkova E, Mizuseki K, Royer S, Buzsáki G. Temporal delays among place cells determine the frequency of population theta oscillations in the hippocampus. Proc Natl Acad Sci U S A 2010;107:7957–62.

Gould AL. A regression technique for angular variates. Biometrics 1969;25:683–700.

Jensen O, Lisman JE. Hippocampal CA3 region predicts memory sequences: accounting for the phase precession of place cells. Learn Mem 1996;3:279–87.

Hafting T, Fyhn M, Bonnevie T, Moser MB, Moser EI. Hippocampus-independent phase precession in entorhinal grid cells. Nature 2008;453:1248–52.

Harris KD, Henze DA, Hirase H, Leinekugel X, Dragoi G, Czurkó A, et al. Spike train dynamics predicts theta-related phase precession in hippocampal pyramidal cells. Nature 2002;417:738–41.

Harvey CD, Collman F, Dombeck DA, Tank DW. Intracellular dynamics of hippocampal place cells during virtual navigation. Nature 2009;461:941–6.

Hasselmo ME, Eichenbaum H. Hippocampal mechanisms for the context-dependent retrieval of episodes. Neural Netw 2005;18:1172–90.

Huxter JR, Burgess N, O'Keefe J. Independent rate and temporal coding in hippocampal pyramidal cells. Nature 2003;425:828–32.

Huxter JR, Senior TJ, Allen K, Csicsvari J. Theta phase-specific codes for two-dimensional position, trajectory and heading in the hippocampus. Nat Neurosci 2008;11:587–94.

Jammalamadaka SR, Sarma YR. A correlation coefficient for angular variables. In: Matusita K, editor. Statistical theory and data analysis II. Amsterdam: North Holland; 1988. p. 349–64.

Jammalamadaka SR, SenGupta A. Topics in circular statistics. World Scientific: Singapore; 2001.

Johnson RA, Wehrly TE. Measures and models for angular correlation and angular-linear correlation. J Roy Stat Soc 1977;39:222–9.

Kamondi A, Acsady L, Wang XJ, Buzsáki G. Theta oscillations in somata and dendrites of hippocampal pyramidal cells in vivo: activity-dependent phase-precession of action potentials. Hippocampus 1998;8:244–61.

Leibold C, Gundlfinger A, Schmidt R, Thurley K, Schmitz D, Kempter R. Temporal compression mediated by short-term synaptic plasticity. Proc Natl Acad Sci U S A 2008;105:4417–22.
Lenck-Santini PP, Holmes GL. Altered phase precession and compression of temporal sequences by place cells in epileptic rats. J Neurosci 2008;28:5053–62.
Lengyel M, Szatmary Z, Erdi P. Dynamically detuned oscillations account for the coupled rate and temporal code of place cell firing. Hippocampus 2003;13: 700–14.
Lisman JE, Talamini LM, Raffone A. Recall of memory sequences by interaction of the dentate and CA3: a revised model of the phase precession. Neural Netw 2005;18:1191–201.
Losonczy A, Zemelman BV, Vaziri A, Magee CM. Network mechanisms of theta related neuronal activity in hippocampal CA1 pyramidal neurons. Nat Neurosci 2010;13:967–72.
Lund U. Least circular distance regression for directional data. J Appl Stat 1999;26:723–33.
Maurer AP, Cowen SL, Burke SN, Barnes CA, McNaughton BL. Phase precession in hippocampal interneurons showing strong functional coupling to individual pyramidal cells. J Neurosci 2006;26:13485–92.
van der Meer MAA, Redish AD. Theta phase precession in rat ventral striatum links place and reward information. J Neurosci 2011;31:2843–54.
Mehta MR, Lee AK, Wilson MA. Role of experience and oscillations in transforming a rate code into a temporal code. Nature 2002;417:741–6.
Mizuseki K, Sirota A, Pastalkova E, Buzsáki G. Theta oscillations provide temporal windows for local circuit computation in the entorhinal–hippocampal loop. Neuron 2009;64:267–80.
Mizuseki K, Royer S, Diba K, Buzsáki G. Activity dynamics and behavioral correlates of CA3 and CA1 hippocampal pyramidal neurons. Hippocampus 2012; doi:10.1002/hipo.22002.
O'Keefe J, Recce ML. Phase relationship between hippocampal place units and EEG theta rhythm. Hippocampus 1993;3:317–30.
Reifenstein ET, Stemmler MB, Herz AVM. Single-run phase precession in entorhinal grid cells. Front Comput Neurosci 2010; doi:10.3389/conf.fncom.2010.51.00093.
Reifenstein ET, Kempter R, Schreiber S, Stemmler MB, Herz AVM. Grid cells in rat entorhinal cortex encode physical space with independent firing fields and phase precession at the single-trial level. Proc Natl Acad Sci U S A 2012., doi:10.1073/pnas.1109599109.
Robbe D, Buzsáki G. Alteration of theta timescale dynamics of hippocampal place cells by a cannabinoid is associated with memory impairment. J Neurosci 2009;29:12597–605.
Schmidt R, Diba K, Leibold C, Schmitz D, Buzsáki G, Kempter R. Single-trial phase precession in the hippocampus. J Neurosci 2009;29:13232–41.
Skaggs WE, McNaughton BL, Wilson MA, Barnes CA. Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. Hippocampus 1996;6:149–72.
Thurley K, Leibold C, Gundlfinger A, Schmitz D, Kempter R. Phase precession through synaptic facilitation. Neural Comput 2008;20:1285–324.
Tsodyks MV, Skaggs WE, Sejnowski TJ, McNaughton BL. Population dynamics and theta phase precession of hippocampal place cell firing: a spiking neuron model. Hippocampus 1996;6:271–80.
Wallenstein GV, Hasselmo ME. GABAergic modulation of hippocampal population activity: sequence learning, place field development and the phase precession effect. J Neurophysiol 1997;78:393–408.
Yamaguchi Y, Aota Y, McNaughton BL, Lipa P. Bimodality of theta phase precession in hippocampal place cells in freely running rats. J Neurophysiol 2002;87:2629–42.
Yin TC, Kuwada S. Binaural interaction in low-frequency neurons in inferior colliculus of the cat. III. Effects of changing frequency. J Neurophysiol 1983;50: 1020–42.
Zar JH. Biostatistical analysis. Upper Saddle River, NJ: Prentice-Hall; 1999.
Zugaro MB, Monconduit L, Buzsáki G. Spike phase precession persists after transient intrahippocampal perturbation. Nat Neurosci 2005;8:67–71.