Detection and assessment of near-zero delays in neuronal spiking activity

G. Schneider a,∗, D. Nikolić b, c

a Department of Computer Science and Mathematics, Johann Wolfgang Goethe University, Robert-Mayer-Str. 10, 60325 Frankfurt (Main), Germany
b Department of Neurophysiology, Max-Planck-Institute for Brain Research, Frankfurt (Main), Germany
c Frankfurt Institute for Advanced Studies, Johann Wolfgang Goethe University, Frankfurt (Main), Germany

Received 23 May 2005; received in revised form 13 August 2005; accepted 22 August 2005

Abstract

Cross-correlation histograms (CCHs) have been widely used to study the temporal relationship between pairwise recordings of neuronal signals. One interesting parameter of a CCH is the time position of the central peak which indicates delays between signals. In order to study the potential relevance of these delays which can be as small as 1 ms, it is necessary to measure them with high precision. We present a method for the estimation of the central peak’s position that is based on fitting a cosine function to the CCH and show that the precision of this estimate can be tracked analytically. We validate the resulting formula by simulations and by the analysis of a sample dataset obtained from cat visual cortex. The results indicate that the time position of the center peak can be estimated with submillisecond precision. The formula allows one also to develop a test of statistical significance for differences between two sets of measurements.

Keywords: Cross-correlation; Central peak; Phase offset; Estimation precision; Spatio-temporal organization; Action potential; Visual cortex

1. Introduction

Cross-correlation histograms (CCHs) (Moore et al., 1966; Perkel et al., 1967; Abeles, 1982; Aertsen and Gerstein, 1985) have been widely used to study the temporal relations between pairwise recordings of neurophysiological signals (Gray et al., 1989; Roelfsema et al., 1997; Castelo-Branco et al., 2000). The CCH for two spike trains is computed by counting the number of coincident events that occur at different delays between the signals up to a maximal delay, T (for an example see Fig. 1A). A center peak indicates that two cells tend to fire spikes simultaneously. Center peaks can be also shifted up to several tens of milliseconds (König et al., 1995), which then indicates that the spikes of one cell tend to be delayed relative to the spikes of the other cell. As these shifts are often associated with oscillatory activity (indicated by the satellite peaks in Fig. 1A), we refer to them as “phase offsets”. The goal of the present study was to develop a method with which one could measure phase offsets with high precision (i.e., 1 ms or smaller) and to equip this method with an estimate of the error of measurement.

In the past, phase offsets smaller than one or two milliseconds have been considered unimportant and thus practically equivalent to zero offsets (Roelfsema et al., 1997). However, the finding that neurons can coordinate their responses with millisecond accuracy (Reinagel and Reid, 2002; Jones et al., 2004; Ikegaya et al., 2004) suggests that even small phase offsets in CCHs may reflect important aspects in cortical information processing. It is therefore important to determine whether small phase offsets can be distinguished confidently from zero offsets and whether two sets of phase offsets differ from each other significantly.

To this end, we propose to fit the central peak of the CCH with a simple cosine function (Fig. 1B and D) and to use the time at which this function reaches its maximum as the estimate of the most frequent delay. This method is related to the technique by König (1994), which is based on more complex damped cosine (Gabor) functions. Those Gabor functions can fit oscillatory CCHs very well (Fig. 1A and C) but do not allow one to directly assess the precision of the phase estimate. The simple cosine function proposed here has fewer parameters, which allows standardizing the starting values and makes the fitting procedure objective. Most importantly, the error of measurement can be tracked analytically.

We first describe the method used for fitting the central CCH, the assumptions used for this fit and the rules for the appropriate choice of the starting values (Section 2). Also, we compare the phase offsets obtained by the cosine fit to those obtained by the Gabor function (König, 1994). The main part of the paper
Fig. 1. Example CCHs and fitted Gabor and cosine functions. (A) An example of a CCH between two multi-unit recordings computed over 20 trials, each 2000 ms in duration (T = 80 ms). Black: original time resolution (1/32 ms); grey: CCH smoothed with a 2 ms kernel. (B) The central part of the CCH in A for T = 10 ms. (C) Gabor function fitted to the CCH in A. (D) Cosine function fitted to B. Vertical lines: t = 0.

Section 3 deals with the statistical properties of the phase estimate derived from the cosine fit. We first present an analytically derived formula to estimate the variance with which the phase offset can be measured. We then compare the theoretical formula with the variance obtained from simulations (Section 4), investigate whether the formula can describe the empirical variability of the phase estimate and examine the possibility to create confidence intervals for the estimated phase shifts. The validation of the formula also includes an analysis performed on a sample dataset based on neuronal responses to different visual stimuli recorded simultaneously from 14 channels of multi-unit activity in cat area 17 (for details on experimental methods see Appendix A). We estimate the error with which phase offsets are measured in this dataset and evaluate the precision with which our formula can assess this error. Finally, we use the formula to design a significance test that enables investigating changes in phase offsets across two groups of paired measurements.

2. Estimating phase offsets with the cosine fit

The phase offset \( \phi \) is estimated by fitting a cosine function to the central part of the CCH. The cosine function was chosen due to the clear interpretation of its parameters \( \phi \), \( \beta \) and \( \omega \), the phase offset, the amplitude and the frequency, and \( \beta_0 \), the offset at the ordinate of the CCH (for an illustration see Fig. 2A).

We assume that an experimentally observed CCH results from an additive composition of a cosine and random error. This error is assumed to be normally distributed with zero mean and the same standard deviation \( \sigma \) at each data point in the CCH (illustrated in Fig. 2B), and independent across different data points. Thus, the observed number of coincidences, CCH, at time shift \( t \) can be described as

\[
CCH = \beta_0 + A \cos(\omega(t - \phi)) + \alpha Z_t, \quad \forall t = -T, \ldots, T,
\]

where all \( Z_t \) are independent and standard normally distributed random variables scaled by \( \alpha \). Independence and normal distribution of errors seemed to be suitable assumptions because the empirically derived residuals were largely uncorrelated and approximately Gaussian (distribution of residuals shown in Appendix B).

To find the parameter values that fit a particular CCH best, we use the Gauss–Newton method implemented in the statistical analysis tool R (http://www.r-project.org). This method finds the values for all four parameters of the cosine \((\alpha, \omega, \phi, \beta_0)\) that minimize the least-squares error. The parameter \( \sigma^2 \) represents the variance of the residuals, i.e., the mismatch between the CCH and the fitted cosine function and is therefore estimated after the fitting procedure.

We applied this method to our dataset and used standardized starting values for the iterative algorithm. As phase shifts were usually small (< 2 ms), the initial value of \( \phi \) was always set to 0. The initial value of \( \alpha \) was always set to 1, and that of \( \omega \), to the oscillation frequency that roughly corresponds to the width of the center peaks in the CCHs (in our case 45 Hz). The initial value of \( \beta_0 \) was chosen to be the mean count of coincidences in the CCH. This value was sufficiently close to \( \beta \) because the fitted parts of the CCHs usually covered about one cosine period.

The fitting procedure was highly robust against changes in the starting values of the parameters. The starting value of \( \alpha \) needed to be sufficiently different from half or double the underlying oscillation frequency. The starting value for \( \beta_0 \) needed to be far enough from the minimal observed height in the CCH because the procedure may otherwise reach a local minimum and return a cosine function with double the period and amplitude. Overall, our experience was that CCHs are easier to fit with cosine than with Gabor functions.

We also investigated whether the shifts obtained by the cosine fit are comparable to those resulting from a Gabor fit with the method of König (1994). The phase estimates obtained by the two methods when fitting 91 CCHs from the sample dataset are shown in Fig. 3. Phase estimates based on cosine and Gabor fit agreed to a high degree as indicated by the scatter along the main diagonal and by the high correlation between the two sets of measurements (\( r = 0.97 \)).

3. Precision of measurement

A thorough investigation of phase offsets requires an estimate of the precision with which these phase offsets can be
where and $G$ (i.e., CCHs computed from the sample dataset.

$\omega$ the parameters

Thus, the approximate variance of $\hat{\phi}$ can be described by two components. The first part, $1/\omega^2 V$, describes its dependence on the parameters $\omega$, $N$, $\sigma$ and $A$. According to the formula, $\text{Var}(\hat{\phi})$ decreases (i.e., the precision increases) with the frequency, $\omega$, because a shorter period of the cosine leads to a narrower center peak. The precision also increases with an increase in the number of data points within the fitted part of the CCH, $N$, and with a decrease in the variability, $\sigma$, relative to the amplitude of the cosine function, $A$.

The second part, $G$, describes how the variance of the phase estimate depends on the geometric parameters $s$ and $f$. To understand the dependence of $G$ on $f$, we first investigate $G$ for $s = 0$, i.e., for CCHs centered at zero (Fig. 4A). The graph of $G(0, f) = 1/D_1(f)$ oscillates around 1 and approaches 1 in the limit. The oscillation occurs because the information about the position of the peak carried by each data point is higher at the edges of the cosine function than at its maxima or minima. If $s$ is bigger, then the geometric term $G$ accounts for the fact that the visible part of a shifted cosine function is asymmetric with respect to the window defined by $\pm T$. Hence, the right and the left side of the window contribute different amounts of information about the size of the shift. This asymmetry is accounted for by Eqs. (4) and (5).

One should note that Fig. 4A describes only the dependence of the geometric term, $G$, on $f$. If one wants to infer about the dependence of the variance $\text{Var}(\hat{\phi})$ on $f$, one has to take into account that the term $V$ also changes with $f$ because an increase in the visible fraction of the period also increases the number of data points, $N$, if the time resolution is kept constant. The dependence of the variance $\text{Var}(\hat{\phi})$ on $f$ is shown in Fig. 4B for constant $\sigma$, $\omega$ and $A$ and for $\omega = 0$. $\text{Var}(\hat{\phi})$ does not oscillate but decreases continuously with growing $f$, the slope indicating the amount of information that is added by the respective data points. Thus, although the extremes of the cosine function contribute less information than the edges, every data point adds information about the position of the peak. Unfortunately, one cannot take advantage of this property by using the maximal possible analysis window because the side peaks of a CCH usually do not continue the cosine function fitted to the central peak. Thus, the fitted area of the CCH should in practice not be much larger than one period of a cosine. In our dataset, the choice of the analysis window was limited by the maximal possible analysis window because the side peaks of a CCH usually do not continue the cosine function fitted to the central peak.

The fitted area of the CCH should in practice not be much larger than one period of a cosine. In our dataset, the choice of the analysis window is limited by the maximal possible analysis window because the side peaks of a CCH usually do not continue the cosine function fitted to the central peak.

The percentage of data points within the fitted part of the CCH, $N$, and with a decrease in the variability, $\sigma$, relative to the amplitude of the cosine function, $A$.

4. Scope of the variance formula

We investigate the applicability of the formula for the variance of the phase estimate in situations in which two assumptions used in its computation are not necessarily met. First, the simplicity of Eq. (2) is based on the assumption that the oscillation frequency, $\omega$, is known beforehand, which is not the case in practice. Second, it is not clear whether the number of data points,
N, that result from the chosen \( T \) and the given time resolution is sufficiently large to justify the asymptotic approximations underlying the formula. Finally, the simulations are also performed to investigate whether the practical use of Eq. (2) is affected by the resolution (1/32 ms) and \( s \) by the nonlinear least squares method described in Section 2. Finally, use Eq. (2) to estimate the variance of the phase estimate, \( \sigma \).

The applicability of the formula to center peaks that are known to originate from noisy cosine functions is tested by simulations with experimentally derived parameter sets for the original time resolution (Section 4.1) and for lower time resolutions resulting from binning (Section 4.2). Finally, we compare the experimentally obtained variability in our sample dataset with the predictions of the formula (Section 4.3) in order to investigate whether it can be used to assess the expected variability of the estimators \( \hat{\sigma} \) and to determine confidence intervals for \( \psi \).

4.1. Simulations

The simulation procedure was as follows: For a cosine function with the parameters \( s \), \( f \), and \( \sigma \), add independent and normally distributed noise with mean 0 and variance \( \sigma^2 \) to each of the \( N \) data points. Fit a cosine to the resulting function by estimating \( \Psi \), \( A \), \( \sigma \), and \( s \) by the nonlinear least squares method described in Section 2. Finally, use Eq. (2) to estimate the variance of the phase estimates, \( \sigma^2 \).

The parameters used to generate surrogate CCHs were derived from the experimental dataset. We used the original time resolution (1/32 ms) and \( T = 10 \) ms, resulting in \( N = 640 \). Fig. 5 shows the distributions of \( f \), \( s \), and \( \sigma \) as estimated from all 91 CCHs of the responses to one stimulation condition (condition 1, see Appendix A). These empirically derived parameter values were used to simulate surrogate data by using all possible combinations of the values from the following sets:

\[
\sigma \in \{0.5, 1, 1.5, 2\}, \quad f \in \{0.9, 1, 1.1, 1.2\}, \quad s \in \{0, 4\%\%, 8\%\%\}.
\]

while \( A \) was always set to 1.

We performed 10,000 simulations for each parameter set.

The resulting distribution of the estimates \( \hat{\psi} \) for a typical parameter set \((f = 1.1, s = 0, \sigma = 1)\) is shown in Fig. 6A. The empirical standard deviation \( \sigma \) of the phase estimate for this parameter set was 0.17 ms. This value was used to validate the approximate formula (2). We used \( \hat{\sigma} = \left( \text{Var}(\hat{\psi}) \right)^{1/2} \) to estimate the standard deviation of the phase estimate and computed the root mean squares (RMS) error, i.e., the typical deviation of the estimates \( \hat{\sigma} \) from the empirical \( \sigma \), expressed in % of \( \sigma \).

\[
\frac{100\%}{\sigma} \sqrt{\frac{\sum (\sigma_i^2 - \sigma^2)^2}{(10000 - 1)}}
\]

The estimates of \( \sigma \) obtained by Eq. (2) deviated from the empirical derived value by about 6.5% on average (Fig. 6B). Thus, the formula predicted well the empirically obtained variation of the phase estimate for the typical parameter set.

Analogous results can be obtained when varying \( \sigma \) in the set \([0.5, 1, 1.5, 2]\). Fig. 6C shows the values of the empirical \( \sigma \) (points) and the RMS error of the estimates \( \hat{\sigma} \) (error bars) for the parameters \( f = 1.1, s = 0 \) and \( \sigma/A \in \{0.5, 1, 1.5, 2\} \). As predicted by Eq. (3), the standard deviation of the phase estimate grows linearly with \( \sigma/A \), the slope being dependent on the particular set of parameters. For the chosen value of \( f = 1.1 \), the slope is 0.17 ms per unit of \( \sigma/A \), which results in a typical measurement error of \( \sigma \approx 0.2 \) ms for a typical value of \( \sigma/A = 1 \).

The slopes for \( f = 0.9 \) and 1.2 were 0.19 and 0.16 ms per unit of \( \sigma/A \), indicating that the formula (2) is robust with respect to \( f \).
σ/A, respectively (simulation results not shown). In addition to the increase in the variability of the phase estimate, an increase in σ/A expectedly increased the typical error between the theoretically predicted variability, \( \hat{\sigma}_s \), and the variability obtained empirically, \( \sigma_s \) (indicated by error bars in Fig. 6C). Similar results were obtained for the other values of \( \sigma/A \), but also the variability of data under the increase of the variability of the phase estimate, an increase in \( \sigma/A \) is harder to estimate. Thus, to analyze phase offsets with the time resolution with which the data are indicated by dots, whereas the expected values 68% and 94% are indicated by solid lines.

To understand how the binning process affects the precision with which the variability of the measured phase offset in the data by binning but should, whenever possible, analyze phase offsets with the time resolution with which the data were acquired.

In general, binning reduces the number of data points, only the number of data points, 1–2 ms. This is because binning that averages counts reduces not remain unaffected if CCHs are binned with a bin size of up to 0.5 ms. It is therefore important to understand how the binning process affects the precision with which the phase shift and its variability can be estimated.

To investigate the degree to which the approximate formula can describe the variability of phase shifts obtained experimentally. To this end, we subdivided the responses to 20 presentations (trials) of the same visual stimulus into two subsets (odd and even trials). For all 91 pairs \((i, j)\) of recorded channels, CCHs were computed separately for each of the two subsets of trials and cosine functions were fitted to estimate phase delays. If phase delays remain stable with repeated stimulus presentations, the difference of the two estimates \(D_{ij} = \hat{\phi}_{ij}^{(1)} - \hat{\phi}_{ij}^{(2)}\) should be distributed normally with mean 0 and variance \(2\sigma^2\).

The deviation of \(\delta_{ij}\) from \(\sigma_s\) increased with the bin size, from 6.5% for the original time resolution of 1/32 ms to 18% for a bin size of 1 ms and to 61% for a bin size of 5 ms (Fig. 8). Large deviations of \(\delta_{ij}\) from \(\sigma_s\) may result in inappropriate confidence intervals and thus in erroneous conclusions about the statistical significance of differences between phase estimates. These results suggest that one should not reduce the time resolution of the data by binning but should, whenever possible, analyze phase offsets with the time resolution with which the data were acquired.

4.3. Agreement with variability in experimental data

As the assumptions associated with the cosine model do not necessarily hold true for the experimental data, it is necessary to investigate the degree to which the approximate formula can describe the variability of phase shifts obtained experimentally.

\[
\frac{1}{N} \sum_{i,j} D_{ij}^2 = \sigma^2 + \frac{1}{N} \sum_{i,j} (\hat{\phi}_{ij}^{(1)} - \hat{\phi}_{ij}^{(2)})^2
\]

Eq. (7). The deviation of \(\delta_{ij}\) from \(\sigma_s\) increased with the bin size, from 6.5% for the original time resolution of 1/32 ms to 18% for a bin size of 1 ms and to 61% for a bin size of 5 ms (Fig. 8). Large deviations of \(\delta_{ij}\) from \(\sigma_s\) may result in inappropriate confidence intervals and thus in erroneous conclusions about the statistical significance of differences between phase estimates. These results suggest that one should not reduce the time resolution of the data by binning but should, whenever possible, analyze phase offsets with the time resolution with which the data were acquired.

4.3. Agreement with variability in experimental data

As the assumptions associated with the cosine model do not necessarily hold true for the experimental data, it is necessary to investigate the degree to which the approximate formula can describe the variability of phase shifts obtained experimentally. To this end, we subdivided the responses to 20 presentations (trials) of the same visual stimulus into two subsets (odd and even trials). For all 91 pairs \((i, j)\) of recorded channels, CCHs were computed separately for each of the two subsets of trials and cosine functions were fitted to estimate phase delays. If phase delays remain stable with repeated stimulus presentations, the difference of the two estimates \(D_{ij} = \hat{\phi}_{ij}^{(1)} - \hat{\phi}_{ij}^{(2)}\) should be distributed normally with mean 0 and variance \(2\sigma^2\).

To investigate whether the estimates of the phase, \(\hat{\phi}_{ij}^{(1)}\), and of its measurement error, \(\hat{\sigma}_s\), derived from the odd trials yield a reliable confidence interval for the phase measurement \(\hat{\phi}_{ij}^{(1)}\) in the even trials, we estimated \(\hat{\sigma}_{ij}^2\) from the odd trials for every channel pair \((i, j)\) and, as the sign of \(D_{ij}\) is irrelevant, compared the distribution of the absolute values

\[
|Z_{ij}| = \frac{|\hat{\phi}_{ij}^{(1)} - \hat{\phi}_{ij}^{(2)}|}{\sqrt{2\hat{\sigma}_s}}
\]

for all 91 CCHs to the non-negative part of the standard normal distribution.
measurements have the same expected values, the test statistic can be estimated with Eq. (2). If the associated pairs of sample dataset, (histogram) distributions of normalized absolute phase measurement errors in the 102 G. Schneider et al. / Journal of Neuroscience Methods 152 (2005) 97–106

predicted and the observed distribution (Fig. 9, Kolmogorov–Smirnov test, p > 0.5). This suggests that Eq. (2) grasps the variability of $\hat{\phi}$ in the experimentally obtained CCHs very well and thus that this formula can be used to estimate the error with which phase offsets are measured.

5. Application: a significance test

In the preceding sections we showed that Eq. (2) reasonably describes the measurement error of the phase estimate not only for the simulated data but also for CCHs obtained from the experimental dataset. Therefore, we can use these error estimates to determine the statistical significance of the differences in phase shifts that come from different pairs of channels or from the same pair under different stimulation conditions. We propose a statistical test that can be used to investigate changes in phase offsets for entire groups of measurements simultaneously, thus avoiding multiple comparisons.

Consider two sets of measurements $S_1 := [\psi_1^{(1)}, \ldots, \psi_n^{(1)}]$ and $S_2 := [\psi_1^{(2)}, \ldots, \psi_n^{(2)}]$ in which the measurements $(\psi_1^{(1)}, \psi_2^{(2)})$ are paired exactly like the measurements $(\psi_1^{(2)}, \psi_2^{(1)})$. The latter can be estimated with Eq. (2). If the associated pairs of measurements have the same expected values, the test statistic $X := \sum_{i=1}^n \frac{(\psi_i^{(1)} - \psi_i^{(2)})^2}{\sigma_i^{(1)} + \sigma_i^{(2)}}$ is (approximately) the sum of squares of $n$ normally distributed random variables and thus has an approximate $\chi^2$-distribution with $n$ degrees of freedom ($\chi^2(n)$). In contrast, if some pairs of measurements differ more than one would expect from the size of their variances, $X$ grows, which then results in smaller $p$-values.

As an example, we applied this test to our experimental dataset (conditions 3 and 5; see Appendix A) to investigate whether the test indicates changes in the sets of $n = 91$ measured phase shifts when neuronal responses are evoked by identical or different stimuli (Table 1). To investigate responses to identical stimuli the 20 presentations of the same stimulus were again divided into two sets of odd and even trials, while the comparison between stimulation conditions was based on all 20 trials.

The results of this analysis are shown in Table 1. The large $p$-values (0.516 & 0.165) indicate that phase shifts do not change across odd and even trials of the same stimulation condition by more than is accounted for by their measurement errors. In contrast, the small $p$-value ($p < 0.0001$) indicates that the observed differences between measured phase shifts in conditions 3 and 5 are unlikely to originate from identical distributions. This application illustrates how the proposed test can be used to investigate temporal dynamics in neuronal activity by making statistical inferences about the stability and changes of phase shifts across stimulation conditions.

6. Discussion

We presented a method that can be used to estimate spiking delays between pairs of units based on fitting a simple cosine function to the central part of a CCH. The simplicity of the fitted function allows the derivation of a formula for the variability of the shift estimate. This formula can be computed easily and explains how different parameters of the CCH affect the precision with which phase offsets are measured. We could show that the formula accounts very well for the variability of the phase estimates in simulated CCHs and also that it predicts accurately the error with which phase offsets are measured in the experimental dataset. Therefore, our results suggest that the proposed method can reliably estimate phase offsets together with their measurement errors.

The application of the method to the experimental dataset indicated that phase offsets can be measured with high precision. We obtained an average standard deviation of a phase estimate of about 0.2 ms, which yields 95% confidence intervals of about 0.8 ms. This suggests that phase offsets can be measured with submillisecond precision and thus, that small phase offsets (e.g., 1 ms) can often be distinguished from zero-offsets with confidence. Therefore, the proposed method can be used to investigate whether these offsets change dynamically in a task dependent way and thus, whether they contain information that might be relevant for cortical processing. The proposed statistical test can assess the difference between two paired sets of phase measurements and thus, can be helpful in such investigations.

Table 1

<table>
<thead>
<tr>
<th>Comparison</th>
<th>X</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odd–even trials, condition 3</td>
<td>89.8</td>
<td>0.516</td>
</tr>
<tr>
<td>Odd–even trials, condition 5</td>
<td>104.1</td>
<td>0.165</td>
</tr>
<tr>
<td>Conditions 3–5</td>
<td>285.0</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

As an example, we applied this test to our experimental dataset (conditions 3 and 5; see Appendix A) to investigate whether the test indicates changes in the sets of $n = 91$ measured phase shifts when neuronal responses are evoked by identical or different stimuli (Table 1). To investigate responses to identical stimuli the 20 presentations of the same stimulus were again divided into two sets of odd and even trials, while the comparison between stimulation conditions was based on all 20 trials.

The results of this analysis are shown in Table 1. The large $p$-values (0.516 & 0.165) indicate that phase shifts do not change across odd and even trials of the same stimulation condition by more than is accounted for by their measurement errors. In contrast, the small $p$-value ($p < 0.0001$) indicates that the observed differences between measured phase shifts in conditions 3 and 5 are unlikely to originate from identical distributions. This application illustrates how the proposed test can be used to investigate temporal dynamics in neuronal activity by making statistical inferences about the stability and changes of phase shifts across stimulation conditions.

6. Discussion

We presented a method that can be used to estimate spiking delays between pairs of units based on fitting a simple cosine function to the central part of a CCH. The simplicity of the fitted function allows the derivation of a formula for the variability of the shift estimate. This formula can be computed easily and explains how different parameters of the CCH affect the precision with which phase offsets are measured. We could show that the formula accounts very well for the variability of the phase estimates in simulated CCHs and also that it predicts accurately the error with which phase offsets are measured in the experimental dataset. Therefore, our results suggest that the proposed method can reliably estimate phase offsets together with their measurement errors.

The application of the method to the experimental dataset indicated that phase offsets can be measured with high precision. We obtained an average standard deviation of a phase estimate of about 0.2 ms, which yields 95% confidence intervals of about 0.8 ms. This suggests that phase offsets can be measured with submillisecond precision and thus, that small phase offsets (e.g., 1 ms) can often be distinguished from zero-offsets with confidence. Therefore, the proposed method can be used to investigate whether these offsets change dynamically in a task dependent way and thus, whether they contain information that might be relevant for cortical processing. The proposed statistical test can assess the difference between two paired sets of phase measurements and thus, can be helpful in such investigations.
Eq. (2) indicates that the precision of phase estimates depends on the parameters of the CCH. We investigated its performance for a range of parameters that was derived from our dataset. This range does not necessarily cover all possible values. For example, since our dataset included multiunit activity with relatively high firing rates, an application of the method to single units with low firing rates might involve different parameter ranges, and the noise around the cosine might not be Gaussian distributed. It might also be necessary to bin CCHs prior to the fitting procedure, in which case one should take into consideration the decrease in the precision with which the measurement error can be estimated.

The high precision with which phase offsets can be measured does not imply that phase offsets are stationary over the entire response window that is used to compute the CCH. For example, summation of two CCHs with different phase offsets can result in an intermediate offset that could also be determined with high precision. The presence of central peaks in un-normalized CCHs also does not imply precise neuronal synchronization but can reflect rate covariation between two signals (Perkel et al., 1967; Brody, 1999). In this case, the time offset of the peak represents the temporal delay of the rate covariation and is likely to require a window size, 2\(T\), that is much larger than 20 ms. Finally, the present methods are not exclusively applicable to CCHs but can be useful for any data analysis that requires assessing the positions of peaks in any dataset that complies with the assumptions of the present model.

In conclusion, the cosine fit and the associated formula for the measurement error of the phase shift can provide a useful tool for the analysis of phase offsets in CCHs. Thus, the proposed methods can be used to investigate temporal properties of neuronal responses and their role in cortical information processing.

Acknowledgements

We are thankful to Martha N. Havenith for help with the analysis in Fig. 3 and many valuable comments, and to Julia Biederlack for help with data acquisition. We also wish to thank Brooks Ferebee for important inputs at many stages of the project, including the derivation of Eq. (2), and Jörg Berns-Müller for help with the Matlab code. Finally, we thank Anton Wakolbinger and Wolf Singer for fruitful discussions and support.

Appendix A. Methods for data acquisition

A.1. Preparation

The cat was initially anesthetized with ketamine, and the anaesthesia was maintained with a mixture of 70% \(\text{N}_2\text{O}\) and 30% \(\text{O}_2\) supplemented with halothane (0.4–0.6%). The animal was paralysed with pancuronium bromide (Pancuronium, Organon, 0.15 mg kg\(^{-1}\) h\(^{-1}\)). All the experiments were conducted according to the guidelines of the Society for Neuroscience and German law for the protection of animals, approved by the local government’s ethical committee and overseen by a veterinarian.

A.2. Recording

Multi-unit activity (MUA) was recorded from a region of area 17 corresponding to the central part of the visual field by using a SI-based multielectrode probe (16 channels per electrode) supplied by the Center for Neural Communication Technology at the University of Michigan (Michigan probes) with inter-contact distance 200 \(\mu\text{m}\) (0.3–0.5 MΩ impedance at 1000 Hz). Signals were filtered between 500 and 3.5 kHz for extracting multi-unit activity (MUA), digitized with 32 kHz sampling frequency and stored in computer memory. All analyses were made on the basis of discrete spike events detected by a threshold that was set to a value of about two times the noise level. The probe was inserted in the cortex approximately perpendicular to the surface and allowed simultaneous recording from neurons at different cortical depths and along an axis tangential to the cortical surface. Fourteen MUA signals showed good responses to visual stimuli, orientation selectivity and overlapping receptive fields (RF). This resulted in a cluster of overlapping RFs that were stimulated simultaneously by a single visual stimulus.

A.3. Visual stimulation

Stimuli were presented binocularly on a 21″ computer screen (HITACHI CMH13ET) with 100 Hz refresh rate. To obtain binocular fusion the optical axes of the two eyes were first determined by mapping the borders of the respective RFs and then aligned on the computer screen with adjustable prisms placed in front of one eye. The software for visual stimulation was a commercially available stimulation tool, ActiveSTIM (http://www.ActiveSTIM.com). The stimuli consisted either of one white bar moving over a black background or consisted of two bars moving in different directions (60° difference). The bars always appeared at about 3° eccentricity of the center of the cluster of RFs and moved with a speed of 1/\(s\) such that they completely covered the cluster of RFs. In the stimuli with two bars, the bars crossed their paths at the center of the RF cluster. At each trial the stimulus was presented in total for 5 s, but only 2 s with strongest rate response were used for the analysis. In the six stimulation conditions the bars moved in the following directions (1) 30° and 330°; (2) 90°; (3) 150° and 210°; (4) 180°; (5) 30° and 150°; (6) 210° and 330°. Each stimulation condition was presented 20 times, different conditions being presented in a randomized order. In the present study only the conditions 1, 3 and 5 were used for the analyses.

Appendix B. Distribution of residuals

To investigate the distribution of the residuals, a cosine function was fitted to each CCH as shown in Fig. B.1A, and the distribution of the residuals (Fig. B.1B) was compared to a normal distribution by application of a Kolmogorov–Smirnov test. The distribution of the resulting \(p\)-values for all 91 CCHs in stimulation condition 1 is shown in Fig. B.1C. For data that conform with the model assumptions, the distribution of \(p\)-values is expected to be uniform on the interval between 0 and 1. The obtained \(p\)-values range within the whole interval from 0 to 1,
with only a slight tendency towards smaller values. Thus, the
normal distribution was considered appropriate to model the
residuals in the present dataset. Independence of residuals was
investigated with serial correlation up to a lag of 10 ms. Absolute
values of correlation coefficients stayed mostly below 0.2 (data
not shown).

Note that residuals can be considered independent only if the
CCH is computed on the basis of counts of discrete spike events
determined by a threshold criterion. If in contrast, CCHs are
computed directly on the filtered MUA signals, dependencies
determined by a threshold criterion. If in contrast, CCHs are
computed directly on the filtered MUA signals, dependencies
between residuals could be introduced. This is because, if the
low-pass filtering frequency is below the Nyquist frequency that
results from the digitization rate, adjacent data points of the
sampling process could be correlated.

Appendix C. Mathematical derivation of the asymptotic
variance

In step 1 we use trigonometric identities to transform the
expression in Eq. (1) into a linear combination and express \( \phi \) in terms
of the coefficients \( \beta_1 \) and \( \beta_2 \). In step 2, the variance of these
coefficients is derived with linear regression analysis, which in
the final step 3 allows the derivation of the variability of the
coefficients from the digitization rate, adjacent data points of the
MUA signals are dependent.

Step 1: \( \phi \) as a function of \( \beta_1 \) and \( \beta_2 \)

\[
A \cos(\phi - \psi) = A [\cos(\phi) \cos(\psi) + \sin(\phi) \sin(\psi)]
= \beta_1 \cos(\psi) + \beta_2 \sin(\psi)
\]  
(C.1)

with \( \cos(\phi) = \beta_1/A \) and \( \sin(\phi) = \beta_2/A \).  
(C.2)

This yields

\[
\phi = \frac{1}{A} \arcsin \left( \frac{\beta_1}{A} \right) \quad \text{for} \quad |\phi| < \frac{\pi}{2}
\]  
(C.3)

and

\[
\psi = \text{sgn}(\psi) \left( \frac{\pi}{2} + \frac{1}{A} \arcsin \left( \frac{\beta_2}{A} \right) \right) \quad \text{for} \quad |\phi| \geq \frac{\pi}{2}
\]  
(C.4)

We only compute the variance of the term in Eq. (C.3), because
the variances of the estimates in (C.3) and (C.4) are identical.

Step 2: Approximate variances of \( \hat{\beta}_1 \) and \( \hat{\beta}_2 \)

The approximate variances \( \sigma_1^2 \) and \( \sigma_2^2 \) of \( \hat{\beta}_1 \) and \( \hat{\beta}_2 \) are given by

\[
\sigma_1^2 := \text{Var}(\hat{\beta}_1) = \frac{2 \sigma^2}{N} \left[ 1 + \frac{\sin^2(\omega T)}{\omega^2 T^2} \right]^{-1},
\]  
(C.5)

and

\[
\sigma_2^2 := \text{Var}(\hat{\beta}_2) = \frac{2 \sigma^2}{N} \left[ 1 + \frac{\sin^2(\omega T)}{2 T} \right]^{-1}.
\]  
(C.6)

Proof. We use linear regression analysis to determine the coef-
ficients \( \hat{\beta}_1 \) and \( \hat{\beta}_2 \) and their asymptotic variances. Let \( m \) denote the time resolution (number of bins per time unit), \( -T \) and \( T \) the
borders of the CCH, and \( N := 2mT \) the number of data points
in the CCH. Let furthermore

\[
x_i := -T + (i - 1)/m \quad (i = 1, \ldots, N)
\]

denote the time shifts for which coincidences are counted in the
CCH. Thus, the continuous variable \( t \) is replaced by discrete
steps \( x_i \). Combining Eqs. (1) and (C.1), we get that the value of
the CCH at \( x_i \) can be written as

\[
Y_i := \text{CCH}(x_i) = \hat{\beta}_0 + \hat{\beta}_1 \cos(ax_i) + \hat{\beta}_2 \sin(ax_i) + \sigma Z_i,
\]  
(C.7)

where \( Z_i, i = 1, \ldots, N \) denote independent random variables with
standard normal distribution. We express Eq. (C.7) in ma-
trix notation

\[
Y = X \hat{\beta} + \sigma Z,
\]

where \( Y = (Y_1, \ldots, Y_N)^T, Z = (Z_1, \ldots, Z_N)^T, \hat{\beta} = (\beta_0, \beta_1, \beta_2)^T, \)
and the \( i \)th row of \( X \) is given by \( X_i := (1, \cos(ax_i), \sin(ax_i)) \)
for all \( i = 1, \ldots, N \). The distribution of \( \hat{\beta} \) is known from linear
regression analysis:

\[
\hat{\beta} \sim \mathcal{N}(\beta, (X'X)^{-1} \sigma^2).
\]

We compute \((X'X)^{-1}\) and get

\[
\sigma_1^2 := \text{Var}(\hat{\beta}_1) = \frac{\sigma^2 N}{\sum \cos^2(ax_i) - \left( \sum \cos(ax_i) \right)^2},
\]  
(C.8)

and

\[
\sigma_2^2 := \text{Var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum \sin^2(ax_i)}.
\]  
(C.9)
As $N$ is usually large, we replace the sums in (C.8) and (C.9) with their asymptotic integrals:

$$\sum_{i=1}^{N} \cos^2(\omega x_i) \to \int_{-T}^{T} \cos^2(\omega x)dx = \frac{N}{2} \left[ 1 - \frac{\sin(2\omega T)}{2\omega T} \right]$$

and

$$\sum_{i=1}^{N} \cos(\omega x_i) \to \int_{-T}^{T} \cos(\omega x)dx = \frac{N}{\omega T} \sin(\omega T)$$

and

$$\sum_{i=1}^{N} \sin^2(\omega x_i) \to \int_{-T}^{T} \sin^2(\omega x)dx = \frac{N}{2} \left[ 1 - \frac{\sin(2\omega T)}{2\omega T} \right].$$

The proof is then concluded by replacing the respective terms in Eqs. (C.8) and (C.9) with their approximate values. □

**Step 3: Approximate variance of $\hat{\omega}$**

$$\text{Var}(\omega^\prime) = \frac{2\sigma^2}{N\lambda^2} \left( \cos^2(\omega^\prime) \left( 1 - \frac{\sin(2\omega T)}{2\omega T} \right)^{-1} + \sin^2(\omega^\prime) \left( 1 - \frac{\sin(2\omega T)}{2\omega T} - \frac{2\sin^2(\omega T)}{a^2T^2} \right)^{-1} \right)$$

This implies

$$\text{Var} \left( \frac{\hat{\beta}_2}{\lambda} \right) = \frac{\beta_2^2}{\beta_1^2 + \beta_2^2} \text{Var} \left( \frac{1 + \sigma_Z \beta_1}{\beta_2 + \sigma_Z \beta_1} \right)$$

and

$$\text{Var} \left( \frac{\hat{\beta}_1}{\lambda} \right) = \frac{\beta_1^2}{\beta_1^2 + \beta_2^2} \text{Var} \left( \frac{1 + \sigma_Z \beta_2}{\beta_1 + \sigma_Z \beta_2} \right).$$

Inserting the right hand side of Eq. (12) into Eq. (10) and replacing $\beta_1, \beta_2, \sigma_1$ and $\sigma_2$ by the terms in Eqs. (C.2), (C.5) and (C.6) yields the claim. □

**Computer code for fitting a cosine**

In order to apply Eq. (2) to a CCH, one can use the following code. The CCH must be stored as an array of integers, CCH, of length $N$, containing the coincidence counts. The time lags for which coincidences are counted are stored in the array $x$. With the publicly available statistical analysis tool R, the commands

```r
#starting values
st <- c(A=1,phi=0,w=3,b0=mean(CCH))
#computation of estimates
estimates <- -coef(nls(CCH~A*cos(x-phi)+b0),start=st))
```

Proof. Eq. (C.3) tells us that $\omega^\prime = \arcsin(\beta_2/A)$. We use the fact that one can approximate $\arcsin(x)$ locally, i.e., around $\beta_2/A$, by a linear function

$$\arcsin(x) \approx \arcsin(\beta_2/A) + (\arcsin(\beta_2/A))/(x - \beta_2/A)$$

Thus, the variance of $\arcsin(\beta_2/A)$ can be approximated by

$$\text{Var} (\arcsin(\beta_2/A)) \approx \frac{1 - \beta_2^2/A^2}{(1 - \beta_2^2/A^2)^2} \text{Var} (\beta_2/A).$$

To approximate the variance of $\hat{\beta}_2/A$, we rewrite

$$\hat{\beta}_2 = \beta_2 + \sigma Z_i,$$

where $Z_i$ is standard normally distributed $(i=1,2)$. Due to (C.2), $(\beta_2/A)^2 + (\beta_2/A)^2 = 1$ and thus, $A = (\beta_2^2 + \beta_2^2)^{1/2}$. Therefore,

$$\text{Var} \left( \frac{\hat{\beta}_2}{\lambda} \right) = \text{Var} \left( \frac{\beta_2 + \sigma Z_i}{\sqrt{(\beta_1 + \sigma Z_1)^2 + (\beta_2 + \sigma Z_2)^2}} \right).$$

Now we transform the denominator into a term of the form $\sqrt{1 + x}$ and approximate $1/\sqrt{1 + x}$ in $x = 0$ by the first two terms of the Taylor expansion, $1 - \frac{1}{2}x$. For large $N$, $\sigma^2$ decreases at the order of $1/N$ and thus, is negligible as compared to $\sigma$, which decreases only at the order of $1/\sqrt{N}$.

References


