LETTER Communicated by Peter König

Spatiotemporal Structure in Large Neuronal Networks Detected from Cross-Correlation

Gaby Schneider
gaby.schneider@math.uni-frankfurt.de
Department of Computer Science and Mathematics, Johann Wolfgang Goethe University, Frankfurt (Main), Germany

Martha N. Havenith
havenith@mpih-frankfurt.mpg.de
Department of Neurophysiology, Max-Planck-Institute for Brain Research, Frankfurt (Main), Germany

Danko Nikolić
danko@mpih-frankfurt.mpg.de
Department of Neurophysiology, Max-Planck-Institute for Brain Research, Frankfurt (Main), and Frankfurt Institute for Advanced Studies, Johann Wolfgang Goethe University, Frankfurt (Main), Germany

The analysis of neuronal information involves the detection of spatiotemporal relations between neuronal discharges. We propose a method that is based on the positions (phase offsets) of the central peaks obtained from pairwise cross-correlation histograms. Data complexity is reduced to a one-dimensional representation by using redundancies in the measured phase offsets such that each unit is assigned a "preferred firing time" relative to the other units in the group. We propose two procedures to examine the applicability of this method to experimental data sets. In addition, we propose methods that help the investigation of dynamical changes in the preferred firing times of the units. All methods are applied to a sample data set obtained from cat visual cortex.

1 Introduction

Cortical neurons can fire in precise temporal relation to each other, producing repeatable spatiotemporal patterns (Mainen & Sejnowski, 1995; Lestienne, 1996; Singer, 1999; Abeles, Bergman, Margalit, & Vaadia, 2000; Reinagel & Reid, 2002; Ikegaya et al., 2004). A variety of methods has been proposed and used for detecting and investigating such precise firing patterns among large networks of neurons (for a review, see, e.g., Brown, Kass, & Mitra, 2004). Each of these methods is tuned to detect specific types of spatiotemporal relations, operating on different timescales, and

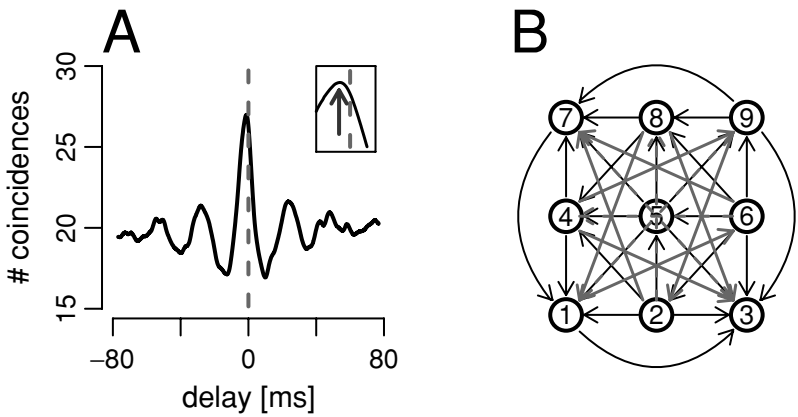

Figure 1: Definition and data structure of phase offsets. (A) A phase offset is defined as the time delay of the central peak's maximum in a CCH. Inset: Enlarged image of the CCH peak. (B) An illustration of the representational complexity that originates from pairwise analysis. For only nine simultaneously recorded units, 36 phase offsets are already measured.

each uses a different definition of what constitutes a pattern. For example, some methods focus on coincident events (Gerstein, Perkel & Dayhoff, 1985; Martignon, von Hasseln, Grün, Aertsen & Palm, 1995; Grün, 1996; Johnson, Gruner, Baggerly, & Seshagiri, 2001; Grün, Diesmann & Aertsen, 2002; Amari, Nakahara, Wu & Sakai, 2003; Schneider & Grün, 2003), while others register events that are delayed (Abeles & Gerstein, 1988; Abeles, 1991; Ikegaya et al., 2004).

In this study, we propose a method for detecting a particular type of spatiotemporal relations occurring between neurons with synchronized discharges, as defined by the presence of a center peak in their raw cross-correlation histograms (CCHs; Moore, Perkel, & Segundo, 1966; Perkel, Gerstein, & Moore, 1967). The method starts with the computation of CCHs between all pairs of spike trains and with the extraction of the center peak's positions by fitting cosine functions to the CCHs (for details, see Schneider & Nikolić, 2006; for an alternative method, see König, 1994). A shift of the center peak indicates that the two neurons tend to fire with a delay (see Figure 1A). If CCHs are associated with oscillatory activity (e.g., Gray, König, Engel & Singer, 1989; Engel, Kreiter, König, & Singer, 1991), these delays are also called *phase offsets*. Such delays are often small (2 ms or less) and thus have been considered equivalent to zero delays (Roelfsema, Engel, König, & Singer, 1997). However, we have shown in Schneider and Nikolić (2006) that these small delays can be statistically different from zero and thus that they can represent real temporal offsets between the firing events of pairs

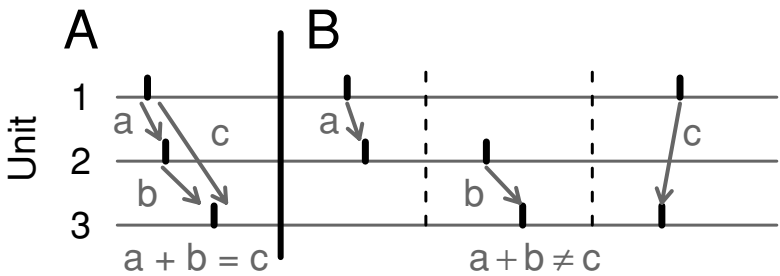

Figure 2: Additive and nonadditive spiking delays. (A) If the same spatiotemporal pattern among three units occurs repetitively (small jitter allowed), phase offsets in the corresponding CCHs (a–c) are additive. (B) An example in which spatiotemporal patterns are restricted to pairs of units and phase offsets are not additive.

of neurons. As a consequence, it appears worthwhile to investigate such delays within large data sets obtained in highly parallel recordings (compare also to König, Engel, Roelfsema, & Singer, 1995; Traub, Whittington, & Jefferys, 1997; Wennekers & Palm, 2000).

As Figure 1B illustrates, analyses of pairwise relations can become cumbersome when the number of neurons (units) in the data set increases. It is thus necessary to decrease the representational complexity of such data sets. The method presented here achieves this in the following manner. Pairwise phase shift measurements ($\binom{n}{2}$ pairs for n neurons) are collapsed into a one-dimensional representation that indicates the preferred time at which each neuron fires an action potential relative to the firing times of all other neurons. As a result, a complex data set is represented by a simple one-dimensional temporal map, also called a *linear configuration*.

This procedure is appropriate only when certain conditions are satisfied. The main precondition for a reduction of representational complexity is that the phase offsets obtained from different pairs of units have the property of additivity. That is, the offset that was measured between units A and C should correspond to the sum of offsets between units A and B and units B and C (illustrated in Figure 2A). If this condition is satisfied across the entire data set, a simple summation method can be applied to estimate the relative temporal positions of all units (see section 2). One should, however, be cautious when using such an analysis because, as illustrated in Figure 2B, additivity of phase offsets is not given by default. Therefore, the suitability of the additivity assumption should be investigated for each data set; the issues related to such tests are discussed in sections 2 and 3. Finally, phase offsets might change due to functionally relevant variables such as changes in stimulation conditions (e.g., König et al., 1995; Schneider & Nikolić, 2006) or due to the shift in the focus of attention or a behavioral event. In section 4

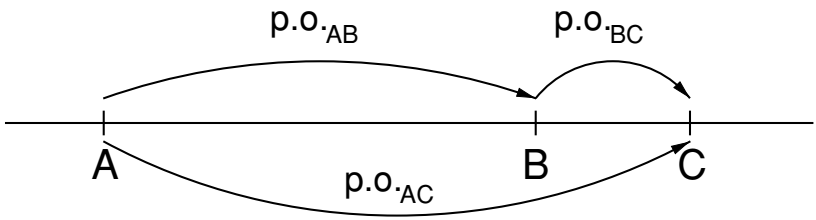

Figure 3: Linear configuration of three units on the time axis. If phase offsets (p.o.) are additive, the relative times at which units prefer to fire can be represented by delays between points in one temporal dimension.

we propose methods that can be used to compare two sets of phase offsets and investigate whether phase offsets change consistently across different measurements.

2 Stochastic Model for the Extraction of Linear Structure

2.1 Assumptions and Estimates. We will first show how additivity can be used to reduce the representational complexity of a set of phase offsets. Phase offsets between the units A, B, and C are additive if the following condition is satisfied:

$$\varphi_{AC} = \varphi_{AB} + \varphi_{BC}.$$

If this holds true for all subsets of pairs of units, the temporal relations between all units in the data set can be represented precisely by positioning all units on one temporal dimension (time axis). The position of a unit then indicates its "preferred firing time" relative to the other units. The absolute value of a phase offset can be read from the time axis as the distance between two preferred firing times (see also Figure 3), and its sign indicates the temporal order of the units on this axis.

Perfect additivity is not likely to be achieved in practice because even additive phase offsets would be measured with an error. It is therefore necessary to investigate how well one can represent the original phase offsets by positioning the units on one time axis. To this end, we first identify the most likely positions of the units on the time axis by using a maximum-likelihood (ML) approach. Thus, the units are positioned such that their pairwise distances (model distances) resemble the measured phase offsets as closely as possible.

A canonical set of assumptions needed for such positioning is illustrated in Figure 4. Here, we assume that n units $(1, 2, \ldots, n)$ have positions $\{x_1, x_2, \ldots, x_n\}$ on the time axis. The positions have mean zero, and the real delays between all unit pairs (i, j) are denoted by $\delta_{ij} := (x_j - x_i)$. The

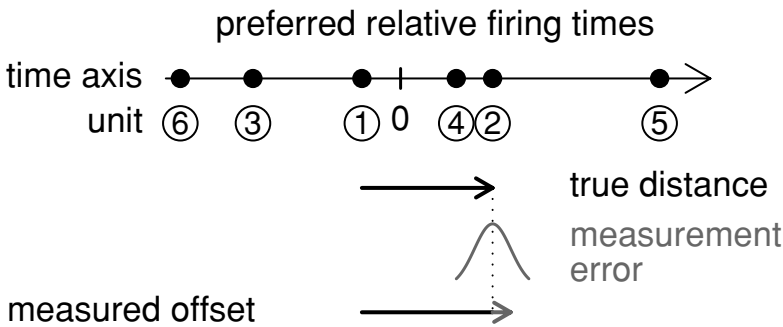

Figure 4: Model assumptions. Units are represented as points on the time axis with mean zero. The measured phase offset φ_{ij} of the distance $\delta_{ij} = (x_j - x_i)$ between units i and j is associated with a normally distributed measurement error σZ_{ij} with zero mean and variance σ^2.

measured phase offsets φ_{ij} are assumed to result as a sum of the real delays, δ_{ij}, and normally distributed measurement errors, which are independent across different phase offsets, have mean 0 and equal variances σ^2. Thus, for all unit pairs (i, j) with $1 \leq i < j \leq n$,

$$\varphi_{ij} = (x_j - x_i) + \sigma Z_{ij},$$

with independent $Z_{ij} \sim \mathcal{N}(0, 1)$. Note that there is only one phase offset for each pair of units because $\varphi_{ij} = -\varphi_{ji}$ (i.e., the CCH between units A and B is the mirror image of the CCH between units B and A).

Given those assumptions, the ML estimate of position x_k of unit k on the time axis can be computed as the normed sum of phase offsets between this unit and all other units under investigation (for a proof, see appendix B.1):

$$\hat{x}_k = \frac{1}{n} \sum_{\ell \neq k} \varphi_{\ell k}. \tag{2.1}$$

These estimates of unit positions remain unbiased even if the assumptions of normality and independence of measurement errors are violated or if the variances of different phase offsets are unequal. This is because the expected value of the sum is the sum of expectations. Thus, the only effect of the violations of model assumptions is that the estimates can no longer be interpreted as ML estimates.

From equation 2.1, it follows that the estimated model distance $\hat{\delta}_{ij}$ between the units i and j is

$$\hat{\delta}_{ij} = \hat{x}_j - \hat{x}_i = \frac{1}{n}\left(2\varphi_{ij} + \sum_{\ell \neq i,j} (\varphi_{i\ell} + \varphi_{\ell j})\right). \tag{2.2}$$

Thus, the model distance between the units i and j is estimated by a weighted average of their direct distance and all indirect distances of paths of length two. The direct measurement φ_{ij} contributes twice as much to this weighted average as each indirect measurement $\varphi_{i\ell} + \varphi_{\ell j}$. This accounts for the fact that the direct estimate is affected by only one error of measurement, while every indirect estimate is contaminated by two errors of measurement.

The estimates $\hat{x}_i$ in the resulting one-dimensional temporal map, $\{\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n\}$, minimize the sum of squares of differences between the measured phase offsets φ_{ij} and the offsets estimated by the model, $\hat{\delta}_{ij}$, that is,

$$Q := \sum_{i<j} (\varphi_{ij} - \hat{\delta}_{ij})^2 = \sum_{i<j} (\varphi_{ij} - (\hat{x}_j - \hat{x}_i))^2 \stackrel{!}{=} \min.$$

This can be used to estimate σ^2:

$$\hat{\sigma}^2 = \frac{1}{\binom{n-1}{2}} \cdot Q. \tag{2.3}$$

It can be shown that this estimate is unbiased—that it neither over- nor underestimates σ^2 in expectation (see appendix B.2).

With equation 2.3, the measurement error is computed on the basis of the agreement across phase offsets: the higher the degree of additivity among a set of phase offsets, the smaller the estimated measurement error. We will therefore also refer to the measurement error as the *error of additivity*. This quantity can also be used to compute

$$\sigma^2_{\hat{x}} := \mathrm{Var}(\hat{x}) = \frac{(n-1)}{n^2} \cdot \sigma^2, \tag{2.4}$$

which takes into account the error of additivity in order to indicate the precision with which a unit can be positioned on the time axis.

Equation 2.4 still holds even if measurement errors are not distributed normally. However, its utility can be affected if measurement errors are dependent or if they have different variances. Dependence of measurement errors will be discussed in section 4.2. If measurement errors have different

variances, each unit has its own estimation error, and thus the global σ from equation 2.3 represents only the average variability. In this case, an estimate of the error can be computed for each individual unit k by using the following equation:

$$\mathrm{Var}(\hat{x}_k) = \frac{1}{n^2} \cdot \sum_{\ell \neq k} \sigma_{\ell k}^2. \tag{2.5}$$

This equation cannot be computed directly from a set of phase offsets but requires that the individual variances $\sigma_{\ell k}^2$ of the measurement errors are estimated from a different source of information than the additivity error. Thus, if the units do not differ strongly in their individual variances, it is more practical to use the global estimate rather than the individual estimates of the positioning error.

2.2 Application to a Sample Data Set. In this section we apply the model to a sample data set that consists of multiunit activity recorded simultaneously from 14 electrodes in the cat primary visual cortex in response to six different stimulation conditions (for details on experimental methods, see appendix A).

For all CCHs in stimulation conditions 1 to 3, the positions of the center peaks in the CCHs were estimated with high precision ($\sim$1 ms) by using the methods described in Schneider and Nikolić (2006). The position of each central peak was determined by fitting a cosine function and extracting the point at which it reaches its maximum. The resulting distribution of phase offsets for all 91 pairs in stimulation condition 1 is shown in Figure 5A. The corresponding distributions in conditions 2 and 3 were similar (data not shown).

We positioned the units on the time axis by using equation 2.1 (see Figure 6A). The resulting positions span a total distance slightly larger than 2 ms. The consistency across the phase offsets, indicated by the global estimate in equation 2.4 (additivity error: $\hat{\sigma}^2 \approx 0.04$; $\hat{\sigma}_{\hat{x}}^2 \approx 0.0026$), is illustrated by the width of the normal distributions drawn in black above each position. These suggest a high degree of separability between the preferred times at which the units fire action potentials. For example, unit 2 is likely to fire before unit 11, while the preferred firing times of units 7 and 10 seem indistinguishable.

The precision with which the units' positions can be estimated may differ across units if the variances of the phase measurements differ across unit pairs. Therefore, we also estimated the measurement errors for each individual position by using equation 2 from Schneider and Nikolić (2006). This equation describes analytically the precision with which a phase offset of a CCH can be determined with respect to the variability of the coincidence counts in the CCH, also taking into account the oscillation frequency and

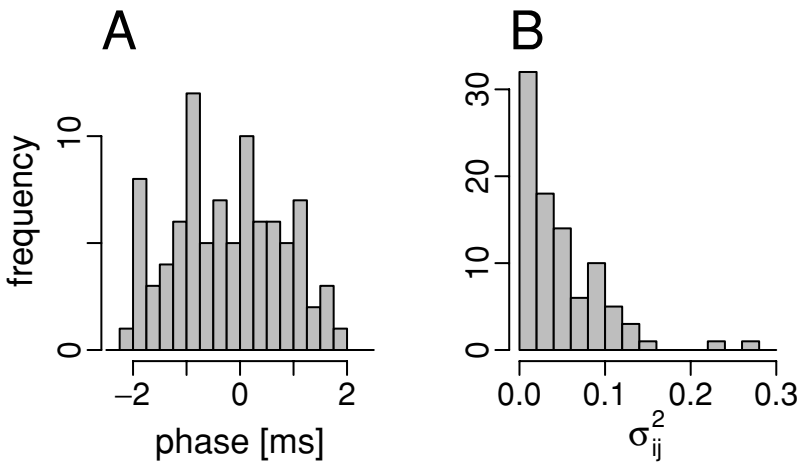

Figure 5: Distributions of 91 phase offsets (A) and the variances of their measurement errors (B) obtained in response to two moving bars that cross at the center of the receptive field (stimulation condition 1). Detailed description of the stimuli is provided in appendix A, and the sketches of the stimuli are provided in Figure 6.

amplitude of the central peak. The resulting histogram of estimated measurement errors of phase offsets is shown in Figure 5B. Separate estimates of σ^2_{ij} in equation 2.5 resulted in a small variability of the precision estimates of the unit positions (minimum, $\hat{\sigma}^2_{\hat{x}} \approx 0.001$ for units 5 and 8; maximum, $\hat{\sigma}^2_{\hat{x}} \approx 0.007$ for units 6 and 9). The normal distributions with their widths and heights adjusted to correspond to the variances of individual positions are added in gray above each unit position in Figure 6A. One can see that the additional consideration of the differences in variances does not affect strongly the separability between the units. Therefore, the general model that uses only the global variance provides a reasonably accurate representation of the precision with which a unit's position can be estimated. Hence, for the remaining stimulation conditions 2 and 3, we show only the global estimates of the variability (see Figures 6B & 6C).

To obtain further understanding of the relations between the phase offsets, we propose to investigate how faithfully the resulting linear configuration represents the original data set. This can be assessed by plotting the model distances—the pairwise delays estimated within the model, $\hat{\delta}_{ij}$—against the measured phase offsets, φ_{ij}, for all unit pairs. Figure 7A shows the corresponding plot for stimulation condition 1. The close scattering around the diagonal and the high correlation ($r = 0.98$) indicate that the model represents the measured phase offsets well ($r = 0.97$ and $r = 0.93$ in conditions 2 and 3, respectively). This suggests that the concept of

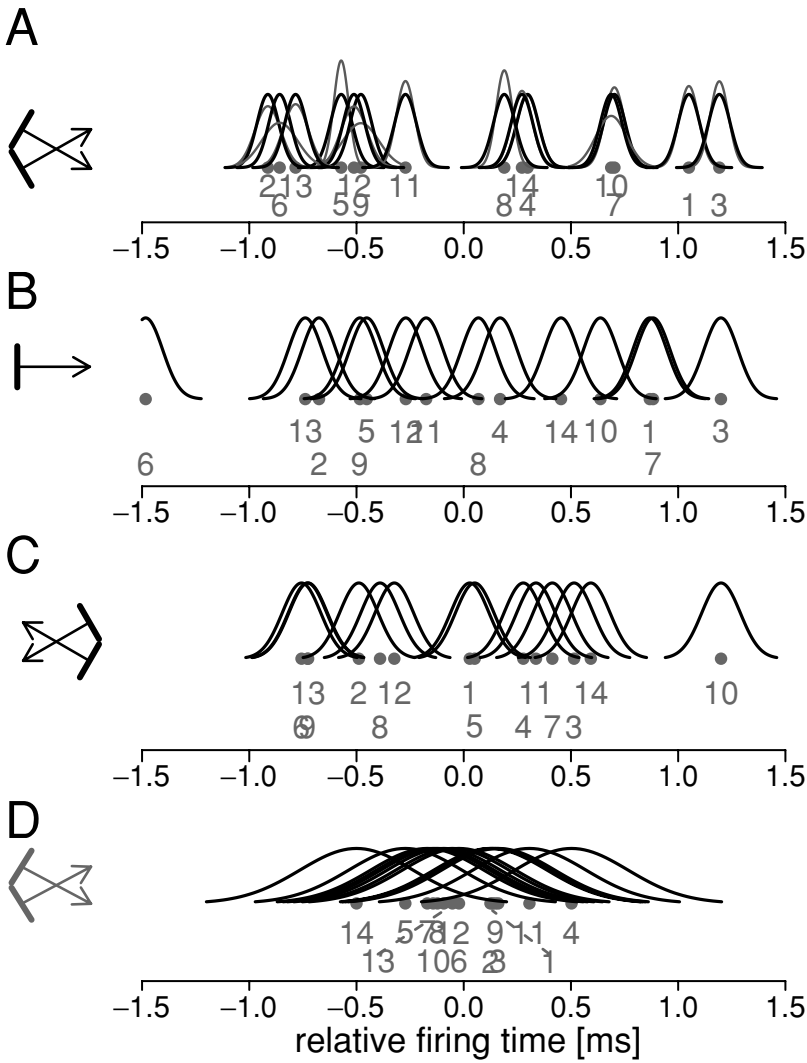

Figure 6: Linear configurations on the time axis. The dots denote the estimated positions of the units, and the black curves indicate localization errors (see equation 2.4). Gray curves indicate localization errors under heteroscedasticity. (A–C) Original data sets obtained from stimulation conditions 1 to 3. Stimulus configurations are indicated on the left side of each panel. (D) Phase offsets in the data set of stimulation condition 1 are permuted randomly.

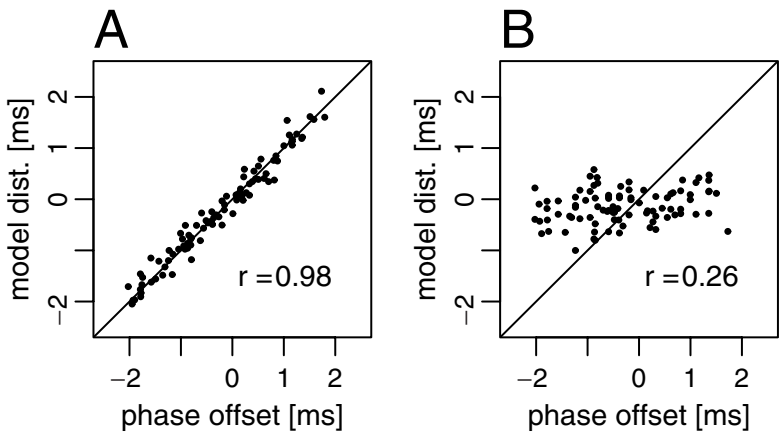

Figure 7: Comparison between measured phase offsets and the corresponding model distances derived from those phase offsets. (A) Original data set. (B) Original phase offsets permuted randomly.

additivity can be useful in decreasing the representational complexity of the temporal structure within large data sets.

3 Consistency Analysis

A high correlation between phase offsets and model distances indicates that the method presented here provides a reasonable representation of the data structure. However, the interpretation of the data representation has to take into account two distorting effects that could be caused by assumptions of the model. First, phase offsets are assumed to be perfectly additive except for measurement errors. As a consequence, a certain degree of additivity is also extracted from data sets that are not inherently additive. To address this issue, we compare the linear configuration obtained from the original data set to a configuration derived from randomly permuted data sets (see section 3.1). Second, the model assumes that phase offsets from different subsets of units agree on the global linear configuration derived from the entire data set. This could hide inconsistencies between different subnetworks. We therefore present a method that can be used to investigate the consistency of phase offsets across different subnetworks of neurons (see section 3.2).

3.1 Consistency in Permuted Data Sets. We permuted the phase offsets estimated in stimulation condition 1 by randomly assigning phase offsets to pairs of units. This procedure destroys additive structure but maintains the empirical distribution of phase offsets. If the linear configuration in

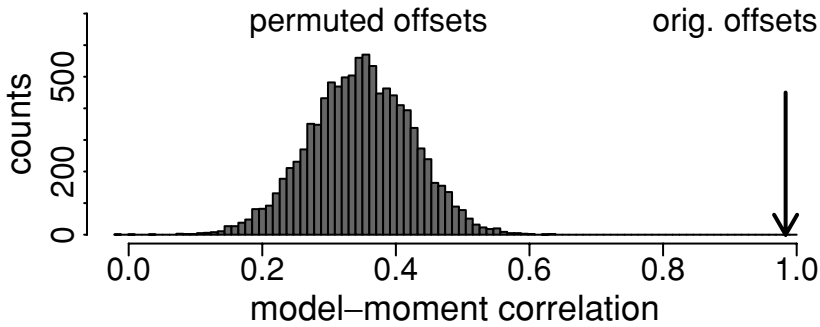

Figure 8: Distribution of correlation coefficients between model distances and phase offsets obtained from 10,000 permutations of the data set obtained in stimulation condition 1. Arrow: correlation coefficient obtained in the original data set ($r = 0.98$).

Figure 6A is entirely imposed by our method and not present in the data set, application of the method to a permuted data set should result in similar distances of units and similar measurement errors. Vice versa, if the latter does not apply, the additivity model grasps important aspects of the data structure. In this case, the positions of the units for the permuted data set should be much closer to each other because a sum of randomly assigned positive and negative phase offsets tends to average out (compare equation 2.1).

The results of such a permutation test are shown in Figures 6D, 7B, and 8. The units are positioned closer to zero for the permuted phase offsets (see Figure 6D) than they were in the original data set (see Figure 6A), and the distributions that indicate the precision of the estimates are much broader and overlap to a higher degree. The large width of these distributions indicates a high disagreement between measured phase offsets and model distances (see equation 2.4) and, thus, poor representation of permuted phase offsets by the one-dimensional model. This can also be seen in the corresponding scatter plot (see Figure 7B). The points are no longer scattered along the diagonal, and the range of model distances is much narrower. This makes the representation of larger phase offsets impossible. Accordingly, the resulting correlation coefficient between permuted phase offsets and model distances is small ($r = 0.26$) but positive, reflecting a small fraction of additivity that is imposed by the estimation method.

The permutation procedure was repeated 10,000 times, and the resulting distribution of correlation coefficients between the permuted phase offsets and their corresponding model distances is shown in Figure 8. All correlation coefficients were far below the value $r = 0.98$ obtained for the original data. Thus, by far the strongest linear structure among all 10,001

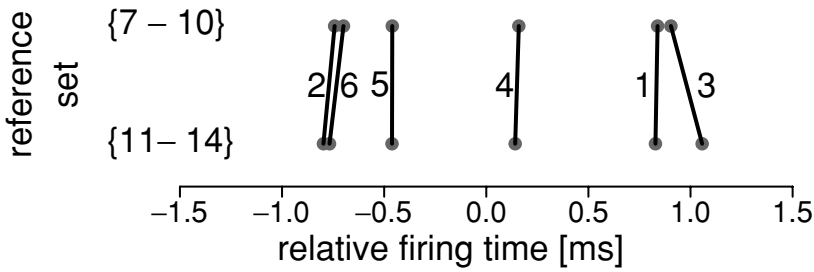

Figure 9: Linear configurations of target units 1 to 6 determined from the phase offsets to the reference units 7 to 10 and 11 to 14, separately.

investigated constellations was obtained from the original constellation, a structure that is therefore highly unlikely to be obtained by chance.

3.2 Consistency Across Subnetworks. To investigate the consistency of phase offsets across different subnetworks, we propose the following procedure. Choose one subset of units (e.g., units 1–6) as the target set that will be positioned on the time axis. Subdivide the remaining units into two reference sets of similar sizes (e.g., units 7–10 and 11–14), and use those sets to separately estimate two linear configurations for the targets, each configuration being obtained from a different reference set. High agreement between the two linear configurations then indicates consistency across subnetworks. To derive linear configurations from a subset of units, only the offsets to one reference set are summed up in equation 2.1. Thus, the position of target unit i derived from ℓ reference units $k_1, \ldots, k_\ell$ is $\frac{1}{(\ell+1)}(\varphi_{k_1,i} + \varphi_{k_2,i} + \cdots + \varphi_{k_\ell,i})$. Note that the resulting linear configurations are derived in relation to different reference sets and therefore do not have the same means. Thus, they need to be recentered at zero prior to further comparisons.

Figure 9 shows the estimated positions of units 1 to 6 in stimulation condition 1 derived from the two reference sets comprising units 7 to 10 and 11 to 14, respectively. Although each position is estimated from only four measurements, the order of units does not change, and distances between temporal positions are conserved to a high degree. Investigations of other combinations of target and reference sets led to similar results (data not shown). Thus, already small data sets can contain highly consistent information about the temporal structure in neuronal activity. Consequently, estimates of temporal maps resulting from the entire data set should yield an interpretable linear configuration that faithfully reflects the data set structure for all units.

4 Comparison of Two Linear Configurations

Phase offsets could represent a particularly interesting neuronal code if their configuration depends on the stimulation condition or behavior (Hopfield, 1995; VanRullen & Thorpe, 2002). The first question that needs to be addressed in this respect is whether linear configurations change across stimulation conditions. If the changes in the phase offsets occur solely due to independent errors of measurement, these changes are also inconsistent across phase offsets, and thus the size of the changes can be evaluated with the estimate of the measurement error in equation 2.3—the error in additivity. In contrast, if the changes result from changes in the relative position of the units, phase offsets change in a consistent manner, preserving additivity. In this section, we propose methods for investigating the consistencies in the changes of phase offsets and thus investigating whether there is a statistically significant amount of change that exceeds the variability resulting from the additivity errors.

To this end, we propose two methods. First, one can obtain a graphical representation of the differences between two configurations and include an indicator of the positioning error based on the error of additivity (see section 4.1). This graph can provide information about the extent of the changes and the identity of the units involved in these changes. The graphical analysis can then be complemented by a statistical test. To this end, we propose an ANOVA approach (see section 4.2). (For an alternative test independent of the additivity assumption, see Schneider & Nikolić, 2006.)

Both the graphical analysis and the ANOVA are applied to the sample data set. To investigate the stability of responses to identical stimuli, the 20 presentations (trials) of the same stimulus were subdivided into two subsets (odd and even trials with respect to the order of stimulus presentation). The comparisons between different stimulation conditions were based on all 20 trials (for details, see appendix A).

4.1 Displaying Differences Between Two Linear Configurations. One can compare two linear configurations of the same set of units by plotting the estimated positions derived from two separate data sets against each other (Figures 10 and 11A–C). Close clustering of the points around the main diagonal indicates little or no differences between the linear configurations. To this graph we add the following estimate of the positioning error. For one unit k, the size of the difference between the two positions, $(\hat{x}_k^{(1)}\ \&\ \hat{x}_k^{(2)})$, can be measured in terms of its variance, $\sigma_D^2 := \mathrm{Var}(\hat{x}_k^{(1)} - \hat{x}_k^{(2)})$. If the unit positions do not change more than is accounted for by the error of additivity, then

$$\sigma_D^2 = \frac{(n-1)}{n^2} \cdot \left(\sigma_1^2 + \sigma_2^2\right). \tag{4.1}$$

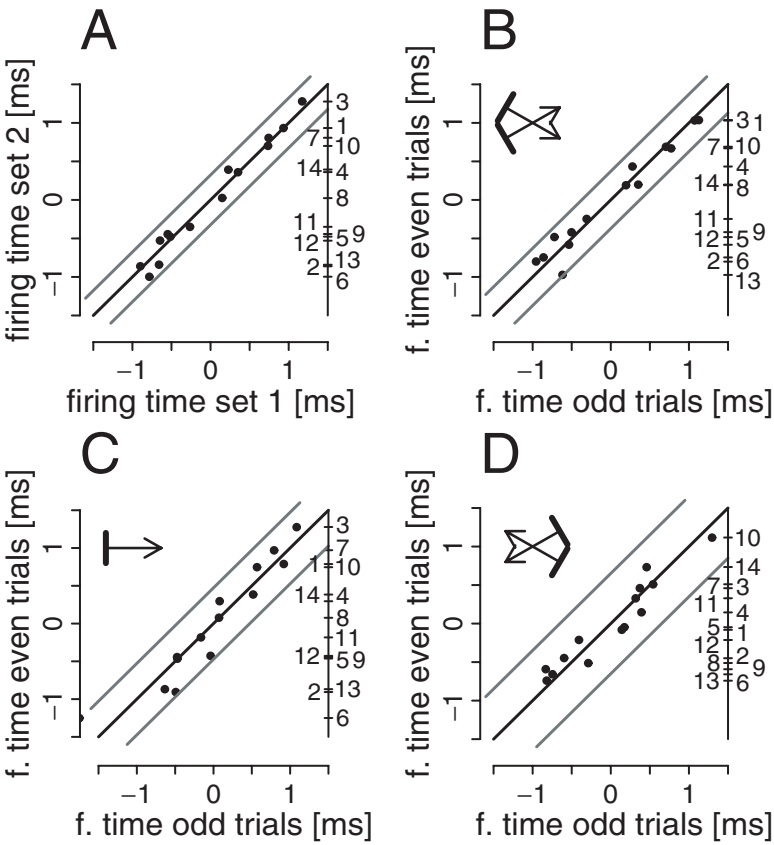

Figure 10: Comparisons of configurations of preferred firing times (f. times). (A) Configurations derived from two simulated sets of phase offsets that differed only in the random noise added to the measurements but not in the underlying temporal structure. (B–D) Configurations obtained from odd and even trials in stimulation conditions 1 to 3. Stimulus configurations are indicated on the upper left side of each panel. Individual units are represented by points and labeled on the right of each graph. Lines parallel to the diagonal indicate error bands ($\pm 2\hat{\sigma}_D$) computed with equation 4.1.

The terms σ_1 and σ_2 indicate the additivity errors in the two data sets (see equation 2.3) and do not have to be equal. The borders indicating approximate 95% confidence intervals (i.e., $2\sigma_D$) are then plotted parallel to the diagonal. Those points that lie outside this band indicate that the corresponding units changed their positions to a higher degree than what would be expected by the additivity errors.

To illustrate the applicability of such graphs, we simulated two sets of phase offsets. We used the linear configuration obtained from stimulation condition 1 and added to each pairwise distance an artificial measurement error: independent and normally distributed random noise with the same variance as estimated from the data set with equation 2.3 ($\hat{\sigma}^2 \approx 0.04$). The graphical approach was then used to compare the two linear configurations derived from the two simulated sets of phase offsets. Figure 10A shows that the units remain within the error borders. Thus, the graphical method does not indicate changes in the units' positions when the linear configuration remains stable.

By using this method, we compared linear configurations that were obtained under both identical and different stimulation conditions. The positions of the units across odd and even trials of the same stimulation condition (conditions 1–3) are shown in Figures 10B to 10D, together with the error lines. The positions of almost all units remained within the error bands, indicating that the variabilities of the positions for the most part did not exceed the variabilities expected by the measurement errors.

In Figures 11A to 11C, we show three pairwise comparisons in linear configurations across different stimulation conditions (conditions 1–2, 1–3 and 2–3, respectively). One can see that in the comparisons involving stimulation condition 3 (Figures 11B and 11C), about half of the units lie outside the error bands. Thus, in contrast to the comparisons between identical stimulation conditions, the changes in the positions across different stimulation conditions can be much larger than what would be expected by the measurement errors. This, however, does not have to apply to all cases, as the comparison between conditions 1 and 2 indicates. With the exception of one unit, there are no changes in positions that exceed the additivity error.

Changes between different stimulation conditions are often also visible in the original CCHs. For one pair of units (3–5), we show the CCHs obtained in stimulation conditions 1–3 in Figures 11D to 11F, respectively. One can see that the point at which the fitted cosine function reaches its maximum is similar in conditions 1 and 2 but changes for condition 3 (for a direct statistical analysis of raw phase offsets, see Schneider & Nikolić, 2006).

In conclusion, the graphical method can be a useful tool for comparing linear configurations. The method illustrates the degree to which the unit positions change relative to the errors in additivity and visualizes the directions of these changes. In the data set, application of the method indicates that temporal positions change much less in response to identical stimuli than in response to different stimuli.

However, the graphical approach cannot be used directly for statistical inferences. This is because the method does not correct for multiple comparisons among different conditions (type I error) or for the dependencies among position estimates of different units introduced by setting the sum of unit positions to zero. Thus, if one unit moves in one direction, the temporal positions of all remaining units move in the opposite direction. As

a consequence, the graphical method cannot provide a rigorous statistical tool to decide whether the changes between two configurations exceed the error in additivity. This question can be addressed much more accurately by an ANOVA approach, which we discuss in the following section.

4.2 Statistical Test for Two Linear Configurations. We now introduce a test with which the graphical representations discussed in the preceding section can be analyzed statistically. Changes across two linear configurations can be evaluated with respect to the errors of additivity by using a general ANOVA approach. Mathematical details are provided in appendix B.3; here we outline the main computational steps.

For each of the two sets of phase offsets ($k = 1, 2$), one has to estimate the linear configuration of units $\{x_1^{(k)}, \ldots, x_n^{(k)}\}$ (see equation 2.1), the pairwise model distances, $\hat{\delta}_{ij}^{(k)}$ (see equation 2.2), and the measurement error, $\hat{\sigma}_{(k)}^2$ (see equation 2.3). Under the null hypothesis, the two configurations are identical, and model distances and phase offsets differ only due to measurement error. Then the test statistic

$$F = \frac{1}{\hat{\sigma}_{(1)}^2 + \hat{\sigma}_{(2)}^2} \cdot \frac{1}{n-1} \sum_{i<j} \left(\hat{\delta}_{ij}^{(1)} - \hat{\delta}_{ij}^{(2)} \right)^2 \tag{4.2}$$

is Fisher distributed with $(n-1)$ and $(n-1)(n-2)$ degrees of freedom (see appendix B.3). In contrast, if the two configurations differ, then F is increased by larger differences in model distances between the two data sets, $(\hat{\delta}_{ij}^{(1)} - \hat{\delta}_{ij}^{(2)})$. One can intuitively interpret F as the squared difference between the linear configurations measured in units of the measurement error, $\hat{\sigma}_{(1)}^2 + \hat{\sigma}_{(2)}^2$.

We applied this ANOVA to the data sets presented in Figures 10 and 11; the results are shown in Table 1. Qualitatively, the F- and the p-values indicate the same relations as the graphical representations in Figures 10 and 11: the changes in the relative firing times are much stronger between responses obtained to different stimulation conditions than between odd and even trials of the same stimulation condition.

Two comparisons of identical stimulation conditions (conditions 1 and 2) show significant p-values (at $\alpha = 0.05$), indicating that the changes across odd and even trials are larger than what is accounted for by the additivity

Figure 11: Comparison of configurations across stimulation conditions. (A–C) Three pairwise comparisons between different stimulation conditions: 1 to 2, 1 to 3, and 2 to 3 (notation as in Figure 10). (D–F) Raw CCH counts for the unit pair 3 to 5 in stimulation conditions 1 to 3 in the original time resolution of 1/32 ms. Black: Cosine functions fitted to the central peak (i.e., ±10 ms) of the CCH counts.

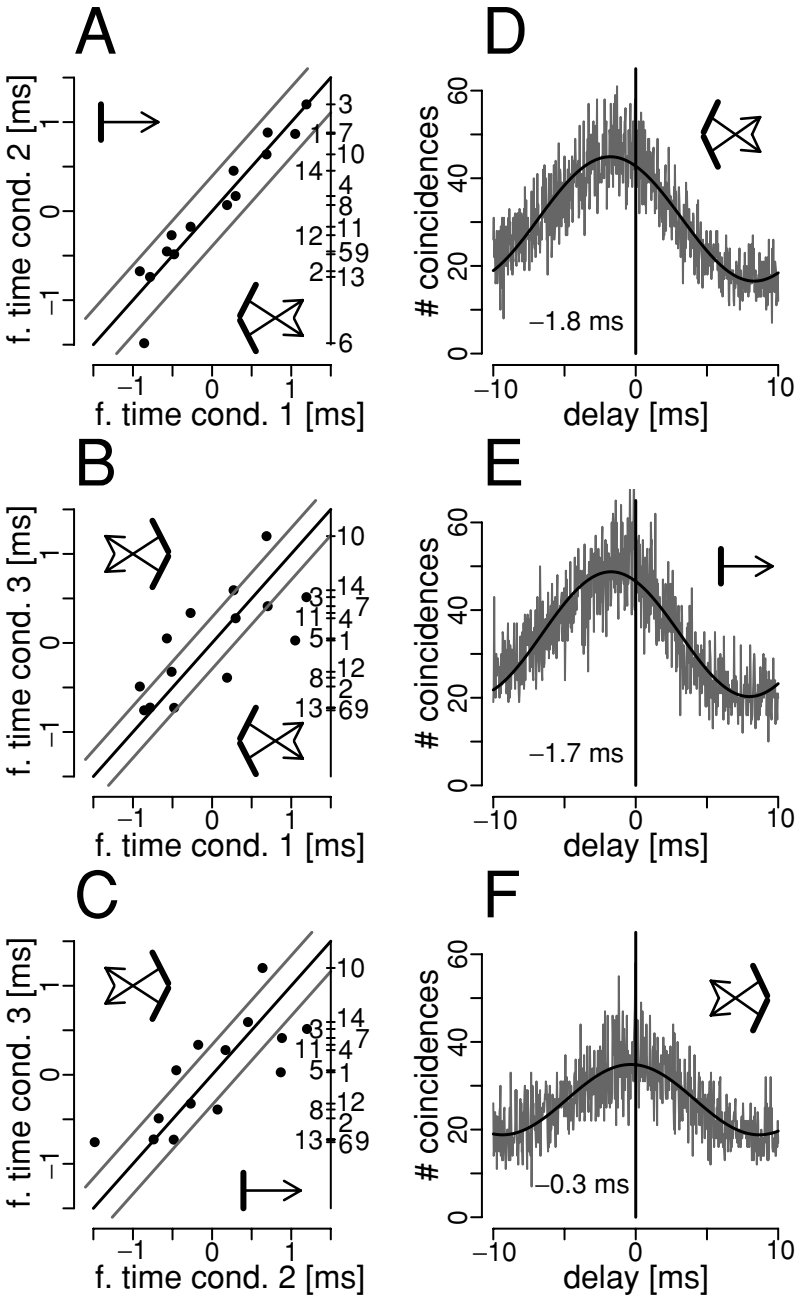
A
f. time cond. 2 [ms]
f. time cond. 1 [ms]
B
f. time cond. 3 [ms]
f. time cond. 1 [ms]
C
f. time cond. 3 [ms]
f. time cond. 2 [ms]
D
coincidences
delay [ms]
−1.8 ms
E
coincidences
delay [ms]
−1.7 ms
F
coincidences
delay [ms]
−0.3 ms

Table 1: Results of ANOVA Applied to a Sample Data Set Tested for Changes in Linear Configurations.

Identical Stimuli				*Different Stimuli*			
Cond.	*F*	*p*	*p**	*Cond.*	*F*	*p*	*p**
1 vs. 1	2.0	0.027	0.047	1 vs. 2	5.3	0	0
2 vs. 2	2.7	0.002	0.006	1 vs. 3	22.5	0	0
3 vs. 3	1.2	0.283	0.288	2 vs. 3	15.4	0	0

Notes: For $n = 14$ units, every data set contains $\binom{n}{2} = 91$ phase offsets, resulting in $(n-1) = 13$ and $(n-1)(n-2) = 156$ degrees of freedom. $p = 0$ indicates that the respective values were smaller than 10^{-7} (1-2: $p < 10^{-7}$, 1-3 and 2-3: $p < 10^{-16}$). p^*-values are derived from 10,000 simulations to correct for inequality of variances. $p^* = 0$ indicates that not a single simulation, out of 10,000, showed an F-value larger than the one obtained in the experimental data set.

errors computed within each of the two configurations. This result suggests that violations of the model assumptions might have increased the type I error. It is therefore necessary to discuss whether the results might be affected by violations of the model assumptions.

A violation of the normality assumption is unlikely to have affected the results. In the data set, measurement errors showed no deviations from the normal distribution (Schneider & Nikolić, 2006), and ANOVA is highly robust to the violations of the normality assumption (e.g., Pearson, 1931).

ANOVA is also highly robust to the violation of the assumption that the variances are equal (e.g., Horsnell, 1953). We nevertheless estimated the extent to which unequal variances could have affected our findings because this assumption was violated to a certain degree in the data set (see section 2.2 and Figure 5B). The p-values that correct for inequality of variances can be obtained by the following simulation procedure. For two sets of estimated phase offsets $\{\varphi_{ij}^{(1)}\}_{i<j}$, $\{\varphi_{ij}^{(2)}\}_{i<j}$ and their individual variances $\{\sigma_{ij}^{(1)}\}_{i<j}$, $\{\sigma_{ij}^{(2)}\}_{i<j}$, estimate first the global linear configuration as the mean of the configurations of the compared sets. Next, phase offsets are generated by adding independent and normally distributed measurement error with variance $(\sigma_{ij}^{(k)})^2$ to each model distance $\delta_{ij}^{(k)}$, $k = 1, 2$. With this procedure, we simulated 10,000 data sets and performed ANOVA by using equation 4.2. The percentage of simulations with a larger F-value than what was found in the original data set was used as the corrected p-value.

The results are shown in Table 1 (column indicated with p^*). The correction for unequal variances increased the p-values in the comparisons between odd and even trials. However, this increase was relatively small, and the changes within stimulation conditions 1 and 2 remained significant. This suggests that in this data set, inequality of variances was not the main

factor responsible for the significant changes within stimulation conditions 1 and 2.

Finally, the model assumes that measurement errors are independent. With regard to this, it is necessary to discuss possible sources of variability that are dependent such that the changes across phase offsets are additive. Dependent variability of phase offsets emerges if the primary source of variability is the position of units rather than the individual phase offsets. The analysis methods intepret such dependent changes in phase offsets as differences in the linear structure, not as measurement error. Any additive variability in phase offsets is included in the position of the units by equation 2.1 and thus not included in the measurement error by equation 2.3. Thus, consistent and dependent changes in phase offsets are the most likely reason for the significant ANOVA comparing odd and even trials in Table 1.

This property of the test does not constitute a violation of the assumption of independent measurement errors and thus does not invalidate the results of the ANOVA. This is because the methods are designed to investigate only the consistency between phase offsets, not the variability of the units' positions across repeated measurements. Therefore, the results obtained should be interpreted accordingly. The proper interpretation is that the rate of type I error does not increase, but instead, the positions of the units vary slightly more than expected from the error in additivity. The questions related to the variability of the unit positions across repeated measurements should be addressed by conventional statistical methods. The methods used here provide the means to obtain these temporal positions and test whether they change in a sufficiently consistent way to warrant further investigation.

5 Discussion

In this letter, we propose a method for analyzing temporal relations between the firing events of large groups of simultaneously recorded neurons. The method uses pairwise temporal relations defined as the positions of center peaks in CCHs and assumes that these relations are additive across all pairs. If the data set complies with this assumption, it is possible to represent the underlying spatiotemporal relations on a single temporal dimension, which then indicates the preferred times at which neurons fire action potentials relative to each other. We present a graphical tool as well as a statistical test that can be applied to data in order to investigate whether changes of such spatiotemporal maps across different measurements are consistent across phase offsets.

The method we have presented requires the existence of prominent central peaks in CCHs and is therefore highly related to the concept of synchronized firing events that occur across groups of neurons (Abeles, 1982; Diesmann, Gewaltig, & Aertsen, 1999; Singer, 1999). Synchronous events

have been found to occur with a precision of up to a few milliseconds (Riehle, Grün, Diesmann, & Aertsen, 1997; Grün, Diesmann, Grammont, Riehle, & Aertsen, 1999), and linear configurations may thus be extracted on a timescale that is finer than the maximal delay up to which events have been characterized as coincident. Another property of the method presented here is that it does not require the delays between firing events to repeat as exact replicas of each other. Instead, it is sufficient that the delays cluster around a certain value that can be obtained from a CCH. Thus, the method detects spatiotemporal patterns on the basis of events that are allowed to jitter over time and may therefore be invisible in raw spike trains. Also, temporal relations between unit pairs do not have to be stable over time. Instead, phase offsets are defined as those delays observed predominantly between a pair of units in the chosen analysis window and may thus result from temporally inhomogeneous processes (Vaadia et al., 1995).

The model that extracts the spatiotemporal relations assumes homoscedasticity and independent and normally distributed measurement errors and is thus compatible with a standard ANOVA approach. This allows for high flexibility in investigating various variables that might affect the relative firing times of neuronal cells. Finally, the method has the advantage of being applicable to the activity of both single units and multiple units, as well as to any set of continuous signals that show preferred delays in their cross-correlation, prominent examples being the local field potential (e.g., Gray, Engel, König, & Singer, 1992; Roelfsema et al., 1997) and the electroencephalogram (e.g., Sauseng et al., 2005).

Perhaps most surprising is our finding that the temporal map indicating the preferred times at which units fire action potentials relative to each other can be determined with a precision much higher than previously reported for the visual cortex (König et al., 1995; Roelfsema et al., 1997). This precision can take values smaller than 1 millisecond and is achieved partially by the fitting procedure used to estimate the positions of the center peaks (Schneider & Nikolić, 2006) and partially due to the integration of multiple pieces of information obtained from different CCHs. As we could show, this allows detecting unusually small changes in the preferred firing times.

The question whether small spiking delays and their resulting spatiotemporal relations have functional significance is beyond the scope of this study. However, we believe that the proposed methods will help to address these issues as they provide important tools for the detection, display, and analysis of temporal relationships between spiking events of a large number of neurons.

Appendix A: Methods for Data Acquisition

A.1 Preparation and Recordings. Anesthesia was induced with ketamine and after trachiotomy, was maintained with a mixture of 70%

N_2O and 30% O_2, and with halothane (0.4–0.6%). The cats were paralyzed with pancuronium bromide applied intravenously (Pancuronium, Organon, 0.15 mg kg-1 h-1). Multiunit activity (MUA) was recorded by using an SI-based 16-channel probe (organized in a 4×4 spatial matrix), which was supplied by the Center for Neural Communication Technology at the University of Michigan (Michigan probes). The probe had intercontact distances of 200 μm (0.3–0.5 MΩ impedance at 1000 Hz). Signals were amplified 1000 times, filtered between 500 Hz and 3.5 kHz, and digitized with 32 kHz sampling frequency. The probe was inserted into the cortex approximately perpendicular to the surface, which allowed us to record simultaneously from neurons at different depths and along an axis tangential to the cortical surface. Fourteen MUA signals responded well to visual stimuli and had good orientation selectivity. This resulted in a cluster of overlapping receptive fields (RF), all being stimulated simultaneously by the same stimulus.

A.2 Visual Stimulation. Stimuli were presented on a 21 inch computer monitor (HITACHI CM813ET, 100 Hz refresh rate). The software for visual stimulation was a commercially available stimulation tool, ActiveSTIM (www.ActiveSTIM.com). Binocular fusion of the two eyes was achieved by mapping the borders of the respective RFs and then aligning the optical axes with an adjustable prism placed in front of one eye. The stimuli consisted of either one white bar or two bars moving in different directions (60 degree difference in orientation). 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 seconds, but only 2 seconds with the strongest rate responses were used for the analysis. The bars appeared at about 3 degrees eccentricity of the center of the RF cluster and moved with a speed of 1 degree per second, covering the cluster of RFs completely. In the six stimulation conditions the bars moved in the following directions: 1: 30 and 330 degrees; 2: 0 degrees; 3: 150 and 210 degrees; 4: 180 degrees; 5: 30 and 150 degrees; 6: 210 and 330 degrees. Each stimulation condition was presented 20 times, and the order of conditions was randomized across trials. In our study, we used only conditions 1, 2, and 3 for the analysis.

Appendix B: Mathematical Proofs

B.1 ML-Estimates of Unit Positions (Eq. 2.1). Consider the vector $\mathbf{x} := (x_1, \ldots, x_n)$ of real numbers representing the positions of the units with $\sum x_i = 0$. The $\binom{n}{2}$ distance measurements are denoted by

$$\varphi_{ij} = (x_j - x_i) + \sigma Z_{ij}, \qquad 1 \le i < j \le n,$$

with independent measurement errors $\sigma Z_{ij} \sim \mathcal{N}(0, \sigma^2)$. Define $\varphi_{ji} := -\varphi_{ij}$ for $j < i$, implying $\varphi_{ii} = 0$. Then the ML estimate of $\mathbf{x}$ is given by

$$\hat{x}_k := \frac{1}{n} \sum_{\ell} \varphi_{\ell k} \quad k = 1, \ldots, n.$$

Proof. As $\varphi_{ij} \sim \mathcal{N}(x_j - x_i, \sigma^2)$ $\forall 1 \leq i < j \leq n$, the likelihood function is given by

$$L(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^{\binom{n}{2}}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i<j} (\varphi_{ij} - (x_j - x_i))^2 \right).$$

Maximizing L requires minimizing the sum of squares:

$$Q(\mathbf{x}) = \sum_{i<j} (\varphi_{ij} - (x_j - x_i))^2 = \sum_{i \leq j} (\varphi_{ij} - (x_j - x_i))^2.$$

With $\frac{\partial Q(\mathbf{x})}{\partial x_k} = 2 \sum_{\ell=1}^{n} (\varphi_{k\ell} + x_k - x_\ell) = 2 \sum_{\ell=1}^{n} \varphi_{k\ell} + 2n x_k$ and $\frac{\partial^2 Q}{\partial^2 x_k^2} = 2n$ and $\frac{\partial^2 Q}{\partial x_k \partial x_l} = 0$ for $k \neq \ell$, the estimates $\hat{x}_k = \frac{1}{n} \sum_{\ell} \varphi_{\ell k}$ minimize $Q(\mathbf{x})$.

B.2 Unbiasedness of $\hat{\sigma}^2$ (Eq. 2.3). We want to show that $\mathbb{E}(\hat{\sigma}^2) = \sigma^2$, that is, that $\hat{\sigma}^2 := \binom{n-1}{2}^{-1} Q(\hat{\mathbf{x}})$ is an unbiased estimate of σ^2. $\hat{\mathbf{x}} := (\hat{x}_1, \ldots, \hat{x}_n)$ denotes the vector of position estimates. The expectation of each summand in $Q(\hat{\mathbf{x}})$ is

$$\mathbb{E}\left[(\varphi_{ij} - \hat{\delta}_{ij})^2\right] = \mathrm{Var}(\varphi_{ij} - \hat{\delta}_{ij}) + (\mathbb{E}(\varphi_{ij} - \hat{\delta}_{ij}))^2,$$

$$\text{with } \mathrm{Var}(\varphi_{ij} - \hat{\delta}_{ij}) = \frac{n-2}{n^2} \sigma^2 [(n-2) + 2] = \frac{n-2}{n} \sigma^2$$

$$\text{and } \mathbb{E}(\varphi_{ij} - \hat{\delta}_{ij}) = \frac{1}{n} \sum_{k \neq i,j} \mathbb{E}(\varphi_{ij} + \varphi_{jk} + \varphi_{ki}) = 0.$$

$$\text{Thus, } \mathbb{E}(\hat{\sigma}^2) = \frac{1}{\binom{n-1}{2}} \cdot \binom{n}{2} \frac{(n-2)}{n} \sigma^2 = \sigma^2.$$

B.3 Distribution of Test Statistic F (Eq. 4.2). Let $C_1 := \{x_1^{(1)}, \ldots, x_n^{(1)}\}$ and $C_2 := \{x_1^{(2)}, \ldots, x_n^{(2)}\}$ be two sets of real numbers representing linear configurations on the time axis, with $\sum_i x_i^{(1)} = \sum_i x_i^{(2)} = 0$. For all pairs $(i, j)_{1 \leq i < j \leq n}$, let

$$\varphi_{ij}^{(1)} := x_j^{(1)} - x_i^{(1)} + \sigma Z_{ij}^{(1)} \quad \text{and} \quad \varphi_{ij}^{(2)} := x_j^{(2)} - x_i^{(2)} + \sigma Z_{ij}^{(2)} \tag{B.1}$$

be the raw distance measurements with independent $(Z_{ij}^{(k)})_{1\le i<j\le n}^{k=1,2} \sim \mathcal{N}(0,1)$.

We want to test the null hypothesis, H_0, that C_1 and C_2 are identical, against the alternative, H_1, that the two configurations differ:

$$H_0: \quad x_i^{(1)} = x_i^{(2)} \quad \forall i = 1,\dots,n, \qquad H_1: \quad \exists i \in \{1,\dots,n\}: \quad x_i^{(1)} \neq x_i^{(2)}.$$

Let $\mathcal{D}$ denote the merged data vector of all $2\binom{n}{2}$ phase offset measurements:

$$\begin{aligned}\mathcal{D} &:= \left((\varphi_{ij}^{(1)})_{i<j}, \left(\varphi_{ij}^{(2)}\right)_{i<j}\right) \overset{\text{(B.1)}}{=} \left(\left(x_j^{(1)} - x_i^{(1)}\right)_{i<j}, \left(x_j^{(2)} - x_i^{(2)}\right)_{i<j}\right) + \vec{\sigma}\mathfrak{Z} \\ &=: \vec{\mu} + \vec{\sigma}\mathfrak{Z},\end{aligned}$$

where $\vec{\sigma}$ is a vector all of whose $2\binom{n}{2}$ entries equal σ and $\mathfrak{Z}$ is standard normal in $\mathbb{R}^{2\binom{n}{2}}$. The systematic term, $\vec{\mu}$, represents the additivity assumption that phase offsets are pairwise distances between points on a line. This holds true for both H_0 and H_1 and is represented by the model space M and the model assumption $\vec{\mu} \in M$. Let H denote a subspace of M such that $\vec{\mu} \in H$ describes the null hypothesis. If the alternative hypothesis is true, then $\vec{\mu} \notin H$, which means that $\vec{\mu}$ has a component in the orthogonal complement of H in M, called E. Note that $\dim(M) = 2(n-1)$, $\dim(M^\perp) = (n-1)(n-2)$, and $\dim(H) = \dim(E) = n-1$.

We decompose the data vector $\mathcal{D}$ by orthogonal projection onto H and M,

$$\mathcal{D} = \mathcal{P}_H\mathcal{D} + \mathcal{P}_E\mathcal{D} + \mathcal{P}_{M^\perp}\mathcal{D},$$

and compare the lengths of $\mathcal{P}_E\mathcal{D}$ and $\mathcal{P}_{M^\perp}\mathcal{D}$. The vector $\mathcal{P}_E\mathcal{D}$ represents the differences between the two linear configurations, and $\mathcal{P}_{M^\perp}\mathcal{D}$ represents the residual error:

Under $H_0: \mathcal{P}_E\mathcal{D} = \mathcal{P}_E\vec{\sigma}\mathfrak{Z}$, and therefore $1/\sigma^2\|\mathcal{P}_E\mathcal{D}\|^2 \sim \chi^2(\dim(E))$
Under $H_1: \|\mathcal{P}_E\mathcal{D}\|^2 = \|\mathcal{P}_E\vec{\mu} + \mathcal{P}_E\vec{\sigma}\mathfrak{Z}\|^2 > \|\mathcal{P}_E\vec{\sigma}\mathfrak{Z}\|^2$

In both cases, $\mathcal{P}_{M^\perp}\mathcal{D} = \mathcal{P}_{M^\perp}\sigma\mathfrak{Z}$, hence, $1/\sigma^2\|\mathcal{P}_{M^\perp}\vec{\sigma}\mathfrak{Z}\|^2 \sim \chi^2(\dim(M^\perp))$. Thus,

$$F = \frac{\|\mathcal{P}_E\mathcal{D}\|^2/\dim(E)}{\|\mathcal{P}_{M^\perp}\mathcal{D}\|^2/\dim(M^\perp)} \sim F((n-1),(n-1)(n-2)) \tag{B.2}$$

is, under H_0, Fisher distributed with $(n-1)$ and $(n-1)(n-2)$ degrees of freedom, whereas under H_1, F is increased systematically.

It remains to compute the lengths of $\mathcal{P}_E\mathcal{D}$ and $\mathcal{P}_{M^\perp}\mathcal{D}$:

$$\mathcal{P}_M\mathcal{D} = \left(\left(\hat{\delta}_{ij}^{(1)}\right)_{i<j}, \left(\hat{\delta}_{ij}^{(2)}\right)_{i<j}\right) = \left(\left(\hat{x}_j^{(1)} - \hat{x}_i^{(1)}\right)_{i<j}, \left(\hat{x}_j^{(2)} - \hat{x}_i^{(2)}\right)_{i<j}\right),$$

where $\hat{x}_i^{(k)}$ are the ML estimates derived separately for each of the two data sets. These estimates minimize the term $\sum_{i<j}(\varphi_{ij}^{(k)} - \hat{\delta}_{ij}^{(k)})^2$ for $k = 1, 2$ and thus minimize also $\sum_{k=1,2}\sum_{i<j}(\varphi_{ij}^{(k)} - \hat{\delta}_{ij}^{(k)})^2$. Thus,

$$\mathcal{P}_{M^\perp}\mathcal{D} = \mathcal{D} - \mathcal{P}_M\mathcal{D} = \left(\left(\varphi_{ij}^{(1)} - \hat{\delta}_{ij}^{(1)}\right)_{i<j}, \left(\varphi_{ij}^{(2)} - \hat{\delta}_{ij}^{(2)}\right)_{i<j}\right)$$

$$\text{and} \qquad \|\mathcal{P}_{M^\perp}\mathcal{D}\|^2 = \sum_{k=1,2}\sum_{i<j}\left(\varphi_{ij}^{(k)} - \hat{\delta}_{ij}^{(k)}\right)^2. \tag{B.3}$$

To derive $\mathcal{P}_E\mathcal{D}$, note that under H_0, the estimates of the model distances are averages of the estimates derived separately in the two data sets because the measurement error is assumed to be of the same size in the two data sets. Thus,

$$\mathcal{P}_H\mathcal{D} = \frac{1}{2}\left(\left(\hat{\delta}_{ij}^{(1)} + \hat{\delta}_{ij}^{(2)}\right)_{i<j}, \left(\hat{\delta}_{ij}^{(1)} + \hat{\delta}_{ij}^{(2)}\right)_{i<j}\right)$$

$$\text{and} \qquad \|\mathcal{P}_E\mathcal{D}\|^2 = \|\mathcal{P}_M\mathcal{D} - \mathcal{P}_H\mathcal{D}\|^2 = 1/2\sum_{i<j}\left(\hat{\delta}_{ij}^{(1)} - \hat{\delta}_{ij}^{(2)}\right)^2. \tag{B.4}$$

With equations B.2 to B.4, we can conclude

$$\begin{aligned} F &= \frac{\frac{1}{2}\sum_{i<j}\left(\hat{\delta}_{ij}^{(1)} - \hat{\delta}_{ij}^{(2)}\right)^2/(n-1)}{\sum_{k=1,2}\sum_{i<j}\left(\varphi_{ij}^{(k)} - \hat{\delta}_{ij}^{(k)}\right)^2/(n-1)(n-2)} \\ &= \frac{1}{\hat{\sigma}_1^2 + \hat{\sigma}_2^2} \cdot \frac{1}{(n-1)} \cdot \sum_{i<j}\left(\hat{\delta}_{ij}^{(1)} - \hat{\delta}_{ij}^{(2)}\right)^2, \end{aligned}$$

where $\hat{\sigma}_k^2$ $(k = 1, 2)$ is the estimate of σ^2 derived in data set k with equation 2.3.

Acknowledgments

We thank Brooks Ferebee for important inputs to several mathematical aspects of the project. We also thank Julia Biederlack, Sergio Neuenschwander,

and Nan-Hui Chen for help with data acquisition and conditioning. Finally, we thank Anton Wakolbinger and Wolf Singer for their helpful comments and support. This work was partially supported by Hertie Foundation.

References

Abeles, M. (1982). *Local cortical circuits: An electrophysiological study*. Berlin: Springer.

Abeles, M. (1991). *Corticonics: Neural circuits of the cerebral cortex*. New York: Cambridge University Press.

Abeles, M., Bergman, H., Margalit, E., & Vaadia, E. (2000). Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *Journal of Neurophysiology*, *70*, 1629–1638.

Abeles, M., & Gerstein, G. L. (1988). Detecting spatiotemporal firing patterns among simultaneously recorded single neurons. *Journal of Neurophysiology, 60*, 909–924.

Amari, S., Nakahara, H., Wu, S., & Sakai, Y. (2003). Synchronous firing and higher-order interactions in neuron pool. *Neural Computation, 15*, 127–142.

Brown, E. N., Kass, R. E., & Mitra, P. P. (2004). Multiple neural spiketrain analysis: State-of-the-art and future challenges. *Nature Neuroscience, 7*, 456–461.

Diesmann, M., Gewaltig, M.-O., & Aertsen, A. (1999). Stable propagation of synchronous spiking in cortical neural networks. *Nature, 402*, 529–533.

Engel, A. K., Kreiter, A. K., König, P., & Singer, W. (1991). Interhemispheric synchronization of oscillatory responses in cat visual cortex. *Science, 252*, 1177–1179.

Gerstein, G. L., Perkel, D. H., & Dayhoff, J. E. (1985). Cooperative firing activity in simultaneously recorded populations of neurons: Detection and measurement. *Journal of Neuroscience, 5*, 881–889.

Gray, C. M., Engel, A. K., König, P., & Singer, W. (1992). Synchronization of oscillatory neuronal responses in cat striate cortex: Temporal properties. *Visual Neuroscience, 8*, 337–347.

Gray, C. M., König, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature, 338*, 334–337.

Grün, S. (1996). *Unitary joint-events in multiple-neuron spiking activity: Detection, significance and intergration*. Frankfurt am Main: Deutsch.

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

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

Hopfield, J. J. (1995). Pattern recognition computation using action potential timing for stimulus representation. *Nature, 376*, 33–36.

Horsnell, G. (1953). The effect of unequal group variances on the F-test for the homogeneity of group means. *Biometrika, 40*, 128–136.

Ikegaya, Y., Aaron, G., Cossart, R., Aronov, D., Lampl, I., Ferster, D., & Yuste, R. (2004). Synfire chains and cortical songs: Temporal modules of cortical activity. *Science, 204*, 559–564.

Johnson, D. H., Gruner, M. C., Baggerly, K., & Seshagiri, C. (2001). Information-theoretic analysis of neural coding. *Journal of Computational Neuroscience, 10*, 47–69.

König, P. (1994). A method for the quantification of synchrony and oscillatory properties of neuronal activity. *Journal of Neuroscience Methods, 54*, 31–37.

König, P., Engel, A. K., Roelfsema, P. R., & Singer, W. (1995). How precise is neuronal synchronization? *Neural Computation, 7*, 469–485.

Lestienne, R. (1996). Determination of the precision of spike timing in the visual cortex of anaesthetised cats. *Biological Cybernetics, 74*, 55–61.

Mainen, Z. F., & Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science, 268*, 1503–1506.

Martignon, L., von Hasseln, H., Grün, S., Aertsen, A., & Palm, G. (1995). Detecting higher-order interactions among the spiking events in a group of neurons. *Biological Cybernetics, 73*, 69–81.

Moore, G. P., Perkel, D. H., & Segundo, J. P. (1966). Statistical analysis and functional interpretation of neuronal spike data. *Annual Review of Physiology, 28*, 493–522.

Pearson, E. S. (1931). The analysis of variance in cases of non-normal variation. *Biometrika, 23*, 114–133.

Perkel, D. H., Gerstein, G. L., & Moore, G. P. (1967). Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. *Biophysical Journal, 7*, 419–440.

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

Reinagel, P., & Reid, C. (2002). Precise firing events are conserved across neurons. *Journal of Neuroscience, 22*, 6837–6841.

Roelfsema, P. R., Engel, A. K., König, P., & Singer, W. (1997). Visuomotor integration is associated with zero time-lag synchronization among cortical areas. *Nature, 385*, 157–161.

Sauseng, P., Klimesch, W., Doppelmayr, M., Pecherstorfer, T., Freunberger, R., & Hanslmayr, S. (2005). EEG alpha synchronization and functional coupling during top-down processing in a working memory task. *Human Brain Mapping, 26*, 148–155.

Schneider, G., & Grün, S. (2003). Analysis of higher-order correlations in multiple parallel processes. *Neurocomputing, 52–54*, 771–777.

Schneider, G., & Nikolić, D. (2006). Detection and assessment of near-zero delays in neuronal spiking activity. *Journal of Neuroscience Methods, 152*, 97–106.

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

Traub, R. D., Whittington, M. A., & Jefferys, J. G. R. (1997). Gamma oscillation model predicts intensity coding by phase rather than frequency. *Neural Computation, 9*, 1251–1264.

Vaadia, E., Haalman, I., Abeles, M., Bergman, H., Prut, Y., Slovin, H., & Aertsen, A. (1995). Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. *Nature, 373*, 515–518.

VanRullen, R., & Thorpe, S. J. (2002). Surfing a spike wave down the ventral stream. *Vision Research, 42*, 2593–2615.

Wennekers, T., & Palm, G. (2000). Cell assemblies, associative memory and temporal structure in brain signals. In R. Miller (Ed.), *Time and the brain* (pp. 251–273). London: Harwood Academic Publishers.

Received July 8, 2005; accepted April 13, 2006.