J Comput Neurosci
DOI 10.1007/s10827-006-9441-7

Non-parametric detection of temporal order across pairwise measurements of time delays

Danko Nikolić

Received: 11 November 2005 / Revised: 19 May 2006 / Accepted: 12 June 2006 / Published online: 19 September 2006

Abstract Neuronal synchronization is often associated with small time delays, and these delays can change as a function of stimulus properties. Investigation of time delays can be cumbersome if the activity of a large number of neurons is recorded simultaneously and neuronal synchronization is measured in a pairwise manner (such as the cross-correlation histograms) because the number of pairwise measurements increases quadratically. Here, a non-parametric statistical test is proposed with which one can investigate (i) the consistency of the delays across a large number of pairwise measurements and (ii) the consistency of the changes in the time delays as a function of experimental conditions. The test can be classified as non-parametric because it takes into account only the directions of the delays and thus, does not make assumptions about the distributions and the variances of the measurement errors.

Keywords Cross correlation · Phase offset · Temporal-order code · Transitivity · Additivity

1 Introduction

Synchronized cortical neurons show a prominent center peak in cross-correlation histograms (CCH; Perkel et al., 1967), and such synchronization can be associated with time delays, one neuron having a tendency to fire earlier than the other (König et al., 1995; Schneider and Nikolić, 2006). A time delay, which is indicated by a shift of the centre peak away from the center of the CCH (Fig. 1(A)) and can be estimated by fitting either a damped Gabor function (König, 1994) (Fig. 1(A)) or a cosine function (Schneider and Nikolić, 2006), is also referred to as a 'phase shift' or a 'phase offset'. In the visual cortex, the magnitude and direction of phase offsets can depend on stimulus properties (König et al., 1995; Schneider and Nikolić, 2006; Schneider et al., 2006), and this opens the possibility that phase offsets play a functional role in cortical computations. In this case, the stimulus-related information carried by phase offsets would need to be extracted by the readout mechanisms that are sensitive to such small delays. Such readout mechanisms have been already discussed in the realm of sensory responses and the spike latencies that vary as a function of stimulus properties (Hopfield, 1995; Van Rullen and Thorpe, 2001).

Offsets extracted from a larger number of simultaneously recorded units (single cells or multi-units) are often mutually dependent, as they adhere to the principle of additivity: for any three units A, B, and C, the offset between units A and C corresponds to the sum of the offsets between units A and B and units B and C (Schneider et al., 2006) (Fig. 1(B)). The additivity of phase offsets can be used to determine the temporal order in which individual units tend to fire action potentials. Thus, for n units, the information from $n(n-1)/2$ pairwise CCHs can be condensed into a linear arrangement of n positions on a single time axis, such that each position indicates the preferred times at which each unit delivers action potentials relative to the timing of the others (Fig. 1(C) and (D)). This considerably decreases the representational complexity of the data.

Action Editor: M. Wiener

D. Nikolić (✉)
Max-Planck-Institute for Brain Research,
Deutschordenstr. 46,
D-60528 Frankfurt am Main, Germany
e-mail: danko@mpih-frankfurt.mpg.de

D. Nikolić
Frankfurt Institute for Advanced Studies, Johann Wolfgang
Goethe-University, Frankfurt, Germany

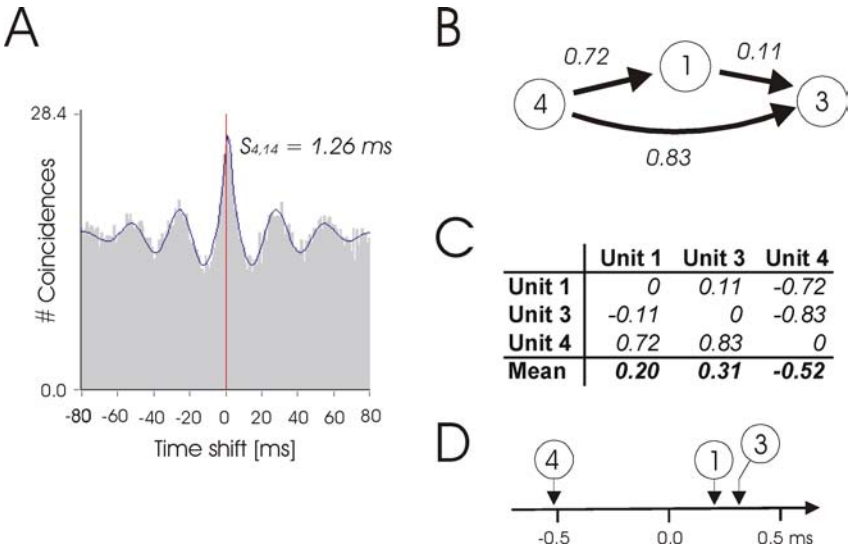

Fig. 1 Examples of a CCH with a phase offset and an additive relation between multiple phase offsets. (A) An example CCH taken from a set depicted in Fig. 5 (units 4 and 14, stimulation condition 1) that is fitted with a Gabor function. The phase offset is indicated by the shift of the center peak away from the center of the CCH. (B) A rare case of additivity precise to two tenths of a millisecond across three phase offsets (a triple). This example is used to illustrate the computation of units' positions on a time axis and is extracted from a larger network shown in Fig. 6(A). (C) Computation of the relative time positions of the three units in (B) and the graphical representation of the results in (D). The relative time positions are obtained by averaging the delays across the columns of the delay matrix (for details see Schneider and Nikolić, 2006)

A parametric method by Schneider et al. (2006) can test whether the arrangement on the time axis is not arbitrary but emerges from genuine temporal structure reflected in consistent relationships between phase offsets. Here, the description of the data is based on the mean and variance, offering high test-power and relying on standard and well-understood statistical concepts (i.e., analysis-of-variance). The potential disadvantage of this and any other parametric method is that it is not always robust against the violation of assumptions about the data properties (e.g., normal distribution of measurement errors, homoscedasticity of variance and interval scaling) (Ramsey, 1980; Boneau, 1960). Thus, it is sometimes necessary to use so-called non-parametric statistical methods, which pose fewer requirements on the data properties (e.g., using a, Mann-Whitney *U*-test instead of a Student *t*-test). This article presents a method for non-parametric investigation of additivity across phase offsets. The method considers only the directions of phase offsets (i.e., the signs of the measured delays) rather than their magnitudes, investigates whether the directions of phase offsets are consistent across pairs of units, and investigates the degree to which the resulting networks of phase offsets are *transitive*.

2 Transitivity

In the presently considered networks, the investigated units of neuronal activity are represented by the nodes (i.e., single neurons or multi-unit activity), and the directions of the phase offsets are indicated by arrows oriented according to the signs of the delays, s_{ij}, which are estimated for pairs of units i and j. Thus, the arrows in the network indicate the flow of time, the units at the 'dull' ends of the arrows firing action potentials earlier than the units at the 'sharp' ends. Unless stated otherwise, it is assumed that networks are fully connected and thus, that for every pair of nodes a CCH is computed and a phase offset is estimated.

By convention, a positive value of s_{ij} indicates that unit i fires on average earlier than j. If the offsets are additive, for any pair of units, ij the following holds:

$$s_{ij} = s_{ix} + s_{xj}, \tag{1}$$

where x is the index of any other unit in the network. Note that

$$s_{p,q} = -s_{q,p} \tag{2}$$

and thus, in Eq. (1) the order of the indexes needs to be taken in account.

If a network is transitive (in graph theory also known as a tournament), no circular path exists. Thus, if one travels through such a network by following the directions of arrows, each node can be visited at most once, and the travel is possible in one direction only. Examples of transitive and non-transitive networks are shown in Fig. 2(A)–(D). In Appendix A one can see various equations describing the properties of transitive networks (see also Moon (1968) for more information on tournaments).

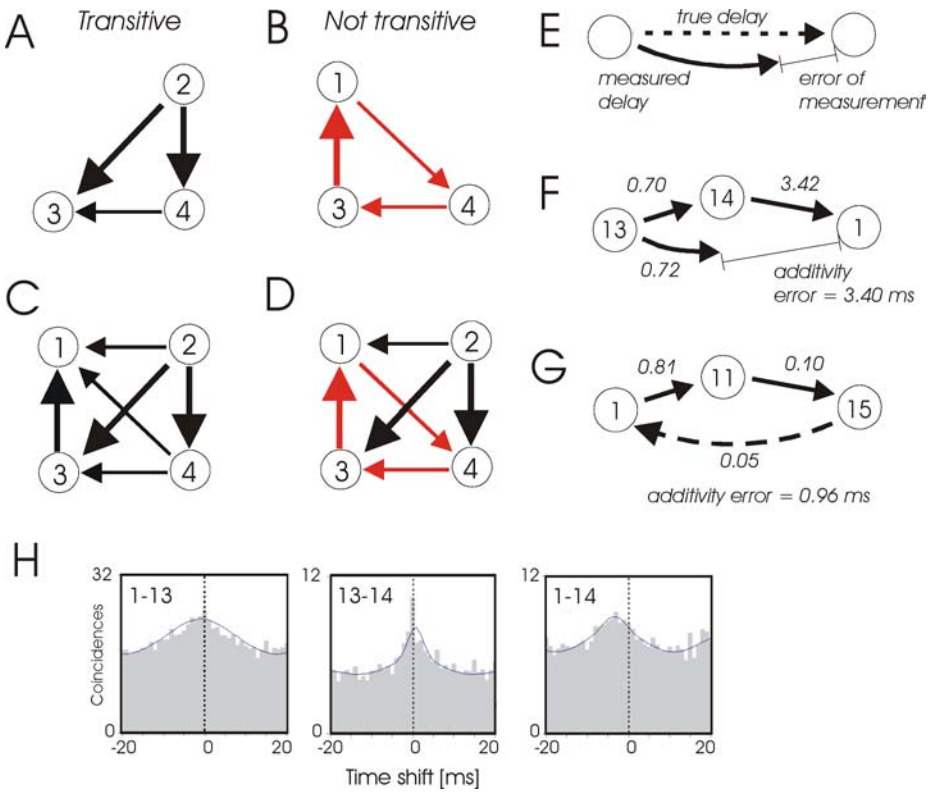

Fig. 2 Transitivity across phase offsets. (A) to (D) Examples of transitive and non-transitive networks that consist of three and four nodes. Non-transitive triples in (B) and (D) are drawn in red. (E) Definition of the error in the measurement of time delays. The magnitude of each measured delay is assumed to be a function of the true delay and the error of measurement. (F) and (G) Examples of the errors in additivity. (F) An atypically large error in additivity, where the network remains nevertheless transitive because the magnitude of the error in the directly-estimated delay between units 13 and 1 is smaller than the indirect delay computed as a sum of the two delays via unit 14. (G) Example of an error in additivity that produces a non-transitive relation. Note that the additivity error is larger than either of the individual delays. The examples are taken from the dataset discussed in more detail in Section 6. (H) CCHs illustrating in more detail the large additivity error shown in (F)

In a transitive network, each node has a unique number of in-going (and also out-going) arrows. Thus, for a network of n nodes, the first node is connected to the rest of the network exclusively by out-going arrows, whose number is $n-1$. The next node in the sequence has one in-going arrow and $n-2$ out-going arrows. This continues until the last node where all the $n-1$ arrows are in-going (see Fig. 2(C) for an example). Thus, in transitive networks the nodes can be ordered according to the number of in-going (out-going) arrows, and this order indicates the order in which the units fire their action potentials. For example, in Fig. 2(C), the transitive relations in a four-node network indicate the following firing order: 2, 4, 3 and 1. In contrast, in the non-transitive network in Fig. 2(D), three of the units share the same number of in-going arrows (units 1, 4, & 3), and thus, the firing order cannot be established unambiguously.

A network that is perfectly additive (i.e., the delays sum up exactly) is also transitive. In contrast, a transitive network is not necessarily perfectly additive. This is because, even if the directions of delays are consistent, their magnitudes might jitter slightly (e.g., due to the measurement errors of phase offsets), and thus, the network lacks perfect additivity. Nevertheless, the transitivity of a network indicates also a high degree of additivity. To illustrate this point, consider a network in which the delays are in reality perfectly additive, but due to non-negligible measurement errors of phase offsets, the delays fail to sum up exactly (i.e., additivity errors). Such a network can remain transitive only if the measurement errors are sufficiently small, so as not to exceed the magnitudes of the delays. Otherwise, the errors would change the directions of arrows unsystematically, resulting in a loss of transitivity (see illustrations in Fig. 2(E)–(H)). Thus, the errors of additivity must be, in a transitive network, smaller than the sizes of these delays. In addition, if the magnitudes of the measurement errors, with which phase offsets are estimated, are similar, irrespective of whether the delays are short or long, transitivity can be achieved only if these errors are smaller than the smallest delays in the network. As the estimated delays are, in cortex, often shorter than 1 ms (Schneider et al., 2006) and can be short even over large cortical distances (Roelfsema et al., 1997), it follows that the transitivity of a network is also a strong indicator of additivity.

As it will be seen shortly, transitivity is, in practice, too strong a requirement because the networks obtained experimentally are likely to contain at least a few additivity errors that exceed the sizes of the delays and that thus, render the networks non-transitive (see Fig. 2(G) for an example). For this reason, it is also necessary to consider the *partial* transitivity of a network and investigate whether the degree to which a network is partially transitive indicates also the degree to which the network is additive.

3 Partial transitivity

Partial transitivity can be defined as the transitivity of sub-networks that constitute the global network. The larger the number of transitive sub-networks, the higher is the degree of partial transitivity in the global network. In a partially transitive network, some units will have a unique position in the firing order, while the positions of others will be ambiguous. Thus, it is possible that the network is highly structured but that the relative firing times cannot be resolved unambiguously for every single pair of units (this was the finding from the application of parametric methods in Schneider et al., 2006). A measure of partial transitivity is proposed in Section 5, and here we motivate the use of statistical analysis to determine the likelihood that partial transitivity within a network is obtained by chance.

Neuronal spiking events do not necessarily have to result in additive positions of center peaks in CCHs. Instead, depending of the particular arrangements of action potentials, the positions of center peaks in CCHs can express relations other than additivity (Schneider et al., 2006). Thus, additivity is not given by default, and hence, partial transitivity may, in principle, arise also by chance. The simulations presented in Section 5 show that, even in networks in which the directions of arrows are assigned randomly, one can detect a certain degree of partial transitivity. Therefore, it is necessary to test statistically the null-hypothesis stating that the directions of arrows in the network are assigned randomly and to contrast it to the alternative hypothesis stating that the degree of transitivity between the directions of arrows exceeds the chance level. Only if it is highly unlikely that the observed degree of partial transitivity is obtained by chance (e.g., $p < 0.05$), can the network be considered as structured and containing additive relations that exceed the chance level.

4 Detecting changes in firing order

As mentioned, when neuronal activity is evoked by different stimulation conditions, neurons can change their order of firing (König et al., 1995; Schneider and Nikolić, 2006). To detect such changes within large networks, one needs a statistical test to investigate whether the phase offsets are obtained from networks with identical firing sequences (null-hypothesis) or alternatively, whether the firing sequences across the compared networks are likely to be different (alternative hypothesis).

Akin to the commonly used non-parametric tests, such as the sign-test or Wilcoxon signed-ranks test, a non-parametric test for changes in the firing order can be created by applying the present test of partial transitivity to a *difference network*: if the network obtained under one stimulation condition is the reference and the one under the other condition is the target, then for each pair of nodes, ij, we subtract the magnitude of the delay observed in the target, s_{ij}, from the corresponding magnitude of the delay in the reference network, $\hat{s}_{ij}$. Note that the order of indexes needs to be matched because $s_{pq} = -s_{qp}$. Hence, if $\boldsymbol{R}$ and $\boldsymbol{T}$ represent matrices of delays in the reference and the target networks, respectively, the difference matrix, $\boldsymbol{D}$, is computed as:

$$\boldsymbol{D} = \boldsymbol{R} - \boldsymbol{T}. \tag{3}$$

The difference networks have the following interesting properties: if two networks are perfectly additive, the differences (or sums) between the respective delays of the networks will also be perfectly additive (see Appendix B.1 for a proof). Thus, if the target and the reference networks originate from different firing sequences of the units, the difference network should contain additive structural changes in the delays. In contrast, if the target and the reference originate from the same firing sequences, the values in the difference network will result solely from the random variations in the measurement errors of phase offsets. This will result in a random assignment of the directions of delays and thus, in a lack of additivity. The partial transitivity within the difference network should also not exceed the chance level (see examples of additive and non-additive difference networks in Fig. 3). Therefore, the previously postulated transitivity test can be also applied to difference networks to investigate whether these networks indicate transitive changes in the temporal structures or alternatively, whether these changes consist solely of random variations of the measurement errors.

It is important to note that transitivity (or additivity) of the difference network does not imply transitivity of the original networks: two non-transitive networks can produce a transitive difference. Therefore, whenever investigating the transitivity of difference networks, it is desirable to determine also the transitivity of the original networks.

5 Transitivity test

To specify a statistical test of partial transitivity, two steps are required: first, one needs to define an appropriate measure of

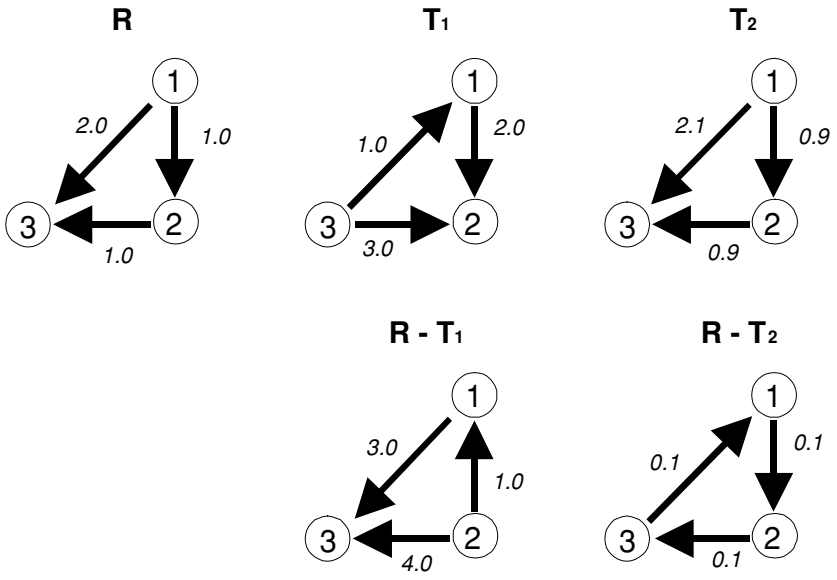

Fig. 3 The properties of difference networks. Two difference networks are illustrated, $\mathbf{R} - \mathbf{T}_1$ and $\mathbf{R} - \mathbf{T}_2$, that are computed by subtracting two target networks, $\mathbf{T}_1$ and $\mathbf{T}_2$, from the same reference network, $\mathbf{R}$. The firing order in $\mathbf{T}_1$ is different than in $\mathbf{R}$, while the firing order in $\mathbf{T}_2$ is the same as in $\mathbf{R}$. The delays in $\mathbf{T}_2$ are also the same as in $\mathbf{R}$ with the exception that small changes in random directions are added in order to mimic the measurement errors (additivity in $\mathbf{T}_2$ remains high because the error does not exceed 15%). The difference network $\mathbf{R} - \mathbf{T}_1$ is additive and thus, also transitive. In contrast, the difference $\mathbf{R} - \mathbf{T}_2$ consists of small delays that are neither additive nor transitive

the degree to which a particular network is partially transitive. Second, probabilities need to be computed for obtaining the given degree of partial transitivity by chance.

5.1 Measure of partial transitivity

The measure of partial transitivity proposed here is based on testing the transitivity of the smallest sub-networks within the global network, which consist always of three nodes, being thus referred to as triples. The transitivity of sub-networks larger than triples is investigated indirectly because, as the proof in Appendix B.2 shows, if all of the triples of a given network are transitive, then the entire network is also transitive. Thus, the present measure assumes that the larger the number of transitive triples, the larger the degree of partial transitivity. The actual measure that is used here is expressed as the corresponding count of non-transitive triples because, as we will see shortly, this number is always much smaller than the number of transitive triples and is, thus, easier to handle. Therefore, for a network of n nodes, all the $\binom{n}{3}$ triples are tested, and the count of the non-transitive ones is used as an indicator of the degree to which the network is partially transitive (see Appendix E.1 for a computational algorithm that counts the non-transitive triples).

5.2 Probabilities to obtain partial transitivity by chance

In Appendix C, analytical computations are provided for the probabilities that fully-connected networks, with zero and with one non-transitive triple, will arise by chance. As these calculations were too complex for larger numbers of triples, the remaining probabilities were obtained by simulations. Thus, up to 10,000,000 random networks were generated with sizes of up to 128 nodes, where the direction of each arrow was decided randomly with a probability of 0.5 and independently of the other directions. The number of non-transitive triples was then counted in every network, and these counts were subsequently used to compute the probability distributions for observing a certain number of non-transitive triples in a random network of a given size. The results are shown in Fig. 4(A) for network sizes of up to 10 nodes. Note the saw-shaped distributions indicating that an even number of non-transitive triples is slightly more likely to be obtained than an odd number.

We next determined the maximal count of non-transitive triples that satisfies a certain criterion of statistical significance (i.e., the critical count of non-transitive triples). To this end, the probability distributions were integrated into cumulative distributions (shown in Fig. 4(B)), and the critical counts of non-transitive triples were determined at the left tails of the cumulative distributions. The resulting counts are given in a tabular form in Appendix D for the following significance criteria: alpha $=$ 0.05, 0.01 and 0.001. From these results, one can see that the critical number of non-transitive triples depends strongly on the network size. For example, transitivity within a network of 6 nodes (20 triples in total) is significant at an alpha level of 0.05 only if the network does not contain a single non-transitive triple (i.e.,

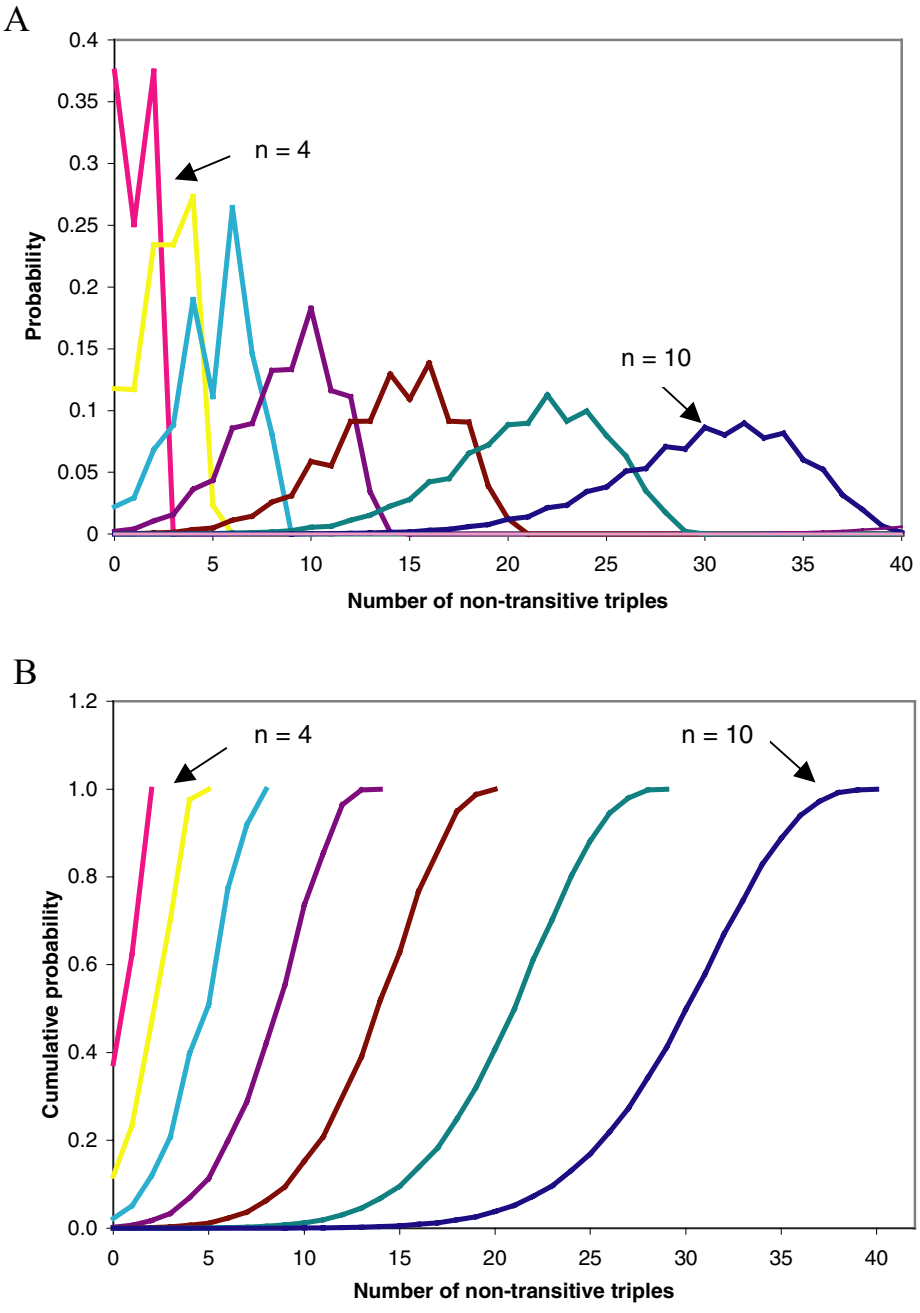

Fig. 4 Results of simulating a large number of networks with randomly-assigned directions of delays. (A) Distributions of the probability that a particular count of non-transitive triples will be observed for networks of different sizes, n (only the sizes of 4 to 10 nodes are shown). (B) Cumulative distributions obtained by integrating the probabilities in A

the whole network must be transitive). In contrast, a network of 16 nodes can contain up to 96 non-transitive triples (out of the total of 560) and is still partially transitive to a significant degree, at a rigorous alpha value of 0.001.

5.3 Missing arrows

In practice, networks do not need to be always fully connected but some of the measurements of phase offsets might be missing. The present statistical tables can be used with missing data if one takes a statistically conservative stand and considers every triple, for which transitivity cannot be determined, as non-transitive. A programming algorithm for such count is provided in Appendix E.2. Note that the test power decreases quickly with every additional missing value and that, thus, one should make every effort to obtain the measurements of phase offsets, the measurements with poor precision having much more value than a lack of measurement. For factors that affect the precision of phase-offset measurements and for the type of noise that is likely to be encountered while extracting these measurements, see Schneider and Nikolić (2006).

5.4 Detecting transitive sub-networks

In theory, the brain activity can be organized such that small sub-networks are transitive while the global networks are largely non-transitive. Consequently, transitivity could serve as a mechanism for the formation and segregation of cell assemblies, where only the cells with transitive relations would define, functionally, a neuronal assembly. To investigate such hypotheses and to detect transitive sub-networks in large non-transitive networks, an exhaustive search for transitive networks is not a satisfactory option. Besides the computational difficulties, exhaustive search would result with a large number of multiple comparisons that would increase sharply the rate of type I error. A much better strategy would be to test specific hypotheses that are formulated on a theoretical ground (e.g., cells that are stimulated by the object in the focus of attention form unique transitive assemblies). As usual, the alpha values should be adjusted according to the number of multiple comparisons made.

6 Applying the test to a sample dataset

The present test was applied also to a sample dataset of nine units and six stimulation conditions obtained from the cat visual cortex. This dataset was acquired by using similar methods to those reported previously in Schneider and Nikolić (2006) and in Schneider et al. (2006) (see Appendix F for more details). The original CCHs computed across nine units and two stimulation conditions are shown

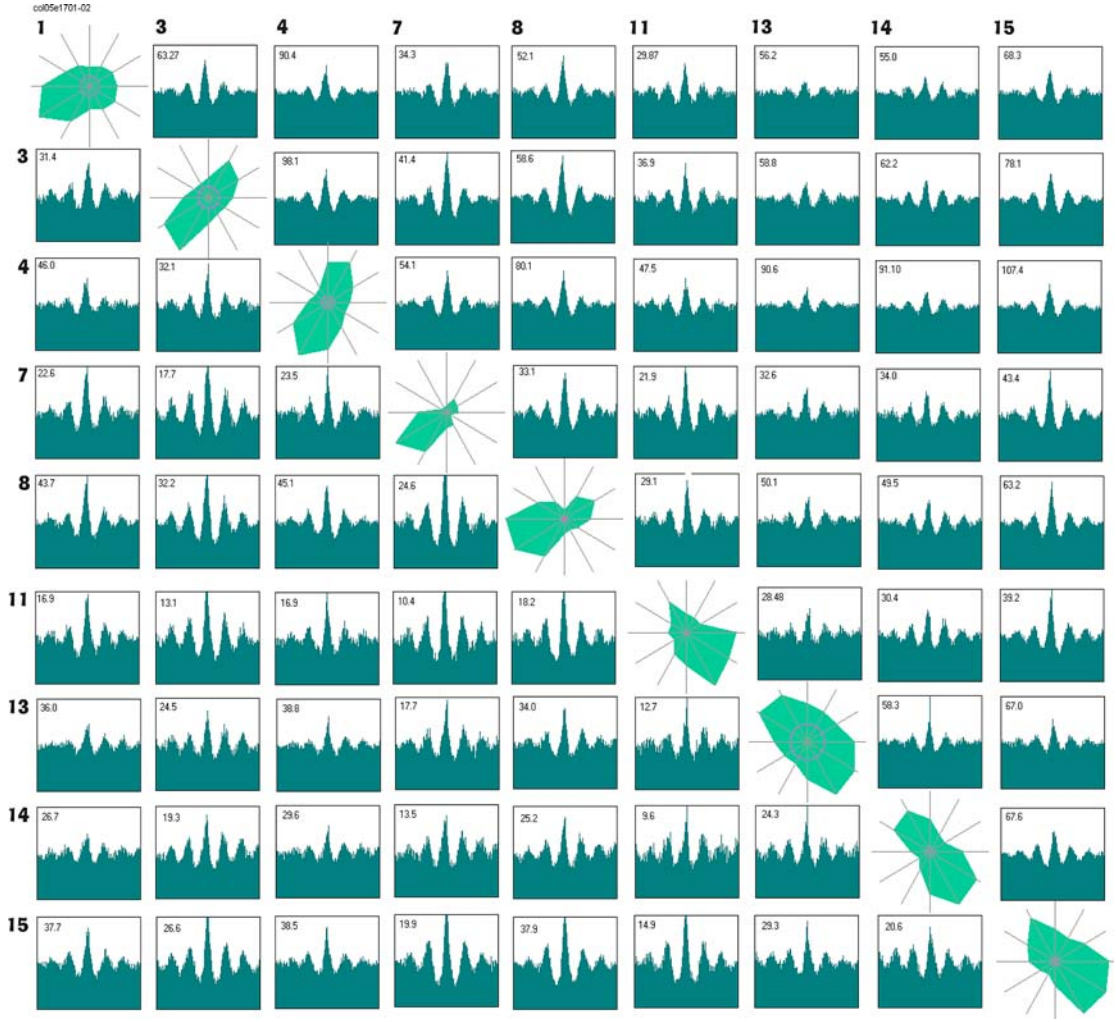

Fig. 5 CCHs computed for all pairs of multi-unit-activity spike trains that entered the analysis, shown for two stimulation conditions. The matrix is organized such that the CCH for each particular pair is located at the intersection of the column and the row of the corresponding unit (unit labels are indicated in the first column/row). CCHs obtained in the stimulation condition 1 are shown in the upper-right and those obtained in the condition 2 in the lower-left triangle. All CCHs are computed for the time period of ± 80 ms, and the numbers in the upper left corners indicate the maximum number of coincidences observed per stimulation trial. Diagonal: orientation tuning of the units as defined by the firing-rate responses to sinusoidal gratings drifting in twelve different directions

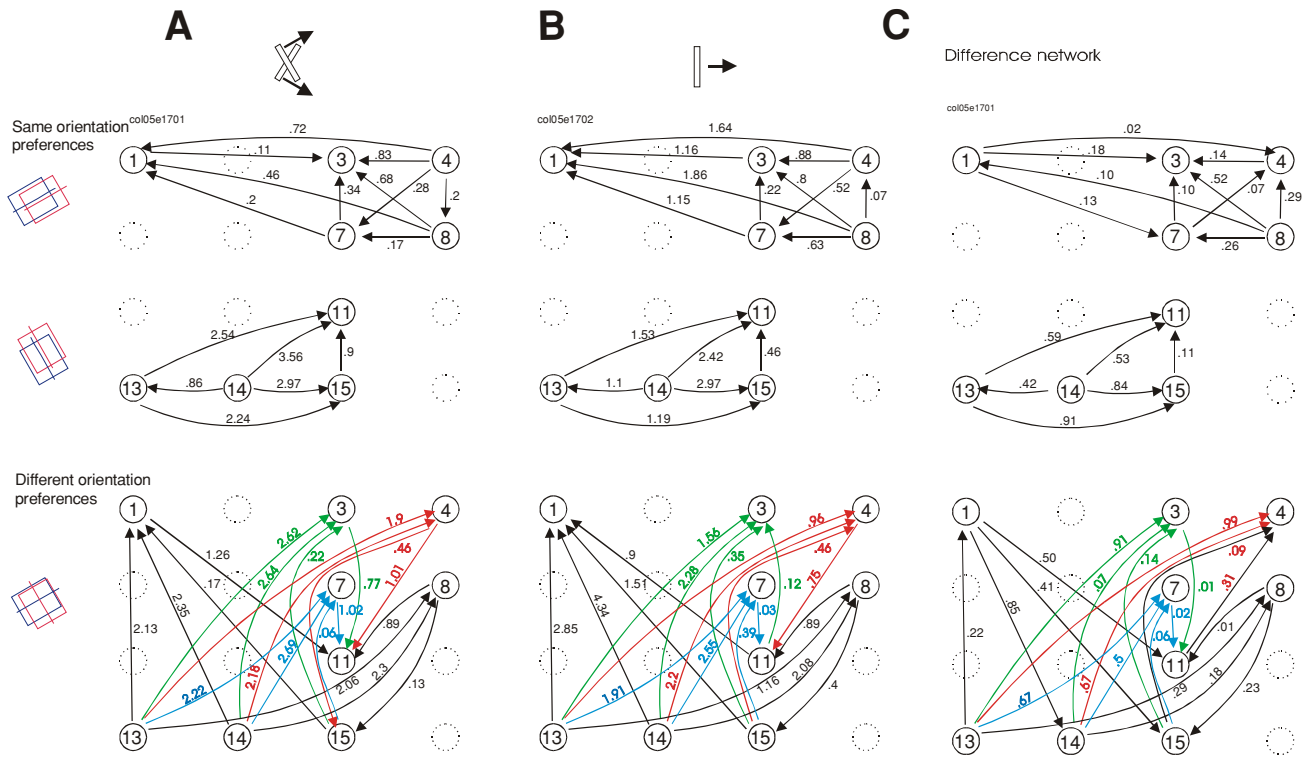

Fig. 6 The directions (arrows) and the magnitudes of phase offsets (in milliseconds) that are estimated by the method of König (1994), applied to the CCHs shown in Fig. 5. The arrangements of the units within the 4×4 matrices reflect the spatial positions of the recording sites on the Michigan probes, units 1 and 13 being positioned in the deeper cortical layers than units 4 and 8, which were located more superficially (see Appendix F for more details). If all 36 phase offsets resulting from 9 units are drawn in the same graph, the resulting representation is crowded. For this reason, the phase offsets are segregated into two panels. The upper panel indicates only the phase offsets between the pairs of units that preferred similar stimulus orientations and the lower panel between the pairs that preferred different stimulus orientations. The orientations of the corresponding receptive fields are schematized on the far left (for details on orientation preferences see Fig. 5). (A) and (B) Phase offsets obtained in the stimulation conditions 1 and 2 respectively. Stimuli are sketched on top of each figure. (C) A difference network that was computed by subtracting the network (A) from a target network, which was extracted from all six stimulation conditions by averaging the corresponding delays (see text for details)

in Fig. 5. The networks of phase offsets, obtained after fitting Gabor functions, are shown in Fig. 6. Visual inspection of the networks in Fig. 6 suggests both additivity of phase offsets and a change in the temporal structure across the stimulation conditions. Thus, the present non-parametric method was applied to investigate (i) whether the partial transitivity of phase offsets in these networks exceeds the chance level and (ii) whether these networks indicate different firing orders, i.e., whether the partial transitivity of their differences networks exceeds the chance level.

A network of 9 nodes consists of 84 triples and, for the alpha levels of 0.05, 0.01 and 0.001, the critical counts of non-transitive triples are 13, 9, and 5, respectively (taken from Appendix D). Table 1 shows the resulting counts of the non-transitive triples. In three stimulation conditions, the count of non-transitive triples was zero, indicating that the network was transitive. In the remaining three conditions, the number of non-transitive triples did not exceed two. Thus, the networks showed highly-significant levels of partial transitivity in all six stimulation conditions, indicating that the structures of the networks were highly additive. These results are consistent with those obtained by application of the parametric methods on a similar dataset (Schneider and Nikolić, 2006; Schneider et al., 2006).

The counts of non-transitive triples resulting from all of the pairwise comparisons based on difference networks

Table 1 The counts of non-transitive triples and the corresponding *p*-values resulting from the analysis of transitivity in six stimulation conditions (Original networks) and the changes in temporal structure relative to the network with delays averaged across the six stimulation conditions (Difference networks)

	Original networks		Difference networks	
Stim. Cond.	Count	*p*-value	Count	*p*-value
1	0	<0.001	8	<0.01
2	0	<0.001	2	<0.001
3	0	<0.001	4	<0.001
4	2	<0.001	10	<0.05
5	1	<0.001	9	<0.01
6	2	<0.001	3	<0.001

Table 2 The counts of non-transitive triples in difference networks resulting from 15 pairwise comparisons between the original six networks

Stim. Cond.	2	3	4	5	6
1	2	4	12	3	5
2		2	7	2	0
3			6	6	1
4				10	6
5					2

($n = 15$) are shown in Table 2. Most of these counts were larger than those obtained from the original networks (ranging between 0 and 12). This increase in the number of non-transitive triples is expected because the delays in the difference networks are much smaller than the delays in the original networks (see the example in Fig. 6). Consequently, the errors of measurement are more likely to affect the directions of the small differences between delays than the directions of the much larger original delays. Nevertheless, all 15 comparisons resulted in counts of non-transitive triples that were within the range of statistical significance at the alpha level of 0.05. Therefore, the test indicates that the directions of phase offsets changed consistently and suggests that the units might be changing their relative firing times. It is important to be aware of the limitations in generalization of the present findings, as the present method does not consider the variance in the positions of the units. This issue is discussed in Section 7.

By performing 15 pairwise comparisons arising from six stimulation conditions, the chances of the type I error in statistical inference increased considerably. To address this issue and to minimize the number of comparisons, one can compare the six original target networks always to the same reference network, resulting in six rather than 15 comparisons. To obtain a reference network that would be suitable for such six comparisons, an average network was computed. Thus, for each pair of units, ij, the delays, s_{ij}, were averaged across all six target networks. This network was transitive (result not shown) and one of the difference networks is shown in Fig. 6(C).

The counts of non-transitive triples in the resulting six difference networks are shown in the rightmost column of Table 1. The counts range between 2 and 10, and consistent with the preceding implementation of the test, all of the comparisons indicate significant changes. Thus, in each of the six datasets, neurons fired action potentials in a sequence that was different from that in the average sequence. These results are also consistent with those reported in Schneider and Nikolić (2006) and in Schneider et al. (2006), suggesting that the transitivity test can be a useful tool for the investigation of the additivity of phase offsets. It is important to note that all of the phase delays investigated here originated from the intrinsic sources of correlation and not from the timing relationships induced by the stimulus dynamics. This was tested by computing shift-predictors and subtracting them from the original CCHs. Shift-predictors are CCHs computed across randomly-shuffled experimental trials and contain, thus, only extrinsic correlations induced by the temporal properties of the stimulus. This subtraction did not change the sizes of the measured phase offsets.

7 Variability of preferred firing times

It is important to understand the limitations of the present method. The present test investigates the consistencies across the phase delays and has no access to the information about the variance of the unit's temporal positions over repeated measurements, i.e., each phase-offset network is measured only once. Thus, the present method investigates whether the firing time of a unit has changed relative to the errors of additivity/transitivity but does not address conventional statistical questions of how large is this change relative to the variability of the units' firing times across repeated measurements.

The units' firing times can vary with repeated measurements. To illustrate this point we show here first the variability of the center-peak positions in CCHs computed separately for single-stimulus presentations (single trials). This analysis is usually difficult because of the insufficient number of coincident events in such CCHs, but for a pair of units with about the highest firing rates (122.7 spikes/s, unit 1 and 80.0 spikes/s, unit 4; the original phase offset, -0.72 ms; stimulation condition 1), the position of center peaks in CCHs could be estimated in 17 out of the total of 20 CCHs. The estimated phase offsets varied from trial to trial with a standard deviation of 0.79 ms (Fig. 7(A)), and 3 out of 17 phase offsets showed positive values (i.e., 18% chance of reversal in the offset direction).

The next question is whether this variability represents the measurement error of phase offsets (e.g., due to the small number of entries in a CCH) or the jitter in the units' relative firing times. Only in the latter case should the changes in phase offsets be consistent across all the pairs of units. To address this question, phase offsets were estimated from CCHs computed for two groups of ten trials (split by odd and even trials according to the order of presentation). The scatter of phase offsets for all the 36 pairs shows high correlation (Fig. 7(B)) but also small deviations from the diagonal. The three largest deviations involve the same channel 1, which, with the three channels 13, 3 and 4, makes differences in phase offsets of 1.65, 1.02 and 0.78 ms, respectively. This suggests that phase offsets might be changing consistently across the pairs of units. This consistency was confirmed by computing the difference network that had 10 non-transitive triples, indicating a temporal structure that is significant at the alpha

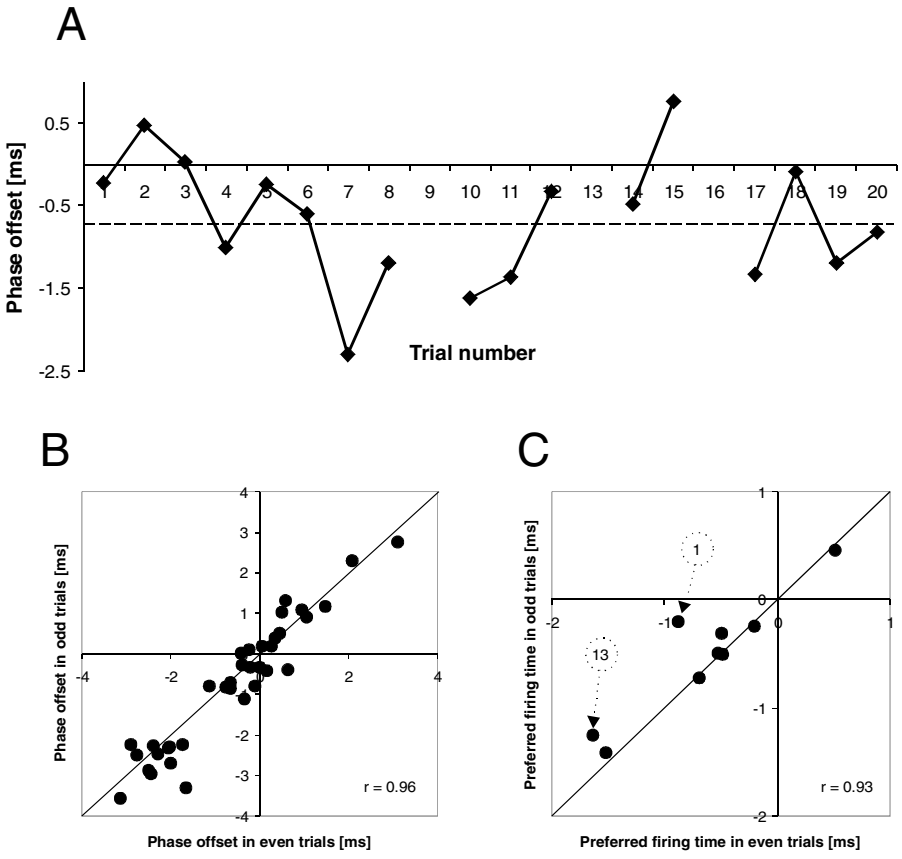

Fig. 7 Variability of units' relative firing times across repeated measurements. (A) Trial-by-trial analysis of the phase offsets between a pair of units 1 and 4. In three trials the position of the center peak could not be estimated from CCHs. Dashed line: The average phase offset of -0.68 ms, a value that is close to the -0.72 ms obtained from a CCH computed on all 20 trials. (B) Scatter plot of phase offsets obtained for odd and even trials in condition 1 and for all 36 pairs resulting from 9 units. (C) Scatter plot of the relative preferred firing times for the 9 units obtained from the data in (B) and by applying the method illustrated in Fig. 1(C) and (D). The positions of units number 1 and 13 are indicated by the dashed circles with the lead lines. *r*: Pearson's correlation coefficients

level of 0.05 (both original networks had zero non-transitive triples). The changes in the relative firing order could be also shown as a scatter plot in Fig. 7(C). As one would expect, the two largest deviations from the diagonal involved channels 1 and 13.

In conclusion, the variability of units' firing times should be taken into account, and this can be achieved only by measuring units' positions repeatedly and by applying conventional statistics. Therefore, a more complete understanding of the relation between the stimulation conditions and dynamics of neuronal firing sequences can be obtained only by corroborating the present analysis methods with other, more conventional, statistical analyses.

8 Discussion

In the present study, a statistical method is developed for testing whether the directions of phase offsets, obtained from pairwise CCHs and computed across a large number of simultaneously-recorded units, are organized into transitive temporal structures. With this test, one can investigate whether the directions of phase offsets are consistent and thus, whether neurons tend to fire their action potentials in a particular temporal sequence. In addition, the same method can be used to investigate whether phase offsets change consistently across different measurements and thus, whether there has been a change in the order of neuronal firing. Therefore, the present method can be a useful tool in the investigation of time delays in neuronal activity. The main interest for such small delays arises from the possibility that they serve as a coding mechanism that carries stimulus-related information in a fashion similar to that found in hippocampus (Mehta et al., 2002).

The present method requires that the data are measured on the ordinal scale and does not pose assumptions about the distributions of the errors or the equality of the variances. The method can be thus classified as non-parametric. Therefore, the method can be used in the cases in which one doubts the outcome of the parametric methods, due to violations of assumptions or because the data have been originally collected on the ordinal scale.

The present statistical test is easy to implement once the CCHs are computed and the phase offsets are extracted. The count of the transitive triples is straightforward, and the lookup table provides a convenient way to test the statistical significance of the observed counts of non-transitive triples. The method does not need to be applied exclusively to CCHs

but can be used as well with any other pairwise measure of time relationships.

It is important to note that non-parametric methods have lower test power than the corresponding parametric ones (e.g., Runyon and Haber, 1980). For that reason, whenever conditions allow the, application of a parametric method should be preferred.

In conclusion, methods for the analysis of large datasets of time delays in neuronal activity are currently scarce, and the apparent complexity of data resulting from multiple pairwise computations might discourage further analyses. The concepts of additivity and transitivity help alleviate these problems because they indicate redundancy across pairwise delays. Therefore, these methods enable us to reduce the representational complexity of the data and to simplify further analyses.

Appendix A

The following equations can be used to compute several quantities related to fully-connected networks.

If n is the number of the nodes in the network, then the number of arrows in the network, m, is given by

$$m = \frac{n(n-1)}{2} = \binom{n}{2}. \tag{A1}$$

The number of sub-networks with three nodes (triples) is

$$N_3 = \frac{n(n-1)(n-2)}{6} = \binom{n}{3}. \tag{A2}$$

In general, the number of sub-networks with c nodes, N_c, is

$$N_c = \frac{n!}{c!(n-c)!} = \binom{n}{c}. \tag{A3}$$

Thus, the total number of sub-networks, t, that includes networks of all different sizes and that can be extracted from a network with n nodes is

$$t = \sum_{c=3}^{n} N_c. \tag{A4}$$

The total number of different networks, d, which can be constructed from all of the possible combinations in which the directions of arrows can be assigned (i.e., the number of all possible transitive and non-transitive networks), can be expressed simply as a function of the number of arrows, m, as follows:

$$d = 2^m. \tag{A5}$$

For example, a network of 9 nodes has 36 connections, 84 triples, and 466 sub-networks, and a total of $6.87 * 10^{10}$ different networks can be constructed.

Appendix B

Appendix B.1

Claim: A sum of the respective delays between two additive networks results in an additive network.

Proof: Consider two additive networks A and B of equal size. Additivity requires that the delays s_1, s_2 and s_3 of all sub-networks of size three (triples) have the following relation:

$$s_1^A + s_2^A = s_3^A \quad \text{and} \quad s_1^B + s_2^B = s_3^B. \tag{B1}$$

The claim states that

$$\left(s_1^A + s_1^B\right) + \left(s_2^A + s_2^B\right) = s_3^A + s_3^B. \tag{B2}$$

Due to the associativity of the sum, the left-side term of Eq. (B2) can be expressed as

$$\left(s_1^A + s_2^A\right) + \left(s_1^B + s_2^B\right). \tag{B3}$$

Thus, from (B1) it follows that Eq. (B3) equals $s_3^A + s_3^B$. The same holds for the difference between networks. □

Appendix B.2

A fully-connected network will be called *transitive* (or a *tournament*) if a circuit (i.e. a non-transitive path) does not exist.

Claim: If in a fully-connected network there exist no circuits of length 3, then no circuits exist.

Proof: For $n = 3$, this is the assumption of the theorem. One can then use mathematical induction and show that no circuits of length $n+1$ exist. Suppose that no circuit of length n exists. If there is a circuit of length $n+1$, by the nonexistence of circuits of lengths 3, there must be an arrow going from the nth node to the 1st node, but this then makes a circuit of length n, contradicting thus the nonexistence of

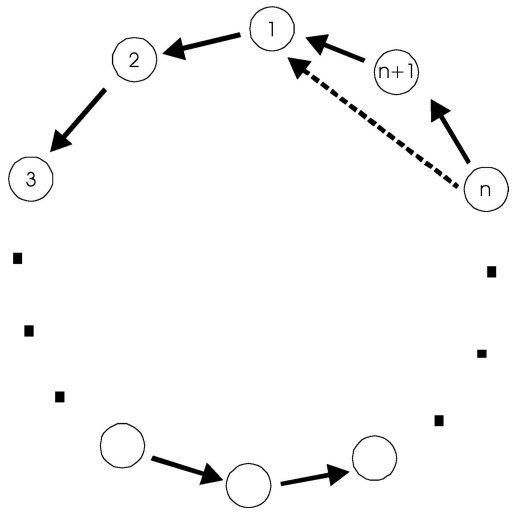

Fig. B.1 Illustration for the proof that, if no circuit of length 3 exists, then no circuit exists in a network. If circuit $n+1$ exists then the arrow indicated by the dashed line contradicts either the nonexistence of the path of length 3 or a path of the length n (depending on the direction of the arrow)

circuits of length n. Therefore, no circuit of length $n+1$ can exist (Fig. B.1). □

It follows that each higher-order non-transitive path must be associated with a non-transitive triple. Therefore, to prove that a network is transitive, it is sufficient to prove that all the triples are transitive.

given by

$$T(0) = n!. \tag{C1}$$

Then, the probability that a network with a random assignment of arrows (the direction of each arrow chosen randomly with $p = 0.5$) will be transitive is

$$p(0) = T(0)/d. \tag{C2}$$

d being defined in Appendix A. Number of networks with one non-transitive triple, $T(1)$, is given by

$$T(1) = 2(n-2)!\binom{n}{3} = 2\frac{(n-2)!n!}{3!(n-3)!} = \frac{(n-2)n!}{3} = T(0)\frac{n-2}{3}. \tag{C3}$$

Thus, the probability that a random network will have one non-transitive triple, $p(1)$, is

$$p(1) = \frac{T(1)}{d}. \tag{C4}$$

The results of applying the analytically-obtained formula to networks of sizes up to nine nodes are given in the following table. The resulting probabilities $p(0)$ and $p(1)$ are also compared to those obtained by simulations ('Sim').

n	m	N	d	$T(0)$	$p(0)$	$p(0)$, Sim.	$T(1)$	p (1)	$p(1)$, Sim.
3	3	1	8	6	0.750000	*0.750822*	2	0.250000	*0.249178*
4	6	4	64	24	0.375000	*0.374744*	16	0.250000	*0.250611*
5	10	10	1,024	120	0.117188	*0.117641*	120	0.117188	*0.117000*
6	15	20	32,768	720	0.021973	*0.021893*	960	0.029297	*0.029500*
7	21	35	2,097,152	5,040	0.002403	*0.002328*	8,400	0.004005	*0.003985*
8	28	56	268,435,456	40,320	0.000150	*0.000180*	80,640	0.000300	*0.000290*
9	36	84	68,719,476,736	362,880	0.000005	*0.000004*	846,720	0.000012	*0.000010*

Appendix C

Presented is an analytical computation of the probabilities to obtain, randomly, a network with zero and with one non-transitive triple. These probabilities are also compared to those obtained by simulations (see main text for more detail on simulations).

The total number of possible transitive networks (networks with zero non-transitive triples), $T(0)$, with n-nodes is

Appendix D

Presented are critical counts for the non-transitive triples. If the number of non-transitive triples for a network of given size (# Nodes) exceeds the provided count, the likelihood that the transitivity within the network is obtained by chance is higher than the indicated alpha level.

# Nodes	Alpha		
	0.05	0.01	0.001
1	–	–	–
2	–	–	–
3	–	–	–
4	–	–	–
5	–	–	–
6	0	–	–
7	3	1	–
8	7	4	1
9	13	9	5
10	20	16	11
11	30	25	19
12	42	36	29
13	57	50	42
14	75	67	58
15	96	87	77
16	121	111	100
17	149	138	125
18	181	169	155
19	217	205	189
20	258	244	228
21	303	289	271
22	354	338	319
23	409	393	372
24	470	453	431
25	537	518	495
26	610	590	565
27	689	668	642
28	774	752	725
29	866	843	814
30	965	941	911
31	1071	1046	1014
32	1185	1158	1125
33	1307	1278	1244
34	1436	1406	1370
35	1574	1543	1505
36	1720	1687	1648
37	1874	1841	1800
38	2038	2003	1961
39	2211	2175	2131
40	2393	2356	2311
41	2586	2547	2500
42	2788	2747	2699
43	3000	2958	2908
44	3223	3180	3128
45	3456	3412	3358
46	3701	3655	3599
47	3956	3909	3852
48	4223	4175	4116
49	4502	4452	4392
50	4793	4741	4679
51	5096	5043	4978
52	5412	5357	5291
53	5740	5684	5616
54	6081	6024	5954
55	6436	6376	6305
56	6803	6743	6670
57	7185	7123	7047
58	7581	7517	7440
59	7990	7925	7846
60	8415	8347	8266
61	8854	8785	8702
62	9308	9237	9152
63	9777	9705	9617
64	10262	10187	10099
65	10762	10686	10596
66	11278	11201	11109
67	11811	11732	11636
68	12360	12279	12183
69	12926	12843	12745
70	13509	13425	13324
71	14109	14023	13921
72	14726	14638	14534
73	15361	15272	15165
74	16014	15923	15813
75	16686	16593	16482
76	17375	17281	17168
77	18084	17988	17873
78	18812	18714	18595
79	19558	19458	19340
80	20325	20223	20102
81	21111	21007	20883
82	21917	21811	21684
83	22743	22636	22508
84	23589	23480	23350
85	24457	24346	24214
86	25345	25232	25098
87	26255	26140	26003
88	27186	27069	26930
89	28138	28019	27878
90	29113	28993	28850
91	30110	29988	29842
92	31129	31005	30857
93	32172	32046	31896
94	33237	33109	32957
95	34325	34195	34039
96	35437	35305	35148
97	36573	36438	36278
98	37733	37596	37434
99	38916	38777	38613
100	40125	39984	39819
101	41357	41215	41046
102	42615	42471	42299
103	43899	43752	43578
104	45208	45059	44884
105	46542	46391	46211
106	47902	47749	47568
107	49289	49134	48950
108	50702	50545	50357
109	52142	51983	51794
110	53609	53448	53257
111	55103	54940	54747
112	56624	56459	56263
113	58173	58006	57809
114	59751	59581	59379
115	61356	61185	60981
116	62990	62816	62611
117	64653	64477	64269
118	66345	66166	65955
119	68066	67885	67672
120	69816	69634	69416
121	71596	71410	71191
122	73406	73219	72997
123	75246	75057	74832
124	77117	76925	76699
125	79019	78825	78595
126	80951	80755	80523
127	82915	82717	82482
128	84910	84709	84472

Appendix E

Appendix E.1

Here is provided the simple programming code that can be used to count all of the non-transitive triples within a network. The code is written in a pseudo programming language and can be easily translated into any other real programming language.

The directions of arrows (values 1 or -1) are assumed to be stored in a two-dimensional array `Dir` of a size $n \times n$, where n is the number of nodes in the network, and the indices of the array run from 1 to n. The routine requires three nested `FOR` loops.

```
BadTripleCount = 0

FOR  i = 1 TO n
 FOR j = i+1 TO n
  FOR k = j+1 TO n
   IF (Dir [i,j] = Dir [j,k])
    AND IF (Dir [i,j] <> Dir [i,k])THEN
     Increment (BadTripleCount)

  ENDFOR
 ENDFOR
ENDFOR
```

Appendix E.2

The following expanded algorithm deals with missing values and assumes that a triple is not transitive if its transitivity cannot be determined. The missing values need to be indicated by value 0 in the array `Dir`.

```
BadTripleCount = 0

FOR i = 1 TO n
 FOR j = i+1 TO n
  FOR k = j+1 TO n
   Comment: Count first the number of missing values
   MissingValueCount = 0
   IF (Dir [i,j] = 0) THEN Increment(MissingValueCount)
   IF (Dir [j,k] = 0) THEN Increment(MissingValueCount)
   IF (Dir [i,k] = 0) THEN Increment(MissingValueCount)

   Comment: Two or more missing values
   IF (MissingValueCount >= 2) THEN Increment
     (BadTripleCount)

   Comment: One missing value
   ELSE IF (MissingValueCount = 1) THEN
   BEGIN
    IF (Dir [i,j] = 0) AND IF (Dir [i,k] <> Dir [j,k])
THEN Increment(BadTripleCount)
    IF (Dir [j,k] = 0) AND IF (Dir [i,j] <> Dir [i,k])
      THEN Increment(BadTripleCount)
    IF (Dir [i,k] = 0) AND IF (Dir [i,j] = Dir [j,k])
      THEN Increment(BadTripleCount)
   END

   Comment: No missing values
   ELSE IF (Dir [i,j] = Dir [j,k])
    AND IF (Dir [i,j] <> Dir [i,k])
     THEN Increment(BadTripleCount)
  ENDFOR
 ENDFOR
ENDFOR
```

Actual implementations of these algorithms are prepared as a Matlab m-file and as a class in C++ and are freely available for download from `http://www.mpih-frankfurt.mpg.de/download`.

Appendix F

Experimental methods

The experimental methods and the design are similar to those reported in two other studies (Schneider and Nikolić, 2006; Schneider et al., 2006). The results from the present dataset have not been reported previously.

A.1 Preparation and recordings

Anesthesia was induced with ketamine and, following the trachiotomy, was maintained with a mixture of 70% N_2O and 30% O_2 and with halothane (0.6%). To prevent eye movements, the cat was paralysed with pancuronium bromide applied intravenously (Pancuronium, Organon, 0.15 mg kg^{-1} h^{-1}). All of 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.

Multi-unit activity (MUA) was recorded by using a silicon-based 16-channel probe (organized in a 4×4 spatial matrix) supplied by the Center for Neural Communication Technology at the University of Michigan. The probe had minimal inter-contact distances of 200 μm (0.3–0.5 M impedance at 1000 Hz). Signals were amplified 1000×, 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

recording simultaneously from neurons at different depths and with different orientation preferences. Nine MUA signals responded well to visual stimuli and had orientation selectivity that was appropriate for eliciting strong responses with the presently used stimuli. All of the receptive fields (RF) were overlapping and were, thus, all stimulated simultaneously by a single stimulus.

A.2. Visual stimulation

Stimuli were presented on a 21″ computer monitor (HITACHI CM813ET) with a 100 Hz refresh rate. The software for visual stimulation was a commercially available stimulation tool, ActiveSTIM (www.ActiveSTIM.com). The stimuli were presented binocularly, and the eyes were fused by mapping the borders of the respective RFs and then aligning the optical axes of the eyes with an adjustable prism placed in front of one eye. The orientation preferences of the multiunits used in the present analysis could be separated in two groups. One group preferred stimuli oriented along the axis 60–240° (units 1, 3, 4, 7 & 8) and the other 120–300° (units 11, 13, 14 & 15) (see Fig. 5). The stimuli consisted either of one white bar or of two bars moving in different directions (60° difference in orientation). In the stimuli with two bars, the bars crossed their paths at the center of the cluster of overlapping RFs. At each trial the stimulus was presented in total for 5 s, but only 2 s with the strongest rate responses were used for the computation of CCHs. The bars appeared at about 3° eccentricity from the center of the RF cluster and moved with a speed of 1°/s. In the 6 stimulation conditions, the bars moved in the following directions: 1: 0° & 60°; 2: 30°; 3: 180° & 240°; 4: 210°; 5: 0° & 240°; 6: 60° & 180°. Each stimulation condition was presented 20 times, and the order of conditions was randomized across presentations.

Acknowledgments The author wishes to thank Julia Biederlack for help with data acquisition, Dubravko Ivanšić for the verification of the mathematical proofs, Martha Havenith, Diek W. Wheeler, Sanja Zolotić-Nikolic and Gaby Schneider for the help during the preparation of the manuscript, and Wolf Singer for support. This work was supported by the Alexander von Humboldt-Stiftung, Hertie Stiftung, and the Max-Planck Society.

References

Boneau CA (1960) The effects of violations of assumptions underlying the *t*-test. Psychol Bull 57:49–64

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

König P, Engel AK, Roelfsema PR, Singer W (1995) How precise is neuronal synchronization? Neural Comput 7:469–485

König P (1994) A method for the quantification of synchrony and oscillatory properties of neuronal activity. J Neurosci Methods 54:31–37

Mehta MR, Lee AK, Wilson MA (2002) Role of experience and oscillations in transforming a rate code into a temporal code. Nature 417:741–746

Moon JW (1968) Topics on tournaments. Holt, Rinehart and Winston, New York, NY

Perkel DH, Gerstein GL, Moore GP (1967) Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. Biophys J 7:419–440

Ramsey PH (1980) Exact Type I error rates for robustness of Student's *t* test with unequal variances. J Educ Stat 5:337–349

Roelfsema PR, Engel AK, König P, Singer W (1997) Visuomotor integration is associated with zero time-lag synchronization among cortical areas. Nature 385:157–161

Runyon P, Haber A (1980) Fundamental of behavioral statistics. Addison Wesley Publishing Company, Reading, MA

Schneider G, Nikolić D (2006) Detection and assessment of near-zero delays in neuronal spiking activity. J Neurosci Methods 152(1–2):97–106

Schneider G, Havenith MN, Nikolić D (2006) Spatio-temporal structure in large neuronal networks detected from cross-correlation. Neural Comput 18(10):2387–2413

Van Rullen R, Thorpe SJ (2001) Rate coding versus temporal order coding:What the retinal ganglion cells tell the visual cortex? Neural Comput 13:1255–1283