NeuroXidence: reliable and efficient analysis of an excess or deficiency of joint-spike events

**Gordon Pipa · Diek W. Wheeler ·
Wolf Singer · Danko Nikolić**

Abstract We present a non-parametric and computationally efficient method named NeuroXidence that detects coordinated firing of two or more neurons and tests whether the observed level of coordinated firing is significantly different from that expected by chance. The method considers the full auto-structure of the data, including the changes in the rate responses and the history dependencies in the spiking activity. Also, the method accounts for trial-by-trial variability in the dataset, such as the variability of the rate responses and their latencies. NeuroXidence can be applied to short data windows lasting only tens of milliseconds, which enables the tracking of transient neuronal states correlated to information processing. We demonstrate, on both simulated data and single-unit activity recorded in cat visual cortex, that NeuroXidence discriminates reliably between significant and spurious events that occur by chance.

Action Editor: Eberhard Fetz

G. Pipa · D. W. Wheeler · W. Singer · D. Nikolić
Frankfurt Institute for Advanced Studies,
Johann Wolfgang Goethe University,
Ruth-Moufang-Straße 1,
60438 Frankfurt am Main, Germany

G. Pipa (✉) · D. W. Wheeler · W. Singer · D. Nikolić
Department of Neurophysiology,
Max-Planck Institute for Brain Research,
Deutschordenstraße 46,
60528 Frankfurt, Germany
e-mail: mail@g-pipa.de

1 Introduction

It is commonly held that neurons encode information by modulations of their discharge rate. A complementary hypothesis is that information is also encoded in the precise relation between the discharges of spatially distributed neurons (Hebb 1949; von der Malsburg 1981; Fetz 1997; Singer 1999). These complementary views are addressed in the literature as the rate coding and the temporal coding hypotheses. Multiple methods have been developed to detect temporal relations between spiking events and to investigate whether these relations that are forming a spike pattern are correlated with stimulus configurations, behavior, or particular states of neuronal systems (Gerstein and Perkel 1969, 1972; Abeles and Gerstein 1988; Gray et al. 1989; Brown et al. 1998, 2004; Grün et al. 1999, 2002a, b; Martignon et al. 2000; Tetko and Villa 2001; Nakahara and Amari 2002; Pipa and Grün 2003; Barbieri et al. 2004, 2005; Sharpee et al. 2004; Samonds and Bonds 2004; Ikegaya et al. 2004; Czanner et al. 2005; Okatan et al. 2005; Kass et al. 2005; Pipa et al. 2006; Oweiss et al. 2007; Schneider et al. 2006). The methods differ in the definitions of the spike patterns, the techniques to detect these patterns, and the approaches to analyze the resulting data (descriptive, statistical hypothesis testing, maximum likelihood, and Bayesian approaches). Even though the temporal coding hypothesis formulates precisely what constitutes a spike pattern, it turns out to be a non-trivial problem to design a

method that detects the existence of such patterns and to investigate their information content, without being confounded by other properties of the data (Oram et al. 1999; Brody 1999; Roy et al. 2000; Baker and Gerstein 2001; Grün et al. 2003).

In this article we present the novel method Neuro-Xidence[1] (Pipa 2003; Pipa et al. 2006) for analyzing precise and consistent temporal relations between discharges of simultaneously recorded neurons that are referred to as coordinated firing events. NeuroXidence (see www.NeuroXidence.com) has been designed to deal with four main properties of neuronal spike trains that make the analysis of coordinated firing events difficult: variability, short timescales, history dependencies and auto-structures of the spiking activities, as well as the rareness of coordinated firing events. First, neuronal recordings show a high degree of variability, which is partially caused by noise, changes in the properties of neuronal responses, and by the dynamics of the information processing. Any analysis of coordinated firing events has to consider all three kinds of variability. Second, evidence indicates that neuronal states linked to the processing of information can last for only a short period of time (e.g., a few tens of milliseconds; Oram and Perrett 1992; Thorpe et al. 1996). This implies that analysis methods need to operate on short time series containing a very limited number of samples. Third, neuronal activity might have strong history dependencies, where the likelihood that a spike will occur at a certain point in time depends on the times at which previous spikes have occurred. Therefore, the auto-structures of the spike trains need to be considered in the analysis of potentially meaningful coordinated firing events. Fourth, past investigations of coordinated firing events indicated that these events are rare (Abeles and Gerstein 1988; Ikegaya et al. 2004). Thus, analysis of coordinated firing events might be difficult, since even highly parallel recordings from up to a hundred simultaneous recording sites might still grossly under-sample the neurons that participate in the actual temporal code.

In the present article we first describe the statistical hypothesis test and the detection algorithm used by Neuro-Xidence. Next, we utilize simulated data to illustrate the statistical properties, performance, and applicability of NeuroXidence based on calibration studies for false positives and for test-power. To this end we use simulated data with and without spike-rate modulations as well as epochs in which spike-rate modulations coincide with epochs of coordinated firing. We next compare the performance and concept of NeuroXidence to related methods such as the unitary events analysis (Grün et al. 2002a, b), the conventional cross-correlation analysis (König 1994), and

maximum entropy approaches (Schneidman et al. 2006). Finally, we apply NeuroXidence to an example dataset consisting of 42 single units recorded simultaneously from visual area 17 of an anesthetized cat.

2 Methods

2.1 Coordinated firing events

Coordinated spiking has been defined in at least two different ways. The first definition is based on the synchronous spiking of at least two neurons with millisecond precision. These coordinated events have been referred to as synchronous firings (Gray et al. 1989), temporal-patterns (Vaadia and Abeles 1987), or joint-spike events (Grün et al. 2002a, b). The second type of coordinated firing is defined as two or more neurons spiking in sequence over a finite period of time. Such a sequence has been referred to as a spatiotemporal-pattern (Abeles and Gerstein 1988) or, more simply, a pattern (Ikegaya et al. 2004). NeuroXidence is designed only for the analysis of coordinated spiking of the first type.

In the present article, each incidence of a coordinated firing event is referred to as a joint-spike event (JSE), while the identity of a JSE is defined by the set of neurons that exhibited the coordinated firing. This set of neurons is referred to as a joint-spike pattern (JS-pattern). The complexity of a JSE (Baker and Lemon 2000) is defined by the number of neurons participating in the event. The complexity of a JSE has to be separated from the order of correlation (Martignon et al. 2000; Nakahara and Amari 2002; Schneidman et al. 2003). The order of correlation is defined by the number of neurons that are directly coupled and are the primary cause of the coordinated firing. Due to one or more spurious or lacking events (Fig. 1(b), *order 2* and (c), *order 3*)), the order is not necessarily the same as the complexity, which is the total number of neurons participating in the resulting spike pattern.

Methods that were developed to identify the complete or partial structure of neuronal coupling (Martignon et al. 2000; Nakahara and Amari 2002; Okatan et al. 2005; Schneidman et al. 2006) needed large datasets, in which at least the structure of neuronal coupling had to be assumed to be fixed. To the best of our knowledge, currently available methods fail to extract the full coupling structures from massive parallel recordings (activity of up to hundreds of neurons recorded simultaneously) if the activity involves fast changes in the properties of the responses and a high degree of trail-by-trial variability or the amount of data is very limited. NeuroXidence addresses the problems of under-sampling and variability and does not try to derive a model of the coupling structure, but

---

[1] Pronunciation of NeuroXidence : NeuroKsedence.
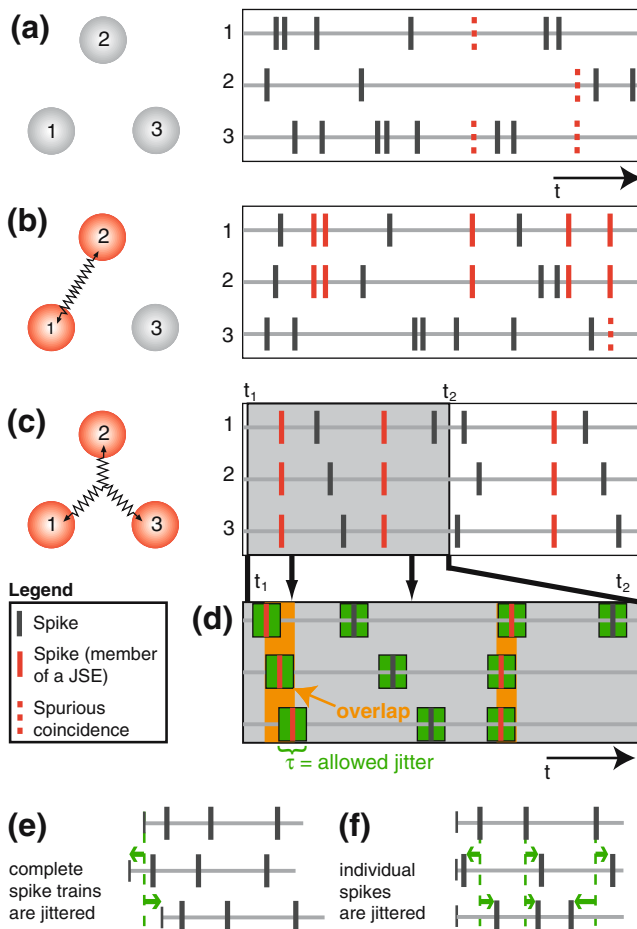
Fig. 1 Synchronization of spikes by neuronal coupling. (**a**–**c**) Three units, indicated by *1–3*, represent either single neurons or three neuronal populations that are mutually independent. The *right sub-panel* shows the resulting spike trains. In the case that the units represent single neurons, each spike train contains single-unit activity, while each spike train is composed of multi-unit activity if the units represent small populations of coupled neurons (**b** and **c**). Synchronized spikes (indicated by *red color*) are referred to as joint-spike events (*JSE*), and their complexities are indicated on top of the spike trains. (**a**) Since the units are not coupled, they are mutually independent and exhibit spike trains that do not share synchronized spikes beyond chance level (indicated by *dashed red color*). (**b**) A second-order coupling between units 1 and 2 results in a JSE of complexity 2 (indicated by *solid red color*) and one spurious JSE of complexity 3 (indicated by *dashed red color*). (**c**) A third-order coupling between units 1, 2, and 3. (**d**) A schematic of the definition of a JSE. Spikes that belong to a JSE share overlapping regions of the maximally allowed jitter, τ_c. Regions of allowed jitter represent either a maximally allowed offset, a maximal random jitter, or both together. (**e**) and (**f**) Two different ways of jittering spikes. (**e**) In NeuroXidence, each complete spike train is jittered on a timescale given by τ_r. This jittering destroys coordinated spiking on all shorter timescales, while preserving the auto-structure of each spike train on all timescales. (**f**) Each spike of each spike train is jittered individually. Therefore, the auto-structure, as well as the cross-structure, is modified

rather it tests whether coordinated firing is random or likely to be induced by, as well as correlated to, information processing or neuronal states.

2.2 Detection of joint-spike events

NeuroXidence analyzes coordinated neuronal firing utilizing a novel and computationally efficient algorithm for identifying and counting JSEs that are based on joint-spike activity that is not precisely synchronous. The central feature of this method is the supposition of a timescale τ_c, which defines the required temporal precision of fine temporal cross-structure that is typically assumed to be in the range between 1–10 ms (Markram et al. 1997; Bi and Poo 1998; Grün et al. 1999; Hopfield and Brody 2000; Aertsen et al. 2001; Hopfield and Brody 2001; Sjostrom et al. 2001). NeuroXidence utilizes τ_c as the maximum temporal difference between the first and the last spike of a JSE (Fig. 1(d)).

By using τ_c in this way, NeuroXidence avoids the problems of previous approaches that utilized τ_c for exclusive binning (Grün et al. 2002a; Pipa and Grün 2003). Exclusive binning allows for detecting JSEs that involve random or systematic deviations from precise synchronization, also referred to as jitter. However, exclusive binning does not consider the exact difference between the spike times of the neurons, rather it reduces the precision of the sampling to τ_c (Fig. 2(g)). This poses the problem that spikes, which are less than τ_c apart and should be considered as a JSE, can be detected as non-coincident if they happen to fall across separate bins (Grün et al. 2002a). For that reason, exclusive binning likely does not detect all joint-spike events. This problem is further aggravated by increases in either the complexity of the JSE and the bin size τ_c (Grün et al. 2002a). One can overcome the problems of exclusive binning with the multiple-shift method (Grün et al. 1999), that shifts entire spike trains in relation to one another, for all existing combinations of shifts. However, this method can be used only on datasets with a small number of neurons because the number of combinations of shifts and the resulting computational complexity increase exponentially with the number of neurons.

NeuroXidence is more efficient because it takes advantage of the sparseness of spiking activity to reduce the computational complexity of detecting JSEs. The algorithm also returns an exact solution and allows for variations in the timing of each individual spike that are less than the defined amount of jitter, τ_c. This is accomplished by the use of binning to represent the spiking data as a time-discrete process, where the bin size, b, is considerably shorter than τ_c (e.g., $b <= 1$ ms and $\tau_c = 5$ ms). Thus, this method preserves the precision of the spike times in the original recording when detecting JSEs (Fig. 1(d)).

There are three key steps in the detection of JSEs: the pre-processing of spike trains (see Appendix 1: Pre-processing and Fig. 12, Appendix 1), the detection of jittered JSEs, and the identification of JS-patterns and their

Fig. 2 Detection of joint-spike events (*JSEs*). (**a**) A spike train is defined as a binary process (1=spike, 0=no spike) with a bin width small enough to resolve an individual spike. (**b**) Each spike is replaced by the kernel $G^{+/-}$ that defines the maximally allowed jitter. The *red ones* ('1') indicate the position of real spikes, while the *black ones* ('1') indicate the maximally allowed jitter given by the kernel $G^{+/-}$. (**c**) The results of the replacements are JSEs, which are defined by overlapping regions of allowed jitter (*dark green*) across different spike trains. (**d**) To determine the complexity of JSEs, the number of spikes ('1s', *black* and *red*) per time bin is computed. The local maxima of the number of synchronous spikes determine the complexity of joint spiking. (**e**) and (**f**) Joint spikes belonging to one period, defined by a local maximum of the complexity, are mapped to one JSE. (**g**) The difference between the exclusive binning used by the unitary-event method (UE-method) and the centered expansion of spikes with a kernel $G^{+/-}$ used by Neuro-Xidence. Since the UE-method is based on exclusive binning, it is likely that JSEs are not detected
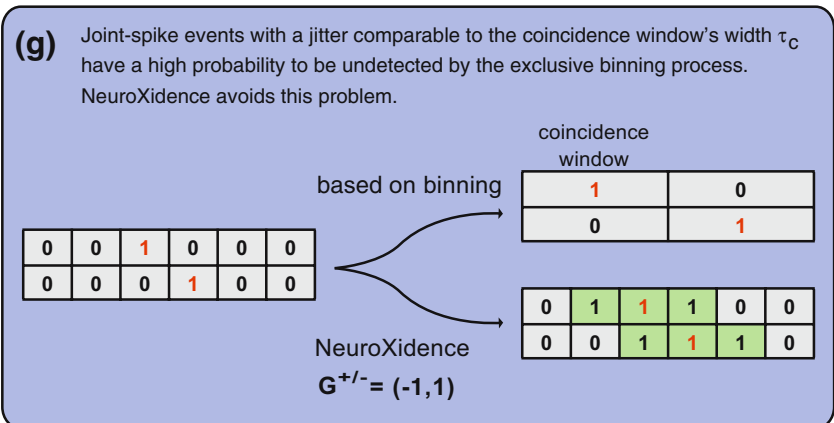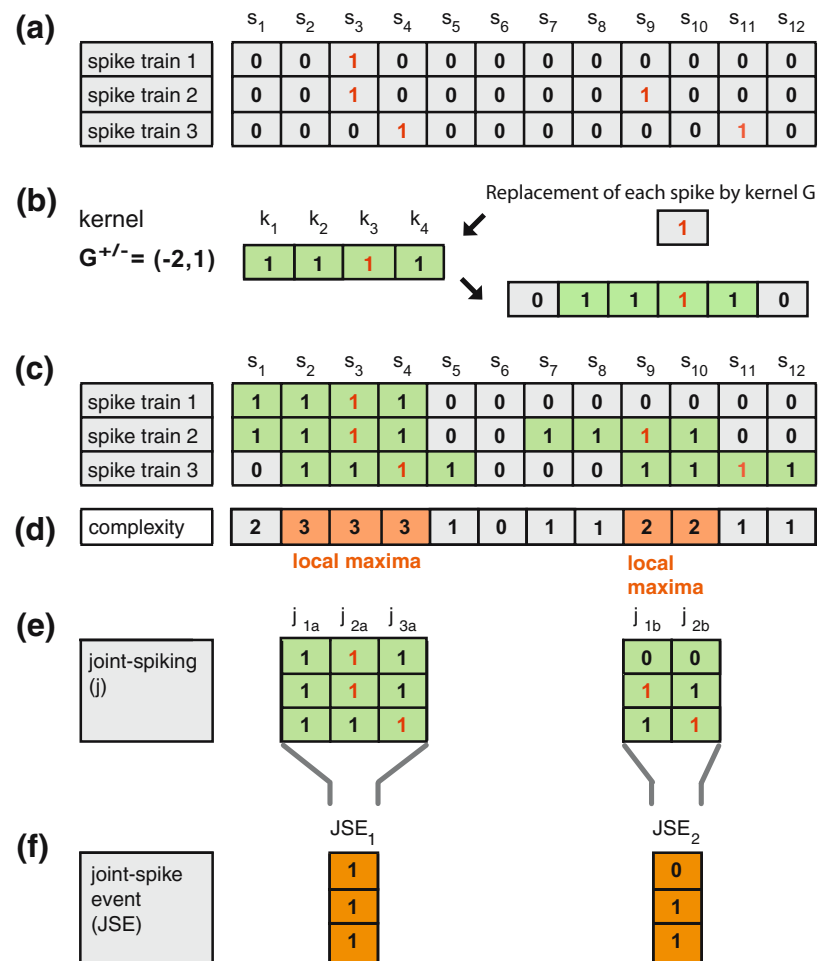
complexity. The pre-processing step ensures that the minimal inter-spike interval of any unit is larger than τ_c, corresponding to the kernel $G^{+/-}$ defined in units of the bin size, b. This step prevents temporally overlapping JSEs, while preserving each individual JSE and its exact frequency of occurrence. In the second step, NeuroXidence detects jittered JSEs by extending the temporal extent of each individual spike to the number of bins defined by $G^{+/-}$. For

that purpose, each spike and its neighborhood are replaced by a kernel, which contains $G^{+/-}$ bins of ones (Fig. 2(a–c)). The third step in JSE detection consists of identifying JS-patterns and determining their complexity. This is accomplished by evaluating the local complexity (i.e. the number of synchronous spikes) per time bin (Fig. 2(d)). Each local maximum of the local complexity corresponds to one instance of joint-spiking (Fig. 2(e)). Thus, zones of

maximal local complexity are used to derive the corresponding JSE within an identified zone of joint-spiking (Fig. 2(f)).

The determination of the total frequency of occurrences of one particular JS-pattern in the dataset involves the sum of two pattern frequencies. One is the frequency of JSEs that are identical to the JS-pattern of interest. The other is the frequency of all supra-patterns in which the JS-pattern is included as a sub-pattern.

2.3 Definition of the hypothesis test

To test whether JSEs of one type of JS-pattern occurs more or less often than expected by chance, a statistical hypothesis test is used. The null hypothesis (H0) states that JS-patterns occur by chance. Therefore, H0 assumes that the precise timing of neuronal firing, on a timescale that is defined by the precision of JS-pattern τ_c, is mutually independent across neurons (Fig. 1(a)). The alternative hypothesis (H1) assumes that there are more (or less) coordinated firing events than expected by chance and thus, that the spiking events are dependent. Consequently, the observed frequency of JSEs has to be compared to the expected frequency of JSEs in the case where the neurons are assumed to be independent. Therefore, the core of the statistical hypothesis test is the estimate of the expected frequency of JSEs.

This estimation can be based on either a parametric or a non-parametric approach. Parametric approaches hypothesize an explicit model of the data, while non-parametric methods do not pose such assumptions. Since, violations of the assumptions of the model often lead to false rejection of H0, a non-parametric method might provide considerable advantages in detecting JSEs. An example of a parametric method that has been discussed as yielding increased false positive levels is the unitary events analysis (Grün et al. 2002a, b), which hypothesizes that spike trains are stationary Bernoulli processes. Such a process is binary and discrete, with constant probability of spike generation across time as well as with no history dependency. Therefore, several factors might induce false positives: (1) low firing rates (Roy et al. 2000), (2) fast rate modulations (Grün et al. 2002b), (3) auto-structure in the spike trains (Baker and Lemon 2000; Pipa et al., submitted), and (4) trial-by-trail variability in firing rates and in latencies of responses (Brody 1999; Baker and Gerstein 2001; Grün et al. 2003).

2.4 Generation of surrogate data

The non-parametric approach of NeuroXidence estimates the chance frequency of JSEs (H0) based on a surrogate dataset, a technique that has been widely used (Schreiber and Schmitz 2000; Baker and Gerstein 2001). The art

of developing a non-parametric method is to produce surrogate data that differs from the original dataset in only one property. This property defines the alternative hypothesis, H1. Hence, the ideal surrogate dataset is derived from the original dataset, such that it maintains all of the original properties of the data but does not contain any coordinated firing. The surrogate dataset should preserve the auto-structure of each individual spike train, trial-by-trial variability, rate modulations, and all history dependencies.

NeuroXidence generates surrogate data by discriminating two disparate timescales. The first timescale, τ_c, has been already described and corresponds to fine temporal cross-structures of interest for JSEs. This range is typically between 1–10 ms (Markram et al. 1997; Bi and Poo 1998; Grün et al. 1999; Hopfield and Brody 2000; Aertsen et al. 2001; Hopfield and Brody 2001; Sjostrom et al. 2001). The second timescale, τ_r, corresponds to the scale of the rate coding and is defined to be distinguishably slower than τ_c; therefore, a factor, η, is introduced, which is the ratio between the two timescales:

$$\tau_r = \eta \cdot \tau_c, \text{ with } \eta > 1. \tag{1}$$

In the present analyses, η is chosen to be in the range of 2–5, meaning that τ_r is two to five times slower than τ_c. As we will show, this difference in timescales is sufficient to test reliably for the presence of fine temporal cross-structure defined by JSEs, even in data in which the auto-structure of the spike trains changes on the timescale of τ_c.

The surrogate dataset is derived from the original spike trains by jittering spikes by an amount given by the slow timescale, τ_r (Fig. 1(e and f)). In contrast to random jittering, also referred to as 'dithering,' of each individual spike (Fig. 1(f)), which destroys the fine temporal auto-structure, NeuroXidence jitters each spike of the same spike train by the same amount (Fig. 1(e)). Thus, the jittering is equivalent to random shifts of an entire spike train, which destroys fine temporal cross-structure on timescales faster than τ_r, but preserves auto-structures on all timescales. Hence, random shifting of entire spike trains is an optimal strategy to produce surrogate data, and it minimizes the chances of generating inappropriate surrogates.

2.5 Implementation of the statistical test

To assess a deficiency or an excess in the frequency of JSEs within a particular JS-pattern, we compare the frequency of their occurrence in the original dataset (org) to their frequency in the surrogate dataset (sur). For a particular JS-pattern, k, and trial, t, we compute the difference, Δf_t^k, between the two frequencies:

$$\Delta f_t^k = f_t^k(org) - f_t^k(sur), \tag{2}$$

where $f_t^k(\text{org}) \in \mathbb{N}_0$, for trial $t=1...T$ and JS-pattern $k=1...K$. Thus, the variability of Δf_t^k is assessed on a trial-by-trial basis.

To improve the estimate of the chance frequency of JSEs occurring in surrogate data (H0), NeuroXidence allows the use of multiple surrogates (S). Each surrogate sample is based on random and independent shifts to derive the average chance frequency of JSEs per trial. Thus, for $S>1$, the difference Δf_t^k becomes the average difference between the original data and the S surrogates $\overline{\Delta f_t^k}$:

$$\overline{\Delta f_t^k} = \frac{1}{S} \sum_{s=1}^{S} \Delta f_{t,s}^k = f_t^k(\text{org}) - \frac{1}{S} \sum_{s=1}^{S} f_{t,s}^k(\text{sur})$$

with $\overline{\Delta f_t^k} \in \mathbb{Q} \wedge f_t^k(\text{org}) \in \mathbb{N}_0$ \hfill (3)

A value of $\overline{\Delta f_t^k}$ larger than zero indicates an excess of JSEs within one particular JS-pattern, k, in the original dataset. Similarly, a negative value of $\overline{\Delta f_t^k}$ indicates a deficiency of JSEs in the original dataset. The differences $\overline{\Delta f_t^k}$ for one particular JS-pattern k computed for all available trials form the set $\overline{\Delta F^k}$:

$$\overline{\Delta F^k} = \left\{ \overline{\Delta f_1^k}, \overline{\Delta f_2^k}, \dots, \overline{\Delta f_T^k} \right\}.$$ \hfill (4)

NeuroXidence uses the set $\overline{\Delta F^k}$ to test whether the excess or deficiency of JSEs is significant and consistent across trials.

2.6 Rare events versus significant events

The estimation of the statistical significance utilized by NeuroXidence is derived from the trial-based evaluation of $\overline{\Delta f_t^k} \neq 0$. Therefore, NeuroXidence requires that a significant increase or decrease of the frequency of JSEs has to be statistically consistent across trials. This dissociates NeuroXidence from other methods that consider only the total number of JSEs across trials and do not take trial-by-trial variability into consideration (Abeles and Gerstein 1988; Grün et al. 1999, 2002a, b; Ikegaya et al. 2004). To test if $\overline{\Delta f_t^k}$ is a significant increase or decrease of the frequency of JSEs, NeuroXidence allows testing whether either the mean (t test) or median (Wilcoxon rank test) of $\overline{\Delta f_t^k}$ is different from zero. The Wilcoxon test is, in many cases, a much better choice, especially when spike rates are low and complexities of JSEs are greater than two (see Appendix 2).

2.7 Parameters of the NeuroXidence significance estimation

We next investigate the properties of the parameters η and S utilized by NeuroXidence to estimate the statistical significance of JSEs. To this end, we discuss how both affect the false-positive rate and test-power of the method.

The scaling factor η As mentioned, NeuroXidence differentiates between two timescales at which neural processes operate. The first, τ_c, defines the timescale of JSEs. The second timescale, τ_r, is slower and is relevant for the other processes considered by the significance evaluation. The ratio between the two timescales is given by $\eta = \tau_r / \tau_c$ (Eq. (1)). The selection of η is based on a trade-off between the information maintained in the surrogate data and the information destroyed by jittering the spike trains. On the one hand, τ_r should be as fast as possible to account for rapid changes in the cross-structure, such as co-variations in firing rates (Brody 1999; Baker and Gerstein 2001). On the other hand, a large value of τ_r (i.e. large η) ensures that jittering destroys most of the coordinated firing in the surrogate data, which is necessary for a high test-power (probability to reject H0 if H1 is true). The probability that jittering destroys JSEs increases exponentially with the complexity of JSEs, since the number of existing, different jitter configurations is exponentially growing with the complexity of the JS-pattern (see Electronic Supplementary Fig. 2A). Therefore, large η improves mainly the test-power for JS-patterns, whose complexities do not exceed two or three. In conclusion, η should be chosen to be in the range of 2–5, which provides a good balance between the test-power and the dynamics of the cross-structure.

Number of surrogates S As mentioned, Using $S>1$ random surrogates per trial improves the estimation of the frequency, $f_t^k(\text{sur})$, of JSEs by using the average frequency, $\overline{f}_t^k(\text{sur})$, for each particular JS-pattern, k. This improves the estimate of the expected number of JSEs under H0 and increases the reliability of the estimated p value. To understand how multiple surrogates influence the significance estimation in detail, we next derive the impact of $S>1$ on the probability distribution of the average difference, $\overline{\Delta f} = f_{org} - \overline{f}_{sur}$, between the frequency of JSEs in the original dataset, f_{org}, and the average frequency in the surrogate dataset, $\overline{f}_{sur}$. The distribution of the average difference, $\overline{\Delta f}$, is given by a convolution of the distributions $p_{\overline{f}_{sur}}$ and $p_{f_{org}}$:

$$p_{\overline{\Delta f}}(f) = \int df' \, p_{\overline{f}_{sur}}(f') \, p_{f_{org}}(f + f').$$ \hfill (5)

Since $p_{f_{org}}$ is not affected by different numbers of surrogates, we focus solely on changes of $p_{\overline{f}_{sur}}$ and assume that events are rare such that the distribution of $p_{\overline{f}_{sur}}$ is expected to be skewed, with a median smaller than the mean. However, with increasing numbers of surrogates, the distribution of the average frequency, $\overline{f}_{sur}$, becomes more normally distributed than f_{sur} because of the central limit theorem. Therefore, the median of the distribution $p_{\overline{f}_{sur}}$ becomes larger with increasing S and approaches the value of the mean.

To illustrate the impact of this mapping on the significance estimation, we assume the null hypothesis, H0, and define that both $\overline{f}_{sur}$ and f_{org} are sampled from the same distribution $\left(H0 : p_{f_{sur}} \equiv p_{f_{org}}\right)$. In this case, the median and the mean of $p_{f_{sur}}$ and $p_{f_{org}}$ are, by definition, identical. However, as described above, multiple surrogates, first, make the distribution $p_{\overline{f}_{sur}}$ more normal than $p_{f_{org}}$ and, second, make the median of $p_{\overline{f}_{sur}}$ larger than the median of $p_{f_{org}}$. Hence, it follows that in the case of H0, multiple surrogates, i.e., $S>1$, bias the median of $\overline{\Delta f}$ (see Eq. (5)) towards negative values (see Electronic Supplementary Fig. 1A). This makes the reliance of NeuroXidence on a median test more conservative when testing for an excess of JSEs and more liberal when testing for a deficiency of JSEs (i.e. increased numbers of false-positive events). Therefore, it is recommended that one use larger values of S only for testing for an excess of JSEs, while one needs to be cautious when testing for a deficiency of JSEs. In the latter case, $S=1$ is a more proper choice to minimize the risk of type-1 errors.

2.8 False-positive rates

False-positive events are defined as false rejections of the null hypothesis (H0). If a statistical hypothesis test is applied to a recording of large sets of neurons, one has to differentiate between two different scenarios leading to different definitions of false error rates.

Scenario one The test is applied to only one JS-pattern of interest. In this case, the rate of false positives is defined by the percentage of p values lower than the test-level (e.g., $p<0.05$). This percentage of false positives is referred to as the *individual-false-positive-rate*, since it is defined for a single JS-pattern. The inclusion of more than one JS-pattern requires a multiple-comparison correction (e.g. Bonferroni).

Scenario two The test is applied only to those JS-patterns that have occurred in the dataset at least once. In this case, the grand-average false error rate is referred to as the *actual-false-positive-rate*. A test that ensures a conservative actual-false-positive-rate has to ensure that the average false-positive-rates of all occurring JS-patterns are below the test-level. This implies that the statistical significance must be independent of the occurrence of the JS-patterns, which is achieved by evaluating the significance on the basis of multiple trials. The consideration of the trial-by-trial variability of the frequencies of occurrences prevents JS-patterns, which are by chance occurring in a few or even only one trial, from being considered to be significant. Thus, significant JS-patterns are required to occur reliably across trials.

In the context of simultaneous recordings of large sets of neurons, only scenario two is of practical usage, since, for scenario one, a multiple-comparison correction has to be performed (e.g. Bonferroni). The latter corrects the test-level for the number of JS-patterns that were tested (K). Since the number of JS-patterns can be very large, the Bonferroni correction can lead to a corrected test-level that is extremely small (e.g. for 20 neurons $K>10^6$, and a 1% level leads to a corrected level of $<10^{-6}$), reducing thus considerably the statistical power.

3 Results

In the following sections we report the statistical and computational properties of NeuroXidence. To this end we utilize simulated data as well as an example dataset consisting of 42 single units simultaneously recorded from cat visual area 17 of an anaesthetized cat.

3.1 False positives for stationary processes

Simulated data, generated by independent and stationary Poisson processes, were used to assess the individual-false-positive-rate of NeuroXidence for one individual JS-pattern of interest (Fig. 3). The individual-false-positive-rate for each individual JSE was derived from 100 independent hypothesis tests, based on independent realizations of the same simulated-data model. The standard set of parameters used to generate the simulated data was defined by 50 trials (T), a mean spike rate of 15 ap/s (r), 20 surrogates samples (S), and η equal to 3. From this standard set, 15 different combinations of parameters were derived by varying the number of trials ($T=20, 50, 100, 200$), the number of surrogates ($S=1, 50, 250$), the mean spiking rate ($r=7, 10, 30, 60, 90$ ap/s), and the scaling factor ($\eta=2, 5, 7$). NeuroXidence was applied to each simulated dataset using a sliding window with durations of $l=200, 400$, and 800 ms. In total, 48 different simulated-data models were used to derive the individual-false-positive-rate for 4 JS-patterns of complexities 2–5. None of the individual-false-positive-rates for any of the parameter sets was above chance-level, for either test-level of 5 or 1% (Fig. 3), demonstrating that NeuroXidence is a conservative statistical hypothesis test. Furthermore, NeuroXidence becomes more conservative with increasing complexity.

3.2 Test-power for stationary processes

Test-power for individual JS-patterns was derived from correlated Poisson processes, generated by a single-interaction process (Kuhn et al. 2003). Thus, correlated spike trains were characterized by a background rate, corresponding to
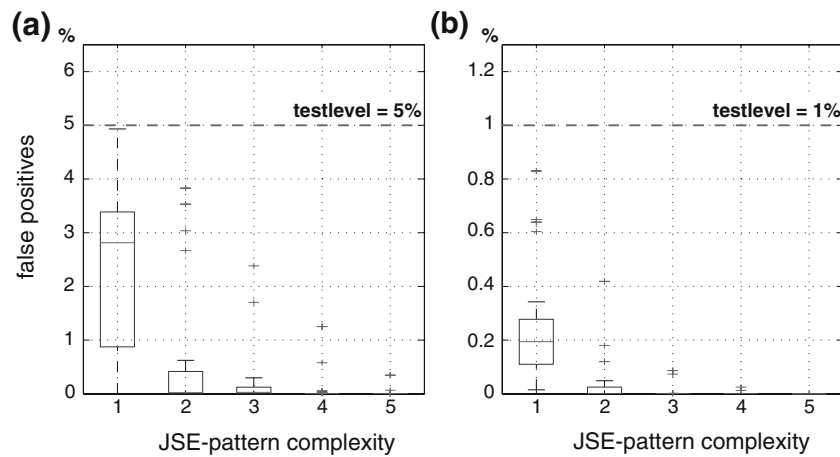
Fig. 3 Percentile plots of the percentage of false-positives estimates by NeuroXidence used to detect joint-spike patterns of complexities 2 to 5. Independent and homogenous Poisson processes were used to generate simulated data that were analyzed by NeuroXidence to estimate the false-positive rate based on 100 independent realizations of each simulated-data model. Each individual percentile plot represents 48 parameter sets describing, three window lengths (1=200 ms, 400 ms, 800 ms) and 16 simulated-data models, which differed, compared to the standard-data model (50 trials, spike-rate 15 ap/s, 20 surrogates, and $\eta=3$), by variation of the mean spike rate (r=7, 10, 30, 60, 90 ap/s), the number of trials (T=20, 50, 100, 200), the number of surrogates (S=1, 50, 250), and η=2, 5, 7. The percentage of false-positives for a test-level of 5% (**a**) and for a test-level of 1% (**b**)

the independent spiking of neurons, and by a JSE rate, defining the expected frequency of the JS-pattern of interest beyond chance-level. Sixteen simulated datasets were used to study how the test-power of NeuroXidence is influenced by the length of the sliding window, the number of trials, the background spike rate, and the number of surrogates (Fig. 4). Dataset 1 contained 50 trials, each 800 ms long, the background spike rate was 15 ap/s, the maximum allowed jitter was 5 ms, and η was 3. Each of the remaining 15 sets was derived from the first set by variations of either the sliding window length (l=200, 400, 800 ms, Fig. 4(a, rows *1–4*)), the number of trials (T=20, 50, 100, 200, Fig. 4(b, rows *1–4*)), the background spike rate (r=10, 15, 30, 60, 90 ap/s, Fig. 4(c, rows *1–4*)), or the number of surrogate datasets (S=1, 20, 50, 250, Fig. 4(d, rows *1–4*)).

Based on these 16 simulated datasets, the test-power was derived for four JS-patterns of complexity 2–5. The test-level used by NeuroXidence was 1%. Increasing the length of the analysis window, as well as increasing the number of trials, led to increased test-power. Given the same frequency of excess JSEs, the test-power tended to increase along with JSE complexities. This was expected since the probability of a chance occurrence of a JSE exponentially decreased with its complexity. The number of surrogates used for the estimation of the average frequency $\overline{\Delta f_t^k}$ changed the test-power only very subtly in comparison to the changes observed when varying the window length, the number of trials, and the spike rate. In particular, the differences in test-power for $S>20$ were negligible, indicating that going beyond 20 surrogates does not substantially improve the reliability of the estimates (Fig. 4(d, rows *1–4*)).

3.3 Sub-patterns and supra-patterns of induced JS-patterns

To study the test-power of sub-patterns and supra-patterns of individual JS-patterns induced by a single-interaction process (Fig. 5), we performed a statistical evaluation of the frequency of occurrence of any potentially existing JS-patterns in 4 simulated datasets. Each dataset comprised 50 trials of 16 simultaneous spike trains. The 4 simulated datasets differed in the complexities (2–5) of the mother-patterns, which are the JS-patterns induced by the single-interaction process (mother-process; Fig. 5).

Any JS-pattern of complexity higher than 2 includes sub-patterns that are expected to be detected at least as frequently as the mother-pattern. On the other hand, the test-power of each of these sub-patterns is expected to be lower than the test-power of the mother-pattern, since the lower complexity leads to a higher chance frequency under H0. Thus, when sub-patterns are only induced by a single mother-process and not by additional correlations with orders equal to or smaller than the complexity of the sub-patterns, the test-power is expected to decrease with the decreasing complexity of the sub-pattern.

Supra-patterns are composed of the mother-pattern itself and spikes from additional neurons, which by chance coincide with the mother-pattern. Therefore, supra-patterns are of higher complexity than the induced mother-pattern. The maximal frequency of a supra-pattern is bound by the maximal frequency of any sub-pattern, including the induced mother-pattern. Furthermore, since the additional spikes are coinciding by chance, the frequency of the supra-pattern occurring is expected to be smaller than that for the mother-pattern. Thus, as long the excess frequency of the
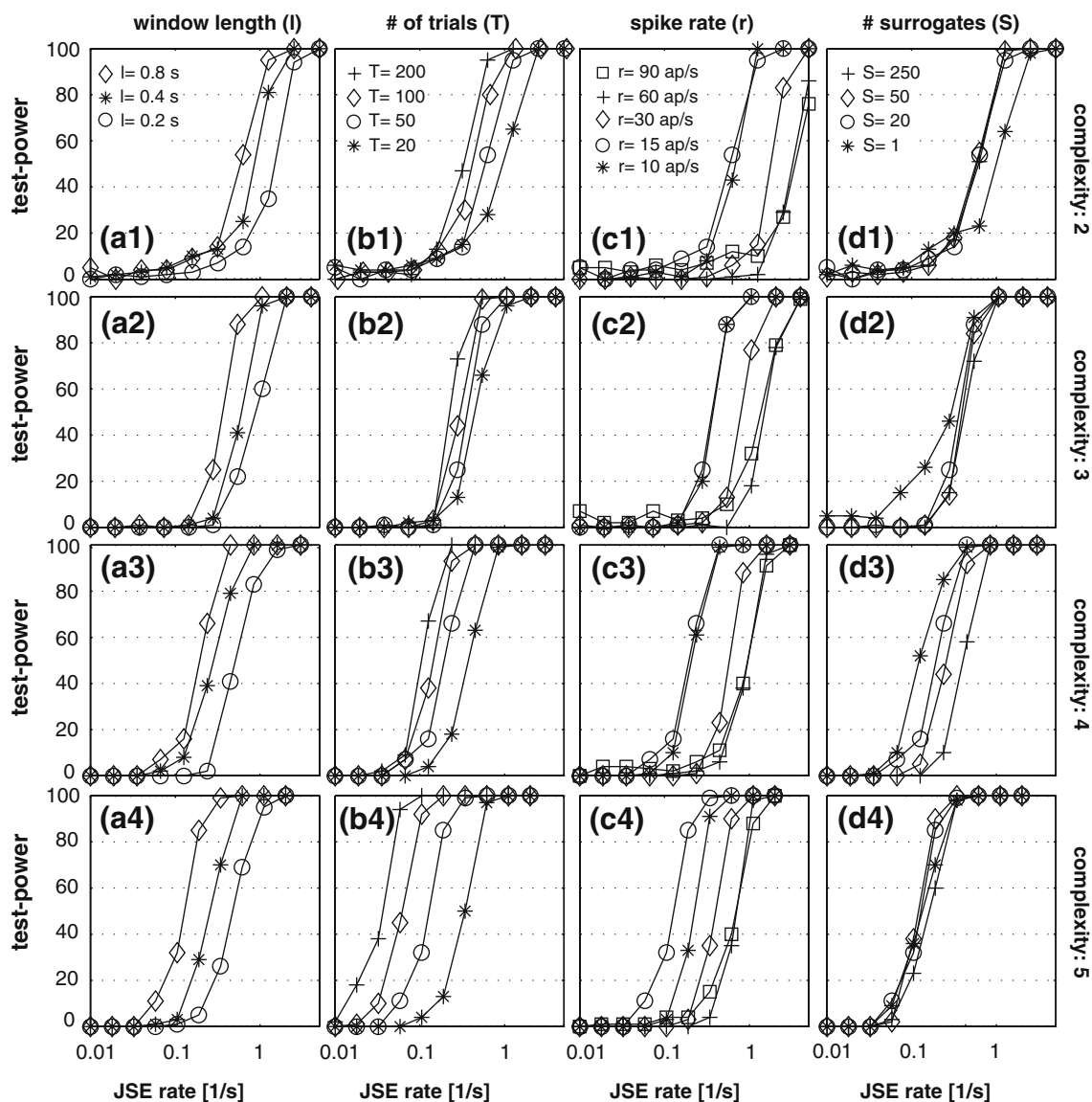
Fig. 4 Test-power of NeuroXidence in relation to the window length (*l*), number of trials (*T*), spike rate (*r*), and number of surrogates (*S*). Subfigures plot the test-power of NeuroXidence as a function of the frequency of joint-spike events (*JSEs*) beyond chance level. Spike trains were analyzed by NeuroXidence and modeled as a single-interaction process based on Poisson processes. Rows *1–4* show the

test-power dependencies on the complexities of the analyzed joint-spike pattern ranging from 2 to 5. (**a–d**) Variations in (**a**), the length of the analysis window *l*; (**b**), the number of trials *T*; (**c**), the spike rate *r*; and (**d**), the number of surrogates *S* from the standard parameter set (*T*=50, *r*=15 ap/s, *S*=20, *η*=3) are presented

mother-pattern is not so high that the supra-pattern still occurs reliably across trials, the test-power of the supra-pattern is expected to be substantially reduced in relation to the mother-pattern.

3.4 False positives and test-power for a non-stationary process

Using data generated by a non-stationary process, we derived the percentage of false positives and the test-power of NeuroXidence (Fig. 6(c and d)) and compared the results with the widely used shuffle-corrected cross-correlogram

(Fig. 7(a)) and the UE-method (Fig. 7(b–d), Appendix 3). The simulated data comprised 18 simultaneous units and 50 trials, which were based on 15 periods, each 2 s long. The simulated dataset also contained features that are frequently observed in real datasets (Fig. 6(a and b)), such as low rates (periods 5, 6), rate modulation (periods 7–12), latency co-variation of rate responses across neurons (periods 10, 12), bursty (period 2,3) and regular spiking (period 4) modeled by a γ-process with three different shape factors, and trial-by-trial variability of rate levels (period 13). Only periods 14 and 15 contained JSEs above chance-level, which were based on correlations modeled by a multiple-interaction
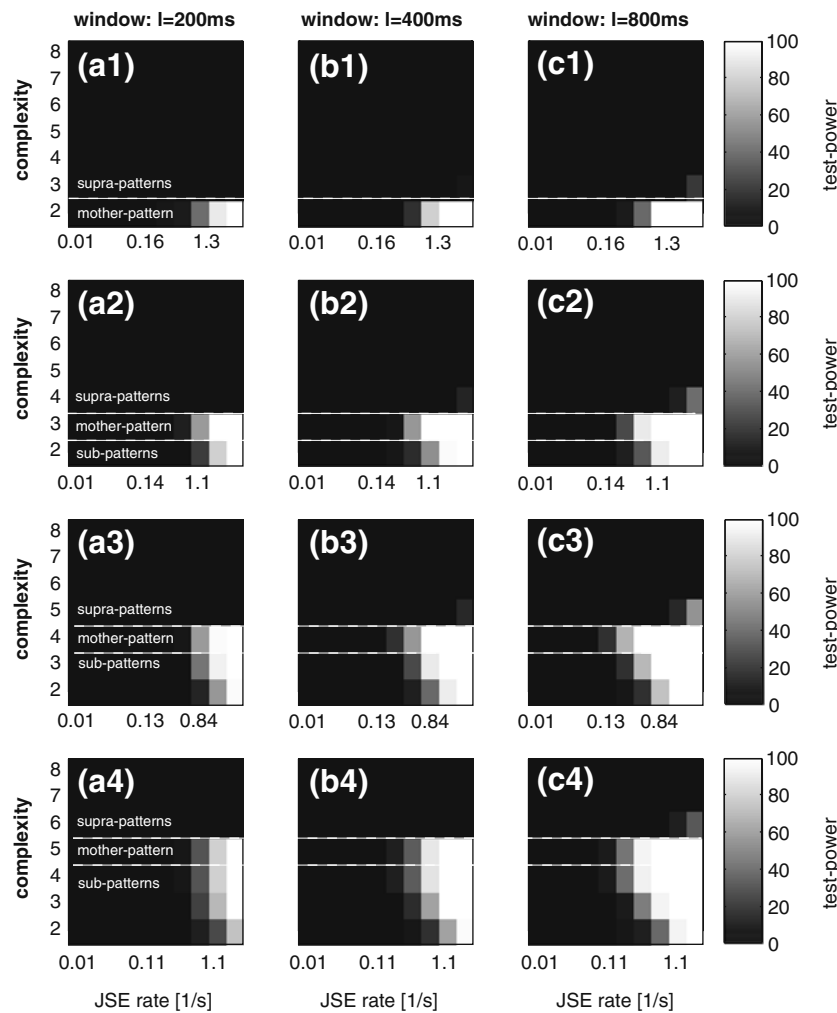
Fig. 5 Test-power of NeuroXidence for an induced mother-pattern and its supra-patterns and sub-patterns. Each subfigure shows the gray-coded test-power of a certain mother-pattern, all sub-patterns of lower complexities, and all supra-patterns of higher complexities. The excess rate of the JSEs beyond the chance-level, which corresponds to the mother-pattern, is given on the x-axis. In row *1* the mother-pattern was of complexity 2, in row *2* of complexity 3, in row *3* of complexity 4, and in row *4* of complexity 5. The data used to derive the test-power consisted of 50 trials of 16 spike trains, modeled as a single-interaction process that is based on homogeneous Poisson processes. A spike rate $r=15$ ap/s, $S=20$ surrogates, and $\eta=3$ were used by NeuroXidence for deriving the statistical significance. (**a**–**c**) Variations in the length of the analysis window l. $l=200$ ms (**a**), $l=400$ ms (**b**), $l=800$ ms (**c**) are presented

process (Kuhn et al. 2003). NeuroXidence was applied to the data in a sliding window of length 800 ms, and significance tests were performed on excesses of JSEs with complexities 2–6. The statistical significance was evaluated for each occurring JS-pattern. We derived the percentage of JS-patterns that occurred significantly more frequently than expected by chance (Fig. 6(d)). To make comparisons across complexities, the number of significant JS-patterns per complexity was normalized by the total number of identified JS-patterns for that complexity level (Fig. 6(c)).

Spike trains during periods 1–13 were independent, which implied that H0 should not be rejected more often than the test-level (5%), if the actual-false-positive-rate were conservative. The percentage of false rejections is clearly below 5% for all complexities and throughout all sliding

windows during periods 1–13. Therefore, NeuroXidence is a hypothesis test with a conservative actual-false-positive-rate that is effected neither by low rates, rate modulation, latency variability, and cross-trial rate changes nor by the different model processes (γ-process, Poisson) used to generate the spike train. High test-power during periods 14 and 15, which were characterized by increased frequencies of JSEs based on multiple-interaction process correlations with $\zeta=0.12$ (period 14) and $\zeta=0.3$ (period 15), demonstrates the applicability of NeuroXidence for the detection of excesses of JSEs.

Unlike NeuroXidence, the shuffle-corrected cross-correlogram (bin width=20 ms, data segments=800 ms) indicated the existence of cross-structure due to its modulation in periods 10 and 12 (Fig. 7(a)). The reasons for these modu-
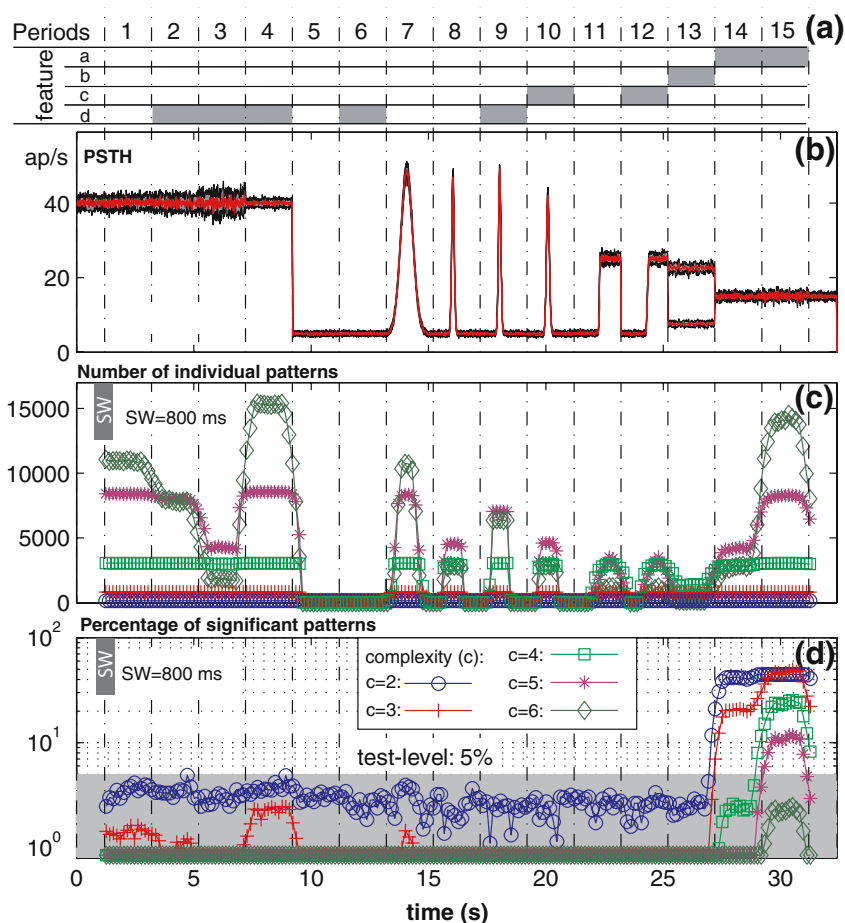
Fig. 6 Statistical significance of an excess of JSEs evaluated by NeuroXidence. Generated simulated data consisted of 50 trials of 18 'simultaneous' spike trains. (**a**) The simulated data contained 15 periods each 2 s long. Each period was described by one set of features used to describe the simulated-data model and to generate the spike trains. An inhomogeneous and independent Poisson process served as a standard model. Four additional features were used as modifications of this standard model. Only periods *14* and *15*, indicated by feature *a* spike trains, were inter-correlated and thus exhibited a statistical excess of JSEs (multiple-interaction process, two groups each with nine neurons are correlated, period *14* correlation $\zeta = 0.12$, period *15* $\zeta = 0.3$). Feature *b*, which was used for period *13*, indicates changing rates across trials and across neurons. Neurons *1–9* were modeled by a homogenous Poisson process with a rate=15 ap/s, while the rate of neurons *10–18* changed from trial to trial between 15–30 ap/s. Feature *c* (periods *10*, *12*) represents latency covariations. From trial to trial, the latency for all

18 neurons varied randomly by the same amount between 0–100 ms. During the periods characterized by feature *d*, inhomogeneous gamma processes (shape factor $\gamma = 7$ (regular) for periods *4*, *6*, and *9*, $\gamma = 0.7$ (bursty) for period *2*, and $\gamma = 0.3$ (bursty) for period *3*) were used instead of inhomogeneous Poisson processes. (**b**) PSTH displays of the rate profiles of the used inhomogeneous processes. During period *7*, rate had been modulated between 5 and 50 ap/s with a Gaussian shape with $\sigma_t = 250$ ms, while during periods *8*, *9*, and *10*, $\sigma_t = 50$ ms. The rates in periods *11* and *12* were modulated between 5 and 30 ap/s by a step function. (**c**) The number of individual and unique JS-patterns of complexities 2–6 that were detected in each sliding window ($\tau_c = 5$ ms, 'SW'=sliding window, 800 ms). (**d**) The percentage of JS-patterns that could not be explained by chance (test-level 5%). In the case of independent processes (periods *1–13*), this percentage corresponds to the percentage of false positives, while it corresponds to the test power during periods *14* and *15*

lations are the shuffle-correction and latency co-variations across neurons that characterize periods 10 and 12 (Brody 1999; Baker and Gerstein 2001). This indicates that modulation in the shuffle-corrected cross-correlogram does not necessarily indicate an excess of fine temporal cross-structure.

We also applied the UE-method (see Appendix 3) to the same dataset (Fig. 7(b–d)). Since the method assumes that data in every sliding window can be described by a stationary Bernoulli processes, the sliding window was chosen to be rather short (100 ms) compared to the 800 ms-long Neuro-

Xidence window. This window length of 100 ms justifies the assumption of a stationary Bernoulli process for a slower modulation of the rate profiles. Therefore, 100 ms is a good trade-off between the danger of false-positive events induced by non-stationarity and the reduced test-power due to the small amount of data contained in each window.

For the UE-method, we derived the actual-false-positive-rate per JS-pattern complexity, which ranged from 2–6 during periods 1–13. As with NeuroXidence, the actual-false-positive-rate was defined as the rate of false rejections of H0, normalized by the number of occurred JS-patterns.
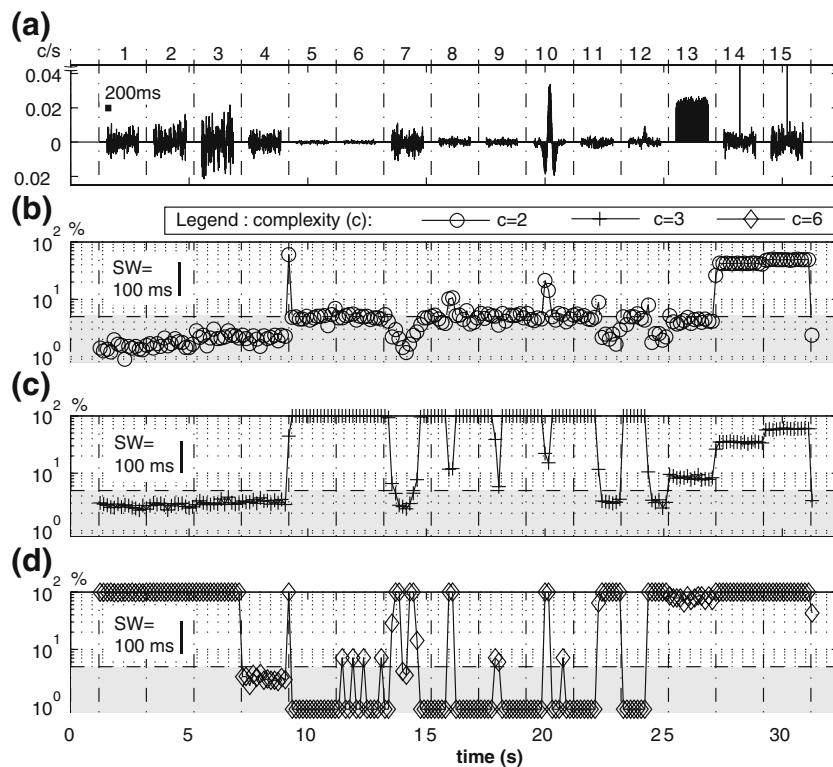
Fig. 7 Comparison of cross-correlograms and the unitary-event method for detecting statistical excesses of joint-spike events. The comparison was based on the same simulated dataset as in Fig. 6. (**a**) Cross-correlograms for the respective periods (the scale bar in period 1 indicates 200 ms). (**b–d**) The percentage of JS-patterns detected by the UE-method, with complexities ranging from 2–6, that could not be explained by chance (5%) [(**b**): complexity=2, (**c**) complexity=3, (**d**) complexity=6]. In the case of independent processes (periods 1–13), this percentage corresponds to the percentage of false positives, while it corresponds to the test-power during periods 14 and 15 (binning=5 ms, 'SW'=sliding window with length 100 ms)

As expected, fast changing rates were nearly fully compensated by the short sliding window (100 ms) for complexity 2 JSEs (Fig. 7(b)), while high-complexity JS-patterns led to an increased actual-false-positive-rate during periods with fast changing rates (periods 7–12; Fig. 7(d)). The second source for increased actual-false-positive-rates was low spike rates. During periods 5 and 6 and all parts of other periods in which the rate was low (5 ap/s), the actual-false-positive-rate increased dramatically for complexity 2 JS-patterns and up to 100% for complexity 3 patterns (Fig. 7(c)). In contrast, the actual-false-positive-rate for complexity 6 JSEs was close to zero because their existence was not detected due to the usage of exclusive binning. The reason for the dramatic increase in the actual-false-positive-rate is that the UE-method does not consider any trial-based variation in the frequency of occurrences. Instead, it tests if the total number of a particular JSE across all trials is significantly different from the expected total number, which is based on the assumption that the expected number is Poisson distributed. Since the significance estimation of the UE-method is only based on the variability of the expected total number and not on the variability of frequency of occurrences across trials, a few or even only one JSE might be evaluated as a significant excess if the expected total number is low enough. The latter is the case if spike rates are low or the complexity is large.

3.5 Test-power and false-positives rate changes co-occurring with JS activity

To demonstrate that NeuroXidence detects epochs containing coordinated firing, even in the presence of rate modulation of the same neurons that are firing synchronously, we generated simulated data that contained rate modulations (Fig. 8(b)) as well as short epochs of coordinated firing (Fig. 8(a)). The simulated data was comprised of four periods. Periods 1–3 contained rate modulation modeled by the same Gaussian waveform (σ_t=200 ms, peak rate 70 ap/s), while period 4 contained a rate modulation with a steeper rising than falling slope (composed with two Gaussian waveforms, rising σ_t=100 ms and falling σ_t=300 ms, peak rate 70 ap/s). In addition, periods 2, 3, and 4 contained short epochs (Fig. 8(a)) with coordinated firing modeled by an MIP process (Kuhn et al. 2003). NeuroXidence was applied to the data with a sliding-window length of 100 ms. We evaluated the
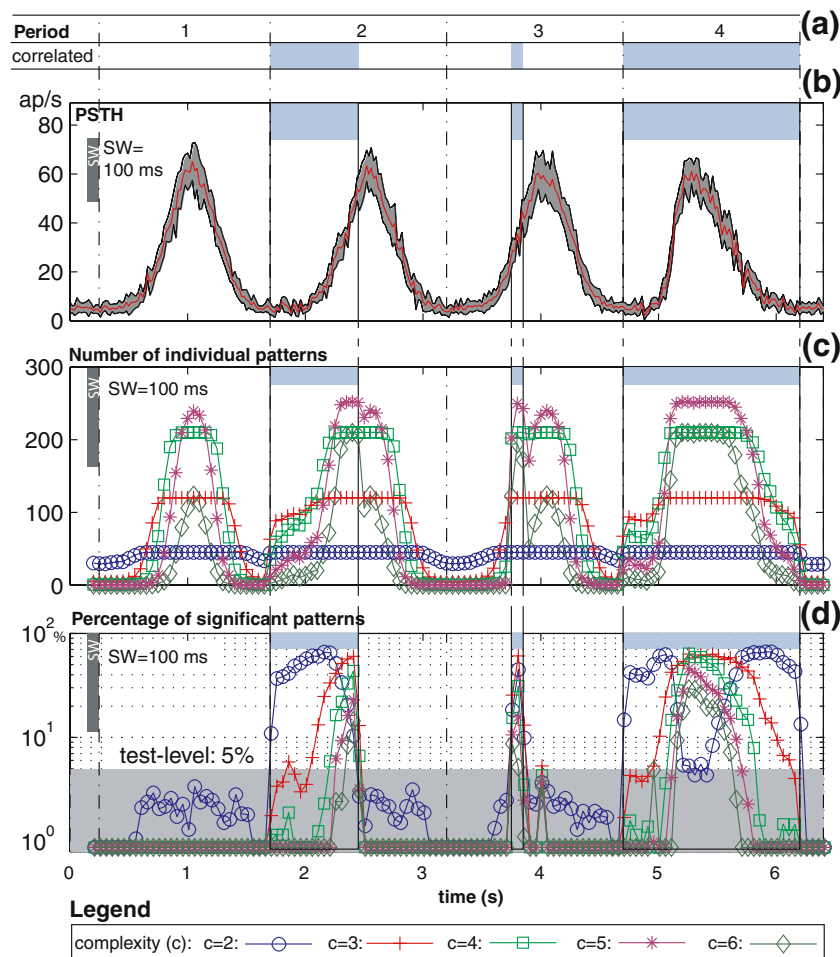
Fig. 8 Statistical significance of an excess of JSEs evaluated by NeuroXidence. Generated simulated data consisted of 50 trials of ten 'simultaneous' spike trains. (**a**) The simulated data contained four periods each 1.5 s long. Inhomogeneous and independent Poisson processes with a spike rate of 5 ap/s served as a standard model. During periods *1*, *2*, and *3* the spike rates of all ten neurons changed following a Gaussian modulation (peak rate=70 ap/s, σ_t=200 ms). Rate modulation in period *4* was skewed since its first half was composed of a Gaussian modulated with σ_t=100 ms and its second half with σ_t=300 ms. *Gray-shaded windows* at the beginning of period *2* (window length=0.75 s from 1.7–2.45 s), during the rising slope of period *3* (window length=0.1 s from 3.75–3.85 s), and for the whole of period *4* (window length=1.5 s from 4.7–6.2 s) were inter-correlated and thus exhibited a statistical excess of JSEs (multiple-interaction process, ζ=0.2). (**b**) PSTH displays the rate profiles of the used inhomogeneous processes. (**c**) The number of individual and unique JS-patterns of complexities 2–6 that were detected in each sliding window (τ_c=5 ms, 'SW'=sliding window: 100 ms, η=3). (**d**) The percentage of JS-patterns of complexities 2–6 that could not be explained by chance (test-level 5%). In the case of independent processes, this percentage corresponds to the percentage of false positives, while it corresponds to the test-power during periods based on correlated multiple-interaction processes

statistical significance for each occurring JS-pattern with complexities 2–6, derived the frequency of significant JS-patterns (Fig. 8(c)), and computed the percentage of significant JS-patterns per complexity (Fig. 8(d)). During periods with rate modulation, but lacking fine temporal structure, the percentage of false rejections of H0 was below the test level of 5% for all complexities. In contrast, periods that contained coordinated firing lead to an increased frequency of significant JS-patterns (Fig. 8(d)). However, in the case of an MIP process, since the rate of JSEs increases for an increasing spike rate, the test power itself was also modulated. Hence, the precise relation of the test-power and spike rate depends on the chosen model of

coordinated firing. Irrespective of this precise relation, the simulated data demonstrates that NeuroXidence is capable of detecting periods of coordinated firing even in the presence of rate modulation of the same neurons that are firing synchronously.

3.6 Test-power of NeuroXidence for oscillatory processes

Since oscillatory modulations of the spike rate are frequently observed in experimental data (Gray et al. 1989; Fries et al. 1997, 2001a, b, 2002; Rodriguez et al. 1999; Tallon-Baudry et al. 1999, 2004; Lachaux et al. 2000), we studied if such modulations lead to rejections of H0 and to

support of H1. To this end we assumed as a model a continuous instantaneous-rate function, $r_i(t)$, that is coherently-modulated across neurons by a sinusoidal oscillation with frequency v:

$$r_i(t) = \frac{A}{2} * [1 + \sin(vt)] \text{ with } i = 1 \dots N \qquad (6)$$

We used the instantaneous rate function (A=10 ap/s) to model spike trains by inhomogeneous Poisson processes and estimated numerically the test-power of complexity 2 JS-patterns. NeuroXidence, based on 50 trials of 800 ms-long analysis windows, was applied to 100 independent realizations of the data. The oscillation frequency ranged between 1 and 80 Hz, while the jitter-kernel width was kept fixed [τ_r=21 ms in Fig. 9(a) and τ_r=11 ms in Fig. 9(b)]. For oscillation frequencies below 6 Hz in the case of τ_r=21 ms and 9 Hz for τ_r=11 ms, the percentage of significant events stayed below the test-level of 5%. With increasing oscillation frequencies, the test-power increased and reached up to 100% at 50 Hz for τ_r=21 ms and at 80 Hz for τ_r=11 ms.

In addition to this numerical estimation of the NeuroXidence test-power, we used an analytical description of the expected difference, Δf, between the frequency of JSEs in the original, $f(org)$, and the jittered, $f(sur)$, datasets to approximate the theoretically predicted test-power $A'(v)^c$ (see Appendix 4). The agreement between the theoretical

prediction of the test-power, in percent given by $100\,A'(v)^c$, and the numerical estimation based on simulated data, both for c=2, indicates that the modulation of the test-power is well described by the analytical solution (Fig. 9). Therefore, $A'(v)^c$ can be used to discuss the impact of coherent oscillations on the probability of rejecting H0.

3.7 Application to real data

We applied NeuroXidence to neuronal activity of 48 single units recorded in area 17 of an anesthetized cat in response to a drifting sinusoidal grating. Gratings were presented for 3.5 s and 20 trials with movement in directions of 30° and 120°. The direction of 30° was the optimal direction for eliciting strong responses in most of the cells, while 120° was chosen to be perpendicular (see Appendix 5 for details on experimental methods).

We computed cross-correlation histograms (König 1994) for all pairs of neurons for the period during which neurons showed sustained responses to the presented stimuli. After the subtraction of shift-predictors (König 1994), cross-correlograms indicated that a number of spikes were synchronized with a precision of several milliseconds (Fig. 10), raising two questions: first, whether higher complexity JSEs might exist in this dataset, and second, whether JSEs are direction specific.

The NeuroXidence algorithm was applied to each set of trials for both stimulus conditions and the entire duration of the recording, which also included the 1 s pre-stimulus period and the 1 s of post-stimulus activity (Fig. 11). The average computation time per sliding window amounted to 35 min on a 1.4 GHz CPU with 1 GB RAM. We investigated JSEs of complexities 2–48 and used 1% as a test-level (τ_c=5 ms, τ_r=20 ms, S=50, bin size=1 ms, Wilcoxon test). Up to 14,769 different and individual JS-patterns per sliding window (length 200 ms, window shift 100 ms) were detected (Fig. 11(a, 2 and b, 2)). JS-patterns occurred significantly more often than expected by chance only during the period of stimulation, and their complexities ranged between 2 and 8. No significant patterns were found of complexity 9 or higher (Fig. 11(a, 3 and b, 3)).

Interestingly, the time-course of the significant JS-patterns partially followed the temporal rhythm of the stimulation. The number of patterns recurred with a period of 820 ms, which corresponded with the timing of the stripes of the grating crossing the receptive fields (Fig. 11(a, 2 and 3, b, 2 and 3), compare to PSTHs in Electronic Supplementary Fig. 3). Also remarkable is the steep increase and decrease at the beginning and end of the period of stimulation. Notably, both the number of JS-patterns as well as the number of significant JS-patterns indicate stimulus specificity of joint-spike activity. This confirms previous findings restricted to pairwise analysis
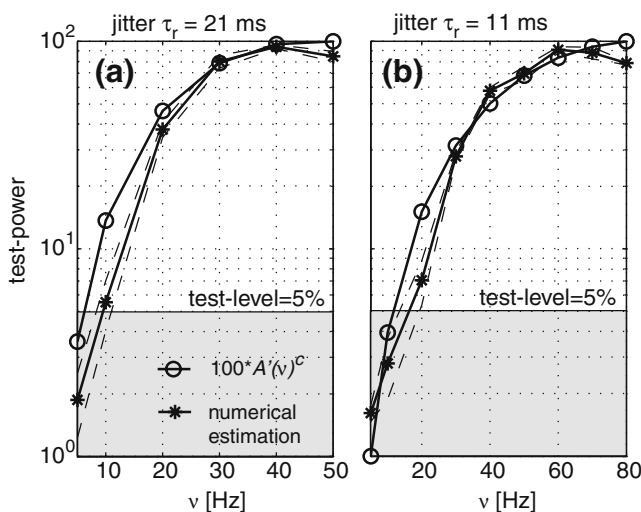

Fig. 9 Test-power of NeuroXidence based on data modeled as independent and inhomogenous Poisson processes. The instantaneous rate functions were modulated by a sinusoid with frequency v and a peak rate of 20 ap/s. Twenty surrogates and 21 ms of jitter were used in (**a**), while in (**b**) 11 ms of jitter were used for deriving statistical significance. The *black solid curve* (*asterisk*) shows the mean test-power (*dashed black*: confidence interval of 1σ) based on 100 independent realization of the same data model, while the other curve (*open circle*) shows the theoretical prediction given by $A'(v)^c$

Fig. 10 All 15 shift-corrected cross-correlograms. The cross-correlograms were derived for the period between 0 and 3.5 s with 5 ms-wide bins. Each subplot represents the shift-corrected cross-correlograms for one combination of two channels. The *y*-axis gives the number of excess coincidences per second. The *x*-axis gives the time delay in milliseconds
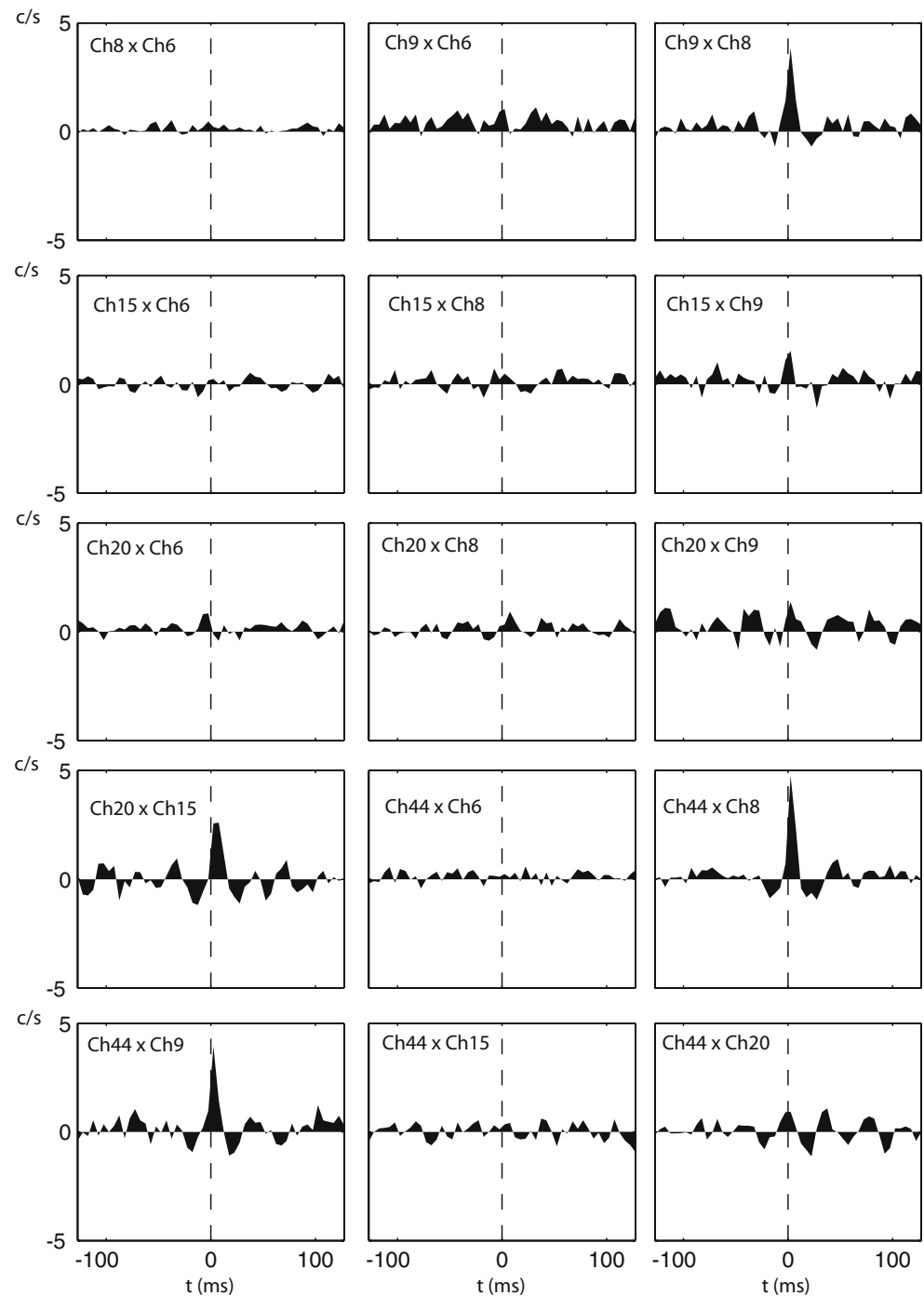
(Samonds et al. 2003). To express this in numbers, we averaged the number of patterns per complexity between 0.1 and 3.5 s after stimulus onset for each condition. Then, we computed the difference, for each individual complexity, between the mean across times for both movement directions and normalized the difference to the average for movement at 30°. The mean number of JS-patterns for 30°

movement exceeds that for 120° by an average of 4.16 and by a maximum of 5.43, while the corresponding numbers for significant JS-patterns are 1.71 and 3.03. Compared to the modulation of the rate response for the same period, which amounts to 0.18, the average modulation of 1.73 of the frequency of significant JS-patterns indicates that joint-spike activity has higher discriminative power than rate.
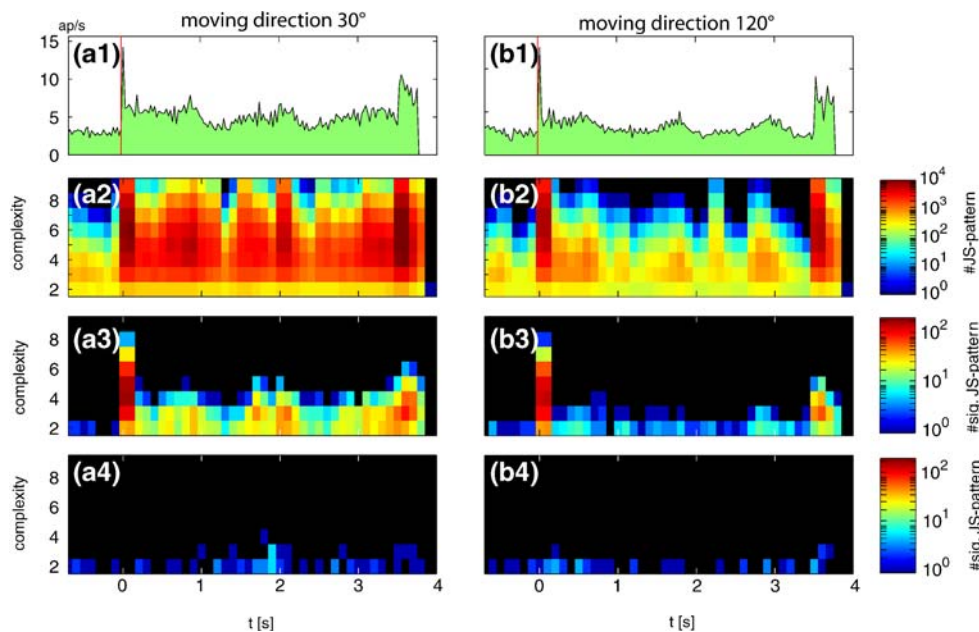
Fig. 11 Analysis of simultaneously recorded single-unit activity from 42 neurons from an anesthetized cat. The recordings were made in area 17 during visual stimulation with moving gratings, which were presented for 3.5 seconds. Stimulus onset occurs at 0 s. Data comprise 20 trials. In (**a**) the stimulus direction was 30° while it was 120° in (**b**). (**a** *1*, **b** *1*) Average PSTHs across all 42 neurons for both stimulus directions. NeuroXidence was applied on sliding windows (0.2 s) that were shifted by 0.1 s (τ_c=5 ms, τ_r=20 ms, S=20). (**a** *2*, **b** *2*) The numbers of unique JS-patterns detected per sliding window are color coded and sorted according to their complexities (y-axis) (no significance criteria applied, color code is logarithmic). (**a** (*3*, *4*)), (**b** (*3*, *4*)) The frequencies of significant JS-patterns (test-level 1%, tested on significant excess) for each sliding window (*x*-axis) are color coded and sorted according to their complexities (*y*-axis). Frequencies shown in (**a** (*3*)), (**b** (*3*)) are based on the original dataset, while in (**a** (*4*)), (**b** (*4*)) the spikes in the original data were jittered by ±10 ms to destroy fine temporal cross-structure

To validate the results given by the number of JS-patterns that exhibited a significant increase of JSEs, we also applied NeuroXidence to a dataset derived by jittering the original dataset. Thus, we destroyed any coordinated-firing by jittering spikes with ±10 ms (Fig. 11(a, *4* and b, *4*)). This jittering of the data is the same as that used by NeuroXidence to derive the difference between the real data and H0 (j=±10 ms, η=4). Hence, the application of NeuroXidence on the jittered dataset allows the derivation of the chance frequency of significant JS-patterns (H0 against H0). The number of significant JS-patterns that occurred in the jittered dataset amounted to 0.2% of the number observed in the original dataset. Thus, the error rate was much smaller than the nominal alpha value of 0.01, indicating that the test is conservative. The positive side is that 99.8% of the significant JS-patterns could not be explained by chance (Fig. 11(a, *3* and b, *3*)).

4 Discussion

The distinction between rate codes and temporal codes in the cerebral cortex has been the origin of heated discussions over recent years (von der Malsburg 1981; Aertsen et al. 1989; Gerstein et al. 1989; Shadlen and Newsome 1994;

Riehle et al. 1997; Fetz 1997; Brody 1999; Baker and Lemon 2000; Aertsen et al. 2001; Grün et al. 2003; Ikegaya et al. 2004). The rate hypothesis assumes that information is encoded in the discharge frequency of individual cells, while the temporal code assumes that information is encoded in the precise timing relations between individual discharges of distributed neurons. The latter can be tested by a statistical hypothesis test that checks whether fine temporal cross-structure represented by joint-spike events (JSE) exists and is beyond chance level. Therefore, the statistical hypothesis test is crucially dependent on the statistical formulation of the chance level of JSEs under the null hypothesis (H0: rate-hypothesis). To distinguish the temporal from the rate code hypothesis, H0 has to consider all features of the data that might be causally related to the rate hypothesis. Only then is a rejection of H0 interpretable in favor of the temporal coding hypothesis and provides evidence for cooperative coding based on joint-spike activity.

We next outline the properties of the NeuroXidence method in detail, starting with statistical testing and the basic idea of timescale separation for NeuroXidence. Then we describe the distinction between rare, spurious, and reliably reoccurring events, the choice of appropriate window lengths, the impact of oscillatory and coherent rate modulations, and the differences between NeuroXidence

and other methods that are used for detecting coordinated firing. Last, we discuss results obtained form real data recorded from area 17 in an anesthetized cat.

4.1 Statistical testing and timescale separation of NeuroXidence

NeuroXidence implements a statistical hypothesis test for the analysis of spiking activity utilizing a large number of neurons, which allows for a dissociation of the temporal-code and the rate-code hypotheses. In contrast to earlier methods (Abeles and Gerstein 1988; Grün et al. 1999, 2002a), NeuroXidence is a non-parametric statistical hypothesis test and does not assume any model about the spike-train generating process. The estimation of the statistical significance of JS-patterns is based on a comparison of the frequency of JSEs between the recorded dataset and one or more surrogate datasets, which are derived from the recorded dataset but lack any coordinated-firing. Random shifting of spike trains was utilized to destroy fine temporal cross-structure while keeping the complete auto-structure in the surrogate dataset intact. Therefore, Neuro-Xidence estimates the chance frequency of JSEs under H0 by allowing for the consideration of rate modulations or for history dependencies, which might exist in each individual spike train.

To this end, NeuroXidence defines two timescales, τ_r and τ_c. The fast timescale, τ_c, corresponds to fine temporal joint-activity across neurons and is believed to be in the millisecond range. The second timescale, τ_r, is a lower bound of rate modulations that are coherent across neurons and that should be considered as rate, not as fine temporal joint-activity. The scale between τ_r and τ_c is given by η. As we have shown, small amounts of shifting with $\eta=3$ lead to the considerable test-power of the NeuroXidence method. Thus, even in cases of very small differences between the timescales τ_r and τ_c (e.g. $\tau_c=15$ ms for $\tau_r=5$ ms and $\eta=3$), NeuroXidence can detect a significant excess or deficiency of JSEs and, therefore, allows for a clear distinction between the rate- and the temporal-code hypotheses.

In conclusion, the concept of timescale separation, the reliability and sensitivity of NeuroXidence, and the high temporal precision for detecting periods of coordinated firing in short sliding windows all make NeuroXidence ideally suited for the analysis of neuronal activity that might exhibit both rate changes and coordinated firing.

4.2 Synchronous events versus spatiotemporal patterns

NeuroXidence is optimized to detect synchronous firings. However, it can also be applied to the analysis of spatiotemporal patterns (Abeles and Gerstein 1988), each consisting of one spike per individual neuron. This requires mapping each spatiotemporal pattern onto a corresponding synchronous event by shifting each spike train by its corresponding temporal delay. Since this mapping is required for each individual spatiotemporal pattern, the computational complexity increases linearly with the number of spatiotemporal patterns of interest.

4.3 Complexity of pattern versus order of coupling

NeuroXidence cannot investigate the nature or the cause of observed JS-patterns. Hence, higher-complexity JS-patterns detected by NeuroXidence can be either due to pairwise coupling or to higher-order interactions of more than pairs. To infer to what degree pairwise interactions are sufficient, Schneidman et al. (2006) introduced an analysis based on the maximum entropy concept and applied it to data recorded from the retina. They concluded that most of the observed patterns are explained by pairwise couplings between the neurons (Schneidman et al. 2006). However, this method assumes for the independence assumption that the recorded spike trains can be described by stationary Bernoulli processes. Both this assumption and the assumption of the independence of firing of a given neuron over time, are likely to be false (Brody 1998; Baker and Gerstein 2001).

The importance of the auto-structure on the joint probability of sets of neurons firing coincidently is highlighted by data presented in this paper. We have shown that the number of individual patterns existing in sets of spike trains can be different by a factor of 10, given the same rate but different regularities of the spiking process (Fig. 6, periods 4 and 5). Therefore, any method based on the assumption of independent Bernoulli processes, such as the maximum entropy method used in Schneidan et al. (2006), is likely to fail since the estimated probability of a JS-pattern does not account for auto-structure such as oscillations, regularity, or burstiness of the individual spiking activity. In addition, the maximum entropy approach requires much more data than NeuroXidence. The study reported in (Schneidman et al. 2006) was based ~1 h of recording to estimate interactions among a set of ten neurons. Given the amount of data that can be recorded for behavioral experiments as well as stability and stationarity issues, it appears to be difficult to obtain sufficient amount of data to allow reliable application of these methods in behavioral paradigms. In contrast, as demonstrated by the analysis of real data in this paper, 20 trials, each with 500 ms sliding windows (in total 10 s), are sufficient to estimate the statistical significance of JS-patterns involving 9 of the 42 recorded cells.

4.4 Induced versus evoked JS-patterns

One might want to distinguish two kinds of JS-patterns: evoked and induced. The first kind consists of JS-patterns

that are evoked by a stimulus and, therefore, are tightly time-locked to it with a certain delay. JS-patterns of the second kind, reflect the internal dynamics of the system and thus, are not time-locked to the stimulus. Neuro-Xidence cannot distinguish between evoked and induced JS-patterns as long as their temporal precision is of the same order. However, since the required precision of JS-patterns, τ_c, is usually on the order of a few milliseconds, this distinction is of a rather theoretical nature. Evoked JS-patterns would require, first, that changes of the stimulus are on the same order as or faster than τ_c, and, second, that the temporal precision is maintained over the whole period from stimulus onset until the occurrence of JS-patterns. Hence, the distinction is of practical use only for a very few stimuli such as flickered or flashed visual stimuli and then only shortly after stimulus onset. In those few cases, trial shuffling, rather than the timescale separation, can be used as an alternative for generating surrogate data to distinguish between induced and evoked JS-patterns. However, since trial shuffling destroys the correspondence of simultaneously recorded trials, it also destroys rate co-variations in the range of the temporal precision of stimulus locking. As discussed by Brody and others (Brody 1999; Grün et al. 2003; Pipa and Grün 2003), this might lead to false positives. Therefore, trial shuffling should be used as an alternative for the concept of timescale separation only if the changes of the stimulus are of the same order as τ_c. If this is not the case, it can be inferred that JS-patterns detected by NeuroXidence are induced rather than evoked.

4.5 Rare and spurious events versus reliably reoccurring events

NeuroXidence differentiates between rare and spurious events and reoccuring events. Only if events are reliably reoccurring across trials might they be classified as JS-patterns with a significant excess or diminution of JSEs. This allows NeuroXidence to estimate the statistical significance reliably even for JS-patterns of complexities higher than 2, which are expected to be rare. In contrast, other methods, which are based only on the total frequency across trials or the frequency in a single trial, might reject H0 if the expected frequency is so low that a few occurrences or even a single JSE is sufficient to be rated as a significant event. This is only likely to be the case for higher complexity JSEs, for which the expectancy is exponentially decreasing. Therefore, methods based on single trials or on pooled data from all trials are likely to fail to differentiate between rare and spurious events, on one side, and events that are reliable and consistent, on the other side.

4.6 Analysis window length

NeuroXidence can be applied to short time windows to track transient neuronal states. In contrast to other methods, the NeuroXidence window length is not restricted by any assumption on the stationarity of data, since the full auto-structure including any rate modulation is considered by the statistical hypothesis test. Thus, the length of the Neuro-Xidence analysis window can be freely chosen by the operator to match the assumed duration of neuronal states of interest.

4.7 Oscillatory processes

Since oscillatory processes are frequently observed in neuronal data, we studied the impact of oscillatory and, across neurons, coherent rate modulation by using simulated data based on inhomogeneous Poisson processes, whose rates were modulated sinusoidally. Based on this simulated data, we estimated test-power numerically and also derived an approximate analytical solution of the test-power (see Appendix 4). Both the analytical approximation and the numerical estimation were in agreement and demonstrate that the test-power increases with increasing frequency of the coherent rate modulations. This behavior is expected since NeuroXidence discriminates between fast processes that are defined on a timescale τ_c, corresponding to the precision of JSEs, and processes that are slower than a timescale τ_r. Hence, if coherent rate modulations occur on a timescale that is comparable with τ_c, NeuroXidence is expected to reject H0 and to support H1. This illustrates that the definition of coordinated firing and fine temporal cross-structure includes explicitly fast and, across neurons, coherent changes of the instantaneous spike rate, which occur on a timescale comparable to τ_c.

4.8 NeuroXidence compared to other methods

In contrast to other methods like the cross-correlation analysis (Gerstein and Perkel 1969; König 1994) and the joint peri-stimulus time histogram (JPSTH; Aertsen et al. 1989), NeuroXidence allows for the analysis of interactions between larger sets of neurons than just pairs. To compare the statistical properties of NeuroXidence with the UE-method and the shuffle-corrected cross-correlogram, we applied all three methods to the same simulated dataset based on 18 simultaneous Poisson or Gamma processes. The simulated data contained epochs with constant rates between 5 and 40 ap/s, rate modulations between 5 and 45 ap/s, latency variability, and trial-by-trial variability of rate-levels. Only NeuroXidence did not exhibit an increased false-positives rate for any of the epochs of the simulated

data, indicating that a consideration of the auto-structure, as well as a trial-based evaluation, is crucial for dissociating the rate and the temporal code hypothesis.

However, if NeuroXidence and the UE-method use analysis windows of equal length, the test-power of the UE-method is higher. This has two reasons. First, the UE-method is a parametric approach and assumes a model, which improves the test-power if the model is correct. The disadvantage is that a model-based approach is only valid as long as the model is describing approximately well the real observed data. Second, NeuroXidence incorporates the trial-by-trial variability of the frequency of occurrence of JSEs. Therefore, it requires that the frequency of JSEs across all trials is reliable and, for most of the trials, is higher in the real than in the surrogate dataset. In contrast, the UE-method considers only the total frequency of JSEs across trials. If the expected frequency is very low, a few occurrences or even one single JSE across all trials might lead to a rejection of H0. This increases the sensitivity, as reflected by the higher test-power, but on the other hand, this also increases the probability of false rejections in the case of rare events.

4.9 Significant joint-spike activity in real data

We applied NeuroXidence to real data recorded in visual area 17 in an anesthetized cat. Our results demonstrate that JSEs, with a precision of 5 ms, occurred significantly more often than expected by chance for complexity 2–8 JS-patterns. We investigated the strength of neuronal synchrony across two stimulation conditions and compared the measures obtained by NeuroXidence (Fig. 11) to the measures obtained by a classical pairwise method, which employs the computation of cross-correlation histograms (Fig. 10; König 1994; Samonds et al. 2003). For pairwise interactions (complexity 2), both measures produce similar results. Neuronal synchrony is strongest in stimulation conditions in which the orientation of the sinusoidal grating matches the preferred orientation of most of the multi-units included in the analysis (condition 30°; orientation preferences are shown in Electronic Supplemental Fig. 6), while synchrony is weakest under the condition of 120° when the neurons are stimulated least optimally. Results obtained with NeuroXidence indicate that interactions of higher complexities show even stronger task- and stimulus-specific changes in the strength of neuronal synchrony than pairwise interactions (Fig. 11). Thus, the comparison of the relative modulation of rate responses between conditions 30° and 120°, reveals that the JS-activity has the higher discriminative power.

We validated our result by analyzing a jittered version of the original data. This validation was equivalent to a comparison of H0 against H0 and served as the estimation of the chance frequency of JS-patterns with excess JSEs. Since only 0.2% of JS-patterns with excess JSEs remained significant after jittering, we conclude that the increase of the frequency of JS-patterns with an excess of JSEs is highly significant for complexities ranging between 2–8. This is, to our knowledge, the first demonstration of JS-activity and its stimulus specificity that considers the complete auto-structure and trial-by-trial variability of spike trains when calculating JSE frequencies in real data.

These results illustrate that NeuroXidence allows for an evaluation of JS-patterns independent of rate modulations, auto-structure, and trial-by-trial variability. This highlights the strength of NeuroXidence for investigating the complex interactions among neurons in terms of fine temporal structure that might coexist with rate modulations of individual neurons.

5 Conclusion

In conclusion, we have presented the new analysis tool NeuroXidence (www.NeuroXidence.com) for detecting temporally coordinated firing events in spike trains. We have demonstrated its performance, reliability, and applicability, and compared it to currently used methods. The results obtained in the analysis of real data provide strong evidence that precise temporal relations between the discharges of distributed neurons contain information about the stimulus configurations and, therefore, support the temporal coding hypothesis.

Appendix 1: Pre-processing

The pre-processing prevents temporally overlapping JSEs, while preserving each individual JSE and its exact frequency of occurrence. Afterwards, the minimal interval between JSEs occurring within any single spike train is larger than the amount of allowed jitter. Pre-processing a dataset requires two steps. First, an array of bins that contain zeros is padded to both ends of the spike trains to prevent border problems at later processing steps (see Fig. 12(a), Appendix 1). Second, NeuroXidence applies a recursive algorithm sequentially to each individual spike train to isolate all JSEs that are included in overlapping JSEs and to represent them as individual and isolated JSEs in the dataset for later processing (see Fig. 12(b), Appendix 1).
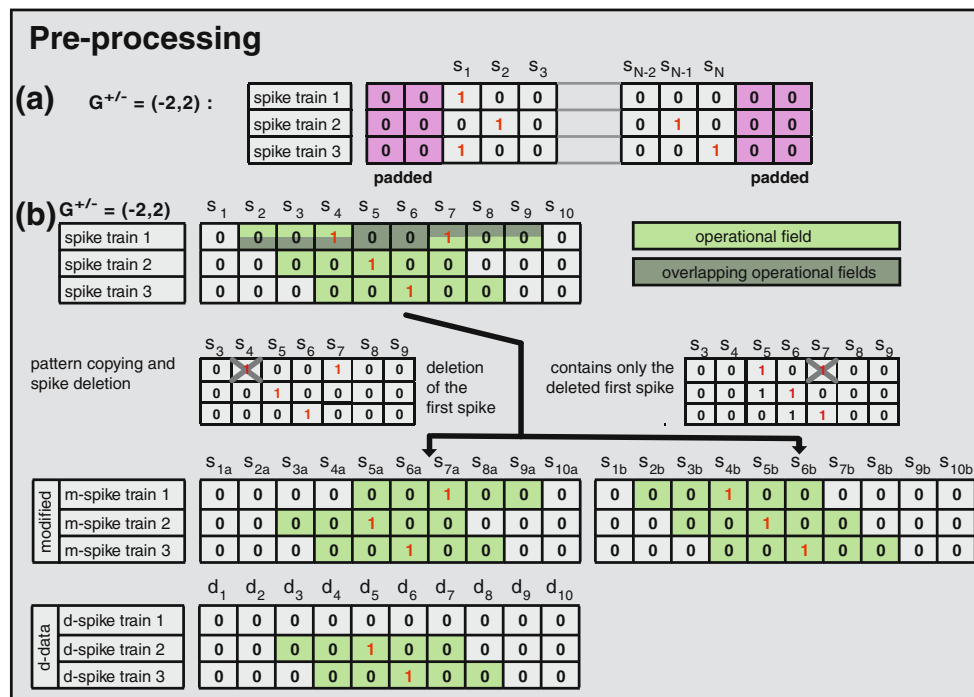
Fig. 12 (**a**, **b**) Preprocessing of spike trains in advance of the detection of joint-spike events (*JSEs*). A spike train is defined as a binary process (1=spike, 0=no spike) with a bin width small enough to resolve an individual spike. (**a**) Padding zeros to the beginning and the end of the spike trains ensures that the replacement of spikes by the kernel G$^{+/-}$ is possible. (**b**) Overlapping JSEs are isolated. Overlapping regions of allowed jitter are identified, which correspond to overlapping operational fields (*dark green*). The algorithm uncovers overlapping patterns in two steps: first, the copying of patterns, and second, the deletion of single spikes. After revealing the overlapping patterns, the 'modified' dataset contains all isolated patterns, while the 'd-data' dataset contains all patterns that had been multiplied. The total frequency of occurrence of a test-pattern in the original data is given by the frequency in the 'modified' dataset, minus the frequency in the 'd-data' set

The resolution of overlapping JSEs might multiply sub-patterns that are included in the resulting isolated JSEs. The recursive algorithm has to correct this multiplication to allow for an exact derivation of the frequency of occurrence of each JSE. The algorithm first identifies, one spike train at a time, all sets of spikes that have an inter-spike interval smaller than G$^{+/-}$, a kernel defining the amount of allowed jitter. Operational fields (see Fig. 12(b), Appendix 1) describe the temporal windows spanning such sets. A field includes the G$^-$ bins preceding the first and the G$^+$ bins following the last spikes of a set. To separate the JSEs localized by the overlapping operational fields (see Fig. 12(b), Appendix 1, *dark green*), a modified copy of the overlapping JSEs is appended to the end of the dataset. Note, that only the data covered by operational fields is replicated and not the entire dataset. The copy is missing the second spike in the overlapping field, leaving an intact and isolated JSE. Likewise, in the original of the over-lapping JSEs, the first spike is deleted, producing the second isolated JSE. Next, the sub-patterns that were doubled by the copying process, i.e. those spikes that did not have overlapping operational fields, are copied into a d-dataset. Since the algorithm is applied sequentially in time, there is no difficulty in separating three or more

spikes that occur in an interval smaller than G$^{+/-}$. For uncovering all overlapping JSEs in the original dataset, as well as in the d-dataset, the recursive algorithm is applied sequentially to all units in the dataset and in the d-dataset. The algorithm first processes unit *n* in the original dataset and then it processes unit *n*+1 in the d-dataset. This ensures that JSEs that are copied from the original dataset into the d-dataset no longer contain any overlapping operational fields for units 1 to *n*. To accurately count the total number of isolated JSEs, one has to consider that JSEs contained in the d-dataset have been doubled. Hence, the number of individual JSEs in the d-dataset has to be subtracted from the number of matching JS-patterns in the original data. In contrast, JS-patterns that contain overlapping operational fields in the d-dataset lead to an overcorrection of the actual frequency in the original dataset. Therefore, the number JSEs that were doubled in the d-dataset should be tracked in the original dataset and added to the total. The computational complexity of resolving overlapping JSEs increases approximately exponentially with the number of spikes that are less than G$^{+/-}$ apart. Thus, a higher computational effort is required if spike trains contain a lot of short intervals, as in the case of bursting cells or Poissonian firing.

Appendix 2: Parametric *t* test versus non-parametric Wilcoxon rank test

In practical applications, the Wilcoxon rank test (Wilcoxon test) is more likely to be appropriate for hypothesis testing than Student's *t* test. The *t* test assumes that the mean value of $\overline{\Delta F}^k$ is normally distributed, while the Wilcoxon test does not make any assumptions about the distribution of the mean. This difference between the two tests is important since JSEs are expected to be rare when spike rates are low and JSE complexities are high (i.e., large numbers of neurons are recorded in parallel). With such rare events, the probability distribution of the total number of JSEs is expected to be better approximated by a Poisson than by a normal distribution. Thus, the distribution of JSEs is likely

to violate the assumptions of the *t* test. Although the *t* test is to some degree robust against skewed distributions (Boneau 1960), the Wilcoxon test may, in many cases, be a much better choice, especially when spike rates are low and complexities of JSEs are greater than two.

Appendix 3: Comparison of the test-power of NeuroXidence and of the UE-method

To compare the test-power of NeuroXidence with that of the UE-method, we applied both methods to two simulated datasets, which were characterized by the same set of parameters. Simulated dataset one contained JSEs that were absolutely synchronous, while simulated dataset two contained JSEs that were deduced from simulated dataset
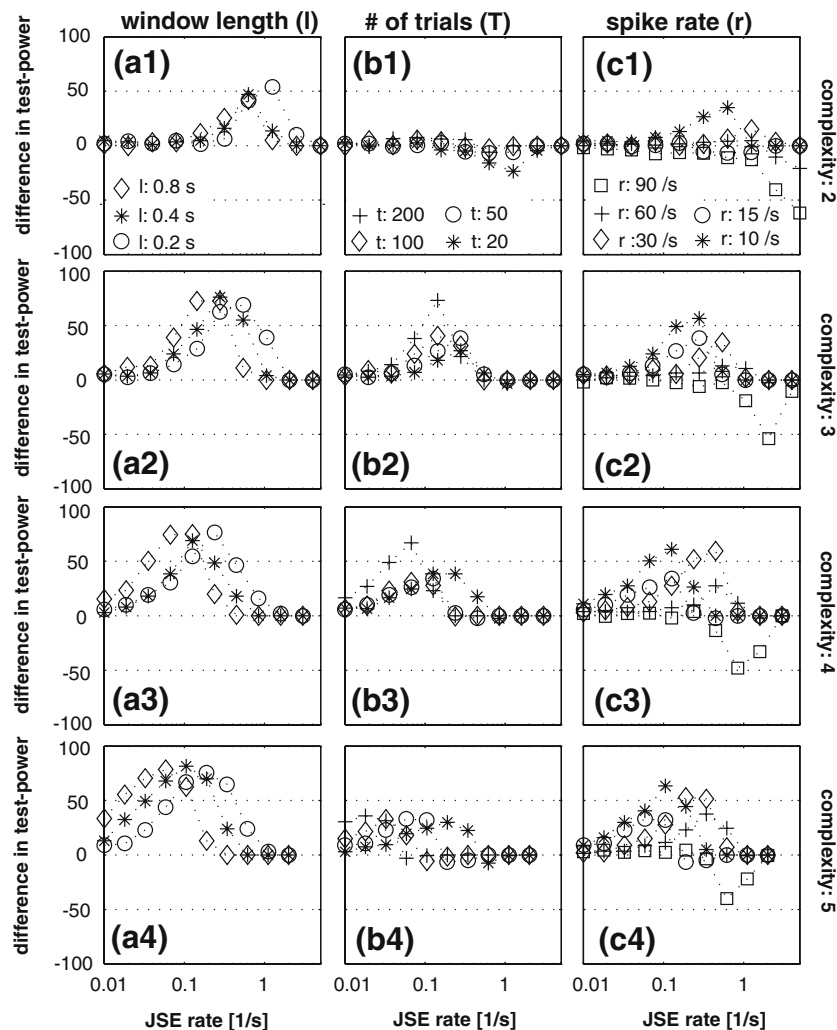
Fig. 13 Difference in the test-power between NeuroXidence and the unitary event method (UE-method). (**a–c**) Variation of length *l* of the analysis window (**a**), the number of trials *T* (**b**), and the spike rate *r* (**c**) from the stand parameter set (*T*=50, *r*=15 ap/s, *S*=20, *η*=3) are presented. Spike trains were modeled as a single-interaction process (*SIP*) based on homogenous Poisson processes. The frequency of excess joint-spike events (*JSEs*) beyond the chance level is given by the x-axis. (**a**) The length of the analysis window used for both NeuroXidence and the UE-method varied between 0.2 s, 0.4 s, and 0.8 s. (**b**, **c**) The NeuroXidence analysis window was 0.8 s while and the UE window was 0.2 s long. Rows (*1–4*) show the difference in test-power for test-patterns of complexity 2–5

one by jittering each individual spike by an amount smaller than the allowed jitter τ_c (Fig. 13, Appendix 3). Positive values indicate a higher test-power for the UE-method. We used absolutely precise JS-patterns to derive the test-power of the UE-method, since it is based on exclusive binning, which limits its capability to detect jittered JS-patterns of complexities higher than 2 (Grün et al. 1999). The combination of parameters, given by 50 trials, a spike rate of the background process of 15 ap/s, 20 surrogates, and $\eta=3$, served as a standard parameter set, from which 11 different simulated datasets were derived by changing either the window length ($l=200, 400, 800$ ms), the number of trials ($T=20, 50, 100, 200$), or the spike rate ($r=10, 15, 30, 60, 90$ ap/s).

First, we studied the difference in the test-power as a function of analysis window length. Initially, we generated simulated datasets where the NeuroXidence analysis window had the same length as that of the UE-method (Fig. 13, Appendix 3 (a, 1–4)). In that case, the test-power of the UE-method was, for all tested parameters, higher than for the NeuroXidence method and the difference increased with increasing JS-pattern complexity. However, since the length of the NeuroXidence analysis window is not constrained by the dynamics in the auto-structure, we increased the window length of the NeuroXidence window up to 0.8 s, while we kept the length of the UE-method analysis window at 0.2 s (Fig. 13(b and c), Appendix 3). The latter was selected to be compatible with the window length used to analyze real data and reflects the constraint that data inside a window have to be approximately described by a stationary Bernoulli process. The longer NeuroXidence window reduced the test-power advantage of the UE-method. For JS-patterns of complexities higher than 2, the difference between the test-power of both methods was substantially reduced and dropped for complexity 2 JS-patterns below 15% (Fig. 13(b, 1), Appendix 3).

Only in the case of high rates was the NeuroXidence test-power higher than the test-power of the UE-method. The reason is that the UE-method assumes that data can be approximately described by a binary Bernoulli process. Using this assumption, the method detects JSEs based on exclusive and binary binning. To ensure a binary process, the UE-method has to utilize clipping, which only counts one spike, no matter how many spikes actually fell in the bin. Thus, the detection scheme of the UE-method changes the data structure and does not detect all existing JS-patterns. Since the simulated data were based on Poisson spike trains, increasing the spike rates led to an increased probability of more than one spike per bin. This led to an increasing number of JS-patterns that went undetected by

the UE-method. It also explains the relative drop in the test-power compared to NeuroXidence, which, in contrast, detects all existing JS-patterns due to the pre-processing (Appendix 1) that uncovers overlapping JSEs.

Appendix 4: Analytical description of the test-power for coherent oscillatory rate changes

We derived an analytical description that allows for the approximation of the theoretically predicted test-power $A'(v)^c$. To this end, we first derived an analytical description of the expected difference, Δf, between the frequency of JSEs in the original, $f(org)$, and the jittered, $f(sur)$, datasets. As a model we assumed an inhomogeneous Poisson process with a rate function, $r_i(t)$, that is coherently modulated across neurons by a sinusoidal oscillation with frequency v:

$$r_i(t) = \frac{1}{2}A[1 + \sin{(v \cdot t)}] \quad \text{with i} = 1 \ldots \text{N}. \tag{7}$$

The analytical description of Δf is based on the description of the impact of jittering. Jittering of spikes is equivalent to a convolution of the instantaneous rate, $r_i(t)$, with a jitter-kernel that is defined by the corresponding probability distribution of the jitter values. We assumed a rectangular jitter-kernel $p_{\text{rect}}(j)$, corresponding to a uniform distribution of random jitter values bounded between 0 and τ_c. The Fourier transformation of $p_{\text{rect}}(j)$ is $P_{\text{rect}}(v')$,

$$p_{rect}(j) = \begin{cases} 1 & \text{for} \quad 0 \leq j \leq \tau_c \\ 0 & \text{else} \end{cases} \Leftrightarrow P_{\text{rect}}(v') = \frac{\sin(\pi \tau_c v')}{\pi \tau_c v'}. \tag{8}$$

Therefore, the convolution of $r_i(t)$ with $p_{\text{rect}}(j)$ in the time domain is equivalent to a modulation of the Fourier spectrum of the instantaneous rate function $R_i(v)$ by $P_{\text{rect}}(v')$ in the frequency domain. In the case that the rate function is continuous, has infinite length, and is a sinusoid with frequency v, the Fourier spectrum of $r_i(t)$ (see Eq. (6) in main text) is a delta-peak at v. Thus, the jittering of spikes maps the instantaneous rate $r_i(t)$ onto $r_i^{jit}(t)$:

$$r_i^{jit}(t) = \frac{1}{2}A'(v)[1 + \sin{(vt)}] \text{ with } A'(v)$$

$$= A\frac{\sin{(\pi \tau_c v)}}{\pi \tau_c v}. \tag{9}$$

Based on the instantaneous rates $r_i(t)$ and $r_i^{jit}(t)$, the instantaneous joint probabilities of a certain kind of JS-

pattern of complexity c is defined by $jp_c(t)$ for the original and $jp_c^{jit}(t)$ for the jittered datasets:

$$jp_c(t) = \prod_{i=1}^{i=c} r_i(t) \quad \text{and} \quad jp_c^{jit}(t) = \prod_{i=1}^{i=c} r_i^{jit}(t) = A'(\nu)^c \prod_{i=1}^{i=c} r_{i(t)}(t). \tag{10}$$

Thus, the expected frequency $\langle f \rangle_c$ of a JSE of complexity c in the time-interval between 0 and l is given by

$$\langle f(org) \rangle_c = \int_0^l jp_c(t) \, dt \text{ and } \langle f(sur) \rangle_c = A'(v)^c \langle f(org) \rangle. \tag{11}$$

Accordingly, the expected difference between the expected frequency of JSEs in the original and the jittered datasets is given by:

$$\langle \Delta f \rangle_c = \langle f(org) \rangle_c - \langle f(sur) \rangle_c = (1 - A'(v)^c) \langle f(org) \rangle_c. \tag{12}$$

Note that the modulation is independent of the duration of the interval (l) used to derive the expected frequency of JSEs.

Appendix 5: Experimental methods

Preparation

Anesthesia was induced with ketamine and maintained with a mixture of 70% N_2O and 30% O_2, supplemented with halothane (0.4–0.6%). The animal was paralyzed with pancuronium bromide (Pancuronium, Organon, 0.15 mg kg^{-1} h^{-1}). All the experiments were conducted according to the guidelines of the American Physiological Society and German law for the protection of animals, approved by the local government's ethical committee and overseen by a veterinarian.

Recordings and visual stimulation

Multi-unit activity was recorded from a region of area 17 corresponding to the central part of the visual field. Single-unit activity was extracted by offline sorting. For recording we used two SI-based multi-electrode probes (16-channels per electrode) supplied by the Centre for Neural Communication Technology at the University of Michigan (Michigan probes) with an inter-contact distance of 200 μm (0.3–0.5 MΩ impedance at 1,000 Hz). The probes were inserted in the cortex approximately perpendicular to the surface to record from neurons at different cortical depths and along an axis tangential to the cortical surface.

The software used for visual stimulation was Active-STIM (www.ActiveSTIM.com). The stimulus consisted of a drifting sinusoidal grating, spanning 15° of visual angle (spatial frequency, 3°/cycle; temporal frequency of the drift, 3.6°/s), which was sufficient to cover the receptive fields of all the recorded cells simultaneously and to stimulate also the regions surrounding the receptive fields.

Appendix 6: Standard parameter set for NeuroXidence

The standard choice for analysis parameters is τ_c=5 ms and τ_r=20 ms, equivalent to η=4. The duration of the analysis window should be around 100 ms or longer. The number of surrogates should be S=20. Only if NeuroXidence is used to identify a deficiency of JSEs S should be 1. To allow for a reasonable test-power, not less than 20 trials should be used. The bin length corresponding to the timing resolution in the spike train should be 1 ms. The statistical evaluation should be based on the Wilcoxon test. See also (www.NeuroXidence.com)

References

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

Aertsen, A., Diesmann, M., Gewaltig, M. O., Grün, S., & Rotter, S. (2001). Neural dynamics in cortical networks-precision of joint-spiking events. *Novartis Foundation Symposium, 239,* 193–204.

Aertsen, A. M., Gerstein, G. L., Habib, M. K., & Palm, G. (1989). Dynamics of neuronal firing correlation: Modulation of "effective connectivity". *Journal of Neurophysiology, 61,* 900–917.

Baker, S. N., & Gerstein, G. L. (2001). Determination of response latency and its application to normalization of cross-correlation measures. *Neural Computation, 13,* 1351–1377.

Baker, S. N., & Lemon, R. N. (2000). Precise spatiotemporal repeating patterns in monkey primary and supplementary motor areas occur at chance levels. *Journal of Neurophysiology, 84,* 1770–1780.

Barbieri, R., Frank, L. M., Nguyen, D. P., Quirk, M. C., Solo, V., Wilson, M. A., et al. (2004). Dynamic analyses of information encoding in neural ensembles. *Neural Computation, 16,* 277–307.

Barbieri, R., Wilson, M. A., Frank, L. M., & Brown, E. N. (2005). An analysis of hippocampal spatio-temporal representations using a Bayesian algorithm for neural spike train decoding. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 13,* 131–136.

Bi, G. Q., & Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of Neuroscience, 18,* 10464–10472.

Boneau, C. A. (1960). The effects of violations of assumptions underlying the *t*-test. *Pyschological Bulletin, 57,* 49–64.

Brody, C. D. (1998). Slow covariations in neuronal resting potentials can lead to artefactually fast cross-correlations in their spike trains. *Journal of Neurophysiology, 80,* 3345–3351.

Brody, C. D. (1999). Correlations without synchrony. *Neural Computation, 11,* 1537–1551.

Brown, E. N., Frank, L. M., Tang, D., Quirk, M. C., & Wilson, M. A. (1998). A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *Journal of Neuroscience, 18,* 7411–7425.

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

Czanner, G., Grün, S., & Iyengar, S. (2005). Theory of the snowflake plot and its relations to higher-order analysis methods. *Neural Computation, 17,* 1456–1479.

Fetz, E. E. (1997). Temporal coding in neural populations? *Science, 278,* 1901–1902.

Fries, P., Neuenschwander, S., Engel, A. K., Goebel, R., & Singer, W. (2001a). Rapid feature selective neuronal synchronization through correlated latency shifting. *Nature Neuroscience, 4,* 194–200.

Fries, P., Roelfsema, P. R., Engel, A. K., König, P., & Singer, W. (1997). Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proceedings of the National Academy of Sciences of the United States of America, 94,* 12699–12704.

Fries, P., Schröder, J. H., Roelfsema, P. R., Singer, W., & Engel, A. K. (2002). Oscillatory neuronal synchronization in primary visual cortex as a correlate of stimulus selection. *Journal of Neuroscience, 22,* 3739–3754.

Fries, P., Schröder, J. H., Singer, W., & Engel, A. K. (2001b). Conditions of perceptual selection and suppression during interocular rivalry in strabismic and normal cats. *Vision Research, 41,* 771–783.

Gerstein, G. L., Bedenbaugh, P., & Aertsen, M. H. (1989). Neuronal assemblies. *IEEE Transactions on Biomedical Engineering, 36,* 4–14.

Gerstein, G. L., & Perkel, D. H. (1969). Simultaneously recorded trains of action potentials: Analysis and functional interpretation. *Science, 164,* 828–830.

Gerstein, G. L., & Perkel, D. H. (1972). Mutual temporal relationships among neuronal spike trains. Statistical techniques for display and analysis. *Biophysical Journal, 12,* 453–473.

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

Grün, S., Diesmann, M., & Aertsen, A. (2002b). Unitary events in multiple single-neuron spiking activity: II. Nonstationary data. *Neural Computation, 14,* 81–119.

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.

Grün, S., Riehle, A., & Diesmann, M. (2003). Effect of cross-trial nonstationarity on joint-spike events. *Biological Cybernetics, 88,* 335–351.

Hebb, D. O. (1949). *Organization of behavior. A neurophysiological theory.* New York: Wiley.

Hopfield, J. J., & Brody, C. D. (2000). What is a moment? "Cortical" sensory integration over a brief interval. *Proceedings of the National Academy of Sciences of the United States of America, 97,* 13919–13924.

Hopfield, J. J., & Brody, C. D. (2001). What is a moment? Transient synchrony as a collective mechanism for spatiotemporal integra-tion. *Proceedings of the National Academy of Sciences of the United States of America, 98,* 1282–1287.

Ikegaya, Y., Aaron, G., Cossart, R., Aronov, D., Lampl, I., Ferster, D., et al. (2004). Synfire chains and cortical songs: Temporal modules of cortical activity. *Science, 304,* 559–564.

Kass, R. E., Ventura, V., & Brown, E. N. (2005). Statistical issues in the analysis of neuronal data. *Journal of Neurophysiology, 94,* 8–25.

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

Kuhn, A., Aertsen, A., & Rotter, S. (2003). Higher-order statistics of input ensembles and the response of simple model neurons. *Neural Computation, 15,* 67–101.

Lachaux, J. P., Rodriguez, E., Martinerie, J., Adam, C., Hasboun, D., & Varela, F. J. (2000). A quantitative study of gamma-band activity in human intracranial recordings triggered by visual stimuli. *European Journal of Neuroscience, 12,* 2608–2622.

Markram, H., Lubke, J., Frotscher, M., & Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science, 275,* 213–215.

Martignon, L., Deco, G., Laskey, K., Diamond, M., Freiwald, W., & Vaadia, E. (2000). Neural coding: Higher-order temporal patterns in the neurostatistics of cell assemblies. *Neural Computation, 12,* 2621–2653.

Nakahara, H., & Amari, S. (2002). Information-geometric measure for neural spikes. *Neural Computation, 14,* 2269–2316.

Okatan, M., Wilson, M. A., & Brown, E. N. (2005). Analyzing functional connectivity using a network likelihood model of ensemble neural spiking activity. *Neural Computation, 17,* 1927–1961.

Oram, M. W., & Perrett, D. I. (1992). Time course of neural responses discriminating different views of the face and head. *Journal of Neurophysiology, 68,* 70–84.

Oram, M. W., Wiener, M. C., Lestienne, R., & Richmond, B. J. (1999). Stochastic nature of precisely timed spike patterns in visual system neuronal responses. *Journal of Neurophysiology, 81,* 3021–3033.

Oweiss, K., Jin, R., & Suhail, R. (2007). Identifying neuronal assemblies with local and global connectivity with scale space spectral clustering. *Neurocomputing, 70,* 1728–1734.

Pipa, G. (2003). Patent application DE10008251A1: Verfahren zum Erfassen und Auswerten von Korrelationen. In 2003.

Pipa, G., & Grün, S. (2003). Non-parametric signicance estimation of joint-spike events by shuffling and resampling. *Neurocomputing, 52–54,* 31–37.

Pipa, G., Riehle, A., & Grün, S. (2006). Validation of task-related excess of spike coincidences based on NeuroXidence. *Neurocomputing*, doi:10.1016/j.neucom.2006.10.142.

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

Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B., & Varela, F. J. (1999). Perception's shadow: Long-distance synchronization of human brain activity. *Nature, 397,* 430–433.

Roy, A., Steinmetz, P. N., & Niebur, E. (2000). Rate limitations of unitary event analysis. *Neural Computation, 12,* 2063–2082.

Samonds, J. M., Allison, J. D., Brown, H. A., & Bonds, A. B. (2003). Cooperation between area 17 neuron pairs enhances fine discrimination of orientation. *Journal of Neuroscience, 23,* 2416–2425.

Samonds, J. M., & Bonds, A. B. (2004). Real-time visualization of neural synchrony for identifying coordinated cell assemblies. *Journal of Neuroscience Methods, 139,* 51–60.

Schneider, G., Havenith, M. N., & Nikolić, D. (2006). Spatiotemporal structure in large neuronal networks detected from cross-correlation. *Neural Computation, 18,* 2387–2413.

Schneidman, E., Berry, M. J., Segev, R., & Bialek W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature, 440,* 1007–1012.

Schneidman, E., Bialek, W., & Berry, M. J. (2003). Synergy, redundancy, and independence in population codes. Testing non-linearity and directedness of interactions between neural groups in the macaque inferotemporal cortex. *Journal of Neuroscience, 23,* 11539–11553.

Schreiber, T., & Schmitz, A. (2000). Surrogate time series. *Physica D-Nonlinear Phenomena, 142,* 346–382.

Shadlen, M. N., & Newsome, W. T. (1994). Noise, neural codes and cortical organization. *Current Opinion in Neurobiology, 4,* 569–579.

Sharpee, T., Rust, N. C., & Bialek, W. (2004). Analyzing neural responses to natural signals: Maximally informative dimensions. *Neural Computation, 16,* 223–250.

Singer, W. (1999). Neuronal synchrony: A versatile code for the definition of relations? *Neuron, 24,* 25–49.

Sjostrom, P. J., Turrigiano, G. G., & Nelson, S. B. (2001). Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron, 32,* 1149–1164.

Tallon-Baudry, C., Kreiter, A., & Bertrand, O. (1999). Sustained and transient oscillatory responses in the gamma and beta bands in a visual short-term memory task in humans. *Vision Neuroscience, 16,* 449–459.

Tallon-Baudry, C., Mandon, S., Freiwald, W. A., & Kreiter, A. K. (2004). Oscillatory synchrony in the monkey temporal lobe correlates with performance in a visual short-term memory task. *Cerebral Cortex, 14,* 713–720.

Tetko, I. V., & Villa, A. E. (2001). A pattern grouping algorithm for analysis of spatiotemporal patterns in neuronal spike trains. 1. Detection of repeated patterns. *Journal of Neuroscience Methods, 105,* 1–14.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381,* 520–522.

Vaadia, E., & Abeles, M. (1987). Temporal firing patterns of single units, pairs and triplets of units in the auditory cortex. *Israel Journal of Medical Sciences, 23,* 75–83.

von der Malsburg, C. (1981). The correlation theory of brain function (Int. Rep. No. 81-2). Report.