J Comput Neurosci (2006) 21:329–342
DOI 10.1007/s10827-006-8899-7

Robustness of the significance of spike synchrony with respect to sorting errors

Antonio Pazienti · Sonja Grün

Received: 21 November 2005 / Revised: 4 March 2006 / Accepted: 21 April 2006 / Published online: 14 August 2006

Abstract The aim of spike sorting is to reconstruct single unit spike times from extracellular multi-unit recordings. Failure in the identification of a spike (false negative) or assignment of a spike to a wrong unit (false positive) are typical examples of sorting errors. Their influence on cross-correlation measures has been addressed and it has been shown that correlation analysis of multi-unit signals may lead to incorrect interpretations. We formulate a model to study the influence of sorting errors on the significance of synchronized spikes, and thus are able to study if and how the significance changes in case of imperfect sorting. Here we explore the case of pairwise analysis of simultaneously recorded neurons. Interestingly, a decrease in the significance is observed in the presence of false positives, as well as for false negatives. Furthermore, false negative errors reduce the significance of synchronized spikes more strongly than false positives. Thus, conservative sorting strategies have a stronger tendency to lead to a loss of the significance of synchronization. We demonstrate that a detailed understanding of sorting techniques and their possible effects on subsequent data analyses is important in order to rule out inconsistencies in the interpretation of results.

Keywords Spike sorting · Synchronization · Multiple single neurons · Statistical analysis · Unitary events

Action Editor: John P. Miller

A. Pazienti (✉) · S. Grün
Neuroinformatics, Institute for Biology - Neurobiology,
Free University, Königin-Luise Str. 28/30, 14195 Berlin, Germany
e-mail: antonio.pazienti@neurobiologie.fu-berlin.de

A. Pazienti · S. Grün
Bernstein Center for Computational Neuroscience,
Berlin, Germany

1. Introduction

In the field of electrophysiology, spike sorting is the procedure of extracting single unit activity from a recorded extracellular "multi-unit" signal. Multiple single-unit recordings offer the chance to detect assembly activities, and to identify the network composition and functions. The spike sorting procedure typically consists of two stages: first, spike waveforms originating from different neurons are identified and separated from the background "extra-cellular noise", which is presumably composed of weaker neural signals and measurement noise (Lewicki, 1998); second, spikes are then classified, i.e. assigned, to putative neurons. Spike sorting failures may occur in any of these two stages.

A variety of techniques and algorithms are meanwhile available for spike sorting (see e.g. Lewicki (1998) for the most recent review). The choice of the best algorithm for this problem depends on a number of factors, like e.g. the type of electrode, the brain area, the cell type of interest, etc. Additional constraints for sorting methods are reasonable computation time, available hard-disk space, compatibility with the recording setup and software, and finally also the intended subsequent data analysis. For offline sorting, standard approaches to the classification problem are cluster cutting in a feature space (see e.g. Harris et al., 2000). Unsupervised (at least partially) statistical algorithms are available, such as independent component analysis (ICA, Hyvärinen and Oja, 2000) for multi-trode (stereotrodes, tetrodes, etc.) recordings, and superparamagnetic clustering (Quiroga et al., 2004) for independent multi-electrode recordings. For resolving more neurons than available electrodes in the multitrode, Takahashi et al. (2003) suggested to combine ICA and *k*-means clustering. However, also less sophisticated techniques, but with the advantage of online applicability, like threshold crossing, window discrimination, or multiple-point template-matching procedures are in use. Major problems in spike sorting are the difficulty to resolve spikes from different neurons which overlap in time, and the variability of the spike waveform (Fee et al., 1996; Quirk and Wilson, 1999; Harris et al., 2000). A solution for spike sorting for the latter problem was suggested by Pouzat et al. (2004), solutions for the problem of overlapping spikes had been suggested by Pouzat et al. (2004) and Zhang et al. (2004).

As this wide interest in finding the "ultimate" algorithm suggests, there is not a unique solution and all existing ones are subject to errors. Objective controls for sorting quality have been proposed by Pouzat et al. (2002) and recently by Schmitzer-Torbert et al. (2005), with the aim to allow for comparability of sorting results from the different methods.

Sorting errors appear either as failures in the identification of spikes, or as assignments of spikes to wrong units, referred to as false negative errors (FN) or false positive errors (FP), respectively. Taking the perspective of a particular neuron, a false positive spike is a spike which is assigned to that neuron despite having originated from another neuron or from extra-cellular noise. Conversely, a false negative spike was emitted by the neuron and successively unclassified or assigned to another neuron. A few studies have quantitatively shown the amount of errors introduced by sorting. Wehr et al. (1999) and Harris et al. (2000) made use of simultaneous intracellular and extracellular recordings in vivo (from adult locust and rat hippocampus, respectively). They reported average error rates of 6.2% for FPs and 15.9% for FNs (Harris et al., 2000), and 3.5% for FPs and 2.8% for FNs (Wehr et al. 1999). Wood et al. (2004) estimated average error rates of 23% FP and 30% FN based on simulated data.

However, studies on how such errors affect subsequent analyses of these data are surprisingly lacking. Bedenbaugh and Gerstein (1997) and Gerstein (2000) investigated the consequences of unresolved spike trains on cross-correlation analysis, and found that the correlation coefficients calculated between spike trains that contain wrongly assigned spikes can be strongly biased, depending on the degree of mixing spike trains and also depending on correlation structures between local and/or remote groups of neurons. In contrast, Bar-Gad et al. (2001) concentrated on the influence of correlated false negative spikes on the cross-correlation analysis. They also found that the cross-correlation function can be heavily biased due to the errors. Quirk and Wilson (1999) showed for the case of neuronal bursting activity that spikes occurring later in the burst may be assigned to another neuron due a change in its spike shapes. The cross-correlation analysis between such neurons also revealed a strong bias by indicating artificial delayed coincidences.

Here we present a study that combines the occurrence of FN and FP sorting errors and that evaluates their influence on unitary events (UE; Grün et al., 1999, 2002a, b) here for the case of pairwise analysis only. UE analysis detects the presence of conspicuous spike coincidences in multiple parallel spike recordings and evaluates their statistical significance. The UE method enabled to study the relation between spike synchronization and behavioral events (Riehle et al., 1997, 2000; Grün et al., 2002b).

A brief introduction to the method is provided in Section 2.1. In Section 2.2 we introduce a simple statistical model for spike sorting errors, which allows us to study how FP and FN errors influence the significance of joint-spike events, as well as the measures entering the significance estimation (Section 3.1). On the basis of the derived analytical expressions and numerical simulations, we demonstrate that the firing rates and the number of coincidences (empirical as well as expected) may be increased or decreased depending on the error rate combination, but the resulting significance is always reduced irrespective of the error type. A rigorous analysis reveals that the origin of significance reduction is different for the two error types. In Section 3.2 we illustrate that variation of physiological parameters, such as firing rates of the neurons and their coincidence rate, influence the resulting significance only weakly as compared to the error rates. Finally we discuss our results and further steps (Section 4).

2. Methods

2.1. Unitary event analysis

Unitary event analysis, introduced by Grün et al. (2002a), measures the significance of joint-spike events occurring amongst multiple, simultaneously recorded neurons. When their synchronization exceeds the chance level by a significant amount, the coincident patterns are called unitary events. UE analysis evaluates the significance based on the null-hypothesis of independent firing. In the case of Poissonian spike trains, the probability distribution of the number of coincident patterns can be analytically derived. For the case of spike trains deviating from Poisson or non-stationary data we suggested methods for correction (Grün et al., 2003; Pipa and Grün, 2003; Pipa et al., 2003). Here we restrict ourselves to the assumptions of stationary and Poissonian spike trains, i.e. the situation described in Grün et al. (2002a), which we briefly introduce below.

Let us consider a stationary process (of rate λ) in the observed time window T containing $N = T/\delta$ bins (with δ the bin width in seconds). The probability of neuron i to fire within the time interval of bin size δ is

$$p_i = \lambda_i \delta, \tag{1}$$

here for neurons $i = 1, 2$. Under the null-hypothesis of statistically independent firing of the neurons, we derive the

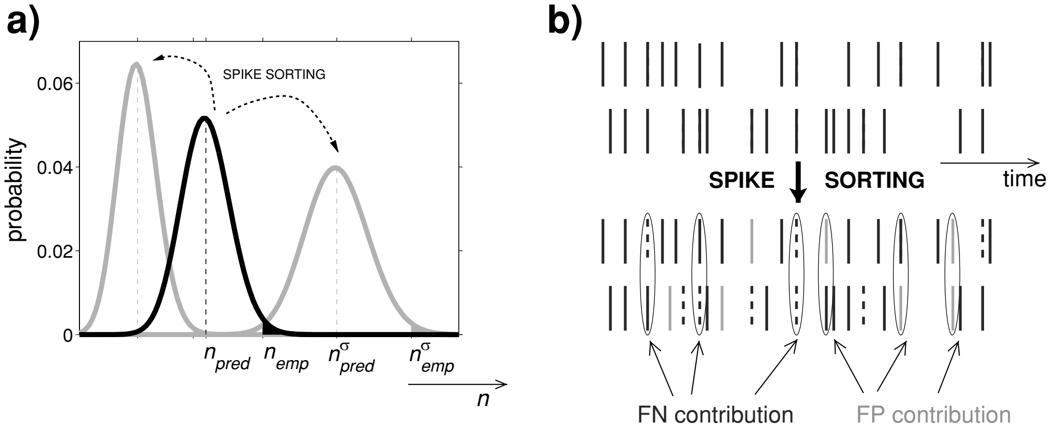

Fig. 1 (a) Probability density function of number of coincident events before and after spike sorting. The distribution in the middle illustrates the density function for the original case before sorting. Its mean n_{pred} is the expected number of coincidences based on the firing rates of the original spike trains. The empirical number of coincidences n_{emp} is shown to be higher than expected, i.e. on the tail of the distribution. The black area represents the *p*-value for significance estimation. The two gray distributions to the left and to the right, show examples of coincidence distributions, whose mean n_{pred} is derived from the firing rates after sorting. The distribution on the left mimicks the case of more FN errors than FP ($\sigma^+ < \sigma^-$), and the one on the right the case for more FP than FN ($\sigma^+ > \sigma^-$). The number of coincidences after sorting (n^{σ}_{emp}) does not necessarily generate the same *p*-value (gray area) as n_{emp}. (b) Sketch of the spike sorting model. (Top) Original spiking activity of two neurons before sorting. Each time a spike occurred is marked by a black line. (Bottom) The spike trains after having experienced sorting errors. Black solid lines show the original and correctly classified spikes. Black dashed spikes are missed (FNs), whereas gray spikes are falsely assigned (FPs). Marked coincidences indicate all possible coincidence error types (FN and FP coincidences) entering the number of coincidences after sorting (n^{σ}_{emp})

probability of joint-firing in one bin as $P = p_1 \cdot p_2$. The probability distribution of the coincident events can be approximated by a Poisson distribution (for a derivation see Grün et al., 2002a):

$$\psi(n; p_i; N) = \frac{(n_{\text{pred}})^n}{n!} \cdot e^{n_{\text{pred}}}, \tag{2}$$

with n_{pred} being the expected number of coincidences, given the firing rates:

$$n_{\text{pred}} = p_1 p_2 N = \lambda_1 \lambda_2 \delta^2 N. \tag{3}$$

The empirically found number of coincidences n_{emp} is then compared to the predicted value n_{pred} using the Poisson distribution (Eq. (2)). Significant deviation from the expected value is estimated by the joint-*p*-value, i.e. the cumulative probability of having n_{emp} or even more coincident events (gray area from n_{emp} to ∞ under the probability density curve in Fig. 1(a), middle distribution). For better visibility of the relevant significance values we express the significance as a non-linear (log-)transformation of the joint-*p*-value Ψ, resulting in the significance measure 'joint-surprise' *js*:

$$js(\Psi) = \log \frac{1 - \Psi}{\Psi}, \tag{4}$$

with

$$\Psi(n_{\text{emp}} | n_{\text{pred}}) = \sum_{n=n_{\text{emp}}}^{\infty} \frac{(n_{\text{pred}})^n}{n!} e^{-n_{\text{pred}}}. \tag{5}$$

When the value of *js* exceeds an a priori threshold, e.g. 1% or 5%, the synchrony is classified as significant.

2.2. Statistical model of spike sorting errors

We shall now formulate a statistical model on how spike sorting errors may affect neuronal spike trains. Given simultaneous spike trains of two neurons (Fig. 1(b), top), we assume a uniformly distributed probability for spikes to be missed as false negatives, or that spikes may be added with uniform probability as false positives to the spike trains (for illustration see Fig. 1(b), bottom). Both error types are applied independently to each neuron. Effectively we assume the neurons to be recorded from different electrodes, such that neuronal activity of one neuron may only be registered at one electrode, and that errors do not interact across electrodes. Thus, for a single neuron the different errors are assumed to be applied independently, however excluding the unrealistic case that an introduced FP spike is considered as FN.

Errors are expressed as fractions σ_i^+ for FPs and σ_i^- for FNs of the original firing rates of neuron $i = 1, 2$. As a result, the firing rate of neuron *i* after sorting ($\lambda_i{}^{\sigma}$) is the sum of three terms, i.e. the original rate (λ_i), the FP rate and the FN rate:

$$\lambda_i \longrightarrow \lambda_i^{\sigma} = \lambda_i + \lambda_i \cdot \sigma_i^+ - \lambda_i \cdot \sigma_i^- = \lambda_i \cdot (1 + \sigma_i^+ - \sigma_i^-). \tag{6}$$

Thus sorting errors alter the original firing rates λ_i by contaminating the spike trains. False negatives reduce the rate, whereas false positives enhance the rate, and they may compensate for equal error rates. Thus, the resulting firing rate after sorting may be larger or smaller than the original rate depending on the relative contributions of errors.

In the following, the index σ will indicate a variable after sorting. For simplifying reasons we define the variable σ_i for the combined influence of errors as:

$$\sigma_i = \sigma_i^- - \sigma_i^+. \tag{7}$$

Inserting Eq. (6) and using Eq. (7) in Eq. (1), we similarly get new values for the firing probability per time bin:

$$p_i \longrightarrow p_i^\sigma = p_i \cdot (1 + \sigma_i^+ - \sigma_i^-) = p_i \cdot (1 - \sigma_i). \tag{8}$$

2.3. Simulations

For understanding the influence of spike sorting errors on significance estimation by the unitary events method, we make use of two approaches: analytical derivation and numerical simulations. For the simulations we follow the approach introduced in Grün et al. (1999). Neuronal spiking activities are realized as a combination of independent background activity and correlated spiking activities. The background activity is generated as realizations of stationary Poisson processes with rate λ'_i for each neuron $i = 1{,}2$. Correlations between the neurons are introduced by inserting coincident events of rate λ_c simultaneously to both processes. Then the total rate of the neurons reads

$$\lambda_i = \lambda_c + \lambda'_i, \qquad i = 1, 2. \tag{9}$$

As a result, each of the simultaneous spike trains contain intermixed independent background spikes and correlated spikes. The consecutively applied sorting errors affect spikes irrespective of their origin.

3. Results

This section presents results for the expected and empirical number of coincidences after sorting as functions of the unperturbed ("original") values and of the error rates (Sections 3.1.1 and 3.1.2) and under variation of physiological parameters, i.e. firing rates of the neurons and degree of correlation (Sections 3.2.1 and 3.2.2). These results find entry in the evaluation of the significance of joint-spike events. The influence of sorting onto the latter is explored in Sections 3.1.3 and 3.2.3.

Analytical descriptions serve to derive the expected and empirical number of coincidences after sorting given the original values and to evaluate their effects on the significance. In addition, numerical simulations of simultaneous spike trains of controllable firing rates and degree of correlation allow to study the influence of physiological parameters on the various measures under sorting errors.

3.1. Dependence on error rates

3.1.1. Number of expected coincidences

In the following we derive the relation between the expected number of coincidences before (n_{pred}, original value) and after sorting (n^σ_{pred}). Their relation results directly from the sorting model introduced in Section 2.2. By substituting in the expression for the original expected number of coincidences (Eq. (3)) the expression for the firing probabilities after sorting (Eq. (8)), we obtain for the expected number of coincidences after sorting:

$$\begin{aligned} n_{\text{pred}} \longrightarrow n^\sigma_{\text{pred}} &= (1-\sigma_1)p_1 \cdot (1-\sigma_2)p_2 \cdot N \\ &= n_{\text{pred}} \cdot (1 + \sigma_1^+ - \sigma_1^-) \cdot (1 + \sigma_2^+ - \sigma_2^-). \end{aligned} \tag{10}$$

The expected number of coincidences after sorting n^σ_{pred} is proportional to the expected number of coincidences before sorting, multiplied by scaling factors that contains the error rates σ_i^+ and σ_i^-. In Fig. 2 (top) it can be seen that the predicted number of coincidences after sorting n^σ_{pred} is an increasing function of σ_1^+ and σ_2^+ and a decreasing function of σ_1^- and σ_2^-. Thus, depending on their relative amounts, n^σ_{pred} may be larger (for $\sigma_i^+ > \sigma_i^-$, $i = 1, 2$) or smaller than n_{pred} (for $\sigma_i^+ < \sigma_i^-$, $i = 1, 2$) as a direct consequence of the sorting errors onto the firing rates (Eq. (6)). This in turn leads to a different position of the probability distribution used for significance estimation (cf. Fig. 1(a)).

3.1.2. Number of empirical coincidences

The empirical number of coincidences after sorting is not a mere function of altered firing rates as the expected number of coincidences, but is also a function of changes in the number of joint-events across neurons. New coincidences may be created and/or deleted by sorting:

$$n_{\text{emp}} \longrightarrow n^\sigma_{\text{emp}} = n_{\text{emp}} - n^{\text{FN}} + n^{\text{FP}} \quad . \tag{11}$$

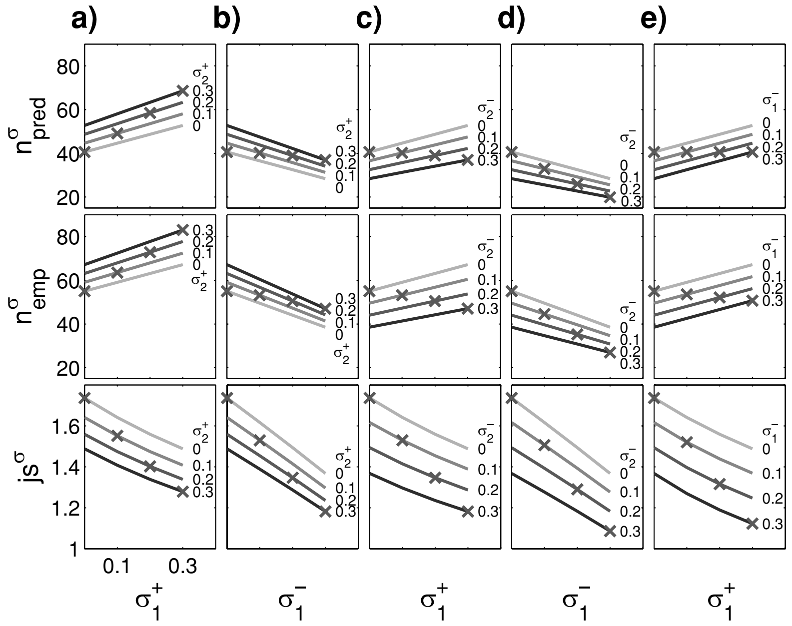

Fig. 2 Number of coincidences and resulting significance as a function of various combinations of error rates. Number of expected (top row) and empirical (middle row) coincidences after sorting and joint-surprise after sorting (bottom row), as a function of errors applied to neuron 1 (either FPs (σ_1^+) or FNs (σ_1^-)) in combination with different error types and rates applied to neuron 2 (a–d, either FPs (σ_2^+) or FNs (σ_2^-)). (e) neuron 1 experiences the two different types of errors, while neuron 2 is perfectly sorted (no errors). For each curve the second error rate is constant. Curves for n^{σ}_{pred} (top), n^{σ}_{emp} (middle) and js^{σ} (bottom) with the same gray level correspond to the same parameters. In each of the cases, only two error types (out of four possible) are applied, the other two are set to zero. Crosses show the symmetrical cases where the two applied errors are of equal absolute amount. Additional parameters are the same for all cases and set to $\lambda_1' = \lambda_2' = 20\,\text{Hz}$, $\lambda_c = 2\,\text{Hz}$, $T = 100$ seconds, bin size $\delta = 1\,\text{ms}$

As illustrated in Fig. 1(b) (bottom) the number of coincidences deleted by false negative spikes (n^{FN}) may be composed by three contributions: spikes originally involved in coincidences may be deleted from neuron 1, or from neuron 2, or simultaneously from both neurons. This is expressed as:

$$n^{\text{FN}} = n_{\text{emp}} \cdot [\sigma_1^-(1-\sigma_2^-) + \sigma_2^-(1-\sigma_1^-) + \sigma_1^- \cdot \sigma_2^-]. \tag{12}$$

On the other hand, wrongly assigned spikes may lead to false positive coincidences (n^{FP}). By inserting spikes either in one or in the other neuron at the very same time when the corresponding other neuron contributes with an original spike, or inserting two false positive spikes in the two neurons coincidently in time, new coincidences are created:

$$n^{\text{FP}} = n_{\text{pred}} \cdot [\sigma_1^+(1-\sigma_2^-) + \sigma_2^+(1-\sigma_1^-) + \sigma_1^+ \cdot \sigma_2^+], \tag{13}$$

(see appendix for a formal derivation of Eqs. (12) and (13)). Inserting these expressions in Eq. (11) and rearranging the terms, leads to the following relation for the empirical number of coincidences after sorting:

$$\begin{aligned} n^{\sigma}_{\text{emp}} &= (1-\sigma_1^-)(1-\sigma_2^-)\cdot[n_{\text{emp}} - n_{\text{pred}}] \\ &\quad + (1+\sigma_1^+ - \sigma_1^-)(1+\sigma_2^+ - \sigma_2^-)\cdot n_{\text{pred}}. \end{aligned} \tag{14}$$

As a result, the empirical number of coincidences n^{σ}_{emp} after sorting can be expressed as a function of the original expected and empirical number of coincidences, multiplied by factors containing the error rates. The first term consists of the difference of the empirical and the expected number of coincidences, i.e. the original "excess" coincidences, multiplied by a factor consisting of false negative errors only. The more false negatives occur, the more the factor is deviating from 1 to smaller values in a nonlinear fashion. Thus, excess coincidences may only be reduced due to sorting errors. The second term is actually the expected number after sorting (cf. Eq. (10)). Interestingly, only here false positive errors enter the expression for the empirical number of coincidences after sorting. As discussed in Section 3.1.1, the errors may –depending on the relation of the errors rates– compensate to 0, or may lead to a factor larger or smaller than 1.

Two important results may be extracted from this discussion of Eq. (14) and are depicted in Fig. 2 (middle row):

- False positive spikes contribute to chance coincidences only and increase them.
- False negatives are the only error type that affects excess coincidences and may only lead to a reduction of their number.

3.1.3. Significance

In the previous two sections we learned that the empirical as well as the expected number of coincidences are affected by both types of sorting errors. These two measures enter the significance estimation of the empirically found number of coincidences given the number of coincidences expected by chance (Eq. (4)), i.e. we obtain now a value js^{σ}, as a function of n^{σ}_{emp} and n^{σ}_{pred} :

$$js^{\sigma} = js\left(n^{\sigma}_{\text{emp}}, n^{\sigma}_{\text{pred}}\right), \tag{15}$$

where n^{σ}_{pred} defines the mean of the distribution that enters the significance measure (Eq. (2)). In addition, this number also affects the shape of the distribution: the larger the mean, the wider the distribution (a feature of the Poisson distribution, Fig. 1(a)), and thus the larger the empirical number of coincidences required for significance. Since the errors enter both the empirical and the expected number, it is not obvious how the significance measure is affected.

Figure 2 (bottom row) shows that the value of the joint-surprise after sorting js^{σ} is always decreasing, irrespective of the combination of error types applied to the two neurons. If only one error rate is modified, the offset of the decreasing joint-surprise is higher (light gray lines) than if another error is also applied (darker lines). Identical levels of the error rates are marked additionally by crosses. Their slope is always steeper than for non-identical levels (i.e. when one neuron experiences less errors), indicating an even stronger effect for identical error levels. Therefore we restrict ourselves in the following to the worst case scenario, i.e. to the case $\sigma_1^- = \sigma_2^- = \sigma^-$ and $\sigma_1^+ = \sigma_2^+ = \sigma^+$, without loss of generality but thereby lightening the formalism.

In order to investigate more deeply where the overall decrease of significance originates from, we jointly observe n^{σ}_{pred}, n^{σ}_{emp} and js^{σ} as functions of (symmetrical) false negatives and false positives error rates (Fig. 3). As expected (Eqs. (14) and (10) and Fig. 2 (bottom)), both measures decrease with σ^-, however n^{σ}_{emp} reduces more rapidly with increasing σ^- than n^{σ}_{pred} (solid gray lines). As a result, the significance also decreases with increasing σ^- (solid black line). For FPs only ($\sigma^- = 0$ and increasing σ^+), i.e. adding chance coincidences only, n^{σ}_{emp} and n^{σ}_{pred} increase in parallel (dashed gray lines). Again, the joint-surprise decreases with increasing errors (dashed black line), however for a different reason. Although the difference of the empirical and the expected number stays constant, due to increasing the mean and thus the width of the distribution the empirical number becomes less significant. As a consequence, both types of errors tend to reduce the significance of the observed coincidences after sorting, although the absolute number of coincidences may even increase with false positive errors.

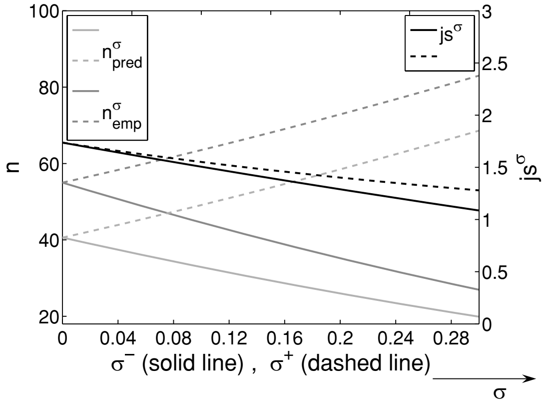

Fig. 3 Number of coincidences and resulting significance as a function of error rates. Two situations: increase of σ^- while $\sigma^+ = 0$ (solid lines), and increase of σ^+ while $\sigma^- = 0$ (dashed lines). For each of the situations, the empirical (n^{σ}_{emp}, dark gray) and the expected number of coincidences (n^{σ}_{pred}, light gray), as well as the resulting joint-surprise (js^{σ}, black) are shown. Both $js^{\sigma}(\sigma^-)$ and $js^{\sigma}(\sigma^+)$ decrease with error increase, although for increasing σ^+ the underlying number of coincidences increases. Parameters: $\lambda_1' = \lambda_2' = 20\,\text{Hz}$, $\lambda_c = 0.15\,\text{Hz}$, 100 seconds, bin size 1 ms

Figure 4(a) illustrates the relation of the joint-suprise after sorting in relation to the original value. Interestingly, this

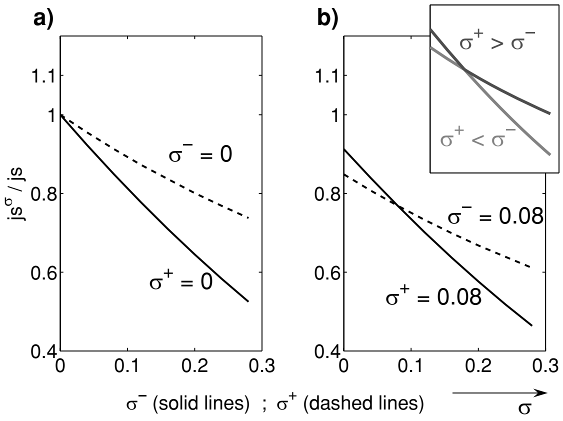

Fig. 4 Underestimation of significance of synchronization after spike sorting. The ratio between the joint-surprise after sorting js^{σ} and the original significance js is shown under variation of σ^+ for fixed values of σ^- (dashed lines) and under variation of σ^- while σ^+ has a fixed value (solid lines). (a) $\sigma^- = 0$ (dashed) and $\sigma^+ = 0$ (solid); (b) $\sigma^- = 0.08$ (dashed) and $\sigma^+ = 0.08$ (solid). (b, inset) Same data and parameters as in b), but now colored according to the relation of the error rates. Portions of the curves for which holds $\sigma^+ < \sigma^-$ are marked in light gray, for the condition $\sigma^+ > \sigma^-$ they are marked in dark gray

ratio (js^{σ}/js) is always smaller than 1, and decreases with increasing sorting errors, irrespective of error type. For false positive errors only (dashed line), the significance is less drastically reduced than for the same amount of false negative errors (solid line). This is particularly surprising, since in case of FP errors new coincidences are created and no coincidences are deleted.

Figure 4(b) shows the same curves as in panel (a) but for different combinations of error rates: instead of the respective other error to be set to 0, here it is set to 0.08. Now both curves start at values lower than 1, and cross each other. The crossing point is exactly at $\sigma^- = \sigma^+$. It is worth noticing that although the case $\sigma^- = \sigma^+$ balances the effect of errors on the firing rates (Eq. (6)), it still affects (and reduces) the significance after sorting. The inset shows that for the part of the curves always being largest (black line) holds the condition $\sigma^+ > \sigma^-$, again indicating that FNs reduce the joint-surprise more strongly than FPs do.

3.2. Dependence on physiological parameters

In this section we evaluate the dependencies of the various measures on physiological parameters, such as background rates λ'_i of the neurons $i = 1, 2$ and injected coincidence rate λ_c. We shall make use of simulations as introduced in Section 2.3. Since we are interested in the change of the coincidence numbers after sorting in relation to before sorting, we again express them as ratios. Also here, we only consider symmetrical error rates and use now Eq. (10) and Eq. (14) in which we have replaced $\sigma_1^- = \sigma_2^-$ by σ^- and $\sigma_1^+ = \sigma_2^+$ by σ^+.

3.2.1. Expected coincidences

The relation of the expected number of coincidences after sorting in relation to before sorting is given by Eq. (10) divided by n_{pred}:

$$\frac{n^{\sigma}_{\text{pred}}}{n_{\text{pred}}} = (1 + \sigma^+ - \sigma^-)^2. \tag{16}$$

Thus, $n^{\sigma}_{\text{pred}}/n_{\text{pred}}$ does not show any dependence on rates, but only a quadratic dependence on error rates.

3.2.2. Empirical coincidences

We analyse here the ratio between n^{σ}_{emp} (Eq. (14)) and n_{emp}. As introduced by Grün et al. (2002b) the latter can be expressed as the expected number given the uncorrelated neuronal activity plus the additionally injected coincidences (of rate λ_c):

$$n_{\text{emp}} = (\lambda_c \delta + \lambda'_1 \delta \cdot \lambda'_2 \delta) \cdot N. \tag{17}$$

For the expression of the expected number of coincidences we replace in Eq. (3) the rates by Eq. (9):

$$n_{\text{pred}} = \left(\lambda_c + \lambda'_1\right) \delta \cdot \left(\lambda_c + \lambda'_2\right) \delta \cdot N. \tag{18}$$

Inserting Eqs. (17) and (18) into Eq. (14) and rearranging leads to:

$$\frac{n^{\sigma}_{\text{emp}}}{n_{\text{emp}}} = (1 - \sigma^-)^2 + (2\sigma^+ - 2\sigma^-\sigma^+ + \sigma^{+2})\frac{n_{\text{pred}}}{n_{\text{emp}}} \tag{19}$$

$$= (1 - \sigma^-)^2 + (2\sigma^+ - 2\sigma^-\sigma^+ + \sigma^{+2}) \times \frac{\left(\lambda_c^2 + \lambda'_1\lambda_c + \lambda'_2\lambda_c + \lambda'_1\lambda'_2\right)\delta^2}{(\lambda_c\delta + \lambda'_1\lambda'_2\delta^2)}. \tag{20}$$

Figure 5(a) illustrates the two variables and their relation ($n^{\sigma}_{\text{emp}}/n_{\text{emp}}$) as a function of firing rate (top) and as a function of coincident firing rate (bottom) for fixed error rates. The graph shows that both, n^{σ}_{emp} and n_{emp}, increase with coincidence rate as well as with firing rate, with n^{σ}_{emp} always being lower than n_{emp}. For increasing λ_c (bottom), n^{σ}_{emp} increases with smaller slope than n_{emp}. As a consequence, the relation $n^{\sigma}_{\text{emp}}/n_{\text{emp}}$ decreases rapidly with λ_c. An increase in firing rate (top) leads to a slightly smaller increase of n^{σ}_{emp} as compared to n_{emp}, such that $n^{\sigma}_{\text{emp}}/n_{\text{emp}}$ grows rather slowly.

Figure 5(b) shows the dependence on error rates. $n^{\sigma}_{\text{emp}}/n_{\text{emp}}$ increases with σ^+ approximately linearly, with changes in firing and coincidence rates. Increasing firing rate increases the slope (top), whereas increasing coincidence rate decreases it (bottom). Taking the dashed line (corresponding to the parameters used in (a)) as a reference, we notice that the slope varies non-linearly with firing and coincidence rates. $n^{\sigma}_{\text{emp}}/n_{\text{emp}}$ decreases approximately linearly with σ^- (Fig. 5(c)). Here, changes in rate or coincidence rate do not influence the slope of the relation but rather the intersection with the vertical axis, again in a non-linear way. With increase in rate the vertical offset of $n^{\sigma}_{\text{emp}}/n_{\text{emp}}$ slightly increases, with increasing coincidence rate it slightly decreases. Note that the ratio may be larger than 1 for high firing rates or low coincidence rates, i.e. after sorting the empirical number of coincidences may be increased.

3.2.3. Significance

Here we analyze how the significance of synchronized spikes changes due to sorting errors in combination with the changes of physiological parameters explored in the foregoing subsections. We have seen above that the empirical number of coincidences may increase or decrease depending

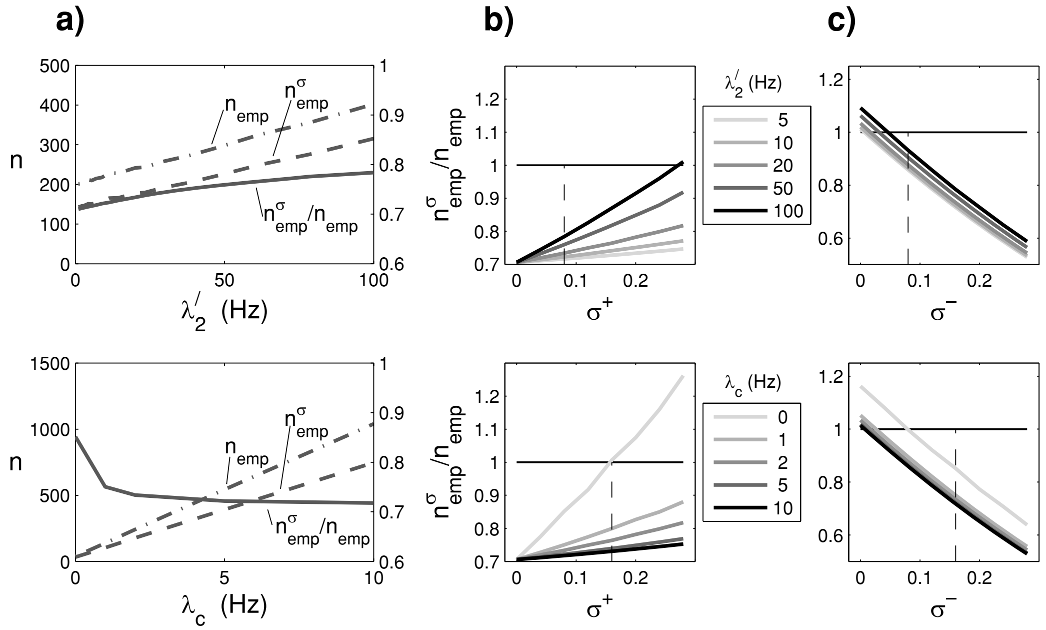

Fig. 5 Influence of firing, coincidence and error rates on the number of coincidences. (a) Empirical number of coincidences before (n_{emp}, dashed-dotted line) and after sorting ($n^{\sigma}_{\mathrm{emp}}$, dashed line) and their relation $n^{\sigma}_{\mathrm{emp}}/n_{\mathrm{emp}}$ (solid line; corresponding axis on the right) as a function of backgound firing rate λ_2' (top) and of injected coincidence rate λ_c (bottom). Values of error rates: $\sigma^{+} = 0.08$, $\sigma^{-} = 0.16$. (b,c) $n^{\sigma}_{\mathrm{emp}}/n_{\mathrm{emp}}$ as a function of error rates, in (b) as a function of σ^{+}, and in (c) as a function of σ^{-}. In the top row, the background rate of neuron 2 is varied (see legend) while $\lambda_c = 2\,\mathrm{Hz}$. In the bottom row, the coincidence rate is varied (see legend) while $\lambda_2' = 20\,\mathrm{Hz}$. In all plots (a,b,c) $\lambda_1' = 20\,\mathrm{Hz}$. The thin black line depicts for reference the identity relation $\frac{n^{\sigma}_{\mathrm{emp}}}{n_{\mathrm{emp}}} = 1$. The dashed thin lines mark the parameters used in (a). The values chosen for σ^{+} and σ^{-} in (a) are (about) the average error rates extracted by Harris et al. (2000)

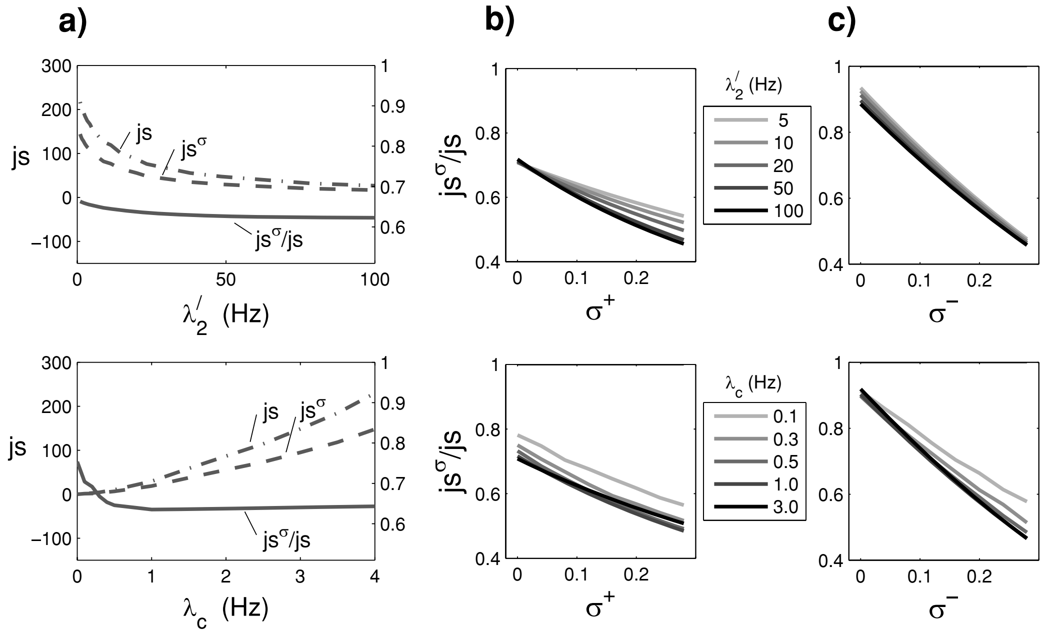

Fig. 6 Significance of synchronization: before and after spike sorting and their relation. Same parameters as in Fig. 5. (a) Joint-surprise before (js, dashed-dotted line) and after sorting (js^{σ}, dashed line) and their ratio (js^{σ}/js, axis on the right) as a function of backgound firing rate λ_2' (top) and of injected coincidence rate λ_c (bottom). (b) js^{σ}/js as a function of σ^{+} and (c) as a function of σ^{-}, both for different λ_2' (top) and λ_c (bottom)

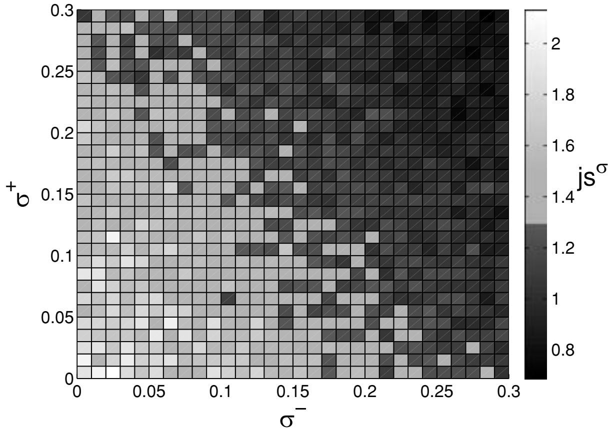

Fig. 7 Example of loss of significance due to sorting errors. Two simultaneous spike trains are simulated with background rates $\lambda_1' = \lambda_2' = 20\,\text{Hz}$ and a coincident rate of $\lambda_c = 0.15\,\text{Hz}$. Duration of the simulation $T = 10^5$ ms, time resolution $\delta = 1$ ms and 50 repetitions for each parameter setting of the error rates. Each bin in the matrix corresponds to a set of parameters σ^- (horizontal axis) and σ^+ (vertical axis), both varied from 0 to 0.3 in steps of 0.01. The original joint-surprise value without any sorting errors (bottom left corner) is $\bar{js} = 1.9 \pm 1.2$, i.e. well above the 5% level which corresponds to a value of $js = 1.28$ (switch dark to light in color bar). Increasing levels of sorting errors cause the significance js^σ to progressively decrease (from left to right and bottom to top). Light values represent js^σ above, dark values below significance threshold of 5%

on the specific parameter combinations. The change of the predicted numbers depends only on the error rates, i.e. no dependence on firing and coincidence rate is present. Now it needs to be evaluated how the significance of empirical coincidences given the predicted numbers changes with sorting. Figure 6(a) illustrates the changes of the joint-suprise relation js as a function of firing rate (top) and coincidence rate (bottom). With increasing λ_2', both components decrease. js^σ is always smaller than js, but js decreases faster and thus the relation of the two decreases. Although both components increase with increasing λ_c, the ratio also decreases, since js^σ is always smaller than js and does not increase as much as js.

Figure 6(b, c) shows the dependence of the ratio js^σ/js on the error rates (in b) for σ^+, in c) for σ^-) under variation of the firing rate (top) and variation of the coincidence rate (bottom). As already shown in Fig. 4, the relation decreases with error rates, and we see here that the firing and coincidence rates are only slightly modifying the slopes. The dominating parameters are the error rates.

In summary, as a general result we observe that sorting errors always reduce the significance, irrespective of the error type, and mostly independent of the physiological parameters. Figure 7 shows an example where the originally significant synchronization is changed due to the sorting errors to an unsignificant result.

4. Discussion

4.1. Sorting errors reduce significance of spike synchrony

We studied the influence of spike sorting errors on the analysis of unitary events. Our statistical model of spike sorting is based on the assumptions that false positive and false negative errors are likely to occur at any instant in time and independently for the two neurons. As a consequence, sorting errors lead to erroneous statistical rate estimations, as well as to erroneous coincidence counts and erroneous significance of the latter. In order to understand the influence of these errors on the significance estimation of coincident spikes, we derived analytically the predicted and the empirical number of coincidences after sorting as compared to their original values before sorting. In addition we also tested, both analytically and through simulations, the influence of physiological parameters like the firing rate of the neurons and the coincidence rate. We showed that the joint-surprise reduces with error rates, finally leading to a loss of significance of originally significant spike synchronization. It turned out, that the significance of spike synchronization is always reduced by sorting errors irrespective of the error type. This also holds for non-symmetrical combinations of error rates experienced by the two neurons, however with less strong reduction of the significance as compared to the symmetrical case. The reduction of significance is mainly due to the "normalization" of the observed coincidence counts by comparison to their expected number. However, the reason for significance reduction is different for the different error types.

False positive spikes lead to an increase of the empirical as well as of the expected number of coincidences. Most interestingly they even increase by the same absolute amount, which is due to additional chance coincidences only. The difference between the two measures are the "excess coincidences", whose relative amount reduces, thus leading to a decrease in significance.

Missing spikes (FNs) reduce the number of expected coincidences as well as the empirical number. The expected number is reduced due to decreased firing rates. The empirical number in addition experiences a reduction of the excess coincidences (first term in Eq. (14)). Therefore the significance reduces (Fig. 3, solid lines). The impact of FNs on synchronized spikes is strong (for error rates found by Harris et al. (2000) the joint-surprise may be reduced by 40%, cf. also Fig. 4), which is explained by the fact that the loss of only one spike of a synchronized event already leads to the loss of the coincidence. Furthermore, the higher the coincidence rate, the larger the chance to miss a spike, and thus the larger the reduction of the empirical coincidences (see Fig. 5(a,b), bottom). In summary, a conservative sorting strategy, i.e. rather missing a spike than capturing a wrong one, might lead to a stronger loss of detectability of originally existing significant synchronization.

Co-occurrence of both error types may lead to partial (for $\sigma^- \neq \sigma^+$) or even to full cancellation (for $\sigma^- = \sigma^+$) of change of the expected number of coincidences after sorting. Depending on the net excess of errors the expected number may be increased (for $\sigma^- < \sigma^+$) or decreased (for $\sigma^- > \sigma^+$). For the empirical number, the errors may not cancel in respect to the number of excess coincidences. Excess coincidences may be deleted by FNs, but cannot be created by FPs, since newly created coincidences enter as chance coincidences only. In addition, the probability for generating chance coincidences is much smaller than the impact of injected coincidences (cf. Eqs. (17) and (18)), and thus the effects on the significance are not canceled in case $\sigma^- = \sigma^+$ (see also Fig. 4). Interestingly, the available literature on the relative amounts of both errors documented a tendency towards higher values of FNs as compared to FPs (Harris et al., 2000; Wood et al., 2004), indicating a preference for conservative strategies. Thus one may conclude that neural synchronization is typically underestimated.

4.2. Relation to cross-correlation

The cross-correlogram is a widely used method for estimating the correlation between the spiking activities of two neurons (Perkel et al. 1967b). The method allows to study zero-lag and delayed coincidences. The unitary event analysis method concentrates on zero-lag (or near-coincident) events and may be directly compared to the zero-lag (or near-by) bins of the cross-correlogram. Cross-correlograms are often used without any normalization and just provide the empirical coincidence counts (raw cross-correlation). In addition, normalization procedures are availabe, that account for the expected number of coincidences given the firing rates of the neurons (cross-covariance) by subtracting the latter from the empirical counts. Normalizing this number to the product of the variances of the single processes yields the correlation coefficient (for an extensive discussion and comparison of such measures see Aertsen et al. (1989)). Thus, the zero-bin correlation coefficient in our variables reads $r = \frac{n_{\text{emp}} - n_{\text{pred}}}{\sqrt{T \cdot \lambda_1}\sqrt{T \cdot \lambda_2}}$. This measure behaves very similarly to the joint-surprise, however with different abolute numbers.

Let us emphasize that correlation analyses which only consider the raw (i.e. empirical) coincidence counts, e.g. like in an uncorrected cross-correlogram, would find increased synchronization in the case of FPs. However relating the observed number of coincidences to the expected ones, as done in the UE analysis (Grün et al., 2002a, b, 1999) or by the correlation coefficient, corrects for that.

4.3. Experimental situations that violate UE assumptions

Inherent assumptions of the unitary event analysis are stationary Poisson processes. However, experimental data typically do not fulfill these assumptions and methods to account for non-stationary (Grün et al., 2002b; Grün 2003) and non-Poisson processes (Pipa and Grün, 2003; Pipa et al., 2003; Pipa 2001; Pipa et al., 2006) have been worked out.

In case of non-stationary processes in time the solution to avoid false positives is to perform the analysis in sliding window fashion to account for the change in rate (Grün et al., 2002b). In case of non-stationarity across trials, the expected number of coincidences is calculated within the sliding window in a trial-by-trial manner and then summed yields the correct estimate for the expected number of coincidences (Grün et al., 2003). By doing that we account for non-stationary firing rates. A change in brain state or different behavioral variables could lead to coherent change of neuronal firing rates, e.g. by a common increase in the membrane potentials as discussed in Brody (1998). In the same way as the firing rates co-vary the amount of chance coincidences do: the larger the firing rates the more chance coincidences. As discussed before such a case is well treated by the unitary event method which is specifically designed to evaluate the degree of excess synchrony as compared to chance synchrony and to estimate its significance, also under non-stationary conditions.

Sorting errors like the loss of coincident spikes due to an overlap of the spike waveforms, would trivially covary with covariation of the rates. However, treating the data in segments of stationary firing rates, we are per segment in a stationary situation which is discussed below (Section 4.4.2).

In Grün et al. (2002a) we have shown that the unitary event analysis method reacts robustly to non-Poissonian point processes that were simulated as Gamma processes. Although the auto-correlation structure of the spike trains affects the shape of the coincidence distribution, it turned out that assuming a Poisson distribution -as analytically derived for Poisson spike trains (Grün et al., 2002a)- leads to conservative estimates for experimentally found ranges of the shape factor (Pipa, 2001; Pipa et al., 2006).

We also performed the analysis for the dependence of FN and FP sorting errors for the case of Gamma processes (not shown here) and found no differences in the results. FN and FP errors applied to Gamma processes also lead to a reduction of the sigificance as in the case of Poisson spike trains.

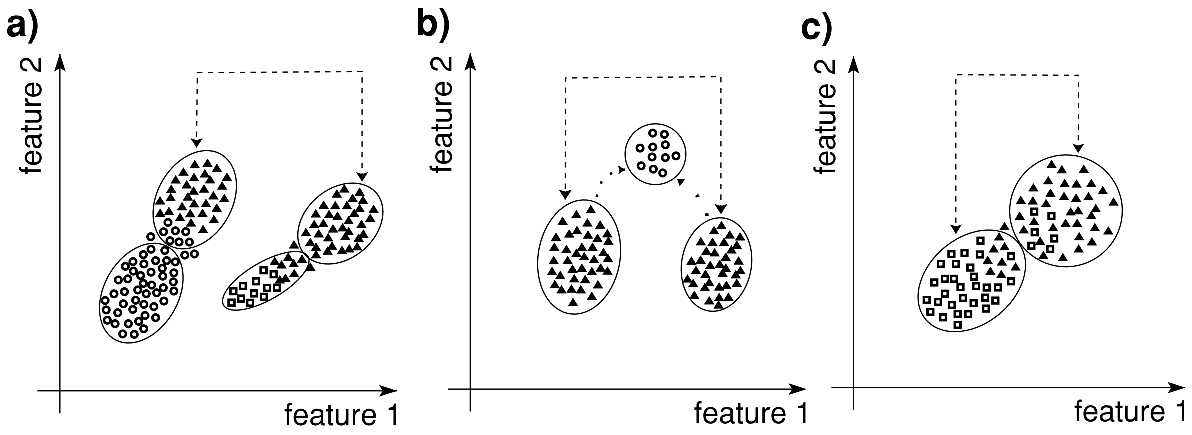

Fig. 8 Experimentally relevant cases of spike sorting. Sketch of possible origins of spike sorting errors due to specific cluster configurations in the space of two spike-waveform features. Triangles, squares and circles represent spikes, contours represent putative clusters, and dashed lines indicate the two neurons whose correlation is analyzed. (a) Due to the presence of other neurons they may erroneously provide (FP error; circles in upper left cluster) or "steal" (FN error; triangles in lower right cluster) spikes to/from the neurons under consideration. (b) Ghost cluster. A significant amount of overlapping spikes originating from the two neurons under consideration for correlation analysis, gives rise to a ghost cluster (middle cluster containing circles), which "steals" coincident spikes from the two neurons. (c) Correlated errors. The neurons being analyzed are recorded from the same electrode and their cluster separation is poor, indicated by squares from the left cluster mixing with triangles in the right cluster, and vice versa. Therefore a number of spikes are assigned from one neuron to the other and vice versa, resulting in false negatives of one neuron that become false positives in the other

4.4. Typical errors in sorting

Now we shall relate our considerations to typical experimental cases, thus eventually abandoning the model hypotheses of independent errors and neurons recorded from different electrodes. Furthermore we will also discuss this in the framework of cluster analysis in spike-waveform feature space.

4.4.1. Independent sorting errors

Here we investigate the correlation between two neurons, each of which experiences FN or FP errors from other, unrelated neurons. This is the case if we consider two neurons that were recorded and sorted from two different electrodes. It may also be the case for two neurons that are recorded from the same electrode and are well sorted, but one or both neurons experience errors due to yet other neurons (see Fig. 8(a)). This is actually the case we studied in the framework of our statistical sorting model, and learned that irrespective of the error type the resulting significance is reduced. The reduction is strongest for the case that both neurons experience the same degree of errors, but not as pronounced if one of the neurons experiences a smaller amount of errors than the other. If only one neuron is exposed to errors, the reduction of significance is even smaller (Fig. 2, all upper joint-surprise curves).

In the case of a rather conservative sorting strategy, i.e. if one tries to catch as little FPs as possible thereby accepting the danger of loosing spikes (i.e. having more FNs), the significance is more reduced than for the case of a rather tolerant sorting strategy. The latter would represent the case if one tries to catch as many spikes as possible while accepting false positives. In the cluster space of a given sorting algorithm, a conservative sorting strategy would correspond to accepting only spikes from within a radius smaller than the outer bound of the cluster of points, whereas a tolerant sorting strategy would take spikes from a radius catching all the points in the cluster of points. Schmitzer-Torbert et al. (2005) suggested measures for cluster quality. The L_{ratio} measures the amount of noise observed in the vicinity of the cluster. The authors showed that a high value of the L_{ratio} correlates with a high number of FNs, the case we would call conservative sorting.

Directly related to our study is the work by Bedenbaugh and Gerstein (1997) and Gerstein (2000) who evaluated the correlation coefficient as a measure for synchrony between unresolved multi-neuron recordings. They assumed the activities recorded on each of the two electrodes to be mixtures of spiking activities of up to three neurons which may fire independently or include correlations. The focus is on the resulting correlation between the multi-neuron recordings from two electrodes as functions of the original correlation of neurons across electrodes and as functions of correlations between neurons at a single electrode. Thus mixtures of spikes trains are studied which in our scenario may be interpreted as false positives, but the framework neglects the problem of missing spikes. Under specific parameter settings, Gerstein's (2000) results are directly comparable to ours. He also found that the resulting correlation between spike trains that experienced contamination with uncorrelated spikes is reduced as compared to the original, distant correlation between two neurons.

4.4.2. Overlapping spikes

Another relevant situation in spike sorting is that coincident spikes of two neurons recorded at the same electrode may overlap in time such that the resulting waveform may result in the formation of a ghost cluster (Fig. 8(b), cluster that includes circles). As a result, the simultaneous spikes are systematically removed from the two spike trains under consideration.

The work by Bar-Gad et al. (2001) studied the influence of missing coincident spikes onto the cross-correlation measure, under the specific condition that a spike is followed by a noticable dead time, and neurons have high firing rates. Under these conditions they found that the shape of the cross-correlation function alters: zero-delay coincidences were lacking, but close delayed coincidences were enhanced as compared to background. Corrections for expectancy were not performed. Thus, in our words, they studied the influence of false negative coincidences and found for the case of zero-delayed coincidences a reduction in their numbers.

In our framework this case can be represented as a specific deletion of coincident spikes. As a consequence, the empirical number of coincidences is reduced by a fraction σ^-_{coinc}: $n^\sigma_{\text{emp}} = (1 - \sigma^-_{\text{coinc}}) \cdot n_{\text{emp}}$. Compared to the case where FN errors are independently applied to both neurons, case which leads to quadratic dependence on σ^-_{coinc} (substitute in Eq. (14) σ^-_i by σ^-_{coinc} and set $\sigma^+_i = 0, i = 1, 2$), here the deletion of coincident events leads also to a reduction of the empirical coincidences, but with a slope that is less steep. On the other hand, with increasing σ^-_{coinc} the expected number is also decreased, however by a smaller amount as compared to independently applied FN errors (Eq. (10)), since here only spikes that are involved in coincidences may be deleted. The expected number of coincidences then reads: $n^\sigma_{\text{pred}} = (\lambda_1 - \lambda_{\text{del}}) \cdot (\lambda_2 - \lambda_{\text{del}}) \cdot \delta^2 \cdot N$, with $\lambda_{\text{del}} = \frac{n_{\text{emp}} \cdot \sigma^-_{\text{coinc}}}{T}$ the spike reduction rate. As a result, also in the case of a joint deletion of coincident spikes, the joint-surprise decreases, even with a similar slope as for the case of independent FN errors.

4.4.3. Overlapping clusters

Another problem that may arise if different neurons are recorded and sorted from the same electrode is, that spikes from one neuron may be assigned to another neuron. Quirk and Wilson (1999) discussed such a case for bursting neurons, where late spikes in a burst were different in shape and thus were wrongly assigned to the other neuron. Using cross-correlation analysis for the identification of correlation between the spiking activities of the two neurons, they found artificially increased delayed coincidences. Although their case does not apply to zero-bin correlations, the case that spikes of one neuron are assigned to another is of general interest.

This type of correlation errors can be expressed in our framework as FNs of one neuron become FPs of the other. The empirical number of the coincidences will stay the same, since coincidences will not be created by moving spikes to the other spike train, nor will spikes of coincidences be moved to the other neuron since that would correspond to another waveform (overlapping spikes). Thus, only the expected number of coincidences is influenced: $n^\sigma_{\text{pred}} = (1 - \sigma^1)\lambda_1 \cdot (1 + \sigma^1)\lambda_2 \cdot \delta^2 \cdot N$, with σ^1 the probability of assigning spikes to a wrong cluster. This leads to a decrease of the expected number of coincidences, and consequently the significance of the empirical coincidences increases with error rate.

A related case is if waveforms lead to overlapping clusters of entries in the feature space (Fig. 8(c)), a case indicated as of bad separation of clusters using the isolation distance measure suggested by Schmitzer-Torbert et al. (2005). If anyway spikes are assigned to distinct clusters, there is a high likelyhood to get false negative and false positive errors. Spikes from one cluster may be assigned to the other, i.e. FNs of the first become FPs of the other, and also vice versa. Thus errors will occur in both directions. The empirical number of coincidences will not be affected by that, for the same reasons as discussed for the one-way case above. However, the expression for the expected number of coincidences will get additional terms expressing the probability for FNs from the second cluster becoming FPs in the first (via σ^2): $n^\sigma_{\text{pred}} = (1 - \sigma^1 + \sigma^2)\lambda_1 \cdot (1 + \sigma^1 - \sigma^2)\lambda_2 \cdot \delta^2 \cdot N$. As a result the error rates may partly compensate each other and thereby leading to a less strong increase of the resulting joint-surprise as compared to the one-way case.

In summary, correlated errors specifically may occur if different neurons are recorded and sorted from a single electrode. Overlapping spikes may be assigned to a third cluster, thereby leading to a loss of empirical coincidences. As a consequence the measured significance is reduced and the underlying existing correlation between neurons may be overseen. Correlation analysis of the activities of two neurons that experienced wrong assignments of spikes from one neuron to the other (and/or vice versa) is dangerous. The detected significance is prone to indicate false positive correlation. Specifically the case when only one of the neurons is giving spikes to other.

4.5. Conclusions

Spike sorting introduces two types of errors into the coincidence count: false positive and false negatives. If errors were experienced independently, they both lead to a reduction of the measured significance as compared to the original correlation between the two neurons. The reason for the reduction is different for the two cases. False negatives simply partly delete correlated spikes. However, false positives may increase the correlation, and correlation measures that do not correct for expectation by chance may conclude false positive correlation.

If spike trains are sorted from a single electrode, also correlated sorting errors may occur, e.g. coincident spikes that are ignored due to overlapping waveforms, or spikes that are assigned from one neuron to another. In particular the latter is problematic if one is interested to detect spike correlation between the two neurons involved, since there is the danger of detecting false positive correlation. Therefore we rather suggest to avoid correlating neuronal spiking activity recorded from the same electrode combined with low quality of cluster separation. However, for errors resulting from a neuron not considered, this scenario leads to a reduction of significance and thus rather to an underestimation of the underlying correlation.

One main aspect is guiding our future work in this context. The unitary event analysis was designed to allow for the detection of neuronal assemblies and thus for correlation analysis of more than two neurons at a time. As a consequence we aim to understand how spike sorting errors influence the analysis of higher-order spike patterns. This includes to evaluate systems with more than two simultaneous recordings, which involves a combinatorial increase of cases to consider.

In conclusion, if one is interested in neuronal interaction in the brain, the activities of single units need to be simultaneously observed and analyzed for correlated activity. Therefore spike sorting is an important, intermediate statistical analysis for extracting the single unit activities. Subsequent analysis may be considerably influenced by sorting errors, and may arrive at wrong conclusions. In this work we have shown that for the unitary event analysis, independent sorting errors lead to reduced significance, but not to artificial correlation.

Appendix

Rewriting Eq. (14) yields:

$$n^{\sigma}_{\mathrm{emp}} = (1-\sigma_1^-)(1-\sigma_2^-)\cdot[n_{\mathrm{emp}} - n_{\mathrm{pred}}] \\ +(1+\sigma_1^+ - \sigma_1^-)(1+\sigma_2^+ - \sigma_2^-)\cdot n_{\mathrm{pred}}. \qquad (14)$$

After sorting, the experimental spike trains contain FP and miss FN spikes (cf. Fig. 1(b), bottom). The former occupy bins that were empty, whereas a fraction of currently empty bins was previously occupied by the latter. We introduce the probability $p_i(+)$ to find a FP spike in a bin and the joint probability $p_i(-, \mathrm{spike})$ to miss a spike as FN:

$$p_i(+) = \sigma_i^+ \cdot \lambda_i \delta \qquad p_i(-, \mathrm{spike}) = \sigma_i^- \cdot \lambda_i \delta,$$

where λ_i are the original firing rates of neuron i and δ the bin width. We will refer to the two neurons through the indices i, j, where $i \neq j$. These expressions are not symmetrical because they refer to intrinsically different generating processes: insertion of new spikes (FPs) versus deletion of already existing spikes (FNs). In the following we also need the expression for the conditional probability of missing a given spike $p_i(-|\mathrm{spike})$. Using Bayes theorem we obtain $p_i(-|\mathrm{spike}) = \frac{p_i(-,\mathrm{spike})}{p_i(\mathrm{spike})} = \sigma_i^-$. It follows that the probability for a (given) spike not to be missed equals $(1 - p_i(-|\mathrm{spike})) = 1 - \sigma_i^-$.

The number of coincidences after sorting can be expressed as the sum of the original number of coincidences plus two additional terms, which correct for the erroneously missed and produced coincidences due to sorting:

$$n^{\sigma}_{\mathrm{emp}} = n_{\mathrm{emp}} - n^{\mathrm{FN}} + n^{\mathrm{FP}}. \qquad (\mathrm{A1})$$

The first of these additional terms is composed of three contributions. The first two are given by the probability that either one of the two coincident spikes is missed: $p_i(-|\mathrm{spike}) \cdot (1 - p_j(-|\mathrm{spike})) = \sigma_i^-(1-\sigma_j^-)$. The last term is the probability that both spikes are missed as FNs: $p_i(-|\mathrm{spike}) p_j(-|\mathrm{spike}) = \sigma_i^- \sigma_j^-$. This results in:

$$n^{\mathrm{FN}} = n_{\mathrm{emp}} \cdot [\sigma_i^-(1-\sigma_j^-) + \sigma_j^-(1-\sigma_i^-) + \sigma_i^- \cdot \sigma_j^-]. \qquad (\mathrm{A2})$$

Similarly, the number of false coincidences introduced by sorting errors (n^{FP}) receives contributions from three terms. Again, the first two terms are given by the probability that a FP is synchronous to an already existing spike, i.e. the product of the probability of a false positive in neuron i ($p_i(+)$) and the probability that in neuron j an original spike existed and survived

$$p_j^{\sigma^-} = p_j(1 - p_j(-|\mathrm{spike})) = \lambda_j \delta (1 - \sigma_j^-),$$

with p_i defined in Eq. (1). This yields: $p_i(+) p_j^{\sigma^-} = \lambda_i \lambda_j \delta^2 \sigma_i^+ (1 - \sigma_j^-)$. The last term consists of the probability of a coincidence of FP spikes: $p_i(+) p_j(+) = \lambda_i \lambda_j \delta^2 \sigma_i^+ \sigma_j^+$.

In total we obtain:

$$n^{\mathrm{FP}} = n_{\mathrm{pred}} \cdot [\sigma_i^+(1-\sigma_j^-) + \sigma_j^+(1-\sigma_i^-) + \sigma_i^+ \cdot \sigma_j^+], \tag{A3}$$

We notice here that $\lambda_i \lambda_j \delta^2 N = n_{\mathrm{pred}}$ (Eq. (3)) and that the factor at the right-hand side of Eq. (A3) can be expressed as:

$$\begin{aligned} &\sigma_i^+\sigma_j^+ + \sigma_i^+(1-\sigma_j^-) + \sigma_j^+(1-\sigma_i^-) \\ &\quad = (1-\sigma_i)(1-\sigma_j) - (1-\sigma_i^-)(1-\sigma_j^-), \end{aligned} \tag{A4}$$

where we used again $\sigma_i = \sigma_i^- - \sigma_i^+$. After substituting Eqs. (A2) and (A3) in Eq. (A1) and rearranging we obtain the relation for $n^{\sigma}_{\mathrm{emp}}$ in Eq. 14.

Acknowledgments Supported in part by the Volkswagen Foundation, the Stifterverband für die Deutsche Wissenschaft and the BMBF Bernstein Center for Computational Neuroscience Berlin (grant 01GQ0413). We also thank Markus Diesmann for useful comments on an earlier version of the manuscript. Also we thank the anonymous reviewer for useful comments for further improvements of the manuscript.

References

Aertsen AMHJ, Gerstein GL, Habib MK, Palm G (1989) Dynamics of neuronal firing correlation: Modulation of effective connectivity. J. Neurophysiol. 61(5): 900–917.

Bar-Gad I, Ritov Y, Vaadia E, Bergman H (2001) Failure in identification of overlapping spikes from multiple neuron activity causes artificial correlations. J. Neurosci. Methods 107(1–2): 1–13.

Bedenbaugh P, Gerstein GL (1997) Multiunit normalized cross correlation differs from the average single-unit normalized correlation. Neural Comput. 9(6): 1265–1275.

Brody CD (1998) Slow covariations in neuronal resting potential can lead to artefactually fast crosscorrelations in the spike trains. J. Neurophysiol. 80: 3345–3351.

Fee MS, Mitra PP, Kleinfeld D (1996) Variability of extracellular spike waveforms of cortical neurons. J. Neurophysiol. 76(6): 3823–3833.

Gerstein GL (2000) Cross-correlation measures of unresolved multineuron recordings. J. Neurosci. Methods 100(1–2): 41–51.

Grün S, Diesmann M, Aertsen A (2002a) Unitary events in multiple single-neuron spiking activity: I. Detection and significance. Neural Comput. 14(1): 43–80.

Grün S, Diesmann M, Aertsen A (2002b) Unitary events in multiple single-neuron spiking activity: II. Nonstationary data. Neural Comput. 14(1): 81–119.

Grün S, Diesmann M, Grammont F, Riehle A, Aertsen A (1999) Detecting unitary events without discretization of time. J. Neurosci. Methods 94(1): 67–79.

Grün S, Riehle A, Diesmann M (2003) Effect of cross-trial nonstationarity on joint-spike events. Biol. Cybern. 88(5): 335–351.

Harris K, Henze D, Csicsvari J, Hirase H, Buszàki G (2000) Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. J. Neurophysiol. 84(1): 401–414.

Hyvärinen A, Oja E (2000) Independent component analysis: Algorithms and applications. Neural Networks 13(4–5): 411–430.

Lewicki MS (1998) A review of methods for spike sorting: The detection and classification of neural action potentials. Network 9(4): R53–R78.

Perkel DH, Gerstein GL, Moore GP (1967b) Neuronal Spike Trains and Stochastic Point Processes. II. Simultaneous spike trains. Biophys. J. 7(4): 419–440.

Pipa G, Diesmann M, Grün S (2003) Significance of joint-spike events based on trial-shuffling by efficient combinatorial methods. Complexity 8(4): 79–86.

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

Pipa G (2001) Entwicklung und Untersuchung einer nicht-parametriscchen Methode zur Schätzung der Signifikanz zeitlich koordinierter Spike-Aktivität. Diploma thesis, Physics, J-W Goethe Univ, Frankfurt/M, Germany.

Pipa G, van Vreeswijk C, Grün S (2006) Impact of spike-train autostructure on probability distribution of joint-spike events (in preparation).

Pouzat C, Delescluse M, Viot P, Diebolt J (2004) Improved spike-sorting by modeling firing statistics and burst-dependent spike amplitude attenuation: A Markov chain Monte Carlo approach. J. Neurophysiol. 91(6): 2910–2928.

Pouzat C, Mazor O, Laurent G (2002) Using noise signature to optimize spike-sorting and to assess neuronal classification quality. J. Neurosci. Methods 122(1): 43–57.

Quirk MC, Wilson MA (1999) Interaction between spike waveform classification and temporal sequence detection. J. Neurosci. Methods 94(1): 41–52.

Quiroga RQ, Nadasdy Z, Ben-Shaul Y (2004) Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Comput. 16(8): 1661–1687.

Riehle A, Grammont F, Diesmann M, Grün S (2000) Dynamical changes and temporal precision of synchronized spiking activity in monkey motor cortex during movement preparation. J. Physiol. (Paris) 94(5–6): 569–582.

Riehle A, Grün S, Diesmann M, Aertsen A (1997) Spike synchronization and rate modulation differentially involved in motor cortical function. Science 278(5345): 1950–1953.

Schmitzer-Torbert N, Jackson J, Henze D, Harris K, Redish AD (2005) Quantitative measures of cluster quality for use in extracellular recordings. Neuroscience 131(1): 1–11.

Takahashi S, Anzai Y, Sakurai Y (2003) Automatic sorting for multineuronal activity recorded with tetrodes in the presence of overlapping spikes. J. Neurophysiol. 89(4): 2245–2258.

Wehr M, Pezaris JS, Sahani M (1999) Simultaneous paired intracellular and tetrode recordings for evaluating the performance of spike sorting algorithms. Neurocomputing 26–27: 1061–1068.

Wood F, Black MJ, Vargas-Irwin C, Fellows M, Donoghue JP (2004) On the variability of manual spike sorting. IEEE Trans. Biomed. Eng. 51(6): 912–918.

Zhang PM, Wu JY, Zhou Y, Liang PJ, Yuan JQ (2004) Spike sorting based on automatic template reconstruction with a partial solution to the overlapping problem. J. Neurosci. Methods 135(1–2): 55–65.