RESEARCH ARTICLE

Significance of Joint-Spike Events Based on Trial-Shuffling by Efficient Combinatorial Methods

GORDON PIPA,[1] MARKUS DIESMANN,[2] AND SONJA GRÜN[1,*]

[1]*Department of Neurophysiology, Max-Planck-Institute for Brain Research, Frankfurt/Main, Germany; e-mail: pipa@mpih-frankfurt.mpg.de*

[2]*Department of Nonlinear Dynamics, Max-Planck-Institut für Strömungsforschung, Göttingen, Germany*

Received 20 August 2002; revised 5 May 2003; accepted 5 May 2003

The assembly hypothesis suggests that information processing in the cortex is mediated by groups of neurons expressed by their coordinated spiking activity. Thus, the unitary events analysis was designed to detect the presence of conspicuous joint-spike events in multiple single-unit recordings and to evaluate their statistical significance. The null hypothesis of the associated test assumes independent Poisson processes and leads to parametric significance estimation. In order to allow for arbitrary processes here we suggest to base the significance estimation on trial shuffling and resampling. In this scheme the null hypothesis is implemented by combining spike trains from nonsimultaneous trials and counting the joint-spike events. The coincidence distribution serving for the significance estimation is generated by repetitive resampling. The number of all possible recombinations, however, grows dramatically with the number of trials and neurons and thus is not practical for a user-interactive implementation of the analysis. We have suggested a Monte-Carlo-based resampling procedure and demonstrated that the procedure yields an appropriate estimate of the distribution and reliable significance estimation. In contrast, here, we present an exact solution. Rewriting the statistical problem in terms of certain macrostates, we are able to systematically sample the coincidence counts from all trial combinations. In addition we restrict the generating process to those counts forming the relevant tail of the distribution. The computationally effective implementation uses the concept of partitions. © 2003 Wiley Periodicals, Inc.

Key Words: spike synchronization; trial shuffling; significance; boot-strapping

1. INTRODUCTION

The assembly hypothesis postulates[1] that information processing in the cortex is mediated by groups of neurons, meanwhile supported by a number of experimental studies (see[2] and references therein). This hypothesis is addressed by unitary event (UE) analysis[3,4] enabling the study of the relation of spike synchronization to behavioral events.[4,5–7] In the UE-analysis, empirically observed coincidence counts are evaluated for their significance based on a Poisson distribution (see[4,8] for discussion). The expectation value parameterizing the distribution is calculated by the product of the firing rates of the contributing

*Present address: Institute for Biology, Neurobiology, Free University Berlin, Berlin, Germany.

neurons. Implicit in this procedure is the assumption that spike trains are generated by Poisson processes. However, experimental data often fail to be compatible with this assumption (e.g.,[9–11]). Thus, in order to consider the original, temporal structures of the spike trains, we suggest to replace the significance test described above by a nonparametric test. The basic idea is to base the null hypothesis of independence on trial shuffling. For the significance estimation of the empirically found coincidences in M trials, we determine the probability distribution of chance coincidences on the basis of the coincidence counts occurring in the set of potential M trial combinations. The derivation of the probability distribution employs resampling procedures, typically leading to large numbers of different trial combinations. However, our goals are (1) to use the UE-analysis in user-interactive manner, (2) to perform the UE analysis in time-resolved fashion (sliding window procedure,[4]), which requires repetitive generation of coincidence distributions, and (3) to carry out the analysis for many different coincidence patterns. Taken together, the procedure for generating the coincidence count distribution is subject to strong time constraints.

One possible solution is to use an approximate Monte-Carlo method for successive random resampling.[12] This is contrasted by the analytical approach for *exact* significance estimation presented here, based on a concept (macrostates versus microstates) borrowed from statistical physics. Although the resulting combinatorics is still demanding, it permits the construction of computationally feasible solutions.

In section 2 we introduce the required notation and show how, in principle, significance can be estimated in exact fashion (section 2.1.), followed by a brief description of the earlier Monte-Carlo approach (details can be found elsewhere[12]). In section 3 we introduce our new approach, which derives the significance measure of joint-spike events on the basis of combinatorial resampling. In section 4 we present an effective implementation of this approach using the idea of *limited partitions* where standard algorithms for the partitions of an integer are supplied with additional constraints. Finally, the new scheme is discussed in the light of earlier approaches.

2. NONPARAMETRIC SIGNIFICANCE ESTIMATION OF JOINT-SPIKE EVENTS

To incorporate the null-hypothesis of independence, trials of the original data set are shuffled to combine nonsimultaneously recorded spike trains. Each tuple of shuffled spike trains yields a count ω^s of joint-spike events. The S possible shuffled combinations of spike trains (all trial indices being different) yield S counts ω composing the list $\Omega_0 = [\omega]$ (see box A in Figure 1). F1

For $N \geq 2$ neurons we have

$$\Omega_0 = [\omega_{j(i_1,\ldots,i_N)} | i_k \neq i_l, \quad \text{with } i_k, i_l \in 1 \ldots M], \quad (1)$$

where the i_k are trial numbers and M is the total number of trials in the data set (the arbitrary bijective function j serves to simplify the index of ω by mapping the index vector $(i_1, \ldots, i_N)$ to an integer in the range $1, \ldots, S$). Therefore Ω_0 is composed of a nonunique, nonordered list of integers $0 \leq \omega_j \leq \omega_{\max}$ with $\omega_{\max} = \max(\Omega_0)$ consisting of

$$S = |\Omega_0| = \frac{M!}{(M-N)!} \quad (2)$$

elements [see Figure 2(A) for the dependence of S on M]. F2

In order to estimate the significance of the empirical number of joint-spike events n_{emp} observed in M trials of simultaneously recorded spike trains, we need to know the probability distribution P_s of the total number of joint-spike events $\bar{\omega}$ occurring in M trials under the assumption of independent processes. P_s can be constructed by drawing multiple sets of M samples from Ω_0 using resampling procedures[13] [see Figure 1(B)]. Two available procedures are explained in sections 2.1 and 2.2. Having derived the distribution of coincidence counts P_s, we can determine the significance of n_{emp} by calculating the joint p-value,[3] i.e., the probability α of observing n_{emp} or an even larger count [see Figure 3(A) for an illustration]: F3

$$\alpha = p(\bar{\omega} \geq n_{\text{emp}}) = 1 - p(\bar{\omega} < n_{\text{emp}}) = 1 - \sum_{k=0}^{n_{\text{emp}}-1} P_s(k) \quad \text{with } n_{\text{emp}} = \sum_{i=1}^{M} \omega_i^s. \quad (3)$$

The ω_i^s denote the M coincidence counts of the original (simultaneous) data set.

2.1. Ideal Estimation of the Probability Distribution

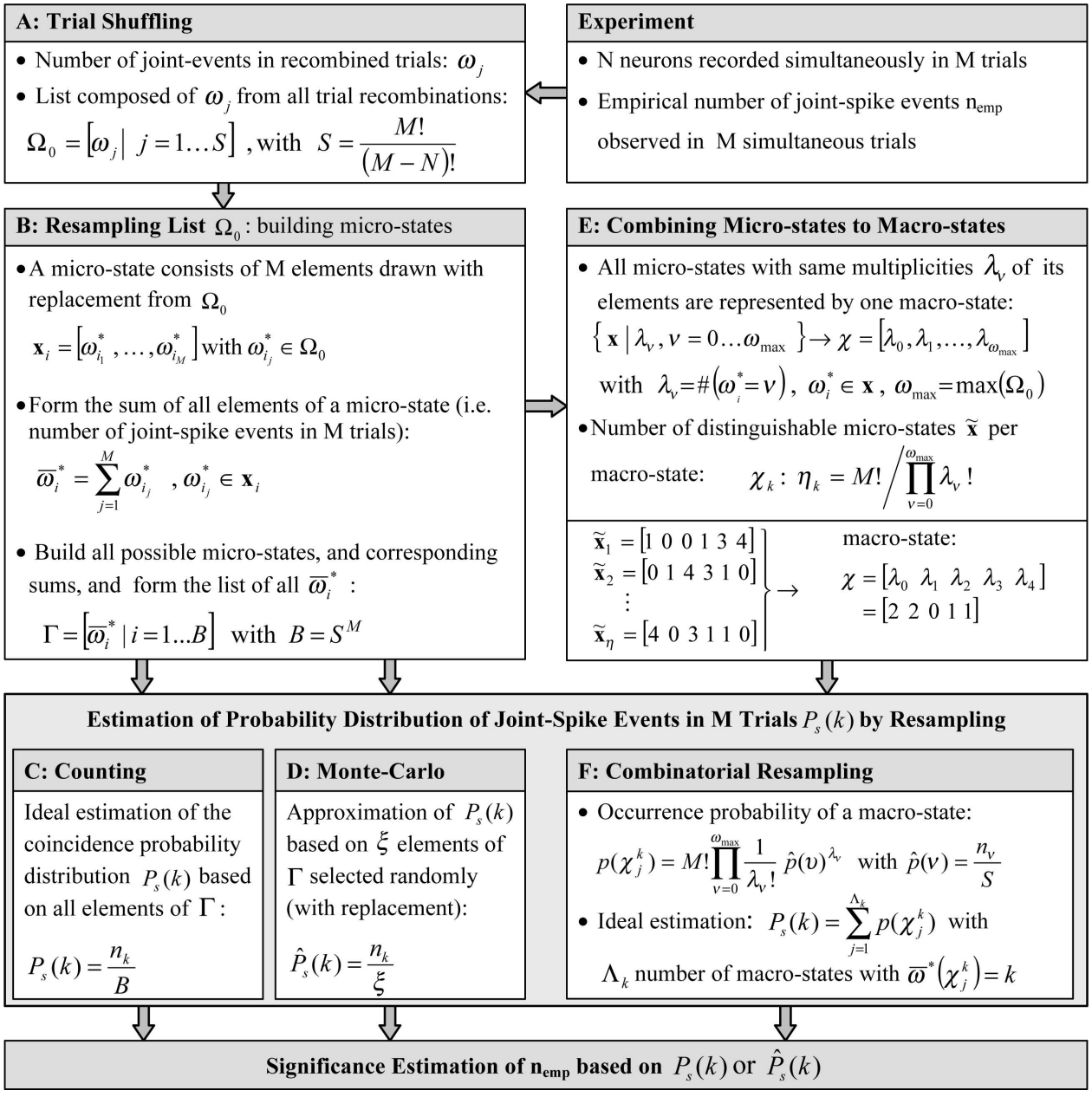
The most simple, and conceptually straightforward way of calculating P_s is to systematically resample (with replacement) *all* possible sublists $\mathbf{x}$ of Ω_0 with exactly M elements, replace each list $\mathbf{x}$ by the sum of its elements, and compute the relative frequencies of identical counts. In the following we introduce some definitions to describe this process more formally.

*Definition of Microstates (**x**)*

All possible nonunique lists, consisting of M elements from Ω_0 using replacement, form microstates:

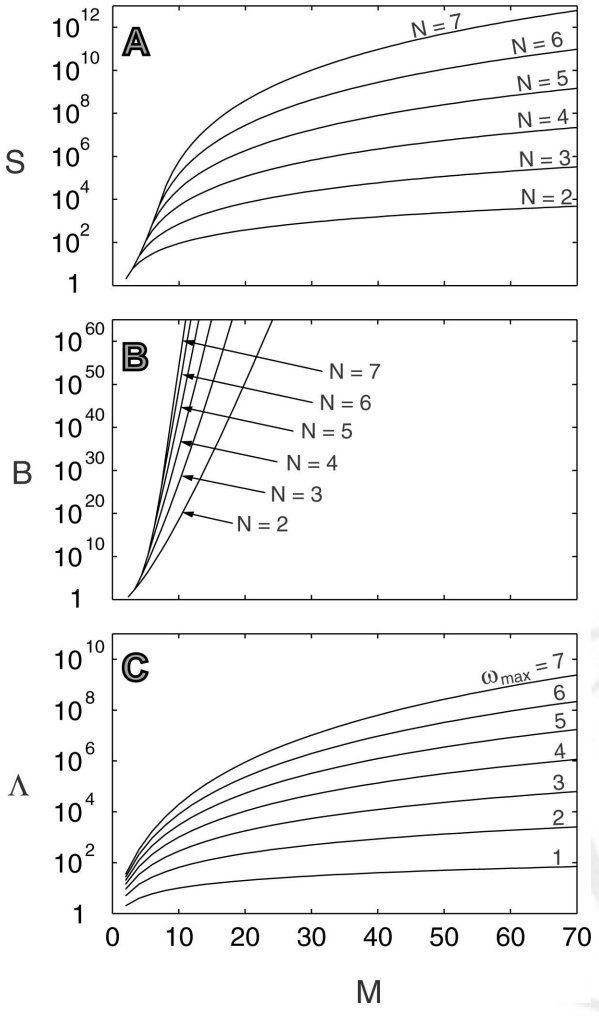
$$\mathbf{x}_i = [\omega_{i_1}^*, \ldots, \omega_{i_M}^*] \quad \text{with } \omega_{i_j}^* \in \Omega_0, \quad (4)$$

FIGURE 1



Flow chart illustrating the different procedures of combined shuffling and resampling (CSR) for significance estimation of joint-spike events. Spike train data from electrophysiological experiments enter the initial steps of trial shuffling (A) and resampling (B), common to all of the following three procedures (C, D, E). Ideal estimation of the coincidence probability distribution is achieved by the counting process (C) based on all possible *M* trial combinations (systematic resampling), and equally by the procedure of combinatorial resampling (F) using the concept of macrostates. (D) An approximative estimate of the distribution by Monte-Carlo resampling. The counting process (ideal estimation, C) as well as the Monte-Carlo approximation (D) operate directly on the list Γ containing the coincidence counts from all possible *M* trial combinations. In F, Γ is expressed in terms of macrostates, collective representations of a set of microstates with the same multiplicities λ_ν of the elements (coincidence counts) *v*. All three procedures lead to an estimation of the probability distribution P_s or $\hat{P}_s$, respectively, that reflects the H_0-hypothesis of independent parallel spike trains, which is used to estimate the significance of the empirically observed number of joint-spike events in *M* trials. Combinatorial resampling (F) avoids the need to access the large number of elements occurring in the definition of Γ (B) by combining microstates to macrostates thereby dramatically reducing the computational costs.

FIGURE 2



(A) Size S of the list Ω_0 that contains coincidence counts resulting from all possible single trial compositions. S is a function of the number of trials M and the number of neurons N. (B) Total number B of individual states (microstates) that can be composed of M elements from the list Ω_0. B increases dramatically with the number of trials M, and with the number N of neurons. (C) Total number of macrostates Λ that can be constructed from the list Ω_0, with $\Omega_0 = \lfloor \omega | \omega \in \{0, \ldots, \omega_{\max}\} \rfloor$. Λ increases with the number of trials M. In contrast to the list Ω_0 and to the list of possible microstates, the size of the set is independent of the number of elements in Ω_0 and of the number of neurons. The number of macrostates that can be constructed is dependent on the value of the maximal element $\omega_{\max}$ of Ω_0. Λ as a function of M is shown for $\omega_{\max} = 1, \ldots, 7$. The values are representative for independent data sets of about 100-ms duration with spike rates in the physiological regime.

where the asterisk labels measures based on resampling or boot-strapping.

Definition of List Γ

For a microstate $\mathbf{x}_i$ we denote the total number of joint-spike events as

$$\bar{\omega}_i^* = \sum_{j=1}^{M} \omega_{ij}^*, \ \omega_{ij}^* \in \mathbf{x}_i$$

Γ is then defined as the (nonordered and nonunique) list of all $\bar{\omega}_i^*$:

$$\Gamma = \lfloor \bar{\omega}_i^* | i = 1, \ldots, B \rfloor \quad \text{with } B = S^M. \tag{5}$$

The ideal estimate of the probability distribution of the total number of joint-spike events in M trials $P_s(\bar{\omega}^*)$ is then given by the normalized histogram of occurrence frequencies n_k of identical elements k in Γ [see Figure 1(c)]:

$$P_s(k) = \frac{n_k}{B}. \tag{6}$$

Unfortunately, aiming at a user-interactive analysis of experimental data, the feasibility of this approach is limited, because the length of Γ dramatically increases with increasing number of trials and/or neurons [Figure 2(B)] and thus is computationally impractical.

2.2. Approximative Estimation of the Probability Distribution

As an alternative approach to the systematic resampling illustrated above, the probability distribution (Eq. 6) can be estimated using a Monte-Carlo approach.[12] Instead of generating P_s based on all elements of Γ, a random sublist Γ_ξ of Γ consisting of $\xi \ll B$ elements is used to obtain an estimator $\hat{P}_s$ of P_s [see Figure 1(D)]. To avoid the approximate nature of Γ_ξ below, we present a procedure to directly compute the ideal estimator of P_s on the basis of Γ with the help of combinatorics.

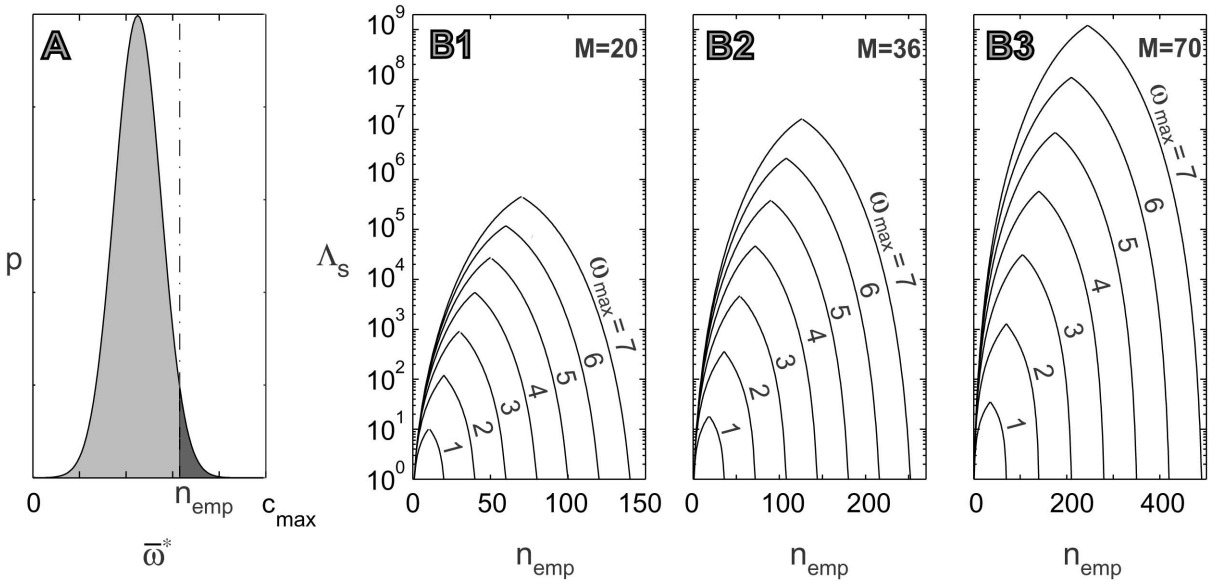
3. ANALYTICAL DERIVATION OF THE SIGNIFICANCE OF JOINT-SPIKE EVENTS BASED ON COMBINATORIAL RESAMPLING

In this section we describe how the combinatorial explosion underlying the ideal estimate of the probability distribution (section 2.1) can be considerably reduced by using a collective description: macrostates instead of microstates. First we introduce appropriate macrostates; then we derive P_s using the occupation probabilities of these macrostates.

3.1. Compact Description by Macrostates

Motivated by approaches used in statistical mechanics we intend to describe a list of microstates distinguished by a common property by a single macrostate. Obviously, the sum of coincidence counts from M trials $\bar{\omega}_i^*$ (Eq. 5) is independent of the arrangement of operands. With all list positions being indistinguishable we can collectively describe all microstates exhibiting the same sequence of coincidence

FIGURE 3



(A) The significance of the empirical number of joint-spike events n_{emp} is illustrated as the black area of the probability distribution for $\bar{\omega}^* \geq n_{emp}$. We can constrain the computational costs of calculating the significance by computing only the macrostates and their probability of occurrence of the corresponding tail of the distribution (black). Depending on the minimal distance of n_{emp} to the maximal or minimal possible number of coincidences bounding the distribution, either the macrostates for $\bar{\omega}^* < n_{emp}$ or for $\bar{\omega}^* \geq n_{emp}$ are computed to derive the significance. (B1–B3) The number of required macrostates Λ_s for the calculation of the significance depends on the number of trials (B1, $M = 20$; B2, $M = 36$; B3, $M = 70$) and on the maximal element ω_{max} in Ω_0. For a given parameter set M and ω_{max}, the number of required macrostates Λ_s shows a symmetrical dependence on n_{emp}: the closer n_{emp} to either the minimal or the maximal extend of the distribution, the less macrostates are required. The more central the position of n_{emp} to the distribution, the more macrostates are required. The larger ω_{max}, the larger the mean of the distribution and their extend, and thus the more macrostates need to be computed.

counts or a permutation of this sequence by a single macrostate χ_j [see Figure 1(E)]:

$$\{\mathbf{x}|\mathbf{x} = \lfloor \omega^*_{\eta 1(i1)}, \ldots, \omega^*_{\eta M(iM)} \rfloor\} \rightarrow \chi_{j(i1,\ldots,iM)} = [\lambda_0, \lambda_1, \ldots, \lambda_{\omega_{max}}], \quad (7)$$

with η denoting all possible permutations of the elements of a distinguished $\mathbf{x}_i$ and j an arbitrary bijective index function mapping the set represented by $\mathbf{x}_i$ to a unique macrostate χ_j. Because in most experimental conditions the maximal number of coincidences ω_{max} is smaller than the number of trials (M), a macrostate can be represented in a compact way by representing the elements $\omega \in 0, \ldots, \omega_{max}$ in a macrostate by their multiplicity (λ_ω). Consequently we get

$$\chi_k \neq \chi_l \quad \text{for } k \neq l \in 1, \ldots, \Lambda. \quad (8)$$

with Λ different macrostates.

Note that different macrostates χ_j as well as different microstates $\mathbf{x}_i$ may have an identical total number of coincidences $\bar{\omega}^*_i$. To simplify the notation we call $\mathbf{x}^k_i$ a microstate with $k = \bar{\omega}^*_i$ and correspondingly χ^k_j (with $j = 1, \ldots, \Lambda_k$) the Λ_k macrostates with identical associated $\bar{\omega}^*_i$. By using macrostates for the description instead of microstates the number of states that need to be considered for the construction of P_s reduces considerably from B elements in Γ to Λ different macrostates [Figure 2(C)]. A worst case estimate of the number of macrostates is dependent on the total number of trials and the maximal number of spikes n found in the recorded data set: $\Lambda = (M + 1)^{(n+1)}$. The maximal number of spikes n (e.g., a function of the firing rate, the total duration of the considered data set, time resolution of the data) gives an upper limit for the coincidence count within a trial, i.e., n.

3.2. Probability of Macrostates

In order to use macrostates for deriving P_s, we calculate the occurrence probability of an individual macrostate on the basis of the occurrence probabilities of the associated microstates. The probability to observe a microstate $\mathbf{x}^k_i$ simply is $1/S^M$. However, many microstates $\mathbf{x}^k_i$ may exhibit the identical sequence of coincidence counts because of the limited number of different coincidence counts in Ω_0. Microstates are defined by the identity of the single neuron trials they are composed of (Eq. 1), not by the sequence of

coincidence counts. Let us denote the collection of microstates with identical sequence by $\tilde{\mathbf{x}}_i^k$ to remove this ambiguity. The probability to observe $\tilde{\mathbf{x}}_i^k$ is given by the product of the occurrence probabilities $p(\upsilon)$ of the elements $\upsilon(\upsilon \in \Omega_0)$ composing the list $\tilde{\mathbf{x}}_i^k$. The $p(\upsilon)$ can be estimated from the number of occurrences n_υ in Ω_0 normalized to the total number of elements in Ω_0, i.e., $\hat{p}(\upsilon) = n_\upsilon / S$. Hence,

$$p(\tilde{\mathbf{x}}_i^k) = \prod_{l=1}^{M} \hat{p}(\upsilon_{il}) = \prod_{l=1}^{m_i} \hat{p}(\upsilon_{il})^{\lambda(\upsilon_{il})}, \tag{9}$$

with $\{\upsilon_{il}\}$ the set of unique elements in $\tilde{\mathbf{x}}_i^k$ with multiplicities λ_l and $m_i = |\{\upsilon_{il}\}|$.

For deriving the occurrence probability of the corresponding macrostate, the possible number of permutations of $\tilde{\mathbf{x}}_i^k$ (i.e., distinguishable microstates), given by the multinomial coefficient, has to be taken into account:

$$p(\chi_j^k) = M! \prod_{l=1}^{m_i} \frac{1}{\lambda(\upsilon_{il})!}\, p(\tilde{\mathbf{x}}_i^k). \tag{10}$$

Finally, the probability to obtain k coincidences $P_s(k)$ is given by the sum of the occurrence probabilities of all individual macrostates (Eq. 10) that have identical sums k:

$$P_s(k) = \sum_{j=1}^{\Lambda_k} p(\chi_j^k), \tag{11}$$

with Λ_k the number of different macrostates with identical k [summarized in Figure 1(F)]. Again, we are in the position to calculate the significance of an empirical number of joint-spike events using Equation 3. However, now this does not require the construction of all B microstates but restricts the computation to only Λ macrostates [see Figure 2(B,C) for comparison].

3.3. Significance Estimation

The computational effort of computing the joint p-value (Eq. 3) can further be reduced by only considering the macrostates forming the respective tail of the distribution [e.g., Figure 3(A)]. Typically, the relevant tail of the distribution is not known beforehand. However, we can make use of the following constraints. First, we calculate the minimal $[M \cdot \omega_{\min} = M \cdot \min(\Omega_0)]$ and the maximal $[M \cdot \omega_{\max} = M \cdot \max(\Omega_0)]$ number of coincidences that bound P_s. Second, the relevant tail of the distribution is determined by the minimal distance of n_{emp} to the two bounds. Thus, Λ_s macrostates are required for the significance test $p(\bar{\omega}^* \geq n_{\text{emp}})$:

$$\Lambda_s = \min(\Lambda_{1s}, \Lambda_{2s}) \quad \text{with} \quad \Lambda_{1s} = \sum_{k=n_{\text{emp}}}^{M\cdot\omega_{\max}} \Lambda_k;\ \Lambda_{2s} = \sum_{k=M\cdot\omega_{\min}}^{n_{\text{emp}}-1} \Lambda_k \tag{12}$$

$$\Lambda_{k=0} = 1 \quad \text{and} \quad \omega_{\max} = \max(\Omega_0). \tag{13}$$

Obviously, the number of elements Λ_s required to compute the statistical significance depends on the empirical number of coincidences n_{emp}, the number of trials M, and the maximum value $\omega_{\max}$ in Ω_0, as illustrated in Figure 3(B1–B3), for the case of $\omega_{\min} = 0$. Again, the number of elements that need to be computed reduces compared with the total number of macrostates Λ [Figure 2(C)].

4. EFFECTIVE IMPLEMENTATION OF THE SIGNIFICANCE TEST

4.1. Limited Partitions

The basic idea of a computationally effective implementation of the above described scheme is to use efficient standard algorithms to compute all macrostates associated with a total number k of joint-spike events. Lists of integers that have identical sums remind us on the concept of "partitions"[14,15] [see Figure 4(G)]. A partition of an integer k is defined as a list of m strictly positive integers that sum up to k. The elements of a partition are composed of $\nu_l \in 1, \ldots, k$ with replacement. The element 0 is excluded and individual partitions of the same integer in general do not have the same number of elements. F4

In order to compute macrostates χ_j by making use of the concept of partitions, we define "limited partitions" that consider additional constraints given by our data sets. First, the number of trials M determines the number of elements in a partition to $m \leq M$. Second, the set Q_0, i.e., the unique set of elements in Ω_0, from which we form partitions is limited in the sense that Q_0 does not necessarily contain all elements $1, \ldots, k$. In addition, we need to consider that the set Q_0 may contain the zero element, which is excluded from partitions by definition. Thus, in order to match macrostates χ_j with M elements onto partitions with m_j operands, $m_j^0 = M - m_j$ zeros are being added. Trivially, Equations 9 and 11 still hold.

Based on the concept of partitions, we now have a tool at hand to compute the above introduced macrostates. Because partitions can be computed for individual k, we are able to generate macrostates to derive the complete probability distribution P_s (Eq. 11) or to restrict the derivation of macrostates to the ones that are needed for the relevant tail. Efficient but computationally still demanding algorithms, to *sequentially* generate all existing partitions of an integer k do exist.[14,15]

4.2. Look-up Tables for Macrostates

Instead of computing the partitions online, one can still save computation time during user interactive application by precomputing the required limit partitions. A single look-up table contains the data for one limited partition,

FIGURE 4

G: Partition of k: $\mathrm{P}(k)$

- Def: list of m strictly positive integers which sum up to k
- A partition is composed of non-unique elements $v_l \in 1 \ldots k$

e.g. $\mathrm{P}_1(5)=5$ $\mathrm{P}_3(5)=3+2$ $\mathrm{P}_5(5)=2+2+1$ $\mathrm{P}_5(5)=1+1+1+1+1$
$\mathrm{P}_2(5)=4+1$ $\mathrm{P}_4(5)=3+1+1$ $\mathrm{P}_6(5)=2+1+1+1$

H: Limited Partition of k $\mathrm{P}^{M,\omega_{\max}}(k)$**:** restricted in M and $\omega_{\max}$

- Number of elements m in each limited partition are restricted to $1 \le m \le M$
- A limited partition is composed of non-unique elements $v_l \in 1 \ldots \omega_{\max}$

e.g. $\mathrm{P}_1^{M=3,\omega_{\max}=4}(5)=4+1$ $\mathrm{P}_3^{3,4}(5)=3+1+1$
$\mathrm{P}_2^{M=3,\omega_{\max}=4}(5)=3+2$ $\mathrm{P}_4^{3,4}(5)=2+2+1$

I: Look-up Table: multiplicity of elements in limited partition $\mathrm{P}^{M^L,\omega_{\max}^L}(k)$

- Multiplicity of element l in $\mathrm{P}^{M^L,\omega_{\max}^L}(k)$ is λ_l with $\lambda_l \in 1 \ldots M^L$ and $l = 1 \ldots \omega_{\max}^L$
- Look-up table consists of the multiplicities of all J existing limited partitions of k:

$$\chi_j^{k,M^L,\omega_{\max}^L} = \left[m \;\; \lambda_1 \;\; \ldots \;\; \lambda_{\omega_{\max}^L} \right] \text{ with } \mathrm{m} = \sum_{l=1}^{\omega_{\max}^L} \lambda_l \text{ and } j = 1 \ldots J$$

- To map $\chi_j^{k,M^L,\omega_{\max}^L}$ on macro-states based on M trials: add $\lambda_0 = M - m$ zeros

Flow chart to illustrate the computation of macrostates based on the concept of partitions and limited partitions. (G) Concept of a partition of an integer k, here demonstrated for $k = 5$. There are 7 different ways of computing the sums of integers that lead to a total of 5 given the elements 1, . . . , k. (H) Concept of the limited partition of an integer k constrained by the number of allowed operands M and the upper boundary $\omega_{\max}$ for each operand. Because in the example the number of trials is limited to 3, only partitions with 3 elements are allowed to be formed. In addition, the values of the elements are limited to $\omega_{\max} = 4$ here, and thus, e.g., the partition P_1 (5) from box G is excluded. (I) illustrates the mapping of limited partitions to macrostates and the constraints for precomputed macrostates given by the maximal number of operands per partition $M^L = M$ and the maximal value $\omega_{\max}^L$. Compared with the concept of partitions, zero elements have to be added to ensure the existence of M elements.

with parameters number of trials M and maximal number of coincidences $\omega_{\max}$ in the shuffled trials. The number of trials constrains the maximal number of operands per partition, i.e., $M^L = M$. The maximal value $\omega_{\max}$ in Q_0 constrains the maximum value $\omega_{\max}^L$ to be taken into account in the limited partition. The limited partition is formed from operands in the range of $Q_0^L = 1, \ldots, \omega_{\max}^L$, with L indicating the parameters for a limited partition [see also Figure 4(H)]. Thus, to obtain the complete probability distribution P_s, we have to precompute all limited partitions $P(k)^{M^L,\omega_{\max}^L}$ for $k = 0, \ldots, M \cdot \omega_{\max}^L$. For a specific data set, we can use the corresponding look-up table (for a given M^L and $\omega_{\max}^L$) restricted to the multiplicities available. This also implies that Q_0^L does not have to correspond to the experimentally observed Q_0. In case $Q_0 \neq Q_0^L$ and $\omega_{\max} \le \omega_{\max}^L$, the Equations 10, 11, and 12 still hold with probability $\hat{p}(v) = 0$ because of $v \notin Q_0$. In case of $\omega_{\max}^L < \omega_{\max}$, the precomputed look-up tables do not include all required macrostates, and the additionally required macrostates have to be computed online. The latter usually is computationally too demanding for an interactive analysis and needs to be avoided, by choosing $\omega_{\max}^L$ quite conservatively at the time the look-up tables are generated.

For a rough estimate of the probability that $\omega_{\max}^L < \omega_{\max}$, one may assume that the spike trains can be described by independent Poisson processes, leading to a description of the coincidence distribution by a Poisson distribution parameterized by the expected number of joint-spike events. The expectation value in turn in determined by the length of the analysis window, the spike rates, and the number of observed neurons.[3] Given the distribution, we can compute the probability to observe a value $\omega_{\max}$ greater than an available $\omega_{\max}^L$.

5. DISCUSSION

The significance test in UE analysis can be formulated using a nonparametric approach based on trial shuffling and subsequent resampling of the coincidence counts.[12] However, as shown here, the combinatorics of all possible single trial compositions increases dramatically with the number of neurons and the number of trials. This contravenes our goal of a user-interactive usability of the UE analysis, performed in time-resolved fashion to investigate the dynamics of cortical processes. As one possible solution we suggested[12] a Monte-Carlo resampling procedure. The method allows for reliability tests of the significance estimation and is well calibrated.

In this article we derived an effective procedure to calculate an *exact* significance estimate for joint-spike events in the context of combined trial shuffling and resampling (CSR). We have shown that the required number of terms can considerably be reduced by reformulating the problem in terms of certain macrostates instead of microstates. Further restrictions can be made by limiting the calculation to the subset of macrostates, which comprise the area of the coincidence distribution relevant for significance estimation. Finally, for the implementation of the calculation of macrostates, we make use of the concept of partitions. By defining limiting partitions we are able to adjust our problem to algorithms for the sequential computation of partitions available in the literature.[14,15] Usage of look-up tables for the limited partitions helps to further speed up the computation. Computationally a macrostate is about twice as expensive as a microstate in the Monte-Carlo approach. As shown in[12] in the order of 10,000 Monte-Carlo steps (microstates) are appropriate to estimate P_s for a test level of 5%. For more strict levels, the number of steps needs to be increased. Thus, the exact combinatorial method has comparable costs [Fig. 3(B)]. The main advantage is its ideal nonapproximative nature, becoming more relevant at strict test levels.

Because of the combination of the improvements presented here we are now able to apply the nonparametric significance test based on the exact coincidence distribution in user-interactive manner.

Acknowledgments

We thank Stefan Rotter and an anonymous referee for helpful comments on an earlier version of the manuscript. This work was supported in part by the Körber Foundation and by the Volkswagen Foundation.

REFERENCES

1. Grün, S.; Diesmann, M.; Aertsen, A. 'Unitary Events' in multiple single-neuron activity. I. Detection and significance. Neural Comput 2002, 14(1), 43–80.
2. Grün, S.; Diesmann, M.; Aertsen, A. 'Unitary Events' in multiple single-neuron activity. II. Non-stationary data. Neural Comput 2002, 14(1), 81–119.
3. Pipa, G.; Grün, S. Non-parametric significance estimation of joint-spike events by shuffling and resampling. Neurocomputing, forthcoming. AQ: 1
4. Hebb, D.O. Organization of Behavior. A Neurophysiological Theory; John Wiley & Sons: New York, 1949.
5. Singer, W.; Engel, A.K.; Kreiter, A.K.; Munk, M.H.J.; Neuenschwander, S.; Roelfsema, P.R. Neuronal assemblies: necessity, signature and detectability. Trends Cog Sci 1997, 1(7), 252–261.
6. Riehle, A.; Grün, S.; Diesmann, M.; Aertsen, A. Spike synchronization and rate modulation differentially involved in motor cortical function. Science 1997, 278, 1950–1953.
7. Grammont, F.; Riehle, A. Precise spike synchronization in monkey motor cortex involved in preparation for movement. Exp Brain Res 1999, 128, 118–122.
8. Riehle, A.; Grammont, F.; Diesmann, M.; Grün, S. Dynamical changes and temporal precision of synchronized spiking activity in motor cortex during movement preparation. J Physiol (Paris) 2000, 94(5–6), 569–582.
9. Grün, S.; Riehle, A.; Diesmann, M. Effect of cross-trial non-stationarity on joint-spike events. Biol Cybern, forthcoming. AQ: 1
10. Riech, D.S.; Mechler, F.; Purpura, K.P.; Victor, J.D. Interspike intervals, receptive fields, and information encoding in primary visual cortex. J Neurosci 2000, 20(5), 1964–1974.
11. Baddeley, R.; Abbott, L.F.; Booth, M.C.A.; Sengpiel, F.; Freeman, T.; Wakeman, E.A.; Rolls, E. Responses of neurons in primary and inferior temporal visual cortices to natural scenes. Proc R Soc Lond B 1997, 264, 1775–1783.
12. Oram, M.W.; Wiener, M.C.; Lestienne, R.; Richmond, B.J. Stochastic nature of precisely timed spike patterns in visual system neuronal responses. J Neurophysiol 1999, 81, 3021–3033.
13. Efron, B.; Tibshirani, R.J. An Introduction to the Bootstrap; Chapmann and Hall/CRC: London, 1998.
14. Nijnhuis, A.; Wilf, H.S. Combinatorial Algorithms: For Computers and Calculators, 2nd Ed.; Academic Press: New York, 1978.
15. Skiena, S.S. Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica; Addison-Wesley: Redwood City, CA, 1990.

AQ1 Update.	