A pattern grouping algorithm for analysis of spatiotemporal patterns in neuronal spike trains. 1. Detection of repeated patterns

Igor V. Tetko a,b, Alessandro E.P. Villa a,*

a Laboratoire de Neuro-heuristique, Institut de Physiologie, Université de Lausanne, Rue du Bugnon 7, CH-1005 Lausanne, Switzerland
b Department of Physiology of Brain, Bogomoletz Institute of Physiology and Department of Biomedical Applications, IBPC, National Ukrainian Academy of Sciences, Kyiv, Ukraine

Received 14 July 2000; received in revised form 27 September 2000; accepted 29 September 2000

Abstract

The existence of precise temporal relations in sequences of spike intervals, referred to as ‘spatiotemporal patterns’, is suggested by brain theories that emphasize the role of temporal coding. Specific analytical methods able to assess the significance of such patterned activity are extremely important to establish its function for information processing in the brain. This study proposes a new method called ‘pattern grouping algorithm’ (PGA), designed to identify and evaluate the statistical significance of patterns which differ from each other by a defined and small jitter in spike timing of the order of few ms. The algorithm performs a pre-selection of template patterns with a fast computational approach, optimizes the jitter for each spike in the template and evaluates the statistical significance of the pattern group using three complementary statistical approaches. Simulated data sets characterized by various types of known non stationarities are used for validation of PGA and for comparison of its performance to other methods. Applications of PGA to experimental data sets of simultaneously recorded spike trains are described in a companion paper (Tetko IV, Villa AEP. A pattern grouping algorithm for analysis of spatiotemporal patterns in neuronal spike trains. 2. Application to simultaneous single unit recordings. J Neurosci Methods 2000; accompanying article). © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Temporal firing patterns; Simulated spike trains; Neural code; Synchronization; Synfire chain

1. Introduction

Multiple dimensions of behaviorally relevant stimuli are processed by thousands of neurons distributed over many areas of the brain. This processing is reflected by changes in firing rate and in the temporal structure of the spike trains — the time series formed by the sequences of time intervals between spikes. There is an active debate about neural coding theories (Georgopoulos, 1995; de Ruyter van Steveninck et al., 1997; Deadwyler and Hampson, 1997; Fujii et al., 1997; Eggermont, 1998; Singer, 1999) and Abeles’ synfire chain theory (Abeles, 1982, 1991) has received particular attention during the last decade (MacGregor et al., 1995; Arnoldi and Brauer, 1996; Miller, 1996; Postma et al., 1996; Arnoldi et al., 1999).

It has been hypothesized that a stronger synaptic influence is exerted by multiple converging neurons firing in coincidence (Abeles, 1982). To achieve this synchrony of firing would require a high degree of temporal precision in spike trains distributed across many neurons. In cell assemblies exhibiting such temporal precision, it has been predicted (Abeles, 1982; Abeles et al., 1993a,c) that if recordings are made from a random selection of neurons from within the network, then some ordered sequences of spike intervals, referred to as ‘spatiotemporal patterns’ of discharges, will recur. The development of analytical methods to reliably detect transient temporal relations in sequences of spike intervals therefore represents a critical tool for the analysis of functional temporal coding in neural networks.

A number of methods have been developed in the past to estimate the significance of patterns of spikes detected in real brains during experimental tests (Day-

Hoff and Gerstein, 1983; Abeles and Gerstein, 1988; Frostig et al., 1990b; Prut et al., 1998). However, these methods were partly characterized by limitations to provide separate estimates of significance for individual patterns and by their inaccuracy to detect significant patterns of spikes with variable jitters. This study proposes a new method called ‘pattern grouping algorithm’ (PGA), designed to identify and evaluate the statistical significance of temporal patterns of spikes formed by three or more different events with slight differences in spike timing. In this paper we describe the new method and its validation by applying it to simulated spike trains characterized by various types of known non-stationarities. The companion article (Tetko and Villa, 2000) provides applications of PGA algorithm to simultaneously recorded spike trains in animal experiments and in a large network of simulated spiking neurons.

2. Materials and methods

The pattern grouping algorithm described here is based on the pattern detection algorithm, PDA of Abeles and Gerstein (1988). This original method is formed by two parts: firstly, the detection, i.e. the algorithm finds all single or multineuron patterns that repeat two or more times within a block of data (Fig. 1a); secondly, the algorithm computes how many of these patterns can be expected by chance and provides the significance level of such estimation (Fig. 1b). A common way to plot the pattern occurrences is provided by a raster display of the spike trains of the simultaneously recorded neurons. Each line of the raster corresponds to the occurrence of patterned activity and the rasters are aligned by displaying the first spike in the pattern at time zero (Fig. 1c). The PDA algorithm estimates the excess of total detected patterns (here 0.04 patterns were expected and two were found, that is significant at level $P < 0.01$) but it does not answer the question which individual patterns are significant.

The PGA includes both steps indicated above, but in addition it groups together patterns with slight differences (‘jitter’) in spike timing, i.e. following the ideas introduced by Dayhoff and Gerstein (1983). The grouping operation is based on an optimization of the allowed jitter for spike intervals in patterns that become united into the same group. This optimization applies to each interval within each pattern, so the allowed jitters for the first, second, etc. intervals may be different from each other. By grouping the detected patterns in this way it becomes possible to estimate the significance of a group and to select the most salient groups. The final output of PGA is one or many pattern groups. Each pattern group consists of all the individual patterns and corresponds to a template pattern with optimized maximal jitters for each interspike interval.

There are three adjustable parameters in PGA: the maximal duration of the pattern, $w$, measured as a delay between the first and the last spike in the sequence of spikes (i.e. the window duration), the level of significance to be used for detection of significant groups and the upper bound of allowed jitter. In the next subsections the symbols and formulae follow notations used elsewhere (Abeles and Gerstein, 1988; Tetko and Villa, 1997b).

2.1. Estimation of patterns with variable time accuracy

Let the list of cell labels that appear in a pattern of $c$ spikes be noted as $S_c = (i_1, ..., i_j, ..., i_c)$. Here any label

Fig. 1. Outline of the general procedure followed by pattern detection algorithms. (a) Analysis of a set of simultaneously recorded spike trains. Three cells, labeled A, B, and C, participate to a patterned activity. Three occurrences of two precise patterns are detected. Each occurrence of the first pattern has been labeled by a specific marker in order to help the reader to identify the corresponding spikes. The spikes belonging to the second pattern are indicated by arrows. (b) Estimation of the statistical significance of the detected patterns. Two patterns, $\langle A, C, B; \Delta t_1, \Delta t_2 \rangle$ and $\langle C, C, C; \Delta t_1, \Delta t_2 \rangle$, were found. The expected number of patterns of this complexity and repetition number was $N_{c} = 0.04$. The probability to observe 2 or more patterns when 0.04 patterns are expected is noted as $p_{null} = 0.02$. (c) Display of the pattern occurrences as a raster plot aligned on the pattern start.
of the recorded neurons can be assigned to \( i_j \). In particular, for a 3-tuple pattern of spikes (triplet) obtained from a simultaneous recording of ten spike trains, let us assume that we find \( S_3 = (3, 5, 4) \), meaning all patterns of complexity \( c = 3 \) formed by a spike recorded from cell no. 3 followed by a spike recorded from cell no. 5, which in turn is followed by a spike recorded from cell no. 4. Note that that the term complexity is used here to refer to the total number of spikes in a pattern. When referring to one specific pattern involving a particular set of intervals between these cells, the cells involved, the template interval and the optimized time jitters are noted within ‘(’ and ‘)’ symbols. For example, a pattern involving cells nos. 3, 5 and 4 with delays of 299 ms between cell nos. 3 and 5 with a time jitter equal to 1 ms (i.e. \( \pm 0.5 \) ms) and 449 ms delays between cell nos. 3 and 4 and with the time jitter equal to 3 ms (i.e. \( \pm 1.5 \) ms) is noted as \( \langle 3, 5, 4; 299 \pm 0.5, 449 \pm 1.5 \rangle \) (see Fig. 2).

We assume in this study that the minimal jitter (i.e. the maximum accuracy) of the recorded spike trains is equal to the resolution of the data acquisition device used for recording spikes, and is noted as \( \Delta_{\text{exp}} \). This value is usually equal to 1 ms. The general form of a pattern of complexity \( c \) can be noted as \( \langle S_c; t_1 \pm (\Delta t)/2, \ldots, t_j \pm (\Delta t)/2, \ldots, t_{c-1} \pm (\Delta t)/2 \rangle \) where \( t_1, \ldots, t_j, \ldots, t_{c-1} \) are the time delays between the first and the \( j \)th spike forming the pattern and \( \Delta t \) are the corresponding time jitters. The pattern is simply noted as \( \langle S_c; t_1, \ldots, t_j, \ldots, t_{c-1} \rangle \) if all jitters are equal, e.g. to the maximum accuracy \( \Delta_{\text{exp}} \). We will refer to such patterns as precise patterns. Patterns of such kind are detected by traditional application of PDA. The terms pattern group or pattern with variable jitters will be used whenever the jitters related to the pattern of spikes are unequal, thus assuming that such pattern is formed by a combination of patterns detected by PDA.

A restriction exists on the maximal duration of the pattern, \( w \), that is in practice may be due to epoch duration in behavioral tasks. The expected number of patterns \( N^{(c)} \) of complexity \( c \) that repeated \( r \) times is expressed according to the original PDA, (see Eq. A6 in Abeles and Gerstein, 1988) by:

\[
N^{(c)}(f_{\Delta}) = K \sum_{\Delta \in \zeta_c} e^{-w_{\Delta c}} \frac{x(S_c)^r}{r!}
\]

\[
K = \frac{(w/\Delta)^{r-1}}{(c-1)!} = \frac{w^{c-1}}{t_{\Delta c}(c-1)!}
\]

where \( x(S_c) \) is the expected number of repetitions of a sequence \( S_c \), \( \zeta_c \) is the set of all possible \( S_c \), \( \Delta \) is the timing accuracy (time jitter), \( K \) is the maximum number of different time delays and \( f_{\Delta} = \prod_{j-1,c-1} \Delta_j = \Delta^{c-1} \) corresponds to the product of time jitters around the spike intervals \( j = 1, \ldots, c-1 \) forming the pattern. The calculations of the expected number of repetitions \( x(S_c) \) account for the duration of the analyzed record (Abeles and Gerstein, 1988).

The patterns detected by PDA are characterized by equal time accuracy, i.e. \( \Delta = \Delta_j = \Delta_{\text{exp}} \). Usage of a larger jitter \( \Delta_j \), in the order of several ms, could be more appropriate for studying recurrent patterns of firing in reverberating neural circuits (Abeles et al., 1993a). In such a case it is possible to apply the PDA by dividing

\[\langle 3,5,4; 299\pm0.5,449\pm1.5 \rangle\]
all timing information by jitter $J$ and to use the above Eq. (1), given $\Delta_j = J$. Let us note that according to Abeles and Gerstein (1988) validity of Eq. (1) does not rest on the assumption that the spike trains behave like Poisson processes, but just on the assumption that at any time instance $t$ the probability of getting one pre-specified pattern is very low. The second assumption used in the deriving of this equation is that the number of repetitions $x(S_c)$ is not depending from the possible different time delay combinations of $S_c$. Thus, the authors assumed that the Poisson approximation holds for the occurrences of patterns, not necessarily for the spike occurrences themselves. The same assumptions are used in the current study. These assumptions are generally true for data recorded from cortical neurons that are characterized by relatively low firing rate of about several spikes per second, except for spikes occurring in a burst. In the current study the spikes within a burst were filtered out on the basis of the comparison of interspike intervals with those expected for a spike train following a Poisson distribution with $f = 400$ spikes/s, as described elsewhere (Tetko and Villa, 1997b). The filtering was repeated until no more than 1% of spikes were deleted after one run of the filter. In order to avoid patterns formed by spikes belonging to high frequency bursts, patterns lasting less than 20 ms were not considered.

Application of Eq. (1) requires calculation of the value $x(S_c)$. The ad-hoc method (Abeles and Gerstein, 1988) estimates this value as $x(S_c) \approx m(S_c)/K$, where $m(S_c)$ is the total number of cases in which the spike sequence $S_c$ appeared in the analyzed spike train with all possible delays less than the maximally allowed window duration $w$.

The value $m(S_c)$ does not depend on the value of the time jitter if the total number of spikes used for its computation does not change. However, because of the filtering procedure, the number of spikes that are filtered out tends to increase as the value of jitter increases so that $m(S_c)$ and $N^{(\xi)}_v$ tend to decrease. Then, according to Eq. (1) the value $m(S_c)$ calculated for maximal accuracy ($\Delta_{exp} = 1$ ms) provides an overestimation of $N^{(\xi)}_v$ for any analysis performed at lower time resolution. The faster and improved estimation of the expected number of patterns can be done according to the combinatorial methods (Tetko and Villa, 1997a,b).

The PDA uses constant jitter $\Delta$ for all intervals in a pattern, but in Eq. (1) only the product of time jitters $f_a$ is important for the calculation. Hence, Eq. (1) can be used also for the estimation of the expected number of patterns with variable accuracy of time jitters. Unfortunately, the correct estimation of $m(S_c)$ for such patterns is more difficult. A criticism often reported about interpretation of complex patterns is that they may occur at random and represent a statistical artifact (Oram et al., 1999). Thus, for practical application of the algorithm it is preferable to keep a conservative attitude and to underestimate the pattern significance. Therefore, we estimated the significance of a pattern group using the upper bound of the value $m(S_c)$ calculated for maximal time accuracy. Such procedure provides an overestimation compared to analysis using fixed jitter $J$ (Abeles et al., 1993a), in which the values $m(S_c)$ is based on the number of spikes after the jittering procedure.

The summation in Eq. (1) is done over all possible combinations of all recorded spikes. This is necessary if one is interested in the computation of the excess of the total number of detected patterns and to compare it over the count of patterns occurring by chance (Abeles and Gerstein, 1988). In the present study, we are interested only in the significance of specific patterns formed by the selected neurons. Thus, in the PGA we restricted the summation in Eq. (1) to the combinations of patterns formed solely by the spikes of the selected neurons and calculated $N^{(\xi)}_v$ as:

$$
\begin{align*}
N^{(\xi)}_v = & Q(a, c, \Omega) \cdot \tilde{N}^{(\xi)}_v, \\
\tilde{N}^{(\xi)}_v = & K \sum_{S_c} e^{-m(S_c)/K} \frac{1}{f_a} \left( \frac{m(S_c)}{K} \right)^r \\
Q(a, c, \Omega) = & F(c - a, a) \cdot \frac{\Omega!}{(\Omega - a)! a!} \\
K = & \left( \frac{w}{\Delta} \right)^{-1} \frac{(c - 1)!}{f_a(c - 1)!}.
\end{align*}
$$

(2)

where $\xi_v \in \xi_c$ is a subset formed by the analyzed neurons with all possible delays between these neurons, $\tilde{N}^{(\xi)}_v$ estimates the expected number for the subset $\xi_v$ and $Q$ corresponds to the number of different subsets $\xi_v$ that can be formed from the pool of $\Omega$ neurons recorded simultaneously. The Fibonacci (or Figurate series) number $F(i, j)$ corresponds to the number of different combinations formed by $i$ out of $j$ neurons, including repetitions of the same neuron (Hogben, 1950). The number of subsets $\xi_v$ is different for different patterns and it depends on how many different cells, $a$, participate to the analyzed pattern.

The number of subsets formed by one cell only (i.e. $\langle 1, 1, 1 \rangle$, $\langle 1, 2, 2 \rangle$, $\langle 1, 1, 1, 1 \rangle$, etc.) does not depend on the complexity of an analyzed pattern and is equal to the number of simultaneously recorded cells, i.e. $Q(1, c, \Omega) = \Omega$. The number $Q$, however, dramatically increases if both complexity and number of different cells are increased, i.e. $Q(2, 3, 10) = 90$, $Q(3, 4, 10) = 360$, etc. Let us consider the example of a pattern $\langle 1, 1, 2; t_1, t_2 \rangle$ detected within a record of $\Omega = 15$ spike trains. The set $\xi_v$ comprises the list $\langle 1, 1, 2 \rangle$ with all possible delays among these neurons. The patterns $\langle 1, 1, 1 \rangle$, $\langle 2, 2, 2 \rangle$, $\langle 1, 2, 2 \rangle$, $\langle 1, 1, 3 \rangle$, $\langle 1, 1, 4 \rangle$, .... $\langle 14, 15, 15 \rangle$, $\langle 15, 15, 15 \rangle$ are not considered to estimate $N^{(\xi)}_v$ because
any such combination of neurons does not change the probability to detect the analyzed pattern. The total number of different subsets of complexity $c = 3$ formed by $a = 2$ different neurons is $Q(2, 3, 15) = 210$. Then, the expected number of patterns $N^{(r)}$ is based on the actual count of detected patterns in a given record. This estimation accounts for the chance of false detection that increases with the total number of analyzed combinations. However, $N^{(r)}$ is not influenced by the firing rates of other simultaneously recorded neurons, some of which could be characterized by high firing rates and could mask significant patterns formed by spikes belonging to cells firing at low frequency.

### 2.2. Estimation of the significance of a single pattern

Application of PDA (Abeles and Gerstein, 1988) is aimed to detect if there is an excess of precise patterns over the number expected by chance. Assuming a Poisson distribution of the point processes, the significance of the excess of detected patterns $Z$ of complexity $c$ and repetition number $r$ over the expected number $X = N^{(r)}$ (one tail $P$-value) is given by the cumulative Poisson probability function (Cox and Isham, 1980; Press et al., 1994)

$$p = pr\{Z,X\} = \gamma(Z,X) \equiv \frac{1}{(Z-1)!} \int_0^X e^{-t^Z-1} dt$$

(3)

This equation refers to probability of seeing $Z$ or more patterns when $X$ is expected and it is reduced to

$$pr\{Z, X\} = 1 - e^{-X}$$

(4)

if only one such pattern is detected. Let us fix the level of significance to be $p_0 = 0.01$ and consider three spike trains recorded simultaneously. For example, if the number of patterns detected by PDA is $Z = 2$ and only 0.04 patterns (i.e. $X = 0.04$) are expected by chance, then Eq. (3) indicates a significant excess of detected patterns, i.e. $pr\{2, 0.04\} = 0.006 < p_0$ (Fig. 1). However, this estimation cannot indicate which ones out of the two detected patterns are significant at the considered level of significance. Indeed, the probability to detect each pattern by chance is $pr\{1, 0.04\} = 0.04 > p_0$. Patterns of very different structure (i.e. formed by one or by several cells) are contributing to $m(S_s)$ as long as they have the same complexity (number of spikes) and are found the same number of times within the record.

Consider a case where the expected number of patterns by PDA is as low as $X = 0.01$. If one pattern is found, the probability to detect it by chance is equal to $pr\{1, 0.01\} = 0.01$, thus indicating that the pattern is significant. If the algorithm detected two patterns instead of one, then both patterns can be considered as highly significant (in such a case each pattern can be considered significant by itself). This situation, however, is rather exceptional in usual application of the PDA as well as other algorithms, mainly due to round-off effect problem analyzed in the Section 3.1.

On the contrary, the PGA always estimates significance of each particular pattern group. This algorithm analyzes sequences of intervals with slight differences in spike timing and clusters them as one single pattern group. The significance of each such group is estimated according to Eq. (4), that is the test based on the PDA algorithm (for more details see Appendix A). In addition, if applicable, the significance of each group is also estimated according to the modified versions of Favored Pattern Detection (FPD) (Dayhoff and Gerstein, 1983), and Joint Triplet Histogram (JTH) (Prut et al., 1998) (Appendices B and C). The pattern group is discarded from further analysis whenever the probability estimation of its occurrence is not significant by any of these methods.

### 2.3. Implementation of PGA

For practical use the PGA implementation is done in three successive steps. Firstly, a pre-selection is performed of sequences of intervals detected by PDA that represent candidate template patterns. Secondly, based on the upper bound allowed for all jitters the significant template patterns are selected. The significance of the templates on first two steps is estimated with a rough fast algorithm that is based on the PDA approach. Thirdly, an iterative procedure gathers all repetitions falling into maximum jitter around the templates, optimizes the jitter for each event and evaluates the statistical significance of the pattern groups using all available statistical algorithms (Appendices A, B and C).

### 3. Results

#### 3.1. Comparative performance of PGA

Ten spike trains were generated according to a non-stationary correlated Poisson distribution with average firing rate equal to 4.1 spikes/s and time constant of the firing rate fluctuation equal to 2 s (dataset E described elsewhere (Tetko and Villa, 1997b)). The firing rates fluctuated between 0 and 33 spikes/s. Spontaneous activity was continuously simulated during 200 s at a time resolution of 1 ms. Ten precise patterns, namely $\langle 3, 5, 4; 299, 448 \rangle$ and $\langle 3, 5, 4; 299, 450 \rangle$ repeating three times and $\langle 3, 5, 4; 299, 449 \rangle$ repeating four times, corresponding to the pattern group $\langle 3, 5, 4; 299 \pm 0.5, 449 \pm 1.5 \rangle$ were inserted arbitrarily in the record (Fig. 2). One complete run of the analysis ended with the search and estimation of significant patterns over a maximum window duration equal to 500 ms. In order to assess the reliability of the method we generated 100 simulated data sets using different seed values. The
same ten patterns were inserted each time at different starting times. The level of significance in our analysis was set to \( p_0 = 0.01 \).

The expected number of patterns was estimated with the probabilistic combinatorial algorithm (Tetko and Villa, 1997b). At the maximum accuracy, i.e. jitter \( J = 1 \), no significant pattern was detected by PDA in any run. The inserted sequences were characterized by a jitter equal to 3 ms, i.e. \( (450 - 448) + 1 \), around the time delay between spike nos. 5 and 4. Note that the jitter value always includes the time accuracy corresponding to the data acquisition resolution. It might appear that using a jitter \( J = 3 \), i.e. dividing all timing information in spike train by the value of the largest jitter (Abeles et al., 1993a) would be the optimal choice to ensure the detection of the ten inserted sequences as a significant pattern. However, this is not necessarily true. Under the PDA settings the analyzed pattern would be detected as significant whenever it was found to repeat seven to eight or more times in one run (depending on the data set) and this occurred only in three out of 100 runs. This flaw arises because the starting times of the inserted patterns in the spike train were randomly generated. After division of raw delays of all spikes by jitter \( J = 3 \) round-off effects appear so that the exact starting times of the inserted patterns would matter. For example, if patterns \( \langle 3, 5, 4; 299, 449 \rangle \) started at absolute time 0 and 1 ms (or 99 and 154 ms, and in general at \( 3 \times n \) and \( 3 \times m \) + 1 ms, where \( n \) and \( m \) are arbitrary integer numbers), the same patterns would be detected as \( \langle 3, 5, 4; 99, 149 \rangle \) and \( \langle 1, 3, 2; 100, 150 \rangle \) after dividing the spike delays by \( J = 3 \) ms. In this example we used the rule of integer division of the C++ programming language, but a similar example can be found for any other sophisticated rounding utility.

The round-off effect also creates difficulties for detecting significant patterns in the recently proposed JTH method (Prut et al., 1998). In our implementation with jitter \( J = 3 \) this algorithm detected the presence of a significant pattern in two out of 100 runs. Conversely, the application of PGA to the same record detected significant pattern \( \langle 3, 5, 4; 299 \pm 0.5, 449 \pm 1.5 \rangle \) in 98 out of 100 runs. The number of detected repetitions was usually nine to 11, thus indicating that some simulated spike train data sets contained patterns that fit the template. It should be noted that in two out of 100 runs the analyzed pattern was lost since it was not identified as significant at the two first steps of the pre-selection procedure (see Section 4.3 and 4.4). This could be avoided if all patterns detected by PDA would be used as templates at the third step of the algorithm, but such calculations would require a much longer time.

The performance of PGA did not change over ranges of jitter \( J \in \{3, 21\} \) (i.e. from \( \pm 1 \) to \( \pm 10 \) ms plus 1 ms corresponding to the experimental accuracy). The significance of the detected pattern was \( p_{\text{PDA}} \in [10^{-6}, 10^{-9}] \) and \( p_{\text{JTH}} \in [10^{-7}, 10^{-11}] \) according to the tests based on PDA and JTH methods respectively. However, the round-off problem did not allow detection of the analyzed pattern using the original versions of these methods even if it was very significant.

The performance of PGA has been tested further on seven simulated records corresponding to stationary (A, B), nonstationary independent (C, D) and nonstationary correlated Poisson processes (E–G) generated as described elsewhere (Abeles and Gerstein, 1988; Tetko and Villa, 1997b). Each record included ten simultaneous spike trains whose activities were collected during 200 s with experimental accuracy of 1 ms. The Fano Factor \( c_e \) defined as \( c_e = \sigma_{\text{STD}}^2/\bar{\lambda} \) where \( \bar{\lambda} \) is the mean value of the interspike intervals and \( \sigma_{\text{STD}} \) the standard deviation of that mean was calculated according to (Softky and Koch, 1993). This factor may be used to characterize the variability of the spike train and it is equal to 1 for the time series generated according to Poisson processes. The average \( (\pm \text{S.E.M.}) \) values of the Fano Factor for the 10 spike trains of files A–G were respectively \( 0.96 \pm 0.04, 1.01 \pm 0.04, 1.8 \pm 0.2, 1.4 \pm 0.1, 1.9 \pm 0.1 \) and \( 1.1 \pm 0.1 \). The algorithm settings were window size of 500 ms and maximum jitter of 7 ms. Three spurious patterns of complexity \( c = 4 \) were detected, namely pattern \( \langle 1, 1, 1, 1; 92 \pm 2.5, 259 \pm 2.0, 456 \pm 3.0 \rangle \) repeating seven times with \( p_{\text{PDA}} = 3 \times 10^{-3}, p_{\text{JTH}} = 8 \times 10^{-8} \) and \( p_{\text{FPD}} < 0.01 \) in dataset C, pattern \( \langle 4, 4, 4, 4; 31 \pm 2.0, 109 \pm 1.5, 151 \pm 3.5 \rangle \) repeating five times with \( p_{\text{PDA}} = 3 \times 10^{-3}, p_{\text{JTH}} = 7 \times 10^{-8} \) and \( p_{\text{FPD}} < 0.01 \) in dataset E and pattern \( \langle 7, 8, A, 9; 10 \pm 2.0, 428 \pm 3.5, 481 \pm 2.0 \rangle \) repeating six times in dataset G with \( p_{\text{PDA}} = 1 \times 10^{-3}, p_{\text{JTH}} = 5 \times 10^{-7} \). The three datasets with spurious patterns were generated according to non-stationary Poisson processes and two of them, corresponding to files C and E, were characterized by Fano Factors close to 2. All three spurious patterns were formed by four cells and did not include any significant subpattern formed by three cells. The absence of significant subpatterns hardly be expected by the synfire chain theory, because the probability to detect a complex pattern in a distributed neural network is much smaller than that to detect a simpler pattern. Thus, the rationale is to require that significant patterns of higher complexity should include at least one significant subpattern. This additional criterion would exclude all spurious patterns mentioned above. Moreover, the spurious patterns were significant at the level \( p_0 = 0.01 \) and would also be excluded from further analysis by using a higher significant level, i.e. \( p_0 = 0.001 \).

If the significance of patterns in PGA was estimated only according to PDA (see Appendix A and Section 2) nine significant spurious patterns of complexity three were detected (Table 1). Five of these patterns were formed by one cell only and seven patterns included at
least one event with the maximum possible allowed jitter, i.e. ± 3.5 ms, thus indicating that they were at the limit of the detection accuracy. The patterns formed by one cell were tested for significance by JTH and FPD methods. Two out of five patterns remained significant according to JTH, but none was selected after applying the FPD test. The spurious patterns formed by two cells were all found in the data set corresponding to the non-stationary correlated Poisson process with a large range of firing rate fluctuation (0–58 spikes/s). Moreover, the cells belonging to the same pattern were characterized by the same rate of fluctuations. In this case the pair of cells tended to fire synchronously and their cross-correlogram was characterized by a hump centered on time zero (Fig. 3a). This type of pattern was also characterized by a time span relatively short, dependent on the time constant of the synchronization (Fig. 3b). For the cases shown in Table 1 the maximum duration of the pattern was 75 ms. The FPD test was not applicable to patterns formed by more than one cell, but application of JTH to our sample discarded all four patterns detected in the data set G. Thus, JTH proved to be a good method to avoid biases in estimation due to correlated firing. This method also discarded spurious patterns due to significant autocorrelation peaks (e.g. patterns formed by one cell for simulated data set G, Table 1). However, the estimation of significance by PDA was by an order of magnitude faster than by JTH and several orders faster than by FPD. Due to this reason both JTH and FPD estimations were applied for significance tests only at the last step of the analysis. The above experiment with simulated spike trains shows that a preliminary fast selection of patterns with PDA estimation followed by JTH and FPD tests of significance is likely to constitute a fast and valid approach to reliably detect spatiotemporal patterns of firing.

3.2. Sensitivity of PGA to time accuracy and number of simultaneously recorded spike trains

Three factors are of particular importance in the practical use of PGA. These are the jitter of the spike intervals, the number of repeating occurrences of a pattern group within a data set and the number of simultaneously recorded neurons. In this paper we analyzed in detail the effect of these factors using dataset E mentioned above. This dataset corresponds to non-stationary correlated Poisson processes and is likely to be representative of a difficult case study.

On each run of the analysis we inserted ten triplets repeating the same number of times and with the same jitter. The jitter was calculated as the standard deviation around each spike interval. The algorithm was set to detect patterns lasting no more than 500 ms (corresponding to the maximal duration of the inserted patterns) and with a maximum allowed jitter of 7 ms. Ten such runs were repeated and the results were averaged. Two approaches were applied to analyze the significance of the inserted patterns. In the first approach, PGA was applied to detect significant patterns in the usual way. No information about the inserted patterns was provided to PGA and the algorithm itself was used to detect pattern templates and estimate their significance using the full protocol as described in Appendix A. In the second approach, the exact templates of all inserted patterns were directly used in the iterative procedure for optimization of jitter and exact calculation of significance. Thus, the second procedure estimated the maximum theoretical number of significant patterns that could be retrieved from the analyzed dataset. We found that on average both approaches detected the same number of patterns. This result confirmed that PGA provided a reliable detection of significant patterns in simulated data.

<table>
<thead>
<tr>
<th>Selected pattern</th>
<th>No. of repetitions</th>
<th>Data set</th>
<th>Test of significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formed by one cell</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>〈3, 3, 3; 15 ± 3.5, 47 ± 3.0〉</td>
<td>10</td>
<td>D</td>
<td>9 × 10⁻³</td>
</tr>
<tr>
<td>〈9, 9, 9; 40 ± 3.0, 202 ± 1.0〉</td>
<td>9</td>
<td>E</td>
<td>2 × 10⁻³</td>
</tr>
<tr>
<td>〈2, 2, 2; 10 ± 3.5, 23 ± 2.5〉</td>
<td>14</td>
<td>G</td>
<td>5 × 10⁻⁴</td>
</tr>
<tr>
<td>〈2, 2, 2; 27 ± 2.5, 41 ± 3.5〉</td>
<td>15</td>
<td>G</td>
<td>4 × 10⁻⁴</td>
</tr>
<tr>
<td>〈5, 5, 5; 23 ± 3.5, 40 ± 2.5〉</td>
<td>13</td>
<td>G</td>
<td>2 × 10⁻⁴</td>
</tr>
<tr>
<td>Formed by several cells</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>〈1, 2, 2; 17 ± 2.0, 54 ± 3.5〉</td>
<td>13</td>
<td>G</td>
<td>7 × 10⁻³</td>
</tr>
<tr>
<td>〈2, 1, 2; 10 ± 3.5, 27 ± 3.0〉</td>
<td>19</td>
<td>G</td>
<td>4 × 10⁻⁶</td>
</tr>
<tr>
<td>〈5, 6, 5; 12 ± 3.5, 40 ± 3.0〉</td>
<td>16</td>
<td>G</td>
<td>1 × 10⁻³</td>
</tr>
<tr>
<td>〈7, 8, 8; 1 ± 3.0, 24 ± 3.0〉</td>
<td>17</td>
<td>G</td>
<td>2 × 10⁻⁵</td>
</tr>
</tbody>
</table>

* Data sets: D = non-stationary independent Poisson processes; E–G = non-stationary correlated Poisson processes. PDA: pattern detection algorithm; JTH: joint triplet histogram; FPD: favored pattern detection; n.s.: not significant; N/A: not applicable.

Fig. 3. Spurious spatiotemporal firing pattern elicited by correlated activity in a pair of cells. This pattern was detected as significant with $P_{\text{PDA}} = 7 \times 10^{-4}$ by the significance test based on the PDA method (Appendix A) but it was not significant and was discarded according to the test based on the JTH method (Appendix B). (a) Autocorrelograms and cross-correlogram of two non-stationary correlated spike trains recorded during 200 s. The abscissa is the lag (in ms) and the ordinate is scaled in rate units (spikes/s). The dashed lines indicate the 99% confidence level assuming a Poisson distribution. Note the peak centered on zero delay in the cross-correlogram that indicates a tendency to synchronous firing between the two cells. (b) Fifteen repetitions of the pattern $\langle 1, 2, 2; 17 \pm 2.5; 54 \pm 3.5 \rangle$ are displayed as rasters aligned on pattern start.

Concerning the jitter and number of repetitions of a pattern, we found, that whenever each triplet formed by the same cell repeated at least seven times and was characterized by a standard deviation equal to S.D. = 0.4 ms the chance to detect all patterns was near 70% at a level of significance $p_0 = 0.01$ (upper solid line curve in Fig. 4a) but decreased to 40% at $p_0 = 0.001$ (upper solid line curve in Fig. 4b). With low variability (S.D. = 0.4) all triplets formed by three different neurons (dotted lines in Fig. 4) were detected at either $p_0 = 0.01$ or $p_0 = 0.001$ significance level if they repeated at least ten times. Conversely, with high variability (S.D. = 1.5) more than half of the triplets formed by three neurons were missed for the same number of repetitions. The curves calculated for triplets that were formed by two cells (data not shown) were similar to those for three-cells triplets. If the time accuracy of the patterns was characterized by a standard deviation equal to 1.5 ms the patterns should be repeated 15 and 20 times in order to have almost the certitude of being detected at significance levels 0.01 and 0.001, respectively (Fig. 4).

An increase in the number of simultaneously recorded cells, and, correspondingly, an increase of factor $Q(a, c, \Omega)$ changed the significance level. The results reported for triplets formed by one cell and significance level $p_0 = 0.001$ calculated for ten spike trains correspond to a level of significance $p_0 = 0.01$ with 100 simultaneously recorded cells $Q(1, 3, 100)/Q(1, 3, 10) = 100/10 = 10$. It is important to note that for triplets formed by three different cells a change of one order of magnitude in the significance level corresponds to an increase of the number of simultaneously recorded cells from ten to 20, because $Q(3, 3, 20)/Q(3, 3, 10) = 1140/120 = 9.5$.

4. Discussion

We have described a new analytical method, the PGA, for detecting spatiotemporal firing patterns in neuronal spike trains, and presented evidence that this method can detect patterns reliably under a variety of
conditions in simulated spike trains. The main feature of the PGA is to detect and to estimate significance of patterns with variable jitters notwithstanding its complexity and number of different cells in the pattern. The algorithm automatically searches for precise patterns that are then clustered into pattern groups, estimates and selects the significant pattern groups. The estimation of the significance of the patterns is done using test based on PDA, FPD and JTH methods in order to minimize the possibility of detection of spurious patterns.

The implementation of PGA was done using three-step procedure that was required to speed up calculations. Indeed, the pre-selection of patterns on the first and the second steps decreased an amount of patterns to be considered on the last step by 2–4 orders of magnitude (Appendix A). Since the time to estimate a significance of a pattern on the last step was by 1–2 orders larger (see also Section 4) this approach provided a considerable increase in the total speed of calculations. The final estimation of all patterns by three statistical methods at the third step of algorithm (and this estimation was the same for all template patterns pre-selected for this step) calculated a valid significance for each such pattern. Of course, despite the implemented procedure was very conservative it was possible that some significant patterns were lost. A more complete procedure could consist to use all detected patterns as templates for the iterative optimization of maximum jitters at the third step of calculations. However, such calculations would increase calculation time by 3–6 orders and would not be feasible for the modern computers.

4.1. Implications of burst filtering

The PGA has been developed for spike train data with low-frequency of spike discharges, and without burst activity. An important parameter of this algorithm is therefore the frequency used for burst filtering. We have applied $f = 400$ Hz to filter out spikes found in burst. This value provided a robust filtering over a wide range of jitters and window duration while removing only a limited number of spikes in data (Tetko and Villa, 1997b).

The filtering frequency was determined for spike train data with average firing rates near 5 spikes/s. Under such firing rates $\sim 1$–2% of spikes were removed by the filtering procedure for reported data, except in dataset G where 7% of spikes were filtered. A larger number of spikes was deleted if the filtering was applied to data with higher firing rate. The number of filtered spikes also depended on the non-stationarity of the discharges. An increase in firing rate from 5 to 21 and 9 spikes/s in datasets A and E, respectively, corresponded to $\sim 10\%$ of all spikes discarded by the filtering procedure. If we consider that 10% is the maximum tolerable rate of spikes discarded by the filtering procedure, then the application of PGA is limited to spike trains data with firing rates not exceeding 10–20 spikes/s.

An important feature of PGA algorithm is that even for simulated data generated according to non-stationary Poisson processes and characterized by Fano Factor values near to 2 the algorithm detected one spurious pattern per simulated data set, at most. Such patterns were excluded by setting the algorithm to accept only a high complexity pattern that must include significant subpatterns of lower complexity. This result demonstrates the robustness of our approach and suggests that PGA can also be extended to the analysis of data generated according to non-stationary Poisson processes using deconvolution processing (Tetko and Villa, 2001).
4.2. Comparison with other methods

The PGA favorably compares with other similar algorithms developed to analyze patterns formed by several spike trains. For three and more distinct spike trains Frostig and co-workers proposed a statistical evaluation method based on the use of $2 \times 2$ contingency tables and the application of Fisher’s exact test (Frostig et al., 1990a,b). At the detection stage of this method all possible intervals between three spike trains are tested for significance and at a second step all significant patterns found previously are checked for higher-order associations among spike trains. A hindrance to the practical application of this method is its inability to detect recurring patterns involving less than three different spike trains and patterns of more than three spikes if not all subpatterns formed by three spikes (triplets) are statistically significant.

Several specific analytical methods limited to analyzing patterns of complexity three (i.e. triplets) have been developed partly because the interaction between three events can be analyzed by relatively simple visual tools. It was also justified by the fact that in most experimental preparations the chance of observing patterns of discharges of higher-order complexity is very small. Detection of triplets was firstly achieved using the ‘snowflake’ triple-spike renewal histogram (Perkel et al., 1975; Abeles, 1983). However, this method does not provide a reliable estimation of the significance of the detected patterns. This problem was solved by using the Joint Peri Stimulus Time Histogram (Aertsen et al., 1989) for studies of threefold correlation among spikes and several applications have been reported (Abeles et al., 1993a,b,c). A new version of this algorithm has been improved for statistical evaluation by taking into account the existence of significant cross-correlation between cells (Prut et al., 1998). We refer to this new algorithm as the JTH. Indeed, the incorporation of an extension of JTH for variable accuracy patterns in PGA has avoided the detection of spurious patterns due to the presence of significant correlation between cells. However, the original version of JTH is unable to detect pattern with variable jitters, as it was demonstrated in the Result section. The estimation of significance by this method is slower than by fast computational approaches based on the PDA method and limits its application to pre-selection of templates on the early steps of the PGA.

The PDA developed by Abeles and Gerstein (1988), which provided the starting point of the method described here, allows for comprehensive search of firing patterns and the evaluation whether there is a statistical excess of patterns notwithstanding their complexity. Improved and computationally effective versions of this method, based on probabilistic and combinatorial algorithms, were recently elaborated (Tetko and Villa, 1997b). The main restriction of PDA when compared to the other methods is that it does not allow the detection of individually significant patterns. In addition, PDA is sensitive to the rates of discharges, so that a few units firing at a high rate may ‘mask’ the spatiotemporal patterns formed by other simultaneously recorded units firing at low rates. In contrast, the performance of PGA is not influenced by differential firing rates among simultaneously recorded spike trains. Furthermore, we have demonstrated that PGA, due to the combined use of other methods, particularly JTH, can avoid the detection of spurious patterns even in conditions of correlated non-stationarity.

The advantage of the FPD developed by Dayhoff and Gerstein (1983) is that it can be used to detect patterns in which individual occurrences may have missing spikes. This allows for the possibility of a particular cell, belonging to one synfire chain, being in a refractory period when synchronous firing is evoked in that chain (Abeles, 1982). Such ‘missing’ spikes may explain why patterns with complexity higher than four are very rarely detected in experimental recordings using other methods such as, e.g. PDA or JTH. The major limitation of the FPD is that it can be applied only for analysis of patterns formed by one neuron. Despite the current version of PGA method cannot detect patterns with missing spikes, such an extension of the algorithm represents an interesting topic for future studies.

In summary, the PGA represents a significant advance on available pattern detection methods and provides a promising new tool for the identification and evaluation of the statistical significance of temporal patterns in spike trains. The selection of very restrictive criteria in the application of PGA should make it possible to move the debate away from the mere existence of such patterns towards the study of their physiological significance.

Acknowledgements

We thank M. Abeles and Y. Prut for providing us with a source code of their programs and simulated data and B. Hyland for his helpful comments. Calculations were partly performed at the Swiss Center for Scientific Computing. This study was partially supported by HFSP STF-421/95, Swiss NSF 2053-055753.98/1 and INTAS-OPEN 97-0168 and 97-0173 grants.
Appendix A. Implementation of PGA

A.1. Preselection of pattern templates

The pattern search by PDA provides a list of exact sequences of intervals, i.e. precise patterns, repeating two or more times. The overall number of precise patterns can be in the order of tens of millions for experimental data sets of average complexity. Thus, the use of all precise patterns as templates for search of patterns with variable accuracy in spike timing is not likely to be applied in routine analysis. On the one hand, it seems reasonable to restrict further analysis only to those patterns repeating a large number of times. On the other hand, no objective criteria exist to set a threshold level to the number of repetitions of a precise pattern. Indeed, neurons discharging at high activity rates are likely to generate precise patterns that repeat a higher number of times. Then, a unique threshold level set at this stage of the analysis would necessarily bias the final result in dramatically decreasing the probability to detect spatiotemporal firing patterns contributed by neurons firing at low rates.

In order to speed-up the algorithm we preselected a number of potential templates according to the following procedure. Firstly, the PDA was used with the time jitter \( \Psi \in \{\Delta_{\text{exp}}, \ldots, J\} \) and all detected patterns were sorted according to the label of the participating neurons and time delays. Then, all patterns formed by the same spikes and characterized by differences in time intervals less than the maximum allowed jitter \( J \) were searched throughout the whole set of patterns detected by PDA. Each time it was possible to find several precise patterns according to this criterion, a pattern group was formed by merging together the corresponding precise patterns. A provisional template was then generated selecting the mean delays of the pattern group. The standard deviations \( s_i \) of the time intervals for each spike \( i \) of the pattern group were calculated.

In order to develop a fast method for selection of templates let us assume that the analyzed patterns are distributed following a Gaussian distribution around the provisional template. Then, 95% of the patterns fall within the interval represented by two times the standard deviations on each side of the mean. In addition, we have to consider the time accuracy of the analysis, that is the jitter \( \Psi \). The product of the spike deviations \( f_{\Delta \Psi} \) is calculated as

\[
 f_{\Delta \Psi} = \prod_{j=1}^{t_i} (4s_j + \Psi) \tag{A1}
\]

We can further assume that all patterns are distributed uniformly in the space of \( c - 1 \) dimensions, where \( c \) is the pattern complexity (note, that both assumptions are only used to pre-select patterns). The value defined by the product of the spike deviations \( f_{\Delta \Psi} \) corresponds to the area occupied in the \( c - 1 \) dimensional space by the precise patterns of maximum duration \( w \). Then, the expected number of patterns \( n^{(i)}_{c} \) falling in the area \( f_{\Delta \Psi} \) is estimated as

\[
 n^{(i)}_{c} = \frac{f_{\Delta \Psi}}{w^{c-1}} \cdot N^{(i)}
\]

where \( N^{(i)} \) is the expected total number of all possible combinations of considered patterns repeating \( r \) times given by Eq. (2) (see Section 2). Note that \( N^{(i)} \) can be calculated according to the ‘ad-hoc’ method (Abeles and Gerstein, 1988), combinatorial or probabilistic combinatorial methods (Tetko and Villa, 1997b). For practical implementations we prefer to use the combinatorial method because it is computationally faster (Tetko and Villa, 1997b).

For estimation of significance of the provisional pattern group we count the number \( v(r_0) \) of precise patterns in the group, each pattern repeating at least \( r_0 \) times. These numbers are compared to the corresponding expected numbers

\[
 n^{(i)}_{c} = \sum_{i=r_0}^{\infty} \frac{f_{\Delta \Psi}}{w^{c-1}} \cdot \sum_{i=r_0}^{\infty} N^{(i)}
\]

using the cumulative Poisson probability function \( p = pr\{v(r_0), n^{(i)}_{c} \geq r_0\} \). If a significant excess (i.e. \( p < p_0 \)) is detected, then the procedure is stopped and the corresponding provisional template is kept for the next stage. Otherwise, we consider the next number of repetitions observed for the group pattern and so on until the maximum number of observed repetitions, if necessary. If this last estimation is not significant, then the provisional template is discarded from further analysis.

This procedure tends to overestimate the significance of preselected patterns. Indeed, at this stage it is preferable to keep some non-significant patterns for final selection. If two pattern groups are very similar and their corresponding time intervals overlap in time, the pattern group with the highest number of repetitions is kept. This whole procedure of preselection of pattern templates is run for increasing jitter values \( \Psi \), \( \Psi \in \{\Delta_{\text{exp}}, \ldots, J\} \). At each run the preselected sequences are pooled with those of the previous run. This stage usually decreases the number of provisional templates provided by the preliminary PDA search by 2–4 orders of magnitude.

A.2. Final selection of templates

Each pre-selected template is used for a comprehensive search of all related patterns in the analyzed record which are characterized by differences in time intervals less than the jitter \( J \). All patterns detected for the same template are grouped together and the product of the
spike deviations $f_{d\Psi}$ is calculated according to Eq. (A1) using the best available accuracy, i.e. $\Psi = \Delta_{\text{exp}}$. In Eq. (2) (see Section 2) we substitute $f_\Psi$ by $f_{d\Psi}$ and we compute the corresponding expected number of patterns $N_{\Psi}(f_{d\Psi})$, where $r$ is the total number of repetitions of patterns belonging to the provisional pattern group. The significance of the group is estimated according to $p = pr\{1, N_{\Psi}(f_{d\Psi})\}$. If $p < p_0$ the provisional template corresponds to a significant pattern group and it is used for the final step of the analysis.

A.3. Iterative procedure for optimization of jitter and exact calculation of significance

At this step a new comprehensive search of all sequences of spike intervals matching the template is started. All possible combinations of patterns belonging to the pattern group are analyzed. For each individual combination we calculate the actual jitters of the corresponding spike intervals, defined as the difference between the maximum and minimum values for each time delay added to the data acquisition accuracy, i.e. $\Delta_{\text{exp}}$. Then, the product of jitters $f_{\delta}$ where $f_{\delta} = \Pi_{i=1}^\delta \Delta_{\alpha}$ is calculated and the significance of that combination of patterns is estimated by $pr\{1, N_{\Psi}(f_{\delta})\}$, see Eq. (4). The most significant combination of patterns is selected as the pattern group. A new template is calculated with the means of the actual distribution of delays in the pattern group. This procedure is repeated until no shift of template is detected, thus allowing detection of a configuration of the pattern group with the highest significance.

The speed of PGA critically depends on several parameters of the analyzed data set, i.e. the number of simultaneously recorded spike trains, the duration of the recording, the maximum complexity of the analyzed patterns, and the number of significant groups. We used ANSI C++ software, and Power Macintosh G3 (processor speed 266 MHz) or Hewlett Packard G3 cluster of the Swiss Center for Scientific Computing. The generation of random numbers was done using ran2() function from Press et al. (1994). Setting the algorithm with a maximum window duration of 1000 ms, maximum jitter equal to 11 ms in data sets of 15 simultaneously recorded single units firing at an average rate of 2.5 spikes/s recorded over 1 h, the analysis of firing patterns of complexity 3, 4 and 5 required $\sim 4 - 5$ h of calculation on the Macintosh computer. An on-line version of the algorithm is available from http://www.neuroheuristic.org (Villa et al., 2001).

Appendix B. FPD test of significance for pattern groups

According to the method of ‘favorite patterns’ detection (Dayhoff and Gerstein, 1983), FPD, a template function $f(x) = \Sigma_{k=1}^{c} s_k(x - t_k)$ is defined, where $c$ is the complexity of the pattern, $t_k$ is the delay between the first and the $k + 1$th spike of the template and

$$s_k(x) = \begin{cases} 1 - 2x/b_k & \text{if } 0 \leq x < b_k/2 \\ 1 + 2x/b_k & \text{if } 0 \leq x < b_k/2 \\ 0 & \text{otherwise} \end{cases} \quad (A4)$$

is a triangular function of width $b_k$ and height 1, centered at $x = 0$. The values of $b_k$ should be equal to a reasonable fraction of the average interspike interval (Dayhoff and Gerstein, 1983). The FPD is only applied to patterns formed by sequences of spike intervals from the same spike train. Correspondingly, in our implementation, this test is applied to the templates of the pattern groups formed by the same neuron. In the original FPD (Dayhoff and Gerstein, 1983) all $b_k$ are set identical, whereas we extend this method to variable accuracy patterns by selecting $b_k$ as corresponding to the time jitters determined by PGA, i.e. $b_k = \Delta_k$.

Let us choose an event of the spike train to initialize the comparison and let us note it as $S_1$. Let $h_k$ be the time between spike $S_1$ and the $k$th spike after $S_1$ in the spike train. The function $f(x)$ is used to calculate a match value $M$ defined as $M = \Sigma_{i=1}^{c} f(h_k)$. All events in the spike train can be used for initialization, thus yielding a set of $M$ values. Usually, only a small fraction of $M$ values is not equal to 0. The higher the value of $M$ the better match between the template pattern and the spike train. A threshold level $M_T$ should be set above which the template is said to ‘match’ the spike train. In our implementation $M_T$ is equal to the minimum value of the match values $M$ calculated for all patterns belonging to the same pattern group detected by PGA.

The significance of a pattern group is tested as follows. According to the FPD the spike train is shuffled $N = \Omega/p_0$ times, where $p_0$ is the level of significance and $\Omega$ is number of simultaneously recorded neurons (in this case the number of combinations is $Q(1, c, \Omega) = \Omega$). For each shuffled record we calculate a value $v_i$ equal to the number of matches, i.e. satisfying the condition $M_i \geq M_T$. After having shuffled the spike train $N$ times, the maximum number of matches is determined as $v_{\text{max}} = \max(v_1, v_2, \ldots, v_N)$. For a fixed threshold level $M_T$, the pattern group is significant, at the level $p_0$, if $v_{\text{max}}$ is less than the number of repetitions of the pattern found in the raw spike train data.

Appendix C. JTH test of significance for pattern groups

The ‘Joint Triplet Histogram’ (JTH) (Prut et al., 1998), allows the estimation of the significance of a
pattern formed by three events that can belong to three different spike trains. This method is based on the computation of a threefold correlation matrix \( C_{2\times 2}(i) \) of all triplets \( S_i = (x, y, z) \). The value of the \( c(i,j) \) element of this matrix is equal to the number of occurrences of the pattern \( (x, y, z) \) in the analyzed record, where \( \Delta_x = \Delta_y = \Delta_z = J/2 \) and \( J \) is the time jitter, or the bin size (Prut et al., 1998). The matrix has a maximum dimension of \( N^2 \) where \( N = w/J \), and \( w \) is the window duration (i.e. the maximum allowed duration of the pattern).

The correlation-predictors \( C_{2\times 2}(i) \) and \( C_{2\times 2}(j) \), where \( i \in \{1, 2, \ldots, N\} \), calculate the number of times the reference cell \( x \) or \( y \) respectively, fired at the delay \( i \pm \Delta_i \) after the trigger cell \( z \) fired a spike. The expected count in a bin is estimated as \( e_{i\pm j} = \langle C_{2\times 2}(i) \rangle / n_i C_{2\times 2}(j) / n_j \), where \( C_{2\times 2} \) and \( C_{2\times 2} \) are the correlation-predictors and \( n_i \) is the number of triggers found for cell \( z \). The correlation between the two reference units (i.e. \( x \to y \) for the upper triangular and \( y \to x \) for the lower triangular of the matrix) is assumed to be constant along each diagonal of the matrix. It is possible to correct the expected count of events in a bin by introducing a correction factor \( CR \) that is dependent on the cross-correlation between the reference units. For the \( k \) diagonal elements of the upper triangular matrix the corrected expected count in a bin \( e' \) is estimated as \( e'_{i\pm j+k} = e_{i\pm j+k} + C_{2\times 2}(j) CR_k \), where the correction factor \( CR_k \) is given by \( CR_k = \sum_{i=0}^{N-1} (C_{2\times 2}(i) - e_{i\pm j+k}) / \sum_{i=0}^{N-1} C_{2\times 2}(j) \). As a conservative measure, only positive value of \( CR_k \) are considered. For the central diagonal, i.e. \( k = 0 \), both correlations \( x \to y \) and \( y \to x \) were analyzed and the maximum value of the two corresponding correction factors was used as \( CR_0 \) (Y. Prut, Personal communication). The significance of each bin of the matrix is tested using the cumulative Poisson probability function \( pr \{ e''(i,j), c(i,j) \} \) and was corrected by factor of \( Q(a, c, \Omega) \) in order to compensate for chance effect due to the number of possible combinations for the considered pattern (see Eq. (2)).

We have also extended the original JTH for application to higher order sequences of spikes. For example we consider the estimation for significance of bins corresponding to patterns formed by four spikes, i.e. quadruplets \( S_4 = (x, y, z, w) \). The correlation-predictors \( C_{2\times 2}(i), C_{2\times 2}(j), C_{2\times 2}(k) \) are calculated as described previously for triplets, and the corresponding expected count in a bin \( e_{i\pm j\pm k} \) is defined as \( e_{i\pm j\pm k} = C_{2\times 2}(i) C_{2\times 2}(j) C_{2\times 2}(k) / (n_i \cdot n_j \cdot n_k) \). Let us consider all possible triplets formed with cell \( z \) as the trigger cell, i.e. \( S_1 \), \( S_2 \), \( S_3 \), and \( S_4 \). For each triplet we estimate the expected count in the corresponding threefold correlation matrix, i.e. \( e_{i\pm j\pm k} = C_{2\times 2}(i) C_{2\times 2}(j) C_{2\times 2}(k) / (n_i \cdot n_j \cdot n_k) \). We can express the expected count for a bin \( e_{i\pm j\pm k} \) as follows:

\[
C_{2\times 2}(i) C_{2\times 2}(j) C_{2\times 2}(k) / (n_i \cdot n_j \cdot n_k) = \text{sqrt}(C_{2\times 2}(i) C_{2\times 2}(j) C_{2\times 2}(k) / (n_i \cdot n_j \cdot n_k))
\]

\[
C_{2\times 2}(i) C_{2\times 2}(j) C_{2\times 2}(k) / (n_i \cdot n_j \cdot n_k) = \text{sqrt}(e_{i\pm j\pm k} / n_i \cdot n_j \cdot n_k)
\]

\[(A5)\]

The substitution of \( e_{i\pm j\pm k} \) by the corresponding values \( e''_{i\pm j\pm k} \) leads to the estimation of the expected number of counts corrected for cross-correlation between all cells, i.e. \( e''_{i\pm j\pm k} = \text{sqrt}(e_{i\pm j\pm k} / n_i \cdot n_j \cdot n_k) \). A similar approach can be extended to higher complexities. The estimation of significance of a bin in such higher order complexity matrix is reduced to a number of estimations of the corresponding bins in threefold correlation matrices. The expected number of counts for a pattern \( S_c \) of any complexity \( c \geq 3 \) is derived as:

\[
e''(i_1, i_2, \ldots, i_c) = c - 2 \prod_{l=2, \ldots, c-1}^{l} e''(i_l, i_{l+1})^{(c-l)/2} - (c-1)/2
\]

\[(A6)\]

References


