

Leaders and followers: quantifying consistency in spatio-temporal propagation patterns

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2017 New J. Phys. 19 043028

(<http://iopscience.iop.org/1367-2630/19/4/043028>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 130.192.10.10

This content was downloaded on 05/07/2017 at 10:01

Please note that [terms and conditions apply](#).

You may also be interested in:

[A study of epileptogenic network structures in rat hippocampal cultures using first spike latencies during synchronization events](#)

Mohan Raghavan, Bharadwaj Amrutur, Kalyan V Srinivas et al.

[Spatiotemporal dynamics of continuum neural fields](#)

Paul C Bressloff

[Measure of synchrony in the activity of intrinsic cardiac neurons](#)

Jean-Philippe Longpré, Siamak Salavatian, Eric Beaumont et al.

[Functional clustering in hippocampal cultures](#)

S Feldt, J X Wang, E Shtrahman et al.

[Use of adaptive network burst detection methods for multielectrode array data and the generation of artificial spike patterns for method evaluation](#)

G D C Mendis, E Morrisroe, S Petrou et al.

[Characterization of long-range functional connectivity in epileptic networks by neuronal spike-triggered local field potentials](#)

Beth A Lopour, Richard J Staba, John M Stern et al.

[Tracking burst patterns in hippocampal cultures with high-density CMOS-MEAs](#)

M Gandolfo, A Maccione, M Tedesco et al.

[Fractal dimension analysis for spike detection in low SNR extracellular signals](#)

Mehrdad Salmasi, Ulrich Büttner and Stefan Glasauer



PAPER

Leaders and followers: quantifying consistency in spatio-temporal propagation patterns

OPEN ACCESS

RECEIVED

8 November 2016

REVISED

22 March 2017

ACCEPTED FOR PUBLICATION

23 March 2017

PUBLISHED

28 April 2017

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Thomas Kreuz¹, Eero Satuvuori^{1,2,3}, Martin Pofahl⁴ and Mario Mulansky¹¹ Institute for complex systems, CNR, Sesto Fiorentino, Italy² Department of Physics and Astronomy, University of Florence, Sesto Fiorentino, Italy³ Move Research Institute Amsterdam, Department of Human Movement Sciences, Vrije Universiteit Amsterdam, The Netherlands⁴ Department of Epileptology, University of Bonn, GermanyE-mail: thomas.kreuz@cnr.it**Keywords:** data analysis, spatio-temporal propagation, SPIKE-order, spike train order, neuroscience, climate science, simulated annealing

Abstract

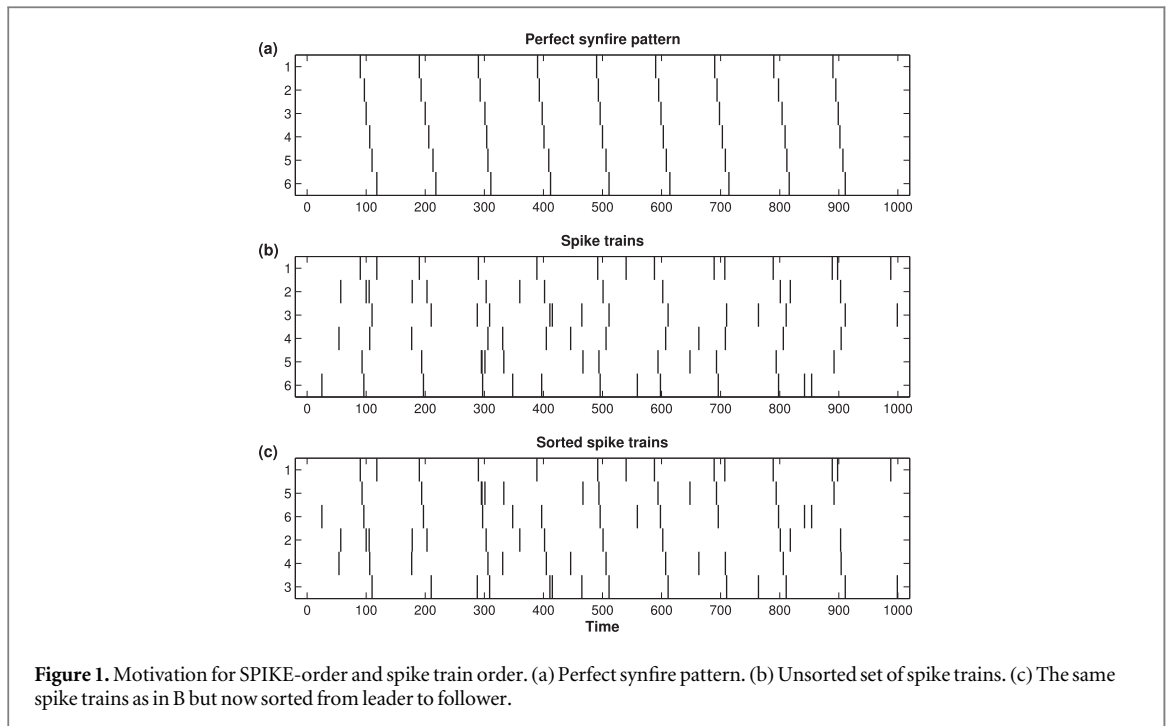
Repetitive spatio-temporal propagation patterns are encountered in fields as wide-ranging as climatology, social communication and network science. In neuroscience, perfectly consistent repetitions of the same global propagation pattern are called a *synfire pattern*. For any recording of sequences of discrete events (in neuroscience terminology: sets of spike trains) the questions arise how closely it resembles such a synfire pattern and which are the spike trains that lead/follow. Here we address these questions and introduce an algorithm built on two new indicators, termed *SPIKE-order* and *spike train order*, that define the *synfire indicator* value, which allows to sort multiple spike trains from leader to follower and to quantify the consistency of the temporal leader-follower relationships for both the original and the optimized sorting. We demonstrate our new approach using artificially generated datasets before we apply it to analyze the consistency of propagation patterns in two real datasets from neuroscience (giant depolarized potentials in mice slices) and climatology (El Niño sea surface temperature recordings). The new algorithm is distinguished by conceptual and practical simplicity, low computational cost, as well as flexibility and universality.

1. Introduction

Recordings of spatio-temporal activity are ubiquitous in many scientific disciplines. Among the most prominent examples are large-scale electrophysiological measurements of neuronal firing patterns in experimental neuroscience [1, 2] and sensor data acquisition in seismology [3], oceanography [4], meteorology [5], or climatology [6]. Other examples include interaction protocols in social communication [7, 8] or monitoring single-node dynamics in network science [9].

In all of these fields recordings often exhibit well-defined patterns of spatio-temporal propagation where some prominent feature first appears at a specific location and then spreads to other areas until potentially becoming a global event. Such characteristic propagation patterns occur in phenomena such as avalanches [10], tsunamis [11], chemical waves and diffusion processes [12], and epileptic seizures [13]. Further examples are the epidemic transmission of diseases [14], and, more recently, the spreading of memes on social networks [15] or in science [16].

In many cases spatio-temporal recordings can be represented as a two-dimensional plot where for each recording site the occurrence of certain discrete events (often obtained from threshold crossings in continuous data) are indicated by time markers. In neuroscience such a plot is known as a raster plot. A sequence of stereotypical neuronal action potentials (*spikes*, [17]) is a *spike train* and a set of spike trains exhibiting perfectly consistent repetitions of the same global propagation pattern is called a *synfire pattern*. In this paper we adapt this terminology and use all of these expressions not only in the context of neuronal spikes but also for any other kind of discrete events. However, note that our use of the term ‘synfire pattern’ differs slightly from the literature (see e.g. [18]). Here we define a synfire pattern as a sequence of global events in which all neurons fire in consistent



order and the interval between successive events is at least twice as large as the propagation time within an event. An example of a rasterplot with spike trains forming a perfect synfire pattern is shown in figure 1(a).

For any spike train set exhibiting propagation patterns the questions arise naturally whether these patterns show any consistency, i.e., to what extent do the spike trains resemble a synfire pattern, are there spike trains that consistently lead global events and are there other spike trains that invariably follow these leaders? Such questions about leader-follower dynamics have been specifically investigated not only in neuroscience [19], but also in fields as wide-ranging as, e.g., climatology [20], social communication [21], and human-robot interaction [22].

In this study we introduce a framework consisting of two directional measures (*SPIKE-order* and *spike train order*) that allows to define a value termed *synfire indicator* which quantifies the consistency of the leader-follower relationships in a rigorous and automated manner. This synfire indicator attains its maximal value of 1 for a perfect synfire pattern in which all neurons fire repeatedly in a consistent order from leader to follower (figure 1(a)).

The same framework also allows to sort multiple spike trains from leader to follower, as illustrated in figures 1(b) and (c). This is meant purely in the sense of temporal sequence. Whereas figure 1(b) shows an artificially created but rather realistic spike train set, in figure 1(c) the same spike trains have been sorted to become as close as possible to a synfire pattern. Now the spike trains that tend to fire first are on top whereas spike trains with predominantly trailing spikes are at the bottom.

We demonstrate the new approach using artificially generated datasets before we apply it to analyze the consistency of propagation patterns in two real datasets from very different fields of research, neuroscience and climatology. The neurophysiological dataset consists of neuronal activity recorded from mice brain slices. These recordings typically exhibit a sequence of global events termed giant depolarized potentials (GDPs) and one of the main questions we investigate is whether it is possible to identify neurons that consistently lead these events (potential hub neurons, see [23]). The climate dataset uses optimum interpolated sea surface temperature (OISST) data provided by the National Oceanic and Atmospheric Administration (NOAA) to follow the El Niño phenomenon in the central pacific region [24] over the range of 35 years, from 1982 to 2016. We employ a threshold criterion to track the El Niño events and then quantify the consistency of the longitudinal movement of the propagation front.

The remainder of the article is organized as follows: in the methods (section 2) we first describe the coincidence detection (section 2.1) and the symmetric measure SPIKE-synchronization (section 2.2). Subsequently, we introduce the new directionality approach consisting of the two measures SPIKE-order and spike train order (section 2.3) as well as the synfire indicator (section 2.4) before we discuss the use of SPIKE-order surrogates to evaluate the statistical significance of the results in section 2.5. The results section 3 consists of three subsections detailing applications of the new approach to artificially generated datasets (section 3.1), neurophysiological data (section 3.2) and climate data from the El Niño phenomenon (section 3.3). Conclusions

are drawn in section 4. Finally, both real datasets are described in the appendix, the electrophysiological recordings in appendix A, and the El Niño dataset in appendix B.

2. Measures

Analyzing leader–follower relationships in a spike train set requires a criterion that determines which spikes should be compared against each other. What is needed is a match maker, a method which pairs spikes in such a way that each spike is matched with at most one spike in each of the other spike trains. This match maker already exists. It is the adaptive coincidence detection first used as the fundamental ingredient for the bivariate measure *event synchronization* [25, 26].

Event synchronization itself is symmetric and quantifies the overall level of synchrony from the number of quasi-simultaneous appearances of spikes. It was proposed along with an asymmetric measure termed *delay asymmetry* which evaluates the temporal order among matching spikes in the two spike trains.

However, unfortunately both event synchronization and delay asymmetry are defined for the bivariate case of two spike trains only, rely on sampled time profiles, and have a very non-intuitive normalization. For the symmetric variant we have already addressed these issues by proposing *SPIKE-synchronization* [27], a renormalized multivariate extension of event synchronization.

The two new measures SPIKE-order and spike train order proposed here improve and extend the asymmetric measure delay asymmetry in the same way. In particular, instead of just quantifying bivariate directionality they open up a completely new application, since they allow us to sort the spike trains according to the typical relative order of their spikes and to quantify the consistency of this order using the synfire indicator.

All four approaches (bivariate/multivariate, symmetric/asymmetric) are time-resolved as well as parameter- and scale-free. Their calculation consists of two steps, adaptive coincidence detection followed by a combination of normalization and windowing. The first step, adaptive coincidence detection, is the same for all of these measures.

2.1. Adaptive coincidence detection

Most coincidence detectors rely on a coincidence window of fixed size τ [28, 29]. However, since in many cases it is very difficult to judge whether two spikes are coincident or not without taking the local context into account (see figure 2(a) for an example), [25] proposed a more flexible coincidence detection. This coincidence detection is scale- and thus parameter-free since the minimum time lag $\tau_{ij}^{(1,2)}$ at which two spikes $t_i^{(1)}$ and $t_j^{(2)}$ of spike trains (1) and (2) are no longer considered to be synchronous is adapted to the local firing rates according to

$$\tau_{ij}^{(1,2)} = \min \{t_{i+1}^{(1)} - t_i^{(1)}, t_i^{(1)} - t_{i-1}^{(1)}, t_{j+1}^{(2)} - t_j^{(2)}, t_j^{(2)} - t_{j-1}^{(2)}\} / 2. \quad (1)$$

For some applications it might be appropriate to additionally introduce a maximum coincidence window τ_{\max} as a parameter. This way additional knowledge about the data (such as typical propagation speed) can be taken into account in order to guarantee that two coincident spikes are really part of the same propagation front. Such a maximum coincidence window will be used in the application to the El Niño climate data in section 3.3.

2.2. SPIKE-synchronization

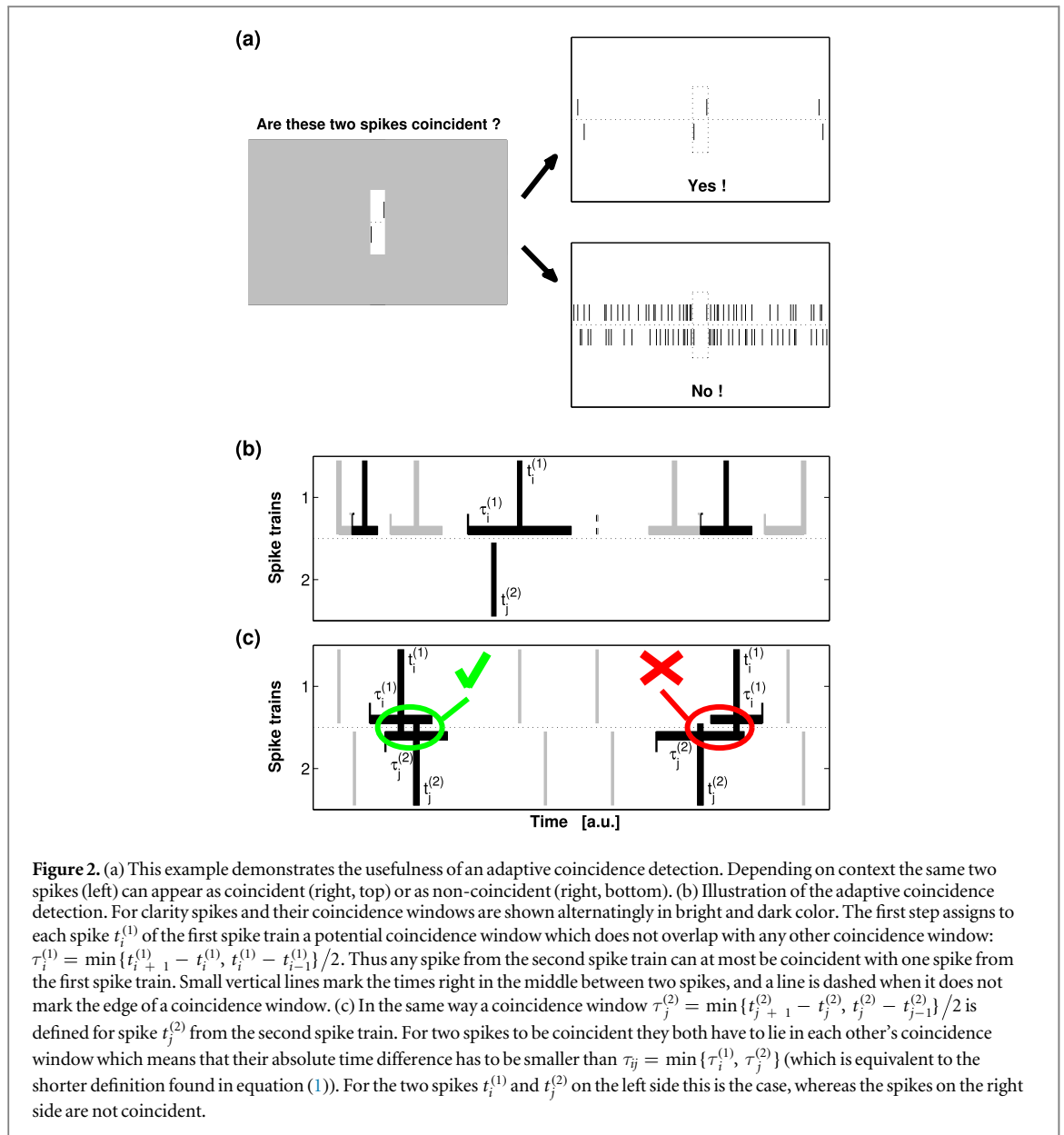
In normalization and windowing SPIKE-synchronization [27] has evolved so substantially from event synchronization that here we refrain from going into any detail on the original measure, but rather just mention the main improvements. For a thorough introduction to event synchronization please refer to the original paper [25], a more detailed comparison of the two measures can be found in [27].

The main difference is that SPIKE-synchronization [27] results in a discrete, not a continuous, spike-timing based profile. The coincidence criterion is quantified by means of a coincidence indicator

$$C_i^{(1,2)} = \begin{cases} 1 & \text{if } \min_j (|t_i^{(1)} - t_j^{(2)}|) < \tau_{ij}^{(1,2)} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

which assigns to each spike either a one or a zero depending on whether this spike is part of a coincidence or not. Note that here, unlike for event synchronization, the minimum function and the ‘<’ guarantee that a spike can at most be coincident with one spike (the nearest one) in the other spike train. In case a spike is right in the middle between two spikes from the other spike train there is no ambiguity since this spike is not coincident with either one of them.

This unambiguity, illustrated in figure 2(b), is the essential property which allows the adaptive coincidence detection to act as a match-maker for the subsequent application of SPIKE-synchronization. Figure 2(c) shows examples, one with two coincident and one with two non-coincident spikes.



A multivariate version of SPIKE-synchronization can be defined by generalizing the bivariate coincidence detection of equation (2) to all pairs of spike trains (n, m) with $n, m = 1, \dots, N$ and N denoting the number of spike trains:

$$C_i^{(n,m)} = \begin{cases} 1 & \text{if } \min_j (|t_i^{(n)} - t_j^{(m)}|) < \tau_{ij}^{(n,m)} \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

Here $\tau_{ij}^{(n,m)}$ is defined equivalent to equation (1). Subsequently, for each spike of every spike train a normalized coincidence counter

$$C_i^{(n)} = \frac{1}{N-1} \sum_{m \neq n} C_i^{(n,m)} \quad (4)$$

is obtained by averaging over all $N - 1$ bivariate coincidence indicators involving the spike train n .

In order to obtain a single multivariate similarity profile we pool the spikes of all the spike trains as well as their coincidence counters:

$$\{C_k\} = \bigcup_n \{C_i^{(n(k))}\}, \quad (5)$$

where we map the spike train indices n and the spike indices i into a global spike index k denoted by the mapping $i(k)$ and $n(k)$.

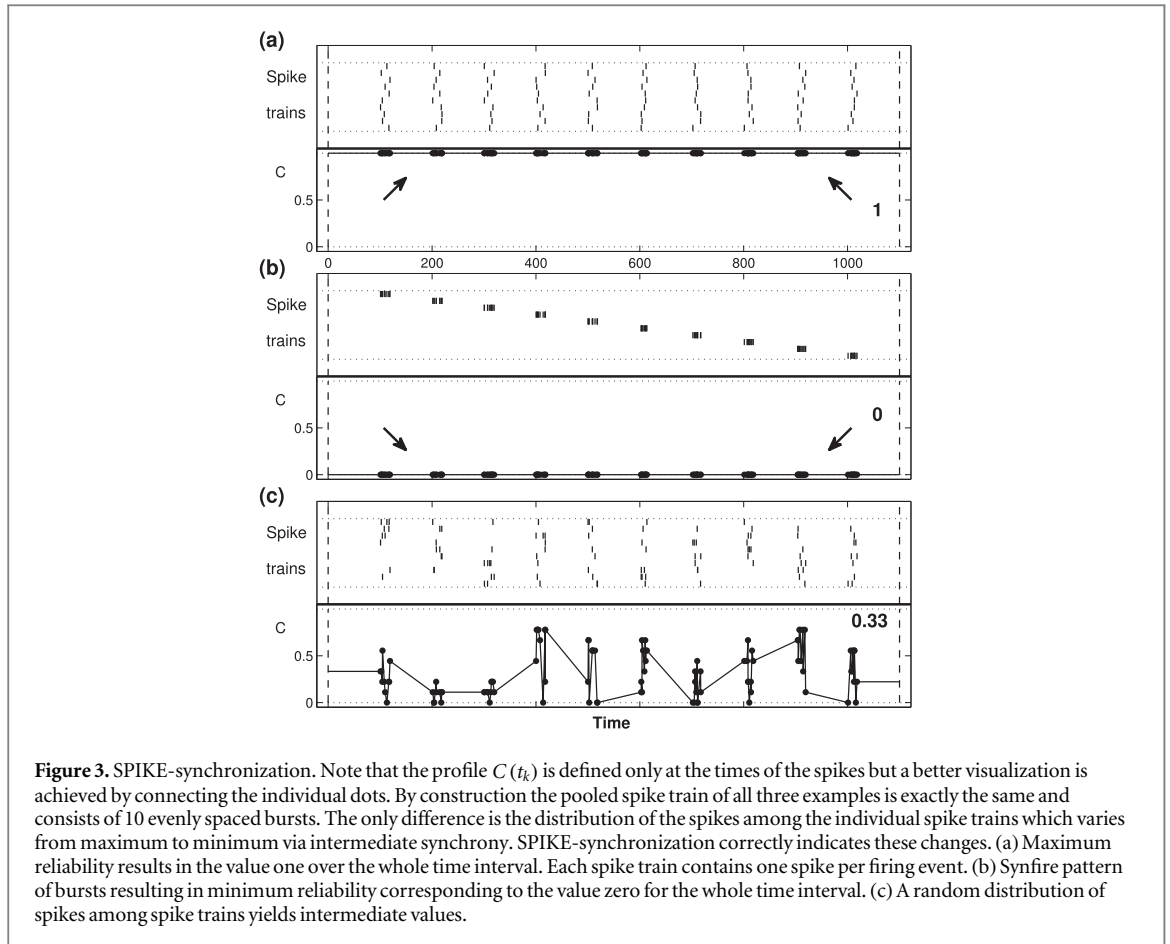


Figure 3. SPIKE-synchronization. Note that the profile $C(t_k)$ is defined only at the times of the spikes but a better visualization is achieved by connecting the individual dots. By construction the pooled spike train of all three examples is exactly the same and consists of 10 evenly spaced bursts. The only difference is the distribution of the spikes among the individual spike trains which varies from maximum to minimum via intermediate synchrony. SPIKE-synchronization correctly indicates these changes. (a) Maximum reliability results in the value one over the whole time interval. Each spike train contains one spike per firing event. (b) Synfire pattern of bursts resulting in minimum reliability corresponding to the value zero for the whole time interval. (c) A random distribution of spikes among spike trains yields intermediate values.

Note that in case there exist perfectly coincident spikes, k counts over all of these spikes. From this discrete set of coincidence counters C_k the SPIKE-synchronization profile $C(t_k)$ is obtained via $C(t_k) = C_k$. Finally, SPIKE-synchronization is defined as the average value of this profile

$$S_C = \begin{cases} \frac{1}{M} \sum_{k=1}^M C(t_k) & \text{if } M > 0 \\ 1 & \text{otherwise} \end{cases} \quad (6)$$

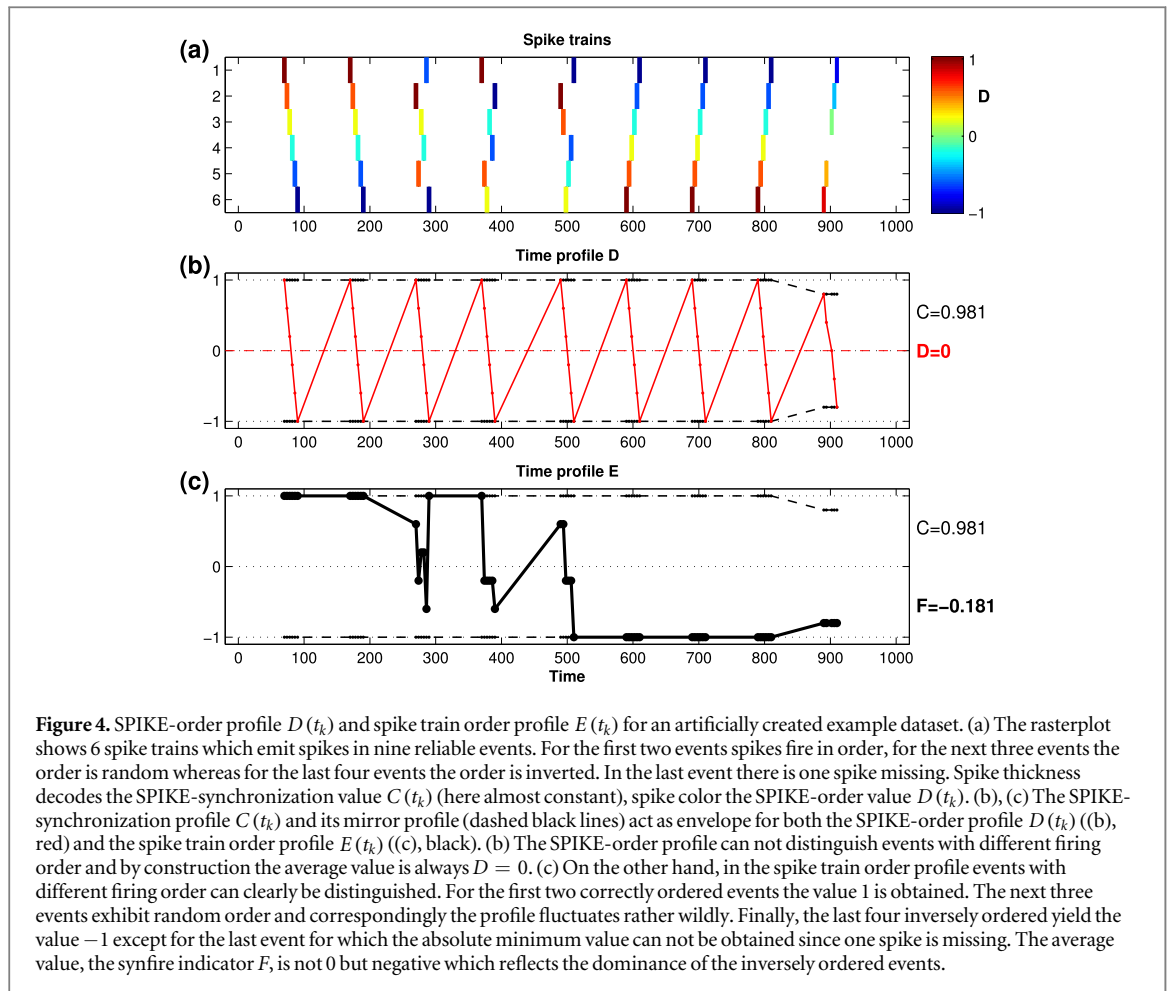
with $M = \sum_{n=1}^N M^{(n)}$ denoting the total number of spikes in the pooled spike train.

This way we have used the same consistent framework for both the bivariate and the multivariate case. The former is just a special case of the latter. The interpretation is very intuitive: SPIKE-synchronization quantifies the overall fraction of coincidences. It reaches one if and only if each spike in every spike train has one matching spike in all the other spike trains (or if there are no spikes at all), and it attains the value zero if and only if the spike trains do not contain any coincidences. Examples for both of these extreme cases can be found in figures 3(a) and (b) and one intermediate example (random distribution of spikes among spike trains) is shown in figure 3(c). For a derivation of the expectation value for Poisson spike trains please refer to [30].

In the multivariate analysis proposed in this paper, SPIKE-synchronization can be used to filter the input to the algorithm. In order to focus on propagation patterns within truly global events it is possible to set a threshold value C_{thr} for the SPIKE-synchronization profile $C(t_k)$. This way only spikes with a coincidence value higher than this parameter C_{thr} are taken into account, all the other noisy background spikes are simply ignored. This kind of filter will be used in the analysis of the neurophysiological datasets in section 3.2.

2.3. SPIKE-order and spike train order

SPIKE-synchronization assigns to each spike of a given spike train pair a bivariate coincidence indicator. These coincidence indicators $C_i^{(n,m)}$, which are either 0 or 1, are then averaged over spike train pairs and converted into one overall profile $C(t_k)$ normalized between 0 and 1. In exactly the same manner SPIKE-order and spike train order assign bivariate order indicators to spikes. Also these two order indicators, the asymmetric $D_i^{(n,m)}$ and the symmetric $E_i^{(n,m)}$, which both can take the values -1 , 0 , or $+1$, are averaged over spike train pairs and converted into two overall profiles $D(t_k)$ and $E(t_k)$ which are normalized between -1 and 1 . The SPIKE-order profile



$D(t_k)$ distinguishes leading and following spikes, whereas the spike train order profile $E(t_k)$ provides information about the order of spike trains, i.e. it allows to sort spike trains from leaders to followers.

First of all, similar to the transition from the symmetric event synchronization to delay asymmetry, the symmetric coincidence indicator $C_i^{(n,m)}$ of SPIKE-synchronization (equation (3)) is replaced by the asymmetric SPIKE-order indicator

$$D_i^{(n,m)} = C_i^{(n,m)} \cdot \text{sign}(t_{j'}^{(m)} - t_i^{(n)}), \quad (7)$$

where the index j' is defined from the minimum in equation (2) as $j' = \arg \min_j (|t_i^{(1)} - t_j^{(2)}|)$.

The corresponding value $D_{j'}^{(m,n)}$ is obtained in an antisymmetric manner as

$$D_{j'}^{(m,n)} = C_{j'}^{(m,n)} \cdot \text{sign}(t_i^{(n)} - t_{j'}^{(m)}) = -D_i^{(n,m)}. \quad (8)$$

Therefore, this indicator assigns to each spike either a 1 or a -1 depending on whether the respective spike is leading or following a coincident spike from the other spike train. The value 0 is obtained for cases in which there is no coincident spike in the other spike train ($C_i^{(n,m)} = 0$), but also in cases in which the times of the two coincident spikes are absolutely identical ($t_{j'}^{(m)} = t_i^{(n)}$).

The multivariate profile $D(t_k)$ obtained analogously to equation (5) is normalized between 1 and -1 and the extreme values are obtained if a spike is either leading ($+1$) or following (-1) coincident spikes in all other spike trains. It can be 0 either if a spike is not part of any coincidences or if it leads exactly as many spikes from other spike trains in coincidences as it follows. From the definition in equations (7) and (8) it follows immediately that C_k is an upper bound for the absolute value $|D_k|$.

While the SPIKE-order profile can be very useful for color-coding and visualizing local spike leaders and followers (figure 4(a)), it is not useful as an overall indicator of spike train order (figure 4(b)). The profile is invariant under exchange of spike trains, i.e. it looks the same for all events no matter what the order of the firing is (in our example only the last event looks slightly different since one spike is missing). Moreover, summing over all profile values, which is equivalent to summing over all coincidences, necessarily leads to an average value of 0, since for every leading spike ($+1$) there has to be a following spike (-1).

So in order to quantify any kind of leader-follower information between spike trains we need a second kind of order indicator. The spike train order indicator is similar to the SPIKE-order indicator defined in equations (7) and (8) but with two important differences. Both spikes are assigned the same value and this value now depends on the order of the spike trains:

$$E_i^{(n,m)} = C_i^{(n,m)} \cdot \begin{cases} \text{sign}(t_j^{(m)} - t_i^{(n)}) & \text{if } n < m \\ \text{sign}(t_i^{(n)} - t_j^{(m)}) & \text{if } n > m \end{cases} \quad (9)$$

and

$$E_{j'}^{(m,n)} = E_i^{(n,m)}. \quad (10)$$

This symmetric indicator assigns to both spikes a +1 in case the two spikes are in the correct order, i.e. the spike from the spike train with the lower spike train index is leading the coincidence, and a -1 in the opposite case. Once more the value 0 is obtained when there is no coincident spike in the other spike train or when the two coincident spikes are absolutely identical.

The multivariate profile $E(t_k)$, again obtained similarly to equation (5), is also normalized between 1 and -1 and the extreme values are obtained for a coincident event covering all spike trains with all spikes emitted in the order from first (last) to last (first) spike train, respectively (see the first two and the last four events in figure 4). It can be 0 either if a spike is not a part of any coincidences or if the order is such that correctly and incorrectly ordered spike train pairs cancel each other. Again, C_k is an upper bound for the absolute value of E_k .

2.4. Synfire indicator

In contrast to the SPIKE-order profile D_k , for the spike train order profile E_k it does make sense to define an average value, which we term the synfire indicator:

$$F = \frac{1}{M} \sum_{k=1}^M E(t_k). \quad (11)$$

The interpretation is very intuitive. The synfire indicator F quantifies to what degree the spike trains in their current order resemble a perfect synfire pattern. It is normalized between 1 and -1 and attains the value 1 (-1) if the spike trains in their current order form a perfect (inverse) synfire pattern. This means that all spikes are coincident with spikes in all other spike trains and that all orders from leading (following) to following (leading) spike consistently reflect the order of the spike trains. The synfire indicator is 0 either if the spike trains do not contain any coincidences at all or if among all spike trains there is a complete symmetry between leading and following spikes.

The spike train order profile $E(t_k)$ for our example is shown in figure 4(c). In this case the order of spikes within an event clearly matters. The synfire indicator F is slightly negative indicating that the current order of the spike trains is actually closer to an inverse synfire pattern.

Given a set of spike trains we now would like to sort the spike trains from leader to follower such that the set comes as close as possible to a synfire pattern. To do so we have to maximize the overall number of correctly ordered coincidences and this is equivalent to maximizing the synfire indicator F . However, it would be very difficult to achieve this maximization by means of the multivariate profile $E(t_k)$. Clearly, it is more efficient to sort the spike trains based on a pairwise analysis of the spike trains. The most intuitive way is to use the anti-symmetric cumulative SPIKE-order matrix

$$D^{(n,m)} = \sum_i D_i^{(n,m)} \quad (12)$$

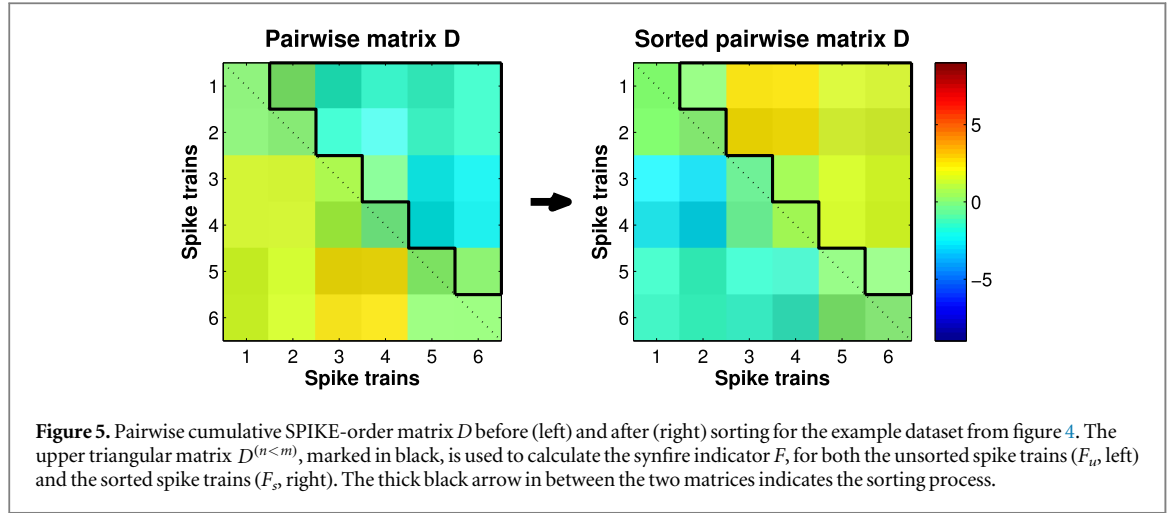
which sums up orders of coincidences from the respective pair of spike trains only and quantifies how much spike train n is leading spike train m (figure 5).

Hence if $D^{(n,m)} > 0$ spike train n is leading m , while $D^{(n,m)} < 0$ means m is leading n . If the current spike train order is consistent with the synfire property, we thus expect that $D^{(n,m)} > 0$ for $n < m$ and $D^{(n,m)} < 0$ for $n > m$. Therefore, we construct the overall SPIKE-order as

$$D_{<} = \sum_{n < m} D^{(n,m)}, \quad (13)$$

i.e. the sum over the upper right tridiagonal part of the matrix $D^{(n,m)}$.

After normalizing by the overall number of possible coincidences, we arrive at a second more practical definition of the synfire indicator:



$$F = \frac{2D_{<}}{(N-1)M}. \quad (14)$$

The value is identical to the one of equation (11), only the temporal and the spatial summation of coincidences (i.e., over the profile and over spike train pairs) are performed in the opposite order.

Having such a quantification depending on the order of spike trains, we can introduce a new ordering in terms of the spike train index permutation $\varphi(n)$. The overall synfire indicator for this permutation is then denoted as F_φ . Accordingly, for the initial (unsorted) order of spike trains φ_u the synfire indicator is denoted as $F_u = F_{\varphi_u}$.

The aim of the analysis is now to find the optimal (sorted) order φ_s as the one resulting in the maximal overall synfire indicator $F_s = F_{\varphi_s}$:

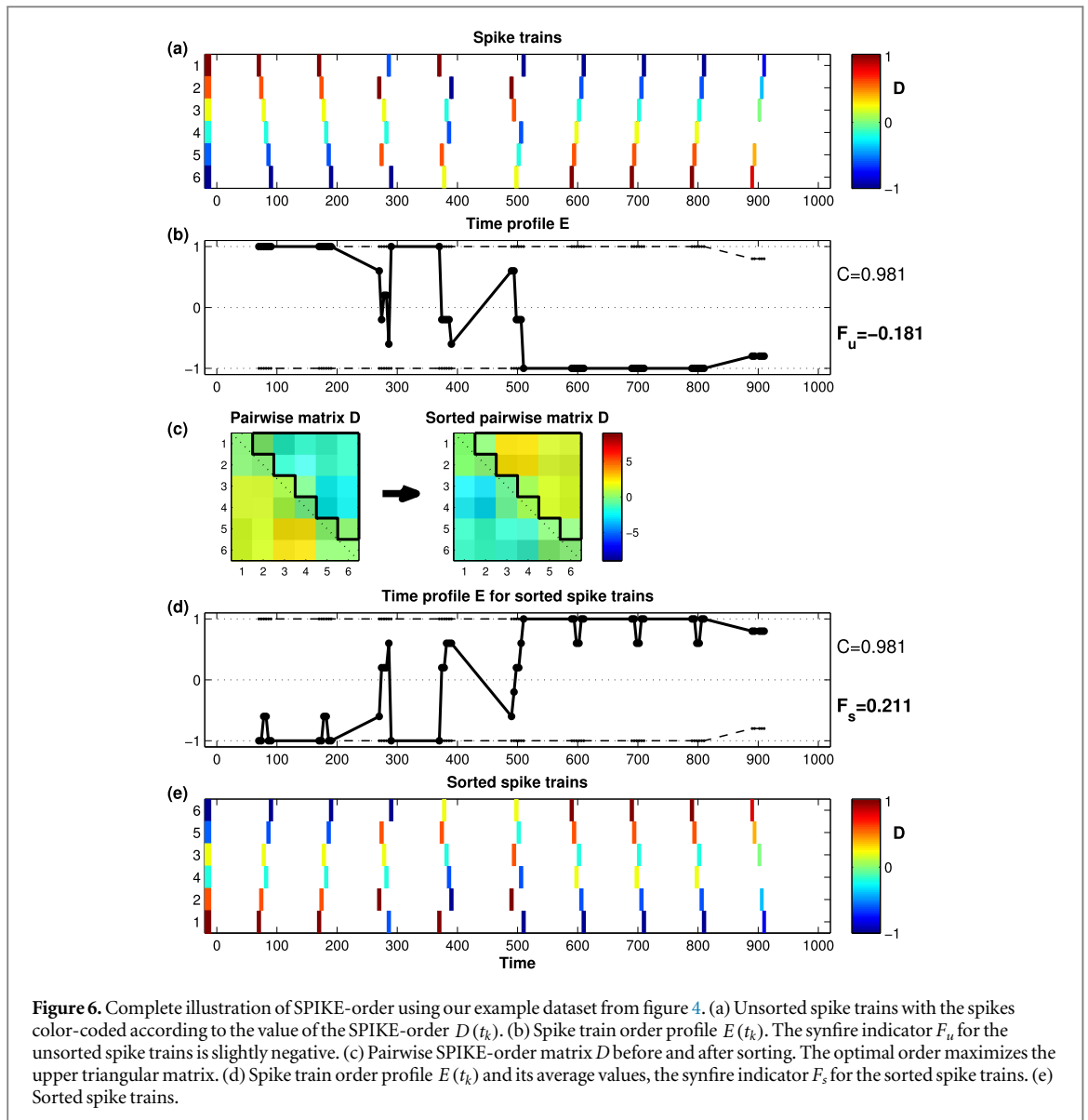
$$\varphi_s: F_{\varphi_s} = \max_{\varphi} \{F_\varphi\} = F_s. \quad (15)$$

This synfire indicator for the sorted spike trains quantifies how close spike trains can be sorted to resemble a synfire pattern, i.e., to what extent coinciding spike pairs with correct order prevail over coinciding spike pairs with incorrect order. Unlike the synfire indicator for the unsorted spike trains F_u , the optimized synfire indicator F_s can only attain values between 0 and 1 (any order that yields a negative result could simply be reversed in order to obtain the same positive value). For a perfect synfire pattern we obtain $F_s = 1$, while sufficiently long Poisson spike trains without any synfire structure yield $F_s \approx 0$.

The complexity of the problem to find the optimal spike train order is similar to the well-known traveling salesman problem [31]. For N spike trains there are $N!$ permutations φ , so for large numbers of spike trains finding the optimal spike train order φ_s is a non-trivial problem and brute-force methods such as calculating the F_φ -value for all possible permutations are not feasible. Instead, one has to make use of methods such as parallel tempering [32] or simulated annealing [33] to search for the optimal order. Here we choose simulated annealing, a probabilistic technique which approximates the global optimum of a given function in a large search space. In our case this function is the synfire indicator F_φ (which we would like to maximize) and the search space is the permutation space of all spike trains. We start with the F_u -value from the unsorted permutation and then visit nearby permutations using the fundamental move of exchanging two neighboring spike trains within the current permutation. The update of the synfire indicator when exchanging the spike trains k and $k+1$ is simply given by $\Delta F = -2D^{(k,k+1)}$. All moves with positive ΔF are accepted while the likelihood of accepting moves with negative ΔF is decreased along the way according to a standard slow cooling scheme. The procedure is repeated iteratively until the order of the spike trains no longer changes or until a predefined end temperature is reached.

In figure 6 we show the complete SPIKE-order analysis including the results for the sorted spike trains. The sorting of the spike trains maximizes the synfire indicator as reflected by both the normalized sum of the upper right half of the pairwise cumulative SPIKE-order matrix (equation (14), figure 6(c)) and the average value of the spike train order profile $E(t_k)$ (equation (11), figure 6(d)). Finally, the sorted spike trains in figure 6(e) are now ordered such that the first spike trains have predominantly high values (red) and the last spike trains predominantly low values (blue) of $D(t_k)$.

The complete analysis returns results consisting of several levels of information. Time-resolved (local) information is represented in the spike-coloring and in the profiles D and E . The pairwise information in the SPIKE-order matrix reflects the leader-follower relationship between two spike trains at a time. The synfire indicator F characterizes the closeness of the dataset as a whole to a synfire pattern, both for the unsorted (F_u)



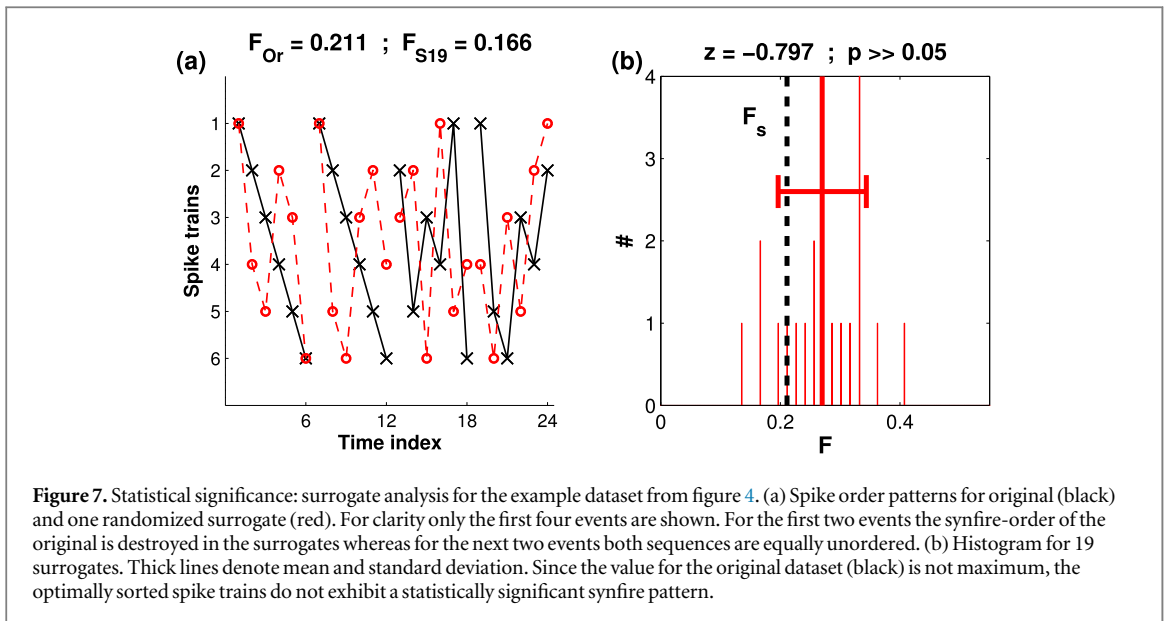
and for the sorted (F_s) spike trains. Finally, the sorted order of the spike trains is a very important result in itself since it identifies the leading and the following spike trains.

2.5. Statistical significance

As a last step in the analysis we evaluate the statistical significance of the optimized synfire indicator F_s . What we would like to estimate is the likelihood that for the given total number of coincidences the prevalence of correctly ordered spike pairs (as quantified by the optimized synfire indicator) could have been obtained by chance. If all coincident spike pairs would be independent, the probability distribution would be strictly binomial and we could calculate this likelihood analytically. However, the pairwise spike orders in coincident events involving multiple spike trains are not independent from each other, and so instead we estimate the likelihood numerically using a set of carefully constructed spike order surrogates.

For each surrogate (figure 7(a)) we maintain the coincidence structure of the original spike trains by preserving the SPIKE-synchronization values of every individual spike. However, we destroy the spike order patterns by swapping the order of the two spikes in a sufficient number of randomly selected coincident spike pairs. Note that the generation of surrogates takes place not on the level of spike times but on the level of order values (the x -axis in figure 7(a) is labeled ‘time index’, not ‘time’). Spike trains with swapped spike times would have different interspike intervals, and this would alter the results of the coincidence criterion in equation (1) and change the value of SPIKE-synchronization. This in turn would make the desired evaluation of pure spike order effects difficult.

In the implementation, from one spike order surrogate to the next the number of spike order swaps is set to the number of coincident spikes in the spike train set, such that all possible spike order patterns can be reached.



Only for the first surrogate, since it starts from the original spike trains, we swap twice as many coincidences in order to account for transients. After each swap we take extra care that all other spike orders that are affected by the swap are updated as well. For example, if a swap changes the order between the first and the third spike in an ordered sequence of three spikes, we also swap both the order between the first and the second as well as the order between the second and the third spike.

For each spike train surrogate we repeat exactly the same optimization procedure in the spike train permutation space that is done for the original dataset. The original synfire indicator is deemed significant if it is higher than the synfire indicator obtained for all of the surrogate datasets (this case will be marked by two asterisks). Here we use $s = 19$ surrogates for a significance level of $p^* = 1/(s + 1) = 0.05$. As a second indicator we state the z -score, e.g., the deviation of the original value x from the mean μ of the surrogates in units of their standard deviation σ :

$$z = \frac{x - \mu}{\sigma}. \quad (16)$$

Results of the significance analysis for our standard example are shown in the histogram in figure 7(b). In this case the absolute value of the z -score is smaller than one and the p -value is larger than p^* and the result is thus judged as statistically non-significant.

In case the initial sorting of the spike trains is used to test a specific hypothesis there also exists a straightforward procedure to test the statistical significance of the synfire indicator F_u for the unsorted spike trains. In this case no optimization of the synfire indicator is required, rather the synfire indicator F_u for the initial sorting is compared against synfire indicators obtained for random permutations of the spike trains. This kind of significance test will be used in section 3.2.

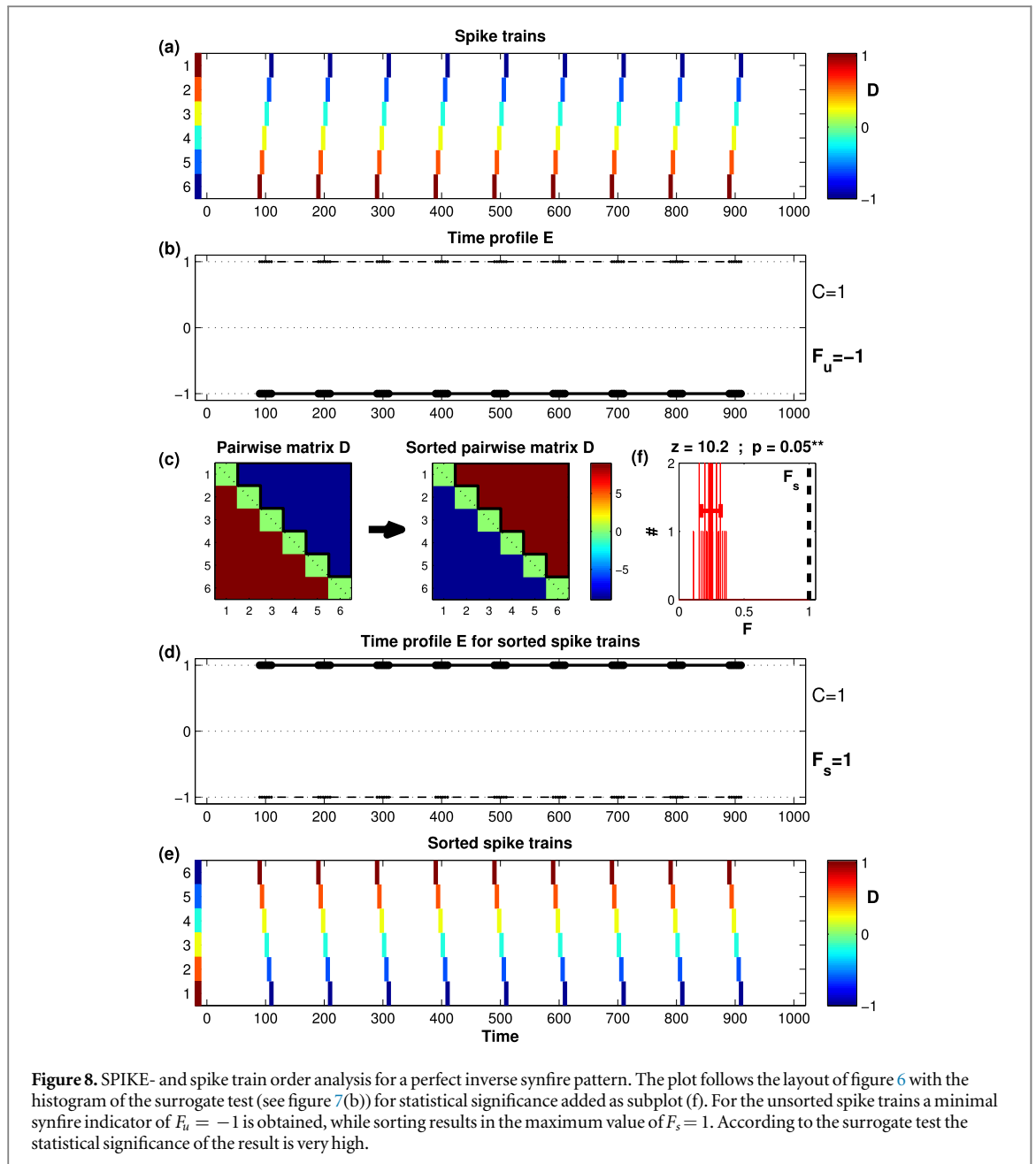
3. Results

In the following we apply our new algorithm to artificially generated datasets (section 3.1), neurophysiological data (section 3.2) and, finally, climate data from the El Niño phenomenon (section 3.3). For the figures we use the same full layout introduced in figure 6 to which we add the significance analysis of figure 7(b).

3.1. Application to artificially generated data

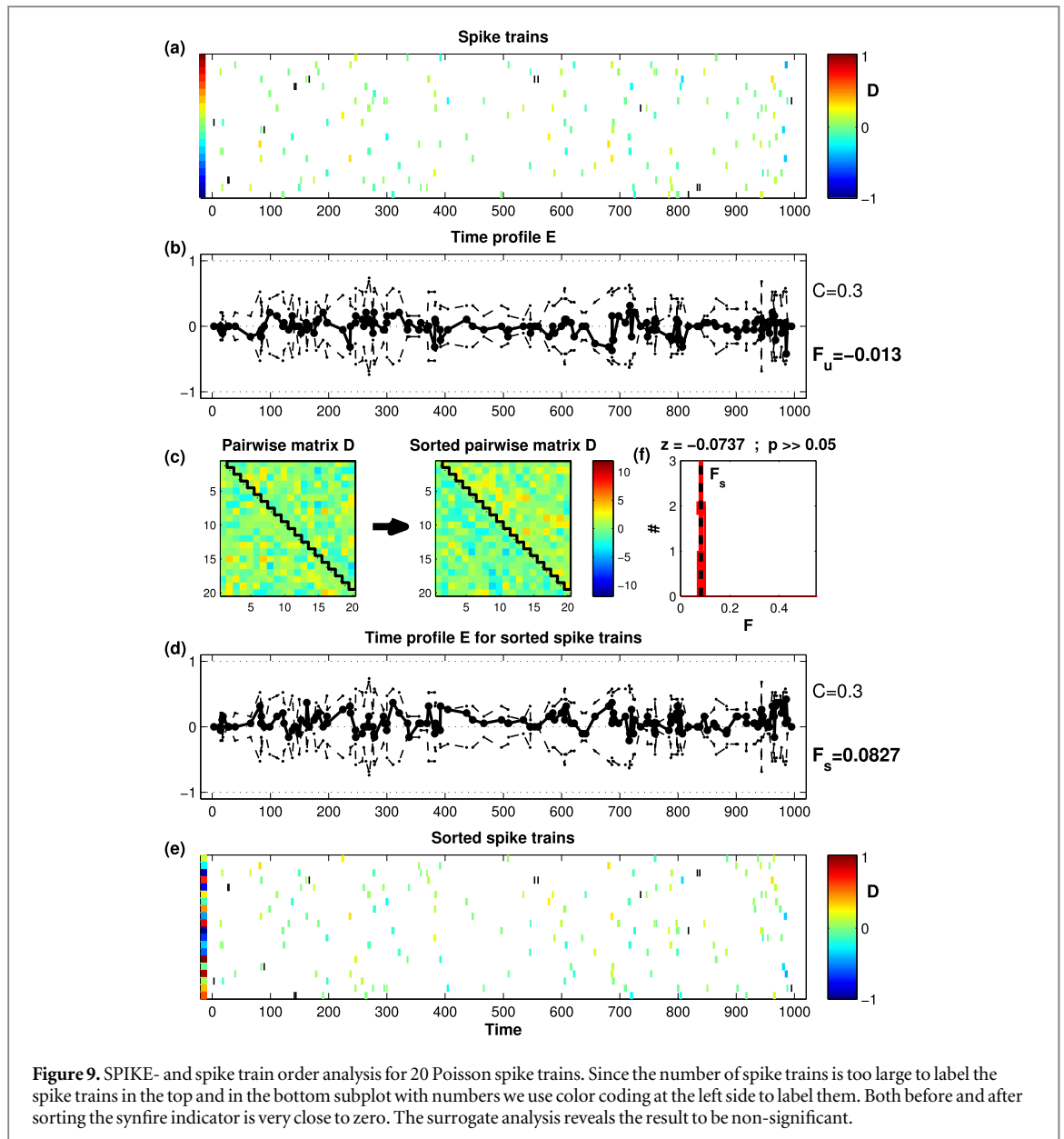
We start with examples covering the two extreme cases of a perfect synfire pattern and a completely random spike train set. First, in figure 8 we apply the algorithm to a perfect inverse synfire pattern for which the spike trains are initially sorted from follower to leader. Therefore, the synfire indicator of the unsorted spike trains yields its minimum value of $F_u = -1$. Sorting just reverses the order of the spike trains and in consequence the maximum value of $F_s = 1$ is obtained. Any shuffling of spike orders necessarily destroys the synfire pattern and thus leads to much lower values of the synfire indicator. Accordingly, the surrogate test (figure 8(f)) shows that the statistical significance of the original synfire indicator is very high.

The other extreme case is Poisson spike trains (figure 9) for which the arrival times of spikes are completely random and without any preferred order. For this realization the synfire indicator F_u for the unsorted spike



trains happens to be slightly negative indicating that the spike trains are closer to an inverse synfire pattern than to a synfire pattern. The absolute value F_s after sorting is higher. The fact that both of these values are non-zero is due to the finite size effect caused by the limited number of spikes. For more and more spike trains and/or spikes the expectation value even for the sorted case would converge towards zero. As expected, the surrogate test shows that the order for the original spike trains is not statistically distinct from the order of the surrogate spike trains (there is no preferred order that can be destroyed by the shuffling) and, accordingly, the value of the original synfire indicator is revealed to be clearly non-significant.

The third example in figure 10 shows a mixture of these two extremes, Poisson spike train interspersed with spike trains that contain a perfect inverse synfire pattern (plus random spikes). Sorting the spike trains restores the correct order of the synfire pattern spike trains within the Poisson spike train. The synfire indicator for the sorted spike trains F_s for this mixed example is actually almost identical to the value obtained for the Poisson spike trains in figure 9, but this time the surrogate test reveals the value to be highly significant. These two examples combined illustrate nicely that the synfire indicator and the surrogate analysis provide complementary information. In the mixture example of figure 10 there are many more random Poisson spikes than ordered synfire pattern spikes. According to the synfire indicator, these two types of spikes together appear to be as ordered as the spikes of the shorter but purely random Poisson spike trains in figure 9. However, the synfire indicator is strongly influenced by the statistics of the dataset and thus is in itself not sufficient to reliably

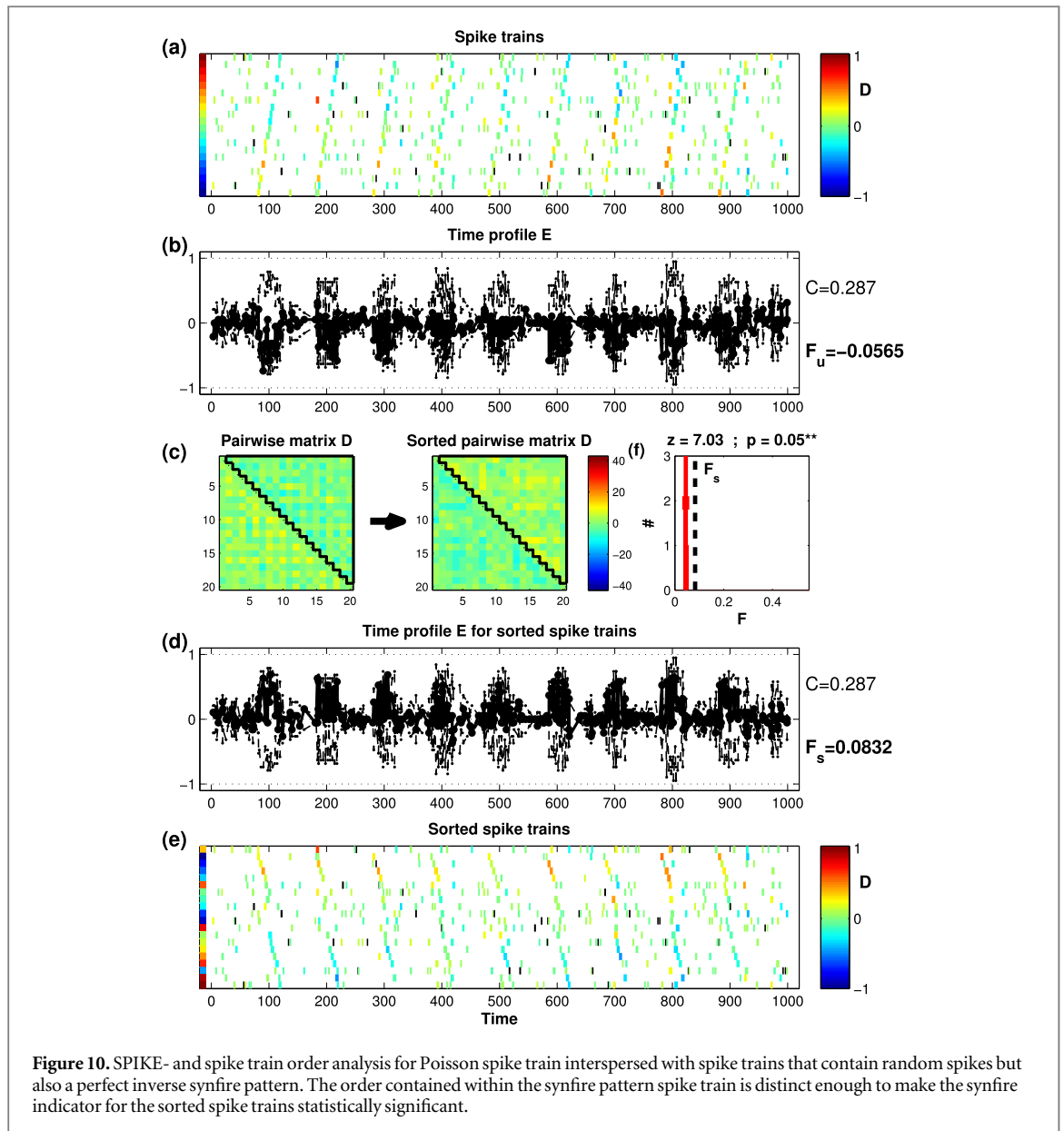


compare two datasets with widely different number of spike trains and spikes. The surrogate analysis, on the other hand, can be used to compare datasets of different size since by preserving the spike numbers in the surrogates it explicitly takes the statistics of each dataset into account.

3.2. Application to neuroscience

In order to apply the spike train order algorithm to real neurophysiological data, we analyzed data recorded via fast multicellular calcium imaging in acute CA3 hippocampal brain slices from juvenile mice. In the juvenile hippocampus, the CA3 region is the origin of a stereotypical network phenomenon of wavelike propagating activity termed giant depolarizing potentials (GDPs [34]). In previous studies, GDPs have been used to investigate the topology of networks and the role of hub cells [23] as well as to reveal the deterministic and stochastic processes underlying spontaneous, synchronous network bursts [35]. Due to the distinct architecture and the repetitive nature of the GDPs this experimental setup offers a very suitable test case for our synfire pattern analysis (for more background and a detailed description of the experimental methods refer to appendix A).

The first dataset analyzed in figure 11 includes 13 GDPs over a bit more than 6 min. Almost all GDPs involve the whole network. Here, as for all other neurophysiological datasets analyzed, initially the spike trains are sorted according to their firing rate such that the sparsely spiking neurons are on top and the most active neurons at the bottom (figure 11(a)). This specific sorting allows us to test the hypothesis that the neurons which fire almost exclusively within the GDPs and are very sparse on background activity might have a stronger role in initiating



GDPs and tend to lead, whereas the more regularly spiking neurons might tend to follow. If this would be the case one would expect a very high value for the initial synfire indicator F_u . However, according to figure 11(b) the actual value is very close to zero and actually slightly negative. A statistical significance test using random permutations of spike trains (see section 2.5) indeed proves the synfire indicator of the unsorted spike trains F_u to be non-significant (result not shown). A further indicator for this is the fact that the order of the sorted spike trains is very different from the initial order, as can be seen by comparing the color bars on the left of figures 11(a) and (e). The color-coding of the GDPs exhibits typically a slightly noisy transition from leader (red) to follower (blue). The synfire indicator for the sorted spike trains F_s is also much higher (figure 11(d)). Finally, the surrogate analysis (figure 11(f)) shows this result to be highly significant.

However, the spiking in figure 11 consists not only of the GDPs. Most neurons exhibit at least to some extent spontaneous background activity, the ones at the top of the initial sorting less than the ones at the bottom. The spikes in this background activity are typically coincident with only few other spikes and do not take part in any propagation patterns (note their green color which indicates SPIKE-order values close to zero). So in the context of our synfire pattern analysis this is just noise that leads to a decrease of the synfire indicator. There is a straightforward way to disregard these background spikes by setting a threshold value C_{thr} for the SPIKE-synchronization profile $C(t_k)$. Only spikes with a coincidence value higher than C_{thr} are taken into account, all other spikes are simply ignored. The result of this background correction can be seen in figure 12 for the same dataset already used in figure 11. Focusing the analysis on the reliable GDPs leads to an increase of the synfire indicator from 0.284 to 0.438.

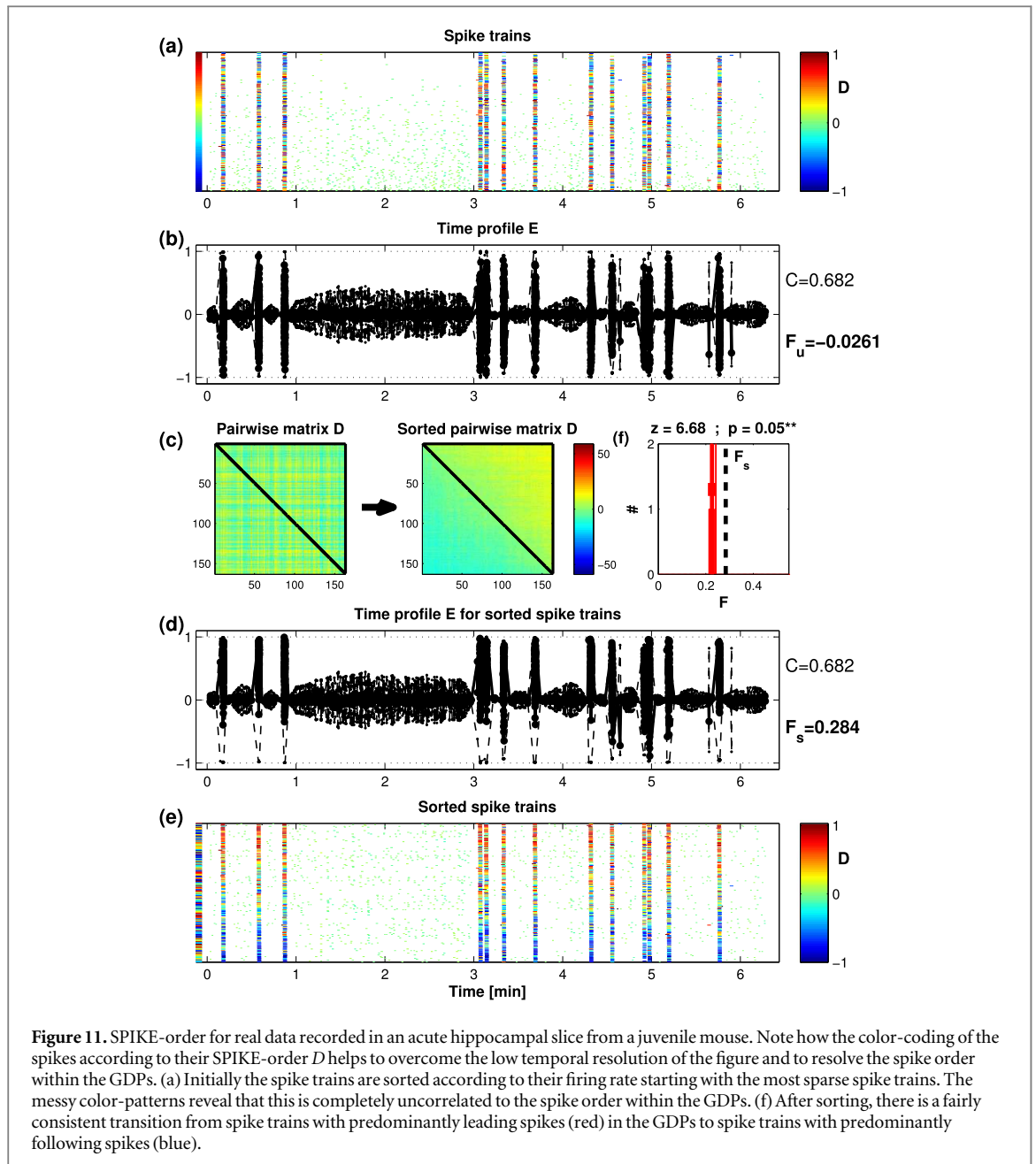


Figure 11. SPIKE-order for real data recorded in an acute hippocampal slice from a juvenile mouse. Note how the color-coding of the spikes according to their SPIKE-order D helps to overcome the low temporal resolution of the figure and to resolve the spike order within the GDPs. (a) Initially the spike trains are sorted according to their firing rate starting with the most sparse spike trains. The messy color-patterns reveal that this is completely uncorrelated to the spike order within the GDPs. (f) After sorting, there is a fairly consistent transition from spike trains with predominantly leading spikes (red) in the GDPs to spike trains with predominantly following spikes (blue).

As already mentioned before, one of the main results of our analysis is the sorted order of the spike trains itself. For these neurophysiological data it allows to identify the leading and the following neurons in the network and to project this information back on the recording setup. This is shown in figure 13 where we have color-coded the optimized spike train order obtained in figure 12 within a 2D-plot of the neurons recorded from the hippocampal slice. For this example there appears to be a clear overall propagation from right to left but there is also a considerable degree of variability which might be due to a non-trivial connectivity within the network.

In figure 14 we apply the SPIKE-order analysis to a second dataset recorded from a different slice, again focusing on the order within the global events only. Here we also added one new feature, the mean value of the spike train order $E(t_k)$ for each global event (we use the maxima and minima of the SPIKE order profile $D(t_k)$ to delineate the GDPs). This again emphasizes the time-resolved nature of the SPIKE order and the spike train order indicators.

Overall, we have analyzed neurophysiological datasets from four hippocampal slices exhibiting an average of 7.75 GDPs. We obtained an average value for SPIKE-synchronization of 0.59 before focusing on the GDPs (as in figure 11) and 0.92 after (as in figure 12). With or without this focus the synfire indicator for the initial spike train sorting F_u was very close to zero and in all cases proved to be non-significant when tested against synfire indicators obtained for random permutations of the spike train order. Since the initial sorting was based on

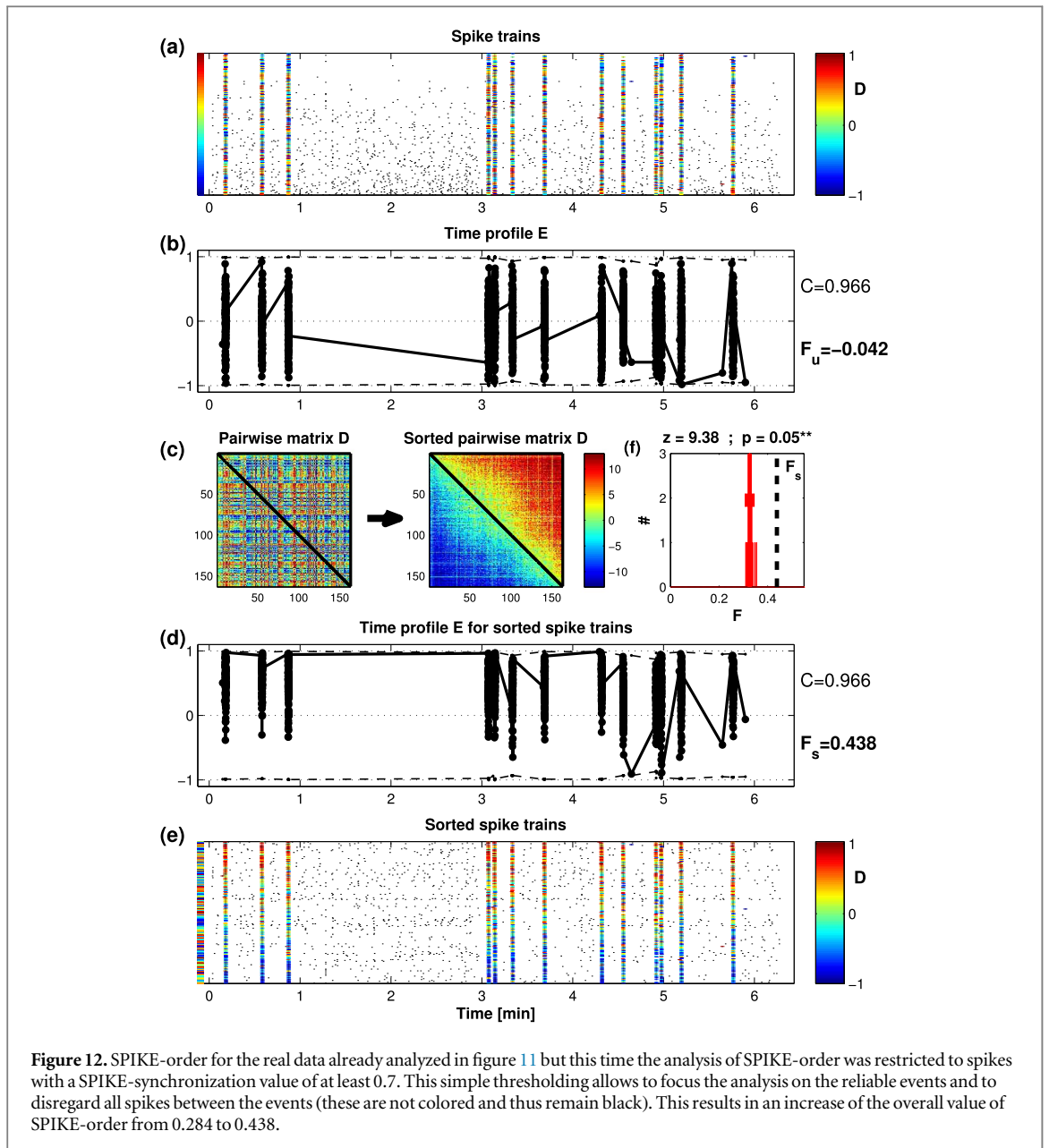
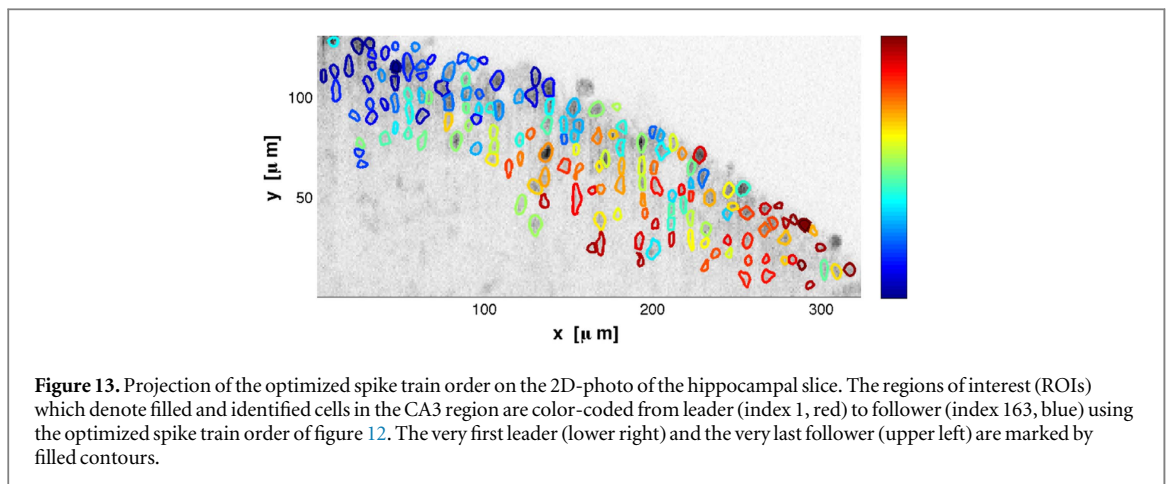
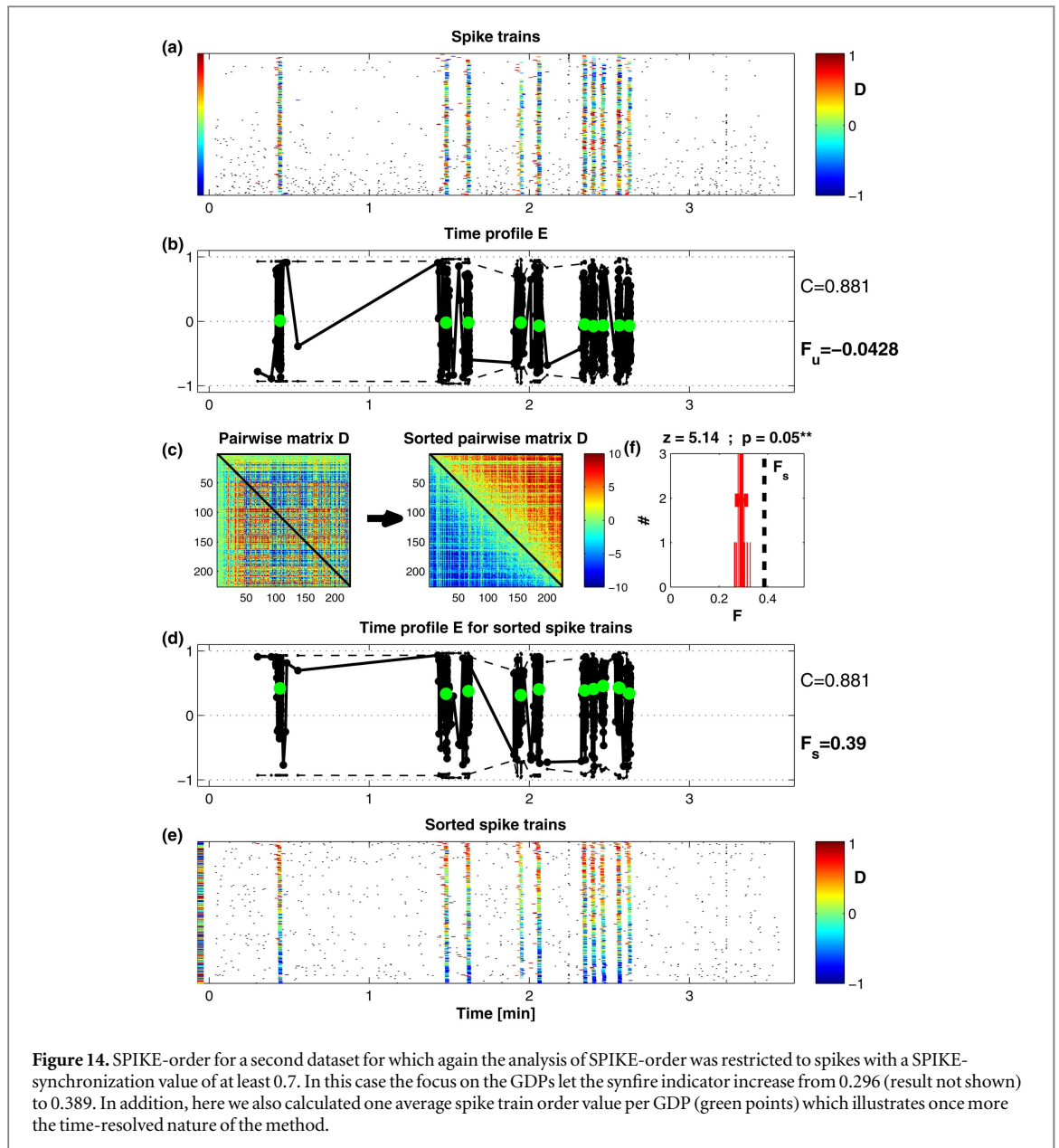


Figure 12. SPIKE-order for the real data already analyzed in figure 11 but this time the analysis of SPIKE-order was restricted to spikes with a SPIKE-synchronization value of at least 0.7. This simple thresholding allows to focus the analysis on the reliable events and to disregard all spikes between the events (these are not colored and thus remain black). This results in an increase of the overall value of SPIKE-order from 0.284 to 0.438.





overall firing rate of the neurons, this signifies that the hypothesis that the low-firing neurons which are basically only active during the GDPs might have a stronger role in initiating GDPs can be rejected. For the sorted spike trains the synfire indicator F_s was 0.20 for all spikes and 0.42 for the spikes within the GDPs only. Suppressing the effect of the noisy background spikes in the analysis thus leads to an average increase of the synfire indicator by about 110%. Finally, according to the surrogate analysis described in section 2.5 the synfire indicator for the sorted spike trains F_s yielded a statistically significant result for all datasets analyzed.

So overall we can conclude that the GDPs recorded in brain slices from juvenile mice are distinguished by a very high consistency of their spatio-temporal propagation patterns. However, it is interesting to note that this consistency does not hold when comparing different slices. In the datasets analyzed in this paper we find examples of both propagation in the direction of CA2 as well as propagation towards the dentate gyrus. This is consistent with results reported in [35].

3.3. Application to climate data

Although being developed mainly for neuroscientific data (spike train recordings), the spike train order approach presented in this paper can be applied in many other contexts as well. One particular field where event-based analysis is employed very frequently is climate science, see e.g. [20, 36].

In the following we will use the new spike train order method to analyze the sea surface temperature (SST) in the central and eastern tropical Pacific Ocean to identify the propagation patterns connected with the warm phase of the El Niño Southern Oscillation (ENSO). Predicting the occurrence and strength of El Niños is very

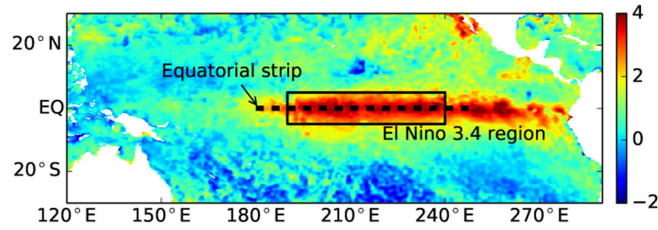


Figure 15. El Niño 3.4 region (rectangle) and the small strip around the equator used for the analysis in this work (dashed line). Color coding represents daily SST anomaly (in °C) on 11 November 2015.

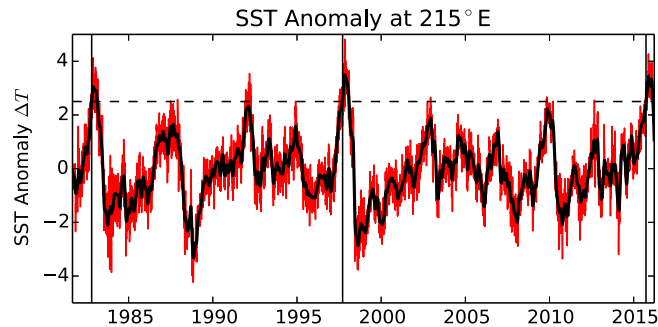
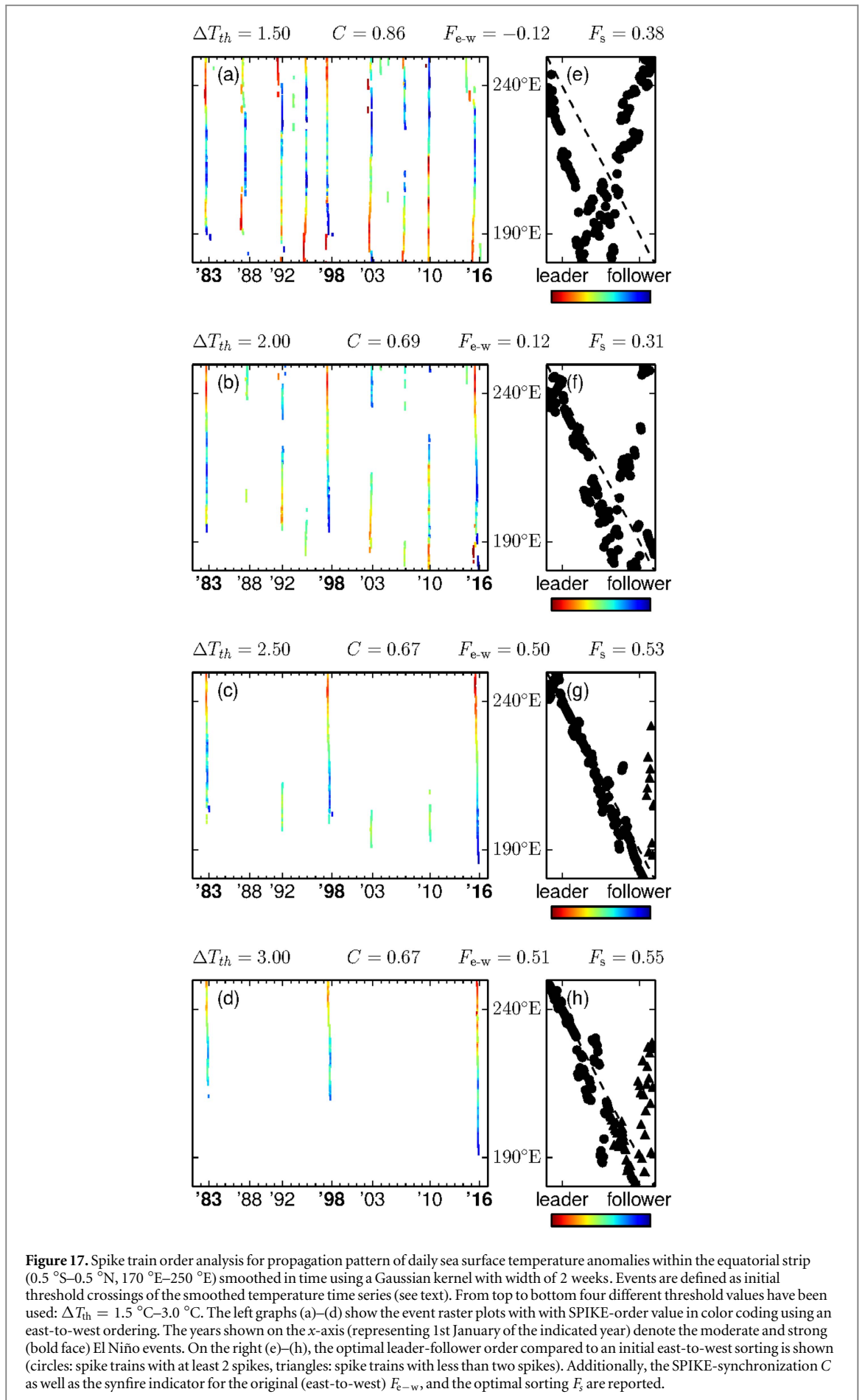


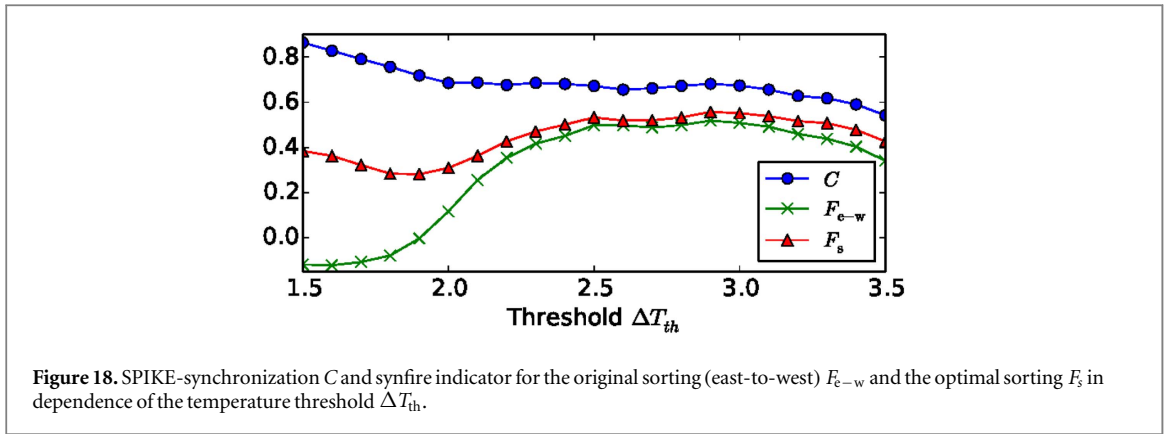
Figure 16. Exemplary time series of the SST anomaly at 215 °E before (red) and after (black) Gaussian smoothing. The horizontal dashed line corresponds to a temperature threshold of $\Delta T_{th} = 2.5$ °C used for event detection and the vertical lines are the resulting events identified from upward threshold crossings of the smoothed time series.

important due to the vast ecological and economical effects [37] and therefore this phenomenon has been studied extensively in terms of both intensity (e.g. [38, 39]) and frequency (e.g. [40, 41]). However, here we will not discuss El Niño prediction, but focus on the longitudinal propagation pattern of the El Niño events within the Pacific Ocean [24, 42]. The region used here to analyze the SST is depicted in figure 15 (dashed line), and an exemplary smoothed time profile for the center of the analyzed region (215 °E) is shown in figure 16. Further details on the region, the dataset and data preprocessing (smoothing) is outlined in appendix B.

Starting from the smoothed SST profiles, we performed an event detection by identifying the onset of the SST anomaly connected with an El Niño in terms of an *upward crossing* of the smoothed time series with a temperature anomaly threshold ΔT_{th} . As we are interested in the initial propagation front for each El Niño event, we additionally introduce an artificial *refractory period* of 11 months, which means threshold crossings occurring within 11 months after previous event are disregarded. Figure 16 depicts this procedure for a threshold value $\Delta T_{th} = 2.5$ °C leading to the identification of three events in this example (vertical lines in figure 16). From this, we finally arrive at $N = 140$ discrete event series corresponding to the onsets of El Niño SST anomaly elevations at the different longitudinal locations along the dashed line in figure 15. The resulting spike trains are depicted in figure 17 on the left for four different threshold values $\Delta T_{th} = 1.5$ °C–3.0 °C (step size of 0.5). The first thing to note is that this event detection procedure is able to correctly identify the El Niño occurrences as seen from the horizontal event structure in accordance with the El Niño years (labeled on the x-axis in the raster plots (a)–(d)) as well as the consistently high SPIKE-synchronization values C for all threshold values. With increasing temperature threshold, however, only the three strongest El Niño events are being captured (1983, 1998 and 2016, bold labels in figure 17) as seen from figures 17(c) and (d).

To quantify the consistency of the propagation patterns, we perform a spike train order analysis on these event sequences. The results are also presented in figure 17. As a central observation, we found that for the three strong El Niño events (1983, 1998, 2016), there is a clear consistent propagation of the SST anomaly elevation at the equator from east to west as confirmed by a synfire indicator of $F_{e-w} = 0.5$ for the east-to-west sorting ($\Delta T_{th} = 2.5$ °C, figure 17(c)). Furthermore, optimizing the synfire indicator by changing the sorting only leads to minor improvements ($F_s = 0.53$) and the resulting optimal order is very close to the original east-to-west sorting indicated in figure 17(g). Finally, this observation is also confirmed by the SPIKE-order based color coding in the raster plots (c) of the spike trains. For each of the three major El Niño events (1983, 1998, 2016) we see a clear tendency of leader (red) to follower (blue) going from east (top) to west (bottom) for all three threshold values. Further increase of the threshold temperature (figures 17(d) and (h)) leads to essentially the





same observation, but parts of the longitudinal propagation might be missed as the SST anomaly maxima stay below threshold.

Moderate El Niño events (1992, 2003, 2010), on the other hand, seem to propagate in eastwards direction, as seen by opposite directed coloring for those years in figure 17(b) ($\Delta T_{th} = 2.0$). Consequently, the synfire indicator for the east-to-west sorting is very low, $F_{e-w} = 0.12$, as some of the propagation events are moving in the opposite direction which reduces the synfire indicator. Therefore, it is also not possible to find a correct ordering, as indicated by a rather low optimized synfire indicator of $F_s = 0.31$ in this case. This is even more pronounced for the smallest threshold temperature $\Delta T_{th} = 1.5$, shown in figure 17(b). There, essentially all moderate El Niños are captured and most of those exhibit an eastwards propagation, opposed to the strong El Niños that propagate westwards. Therefore, the initial synfire indicator for the east-to-west sorting is negative, $F_{e-w} = -0.11$ and the optimal sorting shows the sign of a west-to-east propagation (figure 17(e)) with $F_s = 0.38$.

To better understand the dependency of the spike train order analysis on the temperature threshold, we computed the synfire indicator for the initial east-to-west ordering as well as the value for optimal ordering F_s for varying thresholds $\Delta T_{th} = 1.5 \dots 3.5$ (figure 18). Additionally, figure 18 also shows the SPIKE-synchronization values for these parameters. One first observes that the SPIKE-synchronization remains rather constant over the whole range of threshold values, indicating a consistently good identification of the El Niño events. The synfire indicator for the original east-to-west sorting F_{e-w} on the other hand shows a clear increase saturating in a plateau for a threshold value of around $\Delta T_{th} = 2.5$ °C, where it is then also very close to the optimal value F_s . This is easily understood by the fact that at temperature thresholds above 2.5 °C, only the three strong El Niños are captured that show a consistent westwards propagation, while for values below also weaker El Niños that exhibit eastward propagation enter the analysis.

In summary, we showed that the newly proposed spike train order method is also readily applicable to climate data and is able to identify propagation patterns in recurring events such as the El Niño. Our results indicate that close to the equator, strong El Niño events exhibit a clear westward propagation of an SST front, while for moderate El Niños we found indications for an eastward front propagation. Interestingly, previous results based on zonal current velocity within the El Niño 3.4 region found eastward propagation for the extreme El Niño events in 1983 and 1998 [24], while here we found west-ward propagation (figure 17). This discrepancy is most likely due to the different regions (El Niño 3.4 versus equatorial strip) and methodology (average warming/cooling rates versus front propagation).

4. Discussion

Over the last years a wide variety of measures to quantify the synchrony between spike trains have been introduced. Three recent proposals, *ISI-distance* [43, 44], *SPIKE-distance* [45, 46], and *SPIKE-synchronization* [27, 30], share the desirable property of being time-resolved and parameter-free (time-scale independent). However, their bivariate versions are symmetric and in consequence their multivariate versions are invariant to changes in the order of spike trains. None of these measures is designed to provide information about the directionality of the propagation patterns.

In the present study we address this issue. First we use an adaptive coincidence detection as match maker in order to identify pairs of coincident spikes. Then we define two measures, the asymmetric *SPIKE-order* D and the symmetric *spike train order* E , which are particularly useful in a bivariate representation (pairwise matrix) and as a time-resolved multivariate profile, respectively. From these two measures we can derive the *synfire indicator* F , a condensed scalar value that quantifies the overall consistency of the spatio-temporal propagation

patterns in a rigorous and automated way. Its maximization allows to sort multiple spike trains from leader to follower. This is meant purely in the sense of temporal sequence of the spikes. The question asked is: for which spike trains do spikes tend to occur first and for which do they tend to occur last? We use simulated annealing to search among all permutations of spike trains for the sorting that resembles as closely as possible a synfire pattern, a perfectly consistent repetition of the same global propagation pattern. In a final step we evaluate the statistical significance of the optimized permutation using a set of carefully constructed spike train surrogates.

We first illustrate the merits of our new approach using artificially generated datasets and then apply it to real datasets from two very different fields of research, neuroscience and climatology. In the neurophysiological dataset we analyze GDPs recorded from mice brain slices in order to search for potential hub neurons [23], whereas in the climate application we quantify the consistency of the longitudinal movement of the propagation front of El Niño events [24].

The new algorithm is conceptually simple, of low computational cost and comes with an intuitive and straightforward visualization, including a color-coded rasterplot. It substantially improves on all the bivariate functionalities of its predecessor directional measure *delay asymmetry* (no need for sampled profiles, more intuitive normalization etc) and could thus also be used in the context of the pairwise matrices, both normalized or cumulative, used in complex network theory [20, 47]. However, one of the main advantages of the new algorithm is its multivariate nature which opens up completely new kinds of application such as spike train sorting.

One important advantage that the method shares with other techniques of spike train analysis is the high flexibility in the definition of events. For example, when looking at the synchronization of neuronal bursts instead of individual spikes one can define the events as the onset, the center or the offset of the activity (e.g., the first, the middle or the last spike of each burst). In cases in which a burst of spikes is considered to be equivalent to a single spike one could introduce some kind of meta-events and then look at coincidences between these meta-events.

The application of our measures is also not restricted to truly discrete data. Continuously sampled data can be reduced to a spike train where the only information maintained is the timing of the individual events. Often these event times are obtained in a manner similar to how the neuronal spike times are extracted from recordings of neuronal membrane potentials (usually done via some kind of thresholding). Examples of sampled data to which measures of spike train synchrony have been applied include EEG data [25, 46, 48] and, outside of neuroscience, stock market velocity [49]—and rainfall events [47].

The algorithm is particularly suited for datasets with a high value of SPIKE-synchronization. According to the coincidence criterion (equation (1)) these are spike trains that include sequences of global events for which the interval between successive events is at least twice as large as the propagation time within an event. For these datasets the synfire indicator evaluates the consistency of the order within these well separated global events. The universality of the phenomenon, repetitive propagation patterns, makes our new algorithm applicable in a wide array of fields such as medical sciences, seismology, oceanography, meteorology or climatology. For example, the duration of an epileptic seizure is typically much shorter than the interval between two successive seizures. Also the time it takes a storm front to cross a specific region is typically much smaller than the time to the next storm. Many other repetitive propagation phenomena exhibit similar ratios of characteristic time scales.

In order to understand the scope of our proposed algorithm it is important to understand what it is not designed to achieve. The method deals purely with relative order, it does not consider the length of absolute delays. Moreover, while the instantaneous coincidence criterion makes the method time-scale independent, parameter-free and easy to use, it also renders it insensitive to patterns involving spikes that are not immediately adjacent. Many other, typically more complicated, methods have been designed to characterize the detailed spatio-temporal structure in large neuronal networks [50, 51] or to detect hierarchically structured spike-train communities [52, 53]. The method is also not designed to detect neuronal synfire chains (in the strict sense of the word) in massively parallel data. For this task other statistical methods based on some forms of pattern detection have been developed [54, 55].

Another caveat concerns causality. While a significant value of the synfire indicator F_s in our algorithm clearly shows the presence of a preferred temporal order of some signals with respect to others, it does not necessarily prove a driver-responder relationship. There are other methods that have been developed for this kind of system dynamics analysis (e.g., [56]). But even for such methods causality is always a strong claim. In fact, the two signals might be driven by a common hidden source and a consistent leader (follower) could just indicate a drive with a smaller (larger) delay. Similarly, internal delay loops in one of the two systems can also fool the interpretation.

There are a number of possible directions for future research, both from a methods and from a data point of view. Regarding the algorithm, for the coincidence detection it would be straightforward to limit the range of allowed time lags by incorporating information about the expected speed of propagation [47]. One could introduce a minimum time lag in order to ensure causality and/or limit the maximally allowed time lag in order

to focus on meaningful propagation of activity. In principle the range of allowed time lags could even be selected individually for each pair of spike trains depending on the known properties of the connectivity between the respective two neurons. Importantly, even with such type of time lag restrictions in place, it has still to be guaranteed that each spike can be part of at most one coincidence.

A follow-up task for our neurophysiological data would be to investigate to what extent the neurons that are identified as leading by our analysis are identical to the so-called hub neurons [23], i.e. neurons with a much higher than average degree of connectivity within the network. Regarding the El Niño analysis, the difference in propagation directions observed for the strong and weak El Niño events remains an open question. Furthermore, expanding the analysis to wider regions, i.e. 5°S–5°N and verifying the consistency of the observed propagation patterns is an interesting topic for future research. Finally, the relation to results based on average warming/cooling rates [24] requires further investigations.

The algorithm will be readily applicable for everyone since, together with the existing symmetric measures ISI-distance, SPIKE-distance, and SPIKE-synchronization, SPIKE-order is implemented in three publicly available software packages, the Matlab-based graphical user interface SPIKY⁵ [27], cSPIKE⁶ (Matlab command line with MEX-files), and the Python library PySpike⁷ [57].

Acknowledgments

We are indebted to Heinz Beck from the Department of Epileptology, University of Bonn, Germany, for providing the neurophysiological data. We thank Markus Abel, Roman Bauer, Heinz Beck, Nebojsa Bozanic, Christoph Gebhardt, Marcus Kaiser, Stefano Luccioli, Florian Mormann, Giorgia Paratore, and Alessandro Torcini for useful discussions. Finally, we are grateful to Irene Malvestio and Bernd Mulansky for carefully reading the manuscript. TK thanks Marcus Kaiser and his group for hosting him at the University of Newcastle, UK.

We acknowledge funding support from the European Commission through the Marie Curie Initial Training Network ‘Neural Engineering Transformative Technologies (NETT)’, project 289146 (TK and MM), and through the European Joint Doctorate ‘Complex oscillatory systems: Modeling and Analysis (COSMOS)’, project 642563 (TK and ES). TK also acknowledges the Italian Ministry of Foreign Affairs regarding the activity of the Joint Italian-Israeli Laboratory on Neuroscience.

Appendix A. Neurophysiological dataset

The neurophysiological data analyzed in section 3.2 were recorded via fast multicellular calcium imaging in acute CA3 hippocampal slices from juvenile mice. The CA3 region has a strong recurrent excitatory connectivity [58]. This distinct feature is suggested to be crucial for memory encoding and pattern completion and thus memory retrieval [59]. During memory retrieval in rodents, population bursts of the CA3 lead to high frequency stimulation of the efferent regions, so called sharp wave ripples [60]. In the juvenile hippocampus, due to a higher chloride reversal potential in the CA3 pyramidal cells, the GABA-ergic system is excitatory [34]. GABA-ergic interneurons have been shown to serve as so called hub neurons that trigger the GDPs [23].

The recordings were performed by the group of Heinz Beck at the Department of Epileptology, University of Bonn, Germany, prior to and independently from the design of this study. Transversal acute brain slices (300 μm thick) were prepared from 5 to 10 day old (P5-P10) C57BL/6 mice (Charles River, $n = 19$ slices). Slice preparation, calcium imaging and data analysis were performed as previously described in [61]. For AM-loading of brain slices with OGB1-AM we used a protocol modified from [62]. Multicellular calcium imaging was done using a homemade single planar illumination microscope modified from [63]. Movies were recorded at a frame rate of 200 Hz over a minimal length of 5 min up to 30 min to record a sufficient amount of spontaneous activity. Time points of cell activity from the imaging data were defined as the onsets of Ca^2 events in fluorescence traces of all individual cells using the maximum of the second derivative of each event [64].

In order to test the spike train order algorithm, datasets were chosen that exhibited at least three global GDPs during the recording ($n = 5$). For one dataset the surrogate analysis described in section 2.5 proved to be unfeasible due to its excessive density of spontaneous activity. Therefore this dataset was discarded from further analysis, so the final number of datasets analyzed was $n = 4$.

⁵ <http://fi.isc.cnr.it/users/thomas.kreuz/Source-Code/SPIKY.html>

⁶ <http://fi.isc.cnr.it/users/thomas.kreuz/Source-Code/cSPIKE.html>

⁷ <http://pyspike.de>

Appendix B. El Niño dataset

The data analyzed in section 3.3 describe one of the most well-known global climate effects, the so called El Niño, the warm phase of the ENSO. It is usually characterized by an increased SST in the central and eastern tropical Pacific Ocean typically during the months September–February [65]. According to the US NOAA, an El Niño event is identified by an increased three-month moving average SST by $0.5\text{ }^{\circ}\text{C}$ over at least six months. El Niños can last from nine months up to two years and typically occur in irregular intervals of two to seven years.

The analysis presented in section 3.3 uses the OISST data provided by NOAA. These data result from a high-resolution blended analysis (spatial resolution of 0.25°) of daily SST and ice constructed by combining observations from different platforms (satellites, ships, buoys) on a regular global grid with a time range from 1981 to 2016 [66]. The daily SST data is centralized by the long-term daily mean resulting in daily SST anomaly data (deviations from long-term mean) that form the basis of this analysis.

The area used to define El Niño events, the El Niño 3.4 region, stretches from 5°S to 5°N in latitude and 190°E to 240°E in longitude, as shown in figure 15. However, for the latitudinal direction we here focused only on a small central strip around the equator, i.e. from 0.5°S to 0.5°N over a longitudinal stretch consistent with the El Niño region, i.e. from 180°E to 250°E (dashed line in figure 15). Note, that this focus on the small region around the equator is the main difference to most previous works that studied the propagation of SST anomalies averaged over a much larger latitudinal extent, e.g. 5°S – 5°N in [24, 42]. In latitudinal direction, we average over the whole strip (four grid points), while in longitudinal direction an averaging over two grid points is performed resulting in 0.5° of spatial resolution and a total of $N = 140$ time series of daily SST anomalies. We disregard short-term fluctuation by applying a Gaussian smoothing with width $\sigma = 14$ days. In figure 16 we show the time series resulting from this procedure exemplarily for the center of the observed region, 215°E .

For the SST data we use a threshold criterion to identify the moving high temperature fronts of the El Niño events. The threshold determines the signal-to-noise ratio of the data: for small values even weaker events have an effect on the result, whereas for higher values the analysis is focused on the strongest El Niño events only. Due to the variability of the propagation patterns it might happen that in successive years the threshold is surpassed in different regions. However, since the aim of the analysis is to look at the propagation of individual (seasonal) fronts we suppress coincidences between threshold crossings from different years. We still use the adaptive coincidence detection from equation (1) but define a maximum coincidence window τ_{\max} which in this case is set to 9 months.

References

- [1] Spira M E and Hai A 2013 *Nat. Nanotechnol.* **8** 83
- [2] Lewis C, Bosman C and Fries P 2015 *Curr. Opin. Neurobiol.* **32** 68
- [3] Marañón S, Fäh D and Lu Y M 2014 *Geophys. J. Int.* **197** 1566–81
- [4] Heidemann J, Stojanovic M and Zorzi M 2012 *Phil. Trans. R. Soc. A* **370** 158
- [5] Muller C L, Chman L, Grimmond C, Young D T and Cai X 2013 *Int. J. Climatol.* **33** 1585
- [6] Menne M J, Durre I, Vose R S, Gleason B E and Houston T G 2012 *J. Atmos. Ocean. Technol.* **29** 897
- [7] Lazer D et al 2009 *Science* **323** 721
- [8] Gallos L K, Rybski D, Liljeros F, Havlin S and Makse H A 2012 *Phys. Rev. X* **3** 031014
- [9] Boccaletti S, Bianconi G, Criado R, Del Genio C I, Gómez-Gardeñes J, Romance M, Sendiña-Nadal I, Wang Z and Zanin M 2014 *Phys. Rep.* **544** 1
- [10] Lacroix P, Grasso J-R, Roulle J, Giraud G, Goetz D, Morin S and Helmstetter A 2012 *J. Geophys. Res.: Earth Surf.* **117** F1
- [11] Pelinovsky E 2006 Hydrodynamics of Tsunami waves *Waves in Geophysical Fluids: Tsunamis, Rogue Waves, Internal Waves and Internal Tides* (Vienna: Springer) pp 1–48
- [12] Kuramoto Y 2012 *Chemical Oscillations, Waves, and Turbulence* vol 19 (Berlin: Springer)
- [13] Baier G, Goodfellow M, Taylor P N, Wang Y and Garry D J 2012 *Front. Physiol.* **3** 281
- [14] Belik V, Geisel T and Brockmann D 2011 *Phys. Rev. X* **1** 011001
- [15] Wei X, Valler N C, Prakash B A, Neamtiiu I, Faloutsos M and Faloutsos C 2013 *IEEE J. Sel. Areas Commun.* **31** 1049
- [16] Kuhn T, Perc M and Helbing D 2014 *Phys. Rev. X* **4** 041036
- [17] Rieke F, Warland D, de Ruyter van Steveninck R and Bialek W 1996 *Spikes: Exploring the Neural Code* (Cambridge, Massachusetts: Institute of Technology)
- [18] Abeles M 2009 *Scholarpedia* **4** 1441
- [19] Pereda E, Quiñero R and Bhattacharya J 2005 *Prog. Neurobiol.* **77** 1
- [20] Boers N, Bookhagen B, Barbosa H, Marwan N, Kurths J and Marengo J 2014 *Nat. Commun.* **5** 5199
- [21] Varni G, Volpe G and Camurri A 2010 *IEEE Trans. Multimedia* **12** 576
- [22] Rahbar F, Anzalone S, Varni G, Zibetti E, Ivaldi S and Chetouani M 2015 *Int. Conf. on Social Robotics* 543–553
- [23] Bonifazi P, Goldin M, Picardo M A, Jorquera I, Cattani A, Bianconi G, Represa A, Ben-Ari Y and Cossart R 2009 *Science* **326** 1419
- [24] Santoso A, McGregor S, Jin F-F, Cai W, England M H, An S-I, McPhaden M J and Guilyardi E 2013 *Nature* **504** 126
- [25] Quiñero R, Kreuz T and Grassberger P 2002 *Phys. Rev. E* **66** 041904
- [26] Kreuz T, Mormann F, Andrzejak R G, Kraskov A, Lehnertz K and Grassberger P 2007 *Physica D* **225** 29
- [27] Kreuz T, Mulansky M and Bozanic N 2015 *J. Neurophysiol.* **113** 3432
- [28] Kistler W M, Gerstner W and Leo van Hemmen J 1997 *Neural Comput.* **9** 1015
- [29] Naud R, Gerhard F, Mensi S and Gerstner W 2011 *Neural Comput.* **23** 3016

- [30] Mulansky M, Bozanic N, Sburlea A and Kreuz T 2015 *Int. Conf. on Event-based Control, Communication, and Signal Processing (EBCCSP)* 1–8
- [31] Applegate D L, Bixby R E, Chvatal V and Cook W J 2011 *The Traveling Salesman Problem: a Computational Study* (Princeton, NJ: Princeton University Press)
- [32] Earl D J and Deem M W 2005 *Phys. Chem. Chem. Phys.* **7** 3910
- [33] Dowland K A and Thompson J M 2012 *Handbook of Natural Computing* (Berlin: Springer) pp 1623–55
- [34] Ben-Ari Y, Cherubini E, Corradetti R and Gaiarsa J L 1989 *J. Physiol.* **416** 303
- [35] Takano H, McCartney M, Ortinski P I, Yue C, Putt M E and Coulter D A 2012 *J. Neurosci.* **32** 4743
- [36] Malik N, Bookhagen B, Marwan N and Kurths J 2012 *Clim. Dyn.* **39** 971
- [37] McPhaden M J, Zebiak S E and Glantz M H 2006 *Science* **314** 1740
- [38] Rasmusson E M and Carpenter T H 1982 *Mon. Weather Rev.* **110** 354
- [39] Yeh S-W, Kug J-S, Dewitte B, Kwon M-H, Kirtman B P and Jin F-F 2009 *Nature* **461** 511
- [40] Qian C, Wu Z, Fu C and Wang D 2011 *J. Clim.* **24** 6486
- [41] An S-I and Wang B 2000 *J. Clim.* **13** 2044
- [42] Antico P L and Barros V R 2016 *Theor. Appl. Climatol.* **1** 1–6
- [43] Kreuz T, Haas J S, Morelli A, Abarbanel H D I and Politi A 2007 *J. Neurosci. Methods* **165** 151
- [44] Kreuz T, Chicharro D, Andrzejak R G, Haas J S and Abarbanel H D I 2009 *J. Neurosci. Methods* **183** 287
- [45] Kreuz T, Chicharro D, Greschner M and Andrzejak R G 2011 *J. Neurosci. Methods* **195** 92
- [46] Kreuz T, Chicharro D, Houghton C, Andrzejak R G and Mormann F 2013 *J. Neurophysiol.* **109** 1457
- [47] Malik N, Marwan N and Kurths J 2010 *Nonlinear Process Geophys.* **17** 371
- [48] Rosário R S, Cardoso P T, Muñoz M A, Montoya P and Miranda J G V 2015 *Physica A* **439** 7
- [49] Zhou W X and Sornette D 2003 *Physica A* **330** 543
- [50] Schneider G, Havenith M N and Nikolić D 2006 *Neural Comput.* **18** 2387
- [51] Gansel K S and Singer W 2012 *Front. Neuroinform.* **6** 18
- [52] Billeh Y N, Schaub M T, Anastassiou C A, Barahona M and Koch C 2014 *J. Neurosci. Methods* **236** 92
- [53] Russo E and Durstewitz D 2017 *eLife* **6** e19428
- [54] Schrader S, Grün S, Diesmann M and Gerstein G L 2008 *J. Neurophysiol.* **100** 2165
- [55] Gerstein G L, Williams E R, Diesmann M, Grün S and Trengove C 2012 *J. Neurosci. Methods* **206** 54
- [56] Andrzejak R G and Kreuz T 2011 *Europhys. Lett.* **96** 50012
- [57] Mulansky M and Kreuz T 2016 *Software X* **5** 183–9
- [58] Amaral D G and Witter M P 1995 Hippocampal formation *The Rat Nervous System* ed G Paxinos (San Diego: Academic Press) pp 443–93
- [59] Leutgeb S and Leutgeb J K 2007 *Learn. Memory* **14** 745
- [60] Chrobak J J and Buzsáki G 1996 *J. Neurosci.* **9** 3056–66
- [61] Hedrich U B et al 2014 *J. Neurosci.* **34** 14874
- [62] Crépel V, Aronov D, Jorquera I, Represa A, Ben-Ari Y and Cossart E 2007 *Neuron* **54** 105
- [63] Holekamp T F, Turaga D and Holy T E 2008 *Neuron* **57** 661
- [64] Henze D A, González-Burgos G R, Urban N N, Lewis D A and Barrionuevo G 2000 *J. Neurophysiol.* **6** 2799–809
- [65] Trenberth K E 1997 *Bull. Am. Meteorol. Soc.* **78** 2771
- [66] Reynolds R W, Smith T M, Liu C, Chelton D B, Casey K S and Schlax M G 2007 *J. Clim.* **20** 5473