Extraction and Characterization of Essential Discharge Patterns from Multisite Recordings of Spiking Ongoing Activity

Riccardo Storchi1,2, Gabriele E. M. Biella3, Diego Liberati4, Giuseppe Baselli5

1 Department of Biomedical Sciences, University of Modena, Modena, Italy, 2 Institute of Molecular Bioimaging and Physiology, Milan, Italy, 3 National Research Council, Institute of Molecular Bioimaging and Physiology, Milan, Italy, 4 Department of Electronic and Information, National Research Council, Politechnic School of Milan, Milan, Italy, 5 Department of Biomedical Engineering, Politechnic School of Milan, Milan, Italy

Abstract

**Background:** Neural activation patterns proceed often by schemes or motifs distributed across the involved cortical networks. As neurons are correlated, the estimate of all possible dependencies quickly goes out of control. The complex nesting of different oscillation frequencies and their high non-stationarity further hamper any quantitative evaluation of spiking network activities. The problem is exacerbated by the intrinsic variability of neural patterns.

**Methodology/Principal Findings:** Our technique introduces two important novelties and enables to insulate essential patterns on larger sets of spiking neurons and brain activity regimes. First, the sampling procedure over N units is based on a fixed spike number k in order to detect N-dimensional arrays (k-sequences), whose sum over all dimension is k. Then k-sequences variability is greatly reduced by a hierarchical separative clustering, that assigns large amounts of distinct k-sequences to few classes. Iterative separations are stopped when the dimension of each cluster comes to be smaller than a certain threshold. As threshold tuning critically impacts on the number of classes extracted, we developed an effective cost criterion to select the shortest possible description of our dataset. Finally we described three indexes (C,S,R) to evaluate the average pattern complexity, the structure of essential classes and their stability in time.

**Conclusions/Significance:** We validated this algorithm with four kinds of surrogated activity, ranging from random to very regular patterned. Then we characterized a selection of ongoing activity recordings. By the S index we identified unstable, moderately and strongly stable patterns while by the C and the R indices we evidenced their non-random structure. Our algorithm seems able to extract interesting and non-trivial spatial dynamics from multisource neuronal recordings of ongoing and potentially stimulated activity. Combined with time-frequency analysis of LFPs could provide a powerful multiscale approach linking population oscillations with multisite discharge patterns.

Introduction

In the last twenty years, studies on information encoding in the nervous system have provided fundamental insights into the nature of neural inner dynamics and of sensorimotor representation and coding of the external world. Powerful and flexible statistical techniques have grown in time [1–5], improving the analyses of stimulus-response experimental paradigms along with their corollary complexities.

Scant attention has been paid to other dynamic features such as spontaneous or ongoing activity. Nevertheless, this feature represents 90% percent of the whole metabolic exertion of the brain [6]. An exhaustive description of ongoing activity as a kind of substrate intermingling with signals generated by external sources, could provide fundamental insights into nervous system dynamics.

Spontaneous neuronal population activities from many sites in the central nervous system present complex combinations of different oscillation frequencies [7–9]. Regular and repetitive motifs nest within the frames of these global rhythms. Furthermore, stereotyped patterns of specific neuron subsets have been evidenced in many different experimental conditions such as at the onset of UP states [10] or in *in vitro* recordings with calcium imaging [11,12].

Because of the surging number of items, the statistical evaluation of all possible discharge patterns in multisite recordings of even a few channels quickly goes out of control. Nevertheless, pioneering studies and further advancements have developed algorithms capable of recognizing small, repetitive and well-timed patterns in tonic activity regimes [13–17]. Still, timing is likely not the only encoding means valid in the CNS.

Repetitive spontaneous and evoked neuronal activation patterns evidenced time warping features, namely duration independent activity episodes, in most diverse behavioral conditions [18–20]. These features may represent an additional piece of the
multifarious presence of noise, a pervasive trait throughout the bottom-up (or top-down) scaling of brain analyses [21], that affects the immediate detection of accessible regular and repetitive patterns. In any case the crucial role of noise in brain operations, as it has become evident in recent years, has gone largely unknown and is a current subject of wide debate.

In the context of a lack of comfortable theoretical background for the present scenario, we tried to design an efficient algorithm to detect the presence of regular motifs in all ongoing activity regimes and for most variable numbers of spikes or recording traces. We then developed and evaluated a set of measures to characterize such motifs in terms of spatial structure, complexity and temporal evolution.

Our approach is conceptually different from previous works as we introduced several important novel features. First, in addition to constant time bin sampling, we introduced a conditional sampling based on keeping the number of spikes constant in each sample. This procedure allows the rejection of common mode frequency modulations. Second, we applied a clustering procedure in order to extract a reasonable number of classes from the wide variety of distinct patterns and to reduce the effect of noise. Then, following a cost criterion, we extracted a selection of the most frequently occurring classes that we called essential classes (EC).

Finally, we developed three indexes, labeled \(C, S, \) and \(R\), evaluating respectively the average pattern complexity \((C)\), the structure of the ECs \((S)\) and their stability in time \((R)\). We applied such indexes both to simulated and to real data. We mainly focused on the results obtained by using conditional sampling. However the results obtained with constant time bin sampling will be reported as well and commented for comparison.

In the Methods we describe the whole algorithm flow obtained by cascading constant time bin sampling or conditional sampling, the clustering procedure and the extraction of ECs (respectively subsections 1, 2 and 3). We then describe the indexes \(C, S, \) and \(R\) (subsection 4) and provide a detailed explanatory example of the whole procedure (subsection 5), in order to clarify the most critical steps of the algorithm. In the Results we evaluate the cost criterion performances (subsection 1) and apply the indexes \(C, S, \) and \(R\) to several kinds of simulated activities (subsection 2) and to real data (subsection 3).

In subsections 2 and 3 we also compare the results obtained on simulations and on real data and the information provided by constant time bin and conditional sampling. In the Discussion we briefly discuss advantages and drawbacks of the techniques developed.

Materials and Methods

The sampling procedure

Given \(k\) spikes and \(N\) sources, we define a \(k\)-sequence as an \(N\)-dimensional vector whose values, on each dimension, are the number of spikes counted on each source, the sum over all dimensions being \(k\). By fixing \(k\) and \(N\) it is possible to extract a finite set of \(k\)-sequences from each multisite recording (Fig. 1A).

Consider a set of \(N\) spiking sources and let \(x_1, \ldots, x_N\) be the number of spikes emitted by them in a variable time period. Then setting a constant sum \(\sum_{i=1}^{N} x_i = k\) and collecting sets of \(k\)-sequences, we sample from the conditional probability distribution

\[
P(x_1, \ldots, x_N | \sum_{i=1}^{N} x_i = k)
\]

\(k\)-sequences are sampled along the multichannel recording raster-plot on adjacent non-overlapping windows of variable duration.
interpreted as the number of all its possible distinct sequences and \( D_{\text{tot}} \) as the maximum number for any cluster.

At the first step the whole dataset is bisected. Then, before any successive bisection, clusters are ranked by their dimension \( D \) and bisection is performed on the largest cluster. This criterion allows the obtaining of clusters which are roughly comparable in their cluster dimension.

Once we have chosen the cluster to bisect, we apply PDDP with the following steps. First, we subtract to each \( p \) dimension (of the chosen cluster) its mean value. Then, we compute the covariance matrix of the \( k \)-sequences belonging to the chosen cluster and we extract the eigenvector \( \mathbf{v} \) associated with the largest eigenvalue. Finally, given two subclusters \( G_1 \) and \( G_2 \) we assign each sequence \( x \) to one of them by evaluating the score: \( A = x^T \mathbf{v} \). When \( A \) is larger than 0, we assign \( x \) to \( G_1 \). Otherwise, we assign \( x \) to \( G_2 \). After preliminary bisection by PDDP we compute the centroids \( w_1 \) and \( w_2 \) on respectively, \( G_1 \) and \( G_2 \), and use them to initialize the BK algorithm.

The BK algorithm is iterative and is composed of two steps. First, each item is assigned to the nearest centroid. Then, the centroids are recalculated on the base of the last assignment stage. The distinct clusters obtained by the successive bisection, clusters are ranked by their dimension -means always converge so that after a number of iterations the cost criterion can be used with any kind of sampling procedure (like constant time bin sampling). This approach is equivalent to the use of a common encoder without any a priori knowledge on the sampling procedure used.

To select the best model we start to test the first class by computing

\[ A \epsilon_{1.0}(k,D_\text{tot}) = \epsilon_0(k) - \epsilon_1(k,D_\text{tot}) \]

and, iteratively, test an increasing class number. At each iteration \( i \) we evaluate the differential cost \( A \epsilon_{i,i-1}(k,D_\text{tot}) \) of adding a new \( i \)\(^{th} \) class. Such cost is divided into three components

\[ A \epsilon_{i,i-1}(k,D_\text{tot}) = 2N \log(k+1) \]

\[ A \epsilon_{i,i-1}(k,D_\text{tot}) = (\log(D_\text{max}) + \log(i) - \log(k+1))n(i) \]

\[ A \epsilon_{i,i-1}(k,D_\text{tot}) = (\log((i+1) + 1) - \log((i+1))n_{\text{tot}} \]

and \( n(i) \) indicates the number of the \( i \)\(^{th} \) class occurrences in the dataset.

\( A \epsilon_{i,i-1}(k,D_\text{tot}) \) represents the differential database cost, that is the number of bits we need to store a class in the database. \( A \epsilon_{i,i-1}(k,D_\text{tot}) \) represents the gain, achievable by encoding \( n(i) \) \( k \)-sequences once stored the \( i \)\(^{th} \) class in the database. \( A \epsilon_{i,i-1}(k,D_\text{tot}) \) represents the differential flag cost (see Document S1). The database cost constitutes a regret term for too short datasets and fades away for long sequences, while the flag cost increases linearly with the length of the dataset thus constituting a constant regret term for the number of classes.

For increasing values of \( i \), the \( i \)\(^{th} \) class will be accepted if the differential cost \( A \epsilon_{i-1,i}(k,D_\text{tot}) \) is negative.

Iterations are stopped when this is not verified. The \( i \)\(^{th} \) class is rejected and, setting \( M = i-1 \), \( \epsilon_{\text{max}}(k,D_\text{tot}) = \epsilon_{M+1}(k,D_\text{tot}) \).

The asymptotic upper bound for the average cost of a \( k \)-sequence is (for details see Document S1)

\[ \lim_{n_{\text{tot}}(i) \to \infty} \frac{\epsilon_{\text{max}}(k,D_\text{tot})}{n_{\text{tot}}(k)} = H(p) + \log(D_\text{max}) + \log \left( (k+1)^N + 1 \right) \]

where \( p \) is the probability distribution of the class occurrence (all classes, not only the essentials), \( H(p) \) its entropy and \( \log((k+1)^N + 1) \) an upper limit for the flag cost \( \log((M+1)+1) \).

**Characterization of essential patterns**

The application of the cost criterion described in the previous paragraph enables us to extract, for any \( k \), distinct classes of essential patterns. Such patterns are associated with the \( D_{\text{max}} \) value for which \( \epsilon_{\text{max}}(k,D_\text{tot}) = \epsilon_{\text{max}}(k) \).
To characterize the ECs we use three measures: \( C(k) \), \( S(k) \) and \( R(k) \).

\( C(k) \) stands for complexity, is a general descriptor of a \( k \)-sequence dataset and is given by the ratio

\[
C(k) = \frac{c_{\text{est}}(k)}{c_0(k)}
\]

\( S(k) \) stands for segregation and is computed as follows

\[
S(k) = \sum_{s \in E_k} \sqrt{\sum_{i=1}^{N_k} \left( s_i(z) - \frac{1}{N} \right)^2}
\]

being \( s_i(z) \) the \( i \)th dimension of the \( z \)th \( k \)-sequence in the dataset and, respectively, \( E_k \) and \( n_{\text{est}}(k) \) the set of indexes and the number of \( k \)-sequences belonging to ECs. \( S(k) \) enables us to discriminate \( k \)-sequences where spikes are significantly segregated in some subset of sources, from \( k \)-sequences where spikes are homogeneously distributed among them.

Finally, \( R(k) \) (reiteracy) concerns the way successive \( k \)-sequences shift among different classes. By the evaluation of \( R(k) \) we can detect the presence of stable patterns. Namely, given a symbol string associated with a multichannel recording, we estimate the probability \( P_b(k) \) that successive \( k \)-sequences belong to the same EC

\[
P_b(k) = \sum_{s \in A_{ei}} P(s(z) = y, s(z-1) = y)
\]

where \( z \) is the \( z \)th symbol of a string associated with the \( z \)th \( k \)-sequence of the dataset and \( A_i \) is the set of symbols representing all the distinct ECs (only the essential, not all classes). If \( P_b(k) \) is significantly higher than expected by chance, reiteracy is detected and we set \( R(k) = 1 \), otherwise \( R(k) = 0 \).

An approximated \( P_{\text{est}}(k) \) distribution under the null hypothesis is obtained by \( N_s \) shuffles on the original symbol sequence. We call \( N_s \) the number of estimates for which \( P_{\text{est}}(k) < P_b(k) \). The null hypothesis is rejected at level \( \alpha \) if \( x > (1-N_s)/N_s \).

The indexes \( C, R \) and \( S \) may be meaningfully applied to both \( k \)-sequences and \( t \)-sequences. In case of \( k \)-sequences, \( C(k) \) represents the spatial complexity connected to the number and the diversity of all the possible configurations collected from a multichannel recording regardless of their duration. \( R(k) \), the shift indicator between two different \( k \)-sequence classes, determines the switching properties of the system. Positive reiteracy on \( k \)-sequences may be due to small reverberant circuits, to intrinsic single neuron spiking dynamics or, more generally, to any stable differential activity that emerges over common mode modulations of the considered channels. Finally, \( S(k) \) represents the average euclidean distance from the homogeneous sequence, i.e. that sequence where all spikes are equally distributed among the sources.

The significance of \( C, R \) and \( S \), when dealing with \( t \)-sequences (respectively \( C(t), R(t) \) and \( S(t) \) is less unambiguous. In fact, while \( C(k), R(k) \) and \( S(k) \) values simply evaluate spatial regularities, \( C(t), R(t) \) and \( S(t) \) values mix spatial and temporal information. Thus, \( t \)-sequences may have a low \( C(t) \) because of either stereotyped spatial configurations or regular frequency modulations (or both). The same remark holds true for \( R(t) \) and \( S(t) \). Although positive reiteracy may be due to stable spatial patterns and significant segregation to a sparse source recruitment, both \( R(t) \) and \( S(t) \) may be substantially affected by frequency modulations slower than the sampling frequency.

An Explanatory Example

To gain a better understanding of the whole procedure let’s consider the example data shown in Fig. 1A. After conditional sampling, the associated \( k \)-sequence set, given \( k = 5 \), is displayed in Fig. 1A (bottom). As this dataset is a bit too short for our analysis we build a larger one simply by repeating six times each item.

Then, we add some noise deleting one spike from each \( k \)-sequence and reassigning it randomly to one of the 3 channels of the same \( k \)-sequence. The resulting set is

\[
\begin{align*}
2 & 2 1 1 1 2 4 2 3 3 3 0 0 1 \\
0 & 1 3 4 2 2 1 0 0 0 1 3 3 1 \\
3 & 2 1 0 2 1 0 3 2 2 1 2 2 3 \\
2 & 4 1 3 2 3 0 0 1 3 1 2 2 \\
2 & 1 2 0 0 0 3 4 2 2 3 2 1 0 \\
1 & 0 2 2 3 2 2 1 3 1 0 2 2 3 \\
2 & 1 1 1 2 3 3 3 3 0 0 0 3 \\
1 & 3 3 1 2 2 0 1 1 1 3 4 3 1 \\
2 & 1 1 3 1 0 2 1 3 1 2 1 2 1 \\
3 & 2 2 2 2 0 1 1 2 4 2 3 \\
2 & 2 1 0 1 4 4 2 1 0 2 1 \\
0 & 1 2 3 2 1 0 2 2 1 1 1
\end{align*}
\]

We apply the clustering algorithm with three different values of \( D_{\text{max}} \), respectively 4,12,48,75.

By associating a symbol with each distinct cluster we get the following strings:

\[
\begin{align*}
abedefghillowghjolphjlkcehcnuehbacdfacneiaidfimidflhjdhjhgfj \\
abcddcdnaadvbbacdcaaubbcbdaaanbbbadaddbbddaaahcdadac \\
aabbrbbaaaabahbababahbabahbababahababahbababababa \\
aabbrbbaaaaababababahbababahbabababababababababababa
\end{align*}
\]

Because \( D_{\text{max}} \) value is critical, we select the best description by using the evaluation criterion described in paragraph 3 and in Document S1.

The description associated with \( D_{\text{max}} = 12 \) is selected as the most synthetic, by our criterion, being the complexity \( C(5) = 0.95 \).

The dimension of the clusters \( a,b,c,d \) are respectively 12,12,9,12. The cost \( c_{\text{est}}(5) \) is 418.76 bits, while \( c_{\text{est}}(5) = 398.84 \) bits, being the differential database cost \( \Delta C_{1,2,3,0,(12)} = 2M\log(k+1) = 2*3*3*\log(3) = 46.53 \) bits, the differential gain \( \Delta C_{2,3,0(12)} = \sum(k\log(k+1)+\log(2)) - \log(D_{\text{max}})n(i) = -152.34 \) bits and the differential flag cost \( \Delta C_{2,3,0,1} = t(k\log(\log(M+1)+1)) = 83.59 \) bits.

\( k \)-sequence classes \( a,b,c \) are defined by our criterion as essential, while \( d \) does not occur enough times. Writing \( \ast \ast \) for non-ECs the final string is

\[
abed--aaabbbda--aabb-bbdaadaabbdadaabbbbd-aabab-ad-d
\]

On this string we finally compute, on ECs, reiteracy and segregation (\( S = 2.24 \)). To evaluate \( R(5) \), we calculate \( P_x = 0.31 \) on the original final string, then we shuffle the final string \( N_s = 10000 \) times in order to obtain \( P_{\text{est}} \) (Fig. S1). Having set \( z = 0.05 \), and being \( t(1-N_s)/N_s = 0.046 \) we conclude that reiteracy is significant \( (R = 1) \) at level 0.05 (but not at level 0.01).

We can repeat the same procedure with constant time bin sampling. To obtain a reasonable dataset we replicate 6 times the data in Fig. 1B and add noise as in the previous case. The resulting set is

\[
\begin{align*}
2 & 2 1 0 2 2 0 4 3 4 3 2 3 0 0 1 \\
0 & 2 1 5 2 1 1 1 0 1 2 3 6 1 \\
3 & 2 1 3 1 2 0 2 2 2 1 0 2 3 \\
2 & 0 4 4 1 3 1 0 1 1 0 4 4 2 \\
1 & 0 0 0 11 1 5 3 3 1 0 0 1 \\
2 & 1 3 1 3 2 1 3 1 1 0 3 1 2 \\
2 & 0 0 1 2 0 4 3 2 2 0 1 1 2 \\
2 & 1 5 1 3 1 0 1 1 2 3 6 3 3 \\
2 & 2 3 3 0 0 3 1 2 2 0 1 1 0 \\
1 & 3 4 1 3 0 0 0 2 0 4 2 \\
0 & 2 0 0 0 0 3 5 2 3 1 1 0
\end{align*}
\]
0 2 1 4 3 0 3 0 0 2 3
Then we apply the whole algorithm again. With D_{max} = 4, 12, 48, 75 we obtain the following strings:
ababaabaaa--a---bbaaaaaa----aaaa----aaaaa-----aa
aaaabaaaabaaa---bbaaaaaa----aaaa----aaaaa-----aa
aaaababaabaaa--a---bbaaaaaa----aaaa----aaaaa-----aa
aaaababaabaaa--a---bbaaaaaa----aaaa----aaaaa-----aa

Because of the larger number of possible k-sequences in respect to k-sequences, we expect the optimal D_{max} value to increase. In fact the final string is associated with D_{max} = 75
aaa-aa-aaa--aaaaaaa----aaa-a---aaa----aaaa----aaaa

being the dimension of the class a equal to 75 and the values of C,S equal to respectively 0.9966, 1.6016. To use the cost criterion described, we set k at the largest value collected on a single channel during constant time bin sampling (k = 6).

Being \((I - N_c/N_k) = 0.0406\) we conclude that reiteracy is significant at level \(x = 0.05\).

**Cost criterion performances**

In this paragraph we analyze the relation between Entropy \((H)\) and our cost criterion evaluation. We test the cost criterion on two different sources. We then compute \(c_{max}(k_{max})\) as a function of the message length \(n_{tot}\). A source is composed of a regular and a random part. The regular part is constituted by few sequences that are repeated a number of times. The random part is constituted by random sequences generated by \(N\) independent extractions from a homogeneous distribution: namely, all integers ranging from 0 to \(k_{max}\) have probability \(1/(k_{max}+1)\) of occurrence.

Each message is composed of \(r n_{tot}\) and \((I-r) n_{tot}\) sequences belonging, respectively, to the regular and to the random part.

We define \(n\), the number of distinct sequences of the regular part. Each of them occurs \(n_{occ}/n\) times in a message.

![Figure 2. Cost criterion performances.](image)

**Figure 2. Cost criterion performances.** Upper row (plots A,B,C,D,E,F) related to the first source. \(k_{max} = 10, N = 5\). A) The function \(H/(C(k_{max})N\log(k_{max}+1))\) is estimated, as a function of \(n_{tot}\) at different \(r\) values ranging from 0 to 1. B) The same as in A). Here the cost criterion is modified to start evaluation from \(A_{C1.2}\) instead of \(A_{C1.6}\). C,D)\( M\) values estimated with the original (C) and the modified (D) criterion. E,F) \(D_{max}\) estimated with the original (C) and the modified (D) criterion. Lower row (plots G,H,I,J,K,L) related to the second source \((k_{max} = 10, N = 5)\).

If \(N\) and \(k_{max}\) are not too small

\[
H \approx (1-r)N \log(k_{max}+1) + n \log(n_i)
\]

For an infinitely long message of \(n_{tot}\) sequences, \(H\) represents the smallest possible average description length of a sequence. The asymptotic evaluation of our cost criterion, \(c_{max}(k_{max})\), in Document S1, provides a significantly less parsimonious description. This holds true also for finite length messages, as can be observed in Fig. 2A,B,G,H, where the average values of \(H/(N\log(k_{max}+1)/C(k))\) are plotted as a function of \(n_{tot}\). The difference between entropy and \(c_{max}(k_{max})\) decreases as \(n_{tot}\) increases. Conversely, it strongly increases as a function of \(r\).

The more regular the message, the worse \(c_{max}(k_{max})\) approximates \(H\). The differential database cost \(A_{C1,2,M,0} = 2N\log(k_{max}+1)\), accounting for the lowest \(H/(N\log(k_{max}+1)/C(k))\) values at low \(n_{tot}\) fades away for increasing values of this parameter (see Methods section). The differential flag cost \(A_{C2,2,M,0}\) plays a key role in the choice of the EC number, typically leading to a conservative selection that prevents from overfitting. The contribution of \(A_{C2,2,M,0}\) could be reduced by modifying and expanding the flag and the database structures, in order to use a single flag to encode several sequences. This possibility, under current experimental check, can effectively reduce \(A_{C2,2,M,0}\) and the number of flags, and will be presented in a forthcoming paper. The differential flag cost \(A_{C2,2,M,0}\) is sometimes too large and may impair the detection of regular patterns that are significantly present in the message. To avoid the problem, the cost criterion can be evaluated starting from \(M=1\) instead than from \(M=0\). Results obtained with this modification are drawn in Fig. 2B,D,K,H,J,L, while those relative to the original algorithm...
are presented in Fig. 2A,C,E,G,I,K. The values of $M$ and $D_{max}$ are reported, respectively, in Fig. 2C,D,I,J and in Fig. 2E,F,K,L.

The regular part of the first source (Fig. 2A–F) is composed of the following sequences

- 1 10 5 1
- 3 5 5 2
- 10 0 5 3
- 10 2 5 10
- 2 9 6 10

having set $k_{max} = 10$. For short $r$ or $n_{tot}$ values no ECs are detected from our algorithm (white region in Fig. 2E,F). Instead, increasing these parameters, the algorithm always extracts four distinct ECs with $D_{max} = 1$ (blue region in Fig. 2E,F). The sharp transition between 0 and 4 ECs can be expected because the regular sequences we introduced are really different from each other.

The following sequences

- 1 0 1 0 2 0 0 2
- 3 4 4 4 3 3 3 4
- 10 10 9 10 10 10 10 10
- 10 9 10 10 10 10 10 9
- 2 3 2 2 1 3 2 1

are used to generate the regular part of the second source (Fig. 2G–L). No ECs are detected for low $r$ or $n_{tot}$ values (white region in Fig. 2K,L). With slight increments of these parameters, a single EC is detected containing 7 or 8 distinct sequences of the regular part (violet region in Fig. 2J and red region in Fig. 2K,L). When the EC contains all 8 sequences $D_{max} = 72$. In comparison with those of the first source, these sequences are very similar to each other. Accordingly, when the number of their occurrence is not too high, the algorithm detects them as different noisy versions of a unique essential pattern. For further increments of $n_{tot}$ or $r$, our algorithm returns, as ECs, 7 of the 8 sequences composing the regular part of the message (red region in Fig. 2J and blue region in Fig. 2K,L).

As expected, when the repetition of single sequences becomes significant, the regular sequences cannot be seen as a single noisy pattern. In fact, if it is true, the occurrence of the distinct 72 sequences represented in the EC should roughly follow a homogeneous distribution. This is not the case because only 8 out of 72 sequences occur with high probability ($r$/$8$) while the remaining ones have probability $(1−r)/11^7$. A number of other sources have been tested, by varying $k_{max}$ and $N$, with matching results. In general, the non optimality of the cost criterion leads to conservative choices about the number of the ECs. Moreover, by slightly modifying the cost criterion, the selection of ECs becomes more inclusive. The algorithm typically achieves good performances in the separation of regular components from noise. These include well-tuned generalization capabilities to avoid noise-induced multiplication of the ECs for messages that are too short.

**Results**

**Simulations**

We used four different simulation groups to validate the algorithm.

1. We generated independent geometric processes with the parameter $p_a = \{p_{al} \cdots p_{aN}\}$ for a wide range of $N$ sources and spiking frequencies $f_s = p_{a}/Mt$.

2. We generated $M$ independent geometric processes with the parameter $p_a$ given $M\leq N$. We called these events *activation processes* (Fig. 2A). At each event of activation, processes were assigned to a random subset of $NM$ sources (Fig. 2B). Then, for a period of $T$ time-bins (the length of the activation period), the activated sources were driven by independent geometric processes with parameter $p_{b} = \{f_{b}\}/Mt$.

3. We used the same strategy as simulation group 2, but we assigned the $M$ activation processes to $M$ predefined subsets of $N\backslash M$ sources.

4. All predefined $M$ subsets were driven by a common trigger following a geometric process with parameter $p_{b}$. When an event took place all subsets were activated at successive lags of $T$ time-bins.

Activation processes embody the strong non-stationarity of neuronal activity typical both of conscious and unconscious states. In particular, they closely mimic the transition between ‘up’ and ‘down’ states recorded both intracellularly and extracellularly during recordings of ongoing activity in sleep or anesthesia [26]. The term ‘up’ states commonly indicates a depolarized state of intracellular potential typically constituting the substrate for high frequency discharges. Conversely, the term ‘down’ states indicates a hyperpolarized state of intracellular potential leading the cell firing activity to very low frequency regimes.

Homogeneous poison-like activity is simulated by group 1 (Fig. 3A). Multisite activation processes can randomly distribute over the spiking sources (like in group 2, Fig. 3B), repeat on stereotyped subsets of sources (group 3, Fig. 3C) and even exhibit precise serial patterns among the different subsets (group 4, Fig. 3D). Such serial patterns were effectively observed in hippocampal place cells of rats [19] in recordings performed both during spatially constrained tasks and, soon after that, during sleep. Task-related serial patterns expressed during asleep could be time-compressed up to a factor of 20.

Constant time bin sampling was obtained by dividing each simulation into constant periods containing, on average, $k$ spikes. The x axes in plots H–J represent these progressively augmenting periods called, for brevity, $<k>$.

When applied to $k$-sequences, the evaluation criterion, described in subsection 2 of Methods, works quite well for large $k$ values, typically for $k>5$, while for $k<5$, the flag cost $\alpha_{1,0}$ is often much higher than the gain $\alpha_{1,M}$ and the overall minimum cost is $c_{0}(k)$. In these cases, as suggested in Methods, $R(k)$ and $S(k)$ have been computed by skipping $\alpha_{1,0}(k)$.

A number of simulations were performed for a wide range of $f_a$, $f_b$, $M$, $N$ values. On each simulated dataset we applied the whole algorithm flow described in Methods. Some typical outcomes are displayed for $k$-sequences in Fig. 4E–G,5E–G and for $k$-sequences in Fig. 4H–J,5H–J. The results obtained with conditional and constant time bin sampling can be similar (Fig. 4E,J) or very different (Fig. 5E–J).

These two conditions were obtained by keeping fixed $f_a$, $f_b$, $M$, $N$ values and setting, respectively, $M=3$ (Fig. 4) and $M=1$ (Fig. 5).

In Fig. 4 it is easy to see how, by using $C(k),R(k)$ and $S(k)$ and given the same values of $f_a$, $f_b$, $M$ and $N$, the different groups 2, 3 and 4 are effectively discriminated and suitably different from the simulations belonging to group 1. The distinction between groups 2 and 3–4 and between groups 4 and 2–3 is evident respectively for mean $C(k)$ and $R(k)$ or $S(k)$ values, while group 1 is clearly separated from the others in all the plots. In a way, also $C(0), R(0)$ and $S(0)$ allow for a net discrimination among the different kinds of simulations.

Strong discharge pattern segregation is present in small channel subsets in simulation groups 2,3 and 4 (Fig 4B–D). Accordingly, segregation values $S(k)$ and $S(0)$ (Fig. 4E and 4H) in these simulations are much larger than in group 1 (Fig 4A), where spiking discharges contained in the sequences are, on average, more homogeneously distributed over all the spiking sources.
Figure 3. Structure of simulations 2,3,4. A) Structure of a single simulated source. Spikes are displayed as blue points. (*) Represent an interspike interval during an active state, having set at $f_a$ the mean frequency during the active periods. (**) Represent a period between two successive activation given $f_a$ the mean frequency of activation occurrence. B) Given N = 4 channels, we can observe three successive activations of overlapping subsets of N/M = 2 channels.

doi:10.1371/journal.pone.0004299.g003

Figure 4. Some typical results from simulations. Time-bin resolution $\Delta t$ was set at 1 ms while the total number of spikes on each simulation was 5000. A) Sample plot of a group 1 simulation given $N = 6$, $f_a = 2$ Hz. B,C,D) Sample plot of simulation belonging to group 2,3,4. We set $N = 6$, $f_a = 5$ Hz, $f_b = 50$ Hz, $M = 3$, $T = 50$ ms. E,F,G) Estimation of $S(k)$ (E), $R(k)$ (F) and $C(k)$ (G). H,J,K) Estimation of $S(t)$ (H), $R(t)$ (J) and $C(t)$ (K). Blue, red green and black lines respectively represent group 1,2,3,4.

doi:10.1371/journal.pone.0004299.g004
In group 2, the subsets of $N/M = 2$ sources are randomly assigned across the $N = 6$ channels, so that one among the $N!/(N^2N/M)!/(N/M)! = 15$ different subsets may be selected at each activation occurrence. Instead, in simulation groups 3 and 4, the subsets are predefined and only 3 non-overlapping subsets may be activated. The different spatial complexity between simulation group 2 and simulation groups 3 and 4 is reliably detected by $C(k)$ and $C(t)$ values (Fig. 4F,I).

Reiteracy $R(k)$, like $S(k)$, enables us to distinguish between simulation group 4 and simulation groups 2 and 3. Simulation group 1 displays negligible reiteracy (Fig. 4G,J). The non-null values are due to the decay of the significance level $z$, caused by multiple comparisons. Simulation groups 2 and 3 exhibit substantive reiteracy along the considered $k$ interval. The sharp $R(k)$ decay in simulation groups 4 reflects the regular transition between non-overlapping source subsets. The $k$ value coupled with the decay is proportional to the product $NT/f_a/M$, as shown by additional simulations in Fig. S2. Given that $NT/f_a/M$ is the average number of spikes occurring during a subset activation, this last result is not surprising.

Unlike $R(k)$, $R(t)$ decays in simulation groups 4 is faster than in simulation groups 2 and 3. The higher reiteracy $R(t)$ in simulation group 4 is due to the longer silent periods between activation occurrences.

In simulation group 1, independently from $f_a$, we did not detect significant stable patterns.

In Fig. 5, having only one subset including all channels ($M = 1$), simulation groups 2–4 are equivalent. For this reason in Fig. 5 we reported only the results obtained with simulation groups 1 and 2 (Fig. 5A–C,E,F).

No reiteracy was detected with $k$-sequences in simulation group 2 (Fig. 3J). This is expected because setting $M = 1$ means that channels are all equally modulated. Sample plots from this group are reported in Fig. 5B,C. We also reported, for the sake of comparison, a sample plot from a simulation belonging to the group 4 with $M = 3$ (Fig. 5D). Note that the activation of the different channel subsets follows stereotyped serial patterns. This is not the case for simulation group 2 (Fig. 5B,C). All the channels are simply turned on and off at the same time so that no spatial reiteracy should be detected with conditional sampling. This kind of sampling, as explained in Materials and Methods and Document S1, is insensitive to most of the common mode frequency modulations.

Moreover, no significant difference between simulation groups 1 and 2 can be detected either in terms of segregation $S(k)$ (Fig. 5E) or complexity $C(k)$ (Fig. 5F).

Conversely, constant time bin sampling provides remarkable differences in all the measures (Fig. 5H,I,J). In particular, strong reiteracy is due to the slow frequency modulations represented by the alternation of active and silent periods across all the simulated channels.

Constant time bin sampling mixes temporal and spatial information. Constant time bin and conditional sampling provide matching results when spatial information is salient and represents a major determinant in the clustering procedure (Fig. 4). Otherwise, when the clustering is dominated by frequency modulations, these two sampling procedures can provide incoherent results (Fig. 5).

Real data

Several recordings of ongoing activity were analyzed both in normal and neuropathic isofluorane-anesthetized rats. Experimental methods and general results are reported in a dedicated
paper (Storchi et al. Submitted) while here we just focus on three recording groups to show and discuss some typical outcomes.

The recording groups 1 and 2 were performed in two normal rats while recording group 3 was obtained from a rat neuropathic model (Seltzer model, [27]), known to exhibit neuropathic-deafferentative phenomenologies. The indexes were estimated in each recording group on 25 successive recording epochs, 5000 being the total number of spikes in each epoch (see Fig. 6,7,8).

Each recording epoch was compared to a simulation belonging to the group 1 for which the parameter vector \( p_a \) was set equal to the values estimated in that epoch on real data (see Simulation section).

From a rough observation, the surrogated data do not show evident differences in comparison with the real ones in recording groups 2 and 3 (Fig. 6A,B, 7A,B). This could lead us to conclude that such recordings just reflect unstructured noisy discharge patterns. However, by applying our algorithm it is easy to see how the first impression could be somewhat misleading. In fact the recorded discharge patterns exhibit significant segregation (Fig. 6C,F, 7C,F) and larger complexity (Fig. 6D,G, 7D,G) in comparison with surrogated data. Moreover, reiteracy is also significant, mostly for low \( k \) values in recording group 2 and over all considered \( k \) intervals in recording group 3 (Fig. 6E,H, 7E,H).

The \( R(k) \) decay for increasing values of \( k \), as can be observed in recording group 2, may provide a measure of the mean latency of repetitive configurations. For experimental applications, the detection of the \( k \) value associated with the \( R(k) \) decay \( (k_{\text{latency}}) \) can constitute an interesting tool. The mean \( k_{\text{latency}} \) duration represent the average switching period among different discharge pattern configurations. An altered switching period could represent a marker for altered ongoing activity such as the one we can observe in animal models of neuropathic pain (unpublished data). More generally, spontaneously switching configurations may represent a suitable neural substrate to integrate and contextualize incoming sensory information, amplifying relevant inputs and skipping irrelevant ones. Gain modulations of neural responses driven by an internal scheduling is a well-known general computational principle that enables the performance of a variety of tasks such as attention selection or coordinate transformation [28]. The capacity to switch between different active spatial configurations is embedded in recurrent neural networks [29]. EC and reiteracy can be used in this context of recurrent networks, like small cortical networks, to investigate those switching properties in vivo.

While \( k \)-sequence and \( t \)-sequence processing provide comparable results in recording groups 2 and 3, remarkable differences were obtained in recording group 1. Only indexes \( C(t), S(t), R(t) \) detect significant departures from the surrogated data (Fig. 8F,G,H). In fact, while \( C(k), S(k) \) and \( R(k) \) indexes significantly discriminated between recording groups 2 and 3 and the relative surrogated data, this was not the case for recording group 1 (Fig. 8C,D,E). In this recording group, the presence of generalized activations synchronized across the recording sites is clearly observable (Fig. 8A,B). Fast ‘up’ states occur in the spindle oscillation regime (conventionally in the 7–14 Hz frequency interval). The presence of well-defined spindle-like oscillations could also suggest the presence of structured spatial configurations that match frequency modulations. Again a rough observation in time domain (Fig. 8A,B) is inappropriate. Both \( C(k) \) and \( S(k) \) fail to detect non-random spatial organizations in the \( k \)-sequence dataset (Fig. 8C,E). Moreover, negligible reiteration is detected (Fig. 8D).

The whole set of measures depicts a situation very similar to the

![Figure 6. Recording from normal rat 1 (N=5). A) A sample from the analyzed activity. B) Simulated group 1 activity. The mean frequency on each channel was set at the same value of the associated channel from the recorded activity. C,D,E Estimation of \( S(k) \), \( C(k) \), \( R(k) \) and \( C(t) \) from the recorded (red lines) and the simulated (blue lines) activities. Note that, although generalized oscillations are strong in all channels, the nested organization detected with \( S(k) \), \( C(k) \) and \( R(k) \) is not different from the random simulated activity. F,G,H Estimation of \( S(k) \), \( R(k) \) and \( C(k) \) from the recorded (red lines) and the simulated (blue lines) activities. Note the strong reiteracy \( R(t) \) due to successive silent periods and compare with the negligible reiteracy \( R(k) \).](https://doi.org/10.1371/journal.pone.0004299.g006)
Figure 7. Recording from neuropathic rat (N=5). A) A sample from the analyzed activity. B) Simulated group 1 activity. The mean frequency on each channel was set at the same value of the associated channel from the recorded activity. C,D,E) Estimation of $S(k)$ (C), $R(k)$ (D) and $C(k)$ (E) from the recorded (red lines) and the simulated (blue lines) activities. F,G,H) Estimation of $S(k)$ (F), $R(k)$ (G) and $C(k)$ (H) from the recorded (red lines) and the simulated (blue lines) activities. Discharge patterns are remarkably more stable (see $R(k)$ and, to a less extent, $R(t)$) than in the normal rats 1 and 2.
doi:10.1371/journal.pone.0004299.g007

Figure 8. Recording from normal rat 2 (N=5). A) A sample from the analyzed activity. B) Simulated group 1 activity. The mean frequency on each channel was set at the same value of the associated channel from the recorded activity. C,D,E) Estimation of $S(k)$ (C), $R(k)$ (D) and $C(k)$ (E) from the recorded (red lines) and the simulated (blue lines) activities. F,G,H) Estimation of $S(k)$ (F), $R(k)$ (G) and $C(k)$ (H) from the recorded (red lines) and the simulated (blue lines) activities. Recorded activity strongly differs from the random-like condition.
doi:10.1371/journal.pone.0004299.g008
one described by simulation group 2 with $M=1$ (see Simulations section and compare Fig. 8E–G with Fig. 5E–G). Apparently all the sources are driven by the same process and the time-variance of the driving process is the only difference with the surrogated activity. The last result suggests that within and among single fast ‘up’ states, occurring in the spindle oscillations regime, whose onset and decay are well synchronized across the recording sites, spatial organization of discharge patterns can be negligible or absent. This is consistent with the results of Kurths and coworkers [30], who simulated a biologically plausible network whose nodes were connected by subnetworks of interacting excitable neurons. They found that weak couplings among and within nodes reflected a complex hierarchical structure, well matched with the underlying architectural connectivity, while too strong couplings resulted in an undifferentiated generalized oscillatory activity. In this regard it is important to note how conditional sampling provides results that are independent from generalized frequency modulations. The presence of reiteracy $R(k)$ provides us with additional and complementary information in respect to spiking frequency oscillations. Namely stable active subsets may be present independently of common mode firing frequency modulations.

On the whole, the properties extracted by the application of index $C(k)$, $N(k)$ and $R(k)$ and briefly highlighted here could lead to significant advances in the analysis of ongoing brain activity.

**Discussion**

In this paper we present an original method to extract and characterize sequences with fixed numbers of spikes and distributed on multisite sources in multiple electrode recordings. Being the samples collected on the base of fixed spike counts, their occurrence is detected in a time-independent fashion $k$ across the spiking sources. Time independence yields a basic advantage making insensitive to brain oscillations and distinguishes spatial dynamics of multisite discharge patterns. Starting from the basic observation that the assessment of all possible patterns is unachievable because of the diverging increase of possible configurations as a function of $k$ and $N$, we embedded in our algorithm a procedure, based on clustering, to extract the most salient pattern structures that we called essential classes (EC). To characterize the dynamics within and among the detected classes, we introduced three simple measures ($C$; $R$; $S$) evaluating, respectively, the average pattern complexity, the structure of ECs and their stability in time. Several kinds of surrogate activity, ranging from random to strongly structured, were used to validate the algorithm. Its application to real data significantly revealed both random and non-random spatial structures of ongoing discharge patterns and, in several cases, their remarkable stability.

The main drawback of our conditional sampling technique emerges when spiking frequency, synchrony or the number of sources is high. In these cases, when several spiking sources discharge simultaneously next to the completion of a $k$-sequence, it could happen to collect sequences containing more than the predefined $k$ spikes. For example, given $k=5$ and $N=2$, it is possible that, during the sampling procedure, after 4 spikes there is a simultaneous emission from the two sources, resulting in a sequence of 6 spikes. A criterion to decide which of the two spikes will go in the completing sequence and which in the following could solve this problem. However any criterion will be necessarily arbitrary. Because the number of “overloaded” $k$-sequences is negligible in both the simulations and the data analyzed, we reasonably skipped the problem, potentially more demanding with larger source numbers.

As it concerns the clustering procedure, the algorithm we used [22,23] was selected for its wise initialization and computational efficiency. We also used the farthest-first transversal clustering algorithm [31]. This has the important property of keeping the clustering cost, defined as the largest cluster radius, below a twofold value of the optimal clustering (irrespectively to the number of clusters). The results were qualitatively the same as the ones shown above. In principle, any kind of clustering algorithm could be introduced in the algorithmic flow.

The use of an encoding scheme, inspired by the work of Willems [25], proved to be quite conservative in the extraction of the EC (typically 1–20), exhibiting reasonable generalization performance and avoiding overfitting. The ability to separate repetitive and random occurring sequences and the relation between Entropy and the cost criterion length have been analyzed. The scheme we described is readily implementable as a compression algorithm. We think that, in some cases, we could achieve much better compression by modifying the flag structure. In order to reduce the flag number, a single flag could be modified to encode several blocks, increasing the compression rate of those $k$-sequence datasets with regular serial dynamics. Conceptually, our approach is not very different from the first formulation of Minimum Description Length (MDL) [32], known as the two-part MDL. In adjunct to MDL, the scheme we developed allows the separation of a regular component (composed of the ECs and all their $k$-sequences) from a more random one. The final description length is the shortest achievable by the scheme and is asymptotically bounded by the entropy plus a constant factor.

Several algorithms have been developed in the last decades to evaluate spontaneous and evoked activities. The method we described is original and not straightforwardly comparable with other preexisting algorithms in terms of performance. Instead, given its substantial novelties, our algorithm can be used in addition to other algorithms in order to provide complementary information. The work of Abeles and colleagues [13,15] deals with precisely timed repeating sequences (also called “cortical songs”) while, with conditional sampling, we skip the time dimension and focus on the order of occurrence of spike groups. Thus, when observing tonic activity regimes, our pattern characterization seemingly spatially complements the precise firing reverberation sequences of well-timed activations of the synfire chain model [33].

Some authors took into account the use of relative order of spike occurrence by assigning a distinct symbol to each source [19,34]. Those methods rely on template matching and allow for deciding whether arbitrary, user-chosen patterns occur more often than expected by chance. Such valuable techniques are currently contributing to the discovery of unexpected spiking schemes, such as timed-compressed replays of behavior-related spike sequences in cortex and hippocampus [18–20]. Our algorithm equally uses the relative order of spike occurrences, but, with conditional sampling and clustering, it assigns distinct symbols to different multisource configurations. The method is less detailed (we skip the precise firing order of occurrence of single spikes in a sequence) but provides more general information about all possible recurrent patterns. The algorithm automatically selects the ECs with no need for a priori knowledge of the patterns to be tested.

Conditional sampling might grant complementary spatial information independent from the domain of brain oscillations. Oscillations are a widespread and complex intermingled thread, pervading, at different temporal scales, the whole brain dynamics [35,36]. Local Field Potentials (LFP) or Electroencephalogram (EEG) provided extensive knowledge about the oscillation bands
generated, among others, in the thalamo-cortical loop, and about their neural substrate and their functional significance [7–9].

Recent data support an extended view of neural processing from the sole time-domain measures of oscillating brain dynamics to seemingly time-independent complementary measures as introduced by our work. A recent study highlighted, by simulation runs, the presence of the same dominant discharge patterns at significantly different oscillation frequencies staggered both in wide and in sharply peaked bands [37]. Our method, based on single unit recordings, combined with the time-frequency analysis of EEG and LFP, might provide an interesting multiscale approach aimed to join population oscillatory rhythms with multisite discharge patterns.

The characterization of essential patterns in terms of stability has a solid theoretical background in the great body of work developed to investigate the attractor dynamics of neuronal networks [38–41]. Attractors constitute the essential elements for memory storage and retrieval. Several maintenance mechanisms for attractors have been proposed, ranging from recurrent excitation within cell assemblies to synfire-chains and single-cell bistability.

The presence of stereotyped attractor-like configurations or motifs among the whole set of possible neural combinations is a common finding in works dealing with spatiotemporal characterization of ongoing dynamics. Yuste and coworkers [11], showed the presence of precise and repetitive patterns of discharge in somatosensory thalamocortical slices using calcium imaging. Such patterns could either arise spontaneously at the onset of “up” states or could be evoked by thalamic stimulation [12]. Interestingly, spontaneous and evoked patterns were statistically indistinguishable. Accordingly, they hypothesized that spatiotemporal discharge patterns are predefined in the cortex, the thalamus simply providing a trigger signal. Ordered serial activations of specific neural subsets were also observed with single units in vivo in S1 cortex of rats [10]. The result was quite generalizable because it was obtained both in urethane and ketamine-xylazine anaesthetized and in unanaesthetized rats.

When triggered by incoming sensory stimulations or by pending tasks, the spontaneous subset activations could constitute the neural substrate for gain modulations, a widely analyzed and debated general computational principle [27].

In general the algorithm we developed, thanks to conditional sampling, to its robustness in respect to noise and to the novel cost criterion, could lead to the uncovering of unobserved general properties of spiking network dynamics. We propose its application in characterization of ongoing dynamics in diverse physiological and pathological conditions (sleep, chronic pain, visual-attentive processes, memory based rehearsal of past experiences a.s.o.) Specifically, we briefly showed how our indexes $C(k)/S(k)/R(k)$ could be able to characterize altered ongoing activity in an experimental model of chronic pain.

Beyond the issues already discussed, among future applications and developments, the relation between activity and the underlying anatomical substrate represents a promising field of investigation. Several theoretical works have evidenced the fundamental role exerted by structural motifs on the emergence of variable functional motifs, on different time and spatial scales [30,42,43]. More recently, a paper addressing explicitly the mutual relationships between network architecture and dominant patterns of neural activity stressed how synaptic connections determine the repertoire of spatial patterns in spontaneous activity [36]. Spatial configurations identified by our method could represent a further starting point to address the problem of relation between functional and anatomical connectivity.

Supporting Information

Document S1 In the first paragraph some examples of conditional sampling are provided. In the second the encoding scheme used to develop the cost function is described in detail. In the third an upper bound for $\log Z$ in equation (1) is calculated. Found at: doi:10.1371/journal.pone.0004299.s001 (0.12 MB PDF)

Figure S1 Estimated distribution of $P_{io}$ with $N = 10000$ shuffles of the final string. The estimated value of $P_e$ is represented by the blue vertical line. Found at: doi:10.1371/journal.pone.0004299.s002 (1.41 MB TIF)

Figure S2 $R(k)$ values of a group 4 simulation. We set $N = 6$, $M = 2$, $T = 100$ ms, $f_{io} = 0.5$ Hz and $f_i = 40$ (red), $60$ (green), $80$ (black) and $100$ Hz (cyan). The values $T_f f_i = 4, 8, 12, 16, 20$ Hz, as expected, well reflects the position of $R(k)$ sharp decay. Found at: doi:10.1371/journal.pone.0004299.s003 (1.46 MB TIF)

Acknowledgments

We are indebted to Daniel Freedman for useful discussions and detailed manuscript revision. We also wish to thank the unknown reviewers for their helpful suggestions.

Author Contributions

Conceived and designed the experiments: GEMB. Performed the experiments: RS GEMB. Analyzed the data: RS. Contributed reagents/materials/analysis tools: RS GB DL. Wrote the paper: RS GEMB.

References

1. Rieke F, Warland D, de Ruyter van Steveninck RR, Bialek W (1997) Spikes: Exploring the Neural Code. MIT Press, Cambridge MA.
2. Panzeri S, Treves A (1996) Analytical estimates of limited sampling bias in different information measures. Network 7: 87–107.
3. Nemenman I, Bialek W, de Ruyter van Steveninck RR (2004) Entropy and inference in neural spike trains: progress on the sampling problem. Phys Rev E 69: 056111.
4. Panzeri S, Senatore R, Montemurro MA, Petersen RS (2007) Correcting for the sampling bias problem in spike train information measures. J Neurophysiol 98: 1064–1072.
5. Johnson DH, Gruner MG, Ragghy K (2001) Information-Theoretic Analysis of Neural Coding. Journal of Computational Neuroscience 10: 47–69.
6. Vincent JL, Patel GH, Fox MD, Snyder AZ, Baker JT, et al. (2007) Intrinsic functional architecture in the anaesthetized monkey brain. Nature 447: 83–86.
7. Steriade M, Jones E, Llinas R (1990) Thalamic Oscillations and Signaling. New York: John Wiley & Sons Publishers.
8. Conreras D, Destexhe A, Sejnowski T, Steriade M (1996) Control of Spatiotemporal Coherence of a Thalamic Oscillation by Corticothalamic Feedback Science 274(2208): 771–774.
9. Steriade M, Amzica F (1998) Coalescence of sleep rhythms and their chronology in corticothalamic networks. Sleep Res Online 1(1): 1–10.
10. Artur Luczak, Peter Bartho, Stephan I, Marguet, György Buzsáki, Kenneth D Harris (2007) Sequential structure of neocortical spontaneous activity in vivo. Proc Natl Acad Sci USA, January 2 104: 347–352.
11. Cossart R, Aronov D, Yuste R (2003) Attractor dynamics of network UP states in the neocortex. Nature 423: 283–288.
12. MacLean JN, Watson BO, Aaron GB, Yuste R (2005) Internal dynamics determine the cortical response to thalamic stimulation. Neurosci 48(3): 811–823.
13. Abeles M, Gerstein GL (1988) Detecting spatiotemporal firing patterns among simultaneously recorded single neurons. J Neurophysiol 60: 909–924.
14. Aertsen AM, Gerstein GL, Habib MK, Palm G (1989) Dynamics of neural firing correlation: modulation of “effective connectivity.” J Neurophysiol 61: 900–917.
15. Prut Y, Vaadia E, Bergman H, Haalman I, Skovin H, Abeles M (1998) Spatiotemporal Structure of Cortical Activity: Properties and Behavioral Relevance. J Neurophysiol 79: 2657–2674.
16. Tetko IV, Villa AEP (2001) A pattern grouping algorithm for analysis of spatiotemporal patterns in neuronal spike trains. 1. Detection of repeated patterns. J Neurosci Methods 105: 1–14.
17. Tetko IV, Villa AEP (2001) A pattern grouping algorithm for analysis of spatiotemporal patterns in neuronal spike trains. J Neurosci Methods 105: 15–24.
18. Nadasdy Z, Hirase H, Czurko A, Csicsvari J, Buzsáki G (1999) Replay and time compression of recurring spike sequences in the hippocampus. J Neurosci 19: 9497–9507.
19. Lee AK, Wilson MA (2002) Memory of sequential experience in the hippocampus during slow wave sleep. Neuron 36: 1183–1194.
20. Euston DR, Tatuno M, McNaughton BL (2007) Fast-forward playback of recent memory sequences in prefrontal cortex during sleep. Science 316(5853): 1147–50.
21. Faisal AA, Lac PJS, Daniel MW (2008) Noise in the nervous system. Nat Neurosci 9: 292–303.
22. Abeles M (1991) Corticonics: Neural Circuits of the Cerebral Cortex. Cambridge, New York, USA: Cambridge University Press.
23. Smith AG, Smith P (2006) A set probability technique for detecting relative time order across multiple neurons. Neural Comp 18: 1197–1214.
24. Hoffman KL, Battaglini FP, Harris K, MacLean JN, Marshall L, et al. (2007) The upshot of up states in the neocortex: from slow oscillations to memory formation. J Neurosci 27(44): 11838–41.
25. Galán RF (2008) On how network architecture determines the dominant patterns of spontaneous neural activity. PLoS ONE 3(5): e2148.
26. Honey CJ, Kotter R, Breakspear M, Sporns O (2007) Network structure of cerebral cortex shapes functional connectivity on multiple time scales. Proc Natl Acad Sci U S A 104: 10240–10245.
27. Rissanen J (1978) Modeling by the shortest data description. Automatica 14: 465–471.
28. Hochbaum D, Shamir D (1985) A best possible heuristic for the k-center problem. Math Op Res 10: 180–184.
29. Savaresi SM, Boley DL (2004) A Comparative Analysis on the Bisecting k-means and the PDDP Clustering Algorithms. International Journal on Intelligent Data Analysis 8(4): 345–362.
30. Zhou C, Zemanova L, Zamora G, Hilgetag CC, Kurths J (2006) Hierarchical organization unveiled by functional connectivity in complex brain networks. Phys Rev Lett 97: 238103.
31. Abeles M (1991) Corticonics: Neural Circuits of the Cerebral Cortex. Cambridge, New York, USA: Cambridge University Press.
32. Hochbaum D, Shamir D (1985) A best possible heuristic for the k-center problem. Math Op Res 10: 180–184.
33. Savaresi SM, Boley DL (2004) A Comparative Analysis on the Bisecting k-means and the PDDP Clustering Algorithms. International Journal on Intelligent Data Analysis 8(4): 345–362.
34. Savaresi SM, Boley DL (2004) A Comparative Analysis on the Bisecting k-means and the PDDP Clustering Algorithms. International Journal on Intelligent Data Analysis 8(4): 345–362.
35. Boley DL (1998) Principal Direction Divisive Partitioning. Data Mining and Knowledge Discovery 2(4): 325–344.
36. Rissanen J (1978) Modeling by the shortest data description. Automatica 14: 465–471.
37. Savaresi SM, Boley DL (2004) A Comparative Analysis on the Bisecting k-means and the PDDP Clustering Algorithms. International Journal on Intelligent Data Analysis 8(4): 345–362.
38. Meyer-Lindenberg A, Ziemann U, Hajak G, Cohen L, Berman KF (2002) Transitions between dynamical states of differing stability in the human brain. PNAS 99: 10948–10953.
39. Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. PNAS 79: 2554–2558.
40. Amit DJ, Brunel N (1997) Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. Cerebral Cortex 7: 237–252.
41. Marder E, Abbott LF, Turrigiano GG, Liu Z, Golowasch J (1996) Memory from the dynamics of intrinsic membrane currents. PNAS 93: 13481–13486.
42. Honey CJ, Kotter R, Breakspear M, Sporns O (2007) Network structure of cerebral cortex shapes functional connectivity on multiple time scales. Proc Natl Acad Sci U S A 104: 10240–10245.