# Density-based clustering: A 'landscape view' of multi-channel neural data for inference and dynamic complexity analysis

Gabriel Baglietto[1,2]*, Guido Gigante[3,4], Paolo Del Giudice[1,3]

1 INFN-Roma1, Italian National Institute for Nuclear Research (INFN), Rome, Italy, **2** IFLYSIB Instituto de Física de Líquidos y Sistemas Biológicos (UNLP-CONICET), La Plata, Argentina, **3** Italian Institute of Health (ISS), Rome, Italy, **4** Mperience srl, Rome, Italy

* gabriel.baglietto@gmail.com
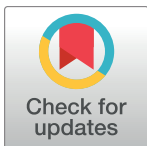
Abstract

Two, partially interwoven, hot topics in the analysis and statistical modeling of neural data, are the development of efficient and informative representations of the time series derived from multiple neural recordings, and the extraction of information about the connectivity structure of the underlying neural network from the recorded neural activities. In the present paper we show that state-space clustering can provide an easy and effective option for reducing the dimensionality of multiple neural time series, that it can improve inference of synaptic couplings from neural activities, and that it can also allow the construction of a compact representation of the multi-dimensional dynamics, that easily lends itself to complexity measures. We apply a variant of the 'mean-shift' algorithm to perform state-space clustering, and validate it on an Hopfield network in the glassy phase, in which metastable states are largely uncorrelated from memories embedded in the synaptic matrix. In this context, we show that the neural states identified as clusters' centroids offer a parsimonious parametrization of the synaptic matrix, which allows a significant improvement in inferring the synaptic couplings from the neural activities. Moving to the more realistic case of a multi-modular spiking network, with spike-frequency adaptation inducing history-dependent effects, we propose a procedure inspired by Boltzmann learning, but extending its domain of application, to learn inter-module synaptic couplings so that the spiking network reproduces a prescribed pattern of spatial correlations; we then illustrate, in the spiking network, how clustering is effective in extracting relevant features of the network's state-space landscape. Finally, we show that the knowledge of the cluster structure allows casting the multi-dimensional neural dynamics in the form of a symbolic dynamics of transitions between clusters; as an illustration of the potential of such reduction, we define and analyze a measure of complexity of the neural time series.

Introduction

Technology nowadays allows neuroscientists to simultaneously record brain activity from increasingly many channels, at multiple scales; indeed, recent years witnessed a kind of 'Moore's Law' for neural recordings [1], and this poses new challenges and opens new opportunities.

One obvious challenge is to devise data representations that easily convey in a compact form the spatio-temporal structure of the recorded data. Various forms of dimensional reduction are now commonly used in the analysis of multiple recordings. In general terms, if one views recorded experimental data as a matrix whose columns are the 'feature vectors' (in the case at hand, the set of recorded activities in a given time bin), and whose rows span the 'sample space' (in the case at hand, the successive time bins), dimensional reduction across the column direction provides a reduced representation in terms of few suitably identified features obtained from the original ones (e.g. principal component analysis); on the other hand, one can view clustering across the rows direction as a way to reduce the dimensionality of the data matrix by lumping together, according to some similarity measure, groups of activity vectors sampled at different times. This latter viewpoint, which we take here, to our knowledge has been much less used in neuroscience (but see [2]).

On the other hand, one recently explored opportunity is to take advantage of multiple recordings to revive old approaches to infer estimates of synaptic couplings from measured correlations between neural activities. Correlations measured from single neuron pairs obviously can only provide ambiguous estimates of the direct synaptic couplings (due to confounding causes like common input to the sampled neurons). However it was noticed in a landmark paper [3] that when many (order 100 for instance) simultaneous recordings are available, even though the underlying biological neural network is still dramatically undersampled, the global pattern of (individually small) pairwise correlations allows to extract meaningful information about the synaptic connectivity. This was achieved by assuming a maximum entropy Ising model, for which an "inverse Ising" problem was solved to infer the parameters (couplings and external input) for the given data. Many efforts were subsequently devoted both to extend the approach to non-equilibrium estimates, and to lighten the computational load of maximum entropy estimates (Boltzmann learning) through various mean-field approximations (see *e.g.* [4] [5] [6]).

In the present work we propose an approach, based on clustering in the multi-dimensional state space of simultaneous recordings, that provides both an advantage for a compact representation of data, also amenable to efficient estimation of the complexity of the system's dynamics, and besides allows to improve inference on the network couplings.

After describing the clustering method (which is a slightly modified version of the 'mean-shift' algorithm [7] [8]), we first illustrate its working on time series generated from the dynamics of a Hopfield network which, since it possesses an energy function, naturally lends itself to density-based clustering in the state space; here we choose a 'hard' regime where the Hopfield network is in a disordered phase and spatio-temporal structures related to the energy landscape are not easily discernible from the time series. At this stage we also formulate a parametrization of the model's synaptic matrix, based on the identified clusters, and show that it allows to obtain an inference of the synaptic couplings which is much more insensitive to noise.

We then move on to the more complex and biologically motivated case of a multi-modular attractor spiking network: its integrate-and-fire neurons are endowed with spike-frequency adaptation (SFA), acting as an activity-dependent self-inhibition that introduces history-dependent effects making the 'landscape' dynamic; modules are approximately bistable

between high and low activity states; within-modules connectivity is stronger than between-modules.

Reasons to choose this particular context include recently published evidence [9] [10] that cortical dynamics can occur in the form of abrupt switches, between an 'Up' and a 'Down' states, of local self-excited modules, such that the whole dynamics appears as the evolution of a 'binary word', each bit being the 'Up' or 'Down' state of each cortical module.

In the chosen setting of weakly coupled bistable modules, the network dynamics is largely determined by inter-modular couplings; in choosing the latter we wanted to preserve some a priori knowledge of key features of the state space, to be later checked against the found cluster structure, and we do this by seeking inter-modular couplings such that the network possesses a prescribed pattern of spatial pairwise correlations, derived in turn from a template Hopfield network.

To achieve this we develop a procedure, inspired by Boltzmann learning but extending its domain of application, which we believe has an interest beyond the contingent purpose.

We check the good performance of this new learning procedure by direct comparison between the prescribed spatial correlations and those generated by the optimal network at the end of learning; by comparing the state-space cluster structure of the optimal network and the template Hopfield network; by inferring effective inter-module synaptic couplings from simulations of the optimal network and comparing them to the synaptic efficacies of the template model.

Knowledge of the cluster structure also allows to cast the multi-dimensional time series generated by the network dynamics in the reduced form of a symbolic dynamics in the clusters' centroid space. We perform such a reduction to show that it allows to expose in a compact way time-dependent features like history-dependent effects; we do this by exploring a large range of characteristic times of SFA, and showing that the Lempel-Ziv complexity measure, applied to the centroid sequence, nicely captures the memory effects induced by SFA.

In summary, what we propose here is a way to use knowledge of the spatial correlations to develop informative and compact representations of multi-dimensional neural data, also allowing for improved inference and useful reduced representation of the multidimensional dynamics, and to develop a strategy for data-driven model building with spiking networks.

Materials and methods

Clustering in the state space: A modified Mean Shift Algorithm

In Fig 1 we outline the logical steps involved in applying Mean Shift Clustering to a multidimensional time series. Although we will be mainly interested in applying the method to time series composed of discrete events (such as spikes, or the result of thresholding population activity), this is not a restriction of the method, which we illustrate here in its generality.

Given a data set $\{x_1, x_2, \ldots x_M\}$ consisting of M (real valued) vectors in a N-dimensional space, the Mean Shift (MS) algorithm [Fukunaga Hostetler] [8] iteratively picks a random i ($1 \le i \le M$) and at step t updates the vector x_i according to the rule:

$$x_i^{t+1} = \frac{\sum_{j \neq i}^{1,M} x_j^t \, w\big(\mathrm{dist}(x_i^t, \, x_j^t)\big)}{\sum_{j \neq i}^{1,M} w\big(\mathrm{dist}(x_i^t, \, x_j^t)\big)}; \tag{1}$$

where $w(\cdot)$ is a non increasing function of its argument, $\mathrm{dist}(x_i^t, \, x_j^t)$ is a distance measure between vectors, and the original vector set is $\{x_i^0\}$. In other words, at each step x_i^t is replaced by a weighted average of the other points x_j^t: points that are closer to x_i^t are given more weight, thus defining a "soft" neighborhood of x_i^t (see Fig 1). Therefore the update rule effectively
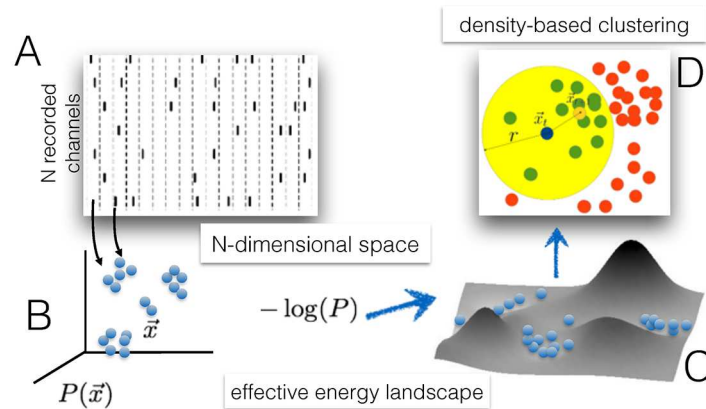
Fig 1. Density-based clustering in the state space. Sketch of the steps involved in the density-based clustering of multi-dimensional time series through the mean shift algorithm. The time-binned *N*-dimensional data ('spikes') in panel A define a density profile in a *N*-dimensional space, interpreted as sampling a stationary probability distribution (panel B), from which an effective energy landscape can be defined (panel C). Panel D illustrates how the mean-shift operates to find the local maxima of the density distribution. Blue dot: current position of a point to be moved by the algorithm. Big yellow area: the chosen neighbourhood of the blue point. Green dots: neighbouring points of the blue point. Red dots: points that are not neighbours of the blue point. Small yellow dot: the new position of the blue point is given by the algorithm as the center of mass of its neighbouring points. See [Materials and methods] for details.

https://doi.org/10.1371/journal.pone.0174918.g001

moves x_i towards regions of the N dimensional space where the (local) density of points is higher at step t. The other $M - 1$ vectors are instead left unchanged ($x_j^{t+1} = x_j^t$ for $j \neq i$); though a parallel, deterministic update rule is in principle possible applying Eq (1) to all the data points at the same time, the random, sequential update just described proved more robust in practice.

The iteration terminates upon a threshold condition: when all $x_i(t)$ belong to a discrete space, the iteration terminates when the fraction of cases (out of the last M) in which $x_i(t + 1) \neq x_i(t)$ is below a fixed threshold; in the continuous case, an additional threshold on a minimal displacement is required. Clusters are finally determined as sets of initial points x_i^0 absorbed by a same final point (their number we call the 'mass' of the centroid); each final point can be taken as representative of a different cluster (its 'centroid'.)

The choice of the weighting function $w(\cdot)$ is central to the MS algorithm. Such function can in general depend both on the step t and the configuration of the neighborhood, thus adapting to different conditions. We used a step weight function:

$$w(x) = \begin{cases} 1 & \text{if } x \leq r \\ 0 & \text{if } x > r \end{cases}.$$ (2)

The choice of the radius r is critical; on the one hand, a small radius can make the algorithm too sensitive to local variation of density (e.g. due to noisy data), leading to many non-representative clusters; on the other, as the radius increases, the reduced sensitivity to local fluctuations can lead to merging different meaningful clusters. We therefore resorted to an adaptive heuristic criterion to select r. At each step t, we compute the standard deviation $\sigma(n)$ of the distances between x_i^t, the point selected for update, and its n nearest neighbors. We then take $n_{\min} = \arg\min_n \sigma(n)|_{n \geq n_0}$, where n_0 is a cutoff number of neighbors (typically set to 10), and set r

as the distance between x_i^t and its n_{min}-th nearest neighbors. We found such heuristics to work well in our case of discrete (hyper cubic) space, for different data sets and for a wide distribution of cluster sizes.

This criterion for determining the radius, together with the additional MS clustering performed on the centroids found in a first run (see [Results]), constitute a modification of the original MS algorithm, which we found advantageous.

In the MS algorithm the initial ordering of data points is irrelevant, as they are treated as samples of a static distribution. For time series this means that the time structure of the data is completely ignored and points that are distant in time can end up being clustered together, *e.g.* as the result of subsequent 'jumps' of the system in a same region of the state-space. Nonetheless, as we will see in [Dynamics in the centroid space], segmenting the original time series in terms of transitions from one cluster to another can highlight relevant dynamical features (non-stationarities, memory effects) in the data.

Inference

Here we describe the method adopted to infer the synaptic couplings from the (binarized) activity of neurons in the network, and how the knowledge of the clusters, extracted by MS from the activity states visited by the network, can be incorporated into a suited parametrization of the synaptic matrix to improve inference.

Specifically, we will assume for the system under consideration an effective Ising-like energy function:

$$H[\boldsymbol{\sigma}] = -\frac{1}{N} \sum_{ij=1}^{N} \sigma_i J_{ij} \sigma_j. \qquad (3)$$

where $\boldsymbol{\sigma} = \{\sigma_1, \sigma_2, \ldots \sigma_N\}$ are binary neural states ($\sigma_i = \pm 1$) for the network comprising N neurons, and J_{ij} is the weight of the synapse connecting neurons j and i. We propose to decompose the coupling matrix J in a set of weighted Hopfield-like terms $c_i^\mu c_j^\mu$ (compare Eq (18)):

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{C} \omega_\mu c_i^\mu c_j^\mu, \qquad (4)$$

where the vectors $\boldsymbol{c}^\mu$ are the C centroids identified by the clustering procedure, and the weights ω_μ are to be found through an optimization procedure.

Notice that, when the ωs are roughly equal, they play a role similar to an inverse temperature; indeed, we found that the value of β used in the Hopfield simulations is very close to the ω values for the centroids corresponding to the patterns.

For the optimization we use a recently introduced technique, Minimum Probability Flow (MPF) [11], with the advantage that it does not rely neither on Monte Carlo (MC) runs (like Boltzmann learning does) nor on mean field approximations (like most Inverse Ising methods do [12]). It is based on the maximization of a function whose maxima are close to the Boltzmann Likelihood ones, but with no needs of performing computationally expensive Gibbs samplings.

Whilst Boltzmann learning aims to minimize the distance (more precisely the Kullback-Liebler divergence) between the distribution of the data and the distribution of the model, MPF aims to minimize the distance between the distribution of the data and the model distribution after an infinitesimal Monte Carlo step, starting from the data. As the name of the algorithm suggests, this amounts to minimizing the outflow, determined by the model, of probability from the data. Since the latter is expressed only through the transition probabilities,

performing the Monte Carlo is not actually needed, and the gradients of the objective function can be explicitly computed.

For completeness, we briefly summarize the main points in the MPF approach. The model probability density **p**, dependent on a set of parameters θ, it is assumed to evolve according to the master equation:

$$\dot{\mathbf{p}} = \mathbf{\Gamma}\,\mathbf{p} \tag{5}$$

The transition probabilities Γ are assumed to satisfy the detailed balance condition:

$$\Gamma_{\alpha \to \beta}\, p_\alpha^{(\infty)}(\theta) = \Gamma_{\beta \to \alpha}\, p_\beta^{(\infty)}(\theta) \tag{6}$$

where $p_\alpha^{(\infty)}$ is the equilibrium probability of state α.

The Maximum Likelihood estimate $\hat{\theta}$ of the parameters, given the data, is:

$$\hat{\theta}_{\mathrm{ML}} = \underset{\theta}{\mathrm{argmin}}\, D_{\mathrm{KL}}(\mathbf{p}^{(0)}||\mathbf{p}^{(\infty)}(\theta)) \tag{7}$$

where $\mathbf{p}^{(0)}$ is the probability distribution on the data.

MPF proposes to consider instead, for infinitesimal ϵ:

$$\hat{\theta}_{\mathrm{MPF}} = \underset{\theta}{\mathrm{argmin}}\, D_{\mathrm{KL}}(\mathbf{p}^{(0)}||\mathbf{p}^{(\epsilon)}(\theta)) \tag{8}$$

It was proven in [11] that:

$$D_{\mathrm{KL}}(\mathbf{p}^{(0)}||\mathbf{p}^{(\epsilon)}(\theta)) \simeq \frac{\epsilon}{N_{\mathrm{Data}}} \sum_{\substack{\alpha\,\in\,\mathrm{Data} \\ \beta\,\notin\,\mathrm{Data}}} \Gamma_{\alpha \to \beta} \tag{9}$$

The MPF strategy is therefore to search for:

$$\min_{\theta} \sum_{\substack{\alpha\,\in\,\mathrm{Data} \\ \beta\,\notin\,\mathrm{Data}}} \Gamma_{\alpha \to \beta} \tag{10}$$

Spiking network and Pseudo-Boltzmann learning

Below we describe the elements and the architecture of the multi-modular spiking neural network whose landscape we explore using MS, and the procedure we devised to assign inter-modular couplings so as to reproduce a given pattern of spatial correlations between the neural activities of pairs of modules (see also Section [Towards a more realistic scenario: a multi-modular network of spiking neurons matching a prescribed spatial correlation structure]).

The single neuron obeys the sub-threshold dynamics:

$$
\begin{aligned}
\dot{V}_i &= \frac{-V_i}{\tau} + \sum_j J_{ij} \sum_{k_j} \delta(t - t_{k_j} - d_{ij}) - g_{SFA} c_i + I_i \\
\dot{c}_i &= \frac{-c_i}{\tau_{SFA}} + \sum_{k_i} \delta(t - t_{k_i}),
\end{aligned} \tag{11}
$$

with the additional condition that if $V_i \geq V^{th}$, the emission of a spike is recorded and V_i is reset to a value H for the duration of a refractory period t^{ref}. In Eq (11), V_i is the membrane potential of neuron i, τ is the membrane time constant, J_{ij} is the synaptic efficacy from neuron j to

neuron i, k_j label spikes emitted by neuron j, d_{ij} is the spike transmission delay from j to i. The neuron model includes spike-frequency adaptation (SFA) in the form of a calcium-dependent inhibitory (potassium) current: g_{SFA} is the strength of SFA, c_i represents the intra-cellular calcium concentration, and τ_{SFA} is a characteristic time of calcium dynamics. The c_i variable acts as an integrator of the spiking activity of the neuron i, such that SFA implements an activity-dependent self-inhibition. I_i is an external current applied to the neuron, implemented as a Poisson process with rate ν_i^{ext} (see e.g. [13–15]).

The network comprises 64 modules. Each module includes $n_E = 32$ excitatory (E) neurons and $n_I = 16$ inhibitory (I) neurons. E/I neurons within a module are connected with probabilities c_{XY}, X, $Y = E/I$; each E/I neuron also receives $C_{E/I}^{ext}$ external spike trains through synapses with mean $J_{E/I}^{ext}$, and with rate ν_{init}^{ext} (*init* refers to the fact that external rates are set to an initial value, then they are optimized during the learning process—see below). Delays d_{ij} are drawn from an exponential distribution between d_{XY}^{min} d_{XY}^{max}, X, $Y = E/I$, where $d_{XY}^{min} = 0.1$ms always. The values of these parameters are given in Table 1. Inter-module synapses are only between excitatory neurons, and can be positive or negative, which can can be viewed as an effective way to also incorporate the excitatory input to inhibitory neurons; their mean efficacies are determined through the learning procedure described below.

All the synapses in the network, given their (learnt or assigned) mean efficacies, are drawn from a Gaussian distribution 25% relative variance.

The probability of connection between two excitatory neurons belonging to different modules is 0.5 while the delays for inter-module communication have $d_{EE}^{max} = 50$ms.

To choose the intra-module synaptic connectivity such that the isolated module is approximately bistable, we used predictions from mean-field theory [9]. The bistable behavior and a narrow range of firing rates for the low and high activity states proved to be robust against quenched noise due to different realizations of the synaptic matrices. This allowed us to define a threshold θ_{bin} on the firing activity to binarize the dynamic states of the modules (the binarization is actually performed on a smoothed version of activity by averaging over a 20 ms window). In this way, the dynamics of the network is represented by a sequence of binary vectors corresponding to the high or low state of all modules. The procedure to construct the inter-module connectivity (see below) preserves to a large extent the modules' bistability for the interconnected network.

Table 1. Parameters of the neural modules.

Parameter	Value
n_E / n_I	32 / 16
τ_E / τ_I	20 / 10 ms
t_E^{ref} / t_I^{ref}	2 / 1 ms
H / V^{th}	15 / 20 mV
c_{EE} / c_{IE} / c_{EI} / c_{II}	0.5 / 0.5 / 1 / 1
J_{EE} / J_{IE} / J_{EI} / J_{II}	0.72 / 1 / −2 / −0.0012 mV
d_{EE}^{max} / d_{IE}^{max} / d_{EI}^{max} / d_{II}^{max}	21 / 21 / 1 / 1 ms
ν_{init}^{ext}	12 kHz
C_E^{ext} / C_I^{ext}	600 / 400
J_E^{ext} / J_I^{ext}	0.320 / 0.111 mV
τ_{SFA}	[62.5 − 4000] ms
g_{SFA}	25/τ_{SFA}
θ_{bin}	40 Hz

Numerical simulations were performed using the event-driven simulator we described in [16].

The procedure to construct inter-module synaptic connectivity aims to "store" prescribed patterns of activities using correlations between the binarized network state vectors to build an effective Hopfield-like synaptic matrix. To store patterns in the modular network we start from a Hopfield model with given P stored patterns, from the simulation of which we estimate the mean single spin magnetization $m_p^{\mathrm{Hopf}} \equiv \langle \sigma_p \rangle_t$, and spatial spin-spin correlation functions $c_{pq}^{\mathrm{Hopf}} \equiv \langle \sigma_p \sigma_q \rangle_t - \langle \sigma_p \rangle_t \langle \sigma_q \rangle_t$, where $\langle \rangle_t$ is the average over the whole time series.

We then iteratively minimize the difference between m_p^{Hopf} and c_{pq}^{Hopf}, and the corresponding quantities m_p and c_{pq} measured from the sequence of binarized states of the multi-modular spiking network (such that p and q are module labels). Such minimization is performed adopting an educated guess inspired by Boltzmann learning that would be appropriate for a binary spin system, and that we verify *a posteriori*. Specifically we implement the following procedure:

- Run a long simulation of the multi-modular system

- Binarize the activity of each module, obtaining a time-series of binary state vectors, from which

- Measure the spatial connected correlations c_{pq} and magnetizations m_p

- Perform a step of the pseudo-Boltzmann iteration over the values of J_{pq} and v_p^{ext}, where J_{pq} is the average synaptic efficacy from neurons belonging to module q to neurons belonging to module p; and v_p^{ext} is the common value of v^{ext} for all neurons in module p

$$
\begin{aligned}
\Delta J_{pq}^{(t)} &= -\mathrm{sign}(c_{pq}^{(t)} - c_{pq}^{\mathrm{Data}}) \, \Delta_{pq}^{(t)} \\
\Delta v_{\mathrm{ext}_p}^{(t)} &= -\mathrm{sign}(m_p^{(t)} - m_p^{\mathrm{Data}}) \, \Delta_p^{(t)}
\end{aligned}
\tag{12}
$$

with the learning rates $\Delta_{pq}^{(t)}$ and $\Delta v^{(t)}$ determined by the Rprop algorithm [17, 18]:

$$
\Delta_{pq}^{(t)} =
\begin{cases}
\min(\eta^+ \Delta_{pq}^{(t-1)}, \Delta_{\max}) & \text{if } (c_{pq}^{(t-1)} - c_{pq}^{\mathrm{Data}})(c_{pq}^{(t)} - c_{pq}^{\mathrm{Data}}) > 0 \\[6pt]
\max(\eta^- \Delta_{pq}^{(t-1)}, \Delta_{\min}) & \text{if } (c_{pq}^{(t-1)} - c_{pq}^{\mathrm{Data}})(c_{pq}^{(t)} - c_{pq}^{\mathrm{Data}}) < 0 \\[6pt]
\Delta_{pq}^{(t-1)} & \text{otherwise} \\[6pt]
& 0 < \eta^- < 1 < \eta^+
\end{cases}
\tag{13}
$$

$$
\Delta_{p}^{(t)} =
\begin{cases}
\min(\eta^+ \Delta_{p}^{(t-1)}, \Delta_{\max}) & \text{if } (m_{p}^{(t-1)} - m_{p}^{\mathrm{Data}})(m_{p}^{(t)} - m_{p}^{\mathrm{Data}}) > 0 \\[6pt]
\max(\eta^- \Delta_{p}^{(t-1)}, \Delta_{\min}) & \text{if } (m_{p}^{(t-1)} - m_{p}^{\mathrm{Data}})(m_{p}^{(t)} - m_{p}^{\mathrm{Data}}) < 0 \\[6pt]
\Delta_{p}^{(t-1)} & \text{otherwise} \\[6pt]
& 0 < \eta^- < 1 < \eta^+
\end{cases}
\tag{14}
$$

In essence, Rprop adapts the learning rate for each parameter only using the sign of the derivatives: when the sign of the derivative does not change between two successive iteration steps, learning should speed up, while when the sign of the derivative changes it should slow down.

We checked that, although every simulation during the training procedure starts from the same exact condition, discarding the first $5 \times \tau_{SFA}$ steps of each run is enough to ensure that no appreciable correlations are detectable between different realizations.

Complexity measure of dynamics in the centroid space

In this section we describe how, building on the knowledge of the state-space cluster structure extracted by MS, the complexity of the states sequence in time can be characterized in a simple form.

After performing MS clustering on the time series generated by a simulation of the multi-modular spiking network with given τ_{SFA} and g_{SFA}, to each of the 64-dimensional vectors containing the modules' firing activities we substitute the label of the centroid that vector belongs to. In this way, the multi-dimensional dynamics is converted into a symbolic sequence of centroid labels.

From such label sequence we measure the matrix of transition probabilities between all pairs of centroids, and generate surrogate label sequences as Markov processes with the same transition probabilities as the actual sequence. While, by construction, the surrogate sequences are memoryless, this may not be the case for the actual one. We focus on such possible memory effects, which are in general interesting to uncover, and for this purpose we adopt a measure typically used to evaluate the complexity of symbolic sequences, the Lempel-Ziv (LZ) complexity, which is a measure of compressibility, suitably normalized in order to eliminate trivial dependence on the length of the sequence.

The LZ complexity is computed as [19]

$$\mathcal{C}_{\mathrm{LZ}} = \frac{S_{\mathrm{LZ}}}{|S| \, \log\left(|A|\right) / \log\left(|S|\right)} \tag{15}$$

where S_{LZ} is the length of the compressed sequence, $|S|$ is the length of the sequence, $|A|$ is the length of the alphabet.

And the relative complexity index is defined as:

$$\mathrm{R} = \frac{\mathcal{C}_{\mathrm{LZ}}^{\mathrm{Markov}} - \mathcal{C}_{\mathrm{LZ}}^{\mathrm{Sample}}}{\mathcal{C}_{\mathrm{LZ}}^{\mathrm{Markov}}} \tag{16}$$

where $\mathcal{C}_{\mathrm{LZ}}^{\mathrm{Sample}}$ is the LZ complexity of the actual centroids sequence, and $\mathcal{C}_{\mathrm{LZ}}^{\mathrm{Markov}}$ is the average LZ complexity of the surrogate Markov centroids sequences.

Results

When the landscape is (partially) known: The Hopfield Model as a first test ground for state-space clustering

We start by providing an example of the information gained from clustering in the state space, *i.e.* finding the local maxima of the density distribution of the configurations generated by a system's dynamics, or—equivalently—the minima of an effective energy landscape (see Fig 1).

Though we want to ultimately apply clustering to data from realistic neural simulations, to illustrate the method we consider here a well known neural network (the Hopfield model) for which an equilibrium distribution of states is defined.

For the Hopfield model [20], the equilibrium probability distribution of the neural states $\boldsymbol{\sigma}$ = $\{\sigma_1, \sigma_2, \ldots \sigma_N\}$ ($\sigma_i = \pm 1$) is given by:

$$p(\boldsymbol{\sigma}) \propto \exp\left(-\beta H[\boldsymbol{\sigma}]\right), \tag{17}$$

where β is interpreted as the inverse of a temperature and H is the energy function:

$$H[\boldsymbol{\sigma}] = -\frac{1}{N} \sum_{i,j=1}^{N} \sum_{\mu=1}^{P} \xi_i^{\mu} \xi_j^{\mu} \sigma_i \sigma_j, \tag{18}$$

where the P discrete vectors $\boldsymbol{\xi}^{\mu}$ are chosen as random uncorrelated configurations of the N spins ($\xi_i^{\mu} = \pm 1$) which, depending on P, N and β can act as 'stored patterns', i.e. the dynamics relaxes to the neighborhood of one of the configurations $\boldsymbol{\xi}^{\mu}$, which are minima of the energy (maxima of the probability distribution) together with their mirror patterns $-\boldsymbol{\xi}^{\mu}$ [21]. Depending on P, N and β, besides the patterns $\boldsymbol{\xi}^{\mu}$ a large number of energy local minima is also present, notably various linear combinations of the stored patterns themselves or states uncorrelated with the patterns in the 'spin glass' phase.

In the following we consider the same Hopfield model (same synaptic couplings, therefore same energy minima) in two operating regimes: low noise (i.e. temperature, where the network is expected to reside most of the time in the proximity of the deepest energy minima), and high noise (where the network explores larger regions of the state space, less constrained by the structure of the underlying energy function). The network comprises $N = 50$ neurons and $P = 4$ stored patterns. Although for such small numbers relying quantitatively on known theoretical estimates of memory capacity is unwarranted, for the chosen values of N and P the model would be predicted to be well above the retrieval phase for both temperatures, and for its non-trivial ('glassy') energy landscape the energy minima are not expected to coincide with the memory patterns.

The neural configuration time series were generated simulating the Hopfield model for 20000 Monte Carlo steps (a suitable number of initial thermalization steps were neglected).

To the N-dimensional resulting time series we applied the (slightly modified) mean-shift (MS) clustering procedure as described in [Materials and methods]: essentially, we iteratively move the points in the N-dimensional space, each of which represents one state generated in the Monte Carlo sequence, towards the center of mass of their neighboring points, thereby identifying at the end the estimated position of the local maxima of the density distribution in the state space.

The states of the Hopfield model are defined over the N-dimensional hypercube, where N is the number of neurons. In order to perform the mean shift displacements we take the sign of the mean value of each coordinate, so that the points remain in the original space. Any time the mean value of a given coordinate is zero, we don't shift that coordinate.

Finally, after reaching a convergence criterion (see [Materials and methods]) of the MS algorithm, we re-run it over the set of found centroids, with a fixed radius given by Hamming distance 2, or overlap 0.92 (the overlap $q^{\alpha\beta}$ between two binary configurations σ^{α} and σ^{β} is defined as $q^{\alpha\beta} = \frac{1}{N} \sum_{i=1}^{N} \sigma_i^{\alpha} \sigma_i^{\beta}$, and it is related with the Hamming distance $h^{\alpha\beta}$ by $q^{\alpha\beta} = 1 - 2\frac{h^{\alpha\beta}}{N}$), and weighting each centroid with its mass in the analogous of Eq (1); we found this further step to make clustering more resistant to noise.

In general, we will consider only clusters containing data points above a minimal fraction ("cutoff") of the whole time series, i.e. minimal *mass* (1% unless specified). In this way we avoid to consider as clusters small bumps of density due to finite sample fluctuations.

Fig 2, left panels, show representative time courses of the neural states for the low-temperature ($\beta = 1.3$, top) and high-temperature ($\beta = 0.83$, bottom) cases. Right panels show, for the two cases, the distribution of the overlaps between all configurations in the time series.

The subsequent Fig 3 illustrates the result of the clustering procedure for the low- and high-temperature cases.
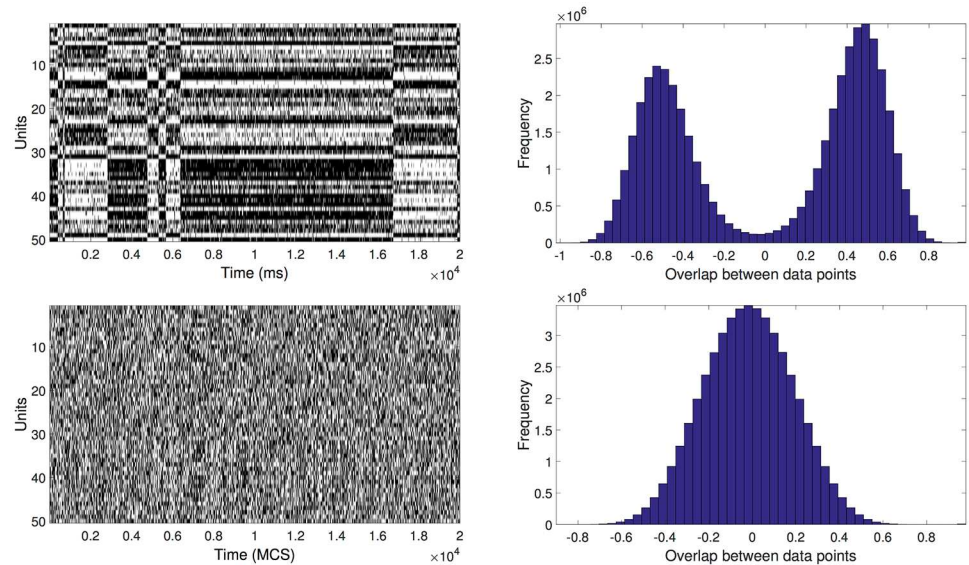
Fig 2. Hopfield network dynamics for two temperatures. Left column: The horizontal axis represents the time in MCS, while the vertical axis identifies the different units. White (black) pixel in the matrix entry (*i, j*) means that the *i* unit had a value -1 (1) at MCS *j*. Right column: Histogram of the overlaps between the MC configurations. Top panels correspond to a case of low temperature ($\beta = 1.3$), and bottom panels to one of high temperature ($\beta = 0.83$).

https://doi.org/10.1371/journal.pone.0174918.g002

In the high-temperature case it is very difficult to discern a structure in the raster plot, and considering also the unimodal overlap distribution, broadly symmetric around zero, it seems a challenging case for a procedure aimed at extracting the energy minima.

This temperature dependence is reflected in the distribution of overlaps between the visited states. We note that, if energy minima would mostly coincide with the stored (uncorrelated) patterns, the overlap distribution would be tri-modal, with a peak in zero and two symmetric peaks at high (positive and negative) overlap, which is not the case even for low temperature (top-right panel in Fig 2), with the distribution of visited states having average overlap about 0.5 in magnitude. For higher temperature (bottom-right panel in Fig 2) the overlap distribution is approximately Gaussian.

From Fig 3 it is seen that for low temperature most centroids are indeed different from the stored patterns; for high temperature more centroids are identified, including the stored patterns.

While we did not perform a thorough analysis of clusters' centroids in terms of energy local minima, we checked that clusters are indeed akin to attractor basins, by measuring the configurations flow at zero temperature: starting the deterministic (zero-temperature Monte Carlo) dynamics from each one of the configurations assigned to a given cluster, the fraction of them for which the final overlap with their assigned centroid is larger than the initial one is 0.90 ± 0.04 (for the clusters found at $\beta = 0.83$), and 0.85 ± 0.05 (for the clusters found at $\beta = 1.3$).

Clustering-aided inference of synaptic connectivity

As mentioned in the [Introduction], several methods have been and are being developed to infer effective synaptic connectivities from the time series of simultaneous neural recordings, some of which fall in the category of so-called "Inverse Ising" problems [3, 5].
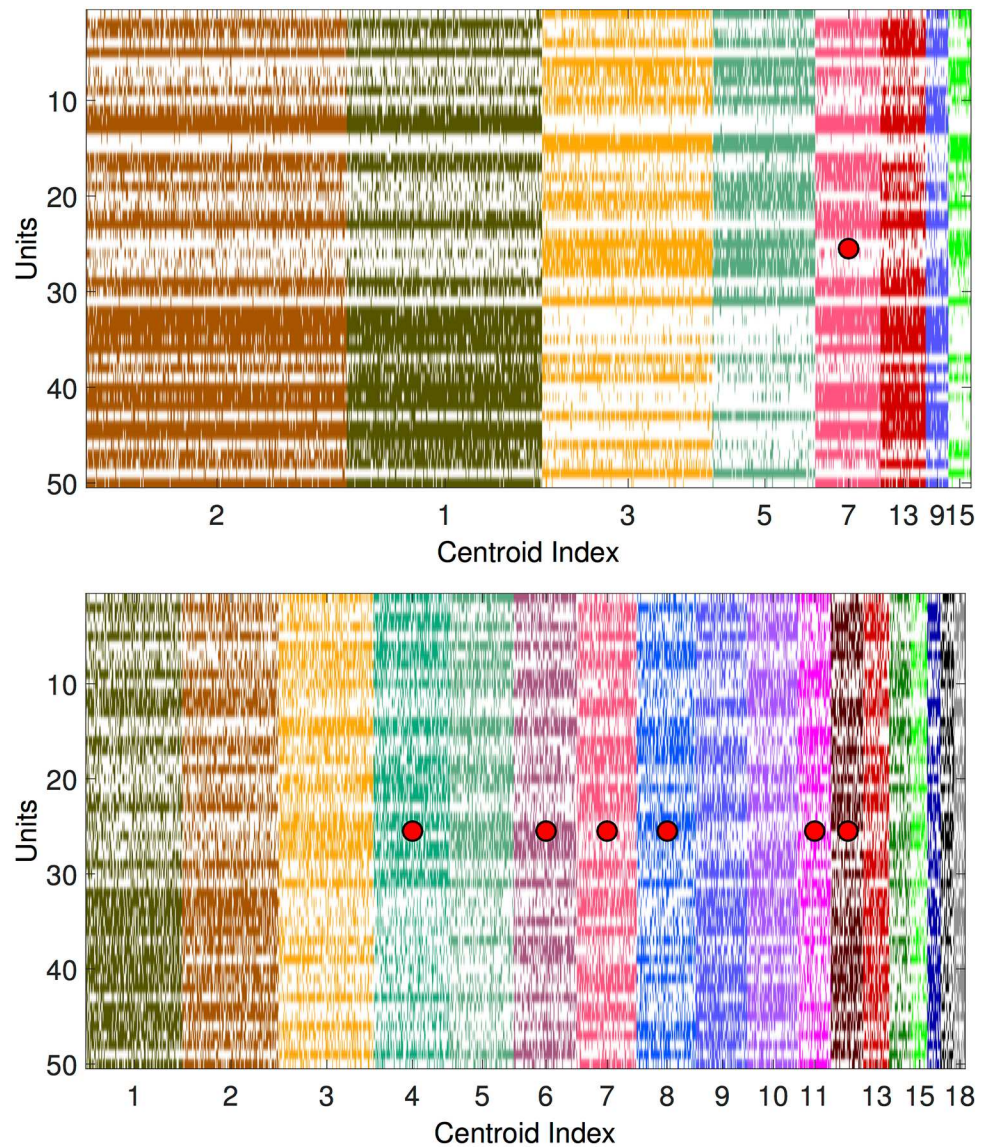
Fig 3. Cluster structure of Hopfield dynamics for two temperatures. Clustered configurations for the raster plots in Fig 2. White is for units with value -1, while in colors (different for different clusters) are indicated units with value 1. The clusters whose centroids correspond to stored patterns or their reflections are marked by red dots. The centroids of the clusters appearing in the upper panel are a subset of those in the lower panel. Equal numbers refer to the same centroid. The clusters' masses can change due to finite size sampling, and clusters in each figure are ordered from biggest to smallest masses.

https://doi.org/10.1371/journal.pone.0174918.g003

Obvious hindrance in the application of such methods is the large number of parameters to be inferred (effective synaptic efficacies, of order N^2 for N neurons), which makes them sensitive to noise in the data.

In the following, we show that the proposed clustering method can be used to formulate the inference problem in a reduced parameter space, using the activity configurations that are the identified centroids to parametrize a coupling matrix (inspired to the construction of the Hopfield connectivity matrix) with a number of parameters equal to the number of centroids.
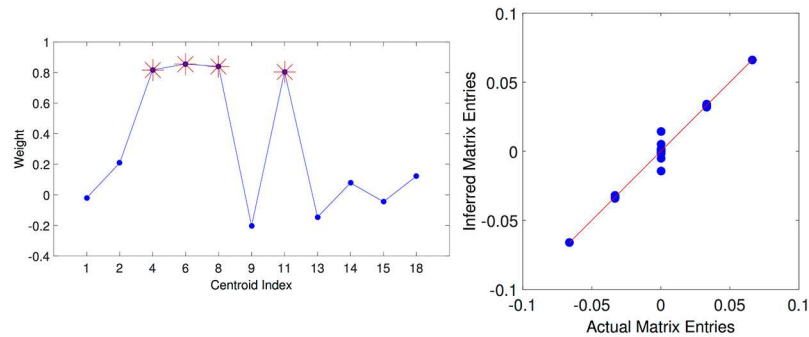
Fig 4. Cluster-aided inference of synaptic couplings for the Hopfield network. Left: Inferred values of the weights ω_α when fitting the reduced model (Eq (20)) to the raster plot shown in the lower column of Fig 2). The red stars mark the weights corresponding to the patterns actually stored in the system. Right: Scatter plot of the inferred (reduced model) synaptic couplings against the actual ones. The red curve shows the identity as a reference.

https://doi.org/10.1371/journal.pone.0174918.g004

Specifically (see also [Materials and methods]), the inference model has the form:

$$H[\boldsymbol{\sigma}] = -\frac{1}{N} \sum_{ij=1}^{N} \sigma_i J_{ij} \sigma_j. \tag{19}$$

where in the following, depending on the context, σ_i will be either the activity of a binary neuron or a suitable binarization of the average activity of a population of spiking neurons. The effective coupling matrix J is the sum of weighted Hopfield-like terms $c_i^\mu c_j^\mu$ (see Eq (4), reproduced here for convenience):

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{C} \omega_\mu c_i^\mu c_j^\mu, \tag{20}$$

c^μ being the C centroids identified by the clustering procedure, and the weights ω_μ are to be inferred.

To test this approach, we first use again the Hopfield model data of Fig 3, bottom panel. Since 14 out of the 18 identified centroids are pairs of reflected patterns and give rise to the same Hopfield-like term in Eq (20), the number of parameters ω to be estimated is reduced to 11.

In Fig 4, left panel, we show the inferred values of ω_μ. Only four of them (marked with an asterisk) are significantly different from zero, and they correspond to the centroids that coincide with the 4 stored patterns.

Notice that, when the ωs are roughly equal, they play a role similar to an inverse temperature; indeed, the value of β used to generate the time series (0.83) is very close to the ω values for the centroids corresponding to the patterns.

We also remark that the obtained values for ω would not be trivially expected from the structure of the state space, in that not only 7 out of 11 centroids (not counting reflections) do not belong to the stored patterns, but the clusters of largest mass are not centered on patterns (remember that the Hopfield system is far from the retrieval phase).

In Fig 4, right panel, we compare the inferred J with the real J ($N \times (N-1) = 2450$ elements, taking only $P + 1 = 5$ values); the inference is actually good (the continuous identity line is drawn to guide the eye), and better than the one obtained inferring directly the full J matrix
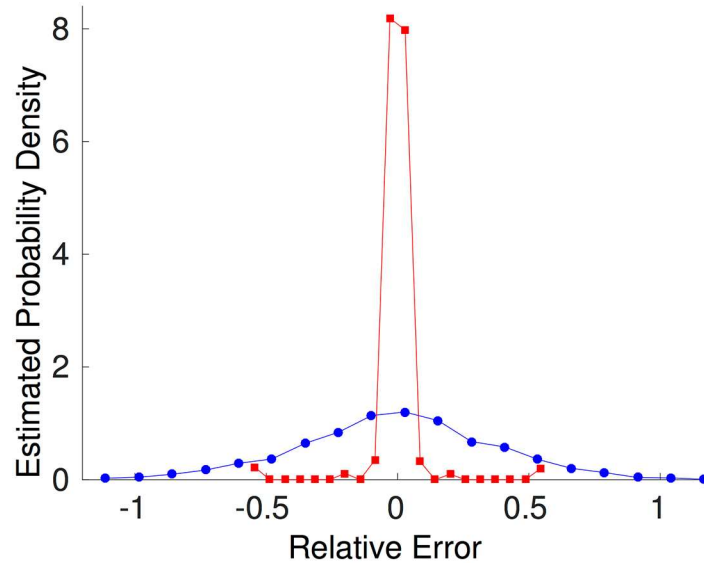
Fig 5. Error comparison for cluster-aided and 'full' inference. A posteriori probabilities of inference errors for the cases of the reduced model (red curve) and the full matrix inference (blue curve). Details in the text.

(i.e. not adopting the parametrization of the synaptic matrix in Eq (20)), as shown in Fig 5, where we compare the relative error $((J_{ij}^{Inferred} - J_{ij})/\langle|J_{ij}|\rangle_{ij})$ for the two cases.

From this simple example, we confirm that the adopted parametrization, besides the obvious greater simplicity and lesser computational load, is effective in reducing the effect of noise in the data and makes the inference less prone to overfitting, while allowing a good match between model and data.

Towards a more realistic scenario: A multi-modular network of spiking neurons matching a prescribed spatial correlation structure

In the previous Section we showed that, from the spatial correlation structure of the configurations sampled by Monte Carlo, MS clustering was effective in reconstructing the main local energy minima, and allowed for a parsimonious parametrization of the synaptic matrix that afforded better inference. In that case, the correspondence between local density maxima in the state space, and local energy minima, was ensured by the existence of a Gibbs equilibrium probability for the Hopfield model.

In the perspective of applicability to real electrophysiology data, in the present Section we extend the approach to networks of spiking neurons, and to a situation where the notion of a static energy landscape is no longer strictly applicable.

To establish a meaningful benchmark, we want to preserve some a priori knowledge of key features of the state space, to be checked against the found cluster structure, and we do this by setting up a spiking network constructed so as to possess a prescribed spatial correlation structure, from which we generate time series to be clustered. To achieve this we develop a method that, we believe, has a wider interest beyond the case at hand.

The chosen network architecture is composed of strongly self-coupled modules of spiking (integrate-and-fire) neurons, with much $(10^{-2} - 10^{-3})$ weaker synaptic couplings between modules. Intra-module synapses are drawn from a Gaussian distribution, with mean and

average chosen (using mean-field predictions) such that each module in isolation is approximately bistable, between states of low (DOWN) and high (UP) firing activity.

The interest in this choice for the architecture of the spiking network stems from accumulating evidence that, not only a modular structure is suggested by the mesoscopic anatomical organization of the cortex, but it also appears to be recognizable in the cortical neural dynamics, which can proceed as the dynamic composition of abrupt jumps between UP and DOWN states ([9, 10]), as mentioned in the Introduction (see also [22] [23] [24]).

We remark that synapses are not be constrained to be symmetric, therefore the clustering procedure will not, strictly speaking, match the minima of an energy function of the system. Besides, the Integrate-and-fire spiking neuron model (see [Materials and methods]) is endowed with spike-frequency adaptation (SFA), a much studied (and pervasively observed) self-inhibitory mechanism depending on the activity history of the neuron; SFA makes the effective landscape, even when it exists, locally dynamic at the point currently corresponding to the network state.

Fig 6 describes the main steps involved in the network construction. The multi-modular network is sketched (panel A) as a collection of 64 neural modules, each composed of 32 excitatory and 16 inhibitory adapting neurons (see [Materials and methods] for details). The approximate bistability of the single modules, which is to a large extent preserved in the interacting network, is illustrated in panel B by the time course of the firing rate of the excitatory neurons from two sample modules; the resulting bimodal distribution of firing rates allows to binarize the modules' activity, as shown in panel C. Once binarized, the time course of the whole network activity can be represented as a sequence of binary vectors (see the 'raster plot' in panel D), to which clustering is applied.

We remark that the binarization step, here instrumental for the construction of the intended benchmark, would not be needed in general, and spate-space clustering can be performed of the raw multi-dimensional time series.

We choose to assign a spatial correlation structure mimicking the one of a Hopfield attractor network, each module corresponding to one binary neuron of the Hopfield network: inter-module synapses, initially drawn from a Gaussian distribution, are subject to an iterative procedure (see below, and [Materials and methods] for details) to match those of the reference Hopfield model.

Were it not for the effects of SFA, the correlation-matched multi-modular network would be expected to behave as an attractor, Hopfield-like system, with the Up and Down states of each module playing the role of the binary values of the Hopfield neurons. However, as the network enters the basin of one attractor, the active modules start lowering their activity because of SFA, thereby destabilizing the state: as noted above, the attractor landscape becomes 'dynamic', in that its depth and curvature around attractor states get lower when the system visits them, promoting transitions to other basins, which are biased by the correlations with other attractor states, due to finite-size effects (the role of SFA in promoting transitions between neural states has been studied in various contexts, see *e.g.* [25–33]).

The new procedure we developed to determine the inter-module synaptic couplings is inspired to the Boltzmann learning strategy where, at each iteration, the change in the couplings is proportional to the difference between the spatial correlations in the model and in the data (and analogously for the external fields). For Boltzmann learning, such difference is the gradient of the Kullback-Leibler distance between the state probabilities in the model and in the data. In our spiking network the explicit form of such a function is lacking; still, intuition suggests, and we assume, that a monotonic relation still holds between the synaptic couplings and the spatial correlations, and this is equivalent to assume that we know the sign of an unknown gradient. Formulated in this way, our 'pseudo-Boltzmann' iterative process to find
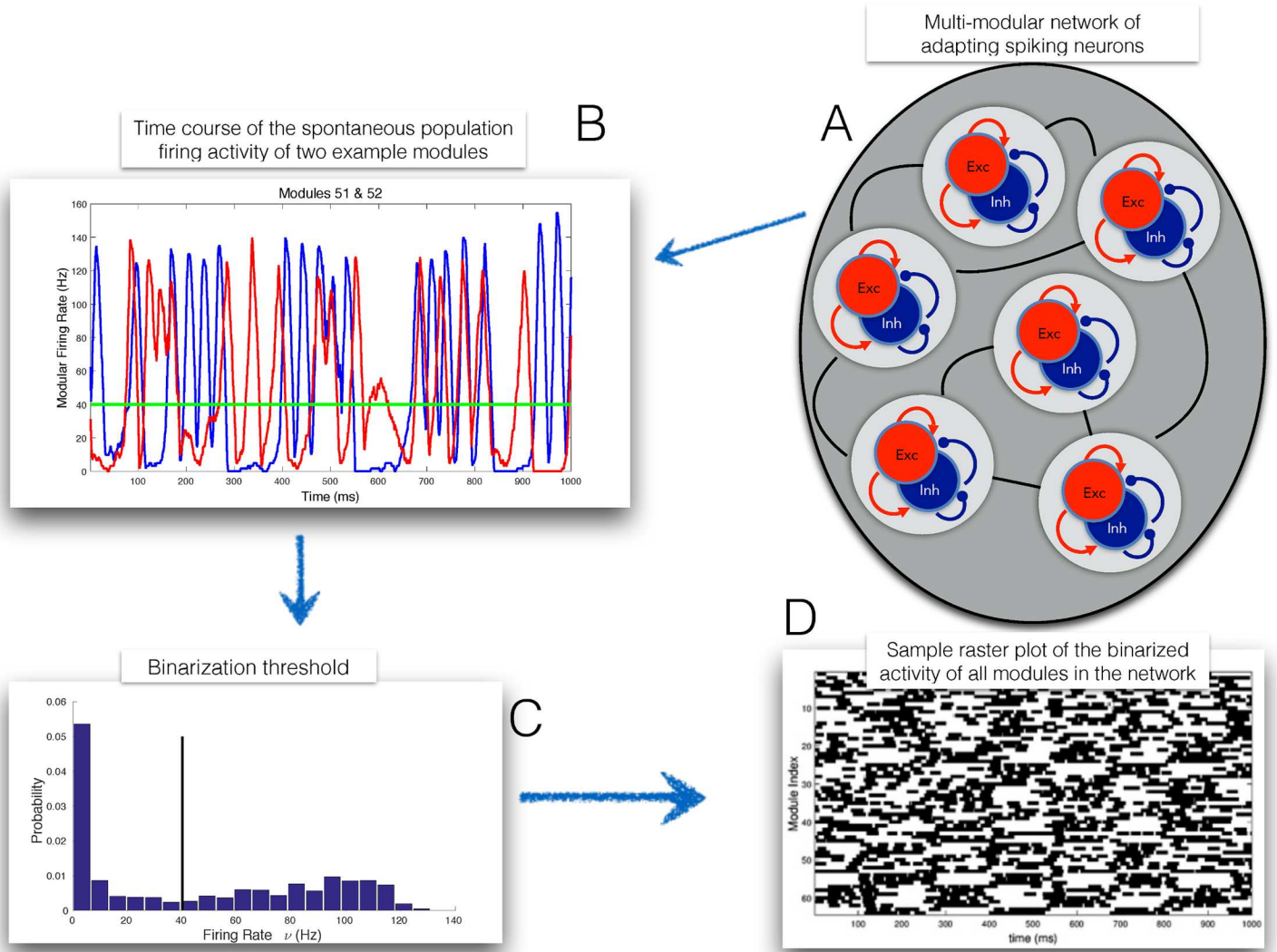
Fig 6. Sketch of the main steps involved in the network construction. See text for details.

the optimal synaptic couplings is naturally mapped onto the Rprop algorithm (see [Materials and methods], Eqs (13) and (14)).

In summary, the sequence of steps we take is the following: 1) store a set of patterns *a la* Hopfield in a network of binary neurons (as in the previous section, the Hopfield network will be in its glassy phase); 2) from the Hopfield network, measure spatial correlations and site magnetizations (average activity); 3) set up a multi-modular spiking network of approximately bistable modules; 4) use pseudo-Boltzmann learning to find inter-modular couplings and external rates to mimic correlations and magnetizations of the Hopfield system; 5) perform MS clustering on the configurations generated by simulations of the spiking multi-modular system; 6) check the quality of the result (see below).

The chosen 'reference' Hopfield model has $N = 64$, $P = 4$, $\beta = 1.1$, for which we generate a long Monte Carlo sequence, and measure the spatial correlations and site magnetization that

are to be matched by the multi-modular spiking network with optimal inter-module couplings J_{pq} and external inputs v_p^{ext} obtained from the pseudo-Boltzmann procedure explained above.

One may ask whether the simplest choice that intuition would suggest, *i.e.* simply taking the inter-module synapses as proportional to the computed Hopfield ones, would work. We checked this option, with poor results. One reason is that the low- and high-firing rate states (which are subject to binarization in the clustering procedure) are not dynamically equivalent (contrary to the binary neurons of the Hopfield case): one state can be deterministically more stable than the other (*e.g.* in the sense of the linear stability of the two corresponding fixed points in the mean-field approximation); furthermore, noise is higher in the high-firing rate state (finite-size noise is activity-dependent and acts multiplicatively). Finally, there are 'quenched noise' effects: the synaptic connectivity in each module is a different (small) realization of the same probabilistic model, and can lead to quite different dynamics between modules.

We checked the success of pseudo-Boltzmann learning in several ways.

First, the correlation between the partial correlations measured from the optimal spiking network and those from the reference Hopfield model is very high ($R^2 = 0.975$); the average 'magnetization' for the optimal spiking network is order 10^{-3}, to be compared with the potential range $[-1, 1]$ and the theoretical null value for the Hopfield network. The found optimal external firing rates have large variations between modules (over ±20% with respect to their average), which is expected, since they contribute to compensate for the heterogeneous excitability of the different modules.

Second, we checked the similarity between the cluster structures emerging from the optimal spiking network and that of the reference Hopfield network, by computing the absolute value $|q|$ of the overlaps between the centroids found for the two networks (c^S and c^H respectively for the spiking and Hopfield networks).

Out of the 12 c^H centroids, 9 of them (75%) have $|q| = 1$ with at least one of the c^S; among the others, 2 of them have $|q| \sim 0.97$ with at least one of the c^S; 1 of them has $|q| \sim 0.84$ with one of the c^S.

Conversely, out of the 13 c^S centroids, 8 of them (62%) have $|q| = 1$ with at least one of the c^H; among the others, 2 of them have $|q| \sim 0.97$ with at least one of the c^S; 3 of them have $|q| \sim 0.75$ with at least one of the c^H.

This shows that the pseudo-Boltzmann iterative procedure, by enforcing approximately equal mean activities and spatial pair correlations between the Hopfield model and the modules of the spiking network results in fact in similar mostly visit regions of the state space for the two systems.

Third, we inferred inter-modular synaptic efficacies (using MPF as in the previous Section) from the spiking network time series, and found that for all modules pairs the inferred synapses are close to corresponding synapses of the reference Hopfield model (the mean absolute value of the relative error, taken over the J entries, is 11%, likely mostly due to finite-sample noise, as suggested by comparison with Fig 5); this confirms that, despite the large differences between the synapses of the Hopfield network and the optimal inter-modular average synaptic efficacies determined by pseudo-Boltzmann learning (not shown), the dynamics of the resulting spiking network effectively embodies inter-module interactions consistent with the reference Hopfield network.

To summarize, the somewhat complex procedure described allowed us to construct a modular spiking system with some control on desired features of the state space, not easily enforced by simple *ad hoc* assignment of the synaptic structure; while in this case the construction was guided by a reference Hopfield model, whose neurons were naturally mapped onto the

network's modules, more in general we believe the procedure is interesting *per se*, as a means to enforce a prescribed pattern of spatial correlations (and associated state-space structure) in relatively complex networks of spiking neurons.

Dynamics in the centroid space

When performing MS clustering in the state space, information about the dynamics of the original time series is—by construction—lost. However, once clustering is done, knowledge of the clusters' centroids allows to go back to the multi-dimensional dynamics, and cast it in a useful compact form: to each one of the vectors expressing the states at successive sampling times, we substitute the label of the cluster that vector was assigned to.

The description of the system's dynamics is reduced to a 'symbolic dynamics' in the centroids space, which opens up options to expose dynamic features that may be difficult to uncover directly from the analysis of the spiking activity, as we illustrate through an example in the present section.

Based on the procedure developed in the previous Section to set up multi-modular spiking networks, we want to generate a family of networks for which different degrees of 'complexity' can be expected, instantiated here in different dynamic memory span, induced by the SFA component that, as discussed, affects locally the dynamics in a history-dependent way.

In the previous section SFA was chosen small, just enough to obtain measurable state transition rates in the simulation time. Here we set up a series of spiking multi-modular networks, each one constructed as in the previous section, but with different parameters for SFA.

We span a range of values for the timescale of SFA (τ_{SFA}), while keeping the product τ_{SFA} g_{SFA} constant (in order to keep the SFA 'strength' constant and have comparable systems, see [Materials and methods]).

We show here that the reduced dynamics in the centroid space lends itself naturally to methods for symbols-oriented measures of complexity, able to easily capture memory effects.

It has long been suggested, and reported in several published works, that the Lempel-Ziv (LZ) complexity measure [19] may be usefully adapted to characterize neural data series [34–36]; in particular, a suitably normalized LZ complexity has been successfully employed as a diagnostic measure of the distance to the conscious state in neurological patients [37]. In essence, the approach is based on the intuitive idea that the more complex the signal, the less its compressibility; in other words, more structure in the signal increases its predictability. Therefore, a memoryless stochastic time series would have maximal complexity, and any memory embedded in the dynamics generating the time series would make it decrease. Such entropic measures provide information beyond what linear correlation analysis can provide.

We therefore measure, for the symbolic sequence reduction of the multi-dimensional dynamics obtained from networks with different τ_{SFA}, LZ complexity and a relative complexity index, as detailed in [Materials and methods].

Expectation is that increasing τ_{SFA} generates multidimensional time series with decreasing complexity. We checked that such expectation is met in our spiking simulation data, and that differences in LZ complexity for high and low τ_{SFA} are statistically significant for the simulated time span.

In order to quantify the difference, for each value of τ_{SFA} and g_{SFA} we also simulated ten realizations of Markov chains in the centroid space, generated by the transition probabilities estimated from the simulation, thereby producing surrogate memoryless sequences with the same transition statistics as the data; we computed the LZ complexity averaged over the ten Markov chains ('c_{Markov}'), and to compare it to the LZ complexity from the simulation data ('c_{sample}'), we computed the ratio $R = (c_{Markov} - c_{sample})/c_{Markov}$ (see [Materials and methods]).
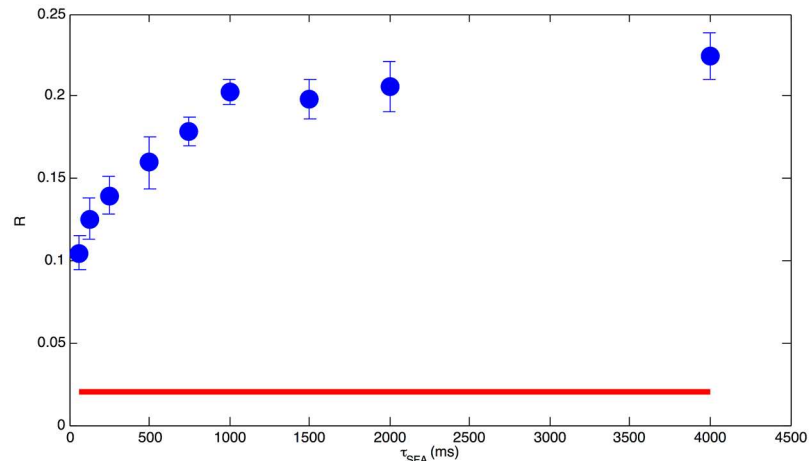
Fig 7. Dependence of the complexity measure R vs τ_{SFA}. Red line: R computed for the Markovian sequence generated by the reference Hopfield system; blue symbols: R vs τ_{SFA} for the actual sequence.

We note that including self-transitions would blur the difference and decrease the LZ complexity, since the corresponding runs of identical centroid labels add to the compressibility of the sequence; we did not consider self-transitions in the analysis.

For a meaningful comparative analysis of the complexity of the dynamics in the centroid space for widely different values of τ_{SFA} (and related g_{SFA}), we should make sure that the corresponding networks are indeed similar between themselves, and with the reference Hopfield network, in terms of the state space structure.

For this purpose we did a preliminary analysis of the clusters found for all the explored values of τ_{SFA}; as τ_{SFA} increases, an increasing number of small clusters appears, with small differences from main ones, which is due in part to the fact that, because of SFA, 'bridge states' appear, with associated small clusters, where the system transits just after SFA has destabilized one major attractor state, on its way to make a transition to another major attractor state. We therefore first clustered the centroids with standard methods (hierarchical clustering using Hamming distance and complete linkage), to obtain a 'fuzzy' version of the whole set of main centroids; for those (76 clusters), we found that: the first 8 fuzzy centroids account for 66.3% of the total configuration mass and are detected, on average, for 84.0% of the τ_{SFA} values; the first 10 fuzzy centroids account for 71.7% of the total configuration mass and are detected, on average, for 73.0% of the τ_{SFA} values; the centroids of the reference Hopfield network are recovered in 87.2% of cases (by 'recovered' we mean that at least one of the centroids found for an Hopfield simulation-clustering has overlap >0.719 with the considered Hopfield centroid, the threshold value 0.719 being determined such that 95% of the overlaps between the centroids found for the considered case are below such threshold.

The results of the analysis are summarized in Fig 7: we see that the ratio R increases with increasing τ_{SFA} or, in other words, that as memory effects increase the complexity of the actual centroid sequences gets increasingly larger than that of surrogate Markov sequences, which provides a quantitative information on the non-Markovian nature of dynamics for higher τ_{SFA} cases (we use the term 'non-Markovian' here in an informal sense, not distinguishing between higher-order Markovian and strictly non-Markovian processes). In the figure we also report for comparison (red line) the R value obtained for the centroid sequence extracted from the time series generated by the reference Hopfield network (which is inherently Markovian); it is

seen that, although the centroid sequence deviates from a Markov process, the relative difference w.r.t. the surrogate sequence is very small, much smaller than the one for the spiking dynamics, where SFA plays a major role.

Of course, deviations from a Markov, memoryless dynamics can take many forms; a generic expectation is that, for a non-Markov process, sequences of states of given length are more, or less, likely to occur than predicted based only on the transition probability matrix.

To gain insight for the case at hand, we considered the centroid sequence for the spiking network with the highest $\tau_{SFA} = 4$s (of length about 1.3×10^3). For each of the triplets of centroid labels (13^3 triplets, since 13 centroids were extracted in this case by MS clustering) the probability of occurrence was estimated from the actual sample, and computed from the Markov transition probabilities estimated from the same sample.

Fig 8, left panel, shows (blue points) a scatter plot of such probabilities (limited to the triplets occurring more than 10 times). With reference to the black identity line (close to which the points would obviously cluster if the actual sequence was Markovian), it is seen that many triplets are over- or under-represented in the actual sequence (particularly in the region of higher probabilities which matters most), as a reflection of its non-Markovian nature.

In order to assess the significance of the observed differences, we also generated a truly Markovian sequence with the same transition probabilities and of the same length as the sample, and from it we estimated the triplets probabilities; the green crosses show the corresponding scatter plot with the computed Markov probabilities, and it is clearly seen that finite-sample effects are much smaller than the spread observed for the actual sequence, confirming its genuine non-Markovian nature.

Finally, given the dependence of R on τ_{SFA} shown in Fig 7, it was natural to ask how the non-Markovian estimated triplets occurrence probabilities would depend on τ_{SFA}; this is illustrated in Fig 8, right panel, where we report the Kullback-Leibler distance between the sampled triplets distributions from the actual sequence and the surrogate Markov sequence, as a function of τ_{SFA}; although in principle for different τ_{SFA} we may expect different contributions to the non-Markov nature of the sequence from sub-sequences of different length, we observe an approximately monotonic dependence of the KL distance on τ_{SFA}.
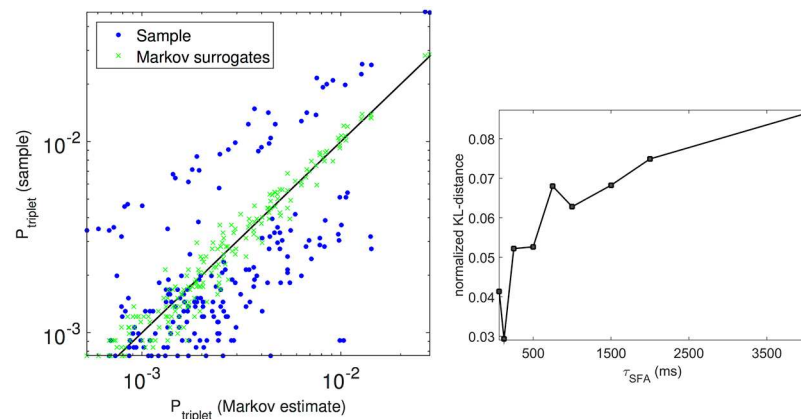

Fig 8. Non-Markovian occurrence of triplets of states. Left: for all triplets occurring more than 10 times, the blue points are the estimated probability from the actual sequence *vs* the computed Markov probabilities from the estimated transition probability matrix; green crosses are the estimated probability from the actual sequence *vs* the estimated probability from the Markov surrogates. Right: the Kullback-Leibler distance between the sampled triplets distributions from the actual sequence and the surrogate Markov sequence, as a function of τ_{SFA}.

https://doi.org/10.1371/journal.pone.0174918.g008

In summary, reducing the multidimensional time series of spiking data to the sequence of labels of the clusters identified by the state-space clustering procedure, casts the dynamics in a form easily suited to capture and quantify traces of memory effects, for instance allowing in principle the comparison between recordings in different experimental conditions.

Discussion

In this work we considered pre-existing strategies and developed new tools that, taken together, compose a methodology with good potential, we believe, in analysis and modeling of neuroscience data.

We started from a simple idea (though, to our knowledge, it was not exploited so far in the analysis of multiple neural recordings): to represent the multidimensional time series of neural activities as a density distribution in a corresponding multidimensional space, and perform a density-based clustering procedure to extract the local density maxima. We strived to show that this type of dimensional reduction is in fact a versatile instrument; in particular we illustrated examples that it can be useful in achieving better inference of synaptic couplings from neural activities, and also to cast the multi-dimensional neural dynamics in a compact form amenable to symbol-oriented methods of complexity analysis.

The clusters, and the associated centroids, have an obvious interpretation when the original time series is generated by an attractor dynamics, but retain an informative value even when (like in the case of SFA in the spiking network) the picture of a static attractor landscape is no longer appropriate; preliminary work in progress on multiple in-vivo recordings during motor tasks makes us confident that the approach can provide compact and usable information even in strongly non-stationary conditions.

We first validated the method on the activity generated by a simple Hopfield network, albeit we choose for it a highly nontrivial working regime in which the energy landscape explored during the network dynamics is 'glassy', with many local minima uncorrelated from the memory pattern embedded in the Hebbian synaptic matrix. We showed that our (modified) Mean-Shift clustering is indeed effective in revealing features of the energy landscape from very noisy time series. Besides, knowledge of the clusters' centroids allowed an efficient parametrization of the synaptic matrix, which in turn allowed a much better result when inferring the synaptic efficacies from the network's sampled activity.

A few remarks are appropriate in this respect. First, it is clearly expected that for an infinite amount of data 'full' inference will be asymptotically better than any parametrization involving fewer parameters than the N^2 elements of the synaptic matrix; however, we already remarked that for limited data samples, a large number of parameters available for inference can lead to overfitting, which can make a reduced parametrization advantageous. Second, of course assuming a form for the inference model inevitably introduces a bias: this is true for our 'Hopfield-like' form, as it is for all works adopting an 'Inverse Ising' inference approach. In a sense, our approach extracts the 'best Hopfield model' compatible with the set of centroids of the found clusters; in this respect, it is interesting to note that in the case we showed, the Hopfield model generating the data is not in the retrieval phase, so that several centroids are found which do not coincide with any of the patterns embedded in the original synaptic matrix, and which also are the largest; nonetheless, the result of inference with the parametrized model gives essentially non-zero weights only to the centroids coinciding with the patterns: this may suggest (but should be further investigated) that forcing a Hopfield-like form on the inference model results in the extraction of the uncorrelated components of the centroids set.

Furthermore, the proposed approach addresses, though admittedly in a specific context, a general issue. Whatever the specific method adopted, the extent to which the inference of

single synapses can be trusted can be severely affected by several factors, like inherent inadequacy of the model used for inference, poor quality of the data and noise, limited data sample. Still, it is usually (explicitly or implicitly) assumed that even when the inference procedure fails to match single synaptic efficacies, if the synaptic matrix has a global structure it should still be captured in the inferred matrix. This would almost inevitably call for some kind of dimensional reduction of the inferred synaptic matrix, such that the informative relevant 'mesoscopic' structure is retained. In a sense, what we propose here can be viewed as a way to formulate a 'mesoscopic' inference problem in the first place; again, we cannot claim full generality here, but our clear success in the case of attractor networks examined in the present work makes the approach, we believe, an interesting option to be further explored.

Wanting to move to more realistic network models, we were naturally led to networks of spiking neurons (Integrate-and-fire, with SFA); in the Introduction and in the Results we provided motivations for a specific choice of the network architecture as composed of weakly coupled, individually bistable neural populations. In order to make contact with the study of clustering in the Hopfield network and, more importantly, to have some *a priori* knowledge of the effective landscape of the spiking system, we also wanted to set up the inter-module couplings such that the spiking network as a whole would share static properties of the state space with an Hopfield attractor network. This need motivated us to develop a procedure that determines the inter-modular couplings such that the network's activity generates a prescribed pattern of pairwise spatial correlations.

The procedure involves a new algorithm that extends the domain of applicability of Boltzmann learning, and uses Rprop learning; as already remarked, we believe this approach has a value beyond the specific purpose it served in the present work.

It rests on the intuition that between excitatory synaptic connection strength and neuronal activity correlation a monotone relationship should hold. On a more general level, this is just an instance of a strategy (which is also a human ability) to identify the relevant variables in a problem and code them in such a way that they have a conditionally monotone relationship with the relevant observables or, more specifically, with the statistic deemed sufficient for the problem under exam [38]. Such ability, coupled with the proven effectiveness in a variety of contexts of optimization algorithms that depend only on the sign of the derivative of the function to be optimized (as Rprop) [17, 18], can make the idea behind the proposed algorithm robustly generalizable to a wide array of problems, dealing both with static and dynamic properties of neuronal networks, *e.g.* by taking into account spatial as well as temporal correlations.

Finally we went back to dynamics. A natural step was to substitute the original multidimensional time series with the corresponding 'symbolic dynamics' of centroid labels, and to ask whether the resulting paths in the centroid space would allow to extract information on the system's original dynamics that would be difficult to directly expose. A case in point, in our context, was to inspect how the sequences of transitions between centroids in the spiking modular network would depend on the strength of SFA. SFA introduces 'memory' in the dynamics, and higher SFA makes the original time series more history-dependent, and the corresponding symbolic dynamics of centroid labels is expected to be less Markovian. As discussed in the text, the pattern of transitions between centroids results from an interplay, in the original time series, between noise, spatial overlaps between attractor states of the multi-modular network, and SFA-dependent effects.

To quantify the memory-related complexity of the network activity, we defined a measure based on Lempel-Ziv (LZ) complexity (inspired to previous work in various scientific domains, including neuroscience): for different time scales of SFA, we compared the LZ complexity of the centroid time series with surrogate Markov sequences with the same transition probabilities, finding that, as expected, longer timescales of SFA correspond to less complex

(and less Markovian) centroid sequences. We also provided insight into such SFA-dependent non-Markovian nature by studying the occurrence probability of selected sub-sequences.

Again, while we illustrated in some detail this reduction to symbolic dynamics and the analysis of its complexity in the specific case under consideration, its value rests with its generic applicability to multidimensional time series.

A few closing remarks, to facilitate comparison with other approaches to the characterization of multidimensional time series in neuroscience (including Hidden Markov Models, that recently have been frequently used in the analysis of neural data (see *e.g.* [39], [23], [24])).

First, the proposed state-space approach is free from bias towards spherical clusters and from a pre-defined number of clusters (or states, in the case of Hidden Markov Models). Such freedom is inherent in the approach taken here, which is also quite easy to implement.

Second, for large data sets the density-based clustering approach can become computationally expensive, and of course it is meaningful to try to speed it up. A recent successful attempt was made in [7], where a preliminary distance-based selection procedure excludes outliers (points with low local density) from the iterative procedure. The validity and performance of the approach is tested in a variety of benchmark data sets, including the Olivetti face dataset [40]. In comparing to the present work, we remark that on the one hand we did not make an effort towards computational efficiency; rather, we wanted to show the potential of density-based clustering for the analysis of neural data from multiple simultaneous recordings. On the other hand, we tried the method proposed in [7], and checked that it performs poorly in several representative situations, due to the high level of noise. A reasonable strategy would probably be a mixed method which first performs a number of iterations (dependent on the noise level) of the mean-shift algorithm (to 'clean up' enough the data distribution in the configuration space), followed by the faster procedure described in [7].

Third, in the very recent paper [2], the authors use clustering to define patterns of populations activity and study how the transitions between such patterns are affected by sensory or behavioral events during trials of a cue-driven decision in a virtual navigation task, and generates an events-dependent evidence accumulation process. Our approach to the complexity analysis of the symbolic dynamics in the centroid space, with the associated estimate of the departure from Markovian behavior, may offer an additional angle from which to look at such dynamically generated memory effects.

Acknowledgments

Author Contributions

Conceptualization: GB GG PDG.

Formal analysis: GB GG PDG.

Methodology: GB GG PDG.

Software: GB.

Writing – review & editing: GB GG PDG.

References

1. Stevenson IH, Kording KP. How advances in neural recording affect data analysis. Nature neuroscience. 2011; 14(2):139–142. https://doi.org/10.1038/nn.2731 PMID: 21270781

2. Morcos AS, Harvey CD. History-dependent variability in population dynamics during evidence accumulation in cortex. Nature Neuroscience. 2016;.

3. Schneidman E, Berry MJ, Segev R, Bialek W. Weak pairwise correlations imply strongly correlated network states in a neural population. Nature. 2006; 440(7087):1007–1012. https://doi.org/10.1038/nature04701 PMID: 16625187

4. Hertz JA, Roudi Y, Tyrcha J. Ising Models for Inferring Network Structure from Spike Data. In: Quiroga RQ, Panzeri S, editors. The Oxford Handbook of Innovation. CRC Press; 2013. p. 527–546.

5. Roudi Y, Dunn B, Hertz J. Multi-neuronal activity and functional connectivity in cell assemblies. Current opinion in neurobiology. 2015; 32:38–44. https://doi.org/10.1016/j.conb.2014.10.011 PMID: 25463563

6. Capone C, Filosa C, Gigante G, Ricci-Tersenghi F, Del Giudice P. Inferring synaptic structure in presence of neural interaction time scales. PloS one. 2015; 10(3):e0118412. https://doi.org/10.1371/journal.pone.0118412 PMID: 25807389

7. Rodriguez A, Laio A. Clustering by fast search and find of density peaks. Science. 2014; 344 (6191):1492–1496. https://doi.org/10.1126/science.1242072 PMID: 24970081

8. Fukunaga K, Hostetler L. The estimation of the gradient of a density function, with applications in pattern recognition. Information Theory, IEEE Tran. 2014; 21(1):32–40. https://doi.org/10.1109/TIT.1975.1055330

9. Mattia M, Pani P, Mirabella G, Costa S, Del Giudice P, Ferraina S. Heterogeneous Attractor Cell Assemblies for Motor Planning in Premotor Cortex. J Neurosci. 2013; 33:11155–11168. https://doi.org/10.1523/JNEUROSCI.4664-12.2013 PMID: 23825419

10. Latimer KW, Yates JL, Meister ML, Huk AC, Pillow JW. Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. Science. 2015; 349(6244):184–187. https://doi.org/10.1126/science.aaa4056 PMID: 26160947

11. Sohl-Dickstein J, Battaglino PB, DeWeese MR. New Method for Parameter Estimation in Probabilistic Models: Minimum Probability Flow. Phys Rev Lett. 2011; 107:220601. https://doi.org/10.1103/PhysRevLett.107.220601 PMID: 22182019

12. Roudi Y, Tyrcha J, Hertz J. Ising model for neural data: Model quality and approximate methods for extracting functional connectivity. Phys Rev E. 2009; 79:051915. https://doi.org/10.1103/PhysRevE.79.051915

13. Liu YH, Wang XJ. Spike-frequency adaptation of a generalized leaky integrate-and-fire model neuron. Journal of computational neuroscience. 2001; 10(1):25–45. https://doi.org/10.1023/A:1008916026143 PMID: 11316338

14. La Camera G, Rauch A, Lüscher HR, Senn W, Fusi S. Minimal models of adapted neuronal response to in Vivo–Like input currents. Neural computation. 2004; 16(10):2101–2124. https://doi.org/10.1162/0899766041732468 PMID: 15333209

15. Gigante G, Del Giudice P, Mattia M. Frequency-dependent response properties of adapting spiking neurons. Mathematical biosciences. 2007; 207(2):336–351. https://doi.org/10.1016/j.mbs.2006.11.010 PMID: 17367823

16. Mattia M, Del Giudice P. Efficient event-driven simulation of large networks of spiking neurons and dynamical synapses. Neural Computation. 2000; 12(10):2305–2329. https://doi.org/10.1162/089976600300014953 PMID: 11032036

17. Riedmiller M, Braun H. RPROP—A Fast Adaptive Learning Algorithm. Proc. of ISCIS VII), Universitat; 1992.

18. Igel C, Hüsken M. Empirical evaluation of the improved Rprop learning algorithms. Neurocomputing. 2003; 50:105–123. https://doi.org/10.1016/S0925-2312(01)00700-7

19. Lempel A, Ziv J. On the complexity of finite sequences. IEEE Transactions on information theory. 1976; 22(1):75–81. https://doi.org/10.1109/TIT.1976.1055501

20. Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences. 1982; 79(8):2554–2558. https://doi.org/10.1073/pnas.79.8.2554

21. Amit DJ. Modeling brain function: The world of attractor neural networks. Cambridge University Press; 1992.

22. Abeles M, Bergman H, Gat I, Meilijson I, Seidemann E, Tishby N, et al. Cortical activity flips among quasi-stationary states. Proceedings of the National Academy of Sciences. 1995; 92(19):8616–8620. https://doi.org/10.1073/pnas.92.19.8616

23. Luczak A, Barthò P, Marguet SL, Buzsàki G, Harris KD. Sequential structure of neocortical spontaneous activity in vivo. Proceedings of the National Academy of Sciences. 2007; 104(1):347–352. https://doi.org/10.1073/pnas.0605643104

24. Mazzucato L, Fontanini A, La Camera G. Dynamics of multistable states during ongoing and evoked cortical activity. The Journal of Neuroscience. 2015; 35(21):8214–8231. https://doi.org/10.1523/JNEUROSCI.4819-14.2015 PMID: 26019337

25. Abdollah-nia MF, Saeedghalati M, Abbassian A. Optimal region of latching activity in an adaptive Potts model for networks of neurons. Journal of Statistical Mechanics: Theory and Experiment. 2012; 2012 (02):P02018. https://doi.org/10.1088/1742-5468/2012/02/P02018

26. Duarte RC, Morrison A. Dynamic stability of sequential stimulus representations in adapting neuronal networks. Frontiers in computational neuroscience. 2014; 8:124. https://doi.org/10.3389/fncom.2014.00124 PMID: 25374534

27. Huguet G, Rinzel J, Hupé JM. Noise and adaptation in multistable perception: Noise drives when to switch, adaptation determines percept choice. Journal of vision. 2014; 14(3):19–19. https://doi.org/10.1167/14.3.19 PMID: 24627459

28. Roach JP, Sander LM, Zochowski MR. Memory recall and spike-frequency adaptation. Physical Review E. 2016; 93(5):052307. https://doi.org/10.1103/PhysRevE.93.052307 PMID: 27300910

29. Deco G, Rolls ET. Sequential memory: a putative neural and synaptic dynamical mechanism. Journal of Cognitive Neuroscience. 2005; 17(2):294–307. https://doi.org/10.1162/0898929053124875 PMID: 15811241

30. Theodoni P, Kovács G, Greenlee MW, Deco G. Neuronal adaptation effects in decision making. The Journal of Neuroscience. 2011; 31(1):234–246. https://doi.org/10.1523/JNEUROSCI.2757-10.2011 PMID: 21209209

31. Mattia M, Sanchez-Vives MV. Exploring the spectrum of dynamical regimes and timescales in spontaneous cortical activity. Cognitive neurodynamics. 2012; 6(3):239–250. https://doi.org/10.1007/s11571-011-9179-4 PMID: 23730355

32. Gigante G, Mattia M, Del Giudice P. Diverse population-bursting modes of adapting spiking neurons. Physical Review Letters. 2007; 98(14):148101. https://doi.org/10.1103/PhysRevLett.98.148101 PMID: 17501315

33. Akrami A, Russo E, Treves A. Lateral thinking, from the Hopfield model to cortical dynamics. Brain research. 2012; 1434:4–16. https://doi.org/10.1016/j.brainres.2011.07.030 PMID: 21839426

34. Kaspar F, Schuster H. Easily calculable measure for the complexity of spatiotemporal patterns. Physical Review A. 1987; 36(2):842. https://doi.org/10.1103/PhysRevA.36.842

35. Abásolo D, Simons S, da Silva RM, Tononi G, Vyazovskiy VV. Lempel-Ziv complexity of cortical activity during sleep and waking in rats. Journal of neurophysiology. 2015; 113(7):2742–2752. https://doi.org/10.1152/jn.00575.2014 PMID: 25717159

36. Amigó JM, Szczepański J, Wajnryb E, Sanchez-Vives MV. Estimating the entropy rate of spike trains via Lempel-Ziv complexity. Neural Computation. 2004; 16(4):717–736. https://doi.org/10.1162/089976604322860677 PMID: 15025827

37. Casali AG, Gosseries O, Rosanova M, Boly M, Sarasso S, Casali KR, et al. A theoretically based index of consciousness independent of sensory processing and behavior. Science translational medicine. 2013; 5(198):198ra105–198ra105. https://doi.org/10.1126/scitranslmed.3006294 PMID: 23946194

38. Dawes RM. The robust beauty of improper linear models in decision making. American psychologist. 1979; 34(7):571. https://doi.org/10.1037/0003-066X.34.7.571

39. Abeles M, Gat I. Detecting precise firing sequences in experimental data. Journal of Neuroscience Methods. 2001; 107(1-2):141–54. https://doi.org/10.1016/S0165-0270(01)00364-8 PMID: 11389951

40. Samaria FS, Harter AC. Parameterisation of a stochastic model for human face identification. In: Applications of Computer Vision, 1994., Proceedings of the Second IEEE Workshop on; 1994. p. 138–142.