



Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth



Characterizing neural phase-space trajectories via Principal Louvain Clustering

Mark M. Dekker^{a,b,*}, Arthur S.C. França^c, Debabrata Panja^{a,b}, Michael X. Cohen^c

^a Department of Information and Computing Sciences, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands

^b Centre for Complex Systems Studies, Utrecht University, Minnaertgebouw, Leuvenlaan 4, 3584 CE Utrecht, The Netherlands

^c Radboud University Medical Center, Donders Centre for Medical Neuroscience, The Netherlands

ARTICLE INFO	ABSTRACT	

Louvain clustering Neural oscillations Cortex Neural synchronization Principal components analysis

Keywords:

Background: With the growing size and richness of neuroscience datasets in terms of dimension, volume, and resolution, identifying spatiotemporal patterns in those datasets is increasingly important. Multivariate dimension-reduction methods are particularly adept at addressing these challenges.

New method: In this paper, we propose a novel method, which we refer to as Principal Louvain Clustering (PLC), to identify clusters in a low-dimensional data subspace, based on time-varying trajectories of spectral dynamics across multisite local field potential (LFP) recordings in awake behaving mice. Data were recorded from prefrontal cortex, hippocampus, and parietal cortex in eleven mice while they explored novel and familiar environments.

Results: PLC-identified subspaces and clusters showed high consistency across animals, and were modulated by the animals' ongoing behavior.

Conclusions: PLC adds to an important growing literature on methods for characterizing dynamics in highdimensional datasets, using a smaller number of parameters. The method is also applicable to other kinds of datasets, such as EEG or MEG.

1. Introduction

The dimensionality of neuroscience measurements has increased drastically over the past few decades (Stevenson and Kording (2011); Danielle and Sporns (2017)), meaning that neuroscientists are now capable of recording simultaneous activity of dozens, hundreds, and even tens of thousands of neurons. Larger and richer datasets provide new opportunities for hypothesis-testing and exploratory discovery, but also challenges in conceptualizing and characterizing the multivariate signals.

Most traditional data analysis methods in neuroscience are univariate or mass-univariate, such as spike counts (in single-unit studies) or spectral power (in LFP studies). Univariate means that each manifest variable – each neuron or electrode – is treated as a unique measurement point, and statistical analyses and neurophysiological interpretations are based on the functions performed by individual neurons.

Univariate methods have been and remain the backbone of neuroscience data analysis; however, larger-scale datasets might benefit from multivariate analyses that are based on identifying patterns embedded in population activity (Bartolo et al., 2020; Kriegeskorte and Kievit, 2013), which might only be weakly represented in a single neuron or electrode. Multivariate dimension-reduction methods seek to identify a low-dimensional subspace in which the most relevant activity patterns exist (Luczak et al., 2009; Trautmann et al., 2021; Van Quyen and Bragin, 2007). Such methods can identify task-related signals that are embedded in the patterns of correlations across the data, which might be undetectable when considering the activity of individual neurons or LFP channels. This is due to neural activity being spatiotemporally synchronized across neurons and circuits, and across multiple spatial scales (Buzsáki, 2010; Varela et al., 2001). The multivariate approach is based on the assumption that neural computations are distributed across ensembles (Buzsáki, 2010), and thus identifying spatiotemporally coherent structure in populations may reveal mechanisms that are not apparent when investigating only individual components of the ensemble.

Although univariate analyses are more common in neuroscience, there are several multivariate analyses that are standard in the field, including principal or independent components analysis, factor analysis, generalized eigendecomposition, and so on (Cohen, 2021; Pang et al.,

https://doi.org/10.1016/j.jneumeth.2021.109313

Received 8 March 2021; Received in revised form 29 July 2021; Accepted 4 August 2021 Available online 9 August 2021

0165-0270/© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

^{*} Corresponding author at: Department of Information and Computing Sciences, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands. *E-mail address*: m.m.dekker@uu.nl (M.M. Dekker).

(a) Experimental setup

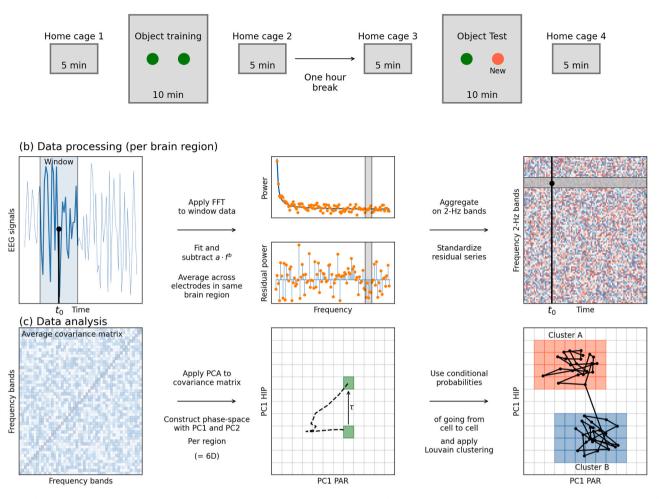


Fig. 1. Panel (a): the experimental setup. Each of 11 mice was held in a home cage for 5 min, and then was placed in an arena with two unknown objects (green dots) for 10 min. We refer to this as the *training phase*. After this, the mouse was returned to its home cage and was given a break for one hour. Then, the process was repeated except that one of the old objects was replaced with a new one (orange circle). We refer to this as the *test phase*. During both the training and test phases, the mouse's LFP signals in the hippocampus, the prefrontal cortex and the parietal cortex were monitored by 32 electrodes. Panel (b): The raw LFP signals were binned into a window surrounding each time step t_0 (blue shaded area), from which the Fourier transform was computed. A 1/*f* function was fitted to the power spectrum, and removed to obtain the residual power spectrum around t_0 . Then, the data were aggregated into 2-Hz bands and normalized, to obtain a plot like in the upperright, in which the value at each element is a residual power value for in 2-Hz frequency bands at each time step. Panel (c): The covariance matrix of the frequency bands was eigen decomposed using principal component analysis (PCA). Utilizing the first two components per brain region led to a six dimensional *phase-space*, which we discretized into equal-volume cells. The multichannel spectra at each time point was localized to a cell in this space, and we computed the conditional probabilities produces clusters, defined as groups of cells where the system is likely to remain within for some time. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2016; Cunningham and Yu, 2014; Makeig et al., 2004). These methods have proven useful at identifying a small number of dimensions in which behaviorally relevant neural dynamics occur.

However, neural data are not static inside those dimensions; they ebb and flow over time with cognitive/behavioral operations, and internal brain states (Mishra et al., 2020; Venkatesh et al., 2019; Zhang et al., 2015; Linderman et al., 2019). Characterizing the time-varying trajectories inside these subspaces is often done visually, which is feasible only in 2 or 3 dimensions.

Here, we introduce a new method for identifying clusters in a lowdimensional data subspace based on spatiotemporal dynamics of continuous trajectories of neural dynamics. The method is based on a combination of principal components analysis (PCA) and Louvain clustering (Dekker et al., 2019), and involves identifying discrete pockets in the principal component (PC) space in which the neural trajectory remains roughly stationary for some period of time. We therefore term this method PLC, for Principal Louvain Clustering. Louvain clustering (Blondel et al., 2008) is a modularity optimization algorithm applicable to weighted graphs, and has been used to characterize spatiotemporal dynamics of railway traffic (Dekker et al., 2019), clusters in air transport networks (Guo et al., 2019), diplomatic structures in formal alliance data (Walentek et al., 2021) and atmospheric states (Tantet et al., 2015).

We modified the Louvain clustering method to multisite LFP data recorded simultaneously from the prefrontal cortex, parietal cortex, and hippocampus, of awake behaving mice during novelty exploration. A spectral decomposition of the LFP via a 1/f-removed short-time Fourier transform was applied, and a six-dimensional (6D) PC space was constructed based on the largest two components from each region. Each time index during the recording is a "brain state point" in this space, and thus successive time points create a trajectory over time. The Louvain clustering method identifies spatial clusters in which these trajectories remain roughly stationary for a period of time τ . Finally, characteristics of these clusters can be quantified and linked to the mice' behavior during the novelty exploration task.

2. Materials and methods

2.1. Animals, electrodes, and behavioral task

Data are from 11 Black57 background male mice. Non-overlapping results from some of these data have been published elsewhere (França et al., 2021; Cohen et al., 2021). The mice had free access to food and water. All experiments were approved by the Centrale Commissie Dierproeven (CCD), and the surgeries and experiments were conducted according to approved indications of the local Radboud University Medical Centre animal welfare body (Approval number 2016–0079).

Custom-designed and self-made electrode arrays (groups of 50 μ m Tungsten wires connected to a PCB) were constructed to target three different regions of the mouse brain. There were 16 electrodes in mPFC (spread in the coordinates AP: 0.5 and 1.5; ML: 0.25 and 0.75; in three columns of electrodes in different depths -2.0, 1.5 and 1.0), eight channels in PAR (AP: -2 and -2.25; ML: 1.0 and 1.75; DV: 0.5) and eight channels in HC (AP -2.5 and -2.75; ML: 1.0 and 1.75; DV: 1.5). Inter-electrode distance was 250 μ m and impedances ranged between 0.1 and 0.9 MOhm. The online reference was a metal screw placed on the skull (interparietal bone - AP: -5, ML: 1.0, DV: 0.5). Additional details of the arrays and the manufacturing process are available in Ref. França et al. (2021). These electrodes recorded data simultaneously.

For surgery, 10–16 week-old mice were anesthetized with isoflurane (induction at 5% isoflurane in 0.5 L/min O_2 ; maintenance at 1–2% in 0.5 L/min O_2). Mice were fixed in a Neurostar Stereotaxic frame. After shaving, the skin was disinfected with ethanol (70%). Local anesthetic xylocaine (2%, adrenaline 1:200,000 [AstraZeneca]) was injected subcutaneously at the incision site before exposing the skull. Peroxide (10–20% H₂O₂; [Sigma]) was applied to the skull with a cotton swab for cleaning and visualization of bregma and lambda. Electrodes and screws were fixed onto the skull with dental cement (Super-Bond C&B). Approximately 40 min prior to the end of the surgery, saline and analgesic (carprofen injected subcutaneous 2.5 mg/Kg) were injected to facilitate recovery.

As illustrated in Fig. 1(a), the experiments involved two sets of 10 min, one where the mice were presented with two objects, and another one where one object was exchanged. More details cvan be found in Ref. França et al. (2021). After the experiments, the mice were euthanized for post-mortem histological confirmation of electrode location. The electrodes in PFC were distributed across anterior cingulate and secondary motor cortex. The PAR electrodes were placed among layers 2–5. HIP electrodes were located in the CA1 region, spread in different mice between *stratum pyramidale* and *stratum lacunosum-moleculare*.

Note that although the anatomical targets were the same in all mice, we do not assume that the electrodes have a one-to-one correspondence across mice (e.g., electrode #x in mouse 1 may not be in the same functional location as electrode #x in mouse 2).

During the recordings, the mice performed a novelty-learning task as depicted in Fig. 1a. For 10 min, they were placed in an arena with two objects that they could explore. Objects were every-day items such as a coffee mug or bath toy. This phase is called the "training phase". After a 60-minute break, they were placed back in the arena for another 10 min, and one of the objects was replaced by a new object (orange circle in Fig. 1a). In between these recording sessions, mice were placed in their home cage; data from those periods are not reported here.

The mice' real-time position was continuously monitored via a webcam sampled at 24 Hz and synchronized with the electrophysiological data. Video data were processed in DeepLabCut (Alexander Mathis et al., 2018), a software package for markerless pose estimation based on convolutional deep neural networks. 200 randomly selected frames were hand-labeled for the left ear, right ear, nose, and the beginning of the tail. The corners of the objects were also labeled. We used the ResNet-101 network with 200,000 iterations, and visual inspection was used to confirm accuracy of the marker labels in test

frames.

Each video frame was given one of three labels with a corresponding variable value ζ : *non-exploration* when all of the mouse's body markers were outside the boundaries of the object ($\zeta = 1$); *exploration* when any of the mouse's head markers were inside the polygon derived from the object corners of the objects ($\zeta = 2$); *novel exploration* when any of the head markers were inside the polygon of the novel object during the test phase (orange dot in Fig. 1a) ($\zeta = 3$). Exploring the familiar object in the test session was labeled $\zeta = 2$.

2.2. Data processing and analysis

Electrophysiology data were acquired using Open Ephys hardware with a sampling rate of 30 kHz. Offline, data were imported into MATLAB, down-sampled to 1000 Hz (by first applying an anti-alias lowpass filter at 500 Hz and then sampling every 30 time points), high-pass filtered at 0.5 Hz, and locally referenced to the average signal from each region. This re-referencing ensured that the signals were locally generated and not volume-conducted from distal brain regions or from the online reference electrode in the skull on top of the cerebellum. The EEGLAB toolbox (Delorme and Makeig, 2004) was used for visual inspection of data quality, and for removing non-neural artifacts (such as mechanical artifacts or line noise) via independent component analysis using the jade algorithm (Cardoso, 1999), as we and others have described previously (Cohen et al., 2021; Khorasani et al., 2019). On average, 1.7 components (out of 32) were removed per recording.

2.3. The PLC method

The overall aim of PLC is to characterize how multivariate brain activity clusters over time. This involves three steps that are detailed below: (1) Preparing the data via spectral time series decomposition, (2) reducing the dimensionality via PCA, and (3) identifying clusters from time-varying trajectories in the low-dimensional subspace via the Louvain method.

All analyses were implemented in Python 3.7.6 using custom-written code that relied on standard Python libraries like numpy, scipy, pandas and matplotlib, and in particular the python-louvain package, which is used for the clustering (https://github.com/taynaud/python-louvain, based on Blondel et al. (2008)). Code will be made available upon acceptance. Below, we continue with a more detailed step by step discussion of the methods.

In the first step, we implemented a time-frequency decomposition via short-time FFT. This is because the time-domain (broadband) signal comprises energy across a range of frequencies, and therefore, a spectral decomposition facilitates identifying the features of the signal.

For the second step, we reduced the dimensionality of these 50 series via PCA, retaining two components per brain region. In total, this produced six principal component time series. The third step involved clustering the 6D PCA space via Louvain method. The clusters represent specific combinations of the activation of frequency bands in the three brain regions that remain pseudo-invariant for some time, which may point to the brain to be in a certain state. In light to the experiment, we will relate these clusters to exploration behavior of the mice.

2.3.1. Spectral time series decomposition

Data preparation involved obtaining a standardized time series of the relative powers of 2 Hz frequency bands from 0 to 100 Hz in 50 steps (Fig. 1b). We began by segmenting the data into 1-second epochs and applied the following procedure: .

- Extract the power spectrum via the Fast Fourier Transform (FFT) to the data in the 1-second segment.
- Fit the function $\rho(f) = a f^b$ to the power spectrum of each 1-second segment, and compute the residual power spectrum after removing this best-fit line. This removes the '1/f' component of the spectrum,

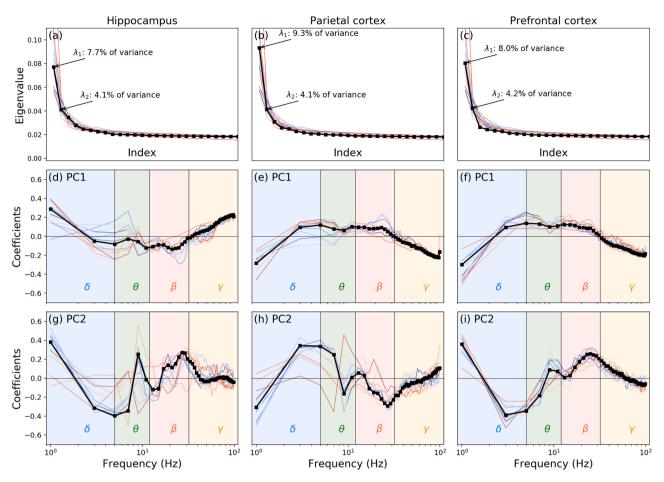


Fig. 2. Panels (a)-(c): eigenspectra of the three brain regions. Panels (d)-(i): the top two PCA eigenvectors. For all panels: black indicates those obtained from the average covariance matrix across all eleven mice, while coloured lines indicate individual mice. The background shading in the bottom two rows indicates approximate canonical spectral boundaries for δ -, θ -, β - and γ -bands. These bands are purely for reference; the analyses was based on the full spectrum up to 100 Hz, not on individual bands. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and thus allows our method to leverage the entire spectrum, rather than being biased by increased energy at low frequencies.

• Crop the resulting 'residual' power spectrum of this window between the frequencies of 2 and 100 Hz and average the values into bands of 2 Hz (resulting in 50 values over the entire frequency range).

The above procedure produced a matrix of time-by-frequency-byelectrodes (per mouse). Next, we averaged the time-by-frequency values across all channels within each brain region, and obtained an estimate of the dominant ("1/f-residualized") spectral dynamics within each brain region. To remove any biases of certain frequency bands having systematically higher power or power-variance, we z-scored the time series for each frequency.

2.3.2. Principal component analysis

We next applied PCA on the residual power spectra over time to obtain a reduced-dimensional representation of the spectral dynamics. We performed the PCA on the time-by-frequency matrices, separately per brain region. The resulting PCs (spectral modes) reflect linear combinations of power across frequencies that maximize the variance of their weighted combinations. In other words, the PC time series are wideband signals, with the contribution of each frequency determined by the PC weights. Also note that these PC time series had a temporal resolution of 1 ms, because the sliding window for the spectral analysis stepped forwards at the same temporal resolution of the LFP data (1 kHz).

In practice, PCA is obtained as the eigendecomposition of the data

covariance matrix (Fig. 1c). In order to get an aggregated result across the 11 mice, we computed the 50-by-50 average covariance matrix of the time series derived above, per brain region. Since the average covariance matrix is symmetric, the PCs constitute an orthonormal basis. The number of PCs to preserve for analyses was selected based on visual inspection of the eigenvalue spectrum (scree plot), which encodes the amount of variance explained by each PC.

Note that although we performed the PCA separately for each brain region, the data for each brain region were acquired simultaneously. This allowed us to localize the data to a reduced PC space that spans all three regions. Using these different dimensions, we constructed a 6D space with the dimensions being $(PC1_{\rm HIP}, PC2_{\rm HIP}, PC1_{\rm PAR}, PC2_{\rm PAR}, PC1_{\rm PFC}, PC2_{\rm PFC})$. While not comprising the whole principal component space, we will refer to this reduced 6D space as the *phase-space*. When the data for each mouse at every time point are projected onto this six-dimensional phase-space, it creates a brain state point corresponding to that time point in the phase space. Connecting these points sequentially in time for a specific mouse then produces the trajectory for the state of its brain in the phase-space.

2.3.3. Louvain clustering

Louvain clustering (Blondel et al., 2008) identifies quasi-stationary pockets of the phase-space—*clusters*—in which the data trajectory remains roughly stable for some period of time. Louvain clustering is a graph clustering method and is based on the concept of modularity, which measures the density of links within the clusters, relative to the density of links between clusters. Weights of links are taken into account in the definition of modularity. The method can be applied to find clusters based on the trajectories in the PC space by discretizing the space into cells. In particular, we gridded the phase-space into 9^6 (531, 441) cells, with 9 cells in each dimension, equally spread in the domain [-12, 12] for all brain regions – HIP, PFC and PAR. (The value 12 as domain boundaries corresponds to the maximum absolute value of the PCs for the entire time series across all brain regions, which is just above 11.5.) We then constructed a matrix *M* with the following entries

$$M_{ij} = P(i \rightarrow j \text{ after } \tau \text{ steps } \mid i), \tag{1}$$

which denotes the conditional probability of the trajectory to move from cell *i* to cell *j* after time τ . In other words, matrix element M_{ij} stores the probability of a jump from cell *i* to cell *j* over a time interval τ . Here, we used $\tau = 30$ ms. *M* is a transfer matrix, and can be used to statistically simulate the time evolution of the state of the mouse brain in phase space (Dekker et al., 2019). It can also be seen as a directed network with weighted edges denoting the conditional probability of transfer from one node to another in a given lag time: nodes in the network are then the phase-space cells in the phase space. Stated differently, Louvain clustering provides an algorithm to identify clusters of phase-space cells in the phase-space in which the mouse brain is (statistically) likely to remain.

The neural trajectories in the phase space are continuous, meaning that exiting one cluster means entering (or passing through) a neighboring cluster. In contrast, a coarse time resolution or larger τ can cause the empirical trajectories to seem discrete and jump from one cell to a distant one. Further, noise can cause a trajectory to exit and rapidly reenter the same cluster. We therefore smoothed the PC time series data using a Gaussian with a full-width-half-maximum of 30-ms (selected to match τ).

The clusters are defined by analyzing spatiotemporal modularity of the neural trajectories in the phase space. In particular, clusters are regions (contiguous cells in the phase space) in which the trajectories remain roughly stable for some period of time, relative to the time it takes to transit to another cluster [Fig. 1(c)]. This is the goal of the Louvain algorithm applied to matrix *M* using time parameter τ . In initial explorations with pilot data, we found that modularity generally and smoothly decreased with increasing τ . Furthermore, the locations and boundaries of clusters were qualitatively comparable with τ 's ranging from 5 to 300 ms. Here we selected $\tau = 30$ based on the approximate time windows of spike smoothing (Lehky, 2010). On the other hand, we acknowledge that this parameter selection is somewhat arbitrary, and it is possible that different features of the data would be highlighted by different values of this parameter. In Sec. 4.3, we discuss this choice in more detail.

2.3.4. Cluster features

The Louvain clusters are defined based solely on the neural trajectory data. In order to characterize these clusters and relate them to mouse behavior, we computed the following attributes for each cluster. Some of these features can also be calculated per mouse, but we computed them as aggregated over all datasets.

• *Exploration bias*, a normalized measure of the percentage of time the mice spent exploring the novel object while the brain was in each cluster. Comparing the exploration percentage while in cluster *i* to the overall exploration percentage yielded its *exploration bias*, defined as

$$\zeta_b(\text{cluster } i) = \frac{\text{percentage exploration in cluster } i}{\text{overall percentage exploration}}.$$
 (2)

If ζ_b (cluster*i*) > 1, then the mice explored the object for longer while in brain-state cluster *i* than they do on average over the whole phasespace. If this number is significantly higher than 1 across the group of mice, the mice explore more often than on average when their brain activity's position in the phase-space places them in cluster *i*, which, via the eigenvectors shown in Fig. 2, can be reconstructed as certain combinations of residual spectra being activated or covarying during exploration.

- Absolute exploration bias, defined as ζ_{abs} = abs(ζ_b 1). This is the distance of ζ_b from one, and depicts the general strength of the bias (exploration or non-exploration).
- Average residence time of state of the mouse's brain within each cluster. Note that since we are working in phase-space, a cell in a given cluster might share boundaries with cells that do not belong to the same cluster. It is possible that noise can make the neural trajectory to rapidly exit-and-reenter a cluster within a few ms. We consider these transient fluctuations to be spurious, and we therefore introduced an additional 30-ms [aligned with the τ in Eq. (1)] allowance period to keep count of the residence time if the state of the brain exits and enters a cluster within this time frame. We also omit residence times of 1 or 2 ms in this statistic.
- *Interregionality*, which indicates whether the cluster has more variance in some brain regions compared to others. This indicates whether the cluster is dominated by one region, or whether it comprises dynamics from two or three regions. We defined interregionality (IR) of cluster *i* as

$$IR(i) = \frac{\max[var(in i, HIP), var(in i, PFC), var(in i, PAR)]}{var(in i, HIP) + var(in i, PFC) + var(in i, PAR)}.$$
(3)

Note that $IR \in [1/3, 1]$; the lower limit 1/3 corresponds to the case where the variance is equally spread across all three brain regions, whereas IR = 1 indicates that the variance is exclusively concentrated within one brain region.

• *Mouse specificity*, which indicates whether a cluster is defined from one or a small number of mice, vs. whether data from most or all mice contributed an equal number of data points. Mouse specificity was computed by first calculating the percentage $p_m(i)$ of cluster *i*'s data points that are from each of the 11 mice m. Then the mouse

specificity is defined as $\sum_{m} |\frac{1}{11} - p_m|$, i.e., as the sum of the absolute differences from 1/11.

• *Cluster magnitude*, the average distance of data points in the cluster to the origin of the phase-space. Distance to the origin was measured by computing the 6D Euclidean distance, weighing each dimension equally.

2.4. Statistics

Statistical robustness was checked in multiple respects. To check whether random exploration data can reproduce the same ζ_b distribution across the clusters, the behavioral time series of exploration were shuffled (without changing the brain or PC data), and the above metrics were recomputed. The shape and number of clusters themselves was checked against random permutations by shuffling the PC data, while preserving the internal 6D coordinates (i.e., only shuffling the time sequence of these 6D data points). A final statistical check was the universality of the results across the mice. To this end, we determined ζ_h for each of the eleven mice and performed a Wilcoxon signed rank test. The Wilcoxon test determines whether any sample \vec{x} is symmetrically distributed around zero. In our case, we test whether it is symmetric around one (testing $\vec{x'} = \vec{x} - 1$). It is different from the regular student's *t*-test in the sense that the Wilcoxon test is non-parametric — it focuses on the signed ranks, rather than the actual values of the differences. In our case, we used it to test the hypothesis that the novel exploration was consistently in the same clusters across animals.

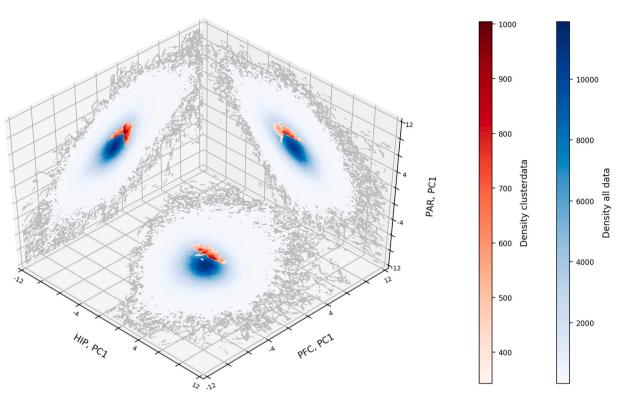


Fig. 3. Example cluster found by applying Louvain clustering to the matrix *M*. The axes are the first PCs from each region. In other words, this shows only 3 of the 6 dimensions of the PC space. Blue indicates the distribution of all data, red indicates the distribution of the data of one example cluster. For visualization purposes, the outer edges of the data clouds are marked in gray. In the background, the resolution of the Louvain clustering is marked (9×9). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Results

3.1. General

Across all eleven mice, the average percentage of time spent exploring the novel object was 16.4%. The individual numbers were 15%, 9%, 16%, 9%, 23%, 16%, 13%, 16%, 13%, 37% and 14%.

The parameters *a* (the scaling factor) and *b* (the exponent) of the fits subtracted from the power spectra in each time step were $a = 0.01 \pm 0.004$ and $b = -0.36 \pm 0.09$. The averages and standard deviations were calculated over all channels, mice, brain regions and parts of the experiment (training + test).

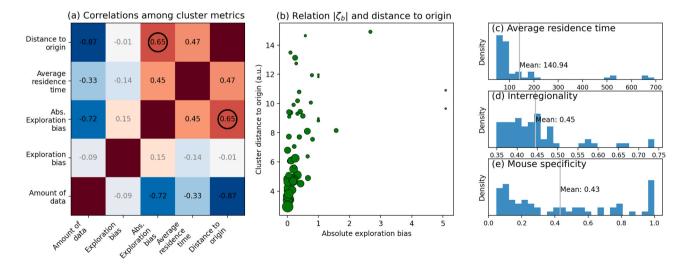


Fig. 4. Panel (a): Spearman's correlation matrix of relevant variables. Negative correlations are denoted in blue, positive correlations are denoted in red. Numerical values of correlations are shown in black for those with correlation significance level p < 0.05, other (non-significant) correlations are shown in gray. Panel (b): Cluster distance to origin versus absolute exploration bias, corresponding to the encircled Spearman correlation of 0.6 in panel (a). Scatter marker size in panel (b) corresponds to amount of data points. For reference, we denoted two clusters with their respective amount of data points. Panel (c)-(e): Histograms and associated mean values of (c) average residence time (in ms), (d) interregionality (between 0.33 and 1, higher means more variance in a particular brain region) and (e) mouse specificity (between 0 and 1, higher means more dominated by a smaller set of mice). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

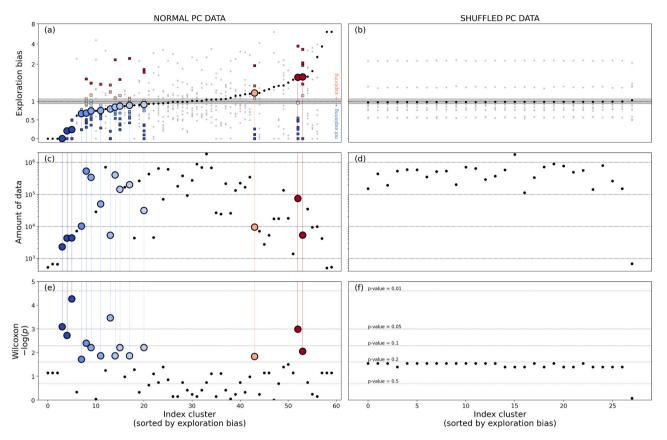


Fig. 5. Attributes of the PC clustering (left column) and the clustering applied to shuffled PC data (right column). Exploration bias for every cluster (panels a-b) is compared to the cluster's respective data amount (panels c-d) and the Wilcoxon *p*-values, that test the consistency of exploration bias values across the 11 animals (panels e-f). Clusters with p < 0.2 on the left hand side are highlighted, all other clusters are denoted by black dots. Squares in panels a-b (Colored and gray) show the respective individual biases per mouse. Grey shaded areas in panels a-b indicate the estimated exploration bias range (mean $+/- 2\sigma$) among 100 times shuffling the exploration data without changing the LFP data. Note that some clusters in panels (a), (c) and (e) on the left only contain few data and non-exploration data, resulting in overlapping points in panel (a). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2. PCA eigenvectors

Inspection of the eigenvalue spectra from the PCA on the residual time series per brain region and per mouse indicates that a large amount of variance in each dataset can be explained only by two PCs (Fig. 2). In the interest of convenience and comparability, upon visual inspection, we retained only the top two PCs from each brain region per mouse.

Remarkably, the PC eigenvectors for each mouse exhibited strong consistency across the mice, as seen in Fig. 2. This motivated us to compute the average covariance matrix for all the mice together, and then rerun the PCA on this average matrix. The resulting PC eigenvectors are displayed in the same figure in black. We used the latter as the reference set of axes onto which we projected the brain state data for all mice. The advantage of this approach is that it provides us with a common phase space for all mice for performing our analyses.

3.3. Clusters

Projecting the brain state data onto the 6D ($PC1_{HIP}$, $PC2_{HIP}$, $PC1_{PAR}$, $PC2_{PAR}$, $PC1_{PFC}$, $PC2_{PFC}$) phase-space produces a time series of the 6D coordinates per mouse. In total, data from eleven mice produced 13,644,532 data points. Gridding the 6D space produced 2417 non-empty cells (c.f. the maximum possible 1,000,000 cells). In the calculation of the conditional probabilities for the brain state to transfer from one cell to another, used as entries in matrix *M*, we made sure that only movements between cells within the data of a single mouse are used. The Louvain clustering on matrix *M* identified 60 clusters.

An example cluster is visualized in Fig. 3. The axes are from the first

PC of each brain region. The blue colors illustrate the distribution of all data on these axes, and the red colors illustrate the density of the example cluster. Note that the data have higher concentration towards the origin of the phase space. This particular cluster is index #4, contains 273 cells, and comprises 529,167 data points (about 4% of the data).

3.4. Cluster attributes

We start investigating the general properties of these clusters by computing the metrics discussed in <u>Section 2.3.4</u>. Fig. 4a shows the correlation matrix (Spearman) for the cluster metrics described in the Methods section.

We focus on the exploration bias and the absolute bias, because these two metrics are linked to the animals' behavior. Exploration bias was largely uncorrelated to the other metrics.

Absolute exploration bias, however, was negatively correlated with the amount of data: clusters with more data had exploration biases closer to 1 (see also Fig. 4b). This is interesting to relate to the positive correlations between absolute exploration bias and residence time and distance to the origin: Clusters in which trajectories remained for longer periods of time — and clusters that were further away from the origin of the PC phase-space — were more likely to be modulated by behavior. One can think of these peripheral clusters are requiring more energy in the PC space to push away from the origin, and may therefore reflect brain states that deviate from the ongoing state.

In contrast, the clusters located towards the center of the PC space contain much more data per cluster, and also more transitions across clusters. Fig. 4c-e shows distributions for three of the cluster properties. The residence times ranged from tens of ms to 700 ms, but were concentrated around 100 ms, with an average of 140 ms. The observation that neural trajectories tend to remain in a stable state for a few hundred ms before transitioning to another stable state is consistent with previous findings in the LFP and EEG literature (Michel and Koenig, 2018; Luczak et al., 2009; Mishra et al., 2020). In the interregionality and mouse specificity, more spread is observed in the histograms, reflecting that some clusters were of a 'universal nature' in terms of how well they were represented across mice and across the three brain regions, but some were not. A relatively high abundance of lower values in both metrics reflects that for the larger fraction of the clusters, dynamics are shared across the animals and reflect coordination across the brain.

3.5. Exploration data and Wilcoxon test

To investigate the significance of the exploration bias, we computed ζ_b for all clusters, both for individual mice and for the group-aggregated data. Results are shown in Fig. 5a (squares for individual data and circles for aggregated data). We sorted the clusters by ζ_b of the full data set. Around half of the clusters had below-average exploration bias. Most clusters showed individual variability across animals (note that the aggregated results can be considered a weighted average, because not all animals had the same amount of data per cluster). Panel (c) shows that clusters also varied in the total amount of data per cluster, with the clusters exhibiting strong exploration biases having overall less data.

We performed three sets of statistical evaluations on the exploration bias results. First, we tested the individual data against a bias score of 1 (corresponding to no exploration bias), which tests the consistency of the sign of the bias across animals. Only animals with at least 2% of their data in the particular cluster are used in this analysis (Wilcoxon ranksign test), as accounting for mice with smaller amounts of data may results in non-representative values of the exploration bias in these clusters. Results are shown in panel (e), with circles indicating data points of p < 0.2. Four clusters exceeded a p < 0.05 threshold. On the one hand, this is a somewhat lenient threshold and was uncorrected for multiple comparisons, but this result is consistent with the cross-animal variability in bias scores seen in panel (a). On the other hand, with N = 11, a Wilcoxon sign-rank test can only achieve p < 0.05 with at least 10 data points in the same direction. We complemented this test using permutation testing. Visually comparing panels (c) and (e) may lead to the conclusion that the exploration significance of the clusters is driven by the amount of data in the cluster. We checked this by performing a biserial rank-correlation between the (dichotomous) significance and the amount of data, resulting in a correlation of 0.22, contradicting this idea.

The second test was based on shuffling the data to determine whether our results could have been observed in random data in the same reduced PC space. To this end, we scrambled the PC time series (which shuffles the links between phase-space cells), a re-applied Louvain clustering. Results are shown in Fig. 5a,d,f. Here we observed only 28 clusters with uniform amount of data and clustering bias. Note that in this process we did not shuffle the exploration data. Because the resulting clusters are now randomly situated in the phase-space, and the amount of data in these clusters is uniform, the exploration data of any mouse is perfectly spread across all clusters, resulting in a near-invariant exploration bias per individual rodent [visible in horizontal dotted 'lines' in panel (b)]. Note also that in all cases, the *p*-values of the shuffled data were larger than the *p*-values of the empirical data labeled as significant.

The final set of statistics involved shuffling the temporal mapping between behavior and brain. We cut the behavior time series at a random time point and swapped the second for the first segment, and then recomputed ζ_b for each cluster. Note that this procedure preserves both the LFP data and the behavior data, only randomizing their behaviour with respect to the LFP signal. This procedure was repeated 100 times, and the gray shaded region in Fig. 5(a) illustrates two standard deviations around the mean shuffled exploration bias.

Overall, these tests confirm that the clustering structure found in the data was not observed when clustering noise, and that there is individual variability in the specific clusters in which animals investigated the novel object that is not accounted for by random reshuffling of the behavioral data.

4. Discussion

4.1. Multivariate neural dynamics

Neural activity is diverse and dynamic over space, time, and frequency (Buzsáki and Llinás, 2017; Varela et al., 2001; Danielle, 2017). Characterizing these dynamics remains a major challenge in neuroscience. Here we adopted a recently developed method (Dekker et al., 2019; Tantet et al., 2015) to define low-dimensional phase-space clusters that allow for characterizing and exploring spatiotemporal dynamics of neural trajectories in multi-site electrophysiology. An advantage of PLC is that it allows for a mix of linear (PCA) and nonlinear (Louvain clustering) methods to identify trajectory-based clustering in a reduced-dimensional space.

4.2. Consistency of PCs across mice

It is striking that the eigenvectors of the top two PCs were so similar across the eleven mice. This is not a trivial result due to biased selection, because we simply took the top two components rather than, e.g., picking the components that maximized a cross-mouse correlation. It is also not a trivial result of the 1/f-like nature of the power spectrum, because this feature was removed prior to the covariance matrix generation.

Instead, we interpret this result to indicate that there are fundamental characteristics of this cortical-hippocampal network that are (1) concentrated in a lower-dimensional space, and (2) conserved across mice. That the dynamics are concentrated in a subspace of the total data space is often reported in multichannel neural recordings (Luczak et al., 2009). This supports the notion that neural computations are distributed over space and frequency, and thus the motivation for investigating weighted combinations of data features (such as electrodes and frequencies). The similarities of the PCs across animals is likely due to intrinsic anatomical and architectural patterns that govern the local neural dynamics. On the other hand, the PCs were not fully identical across mice, and the residual variance likely comprises a combination of unique factors and sampling variability.

Whether to apply multivariate decompositions to data pooled over individuals is debated in the neuroscience literature (Calhoun and Liu, 2009; Parra et al., 2019; Cohen, 2021), and leads to a trade-off between increased generalization vs. increased sensitivity to individuals.

We decided to use the same PC space for all mice in order to have a common (for all mice) phase-space, which facilitated pooling the cluster characteristics across all mice. The mouse specificity in Fig. 4e indicates that there are many clusters in which most or all mice are well represented — indicating that inter-regional brain dynamics of distinct mice are situated in the same clusters — and there were also several clusters with near-1 values of mouse specificity, indicating that some clusters were unique to individual mice.

In this analysis, we have chosen to work with the top two PCs in all three brain regions. For the dataset we use here, these components turned out to be dominant in terms of explained variance, but this is not a trivial or universal result. The PLC method is therefore not strictly bound to the choice of which PCs to use. Depending on the data, it is possible that it is more appropriate to use three in each brain region, for example.

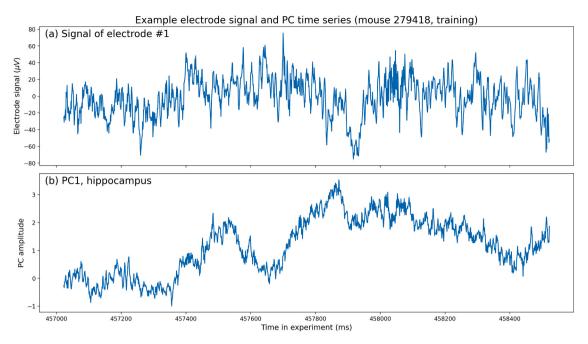


Fig. A.1. Example time series of (a) the raw electrode signal, and (b) a principal component (PC1 Hippocampus) of mouse 278418, the training data set. Only 1500 ms is shown out of approximately 600,000 ms in total.

4.3. Interpretation of clusters and usage of Louvain clustering

Over time, the trajectory of the brain state moves around in the phase-space. These trajectories are not random, and instead remain within relatively confined "provinces" for periods of time. Louvain clustering identifies these periods of relative stability as clusters. A key parameter of Louvain clustering is τ , which is the time-scale at which a trajectory is considered stationary. In our pilot analyses with data from one mouse (not shown here), we found that the clustering was stable for a range of τ values, and thus we selected $\tau = 30$ ms in the interest of consistency. It is possible that different features of the data would be highlighted for larger differences in this parameter, e.g., over hundreds of ms or seconds. Determining how trajectories and clustering are related across temporal scales is an interesting avenue for future research to explore.

Given that the PLC method is several steps away from the raw data, one might wonder whether these clusters are meaningful, or simply reflect residual noise. We addressed this in two ways. First, Louvain clustering applied to shuffled data showed qualitatively distinct patterns of results compared to the real data (e.g., Fig. 4). Indeed, clustering shuffled data produced only 28 clusters, compared to the 60 clusters in the real data, even though the amount of data was the same.

Second, we examined the relationship between the cluster characteristics and exploration behavior. The results shown in Figs. 4 and 5 suggest that clusters close to the center of the PC phase-space were generally not modulated by behavior, whereas clusters further away from the origin had exploration biases that deviated from 1. These findings suggest that the brain space trajectories associated with active behavior are "high-energy" states that are maintained for relatively longer periods of time.

On the other hand, there was also considerable individual variability in the exact exploration bias scores over the different animals, evidenced by the relatively large p-values in the Wilcoxon test (although one must keep in mind that with N = 11, the Wilcoxon test can only return a p < 0.05 result if 10/11 datapoints have the same sign). Some clusters had relatively little or no data for some animals. This finding raises the question of whether the clustering should have been done on individual animals instead of on the group data. As we wrote earlier, group-defined clusters have several advantages in terms of comparability of cluster findings and universality of the spectral-temporal structure of the data. Nonetheless, these are analysis choices that users can determine on their own; our primary purpose here was to illustrate the utility and interpretability of the PLC method.

The quadratic relationship between exploration bias and distance to the phase-space origin is highlighted by the significant correlation with *absolute* exploration bias. Thus, although the direction of the exploration bias score varies across animals and clusters, there is in general more behavioral variability as the data move towards the periphery of the phase-space. It is possible that the center of the space reflects brain states of relatively low energy or of transitions between states of higher energy. On the other hand, one must keep in mind that the outer edges of the phase-space are relatively sparse, which means that the clusters had overall less data. This can become statistically problematic in the extreme case of, for example, a cluster comprising only 50 data points from a single mouse that is exploring the object for the entire time window.

4.4. Limitations and future directions

The PLC method is not without limitations. As mentioned above, clusters with relatively little data, or that are dominated by a single animal, can provide cluster characteristics that may not be fully representative of group dynamics. In the present application, the higher density of data towards the center of the phase-space means that clusters at the outer edges of the data cloud have overall less data. This is apparent in Fig. 5c, which shows differences in densities of clusters spanning four orders of magnitudes. We chose not to remove any clusters based on total amount of data, although we did exclude animals from group-level *t*-tests if those animals had too little data in that cluster.

A second limitation is the operational definition of a cluster in noisy data such as LFP. Small and fast jumps out of, and then back into, a cluster may reflect noise or a true rapid fluctuation between states. This would manifest as low residence times. We chose to filter out jump-but-return trajectories that were faster than 30 ms.

Third, the physiological meaning of PLC-derived clusters remains unknown. We speculate that the brain remains in a particular state while in these clusters, similar to the interpretation of EEG microstates (Michel and Koenig, 2018; Luczak et al., 2009; Mishra et al., 2020). This could be investigated more in the future, for example by investigating spiking dynamics in different LFP-defined clusters. The behavioral relevance of the clusters is also unclear; it is possible that a more easily interpretable link to behavior would result from a simpler experiment (e.g., simple visual responses in anesthetized animals), but we choose for a more naturalistic and therefore more complex environment.

In this paper, we focused on identifying and characterizing the clusters; the data provide additional richness that we did not fully explore. For example, the trajectories exhibit changes in speed, curvature, and direction over time, and may exhibit spatiotemporal oscillations within and across clusters. These dynamics might be related to neural computations and behavior, and might vary systematically during stimulus or motor processing (Linderman et al., 2019; Glaser et al., 2020).

The PLC method is general, and can be applied to many additional tasks or behavior elements. It can also be applied to larger-scale measurements such as EEG and MEG. Depending on the application and the data, it may be the case that the phase-space dimensions have to be adapted, to, for example, focus on a single brain region, or more or fewer principal components.

Author contributions statement

MD, **DP** and **MXC** conceived the method. **AF** performed the experiments. **MD** developed the code and ran the analysis. **MD** wrote the first draft of the manuscript. All authors contributed to the final manuscript.

Declaration of Competing Interest

The authors declare no competing interests.

Acknowledgements

We thank Mihaela Gerova for assistance with data preparation and video processing with DeepLabCut. MXC, ASCF, and this research were funded by an ERC-StG 638589 and a Hypatia fellowship from the Radboud University Medical Center.

Additional information

Accession codes; Data are available upon publication at data. donders.ru.nl. Code is available at github.com/MarkMDekker/Principal _Louvain_Clustering.

Appendix A. Example time series

See Fig. A.1.

References

- Bartolo, Ramon, Saunders, Richard C., Mitz, Andrew R., Averbeck, Bruno B., 2020. Dimensionality, information and learning in prefrontal cortex. PLoS Comput. Biol. 16 (4), e1007514.
- Bassett, Danielle S., Sporns, Olaf, 2017. Network neuroscience. Nat. Neurosci. 20 (3), 353–364.
- Blondel, Vincent D., Guillaume, Jean-Loup, Lambiotte, Renaud, Lefebvre, Etienne, 2008. Fast unfolding of communities in large networks. J. Stat. Mech.: Theory Exp. 2008 (10), P10008.
- György Buzsáki, Neural Syntax: Cell Assemblies, Synapsembles, and Readers, nov 2010. Buzsáki, György, 2010. Neural syntax: cell assemblies, synapsembles and readers. Neuron 68 (3), 362.
- Buzsáki, György, Llinás, Rodolfo, 2017. Space and time in the brain. Science 358 (6362), 482.
- Calhoun, Vince D., Liu, Jingyu, Adalí, Tülay, 2009. A review of group ica for fmri data and ica for joint inference of imaging, genetic, and erp data. NeuroImage 45 (Suppl 1), S163–S172.

- Cardoso, Jean-François, 1999. High-order contrasts for independent component analysis. Neural Comput. 11 (1), 157–192 (2021/02/17).
- Michael X. Cohen, A tutorial on generalized eigendecomposition for source separation in multichannel electrophysiology. arXiv, Apr 2021.
- Cohen, Michael X., 2021. A data-driven method to identify frequency boundaries in multichannel electrophysiology data. J. Neurosci. Methods 347, 108949.
- Cohen, Michael X., Englitz, Bernhard, S. C. França, Arthur, 2021. Large-scale and multiscale networks in the rodent brain during novelty exploration. eNeuro 8 (3).
 Cunningham, John P., Yu, Byron M., 2014. Dimensionality reduction for large-scale
- neural recordings. Nat. Neurosci. 17 (11), 1500–1509.
- Dekker, Mark M., Panja, Debabrata, Dijkstra, Henk A., Dekker, Stefan C., 2019. Predicting transitions across macroscopic states for railway systems. PLoS One 14 (6), 1–26.
- Delorme, Arnaud, Makeig, Scott, 2004. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. J. Neurosci. Methods 134 (1), 9–21.
- França, Arthur S.C., Borgesius, Nils Z., Souza, Bryan C., Cohen, Michael X., 2021. Beta2 oscillations in hippocampal-cortical circuits during novelty detection. Front. Syst. Neurosci. 15, 8.
- França, Arthur S.C., van Hulten, Josephus A., Cohen, Michael X., 2021/02/17 2020. Low-cost and versatile electrodes for extracellular chronic recordings in rodents. Heliyon 6 (9).
- Gao, Weisi, Toader, Bogdan, Feier, Roxana, Mosquera, Guillem, Ying, Fabian, Oh, Se-Wook, Price-Williams, Matthew, Krupp, Armin, 2019. Global air transport complex network: multi-scale analysis. SN Appl. Sci. 1 (7), 680.
- Glaser, Joshua I., Whiteway, Matthew, Cunningham, John P., Paninski, Liam, Linderman, Scott W., 2020. Recurrent switching dynamical systems models for multiple interacting neural populations. bioRxiv (2020.10.21.349282).
- Khorasani, Abed, Shalchyan, Vahid, Daliri, Mohammad Reza, 2019. Adaptive artifact removal from intracortical channels for accurate decoding of a force signal in freely moving rats. Front. Neurosci. 13.
- Nikolaus Kriegeskorte, Rogier A. Kievit.Representational geometry: Integrating cognition, computation, and the brain, aug 2013.
- Lehky, Sidney R., 2010. Decoding poisson spike trains by gaussian filtering. Neural Comput. 22 (5), 1245–1271. PMID: 20028224.
- Linderman, Scott, Nichols, Annika, Blei, David, Zimmer, Manuel, Paninski, Liam, 2019. Hierarchical recurrent state space models reveal discrete and continuous dynamics of neural activity in C. elegans. bioRxiv, 621540.
- Luczak, Artur, Barthó, Peter, Harris, Kenneth D., 2009. Spontaneous events outline the realm of possible sensory responses in neocortical populations. Neuron 62 (3), 413–425.
- Makeig, Scott, Debener, Stefan, Onton, Julie, Delorme, Arnaud, 2004. Mining eventrelated brain dynamics. Trends Cogn. Sci. 8 (5), 204–210.
- Mathis, Alexander, Mamidanna, Pranav, Cury, Kevin M., Abe, Taiga, Murthy, Venkatesh N., Mathis, Mackenzie Weygandt, Bethge, Matthias, 2018. Deeplabcut: markerless pose estimation of user-defined body parts with deep learning. Nat. Neurosci. 21 (9), 1281–1289.
- Michel, Christoph M., Koenig, Thomas, 2018. Eeg microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. NeuroImage 180, 577–593 (Brain Connectivity Dynamics).
- Mishra, A., Marzban, N., Cohen, M.X., Englitz, B., 2020. Dynamics of neural microstates in the vta-striatal-prefrontal loop during novelty exploration in the rat. bioRxiv
- Mishra, Ashutosh, Englitz, Bernhard, Cohen, Michael X., 2020. EEG microstates as a continuous phenomenon. NeuroImage 208.
- Pang, Rich, Lansdell, BenjaminJ., Fairhall, Adrienne L., 2016. Dimensionality reduction in neuroscience. Curr. Biol. 26 (14), R656–R660.
- Lucas C. Parra, Stefan Haufe, Jacek P. Dmochowski, Correlated components analysis extracting reliable dimensions in multivariate data, 2019.
- Stevenson, Ian H., Kording, Konrad P., 2011. How advances in neural recording affect data analysis. Nat. Neurosci. 14 (2), 139–142.
- Tantet, Alexis, Burgt, Fiona R. van der, Dijkstra, Henk A., 2015. An early warning indicator for atmospheric blocking events using transfer operators. Chaos 25 (3), 1–48.
- Trautmann, Eric M., Stavisky, Sergey D., Lahiri, Subhaneil, Ames, Katherine C., Kaufman, Matthew T., O'Shea, Daniel J., Vyas, Saurabh, Sun, Xulu, Ryu, Stephen I., Ganguli, Surya, Shenoy, Krishna V., 2021/02/17 2019. Accurate estimation of neural population dynamics without spike sorting. Neuron 103 (2), 292-308.e4.
- Van Quyen, Michel Le, Bragin, Anatol, 2021/02/17 2007. Analysis of dynamic brain oscillations: methodological advances. Trends Neurosci. 30 (7), 365–373.
- Varela, Francisco, Lachaux, Jean-Philippe, Rodriguez, Eugenio, Martinerie, Jacques, 2001. The brainweb: phase synchronization and large-scale integration. Nat. Rev. Neurosci. 2 (4), 229–239.
- Venkatesh, Manasij, Jaja, Joseph, Pessoa, Luiz, 2019. Brain dynamics and temporal trajectories during task and naturalistic processing. NeuroImage 186, 410–423.
- Walentek, Dawid, Broere, Joris, Cinelli, Matteo, Dekker, Mark M., Haslbeck, Jonas M.B., 2021. Success of economic sanctions threats: coercion, information and commitment. Int. Interact. 1–32.
- Zhang, Xiaofan, Yi, Hu, Bai, Wenwen, Tian, Xin, 2015. Dynamic trajectory of multiple single-unit activity during working memory task in rats. Front. Comput. Neurosci. 9.