**On the methods for reactivation and replay analysis**

David Tingley[1,*], Adrien Peyrache[2,*]

1 Neuroscience Institute, New York University, NYC, NY, USA

[2] Montreal Neurological Institute, McGill University, Montreal, QC, Canada

[*] To whom correspondence should be addressed: davidtingley2@gmail.con; adrien.peyrache@mcgill.ca

Main article published in Philosophical Transactions B.

# SUPPLEMENTARY INFORMATION

## Reactivation

The reactivation methods can be fully described analytically. The starting point is to use a matrix A of binned spike trains:

$$A = \begin{matrix} a_1(1) & a_1(t) & a_1(B) \\ a_i(1) & a_i(t) & a_i(B) \\ a_N(1) & a_N(t) & a_N(B) \end{matrix}$$

Where $a_i(t)$ denotes the number of spikes emitted by neuron $i$ in time bin $t$ for $N$ neurons recorded during $B$ time bins. The population vector of activity will be noted in bold, e.g. $\boldsymbol{a(t)}$ for the population vector of spike counts at time bin $t$.

As firing rate (and spike count variance) are log-normally distributed, it is more convenient and meaningful to work with correlation coefficients. Covariance would give too much weight to high firing neurons. For this reason, the activity is z-scored during each epoch:

$$z_i(t) = \frac{a_i(t) - \mu_a}{\sigma_a}$$

Where $\mu_a$ and $\sigma_a$ denote average and standard deviation of spike count for neuron $i$. Calling $Z$ the z-scored $A$ matrix during a given epoch, the cell-pair correlation matrix $C_{epoch}$ (for a given epoch) is simply:

$$C_{epoch} = \frac{1}{B} Z^T Z.$$

## Explained Variance.

The EV measure can be directly obtained from correlation measures. First, the cell-pair coherence $r_{1,2}$ for epochs 1 and 2 is:

$$r_{1,2} = \frac{1}{N_{pairs}} \rho_1^T \rho_2$$

Where $\rho_{1,2}$ denote the z-scored vector of all cell-pair correlation coefficients during epochs 1 and 2, and $N_{pairs}$ the total number of cell pairs (i.e. *N(N-1)/2*). Then EV is defined as the partial correlation of cell-pair correlations:

$$EV = \left( \frac{r_{W,POST} - r_{W,PRE} r_{PRE,POST}}{\sqrt{\left(1 - r^2{}_{W,PRE}\right)\left(1 - r^2{}_{PRE,POST}\right)}} \right)^2$$

REV is obtained by switching epochs *PRE* and *POST* in the above equation.

## Reactivation strength

To resolve reactivation measure in time, cell-pair coherence can be decomposed on a bin-by-bin basis by projecting instantaneous firing during sleep onto the correlation matrix of the wake epoch $C_{wake}$, thus obtaining the cell-pair coherogram $\chi$

$$\chi_{sleep}(t) = \boldsymbol{z}_{\boldsymbol{sleep}}^{\boldsymbol{T}}(t) C_{wake} \boldsymbol{z}_{\boldsymbol{sleep}}(\boldsymbol{t})$$

Where $z_{sleep}(t)$ is a row of the z-scored matrix of binned spike trains. It is to be noted that, while cell-pair coherence and coherogram are related, they are not exactly similar. The time average of the coherogram is equal to the variance during the *sleep* epoch explained by the correlation of the *wake* epoch. Note that we refer to *sleep* and *wake* epochs only for the sake of clarity and that, obviously, this method applies to any pairs of recording epochs.

PCA can be used at this point to decompose the cell-pair correlation matrix into its eigenvectors $v_i$, each associated with an eigenvalue $\lambda_i$ such that $Cv_i = \lambda_i v_i$. From the eigenvectors, the cell-pair correlation matrix can be decomposed in its projectors **P**:

$$C = \sum_{i=1}^{N} \lambda_i P^i$$

Where projectors are defined by the outer product of the correlation matrix eigenvectors: $P^i = v_i v_i^T$.

The strategy now consists at keeping only a subset of projectors (corresponding to vectors explaining more variance than expected from random data). Furthermore, to exclude all the terms depending on firing of individual neurons in the reactivation measure, the diagonal of the projectors is set at 0: $P_{k,k}=0$

By replacing the cell-pair correlation matrix by one of the *i*th projector in the cell-pair coherogram, one obtains the reactivation strength for the *i*th component:

$$R^i(t) = \boldsymbol{z}_{\boldsymbol{sleep}}^{\boldsymbol{T}} P^i \boldsymbol{z}_{\boldsymbol{sleep}}(\boldsymbol{t})$$

Which by expanding all the terms is equal to:

$$R^i(t) = \sum_{k \neq l} v_k^i v_l^i z_k(t) z_l(t)$$

## Bayesian decoding

For a detailed derivation of this method, see (Zhang et al., 1998). Briefly, the occupancy map is calculated as:

$$P(x) = \frac{N(x)}{\sum_x N(x)}$$

Where $N(x)$ is the number of times the animal occupied position $x$. The average firing rates for each neuron are then calculated as:

$$f(x) = \frac{S(x)}{N(x)\Delta t}$$

Where $S(x)$ is the total number of spike counts as location $x$ and $N(x)\Delta t$ is the total time spent in position $x$. These equations can then be used to calculate the conditional probability of being in position $x$, given the numbers of spikes $n$:

$$P(x|n) = P(x)\left(\prod_{i=1}^{N} f_i(x)^{n_i}\right)\exp\left(-\tau\sum_{i=i}^{N} f_i(x)\right)$$

Let's start with a simple case, where we have two place cells with non-overlapping place fields. One neuron's peak in-field firing rate is 50 Hz, while the other is 25 Hz. Let's assume the reliability (i.e. spatial information) across trials is similar. Now, during a candidate 'replay' event we are presented with a single time bin where both neurons fire one action potential. How will these neurons be weighted? What is the estimated 'position' during such an event? Well, it depends on the temporal window size

used when binning the candidate event. One spike, in a 20-millisecond bin corresponds to a firing rate of 50 Hz, and the Bayesian method will decode to the position of cell 1. Whereas, if that same event is examined with 50-millisecond bins, the firing rates will be estimated to be 20 Hz and the Bayesian method will decode to the position of cell 2. While these neurons carry equivalent spatial information, the estimated position is based solely on the choice of temporal bin size. Thus, average firing rates combined with bin size determines the outcome for this method of replay decoding, while the trial-by-trial spatial reliability of cells is not a factor.

## General replay caveats

The first is a methodological issue that exists for all analyses examining temporal ordering of spiking activity. As a neuron fires multiple action potentials, the shape and amplitude of the action potential can change dramatically. Using extracellular recording methods, such waveforms are the primary source of information when isolating 'single units.' Therefore, it is possible that such bursts of changing waveforms are accidentally misidentified as two or multiple neurons firing in close temporal order. As only a small percentage of CCGs show the temporal asymmetry effect (~5%; (Skaggs and McNaughton, 1996)), even very rare mistakes in clustering waveforms can impact the results greatly (Quirk and Wilson, 1999). One way to address this criticism is to only examine cell pairs recorded across separate tetrodes/shanks, thus limiting the impact of waveform classification errors.

An additional issue with most methods for replay analysis stems from the distribution of baseline firing rates across neurons. As cell pair correlations co-vary with baseline firing rates (De La Rocha et al., 2007), and only a small fraction of neurons display extremely high firing rates (Buzsáki and Mizuseki, 2014), these analyses can be particularly susceptible to bias from a small fraction of the population. Z-scoring or generating bootstrapped firing rate matched distributions for comparison are reasonable approaches to overcome this.

## Template free methods

### *seqNMF*

A recent success in template-free examination of neural sequences was the application of nonnegative matrix factorization (Lawton and Sylvestre, 1971; Paatero and Tapper, 1994) to neural population data (Mackevicius et al., 2019). The matrix $X_{nt}$, where $n$ is the neuron and $t$ is the time bin, can be approximated by:

$$X_{nt} = \sum_{k=1}^{K} \sum_{l=0}^{L-1} W_{nkl} H_{k(t-l)} \equiv (W * H)_{nt}$$

where $W$ is a N x K x L tensor containing 'exemplar' patterns. N is the number of neurons, K is the number of patterns (defined *a priori*), and L is the length of each pattern (defined *a priori*). $H$ is a K x T matrix of temporal loadings for each pattern. By adding a penalty term $R$ that accounts for both correlations between temporal loadings and correlations between different patterns in $W$, the authors were able to achieve an algorithm that reliably converges on a similar result across random initializations while also reducing the correlations between patterns in $W$. This approach, termed 'seqNMF', is capable of detecting reliable sequences without the *a priori* identification of a template or particular spike sequence.

### Dataset details

39 recordings from 5 Long Evans rats were used (3 male, 2 female). All recordings consisted of three continuous sessions (PRE/behavior/POST), each lasting 1-5 hours. Place field heuristics were defined as follows, 1) Minimum peak firing rate of 2 Hz, 2) Minimum field width of 8 cm, 3) Maximum field width of 120 cm, 4) Minimum of 1.5 spikes per trial on average, and 5) at least 10 trials with consistent behavior.

Candidate replay events were initially detected using the CA1 LFP to identify sharp-wave ripples. Events were then screened and only analyzed if at least 5 place cells fired at least one spike each during the event.

## Bibliography


Buzsáki, G., and Mizuseki, K. (2014). The log-dynamic brain: how skewed distributions affect network operations. Nat. Rev. Neurosci. 1–15.

De La Rocha, J., Doiron, B., Shea-Brown, E., Josić, K., and Reyes, A. (2007). Correlation between neural spike trains increases with firing rate. Nature *448*, 802–806.

Lawton, W.H., and Sylvestre, E.A. (1971). Self modeling curve resolution. Technometrics *13*, 617–633.

Mackevicius, E.L., Bahle, A.H., Williams, A.H., Gu, S., Denisenko, N.I., Goldman, M.S., and Fee, M.S. (2019). Unsupervised discovery of temporal sequences in high-dimensional datasets , with applications to neuroscience. Elife *8*, 1–42.

Paatero, P., and Tapper, U. (1994). Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. Environmetrics *5*, 111–126.

Quirk, M.C., and Wilson, M.A. (1999). Interaction between spike waveform classification and temporal sequence detection. J. Neurosci. Methods *94*, 41–52.

Skaggs, W.E., and McNaughton, B.L. (1996). Replay of Neuronal Firing Sequences in Rat Hippocampus During Sleep Following Spatial Experience. Science *271*, 1870–1873.

Zhang, K., Ginzburg, I., McNaughton, B.L., and Sejnowski, T.J. (1998). Interpreting Neuronal Population Activity by Reconstruction: Unified Framework With Application to Hippocampal Place Cells. J.

Neurophysiol. *79*, 1017–1044.