Differential population coding of natural movies through spike counts and temporal sequences

Boris Sotomayor-Gómez[1,2,*], **Francesco P. Battaglia**[3], **and Martin Vinck**[1,2]

[1]Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with Max Planck Society, Frankfurt, Germany
[2]Donders Centre for Neuroscience, Department of Neuroinformatics, Radboud University Nijmegen, Nijmegen, Netherlands
[3]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen
[*]boris.sotomayor@esi-frankfurt.de

ABSTRACT

The traditional view on coding in the cortex is that populations of neurons primarily convey stimulus information through the spike count. However, given the speed of sensory processing, it has been hypothesized that sensory encoding may rely on the spike-timing relationships among neurons. Here, we use a recently developed method based on Optimal Transport Theory called SpikeShip to study the encoding of natural movies by high-dimensional ensembles of neurons in visual cortex. SpikeShip is a generic measure of dissimilarity between spike train patterns based on the relative spike-timing relations among all neurons and with computational complexity similar to the spike count. We compared spike-count and spike-timing codes in up to $N > 8000$ neurons from six visual areas during natural video presentations. Using SpikeShip, we show that temporal spiking sequences convey substantially more information about natural movies than population spike-count vectors when the neural population size is larger than ≈ 200 neurons. Remarkably, encoding through temporal sequences did not show representational drift both within and between blocks. By contrast, population firing rates showed better coding performance when there were few active neurons. Furthermore, the population firing rate showed memory across frames and formed a continuous trajectory across time. In contrast to temporal spiking sequences, population firing rates exhibited substantial drift across repetitions and between blocks. These findings suggest that spike counts and temporal sequences constitute two different coding schemes with distinct information about natural movies.

Introduction

Information in the nervous system is encoded by patterns of spikes fired by large populations of neurons. These spiking patterns can contain information both in the timing of spikes and in the spike count[1–6]. Thus, a major question is what the distinct contributions to population coding are of the spike count (or firing rate code) and the timing of spikes.

While it is well established that neural populations encode information with the spike count, there is some evidence that the timing of spikes can carry additional information[7–11]. As the brain does not have access to absolute time, the time of a spike must necessarily be defined in relation to some other event. One possible coding scenario is that the timing of spikes relative to a stimulus onset encodes information[2,12–17]. For natural sensory inputs, however, there are often no clear stimulus onsets that can be used as a temporal reference. Other studies have examined the timing of spikes relative to ongoing population oscillations[7–11]. Although oscillations may in some cases be regular enough to allow for phase coding, neural activity during active vision may often consist of broadband activity rather than frequency band-limited activity, making it hard to robustly define the phase of each spike[18]. In general, phase coding expresses the relative timing of spike to a (filtered) population average of all other spikes (i.e. one-to-all). However, the relative timing of spikes from different neurons carry a signal that can be detected by synapses[19,20], neurons[21], and neural networks[22]. It should be therefore possible to define a temporal code based on the pattern of relative spike-timing relations among all neurons (i.e. all-to-all combinations), without frequency decomposition or averaging.

Here, we utilize a recently developed method called *SpikeShip*[23] that extracts information from neural populations purely based on the relative timing of spikes alone, with invariance to the overall spike count, and with the same computational complexity as the spike count. The method is based on solving a general optimal transport problem of determining the minimum cost of shifting 'spike mass' such that all relative timing relations become identical. We use this technique to study the encoding of natural movies in a large population (>8000) of neurons in the visual cortex of awake mice, and contrast the information in relative timing of spikes to the firing rate (i.e. spike count). Based on the precise temporal information of spike trains, we show that multi-neuron temporal patterns convey substantially more, and different, information about natural movies than population firing rates. Furthermore, these multi-neuron temporal patterns show high stability, i.e. no representational drift, across presentations. By contrast, firing rate codes show

substantial representational drift but also exhibit memory across frames within a video. Finally, we show that the performance advantage of temporal information increases as a function of population size and the number of active neurons.

Results

Comparing temporal encoding of natural movies to rate coding

Our main question is how natural movies are encoded by the patterns of spiking in populations of neurons in visual cortex. Specifically, we wished to compare two coding schemes: 1) The encoding of natural movies via temporal spiking sequences; and 2) the encoding of natural movies with population vectors of firing rates, i.e. spike counts / second. To this end, we analyzed Neuropixel recordings[24] from 32 mice across six visual areas[25] (for details, see Methods). In total, the dataset yielded population spiking patterns that consisted of $N = 8,301$ neurons, which were pooled across multiple sessions in which the same stimuli were presented. Note that this restricts the analysis to temporal spiking patterns that are time-locked to the natural movie, however we obtained similar results when analyzing single sessions (see below).

For all 32 mice, we analyzed neural recordings from two natural movies: (1) *Natural Movie One* (NM1), a 30-s natural movie with 10 consecutive repetitions, and (2) *Natural Movie Three* (NM3), a 120-s natural movie presented in two blocks with 5 repetitions each (see Fig. 1A). For our first set of experiments, we split the NM1 video into sub-videos of one-second length, similar to previous work[26,27] (See Fig. 1B). Thus, each epoch consisted of a dynamical stimulus of 30 frames with 10 repetitions per epoch (total $M = 30 \times 10 = 300$ epochs) (Fig. 1C). We then analyzed how distinct the neural responses were across the different sub-videos, and how reliable the encoding of the movies was across stimulus repetitions.

For each pair of epochs, we computed two dissimilarity measures that were based on spike counts: (1) the (raw) Euclidean distance between firing rates (FR); and (2) the Euclidean distance between z-scored firing rates. The Z-scoring was performed across epochs, for each neuron separately. The Euclidean distance between raw firing rates tends to emphasize neurons with higher firing rates, whereas the distance between z-scored firing rates weighs each neuron similarly. In addition to these two dissimilarity measures, we also computed a recently proposed dissimilarity measure based on the temporal structure of spike trains, namely SpikeShip[23].

SpikeShip measures the dissimilarity between spike trains of different epochs using the mathematical framework of optimal transport. It considers each spike train as a collection of "masses" (i.e. the spikes). All spikes from each active neuron, together, contribute a unit mass, which ensures the rate invariance of the method. SpikeShip solves the optimal transport problem of finding the minimum cost of shifting the (unit) mass of each neuron's spike train in epoch k, such that the cross-correlations (i.e., sequential firing) between all of the neurons become identical to those in another epoch m (See Methods). Thus, SpikeShip provides a generic dissimilarity measure of how similar all the relative spike-timing relations are between two epochs. Importantly, SpikeShip can be computed with computational cost on the order of number of neurons (times the avg. number of spikes), and can therefore be efficiently computed for high-dimensional neural patterns.

For the three measures (firing rates, z-scored firing rates, SpikeShip), we obtained a dissimilarity / distance matrix for all 300×300 pairs of epochs (Fig. 2A). In addition, we visualized the dissimilarity matrix in a low-dimensional embedding by using the two-dimensional t-SNE algorithm (Fig. 2B). Finally, we computed the differences in clustering performance by applying various clustering algorithms to the dissimilarity matrices and comparing clustering performance using the Adjusted Rand Index (ARI) (Fig. 2C). In addition, we calculated the discriminability index for different sub-videos, which compares the distances within clusters to those between clusters (Fig. 2D).

Compared to firing rates, we found that SpikeShip yielded much lower dissimilarities between repetitions of the same sub-frame as compared to pairs of different sub-frames (Fig. 2B), which conducive to tighter clustering of different sub-frames. Correspondingly, the t-SNE embeddings contained a separate cluster for each sub-video in case of SpikeShip, whereas there was substantial overlap between different sub-videos for the firing rates (Fig. 2C). Consequently, SpikeShip showed almost perfect classification of sub-videos, with much higher performance than in case of firing rates. Likewise, discriminability scores were substantially higher for SpikeShip than for firing rates. Thus, temporal spiking sequences, as quantified with SpikeShip, contained substantially more information about natural-movie content than firing rates.

We wondered whether the performance benefit of temporal sequences depended on the number of neurons that were included in the analysis. To analyze this, we took smaller subsets of neurons and repeated the clustering analysis for different subset sizes. Interestingly, we found that firing rates outperformed SpikeShip for smaller subsets of neurons (< 100), and that the performance benefit of temporal encoding only emerged for subsets of neurons greater than about 100 neurons.

We also performed these analyses for single sessions (Supplementary Fig. S1A-B), which showed comparable results. Specifically, we found that SpikeShip outperformed the firing-rate code, and that this benefit emerged for population sizes of 160 neurons and beyond (See Supplementary Fig. S1C). Given that ARI scores were comparable between single sessions and pooled data across mice, we conclude that the information in temporal spiking sequences emerge

due to recurrent interactions between neurons driven by time-varying natural input.

We also compared coding performances between different visual areas. We found that the result of more precise encoding via temporal sequences held true across visual areas, with the best clustering performance in the primary visual cortex (Supplementary Fig. S2 and S4 for NM1 and NM3. respectively).

Altogether, natural-movie content is substantially more precisely encoded by temporal sequences as compared to firing rates. These sequences can be efficiently detected by SpikeShip with a computational complexity equal to the spike count. The benefits of temporal coding appear for larger population sizes, presumably because this yields a sufficient number of co-active pairs of neurons emitting spikes in the same epoch.

Firing rates reflect ordering of sub-videos

In Figure 2, we had computed the dissimilarity matrix between all pairs of epochs. We wondered whether epochs that were closer in time within the movie would have similar, i.e. correlated representations. Indeed, the t-SNE representations for firing rates suggest that epochs that are close in time lie closer to each in the low-dimensional embedding space. By contrast, it appears that a similar effect was not observed for SpikeShip, with an apparent random ordering of clusters w.r.t. time.

To quantify this, we computed the Euclidean distances (in t-SNE space) between consecutive epochs and compared these with distances to other epochs. For firing rates, the distance between two consecutive epochs is about 3-fold lower than the distances between all epochs (i.e. pairwise comparison) (Fig. 3A-B). For SpikeShip we did not find a difference between consecutive epoch pairs and random epoch pairs. This finding suggests that firing rate representations tend to be similar for consecutive epochs, indicating correlations on longer time-scales. Thus, even though firing rates are less precise, they carry additional information about elapsed time within the movie.

Together, these findings suggest distinct information in spike counts and in temporal sequence information. To further investigate the extent to which they carry distinct information, we computed Spearman correlations between the epoch-to-epoch dissimilarity matrices of firing rates and SpikeShip (as shown in Figure 2). These correlations can be understood as "representational similarity" (Fig. 3C). We ignored (diagonal) entries of the matrix containing dissimilarity between repetitions of the same sub-video. For these correlations, we found indeed that SpikeShip and firing rates showed only weak Spearman correlations of around 0.2, indeed demonstrating largely distinct information in spike counts and temporal sequences.

The latter result held true across different visual areas (Supplementary Fig. S3), with Spearman correlations between 0.02 and 0.292. We also performed representational similarity analyses by comparing the dissimilarity matrices between different visual areas. Interestingly, we found that representational similarity between areas was substantially stronger for firing rates than for SpikeShip (Supplementary Fig. S5). In other words, different visual areas provided more independent information through temporal sequences than through the spike count. A possible explanation for this finding is that firing rates, for all visual areas, show a clear mapping onto the time elapsed within the movie (Fig. S2).

Representational drift of firing rate and temporal code

Next we investigated to what extent the coding of natural movies via spike counts or temporal sequences was stable across time. Several studies have reported that neural representations of stimuli, tasks or contexts can change across time[26,28–32].

To examine this, we computed the correlation of neural activity across different repetitions of the same sub-video. The low-dimensional embedding colored by the repetition order already shows ample variability across repetitions for FR and FRz, and little variability for SpikeShip (Fig. 4A). In particular, firing rates appear to map out a trajectory in the low-dimensional embedding space during each repetition, and this trajectory gradually drifts away across repetitions. To further quantify this, we computed for each repetition a vector of Eucldean distances across the sub-videos, and then correlated these vectors between repetitions (Fig. 2A). This yields a matrix of correlations ordered by repetition number (See Fig. 4B). Firing rate vectors showed a weaker correlation across frames as compared to the SpikeShip measure, indicating that the temporal code has higher stability (i.e. less drift). We furthermore analyzed the representational drift by examining how the Pearson correlation changes as a function of time between two repetitions. The population vector of firing rates decorrelated across time, as shown in Fig. 4C. Yet, for SpikeShip, the Pearson correlation stayed close to 1 for all time-delays.

Together, these analyses indicate that the temporal structure in the spike trains (quantified with SpikeShip) uniquely and reliably encodes the different dynamic stimuli, with essentially no drift in the population code. The spike count on the other hand show systematic drift over longer time-scales within a session.

Analysis of longer natural movies

To generalize these results, we analyzed another dataset (NM3) in which longer movies were presented (5 repetitions), repeated in two separate blocks with 30 minutes between blocks. Similar to the analyses presented above, we found a clear clustering of the different sub-videos for SpikeShip, with no apparent relation of the clusters to time (Figure 5A-C). By contrast, for firing rates we did find not a clear clustering of the sub-videos (Figure 5A-C). Rather, firing rates dissimilarities showed a clear ordering with time, with nearby sub-videos showing more similar representations.

Furthermore, we found a substantial drift in the population firing rates between blocks 1 and 2, which was entirely absent for SpikeShip. This result held true across visual areas (Fig. S4). To quantify drift, we again computed Pearson correlations between the vectors of Euclidean distances. Firing rate representations showed drift both within the blocks (across repetitions) and between blocks (Fig. 5D). SpikeShip, however, showed Pearson correlations close to 1 indicating stable encoding of natural movies based on temporal sequences (Fig. 5D).

Dependence of encoding on number of active neurons

Finally, we investigated to what extent the performance of SpikeShip depended on the overall firing activity in the population.

To this end, we quantified for each sub-video of 3 frames the number of active neurons. There was a strong correlation between the number of active neurons and the curvature across frames[33,34] which measures the variability of movie's pixel across frames (*Global Curvature* (GC)) (Fig. 6A-B). We wondered how the clustering performance of firing rates and SpikeShip depended on the number of active neurons for these shorter sub-videos. We selected epochs that contained either the highest or the lowest number of active neurons for a given pair of epochs k and m ($\mathscr{A}_{k,m}$). When the number of active neurons is high, SpikeShip yields strong clustering even for short sub-videos with an ARI score of 1 (Fig. 6C). By contrast, when the number of active neurons is low, SpikeShip does not yield clear clustering, whereas firing rates yield ARI scores of 1 (Fig. 6D). Hence, temporal coding, as quantified with SpikeShip, is the most reliable coding scheme when the number of active neurons is high, whereas the spike count is more reliable when the number of active neurons is low.

To further quantify this, we correlated the mean dissimilarity between different presentations of the same sub-video as a function of the global curvature, which is a proxy of the number of active neurons. We found a negative correlation for SpikeShip, i.e. representations tended to be more similar when then global curvature is high. By contrast, for firing rates we found a positive correlation, i.e. representations tended to be more dissimilar when the global curvature is high.

Discussion

The classic view on neural computation is that the firing rate, i.e. the number of spikes per second, is the unit of information and computation in the nervous system. Firing rates are typically estimated through spike counts, which are known to show a high degree of (co-)variability across trials, a potential limiting factor to their coding capacity[6,35–37]. Another theoretical perspective is that neural populations encode information via temporal sequences, defined by the timing of spikes relative to other spikes[3,4,38]. Ideally, the quantification of temporal sequences relies on the pattern of relative spike-timing relations among all neurons (i.e. all-to-all comparisons). However, it has remained challenging to quantify such temporal sequences in high-dimensional neural ensembles. Here, we use a recently developed method based on Optimal Transport Theory called SpikeShip to study the encoding of natural movies by high-dimensional ensembles of neurons in visual cortex. SpikeShip is a generic measure of dissimilarity between spike train patterns based on the relative spike-timing relations among all neurons. Remarkably, its computational complexity is similar to the spike count, allowing for its application to large ensembles of neurons.

We compared spike-count and spike-timing codes in up to $N > 8000$ neurons from six visual areas during natural video presentations. We analyzed temporal sequences either by pooling data across sessions, or by analyzing individual sessions, which yielded similar results. Using SpikeShip, we show that temporal spiking sequences convey substantially more information about natural movies than population spike-count vectors when the neural population size is larger than ≈ 200 neurons. Remarkably, encoding through temporal sequences did not show representational drift both within and between blocks. By contrast, population firing rates showed better coding performance when there were few active neurons. Furthermore, the population firing rate showed memory across frames and formed a continuous trajectory across time. In contrast to temporal spiking sequences, population firing rates exhibited substantial drift across repetitions and between blocks. These results suggest that spike counts and temporal sequences constitute two different coding schemes with distinct information about natural movies.

Our findings on the spike count fit well with several recent studies showing so-called representational drift as a function of time[26,28–31,39–42]. Yet, our findings suggest that these previous conclusions may be restricted to the spike count, as we found the representation via temporal sequences to be entirely stable. This may suggest that the spike count encodes additional information, causing correlations across time. It is possible that the decorrelation of spike counts across time reflects the encoding of other variables, e.g. related to behavioral state[43], or stimulus-specific adaptation or facilitation. However, it appears that the information between temporal sequences provides stable encoding, which perhaps solves the conundrum of how sensory information is maintained in the presence of representational drift within a session. We found that spike counts showed correlations also on a short time-scale, in the sense that spike-count representations of nearby movies tended to be similar, whereas this was again not the case for temporal sequences. This suggests that spike counts show autocorrelations on longer time scales, which may be driven e.g. by within-neuron firing adaptation. As a consequence, firing rates contained information about the passage

of time through the movie, while this was not the case for temporal sequences. In sum, it appears that the encoding of natural movie in a recurrent network of neurons is performed through temporal sequences. Spike counts, on the other hand, show less robust encoding but convey additional information on longer time-scales, either forming a continuous trajectory as a function of time during experience or showing representational drift on longer time scales.

We furthermore observed that the advantage of temporal sequences, as quantified with SpikeShip, only become apparent for sufficiently large population sizes of > 200 neurons. Yet, even when analyzing a large number of neurons (i.e. > 8000), we found that during periods when many neurons were inactive, the spike count was more robust than temporal sequences. A likely explanation of this finding is that a sufficient number of active neurons are necessary to form temporal sequences. That is, a neuron can only convey temporal sequence information if it is active. Likewise, we can only compare the temporal structure in two epochs based on neurons that are active in both epochs. When firing rates are low, information may be conveyed by a few active neurons, however between two different epochs there may be only little overlap in which neurons are active.

Together, our findings offer a new view on neural coding that is based on the temporal structure of spike trains and which can be captured by a method based on optimal transport theory that has the same computational complexity as the spike count. The ensuing coding strategy is purely spike-based and does not presuppose a notion of firing rate. Together, our findings suggest that information about natural visual input is encoded robustly and stably by high-dimensional temporal spiking sequences.

Methods

Natural movie's processing from Allen Brain Institute datasets

We used the public available datasets of Allen Brain Institute through AllenSDK (For more details, see http://help.brain-map.org/display/observatory/Documentation). Neuropixels silicon probes[24] were used to record neurons with precise spatial and temporal resolution[25]. We selected the cells of 32 mice during natural scenes presentations. The cells were selected considering a signal-noise ratio (SNR) such that $SNR > 0$. The neural activity from a total of $N = 8,301$ cells was selected from the Primary visual area (VISp), Lateral visual area (VISl), Anterolateral visual area (VISal), Posteromedial visual area (VISpm), Rostrolateral visual area (VISrl), and Anteromedial visual area (VISam).

Finally, for the analyses per brain area we present in Fig. S2, we down-sampled the set of neurons randomly in order to compare the performance of both decoding schemes. Particularly, we used $N_{ds} = 879$ for every brain area.

Computation of dissimilarity matrices of firing rate vectors

For a population of N neurons, we computed the firing rate vectors ($\overrightarrow{FR}$) for each epoch of our analyses as the count of spikes per neuron divided by a window length T. We denote FRz to the normalized firing rate vectors (FRz) across epochs (z-score). Finally, we computed the Euclidean distance between both normalized vectors is $d_{Euc} = \sqrt{\sum_i^N (\overrightarrow{FR_{m,i}} - \overrightarrow{FR_{k,i}})^2}$.

Computation of dissimilarity matrices via SpikeShip

SpikeShip is a dissimilarity measure based on optimal transport theory to extract temporal multi-neuron spike-train patterns. SpikeShip solves the following transport problem:

Suppose a population of neurons in two epochs k and m. For each neuron j in epoch k for which the number of spikes $n_{k,j} > 0$, we define the point process with unit energy

$$\rho_{k,j}(t) = \frac{1}{n_{k,j}} \sum_{u=1}^{n_{k,j}} \delta(t - t_{k,j,u}) \,. \tag{1}$$

This defines for each pair of neurons (i, j) in epoch k the cross-correlation function

$$s_{i,j,k}(\tau) = \sum_{t=0}^{T} \rho_{k,i}(t)\rho_{k,j}(t+\tau) \,, \tag{2}$$

Consider two epochs (k, m). We wish to find for each neuron (in epoch k) a transport of mass from t to t', $[\mathbf{M}]_{j,t,t'}$, such that $s_{i,j,k}(\tau) = s_{i,j,m}(\tau)$ for all (i, j, τ). The mass here consists of the spikes, which have a sum of 1. The objective is then to find a matrix of flows $\mathbf{M}$ that minimizes the total mover cost, i.e.

$$\underset{\mathbf{M}}{\arg\min} \sum_{j,t,t'} \mathbf{M}_{k,t,t'} d(t,t') \tag{3}$$

where $d(t,t') = |t - t'|$ (For more details, see[23]).

Measure of representational drift

From 2D t-SNE embeddings

For each dissimilarity matrix of NM1 and NM3 presentations, we computed the Euclidean distances across consecutive epochs (i.e., trajectory). Then, we measure the representational drift as the Pearson rank correlation between pairs of trajectories, as done in[26]. Such correlations are visualized across the time difference between repetitions of the same movie.

From dissimilarity matrices

We extracted the upper triangle based on the pairwise distances of each scaled dissimilarity matrix (See Fig. S6A). For a dissimilarity matrix raw_diss_m based on metric m with $m \in \{FR, FRz, SS\}$, we compute:

$$diss_m = \frac{raw_diss_m}{max\{raw_diss_m\}}. \quad (4)$$

Finally, we computed the Euclidean distance between dissimilarity matrices from each pair of repetitions, as shown in (See Fig. S6).

Global curvature definition

In order to compute the changes in the pixel domain of the movie, we based on previous studies about curvature computation of pixels in movies[33,34]. We denoted x_t to one pixel of the movie from the frame t. First, we computed the difference of such pixel at time t with the previous one $(t-1)$ as

$$v_t = x_t - x_{t-1}, \quad (5)$$

as shown in Fig. 6A).

We normalized such vector based on its norm as:

$$\hat{v} = \frac{v_t}{||v_t||}. \quad (6)$$

Then, the local curvature is defined as

$$c_i = arccos(\hat{v}_t \cdot \hat{v}_{t+1}). \quad (7)$$

This measure is always a positive number. It reaches its lowest value of zero only when the frames are in a straight line without any bending in the high-dimensional space. Thus, the global curvature correspond to the average of all local curvatures over time, computed as

$$GC = \frac{1}{T}\sum_{i}^{T} c_i \,, \quad (8)$$

in degrees. See Henaff et al. studies[33,34] for further details.

Acknowledgements

This project was funded by the BMBF Computational Life Sciences Grant 031L0167 to M.V.

References

1. Maass, W. & Bishop, C. M. *Pulsed neural networks* (MIT press, 2001).
2. Rieke, F. & Warland, D. *Spikes: exploring the neural code* (MIT press, 1999).
3. Thorpe, S., Delorme, A. & Van Rullen, R. Spike-based strategies for rapid processing. *Neural networks* **14**, 715–725 (2001).
4. Thorpe, S., Fize, D. & Marlot, C. Speed of processing in the human visual system. *nature* **381**, 520–522 (1996).
5. Resulaj, A., Ruediger, S., Olsen, S. R. & Scanziani, M. First spikes in visual cortex enable perceptual discrimination. *Elife* **7**, e34044 (2018).
6. Brette, R. Philosophy of the spike: rate-based vs. spike-based theories of the brain. *Front. systems neuroscience* 151 (2015).
7. Vinck, M., van Wingerden, M., Womelsdorf, T., Fries, P. & Pennartz, C. M. The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization. *Neuroimage* **51**, 112–122 (2010).
8. O'Keefe, J. & Recce, M. L. Phase relationship between hippocampal place units and the eeg theta rhythm. *Hippocampus* **3**, 317–330 (1993).
9. Kayser, C., Montemurro, M. A., Logothetis, N. K. & Panzeri, S. Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron* **61**, 597–608 (2009).
10. Huxter, J. R., Senior, T. J., Allen, K. & Csicsvari, J. Theta phase–specific codes for two-dimensional position, trajectory and heading in the hippocampus. *Nat. neuroscience* **11**, 587–594 (2008).
11. Yiling, Y. *et al.* Robust encoding of natural stimuli by neuronal response sequences in monkey visual cortex. *Nat. Commun.* **14**, 3021 (2023).
12. Victor, J. D. & Purpura, K. P. Nature and precision of temporal coding in visual cortex: a metric-space analysis. *J. neurophysiology* **76**, 1310–1326 (1996).
13. Victor, J. D. & Purpura, K. P. Metric-space analysis of spike trains: theory, algorithms and application. *Network: computation neural systems* **8**, 127–164 (1997).
14. Kreuz, T., Haas, J. S., Morelli, A., Abarbanel, H. D. & Politi, A. Measuring spike train synchrony. *J. neuroscience methods* **165**, 151–161 (2007).
15. Aronov, D. Fast algorithm for the metric-space analysis of simultaneous responses of multiple single neurons. *J. neuroscience methods* **124**, 175–179 (2003).
16. Kreuz, T., Chicharro, D., Houghton, C., Andrzejak, R. G. & Mormann, F. Monitoring spike train synchrony. *J. neurophysiology* **109**, 1457–1472 (2013).
17. Satuvuori, E. *et al.* Measures of spike train synchrony for data with multiple time scales. *J. neuroscience methods* **287**, 25–38 (2017).
18. Ray, S. & Maunsell, J. H. Do gamma oscillations play a role in cerebral cortex? *Trends cognitive sciences* **19**, 78–85 (2015).
19. Brzosko, Z., Mierau, S. B. & Paulsen, O. Neuromodulation of spike-timing-dependent plasticity: past, present, and future. *Neuron* **103**, 563–581 (2019).
20. Andrade-Talavera, Y., Fisahn, A. & Rodríguez-Moreno, A. Timing to be precise? an overview of spike timing-dependent plasticity, brain rhythmicity, and glial cells

interplay within neuronal circuits. *Mol. Psychiatry* 1–12 (2023).

21. Branco, T., Clark, B. A. & Häusser, M. Dendritic discrimination of temporal input sequences in cortical neurons. *Science* **329**, 1671–1675 (2010).

22. Prezioso, M. *et al.* Spike-timing-dependent plasticity learning of coincidence detection with passively integrated memristive circuits. *Nat. communications* **9**, 5311 (2018).

23. Sotomayor-Gómez, B., Battaglia, F. P. & Vinck, M. A method for fast, unsupervised discovery of high-dimensional neural spiking patterns. *bioRxiv* (2021).

24. Jun, J. J. *et al.* Fully integrated silicon probes for high-density recording of neural activity. *Nature* **551**, 232–236 (2017).

25. Siegle, J. H. *et al.* Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature* **592**, 86–92 (2021).

26. Deitch, D., Rubin, A. & Ziv, Y. Representational drift in the mouse visual cortex. *Curr. Biol.* **31**, 4327–4339 (2021).

27. Sadeh, S. & Clopath, C. Contribution of behavioural variability to representational drift. *bioRxiv* (2022).

28. Ziv, Y. *et al.* Long-term dynamics of ca1 hippocampal place codes. *Nat. neuroscience* **16**, 264–266 (2013).

29. Lütcke, H., Margolis, D. J. & Helmchen, F. Steady or changing? long-term monitoring of neuronal population activity. *Trends neurosciences* **36**, 375–384 (2013).

30. Driscoll, L. N., Pettit, N. L., Minderer, M., Chettih, S. N. & Harvey, C. D. Dynamic reorganization of neuronal activity patterns in parietal cortex. *Cell* **170**, 986–999 (2017).

31. Schoonover, C. E., Ohashi, S. N., Axel, R. & Fink, A. J. Representational drift in primary olfactory cortex. *Nature* 1–6 (2021).

32. Marks, T. D. & Goard, M. J. Stimulus-dependent representational drift in primary visual cortex. *Nat. communications* **12**, 1–16 (2021).

33. Hénaff, O. J. *et al.* Primary visual cortex straightens natural video trajectories. *Nat. communications* **12**, 5982 (2021).

34. Hénaff, O. J., Goris, R. L. & Simoncelli, E. P. Perceptual straightening of natural videos. *Nat. neuroscience* **22**, 984–991 (2019).

35. Dayan, P., Abbott, L. F. *et al.* Theoretical neuroscience: computational and mathematical modeling of neural systems. *J. Cogn. Neurosci.* **15**, 154–155 (2003).

36. Dayan, P. & Abbott, L. F. *Theoretical neuroscience: computational and mathematical modeling of neural systems* (MIT press, 2005).

37. Brette, R. Philosophy of the spike: rate-based vs. spike-based theories of the brain. *Front. systems neuroscience* 151 (2015).

38. Singer, W. Time as coding space? *Curr. opinion neurobiology* **9**, 189–194 (1999).

39. Aitken, K., Garrett, M., Olsen, S. & Mihalas, S. The geometry of representational drift in natural and artificial neural networks. *PLOS Comput. Biol.* **18**, e1010716 (2022).

40. Geva, N., Deitch, D., Rubin, A. & Ziv, Y. Time and experience differentially affect distinct aspects of hippocampal representational drift. *Neuron* (2023).

41. Khatib, D. *et al.* Active experience, not time, determines within-day representational drift in dorsal ca1. *Neuron* (2023).

42. Wang, S., de Laittre, E. A., MacLean, J. & Palmer, S. E. Quantifying stimulus-relevant representational drift using cross-modality contrastive learning. *arXiv preprint arXiv:2305.11953* (2023).

43. McGinley, M. J. *et al.* Waking state: rapid variations modulate neural and behavioral responses. *Neuron* **87**, 1143–1161 (2015).

Acknowledgements

This project was financed by the BMF (Bundesministerium fuer Bildung und Forschung), Computational Life Sciences, project BINDA (031L0167).

Author contributions statement

Conception and problem statement: BSG, MV. Data analysis and simulations: BSG. Mathematical analysis: BSG and MV. Supervision: MV, FB. Writing of main draft: BSG, FB and MV.

Additional information

To include, in this order: **Accession codes** (where applicable); **Competing interests** (mandatory statement).

The corresponding author is responsible for submitting a competing interests statement on behalf of all authors of the paper. This statement must be included in the submitted article file.

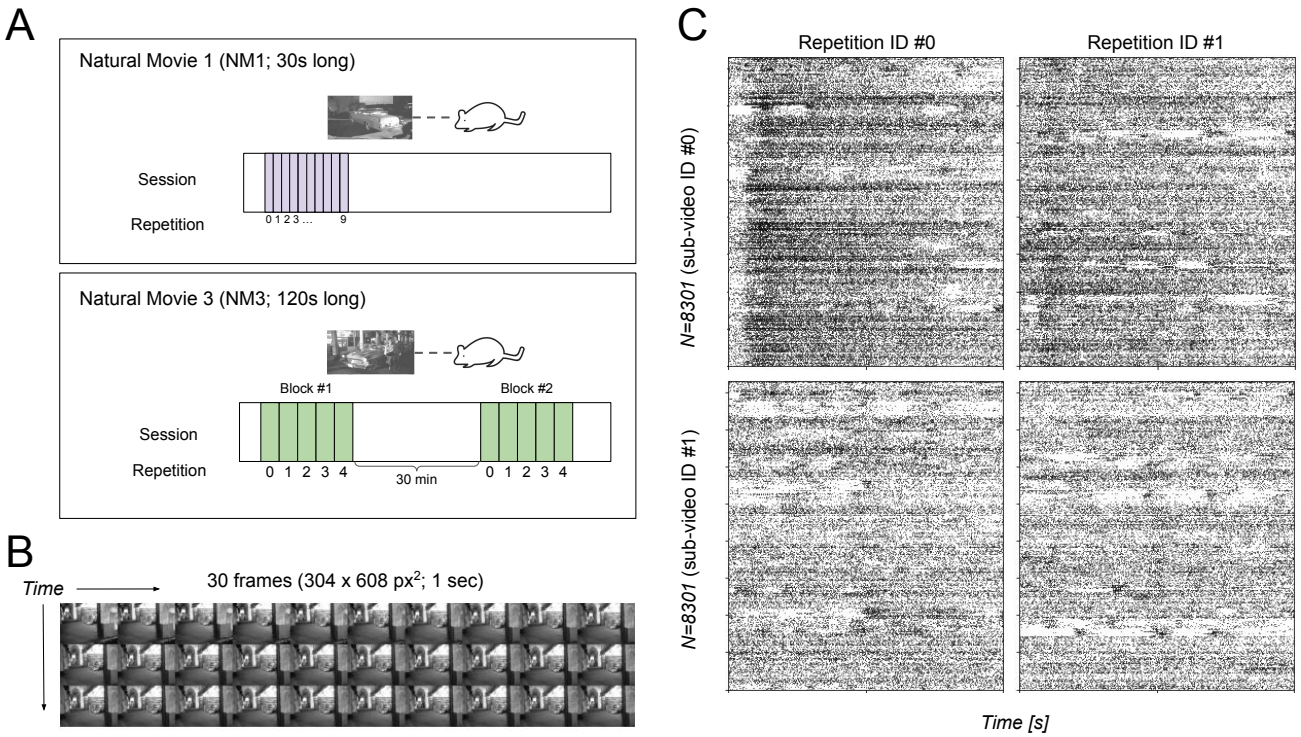

Figure 1. Neural population activity during natural video presentations. A) Schema of natural movies presentations from experiments from Allen Brain Institute. B) Example of frames in natural movie one. Each epoch/trial corresponds to a one-second length sub-video (30 FPS) with 10 repetitions (first epoch). C) Raster plots of first (top) and second (bottom) epochs and their first (left) and second (right) repetition with $N = 8,301$ neurons, which were pooled across 32 sessions. Each epoch/sub-video consider a window length of one second without overlap between epochs.

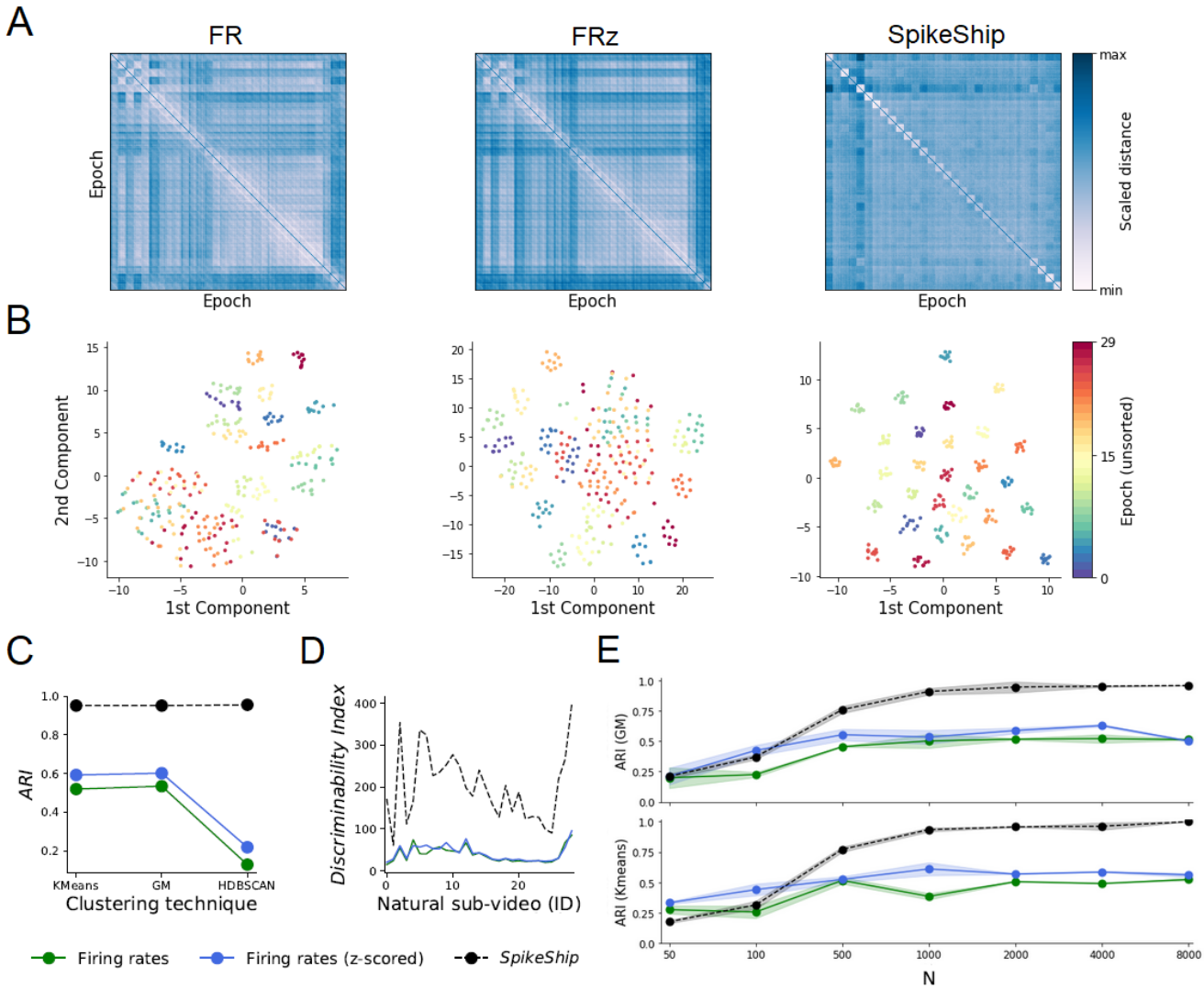

Figure 2. Coding of natural videos via spike count and temporal sequences. A) Dissimilarity matrices for Firing rates (FR), Firing rates z-scored across epochs (FRz), and SpikeShip. SpikeShip quantifies the similarity of epochs in terms of the relative spike-timing relationships among all neurons, based on optimal transport. Epochs are sorted by sub-video's ID. The diagonals of dissimilarity matrices were filled with their maximum values for visualization purposes. B) 2D t-SNE embedding of pairwise distances. Color represents one epoch (unsorted). C) Clustering performance measured via Adjusted Rand Index (ARI) using K-Means, Gaussian mixture clustering model (GM), and HDBSCAN. In the case of K-Means and GM, the number of clusters to determine in the low-dimensional embedding equals the number of distinct sub-videos (i.e., $K = 30$). D) Discriminability index for each measure across sub-videos. E) ARI Score (K-Means) by randomly subsampling N neurons for Natural movie one.

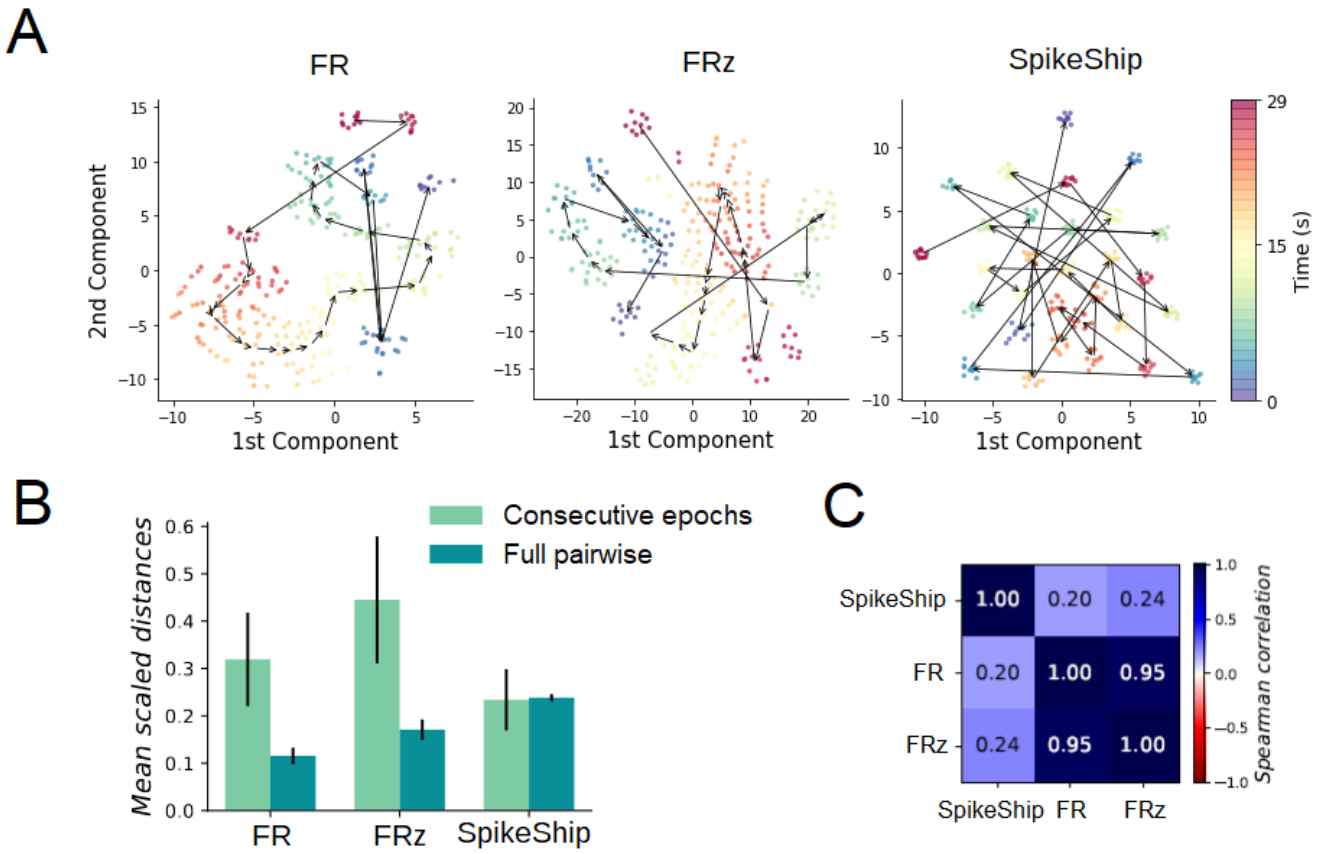

Figure 3. Relation of neural activity to elapsed time in natural movie. Color represents the time of epoch presentation (i.e., second of the movie). Arrows represent the direction of the trajectory between consecutive sub-videos (centroids). Population firing rates form a continuous trajectory, whereas temporal sequence representations are discontinuous. B) Mean scaled (Euclidean) distances across consecutive epochs versus full pairwise distances between epochs. The black line represents the standard deviation. C) Correlation between SpikeShip and firing rate based on the upper triangle of the epoch-to-epoch dissimilarity matrices (excluding comparisons between the same sub-video).

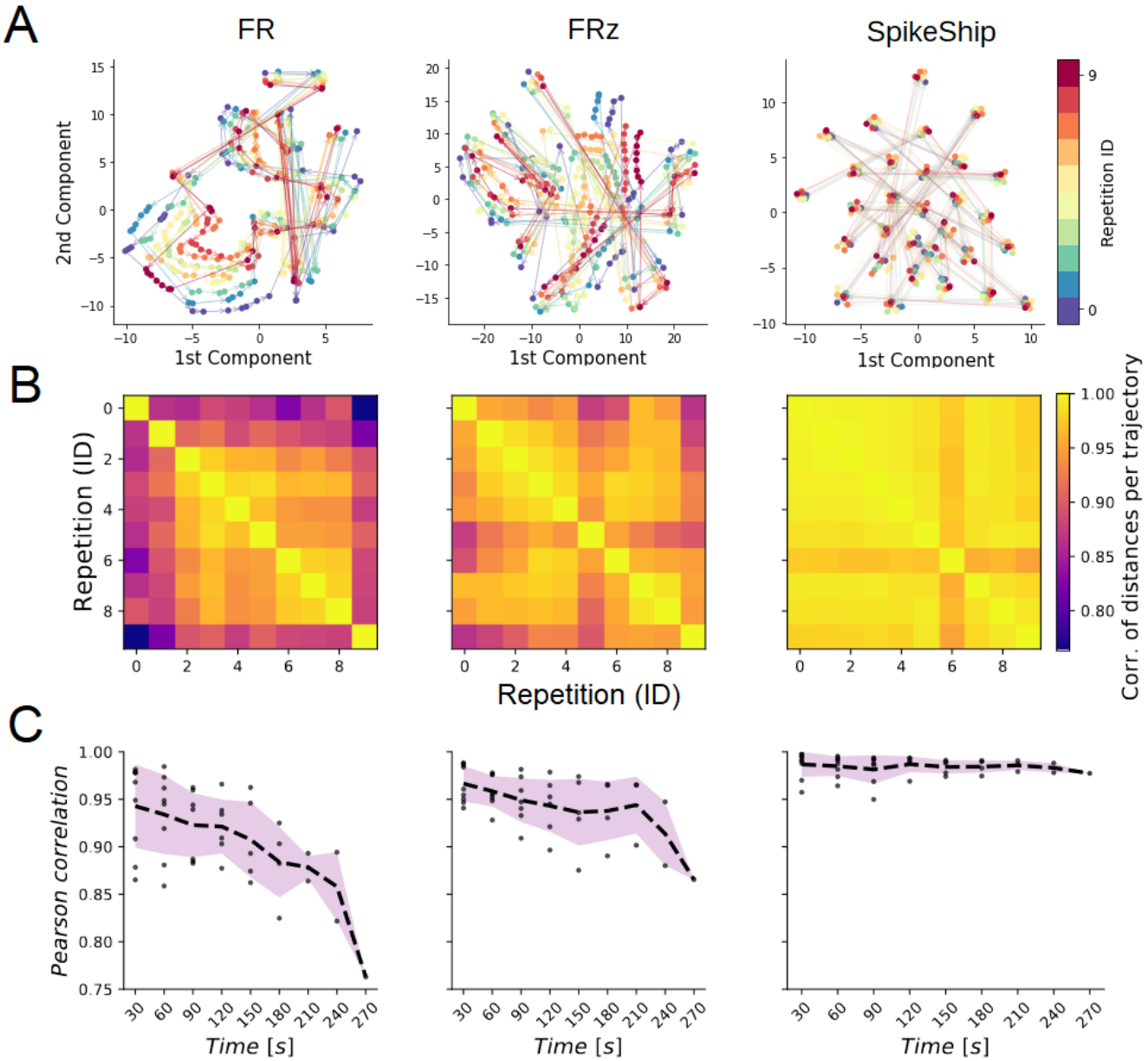

Figure 4. Changes in neural representations across repetitions. A) 2D t-SNE embedding of pairwise distances. Color represents the repetition ID of each Natural Movie presentation. Arrows represent the direction of the trajectory between consecutive sub-videos. Each trajectory corresponds to one repetition. B) For each repetition, the Euclidean distance was computed between sub-videos. Pearson correlations were then computed between these vectors of Euclidean distances. C) Pearson correlation as a function of time.

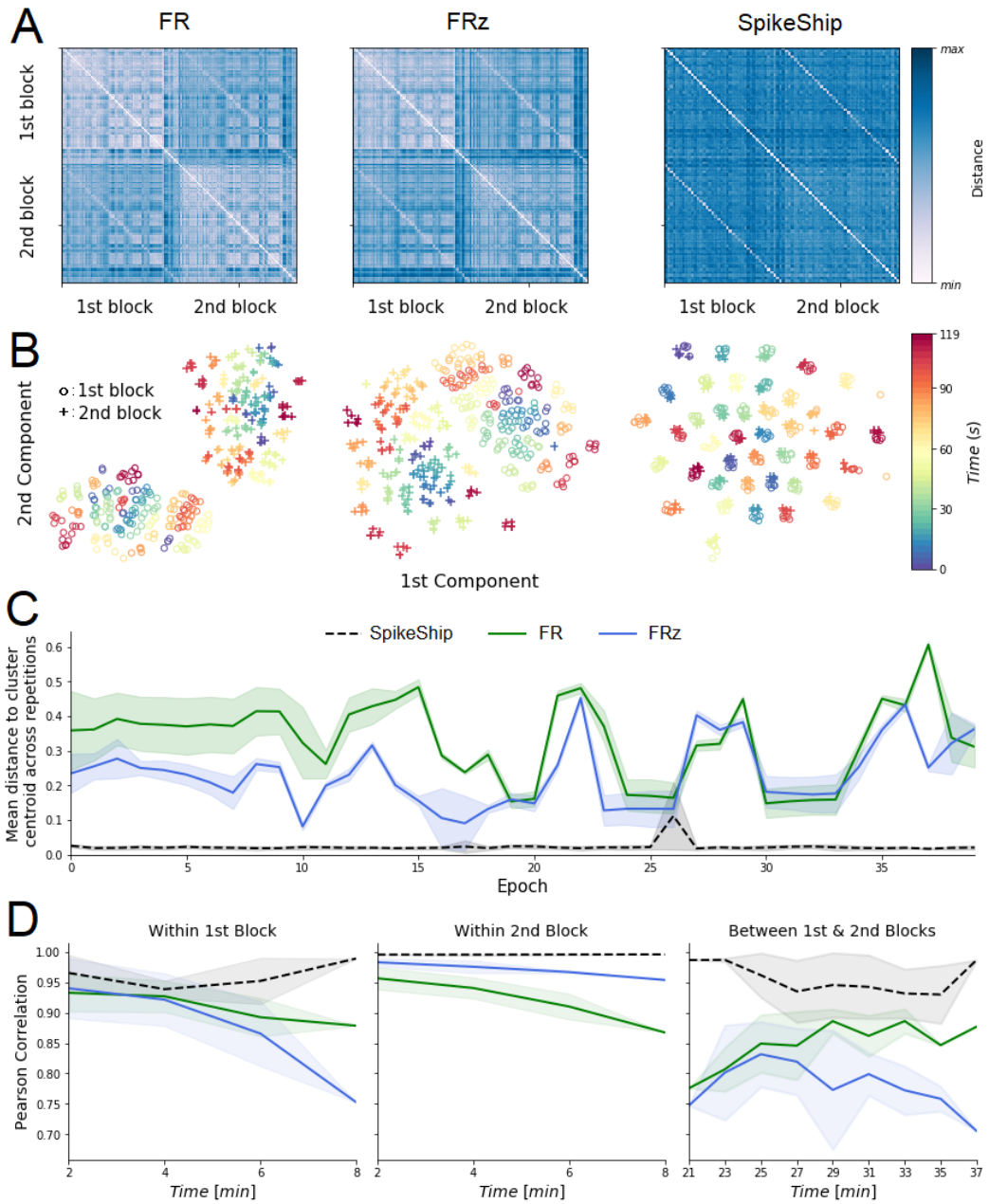

Figure 5. Rate-coding vs temporal coding for long natural movie presentations (NM3). A) Dissimilarity matrix for Firing rates (FR), z-scored Firing rates (FRz), and SpikeShip. Epochs are sorted by sub-video's ID and block. B) 2D t-SNE embedding of pairwise distances. Colors represent sub-video's start time. C) Mean distance to cluster centroids across repetitions. D) Correlation analysis within and between blocks of NM3 presentations, similar to Fig. 4C.

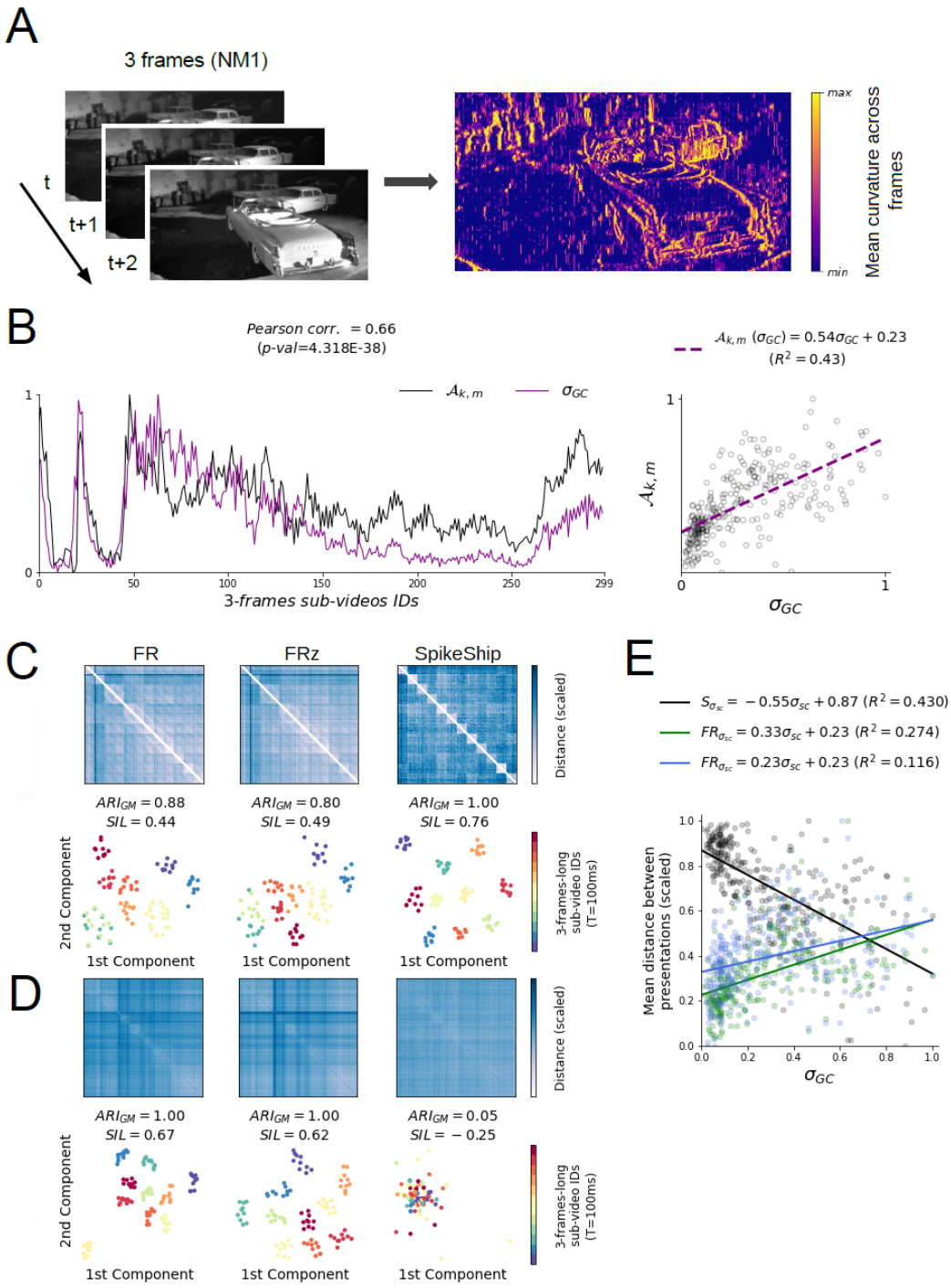

Figure 6. Relationship of population coding to number of active neurons. A) Summary of computation of mean curvature across sub-videos of 3-frames (≈ 100ms). Pixel distance of natural movies frame is computed as in[33,34], mentioned as Global Curvature (GC). B) Raster plot of the proportion of active neurons between all the pairs of epochs k and m ($\mathscr{A}_{k,m}$) and the standard deviation of the global curvature across frames (σ_{GC}). Pixel variability σ_{GC} across sub-videos is correlated with the rate of active neurons $\mathscr{A}_{k,m}$ (i.e., Pearson correlation = 0.66). C) 10 epochs with the highest variability of the global curvature (σ_{GC}). Top: Dissimilarity matrices for firing rates (FR), z-scored firing rates (FRz, normalization across epochs), and SpikeShip. Bottom: 2D t-SNE embeddings with precomputed distance matrix. To measure clustering performance we used ARI_{GM} and SIL which represent Adjusted Rand Index using Gaussian Mixtures and Silhouette, respectively. D) 10 epochs with the lowest variability of the global curvature (σ_{GC}). Top: Dissimilarity matrices for firing rates, z-scored firing rates (normalization across epochs), and SpikeShip. Bottom: 2D t-SNE embeddings with precomputed distance matrix. E) Multi-neuron distance depends on σ_{GC}. Pearson correlation coefficients of $FR_{\sigma_{GC}}$, $FRz_{\sigma_{GC}}$, and $S_{\sigma_{GC}}$ (i.e. SpikeShip) are 0.52 (p-val $< 10^{-220}$), 0.34 (p-val $< 10^{90}$), and -0.66 (p-val $< 10^{-381}$), respectively.

Supplementary information

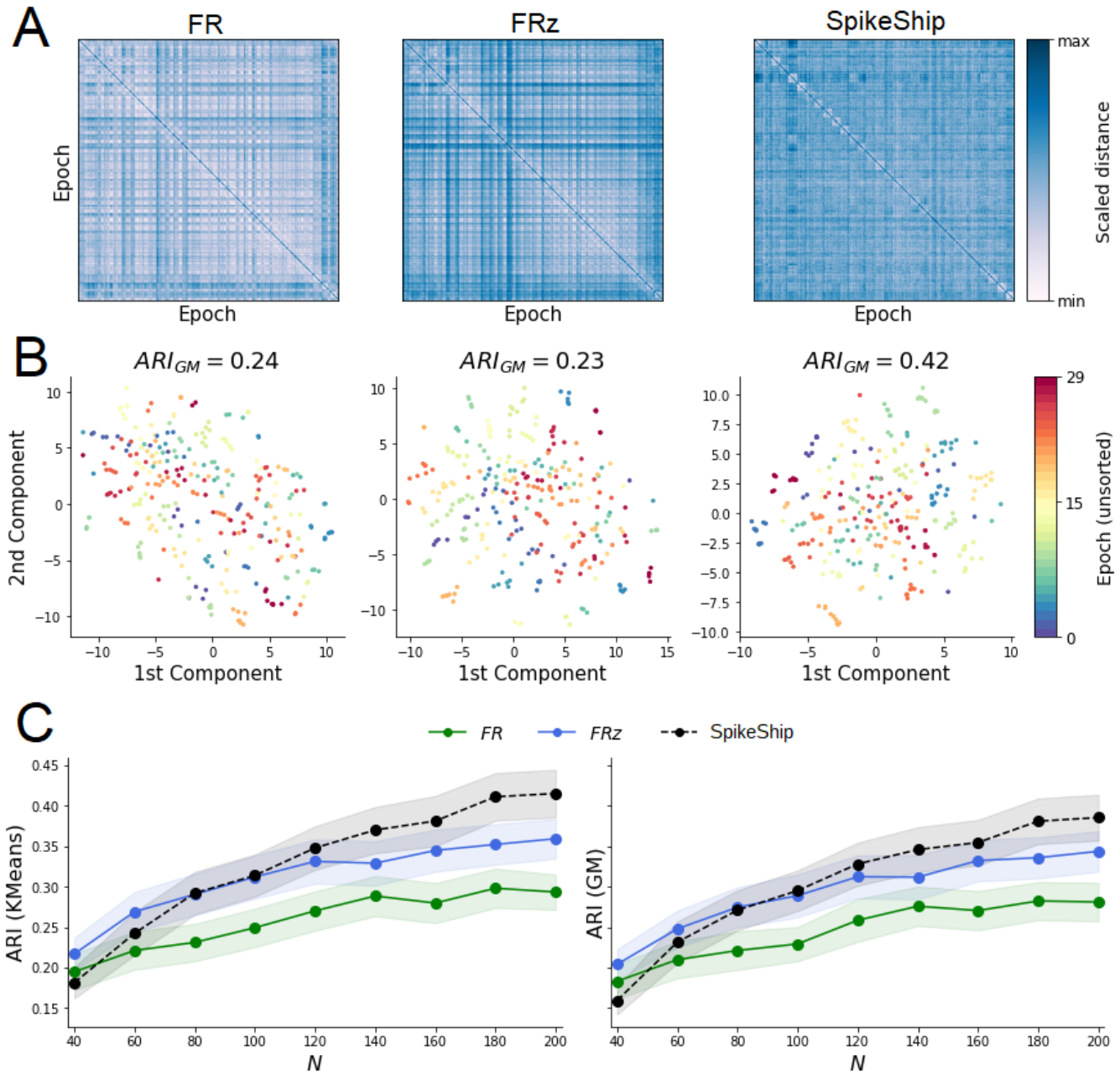

Figure S1. Population coding of natural videos for single sessions. A) Dissimilarity matrices sorted by sub-video presentation time. $M = 300$ epochs (30 sub-videos with 10 repetitions each) and $N = 365$ neurons. The diagonals of each dissimilarity matrix contain the maximum values for visualization purposes. B) 2D t-SNE embeddings from pre-computed values shown in A). Color represent the sub-videos (unsorted). C) Clustering performance for an increasing number of neurons across sessions. We selected sessions with minimally 200 neurons. $ARI(KMeans)$ and $ARI(GM)$ correspond to the Adjusted Rand Index using KMeans and Gaussian Mixtures (GM) as clustering techniques, respectively. Error bars correspond to the standard deviation across sessions divided by the square root of the number of sessions (i.e. $\sigma_{ARI,\,metric}/\sqrt{23}$).

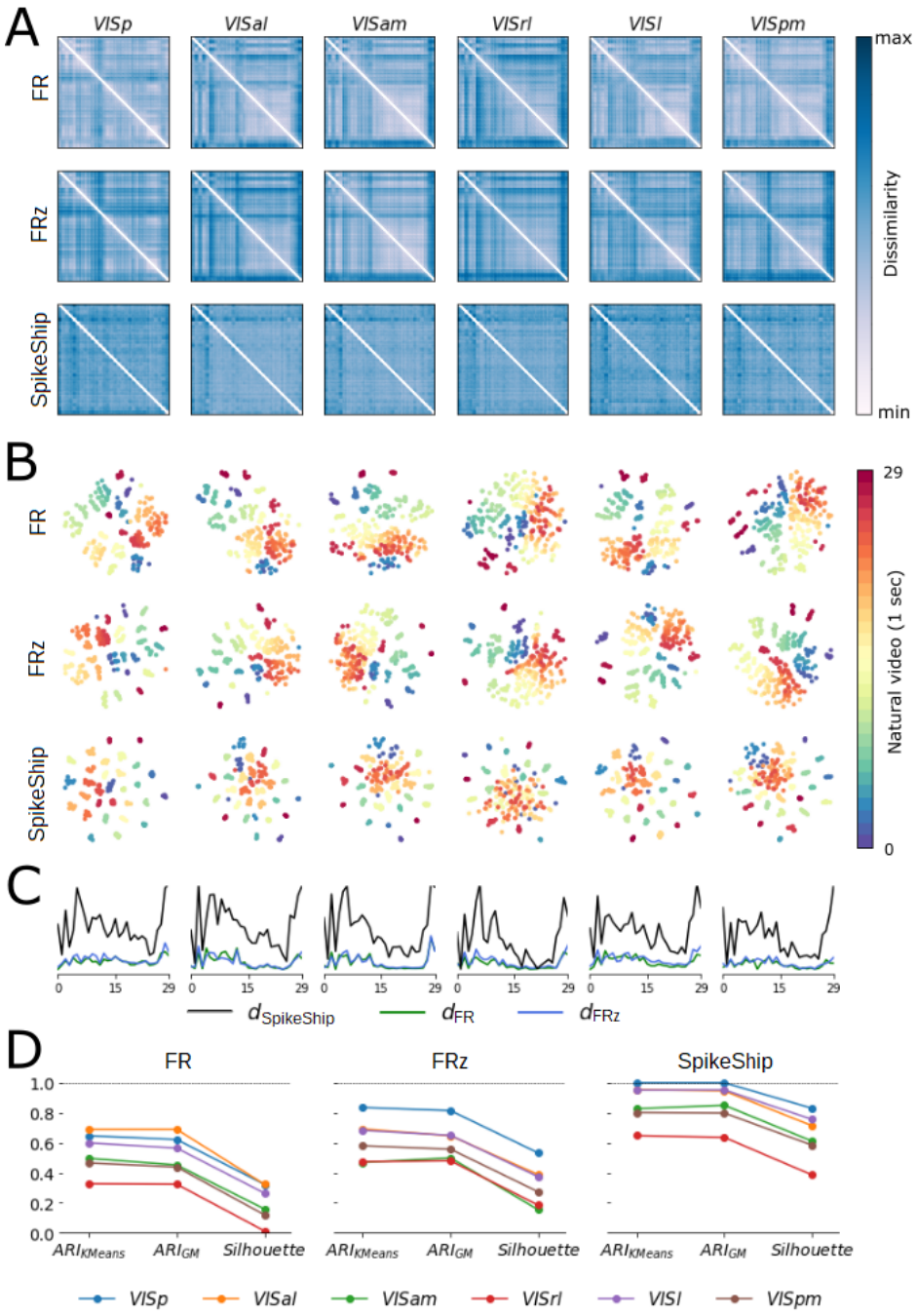

Figure S2. Reliability of time-coding scheme is preserved across brains areas. A) Dissimilarity matrices per brain area (columns) and measures (rows). B) 2D t-SNE embeddings per brain area (columns) and measures (rows). C) Discriminability index for each measure across sub-videos and brain areas. D) ARI and Silhouette score for each measure across sub-videos and brain areas.

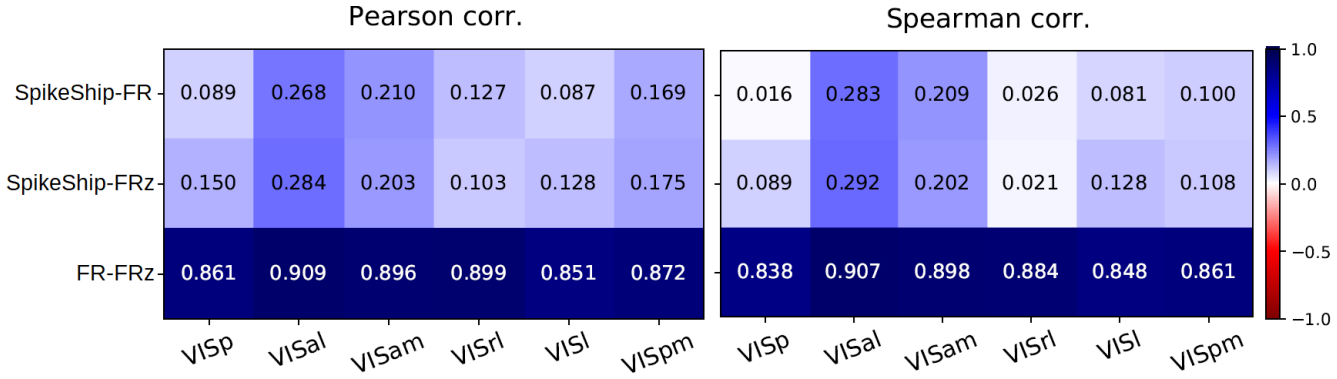

Figure S3. Representational similarity analysis between firing rate and SpikeShip, similar to Fig. 3, but now for separate areas. Left: Pearson correlation coefficient, Right: Spearman correlation coefficient. Both coefficients show that the relation (SpikeShip, FR) and (SpikeShip, FRz) are close to zero across visual brain areas.

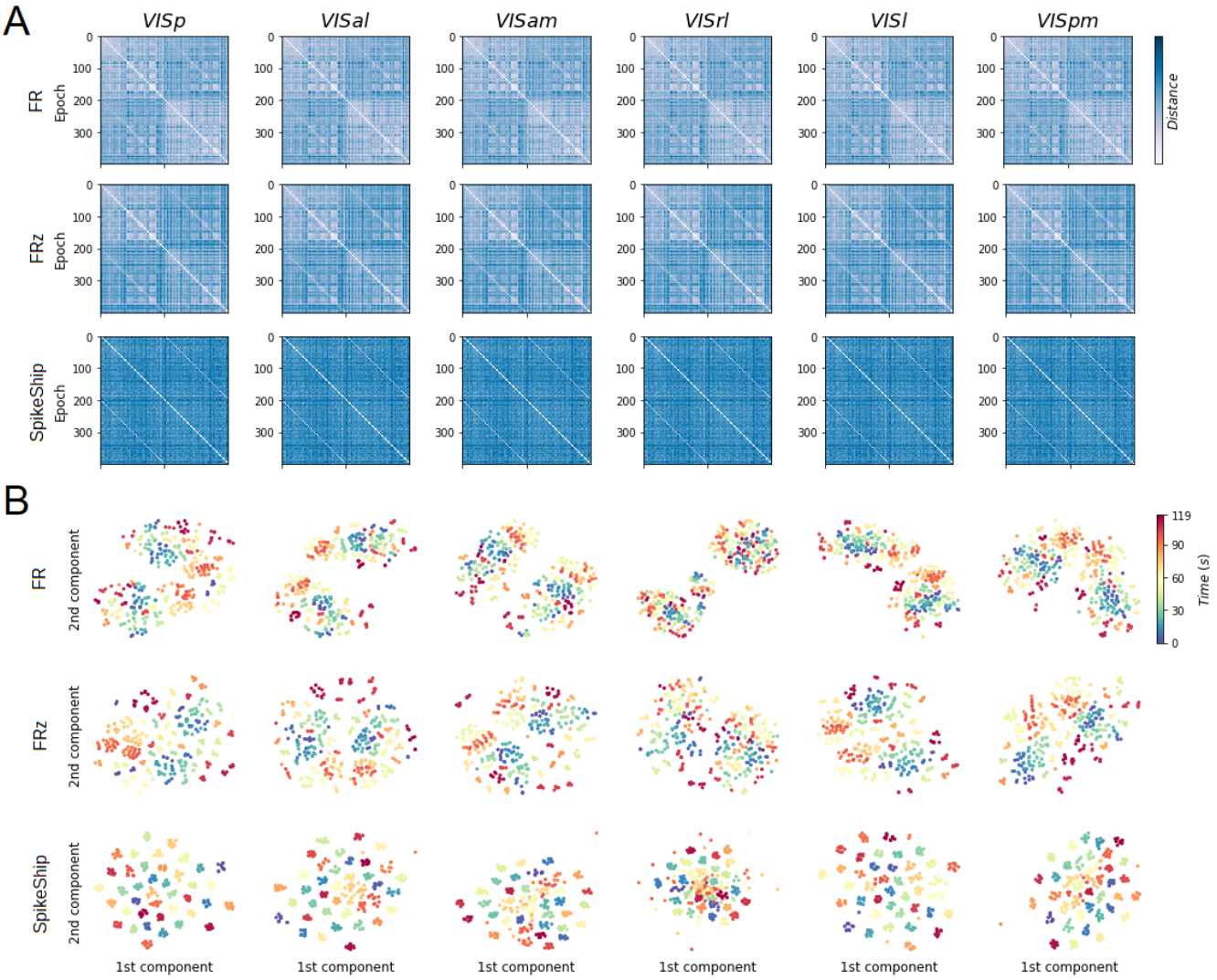

Figure S4. Reliability of time-coding scheme is preserved across brains areas and blocks of natural movie presentations. A) Dissimilarity matrices per brain area (columns) and measures (rows). B) 2D t-SNE embeddings per brain area (columns) and measures (rows).

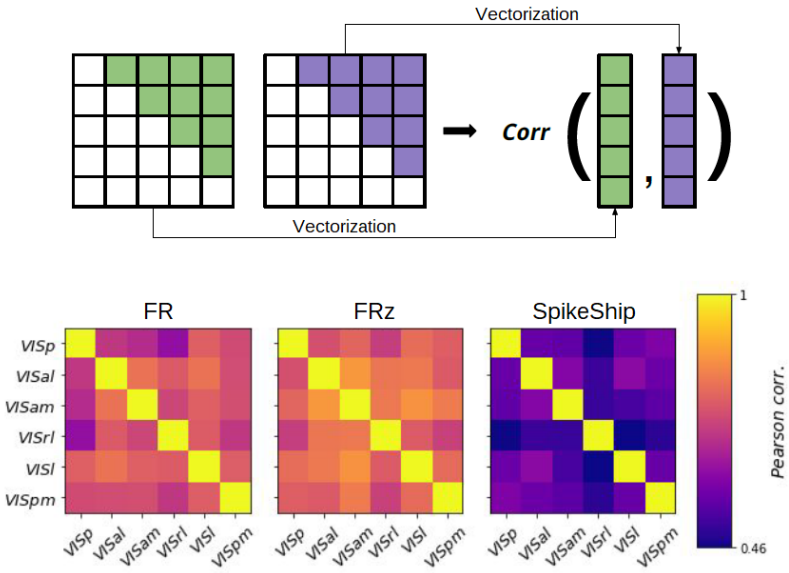

Figure S5. Representational dissimilarity analysis. Top: Summary of vectorization (representational vectors) process for dissimilarity matrices and computation of pairwise correlations. Bottom: Pairwise correlation of representational vectors using Pearson correlation coefficient per visual area across 32 mice ($N = 8301$).

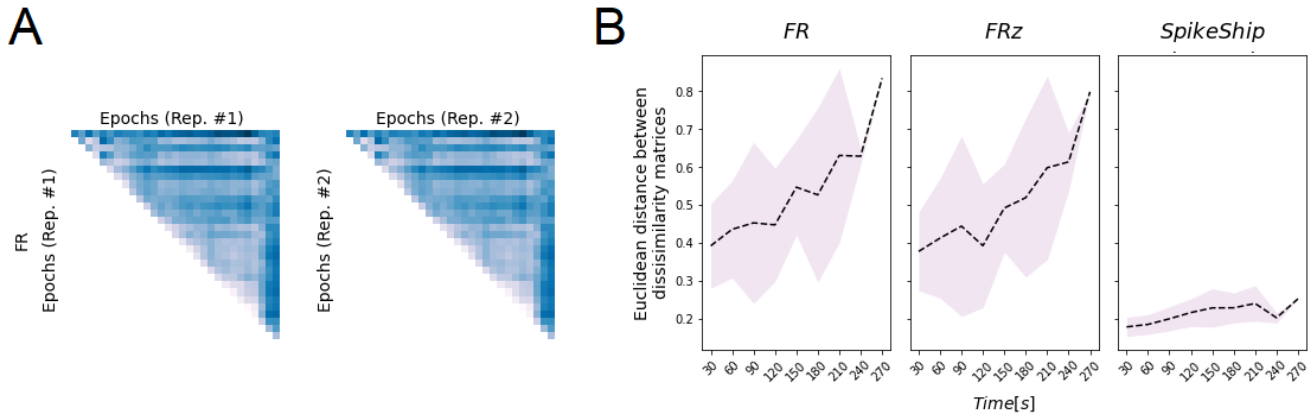

Figure S6. Representational drift as the distance between dissimilarity matrices. A) Example of dissimilarity matrices for two repetitions. Each matrix contains the dissimilarity of multi-neuron spike sequences across the entire movie (M=30 epochs with sub-videos of 1 second). Epochs are sorted by time, as in[26]. B) Euclidean distance between dissimilarity matrices in function of the time difference between two repetitions. Lines and filled regions represent the mean and standard deviation of the correlation for each metric.