A Machine Learning Approach to Investigate Fronto-Parietal Neural Ensemble Dynamics During Complex Behavior

Medorian Gheorghiu[1]
Transylvanian Institute of Neuroscience
Ploiești 33, 400157 Cluj-Napoca, Romania
Technical University of Cluj-Napoca
Barițiu 26-28, 400027, Cluj-Napoca, Romania
Email: gheorghiu@tins.ro

Andrei Ciuparu[1]
Transylvanian Institute of Neuroscience
Ploiești 33, 400157 Cluj-Napoca, Romania
Technical University of Cluj-Napoca
Barițiu 26-28, 400027, Cluj-Napoca, Romania
Email: ciuparu@tins.ro

Bartul Mimica
Kavli Institute for Systems Neuroscience
Norwegian University of Science and Technology
Olav Kyrres Gate 9, 7030 Trondheim, Norway
Email: bartul.mimica@ntnu.no

Jonathan R. Whitlock
Kavli Institute for Systems Neuroscience
Norwegian University of Science and Technology
Olav Kyrres Gate 9, 7030 Trondheim, Norway
Email: jonathan.whitlock@ntnu.no

Raul C. Mureșan
Transylvanian Institute of Neuroscience
Ploiești 33, 400157 Cluj-Napoca, Romania
Email: muresan@tins.ro

***Abstract*—Brain circuits exhibit very complex dynamics, where individual neurons fire action potentials determining coordinated activity patterns. During behavior, a multitude of brain areas are engaged in planning and execution. A particular focus has been recently put on fronto-parietal circuits, which have well-established roles in coordinating complex movements, and which were recently shown to encode whole-body posture. Individual cells in these areas can exhibit striking behavioral tuning, but it is less well known how the collective firing of the neurons correlates with behavior. Here, our objective was to find a way to highlight ensemble firing patterns correlated with behavioral events. To accomplish this, we introduce a machine learning approach for revealing such associations between high-dimensional multi-neuronal firing patterns and behavioral events in freely running rats. We introduce data representations that can extract firing patterns evolving on multiple timescales, and propose an ordered mapping using Kohonen maps to reveal the stereotypically appearing patterns on different timescales in a way that considers the structure of the data space. Furthermore, we apply machine learning on the raw and Kohonen-mapped data and study how and why these approaches can reveal complementary information about the relation between firing patterns and behavior. We conclude that fronto-parietal firing patterns exhibit specific expression in relation to behavior. Depending on the type of data representation, machine learning can reveal patterns with different complexity and evolving on various timescales.**

***Index Terms*—Machine Learning, Posterior-Parietal Cortex, Medial Agranular Cortex, Neural Ensembles, Behavior, Clustering, Kohonen Mapping**

I. INTRODUCTION

Complex behavior in mammals is supported by an intricate network of neural circuits, spanning multiple cortical and subcortical regions [1]. A large body of research has been performed in rodents, where the role of the hippocampus [2] and enthorhinal cortex [3] in navigation has been well established. Less well understood are other, "higher-order" areas, that have been suggested to play an active role in planning and action execution, such as the posterior parietal cortex (PPC) [4] and the medial agranular cortex (AGm), also called frontal motor cortex (M2) [5]. It has been recently demonstrated that neurons in these areas exhibit highly selective firing in relation to the posture of the head, back, and neck of freely moving animals [6].

While the firing properties of individual neurons in PPC and AGm during behavior begin to be deciphered, their collective firing statistics and coordination across circuits remains to be explored. In general, the behavior of multi-neuronal high-dimensional firing patterns, also called neural ensembles, cannot be inferred from the low dimensional behavior of their individual, composing neurons [7], [8]. This phenomenon is called emergence, i.e. the behavior of the whole goes beyond that of individual composing elements [9]–[11]. Here, we argue that specialized methods must be developed to cope with multi-neuronal firing pattern statistics, capable of handling high-dimensional data and able to identify systematic relations between firing patterns and behavioral events.

---

[1] These authors contributed equally.

978-1-7281-7166-1/20/$31.00©2020 IEEE

We have previously developed a framework for the visualization and quantification of neural ensemble firing relying on convolution with exponentially-decaying kernels (causal low-pass filtering) and ordered mapping / clustering using 3D Kohonen maps [12], [13]. Here, we extend this framework with a machine learning approach, whereby we attempt to determine the best strategy to explore how multi-neuronal firing ensembles in PPC and AGm are related to the behavior of the animal. We develop a machine learning strategy and explore how convolution and clustering affect the ability of a classifier to predict relevant behavioral events.

II. MATERIALS AND METHODS

A. Recording of electrophysiology data

Extracellular multi-electrode electrophysiology data was recorded at the Kavli Institute for Systems Neuroscience, Tondheim, Norway. Neural activity was recorded from a single Long-Evans rat (3–5 months old, 400–600 g) using a dual chronic implant with silicon probes (NeuroNexus Inc., MI, USA; custom design based on A8x1-tet-2mm-200-121). The two probes were inserted into PPC (-3.8 to -4.25 mm AP, center shank 2.7 mm ML) and AGm (center shank +0.5 mm AP, 0.7 mm ML), and each consisted of 8 tetrodes (four 160 μm^2 iridium recording sites) disposed on eight, 2 mm long shanks. The H32-to-HS36 custom head stage was connected via unity-gain AC-coupled amplifiers to a Digital Lynx 4SX recording station (Neuralynx Inc. Montana, USA) sampling signals at 32 kSamples/s. The headstage was suspended using elastic string to compensate for the weight of the headstage and cable, enabling animals to effectively explore the recording arena. All experiments were performed in accordance with the Norwegian Animal Welfare Act and the European Convention for the Protection of Vertebrate Animals used for Experimental and Other Scientific Purposes.

B. Behavioral setup design and data segmentation

The rat performed a dynamical foraging task in an open field arena (2 × 2 × 0.8 m) that had 32 conical wells in its floor, such that these wells could be filled with chocolate oat milk. The rat was trained to first seek a reward at the "home" well location (fixed at the beginning of the experiment). After consuming the reward there, a random "target" well was filled, usually prompting a more elaborate search. Upon the consumption of the target well, the home well reward would become available again. This cycle was repeated multiple times during a 20 min recording session, allowing for the expression of both a more “naive” exploration for the target wells and an “informed” navigation to the home well.

Relying on a video recorded by a camera placed on the rat’s headstage, the moments when the rat arrived and departed from home and target wells were identified. These moments were then used as alignment triggers for subsequent data analysis, and included four types of events: start home (SH - the animal arrived to the home well), end home (EH - the animal left the home well), start target (ST - the animal arrived to the target well), end target (ET - the animal left the target well). During analysis, data was segmented by cutting pieces encompassing 8 seconds before and 8 seconds after these four events. Each such 16 seconds data segment was called a "trial".

C. Spike extraction and generation of activity vectors (AV)

Spike extraction from the multi-channel recording was performed using a thresholding operation. The signal was first high-pass filtered (IIR Butterworth order 3, cutoff 300 Hz), in order to remove the low frequencies, and a threshold of 3 standard deviations was chosen. Whenever the signal exceeded this threshold, we considered that a spike was present. Detected spikes were aligned to their peak value and a waveform of 20 samples (0.6 ms) before and 37 samples (1.2 ms) after the peak value was extracted from the signal for each spike. Features of these waveforms were then clustered using K-means clustering based on features including spike amplitude, spike width, the waveform itself, and the first three PCA components. The resulting clusters were manually merged according to waveform shapes obtaining a set of final clusters corresponding to the isolated single units. The timestamps for spikes were then extracted, sorted according to the identity of the firing cell (single unit), and aligned to a sampling rate of 1,000 samples/s (i.e., expressed in milliseconds).

For each single unit we then computed its activity trace using a technique pioneered by Gerstein and Aertsen [14], convolving its spike train with an exponential decay kernel:

$$a_i(t) = \begin{cases} a_i(t-1) + 1, & \text{if neuron } i \text{ fired at time } t \\ a_i(t-1) \cdot e^{-1/\tau}, & \text{otherwise} \end{cases} \quad (1)$$

where, τ is a time constant, set to 4 different levels (20, 50, 100, and 200 ms).

The time constant, τ, regulates the "timescale" of the analysis and can be viewed as a synaptic integration time constant [13]. After convolution, a real-valued signal was obtained for each identified single unit. As described in our previous work [12], [13], the next step was to construct activity vectors (AV) for each time point along the experiment by sampling across activity traces of all units (17 from PPC and 20 from AGm). We obtained AVs with a dimensionality equal to the total number of cells identified across the two areas (37), representing the integrated coordinated co-firing of the observed population of neurons. For small time constants ($\tau \leq 30$ ms) these AVs reflect synchrony patterns [13], while for large time constants ($\tau \geq 50$ ms) they reflect firing-rate covariations [15].

D. Clustering and generation of Kohonen model vectors (KH)

AVs obtained in the previous step are real-valued vectors and, therefore, they span a potentially infinite number of instantiations. In many analyses, it is desirable to identify a limited number of representative vectors, called firing patterns, such that one can correlate the expression of these patterns with various experimental conditions. To this end, a simple solution is to apply a clustering procedure in high-dimensional space to obtain a number of representative cluster centers, also called model vectors [16].

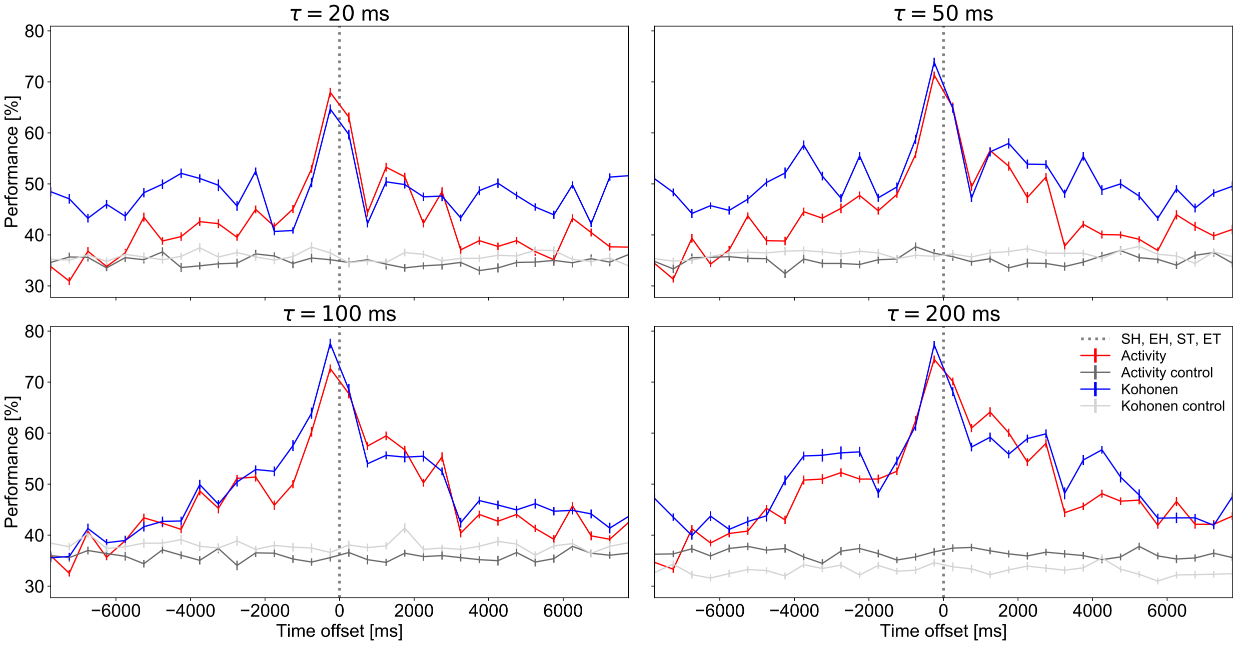

Figure 1. **Machine learning accuracy on different datasets.** The title of each subplot defines the integration time constant.

Clustering can be achieved with multiple techniques, like K-Means [17], but here we opted for a more complex one, based on the Kohonen (KH) self-organizing map [18]. The advantage of using the latter is that one can achieve simultaneously clustering and ordered mapping. We have previously introduced a versatile visualization method for high-dimensional neural data using 3D Kohonen maps [12].

Here we adopted the same strategy and fed the AVs to a 3D Kohonen map (10 × 10 × 10), obtaining 1,000 KH model vectors. We then replaced each activity vector with its closest model vector, obtaining a dataset with a sequence of KH vectors spanning the entire duration of the experiment.

E. Machine learning

The machine learning (ML) paradigm involved a multi-layer perceptron (MLP) [19] trained for 50 epochs to distinguish between the four triggers present in the data (SH, EH, ST, ET), corresponding to events when the animal arrives and leaves the home and target wells. The network's architecture was fixed to 18,500 (37 × 500) input units, two hidden layers of 500 and 100 units, and an output layer of 4 units. All layers had an ELU [20] activation function, except for the output, where we used a softmax layer. The two hidden layers had a dropout of 50%, to prevent overfitting [21]. We used a cross-entropy loss function and trained the networks using the Adam algorithm [22] on batches of 10 samples.

To prepare the data for the ML analysis, the 16 seconds of data (± 8 s) around the four behavioral events was cut into non-overlapping windows of 500 ms. At a certain offset, we obtained a window containing 500 AV or KH vectors for each trial. This window was then flattened and fed to the network as input. Trials were split the into test (23) and training (51) sets randomly. The network was then trained and tested and a performance was obtained for the current offset of the analysis window. The procedure was repeated 50 times in order to control for representative and unrepresentative trials, and we computed the average and standard deviation of classification performance. As an additional control, we also ran the same tests with shuffled labels, such that the correlations between the label and data were destroyed, estimating the chance-level of the classification. Finally, the window was moved to the next time offset and the entire procedure was repeated. Importantly, this setup enables the classifier to learn sequences of patterns expressed at a certain location in the trial.

F. Comparative analyses of AV and KH

In order to determine differences between the classification performance on KH and AV representations, we employed two measures: mean squared error (MSE) and pattern complexity. The MSE was computed for each time point and trial, between the AV and the KH representations, and then averaged over trials. The pattern complexity measure provides information about how many cells are "active" in a certain activity or model vector at a certain time point (for details, see [17]). More specifically, pattern complexity counts the number of cells whose activation in the pattern exceeds 0.36 ($\approx 1/e$). We evaluated complexity at each time point, and then averaged this measure over trials. Complexity was computed separately for the AV and KH representations.

III. RESULTS

We first attempted to determine if coordinated firing of PPC and AGm populations carries informative signatures about the behavioral contingencies while the rat was performing the experimental task. To this end we used a machine learning approach, trying to classify, in a time-resolved manner, the location of the animal with respect to the four behavioral events (SH, EH, ST, ET) - see Figure 1.

A. Profile differences of ML performance on AV and KH

Classifiers reached a significant performance above chance, with a specific peak around the behavioral event (trigger), irrespective of the type of representation, AV or KH (Figure 1). This suggests that the firing patterns expressed across PPC and AGm were more consistently associated to different conditions while the animal was close to the triggers. Indeed, the animal executed different kinds of behaviors in relation to the different events. While in all trials the animal received a reward between start and end events (SH - EH, ST - ET), only prior to SH the animal could anticipate the reward, because only the location of the home well was known. In addition, when departing from the home well, the animal performed a random exploration, searching for the random location of the target well. Classification results indicate that these different behaviors are associated to different sequences of firing patterns across the two areas, with more specificity close to the behavioral events.

Two important effects were revealed by the analysis in Figure 1. First, the maximum classification performance depended on the integration time constant τ: performance peaked already at τ = 100 ms. In addition, at the large τ = 100 and 200 ms (Figure 1, bottom), no significant difference could be found between the classification on AV and KH. Furthermore, for these large integration time constants classification tended to drop towards chance (uncertainty) as one moved away from the behavioral events (at offset 0).

A second interesting effect was found for the lower integration time constants, τ = 20 and 50 ms. For these cases, the KH representation had a higher performance than the AV representation in the uncertain regions. Indeed, especially for τ = 20 ms, the KH representation seemed to contain information about the following / preceding behavioral event, significantly above chance (Figure 1, top-left). These results suggest that the AV and KH representations are not equivalent for lower τ and that Kohonen mapping seems to pick up patterns that are more specifically correlated to the behavior of the animal along the trial.

B. Mean squared error and complexity

To determine the difference between AV and KH representations that may have led to the observed difference in classification performance on the AV and KH sets, we next computed the time-resolved MSE between AVs and KH model vectors, time point by time point for each trial and averaged them across trials (Figure 2, top panel). Similarly, we computed the average complexity of expressed AVs (Figure 2, middle panel) and KH model vectors (Figure 2, bottom panel) in time. All these computations were performed as a function of τ.

We found that the MSE between the AV and KH representations increased in a linear fashion with τ (Figure 2, top panel). The variance and overall value of the MSE also seemed to scale with τ. Regarding complexity, this seemed to scale differently as a function of τ for KH and AV.

By plotting the average complexity of AV and KH representations over the trials on a scatter plot for each τ (Figure 3), we gained more insight into the relationship between complexity and τ. For low values of τ, the AVs had higher complexity than their corresponding model vectors in the KH representation. As τ was increased, the cloud shifted in favor of higher KH complexity. It can also be observed that the complexity of the two representations was correlated. Indeed, this was expected, as the KH representation retained a lot of the information present in the AV representation.

In a final test, we plotted the complexity as a function of τ and tried to fit this relationship with a second degree polynomial (Figure 4). Interestingly, complexity had a more linear dependence on τ for the case of the KH than for the AV representation. The complexity of the AV representation saturated for high τ values, while the KH representation retained higher pattern complexity.

IV. DISCUSSION

We have shown that constructing high-dimensional representations of the firing of multiple neurons in fronto-parietal circuits and application of machine learning enables the detection of non-trivial associations between the expression of firing patterns and behavioral events. This is a significant advantage because, as opposed to experiments targeting sensory areas, where stimuli are presented in a trial-by-trial fashion, behavior is much more variable, being determined by the particular decisions taken by the animal. As a result, defining experimental conditions and cutting the data into segments that are associated with different behaviors is certainly more difficult. We argue that machine learning is particularly valuable for analysis of behavioral experiments because it enables the rapid exploration of the association between neural activity and behavioral events.

Here, we extracted high-dimensional activity patterns by computing activity and model vectors across all identified neurons (37 single units, 17 from PPC and 20 from AGm), thus defining patterns that include the simultaneous activity in PPC and AGm. We found that specific firing patterns tend to occur in relation to different behavioral events and that, as one moves away from these events, patterns on fast and slow timescales seem to exhibit different properties. In particular, stereotypically occurring fast patterns (on timescales of 50 ms and less), identified using Kohonen mapping, provide information about the following or preceding behavioral events. This is not the case for patterns evolving on slower timescales, and cannot be detected by machine learning on fast timescales when the raw activity vectors are used as a feature.

It is unclear why the KH representation remains more specific across time than the AV representation for small timescales. There are several potential explanations. For example, it could be that model vectors computed using Kohonen mapping provide a less noisy representation of the data at

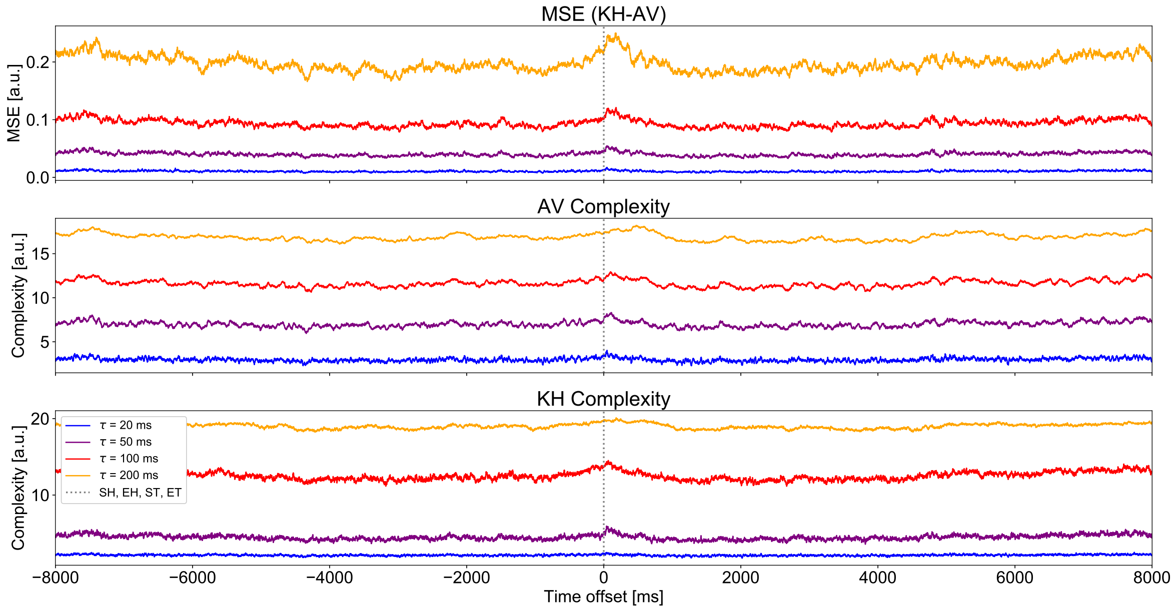

Figure 2. **Mean squared error and complexity** computed on KH and AV representations of the data.

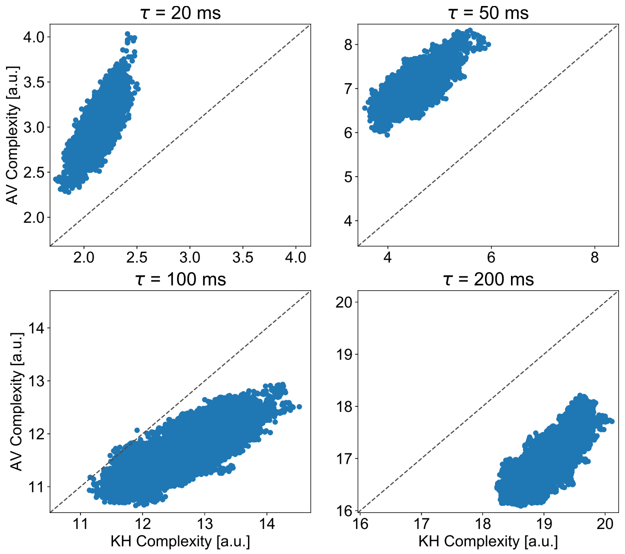

Figure 3. **Comparison of pattern complexity in AV versus KH representations.** The diagonal line represents equal complexity, every point above corresponds to higher complexity in AV, and every point below corresponds to higher complexity in KH. The panels are titled with the time constant used to create the datasets.

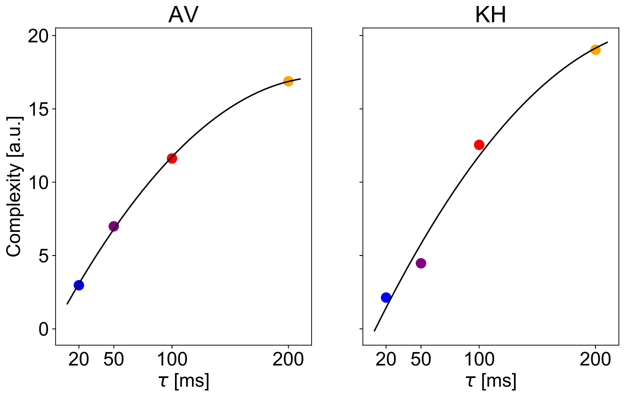

Figure 4. **Pattern complexity in the AV and KH representations** as a function of τ. The black line represents a second degree polynomial fit.

small integration time constants, where random, unrelated spikes, tend to produce a large number of "noise" patterns. Therefore, applying clustering before the machine learning step may be increasing signal-to-noise ratio and favor the detection of specific patterns.

Another potential explanation for the more sustained performance on the KH set has to do with the information fed to the classifier: sequences of patterns in windows of 500 ms duration. On fast timescales, these sequences are particularly sensitive to unrelated spikes, fired by neurons whose activity is not correlated to a particular behavioral event. Because Kohonen mapping creates representations that favor frequently appearing patterns, it is possible that KH vector sequences are less disturbed by event-unrelated spikes. This however has important implications, because it suggests that firing patterns exist, at least across a fraction of neurons in the observed population, which predict behavioral events that will follow / have passed at a temporal distance of seconds. This result expands previous observations on individual PPC cells, which have been shown to fire selectively in anticipation to particular head movements [23].

A potential confirmation for the above scenario comes from complexity analysis. The latter revealed that on fast timescales the KH representation produces patterns with less participating neurons, likely detecting subsets of neurons which fire systematically in relation to behavioral events. On the other hand, for large timescales, this relation reverses, with the KH retaining higher complexity patterns than the AV representation.

The relation between pattern complexity and timescale is more linear and can preserve higher dimensional patterns at large timescales for the KH than for the AV representation. This suggests that Kohonen mapping suppresses rarely appearing high-complexity patterns on fast timescales and retains the more frequently appearing high-complexity patterns on large timescales. The present results suggest that the Kohonen mapping representation we have introduced earlier [12], [13] has the potential of revealing complementary information with respect to the raw activity vectors, in a timescale dependent manner. On fast timescales, it is possible that the KH representation detects robust subsets of neurons which fire in a condition-specific manner, suppressing the spikes of neurons whose activity is noisy and very sparse.

By including the firing of both PPC and AGm neurons in the high-dimensional vectors, we have quantified how cross-area coordination is associated to behavioral events. In a future study, it will be interesting to address how activity in each area alone is correlated with behavior and to determine the relationship between patterns in PPC and patterns in AGm (correlation, delay, etc). These investigations can be performed by using a machine learning approach similar to that introduced here. Furthermore, future studies should also elucidate the manner in which the representations using Kohonen mapping are different from those that consider just the raw signal traces. Because Kohonen maps perform an ordered mapping, the resulting model vectors represent a high-dimensional alphabet that may reflect some intricate structure within the high-dimensional neural data space.

V. CONCLUSIONS

The application of machine learning techniques to data recorded in complex neuroscience experiments is a fruitful avenue. We have shown that it allows important inferences about how neural populations are engaged during behavior. While the number of potential applications in the analysis of neuroscience data is huge, we would like to argue that conclusions drawn from application of machine learning tools to such data need to be formulated carefully as they may depend on the particular way data has been preprocessed and on how the features were defined and computed.

ACKNOWLEDGMENTS

This work was supported by: two grants from the Romanian National Authority for Scientific Research and Innovation, CNCS-UEFISCDI (Project Numbers PN-III-P4-ID-PCE-2016-0010 and COFUND-NEURON-NMDAR-PSY), an ERC Starting Grant no. 335328 (RAT MIRROR CELL) funded by the European Commission, and a National Science Foundation grant NSF-IOS-1656830 funded by the U.S. Government. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] M. M. Jankowski, K. C. Ronnqvist, M. Tsanov, S. D. Vann, N. F. Wright, J. T. Erichsen, J. P. Aggleton, and S. M. O'Mara, "The anterior thalamus provides a subcortical circuit supporting memory and spatial navigation." *Frontiers in systems neuroscience*, vol. 7, p. 45, Aug. 2013.

[2] H. S. Courellis, S. U. Nummela, M. Metke, G. W. Diehl, R. Bussell, G. Cauwenberghs, and C. T. Miller, "Spatial encoding in primate hippocampus during free navigation." *PLoS biology*, vol. 17, p. e3000546, Dec. 2019.

[3] E. I. Moser, Y. Roudi, M. P. Witter, C. Kentros, T. Bonhoeffer, and M.-B. Moser, "Grid cells and cortical representation." *Nature reviews. Neuroscience*, vol. 15, pp. 466–481, Jul. 2014.

[4] J. R. Whitlock, "Movement coding at the mesoscale in posterior parietal cortex." *Neuron*, vol. 95, pp. 1234–1236, Sep. 2017.

[5] F. Barthas and A. C. Kwan, "Secondary motor cortex: Where 'sensory' meets 'motor' in the rodent frontal cortex." *Trends in neurosciences*, vol. 40, pp. 181–193, Mar. 2017.

[6] B. Mimica, B. A. Dunn, T. Tombaz, V. P. T. N. C. S. Bojja, and J. R. Whitlock, "Efficient cortical coding of 3d posture in freely behaving rats." *Science (New York, N.Y.)*, vol. 362, pp. 584–589, Nov. 2018.

[7] W. Singer, A. K. Engel, A. K. Kreiter, M. H. Munk, S. Neuenschwander, and P. R. Roelfsema, "Neuronal assemblies: necessity, signature and detectability." *Trends in cognitive sciences*, vol. 1, pp. 252–261, Oct. 1997.

[8] S. Klampfl and W. Maass, "Emergence of dynamic memory traces in cortical microcircuit models through stdp." *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 33, pp. 11 515–11 529, Jul. 2013.

[9] M. Bunge, "Emergence and the mind," *Neuroscience*, vol. 2, no. 4, pp. 501–509, 1977.

[10] E. Tognoli and J. A. S. Kelso, "The metastable brain." *Neuron*, vol. 81, pp. 35–48, Jan. 2014.

[11] P. A. Corning, "The re-emergence of "emergence": A venerable concept in search of a theory," *Complexity*, vol. 7, no. 6, pp. 18–30, 2002.

[12] O. F. Jurjuț, D. Nikolić, G. Pipa, W. Singer, D. Metzler, and R. C. Mureșan, "A color-based visualization technique for multi-electrode spike trains," *J Neurophysiol*, vol. 102, pp. 3766–3778, 2009.

[13] O. F. Jurjuț, D. Nikolić, W. Singer, S. Yu, M. N. Havenith, and R. C. Mureșan, "Timescales of multineuronal activity patterns reflect temporal structure of visual stimuli." *PLoS One*, vol. 6, no. 2, p. e16758, 2011. [Online]. Available: http://dx.doi.org/10.1371/journal.pone.0016758

[14] G. L. Gerstein and A. M. Aertsen, "Representation of cooperative firing activity among simultaneously recorded neurons," *Journal of Neurophysiology*, vol. 54, no. 6, pp. 1513–1528, 1985.

[15] D. Nikolić, R. C. Mureşan, W. Feng, and W. Singer, "Scaled correlation analysis: a better way to compute a cross-correlogram." *The European journal of neuroscience*, vol. 35, pp. 742–762, Mar. 2012.

[16] O. F. Jurjuț, M. Gheorghiu, W. Singer, D. Nikolić, and R. C. Mureşan, "Hold your methods! how multineuronal firing ensembles can be studied using classical spike-train analysis techniques." *Frontiers in systems neuroscience*, vol. 13, p. 21, 2019.

[17] M. Gheorghiu, A. Nagy-Dăbâcan, and R. C. Mureşan, "Detecting non-redundant collective activity of neurons," in *Proc. IEEE 15th Int. Conf. Intelligent Computer Communication and Processing (ICCP)*, Sep. 2019, pp. 539–543.

[18] T. Kohonen, *Self-Organizing Maps*, 3rd ed., ser. Springer series in information sciences. Berlin, Springer-Verlag, 2001.

[19] A. Ciuparu, A. Nagy-Dăbâcan, and R. C. Mureşan, "Soft++, a multi-parametric non-saturating non-linearity that improves convergence in deep neural architectures," *Neurocomputing*, vol. In press, 2020.

[20] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (elus)," *arXiv preprint arXiv:1511.07289*, 2015.

[21] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *arXiv preprint arXiv:1207.0580*, 2012.

[22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[23] J. R. Whitlock, G. Pfuhl, N. Dagslott, M.-B. Moser, and E. I. Moser, "Functional split between parietal and entorhinal cortices in the rat." *Neuron*, vol. 73, pp. 789–802, Feb. 2012.