LETTER Communicated by Paul Miller

Hidden Markov Models Predict the Future Choice Better Than a PSTH-Based Method

Encarni Marcos
encarni.marcos@uniroma1.it
Department of Physiology and Pharmacology, Sapienza University of Rome, Rome 00185, Italy, and Instituto de Neurociencias de Alicante, Consejo Superior de Investigaciones Científicas–Universidad Miguel Hernández de Elche, Sant Joan d'Alacant, Alicante 03550, Spain

Fabrizio Londei
fabrizio927@gmail.com
Aldo Genovesio
aldo.genovesio@uniroma1.it
Department of Physiology and Pharmacology, Sapienza University of Rome, Rome 00185, Italy

Uncorrected Proof

Beyond average firing rate, other measurable signals of neuronal activity are fundamental to an understanding of behavior. Recently, hidden Markov models (HMMs) have been applied to neural recordings and have described how neuronal ensembles process information by going through sequences of different states. Such collective dynamics are impossible to capture by just looking at the average firing rate. To estimate how well HMMs can decode information contained in single trials, we compared HMMs with a recently developed classification method based on the peristimulus time histogram (PSTH). The accuracy of the two methods was tested by using the activity of prefrontal neurons recorded while two monkeys were engaged in a strategy task. In this task, the monkeys had to select one of three spatial targets based on an instruction cue and on their previous choice. We show that by using the single trial's neural activity in a period preceding action execution, both models were able to classify the monkeys' choice with an accuracy higher than by chance. Moreover, the HMM was significantly more accurate than the PSTH-based method, even in cases in which the HMM performance was low, although always above chance. Furthermore, the accuracy of both methods was related to the number of neurons exhibiting spatial selectivity within an experimental session. Overall, our study shows that neural activity is better described when not only the mean activity of individual neurons is considered and that therefore, the study of other signals rather than only the average firing rate is fundamental to an understanding of the dynamics of neuronal ensembles.

Neural Computation **31, 1–17** (2019)
doi:10.1162/neco_a_01216
© 2019 Massachusetts Institute of Technology

1 Introduction

As new techniques to simultaneously record neural activity are being developed, the need for analytical methods to take advantage of these huge data sets is increasing in parallel. So far, the use of one or just a few single electrodes to obtain neural data has allowed for only the recording of a small number of neurons within a session. The most common approach to identifying behavioral neural correlates was to replicate the same conditions several times because since behavior is stochastic, even for the same conditions the responses that were obtained varied. The average neural activity in the different behavioral conditions was then compared. However, it has been shown that not only the mean firing rate but other measures, such as across-trials variability, carry relevant information of a task (Marcos et al., 2013). In past decades, classification procedures have been applied to neural data to decode stimulus-related information (Ghazanfar, Stambaugh, & Nicolelis, 2000; Gochin, Colombo, Dorfman, Gerstein, & Gross, 1994; Nicolelis et al., 1998; Nicolelis, Lin, & Chapin, 1997; Rolls, Treves, Robertson, Georges-François, & Panzeri, 1998; Rolls, Treves, & Tovee, 1997). Moreover, with the aim of reducing computational complexity, a minimum distance method solely based on the peristimulus time histogram (PSTH) of neurons has been proposed and presented as an alternative to more sophisticated classification methods (Foffani & Moxon, 2004) that allow for varying contributions of the neurons, beyond the average firing rate of the population.

In parallel with the classification problem, hidden Markov models (HMMs) have been applied to describe the dynamics of neural ensembles (Jones, Fontanini, Sadacca, Miller, & Katz, 2007; Mazzucato, Fontanini, & La Camera, 2015). Assuming that state transitions occur following a Markov stochastic process, HMMs can be applied to single trials to identify the different latent states and estimate the state of neural ensembles at each time point (Escola, Fontanini, Katz, & Paninski, 2011; Ponce-Alvarez, Kilavik, & Riehle, 2008; Ponce-Alvarez, Nácher, Luna, Riehle, & Romo, 2012; Seidemann, Meilijson, Abeles, Bergman, & Vaadia, 1996). For instance, using HMMs, it has been shown that after taste delivery, neurons in the gustatory system go through sequences of metastable states (Jones et al., 2007). Here, we aimed to use a HMM to classify the neural activity observed in a single trial while two monkeys performed a strategy task (Genovesio, Brasted, Mitz, & Wise, 2005) and then to compare the classification performance of the model with the recently developed PSTH-based method. In the strategy task, the monkeys had to select one of three spatial targets based on an instruction cue. We assessed the accuracy of the two methods using the neural activity observed in the period after the cue delivery. The action could be performed only after a "go" signal, and thus, in such a premovement period, the neural activity is predictive of future choice. Our data set

Uncorrected Proof

consisted of the activity of prefrontal neurons recorded using single electrodes in sessions with multiple neurons recorded simultaneously.

2 Materials and Methods

2.1 Procedures. Two adult male rhesus monkeys (*Macaca mulatta*; 7.7 and 8.8 kg) were trained to perform a strategy task. All procedures conformed to the *Guide for the Care and Use of Laboratory Animals* (National Research Council, 1996) and were approved by the National Institute of Mental Health (NIMH) Animal Care and Use Committee (IACUC).

2.2 Experimental Task. Each trial started with the presentation of a central cue (0.7° white circle) at the center of the screen, together with three potential targets (2.2° unfilled white squares), 14° left, up or right from the center (see Figure 1a). The monkeys had to fixate their gaze on the central cue (±7.5°) for 1.0 s. Then the central cue disappeared, and an instruction stimulus (IS) appeared instead for a variable delay of 1.0, 1.5, or 2.0 s (pseudo-randomly selected). Subsequently, the IS disappeared, serving as a "go" signal and instructing the monkeys to make a saccade toward one of the three targets within 2.0 s and fixate it (±6.7°) for 1.0 s. Then the targets were filled in white, and after 0.5 ms, a 0.1 ml of fluid reward was delivered if the response was correct. Correct responses depended on the type of trials: repeat-stay, change-shift, or second-chance trials. After a correct response, if the IS was the same as in the previous trial, the monkeys had to select the same spatial target as before (repeat-stay trial). If the IS was different from the previous one, the monkeys had to select one of the two alternative spatial targets (change-shift trial). In this case, the spatial target was rewarded 50% of the time. After an unrewarded response, the monkeys were presented with a second-chance trial in which they had the opportunity to select the alternative target to get a reward. Each IS was selected pseudorandomly from a set of three stimuli; 67% of the trials were change-shift trials, whereas 33% were repeat-stay trials.

2.3 Surgery and Histological Analysis. Aseptic techniques were used together with isofluorane anesthesia (1%–3%, to effect) to implant a 27 × 36 mm recording chamber over the exposed dura mater of the right frontal lobe of each monkey, together with head restraint devices.

Near the end of data collection, electrolytic lesions (15 A for 10 s, anodal current) were made at two depths. After about 10 days, the animals were deeply anesthetized and perfused with buffered formaldehyde. Later, the brains were removed, sectioned on a freezing microtome, mounted on glass slides, and stained for Nissl substance with thionin (Genovesio et al., 2005). The surface projections of the recording sites by the recovered electrolytic lesions and the marking pins inserted during perfusion were plotted.

Uncorrected Proof

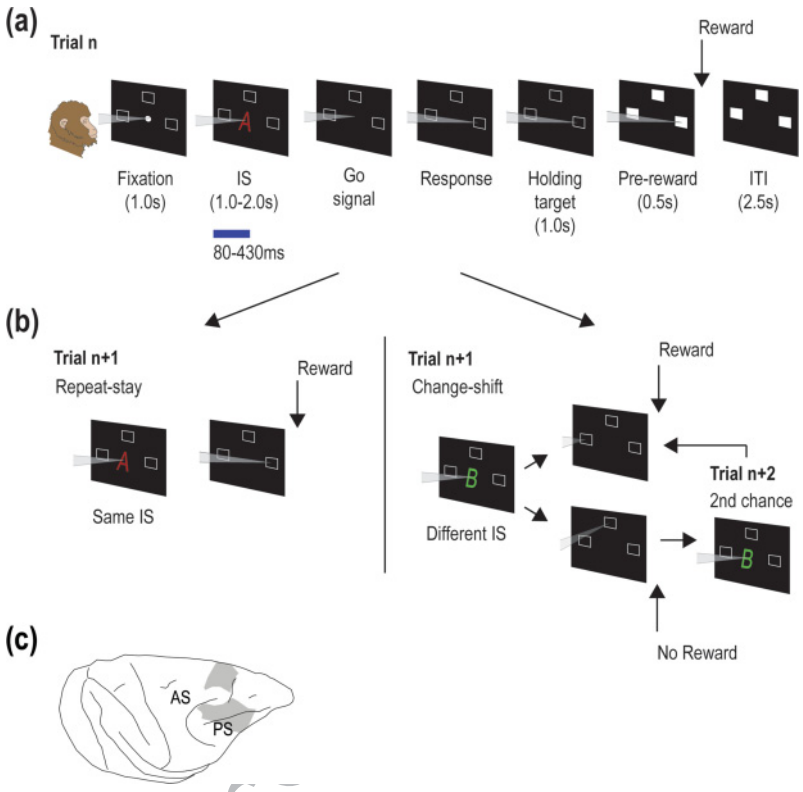

Figure 1: Strategy task and recording sites. (a) Temporal sequence of task events. Each trial started with a cue presented in the center of the screen, together with three potential targets. After a fixation time in which the monkeys were required to keep the gaze on the cue, an instruction stimulus (IS) appeared, and the monkeys were required to keep fixating their gaze until the IS disappeared, serving as a "go" signal. After the go signal, the monkeys responded with a saccade toward one of the three targets (empty squares). After maintaining the gaze on the selected target for 1.0 s (holding target period), the three targets were filled in, and after a subsequent period of 0.5 s (prereward period), the reward was provided if the response was correct. (b) Example trial cases. After each correct trial, a repeat-stay (left) or a change-shift trial (right) could follow. If the IS was the same as in the previous trial, the trial was a repeat-stay trial, and thus the monkeys had to choose the same target as before. In contrast, if the IS was different from the previous one, the trial was a change-shift trial, and the monkeys had to select one of the other two alternatives. Only one of them was rewarded randomly. If the response was unrewarded, the monkeys were presented with the same trial, called a second chance trial, in which they had to select the alternative target in order to get the reward. (c) Recording sites relative to sulcus landmarks: dorsolateral prefrontal cortex and dorsomedial prefrontal cortex (areas 6, 8, and 9). AS, arcuate sulcus: PS, principal sulcus.

2.4 Data Collection. Eye position was recorded with an infrared oculometer (Bouis Instruments, Karlsruhe, Germany). Single-cell potentials were isolated with quartz-insulated platinum-iridium electrodes (0.5–1.5 MΩ at 1 kHz) positioned by a 16-electrode drive assembly (Thomas Recording, Giessen, Germany) with 518 μm electrode spacing. Spikes were discriminated online using a multispike detector (Alpha-Omega Engineering, Nazareth, Israel) or a multichannel acquisition processor (Plexon, Dallas, TX) or offline using an off line sorter (Plexon). Only spike waveforms that clustered clearly in 3D principal component analysis, lacked interspike intervals less than 1 ms, had waveforms clearly coupled with others over time, and had stable and clearly differentiated waveforms were considered for further analyses. The recorded locations, shown in Figure 1c, were located in the dorsolateral prefrontal cortex and dorsomedial prefrontal cortex (areas 6, 8, and 9).

The neural database consisted of 1456 neurons (Genovesio et al., 2005; Genovesio, Brasted, & Wise, 2006; Genovesio, Tsujimoto, & Wise, 2006; Genovesio, Tsujimoto, & Wise, 2008; Marcos & Genovesio, 2016; Tsujimoto, Genovesio, & Wise, 2008). We further selected the neurons that had a mean activity of at least 1 spike/s within the 200 to 800 ms after IS presentation. With this additional constraint, we kept 887 neurons for further analyses.

2.5 Hidden Markov Model. A hidden Markov model (HMM) is a probabilistic model used to describe dynamic systems. Here, we applied it to the analysis of the activity of simultaneously recorded neurons (Abeles et al., 1995; Jones et al., 2007; Ponce-Alvarez et al., 2012; Rabiner, 1989). In particular, we used a discrete-time HMM that consisted of two concurrent stochastic processes: a discrete time Markov chain $(X_t)_{t\in\mathbb{N}}$ that identifies the state at step t and a sequence of observation symbols $(O_t)_{t\in\mathbb{N}}$ that represents the physical output of the system being modeled. An HMM is uniquely determined by the triplet (π, A, E). The initial state distribution $\pi = \{\pi(i)\}$ is an array of size n, where $\pi(i) = P(X_k = i)\ \forall k \leq i \leq n$. The state transition probability distribution $A = \{a_{i,j}\}$ is an $n \times n$ matrix (called a transition matrix) whose elements $a_{i,j}$ correspond to the probability of changing from state i to state j, $a_{i,j} = P(X_{t+1} = j \mid X_t = i)\ \forall t \in \mathbb{N}$. Finally, the observation symbol probability distribution $E = \{a_{i,j}\}$ is an $n \times m$ matrix (called an emission matrix) that represents the probability that in state i, the jth output is emitted, $e_{i,j} = P(O_t = j | X_t = i)\ \forall t \in \mathbb{N}$.

We define the likelihood of M, a generic HMM, as the probability that the observation sequence $O_k^T = \{O_1, O_2, \ldots, O_T\}$ has been generated by the model M: $P(O_k^T \mid M)$.

To calculate it, we used the forward-backward procedure (Rabiner, 1989). Moreover, we used the algorithm of Baum-Welch (Baum, Petrie, Soules, & Weiss, 1970), which iteratively estimates the initial parameters of the HMM using real observations in order to maximize the likelihood $P(O_k^T \mid M)$.

Uncorrected Proof

2.5.1 Choice of Initial Parameters. For each session of neurons analyzed, we built a matrix of observations $O = \{o_{i,j}\}$, where $o_{i,j}$ is the symbol observed in the jth millisecond of the ith trial. In that position, we use 1 if no neuron was active at that millisecond, or $n + 1$ if the nth neuron of the group emitted a spike. This matrix was used for training of the HMM in the Baum-Welch algorithm, together with the initial transition and emission matrices. These two matrices were generated randomly. Then the values of the diagonal elements of the transition matrix were set much larger than the others—all ≥ 0.9 (Seidemann et al., 1996). Similarly, we placed higher values on the first column of the emission matrix, respecting the fact that no neuron was active for more than 90% of the time on average in a trial. The algorithm is very sensitive to initial estimates, and for this reason, in the calculation of each HMM, we started with groups of five matrix pairs, a transition, and an emission matrix and iterated the Baum-Welch algorithm until convergence. For every pair of matrices, we then tested the likelihood of each trial and calculated the average; the pair of matrices that identified the model with a higher average likelihood is the one that was used to obtain the results reported in this letter. The vector of the initial distribution π by convention was set with probability 1 to start in the first state. A Bayesian information criterion (BIC) analysis (Schwartz, 1978) was performed within each session to estimate the optimal number of states. The likelihood of the HMM increases proportionally to the number of states, but by using an excessive number of states, the model can be overfitted. The BIC analysis penalizes the likelihood of the HMM by estimating its complexity. The BIC for an HMM with m states was calculated as follows (Ponce-Alvarez et al., 2012),

$$BIC(m) = \ln[P(O_k^T \mid M)] - \frac{\gamma_m}{2} \ln(T),$$

where O_k^T is the sequence of observations, $P(O_k^T \mid M)$ is the likelihood of the model M with m states obtained through the Baum-Welch algorithm, γ_m is the number of free parameters, and T is the length of the observed data. The number γ_m of free parameters for an HMM with m states and n neurons is given by $m \cdot n + m \cdot (m - 1)$, which takes into account the contribution of the emission and transmission matrices, respectively. Importantly, the initial state distribution π does not add any free parameters, as it is conventionally set to 1 for state 1, without being recalculated during the algorithm. To set the number of states of the HMM, we computed the BIC for each session by varying the number of hidden states from two to eight. The optimal number of states for each session was chosen as the one that resulted in the highest BIC. The number of optimal states varied from two to four (mean with an SEM of 2.3 ± 0.21).

2.5.2 Analyses. We built HMMs through the Baum-Welch algorithm from the matrix of observations and initial transition and emission

Uncorrected Proof

matrices. For each session, we built three different HMMs, one for each spatial position of the choice made by the subject: right, up, and left. To test the prediction power of the model, we eliminated a random trial from the data set and used the remaining trials to train the three HMMs. Next, we assessed the likelihood of each model for the eliminated trial and classified the trial as belonging to the model with the highest likelihood. Finally, we compared the direction identified by this model with the choice actually made by the subject in that trial. This was done once for each trial within a session. The accuracy of the model was calculated as the ratio between the trials in which these directions match and the total number of analyzed trials (Seidemann et al., 1996).

2.6 Classification Based on Peristimulus Time Histogram. We implemented a classification analysis based on the peristimulus time histogram (PSTH; Foffani & Moxon, 2004). First, we sorted the trials by the selected spatial position and then used the mean activity in the interval 80 ms to 430 ms from the IS onset as the predictor variable. As in the HMM, we considered groups of neurons that were registered simultaneously in the same session. In each interaction, we removed one trial (test trial) from the same condition from each neuron and calculated the neural response templates for each condition and neuron without considering that trial. Next, we calculated the Euclidean distance between the test trial and their corresponding neural response templates. The test trial was classified as belonging to the condition with the lowest sum of calculated distances. If the test trial actually belonged to the estimated condition, then the classification was considered correct. This procedure was repeated 10,000 times for each condition, and the reported accuracy corresponds to the proportion of correct classifications.

As a control, we also implemented a normalized PSTH-based method in which the mean activity for each neuron and condition was replaced by a z-score (z), which is defined by

$$z_{ik} = \frac{n_{ik} - \bar{n}_i}{\sigma_i},$$

where n_{ik} is the spike count of a neuron i and condition k, $\bar{n}_i$ is the mean spike count for neuron i, and σ_i is the standard deviation of the spike counts of neuron i.

2.7 Selection of Experimental Sessions. The sessions of neurons analyzed were chosen on the basis of three criteria: number of neurons, number of trials in which all neurons were registered simultaneously, and accuracy achieved by the models. We selected only sessions with at least five neurons, 45 trials, and in which the trials in each spatial condition were equally distributed or without notable differences. Between these sessions,

Uncorrected Proof

we chose the two groups in which the HMM model reached either a high or a low level of accuracy. We selected five sessions of each case. Only trials that performed correctly and belonged to the repeat-stay or change-shift strategy trials were used in the analyses. The time window in which the analyses were carried out started 80 ms after the stimulus presentation and lasted 350 ms. The total number of sessions included in the analyses is the minimum one that ensured a statistical power above 0.95–10 experimental sessions.

2.8 Neural Analyses.

2.8.1 Selectivity Index. For each individual neuron, three selectivity indexes were calculated as

$$\textit{Selectivity index} = \frac{\bar{N}_m - \bar{N}_n}{\bar{N}_m + \bar{N}_n},$$

where $\bar{N}$ is the mean spike count observed in the period of analysis, and m corresponds to one spatial position and n to an alternative one. Thus, for each neuron, we obtained three selectivity indexes: one for left/up, one for left/right, and one for up/right. We normalized the computed values so that the selectivity indexes are within the range of 0 and 1, where 0 corresponds to a lack of selectivity and 1 is the maximum selectivity that can be observed.

2.8.2 Spatial Selectivity. We calculated the statistical significance of the neural activity for the different spatial choices using a one-way ANOVA with the neural activity in the 80 to 430 ms period after IS onset as the dependent variable and left, up, and right choices as factors.

3 Results

Two monkeys were trained to perform a strategy task (Genovesio et al., 2005; see Figures 1a and 1b), while a variable number of neurons were recorded from their prefrontal cortex (see Figure 1c). Based on an instructed stimulus (IS), they had to select one of the three spatial positions: the target position that was previously selected if the IS was the same as in the previous trial (repeat-stay trials) or one of the other two spatial positions otherwise (change-shift trials). Our analyses focused on the first part of the IS period (80–430 ms after IS onset) because all of the information required to make the decision was available at that point. In the selected sessions, the two monkeys performed the task with high accuracy: 97% of trials were correctly performed for monkey 1 and 90% for monkey 2.

We used the activity of simultaneously recorded neurons to compare how well an HMM and a PSTH-based method could classify single trials

Uncorrected Proof

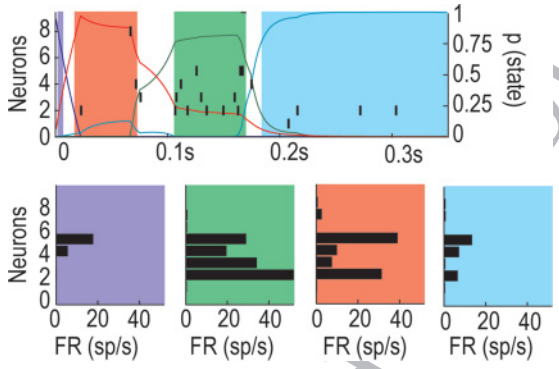

Figure 2: Example of HMM states for an experimental session of eight simultaneously recorded neurons. One trial example of the HMM estimation in the interval between 80 and 430 ms after IS onset. Four states (represented by different colors) are estimated from the firing rate activity of the neurons. (Top) Raster plot showing the activity of each neuron during a trial and the probability of being in each specific state. (Bottom) Mean activity profile of the neurons in each state.

into trials with right, up, or left target choices. The HMM uses the information from all simultaneously recorded neurons within an experimental session and categorizes their activity into a small number of hidden states. The activity of each single trial can then be described using those hidden states. An example of the probability of being in a specific state at each time point during the 80 to 430 ms after IS onset in a single trial is shown in Figure 2. In this case, eight neurons were recorded when monkey 1 performed the strategy task. The different states are represented with colors and correspond to a specific pattern of activation of the neurons (see the bottom panel of Figure 2). At each time point, the neural activity was better described by one unique state.

Using the neural activity in the period of 80 to 430 ms after IS onset, both the HMM and the PSTH-based method could predict the future choice with an accuracy that was notably higher than chance (about 33%). Importantly, at that time, the monkeys had all of the information needed to make their decision, but they were still not allowed to make a saccadic movement toward the choice. The prediction power of the HMM was always higher than that obtained with the PSTH-based method (see Figure 3a and section 2) for experimental sessions with high and low levels of accuracy of the HMM. Thus, even for sessions in which the HMM did not reach high levels of prediction, the accuracy of the PSTH-based method was always at a lower range. The performance of both methods was significantly different ($p < 0.01$, Wilcoxon's matched-pairs signed-rank test), with a mean value of 58% of correctly classified trials for the HMM and a mean value of 52%

Uncorrected

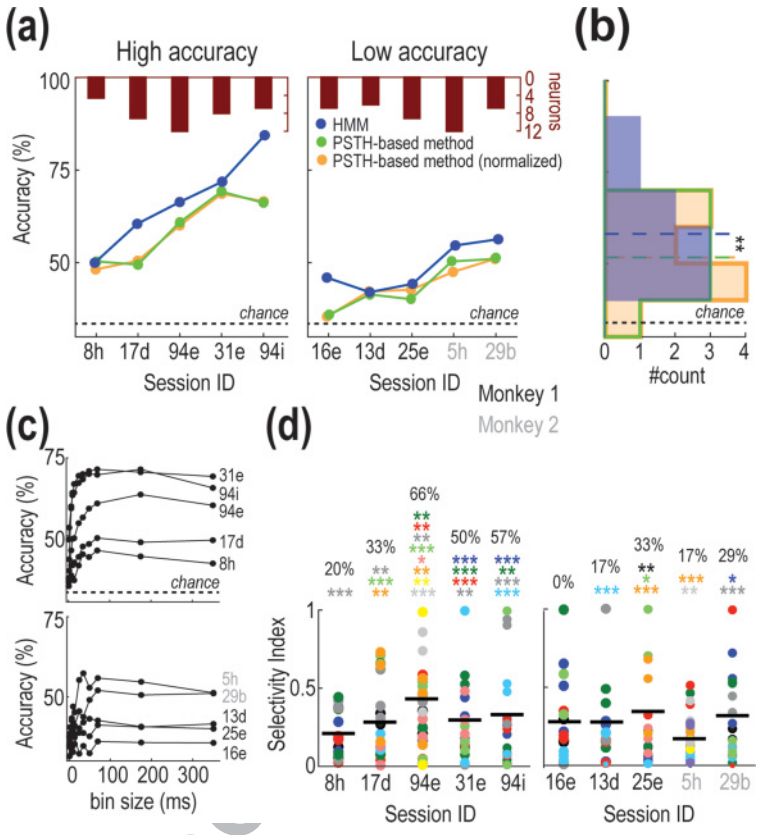

Figure 3: Accuracy of models and neural selectivity. (a) Percentage of correctly classified trials for the HMM (blue), the PSTH-based method (green), and the PSTH-based method with normalization (z-score; orange) for sessions with high (left) and low (right) accuracy of the HMM. (b) Distribution of accuracy for HMM, the PSTH-based method, and the normalized PSTH-based method. The accuracy of the HMM model is significantly higher than the one obtained with the PSTH-based method and the one obtained with the normalized PSTH-based method ($^{**}p < 0.01$, Wilcoxon's matched-pairs signed-rank test). Dashed lines indicate the mean of the distributions. (c) Accuracy of the PSTH-based model for different bin sizes and experimental sessions. (d) Selectivity index of individual neurons within each session. Each colored dot represents the selectivity index of one neuron. The three selectivity indexes from one neuron are represented with the same color. The black line indicates the mean selectivity index of the session. The asterisks indicate whether the neuron represented by the color shows a significant modulation of its activity related to the spatial targets (see section 2: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, One-way ANOVA). The percentages above the colored asterisks show the proportion of neurons with significant spatial selectivity within each session.

for the PSTH-based method (see Figure 3b). To control for possible changes in the accuracy of the PSTH-based method related to the bin size used in the analysis, we calculated the accuracy of the PSTH-based method varying it. The results show that the bin size used (350 ms) provided roughly the best accuracy that could be obtained within the task period (see Figure 3c). Thus, the HMM outperformed the PSTH-based method, and, hence, it could better extract the information on the future choice from the neural activity. To control that the difference in accuracy between the HMM and the PSTH-based method was not just related to differences on how the two methods deal with the variability of the data, we performed a control with the PSTH-based method. In this case, instead of using the mean activity of the neurons for classification, we used the *z*-score. This is a normalized measure of the mean activity that takes into account the variability of the spike counts (see section 2). The accuracy was not significantly different from that obtained when the mean activity was used ($p = 0.70$, Wilcoxon's matched-pairs signed-rank test), and it was significantly lower than the one obtained with the HMM ($p < 0.01$, Wilcoxon's matched-pairs signed-rank test). This confirms that the difference in accuracy between the models was not simply due to the variability of the data.

Furthermore, we tested whether the difference in accuracy within each method was related to the number of neurons or to the neural selectivity within each session (see section 2). We did not observe any correlation between the number of neurons recorded within an experimental session and the accuracy achieved by each method (see the top panel of Figure 3a). Indeed, when we compared the accuracy between two sessions with the same number of recorded neurons ($n = 7$; sessions 94i and 16e), the performance reached within each method was notably different: the accuracy varied from 55% to 90% for the HMM and from 36% to 66% for the PSTH-based method (see Figure 3a). Next, we calculated the individual selectivity indexes for each neuron and checked the proportion of neurons with a significant difference in activity related to the different spatial choices, that is, the proportion of neurons coding the future spatial choice. Experimental sessions with a higher proportion resulted in better accuracy for both the HMM and PSTH-based method (see Figures 3a and 3d). For instance, if we split the experimental sessions into two groups depending on whether they contained a proportion of neurons with spatial selectivity higher or lower than 30%, the difference between the two groups was remarkable for both methods although not significant ($p > 0.05$, Mann-Whitney U-test). In particular, the mean accuracy of the HMM was 65% in the former case and 50% in the latter, whereas the PSTH-based method showed a performance of 57% and 46%, respectively. Hence, the results indicate a possible link between the accuracy of the models and the number of recorded neurons coding the future spatial goal within a session. The specific value of the mean selectivity index within a session was, however, not indicative of the accuracy reached by the methods. Therefore, in the sessions considered, the

Uncorrected Proof

accuracy of the models was not related to the number of recorded neurons per se, but it seems to relate on the proportion of spatial selective neurons. This means that sessions with a lower number of neurons could result in higher accuracy compared to others with a greater number of neurons if they contain a higher proportion of selective neurons.

4 Discussion

We compared two methods to classify neural data into different categories. Our neural data consisted of the activity of some neurons recorded from the prefrontal cortex while two monkeys performed a strategy task. In this task, the monkeys had to choose one of three spatial targets based on the IS displayed on the screen and on the previous IS and their previous choice. We focused our analyses on the period after IS onset and therefore on a period in which all of the information required to make a decision was available. We investigated the accuracy of an HMM and PSTH-based method. The HMM uses the data from all neurons and all trials recorded within a session to infer the hidden states that describe the dynamics of the neural activity. In our task, three models were inferred from each session: one for each group of trials with left, up, or right as selected targets. The estimated models were then used to classify individual trials into the groups that most likely described its dynamics. In contrast, the PSTH-based method classified individual trials solely using the mean activity observed for each spatial choice and neuron and then combined the estimation from each individual neuron to improve the predictive power. In our study, both methods were able to predict the future choice of the monkeys with an accuracy higher than chance. Importantly, we show that the HMM was always more accurate than the PSTH-based method, even for experimental sessions in which the prediction accuracy of the HMM was at a low level. Interestingly, the accuracy of both methods tended to be correlated with the proportion of spatial selective neurons within a session.

Decoding or classification methods for neuronal signals have been of particular interest. One of the reasons is the development of brain-machine interfaces (BMIs; Lebedev, 2014a; Lebedev & Nicolelis, 2006) and therefore the need for accurate and versatile methods to interpret neural activity and relate it to behavior. Large-scale recordings are fundamental to achieving high levels of performance for BMIs (Chapin, 2004; Nicolelis & Lebedev, 2009; Schwarz et al., 2014), as the probability of having neurons with selective activity for a behavioral parameter increases with an increasing number of neurons in the sample (Lebedev, 2014b). Thanks to advances in the decoding and recording techniques, BMIs have been successfully used by monkeys to control the movements of two avatar arms simultaneously (Ifft, Shokur, Li, Lebedev, & Nicolelis, 2013).

HMMs have been used to detect ongoing sequences of states in the neural activity of gustatory neurons in awake rats that emerged after taste

Uncorrected Proof

delivery (Jones et al., 2007) or spontaneously during spontaneous ongoing activity (Mazzucato et al., 2015). Moreover, Mazzucato et al. (2015) showed that most of the neurons exhibited multistability during spontaneous ongoing activity (three or more different firing rates across states) and bistability during taste-evoked activity (a maximum of two different firing rates across states). One hypothesis to account for such a difference in the number of states between ongoing and evoked activity is that ongoing activity serves as a repertoire of representations that is then sampled during evoked activity (Arieli, Sterkin, Grinvald, & Aertsen, 1996; Kenet, Bibitchkov, Tsodyks, Grinvald, & Arieli, 2003; Luczak, Barthó, & Harris, 2009). An alternative explanation is that ongoing activity serves as a Bayesian prior that dynamically adjusts to the statistics of the external stimuli and slowly develops toward the same dynamics observed during evoked responses (Berkes, Orbán, Lengyel, & Fiser, 2011; Fiser, Chiu, & Weliky, 2004).

The PSTH-based method relies on the information carried by the neurons in their mean response. Most of the previous studies have focused on this signal that, for instance, in the somatosensory area describes more than 80% of the information about stimulus location (Panzeri, Petersen, Schultz, Lebedev, & Diamond, 2001). However, reducing the signal carried by many neurons to a mean firing rate can obscure the heterogeneous nature of cell assemblies (Huang & Zeng, 2013; Mattia et al., 2013). Moreover, not only the mean response but also its variability carry additional information such as the recent experience (Marcos et al., 2013) or stimulus onset (Churchland et al., 2010). Recently, new analytical techniques that account for such cell type variety have been applied to measure multineuron activity. For instance, correlations across neurons or machine learning methods such as HMMs are useful to understand the state and connectivity of a network and relate the neural activity to behavior (Cortes & Vapnik, 1995; Doiron, Litwin-Kumar, Rosenbaum, Ocker, & Josić, 2016; Helias, Tetzlaff, & Diesmann, 2014; Pagan, Urban, Wohl, & Rust, 2013; Pernice, Staude, Cardanobile, & Rotter, 2011; Raposo, Kaufman, & Churchland, 2014; Rust & Dicarlo, 2010; Trousdale, Hu, Shea-Brown, & Josić, 2012). Our study shows that although the mean firing rate of neurons alone could be used to classify trials with relatively high accuracy, the specific dynamics of the population could significantly improve accuracy.

Our study shows that with groups of up to 12 neurons, trials were better classified using HMMs than solely using the mean individual firing rates. New recording techniques allow for the simultaneous recording of much larger numbers of neurons, in some cases up to hundreds of neurons and across several brain regions (Dotson, Hoffman, Goodell, & Gray, 2017; Nicolelis et al., 2003). The PSTH-based method has been shown to be very sensitive to the number of neurons used, with an increase in accuracy with an increase in the number of neurons considered (Cirillo, Ferrucci, Marcos, Ferraina, & Genovesio, 2018; Falcone, Cirillo, Ferraina, & Genovesio, 2017; Genovesio, Brasted, et al., 2006; Lebedev, Messinger, Kralik, & Wise, 2004;

Uncorrected Proof

Marcos & Genovesio, 2016; Marcos, Nougaret, Tsujimoto, & Genovesio, 2018; Marcos, Tsujimoto, & Genovesio, 2017). Thus, at least two open questions arise from these new kinds of data sets: (1) Would the two methods reach similar levels of performance for large numbers of neurons, or, even in those cases, would the HMMs always be capable of extracting further information than that obtain from the mean response? and (2) Would the combination of data simultaneously recorded from different areas improve accuracy similarly for both methods, or would the HMM be more sensitive to the dynamics of different areas? It is also important to assess whether the increase in the number of neurons makes the use of models as complex as the HMM unreliable. Thanks to the current use of the new recording techniques for experimental purposes, these questions will soon be answered.

Uncorrected Proof

Acknowledgments

We thank Alex Cummings for the histological material and Andrew Mitz and Steve Wise for their numerous contributions. This research was supported by the Intramural Research Program of the National Institutes of Health, National Institute of Mental Health (NIMH; Z01MH01092–28), and partially supported by Italian grant Avvio alla ricerca di Ateneo 2016 and a Juan de la Cierva–incorporación scholarship (IJCI-2016-27864—Spanish Ministry of Science, Innovation and Universities).

References

Abeles, M., Bergman, H., Gat, I., Meilijson, I., Seidemann, E., Tishby, N., & Vaadia, E. (1995). Cortical activity flips among quasi-stationary states. *Proceedings of the National Academy of Sciences of the United States of America*, *92*, 8616–8620.

Arieli, A., Sterkin, A., Grinvald, A., & Aertsen, A. (1996). Dynamics of ongoing activity: Explanation of the large variability in evoked cortical responses. *Science*, *273*, 1868–1871.

Baum, L. E., Petrie, T., Soules, G., & Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Annals of Mathematical Statistics*, *41*, 164–171.

Berkes, P., Orbán, G., Lengyel, M., & Fiser, J. (2011). Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science*, *331*, 83–87.

Chapin, J. K. (2004). Using multi-neuron population recordings for neural prosthetics. *Nature Neuroscience*, *7*, 452–455.

Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., . . . Shenoy, K. V. (2010). Stimulus onset quenches neural variability: A widespread cortical phenomenon. *Nature Neuroscience*, *13*, 311–316.

Cirillo, R., Ferrucci, L., Marcos, E., Ferraina, S., & Genovesio, A. (2018). Coding of self and other's future choices in dorsal premotor cortex during social interaction. *Cell Reports*, *24*, 1679–1686.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*, 273–297.

Doiron, B., Litwin-Kumar, A., Rosenbaum, R., Ocker, G. K., & Josić, K. (2016). The mechanics of state-dependent neural correlations. *Nature Neuroscience*, *19*, 383–393.

Dotson, N. M., Hoffman, S. J., Goodell, B., & Gray, C. M. (2017). A large-scale semi-chronic microdrive recording system for non-human primates. *Neuron*, *96*, 769–782.

Escola, S., Fontanini, A., Katz, D., & Paninski, L. (2011). Hidden Markov models for the stimulus-response relationships of multistate neural systems. *Neural Computation*, *23*, 1071–1132.

Falcone, R., Cirillo, R., Ferraina, S., & Genovesio, A. (2017). Neural activity in macaque medial frontal cortex represents other's choices. *Scientific Reports*, *7*, 12663.

Fiser, J., Chiu, C., & Weliky, M. (2004). Small modulation of ongoing cortical dynamics by sensory input during natural vision. *Nature*, *431*, 573–578.

Foffani, G., & Moxon, K. A. (2004). PSTH-based classification of sensory stimuli using ensembles of single neurons. *Journal of Neuroscience Methods*, *135*, 107–120.

Genovesio, A., Brasted, P. J., Mitz, A. R., & Wise, S. P. (2005). Prefrontal cortex activity related to abstract response strategies. *Neuron*, *47*, 307–320.

Genovesio, A., Brasted, P. J., & Wise, S. P. (2006). Representation of future and previous spatial goals by separate neural populations in prefrontal cortex. *Journal of Neuroscience*, *26*, 7305–7316.

Genovesio, A., Tsujimoto, S., & Wise, S. P. (2006). Neuronal activity related to elapsed time in prefrontal cortex. *Journal of Neurophysiology*, *95*, 3281–3285.

Genovesio, A., Tsujimoto, S., & Wise, S. P. (2008). Encoding problem-solving strategies in prefrontal cortex: Activity during strategic errors. *European Journal of Neuroscience*, *27*, 984–990.

Ghazanfar, A. A., Stambaugh, C. R., & Nicolelis, M. A. (2000). Encoding of tactile stimulus location by somatosensory thalamocortical ensembles. *Journal of Neuroscience*, *20*, 3761–3775.

Gochin, P. M., Colombo, M., Dorfman, G. A., Gerstein, G. L., & Gross, C. G. (1994). Neural ensemble coding in inferior temporal cortex. *Journal of Neurophysiology*, *71*, 2325–2337.

Helias, M., Tetzlaff, T., & Diesmann, M. (2014). The correlation structure of local neuronal networks intrinsically results from recurrent dynamics. *PLoS Computational Biology*, *10*, e1003428.

Huang, Z. J., & Zeng, H. (2013). Genetic approaches to neural circuits in the mouse. *Annual Review of Neuroscience*, *36*, 183–215.

Ifft, P. J., Shokur, S., Li, Z., Lebedev, M. A., & Nicolelis, M. A. (2013). A brain-machine interface enables bimanual arm movements in monkeys. *Science Translational Medicine*, *5*, 210ra154.

Jones, L. M., Fontanini, A., Sadacca, B. F., Miller, P., & Katz, D. B. (2007). Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 18772–18777.

Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A., & Arieli, A. (2003). Spontaneously emerging cortical representations of visual attributes. *Nature*, *425*, 954–956.

Uncorrected Proof

Lebedev, M. A. (2014a). Brain-machine interfaces: An overview. *Translational Neuroscience*, *5*, 99–110.

Lebedev, M. A. (2014b). How to read neuron-dropping curves? *Frontiers in Systems Neuroscience*, *8*, 102. doi:10.3389/fnsys.2014.00102

Lebedev, M. A., Messinger, A., Kralik, J. D., & Wise, S. P. (2004). Representation of attended versus remembered locations in prefrontal cortex. *PLoS Biology*, *2*, e365.

Lebedev, M. A., & Nicolelis, M. A. L. (2006). Brain–machine interfaces: Past, present and future. *Trends in Neurosciences*, *29*, 536–546.

Luczak, A., Barthó, P., & Harris, K. D. (2009). Spontaneous events outline the realm of possible sensory responses in neocortical populations. *Neuron*, *62*, 413–425.

Marcos, E., & Genovesio, A. (2016). Determining monkey free choice long before the choice is made: The principal role of prefrontal neurons involved in both decision and motor processes. *Frontiers in Neural Circuits*, *10*, 75. doi:10.3389/fncir.2016.00075

Marcos, E., Nougaret, S., Tsujimoto, S., & Genovesio, A. (2018). Outcome modulation across tasks in the primate dorsolateral prefrontal cortex. *Neuroscience*, *371*, 96–105.

Marcos, E., Pani, P., Brunamonti, E., Deco, G., Ferraina, S., & Verschure, P. (2013). Neural variability in premotor cortex is modulated by trial history and predicts behavioral performance. *Neuron*, *78*, 249–255.

Marcos, E., Tsujimoto, S., & Genovesio, A. (2017). Independent coding of absolute duration and distance magnitudes in the prefrontal cortex. *Journal of Neurophysiology*, *117*, 195–203.

Mattia, M., Pani, P., Mirabella, G., Costa, S., Del Giudice, P., & Ferraina, S. (2013). Heterogeneous attractor cell assemblies for motor planning in premotor cortex. *Journal of Neuroscience*, *33*, 11155–11168.

Mazzucato, L., Fontanini, A., & La Camera, G. (2015). Dynamics of multistable states during ongoing and evoked cortical activity. *Journal of Neuroscience*, *35*, 8214–8231.

National Research Council. (1996). *Guide for the care and use of laboratory animals*. Washington, DC: National Academy Press.

Nicolelis, M. A. L., Dimitrov, D., Carmena, J. M., Crist, R., Lehew, G., Kralik, J. D., & Wise, S. P. (2003). Chronic, multisite, multielectrode recordings in macaque monkeys. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 11041–11046.

Nicolelis, M. A. L., Ghazanfar, A. A., Stambaugh, C. R., Oliveira, L. M., Laubach, M., Chapin, J. K., . . . Kaas, J. H. (1998). Simultaneous encoding of tactile information by three primate cortical areas. *Nature Neuroscience*, *1*, 621–630.

Nicolelis, M. A. L., & Lebedev, M. A. (2009). Principles of neural ensemble physiology underlying the operation of brain–machine interfaces. *Nature Reviews Neuroscience*, *10*, 530–540.

Nicolelis, M. A. L., Lin, R. C., & Chapin, J. K. (1997). Neonatal whisker removal reduces the discrimination of tactile stimuli by thalamic ensembles in adult rats. *Journal of Neurophysiology*, *78*, 1691–1706.

Pagan, M., Urban, L. S., Wohl, M. P., & Rust, N. C. (2013). Signals in inferotemporal and perirhinal cortex suggest an untangling of visual target information. *Nature Neuroscience*, *16*, 1132–1139.

Uncorrected Proof

Panzeri, S., Petersen, R. S., Schultz, S. R., Lebedev, M., & Diamond, M. E. (2001). The role of spike timing in the coding of stimulus location in rat somatosensory cortex. *Neuron*, *29*, 769–777.

Pernice, V., Staude, B., Cardanobile, S., & Rotter, S. (2011). How structure determines correlations in neuronal networks. *PLoS Computational Biology*, *7*, e1002059.

Ponce-Alvarez, A., Kilavik, B., & Riehle, A. (2008). *Dynamic sequences of states in ensembles of motor cortical neurons.* Paper presented at the Deuxième conférence française de Neurosciences Computationnelles, "Neurocomp08," Marseille, France.

Ponce-Alvarez, A., Nácher, V., Luna, R., Riehle, A., & Romo, R. (2012). Dynamics of cortical neuronal ensembles transit from decision making to storage for later report. *Journal of Neuroscience*, *32*, 11956–11969.

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*, 257–286.

Raposo, D., Kaufman, M. T., & Churchland, A. K. (2014). A category-free neural population supports evolving demands during decision-making. *Nature Neuroscience*, *17*, 1784–1792.

Rolls, E. T., Treves, A., Robertson, R. G., Georges-François, P., & Panzeri, S. (1998). Information about spatial view in an ensemble of primate hippocampal cells. *Journal of Neurophysiology*, *79*, 1797–1813.

Rolls, E. T., Treves, A., & Tovee, M. J. (1997). The representational capacity of the distributed encoding of information provided by populations of neurons in primate temporal visual cortex. *Experimental Brain Research*, *114*, 149–162.

Rust, N. C., & Dicarlo, J. J. (2010). Selectivity and tolerance ("invariance") both increase as visual information propagates from cortical area V4 to IT. *Journal of Neuroscience*, *30*, 12978–12995.

Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*, 461–464.

Schwarz, D. A., Lebedev, M. A., Hanson, T. L., Dimitrov, D. F., Lehew, G., Meloy, J., . . . Nicolelis, M. A. L. (2014). Chronic, wireless recordings of large-scale brain activity in freely moving rhesus monkeys. *Nature Methods*, *11*, 670–676.

Seidemann, E., Meilijson, I., Abeles, M., Bergman, H., & Vaadia, E. (1996). Simultaneously recorded single units in the frontal cortex go through sequences of discrete and stable states in monkeys performing a delayed localization task. *Journal of Neuroscience*, *16*, 752–768.

Trousdale, J., Hu, Y., Shea-Brown, E., & Josić, K. (2012). Impact of network structure and cellular response on spike time correlations. *PLoS Computational Biology*, *8*, e1002408.

Tsujimoto, S., Genovesio, A., & Wise, S. P. (2008). Transient neuronal correlations underlying goal selection and maintenance in prefrontal cortex. *Cerebral Cortex*, *18*, 2748–2761.

Uncorrected Proof

Received October 1, 2018; accepted May 1, 2019.