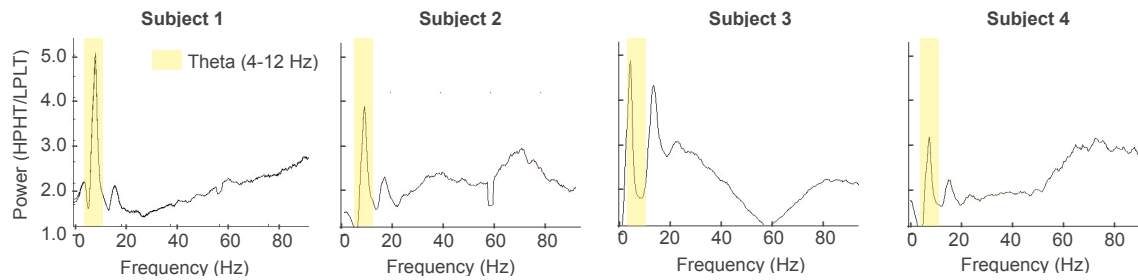


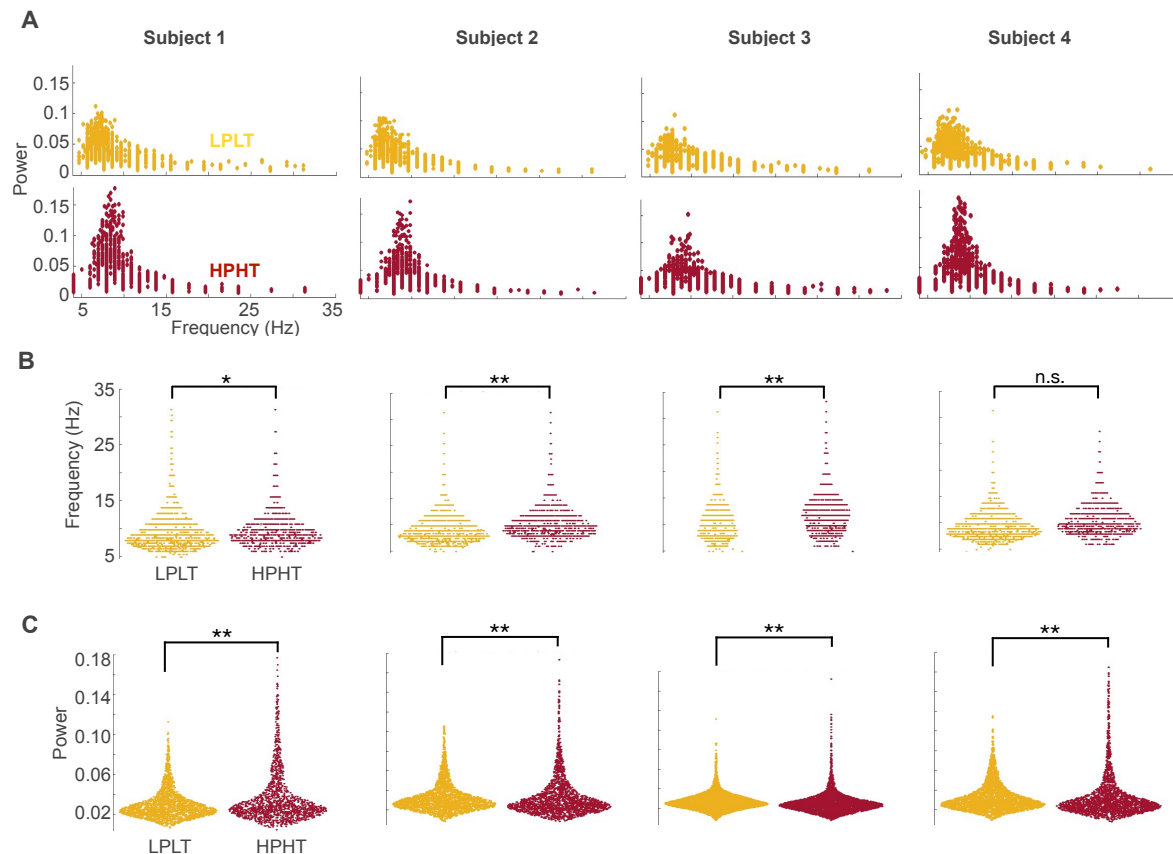
SUPPLEMENTARY FIGURES

SUPPLEMENTARY FIGURE 1



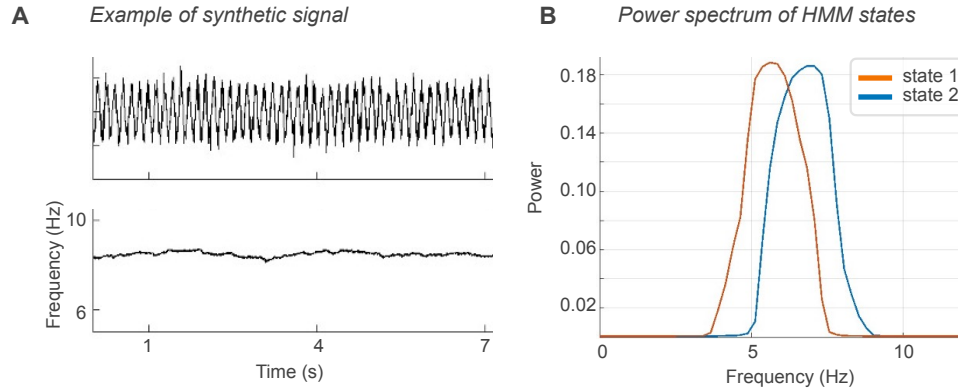
Power difference between HPHT and LPLT states. Here plotted the ratio between the two states power spectrum (HPHT / LPLT), to better visualise the difference in power between the two states proportionally to the overall broadband power spectrum.

SUPPLEMENTARY FIGURE 2



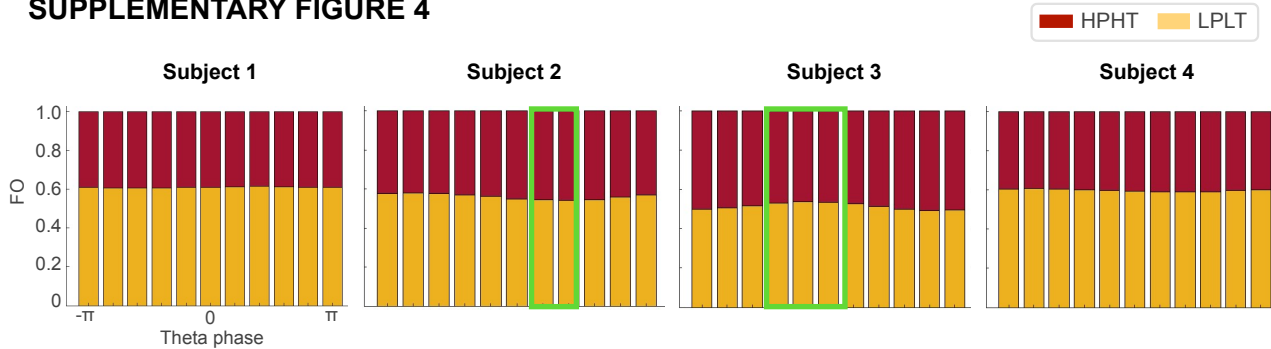
Statistical characterisation of the states. **A.** Power spectrum distributions of the states. For each state visit longer than 50 timepoints, the max power and its corresponding frequency is plotted, for each state, for each subject. **B.** Frequency distribution of state visits. The frequency of each state visit longer than 50 points is plotted as a distribution, per state and per subject. **C.** Power distribution of state visits. Like in **B**, but for the max power of each state visit counted. The two state distributions (frequency distributions in **B** and power distributions in **C**) are then statistically compared by means of the Whitney-Mann U test (for each subject). Here represented with * a p value <0.05, and with ** a p value < 0.0001.

SUPPLEMENTARY FIGURE 3



Testing HMM sensitivity to frequency. **A.** *Example of synthetic signal* used to test the HMM sensitivity on frequency (top panel), together with its instantaneous frequency (bottom panel). The signal is simulated as a simple sinusoid with amplitude $a = 5$, Gaussian noise with standard deviation $= 1$, and instantaneous frequency determined as a semi random walk bounded between 7 and 9 Hz. **B.** *Power spectrum of HMM states.* Here, the HMM settings are identical to those in our experiment: $K=2$, $L=7$. $\delta = 1\text{bn}$. The power spectrum of the states shows a peak frequency difference of about 1Hz.

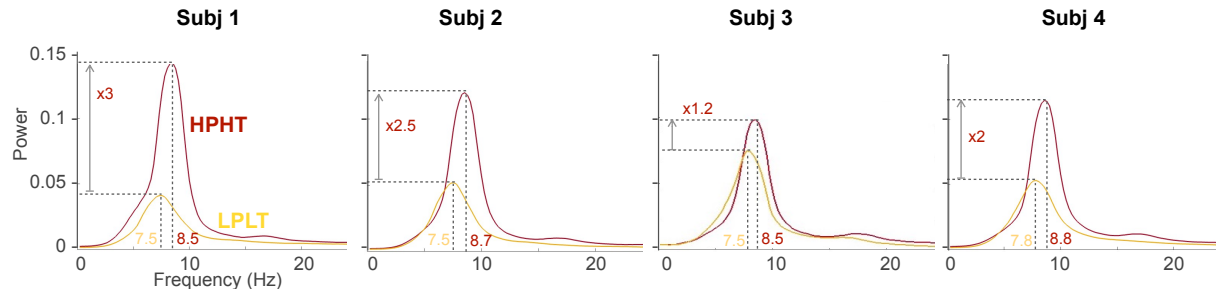
SUPPLEMENTARY FIGURE 4



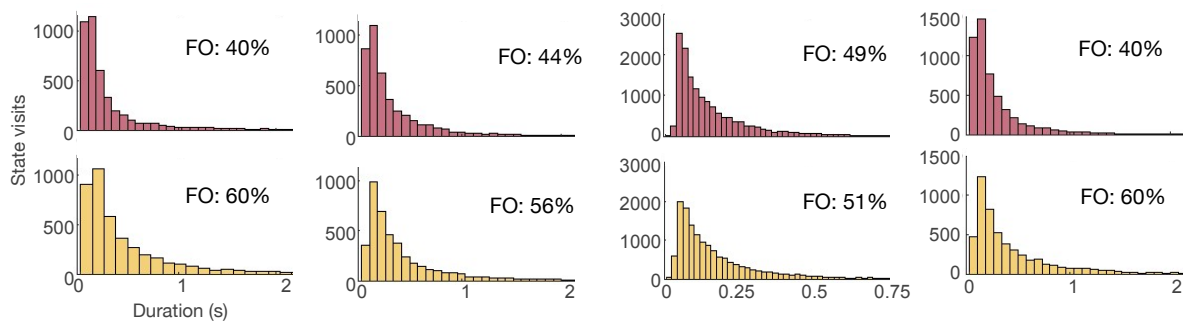
States fractional occupancy (FO) as a function of theta phase. For each theta phase bin of $1/5 \pi$, the states FO is computed and plotted cumulatively. Each bin's distribution has been tested with respect to each others, and the green frame shows the bins where the distribution is significantly different from the others ($p < 0.001$), by means of the Whitney-Mann U test. This serves as a visual representation of the states distribution around theta phase

SUPPLEMENTARY FIGURE 5

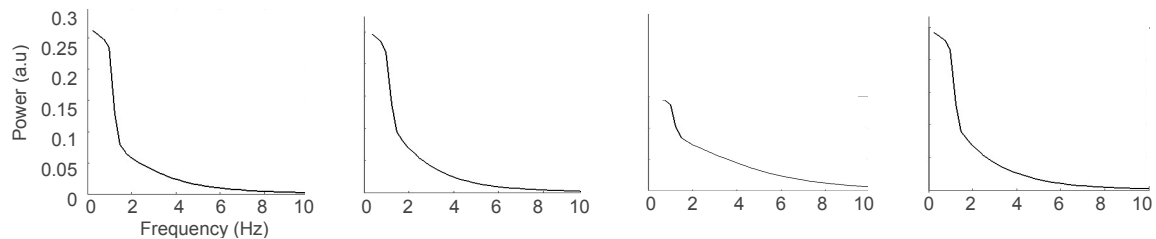
A States power spectra



B States lifetime and fractional occupancy

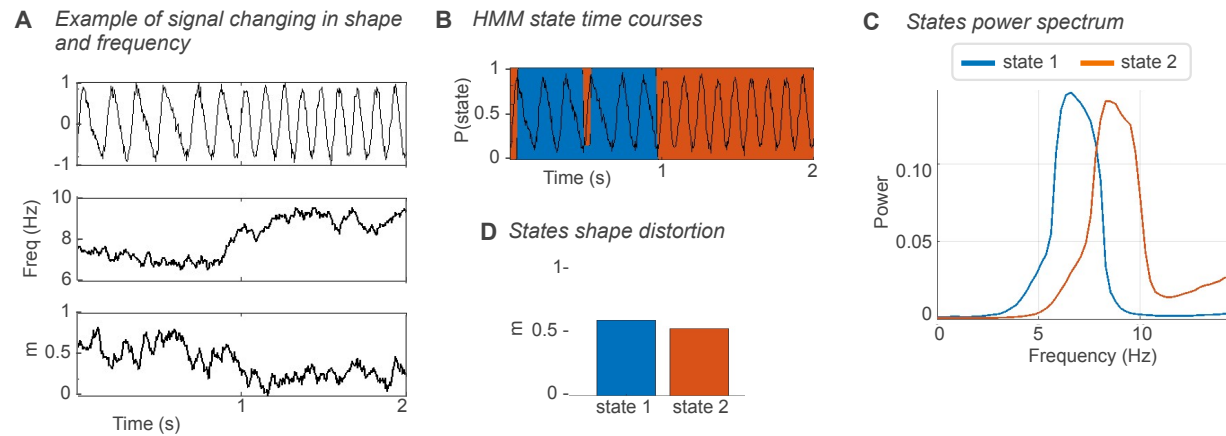


C States periodicity



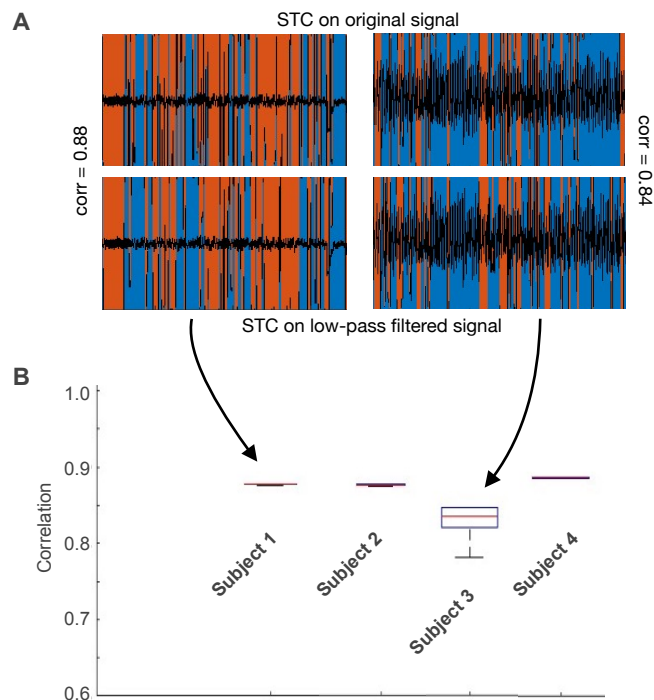
States characterisation per subject. A. States power spectra. Frequency and power content of each state for each subject. The HMM consistently finds a high-power, higher theta (HPHT, dark red) and a low-power, lower theta (LPLT, yellow) state. **B. States lifetime and fractional occupancy.** The count of the time length of each state occurrence, per state and per subject. Fractional occupancy (FO) accounts for the overall presence of a state in the entire dataset. Both FO and states lifetime are comparable across states. **C. States periodicity.** The period of each state activation, per subject. States reoccur typically with a period around 1-2 Hz, meaning each state occurs on average once to twice per second.

SUPPLEMENTARY FIGURE 6



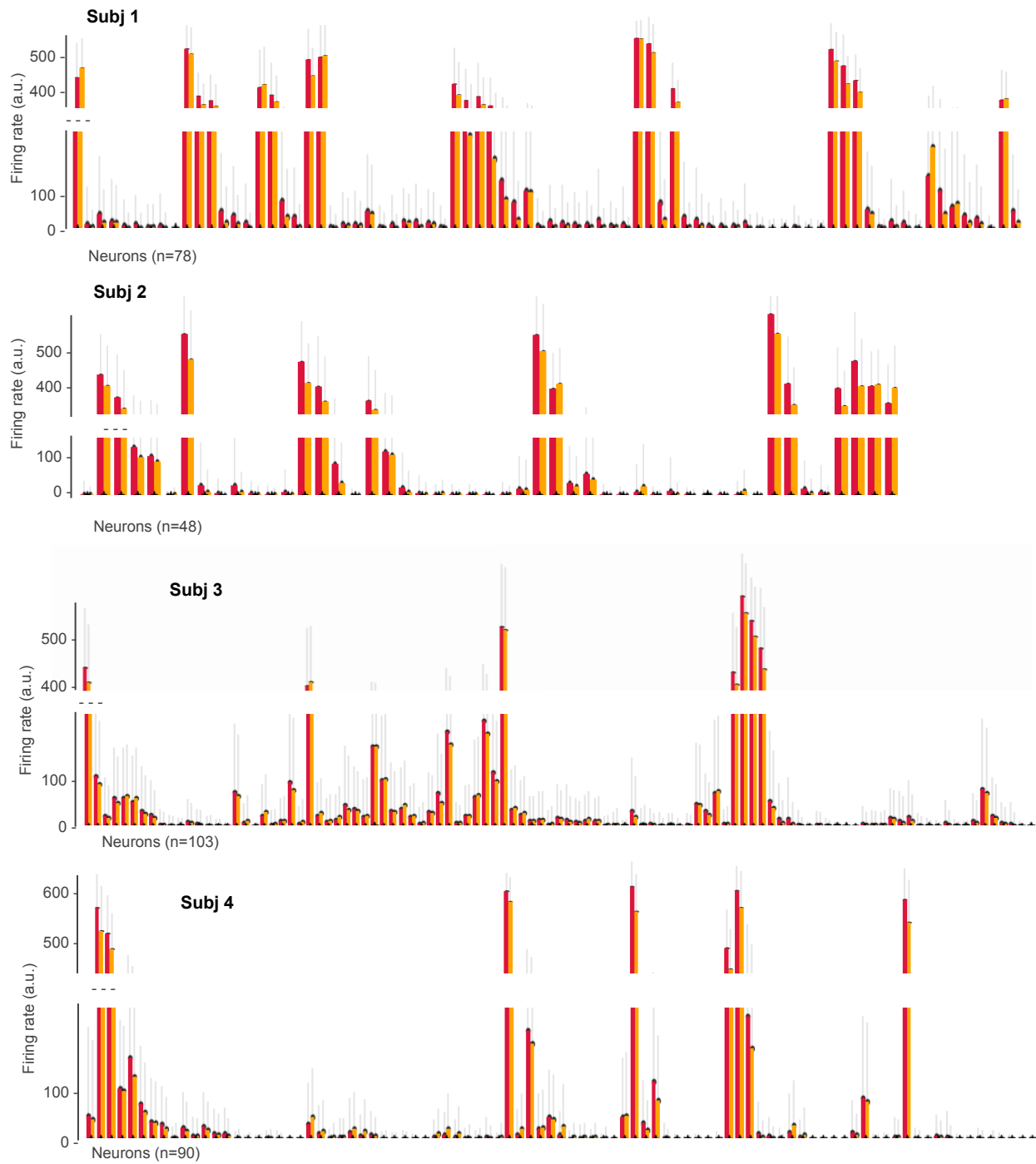
HMM analysis on synthetic data changing in shape and frequency independently. **A.** Example of signal changing in shape and frequency. The synthetic signal (top panel), with its instantaneous frequency (middle panel) and the instantaneous shape distortion coefficient m (bottom panel). **B.** Example of the HMM states time courses on the signal plotted in **A**. **C.** Power spectrum of the HMM states found on the signal in **A**. The states show different frequency peak. **D.** State shape distortion. The average distortion of the signal captured by each state. Each state captures a certain amount of distorted signal, meaning that the HMM state decomposition is driven mostly by frequency.

SUPPLEMENTARY FIGURE 7



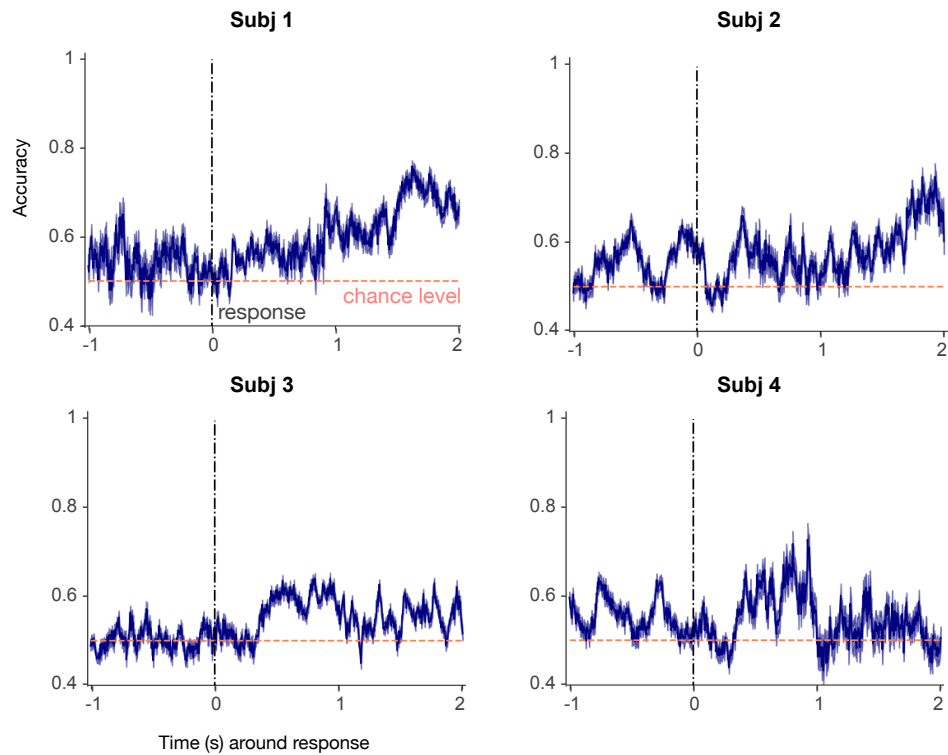
Correlation between the HMM run on the minimally pre-processed LFP signal and the HMM run on the low-pass filtered LFP signal. **A.** Two examples of HMM state time courses (STC) on the minimally preprocessed LFP data (top row) and on the low-pass filtered data (bottom row), and their correlation for Subject 1 (highly correlated) and for Subject 3 (relatively less highly correlated). **B.** Correlation between HMM runs on the minimally pre-processed signal and on the low-pass filtered signal, over 10 different runs. The HMM on the minimally pre-processed is run every time with random initial conditions. The HMM on the low-pass filtered signal is run by setting as initial conditions the HMM on the minimally preprocessed signal (to control for the inherent variability of stochastic inference of HMM). The red line in the box shows the mean, the box indicates the variance, and the eventual extra line indicates outliers.

SUPPLEMENTARY FIGURE 8



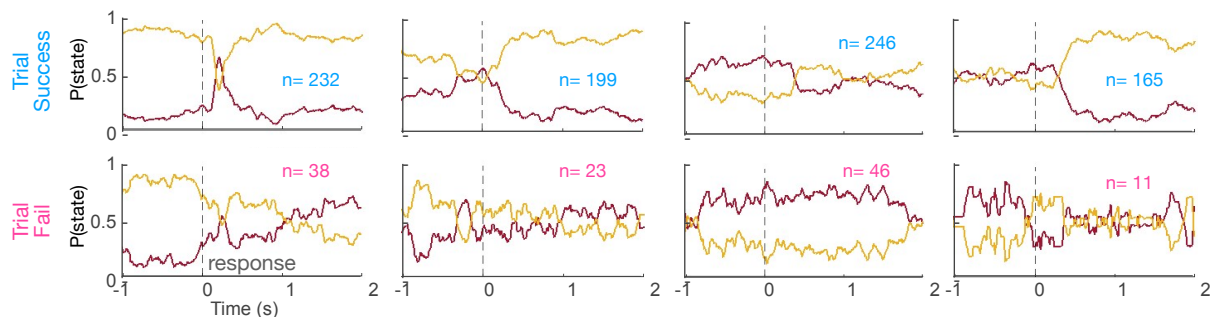
Neuronal average firing rate per state. Average firing rate and its standard deviation per neuron during the different states activation. The HPHT (red) state involves higher neuronal firing rate on almost every neuron.

SUPPLEMENTARY FIGURE 9



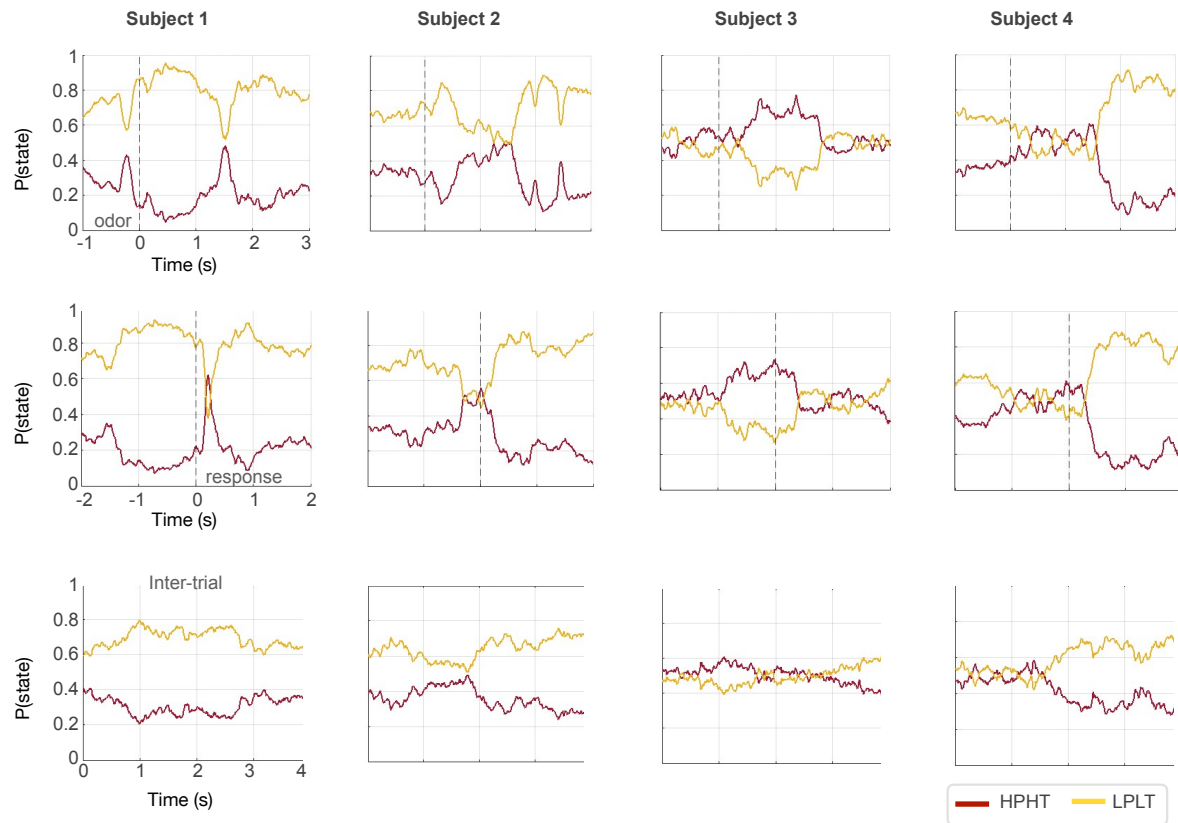
Prediction of state activation from the neural population activity. For each subject, we show the accuracy of decoding the state activation from the neural population activity, standardised across neurons, per trial and per time point, to rule out that specific subpopulations of neurons are preferentially indexed by the states. The standardisation step is needed to ensure that relative activation of each neuron, instead of the average firing rate, is driving the prediction. The prediction accuracy is averaged over 100 repetitions of a class-balanced decoding procedure, and is here reported within its 95% CI. The accuracy, typically around chance level, indicates that the states do not selectively recruit different neurons within the population.

SUPPLEMENTARY FIGURE 10



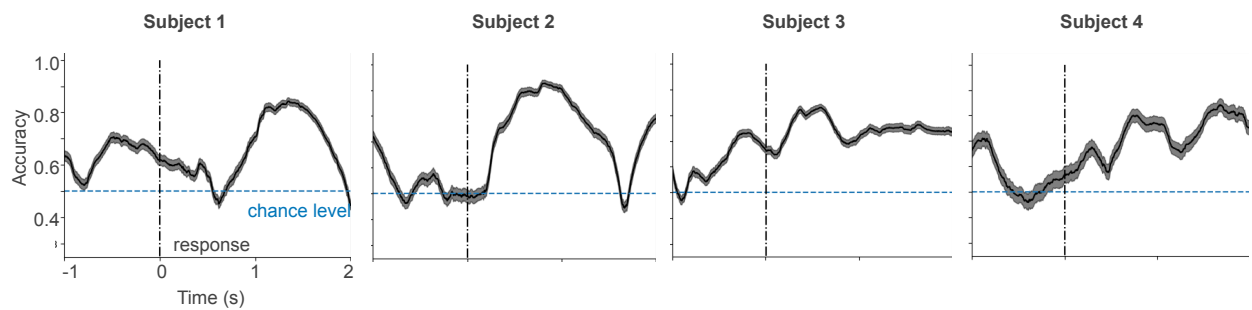
States distribution with respect to trial outcome. For each subject, we show the probability of each state being active around response, for trials that have been successful (top row) or failed (bottom row). The states activation probability is computed as an average across trials (n indicates the number of trials available), aligning trial data at response. Typically, the LPLT (yellow) state tends to be the most active during successful trials, while there is no clear state describing the failed trials.

SUPPLEMENTARY FIGURE 11



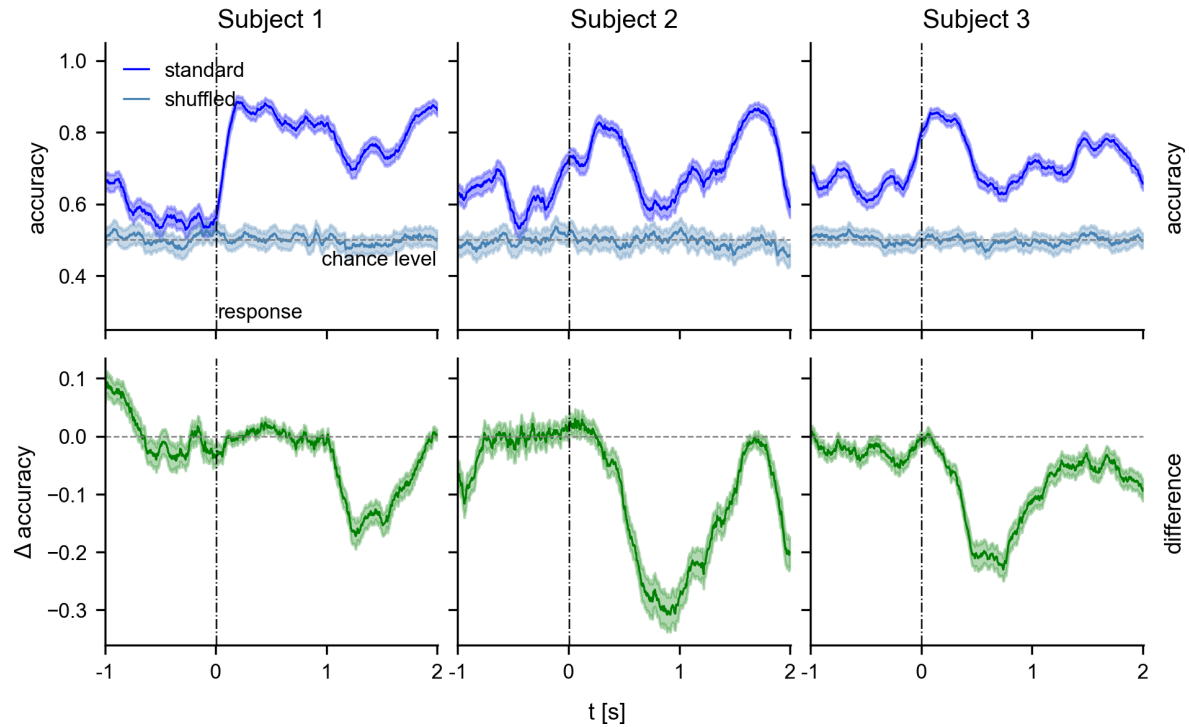
States distributions around task-relevant moments. Distributions around odor presentation (top row), response (middle row) and inter-trial time (bottom row).

SUPPLEMENTARY FIGURE 12



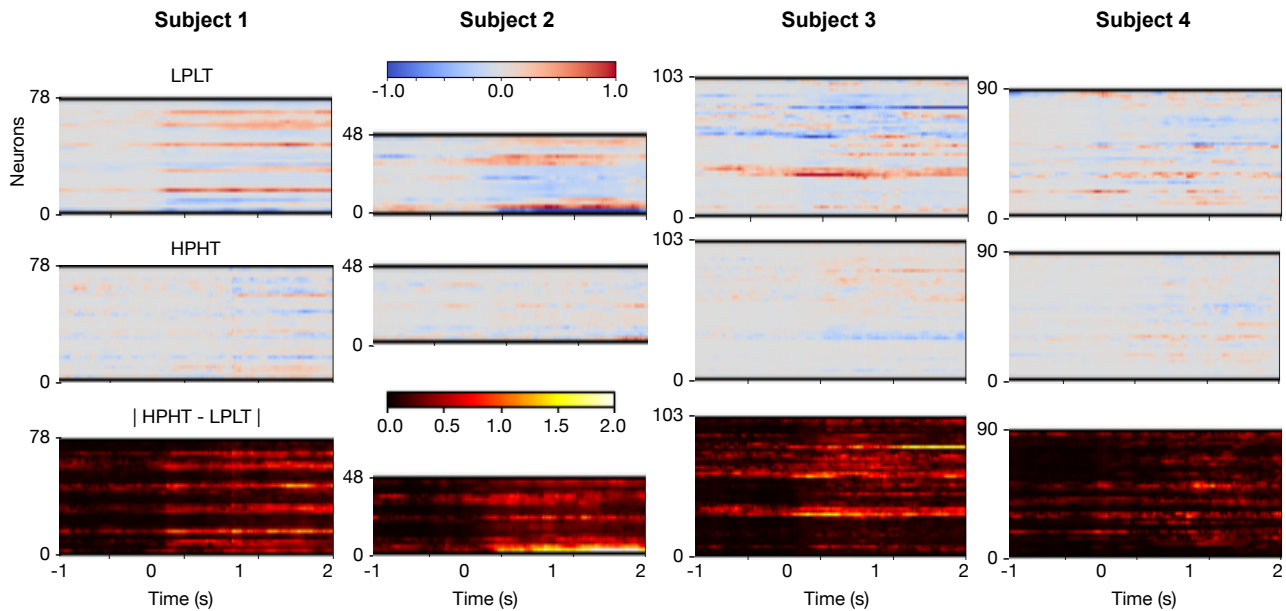
Predicting trial outcome from head position. For each subject, shown the prediction accuracy of trial outcome from X_p and Y_p coordinates, within its CI of 95% as obtained with 200 bootstraps to ensure class balance. Decoding accuracy is high especially after response, likely reflecting the licking-induced movement of rewarded trials.

SUPPLEMENTARY FIGURE 13



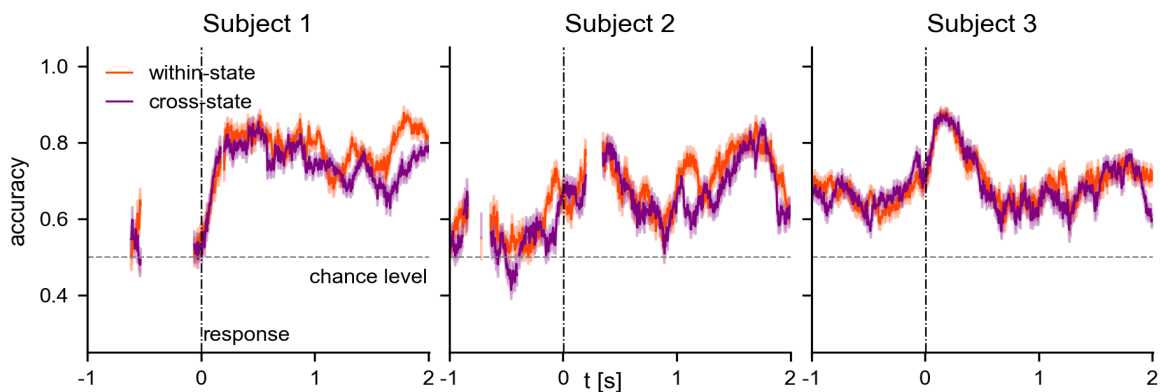
Standard and shuffled decoding from deconfounded neural activity. **Top row:** For each subject, shown the prediction accuracy of trial outcome from the neural activity deconfounded for head motion (X_p and Y_p coordinates), within its CI of 95% as obtained with 200 bootstraps to ensure class balance. For comparison, the decoding is performed also on deconfounded neural activity with shuffled trial outcome labels (cyan, 200 permutations, 95% CI), as expected at chance level. **Bottom row:** the difference between standard decoding performed on the original neural activity and on neural activity deconfounded for head motion.

SUPPLEMENTARY FIGURE 14



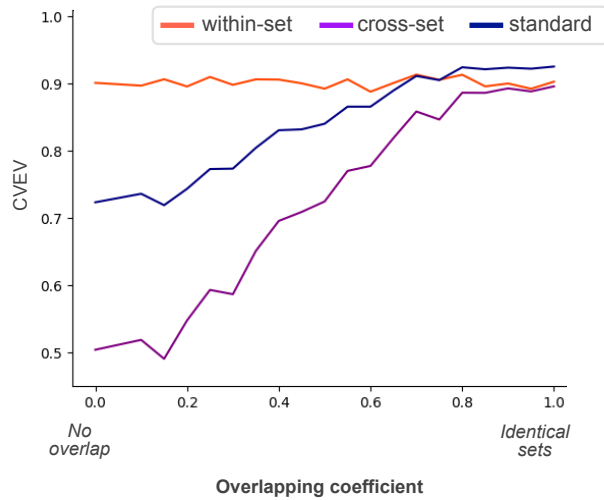
Neural activity representing success vs fail across states. The plots show the difference in neural activity averaged across trials for successful minus failed trials, for those trials in state LPLT (top row), and in HPHT (middle row). The bottom row shows the absolute difference of the activity in the two higher panels.

SUPPLEMENTARY FIGURE 15



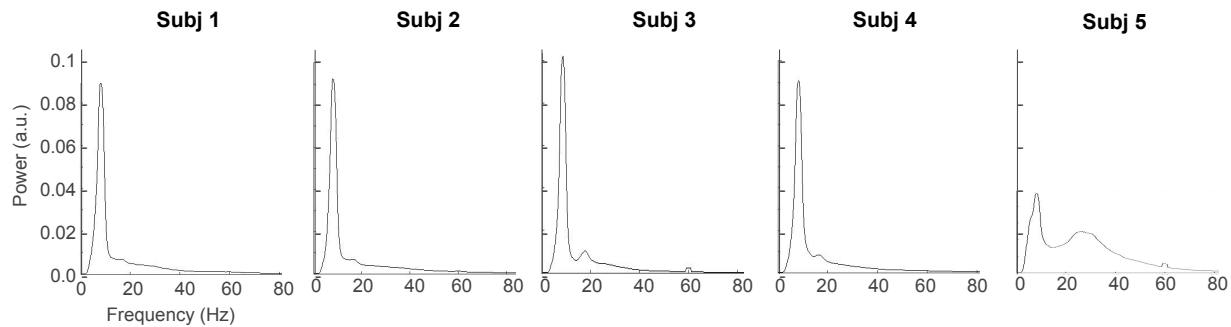
Within-state and cross-state decoding accuracy (signal filtered <100Hz). For each subject, shown the average decoding accuracy within-state (orange) and cross-state (purple) within their 95% CI (over 200 bootstraps to ensure class-balance), with the states obtained by running the HMM on the low-pass filtered signal (sharp filter at 100Hz). If the number of trials for each state at any point is lower than 5, the decoding is not performed (gaps in the plot).

SUPPLEMENTARY FIGURE 16



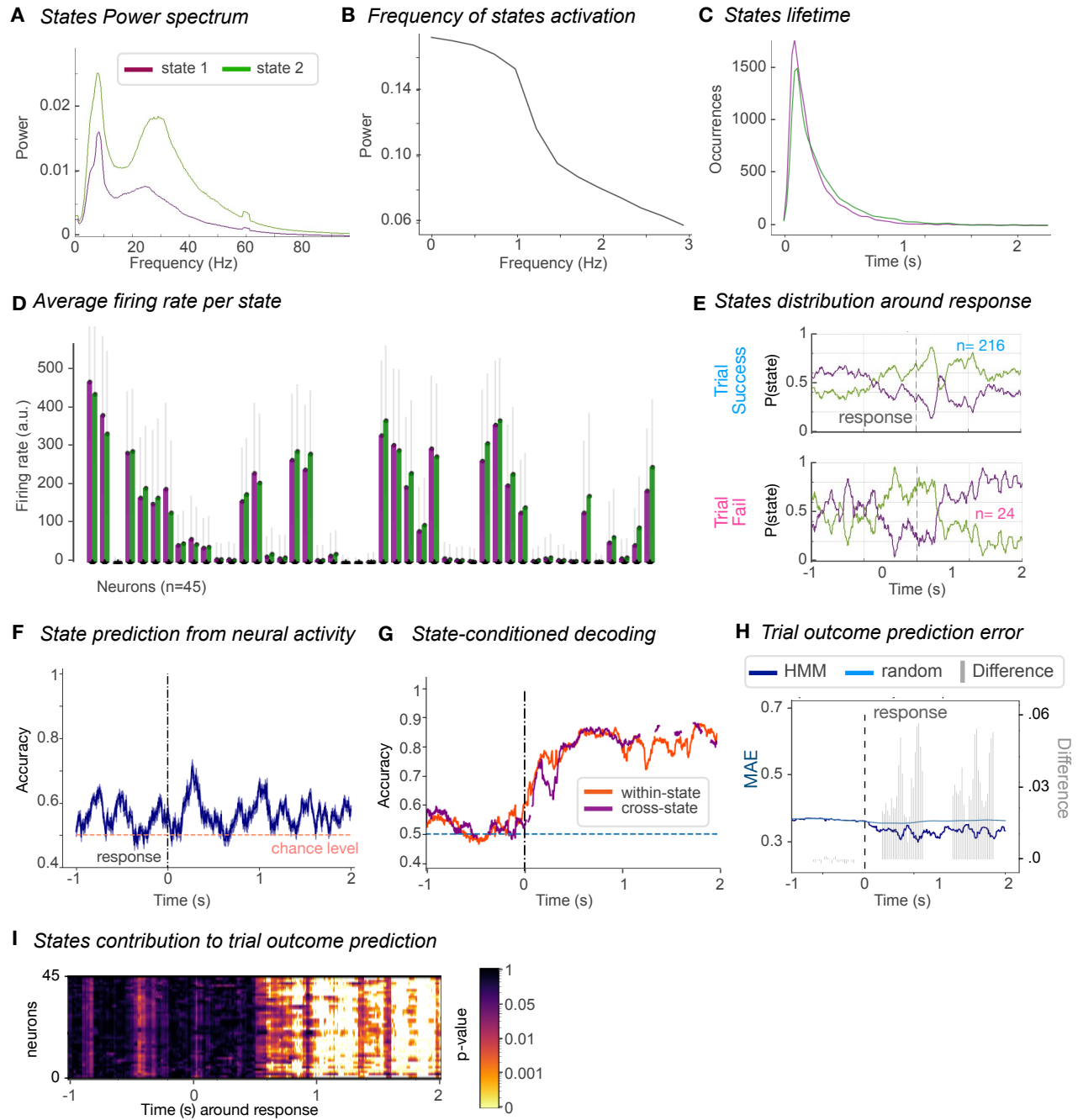
Decoding accuracy from simulated sets. The plot shows cross-validated explained variance (CVEV) of simulated sets, encoding information with various degree of overlap. Reported the CVEV of within-set, cross-set and standard decoding (i.e., decoding from both sets jointly), averaged over 500 simulated set for each value of overlap degree (20 values between 0 and 1). For values between 0 and 0.5 of the overlap degree (no overlap to 50% shared information encoded), within-set decoding works better than standard decoding and cross-set decoding performs much worse than both. For values between 0.6 and 1 of the overlap degree (1 indicating identical encoding weights for the two sets), standard decoding surpasses within-set decoding, and the cross-set decoding performance is increasingly higher. Between 0.7 and 0.8 of overlap degree lies the matching between these simulations and our experimental results.

SUPPLEMENTARY FIGURE 17



LFP data power spectrum of the five rats. Power spectrum of the first principal component (PC) of the LFP data (high-pass filtered at 4Hz) for the five rats. Because of how different its power spectrum was from the other rats, Subject 5 was not included in the main report. We have nonetheless performed all analyses on this animal, reported in **Supplementary Figure 18**.

SUPPLEMENTARY FIGURE 18



Analyses performed on subj 5. The HMM analyses were performed with the exact same settings as in all the other subjects. **A. States power spectra.** The states capture general differences in power across the frequency spectrum, but there is no specific LPLT and HPHT states. **B. States periodicity.** Similarly to all other subjects, the periodicity of the states peaks below 2 Hz, meaning at most two states switch per second. **C. States lifetimes.** Both states have a similar lifetime distribution. **D. Average firing rate per state.** Similarly to the other subjects, the state associated with higher power typically involves a higher firing rate across neurons. **E. State distribution around response.** The state involving higher power across the frequency spectrum is the most active during successful trials, while there is no clear dominance for failed trials. **F. Prediction of state activation from standardised neural activity.** Similarly to the other

subjects, the prediction accuracy is typically around chance level, indicating that the state-neural activation relationship goes beyond the mere average firing rate. **G. State-conditioned decoding.** Within-state vs cross-state decoding accuracy. In some time points after response within-state accuracy is higher than cross-state. The points where cross-state decoding accuracy is missing indicate timepoints where the number of trials for one state is not enough to perform decoding. **H. Trial outcome prediction error** (state-neuron interaction analysis). Like in the other subjects, also here the HMM states help better predicting trial outcome after response. **I. States contribution to prediction of trial outcome.** Shown the significance of the improved prediction accuracy when adding the HMM states as regressors in the prediction of trial outcome.