Supplementary Materials

Neural population partitioning and a concurrent brain-machine interface for sequential motor function

Maryam M. Shanechi, Rollin C. Hu, Marissa Powers, Gregory W. Wornell, Emery N. Brown & Ziv M. Williams

Supplementary Modeling

Model construction using the expectation-maximization (EM) algorithm

We model the activity of each neuron under any given sequence as an inhomogeneous Poisson process whose likelihood function (using the theory of point processes) is given by[49, 50]

$$p(N^c_{1:K} \mid S_i) = \prod_{k=1}^{K} (\lambda_c(k \mid S_i)\Delta)^{N^c_k} \exp(-\lambda_c(k \mid S_i)\Delta) \qquad i = 1{:}12, \qquad (1)$$

where Δ is the time increment taken to be small enough to contain at most one spike, N^c_k is the binary spike event of the c'th neuron in the time interval $[(k-1)\Delta, k\Delta]$, $\lambda_c(k \mid S_i)$ is its instantaneous firing rate in that interval, S_i is the i'th sequence, and K is the total number of bins in a duration $K\Delta$.

For each sequence and neuron, we need to estimate the firing rate $\lambda_c(k \mid S_i)$ using the neuronal data observed. One way to do so is to bin the data into non-overlapping windows of fixed length during which the firing rate is assumed to be constant and estimated using maximum likelihood techniques. This method is equivalent to finding the peristimulus time histogram (PSTH) that simply averages the number of spikes over any given window. The main drawback of this technique is that unless there are a large number of training trials under each sequence, to get a good estimate one has to pick relatively large windows. This in turn masks the fine-scaled evolution of the firing rate. Also, there is no principled way for selecting a window size, which the analysis is dependent on.

One way to avoid these problems in estimating the spike rate function is to use a state-space approach[42, 51] (See also alternative methods using Gaussian processes in prior work[54]). This approach is used in many applications to estimate an unobservable state process and consists of two models: A prior or state model that in general enforces any prior information available about the unobservable states—such as a simple continuity condition—and an observation model that relates the neuronal observations to these states. In the case of estimating the spike rate function, and since it is a non-negative quantity, similar to previous work[42, 51] we take the state at time increment k, x_k, to be the logarithm of the firing rate, i.e., $x_k = \log(\lambda_c(k \mid S_i))$, or equivalently

$$\lambda_c(k \mid S_i) = \exp(x_k), \tag{2}$$

and enforce a continuity condition on it by assuming that it evolves according to a linear first-order Gaussian model[42, 51],

$$x_k = x_{k-1} + \epsilon_k,$$

where ϵ_k is the zero-mean white Gaussian noise with variance σ_ϵ^2. The observation model is in turn given by substituting (2) in (1). Here, $\theta = \sigma_\epsilon^2$ is an unknown parameter of the model and should be estimated jointly with the state. Hence we use the expectation-maximization (EM) iterative algorithm to find the maximum likelihood estimate of θ and in turn estimate the firing rate[42, 51, 52]. Denoting the estimate of θ in the i'th iteration by $\theta^{(i)}$, its estimate in the $i+1$'th iteration after the maximization step is given by

$$\theta^{(i+1)} = \frac{1}{K}\left(\sum_{k=1}^{K} W_k + W_{k-1} - 2W_{k,k-1}\right), \tag{3}$$

where $W_k = E[x_k^2 \mid N_{1:K}^c; \theta^{(i)}]$ and $W_{k,k-1} = E[x_{k-1}x_k \mid N_{1:K}^c; \theta^{(i)}]$ are found from the forward filter, fixed-interval smoothing, and covariance recursive algorithms in the expectation step as follows. Assuming that there are J total trials and denoting the causal filter state estimate by $x_{k|k} = E[x_k \mid N_{1:k}^c; \theta^{(i)}]$ and its variance by $w_{k|k}$, and the smoothed state estimate by $x_{k|K} = E[x_k \mid N_{1:K}^c; \theta^{(i)}]$ and its variance by $w_{k|K}$, the recursions in the E-step are given by the forward filter recursions[42, 51, 55],

$$w_{k|k-1} = w_{k-1|k-1} + \theta^{(i)}$$
$$x_{k|k-1} = x_{k-1|k-1}$$
$$w_{k|k} = (w_{k|k-1}^{-1} + J\exp(x_{k|k-1})\Delta)^{-1}$$
$$x_{k|k} = x_{k|k-1} + w_{k|k}\sum_{j=1}^{J}(N_k^c(j) - \exp(x_{k|k-1})\Delta),$$

for $k = 1,\cdots,K$, where $N_k^c(j)$ is the spike event in trial j, and by the fixed interval smoothing recursions[42, 51, 56],

$$A_k = w_{k|k}w_{k+1|k}^{-1}$$
$$x_{k|K} = x_{k|k} + A_k(x_{k+1|K} - x_{k+1|k})$$
$$w_{k|K} = w_{k|k} + A_k^2(w_{k+1|K} - w_{k+1|k}),$$

for $k = K-1,\cdots,0$ and with initial condition $x_{K|K}$ and $w_{K|K}$ from the filter recursions. We pick the initial conditions for the forward filter at each iteration of the EM algorithm as $x_{0|0}^{(i+1)} = x_{0|K}^{(i)}$ and $w_{0|0}^{(i+1)} = w_{0|K}^{(i)}$. Finally the state-space covariance algorithm gives all the terms needed for the M-step to find $\theta^{(i+1)}$ in (3) using these recursions[42, 51, 57],

$$W_k = w_{k|K} + x_{k|K}^2,$$

for $k = 0,\cdots,K$ and

$$W_{k+1,k} = A_k w_{k+1|K} + x_{k|K}x_{k+1|K},$$

for $k = 0,\cdots,K-1$. The iterations of the EM algorithm are run until convergence. The estimated firing rate at any time bin $k = 1,\cdots,K$ is in turn the smoothed estimate, $\hat{\lambda}_c(k \mid S_i) = \exp(x_{k|K})$ evaluated at the estimate of θ in the final iteration.

Repeating this procedure for all neurons under each sequence and fitting the inhomogeneous Poisson models results in a continuous smoothed estimate of the rate function for each neuron under any given sequence and over the entire length of a trial. Our implementation of the EM algorithm is similar to prior work[42, 51] (for a comparison with PSTH see **Fig. 5** and **Supplementary Figs. 3, 4**).

Testing the decoding performance

Chance level accuracy

For the sequence, the chance level accuracy is simply $1/S$, where S is the number of sequences used. For targets, however, one has to take into account the correlation between the first and second targets when calculating the chance level accuracy. This is because depending on the number of sequences used in the decoding analysis, the first and second decoded targets may not be independent. In the case of 12 sequences, for example, since both targets cannot be at the same location, information about one also implies some information about the other. These correlation effects must therefore be taken into account when calculating the chance level accuracy of the targets.

We define indicator functions for the first and second targets, denoted by I_1 and I_2, that are 1 if the corresponding targets are decoded correctly and 0 otherwise. We show this analysis for the case when 12 sequences are used. In all other cases they can be found similarly. For 12 sequences, using the total law of probability, the probability that the second target is decoded correctly is given by,

$$p(I_2 = 1) = p(I_2 = 1 | I_1 = 1)\, p(I_1 = 1) + p(I_2 = 1 | I_1 = 0)\, p(I_1 = 0).$$

Now if the first target is correctly decoded, the second target could be at one of three possible locations as the two targets cannot be at the same location. Hence the chance level accuracy in this case is given by $p(I_2 = 1 | I_1 = 1) = 1/3$. By a similar argument, if the first target is decoded incorrectly, the chance level accuracy of the second target is $p(I_2 = 1 | I_1 = 0) = 2/9$. Hence the chance level accuracy of the second target is given by

$$p_2^{chance} = \frac{2}{9} + \frac{1}{9} p(I_1 = 1), \qquad (4)$$

and vice versa for the first target as the two targets are selected symmetrically in the choice of sequences. For example, in a session where we observe a first target accuracy of p^* in our decoding analysis, the chance level accuracy for the second target is $p_2^{chance} = 2/9 + p^*/9$ as

opposed to simply $1/4$. Note that if $p^* = 1/4$, i.e., at chance level, then $p_2^{chance} = 1/4$ and also at chance level as expected.

Random permutation test: Testing significance for the divergence in the amount of information held by each cell about the two targets

To determine whether a significant divergence exists in the amount of information held by the premotor neurons about the two targets, we need to show that the absolute difference between the two target accuracies of each cell averaged across the population is significantly larger than that of a population with the same target accuracy values but with no structured relationship between each cell's target accuracies. To find the distribution of this average absolute difference in such a population with no structure, we keep the accuracy values the same but randomly permute them within the population, and repeat this process 100,000 times. This removes any possible structure between the target accuracies of each cell and hence creates a null hypothesis distribution. We can then establish the significance of the divergence by comparing the average absolute difference of the target accuracies of the premotor neurons against this null distribution and calculate a *P*-value.

To do so, we first correct for the correlation effect between the first and second target accuracy values, which is the byproduct of the choice of sequences used in the experimental design. Since the first and second targets cannot be at the same location within the set of 12 sequences, the accuracy of one target has a contribution (even though fairly small) to the accuracy of the other even if the neuronal activity is not encoding that target *per se*. Hence we also need to remove this effect to see the true representation of a target by the neuronal activity, just as we take it into account in calculating the chance level accuracies. This means that we subtract from the first target accuracy value of each cell, the chance level contribution of that cell's second target accuracy, or $P_1^{chance} - 1/4$, and vice versa. We then randomly permute these first target accuracy values among the cells while keeping their second target accuracy values the same, and repeat this process many times. This generates a new population each time with the same accuracy values but no pair-wise structure between the first and second target accuracies of each cell. For each new population, we compute the average difference and after repeating this many times,

find the distribution of this average difference. We then use this distribution to find whether the difference of the first and second target accuracies of the cells averaged over the premotor population is significantly different from a population with no structure (We find that the average divergence in the premotor population is significantly different from that of a population with no structure, with or without applying the correlation correction, $P < 10^{-15}$).

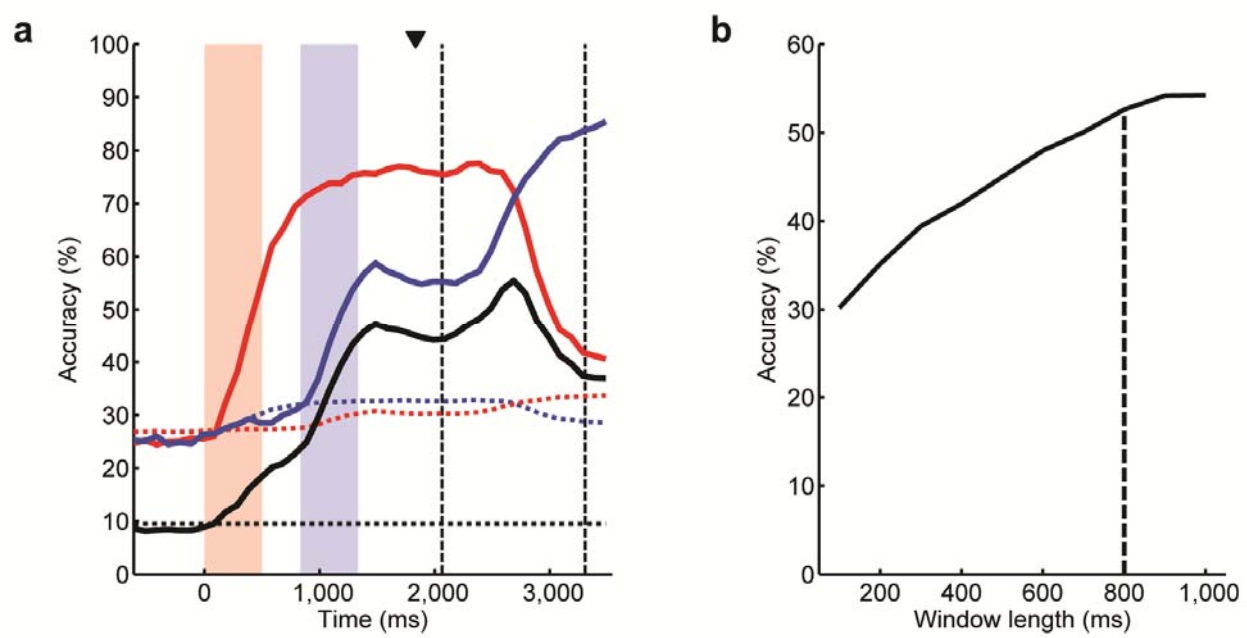
We corrected for the fairly small correlation effect between the first and second target accuracy values in the scatter plots of **Fig. 6** and **Supplementary Fig. 6** as explained for the random permutation test above.

54. Cunningham, J. P., Yu, B. M., Shenoy, K. V. & Sahani, M. Inferring neural firing rates from spike trains using Gaussian processes. In: *Advances in Neural Information Processing Systems 20*, (eds. Platt, J., Koller, D., Singer, Y. & Roweis, S.) 329–336 (MIT Press, Cambridge, MA, 2008).

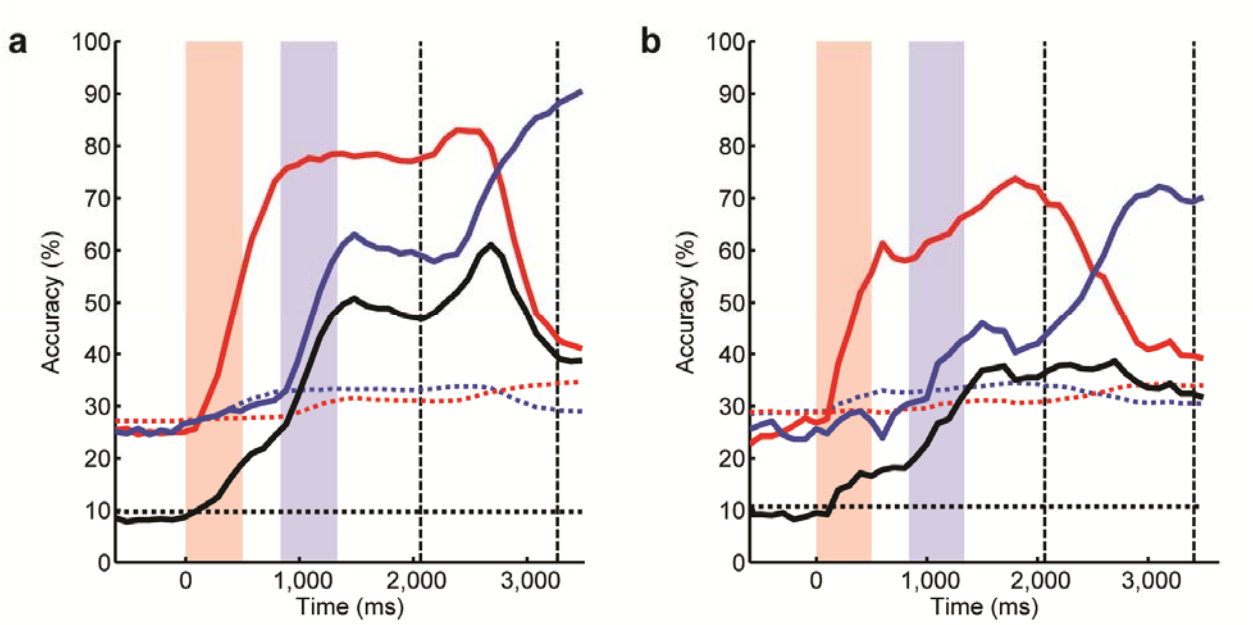
55. Eden, U. T., Frank, L. M., Barbieri, R., Solo, V. & Brown, E. N. Dynamic analysis of neural encoding by point process adaptive filtering. *Neural Comput.* **16**, 971–998 (2004).

56. Brown, E. N., Frank, L. M., Tang, D., Quirk, M. C. & Wilson, M. A. A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *J. Neurosci.* **18**, 7411–7425 (1998).

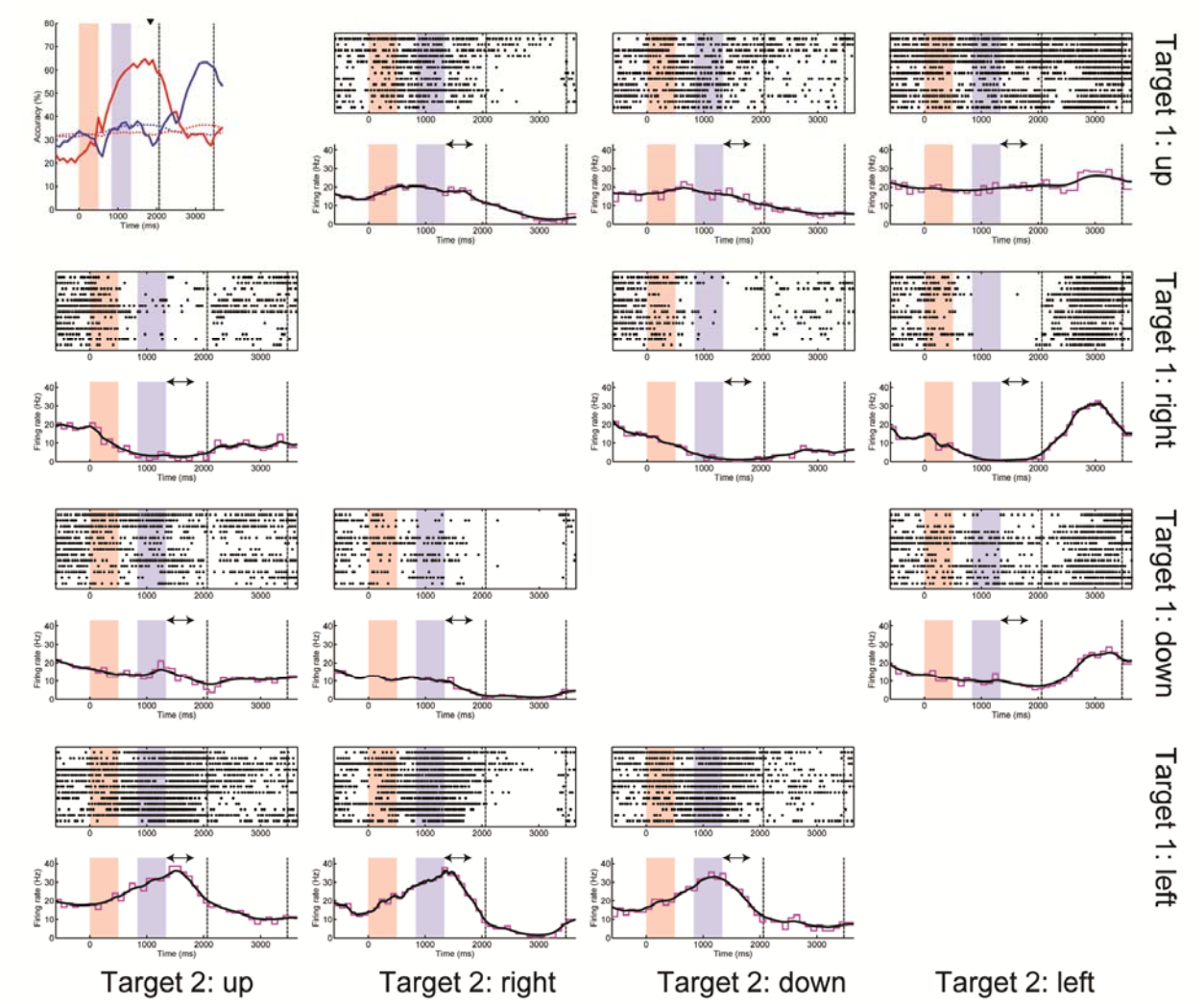
57. Jong, P. D. & Mackinnon, M. J. Covariances for smoothed estimates in state space models. *Biometrika* **75**, 601–602 (1988).



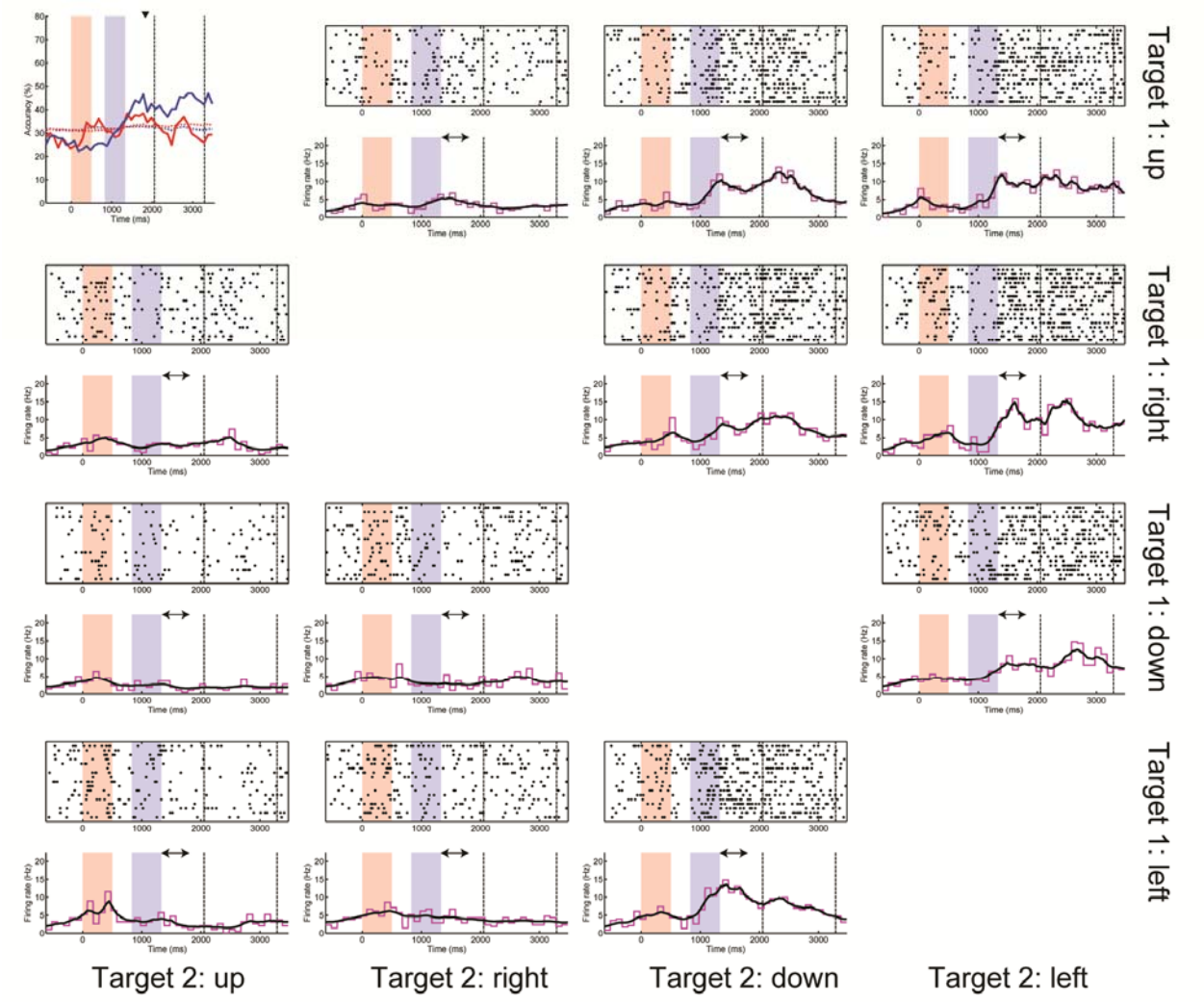
Supplementary Figure 1 Mean population decoding accuracy across all recorded sessions. (a) Mean decoding accuracy for the population across all recorded sessions for the first target (red curve), second target (blue curve) and the full sequence (black curve). The figure has the same convention used in **Fig. 2**. (b) Mean sequence decoding accuracy as a function of the time window length preceding the earliest "go" cue used in decoding. The black curve shows the mean population sequence decoding accuracy (out of 12 possibilities) across all sessions. Using an 800 ms window, the sequence decoding accuracy exceeds 95% of the maximum possible when using the neuronal activity from the start of second target presentation until the earliest "go" cue.



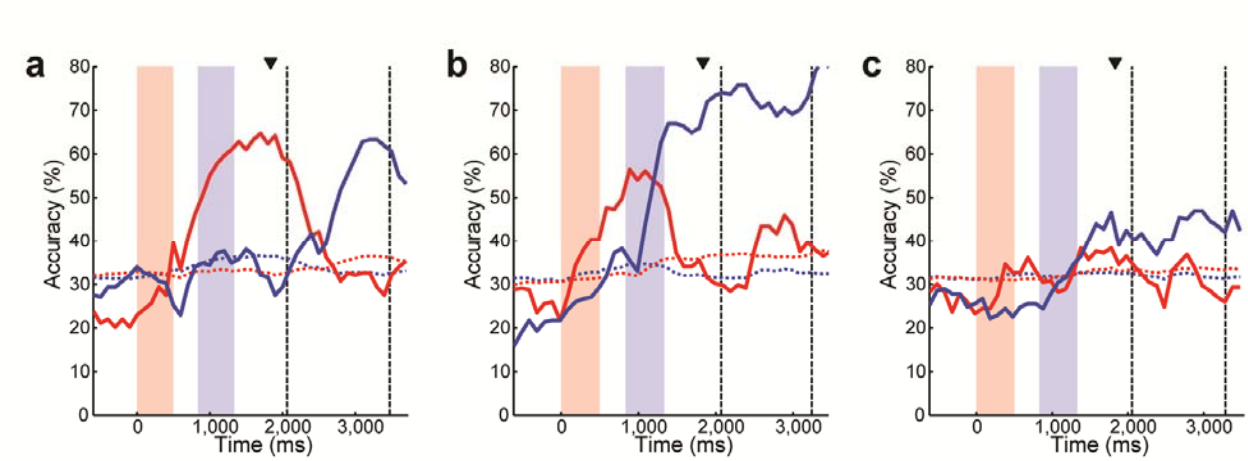
Supplementary Figure 2 Mean population decoding accuracy over time for all recorded sessions in monkey 1 (a) and monkey 2 (b). The figure has the same convention used in **Fig. 2**. Across all standard sessions and during the 500 ms working memory period, the first target, second target, and sequence accuracies were 76 ± 12%, 60 ± 19%, and 48 ± 13% for the first monkey and 74 ± 11%, 43 ± 3%, and 36 ± 3% for the second monkey, respectively (mean ± s.d.).



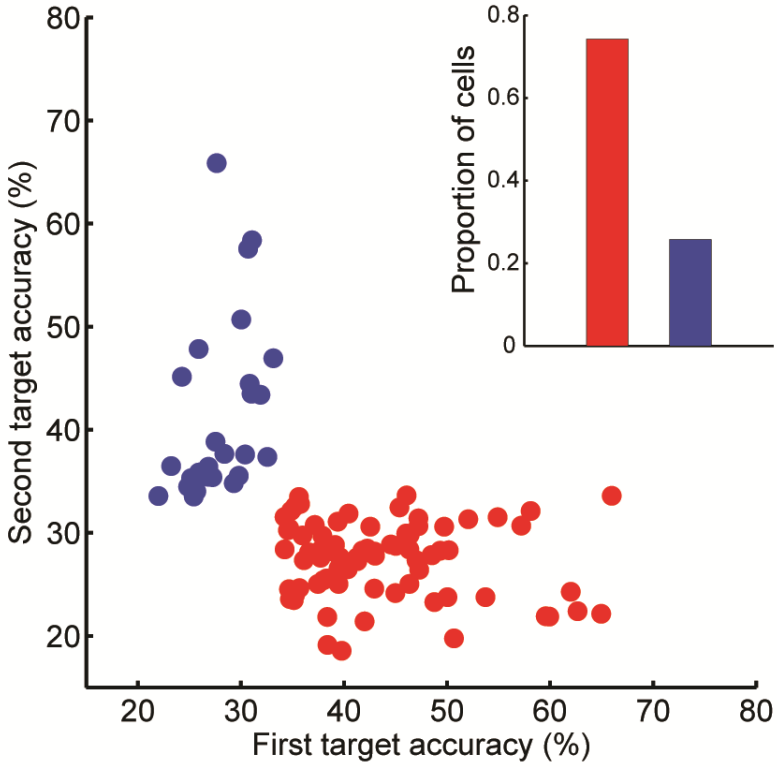
Supplementary Figure 3 Example of a first (currently held) target selective neuron. The subfigure at the upper left corner shows the first and second target accuracies of the cell as a function of time into the trial. The vertical bars/lines and their timings follow the same convention as **Fig. 2**. In all other subfigures, each *top* panel corresponds to a different sequence of movements with each row illustrating the spiking activity during a single trial and the black dots indicating the spike times. Each *bottom* panel indicates the corresponding mean firing rate estimates using the expectation-maximization procedure (black curve) and the corresponding peristimulus time histogram (PSTH) (magenta curve). The arrow indicates the working memory period. The subfigures in the same row correspond to sequences with the same first target location. The subfigures in the same column correspond to sequences with the same second target location. Note that repeated target locations were not used in the sequences and hence there are 3 subfigures per row/column.



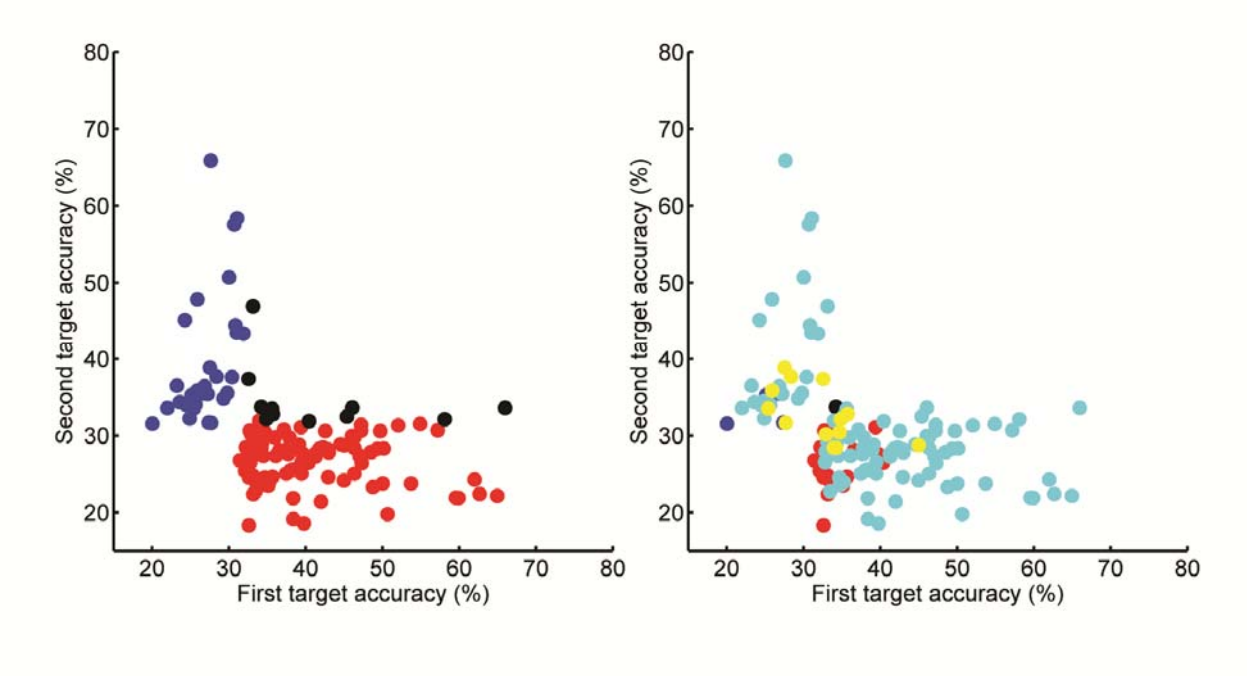
Supplementary Figure 4 Example of a neuron selective for both targets. Figure has the same convention used in **Supplementary Fig. 3**.



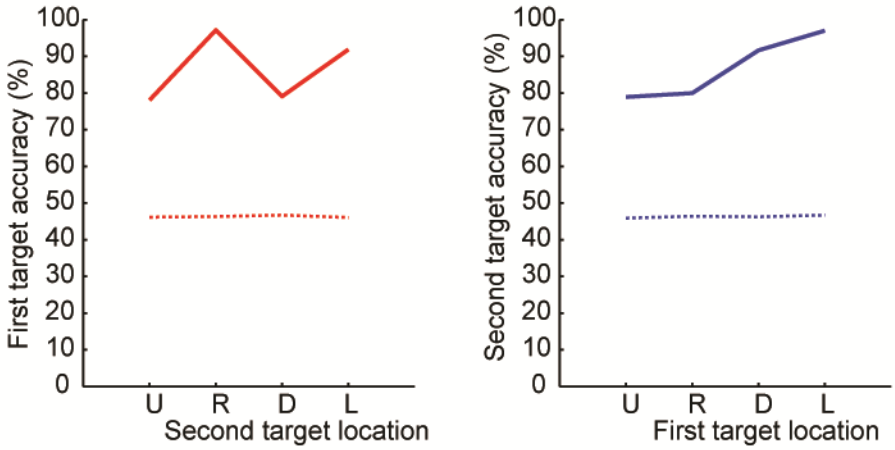
Supplementary Figure 5 Decoding accuracies over time for three sample cells that were selective for the first target only (**a**), second target only (**b**), and both targets (**c**) during the working memory period. Figure conventions are the same as in **Fig. 2**.



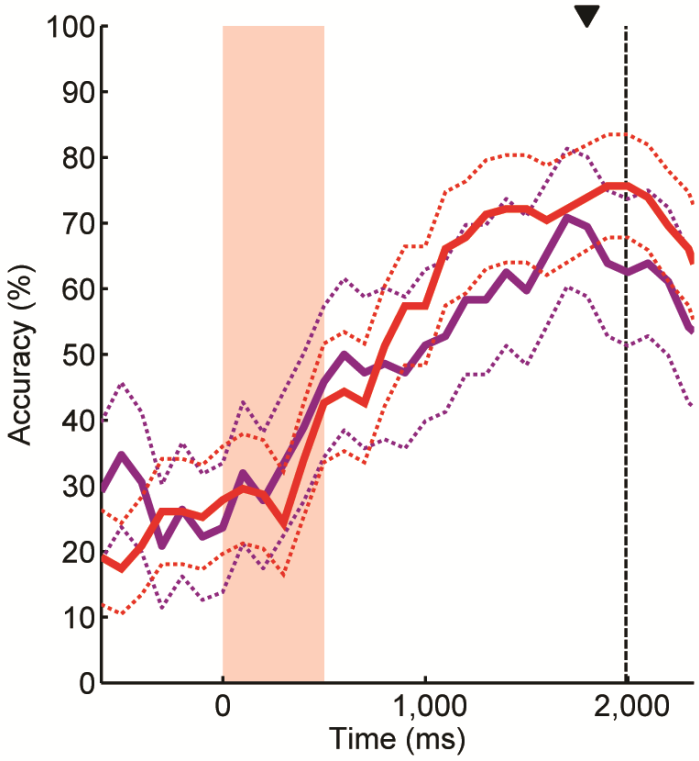
Supplementary Figure 6 Partitioning of the population during working memory. Scatter plot of the first and second target accuracies of the cells that significantly encoded at least one target during the working memory period. Statistical significance of the target accuracies was tested here at a stricter level ($P < 0.001$) than in **Fig. 6**. Red points indicate cells that significantly encoded only the first target and blue points indicate those that significantly encoded only the second target. At this statistical level, no cell had a significant accuracy for both targets. The *inset* indicates the proportion of cells that significantly encoded only the first or only the second target during the working memory period with the same coloring schemes from left to right.



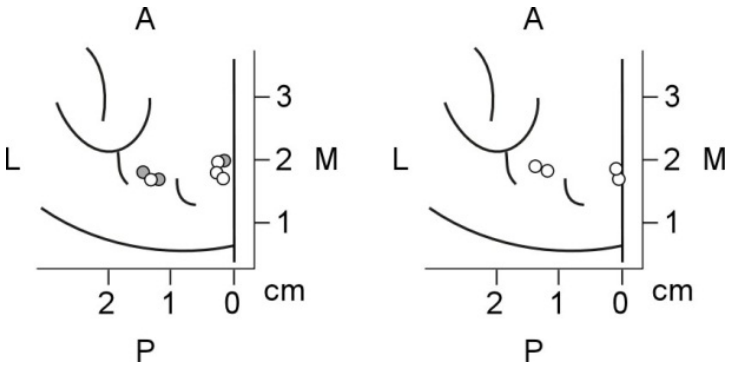
Supplementary Figure 7 Neural partitioning vs. sequence specific selectivity. The original scatter plot (**Fig. 6**) demonstrating the partitioning mechanism is shown on the *left*. On the *right*, cells that are selective to a single specific sequence (i.e., significantly change their firing rate in response to a single sequence) are colored with *yellow* (paired *t*-test, $P < 0.05$, FDR correction for 12 comparisons). We find that these sequence selective cells are few in number (10% of the cells) and are among the least informative cells (i.e., have low accuracies). Cells that show a significant change in firing rate from baseline for *at least one* specific sequence are colored in *cyan*. As evident, this analysis by itself demonstrates that the majority of the cells in the scatter plot display a change in firing rate for at least one sequence (as expected) but it does not reveal the population partitioning. The decoding analysis calculates decoding accuracy as a measure of information when considering all sequence combinations collectively. It also further disambiguates the amount of information held simultaneously about the first and second targets by each neuron and therefore reveals the partitioning mechanism.



Supplementary Figure 8 Conditional decoding accuracies. The conditional first target accuracy given the possible second target locations (i.e., U, R, D, L) and the conditional second target accuracy given the possible first target locations are shown for the population in a sample session (same session as in **Fig. 2**). The dotted lines indicate the 99% chance upper confidence bounds. Across all sessions we found no significant difference in decoding accuracy of the second target based on the location of the first target, and vice versa (repeated ANOVA, $P > 0.15$).



Supplementary Figure 9 Comparison of first target decoding accuracy for the population in interleaved and non-interleaved sessions. In the interleaved dual-target/single-target session, target decoding accuracy on single-target trials is shown in red. In the single-target only session, target decoding accuracy is shown in magenta. Each point on the curves indicates the decoding accuracy for the population over the *preceding* 500 ms window. Dotted lines indicate the 95% confidence bounds for each accuracy curve (rather than chance level). The red vertical bar indicates the time during which the (first) target was presented, and the vertical dotted line indicates the average time of the first "go" cue presentation onset. The arrow indicates the time point corresponding to the decoding accuracy of the preceding working memory period.



Supplementary Figure 10 Electrode recording sites for each of the two monkeys. Top-view schematic of the electrode array positions. Each array (circle) contains 32 electrode contracts. The bar in centimeters is referenced in relation to interaural antero-postero coordinates and midline medio-lateral coordinates. Here, A is anterior, P posterior, M medial and L lateral. In the *left* panel, from monkey 1, the white and gray circles indicate recordings from two separate hemispheres. In the *right* panel, from monkey 2, the white circles indicate recordings from one hemisphere.