ARTICLE

https://doi.org/10.1038/s41586-019-0919-7

Discrete attractor dynamics underlies persistent activity in the frontal cortex

Hidehiko K. Inagaki[1], Lorenzo Fontolan[1], Sandro Romani[1]* & Karel Svoboda[1]*

Short-term memories link events separated in time, such as past sensation and future actions. Short-term memories are correlated with slow neural dynamics, including selective persistent activity, which can be maintained over seconds. In a delayed response task that requires short-term memory, neurons in the mouse anterior lateral motor cortex (ALM) show persistent activity that instructs future actions. To determine the principles that underlie this persistent activity, here we combined intracellular and extracellular electrophysiology with optogenetic perturbations and network modelling. We show that during the delay epoch, the activity of ALM neurons moved towards discrete end points that correspond to specific movement directions. These end points were robust to transient shifts in ALM activity caused by optogenetic perturbations. Perturbations occasionally switched the population dynamics to the other end point, followed by incorrect actions. Our results show that discrete attractor dynamics underlie short-term memory related to motor planning.

Short-term memory is the ability of the brain to maintain information over seconds. Neurons in the frontal cortex and related brain regions show persistent, or slowly varying, changes in spike rate that correlate with the maintenance of short-term memories[1–9]. This neural correlate has been extensively studied in delayed response tasks in non-human primates[3,8] and, more recently, in rodents[6,7,9] (Fig. 1a). In a typical task, an instruction informs the type of action to be performed and a go cue determines the timing of action, but only after a delay epoch, during which animals maintain a memory of the instruction and/or plan a movement. Persistent delay activity that predicts future movements is referred to as preparatory activity[2,7,8,10].

The ALM is part of a multi-regional network that mediates motor planning[10–13]. A large proportion of ALM neurons exhibit preparatory activity that predicts licking direction[6]. The dynamics of ALM neural population can be analysed in activity space, in which each dimension corresponds to the activity of one neuron. ALM activity during the delay epoch is approximately two-dimensional[9]. The coding direction (**CD**) vector discriminates trial types (that is, right or left lick), and activity projected along the **CD** contains almost all direction-selective activity[9,11]. The second dimension corresponds to non-selective slow ramping activity[9,11].

After current is injected into isolated neurons, activity typically decays within milliseconds, reflecting an interaction of rapid repolarizing currents and the neuronal membrane time constant[14] (Fig. 1b, left). Several mechanisms could bridge the gap in time scales between neuronal time constants and the slow dynamics observed during motor planning[15–31]. Neurons could be wired into networks that produce sequential neuronal activation, effectively extending network time constants[18]. However, our previous analysis did not detect sequential activity in ALM during the delay epoch[9]. Here we focus on testing other models of persistent activity. First, specialized cell-autonomous mechanisms could maintain multi-stable persistent activity[22,32,33]. Second, feedback in neural networks can compensate for dissipation of excitation. Depending on the structure of the circuit, and the properties of individual neurons, the network can behave as a continuous attractor (or integrator) with a continuum of stable (or quasi-stable) states[16,23,30,31], or one or more discrete attractors with distinct activity states[17,19,28,29] (Fig. 1b, Extended Data Fig. 1). Continuous and discrete attractor models can explain the neuronal activity that underlies short-term memory and decision-making[15–17,24–26,34], and it has therefore been difficult to distinguish between these models[11]. Here we analyse the mechanisms that underlie persistent activity in mice performing a delayed directional licking task.

Testing for cell-autonomous mechanisms

We performed whole-cell recordings from left ALM neurons during a delayed response task[6,9] (79 cells) (Fig. 2a–c, Extended Data Fig. 2). Twenty of the recorded neurons were selective during the delay epoch (selective cells; spike rate significantly different between correct lick-right and correct lick-left trials; Wilcoxon rank sum test, $P < 0.05$). The membrane potential (V_m) to spike rate relationship was approximately threshold linear[35] (Extended Data Fig. 2). Similar to extracellular recordings[6,9], average selectivity increased during the delay epoch (Fig. 2d, Extended Data Fig. 3b, j). Therefore, persistent changes in V_m are consistent with persistent changes in spike rate during the delay epoch.

The membrane time constant limits how long isolated neurons can maintain activity after a transient input. The membrane time constant, estimated by fitting an exponential function to V_m following current injections, was short for both selective (21.3 ± 19.7 ms, mean ± s.d., $n = 19$) and non-selective cells (20.5 ± 16.9 ms, mean ± s.d., $n = 54$). Moreover, membrane fluctuations were faster in the delay epoch compared to the pre-sample epoch (Extended Data Fig. 3c, d, k, l). Long membrane time constants therefore do not explain preparatory activity.

Other cell-autonomous mechanisms could account for persistent activity[22,32,33]. First, spike bursts can activate voltage-dependent channels to trigger cell-autonomous persistent activity[22,32,33]. However, only a small proportion of neurons (12 out of 79 cells, 4 out of 20 selective cells) showed bursts (at least 1 complex spike per 5 trials; Extended Data Fig. 3e, f, m, n), and bursts did not increase during the delay epoch (Extended Data Fig. 3g, h, o, p). Therefore, spike bursts do not contribute to persistent activity during the delay epoch.

Second, cell-autonomous mechanisms underlying persistent activity are expected to involve conductances activated by depolarization[22,32,33]. Persistent activity should then be perturbed by hyperpolarization[32]. For a subset of recordings with sufficiently long

[1]Janelia Research Campus, HHMI, Ashburn, VA, USA. *e-mail: romanis@janelia.hhmi.org; svobodak@janelia.hhmi.org

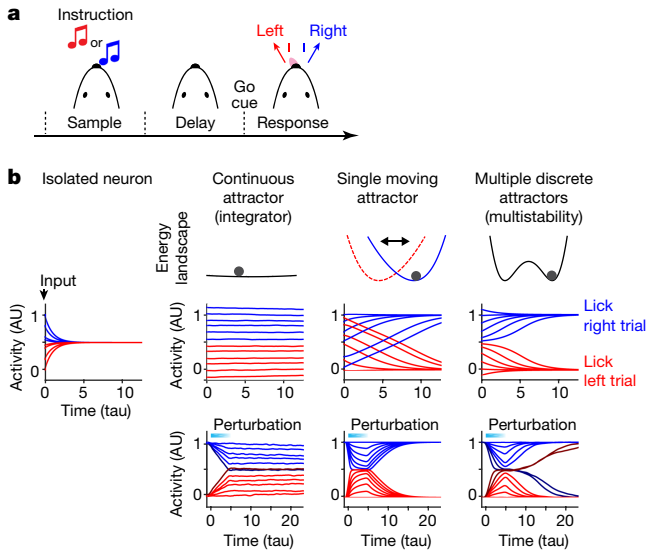


Fig. 1 | Models of persistent preparatory activity. a, Behavioural task. The instruction (tones for the auditory task, Figs. 2–6; pole location for the tactile task, Extended Data Figs. 1–4) was presented during the sample epoch. The mouse reported its decision after the delay epoch by directional licking. The duration of the delay epoch was 1.2 s (Figs. 2–4), 2.0 s (Fig. 5), or randomized (Fig. 6). **b**, Potential mechanisms underlying persistent activity. Left, in an isolated neuron, activity caused by a brief input (arrow, at time 0) decays following the membrane time constant (tau) of the cell. Excitatory feedback can compensate for the decay to produce a continuous or discrete attractor dynamics (Methods). Right, top, energy landscapes of networks; middle, activity trajectories following a brief input at time 0; bottom, activity trajectories following a brief perturbation (cyan bar).

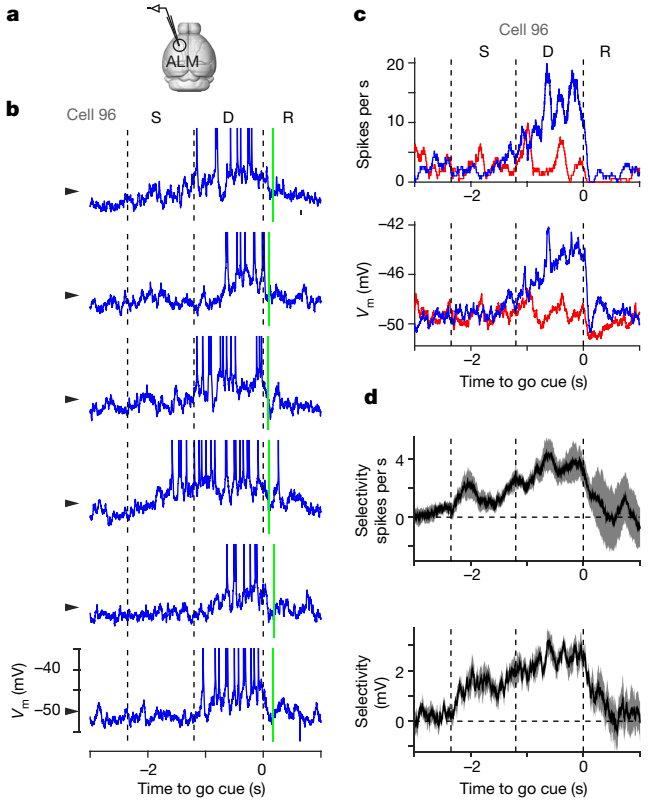


Fig. 2 | Whole-cell recordings in ALM. a, Whole-cell recordings from left ALM. **b**, Membrane potential (V_m), six example trials (lick-right). Time is aligned to the go cue. Dashed lines separate trial epochs. D, delay epoch; R, response epoch; S, sample epoch. Green lines denote first licks. **c**, Top, mean spike rate; bottom, mean membrane potential (same cell as **b**) ($n = 19$ lick-right trials, and 17 lick-left trials). Blue lines denote correct lick-right trials; red lines denote correct lick-left trials. **d**, Selectivity of ALM neurons based on spike rate (top) and V_m (bottom) in the auditory task ($n = 10$ cells). Shaded area denotes s.e.m. (bootstrap).

durations, we hyperpolarized cells by 12.0 ± 4.9 mV (mean ± s.d., $n = 10$). Hyperpolarized cells ceased to fire spikes, but selectivity in V_m was similar after current injection[36] (Fig. 3, Extended Data Fig. 4a, d); differences in selectivity were probably caused by differences in input impedance (Extended Data Fig. 4b, c) and decreased inhibitory currents (hyperpolarization moved the membrane potential closer to the reversal potential for chloride). Although it is unlikely that we controlled membrane potential throughout the dendritic arbors[37], spike bursts—which require dendritic electrogenesis—were eliminated by somatic hyperpolarization (Extended Data Fig. 4e). These results indicate that spiking and conductances activated by depolarization are not necessary for selectivity in V_m. Altogether, cell-autonomous mechanisms by themselves are unlikely to account for persistent activity in ALM. Instead, network mechanisms (Fig. 1b) probably explain the persistent activity, possibly in collaboration with cell-autonomous mechanisms[21,38].

Funnelling of membrane potential

We now focus on network mechanisms (Fig. 1b). In a system that follows continuous attractor dynamics, activity during the delay epoch is expected to diffuse over time[20,39,40], resulting in an increase of across-trial variability (Extended Data Fig. 1). By contrast, fluctuations in V_m decreased during the delay epoch, and reached a minimum immediately after the go cue (Fig. 4a, b, Extended Data Fig. 5). This decrease was stronger in selective neurons, particularly in lick-right trials (Fig. 4c, d, Extended Data Fig. 5) (hierarchical bootstrap comparing across-trial fluctuations in the baseline and the delay epoch; Methods). The reduction in across-trial V_m fluctuations was not caused by a ceiling effect imposed by the spike threshold. First, the threshold linear V_m-to-spike rate relationship (Extended Data Fig. 2) indicates that the activity during the delay epoch was not close to saturation. Second, the reduction in across-trial V_m fluctuations was independent of the distance to spike threshold (Extended Data Fig. 5j).

A reduction in across-trial variability in spike count has been reported during motor planning in non-human primates[41]. However, the interpretation of fluctuations on the basis of spike counts is complex because both the across-trial variance of spike rate and the variance of spike generation contribute to the overall variance[39]. Our whole-cell recording data show a reduction in variability of V_m, which controls spike rate[39]. This is inconsistent with continuous attractor dynamics (except for special conditions, where the variance of input decreases over time; Extended Data Fig. 1x, w).

Robustness to perturbations

We next analysed population activity in ALM. Silicon probes were used to record multiple single units (755 units, 20 sessions, 6 mice; 37.8 ± 16.6 units, mean ± s.d., per behavioural session)[9] (Fig. 5a). For each recording session, we projected population activity of individual trials to the **CD**[9,11] (Fig. 5b, Extended Data Fig. 6a–d, Methods). The distribution of the end points of the trajectories at the go cue was bimodal (Fig. 5b, Extended Data Fig. 6e–g), with one peak corresponding to correct lick-right trials, and the other to correct lick-left trials. Therefore, activity along the **CD** moved towards discrete end points.

To distinguish among models, we performed bilateral optogenetic inactivation (photoinhibition; Methods) of ALM during the beginning of the delay epoch. Following perturbation, activity should remain shifted (continuous attractor), recover to the same end point (single moving attractor and multiple discrete attractors) or occasionally switch to the other end point (multiple discrete attractors) (Fig. 1b, bottom; Extended Data Fig. 1). Consistent with previous work[9,11], strong photoinhibition (1.5 mW per spot; median activity, 2.4% of baseline) resulted in near chance-level performance (Extended Data Fig. 7a, b), confirming that other brain regions cannot rescue preparatory activity

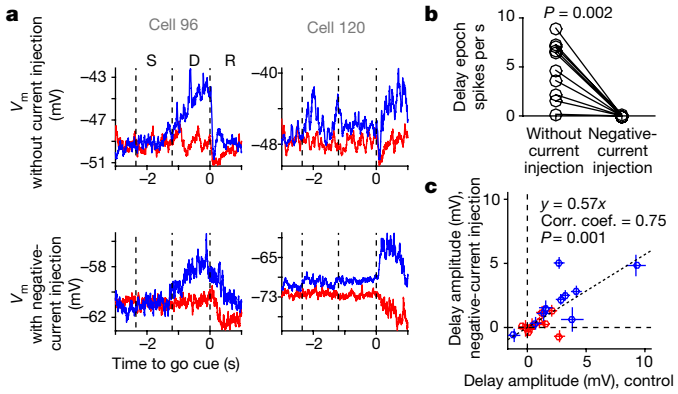


Fig. 3 | Hyperpolarization of ALM neurons. a, Two example ALM neurons. Top, mean V_m without current injection; bottom, mean V_m with negative current injection. Blue, correct lick-right trials; red, correct lick-left trials. **b**, Spike rate during the delay epoch with and without current injections. *P* value determined by two-sided Wilcoxon signed-rank test ($n = 10$ cells). **c**, Delay amplitude of V_m (delay epoch V_m minus pre-sample epoch V_m) with and without current injections ($n = 10$ cells). Correct lick-right trials (blue) and correct lick-left trials (red) are shown separately. Crosses denote s.e.m. (bootstrap). Dashed line denotes linear regression. Slope of linear regression, Pearson's correlation coefficient (corr. coef.), and the *t*-statistic of Pearson's correlation coefficient (*P*) are shown.

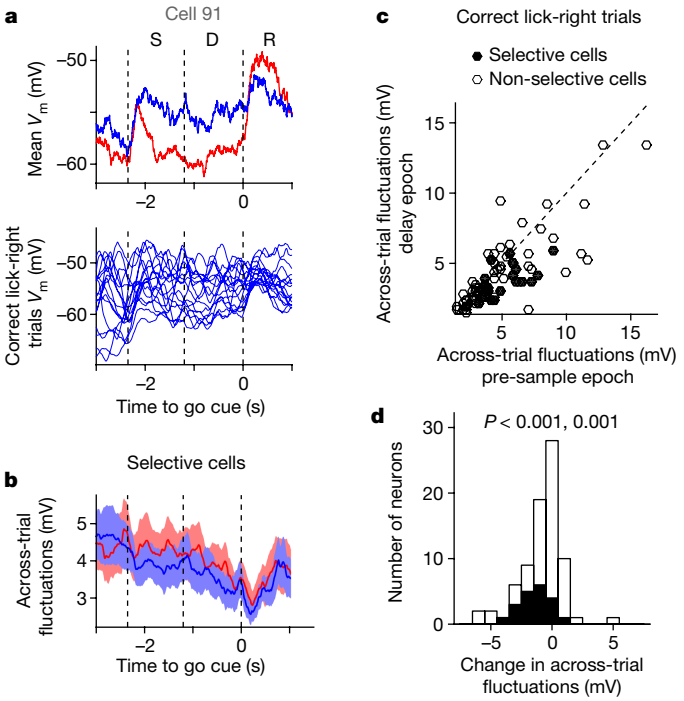


Fig. 4 | Funnelling of the membrane potential. a, Example ALM neuron. Top, mean V_m; bottom, all lick-right trials overlaid. V_m was averaged over 200 ms after spike removal. Blue lines denote correct lick-right trials; red lines denote correct lick-left trials. **b**, Across-trial fluctuations of selective cells in the auditory task ($n = 10$ cells). Line, mean of across-trial fluctuations among cells; shaded area denotes s.e.m. (hierarchical bootstrap, Methods). Across-trial fluctuations are the difference between the first and third quartile of all trials within the same trial type (blue denotes correct lick-right trials; red denotes correct lick-left trials). **c**, Relationship between across-trial fluctuations during the pre-sample epoch and during the delay epoch in correct lick-right trials. In **c** and **d**, $n = 79$ cells (both auditory and tactile tasks are pooled). Black denotes selective cells ($n = 20$ cells). **d**, Distribution of difference in across-trial fluctuations between the pre-sample epoch and the delay epoch (change in across-trial fluctuations). *P* values determined by one-sided Wilcoxon signed-rank test. The first *P* value, all cells (2.8×10^{-5}); second *P* value, selective cells only (5.2×10^{-4}).

in ALM[10,11]. Here we used more modest photoinhibition (0.1–0.3 mW per spot; median activity, 20–60% of baseline; Extended Data Fig. 7d–j). These perturbations transiently reduced selectivity (Fig. 5a, Extended Data Figs. 6d, 8b).

In perturbed trials that resulted in correct movement, selectivity recovered to the trajectories of unperturbed trials (Fig. 5a, Extended Data Fig. 8a, b, e). The recovery was slower for stronger photoinhibition (Extended Data Fig. 8c, Methods); this is predicted by discrete attractor models because the perturbation moves the activity trajectory closer to the separatrix, in which the dynamics are slower[42] (Extended Data Fig. 1t). In perturbed trials that resulted in an incorrect movement direction, neural activity switched trajectories to the other trial type (Fig. 5a, Extended Data Fig. 8e). Activity trajectories thus snapped to one of the two discrete end points for both behavioural outcomes (correct or incorrect). Similar to individual neurons, trajectories along the **CD** recovered in perturbed trials followed by correct movement (Fig. 5c, e, g). Trajectories recovered to the opposite end point in perturbed trials followed by incorrect licks (Fig. 5d, e, g). This bistability is consistent with two discrete attractors (Fig. 1b, right).

We next analysed how the activity along **CD** relates to the drift of trajectories along the **CD** (d(projection to **CD**)/dt in Fig. 5f)[42]. For continuous attractors, drift should be independent of the activity level (Extended Data Fig. 1h). For discrete attractors the drift should be directed towards the end point, faster far from the end points, and slow close to the end points (projection to **CD** near 0 or 1; Extended Data Fig. 1o, v). Therefore, the relationship between activity along **CD** and the drift should show negative slopes around the end points, which indicates stable fixed points. This was apparent for the lick-right end point (Fig. 5f, $P = 0.015$, Methods). This dynamic behaviour is consistent with discrete attractor dynamics.

The switching of **CD** trajectories distinguishes multiple discrete attractors from other models (Extended Data Fig. 1x). Because mice made errors even without perturbation (performance, 87.0 ± 8.4%, mean ± s.d.), some of the perturbed trials may reflect errors that are independent of the perturbation. We decoded the future licking direction on the basis of activity before the perturbation (decoder performance = 78.1 ± 4.5%, mean ± s.d.; Methods). Trials decoded to be correct before the perturbation showed switching in the end points in incorrect trials (Extended Data Fig. 6i–l), which indicates that some switches were caused by the perturbation. The recovery and switching of **CD** trajectories towards discrete end points after perturbation is consistent with multiple discrete attractors dynamics in ALM (Extended Data Fig. 1x).

Discrete attractor dynamics and ramping activity

In a variety of behavioural tasks, preparatory activity ramps up to a movement[3–9]. By contrast, standard discrete attractor models show stationary activity once the fixed points are reached (Fig. 1b, Extended Data Fig. 1). Ramping dynamics in discrete attractor networks can be obtained by the tuning of network parameters to generate slow drift to the fixed points[11] (internal ramping model; Extended Data Fig. 9i–o) or a non-selective ramping input that moves fixed points apart over time (external ramping model; Fig. 6a, Extended Data Fig. 9b–h). The rate of recovery from perturbations is expected to be independent of ramping dynamics in the external ramping model (Extended Data Fig. 9p–r).

We explored these possibilities further. Ramping predicts the timing of movement[5,43,44]. We performed a separate set of recordings with randomized delays (Fig. 6b, Extended Data Fig. 10a) (1,005 units, 23 sessions, 5 mice), which precludes prediction of the timing of movement. Similar to the fixed delay task[9], many neurons (390 out of 867 pyramidal neurons) showed selective persistent activity during the delay epoch. Spike rates and selectivity ramped up rapidly during the sample epoch, before the earliest possible go cue, and then remained near stationary during the delay epoch (Fig. 6c, d, Extended Data Fig. 8k, l). In addition, neither **CD** nor non-selective mode showed ramping-up during the delay epoch, unlike in the fixed delay task (Fig. 6e, Extended Data

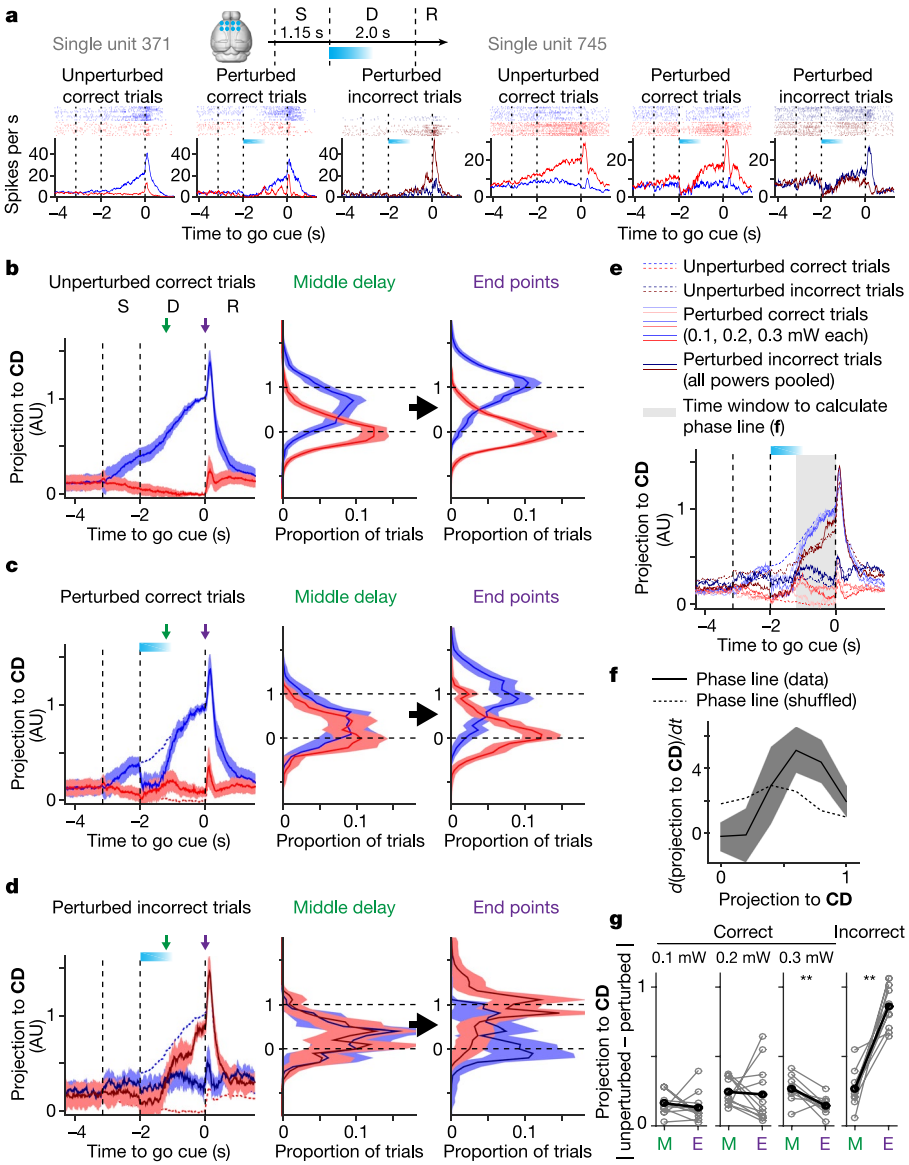


Fig. 5 | Robustness of discrete trajectories. a, Top, ALM was photoinhibited bilaterally during the first 600 ms of the delay epoch with 400 ms ramp down. Bottom, two example single units. For spike rasters, 15 trials per trial type were randomly selected. Solid line denotes mean spike rate. Blue, correct lick-right trials; red, correct lick-left trials; dark blue, incorrect lick-right trials; dark red, incorrect lick-left trials. In the incorrect lick-right trials, mice were instructed to lick right but licked left. Cyan bar denotes photoinhibition. **b**, Left, trajectories along **CD** (trials pooled across sessions, $n = 11$; Methods). Line denotes mean; shading denotes s.e.m. (hierarchical bootstrap). Middle, distribution of projection to **CD** in the middle of the delay epoch (1.2 s before the go cue, green arrowhead). Right, distribution of projection to **CD** at the end point (time 0, purple arrowhead). Shading denotes s.e.m. (hierarchical bootstrap). Blue, correct lick-right trials; red, correct lick-left trials. **c**, As in **b**, for perturbed correct trials. Dashed lines denote trajectories of unperturbed correct trials. **d**, As in **c**, for perturbed incorrect trials. Dark blue, incorrect lick-right trials; dark red, incorrect lick-left trials. **e**, Overlay of trajectories along the **CD**. **f**, Phase line of trajectories along the **CD** (Methods). Solid line, data; dashed line, shuffled data; shading denotes s.e.m. (hierarchical bootstrap, trials pooled across 11 sessions). **g**, Absolute difference in projection to the **CD** between perturbed and unperturbed trials at the middle delay (M; green) and end points (E; magenta). Grey, individual session; black, mean. $P = 0.46$, 0.46, 0.0098 and 0.00098 from left to right (two-sided Wilcoxon signed-rank test, $n = 11$ sessions). $^{**}P < 0.01$.

Figs. 8f–j, 10e–g). We repeated the perturbation experiments during the random delay task. For all conditions, the results were similar to those observed in the fixed delay task (Fig. 6f–h, Extended Data Figs. 6m–x, 10b–d). Furthermore, the speed of recovery from perturbations was similar between the fixed delay task (375 ± 100 ms, mean ± s.e.m.) and the random delay task (367 ± 249 ms, mean ± s.e.m.) (Extended Data Fig. 8c, h; $P = 0.377$, bootstrap). Therefore, ALM activity follows similar bistable discrete attractor dynamics for ramping or stationary dynamics during the delay epoch.

Discussion

The membrane potential dynamics was inconsistent with cell-autonomous mechanisms as a primary mechanism for persistent activity. During the delay epoch, activity funnelled towards one of two discrete end points. After perturbations, activity trajectories recovered to reach one of the end points. These experiments provide direct evidence for multiple discrete attractors as a mechanism that underlies short-term memory (Extended Data Fig. 1x).

ALM neurons with similar selectivity show high spike count correlations[9], suggesting preferential coupling. A nonlinear network with like-to-like coupling can produce discrete attractors without fine-tuning[15]. In addition to local connectivity, ALM is bidirectionally connected with thalamic nuclei that show similar preparatory activity[12]. Given the strong coupling between ALM and thalamus[12], our bilateral perturbations of ALM probably modified activity not only in ALM but also in thalamus. ALM is a crucial part of a multi-regional network that generates bistability[10].

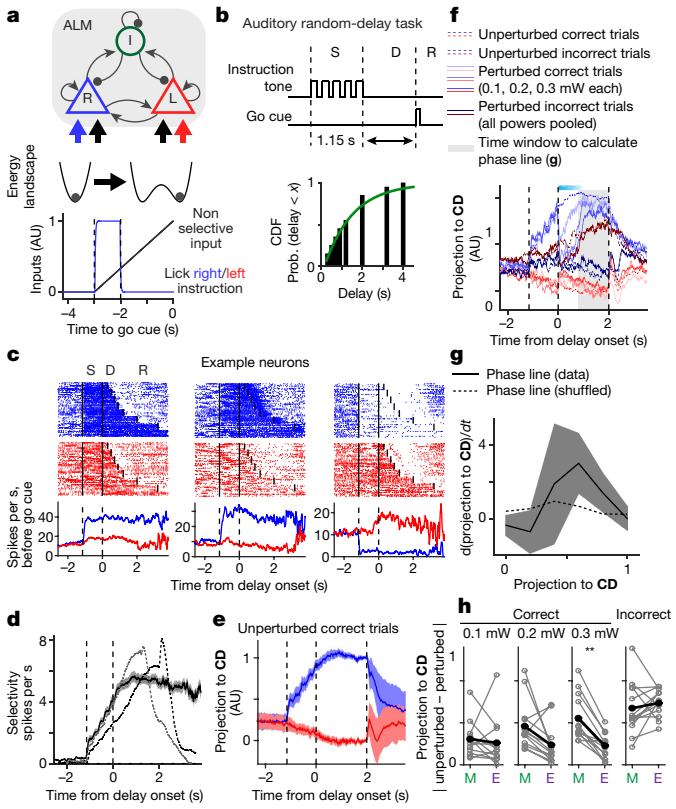


Fig. 6 | Stationary preparatory activity. a, Top, model circuit architecture. R, L and I denote lick-right selective excitatory neurons, lick-left selective excitatory neurons and inhibitory interneurons, respectively. Arrows, inputs; blue and red, lick-right and lick-left instruction; black, non-selective ramping input. Bottom, timing of input and energy landscape at the pre-sample epoch and the end of the delay epoch. **b**, Delays in each trial were randomly selected from eight durations (Methods). The cumulative distribution function (CDF) of delay durations (black bar) approximated the CDF the exponential distribution (green line). **c**, Three example ALM units. Top, spike raster. Lines separate trial epochs. Seven trials per delay duration were randomly selected. Bottom, mean spike rate before the go cue. Because delay durations were different across trials, the time axis is aligned to the onset of the delay epoch (same in **d**). Blue, correct lick-right trials; red, correct lick-left trial. **d**, Solid line, mean selectivity in the random delay task ($n = 390$ cells). Shading denotes s.e.m. (bootstrap). Dashed lines, selectivity in the fixed delay tasks (grey, 1.2 s; black, 2.0 s; data from ref. [9]). **e**, Trajectories along the **CD**. Trials with 2-s delay epoch in the random delay task are shown. Trials from all sessions were pooled (Methods, $n = 13$ sessions). Line denotes mean; shading denotes s.e.m. (hierarchical bootstrap). **f**, Overlay of trajectories along the **CD**. Trials with 2-s delay epoch in the random delay task are shown. **g**, Phase line of trajectories along the **CD**. Solid line denotes data; dashed line denotes shuffled data; shading denotes s.e.m. (hierarchical bootstrap, trials pooled across 13 sessions). **h**, Absolute difference in projection to the **CD** between perturbed and unperturbed trials at the middle delay (green) and end points (magenta). Grey, individual session; black, mean. $P = 0.21$, 0.040, 0.00024 and 0.34 from left to right (two-sided Wilcoxon signed-rank test, $n = 13$ sessions). $^{**}P < 0.01$.

In standard attractor models, the dynamics converge to stationary activity patterns. By contrast, neural activity during the delay epoch shows systematic ramping, creating a challenge for discrete attractor models[24,26]. Four observations suggest that ramping reflects an external input to ALM, which reflects the animal's expectation about timing in the task. First, ramping depends on the task structure. For example, in the random delay, task ramping is absent (Fig. 6)—this indicates that complex dynamics during the delay is not an obligatory component of preparatory activity[43]. Second, recovery from perturbations is faster than ramping itself. Third, despite the change in ramping dynamics, the speed of recovery from perturbations was similar between the two tasks. Fourth, the non-selective mode recovers to ramp even after bilateral silencing of ALM[11]. Ramping might be generated in a cortico-cerebellar loop[13].

In contrast to transient bilateral perturbations of ALM, transient unilateral perturbations early in the delay epoch have no behavioural effect[11]. Consistently, activity and selectivity recover rapidly after unilateral perturbation. Recovery relies on input from the contralateral ALM via the corpus callosum[11]. We created models in which modular discrete attractor networks are distributed across the two hemispheres, coupled via the corpus callosum, and receive non-specific input (Extended Data Fig. 9). These models account for the results presented here and in previous studies[11]. Although our models (Fig. 1 and Extended Data Figs. 1, 9) represent example implementations, the key predictions of the models (Extended Data Fig. 1x) are not expected to depend on the specific implementation.

Our task design only has two behavioural choices. This probably explains two stable end points in ALM dynamics. The attractor model can accommodate a large range of end points[28]. It is possible that each learned movement corresponds to a discrete attractor. Testing this hypothesis represents an important area for future investigation.

Previous studies have reported neural activity and behaviour consistent with discrete attractors[45–47]. For example, selective persistent activity in prefrontal cortex of primates is robust to sensory distractors[48], and remains discrete even with graded sensory stimuli[49,50]. Our perturbation experiments show that this robustness and discreteness are properties of circuits that involve the frontal cortex. Together, discrete attractor dynamics subserve short-term memory in the frontal cortex in a wide-range of behaviours.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-019-0919-7.

Received: 26 October 2017; Accepted: 7 January 2019;
Published online 6 February 2019.

1. Kubota, K. & Niki, H. Prefrontal cortical unit activity and delayed alternation performance in monkeys. *J. Neurophysiol.* **34**, 337–347 (1971).
2. Tanji, J. & Evarts, E. V. Anticipatory activity of motor cortex neurons in relation to direction of an intended movement. *J. Neurophysiol.* **39**, 1062–1068 (1976).
3. Funahashi, S., Bruce, C. J. & Goldman-Rakic, P. S. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* **61**, 331–349 (1989).
4. Romo, R., Brody, C. D., Hernandez, A. & Lemus, L. Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* **399**, 470–473 (1999).
5. Maimon, G. & Assad, J. A. A cognitive signal for the proactive timing of action in macaque LIP. *Nat. Neurosci.* **9**, 948–955 (2006).
6. Guo, Z. V. et al. Flow of cortical activity underlying a tactile decision in mice. *Neuron* **81**, 179–194 (2014).
7. Erlich, J. C., Bialek, M. & Brody, C. D. A cortical substrate for memory-guided orienting in the rat. *Neuron* **72**, 330–343 (2011).
8. Shenoy, K. V., Sahani, M. & Churchland, M. M. Cortical control of arm movements: a dynamical systems perspective. *Annu. Rev. Neurosci.* **36**, 337–359 (2013).
9. Inagaki, H. K., Inagaki, M., Romani, S. & Svoboda, K. Low-dimensional and monotonic preparatory activity in mouse anterior lateral motor cortex. *J. Neurosci.* **38**, 4163–4185 (2018).
10. Svoboda, K. & Li, N. Neural mechanisms of movement planning: motor cortex and beyond. *Curr. Opin. Neurobiol.* **49**, 33–41 (2018).
11. Li, N., Daie, K., Svoboda, K. & Druckmann, S. Robust neuronal dynamics in premotor cortex during motor planning. *Nature* **532**, 459–464 (2016).
12. Guo, Z. V. et al. Maintenance of persistent activity in a frontal thalamocortical loop. *Nature* **545**, 181–186 (2017).
13. Gao, Z. et al. A cortico-cerebellar loop for motor planning. *Nature* **563**, 113–116 (2018).
14. Johnston, D. & Wu, S. M.-S. *Foundations of Cellular Neurophysiology* (MIT Press, Cambridge, 1995).
15. Chaudhuri, R. & Fiete, I. Computational principles of memory. *Nat. Neurosci.* **19**, 394–403 (2016).
16. Lim, S. & Goldman, M. S. Balanced cortical microcircuitry for maintaining information in working memory. *Nat. Neurosci.* **16**, 1306–1314 (2013).

17. Wang, X. J. Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.* **24**, 455–463 (2001).
18. Goldman, M. S. Memory without feedback in a neural network. *Neuron* **61**, 621–634 (2009).
19. Amit, D. J. & Brunel, N. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb. Cortex* **7**, 237–252 (1997).
20. Wimmer, K., Nykamp, D. Q., Constantinidis, C. & Compte, A. Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nat. Neurosci.* **17**, 431–439 (2014).
21. Koulakov, A. A., Raghavachari, S., Kepecs, A. & Lisman, J. E. Model for a robust neural integrator. *Nat. Neurosci.* **5**, 775–782 (2002).
22. Zylberberg, J. & Strowbridge, B. W. Mechanisms of persistent activity in cortical circuits: possible neural substrates for working memory. *Annu. Rev. Neurosci.* **40**, 603–627 (2017).
23. Cannon, S. C., Robinson, D. A. & Shamma, S. A proposed neural network for the integrator of the oculomotor system. *Biol. Cybern.* **49**, 127–136 (1983).
24. Brody, C. D., Romo, R. & Kepecs, A. Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations. *Curr. Opin. Neurobiol.* **13**, 204–211 (2003).
25. Machens, C. K., Romo, R. & Brody, C. D. Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science* **307**, 1121–1124 (2005).
26. Barak, O. & Tsodyks, M. Working models of working memory. *Curr. Opin. Neurobiol.* **25**, 20–24 (2014).
27. Mongillo, G., Barak, O. & Tsodyks, M. Synaptic theory of working memory. *Science* **319**, 1543–1546 (2008).
28. Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl Acad. Sci. USA* **79**, 2554–2558 (1982).
29. Amari, S. I. Learning patterns and pattern sequences by self-organizing nets of threshold elements. *IEEE Trans. Comput.* **C21**, 1197–1206 (1972).
30. Ben-Yishai, R., Bar-Or, R. L. & Sompolinsky, H. Theory of orientation tuning in visual cortex. *Proc. Natl Acad. Sci. USA* **92**, 3844–3848 (1995).
31. Seung, H. S. How the brain keeps the eyes still. *Proc. Natl Acad. Sci. USA* **93**, 13339–13344 (1996).
32. Loewenstein, Y. et al. Bistability of cerebellar Purkinje cells modulated by sensory stimulation. *Nat. Neurosci.* **8**, 202–211 (2005).
33. Egorov, A. V., Hamam, B. N., Fransen, E., Hasselmo, M. E. & Alonso, A. A. Graded persistent activity in entorhinal cortex neurons. *Nature* **420**, 173–178 (2002).
34. Kim, S. S., Rouault, H., Druckmann, S. & Jayaraman, V. Ring attractor dynamics in the *Drosophila* central brain. *Science* **356**, 849–853 (2017).
35. Anderson, J. S., Lampl, I., Gillespie, D. C. & Ferster, D. The contribution of noise to contrast invariance of orientation tuning in cat visual cortex. *Science* **290**, 1968–1972 (2000).
36. Aksay, E., Gamkrelidze, G., Seung, H. S., Baker, R. & Tank, D. W. *In vivo* intracellular recording and perturbation of persistent activity in a neural integrator. *Nat. Neurosci.* **4**, 184–193 (2001).
37. Williams, S. R. & Mitchell, S. J. Direct measurement of somatic voltage clamp errors in central neurons. *Nat. Neurosci.* **11**, 790–798 (2008).
38. Camperi, M. & Wang, X. J. A model of visuospatial working memory in prefrontal cortex: recurrent network and cellular bistability. *J. Comput. Neurosci.* **5**, 383–405 (1998).
39. Churchland, A. K. et al. Variance as a signature of neural computations during decision making. *Neuron* **69**, 818–831 (2011).
40. Burak, Y. & Fiete, I. R. Fundamental limits on persistent activity in networks of noisy neurons. *Proc. Natl Acad. Sci. USA* **109**, 17645–17650 (2012).
41. Churchland, M. M., Yu, B. M., Ryu, S. I., Santhanam, G. & Shenoy, K. V. Neural variability in premotor cortex provides a signature of motor preparation. *J. Neurosci.* **26**, 3697–3712 (2006).
42. Strogatz, S. H. *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry, and Engineering* (Addison-Wesley, Boston, 1994).
43. Janssen, P. & Shadlen, M. N. A representation of the hazard rate of elapsed time in macaque area LIP. *Nat. Neurosci.* **8**, 234–241 (2005).
44. Komura, Y. et al. Retrospective and prospective coding for predicted reward in the sensory thalamus. *Nature* **412**, 546–549 (2001).
45. Wills, T. J., Lever, C., Cacucci, F., Burgess, N. & O'Keefe, J. Attractor dynamics in the hippocampal representation of the local environment. *Science* **308**, 873–876 (2005).
46. Kopec, C. D., Erlich, J. C., Brunton, B. W., Deisseroth, K. & Brody, C. D. Cortical and subcortical contributions to short-term memory for orienting movements. *Neuron* **88**, 367–377 (2015).
47. Piet, A. T., Erlich, J. C., Kopec, C. D. & Brody, C. D. Rat prefrontal cortex inactivations during decision making are explained by bistable attractor dynamics. *Neural Comput.* **29**, 2861–2886 (2017).
48. Suzuki, M. & Gottlieb, J. Distinct neural mechanisms of distractor suppression in the frontal and parietal lobe. *Nat. Neurosci.* **16**, 98–104 (2013).
49. Freedman, D. J., Riesenhuber, M., Poggio, T. & Miller, E. K. Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* **291**, 312–316 (2001).
50. Shadlen, M. N. & Newsome, W. T. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* **86**, 1916–1936 (2001).

Acknowledgements We thank N. Brunel, S. Lim, N. Li, G. Card, M. Economo, K. Daie, J. Yu, T. Wang, L. Liu, A. Ebihara and A. Finkelstein for comments on the manuscript, M. Inagaki for animal training, T. Harris, B. Barbarits, J. J. James and W. L. Sun for help with silicon probe recordings and spike sorting, and D. Hansel and S. Druckmann for discussions. This work was funded by Howard Hughes Medical Institute. H.K.I is a Helen Hay Whitney Foundation postdoctoral fellow.

Reviewer information *Nature* thanks Ila P. Fiete and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions H.K.I. performed experiments and analysed data, with input from all the authors. L.F. and S.R. performed network modelling. H.K.I. and K.S. wrote the paper, with input from all the authors.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-019-0919-7.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-0919-7.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to S.R. or K.S.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

METHODS

Mice. This study is based on data from 31 male mice (aged from postnatal day (P) 60). We used four transgenic mouse lines: PV-IRES-Cre[51], Ai32 (Rosa-CAG-LSL-ChR2(H134R)-eYFP-WPRE, JAX 012569)[52], Gad2-Cre (a gift from B. Zemelman), and VGAT-ChR2-eYFP[53] (see Supplementary Tables 1–3 for details).

All procedures were in accordance with protocols approved by the Janelia Institutional Animal Care and Use Committee. Detailed information on water restriction, surgical procedures and behaviour have been published[6,54]. Surgical procedures were carried out aseptically under 1–2% isofluorane anaesthesia. Buprenorphine HCl (0.1 mg kg^{-1}, intraperitoneal injection; Bedford Laboratories) was used for postoperative analgesia. Ketoprofen (5 mg kg^{-1}, subcutaneous injection; Fort Dodge Animal Health) was used at the time of surgery and postoperatively to reduce inflammation. After the surgery, mice were allowed free access to water for at least three days before start of water restriction. Mice were housed in a 12 h:12 h reverse light:dark cycle and behaviourally tested during the dark phase. A typical behavioural session lasted 1–2 h and mice obtained all of their water in the behaviour apparatus (approximately 1 ml per day; 0.3 ml was supplemented if mice drank less than 0.5 ml). On other days, mice received 1 ml water per day. Mice were implanted with a titanium headpost. For ALM photoinhibition, mice were implanted with a clear skull cap[6]. Craniotomies for recording were made after behavioural training.

Behaviour. Mice were trained using established procedures[54]. For the auditory task (Figs. 2–6, Extended Data Figs. 2–8, 10), at the beginning of each trial, five tones were presented at one of two frequencies: 3 or 12 kHz (for lick-right and lick-left trials, respectively). Each tone was played for 150 ms with 100 ms intertone intervals. The sample epoch (1.15 s total) was the time from onset of the first tone to the end of the last tone. The following delay epoch lasted for another 1.2 s (for whole-cell recordings; Figs. 2–4, Extended Data Figs. 2–5), 2.0 s (for silicon probe recordings; Fig. 5, Extended Data Figs. 6–8), or various durations (for the random delay task; see below). An auditory go cue (carrier frequency 6 kHz, with 360 Hz modulating frequency, to make it distinct from instruction tones) separated the delay and the response epochs (0.1 s). To compensate the sound intensity for tuning curve of C57BL6 auditory system[55], the sound pressure was 80, 70 and 60 dB for 3, 6 and 12 kHz sound, respectively. These frequencies are relatively invulnerable to hearing loss observed in C57BL6 mice[56].

For the random delay task (Fig. 6, Extended Data Figs. 6, 8, 10), delay durations were randomly selected from eight values (0.3, 0.5, 0.7, 0.9, 1.2, 2, 3.2 and 4 (or 5) s; Supplementary Table 3). Probability of the delay durations mimicked cumulative distribution function of the exponential distribution ($\tau = 1.4$ s) with 0.2 s offset (Fig. 6b). Because the hazard function of the exponential distribution is constant, animals cannot predict the timing of the go cue (or reward). Performance in this random delay task was similar to that in a task with a fixed delay duration, or fixed delay task (Extended Data Fig. 10a). In Extended Data Fig. 8l (random delay (short)), we analysed units recorded during a random delay task with shorter delay distributions ($\tau = 0.9$ s with 0.3 s offset), which were randomly selected from six values (0.3, 0.6, 1.2, 1.8, 2.4 and 3.6 s; Supplementary Table 3).

For the tactile task (for whole-cell recordings; Extended Data Figs. 2–5), at the beginning of each trial, a metal pole (diameter, 0.9 mm) moved within reach of the whiskers (0.2-s travel time) for 1.0 s, after which it was retracted (0.2-s retraction time)[6]. The sample epoch (1.4 s total) was the time from onset of the pole movement to completion of the pole retraction. The delay epoch lasted for another 1.2 s after completion of pole retraction. An auditory go cue separated the delay and the response epochs (pure tone, 3.4 kHz, 0.1 s).

The pre-sample (baseline) epoch was 1.2 s long, unless otherwise described. A two-spout lick-port (4.5 mm between spouts) was used to record licking events and deliver water rewards. After the go cue, licking the correct lick-port produced a water reward (approximately 2 μl); licking the incorrect lick-port triggered a timeout (0–5 s). Licking early during the trial (early lick trials) was punished by a timeout (1 s). Trials in which mice did not lick within 1.5 s after the go cue (no-response trials) were rare and typically occurred at the end of behavioural sessions. These no-response and early lick trials were excluded from analyses. For the random delay tasks, mice were first trained with a fixed delay duration of 1.2 s. After criterion performance was reached (80% correct), we switched to the random delay task and trained at least one additional week before recordings.

Photoinhibition. For the fixed delay task, photoinhibition was deployed on 25–33% behavioural trials (Fig. 5, Extended Data Figs. 6, 8). For the random delay task, photoinhibition was deployed in 25% of trials with 2 s delay duration (Fig. 6, Extended Data Figs. 6, 8, 10). To prevent mice from distinguishing photoinhibition trials from control trials using visual cues, a masking flash (1-ms pulses at 10 Hz) was delivered using 470-nm LEDs (Luxeon Star) throughout the trial. Photostimuli from a 473-nm laser (Laser Quantum) were controlled by an acousto-optical modulator (Quanta Tech).

Photoinhibition of ALM was performed through the clear-skull cap (beam diameter at the skull: 400 μm at 4σ). The light transmission through the intact skull is approximately 50%[6]. We stimulated parvalbumin-positive interneurons in PV-IRES-Cre mice crossed to Ai32 reporter mice expressing ChR2. Behavioural and electrophysiological experiments showed that photoinhibition in the PV-IRES-Cre × Ai32 mice was indistinguishable from the VGAT-ChR2-eYFP mice[12].

To silence ALM bilaterally during early or late delay, we photostimulated for 0.6 s (40 Hz photostimulation with a sinusoidal temporal profile) with 0.4 s ramping down, starting at the beginning of the delay epoch or 1.0 s after the beginning of the delay epoch, respectively. We photostimulated four spots in each hemisphere, centred on ALM (anteroposterior (AP) 2.5 mm; medial lateral (ML) 1.5 mm, bregma) with 1 mm spacing (in total eight spots bilaterally) using scanning Galvo mirrors. We photoinhibited each spot sequentially at the rate of 5 ms per step. The laser power noted in the figures and text indicates the time-averaged laser power per spot. The total laser power was therefore eightfold higher. The bilateral manipulation prevents rescue of neural dynamics from unaffected regions[11]. With the modest intensities of laser power we used, rebound spiking[6] was small and did not cause early behavioural responses (Extended Data Fig. 7c, e, g).

Behavioural data analysis. Behavioural performance was the proportion of correct trials, excluding early lick and no-response trials (Extended Data Figs. 7b, 10a). Early lick rate (Extended Data Figs. 7c, 10a) was the proportion of early lick trials excluding no-response trials. For Extended Data Fig. 7c, only the early licks during the last 1 s of the delay epoch (after photoinhibition) were counted. No-response rate (Extended Data Fig. 10a) was the proportion of no-response trial. Photoinhibition decreased performance in both fixed delay and random delay tasks ($P = 0.00074$ and $P = 0.0035$, $n = 20$ and 23 sessions, respectively; two-sided Wilcoxon signed-rank test; comparing unperturbed trials versus trials with 0.3 mW photoinhibition; Supplementary Tables 2 and 3), without affecting the no-response rate ($P = 0.90$ and $P = 0.062$, respectively).

In vivo whole-cell recording. All recordings were made from the left hemisphere. Whole-cell recordings were made using pulled borosilicate glass (Sutter Instrument). A small craniotomy (100–300-μm diameter) was created over ALM under isofluorane anaesthesia and covered with cortex buffer (125 mM NaCl, 5 mM KCl, 10 mM glucose, 10 mM HEPES, 2 mM $MgSO_4$, 2 mM $CaCl_2$; adjust pH to 7.4). Whole-cell patch pipettes (7–9 MΩ) were filled with internal solution (in mM): 135 K-gluconate, 4 KCl, 10 HEPES, 0.5 EGTA, 10 Na_2-phosphocreatine, 4 Mg-ATP, 0.4 Na_2-GTP and 0.3% biocytin (293–303 mOsm, pH 7.3). The membrane potential was amplified (Multiclamp 700B, Molecular Devices) and sampled at 20 kHz using WaveSurfer (http://wavesurfer.janelia.org). Membrane potentials were not corrected for the liquid junction potential. After the recording, the craniotomy was covered with Kwik-Cast (World Precision Instruments). Each mouse was used for two or three recording sessions. Recordings were made 235–818 μm (521.6 ± 120.8 μm, mean ± s.d.) below the pia. Brief current injections (−100 pA, 100 ms) were applied at the end of each trial to measure input resistance, series resistance and membrane time constant[57].

For current injection experiments (Fig. 3), we partially compensated for series resistance and injected a ramping current until action potentials disappeared[58,59]. Actual membrane potential was calculated post hoc based on injected current and series resistance. Mean membrane potential was −48.9 ± 3.4 mV (mean ± s.d., $n = 10$) without current injection. We injected −200 ± 153 pA, resulting in $V_m = -60.9 \pm 4.5$ mV (mean ± s.d., $n = 10$). Series resistance did not change before and after current injections ($P = 0.447$, Wilcoxon rank sum test, $n = 10$). In 5 out of 10 recordings, we were able to release current injections at the end of experiments to confirm that (1) membrane potential returned to spontaneous levels; and (2) neurons still produced action potentials.

Whole-cell recording data analysis. Whole-cell recordings with more than 10 correct trials per direction (lick-right and lick-left trials) were analysed (21.5 ± 8.2 correct trials per direction, mean ± s.d., $n = 79$). Performance during recording was 85.0 ± 11.8% (mean ± s.d., $n = 79$). Only correct trials were analysed. Cells recorded during the tactile task ($n = 42$) have previously been reported[12] (Supplementary Table 1).

To measure membrane potential, spikes were clipped off (Figs. 2–4, Extended Data Figs. 2–5). Neurons that differentiated correct trial-types during the delay epoch based on spike counts were deemed as selective (20 out of 79 in ALM). To compute selectivity, we computed the difference in spike count (100-ms averaging window) or membrane potential between trial types (correct only) for each selective neuron (Fig. 2d, Extended Data Fig. 3b, j).

To obtain spike-triggered median of V_m (Extended Data Fig. 2), we selected spikes that were not preceded by any other spikes in a 50-ms window. To obtain the V_m-to-spike rate relationship, V_m and spike rate were averaged over a 50-ms sliding window. Mean spike rate was calculated for each step of V_m (1-mV step; mean was defined for a step with more than 500 data points). For statistics of spike-triggered median and the V_m-to-spike rate relationship, we tested the null hypothesis that spike-triggered median (or V_m-to-spike rate curves) in the pre-sample epoch and the delay epoch were identical. We performed hierarchical bootstrapping[60–62]: we first randomly selected trials with replacement and then spikes within a trial with

replacement. We measured the Euclidian distance of each bootstrapped spike-triggered median (or V_m-to-spike rate curves) from the mean in the pre-sample epoch. The proportion of bootstrap trials with higher Euclidian distance in the delay epoch compared to that in the pre-sample epoch is shown as the P value. The results were robust to changes in sliding bin size (50, 100 and 200 ms, data not shown).

Delay amplitude of spike rate or V_m (Fig. 3c, Extended Data Figs. 3, 4) was defined as delay epoch spike rate or V_m minus pre-sample epoch spike rate or V_m. Spike bursts (Extended Data Figs. 3, 4) were detected as depolarization events 5 mV higher than spike threshold lasting longer than 20 ms. V_m autocorrelation was calculated after spike clipping (Extended Data Fig. 3). The time constant of membrane fluctuations was based on the autocorrelation function (Extended Data Fig. 3) (time point when the function drops below $1/e$).

To obtain across-trial fluctuations (Fig. 4, Extended Data Fig. 5), we averaged V_m over a 100- or 200-ms sliding window. We used the first and third quartile difference (difference between 75% point and 25% point) or the trimmed standard deviation (standard deviation after trimming off maximum and minimum data point) to calculate the across-trial fluctuations at each time point. This procedure removed the effects of a few outlier trials with spike bursts. For statistics, we performed hierarchical bootstrapping: first we randomly selected cells with replacement, and second we randomly selected trials within a cell with replacement.

Extracellular electrophysiology. A small craniotomy (diameter, 0.5 mm) was made over the left ALM hemisphere one day before the recording session. Extracellular spikes were recorded using Janelia silicon probes with two shanks (250 μm between shanks) (A2x32-8mm-25-250-165). The 64-channel voltage signals were multiplexed, recorded on a PCI6133 board (National Instrument) and digitized at 400 kHz (14 bit). The signals were demultiplexed into 64 voltage traces sampled at 25 kHz and stored for offline analysis. One to eight recording sessions were obtained per craniotomy. Recording depth (between 800 and 1,100 μm) was inferred from manipulator readings. The craniotomy was covered with cortex buffer. The tissue was allowed to settle for at least five minutes before the recording started.

Extracellular recording data analysis. The extracellular recording traces were band-pass filtered (300–6,000 Hz). Events that exceeded an amplitude threshold (4 s.d. above the background) were sorted using JRClust[63].

For the fixed delay task with 2-s delay (Fig. 5, Extended Data Figs. 6, 8), in total 755 single units were recorded across 20 behavioural sessions from 6 mice (Supplementary Table 2). The same dataset has previously been analysed[9]. For the random delay task (Fig. 6, Extended Data Figs. 6, 8, 10), in total 1,307 single units were recorded across 34 behavioural sessions from 9 mice (Supplementary Table 3). For the calibration of photoinhibition without a behavioural task, 316 single units were recorded from 2 mice (Extended Data Fig. 7).

Spike widths were computed as the trough-to-peak interval in the mean spike waveform. The distribution of spike widths was bimodal (data not shown); units with width less than 0.35 ms were defined as putative fast-spiking neurons (261 out of 2,378), and units with width more than 0.5 ms were defined as putative pyramidal (regular-spiking) neurons (2,069 out of 2,378). This classification was verified by optogenetic tagging of GABAergic neurons[6] (Extended Data Fig. 7e). Units with intermediate spike widths (0.35–0.50 ms, 48 out of 2,378) were excluded from our analyses.

Neurons with significant selectivity (two-sided Wilcoxon rank sum test comparing spike counts in two correct trial types, $P < 0.05$) during the delay epochs were classified as selective cells. Selective cells were classified into lick-right preferring versus lick-left preferring, on the basis of their total spike counts during the delay epoch. To compute selectivity, we took the spike rate difference between two correct trial types for each selective neuron.

For peri-stimulus time histograms (Figs. 5, 6, Extended Data Fig. 8k), only correct trials were included. For the peri-stimulus time histograms and selectivity of the random delay task (Fig. 6, Extended Data Fig. 8k), only spikes before the go cue were pooled. Spikes were averaged over 100 ms with a 1-ms sliding window.

For Extended Data Fig. 8a, f, we compared spike rates of all unperturbed trials (both correct and incorrect trials, but not early lick or no-response trials) versus all perturbed trials using two-sided t-test. Cells with $P < 0.05$ was counted as significant cell at each time bin. For this plot, we did not correct for multiple comparisons.

For Extended Data Fig. 8d, i, we calculated selectivity for correct and incorrect trials during 1–2 s after the delay onset. We defined polar coordinates r and θ as below:

$$r = \sqrt{\text{selectivity in correct trials}^2 + \text{selectivity in incorrect trials}^2}$$

$$\theta = \tan^{-1}\left(\frac{\text{selectivity in correct trials}}{\text{selectivity in incorrect trials}}\right)$$

We analysed neurons with more than 10 trials for all trial types (correct lick-right, correct lick-left, incorrect lick-right and incorrect lick-left trials). Many selective neurons in ALM deviate from $\theta = -45°$, indicating mixed coding[9]. We defined cells with $-67.5° < \theta < -22.5°$ with $r > 2$ as preparatory cells because the selectivity changes based on outcomes[9]. We defined cells with $-22.5° < \theta < 22.5°$ with $r > 2$ as non-preparatory cells (Extended Data Fig. 8e, j). The switching of activity trajectories in perturbed incorrect trials was more prominent in preparatory cells.

To quantify the recovery time course of selectivity after photoinhibition (Extended Data Fig. 8c, h), we first calculated a difference in selectivity between unperturbed trials and perturbed trials. The recovery time constant was based on fitting of an exponential decay function to this difference in selectivity. We fitted the last 1.2 s of the delay epoch. Bootstrap was used to calculate the standard error of mean.

To plot the cumulative distribution function of increase in spike rate during the delay epoch (Extended Data Fig. 8l), we calculated changes in spike rate during the delay epoch for each unit. For the fixed delay task, we analysed the whole delay epoch (spike rate at the end of delay minus the spike rate at the delay onset divided by delay duration); for the random delay task, we analysed the first 2 s of the delay epoch (spike rate at 2 s after the delay onset minus the spike rate at the delay onset divided by two) in trials with delay durations longer than 2 s.

CD analysis. To calculate the **CD** for a population of n simultaneously recorded neurons, we identified an n-dimensional vector maximally distinguishing the two trial types (correct lick-right and correct lick-left trials), in the n-dimensional activity space. For each neuron, we calculated average spike rates in the correct lick-right and correct lick-left trials separately (with 100-ms average window and 10-ms sliding step). $\boldsymbol{r}_{\text{lick-right},\,t}$ and $\boldsymbol{r}_{\text{lick-left},\,t}$ are n-dimensional vectors of average spike rate at one time point. The difference in the mean response vector, $\boldsymbol{w}_t = \boldsymbol{r}_{\text{lick-right},\,t} - \boldsymbol{r}_{\text{lick-left},\,t}$, showed high correlation during the delay epoch[9,11] (Extended Data Fig. 6b, m). We averaged $\boldsymbol{w}_t$ during the last 600 ms of the delay epoch and normalized by its own norm to obtain the **CD**. For the fixed delay task with 2-s delay (Fig. 5, Extended Data Fig. 6), we calculated **CD** based on randomly selected 50% of unperturbed correct trials (fit trials). To obtain trajectories along **CD**, we projected spike rate in the remaining trials (test trials) to the **CD** as an inner product. Next, we switched the fit and test trials, and pooled the trajectories along **CD** for analysis. For the random delay task (Fig. 6, Extended Data Figs. 6, 10), we calculated **CD** based on trials with non-2-s delay durations. To obtain trajectories along the **CD**, we projected spike rate in the trials with 2-s delay duration. Therefore, only trials with 2-s delay duration were analysed further for the random delay task.

Approximately half of selective cells show preparatory activity[9], which anticipates upcoming movements, regardless of behavioural outcome (correct or incorrect). We only analysed sessions with more than 5 preparatory neurons (11 out of 20 sessions, 13 out of 23 sessions for the fixed delay task and the random delay task, respectively; Supplementary Tables 2 and 3) in Figs. 5, 6, Extended Data Figs. 6, 10b–g. Preparatory cells were defined as cells with $r > 2$ and $-22.5° > \theta > -67.5°$ (Extended Data Fig. 8d, i; also see above).

In Figs. 5, 6, and Extended Data Figs. 6, 10, projection to **CD** was normalized based on the value at the end point in each session. The distribution of end point values was bimodal (Figs. 5, 6, Extended Data Fig. 6). For Figs. 5b–e, 6e, f and Extended Data Figs. 6, 10, we pooled trials from all sessions (with preparatory cells) after normalizing projections to the **CD** in each session. We calculated median of these pooled trajectories. Hierarchical bootstrapping was used to estimate standard errors of mean: first, we randomly selected session number with replacement and, second, randomly select trial number within a session with replacement.

To construct phase lines (Figs. 5f, 6g, Extended Data Fig. 6k, v), we analysed trajectories of unperturbed and perturbed trial types shown in Figs. 5e, 6f and Extended Data Fig. 6j, u. We analysed the last 1,200 ms of the delay epoch. The drift of a trajectory at time t is defined as $dx_CD_t/dt = (x_CD_{t+1} - x_CD_{t-1})/2$, in which x_CD_t denotes projection to **CD** at time t. The time step was 10 ms. After pooling all time points and trajectories, the mean dx_CD_t/dt was calculated for each step of x_CD (0.2 (AU) step). Hierarchical bootstrapping was used to estimate s.e.m. values. For the shuffling control, we shuffled the relationship between x_CD and dx_CD_t/dt. To test the slope at the lick-right end point ($x_CD = 1$), we constructed phase lines by analysing trajectories of unperturbed and perturbed trial types, in which mice licked right (that is, correct lick-right trials and incorrect lick-left trials). We estimated the slope of phase line using linear regression. We tested the null hypothesis that the slope is higher than 0 (hierarchical bootstrap, $n = 1{,}000$ iterations): $P = 0.015$, 0.020, 0.014 and 0.050 in Figs. 5f, 6g, Extended Data Fig. 6k and 6v, respectively.

To decode future licking direction before the perturbation (Extended Data Fig. 6i–l, t–x), we analysed the values of **CD** projection at the last bin 50 ms before the delay onset ($x_CD_{-2.05}$). We calculated the chance that mice lick right as a function of $x_CD_{-2.05}$ (i) (step between $x_CD_{-2.05}$ (i) is 0.1 (AU); we calculated the probability of lick right for trials with $x_CD_{-2.05}$ value higher than $x_CD_{-2.05}$ (i)) based on unperturbed correct and incorrect trials (Extended Data Fig. 6i, t).

Using this relationship, we defined a decision boundary for lick-right direction as the smallest $x_CD_{-2.05}(i)$ with probability of licking right that was higher than 70%. Perturbed trials with $x_CD_{-2.05}$ values higher than the decision boundary are decoded as lick-right trials. The decision boundary for lick-left direction was based on a similar procedure. We defined the decision boundary and putative correct trials independently for each session. We estimated the decoder performance using leave-one-out procedures. Decoder performance was 78.1 ± 4.5% (mean ± s.d.) for the fixed delay task, and 72.8 ± 10.9% (mean ± s.d.) for the random delay task.

In Extended Data Figs. 6d, o, 10e–g, we analysed modes explaining the remaining activity variance. To calculate the non-selective ramping mode, we first found eigenvectors of the population activity matrix using singular value decomposition at each time point and averaged over the last 600 ms of the delay epoch. The data for singular value decomposition at each time point were an $n \times$ trial-number matrix of spike rate at each time point minus mean spike rate during the pre-sample epoch. The first eigenvectors from the singular value decomposition was rotated using the Gram–Schmidt process to be orthogonal to the **CD**. Because the projection to the first vector resulted in non-selective ramping activity in the fixed delay task (Extended Data Fig. 10f), we referred to this vector as **RM**[11]. To calculate **CD** during early delay (**CDe**), we averaged $\boldsymbol{w}_t$ during the first 1 s of the delay epoch. To calculate, perturbation mode (**PM**), we averaged $\boldsymbol{r}_{\text{perturbed trials, }t} - \boldsymbol{r}_{\text{unperturbed trials, }t}$ (difference in spike rate between perturbed and unperturbed trials at each time point. Here, we pooled both correct lick-right and correct lick-left trials) during the first 1 s of the delay epoch. To calculate, **CD** during stimulation (**CDs**), we averaged $\boldsymbol{r}_{\text{lick-right, perturbed trials, }t} - \boldsymbol{r}_{\text{lick-left, perturbed trials, }t}$ (difference in spike rate between perturbed correct lick-right and correct lick-left trials) during the first 1 s of the delay epoch. All these modes were normalized by their own norm and rotated using the Gram–Schmidt process to be orthogonal to the **CD** and to each other. In Extended Data Fig. 10f, g, projections to non-selective **RM** were normalized by the difference between pre-sample epoch to the end of delay epoch for each session. To calculate the normalized selectivity (Extended Data Fig. 6d, o), we first calculated the total selectivity as a square sum of the selectivity across neurons (square sum of the elements of an n-dimensional vector). Then we calculated the square of selectivity of the projection along each mode in unperturbed and perturbed trials, and divided it by the total selectivity in unperturbed trials.

Network models. We modelled two classes of dynamical systems to simulate the average activity of neurons in ALM: (i) continuous attractor networks (Fig. 1b, left, Extended Data Fig. 1b–h, w); and (ii) discrete attractor networks (Fig. 1b, Extended Data Fig. 1i–v). Furthermore, we constructed two different implementations of discrete attractor networks: a single moving attractor network (Fig. 1b, Extended Data Fig. 1i–o) that exhibits one stable fixed point at different locations in the phase space depending on trial type; a multiple discrete attractors network (Fig. 1b, Extended Data Fig. 1p–v) that exhibits two coexisting stable fixed points, corresponding to left and right licking. To further account for the results of unilateral ALM perturbation experiments that have previously been described[11], we considered a robust modular version of the architecture shown in Fig. 6a (Extended Data Fig. 9). Furthermore, we explored two potential mechanisms that may underlie slow ramping dynamics during the delay epoch: ramping was either caused by non-selective external input (external ramping model; Extended Data Fig. 9b–h) or a consequence of slow internal dynamics (internal ramping model; Extended Data Fig. 9i–o). We used well-established neural implementations of continuous or discrete attractor dynamical systems and reproduced key observables obtained from the recorded neural population. See Supplementary Table 4 for the parameters we chose to build the models described below. Many implementations are possible for each class of model[16,19,21,23,25,28–31,34,38,47,64–66]. We implemented networks with the simplest architecture, but our results (Extended Data Fig. 1x) hold for a large range of parameters and architectures (not shown). Note that our interpretation of the recovery dynamics following optogenetic perturbations might break down in particular situations with fine-tuned connectivity and non-normal dynamics[67]. Because the slow and fast modes are not orthogonal, a transient perturbation along the fast mode would generate a change in **CD** that quickly recovers following the perturbation. However, if the perturbation is even slightly off the fast mode, the slow mode would not recover to the unperturbed trajectory as in our models.

One-hemisphere discrete attractor model. We simulated the average activity of two excitatory populations, one selective to the right licking direction and one selective to the left licking direction, and one inhibitory population (Fig. 1b, Extended Data Fig. 1i–v). A simplified description of the membrane potential dynamics of excitatory population i, $h_i(t)$, was governed by the following nonlinear differential equation:

$$\tau_i \frac{dh_i(t)}{dt} = -h_i(t) + \sum_{j=L,R} W_{ij} G_E(h_j(t)) - W_{iI} G_I(h_I(t)) + I_E^{\text{nonsel}}(t) + I_i^{\text{sel}}(t) + \eta_i(t) \quad (1)$$

W_{ij} denotes the synaptic strength between postsynaptic population i and presynaptic population j, labels L and R indicate the lick-left and lick-right selective excitatory populations, label I indicates the inhibitory population, τ_i is the integration time constant of the excitatory populations, $I_E^{\text{nonsel}}(t)$ is the external non-selective input current, $\eta_i(t)$ is a random Gaussian noise with zero mean and variance $\sigma 2$. $\tau_E = 100$ ms is the integration time constant of the excitatory population. The time-dependent external input $I_i^{\text{sel}}(t)$ consisted of selective input representing the sensory stimulus (a 1-s long boxcar function exponentially filtered with time constant $\tau_{\text{exp}} = 20$ ms) that was delivered during the sample epoch to either lick-left or lick-right excitatory neurons, depending on the trial type. The peak amplitude of $I_i^{\text{sel}}(t)$ was drawn from a Gaussian distribution of mean μ_i^{sel} and standard deviation σ_i^{sel}. $G_E(h_{L,R})$ denotes the firing rates of the excitatory populations (see below). For the inhibitory population, we assumed instantaneous integration of the current received from the excitatory populations ($\tau_I = 0$ ms). We modelled the firing rate of the inhibitory population, denoted by the nonlinear transduction function G_I, as threshold-linear. The inhibitory firing rate can thus be written as:

$$G_I(h_I(t)) = [h_I(t)]_+ = \left[\frac{\sum_{j=L,R} W_{Ij} G_E(h_j(t)) + I_I^{\text{nonsel}} + I_I^{\text{stim}}(t)}{1 + W_{II}} \right]_+ \quad (2)$$

in which $I_I^{\text{stim}}(t)$ denotes the increase in the external input mimicking the optogenetic photostimulation. $I_I^{\text{stim}}(t)$ was modelled as a 600-ms long, exponentially filtered, positive input delivered at the end of the sample period. The value of the baseline input current I_I^{nonsel} was such that r_I was always greater than zero. For this reason, we replaced $G_I(h_I(t))$ in equation (1) with the argument of the threshold linear function in equation (2) to obtain:

$$\tau_i \frac{dh_i(t)}{dt} = -h_i(t) + \sum_{j=L,R} \tilde{W}_{ij} G_E(h_j(t)) + \tilde{I}_i^{\text{nonsel}}(t) + I_i^{\text{sel}}(t) + \eta_i(t), \quad \text{for } i,j = L,R \quad (3)$$

where now the synaptic strengths $\tilde{W}_{ij}$ and the baseline input $\tilde{I}_i^{\text{nonsel}}(t)$ are rescaled according to:

$$\tilde{W}_{ij} = W_{ij} - \frac{W_{iI} W_{Ij}}{1 + W_{II}}, \quad \tilde{I}_i^{\text{nonsel}}(t) = I_E^{\text{nonsel}}(t) - \frac{W_{iI} I_I^{\text{nonsel}}}{1 + W_{II}} \quad (4)$$

The synaptic weights were chosen such that the phase space contained one (single moving attractor) or two (multiple discrete attractor) stable fixed points during the sample and delay epochs, and one fixed point during the pre-sample epoch, corresponding to baseline activity. The network switched between these two regimes approximately 50 ms after the beginning of the sample epoch when the excitatory baseline currents I_E^{nonsel} were significantly increased.

Excitatory transduction function. We chose a sigmoidal transduction function $G_E(h_i)$ for the excitatory populations. This function has an interpretable parametrization, mimicking the effect of short-term synaptic plasticity[68]:

$$G_E(h_i) = \bar{u}_i \bar{x}_i g_i, g_i = g(h_i) \quad (5)$$

$$\bar{u}_i = \frac{U + U\tau_f g_i}{1 + U\tau_f g_i}, \bar{x}_i = \frac{1}{1 + \bar{u}_i \tau_D g_i} \quad (6)$$

Variables $\bar{u}_j$ and $\bar{x}_j$ are steady-state synaptic nonlinearities resulting from, respectively, short-term facilitation and depression of excitatory synapses; U is the synaptic release probability, τ_f and τ_D are the facilitation and depression recovery time constants (Supplementary Table 4). The activation function $g(x)$ displays an exponential behaviour when $e^{x/k} \ll 1$ and a linear one for $x \gtrsim k$:

$$g(x) = k \log(1 + e^{x/k}) \quad (7)$$

The results were qualitatively similar if we used a threshold-linear transduction function $G_E(h_i) = [h_i]_+$ (data not shown).

Modular discrete attractor model. To model robust modular networks, we linked two one-hemisphere attractor networks (Extended Data Fig. 1a) via mutually excitatory coupling (W_C) between populations with the same selectivity in both modules (Extended Data Fig. 9a). Each module thus contained one cell type selective for licking right and one cell type selective for licking left, both connected to an inhibitory cell. During the sample epoch the amplitude of the selective input, randomly varying from trial to trial $I_i^{\text{sel}}(t) = \mathcal{N}(\mu^{\text{sel}}, \sigma^{\text{sel}})$, was delivered to lick-left or lick-right excitatory neurons of both modules, depending on trial type. The particular choice of parameters determined whether—once the dynamics converged to the attractor—across-trial variability increased, decreased or remained constant in an attractor network, contrary to continuous attractor networks in which across-trial variability generally increases (see 'Continuous attractor model'). Our choice of parameters caused a decrease in across-trial variability with time.

Ramping activity. To reproduce the slow ramping of neural activity observed in the electrophysiological recordings (Fig. 5), we considered two potential mechanisms in our simulations: (i) a monotonically increasing nonselective external input (external ramping model; Extended Data Fig. 9b–h); and (ii) the existence of a weak vector field along the direction of input integration, capable of slowing down the network dynamics during sample and delay epochs (internal ramping model; Extended Data Fig. 9i–o). In (i), a nonselective input current $I^{\mathrm{ramp}}(t)$, linearly ramping during the sample epoch and plateauing at the end of the delay epoch, was added to the current $I_E^{\mathrm{nonsel}}(t)$ of all excitatory populations. We set the network parameters, including the peak amplitude of $I^{\mathrm{ramp}}(t)$ such that the network would develop two stable fixed points during the delay period, corresponding to left and right licking directions. A single fixed point, corresponding to the baseline firing rate, was present all along the pre-sample epoch. The nonselective ramping input destabilized the pre-sample fixed point creating two stable fixed points during the sample epoch, corresponding to the left and right licking directions (Fig. 6a, Extended Data Fig. 9b–h). Conversely, in (ii), the network parameters were chosen so that the neural activity would slowly converge towards the fixed points (Extended Data Fig. 9i–o). In this configuration, which demanded substantial fine tuning, the phase space displayed two constant decision-related fixed points (left or right licking) throughout the delay epoch. In the single moving attractor case, a selective input $I_i^{\mathrm{sel,ramp}}(t)$, linearly ramping during sample epoch and plateauing during the delay epoch, was added to the transient selective input to move the fixed point away from baseline activity.

Continuous attractor model. Continuous attractor dynamics was modelled using the negative derivative feedback mechanism (Fig. 1b, left). The firing rates of three populations (Extended Data Fig. 1a), were driven by recurrent synaptic inputs $s_{ij}(t)$:

$$\tau_i \frac{dr_i(t)}{dt} = -r_i(t) + \sum_{j=L,R,I} W_{ij}\, s_{ij}(t) + I_i^{\mathrm{sel}}(t) + \xi_i(t)$$
$$\tau_{ij} \frac{ds_{ij}(t)}{dt} = -s_{ij}(t) + r_j(t), \quad \text{for } i,j = L, R \text{ or } I \tag{8}$$

in which τ_{ij} is the recovery time constant of synaptic state variables connecting presynaptic population j to postsynaptic population i, the noise term $\xi_i(t)$ is a coloured Gaussian noise with zero mean and two-point autocorrelation function $\langle \xi_i(t)\ \xi_i(t+s)\rangle = \frac{\sigma^2}{2\tau_{\mathrm{noise}}} e^{-\frac{|s|}{\tau_{\mathrm{noise}}}}$, in which σ^2 is the variance of the Gaussian distribution and τ_{noise} is the decay constant of the autocorrelation function. The other variables in equation (8) have the same role as in the discrete attractor models. To maintain persistent activity during the delay epoch, the conditions were (i) balanced synaptic strengths; and (ii) time constants of positive feedback are longer than those associated with negative feedback (Supplementary Table 4). The linear system in equation (8) can thus display persistent activity lasting several seconds (Extended Data Fig. 1b, left). During lick-right trials, we delivered a positive input, with the same temporal dynamics as in the discrete attractor model, to the lick-right selective population, while a negative input was delivered during lick-left trials. To obtain the reduction in across-trial variability, we ran an additional set of simulations in which the variance of $\xi_i(t)$ decreased exponentially with time, starting from the end of the sample period (Extended Data Fig. 1w). The same reduction could, in principle, be accounted for in a negative derivative feedback network with nonlinear units close to saturation. Because we did not observe saturation in intracellular recordings, we did not model this situation.

Analysis of simulated activity. Numerical integration was performed using the Euler–Maruyama method for all attractor models and second-order Runge–Kutta method for the continuous attractor model. To mimic the optogenetic activation of inhibitory neurons, we chose—for each model—four values of $I_I^{\mathrm{stim}}(t)$ with increasing peak amplitude (see Supplementary Table 4). We simulated 1,000 trials, of which 500 were lick-right and 500 lick-left trials, for each condition and model. All trajectories were smoothed using a 100-ms average window with 1-ms sliding step. The **CD** was computed using the same procedure adopted in the analysis of extracellular recordings ($\boldsymbol{r}_{\text{lick-right}} - \boldsymbol{r}_{\text{lick-left}}$, then averaging over the last 400 ms of the delay epoch). Neuronal activity in each trial was then projected onto the **CD** to obtain activity traces and subsequently normalized using the distribution of end points in each session (see 'CD analysis'). As with experimental data, we then analysed correct and incorrect trials. Lick-left trials were classified as correct if the neural activity of the lick-left selective neuron was higher than the neural activity of lick-right selective neuron at the end of the delay period. In all attractor models, projected trajectories of trials classified as correct converged to the expected stable fixed point after bilateral or unilateral perturbation (Extended Data Figs. 1, 9). We calculated the drift of projected trajectories for each model using the last 1,200 ms of the delay epoch. We applied the same method as in experimental data. Fluctuations across trials were computed by taking the standard deviation across projected activity traces during correct trials, and normalizing them to their value at the beginning of the delay epoch (Extended Data Fig. 1e, l, s, w). Model performance was computed by dividing the number of correctly classified trials by the total number of trials in each condition (Extended Data Fig. 9h, o).

Statistics. The sample sizes are similar to sample sizes used in the field. No statistical methods were used to determine sample size. We did not exclude any mouse for data analysis. During experiments, trial types were randomly determined by a computer program. During spike sorting, experimenters cannot tell the trial type, so experimenters were blinded to conditions. All comparisons using Wilcoxon signed-rank and rank sum tests were two-sided except for Fig. 4d and Extended Data Fig. 5b–e, in which we tested the decrease in values from the baseline. All bootstrapping was done over 1,000 iterations.

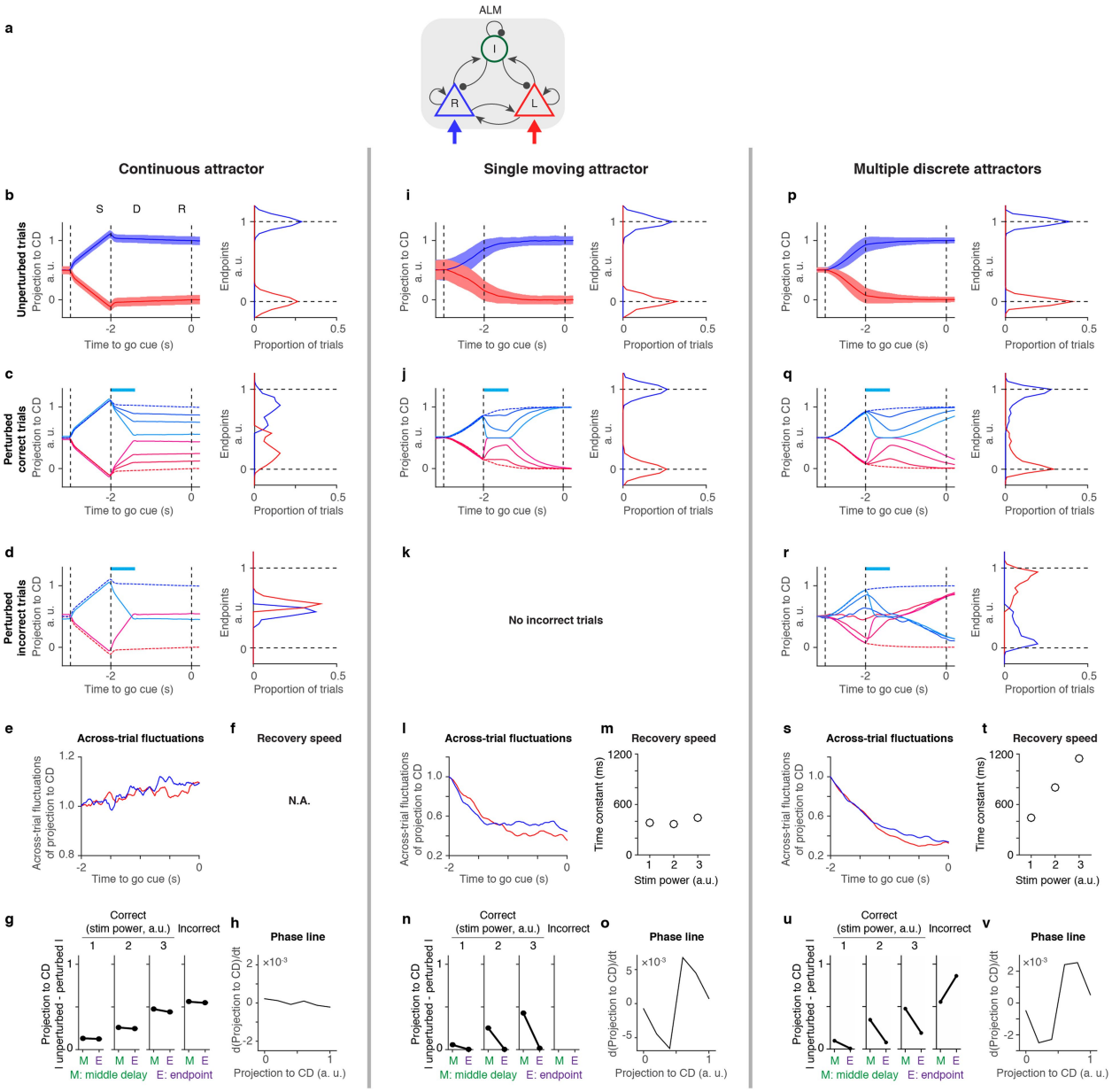
Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Code availability. For network models, Matlab code is available at the Github repository (https://github.com/fontolanl/ALM_attractors).

Data availability

Electrophysiological data are available at FigShare (https://figshare.com/; doi:10.25378/janelia.7489253).

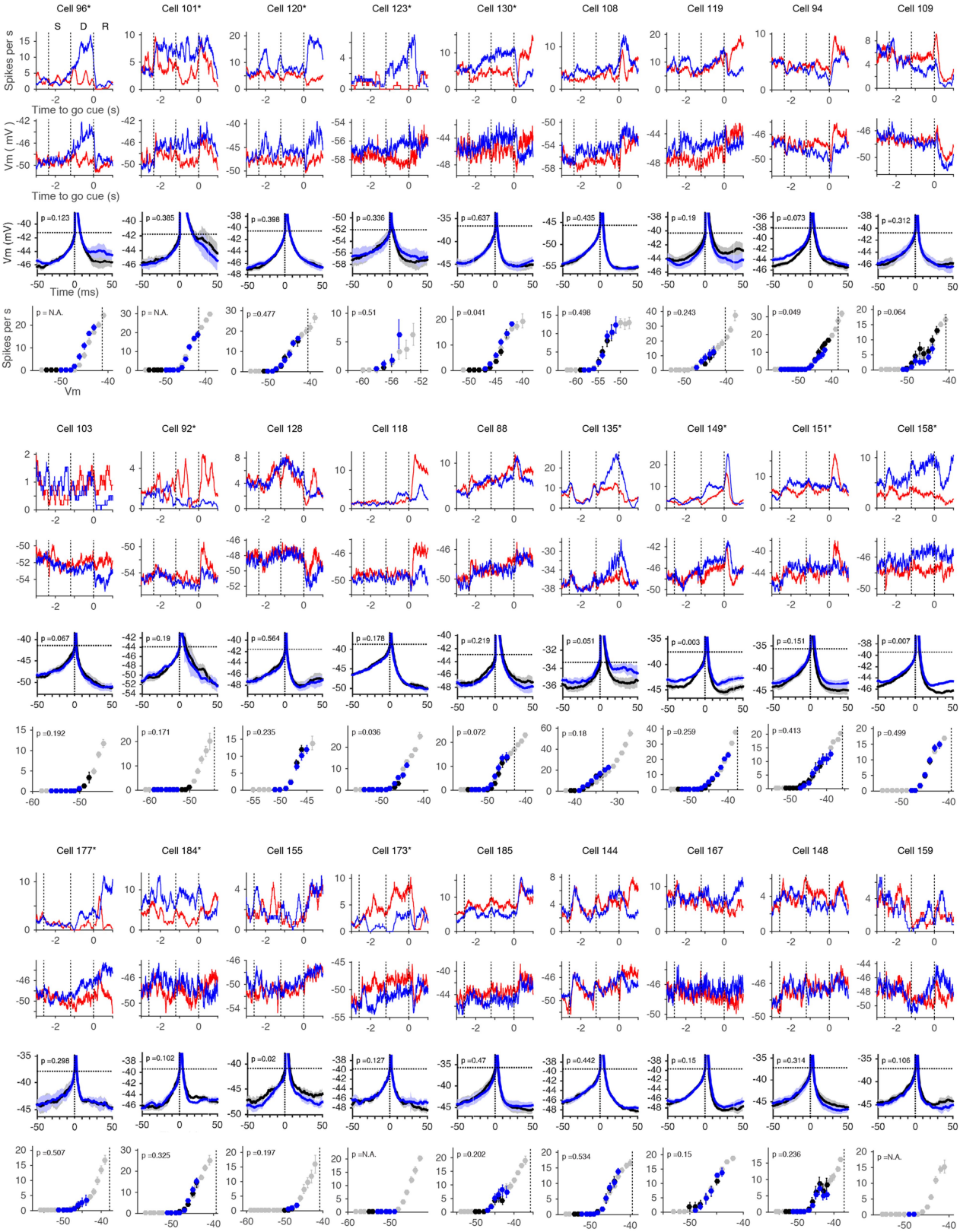
51. Hippenmeyer, S. et al. A developmental switch in the response of DRG neurons to ETS transcription factor signaling. *PLoS Biol.* **3**, e159 (2005).
52. Madisen, L. et al. A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. *Nat. Neurosci.* **15**, 793–802 (2012).
53. Zhao, S. et al. Cell type-specific channelrhodopsin-2 transgenic mice for optogenetic dissection of neural circuitry function. *Nat. Methods* **8**, 745–752 (2011).
54. Guo, Z. V. et al. Procedures for behavioral experiments in head-fixed mice. *PLoS ONE* **9**, e88678 (2014).
55. Taberner, A. M. & Liberman, M. C. Response properties of single auditory nerve fibers in the mouse. *J. Neurophysiol.* **93**, 557–569 (2005).
56. Ison, J. R., Allen, P. D. & O'Neill, W. E. Age-related hearing loss in C57BL/6J mice has both frequency-specific and non-frequency-specific components that produce a hyperacusis-like exaggeration of the acoustic startle reflex. *J. Assoc. Res. Otolaryngol.* **8**, 539–550 (2007).
57. Anderson, J. S., Carandini, M. & Ferster, D. Orientation tuning of input conductance, excitation, and inhibition in cat primary visual cortex. *J. Neurophysiol.* **84**, 909–926 (2000).
58. Yu, J., Gutnisky, D. A., Hires, S. A. & Svoboda, K. Layer 4 fast-spiking interneurons filter thalamocortical signals during active somatosensation. *Nat. Neurosci.* **19**, 1647–1657 (2016).
59. Monier, C., Chavane, F., Baudot, P., Graham, L. J. & Fregnac, Y. Orientation and direction selectivity of synaptic inputs in visual cortical neurons: a diversity of combinations produces spike tuning. *Neuron* **37**, 663–680 (2003).
60. Efron, B. & Tibshirani, R. *An Introduction to the Bootstrap,* 1 edn (Chapman and Hall/CRC, 1994).
61. van der Leeden, R. in *Handbook of Multilevel Analysis* (eds J. de Leeuw & E. Meijer) 401–433 (Springer, 2008).
62. Aarts, E., Verhage, M., Veenvliet, J. V., Dolan, C. V. & van der Sluis, S. A solution to dependency: using multilevel analysis to accommodate nested data. *Nat. Neurosci.* **17**, 491–496 (2014).
63. Jun, J. J. et al. Real-time spike sorting platform for high-density extracellular probes with ground-truth validation and drift correction. Preprint at https://www.bioRxiv.org/content/early/2017/01/19/101030 (2017).
64. Compte, A., Brunel, N., Goldman-Rakic, P. S. & Wang, X. J. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb. Cortex* **10**, 910–923 (2000).
65. Wilson, H. R. & Cowan, J. D. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **12**, 1–24 (1972).
66. Murray, J. D. et al. Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. *Proc. Natl Acad. Sci. USA* **114**, 394–399 (2017).
67. Murphy, B. K. & Miller, K. D. Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron* **61**, 635–648 (2009).
68. Barak, O. & Tsodyks, M. Persistent activity in neural networks with dynamic synapses. *PLoS Comput. Biol.* **3**, e35 (2007).



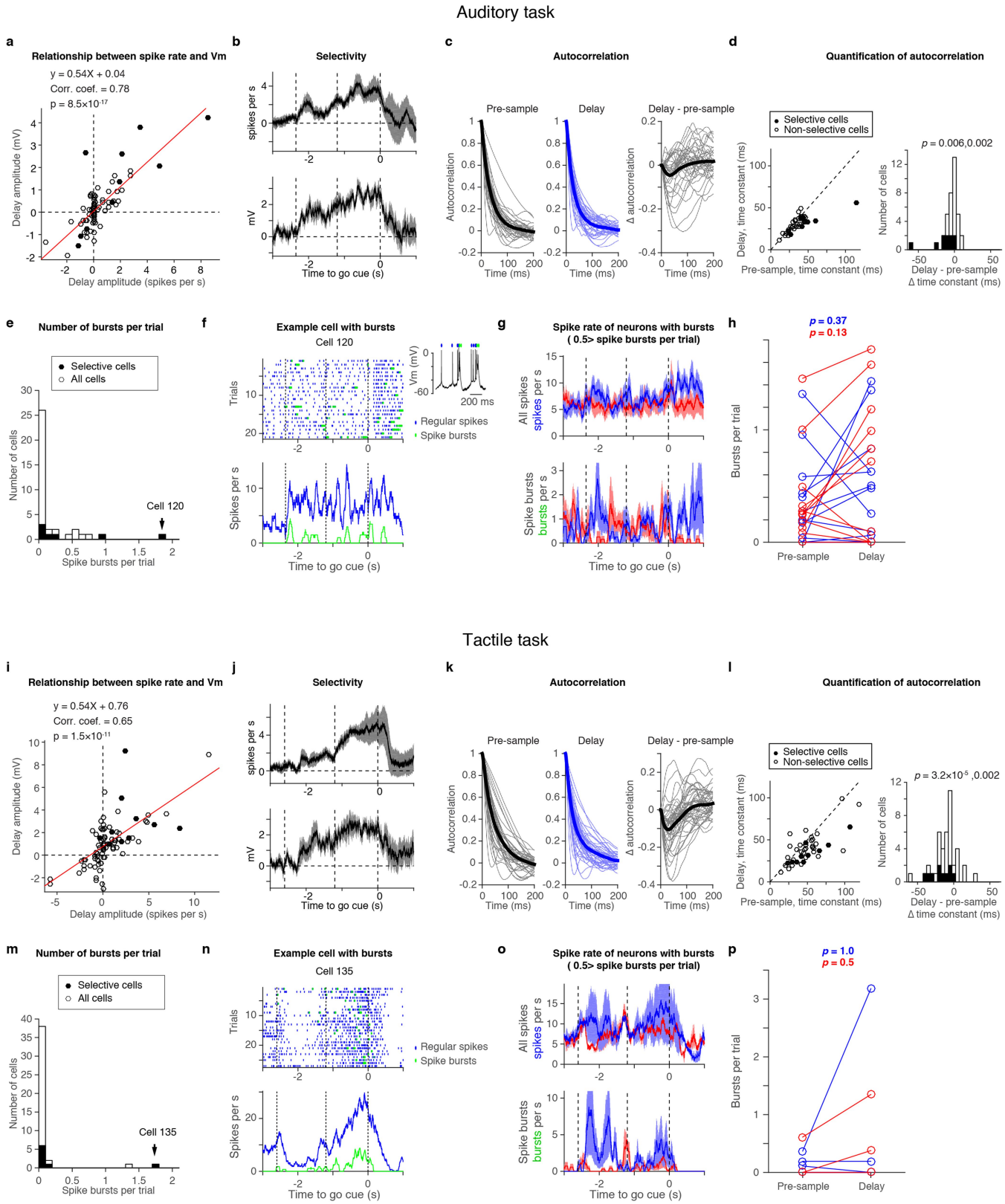
	Model predictions			Data	
	Continuous attractor	Single moving attractor	Multiple discrete attractors	Fixed-delay task	Random-delay task
Funneling	+ (EDF1e)	–/+ (EDF1l)	–/+ (EDF1s)	– (Fig. 4)	N.A.
Recovery	– (EDF1c, g)	+ (EDF1j, n)	+ (EDF1q, u)	+ (Fig. 5c,g)	+ (Fig. 6f,h)
Recovery speed v.s. stim power	N.A. (EDF1f)	-/+ (EDF1m)	+ (EDF1t)	+ (EDF8c)	+ (EDF8h)
Switching	– (EDF1d, g)	– (EDF1k, n)	+ (EDF1r, u)	+ (Fig. 5d,g; EDF6i-l)	+ (Fig. 6f,h; EDF6t-x)
Phase line (slope at 1)	+ (EDF1h)	– (EDF1o)	– (EDF1v)	– (Fig. 5f)	– (Fig. 6g)

Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Continuous and discrete attractor models (one hemisphere). a, Schematic of simulated networks. The same architecture was used for the continuous attractor model (**b**–**h**), single moving attractor model (**i**–**o**) and multiple discrete attractors model (**p**–**v**). L, lick-left selective excitatory neurons; I, inhibitory interneurons, R, lick-right selective excitatory neurons. Blue and red arrows indicate selective external input. **b**, Trajectories projected along the **CD**. Left, activity in unperturbed trials. Right, distribution of end points. Line denotes the mean; shading denotes s.d. Blue, correct lick-right trials; red, correct lick-left trials. **c**, Left, activity in perturbed correct trials. Right, distribution of end points. Cyan bar on top, photoinhibition. Traces with lighter colours correspond to higher intensity of photoinhibition. **d**, Left, activity in perturbed incorrect trials. Right, distribution of end points. **e**, Across-trial fluctuations during the delay epoch in unperturbed correct trials. Across-trial fluctuations are normalized for the value at 2.0 s before the go cue. **f**, Recovery speed of the **CD** in perturbed correct trials following perturbations. **g**, Absolute difference in projection to **CD** between perturbed and unperturbed trials at the middle delay (green) and end points (magenta). **h**, Phase line of trajectories. **i**–**o**, Dynamics of single moving attractor model. **p**–**v**, Dynamics of multiple discrete attractors model. **w**, Across-trial fluctuations in continuous attractor with decreasing external noise (Methods), normalized by their value 2.0 s before the go cue. **x**, A summary table comparing three models and data.

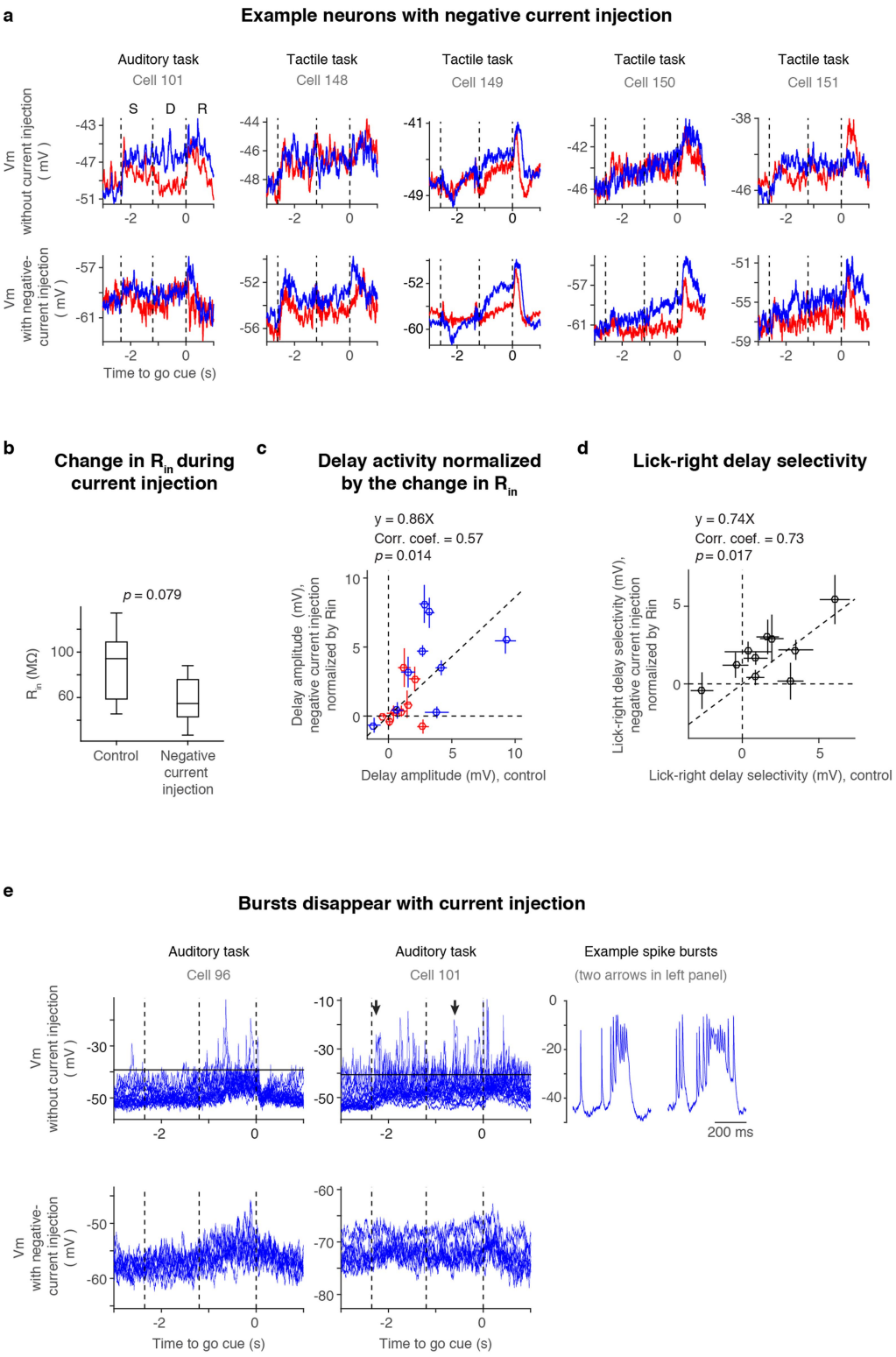


Extended Data Fig. 2 | Whole-cell recordings example cells. Columns represent data from one cell. Top, mean spike rate; second row, mean membrane potential (V_m). Blue, correct lick-right trials; red, correct lick-left trials; third row, spike-triggered median of V_m. Spikes in the pre-sample epoch (black) and the delay epoch (blue) in correct lick-right trials were analysed. Shading denotes s.e.m. (hierarchical bootstrap). *P* value denotes the probability of the null hypothesis that spike-triggered medians are the same between epochs (Methods, hierarchical bootstrap); fourth row, relationship between V_m and spike rate of the pre-sample epoch (black), the delay epoch (blue) and all epochs (grey). Error bar denotes s.e.m. (hierarchical bootstrap). *P* value denotes the probability of the null hypothesis that V_m-to-spike rate curves are the same between the pre-sample epoch and the delay epoch (hierarchical bootstrap, Methods). All statistical tests are two-sided. When there was no overlap between the curves from the two epochs, we did not test (*P* = not applicable (N.A)). Cell 88–130, cells recorded during the auditory task; cell 135–184, cells recorded during the tactile task. Asterisk denotes selective cells. See Supplementary Table 1 for number of trials.



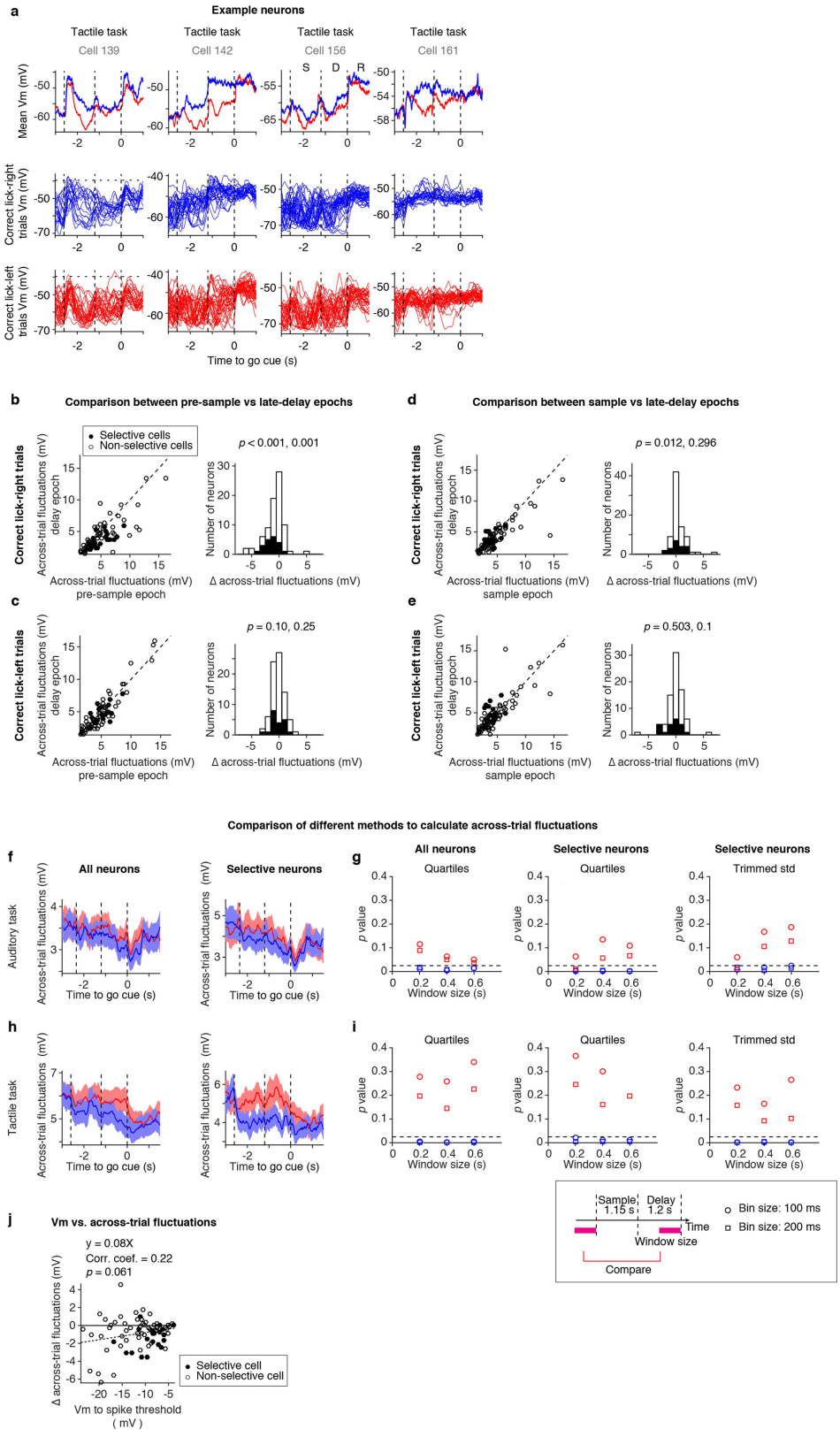
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Analysis of V_m. **a**, Relationship between delay spike rate and V_m (delay epoch spike rate or V_m minus pre-sample epoch spike rate or V_m). For each cell, correct lick-right and correct lick-left trials are shown separately ($n = 37$ cells). Black, selective cells ($n = 10$ cells). Red line, linear regression. Linear regression, Pearson's correlation coefficient, and the t-statistic of Pearson correlation coefficient (P) are shown. **b**, Selectivity of ALM neurons based on spike rate (top) and V_m (bottom). Duplicate of Fig. 2d. Shading denotes s.e.m. (bootstrap, $n = 10$ cells). **c**, Autocorrelation of V_m during the pre-sample epoch (left) and the delay epoch in correct lick-right trials (middle). Subtraction of these two autocorrelation curves is shown on right to emphasize the difference between epochs. Thick line denotes mean across cells; thin lines denote individual cells ($n = 37$ cells). **d**, Comparison of time constant of membrane fluctuations based on autocorrelation between the pre-sample epoch and the delay epoch. Left, scatter plot of the time constant; right, histogram of the difference in time-constant between the delay epoch compared to the pre-sample epoch (change in time constant). P value determined by two-sided Wilcoxon signed-rank test examining a null hypothesis that the change in time constant is 0. The first P value, all cells (left; $n = 37$ cells); second P value, selective cells (right; $n = 10$ cells). The shorter time constant during the delay epoch is presumably due to increase in conductance. **e**, Distribution of number of spike bursts per trial in correct lick-right trials. White, all cells; black, selective cells. **f**, Spike raster (top) and spike rates (bottom) of correct lick-right trials in an example cell with high occurrence of spike bursts (cell 120, arrow in **e**). Blue, regular spikes; green, spikes belonging to spike bursts. Right, example spike bursts. **g**, Grand average spike rate of cells with more than 0.5 spike bursts per trial in correct lick-right trials ($n = 12$ cells). Shading denotes s.e.m. Top, spike rate of all spike types (regular spikes and spike bursts), (bottom) spike rate of spike bursts. **h**, Comparison of number of bursts per trial between the pre-sample epoch and the delay epoch. P value determined by two-sided Wilcoxon signed-rank test examining a null hypothesis that number of bursts per trial are the same between two epochs ($n = 12$ cells). Blue, correct lick-right trials; red, correct lick-left trials. Spike bursts did not increase during the delay epoch. **i**–**p**, The same format as in **a**–**h** for the tactile task. All cells, $n = 42$; selective cells, $n = 10$.



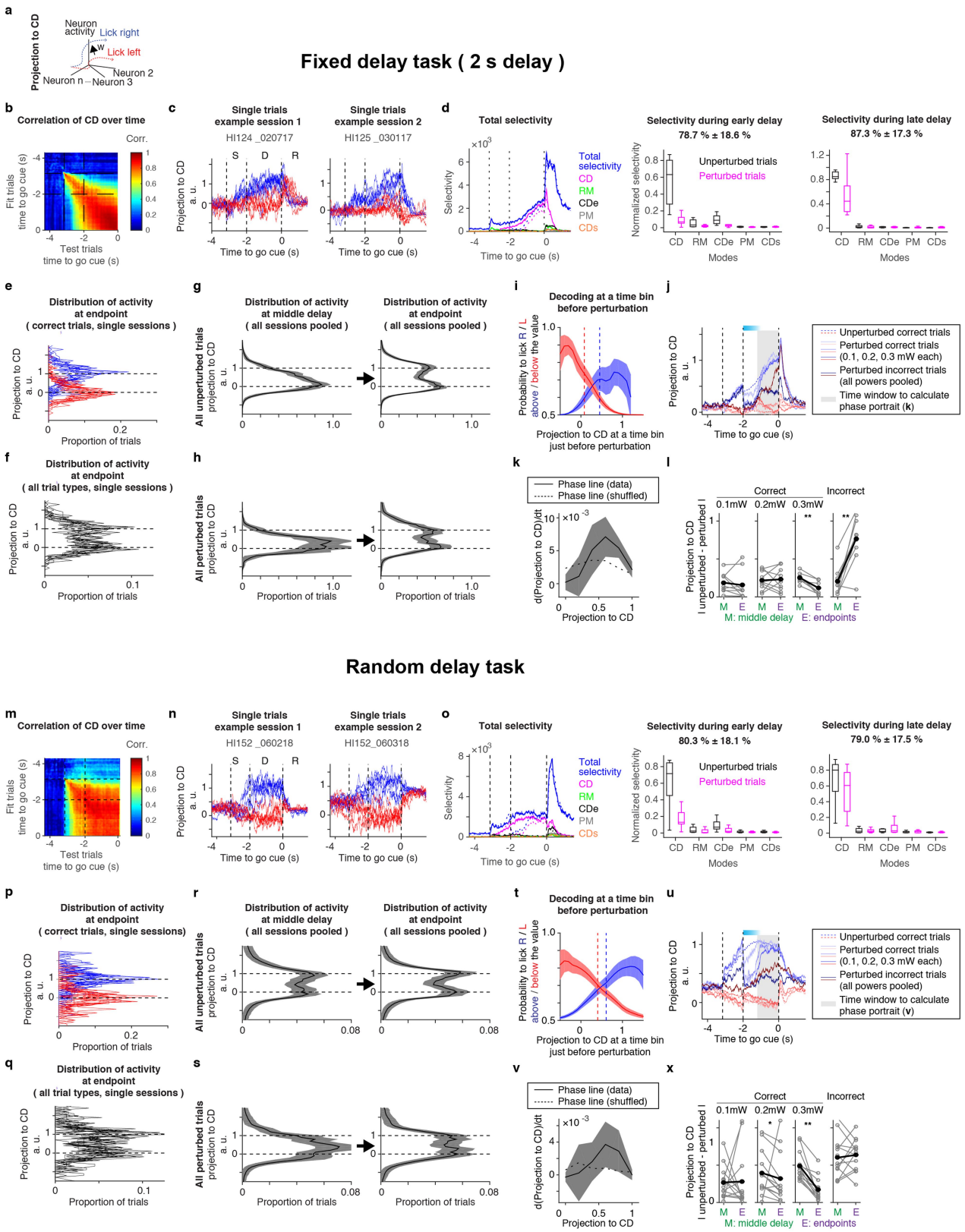
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Negative current injection. a, Five example cells with negative current injection. Top, mean V_m without negative current injection; bottom, mean V_m with negative current injection. Blue, correct lick-right trials; red, correct lick-left trials. **b**, Input resistance (R_{in}) of cells with and without current injection ($n = 10$ cells). Data are pooled from cells analysed in **c**. *P* value determined by two-sided Wilcoxon signed-rank test. Central line in the box plot is the median. Top and bottom edges are the 75% and 25% points, respectively. The whiskers show the lowest datum within 1.5 interquartile range (IQR) of the lower quartile, and the highest datum within 1.5 IQR of the upper quartile. **c**, Delay amplitude of V_m (delay epoch V_m minus pre-sample epoch V_m) with and without current injections ($n = 10$ cells). Delay amplitude during the current injection was normalized by the change in input resistance. Correct lick-right trials (blue) and correct lick-left trials (red) are shown separately. Crosses denote s.e.m. (bootstrap). Dashed line denotes linear regression. Slope of linear regression, Pearson's correlation coefficient and the *t*-statistic of Pearson's correlation coefficient (*P*) are shown. **d**, Lick-right delay selectivity (delay epoch V_m in lick-right trial minus delay epoch V_m in lick-left trial) with and without current injections ($n = 10$ cells). Crosses denote s.e.m. (bootstrap). Dashed line denotes linear regression. Slope of linear regression, Pearson's correlation coefficient), and the *t*-statistic of Pearson correlation's coefficient (*P*) are shown. **e**, Loss of spike bursts after current injection. Top, V_m of two example cells with high occurrence of spike bursts. Overlay of all correct lick-right trials. Black horizontal line, spike threshold. Regular spikes were removed and V_m was averaged over 100 ms. Sharp overshoots above the spike threshold indicate spike bursts. $n = 19$ and 20 trials (cell 96 and 101, respectively). Right, two example spike bursts indicated by arrows (cell 101). Bottom, V_m of the same example cells with negative current injection. There was no spike to remove. V_m was averaged over 100 ms. Note the loss of sharp depolarizing events. $n = 8$, and 9 trials (cell 96 and 101, respectively).



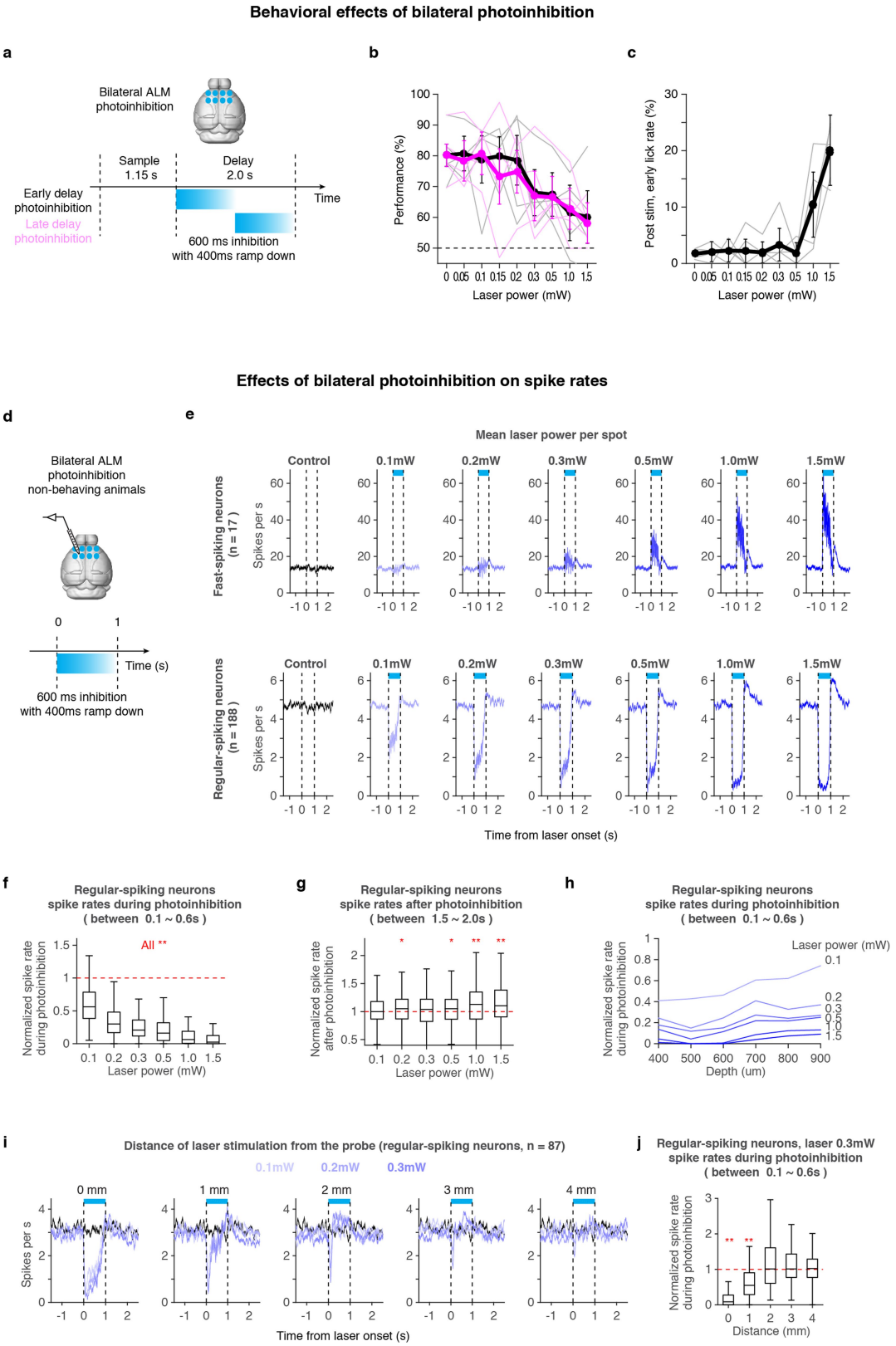
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Funnelling of V_m. a, Four example ALM neurons. Top, mean V_m; middle, all correct lick-right trials overlaid; bottom, all correct lick-left trials overlaid. To remove fast within trial fluctuations, V_m was averaged over 200 ms. **b**, The same plot as in Fig. 4c, d. $n = 79$ cells (both auditory and tactile tasks are pooled). Black, selective cells ($n = 20$ cells). P values were determined by one-sided Wilcoxon signed-rank test (the same applies to **c**–**e**). The first P value, all cells; second P value, selective cells only. **c**, As in **b** for the correct lick-left trials. Statistical methods and n value as in **b**. **d**, As in **b** for comparison between the sample epoch and the delay epoch. Statistical methods and n value as in **b**. **e**, As in **c** for comparison between the sample epoch and the delay epoch. Statistical methods and n value as in **b**. **f**, Across-trial fluctuations of all cells ($n = 37$) (left), and selective cells ($n = 10$) (right) in the auditory task (as in Fig. 4b). Line, mean of across-trial fluctuations; shading denotes s.e.m. (hierarchical bootstrap). Blue, correct lick-right trials; red, correct lick-left trials. **g**, Testing for a decrease in across-trial fluctuations. P values reflect a null hypothesis that across-trial fluctuations during the delay epoch were higher than that during the pre-sample epoch (hierarchical bootstrap, $n = 1{,}000$ iteration). Across-trial fluctuations were measured as the quartile difference (left) or trimmed s.d. (right) of V_m across the same trial type (blue, correct lick-right trials; red, correct lick-left trials) (Methods). Both methods provided similar results. V_m was averaged over 100 ms (triangle) or 200 ms (square) to remove fast within-trial fluctuations. The result was robust to the averaging bin size. The dashed line, $P = 0.025$ ($\alpha = 0.05$ for two-sided test). A schematic of statistical test in **g** and **i** is shown in the box below **i**. Across-trial fluctuations during the pre-sample epoch (0.6 s) was compared with the across-trial fluctuations during the delay epoch (variable durations: window size in **g** and **i**). The window ends at the time t, which is $t =$ (time of the go cue) $-$ (bin size)/2, to exclude the signal after the go cue. **h**, As in **f** for the tactile task. **i**, As in **g** for the tactile task. **j**, Relationship between V_m and the change in across-trial fluctuations (across-trial fluctuations during the delay epoch minus across-trial fluctuations during the pre-sample epoch). V_m during the delay epoch was averaged and normalized to the spike threshold. Dashed line, linear regression. Slope of linear regression, Pearson's correlation coefficient, and the t-statistic of Pearson's correlation coefficient (P) are shown. Pooling both the tactile and auditory tasks, and excluding non-spiking cells ($n = 73$ cells). The slope of regression line is opposite from what is expected for the ceiling effect.



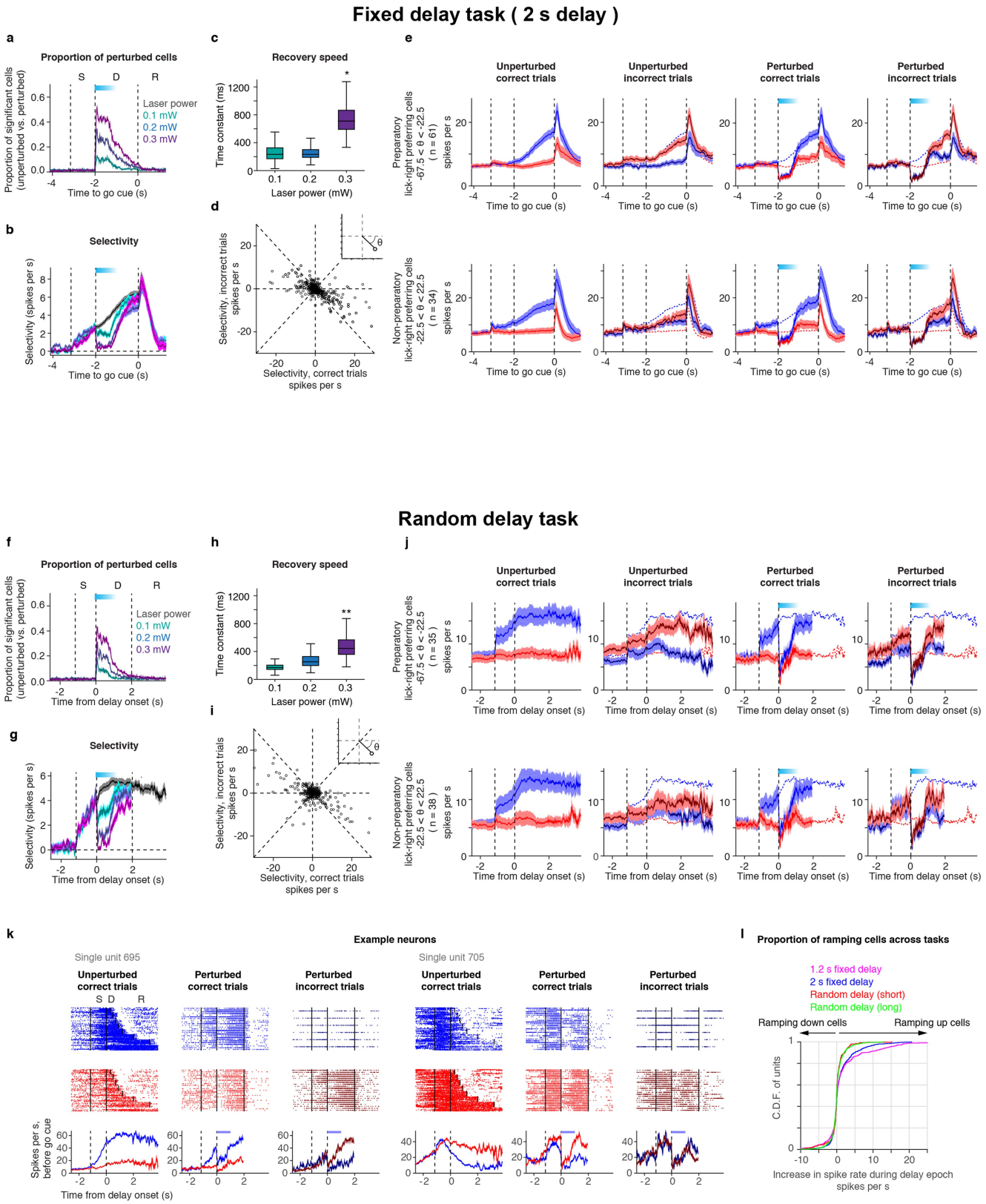
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Robustness of the CD to perturbation. **a**, Schematics. Projection of population activity to the **CD**. **b**, The **CD** is stable during the delay epoch. Pearson's correlation coefficient of the **CD** between half of trials (fit trials) and the other half (test trials) (Methods). Grand average of all sessions ($n = 11$ sessions for **b**–**l**). **c**, Projection of individual trials to the **CD** in two example sessions (seven randomly selected trials per trial type; blue, correct lick-right trials; red, correct lick-left trials). **d**, Selectivity explained by different modes (dimensions in activity space). Left, selectivity over time. Solid line, unperturbed correct trials, dotted line, perturbed correct trials. Blue, total selectivity; other colours, selectivity along different modes. Middle, selectivity explained by each mode during early delay (first 1 s of the delay epoch). Right, selectivity explained by each mode during late delay (last 1 s of the delay epoch) (mean ± s.e.m.). **CDe**, **CD** during early delay; **CDs**, **CD** during stimulation; **PM**, perturbation mode; **RM**, non-selective ramping mode (Methods). Sum of all modes is shown on top. Selectivity did not increase along any of these modes during perturbations. Central line in the box plot is the median. Top and bottom edges are the 75% and 25% points, respectively. The whiskers show the lowest datum within 1.5 IQR of the lower quartile, and the highest datum within 1.5 IQR of the upper quartile. **e**, Distribution of projection to the **CD** at end points in each session is overlaid. Blue, correct lick-right trials; red, correct lick-left trials. **f**, Distribution of projection to **CD** at end points in all trial types (both correct and incorrect trials are pooled, but not early lick or no-response trials). Distribution of end points in each session is overlaid. **g**, Distribution of projection to the **CD** at end points in Fig. 5b, pooling correct and incorrect trials, but not early lick or no-response trials. Shading denotes s.e.m. (bootstrap). **h**, The same format as in **g** for perturbed trial types. **i**, Decoder based on projection to the **CD** at a time bin just before the onset of perturbation (Methods). Probability to lick right in trials with projection to **CD** higher than each value in x axis (blue). Probability to lick left in trials with projection to **CD** lower than each value in x axis (red). Dashed lines indicate decision boundaries for lick-right (blue) and lick-left (red) trials (Methods). Shading denotes s.e.m. (hierarchical bootstrap). **j**, Trajectories along the **CD** in perturbed trials decoded to be correct trials before the onset of perturbation. Dotted line, unperturbed correct trials in Fig. 5e. **k**, Phase line of trajectories along **CD** of trials decoded to be correct trials before the onset of perturbation. Solid line, data; dashed line, shuffled data; shading, s.e.m. (hierarchical bootstrap). **l**, Absolute difference in projection to the **CD** between perturbed and unperturbed trials at the middle delay (green) and end points (magenta). Grey, individual session; black, mean. $P = 0.57$, 0.92, 0.0078 and 0.0039 from left to right (two-sided Wilcoxon signed-rank test). $^{*}P < 0.05$. **m**–**x**, The same format as in **b**–**l** for the random delay task. $n = 13$ sessions. In **x**, $P = 0.79$, 0.027, 0.00024 and 0.64 from left to right (two-sided Wilcoxon signed-rank test). $^{*}P < 0.05$; $^{**}P < 0.01$.



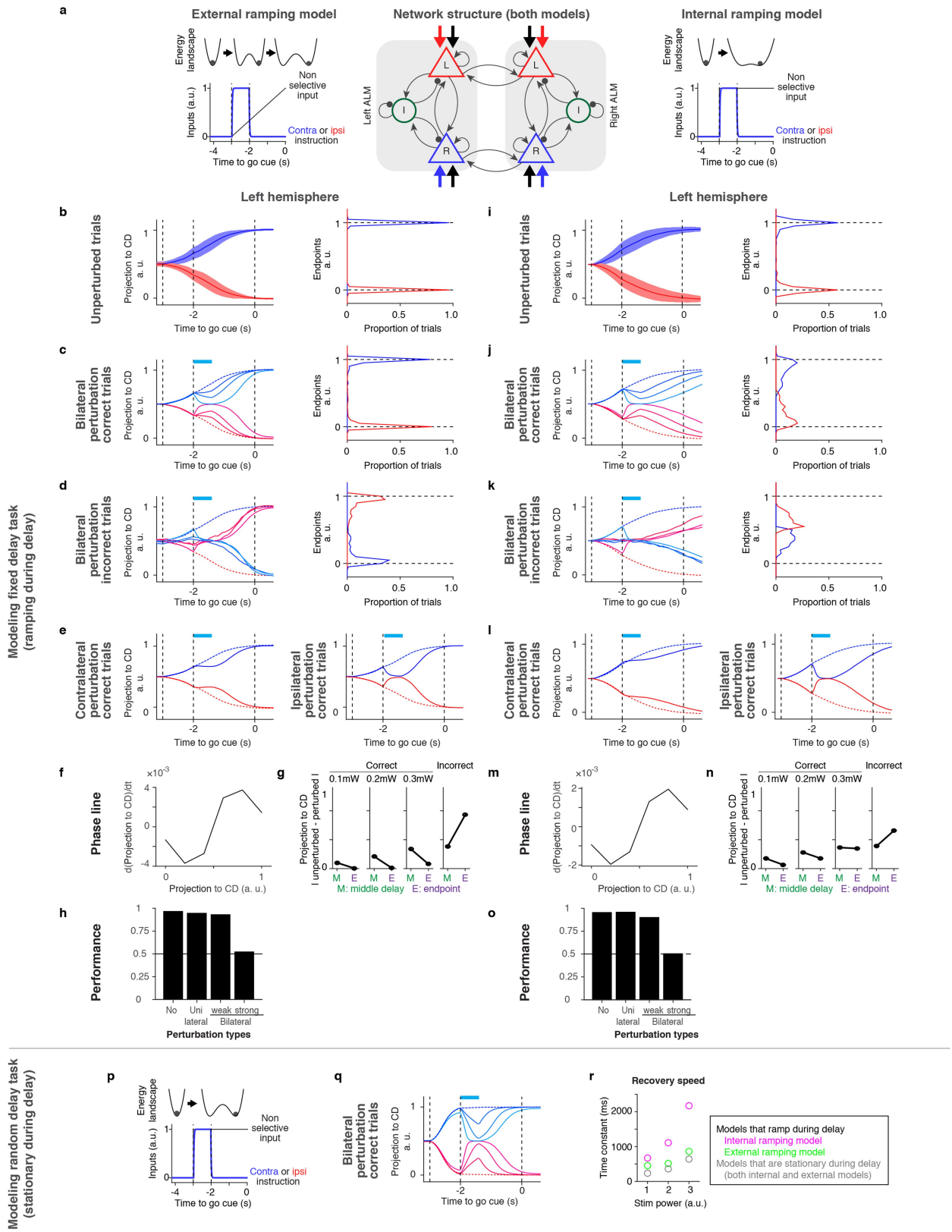
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Characterization of bilateral photoinhibition of ALM. a, Schematic of bilateral photoinhibition of ALM. Eight photostimuli were applied to ALM in both hemispheres (spacing, 1-mm interval). These photostimuli are expected to uniformly silence excitatory neurons in ALM[6]. Photoinhibition started at the onset of the delay epoch (−2 s to the go cue) or at the middle of the delay epoch (−1 s to the go cue) and lasted for 600 ms with additional 400 ms ramping down (Methods). **b**, Effect of the photoinhibition with different laser powers on behavioural performance. Black, early delay inhibition; magenta, late delay inhibition. Thick lines, grand mean performance ($n = 5$ mice); thin lines, mean performance of each mouse. Error bars denote s.e.m. (hierarchical bootstrap). Laser power, time-averaged mean power per spot. **c**, Effect of photoinhibition with different laser powers on early lick rates. Early lick rate after the early delay photoinhibition is shown. The same format as in **b**. **d**, Schematic of silicon probe recording during bilateral photoinhibition of ALM in non-behaving mice (relevant to **e**–**j**). **e**, Spike rates of fast-spiking neurons (top) and regular-spiking neurons (bottom) during photoinhibition of ALM with eight spots. Each column represents data with different laser powers. Mean spike rate is shown. Cyan bar, time of photoinhibition. **f**, Spike rate of regular-spiking neurons during the photoinhibition. For each cell, mean spike rate during photoinhibition (100–600 ms) was divided by mean spike rate before photoinhibition (−1–0 s) to calculate the normalized spike rate (as in **h** and **j**). Increasing laser power resulted in stronger inhibition. $^{**}P < 0.01$, two-sided Wilcoxon signed-rank test with Bonferroni correction. From left to right: $P = 2.1 \times 10^{-19}$, 2.2×10^{-25}, 2.1×10^{-27}, 4.2×10^{-27}, 1.6×10^{-27} and 2.7×10^{-31} (P values without Bonferroni correction), $n = 188$ cells. Central line in the box plot is the median. Top and bottom edges are the 75% and 25% points, respectively. The whiskers show the lowest datum within 1.5 IQR of the lower quartile, and the highest datum within 1.5 IQR of the upper quartile. **g**, Spike rate of regular-spiking neurons after the photoinhibition. For each cell, mean spike rate after photoinhibition (1–2 s) was divided by the mean spike rate before photoinhibition (−1–0 s) to calculate normalized spike rate. Increasing laser power resulted in stronger rebound. $^{*}P < 0.05$, $^{**}P < 0.01$, two-sided Wilcoxon signed-rank test with Bonferroni correction. From left to right: $P = 0.37$, 7.1×10^{-3}, 3.1×10^{-2}, 1.4×10^{-2}, 3.6×10^{-5} and 1.4×10^{-6} (P values without Bonferroni correction), $n = 188$ cells. Box plots follow the same format as in **f**. **h**, Relationship between depth and spike rate of regular-spiking neurons during the photoinhibition. Photoinhibition affected neurons across layers. **i**, Spike rate of regular-spiking neurons during the photoinhibition at different locations. Mean spike rate is shown for each laser power (control, black). Distance of recording site (ALM) and centre of laser stimulation (centre of four spots in the same hemisphere) is shown. Laser spots were moved from ALM to posterior locations. **j**, Spike rate of regular-spiking neurons during the photoinhibition at different locations. Photoinhibition affected neurons 1 mm away from the laser, consistent with a previous report[6]. $^{**}P < 0.01$, two-sided Wilcoxon signed-rank test with Bonferroni correction. From left to right: $P = 1.1 \times 10^{-15}$, 6.0×10^{-7}, 0.61, 0.55, 0.52 (P values without Bonferroni correction), $n = 87$ cells. Box plots follow the same format as in **f**.



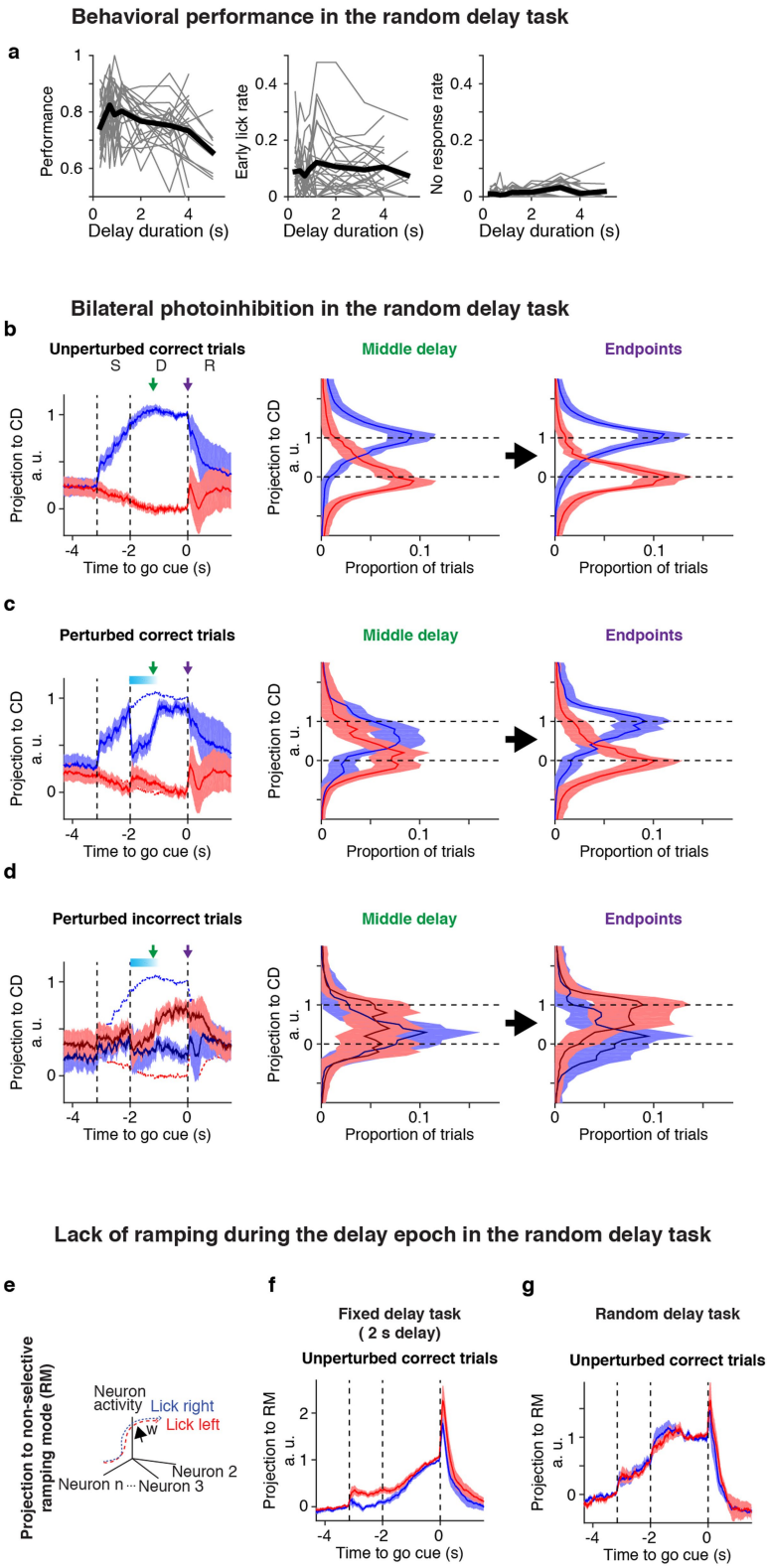
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Effect of photoinhibition on neural populations. **a**, Proportion of cells with significant spike rate difference between unperturbed and perturbed trials at each time point (Methods). Cyan bar, photoinhibition. $n = 667$ cells. **b**, Selectivity during bilateral photoinhibition. Line denotes mean of all cells; shadow denotes s.e.m. (bootstrap). Black, unperturbed trials. Cyan bar, photoinhibition. $n = 667$ cells. **c**, Recovery speed of selectivity following perturbation (approximately −1.2 to 0 s before the go cue; Methods). $^{*}P = 0.032$ (testing a null hypothesis that recovery speed is the same between trials with 0.1 mW and 0.3 mW photoinhibition, bootstrap, two-sided). $n = 667$ cells. Central line in the box plot is the median. Top and bottom edges are the 75% and 25% points, respectively. The whiskers show the lowest datum within 1.5 IQR of the lower quartile, and the highest datum within 1.5 IQR of the upper quartile. **d**, Relationship between selectivity in correct and incorrect trials. Mean selectivity during 1–2 s after the delay onset is shown. Inset, definition of θ (Methods). $n = 465$ cells. **e**, Mean spike rate of lick-right preferring neurons with different θ. Lines, grand mean of peri-stimulus time histogram; shading denotes s.e.m. (bootstrap). Blue, correct lick-right trials; red, correct lick-left trials; dark blue, incorrect lick-right trials; dark red, incorrect lick-left trials. Mean peri-stimulus time histogram of unperturbed correct trials (first row) are also shown as dashed lines in the second to fourth row. **f**–**j**, As in **a**–**e** for the random delay task. To calculate mean spike rate and selectivity, we pooled spikes before the go cue across trials with different delay durations. Because delay durations were different across trials, the time axis is aligned to the onset of the delay epoch. In **h**, $^{**}P = 0.009$ (testing a null hypothesis that recovery speed is the same between trials with 0.1 mW and 0.3 mW photoinhibition, bootstrap, two-sided). $n = 867$ cells for **f**–**h**, and 487 cells for **i**. **k**, Example single units (random delay task). Top, spike raster. Seven and twenty-five trials per trial type were randomly selected for unperturbed trials and perturbed trials, respectively. Bottom, mean spike rate pooling spikes before the go cue across trials with different delay durations. Blue, correct lick-right trials; red, correct lick-left trials; dark blue, incorrect lick-right trials; dark red, incorrect lick-left trials. Cyan bar, photoinhibition. **l**, Cumulative distribution of increase in spike rate during the delay epoch. Data from four different tasks are compared (Methods).



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Modular discrete attractor models. We built models consisting of two modules, each corresponding to one ALM hemisphere, in accordance with the previously described functional architecture[11]. These models are consistent with (1) ramping dynamics during the delay epoch; (2) reproduce the response to transient bilateral inactivation (Fig. 5); and (3) the response to unilateral ALM perturbation[11]. The circuit architecture in each hemisphere is as in Extended Data Fig. 1. Slow ramping was caused by either external non-selective input (**b**–**h**; external ramping model; note ramping external non-selective input in **a**) or internal slow dynamics (**i**–**o**; internal ramping model; note shallower energy landscape in **a**). All panels show results from the left hemisphere. In **b**–**o**, we modelled the fixed delay task (ramping activity during the delay epoch.) In **p, q**, we modelled the random delay task (stationary activity during the delay epoch). To make the activity during the delay epoch stationary, non-selective input was modified to be stationary for the external ramping model, and the energy landscape was modified to be steeper for the internal ramping model. With these modifications the two models are essentially equivalent. Therefore, we show only one model for the random delay task. $n = 1{,}000$ trials per model. **a**, Schematic of two-hemisphere network models. Circuit architecture, common to both models (middle), and time course of external inputs in each model (left and right). R, L and I correspond to lick-right selective excitatory neurons, lick-left selective excitatory neurons and inhibitory interneurons, respectively. Blue and red arrows, selective external input. Black arrows, non-selective external input. **b**, Trajectories projected along the **CD** in model with external ramping input. Activity in unperturbed trials (left), and distribution of end points (right). Line denotes mean; shading denotes s.d. Blue, correct lick-right trials; red, correct lick-left trials. **c**, Projected activity in bilaterally perturbed correct trials (left), and distribution of end points (right). Line denotes mean. Cyan bar on top denotes photoinhibition. Lighter colours correspond to higher intensity of photoinhibition. **d**, Projected activity in bilaterally perturbed incorrect trials. **e**, Projected activity in unilaterally perturbed correct trials. Effect of unilateral ALM photoinhibition contralateral (left) or ipsilateral (right) to the analysed hemisphere (left ALM) is shown. **f**, Phase line of projected trajectories. **g**, Absolute difference in projection to the **CD** between perturbed and unperturbed trials at the middle delay (green) and end points (magenta). **h**, Behavioural performance of the model (Methods). **i**–**o**, As in **b**–**h** for internal slow dynamics model. **p**, Schematic of two-hemisphere network models with stationary delay activity. **q**, Projected activity in bilaterally perturbed correct trials. Same format as in **d**. **r**, Recovery speed of the **CD** in bilaterally perturbed correct trials following perturbation. Note that recovery speed in the external ramping model is similar regardless of delay dynamics (ramping or stationary).



Extended Data Fig. 10 | Performance in the random delay task. **a**, Behavioural performance (left), early lick rate (middle) and no-response rate (right) in the random delay task. Thin lines, individual sessions ($n = 23$); thick line, grand mean among sessions. **b**–**d**, The same format as in Fig. 5b–d for the random delay task. Trials with 2-s delay duration are shown. Line denotes mean; shading denotes s.e.m. (bootstrap). $n = 13$ sessions. **e**, Schematics. Projection of trials to the non-selective ramping mode (**RM**). **f**, Projection of trials to non-selective ramping mode in the fixed delay task with 2-s delay. Line, grand mean across sessions ($n = 11$ sessions); shading denotes s.e.m. (bootstrap). **g**, As in **f** for the random delay task ($n = 13$ sessions). Trials with 2-s delay duration are shown.

nature research

Corresponding author(s): Karel Svoboda

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Matlab_R2007a, R2013b, WaveSurfer release_0.787, SpikeGLX v.20170315
Data analysis	Matlab_R2014b,R2017a, JRClust (2016 May 18)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data sets will be shared at CRCNS.ORG in the NWB format. For network models, Matlab code will be made available for download.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample sizes are similar to sample sizes used in the field. No statistical methods were used to determine sample size.
Data exclusions	We did not exclude any animal for data analysis.
Replication	We performed recordings from multiple animals to confirm reproducibility. Replications were successful.
Randomization	All animals were used for experiments. Trial types were randomly determined by a computer program.
Blinding	During experiments, trial types were randomly determined by a computer program. During spike sorting, experimenters cannot tell the trial type, so experimenters were blind to conditions.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	This study is based on data from 31 male mice (age > P60). We used four transgenic mouse lines: PV-IRES-Cre 70, Ai32 (Rosa-CAG-LSL-ChR2(H134R)-EYFP-WPRE, JAX 012569) 71, Gad2-cre (a gift from Boris Zemelman), and VGAT-ChR2-EYFP 72 (See Supplementary Information Table 1-3 for detail).
Wild animals	This study did not involve wild animals.
Field-collected samples	This study did not involve field-collected samples.