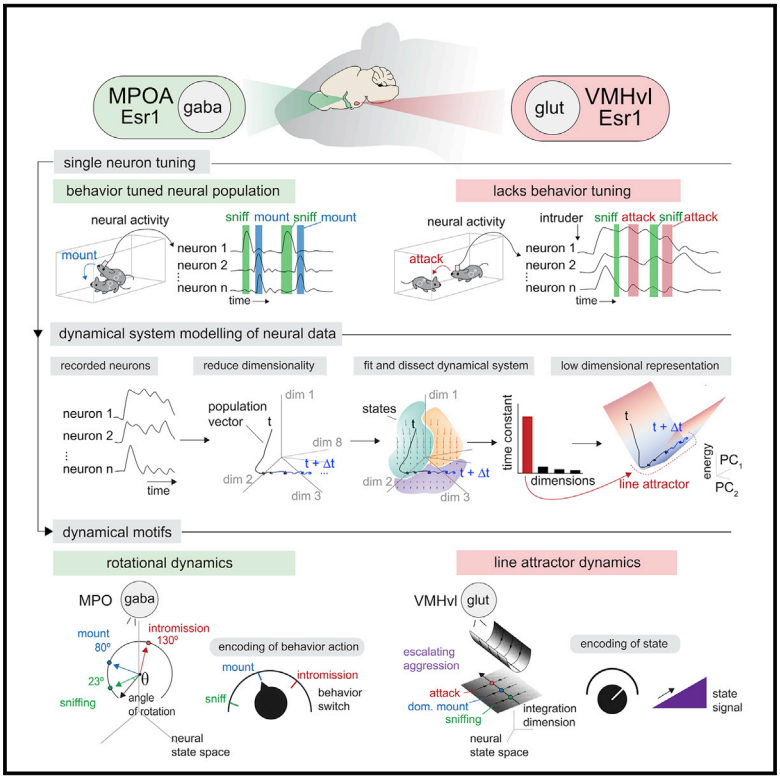
Cell

An approximate line attractor in the hypothalamus encodes an aggressive state

Graphical abstract



Highlights

- Dynamical system analysis reveals a line attractor in VMHvl^{Esr1} neuronal activity
- The line attractor represents a neural correlate of aggressive escalation
- Individual differences in aggression are correlated with line attractor properties
- MPOA^{Esr1} utilizes rotational not line attractor dynamics to encode mating behavior

Authors

Aditya Nair, Tomomi Karigo, Bin Yang, ..., Scott W. Linderman, David J. Anderson, Ann Kennedy

Correspondence

wuwei@caltech.edu (D.J.A.), ann.kennedy@northwestern.edu (A.K.)

In brief

Different hypothalamic regions utilize distinct neural dynamics to encode mating and aggressive behaviors.

Nair et al., 2023, Cell *186*, 178–193
January 5, 2023
https://doi.org/10.1016/j.cell.2022.11.027

CellPress

CellPress
OPEN ACCESS

Cell

Theory

An approximate line attractor in the hypothalamus encodes an aggressive state

Aditya Nair,[1,2,3] Tomomi Karigo,[1,2,3,9] Bin Yang,[1,2,3,10] Surya Ganguli,[4] Mark J. Schnitzer,[2,4,5] Scott W. Linderman,[6,7] David J. Anderson,[1,2,3,11,*] and Ann Kennedy[1,2,3,8,*]

[1]Division of Biology and Biological Engineering, Caltech, Pasadena, CA 91125, USA
[2]Howard Hughes Medical Institute
[3]Tianqiao and Chrissy Chen Institute for Neuroscience, Caltech, Pasadena, CA 91125, USA
[4]Department of Applied Physics, Stanford University, Stanford, CA, USA
[5]Department of Biology, Stanford University, Stanford, CA, USA
[6]Department of Statistics, Stanford University, Stanford, CA 94305, USA
[7]Wu Tsai Neurosciences Institute, Stanford University, Stanford, CA 94305, USA
[8]Department of Neuroscience, Feinberg School of Medicine, Northwestern University, Chicago IL 60611, USA
[9]Present address: Kennedy Krieger Institute, Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD 21205 USA
[10]Present address: Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD 21205 USA
[11]Lead contact
*Correspondence: wuwei@caltech.edu (D.J.A.), ann.kennedy@northwestern.edu (A.K.)
https://doi.org/10.1016/j.cell.2022.11.027

SUMMARY

The hypothalamus regulates innate social behaviors, including mating and aggression. These behaviors can be evoked by optogenetic stimulation of specific neuronal subpopulations within MPOA and VMHvl, respectively. Here, we perform dynamical systems modeling of population neuronal activity in these nuclei during social behaviors. In VMHvl, unsupervised analysis identified a dominant dimension of neural activity with a large time constant (>50 s), generating an approximate line attractor in neural state space. Progression of the neural trajectory along this attractor was correlated with an escalation of agonistic behavior, suggesting that it may encode a scalable state of aggressiveness. Consistent with this, individual differences in the magnitude of the integration dimension time constant were strongly correlated with differences in aggressiveness. In contrast, approximate line attractors were not observed in MPOA during mating; instead, neurons with fast dynamics were tuned to specific actions. Thus, different hypothalamic nuclei employ distinct neural population codes to represent similar social behaviors.

INTRODUCTION

A fundamental problem in neuroscience is to understand how the brain controls innate behaviors. Many such behaviors are governed by the hypothalamus, a deep subcortical brain region present in all vertebrates.[1,2] Classical brain stimulation and lesion experiments have implicated different hypothalamic regions ("nuclei") in diverse innate behaviors (reviewed in Paredes and Baum,[3] Siegel et al.,[4] Canteras,[5] King,[6] Kruk,[7] Swanson,[8] and Simerly[9]). More recently, optogenetic stimulation has identified genetically marked neuronal subpopulations that can evoke such behaviors[10–13] (reviewed in Yamaguchi,[14] Zha and Xu,[15] Augustine et al.,[16] and Sternson[17]). Genetic ablation or reversible silencing has demonstrated that these subpopulations are essential for natural occurrences of these behaviors.[10–12,18]

An important open question is how the activity of these neural subpopulations during naturally occurring behavior reflects their "causative" function. Relatively few single-unit recordings have been performed in hypothalamic nuclei because of their inaccessibility.[13,19–21] Recordings of bulk calcium signals[22] have confirmed that these neuronal subpopulations are active during the natural behaviors they can artificially evoke.[23–25] However, this averaging method obscures individual cell activity patterns.

Miniature head-mounted microscopes allow calcium imaging with single-cell resolution in freely moving animals.[26,27] Application of this approach to the hypothalamus has identified cells exhibiting stimulus-locked activity during natural behavior.[28–30] For example, imaging of estrogen receptor type 1 (Esr1)-expressing neurons in the medial preoptic area (MPOA), whose optogenetic activation can elicit mounting behavior in male mice,[31,32] has revealed cells that respond specifically during spontaneous mounting of females (see also Figure 1E). Such results, together with single-cell transcriptomic analysis, have reinforced the prevailing view that the hypothalamus controls different survival

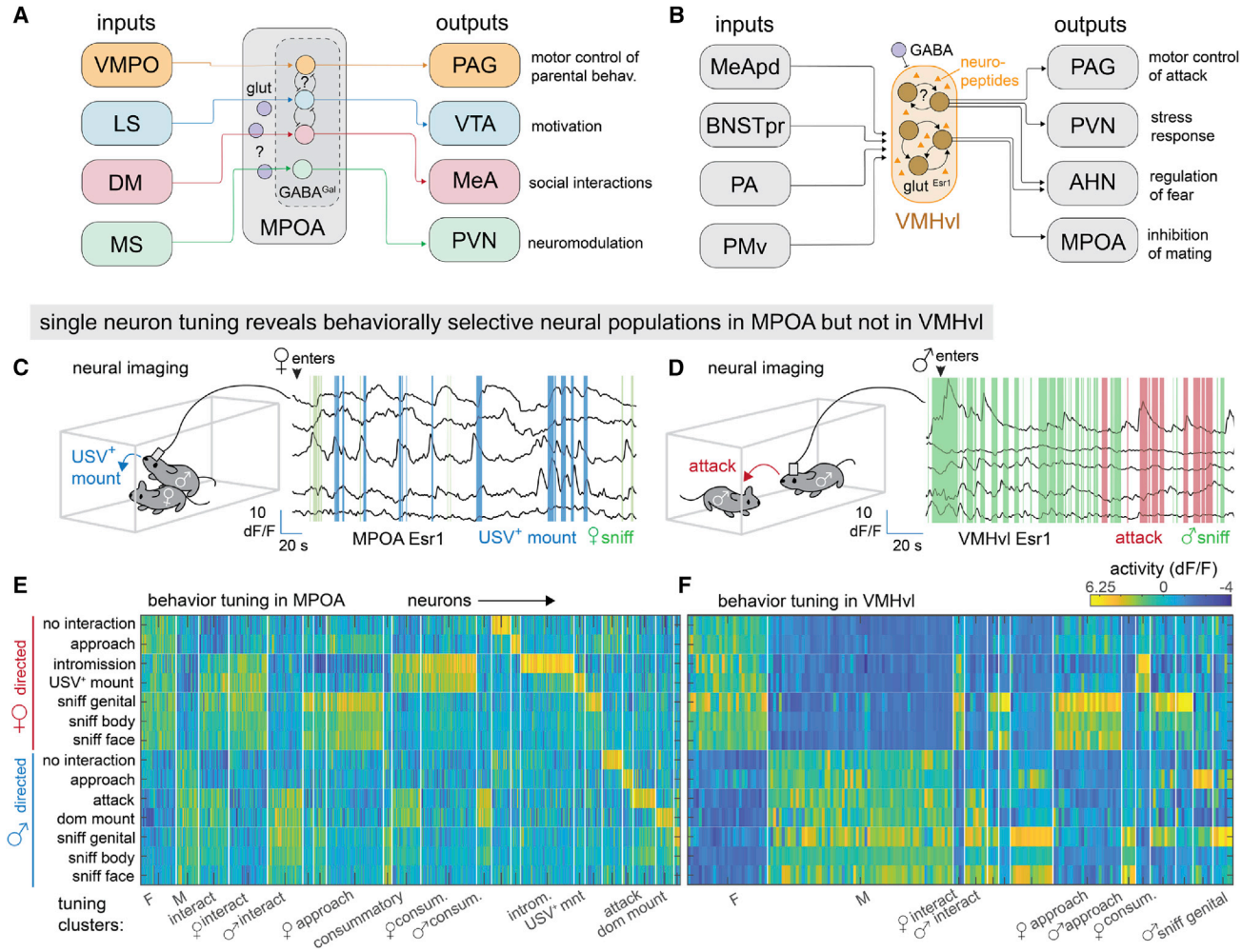


Figure 1. Cytoarchitectures and cellular representations in a neural system regulating social behavior

(A and B) Cytoarchitecture and input-output maps of MPOA[33,35] (A) and VMHvl[31,36–38] (B).

(C and D) Example traces from Esr1$^+$ neurons in MPOA (C) and VMHvl (D).

(E and F) Clustering of recorded Esr1$^+$ neurons in MPOA (E, n = 306 neurons from 3 mice) and VMHvl (F, n = 391 neurons from 4 mice) using a regression model. Rows, hand-annotated behaviors; columns, individual neurons.

behaviors via genetically determined, functionally specific neuronal subpopulations.[33,34]

The case of aggression, however, presents a paradox seemingly at odds with this view. On one hand, optogenetic stimulation of Esr1$^+$ neurons in the ventrolateral subdivision of the ventromedial hypothalamus (VMHvl) neurons triggers attack behavior,[12,39–41] identifying these neurons as the likely cellular substrate of electrical brain-stimulated aggression.[4,7,42] Conversely, genetic ablation of VMHvl neurons expressing the progesterone receptor (PR; co-expressed with Esr1) or optogenetic silencing of VMHvl^{Esr1} neurons blocks natural aggression[12,18]

On the other hand, miniscope imaging of VMHvl^{Esr1} neurons during natural fighting revealed surprisingly few cells that exhibited time-locked, attack-specific activity.[29] Instead, most such neurons exhibited "mixed selectivity," responding during different phases of an aggressive interaction. Different subsets of Esr1$^+$ neurons responded to male versus female conspecifics, suggesting an encoding of conspecific sex.[29,31,43] Nevertheless, decoders trained on VMHvl^{Esr1} neural imaging data could accurately distinguish episodes of attack from sniffing.[29]

Thus, observational versus perturbational studies of VMHvl^{Esr1} neurons yield seemingly inconsistent views: these neurons causally control aggressive behavior; however, very few of them are specifically "tuned" to attack. There are two possible explanations for this paradox. First, the small fraction of VMHvl^{Esr1} neurons that are more active during attack may be the ones responsible for the specific causative influence of this population. Alternatively, the majority of VMHvl^{Esr1} neurons, despite their mixed behavioral selectivity, may control attack through some type of population code.

In other systems where there is no clear correlation between single-unit spiking patterns and behavior, modeling neural populations as a dynamical system[44–46] (reviewed in Vyas et al.[47]) has revealed signals in the dynamics of population activity that

can robustly predict motor actions.[48,49] We have therefore carried out similar modeling of VMHvl^{Esr1} neural activity dynamics during naturalistic social behaviors, using legacy data from previous studies.[29,31,43] Our results reveal line attractor dynamics in VMHvl that correlate with escalating levels of aggressive behavior, suggesting that they may represent or encode an aggressive internal state. Strikingly, line attractor dynamics are absent in MPOA activity during both mating and aggression. This analysis therefore reveals fundamental differences in the neural coding of social behaviors by different hypothalamic nuclei.

RESULTS

Cellular tuning analysis confirms behaviorally selective neural populations in MPOA but not in VMHvl

Calcium imaging of MPOA^{Esr1} or VMHvl^{Esr1} neurons revealed distinct patterns of neuronal activation during social interactions[29,31] (Figures 1A and 1B). To quantify these differences, we re-analyzed calcium imaging data[31] from sexually experienced male C57Bl/6N$^{Esr1\text{-}2A\text{-}Cre/+}$ mice during standard resident-intruder assays, using male or female BalbC intruders (Figures 1C and 1D). We then computed the mean activity of each neuron during each of 14 different hand-annotated actions and clustered them using a regression model (VMHvl: n = 306 neurons from 3 mice; MPOA: n = 391 neurons from 4 mice, see STAR Methods).

Confirming previous observations,[29,31] many MPOA clusters contained neurons only active during specific behavioral actions, such as intromission or mounting toward females (Figure 1E). In contrast, most VMHvl^{Esr1} neurons were activated in response to either males or females, with very few neurons showing behavior-specific activation (Figure 1F).

Unsupervised dynamical systems analysis of neural activity during social behavior

In other systems, population analysis via fit dynamical systems has revealed a neural encoding of behavioral actions that were not apparent in neuron-by-neuron analysis.[47,48,50,51] We therefore investigated whether behavioral representations among VMHvl^{Esr1} neurons might be encoded at a population level, using an unsupervised dynamical systems approach.

To do so, we fit a dynamical model to the population activity of VMHvl^{Esr1} cells from each of multiple mice (n = 6), from two different studies[31,43] in which recordings were made throughout male-male or male-female encounters (average duration 5.1 ± 0.68 min and 11.4 ± 0.68 min, respectively; mean ± SEM). Specifically, we fit a recurrent switching linear dynamical system (rSLDS) model,[52] which approximates a complex non-linear dynamical system as a composite of more easily interpretable linear dynamical systems, or "states" (Figure S1A).

rSLDS first reduces neural activity to a set of latent variables (also called "dimensions" or "factors"), defining a low-dimensional "state space" in which the time-evolving population neural activity vector can be analyzed (Figure 2A①). Population activity in this low-dimensional space is then segmented into a set of discrete states (Figure 2A②) while fitting a linear dynamical system model (Figure 2A③) to neural activity within each state. Each state has a different dynamics matrix, which dictates how neural activity evolves over time from any given point within that state space. Quantitative examination of parameters from this matrix after model fitting can unveil dynamical properties of the neural circuit, such as the time constant of each dimension.[46] Finally, to visualize more easily the dynamical properties of each state, we plotted its "flow field" in 2D using principal component analysis (PCA) (Figure 2A④ right; see STAR Methods).

In fitting the rSLDS model, we chose the minimum number of states and dimensions that could capture 90% of observed variance in neural activity, determined using cross validation in each mouse separately (Figures S1B–S1E; 7–8 dimensions [7.2 ± 0.1, N = 6 mice] and 3–4 states). We evaluated the "goodness of fit" of each model iteration using both the log likelihood of the data[52] and an additional metric that we call the "forward simulation error" (FSE; Figure S1F; see STAR Methods). Plotting the FSE over time allows visualization of periods wherein model performance drops (Figure S1G). By this metric, our best-fit models captured most of the variance in neural data (model performance [1-FSE] = 0.72 ± 0.02, N = 6 mice; Figure S1H).

The rSLDS framework allows the fit dynamical system models to be either autonomous or to receive external input. Since VMHvl neuron firing rates correlate with the distance to another male or to male mouse urine,[53] likely reflecting the concentration of chemosensory cues,[54] we used the distance between animals and their facing angle as a proxy for external sensory input strength[53,55] (see STAR Methods).

rSLDS analysis of VMHvl neural activity discovers an integration dimension that correlates with aggressive escalation

Next, we performed retrospective alignment of the unsupervised neural data model with behavioral annotations over time. This comparison revealed that the probability of attack was elevated during a single rSLDS state (state 3, Figures S1I–S1K). Importantly, attacks were not time-locked to the onset/offset of this state; rather, epochs of this state outlasted individual attack bouts (state 3 epoch duration: 79.5 ± 5.5 s, attack bout duration: 4.86 ± 0.44 s, N = 6 mice, Figures S1I5, S1J3, and S1K3). This suggests that the state did not simply represent motor activity (Figure S1A, cf. case 2 versus 1).

To better understand the neural population dynamics related to attack behavior, we examined the dynamics matrix for this state, which describes how dimensionally reduced neural activity in that state changes over time. The eigenvalues of this matrix reflect the rate at which activity along each of these dimensions decays to zero following external input and can be converted to a time constant for each dimension.[56,57] Input to dimensions with short time constants will quickly decay to zero, whereas input to dimensions with long (large) time constants persists and decays slowly. Strikingly, one of the rSLDS dimensions had an estimated time constant of over 100 s that was significantly higher than that of all other dimensions (Figures 2B red dot, 2C, and 2D, N = 6 mice). Because systems with long time constants approximately integrate their input over time, we refer to the longest time constant dimension as the "integration" dimension.[58,59]

The integration dimension accounted for 19.5% ± 1.9% of the overall variance in neural activity (N = 6 mice). In contrast, a

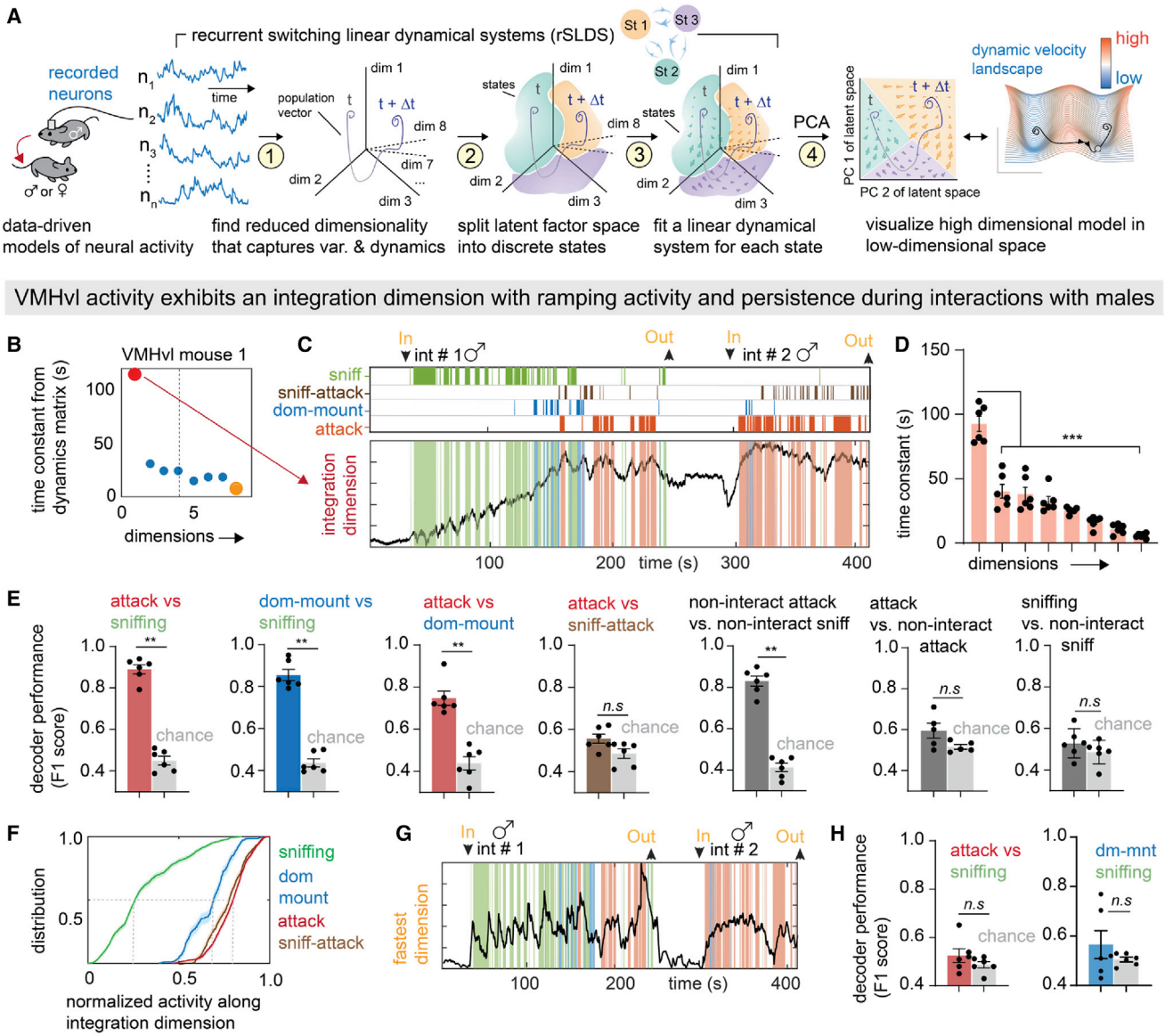


Figure 2. Dynamical analysis of VMHvl neural activity reveals an integrator dimension that correlates with aggressive escalation

(A) Schematic illustrating rSLDS[52] analysis. Steps ①–④ are shown sequentially for illustrative purposes only.

(B) Time constants of rSLDS dimensions (see A①) in attack enriched state from VMHvl mouse 1. Dimensions with longest (red dot) and shortest (yellow dot) time constants are indicated.

(C) Projection onto time axis of integration dimension with overlayed behavior annotations.

(D) Average time constant of all dimensions, arranged in decreasing order. (***$p < 0.001$, n = 6 mice.)

(E) Average F1 score of binary decoder of behavior pairs trained on integration dimension activity (**$p < 0.005$, *$p < 0.01$, n = 6 mice).

(F) Cumulative distribution of integration dimension value (normalized) for different behaviors.

(G) Projection of fastest dimension in example VMHvl mouse 1.

(H) Performance of binary decoder of behavior pairs trained on fastest dimension activity (n = 6 mice). For additional data, see Figures S1, S2, and S3.

support vector machine (SVM) decoder trained on all neural data to distinguish attack from sniffing periods explained much less variance (0.3% ± 0.1%, N = 6 mice, $p < 0.001$, Figure S2B). Examining the activity of individual neurons that were weighted strongly in the integration dimension (Figure S2D) revealed that around 20% of neurons per animal contributed to this dimension, with some showing ramping and persistent activity (Figures S2I, S2J, S2L, and S2N). Moreover, most of these neurons were tuned to male intruders (Figures S3A and S3B). Thus, the integration dimension encapsulates a signal that is present at the level of at least some individual neurons but is also an emergent property of the population.[51]

We next compared the time-varying activity of the integration dimension with the animals' actions during aggressive

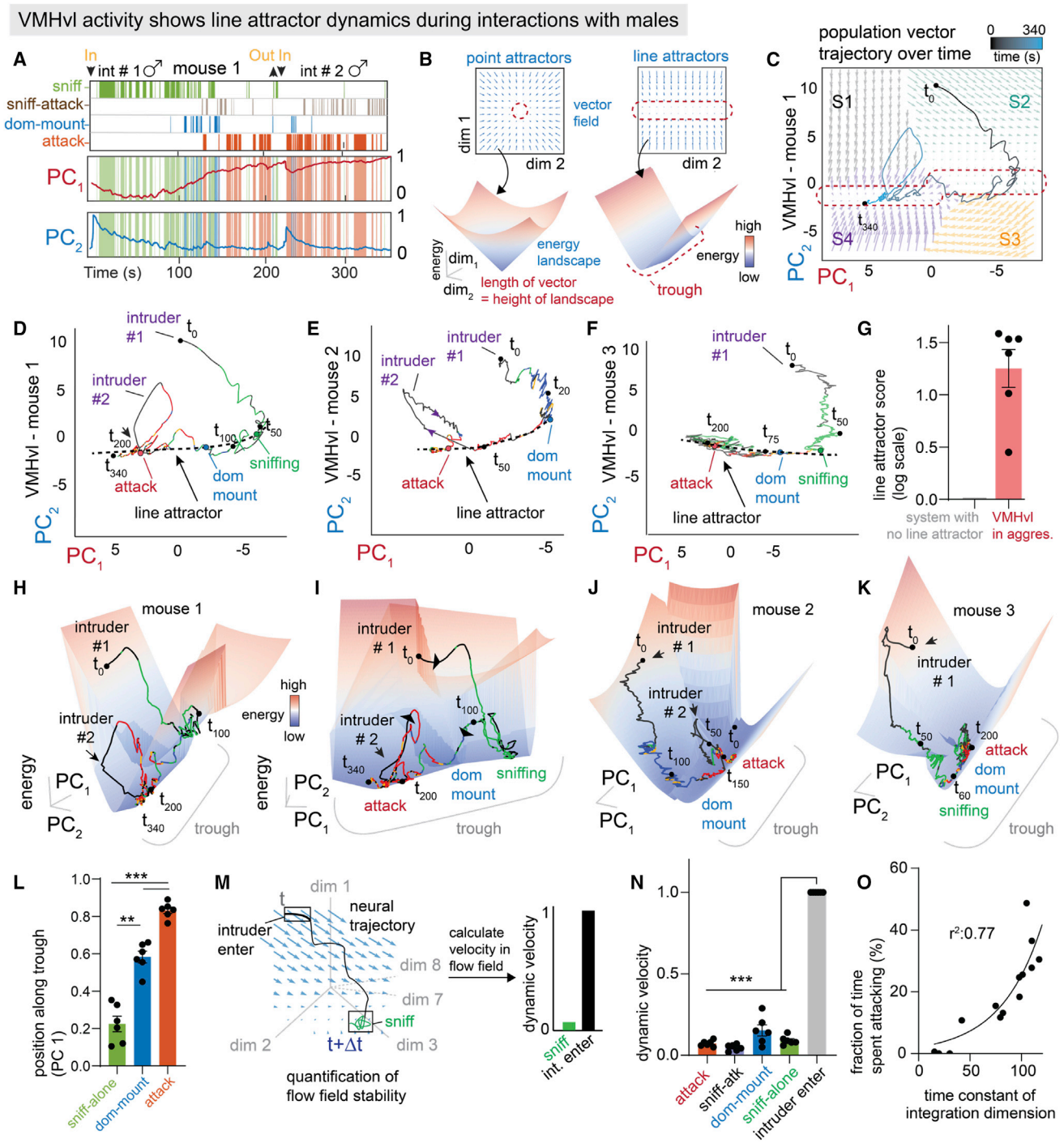


Figure 3. VMHvl contains an approximate line attractor that integrates aggressive escalation

(A) Behavior rasters shown with first two principal components of dynamical system (see STAR Methods) for example VMHvl mouse 1.

(B) Schematic illustrating inferred dynamics shown as idealizedflow fields (with attractors highlighted by red dashed lines) and 3D landscapes for point attractors (left) and line attractors (right).

(C) Neural state space with population trajectories and inferred flow field colored by rSLDS states for VMHvl mouse 1, with line attractor highlighted. The slight shift in line attractor position in state S2 vs S4 accounts for the curvature in the attractor in (D) and (E).

(D–F) Neural state space for VMHvl mouse 1 (D), mouse 2 (E), and mouse 3 (F) with line attractor highlighted (see STAR Methods).

(G) Line attractor score (see STAR Methods) for VMHvl (red bar, n = 6 mice).

(H and I) Inferred 3D dynamic landscape in VMHvl mouse 1. Front and side views of line attractor are shown.

(legend continued on next page)

encounters. In mouse 1, activity along the integration dimension was low during sniffing, ramped up at the onset of dominance mounting (a low-intensity aggressive behavior[31]), and increased further to a stable plateau value as the animal attacked (Figure 2C). A cumulative distribution function (cdf) of the normalized level of activation along the integration dimension during sniffing, dominance mounting, and attack revealed that these three behaviors occurred at low, medium, and high values of this dimension, respectively (Figure 2F; distribution means: sniffing: 0.30, dominance mount: 0.66, attack: 0.82, N = 6 mice).

Remarkably, a binary classifier created by thresholding the value of the integration dimension could distinguish periods of sniffing from attack or from dominance mounting, with a high F1 score (0.89 ± 0.02, N = 6 mice; Figure 2E). The same method could also distinguish dominance mounting versus attack (F1 score 0.74 ± 0.03, N = 6 mice; Figure 2E). However, such classifiers could not distinguish behaviors occurring close together in time, such as attack and sniff-attack (defined as periods of sniffing that occurred within 1 s prior to attack, as described recently[60]), perhaps due to the gradual ramping of activity along this dimension. Remarkably, none of the other seven fit dimensions could be used to distinguish aggressive behaviors from sniffing with above chance accuracy (Figures 2G and 2H; Figure S2C).

The foregoing analysis suggested that a low-dimensional signal in VMHvl represents escalating aggressive behaviors. To account for possible spurious behavioral correlations due to the slow decay of activity in this dimension, we devised a version of session permutation as described recently,[61] by cross validating decoder thresholds between animals (see STAR Methods). This more rigorous paradigm could still decode behaviors with high F1 scores (Figure S2H).

Sniffing, attack, and dominance mounting are performed in bouts separated by short inter-bout intervals (IBIs). Because of its slow ramping and stable plateau, activity in the integration dimension did not decay during such IBIs and therefore could not distinguish behavioral bouts from adjacent IBIs (Figure S1A, case 2). However, decoders trained on this activity could distinguish IBIs from sniffing versus attack epochs, which were behaviorally indistinguishable to a human observer, with a high F1 score (0.83 ± 0.02, N = 6 mice; Figure 2E).

Thus, our unsupervised approach uncovered a one-dimensional signal in VMHvl^{Esr1} neural population activity that closely tracks and scales with an animal's escalating level of aggressiveness and is reflected in the activity of approximately 20% of individual VMHvl^{Esr1} neurons. Different aggressive actions are observed as activity along this dimension reaches different thresholds, suggesting an aggression-intensity code in VMHvl^{Esr1} activity. The level of activity along the integration dimension could not be fully predicted from pose features such as the acceleration, facing angle, or velocity of the resident, or from the distance between mice (mean R^2: 0.28 ± 0.04, N = 6 mice, Figure S2A). Tracking metrics used as inputs to the model were also not predictive of behavior annotations (Figures S3F–S3H). Furthermore, models of VMHvl fit without any tracking inputs also recovered an integration dimension with similar time constants (Figure S3D). These results further highlight that the relationship between the integration dimension and escalating aggressive behavior is not due to the incorporation of inputs such as facing angle and distance between mice. Even the incorporation of additional tracking metrics such as speed and area of the ellipse fit to the resident mouse did not improve rSLDS fits, suggesting that VMHvl was likely not integrating features of these sensory related signals (Figure S3I).

This relationship between VMHvl^{Esr1} activity and aggression is consistent with our observation that increasing the intensity of optogenetic stimulation of VMHvl^{Esr1} neurons progressively evokes sniffing, dominance mounting, and attack,[12] actions that can be decoded from the integration dimension as its activity ramps up.

VMHvl contains an approximate line attractor that represents escalating aggressiveness

We examined next how the integration dimension of the fit model influences the overall topology of neural state space during social behavior (Figure 2A④; see STAR Methods). PCA indicated that the first two PCs accounted for 68.5% ± 1.2% of the total variance in VMHvl activity (N = 6 mice). In all imaged animals, PC1 showed slow ramping dynamics (Figure 3A; Figures S4C; PC_1 [behavior-triggered average, N = 6 mice]). We confirmed that the rSLDS integration dimension makes the largest contribution to this PC (Figure S4A). Activity along PC2 was high when a new intruder was introduced (Figure 3A; Figure S4C; PC_2 [behavior-triggered average, N = 6 mice]) but was otherwise low.

To visualize neural state space dynamics, we next generated a 2D flow field in PC space, whose vectors at each point indicate how neural dynamics evolve according to the fit rSLDS model (see Figure 2A④). This revealed a region of low vector flow that forms an approximate line attractor (Figures 3B, right and 3C), meaning that the neural population activity vector tends to move toward persistent points along a line[62] (Figure 3D, t_{50}-t_{340}). To quantitatively delimit this attractor, we calculated the points in the flow field where vector length is at a minimum ("slow points"; see STAR Methods) and linked these points into a dashed line (Figure 3D, dashed black line). Such approximate line attractors were observed in multiple mice (Figures 3E and 3F; Figures S4D and S4E). Importantly, these line attractors are largely aligned with the PC1 axis, which principally reflects variance in the slow integration dimension identified by rSLDS (Figure 2B; Figure S4A and S4B).

To quantitatively test for the existence of a line attractor in each mouse, we devised a "line attractor score" as the base-2

(J and K) Same as (H) but for VMHvl mouse 2 (J) and mouse 3 (K).
(L) Position of various behaviors along trough, i.e., PC1 in neural state space (n = 6 mice, **p < 0.005, *p < 0.01).
(M) Schematic showing quantification of dynamic velocity. Low dynamic velocity indicates that behavior exists in a stable region of neural state space.
(N) Dynamic velocity for various behaviors in VMHvl (***p < 0.001, n = 6 mice).
(O) Relationship between the time spent attacking and the time constant of the integration dimension of individual mice (r^2: 0.77, n = 14 animals). See Figures S4K–S4P for analysis restricted to GCaMP7f animals.

log of the ratio of the largest to the second-largest time constants of the eight rSLDS dimensions (Figure 2C). According to basic concepts in dynamical systems theory,[56] this ratio has a relatively high value in systems containing a single integration dimension (forming an approximate line attractor) and is otherwise close to zero. We find that all mice with VMHvl recordings possess a line attractor score greater than zero, indicating the presence of a line attractor (Figure 3G, n = 6 mice).

As population activity progressed along the line attractor from low to high values of PC1, behavior progressed from sniffing to dominance mounting to attack (Figures 3D–3F and 3L; Video S1). This reflects the "ramping up" of activity seen in the integration dimension as social behavior progresses through these phases (Figure 2C), suggesting an encoding of an underlying continuous variable, as seen in line attractors in other brain regions.[46,50,63–65]

To visualize the dynamical topology of the rSLDS model, we represented the 2D flow field as a 3D landscape, by converting the length of the flow-field vectors at each position in neural state space into the height (z axis) of the landscape (Figure 3B); the x-y axes are still represented by PC1 and PC2. In this topographic representation, a line attractor appears as a region shaped like a trough or gully, reflecting a slow rate of change (short vectors). A point attractor would appear as a locus of slow rate of change at the base of a cone (Figure 3B, left[63]). We observed a trough-like structure in the 3D dynamics landscape in each imaged animal (Figures 3H–3K; Figure S4P), along which neural activity progressed slowly as aggression escalated (Video S2). Consistent with the persistent, slow-decaying activity characteristic of "leaky" neural integrators,[59,63] VMHvl activity remained high following intruder removal and slowly decayed along the trough of the attractor over tens of seconds (Figures S4H–S4J).

Although the animals' behavior appears to occur while the system is in the line attractor, it could be that other rSLDS dimensions also show a change in their activity during behavior. To test this possibility, we computed each behavior's "dynamic velocity," by calculating the average vector length across all eight rSLDS dimensions at all time points in which a given behavior occurred (Figure 3M; see STAR Methods). Time points associated with initial intruder entry had the highest dynamic velocity and were present on the walls of the trough (Figures 3H–3K), whereas aggressive behaviors exhibited low dynamic velocities and were distributed along the base of the trough (Figure 3N).

Once the system is in the line attractor, input that is not aligned with the attractor should produce a transient excursion of the population activity vector out of the trough; however, once that input decays the vector should move back into the trough close to where it started[46] (Figure S4F). We tested this prediction using a subset of experiments in which one intruder male was removed and a second male introduced 30–60 s later. Strikingly, the introduction of a new intruder male drove a rapid rise in neural firing rates that pushed VMHvl activity away from the trough of the line attractor (Figures 3C–3E, intruder #2). However, this signal decayed relatively quickly and the system re-entered the line attractor at nearly the same point (Figures S4F–S4J; Video S1). Importantly, the system recovered to the point in the attractor where it had been prior to introduction of the second intruder, regardless of when in the trial the first intruder was removed (Figure S4K).

The time constant of the integration dimension in VMHvl predicts levels of aggressiveness across animals

Although VMHvl^{Esr1} imaging data from different mice always revealed a single integration dimension with a long time constant, the magnitude of this time constant varied across individuals. Unexpectedly, we observed a trend in which animals that displayed more aggressive behavior (calculated as the fraction of time spent attacking) also exhibited an integration dimension with a longer time constant (Figure 3O, r^2 = 0.77, n = 14 animals). This relationship held for imaging data from different studies[29,31,43] using different versions of GCaMP (6s versus 7f; Figures S4L–S4O). This striking correlation of integration time constant with time spent attacking suggests that individual differences in aggressiveness may be reflected in the intrinsic dynamics of VMHvl^{Esr1} neurons.

Mating behaviors are represented using rotational dynamics in the MPOA

Since rSLDS was able to uncover evidence for integration in VMHvl, we next examined whether the same analysis would uncover population dynamics important for mating in MPOA, by fitting models to MPOA^{Esr1} neural data recording during interactions with female intruders.[31]

Fit models of MPOA required three rSLDS states in every animal, with mounting and intromission mostly occurring in single but different states (Figures S5A–S5J). Unlike in VMHvl, the bout length of mating behaviors was similar to that of the corresponding state (Figures S5D and S5E). Strikingly, the eigenvalues of the dynamics matrix for such states did not include dimensions with long time constants (Figure 4A). Instead, the first two PCs of the fit model revealed fast dynamics that were highly correlated with specific behaviors (Figure 4B). PC1 peaked at the onset of USV^{+} mounting bouts, whereas PC2 peaked during intromission (Figures 4B and 4C, behavior-triggered average, N = 3 mice).

The 2D flow field in PCA space revealed that neural dynamics were dominated by a rotational flow, with activity during mating epochs exhibiting periodic orbits (Figures 4D and 4F). The phase of the rotations was correlated with progression through sniffing, mounting, and intromission (Figures 4D and 4F; Figures S5K–S5M) and corresponded to the sequential activation of different neurons during these successive behaviors (Figures 4E and 4G; Figure S5A). Accordingly, the "sequentiality index" of the data[66] was significantly greater than shuffled data or random matrices of similar sizes (seq. index = 0.22 ± 0.01, N = 3 mice, shuffle seq. index = 0.10 ± 0.002, N = 3 mice, Figure 4H).

We assessed the relationship between the phase of rotational trajectories and behavior by calculating the angle of the population activity vector relative to its value at the start of sniffing (Figure 4I). This revealed that sniffing, mounting, and intromission occurred at characteristic angles of the population vector (sniffing: 18.6° ± 6.2°, mounting: 79.61° ± 13.6°, intromission: 132.2° ± 8.1°, N = 3 mice; Figure 4J). High dynamic velocities were associated with mounting and intromission, in striking contrast to the low dynamic velocity during attack behavior in VMHvl (Figures 3N and 4K).

MPOA activity shows rotational dynamics during interaction with females

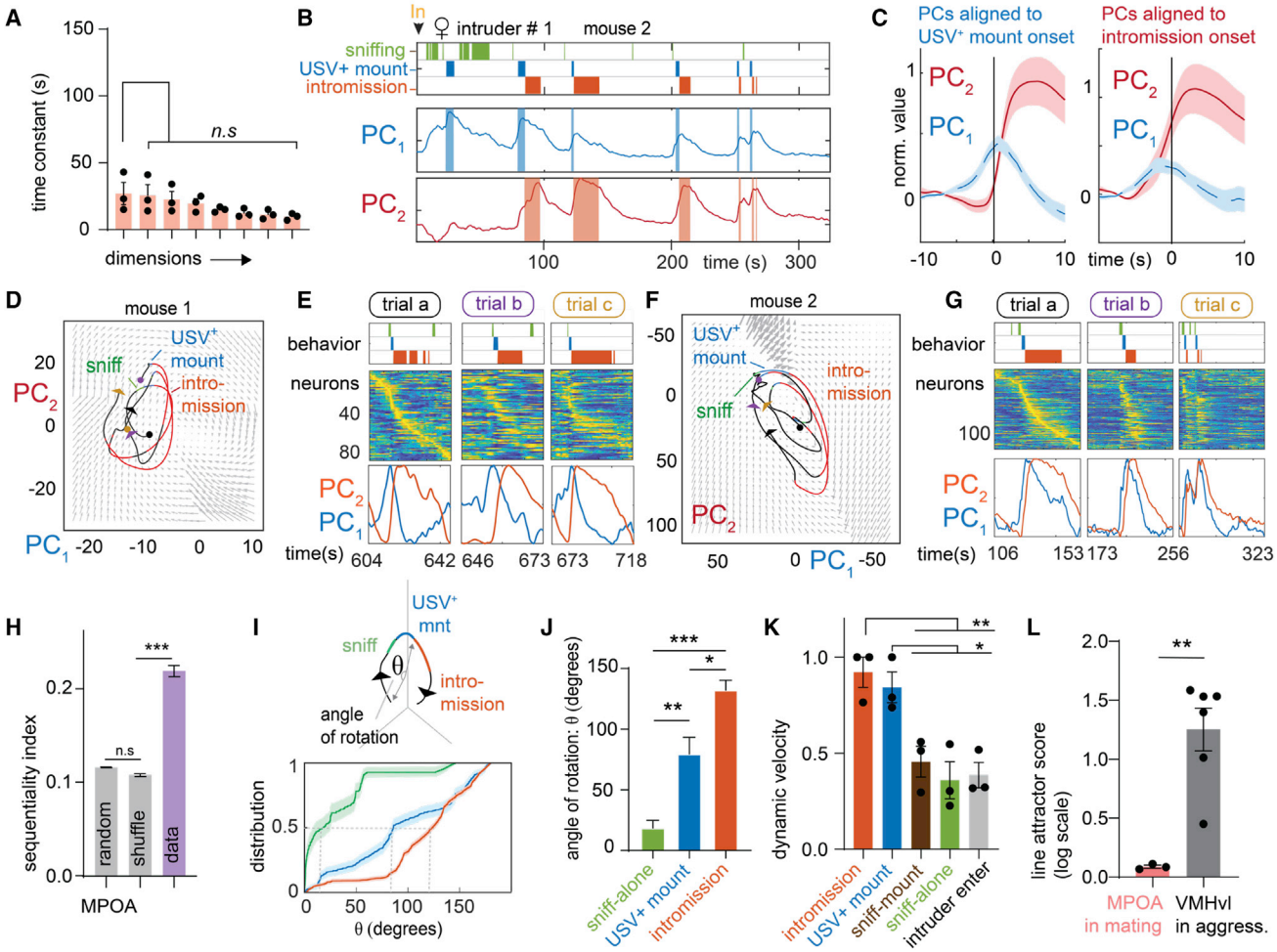


Figure 4. Mating behaviors are represented using rotational dynamics in the MPOA

(A) Time constants of rSLDS dimensions in mating behavior-enriched state in MPOA (n = 3 mice).

(B) Behavior rasters shown with first two principal components of latent factors for example MPOA mouse 2.

(C) Behavior-triggered average of top two principal components aligned to USV+ mount onset (left) and intromission (right) onset (n = 3 mice).

(D) Neural state space with rotational population trajectories from mating episodes shown in (E) of MPOA mouse 1, colored by behaviors performed by resident mouse.

(E) Sequential activity of MPOA neurons during mating episodes whose rotational population trajectories are shown in (D).

(F and G) Same as (D) and (E) but for MPOA mouse 2.

(H) Sequentiality index for MPOA (n = 3 mice, ***p < 0.001).

(I) Calculation of angle of rotation (θ) aligned to the start of sniffing during mating episodes (top). Empirical cumulative distribution of θ for various behaviors (n = 3 mice, bottom).

(J) Quantification of θ for various mating behaviors (n = 3 mice, ***p < 0.001, **p < 0.005, *p < 0.01, top).

(K) Dynamic velocity for mating behavior in MPOA (n = 3 mice,**p < 0.005, *p < 0.01).

(L) Line attractor score for MPOA activity in mating behaviors toward females (left, pink bar, n = 3 mice) and VMHvl activity in aggressive behavior toward males (right, gray bar, n = 6 mice, **p < 0.005, data from Figure 3G reproduced for comparative purposes). For additional data, see Figure S5.

To quantitatively assess the presence of line attractor dynamics, we computed the Line Attractor Score for MPOA. These values were close to zero and significantly different from those for VMHvl during aggression (Figure 4L). Thus, unlike the slow ramping and persistent dynamics identified in VMHvl, rSLDS-discovered fast, sequential, and behaviorally time-locked rotational dynamics in MPOA.

A direct comparison of key quantitative dynamics parameters highlights the key differences between VMHvl and MPOA (Figures 5A–5D, 5G, and 5H). Nevertheless, in both regions, evolving behavior tracks a single continuous variable: the value of the integration dimension in VMHvl and the angle of the orbit in MPOA (Figures 5E and 5F). These variables are instantiated as a line attractor versus rotational flow, respectively (Figures 5I and 5J).

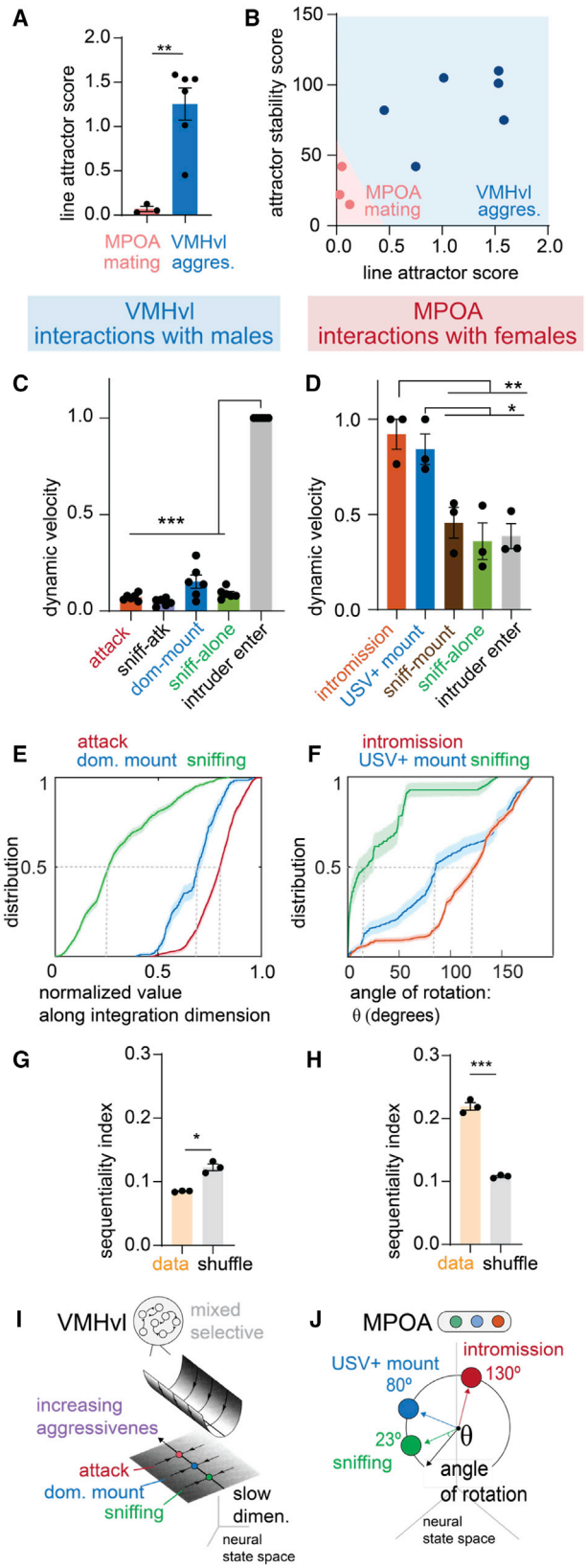


Figure 5. Distinct neural coding schemes for similar behavior in VMHvl versus MPOA

(A) Line attractor score for mating behavior in MPOA and aggressive behavior in VMHvl (n = 3 mice for MPOA, n = 6 mice for VMHvl), reproduced from Figure 4L.

(B) Scatter plot for line attractor score versus attractor stability score (magnitude of largest time constant) separates VMHvl and MPOA.

(C and D) Dynamic velocity score in VMHvl during aggression (C) and MPOA during mating (D), reproduced from Figures 3N and 4K, respectively.

(E) Empirical cumulative distribution of value of integration dimension (normalized) in VMHvl for various aggressive behaviors, reproduced from Figure 2F.

(F) Empirical cumulative distribution of angle of rotation (normalized) in MPOA for various mating behaviors, reproduced from Figure 4I.

(G and H) Sequentiality index in MPOA (n = 3 mice), reproduced from Figure 4E, and in VMHvl (H) in aggression (n = 3 mice).

(I) Summary of line attractor dynamics in VMHvl. Diagram modified from Seung.[63]

(J) Summary of rotational dynamics in MPOA.

VMHvl exhibits an approximate line attractor encoding reproductive behavior

The foregoing findings raised the question of whether the contrasting dynamics in VMHvl versus MPOA reflect differences specific to aggression versus mating or rather generic differences in behavioral coding between these nuclei. To address this, we fit rSLDS models to VMHvl^{Esr1} and MPOA^{Esr1} neuronal activity during mating versus aggression, respectively.

Models fit to VMHvl activity during male-female encounters yielded a single integration dimension with a long time constant, created by neurons that displayed ramping and persistent activity (Figures 6A red dot, 6B, and 6D; Figure S6L). In addition, the duration of the rSLDS-discovered mating states in VMHvl tended to outlast individual bouts of mating actions (Figures S6D and S6H), similar to the case of aggression in VMHvl (Figures S1I–S1K).

The cumulative distribution of the value of the integration dimension during various behaviors revealed that sniffing occurred at the lowest values, USV^{+} mounting at intermediate values, and intromission at the highest values of this dimension (Figure S6J). Strikingly, pairwise decoders trained on this dimension performed with high accuracy (intromission versus sniffing: F1 = 0.92 ± 0.01 N = 4 animals; mounting versus sniffing: F1 = 0.81 ± 0.02 N = 6 animals; Figure 6C). Such decoders could also distinguish periods of non-interaction between mounting bouts from those between sniffing bouts (Figure S6I). Thus, VMHvl^{Esr1} neuronal dynamics during mating resembled those exhibited during aggression. However, the integration dimension seen during mating was biased toward neurons tuned to female intruders,[29] whereas male-tuned neurons primarily contributed to this dimension during aggression (7.73% ± 0.8% overlap, n = 6 mice, Figures S3A, S3B, and S6M).

As for aggression, a single dimension of the rSLDS model for mating exhibited a long time constant, yielding a high line attractor score (Figures 6D and 6H). The first two PCs of the fit model were similar to those seen during aggression, with PC1 exhibiting ramping during the progression from sniffing to mounting to intromission (Figure 6E). Examination of the underlying 2D vector flow field revealed an approximate line attractor

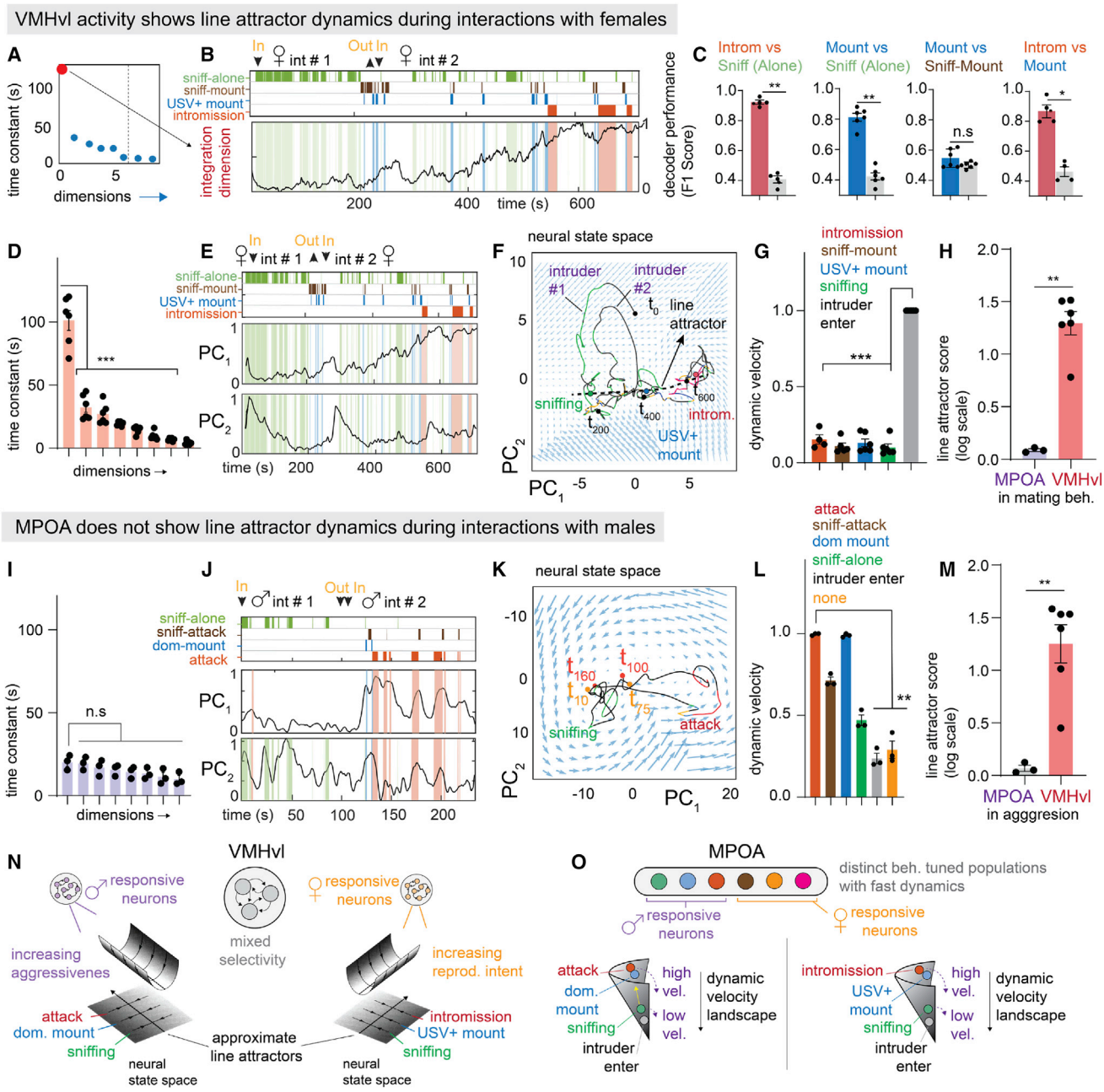


Figure 6. Distinct coding schemes of VMHvl and MPOA are region specific, not intruder specific

(A) (Left) Time constants of rSLDS dimensions of mating enriched state from example VMHvl mouse 1. The red dot highlights the integration dimension.

(B) Projection of integration dimension with overlayed behavior annotations.

(C) F1 score for decoding behavior pairs from integration dimension (**p < 0.005, *p < 0.01, n = 4 mice for comparisons involving intromission as only 4/6 mice showed this behavior. n = 6 mice for all other comparisons).

(D) Time constant arranged in decreasing order. (p < 0.001, n = 6 mice.)

(E) Behavior rasters shown with PCs of dynamical system for example VMHvl mouse 1.

(F) Neural state space with population trajectories for VMHvl mouse 1 colored by behavior annotations and flow field showing a line attractor.

(G) Quantification of dynamic velocity during mating behavior in VMHvl (***p < 0.001, n = 6 mice).

(H) Line attractor score for MPO (n = 3 mice) and VMHvl (n = 6 mice) during mating behavior with females (**p < 0.005).

(I) Time constants of rSLDS dimensions from MPOA during aggression.

(J) Behavior rasters shown with PCs of dynamical system for example MPOA mouse 1.

(K) Neural state space with population trajectories for MPOA mouse 1 colored by behavior annotations and flow field.

(L) Dynamic velocities during aggressive behaviors in MPOA (**p < 0.005, n = 3 mice).

(legend continued on next page)

(Figure 6F) and a corresponding trough shape in the 3D dynamic velocity landscape, with neural activity moving along the trough as the animal progressed from the appetitive to consummatory phases of mating (Figure S6K). Transient movements out of the line attractor occurred only during the introduction of a new intruder and were aligned with PC2 (Figure 6F, intruder #2). Accordingly, periods during intruder entrance had high dynamic velocities, whereas mating behaviors had low dynamic velocities (Figure 6G).

Thus, rSLDS modeling of VMHvl^{Esr1} neuronal activity during mating revealed an approximate line attractor, with many features similar to those observed during aggression. However, the mating and aggression line attractors incorporate primarily female- versus male-selective neurons, respectively (Figure S6M). These data suggest that line-attractor dynamics are a general feature of social behavior coding in VMHvl, rather than a unique signature of aggression per se.

MPOA does not exhibit line attractor dynamics during aggression

Finally, we fit rSLDS models to MPOA^{Esr1} neuronal dynamics during male-male encounters. Analogous to the case of mating behaviors, we found a state (state 3) that is closely aligned to the onset and offset of attack behavior (Figure S6N). No dominant "slow" dimension was apparent in the time constants of the rSLDS dimensions (Figures 6I and 6M). Reflecting this, PC1 of rSLDS state space exhibited a fast increase in activity at the onset and offset of attack (Figure 6J; Figure S6O, blue trace), in contrast to the slow attack-related dynamics in VMHvl (Figure S6O, red trace).

Visualizing the 2D MPOA flow field in PC space revealed little change in the population trajectory during investigation (Figure 6K). During attack bouts, activity showed excursions into a separate region of state space but quickly returned to the "sniffing" region after fighting (Figure 6K), reflecting the activation of different neuronal subsets (Figure 1E). Accordingly, attack and dominance mounting had high dynamic velocities in MPOA, rather than the low dynamic velocities in VMHvl (Figure 6L; Figure S6P). Finally, the line attractor score in MPOA during aggression had a value close to zero and was significantly different from that of VMHvl (Figure 6M, n = 3 for MPOA, n = 6 mice for VMHvl), confirming the absence of line attractor dynamics.

In MPOA, therefore, we find a representation of male-male encounters that alternates between investigatory and aggressive states, with the latter largely time-locked to the onset and offset of attack bouts. Strikingly, MPOA activity during aggression lacks the persistence, ramping, and line attractor dynamics seen in VMHvl. Together with our analysis of VMHvl activity during mating, these results support the conclusion that MPOA and VMHvl exhibit fundamentally different coding of the same social behaviors.

DISCUSSION

MPOA and VMHvl control social behaviors using different population codes

Here, we report that MPOA^{Esr1} and VMHvl^{Esr1} neurons utilize very different schemes for the neural coding of mating and aggression, despite the fact that optogenetic perturbation specifically elicits mating in MPOA^{Esr1} neurons and attack in VMHvl^{Esr1} neurons. GCaMP imaging of Esr1^{+} neurons in MPOA indicates that specific actions can be decoded according to which cells are active,[31] consistent with transcriptomic studies.[33] In contrast, most VMHvl^{Esr1} neurons exhibit mixed behavioral selectivity in both imaging and transcriptomic studies.[29,36] Thus, MPOA represents behavior via a cell identity code, whereas VMHvl apparently does so via population coding.

Our studies suggest a possible mechanism underlying this population code. rSLDS analysis of VMHvl neural activity during male-male social interactions revealed one dimension of neural activity with a long time constant that exhibits progressively increasing activity during escalating aggressive encounters. In a topological representation, these dynamics can be visualized as a progression along a stable "trough" or gully, which has the characteristics of an approximate line attractor.[63] In contrast, rSLDS analysis of MPOA revealed rotational dynamics, generated by the sequential activity of behavior-specific cell types during each bout of mating. Put simply, VMHvl coding of behavior appears to be analog, whereas MPOA coding of behavior appears more digital.

In other neural systems, line attractors often encode a continuous, low-dimensional variable.[46,50] Here, this variable may correspond to the intensity of an aggressive internal state. VMHvl neurons have previously been implicated in the motivation to engage in fighting, as operationalized using instrumental conditioning assays.[23] However, such assays cannot measure aggressive motivation during attack itself, for technical reasons. The escalating (scalable) nature of aggression has ethological relevance as a means of establishing dominance while minimizing the risk of injury.[67] Unexpectedly, in comparing data across multiple animals, we discovered a strong positive correlation between each mouse's level of aggressiveness and the magnitude of the time constant of its integration dimension. This result reveals a neural correlate of individual differences in aggressiveness within VMHvl.

The different neural codes for social behavior we have uncovered in VMHvl and MPOA may reflect their distinct neurochemical and cytoarchitectonic features. VMH neurons are primarily glutamatergic. Recurrent connectivity among excitatory neurons is often invoked as a mechanism to achieve persistent activity.[58,59,63] Indeed, there is evidence that glutamatergic neurons in the dorsomedial subdivision of the ventromedial hypothalamus (VMHdm) that encode persistent defensive

(M) Line attractor score for MPO (n = 3 mice) and VMHvl (n = 6 mice, reproduced from Figure 3G) during aggressive behavior (**p < 0.005).
(N) Schematic[63] illustrating two approximate line attractors discovered in VMHvl encoding scalable states of aggressiveness and mating.
(O) Schematic illustrating dynamics seen in MPOA showing similarity in stability of behaviors during interactions with males and females. For additional data see Figure S5.

behaviors exhibit local connectivity.[68] However, slow dynamics can also be achieved using neuromodulatory signaling, and there is indirect evidence for peptidergic transmission in VMHvl[36,37]

By contrast, MPOA neurons are 85% GABAergic; to our knowledge, there is no way to achieve similar graded and persistent signals within a population of inhibitory neurons. However, GABAergic neurons could provide a substrate for reciprocal inhibitory connections between action-specific subpopulations. Such connectivity could produce winner-take-all dynamics or feedforward dis-inhibitory circuits that control transitions between sequential action phases of mating, e.g., from sniffing to mounting,[43] giving rise to the rotational dynamics observed in neural data. The existence of such circuits in MPOA can be investigated using slice physiology or *in vivo* imaging once genetic access to the appropriate cell types is achieved.

Why should MPOA and VMHvl utilize such different strategies for the coding of closely related social behaviors? It is tempting to attribute this difference in population dynamics to distinct features of reproductive versus aggressive behavior. For example, aggressive encounters can dynamically escalate or de-escalate to avoid serious injury or death to the combatants, whereas male mating must proceed to completion (ejaculation) to be reproductively beneficial. These differences are well-suited to control by ramping and rotational neuronal dynamics, respectively. In this view, the different properties and coding strategies of VMHvl and MPOA may have evolved to be optimally adaptive for fighting and mating, respectively.

However, our analysis also revealed approximate line attractor dynamics in a subset of VMHvl^{Esr1} neurons that is female-tuned and active during mating.[29] This suggests that line attractor-like dynamics are a general property of behavioral coding by VMHvl, not an aggression-specific feature. Conversely, MPOA contains specific Esr1^{+} neurons highly tuned to attack which do not exhibit line attractor dynamics (although there is no evidence that these neurons play a causative role in aggression). These data suggest that MPOA and VMHvl more likely encode different features of a given social behavior, such as action selection versus motive state intensity, respectively. If so, then by extension, the hypothalamus may contain GABAergic populations that control action selection during aggression. Indeed, the anterior hypothalamic nucleus (AHN), which has a similar neurochemical and cytoarchitectonic structure as MPOA, can promote defensive attack[69,70]; it will be interesting to see whether rotational dynamics are observed in this structure. By the same token, PMv which controls aggression and is also primarily glutamatergic,[54,71,72] may utilize population coding like VMHvl.

Potential functions of the VMHvl line attractor

Line attractors have been identified in cortical and hippocampal regions involved in cognitive functions, such as decision-making, spatial mapping, and sensory discrimination.[46,50] They have also recently been shown to encode reward history in the habenula.[65] However, it is unexpected to find such neural dynamics in the hypothalamus, which is widely viewed as controlling innate behaviors via action-specific cell types (as observed in MPOA[33]). What function(s) might such attractor dynamics serve, in the context of innate behaviors? Two explanations are possible, which are not mutually exclusive.

As mentioned earlier, progression along the line attractor may encode the intensity of an internal motive state of aggressiveness. This is supported by our finding that the integration dimension that contributes to this attractor can distinguish periods of non-social interaction during high- versus low-intensity phases of aggressive escalation (Figure 2E). In this view, the line attractor serves to maintain the system in a stable internal motive state that persists continuously during stochastic expressions of observable attack.

Previous studies have indicated that the greatest source of variance in VMHvl^{Esr1} neural activity is intruder sex.[29] Whether VMHvl encodes intruder sex per se or an internal motive state tightly correlated with intruder sex has been difficult to distinguish because males only attack other males and not females. In female mice, however, lactating mothers attack intruders of both sexes. Recently, we identified a subset of VMHvl^{Esr1} neurons in females that express the GPCR gene *Npy2r*, called ß cells, which are both necessary for maternal aggression and sufficient to promote attack in non-aggressive virgins.[41] Bulk calcium measurements revealed that ß cells are strongly active during maternal aggression toward both male and female intruders. However, these cells display low activity in individual females that are non-aggressive.[41] Thus, in females, the encoding of aggressive state by VMHvl^{Esr1} neurons can be decoupled from the encoding of intruder sex. These data reinforce the idea that in males, the VMHvl^{Esr1} line attractor (which reflects a dimension weighted primarily by male-selective neurons) encodes aggressiveness, rather than simply intruder sex.

An alternative function for the line attractor is that it may serve as an integrator that accumulates ''evidence'' used to make behavioral decisions, such as the decision to switch from sniff to dominance mount or from dominance mount to attack. Such a function would require that different behaviors be triggered at different threshold values of the integrator. This type of ramp-to-threshold mechanism has been suggested to control sequential actions during male courtship behavior in *Drosophila*[73] and predator escape in mice.[74] These two hypotheses are not incompatible: the attractor could encode both the intensity of an internal state and (indirectly) the selection of actions at different state intensities.

Line attractor dynamics could also serve useful functions in the context of behavioral plasticity and individual variation. For example, VMHvl^{Esr1} neurons show increased selective tuning for male versus female intruders as a function of social experience[29] and exhibit a form of long-term potentiation that underlies the increase in aggressiveness that occurs when mice win a series of fights.[75] It will be interesting to determine whether changes in flow field dynamics or attractor properties are associated with these forms of experience-dependent plasticity. Finally, we note that differences in line attractor properties were observed among mice which exhibited different and characteristic levels of aggressiveness (Figure 3O). It is possible that individual differences in aggressiveness may reflect, or be caused by, individual constraints on population dynamics in VMHvl.

Testable predictions of the line-attractor model

Our rSLDS model of VMHvl dynamics makes several testable predictions and raises several interesting questions for future investigation. First, it predicts that once in the attractor, the system will return quickly to it following perturbations that move it out of this stable trough. This behavior is suggested by the brief excursion out of the attractor that occurs when a first intruder is removed and replaced by a second one. However, it would be ideal to demonstrate this directly by transiently activating neurons that contribute to the attractor and determining whether the system rapidly returns to it following stimulus offset, as has been demonstrated for point attractors underlying working memory in the anterior lateral motor cortex (ALM).[76] Another prediction is that selectively inactivating the VMHvl^{Esr1} neurons that exhibit slow dynamics should eliminate activity along the line attractor. Such experiments will require combined optogenetic perturbations and calcium imaging in this deep subcortical structure. Such experiments will also be critical to confirm whether line attractor properties indeed play a causative role in controlling levels of aggressiveness.

The results herein show that about 20% of VMHvl^{Esr1} neurons exhibit persistent activity and ramping dynamics, raising the question of whether these cells constitute a genetically determined subpopulation. Single-cell RNA-seq experiments have shown that the Esr1^{+} population in VMHvl can be subdivided into 6–7 distinct transcriptomic subtypes.[36] Whether any of these subtypes selectively contributes to attractor dynamics can be addressed once genetic drivers specific for these subtypes are available. An additional question is whether the slow dynamics observed for some VMHvl^{Esr1} neurons reflects recurrent connectivity between them, as has been demonstrated for fear-encoding neurons in VMHdm,[68] or the release of slow-acting neuromodulators such as neuropeptides. Recurrent connectivity in VMHvl can be investigated by slice electrophysiology[37] and ultimately by EM connectomics. VMHvl^{Esr1} neurons are known to express multiple neuropeptides, as well as receptors for neuropeptides and other neuromodulators. New sensors for detecting neuromodulator release,[77,78] as well as methods for dynamically perturbing neuromodulator function *in vivo*, should help to address these questions in the future.

Limitations of the study

Our discovery of approximate line attractor dynamics in VMHvl derives from quantitative analysis of a dynamical system model fit to neural data. Although this analysis has revealed several conditions required for line attractor dynamics, such as persistence in the absence of input and robustness to behavioral perturbation, a definitive test requires experimental perturbation of neural activity.[62] Perturbations are also required to determine the contributions to line attractor dynamics of region-intrinsic versus extrinsic (i.e., via other nuclei) recurrent dynamics and feedback, as well as whether the attractor is truly "autonomous" and not input driven. The biological line attractor in VMHvl is a leaky approximation of a mathematically defined line attractor, exhibiting slow decay over timescales similar to line attractors discovered in other neural systems.[50,63] Further knowledge of the underlying neural mechanisms is required to understand the extent to which the region of stability identified here approximates a true line attractor.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Neural imaging data
- METHOD DETAILS
  - Tuning rasters for single neurons
  - Computation of pose features for input to dynamical model
  - Dynamical system models of neural data
  - Estimation of time constants
  - Calculation of line attractor score
  - Decoding behavior from integration dimension
  - Low dimensional (PCA) representation of dynamical system
  - Dynamic velocity as a measure of stability in a dynamical system and visualization as 3D landscape
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2022.11.027.

ACKNOWLEDGMENTS

We thank H. Inagaki and L.F. Abbott for critical feedback on this manuscript, C. Chiu for laboratory management, G. Mancuso for administrative assistance, and members of the Anderson and Kennedy labs for helpful comments on this project. A.N. is supported by a National Science Scholarship from the Agency for Science, Technology and Research, Singapore. D.J.A. is an investigator of the Howard Hughes Medical Institute. A.K. is supported by NIH R00MH117264. This work was supported in part by NIH grants NS123916, MH1223612, and MH070053 to D.J.A. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We support inclusive, diverse, and equitable conduct of research. We acknowledge that this research was conducted at Caltech, which is located on the unceded land of the Indigenous Tongva people.

AUTHOR CONTRIBUTIONS

D.J.A., A.K., S.G., M.J.S., and A.N. conceived of the project. A.N., A.K., and D.J.A. wrote the manuscript, with input from T.K., B.Y., and S.W.L. A.K. performed the clustering analysis of single neurons and A.N. performed all dynamical system modeling.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in their field of research or within their geographical location.

While citing references scientifically relevant for this work, we also actively worked to promote gender balance in our reference list.

Received: April 19, 2022
Revised: October 5, 2022
Accepted: November 22, 2022
Published: January 5, 2023

REFERENCES

1. Hahn, J.D., Sporns, O., Watts, A.G., and Swanson, L.W. (2019). Macroscale intrinsic network architecture of the hypothalamus. Proc. Natl. Acad. Sci. USA *116*, 8018–8027. https://doi.org/10.1073/pnas.1819448116.
2. Saper, C.B., and Lowell, B.B. (2014). The hypothalamus. Curr. Biol. *24*, R1111–R1116. https://doi.org/10.1016/j.cub.2014.10.023.
3. Paredes, R.G., and Baum, M.J. (1997). Role of the medial preoptic area/anterior hypothalamus in the control of masculine sexual behavior. Annu. Rev. Sex Res. *8*, 68–101.
4. Siegel, A., Roeling, T.A., Gregg, T.R., and Kruk, M.R. (1999). Neuropharmacology of brain-stimulation-evoked aggression. Neurosci. Biobehav. Rev. *23*, 359–389. https://doi.org/10.1016/s0149-7634(98)00040-2.
5. Canteras, N.S. (2002). The medial hypothalamic defensive system: hodological organization and functional implications. Pharmacol. Biochem. Behav. *71*, 481–491.
6. King, B.M. (2006). The rise, fall, and resurrection of the ventromedial hypothalamus in the regulation of feeding behavior and body weight. Physiol. Behav. *87*, 221–244. https://doi.org/10.1016/j.physbeh.2005.10.007.
7. Kruk, M.R. (2014). Hypothalamic attack: a wonderful artifact or a useful perspective on escalation and pathology in aggression? A viewpoint. Curr. Top. Behav. Neurosci. *17*, 143–188. https://doi.org/10.1007/7854_2014_313.
8. Swanson, L.W. (2005). Anatomy of the soul as reflected in the cerebral hemispheres: neural circuits underlying voluntary control of basic motivated behaviors. J. Comp. Neurol. *493*, 122–131. https://doi.org/10.1002/cne.20733.
9. Simerly, R.B. (2002). Wired for reproduction: organization and development of sexually dimorphic circuits in the mammalian forebrain. Annu. Rev. Neurosci. *25*, 507–536. https://doi.org/10.1146/annurev.neuro.25.112701.142745.
10. Wu, Z., Autry, A.E., Bergan, J.F., Watabe-Uchida, M., and Dulac, C.G. (2014). Galanin neurons in the medial preoptic area govern parental behaviour. Nature *509*, 325–330. https://doi.org/10.1038/nature13307.
11. Atasoy, D., Betley, J.N., Su, H.H., and Sternson, S.M. (2012). Deconstruction of a neural circuit for hunger. Nature *488*, 172–177. https://doi.org/10.1038/nature11270.
12. Lee, H., Kim, D.W., Remedios, R., Anthony, T.E., Chang, A., Madisen, L., Zeng, H., and Anderson, D.J. (2014). Scalable control of mounting and attack by Esr1+ neurons in the ventromedial hypothalamus. Nature *509*, 627–632. https://doi.org/10.1038/nature13169.
13. Lin, D., Boyle, M.P., Dollar, P., Lee, H., Lein, E.S., Perona, P., and Anderson, D.J. (2011). Functional identification of an aggression locus in the mouse hypothalamus. Nature *470*, 221–226. https://doi.org/10.1038/nature09736.
14. Yamaguchi, T. (2022). Neural circuit mechanisms of sex and fighting in male mice. Neurosci. Res. *174*, 1–8. https://doi.org/10.1016/j.neures.2021.06.005.
15. Zha, X., and Xu, X.H. (2021). Neural circuit mechanisms that govern intermale attack in mice. Cell. Mol. Life Sci. *78*, 7289–7307. https://doi.org/10.1007/s00018-021-03956-x.
16. Augustine, V., Lee, S., and Oka, Y. (2020). Neural control and modulation of thirst, sodium appetite, and hunger. Cell *180*, 25–32. https://doi.org/10.1016/j.cell.2019.11.040.
17. Sternson, S.M. (2013). Hypothalamic survival circuits: blueprints for purposive behaviors. Neuron *77*, 810–824. https://doi.org/10.1016/j.neuron.2013.02.018.
18. Yang, C.F., Chiang, M.C., Gray, D.C., Prabhakaran, M., Alvarado, M., Juntti, S.A., Unger, E.K., Wells, J.A., and Shah, N.M. (2013). Sexually dimorphic neurons in the ventromedial hypothalamus govern mating in both sexes and aggression in males. Cell *153*, 896–909. https://doi.org/10.1016/j.cell.2013.04.017.
19. Esteban Masferrer, M., Silva, B.A., Nomoto, K., Lima, S.Q., and Gross, C.T. (2020). Differential encoding of predator fear in the ventromedial hypothalamus and periaqueductal grey. J. Neurosci. *40*, 9283–9292. https://doi.org/10.1523/JNEUROSCI.0761-18.2020.
20. Kato, A., and Sakuma, Y. (2000). Neuronal activity in female rat preoptic area associated with sexually motivated behavior. Brain Res. *862*, 90–102. https://doi.org/10.1016/s0006-8993(00)02076-x.
21. Mandelblat-Cerf, Y., Ramesh, R.N., Burgess, C.R., Patella, P., Yang, Z., Lowell, B.B., and Andermann, M.L. (2015). Arcuate hypothalamic AgRP and putative POMC neurons show opposite changes in spiking across multiple timescales. eLife *4*, e07122. https://doi.org/10.7554/eLife.07122.
22. Gunaydin, L.A., Grosenick, L., Finkelstein, J.C., Kauvar, I.V., Fenno, L.E., Adhikari, A., Lammel, S., Mirzabekov, J.J., Airan, R.D., Zalocusky, K.A., et al. (2014). Natural neural projection dynamics underlying social behavior. Cell *157*, 1535–1551. https://doi.org/10.1016/j.cell.2014.05.017.
23. Falkner, A.L., Grosenick, L., Davidson, T.J., Deisseroth, K., and Lin, D. (2016). Hypothalamic control of male aggression-seeking behavior. Nat. Neurosci. *19*, 596–604.
24. Zhao, Z.D., Yang, W.Z., Gao, C., Fu, X., Zhang, W., Zhou, Q., Chen, W., Ni, X., Lin, J.K., Yang, J., et al. (2017). A hypothalamic circuit that controls body temperature. Proc. Natl. Acad. Sci. USA *114*, 2042–2047. https://doi.org/10.1073/pnas.1616255114.
25. Li, Y., Zeng, J., Zhang, J., Yue, C., Zhong, W., Liu, Z., Feng, Q., and Luo, M. (2018). Hypothalamic circuits for predation and evasion. Neuron *97*, 911–924.e5. https://doi.org/10.1016/j.neuron.2018.01.005.
26. Ghosh, K.K., Burns, L.D., Cocker, E.D., Nimmerjahn, A., Ziv, Y., El Gamal, A.E., and Schnitzer, M.J. (2011). Miniaturized integration of a fluorescence microscope. Nat. Methods *8*, 871–878.
27. Ziv, Y., Burns, L.D., Cocker, E.D., Hamel, E.O., Ghosh, K.K., Kitch, L.J., El Gamal, A., and Schnitzer, M.J. (2013). Long-term dynamics of CA1 hippocampal place codes. Nat. Neurosci. *16*, 264–266.
28. Jennings, J.H., Ung, R.L., Resendez, S.L., Stamatakis, A.M., Taylor, J.G., Huang, J., Veleta, K., Kantak, P.A., Aita, M., Shilling-Scrivo, K., et al. (2015). Visualizing hypothalamic network dynamics for appetitive and consummatory behaviors. Cell *160*, 516–527. https://doi.org/10.1016/j.cell.2014.12.026.
29. Remedios, R., Kennedy, A., Zelikowsky, M., Grewe, B.F., Schnitzer, M.J., and Anderson, D.J. (2017). Social behaviour shapes hypothalamic neural ensemble representations of conspecific sex. Nature *550*, 388–392.
30. Krzywkowski, P., Penna, B., and Gross, C.T. (2020). Dynamic encoding of social threat and spatial context in the hypothalamus. eLife *9*, e57148. https://doi.org/10.7554/eLife.57148.
31. Karigo, T., Kennedy, A., Yang, B., Liu, M., Tai, D., Wahle, I.A., and Anderson, D.J. (2021). Distinct hypothalamic control of same-and opposite-sex mounting behaviour in mice. Nature *589*, 258–263.
32. Wei, Y.-C., Wang, S.-R., Jiao, Z.-L., Zhang, W., Lin, J.-K., Li, X.-Y., Li, S.-S., Zhang, X., and Xu, X.-H. (2018). Medial preoptic area in mice is capable of mediating sexually dimorphic behaviors regardless of gender. Nat. Commun. *9*, 279.
33. Moffitt, J.R., Bambah-Mukku, D., Eichhorn, S.W., Vaughn, E., Shekhar, K., Perez, J.D., Rubinstein, N.D., Hao, J., Regev, A., Dulac, C., et al. (2018). Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. Science *362*, eaau5324.

34. Ishii, K.K., Osakada, T., Mori, H., Miyasaka, N., Yoshihara, Y., Miyamichi, K., and Touhara, K. (2017). A labeled-line neural circuit for pheromone-mediated sexual behaviors in mice. Neuron *95*, 123–137.e8. e128.

35. Kohl, J., Babayan, B.M., Rubinstein, N.D., Autry, A.E., Marin-Rodriguez, B., Kapoor, V., et al. (2018). Functional circuit architecture underlying parental behaviour. Nature, 1–21. https://doi.org/10.1038/s41586-018-0027-0.

36. Kim, D.W., Yao, Z., Graybuck, L.T., Kim, T.K., Nguyen, T.N., Smith, K.A., Fong, O., Yi, L., Koulena, N., and Pierson, N. (2019). Multimodal analysis of cell types in a hypothalamic node controlling social behavior. Cell *179*, 713–728.e17.

37. Shao, Y.Q., Fan, L., Wu, W.Y., Zhu, Y.J., and Xu, H.T. (2022). A developmental switch between electrical and neuropeptide communication in the ventromedial hypothalamus. Curr. Biol. *32*, 3137–3145.e3. https://doi.org/10.1016/j.cub.2022.05.029.

38. Lo, L., Yao, S., Kim, D.-W., Cetin, A., Harris, J., Zeng, H., et al. (2019). Connectional architecture of a mouse hypothalamic circuit node controlling social behavior. Proc. Natl. Acad. Sci. U.S.A. *116*, 7503–7512. https://doi.org/10.1073/pnas.1817503116.

39. Hashikawa, K., Hashikawa, Y., Tremblay, R., Zhang, J., Feng, J.E., Sabol, A., Piper, W.T., Lee, H., Rudy, B., and Lin, D. (2017). Esr1+ cells in the ventromedial hypothalamus control female aggression. Nat. Neurosci. *20*, 1580–1590.

40. Yang, T., Yang, C.F., Chizari, M.D., Maheswaranathan, N., Burke, K.J., Jr., Borius, M., Inoue, S., Chiang, M.C., Bender, K.J., Ganguli, S., and Shah, N.M. (2017). Social control of hypothalamus-mediated male aggression. Neuron *95*, 955–970.e4. https://doi.org/10.1016/j.neuron.2017.06.046.

41. Liu, M., Kim, D.W., Zeng, H., and Anderson, D.J. (2022). Make war not love: the neural substrate underlying a state-dependent switch in female social behavior. Neuron *110*, 841–856.e6. https://doi.org/10.1016/j.neuron.2021.12.002.

42. Hess, W.R., and Brügger, M. (1943). Das subkortikale Zentrum der affektiven Abwehrreaktion. Helv. Physiol. Pharmacol. Acta *1*, 33–52.

43. Yang, B., Karigo, T., and Anderson, D.J. (2022). Transformations of neural representations in a social behaviour network. Nature *608*, 741–749. https://doi.org/10.1038/s41586-022-05057-6.

44. Miller, P. (2016). Dynamical systems, attractors, and neural circuits. F1000Res *5*, 5. https://doi.org/10.12688/f1000research.7698.1.

45. Sussillo, D. (2014). Neural circuits as computational dynamical systems. Curr. Opin. Neurobiol. *25*, 156–163. https://doi.org/10.1016/j.conb.2014.01.008.

46. Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature *503*, 78–84.

47. Vyas, S., Golub, M.D., Sussillo, D., and Shenoy, K.V. (2020). Computation through neural population dynamics. Annu. Rev. Neurosci. *43*, 249–275.

48. Shenoy, K.V., Sahani, M., and Churchland, M.M. (2013). Cortical control of arm movements: a dynamical systems perspective. Annu. Rev. Neurosci. *36*, 337–359.

49. Churchland, M.M., Cunningham, J.P., Kaufman, M.T., Ryu, S.I., and Shenoy, K.V. (2010). Cortical preparatory activity: representation of movement or first cog in a dynamical machine? Neuron *68*, 387–400. https://doi.org/10.1016/j.neuron.2010.09.015.

50. Hulse, B.K., and Jayaraman, V. (2020). Mechanisms underlying the neural computation of head direction. Annu. Rev. Neurosci. *43*, 31–54.

51. Ebitz, R.B., and Hayden, B.Y. (2021). The population doctrine in cognitive neuroscience. Neuron *109*, 3055–3068. https://doi.org/10.1016/j.neuron.2021.07.011.

52. Linderman, S., Johnson, M., Miller, A., Adams, R., Blei, D., and Paninski, L. (2017). Bayesian learning and inference in recurrent switching linear dynamical systems. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 914–922.

53. Falkner, A.L., Dollar, P., Perona, P., Anderson, D.J., and Lin, D. (2014). Decoding ventromedial hypothalamic neural activity during male mouse aggression. J. Neurosci. *34*, 5971–5984. https://doi.org/10.1523/JNEUROSCI.5109-13.2014.

54. Itakura, T., Murata, K., Miyamichi, K., Ishii, K.K., Yoshihara, Y., and Touhara, K. (2022). A single vomeronasal receptor promotes intermale aggression through dedicated hypothalamic neurons. Neuron *110*, 2455–2469.e8. https://doi.org/10.1016/j.neuron.2022.05.002.

55. Segalin, C., Williams, J., Karigo, T., Hui, M., Zelikowsky, M., Sun, J.J., Perona, P., Anderson, D.J., and Kennedy, A. (2021). The Mouse Action Recognition System (MARS) software pipeline for automated analysis of social behaviors in mice. eLife *10*, e63720. https://doi.org/10.7554/eLife.63720.

56. Strogatz, S.H. (2018). Nonlinear Dynamics and Chaos with Student Solutions Manual: With Applications to Physics, Biology, Chemistry, and Engineering (CRC Press).

57. Maheswaranathan, N., Williams, A.H., Golub, M.D., Ganguli, S., and Sussillo, D. (2019). Reverse engineering recurrent networks for sentiment classification reveals line attractor dynamics. Adv. Neural Inf. Process. Syst. *32*, 15696–15705.

58. Major, G., and Tank, D. (2004). Persistent neural activity: prevalence and mechanisms. Curr. Opin. Neurobiol. *14*, 675–684. https://doi.org/10.1016/j.conb.2004.10.017.

59. Goldman, M.S., Compte, A., and Wang, X.J. (2009). Neural integrator models. In Encyclopedia of Neuroscience, L.R. Squire, ed. (Academic Press), pp. 165–178. https://doi.org/10.1016/B978-008045046-9.01434-0.

60. Zhu, Z., Ma, Q., Miao, L., Yang, H., Pan, L., Li, K., Zeng, L.H., Zhang, X., Wu, J., Hao, S., et al. (2021). A substantia innominata-midbrain circuit controls a general aggressive response. Neuron *109*, 1540–1553.e9. https://doi.org/10.1016/j.neuron.2021.03.002.

61. Harris, K.D. (2021). Nonsense correlations in neuroscience https://doi.org/10.1101/2020.11.29.402719.

62. Khona, M., and Fiete, I.R. (2022). Attractor and integrator networks in the brain. Nat. Rev. Neurosci. *23*, 744–766. https://doi.org/10.1038/s41583-022-00642-0.

63. Seung, H.S. (1996). How the brain keeps the eyes still. Proc. Natl. Acad. Sci. USA *93*, 13339–13344.

64. Ganguli, S., Bisley, J.W., Roitman, J.D., Shadlen, M.N., Goldberg, M.E., and Miller, K.D. (2008). One-dimensional dynamics of attention and decision making in LIP. Neuron *58*, 15–25.

65. Sylwestrak, E.L., Jo, Y., Vesuna, S., Wang, X., Holcomb, B., Tien, R.H., Kim, D.K., Fenno, L., Ramakrishnan, C., Allen, W.E., et al. (2022). Cell-type-specific population dynamics of diverse reward computations. Cell *185*, 3568–3587.e27. https://doi.org/10.1016/j.cell.2022.08.019.

66. Zhou, S., Masmanidis, S.C., and Buonomano, D.V. (2020). Neural sequences as an optimal dynamical regime for the readout of time. Neuron *108*, 651–658.e5. https://doi.org/10.1016/j.neuron.2020.08.020.

67. Kennedy, A. (2022). The what, how, and why of naturalistic behavior. Curr. Opin. Neurobiol. *74*, 102549. https://doi.org/10.1016/j.conb.2022.102549.

68. Kennedy, A., Kunwar, P.S., Li, L.Y., Stagkourakis, S., Wagenaar, D.A., and Anderson, D.J. (2020). Stimulus-specific hypothalamic encoding of a persistent defensive state. Nature *586*, 730–734. https://doi.org/10.1038/s41586-020-2728-4.

69. Xie, Z., Gu, H., Huang, M., Cheng, X., Shang, C., Tao, T., Li, D., Xie, Y., Zhao, J., Lu, W., et al. (2022). Mechanically evoked defensive attack is controlled by GABAergic neurons in the anterior hypothalamic nucleus. Nat. Neurosci. *25*, 72–85. https://doi.org/10.1038/s41593-021-00985-4.

70. Nelson, R.J., and Trainor, B.C. (2007). Neural mechanisms of aggression. Nat. Rev. Neurosci. *8*, 536–546. https://doi.org/10.1038/nrn2174.

71. Stagkourakis, S., Spigolon, G., Williams, P., Protzmann, J., Fisone, G., and Broberger, C. (2018). A neural network for intermale aggression to

establish social hierarchy. Nat. Neurosci. *21*, 834–842. https://doi.org/10.1038/s41593-018-0153-x.

72. Soden, M.E., Miller, S.M., Burgeno, L.M., Phillips, P.E.M., Hnasko, T.S., and Zweifel, L.S. (2016). Genetic isolation of hypothalamic neurons that regulate context-specific male social behavior. Cell Rep. *16*, 304–313. https://doi.org/10.1016/j.celrep.2016.05.067.
73. McKellar, C.E., Lillvis, J.L., Bath, D.E., Fitzgerald, J.E., Cannon, J.G., Simpson, J.H., and Dickson, B.J. (2019). Threshold-based ordering of sequential actions during drosophila courtship. Curr. Biol. *29*, 426–434.e6. https://doi.org/10.1016/j.cub.2018.12.019.
74. Evans, D.A., Stempel, A.V., Vale, R., Ruehle, S., Lefler, Y., and Branco, T. (2018). A synaptic threshold mechanism for computing escape decisions. Nature *558*, 590–594. https://doi.org/10.1038/s41586-018-0244-6.
75. Stagkourakis, S., Spigolon, G., Liu, G., and Anderson, D.J. (2020). Experience-dependent plasticity in an innate social behavior is mediated by hypothalamic LTP. Proc. Natl. Acad. Sci. USA *117*, 25789–25799. https://doi.org/10.1073/pnas.2011782117.
76. Inagaki, H.K., Fontolan, L., Romani, S., and Svoboda, K. (2019). Discrete attractor dynamics underlies persistent activity in the frontal cortex. Nature *566*, 212–217. https://doi.org/10.1038/s41586-019-0919-7.
77. Sabatini, B.L., and Tian, L. (2020). Imaging neurotransmitter and neuromodulator dynamics in vivo with genetically encoded indicators. Neuron *108*, 17–32. https://doi.org/10.1016/j.neuron.2020.09.036.
78. Sun, F., Zeng, J., Jing, M., Zhou, J., Feng, J., Owen, S.F., Luo, Y., Li, F., Wang, H., Yamaguchi, T., et al. (2018). A genetically encoded fluorescent sensor enables rapid and specific detection of dopamine in flies, fish, and mice. Cell *174*, 481–496.e19. https://doi.org/10.1016/j.cell.2018.06.042.
79. Linderman, S., Nichols, A., Blei, D., Zimmer, M., and Paninski, L. (2019). Hierarchical recurrent state space models reveal discrete and continuous dynamics of neural activity in C. elegans https://doi.org/10.1101/621540.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
neural imaging data from VMHvl and MPOA Esr1 neurons	Karigo et al.[31]	https://doi.org/10.1038/s41586-020-2995-0
neural imaging data from VMHvl Esr 1 neurons	Yang et al.[43]	https://doi.org/10.1038/s41586-022-05057-6
neural imaging data from VMHvl Esr1 neurons	Remedios et al.[29]	https://doi.org/10.1038/nature23885
Software and algorithms		
Matlab 2020a	Mathworks	N/A
Python	Python Software Foundation	N/A

RESOURCE AVAILABILITY

Lead contact

Requests for resources and reagents should be addressed to lead contact, David J. Anderson (wuwei@caltech.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Source data used in this paper will be shared by the lead contact upon request.
- Code used for analyses in this paper is available in the following repositories: https://github.com/lindermanlab/ssm and https://github.com/DJALab/VMHvl_MPOA_dynamics
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Neural imaging data

We analyzed data from three sets of previous experiments[29,31,43] All experiments were approved by the Institute Animal Care and Use Committee (IACUC) and the Institute Biosafety Committee (IBC) at the California Institute of Technology (Caltech). All experiments utilized heterozygous *Esr1*$^{cre/+}$ knock-in mice on a C457BL6/N background (B6N.129S6(Cg)-*Esr1*$^{tm1.1(cre)And}$/J, JAX strain #017911). Expression of GCaMP6s[29,31] or GCaMP7f[43] was achieved by stereotaxic injection of Cre-dependent GCaMP-expressing adeno-associated viruses (AAVs). Briefly, for data obtained from Karigo et al.,[31] mice expressing GCaMP6s selectively in *Esr1* neurons in either the medial preoptic area (MPOA) or the ventrolateral subdivision of the ventromedial hypothalamus (VMHvl), were allowed to interact with BALB/c male and female intruders in a standard resident intruder assay.[31] Male or female intruders were introduced into the home cage in a random order, with a 5–10 min interval between intruder session. Each session typically lasted 10–20 minutes. Behavior videos of interacting animals were annotated using a custom MATLAB-based interface. A total of 7 behaviors including sniffing, dominance-mount, attack, mount, intromission, interact (periods where animals were close to each other but other behaviors were absent) were annotated with male and female intruders. A head-mounted micro-endoscope (Inscopix, Inc.) was used to acquire Ca^{2+} imaging data at 15Hz from either MPOA^{Esr1} neurons (total of 583 neurons from 3 mice) or VMHvl^{Esr1} neurons (total of 421 neurons from 3 mice) for neural data analysis described in sections below.

For data obtained from Yang et al.,[43] Esr1-Cre mice in which GCaMP7f was expressed selectively in *Esr1* neurons in VMHvl, were allowed to interact with BALB/c male intruders in a standard resident intruder assay. In addition to the behaviors annotated for above, male intruders were also ''dangled'', where the ano-genital region of the dangled intruder is held next to the resident mouse. A head-mounted micro-endoscope was used to acquire Ca^{2+} imaging data at 30Hz from VMHvl^{Esr1} neurons (386 neurons from 3 mice) for neural data analysis described in sections below.

For data obtained from Remedios et al.,[29] Esr1-Cre mice in which GCaMP6s was expressed selectively in *Esr1* neurons in VMHvl were allowed to interact with BALB/c male intruders in a standard resident intruder assay. A head-mounted micro-endoscope was used to acquire Ca^{2+} imaging data at 30Hz from VMHvl^{Esr1} neurons (358 neurons from 3 mice) for neural data analysis described in sections below. rSLDS models were fit to data from n=14 mice to extract the time constant of the integration dimension used for correlation with individual differences in aggressiveness in Figure 3O. However 8 of those mice were excluded from decoder analysis of

sniffing, mounting and attack, either because they were highly aggressive and attacked without any prior sniffing or dominance mounting (5 mice), or because they were non-aggressive and failed to attack (3 mice). Typically 20%–25% of male mice from the C57BL6 background fail to show aggression in resident-intruder assays.[75]

METHOD DETAILS

Tuning rasters for single neurons

We examined the tuning properties of single neurons in VMHvl^{Esr1} or MPOA^{Esr1} by creating behavior tuning rasters (Figures 1C and 1D). We first computed the mean activity of each neuron for each of the 14 manually annotated behavioral actions. To group neurons, we created a set of 40 regressors representing combinations of behavioral actions, and grouped neurons by which single regressor captured the most variance in each cell's activity. In addition to regressors for individual behaviors, example regressors include signals such as all male-directed actions, all female-directed actions, all male-directed/female-directed/sex-invariant investigative behaviors, and all male-directed/female-directed/sex-invariant consummatory behaviors. Neurons for which no single regressor captured at least 50% of variance in behavior-averaged activity were omitted from the visualization (approximately 5% of cells.)

Computation of pose features for input to dynamical model

As external input to the dynamical model (see next section), we selected two features of animal pose estimates produced by the Mouse Action Recognition System (MARS,[55] The first of these is the distance between animals, computed as the distance between centroids of ellipses fit to the poses of the two mice. The second is the facing angle of the resident toward intruder mouse, defined as the angle between a vector connecting the centroids of the two mice and a vector from the centroid to the nose of the resident mouse. In addition we also fit dynamical models with either no input or with additional inputs in the form of the speed of the resident (computed as the mean change in position of centroids of the head and hips, computed across two consecutive frames) and area of ellipse fit to the resident mouse's pose.

Dynamical system models of neural data

We model neural activity using a recurrent switching linear dynamical systems (rSLDS) according to previous methods.[52,79] Briefly, rSLDS is a generative model that breaks down non-linear time series data into sequences of linear dynamical modes. The model relates three sets of variables: a set of discrete states (z), a set of continuous latent factors (x) that captures the low-dimensional nature of neural activity, and the activity of recorded neurons (y). The model also allows for external inputs (u) which consists of extracted pose features including the distance between animals and the facing angle between the resident and intruder mouse.

The model is formulated as follows: At each time t = 1,2,…T_n, there is a discrete state $z_t \in \{1, 2, \ldots, K\}$. In a standard SLDS, these states follow Markovian dynamics, however rSLDS allows for the transitions between states to depend recurrently on the continuous latent factors (x) and external inputs (u) as follows:

$$p(z_{t+1} = k, z_t = j, x_t) \propto exp\{Rx_t + Wu_t + r\} \quad \text{(Equation 1)}$$

where R, W and r parameterizes a map from the previous discrete state, continuous state and external inputs using a softmax link function to a distribution over the next discrete states.

The discrete state z_t determines the linear dynamical system used to generate the continuous latent factors at any time t:

$$x_t = A_{z_t} x_{t-1} + V_{z_t} u_t + b_{z_t} \quad \text{(Equation 2)}$$

where $A_k \in \mathbb{R}^{d \times d}$ is a dynamics matrix, $V_{z_t} \in \mathbb{R}^{d \times m}$ is a matrix that describes the contribution of external inputs (u_t) to each dimension of the latent space and $b_k \in \mathbb{R}^d$ is a bias vector, where d is the dimensionality of the latent space and m is the dimensionality of the external inputs. Thus, the discrete state specifies a set of linear dynamical system parameters and specify which dynamics to use when updating the continuous latent factors.

Lastly, we can recover the activity of recorded neurons by modelling activity as a linear noisy Gaussian observation $y_t \in \mathbb{R}^N$ where N is the number of recorded neurons:

$$y_t = Cx_t + d \quad \text{(Equation 3)}$$

For $C \in \mathbb{R}^{N \times D}$ and $d \sim N(0,S)$, a gaussian random variable. Overall, the system parameters that rSLDS needs to learn consists of the state transition dynamics, library of linear dynamical system matrices and neuron-specific emission parameters, which we write as:

$$\theta = \{A_k, V_k, b_k, C, d, R, W, r\}$$

These parameters are estimated using maximum likelihood using approximate variational inference methods as described in detail in Linderman et al.[52] and Linderman et al.[79]

Model performance is reported as the *evidence lower bound (ELBO)* which is equivalent to the Kullback-Leibler divergence between the approximate and true posterior, $KL(q(x,z;\varphi) \,||\, p(x,z\,|y;\theta))$ using 5-fold cross validation.

Since the ELBO is sensitive to the inclusion of regularizers and the amount of data used during fitting, we also provide an additional "forward simulation error (FSE)" model evaluation metric calculated as follows: given observed neural activity in state space at time t, we predict the trajectory of the population activity vector over an ensuing small time interval Δt using the model, then compute the mean squared error (MSE) between that trajectory and the observed data at time $t+\Delta t$ (Figure S1F). This MSE is computed across all dimensions of the latent space and repeated for all times t. This error metric is normalized to a 0-1 range in each animal across the whole recording to obtain a bounded measure of model performance (Figure S1F). This metric is computed across cross-validation folds and can provide intuition about time segments where model performance drops

Code used to fit rSLDS on neural data is available in the SSM package: (https://github.com/lindermanlab/ssm)

Code to generate flow fields and energy landscapes from fit dynamical systems is available in (https://github.com/DJALab/VMHvl_MPOA_dynamics)

Estimation of time constants

We estimated the time constant of each mode of linear dynamical systems using eigenvalues λ_a of the dynamics matrix of that system, derived by Maheswaranathan et al.[57] as:

$$\tau_a = \left| \frac{1}{\log(|\lambda_a|)} \right|$$

Calculation of line attractor score

To provide a quantitative measure of the presence of line attractor dynamics, we devised a line attractor score defined as:

$$\textit{line attractor score} = \log_2 \frac{t_n}{t_{n-1}}$$

where t_n is the largest time constant of the dynamics matrix of a dynamical system and t_{n-1} is the second largest time constant. This measure would be zero in a system without line attractor dynamics due to the similar magnitudes of the first two largest time constants and would be greater than one for systems that possess a line attractor.

Decoding behavior from integration dimension

We trained a frame-wise decoder to discriminate pairs of behavior (such as sniffing vs attack) from the activity of the integration dimension on individual frames of a behavior (sampled at 15Hz) as described previously.[31] We first created 'trials' from bouts of social behavior by merging all bouts that were separated by less than five seconds. We then trained a linear support vector machine (SVM) to identify a decoding threshold that maximally separates the values of our normalized "integration dimension" signal on frames during which behavior A occurred from values on frames during which behavior B occurred, for the pair-wise behavioral comparison. 'Shuffled' decoder data was generated by setting the decoding threshold on the same "trial", but with the behavior annotations randomly assigned to each behavior bout. We repeated shuffling 20 times for each intruder and each imaged mouse. We report performances of actual and shuffled 1D-threshold "decoders" as the average F1 score of the fit decoder, on data from all other "trials" for each mouse. For significance testing, the mean accuracy of the decoder trained on shuffled data was computed across mice, with shuffling repeated 1000 times for each mouse. Significance is determined by bootstrapping; we considered observed F1 scores significant if they fell above the 97.5th percentile of the distribution of chance F1 scores as done previously.[29]

As a stringent test for spurious correlations due to the slow decay seen in the integration, we performed a variation of session permutation[61] as follows. Consider a neural signal that displays a slow ramp in activity, which can be used to decode attack from sniffing (Figure S2E). If this correlation was spurious and occurred due to slow drift in activity, that decoding threshold would perform poorly if used on the integration dimension from another mouse (Figure S2F). On the other hand, that same threshold would produce a high F1 score if the correlation was not spurious as shown in Figure S2G. To implement this paradigm, we used the decoding threshold obtained in a given mouse on the integration dimension from all other mice and averaged the final performance.

Low dimensional (PCA) representation of dynamical system

Since the latent states are invariant to linear transformations, it is possible to apply a suitable transformation to obtain an equivalent model using rSLDS. We use PCA for this transformation as it allows us to describe our high dimensional rSLDS latent space in a concise manner with few dimensions while capturing the overall dynamics. To perform this the following steps are applied:

1 Given latent factors: $x_1, x_2, \ldots, x_t$ of the raw neural data y_t
2 Compute a whitening transformation W such that Wx is the identity
3 Compute the transformed linear dynamical system $x_t' = Wx_t$ with new emission matrix $C' = CW^{-1}$.
4 Compute the singular value decomposition (SVD) of the new emission matrix $C' = USV^T$. Let $P = SV^T$, such that$P^{-1} = \frac{V}{S}$
5 Compute the final transformed latent states (i.e principal components) $x_t'' = P^{-1}x_t{}' = P^{-1}Wx_t$

In this final transformation, since the singular values are ordered, the first two components of x_t'' accounts for the most variance in the raw neural data y_t. This method of applying PCA also accounts for the emission matrix C of the fit dynamical system.

Dynamic velocity as a measure of stability in a dynamical system and visualization as 3D landscape

We devised a metric termed the "dynamic velocity" to quantify the average intrinsically generated rate of change of the fit dynamical system during a given behavior of interest. We first calculated the average norm of $A_{z_t}x_t$ for every value of x_t associated with a given behavior, for a given state z. We then averaged this value across states, giving a definition of $V_b = \frac{1}{n(Z)}\sum_{z\in Z}\left(\frac{1}{n(T_b)}\sum_{t\in T_b} A_{z_t}x_t\right)$, where Z is the set of states, T_b is the set of all timepoints during which behavior b occurred, $\|\cdot\|$ is the Euclidean norm, and $n(\cdot)$ is the number of elements in a set. Finally, to facilitate comparison across animals, we normalized this value to a 0-1 range, with respect to its maximum across behaviors in each animal. Low values of this measure close to zero indicate regions with high stability while large values indicate unstable regions of neural state space.

We also converted the flow-fields obtained from rSLDS into a 3D landscape for visualization by calculating the dynamic velocity at each point in neural state space and using it as the height of a 3D landscape.

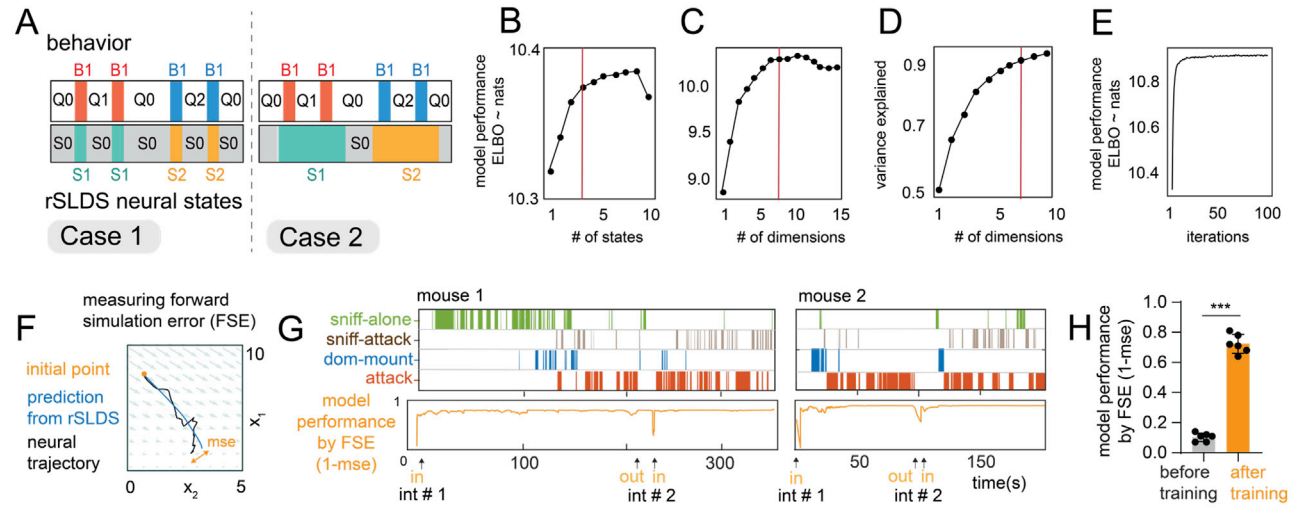
QUANTIFICATION AND STATISTICAL ANALYSIS

Data were processed and analyzed using Python, MATLAB, and GraphPad (GraphPad PRISM 9). All data were analyzed using two-tailed non-parametric tests. Mann-Whitney test were used for binary paired samples. Friedman test was used for non-binary paired samples. Kolmogorov-Smirnov test was used for non-paired samples. Multiple comparisons were corrected with Dunn's multiple comparisons correction. Not significant (NS), $P > 0.01$; *$P < 0.01$; **$P < 0.005$; ***$P < 0.001$; ****$P < 0.0001$.

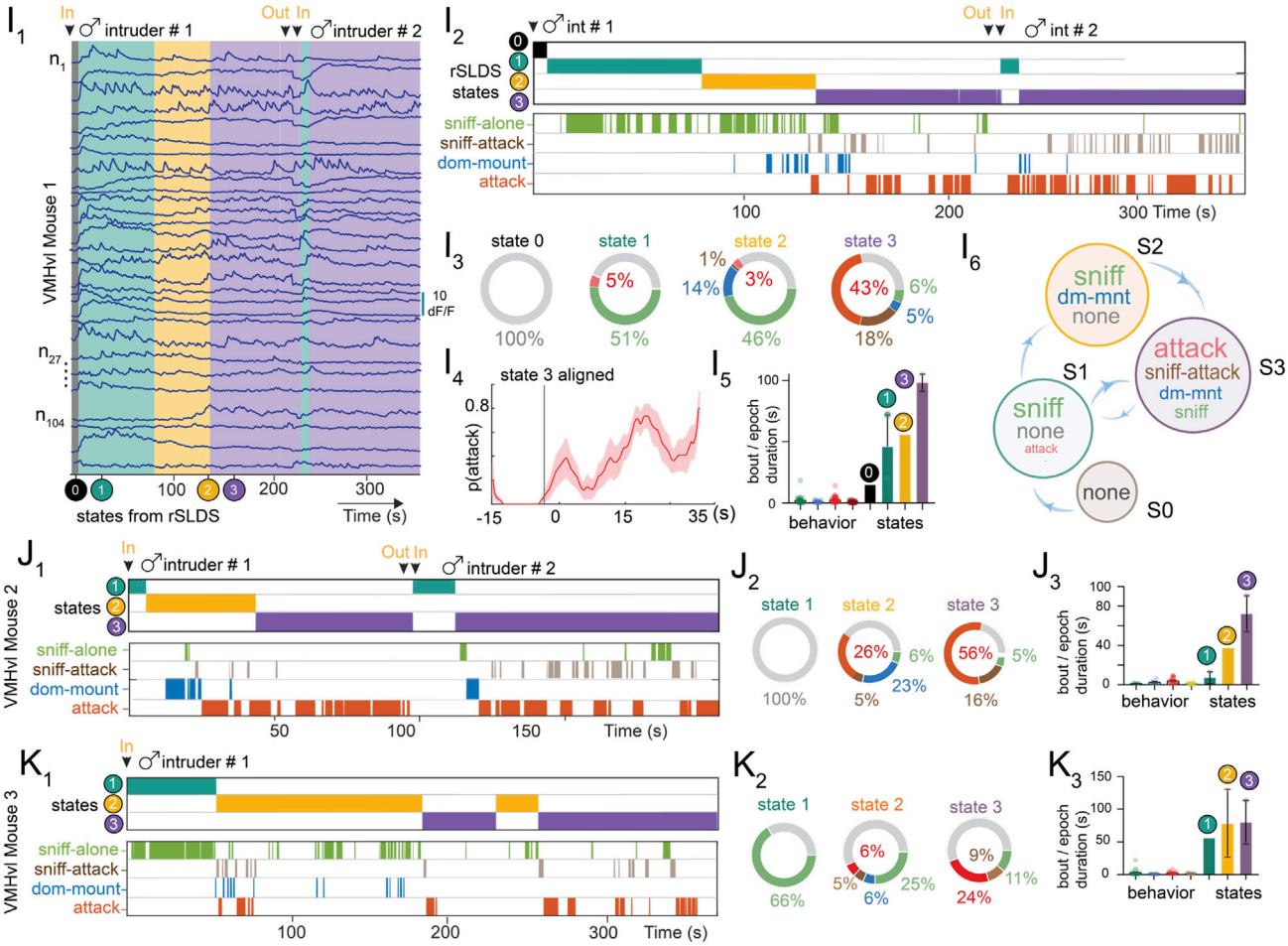
Cell
Theory

CellPress
OPEN ACCESS

Supplemental figures



dynamical analysis of VMHvl reveals aggression enriched states during interactions with males



(legend on next page)

Figure S1. Unsupervised discovery of aggression-enriched states in VMHvl, related to Figure 2

(A) Types of neural states identified by rSLDS. B1, B2: behaviors; Q0, Q1: periods of quiescence between behavior bouts; S0, S1, S2: rSLDS states. Case 1: rSLDS states cannot distinguish behavior versus internal states. Case 2: rSLDS reflects internal state-encoding due to persistence during behavioral quiescence.

(B) Optimization of number of rSLDS states in example VMHvl mouse 1. Model performance is measured as ELBO (see STAR Methods).

(C) Same as (B), but for dimensionality.

(D) Variance explained by dimension chosen in (C).

(E) Convergence of model performance.

(F) Creation of a bounded model performance metric (forward simulation error [FSE]; see STAR Methods).

(G) FSE for VMHvl mouse 1 & 2.

(H) Average model performance (FSE) before and after training (n = 6 mice,***$p < 0.001$)

(I) (I_1) rSLDS states in VMHvl mouse 1. (I_2) Comparison of rSLDS states with behaviors. (I_3) Behavioral composition of rSLDS states. State 3 possesses the highest amount of attack behavior across mice (see J and K). (I_4) Probability of attack aligned to the onset of state 3 (n = 6 mice). (I_5) Timescale of behavior bouts and discovered states epochs. (I_6) State transition diagram from empirical transition probabilities.

(J) Same as F_2, F_3, F_5 but for VMHvl mouse 2.

(K) Same as F_2, F_3, F_5 but for VMHvl mouse 3.

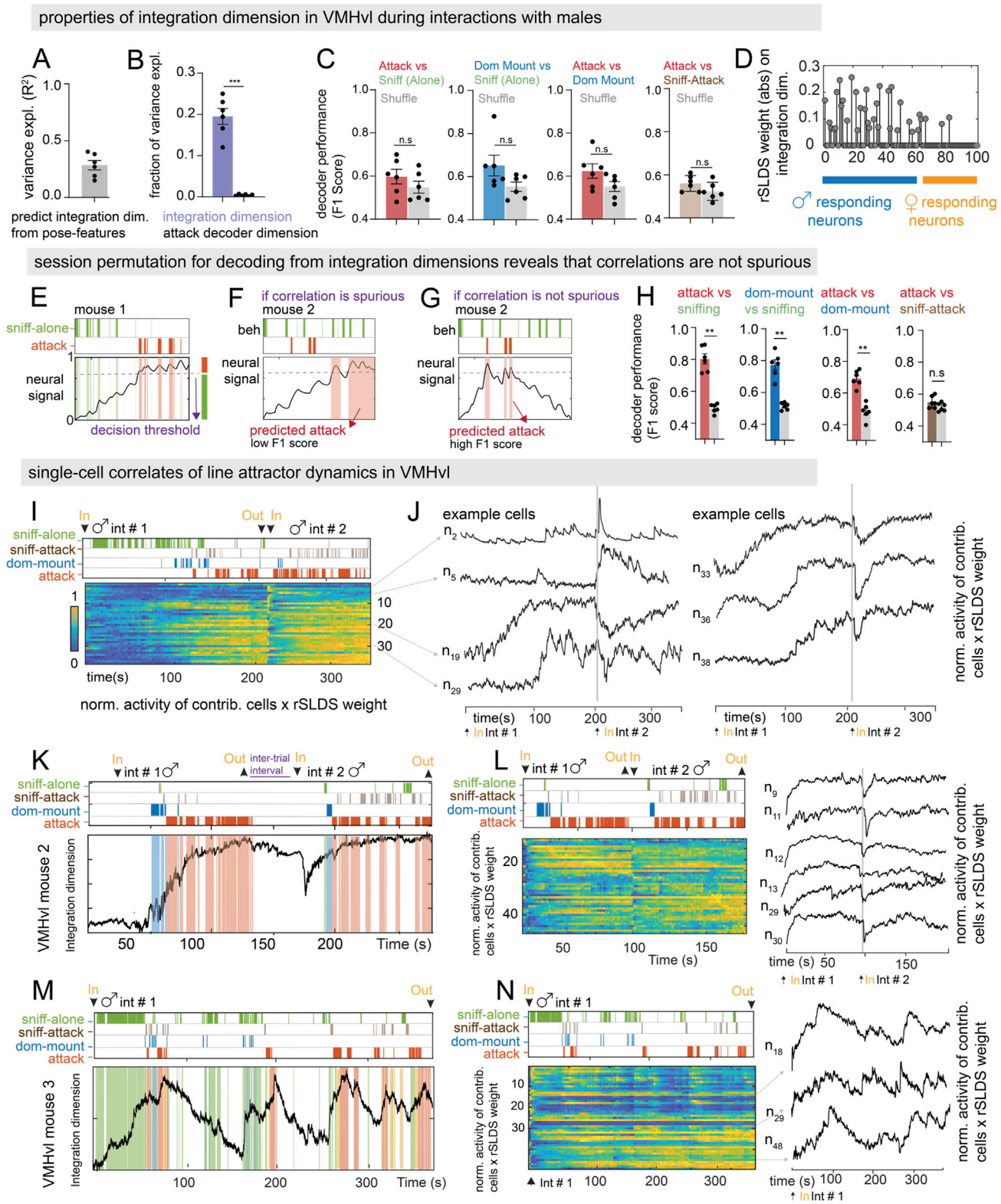


Figure S2. Characterization of aggression-integration dimension, related to Figure 2

(A) Variance explained by a generalized linear model trained to predict integration dimension from pose features including distance between mice, facing angle, speed, acceleration, and velocity of resident mouse (mean: 0.28 ± 0.04 R^2, n = 6 mice).

(legend continued on next page)

(B) Fraction of overall variance explained by integration dimension (purple) compared to variance explained by decoder dimension trained to distinguish attack from sniff bouts (integration dimension mean: 19.5% ± 1.9%, attack decoder mean: 0.3% ± 0.1%, n = 6 mice, ***$p < 0.001$).
(C) Decoding behaviors from non-integration dimensions (average across dimensions, n = 6 mice).
(D) Absolute rSLDS weight on integration dimension of VMHvl mouse 1 (cell number on x axis), sorted by choice probability values for male versus female intruder encounter.
(E–G) Paradigm to account for spurious correlations: decoding threshold obtained using integration dimension of mouse 1 (E, purple line) is used on integration dimension from mouse 2 (F). Spurious correlations lead to low F1 scores (F) while true correlations retain high F1 scores (G).
(H) Decoding behaviors using paradigm described above (**$p < 0.005$, n = 6 mice).
(I) Normalized activity of neurons times rSLDS weight for cells with significant weights for integration dimension of VMHvl mouse 1.
(J) Example cells from (I).
(K) Integration dimension in VMHvl mouse 2.
(L) Same as (I) for VMHvl mouse 2.
(M and N) Same as (K) and (L) for VMHvl mouse 3.

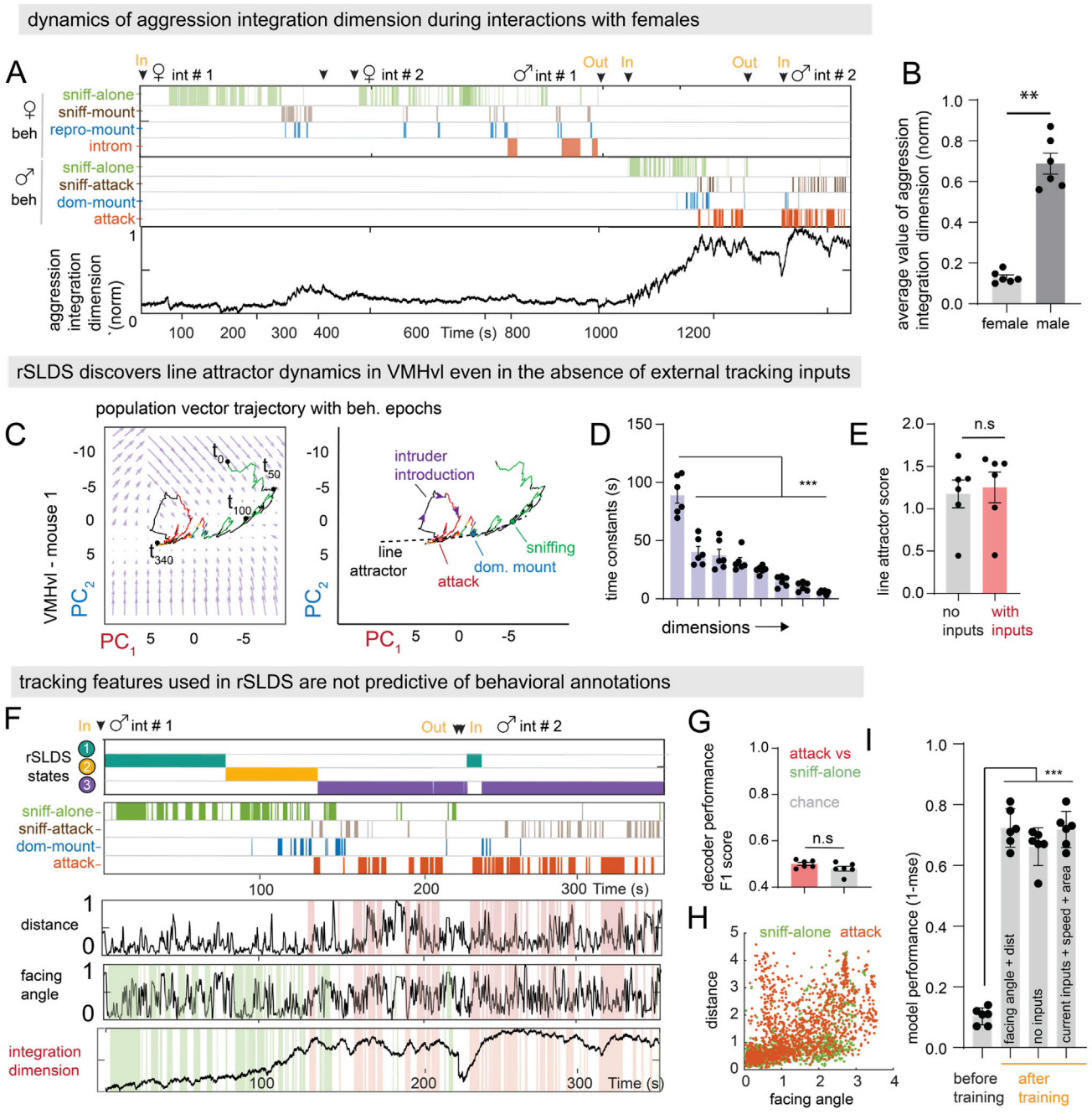


Figure S3. Characterization of aggression-integration dimension and dependence on tracking feature based external inputs, related to Figure 2

(A) Aggression-integration dimension in female and male trials in VMHvl mouse 1.

(B) Mean projection of neural activity from female versus male trials onto the aggression-integration dimension (n = 6 mice, **p < 0.005).

(C) Low dimensional dynamics and flow field from model with no behavioral inputs included with line attractor highlighted.

(D) Time constants from the fit dynamical system (n = 6 mice).

(E) Line attractor score for VMHvl models without input.

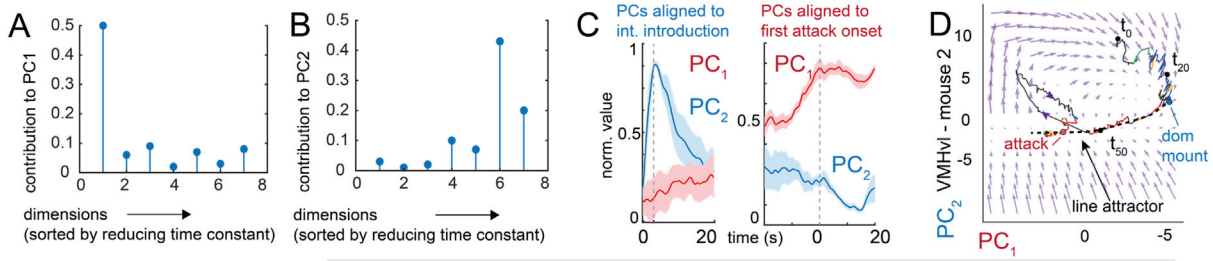
(F) Tracking features used in rSLDS shown alongside discovered states and integration dimension in VMHvl mouse 1.

(G) Performance of decoder used to separate attack frames from sniff-alone frames using the distance between mice and facing angle of the resident.

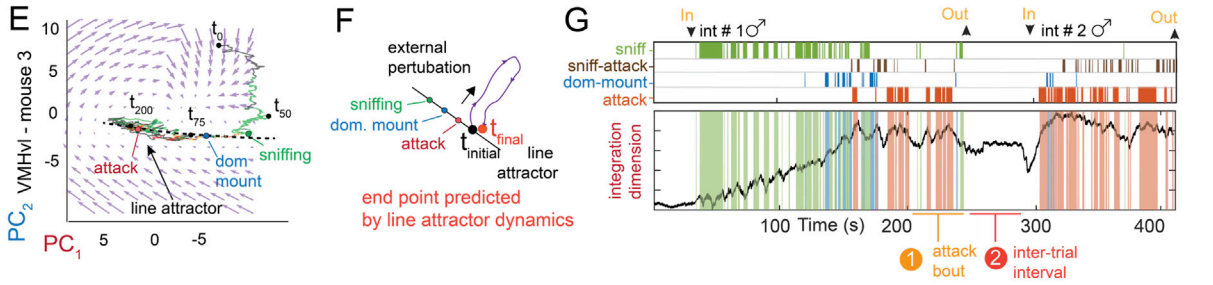
(H) Scatter plot of distance between mice and facing angle of resident.

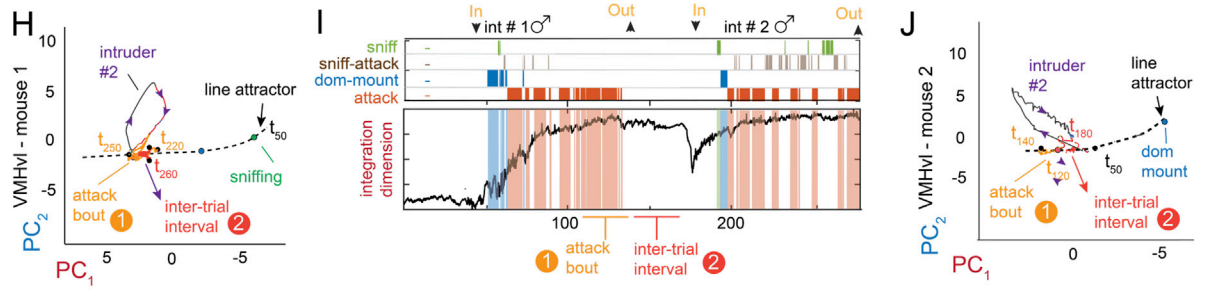
(I) Model performance (1-FSE) for different types of external inputs (n = 6 mice); current inputs = distance between animals, facing angle of resident (***p < 0.1).

relationship between integration dimension and low dimensional (PCA) representation of neural state space

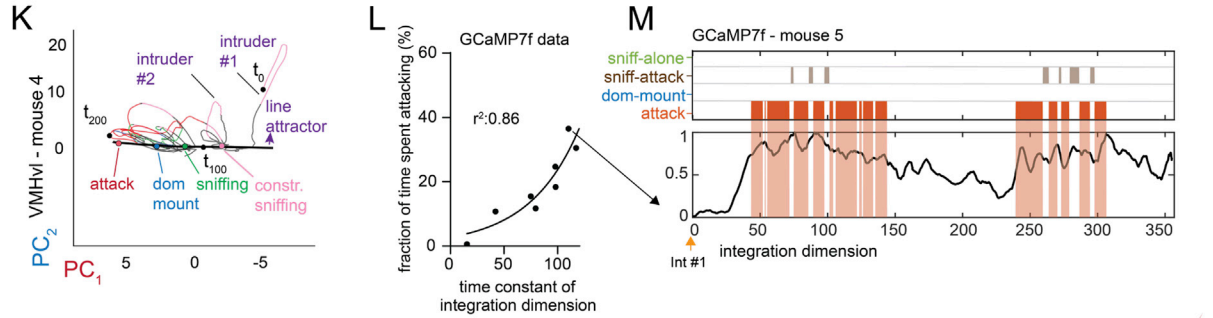


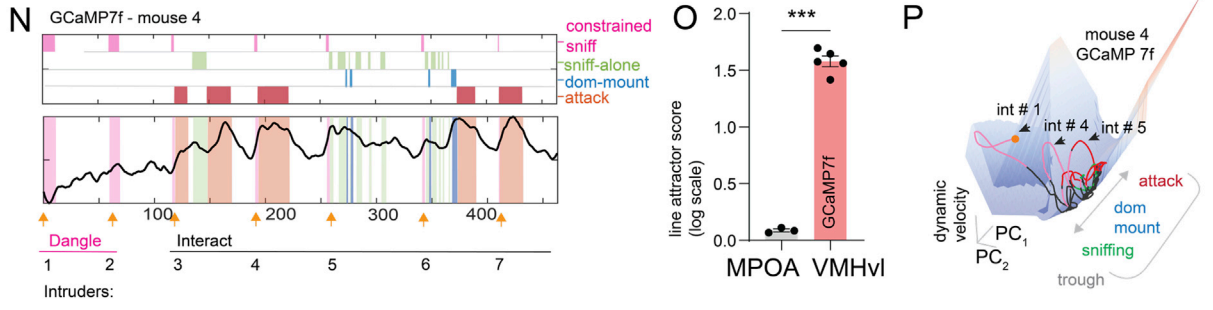
line attractor dynamics in VMHvl are robust to pertubations and show persistence





rSLDS discovers line attractor dynamics even with the faster indicator GCaMP7f





(legend on next page)

Figure S4. Properties of line attractor dynamics in VMHvl, related to Figure 3

(A and B) Absolute PCA weights of PC1 (A) and PC2 (B) on dimensions of dynamical system sorted by decreasing time constant in VMHvl mouse 1.

(C) Behavior-triggered average of top two principal components aligned to introduction of first intruder or first attack onset (n = 6 mice).

(D and E) Low dimensional dynamics and flow field showing line attractor dynamics for VMHvl mouse 2 and mouse 3 with line attractor highlighted.

(F) Schematic showing how perturbations orthogonal to a line attractor do not alter the position of the system.

(G) Integration dimension in VMHvl mouse 1 (reproduced from Figure 2B) with attack bout (1) and inter-trial interval (2) highlighted.

(H) Neural state space with line attractor highlighted in VMHvl mouse 1, showing the persistence of activity during the inter-trial interval shown in (G). The introduction of intruder #2 acts as an orthogonal perturbation and activity returns to the same point along the attractor.

(I and J) Same as (G) and (H) for VMHvl mouse 2.

(K) Neural state space with line attractor highlighted in VMHvl mouse 4. The introduction of intruder #2 occurs earlier in the trial when the animal displays sniffing behavior but results in a similar perturbation as above.

(L) Relationship between fraction of time spent attack versus time constant of integration for animals with GCaMP7f recordings (n = 8 mice).

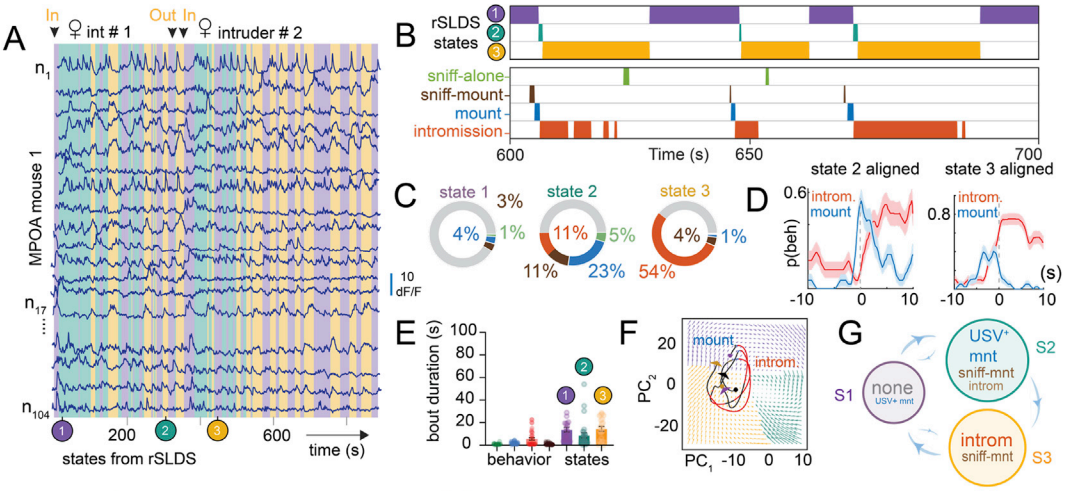
(M) Integration dimension in VMHvl mouse 5 (GCaMP 7f) shows the same persistence and slow decay of activity.

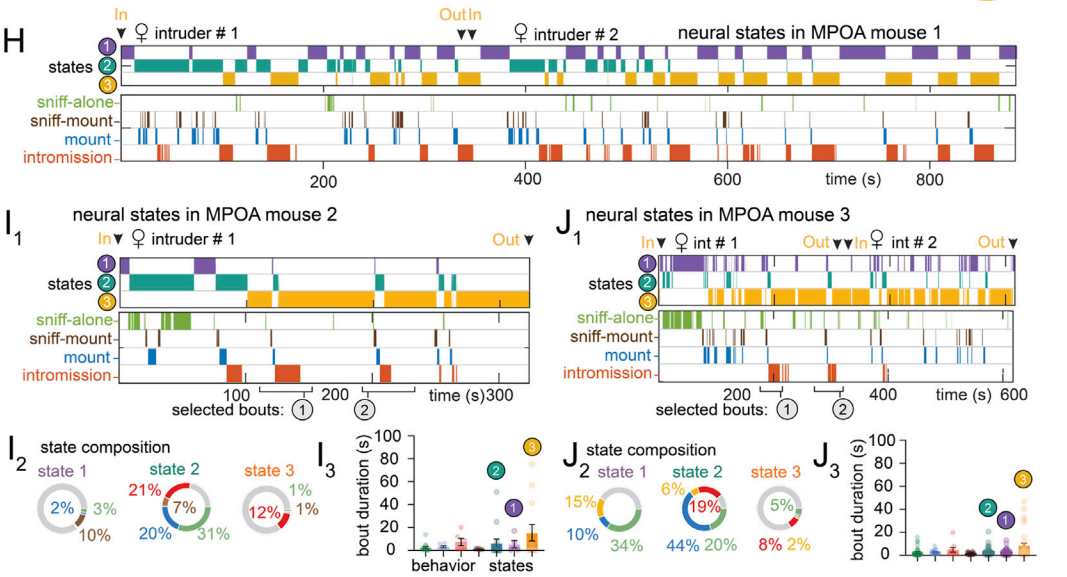
(N) Same as (M) for VMHvl mouse 4 (GCaMP 7f).

(O) Line attractor score for mice with GCaMP7f recordings (***$p < 0.001$).

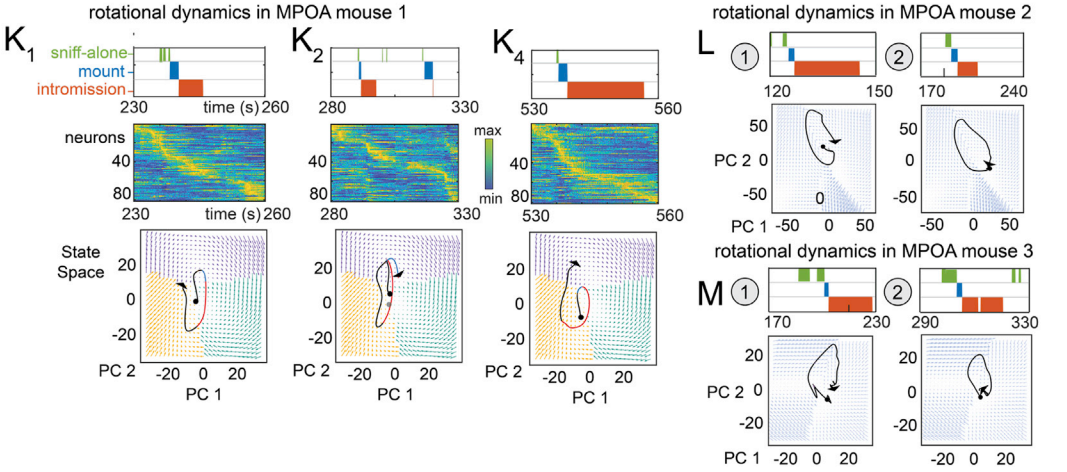
(P) Dynamics landscape for VMHvl mouse 4 (GCaMP 7f) showing a trough shaped landscape.

dynamical analysis of MPOA reveals mating related states in interactions with females





rSLDS discovers rotational dynamics in MPOA



(legend on next page)

Figure S5. Mating enriched states and rotational dynamics in MPOA, related to Figure 4

(A) rSLDS states in MPOA mouse 1.

(B) Comparison of rSLDS states with behavior in MPOA mouse 1 for period from t = 600 s to t = 700 s.

(C) Behavioral composition of rSLDS states.

(D) Probability of intromission and USV+ mounting aligned to the onset of state 2 and state 3 (also see I and J, n = 3 mice).

(E) Timescale of behavioral bouts and states epochs.

(F) Reproduced from Figure 4D but with state-specific inferred flow-field colors.

(G) State transition diagram from empirically calculated transition probabilities.

(H) State and behavior raster for MPOA mouse 1 for entire recording.

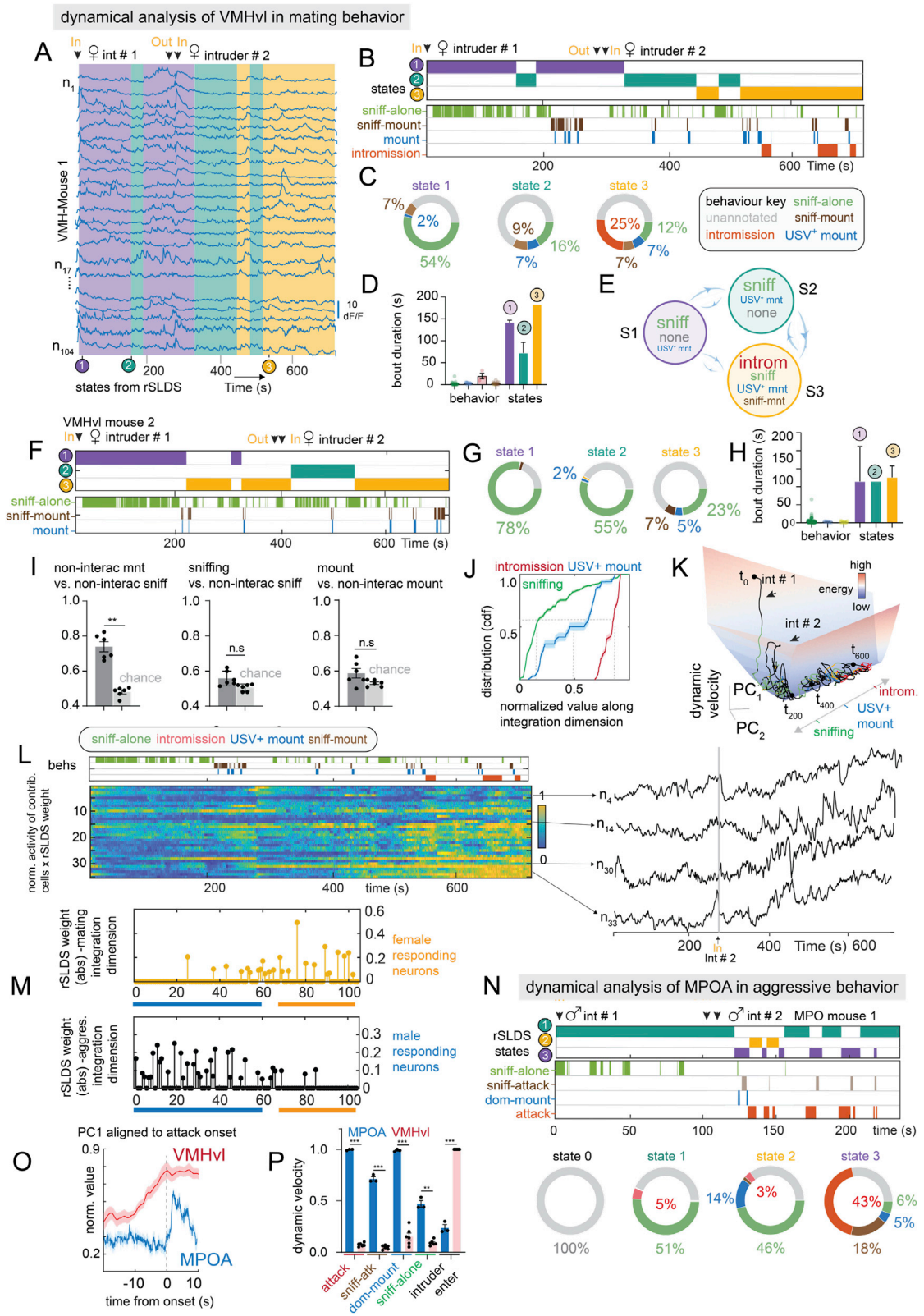
(I) (I_1) Same as (H) for MPOA mouse 2, selected mating bouts highlighted. (I_2) Behavioral composition of rSLDS states (bottom). (I_3) Timescale of behavioral bouts and states epochs.

(J) ($J_{1–3}$) Same as ($I_{1–3}$) for MPOA mouse 3, selected mating bouts highlighted.

(K) Rotational trajectories for 3 mating episodes in MPOA mouse 1.

(L) Same as (K), for mating bouts highlighted in highlighted in (I_1) for MPOA mouse 2.

(M) Same as (K), for mating bouts highlighted in highlighted in (J_1) for MPOA mouse 3.



(legend on next page)

Figure S6. Dynamical analysis of VMHvl activity in mating behavior and MPOA activity in aggression, related to Figure 6

(A) rSLDS states in VMHvl mouse 1 during interactions with female intruders.

(B) Comparison of rSLDS states with behaviors.

(C) Behavioral composition of rSLDS states. State 3 possesses the highest amount of mating behavior across mice (see H).

(D) Timescale of behavior bouts and state epochs.

(E) State transition diagram from empirical transition probabilities.

(F–H) Same as (B)–(D) for VMHvl mouse 2. This mouse did not achieve intromission.

(I) Decoding behaviors from integration dimension (**$p < 0.005$).

(J) Empirical cumulative distribution of value of integration dimension (normalized) for various behaviors.

(K) Dynamics velocity landscape showing a progression of mating behavior along the trough for VMHvl mouse 1.

(L) Normalized activity times rSLDS weight for cells contributing significantly to integration dimension of VMHvl mouse 1.

(M) Absolute rSLDS weight on integration dimension of VMHvl mouse 1 during mating behavior (top, yellow dots) and aggression (bottom, black dots) sorted by choice probability values for male versus female intruder encounter.

(N) Top: state and behavior raster for MPOA mouse 1 during aggressive behavior. State 3 is aligned closely to the onset of attack bouts, bottom: behavioral composition of discovered states.

(O) Behavior-triggered average of principal component 1 in VMHvl (red line) and MPO (blue line) (n = 3 mice for MPOA, n = 6 mice for VMHvl).

(P) Comparison of dynamic velocity for similar behavior between VMHvl and MPOA (reproduced from Figures 6F and 6K) (**$p < 0.005$,***$p < 0.001$) (n = 3 mice for MPOA, n = 6 mice for VMHvl).