Supporting Information

Hudson et al. 10.1073/pnas.1408296111

SI Materials and Methods

Surgery and Animal Care. All use of laboratory animals was consistent with *Guide for the Care and Use of Laboratory Animals* and approved by The Rockefeller University's Institutional Animal Care and Use Committee. Adult Sprague–Dawley rats (weight 250–350 g) were anesthetized with isoflurane in an anesthetizing chamber and secured in a stereotaxic frame with a nose cone for anesthetic delivery by using 2-L/min flows of 100% oxygen. Eyes were protected with ophthalmic ointment. The skull was fixed to the stereotaxic frame using a custom-designed head-holder device to ensure durable head restraint. Ear bars then were withdrawn. A craniotomy was made over the thalamus, centered on anterior–posterior −3.6 mm and medial–lateral 1.3 mm. The dura was removed and the surface of the brain irrigated with sterile saline. For analgesia, the wound edge was infiltrated with bupivacaine 0.125%, and flunixin 2.5 mg/kg was administered s.c. before lightening the anesthesia.

Spontaneous ventilation was maintained throughout the experiment. Respiratory and heart rates were monitored continuously with electrocardiography (ECG). ECG was monitored and streamed to disk for offline computation of heart rate and respiratory rate (derived from respiratory variation in QRS amplitude) to allow time-locked comparison with neuronal recordings. Temperature was maintained at 36 ± 0.5 °C by using a warming blanket controlled via a rectal temperature probe. Intraperitoneal saline injections were administered at regular intervals to maintain adequate hydration.

The animal was taken through a slow stepwise decrease in inhaled anesthetic concentration until awakening, defined as onset of spontaneous purposeful limb movement. Typically, we observed the animal attempting to escape the stereotaxic apparatus. This reliably occurred in all six animals included in this study at 0.75% isoflurane. Isoflurane was delivered using a calibrated isoflurane vaporizer, and the inhaled concentration was monitored (Riken Gas Indicator FI-21; A.M. Bickford).

At the conclusion of the experiment, the rats were anesthetized deeply and euthanized using high-dose isoflurane. Using standard methodology, the animals were intracardially perfused with PBS and paraformaldehyde (4%), followed by brain extraction and postfixation for at least 12 h. The brains were cryosectioned with a microtome (60-μm sections). Sections were stained for acetylcholinesterase to better define intralaminar nuclei and visualize electrode placement/scar.

Although we observed abrupt state transitions in more than 20 rats and 20 mice, here we included only rats that were maintained at each anesthetic concentration continuously for at least 1 h, had artifact-free recording in each of the channels of interest (see below), and were confirmed to have the correct electrode placement into the intralaminar thalamus as well as the cingulate and retrosplenial cortices ($n = 6$).

Dimensionality Reduction. Power spectra for five channels of local field potentials corresponding to the deep cingulate, superficial cingulate, deep retrosplenial, superficial retrosplenial, and centrolateral thalamus were concatenated into a 1,245-dimensional vector (249 independent frequency estimates per channel) at each spectral window. The spectrum in each channel was normalized by the total power and expressed as differences from mean spectrum determined over the entire anesthetic ramp for each animal after log transformation. The matrices so obtained for each of the six animals were then concatenated. Because principal component analysis (PCA) may be strongly affected by outliers (1), before dimensionality reduction, time points in which deviations exceeded 12 dB were removed. This artifact-removed matrix containing more than 30 h of recordings was subjected to PCA.

The first three principal components (PCs) contained ∼70% of the variance (Fig. 3*A*). The distribution of the data projected onto the first three PCs (shown in the plane PC1 vs. PC2 in Fig. 3*B*) clearly was multimodal.

Clustering. Data projected onto the first three PCs was clustered via *k*-means with squared Euclidean distance as the dissimilarity measure. For the clustering analysis, we minimized outlier effects by rejecting artifacts using a 10-dB threshold. To determine the optimal number of clusters, we varied the number of clusters from 2 to 10, starting from 100 random seeds. Clustering quality was assessed using mean of the silhouette value. Silhouette value was computed as $S_i = (b_i - a_i)/\max(a_i,b_i)$, where a_i is the average of the distance from *i*th point to all other points in the same cluster and b_i is the average of the distance from the *i*th point to points in all other clusters minimized over clusters (2) and implemented using the silhouette function in MATLAB. The peak silhouette occurred with eight clusters. The results are not affected strongly by the choice of the similarity measure, or by using a more liberal artifact rejection threshold of 12 dB.

One hundred spectrogram samples closest to the centroids for each cluster then were averaged to generate a representative spectral profile for that cluster for visual inspection (Fig. 4*A*).

Cluster Validation. The probability distribution shown in Fig. 3*B* of the main text shows the distribution of all data concatenated across six animals. To rule out the possibility that the peaks in this probability distribution reflect differences between animals rather than fluctuations within each animal, the probability distribution computed for each animal is shown in Fig. S5*A*. Although, as expected, there is some intersubject variability, all the individual probability distributions show multiple peaks, suggesting they are multimodal. Furthermore, inspection of the probability distributions suggests that the apparent peaks are best separated along the second PC.

To show the existence of multiple states statistically, we tested the null hypothesis that the data can be explained adequately by a unimodal distribution using the Hartigan dip test (3) with 1,000 bootstrap surrogates for each experiment. As this test is suited for univariate distributions, it was applied to the data projected onto PC2 from each experiment individually. This distribution, along with the *p* value for unimodality, is shown in Fig. S5*B*. The null hypothesis of unimodality was rejected ($p < 0.001$) in each experiment, which implies that in each experiment, the distribution of states is multimodal, i.e., contains multiple clusters.

Determining the optimal number of clusters in the data is an ill-posed problem, and a variety of tests and distance measures have been proposed. Here, we used the maximum silhouette value to partition the data into eight clusters. The results are not strongly dependent on the choice of the outlier threshold, distance measure, or cluster quality measure. Thus, in what follows, we assumed eight clusters.

Our method for measuring cluster consistency between experiments relies on the approach and formalism proposed in ref. 4.

Given a finite dataset $X = (x_1, x_2 \ldots, x_n)$, a clustering algorithm $\mathcal{A}$ assigns each point x_i a label (cluster index) $y_i \in \mathbb{L} := \{1, \ldots, k\}$ ($k = 8$ in this case), thus producing a vector of labels $Y := \mathcal{A}(X)$. If we assume that different experiments are realizations of the same probabilistic process, vectors Y_{temp} and Y_{test} produced by

clustering algorithms $\mathcal{A}_{temp}$ and $\mathcal{A}_{test}$ independently derived from subsets *temp* and *test* of the overall dataset should be similar.

To quantify the degree of similarity, we used the normalized Hamming distance defined as $d(Y_{temp},Y_{test})=\frac{1}{n}\sum_{i=1}^{n}1\{Y_{temp}(i)\neq Y_{test}(i)\}$, where $1\{Y_{temp}(i)\neq Y_{test}(i)\}=1$ *if* $Y_{temp}(i)\neq Y_{test}$ and 0 otherwise. This quantifies the fraction of all n points that were assigned to different clusters using algorithms $\mathcal{A}_{temp}$ and $\mathcal{A}_{test}$.

To determine the degree to which clustering is reproducible between experiments, one preparation was chosen as "template" and subjected to k-means clustering. Data from a different experimental preparation, "test," were assigned to clusters based on the proximity (squared Euclidian distance) to the centroids derived based on the template (algorithm $\mathcal{A}_{temp}$). This gives rise to the vector Y_{temp}. The test experiment then was independently subjected to the k-means algorithm, i.e., the centroids and cluster allocations were based solely on the data from the test experiment (algorithm $\mathcal{A}_{test}$). This gives rise to the vector Y_{test}.

Note that even in the perfect classification scheme, two label vectors correspond up to a permutation $\varsigma\in\mathfrak{G}_k$ of the indices (i.e., naming conventions), where $\mathfrak{G}_k$ is the set of all permutations of the elements in cluster labels $\mathbb{L}$. For example, cluster 1 in one label vector may be labeled cluster 2 in the other label vector, and vice versa, without any loss of consistency between classification schemes. Thus, we computed the minimum normalized Hamming distance between Y_{temp} across all permutations of labels in Y_{test} as follows: $d_{\mathfrak{G}_k}(Y_{temp},Y_{test}):=\min_{\varsigma\in\mathfrak{G}_k}\frac{1}{n}\sum_{i=1}^{n}1\{Y_{temp}(i)\neq\varsigma(Y_{test}(i))\}$.

If centroids in the template and test experiments are identical, then $d_{\mathfrak{G}_k}(Y_i,Y_j)$ will be zero. If there is no correspondence between the centroids obtained in the two cases, the distance between the two vectors will be maximized (see below). Thus, this is a useful measure to assess the consistency of clustering.

Using this algorithm, we used each subject as a template and computed the consistency between it and the remaining five subjects. The results shown in Fig. S5*C* show a 1-Hamming distance obtained in each case (i.e., the fraction of all data allocated to the same cluster). It can be seen that in each case, clustering is ∼60% consistent between subjects.

Distance between two label vectors strongly depends on the number of clusters. For instance, concordance of 50% will be equal to chance performance if there are only two clusters but would be highly improbable if the data were partitioned into 100 clusters. A random labeling algorithm that assigns a point x to one of k categories (clusters) with probability $1/k$ will asymptotically achieve the normalized Hamming distance of $1-1/k$ with Y_{temp}. This expected performance of a random labeling algorithm is shown by the black line in Fig. S5*C*. As the number of observations in each experiment was >10,000, this prediction was followed closely in 1,000 surrogate datasets produced by random assignments of the data to eight clusters. The highest similarity between template and randomly assigned surrogate never approached the lowest similarity observed between experiments shown in Fig. S5*C*. Thus, obtaining this degree of concordance between experiments cannot be explained plausibly by the null hypothesis of random cluster assignments.

One limitation of this null hypothesis, however, is that random assignment of the data into clusters may result in two data points with the same coordinates being assigned to different clusters. Furthermore, the data are not expected to be partitioned equally into eight clusters—some clusters have many more points than others. To overcome these limitations, we tested an additional null hypothesis that the position of the centroids obtained by k-means clustering of each experiment is random. For each pair of animals (i,j) where $i\neq j$, we constructed a set of 10 vectors Y_{test} that were assigned to eight clusters based on randomly chosen centroids (constrained to be within the range observed in the data for rat j) and computed the distribution of distances between the vector Y_{temp} and each Y_{test}, as described above. The experimentally observed (based on data in Fig. S5*C*) cumulative distribution of Hamming distances, along with jackknife confidence intervals (red, Fig. S6), and cumulative distribution of Hamming distances obtained using random centroids (blue, Fig. S6) are shown.

To determine the likelihood of obtaining the experimentally observed distribution of Hamming distances under the assumption of null hypothesis, we performed a Mann–Whitney U test ($P<0.001$). Thus, although there is an overlap between the distributions in Fig. S6, it is highly implausible that taken together, the experimentally observed Hamming distances can be obtained under the assumption of null hypothesis.

Hamming distance reflects the fraction of all points that were assigned inconsistently in the two label vectors. The absolute significance of a given difference in the Hamming distance between data and the null hypothesis therefore depends on the length of the dataset. To account for the differences in the total number of points in different experiments, for each pair of rats (i,j) we computed the distribution of the excess number of points consistently assigned to the same cluster when dataset-based centroids were used rather than when random centroids were used. This corresponds to the rat-pair–by–rat-pair difference between the experiment and the null hypothesis that takes into account the differences in the length of the dataset in each animal. If there were no difference between the data and the null hypothesis, the expected value of the difference in the number of consistently assigned points should be zero. The mean of the observed difference in the number of consistently assigned points was 1,012.5 ($P<0.0001$, two-tail t test). Thus, the concordance between experiments is not consistent with a null hypothesis.

Random Walk Simulation. To avoid assumptions concerning the distribution of steps, we computed pairwise differences between spectra. Each random walk was started from a randomly chosen, experimentally observed spectrum. At each time step, a random step was chosen from the distribution of pairwise differences between spectra. To prevent the random walker from exceeding the experimentally observed range, we imposed reflective boundary conditions such that the position of the simulation along each frequency (dimension) was contained in the range defined by the minimum and maximum of the experimentally observed data along this dimension.

Awakening, defined as the onset of spontaneous movement, predominantly occurred within a single cluster (red asterisk, Fig. 4). Thus, a random walker was considered awake when it was found inside this (goal) cluster. Note that in the experimental data, merely arriving into the cluster does not necessarily imply awakening. Many visits into the red asterisk cluster occur before awakening in Fig. 4, which biases the results toward success of the simulation.

The boundaries of the goal cluster were defined as a sphere centered at the cluster centroid with a radius equal to the 95th percentile of the distances of all cluster members to the centroid. The random walker was considered to be inside the cluster when its distance to the cluster centroid was less than this distance. If the simulation failed to converge on the cluster in 10^6 steps, it was aborted.

In the preliminary analysis, none of the simulations constructed on the basis of five concatenated channels (*Dimensionality Reduction*) converged. To further bias the results toward the simulation, we considered a single cortical channel. Furthermore, if the simulation of a single cortical channel was started from a point corresponding to burst suppression, it failed to converge. Thus, we simulated only the lowest anesthetic concentration (0.75% isoflurane), at which burst suppression rarely is observed, and started from a randomly chosen state observed during this concentration (assuming that it itself was not inside the goal

cluster). Pairwise difference spectra observed at this isoflurane concentration were used as steps.

The results obtained from 250 such simulations are shown in Fig. 1*D*. Only simulations that were randomly started close to the cluster managed to converge (note the steep increase in the cumulative probability, followed by a plateau). The red arrows show the experimentally observed times of awakening, defined as the time from the switch to 0.75% isoflurane to the time movement first was observed.

Failure of the simulations to converge despite multiple biases favoring convergence is not surprising. If the total space occupied by brain dynamics is considered a hyperrectangle whose size in each dimension is given by the minimum and maximum experimentally observed values and the goal cluster is considered a multidimensional sphere whose radius is given by the cluster boundary (defined above), then the fraction of the total space occupied by activity patterns consistent with awakening is $\sim 10^{-18}$. This simple calculation illustrates the improbability of recovery from anesthesia by random walk alone.

Network Analysis. Visual inspection of the network formed by the transition probabilities linking different metastable states suggests that two states (marked by blue and green asterisks in Fig. 4*C*) form hubs, i.e., states that link otherwise disconnected nodes. To show this quantitatively, we constructed an 8×8 adjacency matrix A such that

$$\mathrm{A}(i,j) = \begin{cases} 1, & p(i|j) \geq 0.003 \\ 0, & otherwise \end{cases}.$$

(The threshold is the same as in Fig. 4*C*.) Note that because $(i|j) \neq p(j|i)$, the graph is directed.

We then found the shortest path (i.e., the one requiring the smallest number of steps) starting from ith state and culminating in the awake state w (red asterisk in Fig. 4*C*), where $i \neq w$. Centrality of the jth node then was computed as a fraction of the shortest paths starting from each of the seven states $i \neq w$ and culminating in w that include state j. Because state i is guaranteed to be in all paths starting from state i, it was excluded from the calculations. The results of this calculation are shown in Fig. S9. Note that the centrality of the hub marked by the blue asterisk is less than that of the one marked by the green asterisk. This is because we considered all states, including those closer to the awake cluster than the hub marked by the blue asterisk. Also note that the centrality of all other nodes is zero.

Respiratory Rate Analysis. Respiratory rate was computed from the amplitude variation in the ECG. Briefly, R-wave peaks were identified in the raw ECG recordings. The time series consisting of R-wave amplitudes then was resampled by interpolation to assure a fixed sampling rate. The spectrum of this signal then was estimated with a standard multitaper approach using 60-s windows with a window step of 1 s. The peak of the spectrum in window centered at time t was used as the estimate of the respiratory rate at time t. This analysis was performed for each animal and each anesthetic concentration. To determine whether a statistically significant variation in the respiratory rate was present, we subjected these data to repeated measures ANOVA. As we were not interested in the differences in the absolute value of the respiratory rate observed in different experiments, the respiratory rate of each animal was normalized by the mean value observed for each anesthetic concentration for that animal, i.e., all data were expressed as deviations from mean. One hour of recording then was divided into 4 windows (~15 min long), and the respiratory rate was averaged within each window for each animal at each anesthetic concentration. Data from all animals and all concentrations then were submitted to repeated measures ANOVA ($\mathrm{df} = 19$, $F = 0.672$, $P = 0.830$). Thus, we cannot detect any significant deviation from the mean respiratory rate observed during fixed anesthetic exposure.

1. Thomson DJ (1982) Spectrum estimation and harmonic analysis. *Proc IEEE* 70(9): 1055–1096.
2. Kaufman L, Rouseeuw J (1990) *Finding Groups in Data: An Introduction to Cluster Analysis* (Wiley, New York).
3. Hartigan PM (1985) Algorithm AS 217: Computation of the dip statistic to test for unimodality. *J R Stat Soc Ser C Appl Stat* 34(3):320–325.
4. Lange T, Roth V, Braun ML, Buhmann JM (2004) Stability-based validation of clustering solutions. *Neural Comput* 16(6):1299–1323.

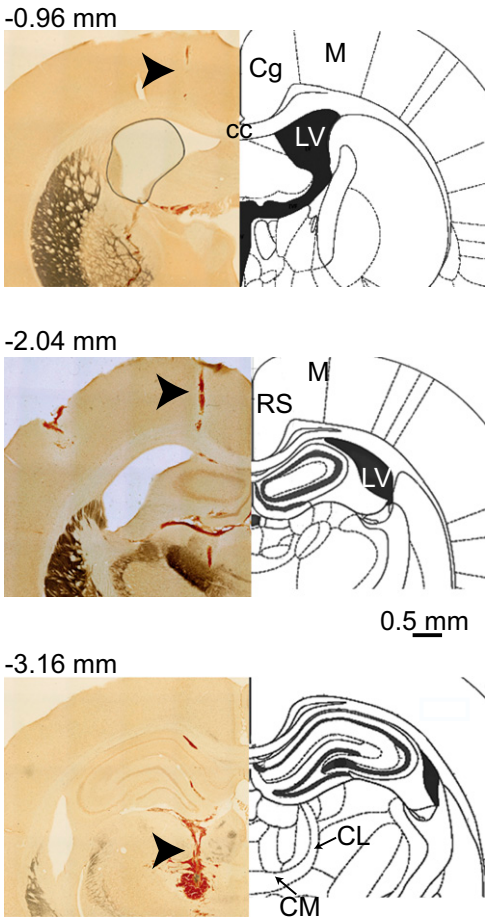


Fig. S1. Histology confirms correct electrode placement. All sections are from the same animal. Sections containing the scar from each of the prongs of the array are shown. The number above each section shows the approximate distance from the bregma along the anterior–posterior axis. Each section is juxtaposed with the corresponding diagram from the rat brain atlas. The electrode track is marked with a large arrowhead. All sections are shown on the same scale. cc, corpus callosum; Cg, cingulate cortex; CL, centrolateral nucleus of the thalamus; CM, centromedian nucleus of the thalamus; LV, lateral ventricle; M, motor cortex; RS, retrosplenial cortex.

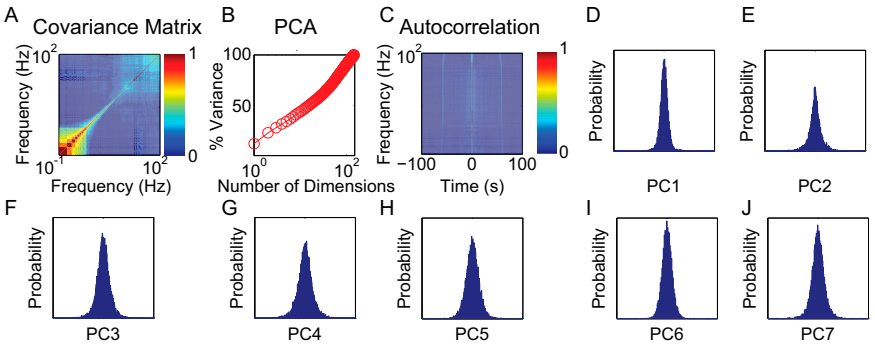


Fig. S2. Fluctuations in neuronal activity on a fast time scale are multidimensional noise. (*A*) Covariance matrix of fluctuations in power computed for a single cortical channel. Although some correlations exist, they are sparse and weak. Note that because in the random walk simulation we sample the distribution of these fluctuations at each time step, we preserve whatever weak covariance structure is present in the data. (*B*) Consistent with the results in *A*, PCA fails to reduce the dimensionality of the fluctuations. It takes >50 dimensions to capture 70% of the total variance (contrast this with Fig. 2*C*). (*C*) The autocorrelation function computed for fluctuations in each frequency is close to zero. Thus, randomly sampling the distribution of steps for the random walk simulation is justified. (*D–I*) Distribution of the fluctuations projected onto the first seven PCs demonstrates that fluctuations, to a good approximation, are noise distributed among many dimensions.

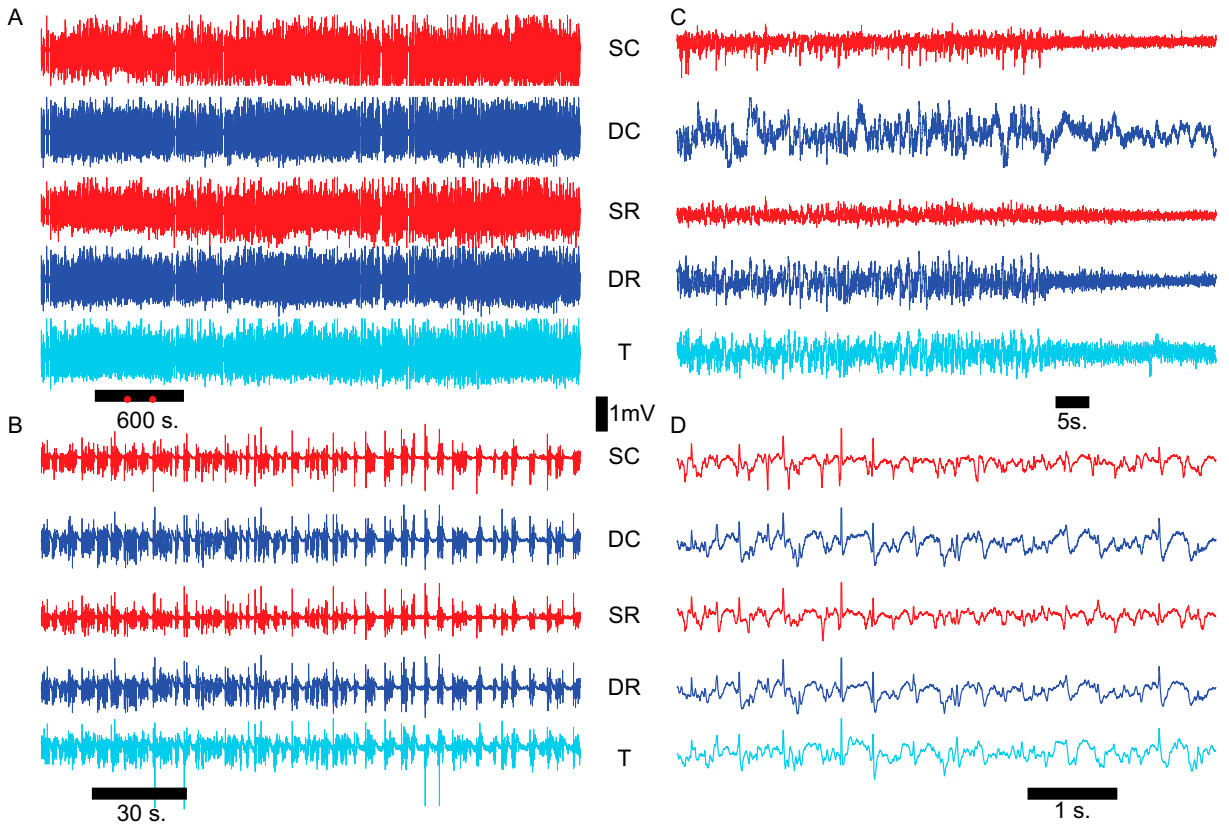


Fig. S3. Examples of abrupt state transitions. All traces are from the same experimental preparation. (*A*) The whole hour of recording obtained at 1.75% isoflurane. Although burst suppression is exhibited throughout much of the recording, the different modes characterizing bursts and periods of quiescence fluctuate among several distinct states. (*B*) Expansion of traces in *A* in the range denoted by two red dots on the time scale bar in *A*. (*C*) Transition to γ-range oscillations simultaneously observed in multiple channels at 1.0% isoflurane. (*D*) Switch from δ-dominant to θ-dominant activity at 1.5% isoflurane. DC, deep cingulate; DR, deep retrosplenial; SC, superficial cingulate; SR, superficial retrosplenial; T, centrolateral thalamus.

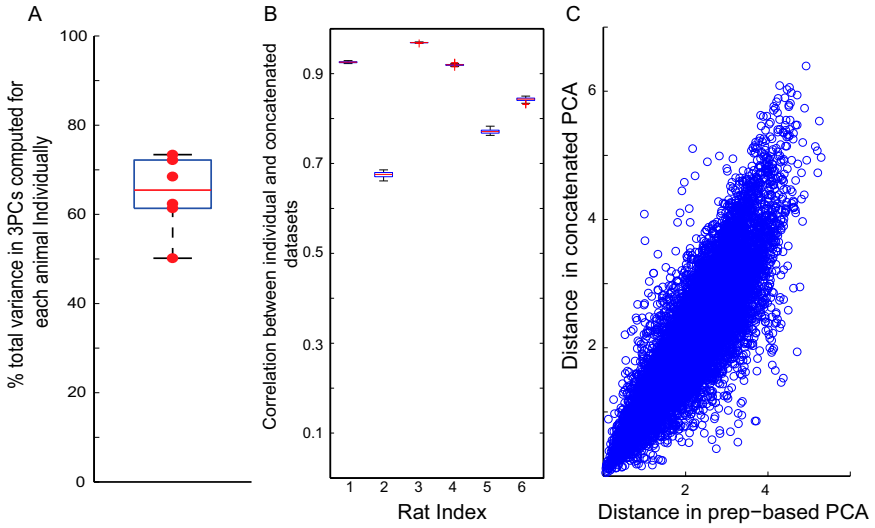


Fig. S4. Concatenation does not strongly influence the structure of the data from each animal. (*A*) Data from each animal were independently subjected to PCA. The fraction of total variance captured in the first three PCs is represented as a box plot. In addition, data from each individual rat is shown as a red ●. With the exception of one outlier, the dimensionality reduction obtained in each rat is similar to that obtained in the concatenated dataset (Fig. 2). (*B*) Standardized Euclidean distance (each coordinate is expressed in the units of SD of the data along each dimension) was computed between all pairs of points from each rat projected in two different ways: (*i*) onto the first three PCs derived solely based on data from that rat and (*ii*) onto the first three PCs derived based on the concatenated dataset. The Pearson correlation coefficient between the two distance matrices was computed using 100 bootstrap replicas, each consisting of 10,000 pairwise distance measures. The distribution of estimates is shown as a box plot for each rat. (*C*) Scatterplot showing that distances in prep-based PCA are similar to those observed when the data are projected onto PCs derived on the basis of the concatenated dataset.

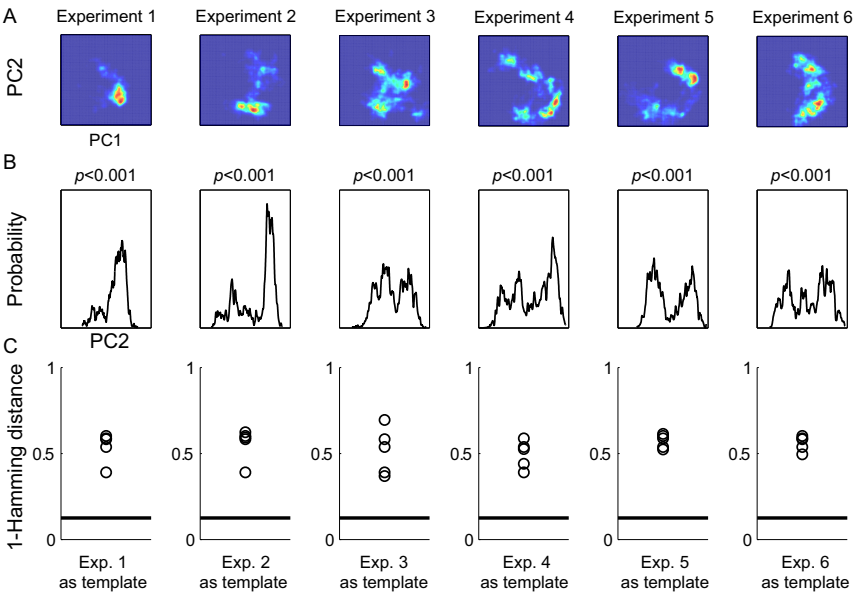


Fig. S5. Clustering is consistent among experiments. (*A*) Projection of the probability distribution onto the first two PCs for each animal individually. Note that in each case, multiple distinct peaks can be observed. (*B*) Projection of the probability distribution of the data in each experiment onto PC2 is multimodal in every animal. Above each plot is the *P* value for unimodality computed using the Hartigan dip test (*SI Materials and Methods*). (*C*) Concordance between clusters shown as the fraction of all data allocated to the same cluster when each experiment was used as a template in turn (*SI Materials and Methods*). Each ○ indicates the degree of concordance between the template and each of the five remaining experiments. The black line marks the chance probability of concordance. The observed degree of concordance is statistically significant (*SI Materials and Methods*) when tested against both random cluster assignments and random centroid locations (distribution of Hamming distances obtained with experimentally derived and random centroids shown in Fig. S6).

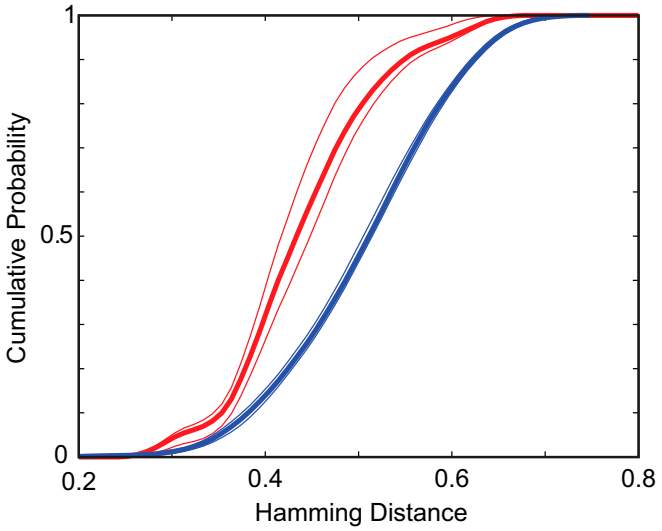


Fig. S6. Distribution of Hamming distances obtained based on the data-derived and randomly chosen centroids. For each pair of animals (*i,j*) where $I \neq j$, we computed the Hamming distance between cluster assignments of rat *j* clustered independently and using centroids derived on the basis of data from rat *i* (template). This gives rise to 30 distance measures. Cumulative distribution is shown by the thick red line. Jackknife confidence intervals (0.05 and 0.95) were computed by leaving out each of the six rats used as a template in turn and recomputing the probability distribution (thin red lines). For each pair (*i,j*), we then used 10 sets of random centroids (constrained to be within range of data in rat *j*) and computed the Hamming distances between rat *j* clustered to the random centroids and those computed on the basis of rat *i* (total of 300 distances). Cumulative distribution of these is shown by the thick blue line. Jackknife confidence intervals (0.05 and 0.95) were computed as above (thin blue line). $P < 0.001$ (Mann–Whitney *U* test). Thus, the concordance between experiments is statistically significant. As different experiments have a different number of points, we computed the distribution of the excess number of points that were allocated consistently into clusters when data-based centroids were used for each pair of rats (*i,j*). We then tested this distribution against the null hypothesis that the excess number of consistently clustered points is zero (mean 1,012.5, $P < 0.0001$, two-tail *t* test). Thus, on average, 1,012.5 more points were allocated to clusters correctly when centroids from another experiment were used rather than randomly chosen centroids.

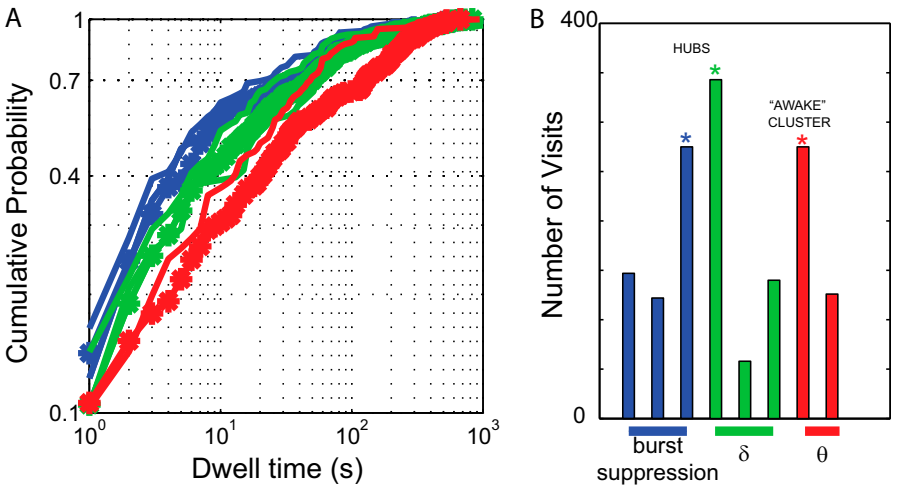


Fig. S7. Cumulative dwell time distributions for each of the eight clusters and the number of distinct visits to each cluster. (*A*) Although some dwell times are short, many extend for several minutes. All distributions are right skewed. Line colors are the same as those in Figs. 4 and 3*C* of the main text. Blue and green asterisks correspond to the hub states; the red asterisk corresponds to the cluster in which awakening most commonly was observed. (*B*) Hub states (blue and green asterisks) are visited more frequently than nonhub states.

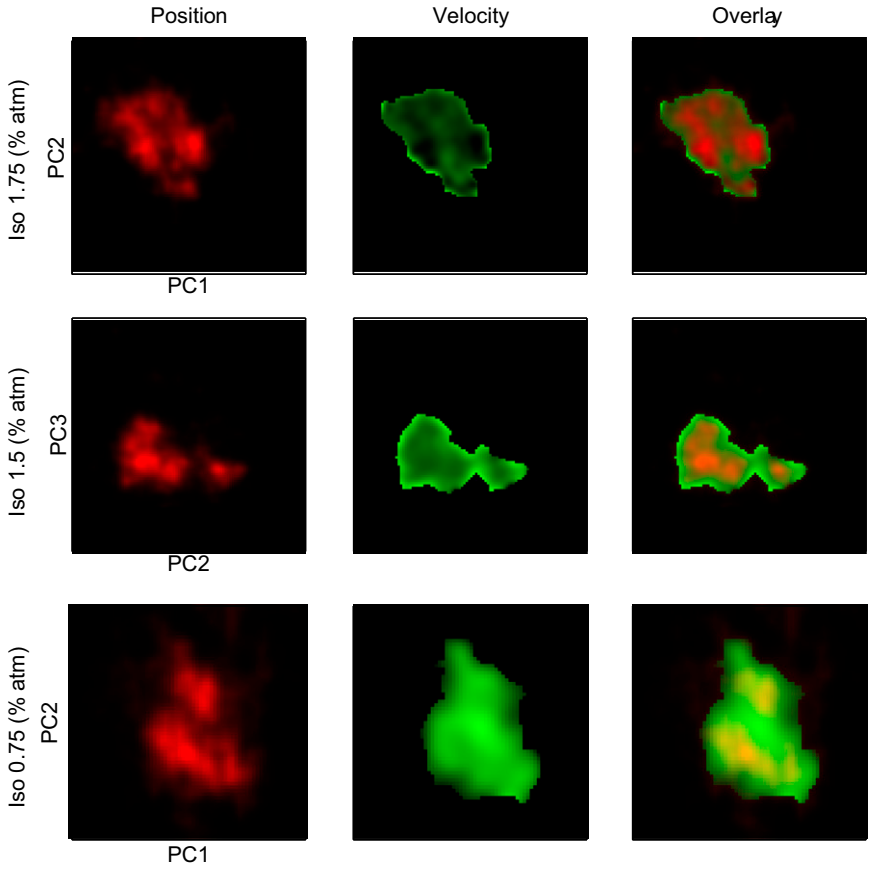


Fig. S8. Highly occupied states are stabilized. (*Position*) Distribution of data in the plane that allows for best visualization of the separation between clusters is shown for three different anesthetic concentrations. The histogram count in each bin is normalized to the range between 0 (not occupied) and 1 (most occupied) and encoded as red in the RGB color scheme. (*Velocity*) Mean fluctuation amplitude as a function of position was computed as $\bar{v}=\frac{1}{n}\sum_{i=1}^{n} d^2(x_{i+1}, x_i)$, where $d^2(\cdot)$ is the squared Euclidean distance between two positions separated by 1 s and $i\in\{1,..,n\}$ are indices of points that lie in a particular bin in the position histogram. This quantity similarly was normalized to lie between 0 and 1 and encoded in green in the RGB color scheme. (*Overlay*) Superposition of *Position* and *Velocity* data is shown. This image is constructed such that if both occupancy and velocity are high, the color appears yellow. Pure red indicates locations with high occupancy and low velocity. Pure green corresponds to high velocity and low occupancy. Note the rim of green (large fluctuations) that surrounds areas of red (large occupancy). Thus, the densely occupied locations are stabilized, whereas sparsely occupied regions between them are relatively unstable.

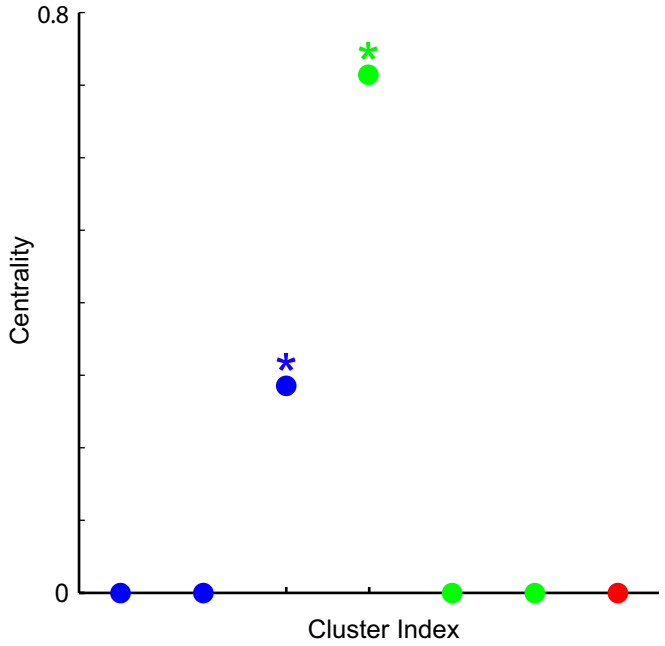


Fig. S9. Analysis of centrality confirms definition of hub states. The centrality of each state was computed as described in *SI Materials and Methods*. Clusters are color coded in the same way as in Fig. 4. Green and blue asterisks indicate hubs shown in Fig. 4. Note that the centrality of all other nodes in the network is zero.

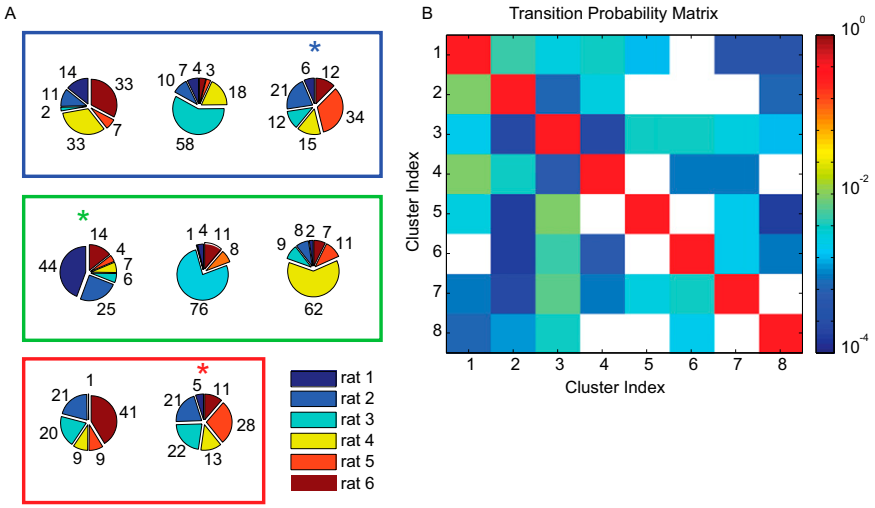
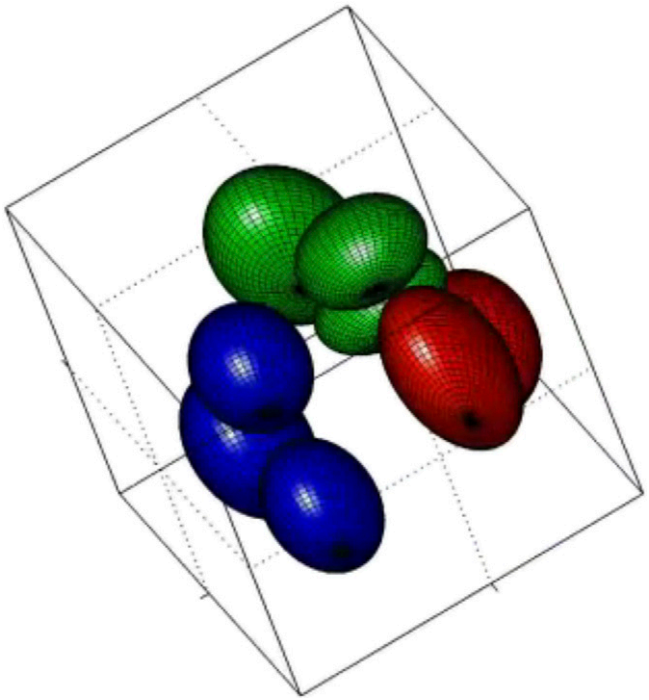
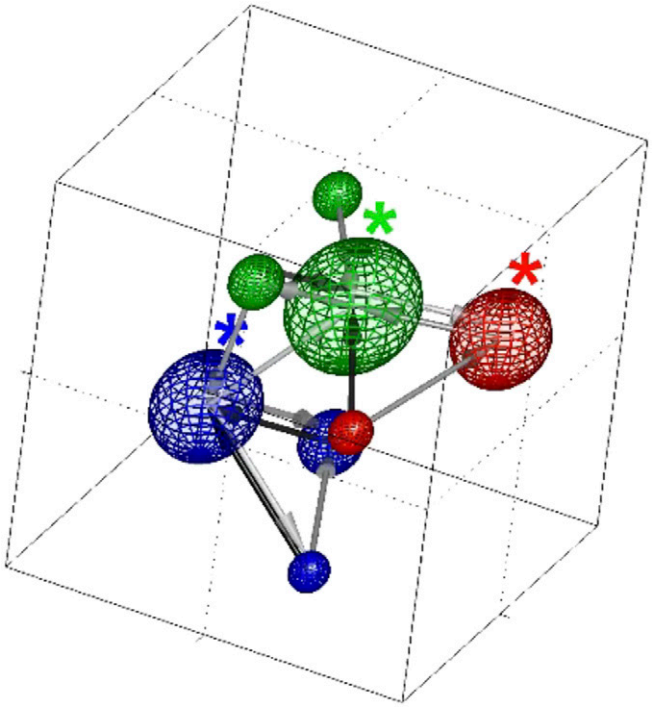


Fig. S10. (*A*) Distribution of points assigned to each cluster among different animals. Each circle represents a cluster (grouped and ordered in the same fashion as in Fig. 4). The area of the circle occupied by each segment is proportional to the percentage (shown around the perimeter) of all data allocated to that cluster from each experiment. Seven of 8 clusters were visited by each of the six rats, whereas one cluster was visited by five of six rats. Data from all six animals contribute to all the hubs (shown by asterisks), whereas some nonhub states appear to be dominated by data from a single animal. Note that different experiments have a different number of points. This measure is not normalized to reflect this difference. (*B*) Transition probability matrix computed on the basis of all animals. Note that the probability of staying in the cluster (diagonal elements) is approximately two orders of magnitude higher than the probability of switching between clusters. Thus, a relatively small difference in the number of unique visits to a cluster may result in a larger difference in the total number of points in the cluster. This is especially likely for the less frequently visited nonhub states.



Movie S1. Rotation of the data shown in Fig. 3*C* to improve visualization of the location of the activity clusters in the space spanned by the first three PCs.

Movie S1



Movie S2. Rotation of the data shown in Fig. 4*C* to improve visualization of the network structure.

Movie S2