Sensory stimuli can be recognized more rapidly when they are expected. This phenomenon depends on expectation affecting the cortical processing of sensory information. However, the mechanisms responsible for the effects of expectation on sensory circuits remain elusive. In the present study, we report a novel computational mechanism underlying the expectation-dependent acceleration of coding observed in the gustatory cortex of alert rats. We use a recurrent spiking network model with a clustered architecture capturing essential features of cortical activity, such as its intrinsically generated metastable dynamics. Relying on network theory and computer simulations, we propose that expectation exerts its function by modulating the intrinsically generated dynamics preceding taste delivery. Our model's predictions were confirmed in the experimental data, demonstrating how the modulation of ongoing activity can shape sensory coding. Altogether, these results provide a biologically plausible theory of expectation and ascribe an alternative functional role to intrinsically generated, metastable activity.

In the present study, we propose a mechanism whereby an anticipatory cue modulates the timescale of temporal dynamics in a recurrent population model of spiking neurons. In our model, neurons are organized in strongly connected clusters and produce sequences of metastable states similar to those observed during both pre-stimulus and evoked activity periods. A metastable state is a vector across simultaneously recorded neurons, which can last for several hundred milliseconds before giving way to the next state in a sequence. The ubiquitous presence of state sequences in many cortical areas and behavioral contexts has raised the issue of their role in sensory and cognitive processing. In the present study, we explain the central role played by pre-stimulus metastable states in processing forthcoming stimuli, and show how cue-induced modulations drive anticipatory coding. Specifically, we show that an anticipatory cue affects sensory coding by decreasing the duration of metastable states and accelerating the pace of state sequences. This phenomenon, which results from a reduction in the effective energy barriers separating the metastable states, accelerates the onset of specific states coding for the presented stimulus, thus mediating the effects of general expectation. The predictions of our model were confirmed in an analysis of the experimental data, also reported here.

Altogether, our results provide a model for general expectation, based on the modulation of pre-stimulus ongoing cortical dynamics by anticipatory cues, leading to acceleration of sensory coding.

Results
Anticipatory cue accelerates stimulus coding in a clustered population of neurons. To uncover the computational mechanism linking cue-evoked activity with coding speed, we modeled the gustatory cortex as a population of recurrently connected excitatory and inhibitory spiking neurons. In this model, excitatory neurons are arranged in clusters (Fig. 1a), reflecting the existence of assemblies of functionally correlated neurons in the gustatory cortex and other cortical areas. Recurrent synaptic weights between neurons in the same cluster are potentiated compared with neurons in different clusters, to account for metastability in the gustatory cortex and other cortical areas. Recurrent synaptic weights between neurons in the same cluster are potentiated compared with neurons in different clusters, to account for metastability in the gustatory cortex and other cortical areas. This spiking network also has bidirectional random and homogeneous (that is, non-clustered) connections among inhibitory neurons and between inhibitory and excitatory neurons. Such connections stabilize network activity by preventing runaway excitation and play a role in inducing the observed metastability.

The model was probed by sensory inputs modeled as depolarizing currents injected into randomly selected neurons. We used four sets of simulated stimuli, wired to produce gustatory responses
Fig. 1 | Anticipatory activity requires a clustered network architecture. Effects of anticipatory cue on stimulus coding in the clustered (a–c) and homogeneous (d–f) network. a, Schematics of the clustered network architecture and stimulation framework. A recurrent network of inhibitory (red circles) and excitatory (triangles) neurons arranged in clusters (ellipsoids) with stronger intracluster recurrent connectivity. The network receives bottom-up sensory stimuli targeting random, overlapping subsets of clusters (selectivity to four stimuli is color-coded), and one top-down anticipatory cue inducing a spatial variance in the cue afferent currents to excitatory neurons. b, Representative single-neuron responses to cue and one stimulus in expected trials in the clustered network of a. Black tick marks represent spike times (rasters), with a peristimulus time histogram (mean ± s.e.m.) overlaid in pink. Activity marked by horizontal bars was significantly different from baseline (pre-cue activity) and could be either excited (top panel) or inhibited (bottom) by the cue. Spks, spikes. c, Time course of cross-validated, stimulus-decoding accuracy in the clustered network. Decoding accuracy increases faster during expected (pink) than unexpected (blue) trials in clustered networks (curves and color-shaded areas represent mean ± s.e.m. across four tastes in n = 20 simulated networks; color-dotted lines around gray-shaded areas represent the 95% confidence interval from shuffled datasets). A separate classifier was used for each time bin, and decoding accuracy was assessed via a cross-validation procedure, yielding a confusion matrix, with a diagonal representing the accuracy of classification for each of four tastes in that time bin (see text and Supplementary Fig. 3 for details). Inset: aggregate analysis across n = 20 simulated networks of the onset times of significant decoding (mean ± s.e.m.) in expected (pink) versus unexpected (blue) trials shows significantly faster onsets in the expected condition (two-sided t-test, P = 0.0017). d, Schematics of the homogeneous network architecture. Sensory stimuli modeled as in a, e, Representative single-neuron responses to cue and one stimulus in expected trials in the homogeneous network of d (same conventions as in b), f, Cross-validated decoding accuracy in the homogeneous network (same analysis as in c). The latency of significant decoding in expected versus unexpected trials is not significantly different. Inset: aggregate analysis of onset times of significant decoding (same as inset of e; two-sided t-test, P = 0.31). b, c, e, f, pink and black horizontal bars, P < 0.05, two-sided t-test with multiple-bin Bonferroni correction. c, **P < 0.01, two-sided t-test. f, NS, non-significant.

reminiscent of those observed in the experiments in the presence of sucrose, sodium chloride, citric acid, and quinine (see Methods). The specific connectivity pattern used was inferred by the presence of both broadly and narrowly tuned responses in the gustatory cortex25,26, and the temporal dynamics of the inputs were varied to determine the robustness of the model (Fig. 2).

In addition to input gustatory stimuli, we included anticipatory inputs designed to produce cue responses analogous to those seen experimentally in the case of general expectation. To simulate general expectation, we connected anticipatory inputs with random neuronal targets in the network. The peak value of the cue-induced current for each neuron was sampled from a normal distribution with zero mean and fixed variance, thus introducing a spatial variance in the afferent currents. This choice reflected the large heterogeneity of cue responses observed in the empirical data, where excited and inhibited neural responses occurred in similar proportions and overlapped partially with taste responses⁵. Figure 1b shows two representative cue-responsive neurons in the model: one inhibited by the cue and one excited by the cue. Cue responses in the model were in quantitative agreement with the observed responses for a large range of cue-induced spatial variance (Supplementary Fig. 1; see Supplementary Fig. 2 for representative unresponsive neurons).

Given these conditions, we simulated the experimental paradigm adopted in awake-behaving rats to demonstrate the effects of general expectation⁶. In the original experiment, rats were trained to self-administer into an intraoral cannula one of four possible tastants following an anticipatory cue. At random trials and time during the intertrial interval, tastants were unexpectedly delivered in the absence of a cue. To match this experiment, the simulated
paradigm interleaves two conditions: in expected trials, a stimulus (out of four) is delivered at \( t = 0 \) after an anticipatory cue (the same for all stimuli) delivered at \( t = -0.5 \) s (Fig. 1b); in unexpected trials the same stimuli are presented in the absence of the cue. Importantly, in the general expectation paradigm adopted here, the anticipatory cue is identical for all stimuli in the expected condition. Therefore, it does not convey any information about the identity of the stimulus being delivered.

We tested whether cue presentation affected stimulus coding. A multiclass classifier (see Methods and Supplementary Fig. 3) was used to assess the information about the stimulus encoded in the neural activity, in which the four class labels correspond to the four tastants. Stimulus identity was encoded well both in conditions, reaching perfect accuracy for all four tastants after a few hundred milliseconds (Fig. 1c). However, comparing the time course of the decoding accuracy between conditions, we found that the increase in decoding accuracy was significantly faster in expected than in unexpected trials (Fig. 1c, pink and blue curves represent expected and unexpected conditions, respectively). Indeed, the onset time of a significant decoding occurred earlier in the expected versus the unexpected condition (decoding latency was 0.13 ± 0.01 s (mean ± s.e.m.) for expected compared with 0.21 ± 0.02 s for unexpected, across 20 independent networks; \( P = 0.002 \), two-sided \( t \)-test, degrees of freedom = 39; inset in Fig. 1c). These measures refer to the decoding accuracy averaged across all tastants; similar results were obtained for each individual tastant separately (see Supplementary Fig. 3b). Thus, in the model network, the interaction of cue response and activity evoked by the stimuli results in faster encoding of the stimuli themselves, mediating the expectation effect.

To clarify the role of neural clusters in mediating expectation, we simulated the same experiments in a homogeneous network (that is, without clusters) operating in the balanced asynchronous regime\(^{23,29}\) (Fig. 1d, intra- and intercluster weights were set equal; all other network parameters and inputs were the same as for the clustered network). Even though single neuron responses to the anticipatory cue are identical for all stimuli in the expected condition, the decoding accuracy averaged across all tastants; similar results were obtained for each individual tastant separately (see Supplementary Fig. 3b). Thus, in the model network, the interaction of cue response and activity evoked by the stimuli results in faster encoding of the stimuli themselves, mediating the expectation effect.

Fig. 2 | Robustness of anticipatory activity to variations in stimulus and cue models. a, Latency of significant decoding increased with stimulus intensity (left: top, stimulus peak expressed as percentage of baseline, darker shades represent stronger stimuli; bottom, decoding latency, mean ± s.e.m. across \( n = 20 \) simulated networks for each value on the x axis; see main text for a–c statistical tests) in both conditions, and it is faster in expected (pink) than in unexpected (blue) trials. Anticipatory activity was present for a large range of network sizes (right: \( J = 5, 10, 20, 30, 40 \) for \( N = 1, 2, 4, 6, 8 \times 10^3 \) neurons, respectively). Network synaptic weights scaled as reported in Tables 1 and 2. b, Anticipatory activity was present when stimuli targeted both excitatory (E) and inhibitory (I) neurons (notations as in Fig. 1c; 50% of both E and I neurons were targeted by the cue; inset: mean ± s.e.m. across \( n = 20 \) simulated networks, two-sided \( t \)-test, \( P = 0.0011 \)). c, Increasing the cue-induced spatial variance in the afferent currents \( \sigma^2 \) (top left: histogram of afferents’ peak values across neurons; \( x \) axis, expressed as percentage of baseline; \( y \) axis, calibration bar: 100 neurons) leads to more pronounced anticipatory activity (bottom left: latency in unexpected (blue) and expected (pink) trials). Anticipatory activity was present for a large range of cue time courses (top right: double exponential cue profile with rise and decay times \( [\tau_r,\tau_d] = (0.1,0.5) \) s, for \( g \) in the range from 1 to 3; bottom right: decoding latency during unexpected (blue) and expected (pink) trials). d, Anticipatory activity was also present when the cue targeted 50% of both E and I neurons (\( \sigma = 20\% \) in baseline units; inset: mean ± s.e.m. across \( n = 20 \) simulated networks, two-sided \( t \)-test, \( P = 0.0034 \)). a–d, \( * P < 0.05, ** P < 0.01, *** P < 0.001, \) post hoc, multiple-comparison, two-sided \( t \)-test with Bonferroni correction. Insets: \( ** P < 0.01, *** P < 0.001, \) two-sided \( t \)-test. c,d, notations as in Fig. 1c.
A defining feature of our model is that it incorporates excited and inhibited cue responses in such a manner as to affect only the spatial variance of the activity across neurons, while leaving the mean input to the network unaffected. As a result, the anticipatory cue leaves average firing rates unchanged in the clustered network (see Supplementary Fig. 4), and only modulates the network temporal dynamics. Our model thus provides a mechanism whereby increasing the spatial variance of top-down inputs has, paradoxically, a beneficial effect on sensory coding.

Robustness of anticipatory activity. To test the robustness of anticipatory activity in the clustered network, we systematically varied key parameters related to the sensory and anticipatory inputs, as well as network connectivity and architecture (Fig. 2 and see Supplementary Fig. 5). First we investigated variation in stimulus features. Increasing stimulus intensity led to a faster encoding of the stimulus in both like model), to further investigate a potential role (if any) of the cue’s spatial coordinates. Given this type of network. However, in a model where the cue recruited inhibitory neurons, in addition to excitatory neurons, one may be tempted to consider the effect of inhibitory neurons. We tested whether or not anticipation was present if inhibitory neurons were included in the network (see Supplementary Figs. 4 and 5b). Two-way ANOVA with factors ‘cue variance’ (P = 0.09, F(7) = 1.8) and ‘condition’ (P < 0.01, F(1) = 90; P(interaction) = 0.07, F(7) = 1.9)). Anticipatory activity was also robust to variations in the time course of the cue-evoked currents (Fig. 2c, two-way ANOVA with factors ‘time course’ (P = 0.03, F(4) = 2.8) and ‘condition’ (P < 0.01, F(1) = 72.7; P(interaction) = 0.12, F(4) = 1.8)). We also considered a model with constant cue-evoked currents (step-like model), to further investigate a potential role (if any) of the cue’s variable time course on anticipatory activity, and found anticipatory coding in this case (see Supplementary Fig. 5c).

Finally, we tested whether or not anticipation was present if inhibitory neurons were included in the network (see Supplementary Figs. 4 and 5b). Two-way ANOVA with factors ‘cue variance’ (P = 0.09, F(7) = 1.8) and ‘condition’ (P < 0.01, F(1) = 90; P(interaction) = 0.07, F(7) = 1.9)). Anticipatory activity was also robust to variations in the time course of the cue-evoked currents (Fig. 2c, two-way ANOVA with factors ‘time course’ (P = 0.03, F(4) = 2.8) and ‘condition’ (P < 0.01, F(1) = 72.7; P(interaction) = 0.12, F(4) = 1.8)). We also considered a model with constant cue-evoked currents (step-like model), to further investigate a potential role (if any) of the cue’s variable time course on anticipatory activity, and found anticipatory coding in this case (see Supplementary Fig. 5c).

We then performed variations in network size and architecture. We estimated the decoding accuracy in ensembles of constant size (20 neurons) sampled from networks of increasing size (Fig. 2a), fixing both cluster size and the probability that each cluster was selective to a given stimulus (50%). Cue-induced anticipation was even more pronounced in larger networks (Fig. 2a, two-way ANOVA with factors ‘network size’ (P < 0.01, F(3) = 49.5) and ‘condition’ (P < 0.01, F(1) = 90; P(interaction) = 0.01, F(3) = 20)). In our main model the neural clusters were segregated in disjoint groups. We investigated an alternative scenario where inhibitory neurons may belong to multiple clusters, resulting in an architecture with overlapping clusters. We found that anticipatory activity was also present in networks with overlapping clusters (see Supplementary Fig. 5d).

Next, we assessed robustness to variations in cue parameters — specifically: variations in the spatial variance $\sigma^2$ of the cue-induced afferent currents; variations in the kinetics of cue stimulation; and variations in the type of neurons targeted by the cue. We found that coding anticipation was present for all values of $\sigma^2$ above 10% (Fig. 2c, two-way ANOVA with factors ‘cue variance’ (P = 0.09, F(7) = 1.8) and ‘condition’ (P < 0.01, F(1) = 90; P(interaction) = 0.07, F(7) = 1.9)). Anticipatory activity was also robust to variations in the time course of the cue-evoked currents (Fig. 2c, two-way ANOVA with factors ‘time course’ (P = 0.03, F(4) = 2.8) and ‘condition’ (P < 0.01, F(1) = 72.7; P(interaction) = 0.12, F(4) = 1.8)). We also considered a model with constant cue-evoked currents (step-like model), to further investigate a potential role (if any) of the cue’s variable time course on anticipatory activity, and found anticipatory coding in this case (see Supplementary Fig. 5c).

Finally, we tested whether or not anticipation was present if inhibitory neurons were included in the network (see Supplementary Figs. 4 and 5b). Two-way ANOVA with factors ‘cue variance’ (P = 0.09, F(7) = 1.8) and ‘condition’ (P < 0.01, F(1) = 90; P(interaction) = 0.07, F(7) = 1.9)). Anticipatory activity was also robust to variations in the time course of the cue-evoked currents (Fig. 2c, two-way ANOVA with factors ‘time course’ (P = 0.03, F(4) = 2.8) and ‘condition’ (P < 0.01, F(1) = 72.7; P(interaction) = 0.12, F(4) = 1.8)). We also considered a model with constant cue-evoked currents (step-like model), to further investigate a potential role (if any) of the cue’s variable time course on anticipatory activity, and found anticipatory coding in this case (see Supplementary Fig. 5c).
Anticipatory cue speeds up the network’s dynamics. Having established that a clustered architecture enables the effects of expectation on coding, we investigated the underlying mechanism.

Clustered networks spontaneously generate highly structured activity characterized by coordinated patterns of ensemble firing. This activity results from the network hopping between metastable states in which different combinations of clusters are simultaneously activated. To understand how anticipatory inputs affect network dynamics, we analyzed the effects of cue presentation for a prolonged period of 5 seconds in the absence of stimuli. Activating anticipatory inputs led to changes in network dynamics, with clusters turning on and off more frequently in the presence of the cue (Fig. 3a). We quantified this effect by showing that a cue-induced increase in input spatial variance ($\sigma^2$) led to a shortened cluster activation lifetime (top panel in Fig. 3b; Kruskal–Wallis one-way ANOVA: $P < 10^{-17}, \chi^2(5) = 91.2$), and a shorter cluster interaction interval (that is, quiescent intervals between consecutive activations of the same cluster, bottom panel in Fig. 3b; Kruskal–Wallis one-way ANOVA: $P < 10^{-18}, \chi^2(5) = 98.6$).

Previous work has demonstrated that metastable states of co-activated clusters result from attractor dynamics. Hence, the shortening of cluster activations and interaction intervals observed in the model could be due to modifications in the sensory coding induced by expectation and that removing clustering impairs this function.
network's attractor dynamics. To test this hypothesis, we performed a mean field theory analysis of a simplified network with only two clusters, therefore producing a reduced repertoire of configurations. These include two configurations in which either cluster is active and the other inactive (A and B in Fig. 3c), and a configuration where both clusters are moderately active (C). The dynamics of this network can be analyzed using a reduced, self-consistent theory of a single excitatory cluster, said to be in focus (see Methods for details), based on the effective transfer function relating the input and output firing rates of the cluster (r and r_out, Fig. 3c). The latter are equal in the A, B, and C network configurations described above—also called 'fixed points' because these are the points where the transfer function intersects the identity line, r_out = \Phi(r_a).

Configurations A and B would be stable in an infinitely large network, but they are only metastable in networks of finite size, due to intrinsically generated variability. Transitions between metastable states can be modeled as a diffusion process and analyzed using Kramers' theory, according to which the transition rates depend on the height, \Delta, of an effective energy barrier separating them. In our theory, the effective energy barriers (see Fig. 3c, bottom row) are obtained as the area of the region between the identity line and the transfer function (shaded areas in top row of Fig. 3c; see Methods for details). The effective energy is constructed so that its local minima correspond to stable fixed points (here, A and B) whereas local maxima correspond to unstable fixed points (C). Larger barriers correspond to less frequent transitions between stable configurations, whereas lower barriers increase the transition rates and therefore accelerate the network's metastable dynamics.

This picture provides the substrate for understanding the role of the anticipatory cue in the expectation effect. Basically, the presentation of the cue modulates the shape of the effective transfer function, which results in the reduction of the effective energy barriers. More specifically, the cue-induced increase in the spatial variance, \sigma^2, of the afferent current, flattens the transfer function along the identity line, reducing the area between the two curves (shaded regions in Fig. 3c). In turn, this reduces the effective energy barrier separating the two configurations (Fig. 3c, bottom row), resulting in faster dynamics. The larger the cue-induced spatial variance \sigma^2 in the afferent currents, the faster the dynamics (Fig. 3d; lighter shades represent larger \sigma values).

In summary, this analysis shows that the anticipatory cue increases the spontaneous transition rates between the network's metastable configurations by reducing the effective energy barrier necessary to hop among configurations. In the following we uncover an important consequence of this phenomenon for sensory processing.

Anticipatory cue induces faster onset of taste-coding states. The cue-induced modulation of attractor dynamics led us to formulate a hypothesis for the mechanism underlying the acceleration of coding: the activation of anticipatory inputs before sensory stimulation may allow the network to enter stimulus-coding configurations more easily while exiting non-coding configurations more
Figure 4a shows simulated population rasters in response to the same stimulus presented in the absence of, or following, a cue. Spikes in red represent activity in taste-selective clusters and show a faster activation latency in response to the stimulus preceded by the cue compared with the uncued condition. A systematic analysis revealed that, in the cued condition, the clusters activated by the subsequent stimulus had a significantly faster activation latency than in the uncued condition (Fig. 4b, $0.22 \pm 0.01$ s (mean $\pm$ s.e.m.) during cued compared with $0.32 \pm 0.01$ s for uncued stimuli; $P < 10^{-6}$, two-sided $t$-test, degrees of freedom $= 39$).

We explained this effect using mean field theory. In the simplified two-cluster network of Fig. 4c (the same network as in Fig. 3d), the configuration where the taste-selective cluster is active (coding state) or the non-selective cluster is active (non-coding state) have initially the same effective potential energy (local minima of the black line in Fig. 4c). The onset of the cue reduces the effective energy barrier separating these configurations (dashed versus full line). After stimulus onset, the coding state sits in a deeper well (lighter lines) compared with the non-coding state, due to the stimulation biasing the selective cluster. Stronger stimuli (lighter shades in Fig. 4c) progressively increase the difference between the wells’ depths breaking their initial symmetry, making a transition from the non-coding to the coding state more likely than a transition from the coding to the non-coding state$^{14}$. The anticipatory cue reduces further the existing barrier and thereby increases the transition rate toward coding configurations. This results in faster coding, on average, of the encoded stimuli.

We tested this model prediction on the data from Samuelsen et al.$^6$ (Fig. 5). To compare the data with the model simulations, we randomly sampled ensembles of model neurons so as to match the sizes of the empirical datasets. As we only have access to a subset of neurons in the experiments, rather than the full network configuration, we segmented the ensemble activity in sequences of metastable states via a hidden Markov model (HMM) analysis (see Methods). Previous work has demonstrated that HMM states can be treated as proxies of metastable network configurations$^{14}$. In particular, activation of taste-coding configurations for a particular stimulus results in HMM states containing information about that stimulus (that is, taste-coding HMM states). If the hypothesis originating from the model is correct, transitions from non-coding to taste-coding HMM states should be faster in the presence of the cue. We did indeed find faster transitions to HMM coding states in cued compared with uncued trials for both model and data (Fig. 5a,c, respectively; coding states indicated by horizontal bars), with shorter latency of coding states in both cases (Fig. 5b–d, mean latency of first coding state in the model was $0.32 \pm 0.02$ s for expected versus $0.38 \pm 0.01$ s for unexpected trials; two-sided $t$-test, degrees of freedom $= 39$, $P = 0.014$; in the data, mean latency was $0.46 \pm 0.02$ s for expected versus $0.56 \pm 0.03$ s for unexpected trials; two-sided $t$-test, degrees of freedom $= 37$, $P = 0.026$).

Altogether, these results demonstrate that anticipatory inputs speed up sensory coding by reducing the effective energy barriers from non-coding to coding metastable states.

**Discussion**

Expectations modulate perception and sensory processing. Typically, expected stimuli are recognized more accurately and more rapidly than unexpected ones$^{13}$. In the gustatory cortex, expectation of tastants has been related to changes in firing activity evoked by anticipatory cues$^7$. What causes these firing rate changes? In the present study, we propose that the effects of expectation follow from the modulation of the intrinsically generated cortical dynamics, ubiquitously observed in cortical circuits$^{4,14,16–18,22,23–37}$. The proposed mechanism entails a recurrent spiking network where excitatory neurons are arranged in clusters$^{4,14}$.

In such a model, network activity unfolds through state sequences, the
Articles

Notably, this anticipatory mechanism is unrelated to changes in network excitability, which would lead to unidirectional changes in firing rates. It relies instead on an increase in the spatial (that is, across neurons) variance of the network's activity caused by the anticipatory cue. This increase in the input's variance is observed experimentally after training, and is therefore the consequence of having learned the anticipatory meaning of the cue. The consequent acceleration of state sequences predicted by the model was also confirmed in the data from ensembles of simultaneously recorded neurons in awake-behaving rats.

These results provide a functional interpretation of ongoing cortical activity and a precise explanatory link between the intrinsic dynamics of neural activity in a sensory circuit and a specific cognitive process, that of general expectation.

Clustered connectivity and metastable states. A key feature of our model is the clustered architecture of the excitatory population. Theoretical work had previously shown that a clustered architecture can produce stable activity patterns. Noise (either externally or internally generated) may destabilize those patterns and ignite a progression through metastable states. These states are reminiscent of those observed in the cortex during both task engagement and inter-trial periods, including those found in rodent gustatory cortex during taste processing and decision-making. Clustered spiking networks also account for various physiological observations such as stimulus-induced reduction of trial-to-trial variability, neural dimensionality, and firing rate multistability.

In the present study, we showed that clustered spiking networks also have the ability to modulate coding latency and therefore explain the phenomenon of general expectation. The uncovered link of generic anticipatory cues, network metastability, and coding speed is dependent on a clustered architecture, because removing the excitatory clusters (that is, having homogeneous connectivity) eliminates the anticipatory mechanism (see Fig. 1d–f).

Functional role of heterogeneity in cue responses. As stated in the previous section, the presence of clusters is a necessary ingredient to obtain a faster latency of coding. Below we discuss the second necessary ingredient, that is, the presence of heterogeneous neural responses to the anticipatory cue (see Fig. 1b).

Responses to anticipatory cues have been extensively studied in cortical and subcortical areas in alert rodents and . Cues evoke heterogeneous patterns of activity, either exciting or inhibiting single neurons. The proportion of cue responses and their heterogeneity develop with training, suggesting a fundamental function of these patterns. In the generic expectation paradigm considered here, the anticipatory cue does not convey any information about the identity of the forthcoming tastant, but rather it just signals the availability of a stimulus. Experimental evidence suggests that the cue may induce a state of arousal, which was previously described as 'priming' the sensory cortex. Here, we propose an underlying mechanism in which the cue is responsible for acceleration of coding by increasing the spatial variance of the pre-stimulus activity. In turn, this modulates the shape of the neuronal current-to-coding by increasing the spatial variance of the pre-stimulus activity. Additional mechanisms and complex patterns of connectivity may be required to explain the suppression effects induced by the cue. However, in the present study we provide a parsimonious explanation of how heterogeneous responses can improve coding without postulating any specific pattern of connectivity other than (1) random projections from thalamic and anticipatory cue afferents and (2) the clustered organization of the intracortical circuitry. Notice that the latter contains wide distributions of synaptic weights and can be understood as the consequence of Hebb-like reorganization of the circuitry during training.

Specificity of the anticipatory mechanism. Our model of anticipation relies on gain reduction in clustered excitatory neurons due to a larger spatial variance of the afferent currents. We have shown that this model is robust to variations in parameters and architecture (see Fig. 2 and Supplementary Fig. 5); what about the specificity of its mechanism? A priori, the expectation effect might be achieved through different means, such as: increasing the strength of feedforward couplings; decreasing the strength of recurrent couplings; or modulating background synaptic inputs. However, when scoring those models on the criteria of coding anticipation and heterogeneous cue responses, we found that they failed to simultaneously match both criteria, although for some range of parameters they could reproduce either one (see Supplementary Figs. 7–9 and Supplementary Table 1 for a detailed analysis).

Although our exploration of the alternative models' parameter space did not produce an example that could match the data, we cannot in principle exclude the possibility that one of the alternative models could match the data in a yet unexplored parameter region. This may be particularly the case for the model shown in Supplementary Fig. 7c, which can produce anticipation of coding but does not match the patterns of cue responses observed in the experiments. We are aware that, due to the typically large number of cellular and network parameters, it is hard to rule out these or other alternative models entirely. We could, however, demonstrate that the main mechanism proposed here (see Fig. 1a) captures the plurality of experimental observations pertaining to anticipatory activity in a robust and biologically plausible way.

Cortical timescales, state transitions, and cognitive function. In populations of spiking neurons, a clustered architecture can generate reverberating activity and sequences of metastable states. Transitions from state to state can be typically caused by external inputs. For instance, in frontal cortices, sequences of states are related to specific epochs within a task, with transitions evoked by behavioral events. In sensory cortex, progressions through state sequences can be triggered by sensory stimuli and reflect the dynamics of sensory processing. Importantly, state sequences have also been observed in the absence of any external stimulation, promoted by intrinsic fluctuations in neural activity. However, the potential functional role, if any, of this type of ongoing activity has remained unexplored.

Recent work has started to uncover the link between ensemble dynamics and sensory and cognitive processes. State transitions in various cortical areas have been linked to decision-making, choice representation, rule-switching behavior, and the level of task difficulty. However, no theoretical or mechanistic explanations have been given for these phenomena.

Here we provide a mechanistic link between state sequences and expectation in terms of specific modulations of intrinsic dynamics, that is, the anticipatory cue triggers a change of the transition probabilities. The modulation of intrinsic activity can dial the duration of states, producing either shorter or longer timescales. A shorter timescale leads to faster state sequences and coding anticipation after stimulus presentation (see Figs. 1 and 5). Other external perturbations may induce different effects: for example, recruiting the network's inhibitory population slows down the timescale, leading to a slower coding (see Supplementary Fig. 6).
The interplay between intrinsic dynamics and anticipatory influences presented here is a mechanism for generating diverse timescales, and may have rich computational consequences. We demonstrated its function in increasing coding speed, but its role in mediating cognition is likely to be broader and calls for further explorations.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41593-019-0364-9.

Received: 6 October 2017; Accepted: 15 February 2019; Published online: 1 April 2019

References
1. Gilbert, C. D. & Sigman, M. Brain states: top-down influences in sensory processing. Neuron 54, 677–696 (2007).
2. Jaramillo, S. & Zador, A. M. The auditory cortex mediates the perceptual effect of acoustic temporal expectation. Nat. Neurosci. 14, 246–251 (2011).
3. Engel, A. K., Fries, P. & Singer, W. Dynamic predictions: oscillations and synchrony in top-down processing. Nat. Rev. Neurosci. 2, 704–716 (2001).
4. Doherty, J. R., Rao, A., Mesulam, M. M. & Nobre, A. C. Synergistic effect of combined temporal and spatial expectations on visual attention. J. Neurosci. 25, 8259–8266 (2005).
5. Niwa, M., Johnson, J. S., O’Connor, K. N. & Sutter, M. L. Active engagement improves primary auditory cortical neurons’ ability to discriminate temporal modulation. J. Neurosci. 32, 9323–9334 (2012).
6. Samuelsen, C. L., Gardner, M. P. & Fontanini, A. Effects of cue-triggered expectation on cortical processing of taste. Neuron 74, 410–422 (2012).
7. Yoshida, T. & Katz, D. B. Control of prestimulus activity related to improved sensory coding within a discrimination task. J. Neurosci. 31, 4101–4112 (2011).
8. Gardner, M. P. & Fontanini, A. Encoding and tracking of outcome-specific expectancy in the gustatory cortex of alert rats. J. Neurosci. 34, 13000–13017 (2014).
9. Vincis, R. & Fontanini, A. Associative learning changes cross-modal representations in the gustatory cortex. eLife 5, e16420 (2016).
10. Amit, D. J. & Brunel, N. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. Cereb. Cortex 7, 237–252 (1997).
11. Deco, G. & Hugues, E. Neural network mechanisms underlying stimulus driven variability reduction. PLoS Comput. Biol. 8, e1002395 (2012).
12. Harvey, C. D., Coen, P. & Tank, D. W. Choice-specific sequences in localised cortical networks with balanced excitatory and inhibitory activity. Science 273, 1724–1726 (1996).
13. Curti, E., Mongillo, G., La Camera, G. & Ami, D. J. Mean field and capacity in realistic networks of spiking neurons storing sparsely coded random memories. Neuron Comput. 16, 2597–2637 (2004).
14. Mascaro, M. & Amit, D. J. Effective neural response function for collective population states. Network 10, 351–373 (1999).
15. Mattia, M. et al. Heterogeneous attractor cell assemblies for motor planning in premotor cortex. J. Neurosci. 33, 11155–11168 (2013).
16. La Camera, G., Giugliano, M., Senn, W. & Fusi, S. The response of cortical neurons to in vivo-like input current: theory and experiment. I. Noisy inputs with stationary statistics. Biol. Cybern. 99, 279–301 (2008).
17. Hänggi, P., Töllner, P. & Borkovec, M. Reaction-rate theory: Fifty years after Kramers. Rev. Mod. Phys. 62, 251 (1990).
18. Renart, A. et al. The asynchronous state in cortical circuits. Nat. Rev. Neurosci. 1, 773–782 (2002).
19. Runyan, C. A., Piasini, E., Panzeri, S. & Harvey, C. D. Context-dependent coding of population states in sensory ensembles. Nature 425, 954–956 (2003).
20. Plasselakhova, E., Itoh, V., Amarasingsha, A. & Buzsáki, G. Internally generated cell assembly sequences in the rat hippocampus. Science 321, 1322–1327 (2008).
21. Luczak, A., Barthó, P. & Harris, K. D. Spontaneous events outline the realm of possible sensory responses in neocortical populations. Neuron 62, 413–425 (2009).
22. Puccini, G. D., Sanchez-Vives, M. V. & Compte, A. Integrated mechanisms of anticipation and rate-of-change computations in cortical circuits. PLoS Comput. Biol. 3, e82 (2007).
23. Miller, P. & Katz, D. B. Stochastic transitions between neural states in taste processing and decision-making. J. Neurosci. 30, 2559–2570 (2010).
24. Jones, L. M., Fontanini, A., Sadacca, B. F., Miller, P. & Katz, D. B. Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles. Proc. Natl. Acad. Sci. USA 104, 18772–18777 (2007).
25. Runyan, C. A., Piasini, E., Panzeri, S. & Harvey, C. D. Context-dependent timescales of population coding across cortex. Nature 548, 92–96 (2017).
26. Sadacca, B. F. et al. The behavioral relevance of cortical neural ensemble responses emerges suddenly. J. Neurosci. 36, 655–696 (2016).
27. Churchland, M. M. et al. Stimulus onset quenches neural variability: A widespread cortical phenomenon. Nat. Neurosci. 13, 569–578 (2010).
28. Liu, H. & Fontanini, A. State dependence of chemosensory coding in the gustatory thalamus (VPMpc) of alert rats. J. Neurosci. 35, 15479–15491 (2015).
29. Grewe, B. F. et al. Neural ensemble dynamics underlying a long-term associative memory. Nature 543, 670–675 (2017).
30. Chow, S. S., Romo, R. & Brody, C. D. Context-dependent modulation of functional connectivity: secondary somatosensory cortex to prefrontal cortex connections in two-stimulus-interval discrimination tasks. J. Neurosci. 29, 7238–7245 (2009).
31. Zenke, F., Agnes, E. J. & Gerstner, W. Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. Nat. Commun. 6, 6922 (2015).
32. Litwin-Kumar, A. & Doiron, B. Formation and maintenance of neuronal assemblies through synaptic plasticity. Nat. Commun. 5, 5319 (2014).
33. Chance, F. S., Abbott, L. F. & Reyes, A. D. Gain modulation from background input on cortical processing of taste. Neuron 13, 773–782 (2002).
34. Engel, T. A. et al. Selective modulation of cortical state during spatial attention. Science 334, 1140–1144 (2010).

Acknowledgements
This work was supported by a National Institute of Deafness and Other Communication Disorders Grant no. K25-DC013557 (L.M.), by the Swartz Foundation Award 66438 (L.M.), by National Institute of Deafness and Other Communication Disorders
Grant nos. R01DC012543 and R01DC015234 (A.E.), and partly by a National Science Foundation Grant no. IIS1161852 (G.L.C.). The authors would like to thank S. Fusi, A. Maffei, G. Mongillo, and C. van Vreeswijk for useful discussions.

Author contributions
L.M., G.L.C., and A.F. designed the project, discussed the models and the data analyses, and wrote the manuscript. L.M. performed the data analysis, model simulations, and theoretical analyses.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41593-019-0364-9.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to G.C. or A.F.

Journal peer review information: Nature Neuroscience thanks Paul Miller and other anonymous reviewer(s) for their contribution to the peer review of this work.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019
Methods
Behavioral training and electrophysiology. The experimental data come from a previously published dataset. Experimental procedures were approved by the Institutional Animal Care and Use Committee of Stony Brook University and complied with university, state, and federal regulations on the care and use of laboratory animals. Movable bundles of 16 microwires were implanted bilaterally in the gustatory cortex and intraoral cannulae were inserted bilaterally. After postsurgical recovery, rats were trained to self-administer fluids through intraoral cannulae by pressing a lever under head-restraint within 3 s presentation of an auditory cue (expected rates of 5–7.5 spks s⁻¹; a frequency of 5 kHz). The intervalal period was progressively increased to 40 ± 3 s. Early presses were discouraged by the addition of a 2-s delay to the intervalal period. During training and electrophysiology sessions, additional tasts were automatically delivered through the intraoral cannulae at random times near the middle of the intervalal period and in the absence of the anticipatory cue (ununexpected trials). Only ensembles with five or more simultaneously recorded neurons were included in the rest of the analyses. Ongoing and evoked activity were defined as occurring in the 5-s-long interval before or after taste delivery, respectively.

Cue responsiveness. A neuron was deemed responsive to the cue if a sudden change in its firing rate was observed during the post-cue interval, detected via a ‘change-point’ procedure described in the literature. Briefly, we built the cumulative distribution function of the spike count (CumSc) across all trials in each session, in the interval starting 0.5 s before, and ending 1 s after, cue delivery. We then ran the change-point detection algorithm on the CumSc record with a given tolerance level P=0.05. If any change point was detected before cue presentation, the algorithm was repeated with lower tolerance (lower P value) until no change point was present before the cue. If a legitimate change point was found anywhere within 1 s after cue presentation, the neuron was deemed responsive. If no change point was found, the neuron was deemed not responsive. For neurons with excited (inhibited) cue responses, change in peristimulus time histogram (ΔPSTH) was defined as the difference between positive (negative) peak responses post-cue and the mean baseline activity in the 0.5 s preceding cue presentation.

Ensemble states detection. An HMM analysis was used to detect ensemble states in both the empirical data and the model simulations. Below, we give a brief description of the method used and we refer the reader the literature already in the literature for more detailed information.

The HMM assumes that an ensemble of N simultaneously recorded neurons is in one of M hidden states at each given time bin. States are firing rate vectors \( r(m) \), where \( m = 1, \ldots, N \) is the neuron index and \( n = 1, \ldots, M \) identifies the state. In each state, neurons were assumed to discharge as stationary Poisson processes (Poisson-HMM) conditional on the state’s firing rates. Trials were segmented in 2 ms bins, and the value of either 1 (spike) or 0 (no spike) was assigned to each bin for each given neuron (Bernoulli approximation for short time bins); if more than one neuron fired in a given bin (a rare event), a single spike was randomly assigned to one of the firing neurons. A single HMM was used to fit all trials in each recording session, resulting in the emission probabilities \( r(m) \) and in a set of transition probabilities between the states. Emission and transition probabilities were calculated using the Baum–Welch algorithm\(^{43}\) with a fixed number of hidden states \( M \), yielding a maximum likelihood estimate of the parameters given the observed spike trains. As the model log-likelihood, LL increases with \( M \), we repeated the HMM fits for increasing values of \( M \) until we hit a minimum of the Bayesian information criterion (BIC, see below and Zucchini and Macdonald\(^{42}\)).

For each of the, L, used in the BIC was the sum of over ten independent HMM fits with random initial guesses for emission and transition probabilities. This step was needed because the Baum–Welch algorithm guarantees reaching only a local rather than a global maximum of the likelihood. The model with the lowest BIC (having \( M^* \) states) was selected as the winning model, where \( BIC = -2 \times \log(\text{LL}) + \alpha \times MN \ln N \ln T \). Being estimated in each session (which equals number of trials \( N \) x number of bins per trials). Finally, the winning HMM model was used to ‘decode’ the states from the data according to their posterior probability, given the data. During decoding, only those states with probability exceeding 80% in at least 25 consecutive 2-ms bins were retained (henceforth denoted simply as states)\(^{43}\). This procedure eliminates states that appear only very transiently and with low probability, also reducing the chance of overfitting. A median of six states per ensemble was found, varying from three to nine across ensembles.

Coding states. In each condition (that is, expected versus unexpected), the frequencies of occurrence of a given state across taste stimuli were compared with a test of proportions (\( q \geq 0.001 \) with Bonferroni correction to account for multiple states). When a significant difference was found across stimuli, a post-hoc Marascuilo test was performed\(^{2} \). A state with a frequency of occurrence significantly higher in the presence of one taste stimulus compared with all other tastes was deemed a ‘coding state’ for that stimulus (see Fig. 5).

Sweeping network model. We modeled the local neural circuit as a recurrent network of N leaky-integrate-and-fire neurons, with a fraction \( n_2 = 80\% \) of excitatory (E) and \( n_3 = 20\% \) of inhibitory (I) neurons\(^{18}\). Connectivity was random with probability \( p_{EI} = 0.2 \) for E-to-E connections and \( p_{II} = p_{EI} = 0.5 \) otherwise. Significant weights \( I_{j} \) from presynaptic neuron \( j \) to postsynaptic neuron \( i \) were defined as:

\[
I_{ij} = I_{c} + \alpha \delta(t - \tau_{ij})
\]

where \( I_{c} \) is the constant background current, \( \alpha \) is a firing-rate-dependent synaptic weight, \( \delta(t) \) is the Dirac delta function, and \( \tau_{ij} \) is the delay between postsynaptic neuron \( i \) and presynaptic neuron \( j \). The firing rate of the network was then clamped to the reset value, \( V_{res} \), for a refractory period, \( t_{ref} \). Thresholds were chosen so that the homogeneous network fired at rates \( r_1 = 5 \) spks s⁻¹ and \( r_2 = 7 \) spks s⁻¹. The recurrent contribution to the postsynaptic current to the 4th neuron was a low-pass filter of the incoming spike trains

\[
I_{rec} = \frac{d}{dt} V - \frac{V}{\tau_{m}} + I_{ext} + I_{synr}
\]

where \( r_1 \) is the membrane time constant and the input currents \( I_{ext} \) are a sum of a recurrent contribution \( I_{rec} \) coming from the other network neurons and an external current \( I_{ext} = I_{ext,E} + I_{ext,I} \) (units of \( V/s \)). Here, \( I_{ext,E} \) and \( I_{ext,I} \) are the incoming stimulus and cue, respectively (see Stimulation protocols below). When \( V \) hits threshold, \( V_{th} \), a spike is emitted and \( V \) is then clamped to the reset value, \( V_{res} \), for a refractory period, \( t_{ref} \). Thresholds were chosen so that the homogeneous network fired at rates \( r_1 = 5 \) spks s⁻¹ and \( r_2 = 7 \) spks s⁻¹. The recurrent contribution to the postsynaptic current to the 4th neuron was a low-pass filter of the incoming spike trains

\[
r_{rec} = \frac{d}{dt} I_{rec} = \frac{d}{dt} \sum_{j=1}^{N} I_{ij} \sum_{k=1}^{N} \delta(t - \tau_{ij})
\]

where \( \tau_{ij} \) is the synaptic time constant, \( I_{ij} \) is the recurrent synaptic weights from presynaptic neuron \( j \) to postsynaptic neuron \( i \), and \( \delta(t) \) is the 4th spike time from the 3rd postsynaptic current. The external constant current \( I_{ext} \) was \( 0 \) for all neurons except those neurons targeted by the stimulus, which had \( I_{ext,E} \) and \( I_{ext,I} \) representing the incoming stimulus and cue, respectively (see Stimulation protocols below). When \( V \) hits threshold, \( V_{th} \), a spike is emitted and \( V \) is then clamped to the reset value, \( V_{res} \), for a refractory period, \( t_{ref} \). Thresholds were chosen so that the homogeneous network fired at rates \( r_1 = 5 \) spks s⁻¹ and \( r_2 = 7 \) spks s⁻¹. The recurrent contribution to the postsynaptic current to the 4th neuron was a low-pass filter of the incoming spike trains

Stimulation protocols. Stimuli were modeled as time-varying, stimulus afferent currents targeting 50% of neurons in stimulus-selective clusters \( I_{stim}(t) \). Each cluster had a 50% probability of being selective to a given stimulus, so different stimuli targeted overlapping sets of clusters. The anticipatory cue, targeting a random 50% subset of E neurons, was modeled as a double exponential with rise and decay times of 0.2 and 1 s, respectively, unless otherwise specified; its peak value for each selective neuron was sampled from a normal distribution with zero mean and s.d. \( \sigma \) (expressed as fraction of the baseline current \( I_{c} \), \( \sigma = 20\% \) unless otherwise specified). The cue did not change the mean afferent current but only its spatial (quenched) variance across neurons.

In both the unexpected and the expected conditions, stimulus onset at \( t = 0 \) was followed by a linear increase \( r_{stim}(t) \) in the stimulus afferent current to stimulus-selective neurons, reaching a value \( r_{stim} \) at \( t = 1 \) s (\( r_{stim} = 20\%, \) unless otherwise specified). In the expected condition, stimuli were preceded by the anticipatory cue \( r_{stim}(t) \) with onset at \( t = -0.5 \) s before stimulus presentation.

Network simulations. All data analyses, model simulations, and mean field theory calculations were performed using custom software written in MATLAB.
Mean field theory. Mean field theory was used in a simplified network with two excitatory clusters (as parameters as in Table 2 in the main text) using the population density approach: the input to each neuron was completely characterized by the infinitesimal mean \( \mu_e \) and variance \( \alpha_e \) of the postsynaptic current for Q+2 neural populations; the first Q populations representing the Q excitatory clusters, the (Q+1)th population representing the background unstructured excitatory population, and the (Q+2)th population representing the inhibitory population (see Table 1 in the main text for parameter values):

\[
\mu_e = \tau_m s \frac{n_e f}{Q} \left( \sum_{i=1}^{Q} \frac{f_{pE}}{J_L} + \frac{J_{inh} r_{inh} + n_i (1-f_i) P_{E/E} f_{inh}}{\tau_{inh}} \right) + n_i (1-f) P_{E/E} f_{inh} \left( r_{inh} + n_i P_{E/E} \right) + n_i \left( 1-f_i \right) \frac{f_{pE}}{J_L} \left( 1 + \delta \right) r_{inh} + \frac{J_{inh} r_{inh} + n_i (1-f_i) P_{E/E} f_{inh}}{\tau_{inh}} \]

\[
\alpha_e^2 = \tau_m s \frac{n_e f}{Q} \left( \sum_{i=1}^{Q} \frac{f_{pE}}{J_L} + \frac{J_{inh} r_{inh} + n_i (1-f_i) P_{E/E} f_{inh}}{\tau_{inh}} \right) \left( 1 + \delta \right) r_{inh} + \frac{J_{inh} r_{inh} + n_i (1-f_i) P_{E/E} f_{inh}}{\tau_{inh}} \left( 1 + \delta \right) \]

where \( r_{inh} \) is the input current to a background E-population neuron reads:

\[
\mu_e^{(inh)} = \tau_m s \frac{n_e f}{Q} \left( \sum_{i=1}^{Q} \frac{f_{pE}}{J_L} + \frac{J_{inh} r_{inh} + n_i (1-f_i) P_{E/E} f_{inh}}{\tau_{inh}} \right) - n_i \frac{J_{inh} r_{inh} + n_i P_{E/E} f_{inh}}{\tau_{inh}}
\]

\[
\alpha_e^{(inh)2} = \tau_m s \frac{n_e f}{Q} \left( \sum_{i=1}^{Q} \frac{f_{pE}}{J_L} + \frac{J_{inh} r_{inh} + n_i (1-f_i) P_{E/E} f_{inh}}{\tau_{inh}} \right) \left( 1 + \delta \right) r_{inh} + n_i \frac{J_{inh} r_{inh} + n_i P_{E/E} f_{inh}}{\tau_{inh}} \left( 1 + \delta \right) r_{inh}
\]

and, similarly, for an I neuron:

\[
\mu_i = \tau_m s \frac{n_e f}{Q} \left( \sum_{i=1}^{Q} \frac{f_{pE}}{J_L} + \frac{J_{inh} r_{inh} + n_i (1-f_i) P_{E/E} f_{inh}}{\tau_{inh}} \right) - n_i \frac{J_{inh} r_{inh} + n_i P_{E/E} f_{inh}}{\tau_{inh}}
\]

\[
\alpha_i^{2} = \tau_m s \frac{n_e f}{Q} \left( \sum_{i=1}^{Q} \frac{f_{pE}}{J_L} + \frac{J_{inh} r_{inh} + n_i (1-f_i) P_{E/E} f_{inh}}{\tau_{inh}} \right) \left( 1 + \delta \right) r_{inh} + n_i \frac{J_{inh} r_{inh} + n_i P_{E/E} f_{inh}}{\tau_{inh}} \left( 1 + \delta \right) r_{inh}
\]

The network fixed points satisfied the Q+2 self-consistent mean field equations:

\[
r_a = E_a \left( r_a, \sigma_a^2 \right)
\]

where \( r = \left[ r_1, \ldots, r_Q, r_{inh} \right] \) is the population firing rate vector (boldface represents vectors). \( E_a \) is the current-to-rate function for population \( a \), which varied depending on the population and the condition. In the absence of the anticipatory cue, the LIF current-to-rate function \( E_a \) was used

\[
E_a^{LIF}(r_a, \sigma_a^2) = \tau_a + r_a (1 + \theta \text{erf}(a) )
\]

where \( \theta_a = \frac{\mu_a - \theta_c}{\sigma_a} + \frac{1}{2} \left( 1 + \frac{1}{\sqrt{\pi}} \right) \) is the baseline afferent current, and \( \theta_c \) is the anticipatory cue's s.d. as a fraction of \( \mu_a \) (see Fig. 3d). For the plots in Fig. 3c we replaced the transfer function \( E_a(r_a, \sigma_a^2) \) of the leaky-integrate-and-fire neuron with the simple function \( f_a = 0.5(1 + \text{tanh}(\mu_a)) \). Note that the latter depends only on the mean input current, because the mean input is the only variable required to implement the effect of the anticipatory cue on the transfer function via \( f_a(\mu_a) \).

\[
F_a^{cue}(r_a, \sigma_a^2) = \int Dz \ F_a^{LIF}(r_a + \sigma_a^2, \mu_a + \sigma_a^2) \]

Fixed points \( r^* \) of equation (1) were found with Newton's method; the fixed points were stable (attractors) when the stability matrix

\[
S = \frac{1}{\tau_m s} \frac{\partial E_a}{\partial r_a} \left( \mu_a, \sigma_a^2 \right) \frac{\partial E_a}{\partial \sigma_a^2} \left( \mu_a, \sigma_a^2 \right) \frac{\partial E_a}{\partial \mu_a} \left( \mu_a, \sigma_a^2 \right) \]

evaluated at \( r^* \) was negative definite. Stability was defined with respect to an approximate linearized dynamics of the mean \( m_t \) and s.d. \( s_t \) of the input currents:

\[
\frac{dm_t}{dt} = -m_t + \mu_e(r)
\]

\[
\frac{ds_t}{dt} = -s_t + \sigma_e^2(r)
\]

\[
r_a(t) = E_a(m_a(t), \sigma_a^2(t))
\]

Effective mean field theory for the reduced network. The mean field equation (1) for the \( P = Q + 2 \) populations may be reduced to a set of effective equations governing the dynamics of a smaller subset of \( q < P \) of populations, henceforth referred to as populations in focus. The reduction is achieved by integrating out the remaining \( P-q \) out-of-focus populations. This procedure was used to estimate the energy barrier separating the two network attractors in Figs. 3d and 4c. Given a fixed set of values \( \bar{r} = [r_1, \ldots, r_q] \) for the in-focus populations, one solves the mean field equations for \( P-q \) out-of-focus populations

\[
r_p(\bar{r}) = E_p(m_p(\bar{r}), \sigma_p^2(\bar{r}))
\]

and, similarly, for an I neuron:

\[
\sigma_i^2 = \tau_m s \frac{n_e f}{Q} \left( \sum_{i=1}^{Q} \frac{f_{pE}}{J_L} + \frac{J_{inh} r_{inh} + n_i (1-f_i) P_{E/E} f_{inh}}{\tau_{inh}} \right) \left( 1 + \delta \right) r_{inh} + n_i \frac{J_{inh} r_{inh} + n_i P_{E/E} f_{inh}}{\tau_{inh}} \left( 1 + \delta \right) r_{inh}
\]

for \( \beta = q+1, \ldots, P \) to obtain the stable fixed point \( \bar{r}^* = [\bar{r}_1, \ldots, \bar{r}_q, \bar{r}_p^*(\bar{r})] \) of the out-of-focus populations as functions of the in-focus firing rates \( \bar{r} \). Stability of the solution \( \bar{r}^* \) is computed with respect to the stability matrix (2) of the reduced system of \( P-q \) out-of-focus populations. Substituting the values \( \bar{r}^* \) into the fixed-point equations for the \( q \) populations in focus yields a new set of equations relating input rates \( \bar{r} \) to output rates \( \bar{r}^* \):

\[
r_a^{out}(r) = E_a \left[ \mu_a (r, \bar{r}), \sigma_a^2 (r, \bar{r}) \right]
\]

where \( \mu_a \) and \( \sigma_a^2 \) are the stationary values given above.
for \( \alpha = 1, \ldots, q \) in-focus populations at the point \( \tilde{r} \). The fixed points \( r^\text{out}_a(\tilde{r}) = r^\text{in}_a \) of the in-focus equation (4) are fixed points of the entire system. It may occur, in general, that the out-of-focus populations attain multiple attractors for a given value of \( \tilde{r} \), in which case the set of effective transfer functions \( E^\text{out} \) is labeled by the chosen attractor; in our analysis of the two-clustered network, only one attractor was present for a given value of \( \tilde{r} \).

**Energy potential.** In a network with \( Q = 2 \) clusters, one can integrate out all populations (out-of-focus) except one (in-focus) to obtain the effective transfer functions for the in-focus population representing a single cluster, with firing rate \( \tilde{r} \) (equation (4) for \( q = 1 \)). Network dynamics can be visualized on a one-dimensional curve, where it is well approximated by the first-order dynamics (see Mascaro and Amit\(^{31} \) for details):

\[
\tau_{\text{res}} \frac{d\tilde{r}}{dt} = -\tilde{r} + r^\text{out}(\tilde{r})
\]

These dynamics can be expressed in terms of an effective energy function \( E(\tilde{r}) \) as

\[
\tau_{\text{syn}} \frac{d\tilde{r}}{dt} = -\frac{dE(\tilde{r})}{d\tilde{r}}
\]

so that the dynamics can be understood as a motion in an effective potential energy landscape, as if driven by an effective force \( \frac{dE(\tilde{r})}{d\tilde{r}} \). The minima of the energy with respect to \( \tilde{r} \) are the stable fixed points of the effective one-dimensional dynamics, whereas its maxima represent the effective energy barriers between two minima, as illustrated in Fig. 3c. The one-cluster network has three fixed points, two stable attractors (A and B in Fig. 3c) and a saddle point (C). We estimated the height, \( \Delta \), of the potential energy barrier on the trajectory from A to B through C as minus the integral of the force from the first attractor A to C:

\[
\Delta = \int_A^C (\tilde{r} - r^\text{out}(\tilde{r})) d\tilde{r}
\]

which represents the area between the identity line \( \tilde{r} = \tilde{r} \) and the effective transfer function \( \tilde{r} = r^\text{out}(\tilde{r}) \) (see Fig. 3c). In the finite network, where the dynamics comprise stochastic transitions among the states, switching between A and B would occur with a frequency that depends on the effective energy barrier \( \Delta \), as explained in the main text.

**Population decoding.** The amount of stimulus-related information carried by spike trains was assessed through a decoding analysis\(^{59} \) (see Supplementary Fig. 3 for illustration). A multiclass classifier was constructed from the population (one neuron from each of the Q clusters for clustered networks, or Q random excitatory neurons for homogeneous networks). Spike counts from all trials of neuron \( i \) taste stimulus in each condition (expected versus unexpected) were split into training and test sets for cross-validation. A template was created for the population peristimulus time histogram for each stimulus, with neuron \( q \) from a given cluster. A cluster was deemed active if its firing rate exceeded 10 spks s\(^{-1} \). This threshold was chosen so as to lie between the inactive and active clusters’ firing rates, which were obtained from a mean field solution of the network\(^{16} \).

**References**

51. Zucchi, W. & MacDonald, I.L. Hidden Markov Models for Time Series: An Introduction Using R (CRC Press, 2009).
52. La Camera, G. & Richmond, B. J. Modeling the violation of reward maximization and invariance in reinforcement schedules. *PLoS Comput. Biol.*, e1000131 (2008).
53. Tuckwell, H. C. *Introduction to Theoretical Neurobiology* (Cambridge Univ. Press, 1988).
54. Lánský, P. & Sato, S. The stochastic diffusion models of nerve membrane depolarization and interspike interval generation. *J. Peripher. Nerv. Syst.*, 4, 27–42 (1999).
55. Richardson, M. J. Effects of synaptic conductance on the voltage distribution and firing rate of spiking neurons. *Phys. Rev. E*, 69, 051918 (2004).
56. Brunel, N. & Sergi, S. Firing frequency of leaky integrate-and-fire neurons with synaptic current dynamics. *J. Theor. Biol.*, 195, 87–95 (1998).
57. Fourcaud, N. & Brunel, N. Dynamics of the firing probability of noisy integrate-and-fire neurons. *Neural Comput.*, 14, 2057–2110 (2002).
58. La Camera, G., Rauh, A., Lüscher, H. R., Senn, W. & Fusi, S. Minimal models of adapted neuronal response to in vivo-like input currents. *Neural Comput.*, 16, 2101–2124 (2004).
59. Rigotti, M. et al. The importance of mixed selectivity in complex cognitive tasks. *Nature*, 497, 585–590 (2013).
60. Breiman, L. Bagging predictors. *Mach. Learn.*, 24, 123–140 (1996).
Reporting Summary

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

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- n/a

- Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement

- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly

- The statistical test(s) used AND whether they are one- or two-sided
  Only common tests should be described solely by name; describe more complex techniques in the Methods section.

- A description of all covariates tested

- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons

- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)

- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
  Give P values as exact values whenever suitable.

- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings

- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes

- Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
  Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

- Single units were isolated using a template algorithm, clustering techniques, and examination of inter-spike interval plots (Offline Sorter, Plexon).

Data analysis

- All data analyses, model simulations and mean field theory
- All calculations were performed using custom software written in MATLAB R2017b (MathWorks), and C.

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

Data

Policy information about availability of data

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

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

A demo code for simulating the network model in the expected and unexpected conditions will be available at publication on GitHub (https://github.com/mazzulab).

Field-specific reporting

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

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences
Life sciences study design

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

| Sample size | No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in previous publications. |
|-------------|----------------------------------------------------------------------------------------------------------------------------------|
| Data exclusions | Only experimental sessions yielding ensembles of five or more simultaneously recorded neurons were included in the analyses, due to the statistical constraints inherent in the hidden Markov model analysis. |
| Replication | We have run our simulations multiple times and we are able to reproduce our conclusions. |
| Randomization | Each network was initialized with random synaptic weights and simulated with random initial conditions. Stimulus and cue selectivities were assigned to randomly chosen sets of neurons in each network. |
| Blinding | Not relevant - there were no experimental groups. |

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

**Materials & experimental systems**

| n/a | Involved in the study |
|-----|-----------------------|
| ☒ Antibodies |
| ☒ Eukaryotic cell lines |
| ☒ Palaeontology |
| ☒ Animals and other organisms |
| ☒ Human research participants |
| ☒ Clinical data |

**Methods**

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