Anticipation is a strategy used by neural fields to compensate for transmission and processing delays during the tracking of dynamical information, and can be achieved by slow, localized, inhibitory feedback mechanisms such as short-term synaptic depression, spike-frequency adaptation, or inhibitory feedback from other layers. Based on the translational symmetry of the mobile network states, we derive generic fluctuation-response relations, providing unified predictions that link their tracking behaviors in the presence of external stimuli to the intrinsic dynamics of the neural fields in their absence.

PACS numbers: 87.19.ll, 05.40.-a, 87.19.lq
Figure 1: (color online). (a) The rescaled neuronal current, \( \tilde{u}(x, t) \), and the rescaled inhibitory variable for SFA, \( \tilde{v}(x, t) \) during a spontaneous motion in the moving frame centered at \( z(t) \). \( \tilde{v} \) plays the role of \( p \) in Eqs. (1) and (2). \( z(t) \) is the center of mass of \( \tilde{u}(x, t) \). The \( \tilde{u}(x, t) \) profile is moving to the direction pointed by the arrow. Parameters: \( \tilde{k} \) (rescaled inhibition) = 0.5, \( \gamma \) (SFA strength) = 0.2, \( \tau_u \) (time constant of neuronal current) = 1 ms and \( \tau_v \) (time constant of SFA) = 50 ms. (b) \( \tilde{u}(x, t) \) and \( \tilde{I}^{ext}(x, t) \), rescaled external stimulus, during a tracking process. Inset: \( z_0(t) \) and \( z(t) \), the centers of mass of \( \tilde{I}^{ext}(x, t) \) and \( \tilde{u}(x, t) \), respectively. The \( \tilde{I}^{ext}(x, t) \) profile is moving in the direction of the arrow with velocity \( v_I \). Parameters: \( \tilde{k} = 0.5, \gamma = 0, \tau_u = 1 \) ms, \( \tilde{A} \) (rescaled magnitude of \( \tilde{I}^{ext} \)) = 1.0 and \( v_I = 0.01 \). (c) Displacement of the \( \tilde{u} \) profile relative to the external stimulus, \( z(t) - z_0(t) \). Parameters: \( \tilde{k} = 0.5, \tau_v = 50 \) ms and \( \tau_u = 1 \) ms. (d) Curve: The anticipation time, \( \tau_{ant} = \frac{[z(t) - z_0(t)]}{v_I} \), for the case with \( \gamma = 0.1 \) in (c). Symbols: Anticipation time in Fig. 4 of [26] with the assumption that \( \tau_i = 50 \) ms and \( \alpha = 22.5^\circ \).

This provides the context for the application of the FRR.

In processing time-dependent external information, real-time response is an important and even a life-and-death issue to animals. However, time delay is pervasive in the dynamics of neural systems. For example, it takes 50 – 80 ms for electrical signals to transmit from the retina to the primary visual cortex, and 10 – 20 ms for the neurons to process and integrate temporal information in such tasks as speech recognition and motor control [10].

To achieve real-time tracking of moving stimuli, a way to compensate delays is to predict their future position. This is evident in experiments on the head-direction (HD) systems of rodents during head movements [11, 12], in which the direction perceived by the HD neurons has nearly zero lag with respect to the true instantaneous position [13], or can even lead the current position by a constant time [14]. This anticipatory behavior is also observed when animals make saccadic eye movements [15]. In psychophysics experiments, the future position of a continuously moving object is anticipated, but intermittent flashes are not [16].

There are different delay compensation strategies, and many of them have slow, local inhibitory
feedback in their dynamics. For example, short-term synaptic depression (STD) can implement anticipatory tracking [17]. Its underlying mechanism is the slow depletion of neurotransmitters in the active region of the network state, facilitating the network state to shift to neighboring positions. For sufficiently strong STD, the tracking state can even overtake the moving stimulus. At the same time, local inhibitory feedbacks can induce spontaneous motion of the localized states in neural fields [18–20]. Remarkably, the parameter region of anticipatory tracking is effectively identical to that of spontaneous motion. Since spontaneous motion sets in when location fluctuation diverges, this indicates the close relation between fluctuations and responses, and implies that such a relation should be more generic than the STD mechanism itself.

Besides STD, other mechanisms can also provide slow, local inhibitory feedback to neurons. Examples include spike-frequency adaptation (SFA) that refers to the reduction of neuron excitability after prolonged stimulation [21], and inhibitory feedback layers (IFL) in multilayer networks [22]. Like STD, such local inhibition can generate spontaneous traveling waves [18]. Likewise, they are expected to exhibit anticipatory tracking [22]. In this Letter, we will consider how FRR provides a unified picture for this family of systems driven by different neural mechanisms. As will be shown, generic analyses based on the translational symmetry of the systems show that anticipative tracking is associated with spontaneous motions, thus providing a natural mechanism for delay compensation.

We consider a neural field in which neurons are characterized by location $x$, interpreted as the preferred stimulus of the neuron, which can be spatial location [7] or head direction [6]. Neuronal activities are represented by $u(x, t)$, interpreted as neuronal current [23]. To keep the formulation generic, the dynamical equation is written in the form

$$
\frac{\partial u(x, t)}{\partial t} = F_u[x; u, p] + I_{\text{ext}}(x, t).
$$

(1)

$F_u$ is a functional of $u$ and $p$ evaluated at $x$. $p$ is a dynamical variable representing neuronal activities with no direct connections with the external environment. In the context of anticipatory tracking, $p$ corresponds to a dynamical local inhibitory mechanism. It could represent the available amount of neurotransmitters of presynaptic neurons for STD [20, 24], or the shift of the firing thresholds due to SFA [21], or the neuronal activities of a hidden neural field layer in IFL [22]. Explicit forms of $F_u[x; u, p]$ for STD, SFA and IFL can be found in Supplementary Material (SM). Besides the force $F_u$, the dynamics is also driven by an external input, $I_{\text{ext}}$. 
Figure 2: (color online). (a) Contours of intrinsic speed in the phase diagram of a neural field model with SFA. (b) Contours of anticipation time of a neural field model with SFA. (c) Same as (a), but for IFL. (d) Same as (b), but for IFL. Color curves: contours of intrinsic speed ((a)&(c)), anticipatory time ((b)&(d)).

Number labels: values of the corresponding contour, in units of (a) $a/\tau_i$, (b) $\tau_i$, (c) $\tau_1/\tilde{J}_{ff}$ and (d) $a/(\tau_1/\tilde{J}_{ff})$.

Black curves: phase boundaries separating the static, moving, repulsive (for IFL only) and silent phases.

Parameters: (a) $\tau_i = 50\tau_s$, (b) $\hat{A} = 0.25$, $v_I = 0.002a/\tau_s$ and $\tau_1 = 50\tau_s$. (c) $\tilde{J}_{ff} = 0.1$ and $\tau_1$ (time constant of the primary layer) = $\tau_2$ (time constant of the hidden layer) = $\tau_s$, (d) $\tilde{J}_{ff} = 0.1$, $\hat{A} = 0.1$, $v_I = 0.002a/\tau_s$, $\tau_1 = \tau_2$. In the shaded areas of (b) and (d), $\hat{A}$ is too small to stabilize the system.

Similar to Eq. (1), the dynamics of $p$ is given by

$$\frac{\partial p(x,t)}{\partial t} = F_p[x; u, p].$$

$F_p$ is also a functional of $u$ and $p$ evaluated at $x$. Explicit expressions of $F_p$ for STD, SFA and IFL can be found in SM. For the present analysis, it is sufficient to assume that (i) the forces are translationally invariant, and (ii) the forces possess inversion symmetry.

Studies on neural field models showed that they can support a profile of localized activity even in the absence of external stimuli [5, 8, 9, 23]. Irrespective of the explicit form of this “bump”, it is sufficient to note that (i) there exists a non-trivial stable solution $\{u_0, p_0\}$ satisfying $F_u[x; u_0, p_0] = F_p[x; u_0, p_0] = 0$, and that this solution is neutrally stable in $x$, that is, $F_u[x-z; u_0, p_0] = F_p[x-z; u_0, p_0] = 0$ for arbitrary bump position $z$.

Studies on neural field models with STD [19, 20], SFA [25] and IFL [22] suggested that the network can support spontaneously moving profiles, even though there is no external moving input. This occurs when the static solution becomes unstable to positional displacement in some parameter region. As shown in SM, the natural speed of the moving profile is given by $v_{\text{nat}} = \varepsilon_0/\tau_{\text{int}}$, where $\varepsilon_0$ is the lag of the inhibitory profile $p$ behind the exposed profile $u$, and $\tau_{\text{int}}$ is an intrinsic time scale. This shows that the spontaneous motion is caused by the inhibitory profile lagging behind the exposed profile. It is further shown in SM that the value of $\varepsilon_0$ depends on an
instability eigenvalue $\lambda$, which is negative in the static phase, and positive in the moving phase. Hence $\varepsilon_0$ vanishes in the static phase, and becomes non-zero in the moving phase.

In the presence of a weak and slow external stimulus, the anticipation time of the exposed profile relative to the stimulus profile is given by $\tau_{\text{ant}} = \lambda \tau_{\text{int}} \tau_{\text{stim}}$, where $\tau_{\text{stim}}$ is referred to as the stimulus time, representing the time scale for the stimulus to produce significant response from the exposed profile. Hence when the intrinsic behavior changes from static to moving, the tracking behavior changes from lagging ($\tau_{\text{ant}} < 0$) to leading ($\tau_{\text{ant}} > 0$) with respect to the moving stimulus.

For the example of the neural field with SFA in Fig. 1(a), the lag of the inhibitory profile $\tilde{v}$ drives the exposed profile $\tilde{u}$ to move in the direction with smaller $\tilde{v}$ (pointed by the arrow), as $\tilde{v}$ inhibits $\tilde{u}$.

In the absence of SFA, the bell-shaped attractor state of $\tilde{u}$ centered at $z(t)$ (shown in Fig. 1(b) as the green dashed line) lags behind a continuously moving stimulus $z_I(t)$ (shown as the blue dotted line). In the inset of Fig. 1(b), the lag of the network response develops after the stimulus starts to move and becomes steady after a while. This lag is due to the time needed for neurons
to build up their activities. In contrast, when SFA is sufficiently strong, the bump can track the stimulus at an advancing position (red solid curve in Fig. [1b]). In this case, this tracking process anticipates the continuously moving stimulus. This behavior for SFA is summarized in Fig. [1c].

Furthermore, the anticipation time is effectively constant in a considerable range of the stimulus speed. There is an obvious advantage for the brain to compensate delays with a constant leading time independent of the stimulus speed, and this contrasts with the intuition that the anticipation time decreases with the stimulus speed. To see this point, we note that $\varepsilon_0 = v_I \tau_{\text{int}}$, implying that $\tau_{\text{ant}} = \lambda \tau_{\text{stim}} \varepsilon_0 / v_I$. This shows that while the stimulus speed in the denominator increases, the lag of the inhibitory profile behind the exposed profile in the numerator also increases, providing an increasing driving force for the bump.

This is confirmed when the SFA strength $\gamma$ is strong enough. As shown in Fig. [1c] for $\gamma = 0.1$, there is a velocity range such that the displacement of the center of mass relative to the stimulus, $z(t) - z_I(t)$, is directly proportional to the stimulus velocity. Thus the anticipation time $\tau_{\text{ant}} \equiv (z - z_I) / v_I$, given by the slope of the curve, is effectively constant. In Fig. [1d], the anticipatory time is roughly $0.3 \tau_\text{I}$ ($\tau_\text{I}$ is the time constant of SFA) for a range of stimulus velocity, and has a remarkable fit with data from rodent experiments [26]. This behavior can also be observed in neural field models with STD [17].

The interdependency of anticipatory tracking dynamics and intrinsic dynamics is further illustrated by the relation between the anticipatory time and the intrinsic speed of spontaneous motions. Near the boundary of the moving phase, it is derived in SM that

$$\tau_{\text{ant}} = K \tau_{\text{stim}} \tau_{\text{int}}^3 \left(v_{\text{nat}}^2 - v_I^2\right) + K_0 \tau_{\text{int}},$$

(3)

where $K$ and $K_0$ are constants. Since all parameters besides $v_{\text{nat}}^2$ and $v_I^2$ (taken to approach 0) are mostly slowly changing functions of system parameters, the contours of $v_{\text{nat}}$ and $\tau_{\text{ant}}$ in the parameter space have very similar patterns. The case for SFA is illustrated in Fig. [2a] and (b).

Since these phenomena depend on the underlying symmetry of the system and its response to weak stimuli, they are expected to be observed in networks with the same symmetry as SFA networks. The correspondence between intrinsic motion and anticipation has been described in the specific case of STD networks [17]. Similar phenomena can be found in Fig. [2c] and (d) for IFL, except that the contours in Fig. [2] are distorted in the proximity of the repulsive phase. A minor correction is that the contour for zero anticipatory time does not coincide perfectly with the phase boundary separating the moving and static phases. This is due to deviations from the weak
input limit, since a finite input amplitude is necessary to prevent the network state from becoming “untrackable” (the region shaded in (b) and (d)).

For non-vanishing stimulus velocities in the moving phase, Eq. (3) predicts another interesting phenomenon linking tracking dynamics and intrinsic dynamics. When the stimulus is moving at the natural speed, i.e. $v_I = v_{nat}$, the anticipation time becomes independent of the strength of the external input which determines $\tau_{stim}$, and the anticipation time curves are confluent at the value $\tau_{ant} = \tau_{con} \equiv \tau_{int}/K_0$. This resonance phenomenon for a particular neural field model with STD has been reported in [17]; here we show that it is generic in an entire family of neural fields.

The physical picture of this confluent behavior is that the stimulus plays two roles in driving the moving bump. First, it is used to drive the bump at the stimulus speed, if it is different from the intrinsic speed. Second, it is used to distort the shape of the bump. In the second role, the distortion is proportional to both the strength of the stimulus and the bump-stimulus displacement, $z(t) - z_0(t)$. Hence when the stimulus speed is the same as the intrinsic speed, the stimulus is primarily used to distort the bump shape. At the steady state, the bump-stimulus displacement is determined by the distortion per unit stimulus strength, which becomes independent of stimulus strength.

Since this phenomenon is based on a generic mechanism, it can be observed in all neural field models considered in this Letter. Fig. 3 shows the simulation results in neural field models with SFA and IFL. Fig. 3(a) shows the displacements in the SFA neural field model with the intrinsic speed $v_{nat} = a/\tau_i$, where $\tau_i$ is the SFA time scale. $\tau_{ant} - v_I$ curves corresponding to different stimulus amplitudes intersect at $\tau_i v_{nat}/a = 1$. Similar behaviors are shown in Fig. 3(b) for $v_{nat} = 2.5a/\tau_i$, and in Fig. 3(c) and (d) for IFL. Remarkably, the confluent behavior remains valid even when the curves deviate from the parabolic shape predicted by Eq. (3).

To further illustrate FRR, we consider the correlation between fluctuations due to noise in the absence of external input and the anticipation time reacting to a weak and slow moving stimulus. This can be done by replacing $I^{ext}$ in Eq. (1) with displacement noise $\xi(x,t) \equiv \eta(t) \partial u_0/\partial x$, where $\langle \eta(t) \rangle = 0$ and $\langle \eta(t) \eta(t') \rangle = 2T \delta(t - t')$. Analysis in SM shows that for weak and slow stimuli,

$$\frac{\langle \delta \varepsilon_0^2 \rangle}{T} = \begin{cases} -\frac{\tau_{stim} \tau_{int}}{\tau_{ant} - \tau_{con}}, & \text{for static phase}, \\ \frac{\tau_{stim} \tau_{int}}{2(\tau_{ant} - \tau_{con})}, & \text{for moving phase}. \end{cases}$$

Here, $\langle \delta \varepsilon_0^2 \rangle$ represents the fluctuations of the lag of the inhibitory profile behind the exposed profile.
in response to the displacement noise.

The behavior predicted by Eq. (4) can be seen from simulations. In Fig. 4 there are two branches in each sub-figure. The branches for \( \tau_{\text{ant}} > \tau_{\text{con}} \) and \( \tau_{\text{ant}} < \tau_{\text{con}} \) correspond to the moving and static phases respectively. Remarkably, data points with different network parameters collapse onto common curves. The fluctuations are divergent at the confluence point predicted by Eq. (3). The regimes of \( \tau_{\text{ant}} > 0 \) and \( \tau_{\text{ant}} < 0 \), corresponding to anticipatory and delayed tracking respectively, effectively coincide with the two branches in the limit of weak stimuli, since at the confluence point the instability eigenvalue \( \lambda = -\tau_{\text{con}} / (\tau_{\text{stim}} \tau_{\text{int}}) \) approaches 0 in that limit.

In summary, we have found a fluctuation-response relation for neural fields having an exposed variable \((u)\) and a slow inhibitory variable \((p)\), relating: (i) the positional stability of the activity states, to (ii) their lagging/leading position relative to external stimuli during tracking, and to (iii) fluctuations due to thermal noises. This relation is entirely generic irrespective of the explicit form of the dynamical equations. It is only based on the following properties of the dynamical equations: (i) the existence of a non-zero solution, and (ii) this solution is translationally invariant and (iii) possesses inversion symmetry. The resulting inference, as a unified property applicable to
neural fields with slow inhibitory feedback mechanisms such as STD, SFA, IFL and other neural fields of the family, is that the regimes of anticipatory and delayed tracking effectively coincide with the regimes of moving and static phases respectively, and that the anticipation time becomes independent of stimulus speed for slow and weak stimuli, and independent of stimulus amplitude when the stimulus moves at the intrinsic speed. Thus, FRR provides us a unified framework to analyze mechanisms of delay compensation in the brain, which have been observed in HD cells and are expected to be an important element in understanding the processing of dynamic information in the brain.

This work is supported by the Research Grants Council of Hong Kong (grant numbers 604512, 605813 and N_HKUST606/12) and the National Foundation of Natural Science of China (No. 31221003, No. 31261160495).

Supplementary Material

S.1. GENERAL FORMULATION

Here, we search for a general relation between the translational stability of neural field models and their tracking behavior. We will start from general dynamical equations without considering their explicit forms. There are two dynamical variables: \( u(x, t) \) and \( p(x, t) \). We call \( u(x, t) \) the exposed variable, which can be the firing rate or neuronal current of neurons with preferred stimulus \( x \). It can be coupled to external stimuli. We call \( p(x, t) \) the inhibitory variable, which can be variables describing short-term synaptic depression (STD), spike frequency adaptation (SFA), or inhibitory feedbacks from higher layer networks (specifically, we consider two-layered networks with an inhibitory feedback layer (IFL)). \( p(x, t) \) is not directly coupled to the outside world. \( u(x, t) \) receives slow inhibitory feedback from \( p(x, t) \), while \( p(x, t) \) receives fast feed-forward excitation from \( u(x, t) \), but our results are not restricted to any explicit forms of these couplings.

The dynamics of \( u(x, t) \) and \( p(x, t) \) are governed by

\[
\frac{\partial u}{\partial t} = F_u [x; u, p] + I^{\text{ext}} (x, t) \quad \text{and} \quad \frac{\partial p}{\partial t} = F_p [x; u, p].
\]

(S.1)

(S.2)

\( F_u \) and \( F_p \) are functionals of \( u(x, t) \) and \( p(x, t) \) to be defined below. \( I^{\text{ext}} (x, t) \) is the external stimulus. Although the analytical result does not depend on the exact definitions of \( F_u \) and \( F_p \),
choices of $F_u$ and $F_p$ are important for us to use numerical simulation to verify the analytical result. In simulations, $u(x, t)$ is chosen to be the neuronal current of neurons having preferred stimulus $x$.

**A. Short-term Synaptic Depression**

For short-term synaptic depression (STD), $F_u$ is defined by [20]

$$\frac{\partial u(x, t)}{\partial t} = F_u[x; u, p] \equiv \frac{1}{\tau_s} \left[ \rho \int dx' J(x, x') p(x', t) r(x', t) - u(x, t) + I^{\text{ext}}(x, t) \right].$$  (S.3)

$\tau_s$ is the neuronal time scale, and is of the order of 1 ms. $\rho$ is the density of neurons in the space spanned by $\{x\}$. $r(x, t)$ is the firing rate of neurons at $x$ defined by

$$r(x, t) = \frac{\max(u(x, t), 0)^2}{1 + k \rho \int dx' u(x', t)^2}.$$  (S.4)

$k$ is the inhibition parameter controlling the global divisive inhibition.

In Eq. (S.3), $J(x, x')$ is the coupling between neurons at $x$ and $x'$, depending on $x - x'$ and hence is translationally invariant:

$$J(x, x') = \frac{J_0}{a \sqrt{2\pi}} \exp \left[ -\frac{(x - x')^2}{2a^2} \right].$$  (S.5)

$J_0$ is the strength of the excitatory coupling. $a$ is the range of the coupling.

$p(x, t)$ is the available fraction of neurotransmitters of presynaptic neurons. Its dynamics is given by

$$\frac{\partial p(x, t)}{\partial t} = F_p[x; u, p] \equiv \frac{1 - p(x, t)}{\tau_d} - \beta p(x, t) r(x, t).$$  (S.6)

$\tau_d$ is the timescale of $p$, which is of the order 100 ms. $\beta$ is the fraction of neurotransmitter consumed per spike, and represents the strength of the STD effect. $I^{\text{ext}}$ is defined by

$$I^{\text{ext}}(x, t) = A \exp \left\{ -\frac{(x - z_0(t))^2}{4a^2} \right\}.$$  (S.7)

$A$ is the magnitude of $I^{\text{ext}}$. $z_0(t)$ is the position of the stimulus.

**B. Spike Frequency Adaptation**

For spike frequency adaptation (SFA), the dynamics of $u$ is given by [27]

$$\frac{\partial u(x, t)}{\partial t} = F_u[x; u, v] \equiv \frac{1}{\tau_s} \left[ \rho \int dx' J(x, x') r(x', t) - v(x, t) - u(x, t) + I^{\text{ext}}(x, t) \right].$$  (S.8)
\( v \) is the suppression due to SFA. Other parameters and variables have the same definitions as mentioned above. Here, for simplicity, we let

\[
\frac{\partial v(x,t)}{\partial t} = F_p [x; u, v] \equiv \frac{1}{\tau} [\gamma \max(u(x,t) , 0) - v(x,t)]. \tag{S.9}
\]

\( \tau \) is the SFA timescale, and is of the order of 100 ms. \( \gamma \) is the magnitude of SFA. Note that \( v \) acts as \( p \) in Eqs. (S.1) and (S.2).

C. Inhibitory Feedback Layer

In a two-layered network with an inhibitory feedback layer, there are two variables representing neuronal currents on different layers \[22\]. They are denoted as \( u_1 \) and \( u_2 \). Their dynamics are specified by

\[
\frac{\partial u_1(x,t)}{\partial t} = F_u [x; u_1, u_2] \equiv \frac{1}{\tau_1} \left[ \rho \int dx' J(x, x') r_1(x', t) 
+ \rho \frac{J_{fb}}{J_0} \int dx' J(x, x') r_2(x', t) - u_1(x, t) + I_{ext}(x, t) \right], \tag{S.10}
\]

\[
\frac{\partial u_2(x,t)}{\partial t} = F_p [x; u_1, u_2] \equiv \frac{1}{\tau_2} \left[ \rho \int dx' J(x, x') r_2(x', t) 
+ \rho \frac{J_{ff}}{J_0} \int dx' J(x, x') r_2(x', t) - u_2(x, t) \right]. \tag{S.11}
\]

\( u_1 \) and \( u_2 \) play the role of \( u \) and \( p \) in Eqs. (S.1) and (S.2). \( r_1 \) and \( r_2 \) are the firing rates given in Eq. (S.4) by replacing \( u \) to be \( u_1 \) or \( u_2 \). \( J_{fb} \) is the magnitude of the feedback connections, and \( J_{ff} \) is the magnitude of the feed-forward connections.

D. Rescaled Variables and Parameters

In the models presented above, the dynamical variables and parameters can be rescaled to simplify the dynamical picture. Using the rescaled parameters can also reduce the parameter-dependence of simulation results.

We first consider STD. In (S.1), we rescale \( u(x,t) \) to \( \tilde{u}(x,t) \equiv \rho J_0 u(x,t) \), and \( k \) to \( \tilde{k} \equiv k / \left[ \rho J_0^2 / (8 \sqrt{2\pi} a) \right] \). Then \( r(x,t) \) can be rescaled by

\[
\tilde{r}(x,t) \equiv \rho^2 J_0^2 r(x,t) = \frac{\tilde{u}(x,t)^2}{1 + \frac{k}{8} \int \frac{dx'}{a\sqrt{2\pi}} \tilde{u}(x,t)^2}, \tag{S.14}
\]
Figure S.1: A snapshot of the network state of a two-layered network in its moving phase. \( \tilde{u}_1 \) and \( \tilde{u}_2 \) are the rescaled neuronal current profile of the first and second layers respectively. The profiles are moving in the direction of the arrow at the top. Parameters: \( \tilde{k} = 0.5, \tilde{J}_{ff} = 0.1, \tilde{J}_{fb} = -0.1 \) and \( \tau_2 = \tau_1 = 1 \).

and \( J(x, x') \) is rescaled to \( \tilde{J}(x, x') \equiv (J_0)^{-1} J(x, x') \). In Eq. (S.2), \( p(x, t) \) is already dimensionless, and the only parameter needs to be rescaled is \( \beta \), which becomes \( \tilde{\beta} \equiv \tau_0 \beta / (\rho^2 J_0^2) \). In Eq. (S.7), the external input \( I_{\text{ext}}(x, t) \) becomes \( \tilde{I}_{\text{ext}}(x, t) \equiv \rho J_0 I_{\text{ext}}(x, t) \) and \( A \) becomes \( \tilde{A} \equiv \rho J_0 A \).

Note that in the absence of STD, the profile of \( \tilde{u} \) has a Gaussian shape with a magnitude of \( \max_x \tilde{u}(x, t) = \sqrt{8(1 + \sqrt{1 - \tilde{k}})} / \tilde{k} \). To study the dynamical properties of the network, we need to maintain the time scale \( \max_x \tilde{u}(x, t) / \tilde{A} \) (referred to as the stimulation time \( \tau_{\text{stim}} \) in Sec. S.6) to be a slowly changing parameter as \( \tilde{k} \) varies, as required in Figs. 2-4 in the main text. Hence we introduce the rescaled variable \( \hat{A} \equiv \tilde{A} / [\sqrt{8} (1 + \sqrt{1 - \tilde{k}}) / \tilde{k}] \), so that \( \hat{A} / \max_x \tilde{u}(x, t) \) is roughly a constant approximating \( \tau_{\text{stim}}^{-1} \).

For SFA, \( u(x, t) \) in Eq. (S.8) can also be rescaled by \( \tilde{u}(x, t) \). Also, \( \tilde{r}, \tilde{J}, \) and \( \tilde{k} \) are applicable in this case. In Eq. (S.9), we let \( \tilde{v}(x, t) = \rho J_0 v(x, t) \) since \( v(x, t) \) has the same dimension as \( u(x, t) \).

For IFL, \( u_1(x, t) \) and \( u_2(x, t) \) in Eqs. (S.11) and (S.13) have the same dimension as \( u(x, t) \). So, we apply the same rescaling: \( \tilde{u}_1(x, t) \equiv \rho J_0 u_1(x, t) \), \( \tilde{u}_2(x, t) \equiv \rho J_0 u_2(x, t) \) and \( \tilde{k} \equiv k / [\rho J_0^2 / (8 \sqrt{2 \pi a})] \). For convenience, we let \( \tilde{J}_{fb} = J_{fb} / J_0 \) and \( \tilde{J}_{ff} = J_{ff} / J_0 \).
A detailed study on CANNs with STD can be found in [20]. The intrinsic behavior of CANNs with SFA is similar. In this section, the intrinsic behaviors of a bump-shaped profile in a two-layered network with an inhibitory feedback layer are summarized.

If the negative feedback strength ($\tilde{J}_{fb}$) is strong enough, the bump in the second layer that provides a negative feedback to the first layer can destabilize the bump in the first layer. At the steady state, the misalignment between two profiles becomes a constant. As shown in Fig. S.1, the two misaligned bumps move spontaneously. Since the neurons in the first layer receive negative feedbacks and neurons in the second layer receives positive feedforwards, the magnitude of $\tilde{u}_2$-profile is larger than $\tilde{u}_1$-profile.

The intrinsic behavior supported by the system is determined by the choice of parameters. Figure S.2 shows the typical cases of the static phase, the moving phase and the repulsive phase. In simulations, the initial conditions of $\tilde{u}_1$ and $\tilde{u}_2$ are misaligned so that the environment of $\tilde{u}_1$ is not symmetric about its center. If the magnitude of $\tilde{J}_{fb}$ is not strong enough, the bump will relax to a static state, see Fig. S.2 (a) and (b). For a sufficiently strong $\tilde{J}_{fb}$, the bump can move spontaneously as in Fig. S.1 and Fig. S.2 (c) and (d). This is the moving phase. In this phase, the
\( \tilde{u}_2 \)-profile repels the \( \tilde{u}_1 \)-profile. However, at the same time, the \( \tilde{u}_1 \)-profile attracts the \( \tilde{u}_2 \)-profile. So, at an equilibrium state, the misalignment between two profiles becomes steady.

If \( \tilde{J}_{fb} \) is too strong, the spontaneous motion will terminate. In this case, initially, the \( \tilde{u}_2 \)-profile repels the \( \tilde{u}_1 \)-profile and the \( \tilde{u}_1 \)-profile attracts the \( \tilde{u}_2 \)-profile. However, in the repulsive phase, the repulsion is so strong that the attraction can no longer balance the repulsive force. As a result, the two profiles move apart out of the interactive range of each other, as shown in Fig. [S.2](e) and (f). The spontaneous motion cannot sustain at the steady state. In general, together with the trivial solution, there are four phases in two-layer CANNs, under the current setting. The phase diagram for these four phases is shown in the main paper.

The slowness of the inhibitory feedback, and hence the existence of the moving phase, arises from the weak coupling between the exposed and inhibitory layers. To see this, we consider the moving bump solution

\[
\tilde{u}_1(x, t) = u_1^0 \exp \left[-\frac{(x - vt)^2}{4a^2}\right], \quad \text{and} \quad \tilde{u}_2(x, t) = u_2^0 \exp \left[-\frac{(x - vt - s)^2}{4a^2}\right]. \quad (S.15)
\]

Substituting into Eq. (S.1), multiplying both sides by \exp[-(x - vt)^2/(4a^2)]/\sqrt{2\pi a^2} and integrating,

\[
\tilde{u}_1 = \frac{\tilde{u}_2^2}{\sqrt{2B_1}} + \tilde{J}_{fb} \frac{\tilde{u}_2^2}{\sqrt{2B_2}} e^{-\frac{x^2}{8a^2}}, \quad (S.16)
\]

where \( B_1 = 1 + \kappa \tilde{u}_2^2/8 \) and \( B_2 = 1 + \kappa \tilde{u}_2^2/8 \).

Substituting into Eq. (S.1), multiplying both sides by \((x - vt)/a\) \exp[-(x - vt)^2/(4a^2)]/\sqrt{2\pi a^2} and integrating,

\[
\frac{v \tau_1}{2a} \tilde{u}_1 = \tilde{J}_{fb} \frac{\tilde{u}_2^2}{\sqrt{2B_2}} \frac{s}{2a} e^{-\frac{x^2}{8a^2}}, \quad (S.17)
\]

Consider the condition for the moving phase boundary with both \( v \) and \( s \) approaching 0 at a finite ratio. The above equations imply that

\[
\frac{v \tau_1}{s} = \frac{\tilde{J}_{fb} \tilde{u}_2^2}{\sqrt{2B_1}} + \tilde{J}_{fb} \frac{\tilde{u}_2^2}{\sqrt{2B_2}} \sim \tilde{J}_{fb}. \quad (S.18)
\]

Similarly, by considering the dynamics of the second layer, we have

\[
\frac{v \tau_2}{s} \sim -\frac{\tilde{J}_{ff}}{1 + J_{ff}}. \quad (S.19)
\]
Hence weak interlayer couplings, $|\tilde{J}_b| \ll 1$ or $\tilde{J}_H \ll 1$ play the same role as the ratio $\tau_s/\tau_d$ in STD [20].

S.3. TRANSLATION INVARIANCE OF NEURAL FIELD MODELS

We will consider the dynamics of any CANN model described by Eqs. (S.1) and (S.2). In the absence of external stimuli, the dynamical equations are given by

$$\frac{\partial u(x)}{\partial t} = F_u [x; u, p], \text{ and}$$

$$\frac{\partial p(x)}{\partial t} = F_p [x; u, p].$$

(S.20)

(S.21)

Here, $F_u$ and $F_p$ are functionals of the functions $u$ and $p$. They are the generalized forces determining the dynamics of the system. In the existence region of the bump state of the CANN, there is a nontrivial solution to the equations

$$F_u [x; u_0, p_0] = 0, F_p [x; u_0, p_0] = 0.$$  

(S.22)

Furthermore, due to the neutral stability of the bump we have, for an arbitrary bump position $z$,

$$F_u [x - z; u_0, p_0] = 0, F_p [x - z; u_0, p_0] = 0.$$  

(S.23)

Consider the dynamics of the fluctuations about the steady state,

$$\frac{\partial}{\partial t} \delta u (x) = \int dx’ \frac{\partial F_u (x)}{\partial u (x')} \delta u (x') + \int dx’ \frac{\partial F_u (x)}{\partial p (x')} \delta p (x'),$$

(S.24)

$$\frac{\partial}{\partial t} \delta p (x) = \int dx’ \frac{\partial F_p (x)}{\partial u (x')} \delta u (x') + \int dx’ \frac{\partial F_p (x)}{\partial p (x')} \delta p (x').$$

(S.25)

Consider the solutions of these equations with time dependence $\exp(-\lambda t)$. Then the eigenvalue equations become the $\Delta x \to 0$ limit of the matrix eigenvalue equation

$$\left( \begin{array}{c} \frac{\partial F_u (x_i)}{\partial u (x_j)} \\ \frac{\partial F_u (x_i)}{\partial p (x_j)} \\ \frac{\partial F_p (x_i)}{\partial u (x_j)} \\ \frac{\partial F_p (x_i)}{\partial p (x_j)} \end{array} \right) \left( \begin{array}{c} f_u (x_j) \\ f_p (x_j) \end{array} \right) = \lambda \left( \begin{array}{c} f_u (x_i) \\ f_p (x_i) \end{array} \right)$$

(S.26)

The left eigenvector with the same eigenvalue is given by

$$\left( \begin{array}{c} g_u (x_j) \\ g_p (x_j) \end{array} \right) \left( \begin{array}{c} \frac{\partial F_u (x_i)}{\partial u (x_j)} \\ \frac{\partial F_u (x_i)}{\partial p (x_j)} \\ \frac{\partial F_p (x_i)}{\partial u (x_j)} \\ \frac{\partial F_p (x_i)}{\partial p (x_j)} \end{array} \right) \Delta x = -\lambda \left( \begin{array}{c} g_u (x_i) \\ g_p (x_i) \end{array} \right)$$

(S.27)
Translational invariance implies that $\partial u_0/\partial x$ and $\partial p_0/\partial x$ are the components of the right eigenfunction of the dynamical equations with eigenvalue 0, satisfying

$$
\int dx' \frac{\partial F_u}{\partial u(x')} u_i(x') + \int dx' \frac{\partial F_u}{\partial p(x')} p_1(x') = 0, \quad (S.28)
$$

$$
\int dx' \frac{\partial F_p}{\partial u(x')} u_i(x') + \int dx' \frac{\partial F_p}{\partial p(x')} p_1(x') = 0. \quad (S.29)
$$

The corresponding left eigenfunctions satisfy

$$
\int dx' g^0_u(x') \frac{\partial F_u}{\partial u(x)} u_i(x') + \int dx' g^0_p(x') \frac{\partial F_p}{\partial u(x)} p_1(x') = 0, \quad (S.30)
$$

$$
\int dx' g^0_u(x') \frac{\partial F_u}{\partial p(x)} u_i(x') + \int dx' g^0_p(x') \frac{\partial F_p}{\partial p(x)} p_1(x') = 0. \quad (S.31)
$$

For stable bumps, the eigenvalues of all other eigenfunctions are negative. Let $u_1$ and $p_1$ be the components of the eigenfunction with the largest negative eigenvalue $-\lambda_1$, satisfying

$$
\int dx' g_{u_1}^1(x') \frac{\partial F_u}{\partial u(x)} u_i(x') + \int dx' g_{p_1}^1(x') \frac{\partial F_p}{\partial u(x)} p_1(x') = -\lambda_1 g_{u_1}^1(x'), \quad (S.32)
$$

$$
\int dx' g_{u_1}^1(x') \frac{\partial F_u}{\partial p(x)} u_i(x') + \int dx' g_{p_1}^1(x') \frac{\partial F_p}{\partial p(x)} p_1(x') = -\lambda_1 g_{p_1}^1(x'). \quad (S.33)
$$

The eigenfunctions corresponding to eigenvalues $\lambda_m \neq \lambda_n$ satisfy the orthogonal condition

$$
\int dx' g_{u_i}^m(x') f^n_u(x') + \int dx' g_{p_i}^m(x') f^n_p(x') = \delta_{mn}. \quad (S.36)
$$

For later use, we first explore the implications of translational invariance. Multiplying both sides of Eq. (S.28) by $g^0_u$ and integrating over $x$,

$$
\int dx g^0_u(x) \int dx' \frac{\partial F_u}{\partial u(x')} u_i(x') + \int dx g^0_u(x) \int dx' \frac{\partial F_u}{\partial p(x')} p_1(x') = 0. \quad (S.37)
$$

Hence,

$$
Q_{uu} + Q_{up} = 0, \quad (S.38)
$$

where, for $i = u, p$,

$$
Q_{ij} = \int dx g_{u_i}^0(x) \int dx' \frac{\partial F_j}{\partial u_j(x')} \frac{\partial u_i}{\partial x'}. \quad (S.39)
$$
with \( u_j^0(x') \) representing the functions \( u_0(x') \) and \( p_0(x') \) for \( j = u, p \) respectively. The function \( u_j(x') \) in the functional derivatives \( \partial F_j(x)/\partial u_j(x') \) refers to \( u(x') \) and \( p(x') \) for \( j = u, p \) respectively. Similarly,

\[
Q_{pu} + Q_{pp} = 0. \tag{S.40}
\]

Similarly, multiplying both sides of Eq. (S.30) by \( \partial u_0(x)/\partial x \) and integrating over \( x \), we obtain

\[
Q_{uu} + Q_{pu} = 0. \tag{S.41}
\]

Similarly,

\[
Q_{up} + Q_{pp} = 0. \tag{S.42}
\]

Next, we consider the implications of with inversion symmetry, that is, \( \partial F_u(x)/\partial u(x') = \partial F_u(-x)/\partial u(-x') \), \( \partial F_u(x)/\partial p(x') = \partial F_u(-x)/\partial p(-x') \), \( \partial F_p(x)/\partial u(x') = \partial F_p(-x)/\partial u(-x') \) and \( \partial F_p(x)/\partial p(x') = \partial F_p(-x)/\partial p(-x') \). Then the dynamics preserve parity. Suppose the bump state \( u_0(x) \) and \( p_0(x) \) has even parity. Then the distortion mode \( \partial u_0/\partial x \) and \( \partial p_0/\partial x \) has odd parity. Note that the corresponding left eigenfunctions \( g_0^u \) and \( g_0^p \) have the same parity as the right eigenfunctions. Further suppose the next distortion mode \( u_1(x) \) and \( p_1(x) \) has even parity.

### S.4. INTRINSIC BEHAVIOR OF PROFILE SEPARATION

We consider perturbations of the form

\[
\delta u (x) = c_0 \frac{\partial u_0}{\partial x} + c_1 u_1 (x), \quad \delta p (x) = \varepsilon_0 \frac{\partial p_0}{\partial x} + \varepsilon_1 p_1 (x). \tag{S.43}
\]

\( c_0 \) and \( \varepsilon_0 \) are considered to be the displacement of the exposed and inhibitory profiles respectively (in the direction opposite to their signs). They are substituted into the dynamical equations (S.24) and (S.25). Multiplying both sides of Eq. (S.24) by \( g_0^u \) and integrating,

\[
\text{LHS} = \frac{\partial c_0}{\partial t} \int dx g_0^u (x) \frac{\partial u_0}{\partial x} + \frac{\partial c_1}{\partial t} \int dx g_0^u (x) u_1 (x) = \frac{\partial c_0}{\partial t} I_u, \tag{S.44}
\]

where, for \( i = u, p \),

\[
I_i = \int dx g_i^0 (x) \frac{\partial u_i^0}{\partial x}, \tag{S.45}
\]
Note that the second term in Eq. (S.44) vanishes since $g_0^0$ and $u_1$ have opposite parity. On the right hand side,

$$\text{RHS}_1 = c_0 \int dx g_0^0 (x) \int dx' \frac{\partial F_u (x)}{\partial u (x')} \frac{\partial u_0 (x')}{\partial x'} + c_1 \int dx g_0^0 (x) \int dx' \frac{\partial F_u (x)}{\partial u (x')} u_1 (x'). \quad (S.46)$$

The second term vanishes due to odd parity. Hence

$$\text{RHS}_1 = c_0 \int dx g_0^0 (x) \int dx' \frac{\partial F_u (x)}{\partial u (x')} \frac{\partial u_0 (x')}{\partial x'} = c_0 Q_{uu} \quad (S.47)$$

Similarly, the second term on the right hand side becomes

$$\text{RHS}_2 = \varepsilon_0 \int dx g_0^0 (x) \int dx' \frac{\partial F_u (x)}{\partial p (x')} \frac{\partial p_0 (x')}{\partial x'} = \varepsilon_0 Q_{up}. \quad (S.48)$$

Hence we obtain

$$I_u \frac{dc_0}{dt} = Q_{uu} c_0 + Q_{up} \varepsilon_0. \quad (S.49)$$

Similarly, from Eq. (S.25),

$$I_p \frac{d\varepsilon_0}{dt} = Q_{pu} c_0 + Q_{pp} \varepsilon_0. \quad (S.50)$$

Using the identities of translational invariance in Eqs. (S.38) and (S.40),

$$\frac{\partial}{\partial t} \begin{pmatrix} c_0 \\ \varepsilon_0 \end{pmatrix} = \begin{pmatrix} Q_{uu}/I_u & -Q_{uu}/I_u \\ -Q_{pp}/I_p & Q_{pp}/I_p \end{pmatrix} \begin{pmatrix} c_0 \\ \varepsilon_0 \end{pmatrix}. \quad (S.51)$$

This implies

$$\frac{\partial}{\partial t} (\varepsilon_0 - c_0) = \left( \frac{Q_{uu}}{I_u} + \frac{Q_{pp}}{I_p} \right) (\varepsilon_0 - c_0), \quad (S.52)$$

$$\frac{\partial}{\partial t} \left( I_u \frac{c_0}{Q_{uu}} + I_p \frac{\varepsilon_0}{Q_{pp}} \right) = 0. \quad (S.53)$$

Eq. (S.52) describes the dynamics of the displacement of the inhibitory profile relative to the exposed profile. The instability eigenvalue in Eq. (S.52) is denoted as

$$\lambda \equiv \frac{Q_{uu}}{I_u} + \frac{Q_{pp}}{I_p}. \quad (S.54)$$

S.5. INTRINSIC SPEED

When the bump becomes translationally unstable, it moves with an intrinsic speed (or natural speed). To investigate the intrinsic speed, we need to expand the dynamical equation beyond first
order. In this case, the translational variables become coupled with the next eigenfunction. To keep the analysis trackable, we choose the coordinate with $c_0 = 0$. Near the phase boundary of the static and moving phases, $v_{nat} \sim \varepsilon_0$ and $c_1 \sim \varepsilon_1 \sim \varepsilon_0^2$, as will be verified in this section. Hence to include third order terms, it is sufficient to consider terms in the dynamical equations containing $\varepsilon_0, c_1, \varepsilon_1, \varepsilon_0^2, \varepsilon_0 \varepsilon_1$, $\varepsilon_0^3$, $v_{nat} \varepsilon_0$, $v_{nat} c_1$, $v_{nat} \varepsilon_1$. Substituting Eq. (S.43) into the dynamical equation (S.24), expanding to third order for a bump moving with natural speed $v_{nat}$, multiplying both sides of Eq. (S.24) by $g_u^0$ and integrating, 

$$- I_u v_{nat} - M_v v_{nat} c_1 = Q_{up} \varepsilon_0 + T_{upa} \varepsilon_0 c_1 + T_{upp} \varepsilon_0 \varepsilon_1 + \frac{Q_{upp} \varepsilon_3}{6} \varepsilon_0^3,$$  

(S.55)

where, for $i, j, k, l = u, p$,

$$M_i = \int dx g_i^0 \frac{\partial u^1(x)}{\partial x},$$  

(S.56)

$$T_{ijk} = \int dx g_i^0(x) \int dx_1 \int dx_2 \frac{\partial^2 F_i(x)}{\partial u_j(x_1) \partial u_k(x_2)} \frac{\partial u^0(x_1)}{\partial x_1} u^1_j(x_2).$$  

(S.57)

$$Q_{ijkl} = \int dx g_i^0(x) \int dx_1 \int dx_2 \int dx_3 \frac{\partial^3 F_i(x)}{\partial u_j(x_1) \partial u_k(x_2) \partial u_l(x_3)} \frac{\partial u^0(x_1)}{\partial x_1} \frac{\partial u^0(x_2)}{\partial x_2} \frac{\partial u^0(x_3)}{\partial x_3},$$  

(S.58)

Substituting Eq. (S.43) into the dynamical equation (S.24), multiplying both sides of Eq. (S.24) by $g_u^0$ and integrating,

$$\text{LHS} = \frac{\partial c_1}{\partial t} \int dx g_u^1(x) u_1(x) = \frac{\partial c_1}{\partial t} J_u,$$  

(S.59)

where, for $i = u, p$,

$$J_i = \int dx g_i^1(x) u_1^1(x),$$  

(S.60)

with $u_i^1(x)$ representing the functions $u_1(x)$ and $p_1(x)$ for $i = u, p$ respectively. On the right hand side,

$$\text{RHS} = c_1 \int dx g_u^1(x) \int dx' \frac{\partial F_u(x)}{\partial u(x')} u_1(x') + \varepsilon_1 \int dx g_u^1(x) \int dx' \frac{\partial F_u(x)}{\partial p(x')} p_1(x')$$

$$+ \frac{\varepsilon_0^2}{2} \int dx_1 \int dx_2 \frac{\partial^2 F_u(x)}{\partial p(x_1) \partial p(x_2)} \frac{\partial p_0(x_1)}{\partial x_1} \frac{\partial p_0(x_2)}{\partial x_2}$$

$$= c_1 P_{uu} + \varepsilon_1 P_{up} + \frac{S_{upp}}{2} \varepsilon_0^2,$$  

(S.61)

where, for $i, j, k = u, p$,

$$P_{ij} = \int dx g_i^1(x) \int dx' \frac{\partial F_i(x)}{\partial u_j(x')} u_j^1(x').$$  

(S.62)

$$S_{ijk} = \int dx g_i^1(x) \int dx_1 \int dx_2 \frac{\partial^2 F_i(x)}{\partial u_j(x_1) \partial u_k(x_2)} \frac{\partial u^0_j(x_1)}{\partial x_1} \frac{\partial u^0_k(x_2)}{\partial x_2}.$$  

(S.63)
Hence we obtain

\[ J_u \frac{\partial c_1}{\partial t} = P_{uu}c_1 + P_{up} \varepsilon_1 + \frac{S_{upp}}{2} \varepsilon_0^2. \]  

(S.64)

Similarly, from Eq. (S.25),

\[ I_p \frac{\partial \varepsilon_0}{\partial t} - I_p v_{n_0} - M_p v_{n_0} \varepsilon_1 = Q_{ppp} \varepsilon_0 + T_{ppu} \varepsilon_0 c_1 + T_{ppp} \varepsilon_0 \varepsilon_1 + \frac{Q_{pppp}}{6} \varepsilon_0^3. \]  

(S.65)

\[ J_p \frac{\partial \varepsilon_1}{\partial t} - K_p v_{n_0} \varepsilon_0 = P_{pu} c_1 + P_{pp} \varepsilon_1 + \frac{S_{ppp}}{2} \varepsilon_0^2. \]  

(S.66)

where, for \( i = u, p, \)

\[ K_i = \int dx g_i(x) \frac{\varphi_i^0(x)}{\partial x^2}. \]  

(S.67)

At the steady state, Eqs. (S.55) and (S.65) simplify to

\[ M_u \frac{v_{n_0} c_1 - M_p v_{n_0} \varepsilon_1 - \left( \frac{Q_{uu}}{I_u} + \frac{Q_{pp}}{I_p} \right) \varepsilon_0 + \left( \frac{T_{ppu}}{I_u} - \frac{T_{ppu}}{I_p} \right) \varepsilon_0 c_1 + \left( \frac{T_{ppp}}{I_u} - \frac{T_{ppp}}{I_p} \right) \varepsilon_0 \varepsilon_1 + \frac{Q_{pppp} I_u}{I_p} \varepsilon_3^6}{6} = 0. \]  

(S.68)

Eliminating \( c_1 \) and \( \varepsilon_1 \) from Eqs. (S.64) and (S.66),

\[ c_1 = \left( P_{uu} P_{pp} - P_{up} P_{pu} \right)^{-1} \left( -P_{pp} S_{upp} \frac{\varepsilon_0^2}{2} + P_{up} S_{ppp} \frac{\varepsilon_0^2}{2} + K_p P_{up} v_{n_0} \varepsilon_0 \right), \]  

(S.69)

\[ \varepsilon_1 = \left( P_{uu} P_{pp} - P_{up} P_{pu} \right)^{-1} \left( P_{pu} S_{upp} \frac{\varepsilon_0^2}{2} - P_{uu} S_{ppp} \frac{\varepsilon_0^2}{2} - K_p P_{uu} v_{n_0} \varepsilon_0 \right). \]  

(S.70)

Near the critical point, we have \( Q_{uu}/I_u + Q_{pp}/I_p \approx 0 \). It is convenient to eliminate \( c_1 \) and \( \varepsilon_1 \) in Eq. (S.68) using Eqs. (S.69) and (S.70) yielding

\[ \left( \frac{Q_{uu}}{I_u} + \frac{Q_{pp}}{I_p} \right) \varepsilon_0 \]
\[ = \left[ \left( \frac{T_{ppu}}{I_u} - \frac{T_{ppu}}{I_p} \right) \varepsilon_0 + \frac{M_u}{I_u} v_{n_0} \right] \frac{-P_{pp} S_{upp} \frac{\varepsilon_0^2}{2} + P_{up} S_{ppp} \frac{\varepsilon_0^2}{2} + K_p P_{up} v_{n_0} \varepsilon_0}{P_u P_{pp} - P_{up} P_{pu}} \]
\[ + \left[ \left( \frac{T_{ppp}}{I_u} - \frac{T_{ppp}}{I_p} \right) \varepsilon_0 - \frac{M_p}{I_p} v_{n_0} \right] \frac{P_{pu} S_{upp} \frac{\varepsilon_0^2}{2} - P_{uu} S_{ppp} \frac{\varepsilon_0^2}{2} - K_p P_{uu} v_{n_0} \varepsilon_0}{P_u P_{pp} - P_{up} P_{pu}} \]
\[ + \left( \frac{Q_{pppp}}{I_u} - \frac{Q_{pppp}}{I_p} \right) \frac{\varepsilon_0^3}{6} \].  

(S.71)
Furthermore, from Eq. (S.65), we have, to the lowest order,

\[ v_{\text{nat}} \approx - \frac{Q_{pp}}{I_p} \varepsilon_0. \]  

(Eq. (S.55) yields the same result if we make use of the translational symmetry relation in Eq. (S.38).) This enables us to express \( \varepsilon_0 \) and \( v_{\text{nat}} \) in terms of the eigenvalue in Eq. (S.54),

\[ \varepsilon_0 = \pm \sqrt{\lambda K}. \]  

where

\[ K = \left( \frac{T_{ppu}}{I_u} - \frac{T_{ppu}}{I_p} - \frac{M_p Q_{pp}}{I_u I_p} \right) (P_{uu} P_{pp} - P_{up} P_{pu})^{-1} \left( - \frac{P_{pp} S_{upp}}{2} + \frac{P_{up} S_{ppp}}{2} - \frac{K_p P_{up} Q_{pp}}{I_p} \right) \]

\[ + \left( \frac{T_{upp}}{I_u} - \frac{T_{ppp}}{I_p} + \frac{M_p Q_{pp}}{I^2_p} \right) (P_{uu} P_{pp} - P_{up} P_{pu})^{-1} \left( \frac{P_{pp} S_{upp}}{2} - \frac{P_{uu} S_{ppp}}{2} + \frac{K_p P_{up} Q_{pp}}{I_p} \right) \]

\[ + \frac{Q_{pppp}}{6 I_u} - \frac{Q_{pppp}}{6 I_p}. \]  

In the static phase, \( \lambda < 0 \), and both \( v_{\text{nat}} \) and \( \varepsilon_0 \) vanish. In the moving phase, \( \lambda > 0 \), and the critical regime is given by \( v_{\text{nat}} \sim \varepsilon_0 \sim \sqrt{\lambda} \). Using Eq. (S.72), the relation between \( v_{\text{nat}} \) and \( \varepsilon_0 \) is given by

\[ v_{\text{nat}} = \frac{\varepsilon_0}{\tau_{\text{int}}}, \quad \tau_{\text{int}} = - \frac{I_p}{Q_{pp}}. \]  

\( \tau_{\text{int}} \) is an intrinsic time scale of the neural system. The existence of a stable non-trivial solution implies that \( \tau_{\text{int}} \) is positive. Otherwise, Eq. (S.52) implies that \( \varepsilon_0 \) grows exponentially in the reference frame that \( c_0 = 0 \), indicating that the exposed and inhibition profiles disintegrate. The example of the two-layer network illustrates the significance of the negativity of \( Q_{pp}/I_p \) for stable solutions. For weak excitatory feedforward and inhibitory feedback couplings between the two layers, we have shown in Sec. S.2 that the inhibitory profile can be repelled from the exposed profile. In this case, \( Q_{pp}/I_p \) becomes positive and there is no solution with the exposed and inhibitory profiles coupled.

Noting that Eq. (S.75) also holds in the static phase with \( v_{\text{nat}} = \varepsilon_0 = 0 \), we infer that the separation of the exposed and inhibition profiles is the cause of the spontaneous motion. The physical picture is that when the inhibition profile lags behind the exposed profile, the neuronal activity will have a stronger tendency to shift away from the strongly inhibited region.
S.6. TRACKING BEHAVIOR

Here we consider the network response to an external stimulus moving with velocity \( v_I \). The dynamical equations are analogous to those in the previous section, except that an external stimulus is present in the dynamical equation for the exposed profile, and the natural velocity is replaced by the stimulus velocity \( v_I \).

\[
\frac{\partial}{\partial t} \delta u(x) - v_I \frac{\partial u_0(x)}{\partial x} = \int dx' \frac{\partial F_u(x)}{\partial u(x')} \delta u(x') + \int dx' \frac{\partial F_u(x)}{\partial p(x')} \delta p(x') + I^{\text{ext}}(x), \quad (S.76)
\]

\[
\frac{\partial}{\partial t} \delta p(x) - v_I \frac{\partial p_0(x)}{\partial x} = \int dx' \frac{\partial F_p(x)}{\partial u(x')} \delta u(x') + \int dx' \frac{\partial F_p(x)}{\partial p(x')} \delta p(x') . \quad (S.77)
\]

Here, \( x \) is the coordinate relative to the moving bump. Now we consider the distortion due to the bump movement in the reference frame that \( c_0 = 0 \),

\[
\delta u(x) = c_1 u_1(x), \quad \delta p(x) = \varepsilon_0 \frac{\partial p_0}{\partial x} + \varepsilon_1 p_1(x) . \quad (S.78)
\]

To make the discussion more concrete, we consider stimuli having the same profile as the bump, and the bump is displaced by \( s \) relative to the stimulus, that is,

\[
I^{\text{ext}}(x) = \frac{u_0(x + s)}{\tau_{\text{stim}}} \approx \frac{1}{\tau_{\text{stim}}} \left[ u_0(x) + s \frac{\partial u_0(x)}{\partial x} \right], \quad (S.79)
\]

where the amplitude of the stimulus is given by the amplitude of \( u_0(x) \) divided by \( \tau_{\text{stim}} \), referred to as the stimulus time. It represents the time scale for the stimulus to produce significant response from the exposed profile.

Multiplying both sides of Eq. (S.76) by \( g_u \) and integrating, the last term in Eq. (S.76) becomes proportional to the displacement \( s \). Following steps similar to those in the previous section, we obtain the following equations

\[
-I_u v_I - M_u v_I c_1 = Q_{up} \varepsilon_0 + T_{upu} \varepsilon_0 c_1 + T_{upp} \varepsilon_0 \varepsilon_1 + \frac{Q_{uopp}}{6} \varepsilon_0^3 + \frac{I_u}{\tau_{\text{stim}}} s . \quad (S.80)
\]

\[
0 = P_{uu} c_1 + P_{up} \varepsilon_1 + \frac{S_{uopp}}{2} \varepsilon_0^2 + \frac{L_u}{\tau_{\text{stim}}}. \quad (S.81)
\]

\[
-I_p v_I - M_p v_I \varepsilon_1 = Q_{pp} \varepsilon_0 + T_{ppu} \varepsilon_0 c_1 + T_{ppp} \varepsilon_0 \varepsilon_1 + \frac{Q_{pppp}}{6} \varepsilon_0^3 . \quad (S.82)
\]

\[
-K_p v_I \varepsilon_0 = P_{pu} c_1 + P_{pp} \varepsilon_1 + \frac{S_{ppp}}{2} \varepsilon_0^2 . \quad (S.83)
\]

In Eq. (S.81), we have introduced

\[
L_u = \int dx g_u^1(x) u_0(x) . \quad (S.84)
\]
Using Eqs. (S.81) and (S.83), we express \( c_1 \) and \( \varepsilon_1 \) in terms of other variables,

\[
c_1 = (P_{uu}P_{pp} - P_{up}P_{pu})^{-1} \left( -P_{pp}S_{upp} \varepsilon_0^2 + P_{up}S_{ppp} \varepsilon_0^2 + K_P P_{uu}v_I \varepsilon_0 - \frac{L_u}{\tau_{stim}} P_{pp} \right), \tag{S.85}
\]

\[
\varepsilon_1 = (P_{uu}P_{pp} - P_{up}P_{pu})^{-1} \left( P_{pu}S_{upp} \varepsilon_0^2 + P_{uu}S_{ppp} \varepsilon_0^2 - K_P P_{uu}v_I \varepsilon_0 + \frac{L_u}{\tau_{stim}} P_{pu} \right). \tag{S.86}
\]

Substituting into Eqs. (S.80) and (S.82) and collecting terms,

\[
-v_I = \frac{\varepsilon_0 Q_{up}}{I_u} + \frac{T_{upu} \varepsilon_0 + M_u v_I}{P_{uu}P_{pp} - P_{up}P_{pu}} \left( \frac{P_{up} S_{ppp} - P_{pp} S_{upp} \varepsilon_0^2}{2 I_u} + \frac{K_P P_{up} v_I \varepsilon_0}{I_u} - \frac{L_u}{\tau_{stim}} P_{pp} \right)
\]

\[
+ \frac{T_{upp} \varepsilon_0}{P_{uu}P_{pp} - P_{up}P_{pu}} \left( \frac{P_{pu} S_{upp} - P_{uu} S_{ppp} \varepsilon_0^2}{2 I_u} - \frac{K_P P_{uu} v_I \varepsilon_0}{I_u} + \frac{L_u}{\tau_{stim}} P_{pu} \right) + \frac{Q_{upp} \varepsilon_0^3}{6 I_u} + \frac{s}{\tau_{stim}}. \tag{S.87}
\]

\[
-v_I = \frac{\varepsilon_0 Q_{pp}}{I_p} + \frac{T_{ppu} \varepsilon_0}{P_{uu}P_{pp} - P_{up}P_{pu}} \left( \frac{P_{pp} S_{ppp} - P_{pp} S_{upp} \varepsilon_0^2}{2 I_p} + \frac{K_P P_{pp} v_I \varepsilon_0}{I_p} - \frac{L_u}{\tau_{stim}} P_{pp} \right)
\]

\[
+ \frac{T_{ppp} \varepsilon_0 + M_p v_I}{P_{uu}P_{pp} - P_{up}P_{pu}} \left( \frac{P_{pu} S_{upp} - P_{uu} S_{ppp} \varepsilon_0^2}{2 I_p} - \frac{K_P P_{uu} v_I \varepsilon_0}{I_p} + \frac{L_u}{\tau_{stim}} P_{pu} \right) + \frac{Q_{ppp} \varepsilon_0^3}{6 I_p} - 2 \tau {\varepsilon_0}. \tag{S.88}
\]

Eliminating \( v_I \),

\[
\frac{s}{\tau_{stim}} = \left( \frac{Q_{uu}}{I_u} + \frac{Q_{pp}}{I_p} \right) \varepsilon_0
\]

\[
- \frac{\varepsilon_0^3}{P_{uu}P_{pp} - P_{up}P_{pu}} \left( \frac{T_{upu}}{I_u} - \frac{T_{ppu}}{I_p} - \frac{M_u Q_{pp}}{I_u I_p} \right) \left( \frac{P_{up} S_{ppp} - P_{pp} S_{upp} \varepsilon_0^2}{2 I_u} - \frac{2}{I_p} \right)
\]

\[
+ \frac{\varepsilon_0^3}{P_{uu}P_{pp} - P_{up}P_{pu}} \left( \frac{T_{ppu}}{I_p} - \frac{T_{upp}}{I_u} - \frac{M_p Q_{pp}}{I_u I_p} \right) \left( \frac{P_{pp} S_{ppp} - P_{pp} S_{upp} \varepsilon_0^2}{2 I_p} + \frac{2}{I_u} \right)
\]

\[
+ \frac{Q_{ppp}}{6 I_p} - \frac{Q_{upp}}{6 I_u} \varepsilon_0^3 + \frac{L_u \varepsilon_0}{\tau_{stim}} \left( \frac{P_{pp} T_{upu}}{I_u} - \frac{P_{pu} T_{upp}}{I_u} - \frac{P_{pp} T_{ppu}}{I_p} \right)
\]

\[
+ \frac{P_{pp} T_{ppp}}{I_u} - \frac{M_u P_{pp} Q_{pp}}{I_u I_p} - \frac{M_p P_{pu} Q_{pp}}{I_p I_p^2}. \tag{S.89}
\]

Using Eq. (S.74), this equation simplifies to

\[
\frac{s}{\tau_{stim}} = \lambda \varepsilon_0 - K \varepsilon_0^3 + \frac{K_0 \varepsilon_0}{\tau_{stim}}, \tag{S.90}
\]

where

\[
K_0 = \frac{L_u}{P_{uu}P_{pp} - P_{up}P_{pu}} \left( \frac{P_{pp} T_{upu}}{I_u} - \frac{P_{pu} T_{upp}}{I_u} - \frac{P_{pp} T_{ppu}}{I_p} + \frac{P_{pp} T_{ppp}}{I_p} \right)
\]

\[
- \frac{M_u P_{pp} Q_{pp}}{I_u I_p} - \frac{M_p P_{pu} Q_{pp}}{I_p I_p^2}. \tag{S.91}
\]
To the lowest order, $\varepsilon_0$ is determined from Eq. (S.88) by

$$v_I \approx -\varepsilon_0 \frac{Q_{pp}}{I_p} = \frac{\varepsilon_0}{\tau_{\text{int}}}. \tag{S.92}$$

This leads to an expression for the anticipation time $\tau_{\text{ant}} \equiv s/v_I$ given by

$$\tau_{\text{ant}} = \tau_{\text{stim}} \tau_{\text{int}} (\lambda - K v_I^2 \tau_{\text{int}}^2) + K_0 \tau_{\text{int}}. \tag{S.93}$$

We first consider the limit of weak and slowly moving stimulus, in which $\tau_{\text{stim}}$ is large and $v_I$ is small. This reduces the anticipation time to the transparent form

$$\tau_{\text{ant}} = \tau_{\text{stim}} \tau_{\text{int}} \lambda. \tag{S.94}$$

Hence $\tau_{\text{ant}}$ and $\lambda$ have the same sign. In the static phase, $\lambda < 0$ implies that the tracking is delayed with $\tau_{\text{ant}} < 0$, whereas in the moving phase, $\lambda > 0$ implies that the tracking is anticipative with $\tau_{\text{ant}} > 0$. At the phase boundary, the system is in the ready-to-go state; here $\tau_{\text{ant}} = 0$ and the tracking is perfect.

Equation (S.94) further implies that the anticipation time is effectively independent of the stimulus velocity for sufficiently weak and slow stimuli. This is contrary to the intuition that the anticipation time decreases with stimulus speed. This can be interpreted by substituting Eq. (S.92) into Eq. (S.94), yielding

$$\tau_{\text{ant}} = \frac{\tau_{\text{stim}} \varepsilon_0 \lambda}{v_I}. \tag{S.95}$$

This shows that while the anticipation time decreases with increasing stimulus speed, the profile separation also increases, so that the tendency for spontaneous motion also increases proportionately.

Note that Eq. (S.94) is already a manifestation of the fluctuation-response relation, since it relates the anticipation time $\tau_{\text{ant}}$, as an extrinsic property, to the instability eigenvalue $\lambda$, as an intrinsic property. The close relation between extrinsic and intrinsic behaviors can be further illustrated by combining Eqs. (S.94), (S.92) and (S.73) to yield

$$\tau_{\text{ant}} = K \tau_{\text{stim}} \tau_{\text{int}}^3 \nu_{\text{nat}}. \tag{S.96}$$

This shows that the contours of contant $\tau_{\text{ant}}$ and those of contant $\nu_{\text{nat}}$ have a one-to-one correspondence, as shown in Fig. 2.
Next, we consider the case that the external stimulus moves with a finite velocity. In particular, we are interested in the moving phase in which the stimulus velocity is comparable to the natural velocity. Eliminating the eigenvalue $\lambda$ in Eq. (S.93) via Eqs. (S.73) and (S.75), we arrive at

$$\tau_{\text{ant}} = K \tau_{\text{stim}} \tau_{\text{int}}^3 \left( v_{\text{nat}}^2 - v_I^2 \right) + K_0 \tau_{\text{int}}.$$  \hspace{1cm} (S.97)

Note that when the network tracks a stimulus moving at the natural velocity $v_{\text{nat}}$, $\tau_{\text{ant}}$ becomes independent of the stimulus time $\tau_{\text{stim}}$. Hence when $\tau_{\text{ant}}$ is plotted as a function of $v_I$, the curves become confluent at the point $v_I = v_{\text{nat}}$ and $\tau_{\text{ant}} = \tau_{\text{con}} \equiv K_0 \tau_{\text{int}}$.

The origin of this independence of the stimulus amplitude can be traced by the following reasoning.

1. At natural tracking, the forces acting on the system are used to drive the system at the natural velocity.

2. If a stimulus is present and moves with a velocity different from the natural velocity, the extra force ($I_u s / \tau_{\text{stim}}$ in Eq. (S.80)) due to the stimulus can be used to drive the system to move at the stimulus velocity rather than the natural velocity. Since the extra force is proportional to the displacement, the displacement at the steady state will adjust itself so that the extra force is totally used to drive the system at the stimulus velocity, as shown in Eqs. (S.80)-(S.83). At the same time, a small amount of the extra force is used to distort the shape of the bump (mainly in its height and width), as shown by the higher order terms in Eqs. (S.81) and (S.83). This component of force is proportional to the stimulus amplitude. For weak stimuli, this component of force is negligible as long as the stimulus velocity is different from the natural velocity.

3. When the stimulus moves at the natural velocity, there is no need to use the extra force due to the stimulus to drive the bump to a different velocity. Hence to the lowest order, the extra force is not balanced. In this case, the extra force is solely used to distort the shape of the bump. Since both the force and the distortion are proportional to the stimulus amplitude, the displacement becomes independent of the stimulus amplitude.
S.8. RESPONSE TO NOISES

From the viewpoint of fluctuation-response relations, we would like to connect our results with thermal fluctuations. Hence we consider the dynamics in the presence of thermal noises by modifying Eq. (S.20),

$$\frac{\partial u(x)}{\partial t} = F_u [x; u, p] - \eta(t) \frac{\partial u_0}{\partial x}, \quad \frac{\partial p(x)}{\partial t} = F_p [x; u, p], \quad (S.98)$$

where

$$\langle \eta(t) \rangle = 0, \text{ and } \langle \eta(t) \eta(t') \rangle = 2T \delta(t - t'). \quad (S.99)$$

This implies that Eq. (S.24) has to be modified to

$$\frac{\partial}{\partial t} \delta u(x) = \int dx' \frac{\partial F_u(x)}{\partial u(x')} \delta u(x') + \int dx' \frac{\partial F_u(x)}{\partial p(x')} \delta p(x') - \eta(t) \frac{\partial u_0}{\partial x}. \quad (S.100)$$

We first consider the behavior in the static phase. Following the analysis in Sec. S.4, we arrive at

$$\frac{\partial}{\partial t} \begin{pmatrix} c_0 \\ \varepsilon_0 \end{pmatrix} = \begin{pmatrix} Q_{uu}/I_u & -Q_{uu}/I_u \\ -Q_{pp}/I_p & Q_{pp}/I_p \end{pmatrix} \begin{pmatrix} c_0 \\ \varepsilon_0 \end{pmatrix} - \begin{pmatrix} \eta(t) \\ 0 \end{pmatrix}. \quad (S.101)$$

This implies

$$\frac{\partial}{\partial t} (\varepsilon_0 - c_0) = \lambda (\varepsilon_0 - c_0) + \eta(t). \quad (S.102)$$

The solution to this differential equation is

$$\varepsilon_0 - c_0 = \int_{-\infty}^{t} dt' \exp \left\{ \lambda (t - t') \right\} \eta(t'). \quad (S.103)$$

Averaging over thermal noises, \( \langle \varepsilon_0 - c_0 \rangle = 0 \) and

$$\langle (\varepsilon_0 - c_0)^2 \rangle = \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t} dt_2 \exp \{ \lambda [(t - t_1) + (t - t_2)] \} \langle \eta(t_1) \eta(t_2) \rangle. \quad (S.104)$$

Using the noise average in Eq. (S.99),

$$\langle (\varepsilon_0 - c_0)^2 \rangle = 2T \int_{-\infty}^{t} dt' \exp \left\{ 2\lambda (t - t') \right\} = -\frac{T}{\lambda}. \quad (S.105)$$

Equation (S.93) can now be cast into the form of a fluctuation response relation. In this case, the response term is the effective anticipation rate, that is, the inverse of the anticipation time minus its value at the confluence point,

$$\frac{\langle (\varepsilon_0 - c_0)^2 \rangle}{T} = -\frac{\tau_{stim}}{\tau_{ant} - \tau_{con}}. \quad (S.106)$$
This shows that the effective anticipation time in the static phase is negative. The relation means that when the fluctuations of the separation between the exposed and inhibition profiles has a faster rate of increase with the temperature, the network becomes more responsive to the moving stimulus by shortening the delay time. At the boundary of the static phase, fluctuations diverge and the bump is in a ready-to-go state.

Next, we consider the behavior in the moving phase. We choose the eigenfunctions to be centered at the instantaneous position of the bump, whose motion is the sum of the spontaneous movement as well as thermal noises. Hence the dynamical variables relevant to the third order expansion are \( \delta v, \varepsilon_0, c_1 \) and \( \varepsilon_1 \), where \( \delta u = u_1 c_1 \) and \( \delta p = (\partial p_0/\partial x)\varepsilon_0 + p_1 \varepsilon_1 \). Following the analysis in Secs. (S.5) and (S.6), the dynamical equations become

\[
-I_u \delta v - M_u v_{nat} \delta c_1 - M_u c_1 \delta v = Q_{up} \varepsilon_0 + T_{upu} \varepsilon_0 c_1 + T_{upp} \varepsilon_0 \varepsilon_1 + \frac{Q_{uppp}}{6} \varepsilon_0^3 - I_u \eta. \tag{S.107}
\]

\[
J_u \frac{\partial \varepsilon_0}{\partial t} = P_{uu} c_1 + P_{up} \delta \varepsilon_1 + S_{upp} \varepsilon_0^2 \delta \varepsilon_0. \tag{S.108}
\]

\[
I_p \frac{\partial \varepsilon_0}{\partial t} - I_p v_{nat} \varepsilon_1 - M_p v_{nat} \varepsilon_1 = Q_{pp} \varepsilon_0 + T_{ppu} \varepsilon_0 c_1 + T_{ppp} \varepsilon_0 \varepsilon_1 + \frac{Q_{pppp}}{6} \varepsilon_0^3. \tag{S.109}
\]

\[
J_p \frac{\partial \varepsilon_1}{\partial t} - K_p v_{nat} \varepsilon_0 = P_{pu} c_1 + P_{pp} \delta \varepsilon_1 + S_{ppp} \varepsilon_0^2. \tag{S.110}
\]

Now consider fluctuations around the noiseless solution of the spontaneously moving bump, obtained from Eqs. (S.55), (S.64), (S.65) and (S.66). Subtracting the noiseless equations, Eqs. (S.107)-(S.110) become

\[
-I_u \delta v - M_u v_{nat} \delta c_1 - M_u c_1 \delta v = Q_{up} \delta \varepsilon_0 + T_{upu} \delta \varepsilon_0 c_1 + T_{upu} \varepsilon_0 \delta c_1 + T_{upp} \varepsilon_1 \delta \varepsilon_0 + T_{upp} \varepsilon_0 \delta \varepsilon_1 \\
+ \frac{Q_{uppp}}{2} \varepsilon_0^2 \delta \varepsilon_0 - I_u \eta. \tag{S.111}
\]

\[
J_u \frac{\partial \delta c_1}{\partial t} = P_{uu} \delta c_1 + P_{up} \delta \varepsilon_1 + S_{upp} \varepsilon_0 \delta \varepsilon_0. \tag{S.112}
\]

\[
I_p \frac{\partial \varepsilon_0}{\partial t} - I_p \delta v - M_p v_{nat} \delta \varepsilon_1 - M_p \varepsilon_1 \delta v = Q_{pp} \delta \varepsilon_0 + T_{ppu} \delta \varepsilon_0 c_1 + T_{ppu} \varepsilon_0 \delta c_1 \\
+ T_{ppp} \varepsilon_1 \delta \varepsilon_0 + T_{ppp} \varepsilon_0 \delta \varepsilon_1 + \frac{Q_{pppp}}{2} \varepsilon_0^2 \delta \varepsilon_0. \tag{S.113}
\]

\[
J_p \frac{\partial \delta \varepsilon_1}{\partial t} - K_p v_{nat} \delta \varepsilon_0 - K_p \varepsilon_0 \delta v = P_{pu} \delta c_1 + P_{pp} \delta \varepsilon_1 + S_{ppp} \varepsilon_0 \delta \varepsilon_0. \tag{S.114}
\]
The expression for $\delta v$ can be obtained from Eq. (S.111),

$$
\delta v = \eta - \frac{Q_{up} + T_{upu}c_1 + T_{upp}^2 + Q_{uppp}^2/2}{I_u + M_u c_1} \delta v_0 - \frac{T_{upu}c_0 + M_u v_{nat}}{I_u} \delta c_1 - \frac{T_{upp}^2}{I_u} \delta \varepsilon_1.
$$

(S.115)

Eliminating $\delta v$ in Eq. (S.113),

$$
\frac{\partial \delta \varepsilon_0}{\partial t} = \eta + \left( \frac{Q_{pp} + T_{ppu}c_1 + T_{ppp}^2 + Q_{pppp}^2/2}{I_p + M_p \varepsilon_0} \right) \delta \varepsilon_0 + \left( \frac{T_{ppu}c_0 + M_p v_{nat}}{I_p} \right) \delta c_1 + \left( \frac{T_{ppp}c_0 + M_p v_{nat}}{I_p} - \frac{T_{upp}^2}{I_p} \right) \delta \varepsilon_1.
$$

(S.116)

Eliminating $\delta v$ in Eq. (S.113), we arrive at

$$
\frac{Q_{up} + T_{upu}c_1 + T_{upp}^2 + Q_{uppp}^2/2}{I_u + M_u c_1} = \frac{Q_{pp} + T_{ppu}c_1 + T_{ppp}^2 + Q_{pppp}^2/2}{I_p + M_p \varepsilon_0}.
$$

Together with Eq. (S.72), this simplifies Eq. (S.116) to

$$
\frac{\partial \delta \varepsilon_0}{\partial t} = \eta + \left( \frac{Q_{pppp}}{I_p} - \frac{Q_{uppp}}{I_u} \right) \frac{\varepsilon_0^3}{3} \delta \varepsilon_0 + \left( \frac{T_{ppu}}{I_p} - \frac{T_{upu}}{I_u} + \frac{M_u Q_{pp}}{I_u I_p} \right) \varepsilon_0 \delta c_1 + \left( \frac{T_{ppp}}{I_p} - \frac{T_{upp}}{I_p} - \frac{M_p Q_{pp}}{I_p^2} \right) \varepsilon_0 \delta \varepsilon_1.
$$

(S.117)

Substituting $\delta v$ into Eq. (S.114), using Eq. (S.72), and retaining the leading terms,

$$
\frac{\partial \delta \varepsilon_1}{\partial t} = \left( \frac{S_{ppp}}{J_p} - \frac{K_p Q_{pp}}{I_p J_p} - \frac{K_p Q_{up}}{I_u I_p} \right) \varepsilon_0 \delta \varepsilon_0 + \frac{P_{pp}}{J_p} \delta c_1 + \frac{P_{pp}}{J_p} \delta \varepsilon_1 + \frac{K_p}{J_p} \varepsilon_0 \eta.
$$

(S.118)

Combining Eqs. (S.112), (S.117) and (S.118), we arrive at the matrix equation

$$
\frac{\partial}{\partial t} \begin{pmatrix} \delta \varepsilon_0 \\ \delta c_1 \\ \delta \varepsilon_1 \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix} \begin{pmatrix} \delta \varepsilon_0 \\ \delta c_1 \\ \delta \varepsilon_1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ K_p \varepsilon_0 \end{pmatrix} \eta.
$$

(S.119)

The matrix $M$ is given by

$$
\begin{pmatrix}
\frac{Q_{pppp}}{I_p} - \frac{Q_{uppp}}{I_u} & \varepsilon_0^3/3 & \left( \frac{T_{ppu}}{I_p} - \frac{T_{upu}}{I_u} + \frac{M_u Q_{pp}}{I_u I_p} \right) \\
S_{pp} \varepsilon_0 & \frac{P_{pp}}{J_p} & \frac{P_{pp}}{J_p} \\
\left( \frac{S_{pp}}{J_p} - \frac{K_p Q_{pp}}{I_p J_p} \right) \varepsilon_0 & \frac{P_{pp}}{J_p} & \frac{P_{pp}}{J_p} \end{pmatrix}
$$

(S.120)
Let $-\lambda_m$ be the largest eigenvalue of $M$, which is expected to be of the order $\varepsilon_0^2$. It is given by

$$
\frac{\lambda_m}{\varepsilon_0^2} = \frac{Q_{uupp}}{3I_u} - \frac{Q_{pppp}}{3I_p} + (P_{uu}P_{pp} - P_{up}P_{pu})^{-1}
\times \left[ \left( \frac{T_{ppu}}{I_p} - \frac{T_{upu}}{I_u} + \frac{M_u Q_{pp}}{I_u I_p} \right) \left( P_{pp}S_{uupp} - P_{up}S_{ppp} + \frac{K_p P_{up} Q_{pp}}{I_p} + \frac{K_p P_{uu} Q_{pp}}{I_u} \right) 
- \left( \frac{T_{ppp}}{I_p} - \frac{T_{upp}}{I_u} - \frac{M_p Q_{pp}}{I_p^2} \right) \left( P_{pu}S_{uupp} - P_{uu}S_{ppp} + \frac{K_p P_{uu} Q_{pp}}{I_p} + \frac{K_p P_{uu} Q_{up}}{I_u} \right) \right].
$$

(S.121)

Note that near the onset of the moving phase, $Q_{up}/I_u \approx Q_{pp}/I_p$. Comparing with Eq. (S.74),

$$\lambda_m \varepsilon_0^2 = 2K. \quad (S.122)$$

If the moving bump is stable, then the eigenvalue $-\lambda_m$ is non-positive. The corresponding right and left eigenvectors are $(1, O(\varepsilon_0), O(\varepsilon_0))^T$ and $(1, O(\varepsilon_0), O(\varepsilon_0))$ respectively.

The second eigenvalue is $-\lambda_1 + O(\varepsilon_0^2)$, with the corresponding right and left eigenvectors $(O(\varepsilon_0), 1, 1)^T$ and $(O(\varepsilon_0), P_{uu}/J_u, P_{up}/J_u)$ respectively. The third eigenvalue is $\lambda_1 + P_{uu}/J_u + P_{pp}/J_p + O(\varepsilon_0^2)$, with the corresponding right and left eigenvectors $(O(\varepsilon_0), P_{up}/J_u, -P_{pu}/J_p)^T$ and $(O(\varepsilon_0), 1, -1)$ respectively. Hence

$$\delta \varepsilon_0(t) = \int_{-\infty}^{t} dt' \exp \left[ -\lambda_m (t - t') \right] \left[ 1 + \frac{K_p}{J_p} O(\varepsilon_0^2) \right] \eta(t'). \quad (S.123)$$

Near the onset of the moving phase, $\varepsilon_0$ is negligible. The noise amplitude is given by

$$\langle \delta \varepsilon_0^2 \rangle = \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t} dt_2 \exp \left\{ -\lambda_m \left[ (t - t_1) + (t - t_2) \right] \right\} \langle \eta(t_1) \eta(t_2) \rangle
= 2T \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t} dt_2 \delta(t_1 - t_2) = \frac{T}{\lambda_m}. \quad (S.124)$$

Combining with Eqs. (S.122) and (S.73),

$$\frac{\langle \delta \varepsilon_0^2 \rangle}{T} = \frac{1}{2\lambda}. \quad (S.125)$$

Finally, from Eq. (S.93), we can write the fluctuation response relation

$$\frac{\langle \delta \varepsilon_0^2 \rangle}{T} = \frac{\tau_{stim} \tau_{int}}{2(\tau_{ant} - \tau_{con})}. \quad (S.126)$$
A. Numerical Measurement of $\langle \delta \varepsilon_0^2 \rangle$

The variance of $\varepsilon_0(t)$ can be easily obtained from simulations, if the set of parameters is chosen to be far from phase boundaries. Those examples for CANNs with STD are shown in Fig. S.3(a) and (c). In Fig. S.3(a), $\tilde{\beta}$ is small enough to have a stable static fixed point solution. In this case, there is only one fixed point solution of $\varepsilon_0 = 0$. The statistics of $\varepsilon_0(t)$ is relatively simple. For a large enough $\tilde{\beta}$, as shown in Fig. S.3(c), the two fixed point solutions to $\varepsilon_0$ have opposite signs and are separated far apart. As a result, $\varepsilon_0(t)$ will mostly stick to one of the fixed point solution. The statistics of $\varepsilon_0(t)$ is similar to that of the static phase.

However, in the moving phase near the phase boundary, e.g. Fig. S.3(b), the statistics may be problematic. The problem is due to the difference between two fixed point solutions being too small, so that $\varepsilon_0(t)$ is fluctuating around two fixed point solutions ($\varepsilon_0^{+\text{fixed}}$ and $\varepsilon_0^{-\text{fixed}}$), even though the noise temperature $T$ is small. Whenever $\varepsilon_0(t)$ is between two fixed point solutions, attractions due to fixed point solutions can affect our estimations of the variance of $\varepsilon_0(t)$ around a single fixed point solution.

To overcome the interference between two fixed point solutions, a trick is needed to filter out some data. In the statistics of Fig. 4 in the main text, we have discarded $\varepsilon_0(t)$ less than $|\varepsilon_0^{\pm\text{fixed}}|$. So, we approximate the variance by

$$\text{Var} \left[ \varepsilon_0(t) - \varepsilon_0^{\pm\text{fixed}} \right] = \sum_{t' \in S} \frac{\left| \varepsilon_0(t') - |\varepsilon_0^{\pm\text{fixed}}| \right|^2}{N_{\text{sample}} - 1},$$

(S.127)

where $S \equiv \{ t' \mid |\varepsilon_0(t')| > |\varepsilon_0^{\pm\text{fixed}}| \}$ and $N_{\text{sample}} \equiv |S|$.

\[ \[ \]

1. A. Einstein, Ann. Phys. (Berlin) 322(8), 549–560 (1905).
2. M. von Smoluchowski, Ann. Phys. (Berlin) 326(14), 756–780 (1906).
3. H. Nyquist, Phys. Rev. 32, 110 (1928).
4. K. Huang, Statistical Mechanics (Wiley, New York, 1987).
5. R. Ben-Yishai, R. L. Bar-Or, and H. Sompolinsky, H. Proc. Natl. Acad. Sci. U.S.A. 92, 3844 (1995).
6. K. Zhang, The Journal of Neuroscience 16, 2112–2126 (1996).
7. A. Samsonovich and B. L. McNaughton, J. Neurosci. 17, 5900–5920 (1997).
8. H. Wilson and J. Cowan, Biophysical Journal 12, 1 (1972).
9. S. Amari, S, Biological Cybernetics 27, 77 (1977).
[10] R. Nijhawan and S. Wu, S. Phil. Trans. R. Soc. A 367, 1063 (2009).
[11] J. O’Keefe and J. Dostrovsky, J. Brain Res. 34, 171 (1971).
[12] J. S. Taube, R. U. Muller, and J. B. Ranck Jr., J. Neurosci. 10, 420 (1990).
[13] J. S. Taube and R. I. Muller, Hippocampus 8, 87 (1998).
[14] H. Y. Blair and P. E. Sharp, J. Neurosci. 15, 6260 (1995).
[15] M. A. Sommer and R. H. Wurtz, Nature 444, 374 (2006).
[16] R. Nijhawan, Nature (370), 256 (1994).
[17] C. C. A. Fung, K. Y. M. Wong and S. Wu, Advances in Neural Information Processing Systems 25, 1097 (2013).
[18] R. Ben-Yishai, D. Hansel and H. Sompolinsky, J. Comp. Neurosci. 4, 57 (1997).
[19] L. C. York and M. C. Van Rossum, J. Comput. Neurosci. 27, 607 (2009).
[20] C. C. A. Fung, K. Y. M. Wong, H. Wang and S. Wu, Neural Comput. 24, 1147 (2012).
[21] A., Treves, Network: Comput. in Neural Systems 4, 259 (1993).
[22] W. Zhang and S. Wu, Neural Comput. 24, 1695 (2012).
[23] C. C. A. Fung, K. Y. M. Wong and S. Wu, Neural Comput. 22, 752 (2010).
[24] M. V. Tsodyks and H. Markram, Proc. Natl. Acad. Sci. U.S.A. 94, 719-723 (1997).
[25] S. Coombes and M. R. Owen, Phys. Rev. Lett. 94, 148102 (2005).
[26] J. P. Goodridge and D. S. Touretzky, J. Neurophysio. 83, 3402 (2000).
[27] Y. Mi, Y. Xia, C. C. A. Fung, S. Amari, K. Y. M. Wong and S. Wu (Unpublished).
Figure S.3: Three samples of $\varepsilon_0(t)$ with different values of $\tilde{\beta}$. (a) $\varepsilon_0(t)$ in the static phase. (b) $\varepsilon_0(t)$ in the moving phase near the static-moving transition. (c) $\varepsilon_0(t)$ in the moving phase. Parameters: $\tilde{k} = 0.3$, $\tau_d = 50\tau_s$ and $T = 1 \times 10^{-6}$.