

## Methods

### *Equilibria and stability analysis of the LDDM*

In **Figs. 3** and **5**, we showed that the LDDM exhibits different pattern of equilibria and stabilities under normalized value coding and WTA competition, mediated through disinhibition. Here we provide detailed mathematical analysis about the equilibria and stability of this dynamic system under different states of disinhibition.

Equilibria of the system were solved by taking the intersection of the nullclines of all units, i.e., the steady states of each unit. This is obtained by setting  $dR_i/dt$ ,  $dG_i/dt$ , and  $dD_i/dt$  all equal to 0 in Eqs. 1-3. The solution of the equilibrium state of  $R$  units ( $R_i^*$ ) can be written as:

$$R_i^* = \frac{V_i + B}{1 + G_0 - \alpha + (\omega - \beta)R_i^* + \omega \sum_{j \neq i} R_j^*} \quad (5)$$

For a binary input system ( $N = 2$ ), the six differential equations can be simplified to two equations with only the  $R$  units explicitly in the expression (Eqs. 6 and 7). Each equation describes the nullcline of a single  $R$  unit.

$$\begin{cases} \frac{V_1 + B}{R_1^*} - (\omega - \beta)R_1^* - (1 + G_0 - \alpha) = \omega R_2^* & (6) \\ \frac{V_2 + B}{R_2^*} - (\omega - \beta)R_2^* - (1 + G_0 - \alpha) = \omega R_1^* & (7) \end{cases}$$

Given that the equilibrium states of the system can be reduced with only  $R$  units explicitly in the expression, these equilibrium points can be visualized in the  $\mathbb{R}_+^2$  space of  $R_1$  and  $R_2$  activities as the intersection of the nullclines of the two  $R$  units (as shown in **Fig. 3** and **Fig. S2**). The stability of each equilibrium point was then examined by checking the eigenvalues of the Jacobian matrix around it. The equilibrium point is attractive and stable when all of the eigenvalues have negative real parts; the equilibrium point is divergent and unstable when there exist any positive real parts of eigenvalues. By denoting  $\mathbf{F} = (F_{R_1}, F_{G_1}, F_{D_1}, F_{R_2}, F_{G_2}, F_{D_2})$  as the differential equations for all units in their steady states, the Jacobian matrix around the point can be written as Eq. 8:

$$J = \begin{bmatrix} \frac{\partial F_{R_1}}{\partial R_1} & \frac{\partial F_{R_1}}{\partial G_1} & \frac{\partial F_{R_1}}{\partial D_1} & \frac{\partial F_{R_1}}{\partial R_2} & \frac{\partial F_{R_1}}{\partial G_2} & \frac{\partial F_{R_1}}{\partial D_2} \\ \frac{\partial F_{G_1}}{\partial R_1} & \frac{\partial F_{G_1}}{\partial G_1} & \frac{\partial F_{G_1}}{\partial D_1} & \frac{\partial F_{G_1}}{\partial R_2} & \frac{\partial F_{G_1}}{\partial G_2} & \frac{\partial F_{G_1}}{\partial D_2} \\ \frac{\partial F_{D_1}}{\partial R_1} & \frac{\partial F_{D_1}}{\partial G_1} & \frac{\partial F_{D_1}}{\partial D_1} & \frac{\partial F_{D_1}}{\partial R_2} & \frac{\partial F_{D_1}}{\partial G_2} & \frac{\partial F_{D_1}}{\partial D_2} \\ \frac{\partial F_{R_2}}{\partial R_1} & \frac{\partial F_{R_2}}{\partial G_1} & \frac{\partial F_{R_2}}{\partial D_1} & \frac{\partial F_{R_2}}{\partial R_2} & \frac{\partial F_{R_2}}{\partial G_2} & \frac{\partial F_{R_2}}{\partial D_2} \\ \frac{\partial F_{G_2}}{\partial R_1} & \frac{\partial F_{G_2}}{\partial G_1} & \frac{\partial F_{G_2}}{\partial D_1} & \frac{\partial F_{G_2}}{\partial R_2} & \frac{\partial F_{G_2}}{\partial G_2} & \frac{\partial F_{G_2}}{\partial D_2} \\ \frac{\partial F_{D_2}}{\partial R_1} & \frac{\partial F_{D_2}}{\partial G_1} & \frac{\partial F_{D_2}}{\partial D_1} & \frac{\partial F_{D_2}}{\partial R_2} & \frac{\partial F_{D_2}}{\partial G_2} & \frac{\partial F_{D_2}}{\partial D_2} \end{bmatrix} \quad (8)$$

$$= \begin{bmatrix} -1 + \frac{\alpha}{1 + G_1^*} & -\frac{V_1 + B + \alpha R_1^*}{(1 + G_1^*)^2} & 0 & 0 & 0 & 0 \\ \omega & -1 & -1 & \omega & 0 & 0 \\ \beta & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 + \frac{\alpha}{1 + G_2^*} & -\frac{V_2 + B + \alpha R_2^*}{(1 + G_2^*)^2} & 0 \\ \omega & 0 & 0 & \omega & -1 & -1 \\ 0 & 0 & 0 & \beta & 0 & -1 \end{bmatrix}$$

We examined the configuration of nullclines and checked the eigenvalues of the Jacobian matrix across a wide range of parameter values  $\alpha$  and  $\beta$ .  $\omega$  was set as a unit value of 1 for the sake of simplicity.  $B$  and  $G_0$  were set as zero in the following visualization.

The property of the system under equivalent inputs is a critical test since it determines whether the system is able to implement a WTA choice and select an option. Thus, we examined the property of the system for WTA under equal inputs. Examining the full space of  $\alpha$  and  $\beta$  revealed five territories distinguished by the number of equilibrium points and their stabilities (**Figure 5-figure supplement 1A**). For each territory, the configuration of nullclines are illustrated in **Figure 5-figure supplement 1** labeled by color. **Dark green region:** When disinhibition is smaller ( $\beta < 1$ ),  $\alpha$  and  $\beta$  show a trade-off in generating WTA competition. When both  $\alpha$  and  $\beta$  are small, the system generates a unique equilibrium point of normalized coding (dark green region in **Figure 5-figure supplement 1A**, nullclines shown in **Figure 5-figure supplement 1B**). Eigenvalues in this regime show all negative real parts on this equilibrium point, indicating it is a stable equilibrium. **Blue region:** As  $\alpha$  values increase (at smaller  $\beta$  values), the system generates three equilibrium points (**Figure 5-figure supplement 1D**), with two high-contrast (stable) attractors at the peripheral and one (unstable) repeller in the center of space  $R_1$ - $R_2$ . Neural activities of  $R_1$  and  $R_2$  with equal initial values bifurcate into the high-contrast attractors to realize WTA competition (example traces shown in red and blue lines). **Green region:** When the strength of disinhibition increases ( $\beta > 1$ ), most of the regimes (yellow and red regions) show the properties of WTA competition except for a small regime when  $\alpha < 1 + G_0$  (green region). In the green region, the nullclines of  $R_1$  and  $R_2$  still intersect on three equilibrium points but, in contrast to the blue region, the two points with high contrast of  $R_1$ - $R_2$  activities are unstable and the equilibrium point in the center is stable (**Figure 5-figure supplement 1C**). The neural activity of  $R_1$  and  $R_2$  is restricted under a value of  $\frac{V_i+B}{1+G_0-\alpha}$ , which is lower than the high-contrast equilibria, therefore, the system maintains normalized coding. **Yellow region:** When disinhibition is large ( $\beta > 1$ ), most of the

parameter regime in the yellow region shows only one repellor at the center (**Figure 5-figure supplement 1E**). The activities of  $R_1$  and  $R_2$  bifurcate from the center repellor to the high-contrast corners. The restriction of maximum activity depends on the value of  $\alpha$ . When  $\alpha < 1 + G_0$ , the model predicts limited value of activity on each  $R$  unit as  $(\frac{V_i+B}{1+G_0-\alpha})$  (vertical and horizontal dashed lines in **Figure 5-figure supplement 1E**). When  $\alpha \geq 1 + G_0$ , the model predicts no boundary on the maximum activities (though a boundary may still need to be considered because of biological constraints). **Red region:** When disinhibition is extremely large ( $\beta > 2$ ), the two nullclines show no intersections (**Figure 5-figure supplement 1F**). Most of the other features in this region are similar to the yellow region. The neural activities of  $R_1$  and  $R_2$  bifurcate from initial values from the center to the corners of high contrast (example traces shown in red and green thin lines). The boundary of neural activity is predicted when  $\alpha < 1 + G_0$  and not accounted when  $\alpha \geq 1 + G_0$ .

Taken together, the five territories can be simplified to two regions based on the properties of the system in implementing either normalized coding or WTA competition as discussed in the main text (**Fig. 5E**). These two regions show clear-cut dichotomous separation in the two-dimensional space of recurrent excitation weight ( $\alpha$ ) and local disinhibition weight ( $\beta$ ).

### *Numerical simulations*

To quantify neural dynamics and behavioral performance (choice/RT), time varying activity was represented by a system of differential equations (Eqs. 1-3) which was solved numerically using the Runge-Kutta method implemented in MATLAB (MathWorks) at time step of 1 ms. Evaluations using smaller time steps (0.1 ms) were examined and produced similar results. At each

time step, the model unit activities were updated based on their values at the previous step according to the differential equations. Considering the biological reality that spike rates cannot be negative, the activities were constrained to be non-negative. For the simulations including noise, we assumed an additive noise term for each unit, which evolved independently based on an Ornstein-Uhlenbeck process (Eq. 9),

$$\tau_{noise} \frac{dNoise(t)}{dt} = -Noise(t) + \eta(t) \sqrt{\sigma^2 dt} \quad (9)$$

where  $\sigma^2$  is the variance of the noise,  $\eta$  is a Gaussian white noise with zero mean and unit variance, and  $\tau_{noise}$  is the time constant for the noise fluctuation process. The time constant for the noise process ( $\tau_{noise}$ ) was set to 2 ms aligned with previous studies (Wang, 2002; Wong & Wang, 2006). Note that this approach assumes for convenience that noise arises in model unit activity; however, similar stochasticity can be implemented assuming noise arises in inputs external to the circuit, generalizing our findings.

All parameters used for visualization were set as the following unless specified elsewhere or fitted as free parameters in **Fig. 6**:  $\tau_R$ ,  $\tau_G$ , and  $\tau_D$  were set as the same value of 100 ms only for non-quantitative visualization purposes and fitted independently as free parameters in the model fittings; the gain control weight  $\omega_{ij}$  was set as a unit value of 1 for simplicity; the self-excitation weight  $\alpha$  was set as 15; the disinhibition weight  $\beta$  was assumed as zero in representation (i.e.,  $\beta_{off} = 0$ ) and set as 1.1 in WTA competition; the input values  $V_1$  and  $V_2$  were set as  $S^*(1+c')$  and  $S^*(1-c')$ , where  $c'$  indicates the motion coherence of the stimulus, with varied values [0, 3.2, 6.4, 12.8, 25.6, 38.4, 51.2]%) and  $S$  indicates the scale of input (set as 250). Baseline input  $B$  was set as 70 in **Figs. 3 and 5** and set as 0 in **Figs. 7-9**. Ornstein-Uhlenbeck noise was set as zero in most of the simulation but  $\sigma = 10$  in **Fig. 5B**. Parameters in **Fig. 7** were adjusted to predict the multi-alternative choice

data:  $\alpha$  was set as 0;  $\beta$  was set as 1.5; scaling parameter was set as 640 for both pre-motion and motion period but set as 427 for the first 190 ms of motion period to replicate the initial dip; all parameters were kept the same between 2 and 4-alternative choices. Parameters in **Fig. 8** were adjusted between 2-item and 5-item cases in order to get comparable scale of activities in visualization: for 2-item case,  $S = 250$ ,  $\alpha = 15$ , and  $\beta_{on} = .4$ ; for 5-item case,  $S = 50$ ,  $\alpha = 37.5$ , and  $\beta_{on} = .1$ .

### *Fitting the LDDM and the DNM to the neural firing rates of normalized value coding*

In order to quantify the performance of the LDDM in fitting to the neural dynamics of normalized value coding and compare with the original DNM, we fit the equilibrium values of the LDDM and DNM to the dataset of normalized value coding (the Fig. 4 in Louie et al., 2011). In this task, monkeys are asked to represent the reward targets (1, 2, or 3) on the corresponding location of the screen. The neural activity in the response field receiving direct input  $V_I$  is recorded. Different combinations of  $V_I$ ,  $V_2$ , and  $V_3$  are provided to the monkeys based on the associated volumes of juice in the presented targets (varying from 50, 100, 200, and 250  $\mu$ l or omitted target marked as 0), resulting in 28 data points.

To fit the DNM, we employed the following differential equations,

$$\begin{cases} \tau_R \frac{dR_i}{dt} = -R_i + \frac{V_i + B}{1 + G_i}; \\ \tau_G \frac{dG_i}{dt} = -G_i + \sum_j \omega R_j. \end{cases} \quad (10)$$

$$(11)$$

To fit LDDM, we employed Eq. 1-3.

The direct input value ( $V_i$ ) to each pool takes the value of the volume of juice reward ( $\mu l$ ) plus a baseline input value  $B$ .  $\omega$  was set as 1. In the LDDM, there are additional terms of self-excitation weighted by  $\alpha$ , baseline gain control input  $G_0$  fed into  $G_i$ , and coupling between  $R_i$  and the disinhibitory neurons  $D_i$  weighted by  $\beta$ .

To fit the predicted activities to the empirical mean firing rates during the sustain phase, we fit the predicted activities during the equilibria of these models. The equilibria of the two models were solved in Eq. 12 and Eq. 13 respectively by taking the differential equations (Eqs. 10-11 and Eqs. 1-3) to zero.

For DNM,

$$R_i = \frac{V_i + B}{1 + \sum_j R_j} \quad (12)$$

For LDDM,

$$R_i = \frac{V_i + B + \alpha R_i}{1 + G_0 + \sum_j R_j - \beta R_i} \quad (13)$$

To fit the empirical activities with normalized scale, we need another scaling parameter  $R_{max}$  to capture the arbitrary rescaling, which result in the following equations (Eq. 14 and Eq. 15).

For DNM,

$$R_i = R_{max} \frac{V_i + B}{1 + \sum_j R_j} \quad (14)$$

For LDDM,

$$R_i = R_{max} \frac{V_i + B + \alpha R_i}{1 + G_0 + \sum_j R_j - \beta R_i} \quad (15)$$

Since we assume the disinhibition modules in LDDM keep silent during representation,  $\beta$  takes zero. For a trinary input system, the equilibria of  $R_i$  the two models can be described by the following equations (Eqs. 16-18 for DNM and Eqs. 19-21 for LDDM).

For DNM,

$$\begin{cases} R_1 + R_2 + R_3 - \frac{R_{max}(V_1 + B)}{R_1} = 1; & (16) \\ R_2 + R_1 + R_3 - \frac{R_{max}(V_2 + B)}{R_2} = 1; & (17) \\ R_3 + R_1 + R_2 - \frac{R_{max}(V_3 + B)}{R_3} = 1. & (18) \end{cases}$$

For LDDM,

$$\begin{cases} R_1 + R_2 + R_3 - \frac{R_{max}(V_1 + B)}{R_1} = 1 + G_0 - \alpha; & (19) \\ R_2 + R_1 + R_3 - \frac{R_{max}(V_2 + B)}{R_2} = 1 + G_0 - \alpha; & (20) \\ R_3 + R_1 + R_2 - \frac{R_{max}(V_3 + B)}{R_3} = 1 + G_0 - \alpha. & (21) \end{cases}$$

From Eqs. 19-21, we realized that  $\alpha$  and  $G_0$  share the same term and cannot be independently identified. Thus, we combined these parameters as one in our model fitting.

Based on the above analyses, two free parameters were estimated for the DNM (Baseline input  $B$  and the scaling parameter  $R_{max}$ ). Three free parameters were estimated for the LDDM ( $B$ ,  $S$ , and a combined parameter  $G_0 - \alpha$ ). The Bayesian adaptive direct search algorithm (BADS) (Acerbi & Ma, 2017a, 2017b) was implemented to minimize the ordinary squared error between the steady state of the predicted neural firing rates on  $R_i$  and the empirical data.



### *Fitting the RNM to the neural firing rates of normalized value coding*

In order to quantify the performance of the RNM in predicting normalized value coding, we fit the reduced form of RNM (Wong & Wang, 2006) with 4 free parameters ( $JN_{i,i,i}$ ,  $JN_{i,j,k(i \neq j \neq k)}$ ,  $I_0$ , and a scaling parameter  $S$  applied to the predicted neural firing rates) to a normalized value coding dataset (Fig. 4 in Louie et al., 2011). Other parameters are set the same as reported in the original paper (Wong & Wang, 2006), except that the noise term  $\sigma$  is set as zero. The RNM is expanded to a trinary choice circuit, with three selective populations wired together based on the same rules specified in the original paper (Wong & Wang, 2006). We study the predicted neural activity on the pool 1 that receives direct input from  $V_1$  and investigate how the activity of pool 1 changes with the values of contextual inputs  $V_2$  and  $V_3$ . The BADS algorithm was used to minimize the mean squared error between the predicted neural firing rates of pool 1 and the empirical neural firing rates data reported in Fig. 4 of Louie et al., 2011. The best-fitting result shows that the RNM explains 89.2% of the variance, worse than the DNM and LDDM we reported in the main text (Best-fitting parameters:  $JN_{i,i,i} = .0055$ ,  $JN_{i,j,k(i \neq j, i \neq k)} = .0861$ ,  $I_0 = .3511$ ,  $S = 1.074$ ).

### *Fitting the LDDM to empirical behavioral data*

The LDDM with seven free parameters (the weights of self-excitation ( $\alpha$ ) and disinhibition ( $\beta$ ), the variance of Gaussian white noise in the Ornstein-Uhlenbeck process ( $\sigma^2$ ), the scaling parameter of input ( $S$ ), and time constants for three types of units  $\tau_R$ ,  $\tau_G$ , and  $\tau_D$ ) was fit to choice behavior (RT and choice accuracy) in a classic perceptual decision-making dataset (Roitman &

Shadlen, 2002). We employed the commonly used quantile maximum likelihood estimation (QMLE) method (Heathcote et al., 2002; Ratcliff & McKoon, 2007). The rationale of QMLE is to minimize the differences between the predicted data and the empirical data on the proportion of number of trials located in each RT bin. Choice accuracy was implicitly estimated because the algorithm accounts for the proportion of number of trials between correct and error trials. Nine quantiles (from .1 to .9 with .1 of step size) were used resulting in ten RT bins, with correct and error trials accounted for separately at each coherence level. Because the LDDM has no closed-form analytic expression for the RT distribution, we evaluated the prediction by Monte Carlo simulations (10240 repetitions for each input coherence). In each simulated trial, the initial values of  $R_1$  and  $R_2$  activities were set as 32 Hz to be comparable to empirical data (Churchland et al., 2008; Roitman & Shadlen, 2002). Visual stimulus (motion) inputs were defined as  $S^*(1+c')$  and  $S^*(1-c')$  for  $V_1$  and  $V_2$ , where the free parameter  $S$  models input scaling and the coherence  $c'$  replicated values in the original experiment (0, 3.2, 6.4, 12.8, 25.6, and 51.2 %)(Roitman & Shadlen, 2002). At visual stimulus onset, a gap period (90 ms) was implemented in order to capture the commonly observed initial dip in empirical firing rates (Roitman & Shadlen, 2002). Gated disinhibition was activated along with inputs after the gap. A decision was reached when either of the  $R$  unit activities reached a decision threshold of 70 Hz, the biological threshold observed in the empirical data (Roitman & Shadlen, 2002). 30 ms was added to the RT of threshold hitting to capture the delay in the down-streaming motor execution. After the decision, the input values,  $\alpha$ , and  $\beta$  were reset to zero. The negative loglikelihood (nLL) of QMLE was minimized using BADS algorithm in Matlab (Acerbi & Ma, 2017a). The estimation was conducted using GPU (NVIDIA Tesla V100) parallel computation on a high-performance cluster (NYU Langone), with 160 chains

of random initial parameter values to prevent local minima. The chain with the smallest nLL in its fitting result was selected as the best-fitting result.

The visualization of the predicted RT distribution (**Fig. 6A**) was calculated based on 60 evenly distributed RT bins, with correct and error trials calculated separately under each coherence. The predicted neural dynamics (**Fig. 6D**) were generated using the model best fit to behavior.  $R$  unit activities were aggregated across correct trials, segregated by units associated with the chosen and unchosen sides. As in the original experiment data visualization (Roitman & Shadlen, 2002), activity early in trials was aligned to stimulus onset and data within 100 ms of boundary crossing were omitted to reduce the impact of decision dynamics on visualizing early-stage ramping dynamics. Early activity traces were cut off at the median value of RT for each coherence level to ensure that the average trace was based on at least half of the trials. Activity late in trials was aligned to the time of the decision, and data within 200 ms of stimulus onset was omitted.

### *Fitting the RNM to empirical behavioral data*

In order to compare the model performance in predicting choice behaviors, we fit the original RNM to the classical perceptual decision dataset (Roitman & Shadlen, 2002). We used the reduced form of the RNM (Wong & Wang, 2006). We set eight parameters in the reduced model (see the Appendix in its original paper) as free parameters to fit: self-excitatory coupling weights  $JN_{1,1} = JN_{2,2}$ , mutual inhibitory coupling weights  $JN_{1,2} = JN_{2,1}$ , non-selective input  $I_0$ , noise amplitude of OU process  $\sigma_{noise}$ , input scale  $\mu_0$ , synaptic kinetic parameter  $\gamma$ , initial value  $H_0$ , and time

constant  $\tau_S$ . The other parameters that describing the input-output relationship of a single cell were set as the same in the paper:  $a = 270 \text{ (VnC)}^{-1}$ ,  $b = 108 \text{ Hz}$ ,  $d = 0.154 \text{ s}$ . The time constant for the AMPA receptor  $\tau_{AMPA}$  was fixed as 2 ms. The task setting and the optimization used were kept the same as in fitting the LDDM (see above). The time step  $dt$  was set as .001 s.

### *Fitting the LCA to empirical behavioral data*

Another widely acknowledged decision circuit model – the leaky competing accumulator model (LCA) (Usher & McClelland, 2001) was fit to the behavioral data (Roitman & Shadlen, 2002). The dynamics of the two nodes in the LCA can be described using the following differential equations (Eq. 22).

$$dx_i = \left( \rho_i - kx_i - \beta \sum_{j \neq i} x_j \right) \frac{dt}{\tau} + \xi_i \sqrt{\frac{dt}{\tau}} \quad (22)$$

where  $x_i$  ( $i = 1 \text{ and } 2$ ) indicates the activity of each node;  $\rho_i$  indicates the excitatory input value to each node;  $k$  indicates the net leakage on each node after the cancellation of recurrent excitation;  $\beta$  weighs the mutual inhibition strength from the other nodes;  $\xi_i$  is a Gaussian random noise on each node with a standard deviation of  $\sigma$ .

The input values  $\rho_i$  were set as  $1+c'$  for option 1 and  $1-c'$  for option 2, with  $c'$  changing over 0 to .512. We fitted the threshold as a free parameter. In that way, the time constant  $\tau$  can be taken as an arbitrary value (100 ms used in our case) since it was not independent from the threshold. Other than the parameters we mentioned above,  $T_0$  was estimated as a non-decision delay. That

gives in a total of five free parameters to estimate  $k$ ,  $\beta$ ,  $\sigma$ ,  $threshold$ , and  $T_0$ . Since the scale of the activities is arbitrarily defined, it would need rescaling when compared to the empirical data of mean firing rates in the unit of Hz. The task setting and the optimization used were kept the same as in fitting the LDDM (see above). The time step  $dt$  was set as .001 s.

### *Analysis of persistent activity*

We showed in *Results* that the LDDM with recurrent excitation predicts persistent activity that maintains input information during delay intervals. Here we provide mathematical analysis of the LDDM differential equations to examine the properties and genesis of this persistent activity. In addition to examining the property of the system with symmetric gain-control weights ( $\omega_{ii} = \omega_{ij(i \neq j)}$ ), we expanded our analysis to allow the gain-control weights to be asymmetric; this allows us to examine the robustness of LDDM properties to asymmetric weights.

Equilibrium states of the differential equations (Eqs. 1-3) after the withdrawal of inputs were considered. The gain control weights  $\omega_{ij}$  were split into two parts, with the local-option weight denoted as  $w$  ( $\omega_{ii} = w$ ) and the cross-option weight denoted as  $v$  ( $\omega_{ij(i \neq j)} = v$ ). The input values were set to zero and local disinhibition was assumed inactive ( $\beta = 0$ ). Equilibria of the system were solved by taking the intersection of the steady states of all units, i.e., when  $dR_i/dt$ ,  $dG_i/dt$ , and  $dD_i/dt$  all equal to 0. When the input terms are set to zero, the solution degrades from Eq. 5 to Eq. 23 as a linear form,

$$wR_i^* + v \sum_{i \neq j} R_j^* = \alpha - 1 - G_0 \quad (23)$$

For a binary choice system, the solution of Eq. 23 is denoted in linear algebra as:

$$\begin{bmatrix} w & v \\ v & w \end{bmatrix} \begin{bmatrix} R_1^* \\ R_2^* \end{bmatrix} = \begin{bmatrix} \alpha - 1 - G_0 \\ \alpha - 1 - G_0 \end{bmatrix} \quad (24)$$

The solutions of the equations depend on the value of recurrent excitation  $\alpha$  and baseline gain control  $G_0$ . When  $\alpha \leq 1 + G_0$ , the equations do not provide a positive solution. This explains why the system without recurrent excitation ( $\alpha = 0$ ) cannot generate persistent activity. When  $\alpha > 1 + G_0$ , the equations provide positive solutions. The model generates persistent activities in three different patterns depending on the symmetry of gain control weights, i.e.,  $v < w$ ,  $v = w$ , and  $v > w$ .

First, by assuming  $v = w$  and  $\alpha > 1 + G_0$ , the nullclines of  $R_1$  and  $R_2$  overlap on a line of attraction, as shown in **Fig. 8B** (the same as **Fig. 8-figure supplement 1B**). Any position on this line is an equilibrium point. This is a special case where the eigenvalues on each point have a real part of zero, therefore, linearization around the equilibrium points cannot tell us their stability. Thus, we checked instead the instantaneous change direction of neural activities across a wide range of initial values to see whether the system converges to the line of attraction. From the differential equations (Eqs. 1-3), the ratio  $dR_1/dR_2$  of the instantaneous change rates of  $R_1$  ( $dR_1/dt = R_1(1 - \frac{\alpha}{1+G_1})$ ) and  $R_2$  ( $dR_2/dt = R_2(1 - \frac{\alpha}{1+G_2})$ ) keeps the same ratio as the ratio of original activities ( $R_1/R_2$ ), given  $G_1 = G_2$  under the assumption of symmetric gain control weights. As a result, for any given initial values,  $R_1$  and  $R_2$  activities change in the direction that preserves

the original ratio until reaching equilibrium on the line of attraction. The instantaneous changes of  $R_1$  and  $R_2$  are shown as a vector field (red arrows) in **Fig. 8B**. Thus, any positive initial values will drop into an equilibrium state with the ratio of  $R_1^*/R_2^*$  maintaining the ratio of initial values, which preserves the ratio of inputs when the activities are inherited from the stage of value representation. **Fig. 8-figure supplement 1E** shows example dynamics of  $R_1$  and  $R_2$  under different ratios of input values (**Fig. 8-figure supplement 1G**). The activities show the characteristic dynamic of divisive normalization during the inputs and preserve this input information after withdrawal of inputs.

However, since the values of  $R_1^*$  and  $R_2^*$  are complementary on the line of attraction, any combinations of values with a constant sum satisfies the equilibrium. Thus, any disturbance to the system (e.g., random noise) will drive  $R_1^*$  and  $R_2^*$  to deviate from their original ratio resulting in a loss of the coded information about the inputs. Noise-driven drift on the line of attraction will cause decaying of the coded value information over time, consistent with the degradation attribute of working memory (Barrouillet et al., 2011; Barrouillet & Camos, 2012; Lee & Harris, 1996; Paivio & Bleasdale, 1974; Portrat et al., 2008).

In addition, under the special condition of symmetric gain control weights ( $v = w$ ), the formula in Eq. 24 can be easily expanded to multiple inputs with the equilibrium delay interval activities defined by:

$$\sum_i^N R_i^* = \frac{\alpha - 1 - G_0}{w} \quad (25)$$

The summed value of all  $R$  units equals to a constant  $\frac{\alpha-1-G_0}{w}$ . When the number of inputs ( $N$ ) increases, the activity shared by each  $R$  unit decreases and leads to a lower signal relative to noise scale. Thus, as the number of coded items increasing, the information kept during persistent activity may become less accurate considering lower signal-to-noise ratio. This may explain another important attribute of working memory – the constraint of working memory span (Cowan, 2010, 2016; Engle, 2001, 2002; Oberauer et al., 2016).

Second, by assuming  $v < w$  and  $\alpha > 1$ , the nullclines of  $R_1$  and  $R_2$  intersect on a unique equilibrium point, where  $R_1$  and  $R_2$  share the same value  $\frac{\alpha-1-G_0}{w+v}$  (**Fig. 8-figure supplement 1A**). The point is confirmed as attractive by linearization. Any positive initial values on the space of  $R_1$  and  $R_2$  will converge into this point, which is visualized in the instantaneous change ranges of  $R_1$  and  $R_2$  (red arrows) for a wide range of given initial values (**Fig. 8-figure supplement 1A**). Thus,  $R_1$  and  $R_2$  will gradually converge to be equal and the original information about input values will be lost. Nevertheless, the dynamic of information losing is based on the level of asymmetry of  $\omega_{ij}$ . For a close-to-symmetric  $\omega_{ij}$  matrix, the input information can be still preserved for a considerable amount of time. We showed example dynamics of information loss in **Fig. 8-figure supplement 1D**. After withdrawal of inputs, the  $R$  unit activities collapse into the same level and the coded ratio information gradually diminishes (simulation parameters:  $\alpha = 10, G_0 = 0, w = 1, v = .7, \beta = 0$ ).

Finally, by assuming  $v > w$  and  $\alpha > 1 + G_0$ , the nullclines of  $R_1$  and  $R_2$  intersect on a unique equilibrium point, which is confirmed as unstable by linearization (**Fig. 8-figure supplement 1C**). Any initial values of activities on the space will diverge into the upper-left or bottom-right corner



of the space generating high contrast between  $R_1$  and  $R_2$ , with the higher activity as  $\frac{\alpha-1-G_0}{w}$  and the lower activity suppressed to zero. The instantaneous change rates of  $R_1$  and  $R_2$  (red arrows) are visualized in the vector field in **Fig. 8-figure supplement 1C**. The instantaneous change direction bifurcates at the line of  $R_1 = R_2$ , biased to the side associated with higher initial activity. As an outcome, the  $R$  unit with higher initial values tends to increase while the opponent unit tends to be suppressed to zero, a process that implements WTA competition before the action stage but with constrained higher activity. Example  $R_1$  and  $R_2$  activity dynamics are shown in **Fig. 8-figure supplement 1F**. After withdrawal of inputs,  $R_1$  activities with different preceding input values collapse onto the same level of high activity, while  $R_2$  activities with lower input values are suppressed to zero. Thus, the system gradually switches from the normalized coding of inputs to a categorical coding of choice over the delay interval.

We also examined whether persistent activity could exist with active local disinhibition. We showed in *Results* that persistent activity in the working-memory task switches to WTA choice under the dynamic control of disinhibition (**Fig. 8D-F**). How does the transition from persistent activity to WTA choice happen? How might disinhibition change the dynamic pattern of persistent activity during a delay interval?

The analysis was based on the differential equations of the system with symmetric gain control weights and without inputs (Eqs. 1-3). The equilibrium solution is given by:

$$(\omega - \beta)R_i^* + \omega \sum_{i \neq j} R_j^* = \alpha - 1 - G_0 \quad (26)$$

With binary inputs, the solution can be thus written as:

$$\begin{bmatrix} \omega - \beta & \omega \\ \omega & \omega - \beta \end{bmatrix} \begin{bmatrix} R_1^* \\ R_2^* \end{bmatrix} = \begin{bmatrix} \alpha - 1 - G_0 \\ \alpha - 1 - G_0 \end{bmatrix} \quad (27)$$

Besides the impact of recurrent excitation and baseline gain control discussed above, equilibrium responses are determined by the relative strength between disinhibition ( $\beta$ ) and the gain control weight ( $\omega$ ). We examined three separate conditions:  $\beta = 0$ ,  $0 < \beta < \omega$ , and  $\beta > \omega$ . We have already shown the analysis for the special case when  $\beta = 0$  above (phase plane analysis and example dynamic shown in **Fig. 8-figure supplement 1B**) and replotted in **Fig. 8-figure supplement 2A** for the sake of comparison with other two conditions.

By assuming  $0 < \beta < \omega$ , the nullclines of  $R_1$  and  $R_2$  intersect on a unique equilibrium point, whose stability was confirmed as unstable after checking the eigenvalues of Jacobian matrix around the point (**Fig. 8-figure supplement 2B**). Any initial values on the space will diverge into the upper-left or bottom-right corner of the space, with the higher activity value as  $\frac{\alpha-1-G_0}{\omega-\beta}$ , and the lower activity value as zero. We show the instantaneous change rates of  $R_1$  and  $R_2$  at given initial values in the vector field (red arrows) (**Fig. 8-figure supplement 2B**). In **Fig. 8-figure supplement 2E**, we show example  $R_1$  and  $R_2$  activity dynamics (value setting kept the same as in **Fig. 8-figure supplement 1G**). All of the  $R_1$  with larger input values converge into the same level of activity after withdrawal of inputs, while all of the  $R_2$  with lower input values are suppressed to zero, implementing a WTA competition. Thus, the system gradually switches from normalized coding of input values to categorical choice from the early to the late stage of persistent activity.

By assuming  $\beta > \omega$ , most of the features are similar to the previous situation, except that the model now predicts no constraints on the maximum activity (**Fig. 8-figure supplement 2C**). The system shows nullclines with an intersection at a unique repeller. The activities of  $R_1$  and  $R_2$  bifurcate at the line of  $R_1 = R_2$ . The example dynamics show that the activity of  $R_1$ , which has higher initial value, increases to an unlimited level and thus will reach a decision threshold. The rising speed of  $R_1$  depends on the advantage of  $R_1$  over  $R_2$  as defined by their initial values.

Taken together, these analyses show that persistent activity is present as normalized coding of input values only with symmetric gain control weights ( $w = v$ ) and inactive disinhibition ( $\beta$ ). When disinhibition has a moderate strength ( $0 < \beta < \omega$ ), the persistent activity gradually transitions from value coding to categorical choice coding but avoids hitting the decision threshold. When disinhibition is strong enough ( $\beta > \omega$ ), the system generates WTA competition and reaches the decision threshold.

### *Simulation of pharmacological manipulation of inhibitory activity*

In **Fig. 10** we tested GABAergic agonist manipulation effects in both the LDDM and RNM (Wong & Wang, 2006) by assuming different levels of enhancement of the inhibitory projections. For LDDM (**Figs. 10A-D**), we assumed  $\omega_{ij} = 1$ ,  $\tau_R = \tau_G = \tau_D = 100 \text{ ms}$ , input scale  $S = 256$ , decision threshold = 70 Hz, and  $dt = 1 \text{ ms}$ . Panel **A** illustrated the temporal dynamic of excitatory pools ( $R_1$  and  $R_2$ ) under input coherence of 25% between control (inhibitory connection weight = 1.0) and agonist (inhibitory connection weight = 3.8) conditions (other parameters used were  $\alpha =$

5,  $G_0 = 0$ ,  $\beta = 1.4$ ,  $\sigma = 0$ ). Panel **B** examined the predicted RT and choice accuracy over different input coherences ( $c' = [0, 3.2, 6.4, 12.8, 25.6, 51.2]$  %) and levels of inhibitory weights (from 1 (control) to 4 (enhanced)) ( $\alpha = 10$ ,  $G_0 = 0$ ,  $\beta = 1.1$ ,  $\sigma = 2$ , and 10000 repetitions). Panel **C** showed the chromomeric and psychometric curves over a number of input coherences (1% – 100%) under the section between control and GABAergic agonist (inhibitory connection weight = 1.8). Panel **D** scanned the full parameter space of  $\alpha$  and  $\beta$  between the contrast of control and GABAergic agonist (inhibitory connection weight = 1.8) ( $c' = 3.2\%$ ,  $G_0 = 0$ ,  $\sigma = 2.0$  and 10000 repetitions). For RNM (**Figs. 10E-G**), we used the parameters specified in Wong and Wang, 2006 for the mean-field rate model. GABAergic weight was manipulated by weighting the inhibitory connection in the model. Panel **E** illustrated the noiseless neural dynamics of RNM using the same input coherences and inhibitory enhancement levels as in panel **A**. Panel **F** was set to compare with panel **B**, thus the input coherences and inhibitory enhancement kept the same as in panel **B**, with noise amplitude set as  $\sigma = .02$  under the framework of Wong and Wang, 2006. Panel **G** showed the chromomeric and psychometric function predicted by RNM under the same input and GABAergic assumptions as in panel **C**.

#### *Motifs tested and compared for normalized coding and winner-take-all choice*

We tested a series of motifs and found local disinhibition is critical for the integration of normalized valuation and choice functions. To do this, we tested four types of modifications that might enhance mutual competition between the option-specific local sub-circuits (**Fig. 2-figure supplement 1A**): a) *Recurrent self-excitation* (loops weighted by  $\alpha$ ), with self-amplification of each  $R$  unit, a property shown to be important for mutual competition in the RNM. b) *Local*

*disinhibition* (loops weighted by  $\beta$ ), which is the focus of the main text, mediated through disinhibitory units ( $D$ ); the function of a  $D$  unit is to inhibit the gain control  $G$  unit in the local sub-circuit therefore release inhibition on the local  $R$  units. c) *Cross inhibition* (loops weighted by  $\eta$ ), which directly inhibits the lateral  $R$  units through inhibitory units ( $I$ ) to implement mutual inhibition. d) *Lateral gain control boost* (loops weighted by  $\gamma$ ), which is mediated through excitatory units ( $E$ ) to boosts the lateral  $G$ , therefore drives higher gain control on the lateral  $R$  than the local  $R$  (i.e., asymmetric gain control) and realizes mutual inhibition.

To see which type of modification(s) is/are critical for integrated value normalization and choice, we tested different combinations of these modifications on the original DNM circuit. The full model with all modifications can be described by a set of differential equations (Eqs. 28-32):

$$\tau_R \frac{dR_i}{dt} = -R_i + \frac{V_i + \alpha R_i - I_j}{1 + G_i} \quad (28)$$

$$\tau_G \frac{dG_i}{dt} = -G_i + \omega \sum_{j=1}^N R_j + E_j - D_i \quad (29)$$

$$\tau_D \frac{dD_i}{dt} = -D_i + \beta R_i \quad (30)$$

$$\tau_I \frac{dI_i}{dt} = -I_i + \eta R_i \quad (31)$$

$$\tau_E \frac{dE_i}{dt} = -E_i + \gamma R_i \quad (32)$$

where  $i = 1, \dots, N$  designates choice alternatives, each of which receive input  $V_i$ , and  $\tau_R, \tau_G, \tau_D, \tau_I$ , and  $\tau_E$  are the time constants for the  $R, G, D, I$ , and  $E$  units. The weights  $\omega$  represent the coupling strength between excitatory units  $R$  and gain control units  $G$ , the parameters  $\alpha, \beta, \eta$ , and

$\gamma$  control the active state of recurrent excitation, local disinhibition, cross inhibition, and lateral gain control boost loops, respectively.

The active and inactive states of the four types of loops can be combined into  $2^4 = 16$  possible models. Example dynamics were shown in **Fig. 2-figure supplement 1B** for each type of model. When local disinhibition ( $\beta$ ) is off (left two columns), the model generates WTA dynamics only when cross inhibition ( $\eta$ ) is on. But the maximum activity in the late stage is still restricted to a value lower than the phasic peak during the early stage, contradicting empirical findings that the late stage decision threshold is usually higher than activity in the early phasic peak (Churchland et al., 2008; Kiani et al., 2008; Kiani & Shadlen, 2009; Louie et al., 2011; Roitman & Shadlen, 2002; Rorie et al., 2010; Shadlen & Newsome, 2001; Sugrue et al., 2004). This restriction arises because, with only cross inhibition, local option gain control is not released; this release requires local disinhibition. With local disinhibition on ( $\beta > 0$ , the right two columns), the models generate WTA dynamics with high activity in the late stage to reach the decision threshold. This is robust even without any other modifications (see the panel with  $\eta$  and  $\gamma$  off), highlighting the role of local disinhibition in generating WTA competition. For the sake of simplicity, we omitted other non-essential modifications and kept only the loop of local disinhibition. Because recurrent excitation is important for persistent activity and exists widely in cortical circuits, we retained it as well. The modified DNM model with local disinhibition and recurrent self-excitation is the primary model (LDDM) characterized in the current work.

*Data and code availability*

The empirical data presented in this paper and MATLAB code used for simulations and fitting the empirical data will be available upon publication at DOI [10.17605/OSF.IO/YGR57](https://doi.org/10.17605/OSF.IO/YGR57).