**Results**

*Local disinhibition decision model (LDDM)*

To develop an integrated circuit model of decision making, we systematically tested a series of models incorporating the core elements of existing models, namely divisive gain control, recurrent excitation, and mutual competition (**Fig. 2-figure supplement 1**; see **Methods** *Motifs tested and compared for normalized coding and winner-take-all choice* for the analysis details). This analysis identified *local disinhibition* as the crucial component that can integrate mutual competition and value normalization within the existing circuit architecture of DNM. In the rest of this paper, we focus on the novel local disinhibition decision model (hereafter LDDM).

In the LDDM (**Fig. 2A**), an option-specific disinhibitory *D* unit receives input from its associated excitatory *R* unit and inhibits the inhibitory *G* unit in the local sub-circuit. Biased disinhibition – via different value inputs to option-selective *R* units – can thus selectively release local circuit gain control, generating an unbalanced gain control between local and opponent circuits and leading to WTA competition. In this model, the network shifts from value coding to WTA competition regimes in response to an onset of disinhibition (controlled by the coupling strength between *R* and *D*).With zero or weak *R-D* coupling, the circuit preserves normalized value coding consistent with the DNM; with strong *R-D* coupling, the circuit switches to a state of WTA selection (**Fig. 2B**). Inhibitory units, as a result, dynamically switch from a non-selective response pattern to a selective response pattern (*G* and *D* units in **Fig. 2B**). This flexible onset of disinhibition is modeled after biological findings, which show that activation of disinhibition in cortical circuits arises from exogenous, long-distance projections (Fu et al., 2014; Kamigaki, 2019; Lee et al., 2013; Pi et al., 2013; S. Zhang et al., 2014) (**Fig. 2C**). This form of top-down control allows for flexibility in the relative timing of the valuation and selection processes, consistent with neural and behavioral data in different task paradigms (see below *Gated disinhibition provides top-down control of choice dynamics*).

Activity dynamics of the LDDM are described by a set of differential equations:

|  |  |  |
| --- | --- | --- |
|  |  | (1) |
|  |  | (2) |
|  |  | (3) |

where *i* = 1, …, *N* designates choice alternatives, each of which is represented by an R unit receiving selective input and non-selective baseline input *BR*. , , and are the time constants for the *R*, *G*, and *D* units. The weights represent the coupling strength between excitatory units and inhibitory (gain control) units , with each *G* unit driven by a weighted sum of excitatory inputs from all *R* units and a non-selective baseline input and inhibited by its local ; the parameter reflects the strength of recurrent self-excitation on *R* units. Finally, weights the coupling strength between the excitatory and the disinhibitory units and is presumed to be under external (task-triggered) control.

[Insert **Figure 2** about here]

[Insert **Figure 2-figure supplement 1** about here]

*Dynamic divisive normalization preserved in the LDDM*

We first examine whether the LDDM retains the dynamics of divisively normalized value coding seen in the DNM (LoFaro et al., 2014; Louie et al., 2014). As discussed above, during initial option evaluation the disinhibitory units are silent () and the sole difference between the LDDM and the DNM is recurrent excitation (controlled by ). Example activity traces in **Fig. 3B** show that the LDDM preserves characteristic early-stage dynamics and contextual modulation seen in both empirical data (**Fig. 3C**) and the original DNM (LoFaro et al., 2014; Louie et al., 2011, 2014). Immediately after stimulus onset, *R1* activities replicate the transient peak observed in a wealth of studies (Andersen & Buneo, 2002; Churchland et al., 2008; Gnadt & Andersen, 1988; Louie et al., 2011, 2014; Platt & Glimcher, 1999; Rorie et al., 2010; Sugrue et al., 2004). Further, the network settles to equilibrium displaying relative value coding: *R1* activity increases with *V1* and decreases with *V2* (**Fig. 3B,** *R1* activityacross *V1* inputs (upper panels) and *V2* inputs (bottom panels)), reflecting a contextual representation of value (see Methods for details of parameters used in visualization).

Taking advantage of its simplified mathematical form, we analytically evaluated the LDDM by conducting phase plane analysis and found that it represents each set of input values () as one unique and stable equilibrium point in its output space () when. Specifically, we solved the equilibrium state of each *R* unit by setting each differential equation (Eqs. 1-3) to zero, which results in the nullcline of each *R* unit as a function of the given activities of the other *R* unit, visualized in **Fig. 3D**. The nullclines of *R1* (solid) and *R2* (dashed) intersect at a unique point, regardless of equal or unequal input values (see different panels for examples of different inputs). This point indicates that the dynamical system, when receiving any positive inputs, can maintain an equilibrium where every unit maintains a steady level of activity. Linearization analysis around this point suggests that this point is attractive: given any initial values to the system, the activities of the units will converge into the unique equilibrium point (see **Methods** *Equilibria and stability analysis of the LDDM* formathematical proof). The solution of the steady state of neural activity at the equilibrium (noted as ) reflects divisive normalization (Eq. 4), inheriting the property from the original DNM (LoFaro et al., 2014; Louie et al., 2014). The only difference at the equilibrium is the constant in the denominator () introduced by baseline gain control and recurrent excitation; this change rescales the activity magnitudes but preserves normalized value coding.

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

We next verified that the normalized value coding produced by the LDDM cannot be implemented by standard recurrent RNM models. **Fig. 4A** compares the activity of as a function of both value inputs (*V1* and *V2*) in the LDDM (left panel), the original DNM (middle panel), and the RNM (right panel). Both the LDDM and the DNM exhibit activities (indicated by color) that monotonically increase with input but decrease with , with a slightly steeper dependence in the LDDM versus the DNM model depending on the rescaling of . In contrast, strong WTA dynamics in the RNM implement categorical (choice) coding rather than relative value representation, with high or low coding of input values (right panel).

[Insert **Figure 3** about here]

To quantitatively test value normalization, we fit the models to observed firing rates of monkey lateral intraparietal (LIP) neurons under varying reward conditions (Louie et al., 2011). In the empirical data (**Fig. 4B**, dots), LIP activity increases with the reward (juice quantity) associated with the target inside the neuronal response field (*Vin*) and decreases with the summed rewards of targets outside the response field (*Vout*). The fitting results show that the DNM captures the rescaled firing rates very well with only two free parameters (baseline input *B­R* = 70.92, and an arbitrary scaling parameter ; see **Methods**; middle panel in **Fig. 4B**, *R2* = .9640). The LDDM with an additional parameter related to self-excitation () and baseline gain control () fitted slightly better than the DNM (*BR* = 71.53, ; see **Methods**; left panel in **Fig. 4B**,*R2* = .9646; parameter recovery analysis shows that the LDDM is highly robust in the data fitting, **Fig. 4-figure supplement 1**). Note that fitting to the current dataset is not able to differentiate the contributions of and to the neural dynamics (see proof in **Methods**); thus more empirical data will be needed to draw conclusions about the role of recurrent self-excitation in value coding. However, we do show below that self-excitation is critical for generating persistent activities (see section *line-attractor persistent activity*).

Fitting the standard RNM with four parameters (see **Methods**) does not capture the neural activities as well as the LDDM and DNM (right panel in **Fig. 4B**) (*R*2 = .8920). The difference of the performance between these models is mild because divisive and subtractive types of inhibition differ in a fine-tuned way. The curvature of neural activity as a function of *Vout* shows a linear type of lateral inhibition, in contrast to the concave curvature predicted by divisive normalization in LDDM (and DNM). Furthermore, fitting the RNM to the data results in a parameter regime that can no longer generate WTA competition; instead, the model predicts mean firing rates in a low-activity regime with maximum value ~ 3.5 Hz (**Fig. 4-figure supplement 2**).

[Insert **Figure 4** about here]

[Insert **Figure 4-figure supplement 1** about here]

*Local disinhibition drives winner-take-all competition*

A key question is whether the LDDM also produces WTA competition. Given the architecture of the LDDM, local disinhibition is hypothesized to break the symmetry between option-specific *R*-*G* sub-circuits, enabling a competitive interaction between sub-circuits. To examine whether this competition produces WTA selection, we simulated model activity in a reaction-time version of a motion discrimination task, a standard perceptual decision-making paradigm in non-human primates (Churchland et al., 2008; Roitman & Shadlen, 2002). The task contains two stages of processing: the pre-motion stage with only the choice targets presented and the motion stage presenting a random-dot motion stimulus simultaneously with a go signal. Animals are allowed to select an option, indicating their percept of the main direction of motion, at any time following motion stimulus/go signal onset (see timeline, **Fig. 5A**). During the pre-motion stage, we simulated equal value inputs, given the equal prior probability of either target being correct. The simulated pre-motion dynamics replicate the characteristic transient peak observed in both perceptual and economic decision-making tasks (Andersen & Buneo, 2002; Churchland et al., 2008; Louie et al., 2011; Rorie et al., 2010). At motion stimulus onset, inputs to the two *R* units are changed according to the task design; disinhibition (i.e., value) is switched on at the go signal, simultaneously with motion inputs.

We find that the LDDM replicates neural and behavioral aspects of WTA competition. In **Fig. 5A**, we show example model activity for five input strengths corresponding to different motion coherence levels. Consistent with electrophysiological recordings in posterior parietal cortex (Churchland et al., 2008; Roitman & Shadlen, 2002; Shadlen & Newsome, 2001), model *R* unit activities bifurcate based on the input strengths, with the unit receiving stronger input ramping-up to an (arbitrary) decision threshold while the activity of the opponent unit is suppressed. The speed of bifurcation depends on the contrast between the inputs, a variable equivalent to motion coherence in the experimental literature (Roitman & Shadlen, 2002; Shadlen & Newsome, 2001). Furthermore, the LDDM predicts the dynamics of the two types of interneurons *G* and *D* governing excitatory neuron computation (**Fig. 5B**). Prior to the go signal, the two *G* units share the same activity, but after the go signal, the activity levels bifurcate because of disinhibition. In contrast to *R* units, the *G* unit in the sub-circuit receiving stronger input shows lower activity, indicating a stronger disinhibition of the associated *R* unit. Thus, the LDDM exhibits mutual competition that generates WTA selection in excitatory neurons, as in the existing RNM; this competition is mediated by a novel disinhibitory control through different interneuron subtypes.

What features of the LDDM are essential to generate WTA competition? We examined the dynamical properties of the system under disinhibition by conducting phase plane analyses. As shown in **Fig. 5C**, the network in the choice regime ( in this example) shows a different configuration of nullcline intersections than the network in the value representation regime (; **Fig. 3D**). Given equal inputs, the nullclines of *R1* and *R2* intersect at three equilibrium points (left panel in **Fig. 5C**), with the central point unstable and the two peripheral points stable. Thus, given an initial configuration of *R1*-*R2* activities (with the presence of noise), the system will converge to the closer peripheral attractor (see example activity traces in blue and red thin lines) and implement WTA competition. Given moderately unequal inputs, the basin of attraction is biased towards the side with higher input, resulting in a higher probability falling into the side with higher input (middle panel in **Fig. 5C**). When inputs are extremely unequal, the stable equilibrium in the middle of the basin and the unstable equilibrium point associated with weaker input no longer exist, leaving only the attractor associated with stronger input (**Fig. 5C**, right). Thus, across varying degrees of input coherences, disinhibition drives the LDDM towards a selection of one of the potential choices. This can be seen in **Fig. 5D** by viewing the output ratio () of the preferred attractor as a function of input ratio (): under active disinhibition () we observe categorical coding (green line), in contrast to inactive disinhibition () where the output ratio faithfully preserves the original ratio of inputs (dark line) (Other parameters used in simulations: , , , , , and ).

To understand the operating regimes of the LDDM, we quantified model behavior across the full parameter space defined by recurrent excitation weight () and local disinhibition weight (), both of which are critical in determining the properties of the system (see **Methods** *Equilibria and stability analysis of the LDDM* for mathematical proof). Decisions with equivalent inputs are a critical test of WTA behavior, since WTA systems should select an option (stochastically) even in these symmetric scenarios (Furman & Wang, 2008; Lo & Wang, 2006; Wang, 2002; Wong & Wang, 2006); we therefore analyzed system behavior under equal value inputs. As shown in **Fig. 5E**, this analysis revealed two distinct territories corresponding to value representation and WTA operating regimes. The value representation regime generates a unique attractor for normalized value representation but no WTA attractors; in contrast, the WTA regime (induced by a change in ) generates no normalization attractor but instead *R1* and *R2* always diverge into high-contrast attractors (see **Figure 5-figure supplement 1** and **Methods** *Equilibria and stability analysis of the LDDM* for a full description of regime parcellation). While the WTA regime asymptotically approaches zero disinhibition when recurrent excitation is extremely strong, local disinhibition is always required to generate WTA choice. Models with a wide range of recurrent excitation can transit from value representation to WTA choice with an increase in local disinhibition strength (for example, red arrow in **Fig. 5E**). These findings emphasize the importance of local disinhibition to WTA choice, and highlight a particular role for a dynamic gating signal in controlling the transition from value coding to option selection.

[Insert **Figure 5** about here]

[Insert **Figure 5-figure supplement 1** about here]

*The LDDM captures empirical choice behavior and neural activity*

While the preceding analyses show that the LDDM can generate value normalization and WTA selection, a critical question is whether this circuit architecture accurately captures behavioral and neural aspects of empirical decision making. Here, we take advantage of the limited number of parameters in this differential equation-based LDDM (compared to more complicated conductance-based biophysical models (Tegnér et al., 2002; Wang, 1999, 2002; Wong & Wang, 2006)), which allows model fitting to empirical data. Specifically, we fit LDDM parameters to nonhuman primate behavior from the reaction-time version of the motion discrimination task described above. These choice and RT data from monkeys align with a reduced form model of decision making (the drift diffusion model) (Ratcliff & McKoon, 2008), and the activity of posterior parietal neurons recorded during this task display characteristic decision-related features (motion-dependent ramping, a common decision threshold, and WTA activity).

To fit the LDDM to behaviorally observed RTs, we employed the standard quantile maximum likelihood method (QMLE) to the RT distributions across input coherence levels (0 – 51.2%), with correct and error trials dissociated (Hawkins et al., 2015; Heathcote et al., 2002; Ratcliff & Tuerlinckx, 2002). We set as 1 and the baseline input *BR* as zero. Because of the collinearity issue between baseline gain control () and self-excitation () mentioned above (see model fitting in **Fig. 4**) also exists in fitting WTA choice behavior (see **Fig. 6-figure supplement 3**), we kept as a free parameter but set to zero (note that this limits the interpretability of fit values as simply the level of recurrence). The model is reduced to seven parameters: recurrent excitation weight , local disinhibition weight , noise parameter , input value scaling parameter *S*, and time constants , , and (see Methods for model-fitting details). Predictions of the best fitting model are shown in **Fig. 6A** (best fitting parameters: , , , *S* = 3251, , , and ). The optimization surfaces visualized across pairs of parameters (**Fig. 6-figure supplement 1**) were consistent with robust parameter fitting, and parameter recovery indicated that the parameters are recoverable and identifiable within the network (**Fig. 6-figure supplement 2**). Model-predicted RT distributions (lines) closely follow the empirical distributions (bars) for both correct (blue) and error (red) trials across different levels of input coherence. The aggregated mean choice accuracy and RT data are shown in **Fig.** **6C**. Model choice accuracy (line) captures the average empirical psychometric function (crosses); model RT captures coherence-dependent changes in the chronometric function, including longer RTs in error trials (dashed line and empty dots) compared to correct trials (solid line and dots). Beyond mean RT data, the LDDM accurately captured aspects of the empirical RT distributions, as evident in the quantile probability plot (QPP) of RT quantiles as functions of chosen ratio (**Fig. 6B**). Given the collinearity issue between and , the fitted value of does not reflect the exact level of recurrence in the circuit, and future empirical data will be needed to differentiate how recurrence and baseline inhibition contribute to LDDM WTA selection.

We compared the performance of the LDDM in fitting this classical dataset with the reduced form of the RNM (Wong & Wang, 2006) (**Fig. 6-figure supplement 4**), as well as another prominent computational decision model with a similar architecture of mutual inhibition – the leaky competing accumulator model (LCA) (Usher & McClelland, 2001) (see **Fig. 6-figure supplement 5**). The performances of the three models were close in predicting averaged RTs and choice accuracy. However, the LDDM captures the skewness and the shape of RT distributions better than the other two, thus showing better goodness of fit (negative log likelihood) and AIC measures (nLLLDDM = 16546, nLLRNM = 16573, nLLLCA = 16948, AICLDDM = 33109, AICRNM = 33165, AICLCA = 33932).

Importantly, the LDDM – fit only to behavior – generates predictions about the underlying neural dynamics that can be compared to electrophysiological findings. We examined *R* unit activity in the best-fitting model, with predicted activity aggregated across trials and aligned to the onset of stimuli and the time of decision as in the original study (Roitman & Shadlen, 2002). Aligned to the onset of stimuli **(Fig. 6D**, left), neural responses are aggregated by coherence level and eventual choice, and truncated at median RT. These data show clear evidence of WTA competition: chosen (solid) and unchosen (dashed) activity traces diverge over time. Moreover, neural activity is stimulus-dependent: the dynamics of both chosen and unchosen units ramp at different, coherence-dependent speeds, consistent with empirical findings consistent with an accumulation process. More quantitatively, we examined the relationship between activity and coherence at the specific time point (arrow points **a** and **b**) reported in the original work (**Fig. 6E**). Model predictions align well with empirical observations: the deviance of the LDDM predicted activities is the smallest among the three alternatives models from the empirical recordings (quantified by root-mean-square error (RMSE); RMSELDDM = 2.74, RMSERNM = 20.10 (**Fig. 6-figure supplement 4E**), RMSELCA = 3.92 (**Fig. 6-figure supplement 5E**)).

Aligned to the onset of decision (**Fig. 6D,** right), model *R* unit activity near the time of choice shows further evidence of WTA competition observed in real neurons: the initial divergence between chosen and unchosen activity traces extends into a categorical coding of choice. The relationship between activity and coherence quantitatively replicates the empirical pattern immediately preceding the decision time (Roitman & Shadlen, 2002): chosen activity (indicated by arrow **c** in **Fig. 6D** and plotted in **Fig. 6E**) no longer shows much difference across coherence conditions, while unchosen activity (indicated by **d** in **Fig. 6D** and plotted in **Fig. 6E**) retains a decrease. The quantification of the model activities shows that LDDM traced the best among the three models on the empirical activities when aligned to the choice onset (RMSELDDM = 6.77 (**Fig. 6E**); RMSERNM = 9.35 (**Fig. 6-figure supplement 4E**); RMSELCA = 7.51 (**Fig. 6-figure supplement 5E**)). Thus, *R* unit activity – in a model with parameters fit only to behavior – replicates the recorded activity of parietal neurons during both initial decision processing and eventual choice selection.

Different from the other two models, the LDDM is the only model that provides the dynamics of the two types of interneurons (**Fig. 6F-I**). The inhibitory (*G*) units selectively code input values and choice but exhibit complex dynamics due to the aggregating method across trials with different RTs: in the early stage (dynamics sorted to the left in **Fig. 6F** and upper panel in **Fig. 6G**), the *G* activities show a trend of increasing because of stronger excitatory inputs via *R* units. Later on, when the inhibition from *D* units increase (**Fig. 6H**), the *G* activities start to decrease. In the late stage (dynamics sorted to the right in **Fig. 6F** and the lower panel in **Fig. 6G**), the chosen *G* units show lower activities than the unchosen side because of stronger inhibition from *D* as an outcome of WTA competition. The competition happens faster and takes a shorter time to reach a decision in the trials with higher coherence. The aggregating method based on RT makes the pattern complicated in the coherence dependency of **Fig. 6G**. The chosen side activities maintain higher in the trials with higher coherence because of shorter duration of inhibition from *D* to *G* in those trials with shorter RT (indicated by arrows **a** and **c**). The unchosen side in the early stage (arrow **b**) shows lower activities in the trials with higher coherence because WTA happened faster in those trials and the *G* units are more rapidly inhibited by *D* (arrow **b**); whereas, in the late stage, the unchosen side activities maintain higher in the trials with higher coherence because of shorter duration of WTA competition (arrow **d**): the twist shown in the dynamics across different coherence levels (unchosen dynamics aligned to the right in **Fig. 6F**) reflects such impacts from the competition speeds and the competition durations waning and the waxing over time, respectively. The dynamics of *D* units rapidly increase in the early stage (dynamics sorted to the left in **Fig. 6H**). The chosen side activities show weak coherence dependency (arrow **a**, **Fig. 6I**) because of the aggregating method. Stronger competition in high-coherence trials boosts the chosen side activity more rapidly but those trials are also the trials with shorter durations of disinhibition. When aggregating based on RTs, the coherence dependency becomes weak. While, the unchosen side activities preserve the trends of decreasing with input coherence (arrow **b**, **Fig. 6I**) since the unchosen side activities increase slower and terminated earlier in the trials with higher coherence. Dynamics in the late stage (dynamics sorted to the right in **Fig. 6H**) shows higher activities in the chosen side than the unchosen side as an outcome of WTA competition. The coherence dependency mainly reflects the duration of competition, in which the trials with higher coherence terminated earlier so that has lower *D* activities (arrows **c** and **d**, **Fig. 6I**). These new predictions distinguish the dynamics of different populations of interneurons under a commonly used strategy of data aggregation based on RT.

[Insert **Figure 6** about here]

[Insert **Figure 6-figure supplement 1** about here]

[Insert **Figure 6-figure supplement 2** about here]

[Insert **Figure 6-figure supplement 3** about here]

[Insert **Figure 6-figure supplement 4** about here]

[Insert **Figure 6-figure supplement 5** about here]

*The LDDM integrates normalized value coding and WTA choices*

While the LDDM separately replicates normalized value coding and WTA dynamics shown in different empirical studies, a key distinguishing feature of the LDDM is that it can capture both phenomena within a single experimental context. Numerous studies using the random-dot motion paradigm show two stages of dynamics: target (action) representation during the pre-motion stage and WTA selection after the go cue following motion stimuli (Churchland et al., 2008; Rorie et al., 2010). Neural activity in the pre-motion stage shows a characteristic phasic-sustained dynamic to the presentation of visual cues; rather than purely sensory information, activity during this stage reflects the magnitude and probability of reward associated with the visual cues (Rorie et al., 2010). After the go cue, WTA dynamics reflects an integration of motion information and implements a transition from initial value coding to a categorical coding of choice in the late stage of decision (Churchland et al., 2008; Ding & Gold, 2010; Kiani et al., 2008; Roitman & Shadlen, 2002; Rorie et al., 2010; Shadlen & Newsome, 2001). Studies of economic choice show a similar set of dynamics, a context dependent valuation followed, after a go-cue, by a shift to WTA (Louie et al., 2011, 2014; Louie & Glimcher, 2010; Pastor-Bernier & Cisek, 2011; Sugrue et al., 2004). Interestingly, the number of alternatives affects the neural dynamics during both representation and choice (Basso & Wurtz, 1997, 1998; Churchland et al., 2008). When the choice set is expanded from two options to four options, early representational activity is lower during pre-motion dynamics (**Fig. 7A**) and the speed of WTA dynamics slows after motion onset (**Fig. 7C**).

Accordingly, in this section we examined whether the LDDM replicates the impact of the number of alternatives on both early and late empirical neural dynamics during both the representation phases and the WTA phases observed in real neurons. Under four (versus two) alternatives, LDDM R unit activity during the representation stage decreases because of increased recurrent inhibition, driven by multiple contextual inputs (left side in **Fig. 7D**). Similarly, the ramping speed after motion onset and disinhibition decreases in the 4-alternative (versus the 2-alternative) condition, despite identical parameters (**Fig. 7E**). These results highlight the LDDM as a potential mechanism of integrating normalized value coding and WTA competition within a single circuit architecture.

[Insert **Figure 7** about here]

*Disinhibition controls point versus line attractor dynamics in persistent activity*

We next examine implications of the local disinhibition architecture for another characteristic of decision-related neural firing: persistent activity. In cortical areas such as parietal (Kiani et al., 2008, 2014; Kiani & Shadlen, 2009; Roitman & Shadlen, 2002; Shadlen & Newsome, 2001), prefrontal (Funahashi et al., 1989; Fuster & Alexander, 1971; Goldman-Rakic, 1995; Rigotti et al., 2013), and premotor cortices (Pastor-Bernier & Cisek, 2011), neurons show elevated firing in the absence of stimulus-driven input over intervals of seconds; such persistent activity is thought to underlie working memory and enable decisions based on internally maintained information. In RNMs, recurrent excitation and feedback inhibition preserve categorical choice information after input withdrawal because of point-attractor dynamics (Furman & Wang, 2008; Wang, 2002; Wong & Wang, 2006). Here, we answer two questions: does the LDDM generate persistent activity, and how does this persistent activity differ from that in standard RNMs?

We found that the LDDM can generate two distinct forms of persistent activity, controlled by the state of disinhibition. **Fig. 8A** shows example dynamics of two *R* units before and after withdrawal of inputs while disinhibition is silent. Following input withdrawal, network activity decreases but still preserves elevated firing rates, governed by the self-excitation parameter (the network loses elevated activity when ). The persistent activity ratio between *R1* and *R2* preserves the ratio between the input values *V1* and *V2*, in contrast to RNMs which only preserve categorical information about the largest value (see **Fig. 8-figure supplement 1** and **Methods** *Analysis for persistent activity* for mathematical proof). Phase plane analysis suggests that relative value coding in persistent activity arises from a line-attractor dynamic in the network during the inactivation of disinhibition, unlike point-attractor dynamics in the RNM (**Fig. 8B**). Like other line-attractor models of persistent activity that store continuous-valued information (Burak & Fiete, 2009; Compte, 2000; Ganguli et al., 2008; Seung, 1996), an unbiased coding of the input ratio requires perfectly balanced gain control weights from *G* to *R*. Unbalanced weights will result in distorted coding of the input ratio and graded coding of the inputs will decay over time (**Figs. 8-figure supplement 1D** and **E**). For perfectly balanced weights, the line attractor state is vulnerable to noise perturbation. A small perturbation can easily drive the activity to drift on the line of attractors, with the summed value of *R1* and *R2* as a constant (). The preserved ratio between *R1* and *R2* drifts stochastically over time, similar to the prediction of other line-attractor circuits and consistent with behavioral and neural variability related to working memory (Seung, 1996; Wimmer et al., 2014).

However, a line attractor is not the only state that the LDDM predicts. If disinhibition is activated during the delay interval, the network switches to a point attractor dynamic (see **Figs. 8-figure supplement 2** and **Methods** *Analysis for persistent activity* for mathematical proof). **Fig. 8D** shows example dynamics of two *R* units before and after withdrawal of inputs. Disinhibition drives a competition between the two *R* units, resulting in a switch between graded coding of the input ratio to a categorical coding of the largest value ( in visualization). Interestingly, a transition of coded information from input values to categorical information has been widely observed in firing rates in decision related regions, such as LIP and superior colliculus, during the delay period of decision making (Rorie et al., 2010; Shadlen & Newsome, 2001; B. Zhang et al., 2021). The point attractor predicted by the circuit under disinhibition (**Fig. 8E**) is highly tolerable to perturbations compared to the line attractor, and choice performance over long delays may require a switch from the value coding to the categorical regimes. As a plausible biological mechanism for mediating top-down control, disinhibition may gate such a transition without changing the network architecture.

The LDDM can be easily expanded to multiple alternatives. Here we show an example of a 5-alternative case, with 5 sets of option-specific *R*-*G*-*D* units. A line attractor network with silent disinhibition (**Fig. 8C**, right) is able to retain input value information of the 5 items simultaneously in the network. Due to normalization, the neural activity representing each alternative decreases with the total number of alternatives, with the summed value as a constant (), leading to a lower signal-to-noise ratio when coding more items; this set-size effect may be related to WM memory span constraints (Cowan, 2010, 2016; Engle, 2001, 2002; Oberauer et al., 2016). When disinhibition is active, the LDDM exhibits a point attractor (**Fig. 8F**, right), and the network only holds the information of the largest item as a categorical code during persistent activity.

[Insert **Figure 8** about here]

[Insert **Figure 8-figure supplement 1** about here]

[Insert **Figure 8-figure supplement 2** about here]

*Gated disinhibition provides top-down control of choice dynamics*

In addition to its crucial role in generating WTA competition, local disinhibition provides an intrinsic mechanism for top-down control of choice dynamics. Decision circuits show remarkable flexibility in timing, with similar neurophysiological data recorded in a variety of task paradigms: in addition to reaction-time tasks, in which subjects can choose at any time immediately after onset of stimulus, decision-rela­ted neural activity has been widely studied in fixed-duration and delayed-response tasks. In fixed-duration tasks, subjects are required to withhold selection of an action until an instruction signal. Neural activity prior to the instruction signal reflects value information, for example about reward characteristics (Dorris & Glimcher, 2004; Louie et al., 2011; Platt & Glimcher, 1999; Sugrue et al., 2004; Watanabe, 1996) or accumulating perceptual evidence (Kiani et al., 2008, 2014; Kiani & Shadlen, 2009; Kim & Shadlen, 1999; Roitman & Shadlen, 2002; Rorie et al., 2010; Shadlen & Newsome, 2001); however, this activity never fully diverges or reaches the decision threshold until after the instruction cue, suggesting a gating of the selection process. In delayed-response (working memory) tasks, subjects must postpone selection for an interval that includes both stimulus presentation and an additional subsequent interval after the stimulus is withdrawn. As in fixed-duration tasks, neural activity in delayed-response tasks typically carries decision–related information (across both the stimulus and delay periods) but WTA selection – and behavioral choice – is withheld until the instruction cue is given (Kiani et al., 2008, 2014; Kiani & Shadlen, 2009; Kim & Shadlen, 1999; Roitman & Shadlen, 2002; Shadlen & Newsome, 2001). Thus, biological decision circuits are able to evaluate choice options while selectively initiating the WTA selection process with variable context-dependent timing.

[Insert **Figure 9** about here]

How neural circuits implement dynamic control of selection – and temporal separation of evaluation and WTA choice – is largely unaddressed in current decision models. In RNM models, neural activity is driven by attractor dynamics; option evaluation and the selection process cannot be disambiguated, and WTA competition is not under top-down control. Here, we examine how the timing of a dynamic top-down control signal – modulating the strength of disinhibition via long-range inputs and neuromodulation – allows the LDDM to capture neural activity in different task paradigms. In these simulations, disinhibition is activated when the instruction cue to choose is presented. **Fig. 9A** shows LDDM activity in a reaction-time task, a standard paradigm in perceptual decision-making (Churchland et al., 2008; Roitman & Shadlen, 2002). As in prior analyses (**Figs. 5** and **6**), LDDM *R* units show simultaneous evaluation (coherence-dependent ramping) and WTA selection (rise to threshold) processes, driven by an immediate activation of disinhibition at motion stimulus onset.

In a fixed-duration task (**Fig. 9B**), disinhibition is activated after a required interval of stimulus presentation. Compared to the reaction-time task, LDDM activity here shows distinct, temporally separated patterns during stimuli viewing and option selection; this temporal segregation is driven by the activation of disinhibition (a step function on in this example), which promotes a transition between value representation and WTA choice.

A further demonstration of this temporal flexibility arises from considering delayed-response tasks (**Fig. 9C**), which include an interval between stimuli offset and onset of the instruction cue. Consistent with its ability to maintain persistent activity (**Fig. 8**), the LDDM shows value coding across the delay interval and implements WTA selection until instruction and accompanying activation of disinhibition. These results show that the LDDM – via modulation in the timing of disinhibition activation - can temporally separate the value representation and selection processes, enabling it to capture the diversity of neural dynamics seen in reaction-time, fixed-duration, and delayed-response tasks.

*Inhibitory potentiation distinguishes LDDM from earlier models*

The architecture of disinhibition employed by the LDDM is more structured than earlier non-selective inhibition used in more standard competition networks. This distinction gives rise to the novel prediction from LDDM that the influence of global changes in inhibitory tone are non-selective during representation, but switch to being input-selective after disinhibition is increased. This reflects a fundamentally novel prediction of this class of model. To empirically test that key prediction, optogenetic/pharmacological manipulation of inhibitory connection weights (e.g., via GABAergic agonist) could be introduced. The LDDM contains two different types of inhibition and thus its reaction to inhibitory potentiation depends on both the state of the disinhibitory network and the intensity of potentiation. To highlight the importance of that prediction, we implemented different levels of inhibitory connection weights in both the LDDM and the standard RNM.

[Insert **Figure 10** about here]

At the neural level, the LDDM predicts a dissociable effect of potentiated inhibition on the primary (i.e., *R*) neuron’s activity (**Fig 10A**). During option representation (cue interval in fixed duration trials), potentiated inhibition increases both recurrent and lateral inhibition, leading to decreased firing rates and a weaker modulation by value in the primary neurons. During option selection (go/choice intervals in fixed duration trials), stimulation of local disinhibition increases WTA activity and simultaneously decreases the late-stage representation of value. As an outcome, these changes produce a speeding up of RTs but a decreased choice accuracy (**Fig. 10B**). The expected differences between the control condition and the inhibitory potentiation condition would be evident in chronometric and psychometric curves across different levels of inputs (**Fig. 10C**). Note that the qualitative predictions for inhibitory potentiation effects on RT and accuracy (i.e., the direction of changes) are robust to specific LDDM parameterizations (**Fig. 10D**). In contrast, in more traditional networks like the RNM that employ non-selective inhibition, potentiated inhibition suppresses the excitatory neural activities during the WTA competition (**Fig. 10E**). The suppression in neural coding will slow down RTs but will not affect the choice accuracy (**Figs. 10F** and **G**). These novel predictions could be readily tested and differentiate models that rely on the structured disinhibition that we propose from models that employ traditional changes in the E/I balance to achieve state changes.