Systems/Circuits

Working Memory and Decision-Making in a Frontoparietal Circuit Model

John D. Murray,[1]* Jorge Jaramillo,[2]* and Xiao-Jing Wang[2,3]

[1]Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut 06510, [2]Center for Neural Science, New York University, New York, New York 10003, and [3]New York University-East China Normal University Joint Institute of Brain and Cognitive Science, NYU Shanghai, Shanghai 200122, People's Republic of China

Working memory (WM) and decision-making (DM) are fundamental cognitive functions involving a distributed interacting network of brain areas, with the posterior parietal cortex (PPC) and prefrontal cortex (PFC) at the core. However, the shared and distinct roles of these areas and the nature of their coordination in cognitive function remain poorly understood. Biophysically based computational models of cortical circuits have provided insights into the mechanisms supporting these functions, yet they have primarily focused on the local microcircuit level, raising questions about the principles for distributed cognitive computation in multiregional networks. To examine these issues, we developed a distributed circuit model of two reciprocally interacting modules representing PPC and PFC circuits. The circuit architecture includes hierarchical differences in local recurrent structure and implements reciprocal long-range projections. This parsimonious model captures a range of behavioral and neuronal features of frontoparietal circuits across multiple WM and DM paradigms. In the context of WM, both areas exhibit persistent activity, but, in response to intervening distractors, PPC transiently encodes distractors while PFC filters distractors and supports WM robustness. With regard to DM, the PPC module generates graded representations of accumulated evidence supporting target selection, while the PFC module generates more categorical responses related to action or choice. These findings suggest computational principles for distributed, hierarchical processing in cortex during cognitive function and provide a framework for extension to multiregional models.

Key words: attractor network; decision-making; NMDA receptor; parietal cortex; prefrontal cortex; working memory

Significance Statement

Working memory and decision-making are fundamental "building blocks" of cognition, and deficits in these functions are associated with neuropsychiatric disorders such as schizophrenia. These cognitive functions engage distributed networks with prefrontal cortex (PFC) and posterior parietal cortex (PPC) at the core. It is not clear, however, what the contributions of PPC and PFC are in light of the computations that subserve working memory and decision-making. We constructed a biophysical model of a reciprocally connected frontoparietal circuit that revealed shared and distinct functions for the PFC and PPC across working memory and decision-making tasks. Our parsimonious model connects circuit-level properties to cognitive functions and suggests novel design principles beyond those of local circuits for cognitive processing in multiregional brain networks.

Introduction

Cognitive functions engage distributed networks of areas in the primate brain, with prefrontal cortex (PFC) and posterior parietal cortex (PPC) as key nodes (Duncan, 2010; Mitchell et al., 2016; Domenech et al., 2017). Working memory (WM) and decision-making (DM) are fundamental building blocks of cognition that recruit a common prefrontal–parietal network, with WM and DM signals partially overlapping at the neuronal level (Meister et al., 2013). Both PPC and PFC exhibit characteristic neural activity of WM and DM. WM is associated with stimulus-selective persistent activity that spans the mnemonic delay (Goldman-Rakic, 1995; Constantinidis and Procyk, 2004). DM is associated with ramping dynamics reflecting the accumulation of

Received Feb. 1, 2017; revised Aug. 24, 2017; accepted Sept. 19, 2017.

Author contributions: J.D.M., J.J., and X.-J.W. designed research; J.D.M. and J.J. performed research; J.D.M., J.J., and X.-J.W. wrote the paper.

This work was supported by National Institutes Health Grant R01-MH-062349, Naval Research Grant N00014-17-1-2041, a Simons Foundation Collaboration on the Global Brain Grant, and Science and Technology Commission of Shanghai Municipality Grants 14JC1404900 and 15JC1400104 (to X.-J.W.). We thank Dr. Braden Purcell and Robert Yang for careful reading of an earlier version of the manuscript.

*J.D.M. and J.J. contributed equally to this work.

The authors declare no competing financial interests.

Correspondence should be addressed to Xiao-Jing Wang, Center for Neural Science, 4 Washington Place, New York University, New York, NY 10003. E-mail: xjwang@nyu.edu.

DOI:10.1523/JNEUROSCI.0343-17.2017

Copyright © 2017 the authors 0270-6474/17/3712167-20$15.00/0

evidence and target selection (Schall, 2001; Gold and Shadlen, 2007). The general similarity of neural activity of PPC and PFC during WM and DM has supported the view that they make comparable contributions to these functions.

Important open questions are to identify how PPC and PFC interact during WM and DM and what their specialized roles may be. For instance, WM-related persistent activity in these areas may be a locally generated phenomenon or, alternatively, the result of distributed interareal interactions (Christophel et al., 2017). Despite general similarities between neural responses in PPC and PFC, important differences, including in distractor processing and evidence accumulation, have been found that provide insight into their unique contributions to WM and DM (Katsuki and Constantinidis, 2012b; Suzuki and Gottlieb, 2013; Hanks et al., 2015). It is unclear to what degree the function specialization of areas may be due to intrinsic differences in local microcircuitry (Katsuki et al., 2014; Murray et al., 2014b). In addition to differences between areas, single-neuron recordings have revealed a diversity of functionally defined cell types within frontoparietal circuits, across both WM and DM (Schall and Thompson, 1999; Ferraina et al., 2002; Lawrence et al., 2005). These findings raise questions of the division of labor among brain areas, or functional cell types, within distributed cortical circuits during cognition.

Biophysically based computational models of cortical circuits have characterized neural circuit mechanisms for WM and DM functions. A class of models called attractor networks can perform these functions through strong recurrent synaptic interactions (Wang, 2008). In the attractor network framework, strong synaptic connections among neurons can provide reverberatory excitation, potentially mediated by slow NMDA receptors (Wang, 1999), that maintains a stimulus-selective persistent activity pattern for WM (Amit and Brunel, 1997; Wang, 2001; Machens et al., 2005). Strong lateral inhibition, mediated by GABAergic interneurons, can enforce selectivity of the WM representation, preventing an unstructured spread of excitatory activity (Compte et al., 2000; Brunel and Wang, 2001). Attractor networks can also perform slow integration and categorical, winner-take-all competition for perceptual DM (Wang, 2002; Wong and Wang, 2006). Indeed, strong recurrent excitation and lateral inhibition are required for winner-take-all DM in these models. Attractor networks therefore constitute a flexible type of "cognitive circuit" capable of performing both WM and DM (Wang, 2013). In contrast to these theoretical advances in characterizing how local microcircuits can support cognitive processing, the computational principles for distributed cognitive processing in multiregional cortical networks remain poorly understood.

To address these issues, we developed a biophysically based computational model of two reciprocally connected modules, potentially representing circuits in PPC and PFC. We found that a single local circuit faces a tradeoff between optimization for WM versus DM function. This performance tradeoff can be ameliorated in the distributed circuit, whose network properties are functionally desirable for both WM and DM. With a single set of network parameters, the distributed model can capture salient empirical observations from single-neuron recordings in PPC and PFC during WM and DM. In the context of WM, the model captures the relative roles of the two areas in distractor filtering. With regard to DM, the model captures key properties of functional cell types and the timing of their activity during visual search tasks. We propose that this cortical circuit model can provide insight into canonical features of distributed cognitive processing.

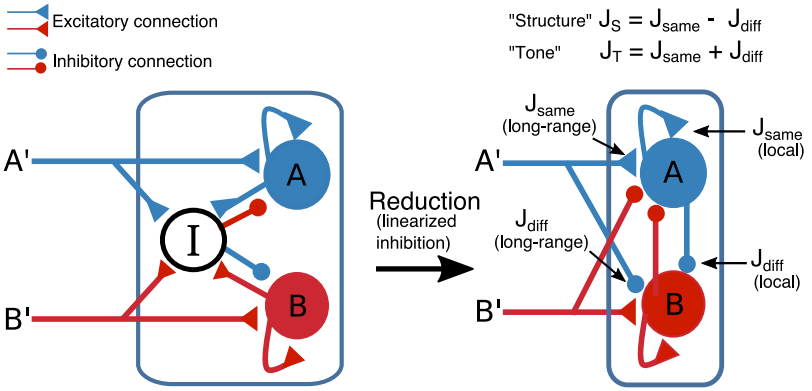

Figure 1. Circuit schematic of the firing-rate model. A module is defined as a set of two excitatory populations where each population is selective to one of two spatial, directional, or object stimuli (left). Each excitatory population is recurrently connected and also receives inhibition from a common pool of interneurons. The effects of inhibition and recurrent excitation are to generate bistability for WM and winner-take-all dynamics and ramping activity through slow reverberation for DM. Population A (B) receives input either from spatially selective stimulus A′ (B′) or from another population A (B) in another module. The circuit dynamics can be simplified (right) by linearizing inhibition, so that effectively inhibition is represented by negative weights. Thus, the effect of the pool of interneurons is implicit in the inhibitory connections between the excitatory populations. In general, synaptic weights J can connect two selective populations of either the same ($J_{same} > 0$) or opposite ($J_{diff} < 0$) stimulus selectivity and can be either local or long range. The structure $J_S = J_{same} - J_{diff}$ denotes the total recurrent strength, while the tone $J_T = J_{same} + J_{diff}$ denotes the net input onto a population. Synapses labeled with triangles and circles denote net-excitatory and net-inhibitory connections, respectively.

Materials and Methods

Model architecture. We constructed a distributed circuit model that is able to perform WM and DM computations. The model is comprised of two reciprocally interacting modules (Fig. 1; also see Fig. 3*A*). Each module contains two selective, excitatory populations, labeled A and B (Wong and Wang, 2006; Wong et al., 2007). Within a module, the two populations have self-excitation and interact through a local inhibitory population that allows for cross-inhibition between the two excitatory populations. Each recurrently connected excitatory population receives inhibition from a common pool of interneurons. Inhibition is linearized so that projections between the two excitatory populations A and B are effectively represented by negative weights (Wong and Wang, 2006). The two modules interact through long-range projections that are structured according to the stimulus selectivity of populations within each module. Long-range projections between modules are structured so that populations with the same selectivity are connected through excitatory projections, whereas populations with different selectivity are connected via net inhibitory projections. The two modules are labeled 1 and 2, and external input related to the stimulus enters into Module 1 (see Fig. 3*A*).

Dynamics and stimuli. We constructed a population firing-rate model for each population $i = A, B$ in the two modules. The firing-rate dynamics of the population i in Modules 1 and 2 are dominated by the dynamics of the average NMDA synaptic gating variable S_i^n ($n = 1,2$). This approximation is based on the fact that the dynamics of the NMDA synaptic gating variable are slower than other timescales in the system (Wang, 2002; Wong and Wang, 2006; Wong et al., 2007). The gating variable S_i^n is described as follows:

$$\frac{dS_i^n}{dt} = -\frac{S_i^n}{\tau} + \gamma(1 - S_i^n)r(I_i^n), \quad (1)$$

where $\tau = 60$ ms is the NMDA time constant, $\gamma = 0.641$ controls the rate of saturation of S, and $r(I_i^n)$ is the firing rate of the population i as a function of the input current I_i^n. The firing rate as a function of input current is given by the following frequency-current (F–I) curve relation (Abbott and Chance, 2005):

$$r(I) = F(I) = \frac{aI - b}{1 - \exp[-c(aI - b)]} \quad (2)$$

with $a = 270$ Hz/nA, $b = 108$ Hz, and $c = 0.154$ s. The input current to population $i = A, B$ in Modules 1 and 2 is given by the following:

$$I_i^n = \sum_{m,j} S_j^m J_{ij}^{(m\to n)} + I_0 + I_{\text{noise,i}}^n + I_{\text{app,i}}^n \quad (3)$$

where $J_{ij}^{(m\to n)}$ is the connection weight from population j in Module m to population i in Module n, $I_0 = 0.334$ nA is the background current, $I_{\text{noise,i}}^n$ is the noise current to population i in Module n, and $I_{\text{app,i}}^n$ is the applied current to population i in Module n from external sources. For simulations of the one-module network in Figure 2, $m = n = 1$ so that all connections are local.

The noise current to each population follows Ornstein–Uhlenbeck dynamics with the time constant of AMPA synapses, as follows:

$$\tau_{\text{AMPA}}\frac{dI_{\text{noise,i}}(t)}{dt} = -I_{\text{noise,i}}(t) + \eta_i(t)\sqrt{\tau_{\text{AMPA}}\sigma_{\text{noise}}^2}, \quad (4)$$

where $\tau_{\text{AMPA}} = 2$ ms, η is Gaussian white noise with zero mean and unit variance, and σ_{noise} sets the strength of the noise. As in Wong et al. (2007), we set $\sigma_{\text{noise}} = 0.009$ nA.

We consider the external applied current I_{app} for the following scenarios: (1) WM and DM in a one-module (local) circuit; (2) WM and distractors in the PPC–PFC circuit; (3) evidence accumulation in the PPC–PFC circuit; and (4) target selection in a visual search task in the selection–action circuit. We will specify the external currents below after the connectivity section.

Connectivity. The connectivity in our model is specified by the sign and magnitude of the connection weights between the selective excitatory populations. The connections can be local (within a module) and, for a two-module circuit, long range (across modules). To this end, it is useful to express the connection weights with the following terms:

$$J_S^k \equiv J_{(\text{same})}^k - J_{(\text{diff})}^k \quad (5)$$

$$J_T^k \equiv J_{(\text{same})}^k + J_{(\text{diff})}^k \quad (6)$$

where $J_{(\text{same})}$ denotes the positive connection weight between same-selectivity populations (e.g., from population A in Module 1 to population A in Module 1 or 2). $J_{(\text{diff})}$ denotes the negative connection weight between different-selectivity populations (e.g., from population A in Module 1 to population B in Module 1 or 2), and $k = 1\to1, 1\to2, 2\to1, 2\to2$ defines whether the connection is local or long range. We define J_S as the structure of the network, since it reflects the magnitude of same-selectivity excitation and different-selectivity cross-inhibition, and thus the total recurrent strength. Analogously, we define J_T as the tone of the network, which reflects the net input onto a particular population.

In Figure 2, we consider a one-module circuit with two structure values ($J_S = 0.35$ and $J_S = 0.4182$ nA and tone $J_T = 0.28387$). A two-module network is constructed by connecting two modules (Module 1 and Module 2) via long-range projections. Local connectivity for Module 1 is set by the following: $J_S^{(1\to1)} = 0.35$ nA and $J_T^{(1\to1)} = 0.28387$ nA. Local connectivity for Module 2 is set by the following: $J_S^{(2\to2)} = 0.4182$ nA and $J_T^{(2\to2)} = 0.28387$ nA. The only difference in local network properties between modules is that Module 2 has an enhanced structure compared with Module 1: $J_S^{(2\to2)} > J_S^{(1\to1)}$.

For both long-range projections between modules, we constrain them to have the following pathway-specific excitation–inhibition (E/I) balance:

$$J_T^{(1\to2)} = 0 \text{ nA}, \quad (7)$$

$$J_T^{(2\to1)} = 0 \text{ nA}, \quad (8)$$

so that the excitatory weight to a given population is counteracted by the inhibition of equal magnitude but opposite sign. If a pathway from one module to another exhibits balance ($J_T = 0$), the impact of one module to another is only nonzero when the populations have unequal activity (see also Vogels and Abbott (2009)). The structure of the projection from Module 1 to Module 2 is set by $J_S^{(1\to2)} = 0.15$ nA. The structure of the projection from Module 2 to Module 1 is set by $J_S^{(2\to1)} = 0.04$ nA.

We can easily translate the structure J_S and tone J_T into individual synaptic weights. For example, $J_{AB}^{(1\to2)}$ denotes the feedforward projection between the population A in the first module onto the population B in the second module and is given by the following:

$$J_{AB}^{(1\to2)} = \frac{J_T^{(1\to2)} - J_S^{(1\to2)}}{2} = -0.075 \text{ nA}.$$

The decomposition of connectivity in terms of structure J_S and tone J_T is useful for capturing the impact of changes in activity that are symmetric or asymmetric between the populations of a module. The input current I to a given population from Module m to Module n can be written as follows (see first term on the right-hand side of Eq. 3):

$$I^{(m\to n)} = \frac{1}{2}\left(J_S^{(m\to n)}S_S^m + J_T^{(m\to n)}S_T^m\right), \quad (9)$$

where S is similarly redefined through $S_S^m \equiv S_{(\text{same})}^m - S_{(\text{diff})}^m$ and $S_T^m \equiv S_{(\text{same})}^m + S_{(\text{diff})}^m$. When activity in a module is equal in the two populations (i.e., $S_S^m = 0$), the net input from that module is determined only by the tone J_T and not by the structure J_S.

For the results in this study, we associate Module 1 with the PPC and Module 2 with the PFC.

Working memory and decision-making in a local network. To characterize WM in a local network (Fig. 2*A*,*B*), we studied the generation of stimulus-selective persistent activity states and their robustness against intervening distractor inputs. The target to be held in WM is the first stimulus presented. We set the target as a current applied to population A, as follows: $I_{\text{app,A}} = I_{\text{target}}$ of 500 ms duration and amplitude 0.0295 nA. Distractors are defined as inputs $I_{\text{app,B}} = I_{\text{distractor}}$ of equal amplitude and duration applied to population B arriving after the target.

In Figure 2, *C* and *D*, we characterized the robustness of WM against distractors in a local one-module circuit, over a two-dimensional parameter space of structure J_S and applied current I_{app}. For these results, bifurcations and continuations were calculated using PyDSTool, a Python-based platform developed for the analysis of dynamic systems (Clewley, 2012).

We modeled perceptual decision-making in Figure 2, *E* and *F*. The strength of evidence is modeled as an external current to the two populations as follows:

$$I_{\text{app,i}} = I_e\left(1 \pm \frac{c'}{100\%}\right), \quad (10)$$

where $I_e = 0.0118$ nA scales the overall strength of the input and c', referred to as the contrast, sets the bias of the input for one population over the other, which is equivalent to the coherence in the study by Wong and Wang (2006).

In Figure 2*H*, we defined the discrimination threshold c'_{disc} as the minimum contrast to achieve 81.6% accuracy in a simulated two-alternative forced choice task. This number follows from the equation commonly used to fit psychometric curves, as follows:

$$P(c') = 1 - 0.5 \cdot \exp(-c'/c'_{\text{disc}})^{\beta}, \quad (11)$$

where P is the accuracy in the task and β determines the slope of the psychometric curve. We varied the recurrent structure J_S from $J_S = 0.35$ to $J_S = 0.42$ nA in steps of 0.01 nA and used Equation 11 to fit the behavior of our simulated model and determine the discrimination threshold.

In Figure 3*D*, to understand how the dynamics of the distributed two-module circuit depends on the structure, we parametrically varied the Module 1 (PPC) local structure $J_S^{(1\to1)}$ from 0.15 to 0.45 nA and the PFC → PPC feedback structure $J_S^{(2\to1)}$ from 0 to 0.08 nA, in steps of 0.02 nA, while keeping all other parameters constant.

Working memory and distractors in the PPC–PFC circuit. For Figures 4 and 5, we simulated a WM task with distractors, based on a primate electrophysiology study using a visuospatial WM task in which a subject must hold in WM the position of a target and filter intervening distractor stimuli appearing at other positions during the delay period (Suzuki and Gottlieb, 2013). We implemented a discrete version of this task with selectivity to two stimuli. A flash of 100 ms appears on one of two positions of a screen indicating the target position. As in the WM simulations for one module, the target to be held in WM is the first stimulus presented. We set the target as a current $I_{\mathrm{app,A}} = I_{\mathrm{target}}$ of 100 ms duration that is applied to population *A* in the PPC module. Distractors are defined as inputs $I_{\mathrm{app,B}} = I_{\mathrm{distractor}}$ of equal duration applied to population *B* arriving after the target and at an opposite location of the visual field. Similar to the study by Suzuki and Gottlieb (2013), we considered four times of target-distractor onset asynchrony (TDOA), which was defined as the onset of the distractor relative to onset of the target, as follows: 100, 150, 200, and 300 ms. Target and distractor amplitudes for Figure 4 were 0.09 nA. To obtain both correct and error trials in Figure 5, targets and distractor amplitudes were sampled randomly and independently from a Gaussian distribution with a mean of 0.09 nA and an SD of 0.04 nA and were, in general, different from each other and from trial to trial.

In Figure 4*B*, we calculated the differences between the two modules in terms of how distractors are suppressed. For this, we fitted the time courses of the firing rates $r(t)$ of the populations selective to the distractor for each of the modules with an exponential function, as follows:

$$r(t) = a_1 \cdot \exp\left[\frac{-(t - t_{\mathrm{target}})}{\tau_{\mathrm{sup}}}\right] + a_2, \tag{12}$$

where t_{target} marks the time of target onset and also suppression of the distractor, τ_{sup} is the timescale of suppression, and a_1 and a_2 are parameters of the fit.

In Figure 4*C*, we performed an autocorrelation on the firing rate to reveal the intrinsic or fluctuation timescales of spontaneous activity. The firing rate was first filtered with a Gaussian function with window $\sigma_{\mathrm{filter}} = 20$ ms. To compute the autocorrelation of the firing rate, we subtracted the mean from the firing rate and then normalized. We then used Equation 12 to fit the normalized firing-rate autocorrelation and extract the respective timescales τ_{fluct}.

We plotted error rates versus the time of distractor presentation relative to the target in Figure 5*C*, where an error is recorded when the population selective to the distractor is at the high memory state at $t > 3000$ ms.

Evidence accumulation in the PPC–PFC circuit. As shown in Figure 6, we simulated a simple version of a discrete evidence accumulation task based on a two-alternative forced choice task with perceptual decisions that rely on evidence accumulation from discrete auditory stimuli (Brunton et al., 2013). The auditory input consists of a sequence of clicks on both sides (left and right), parametrized by click frequency in units of clicks per second. For example, 10:24 constitutes a trial where 10 represents the click frequency for the left side and 24 constitutes the click frequency for the right side. In the task, the subject is rewarded for reporting which side (left vs right) had the higher frequency signal. This task can be solved by integrating evidence, where each click represents a unit (quantum) of evidence (Hanks et al., 2015).

In our model, clicks are represented by a set of Poisson-distributed times, parametrized by the rate and side of origin, either left or right. The rates for each side are such that they add to up to 34 clicks/s in total. For example, a difficult trial is 18:16 clicks/s, while an easy trial is 30:4 clicks/s. The click times t_i for each side are convolved with a current pulse kernel Π of amplitude 0.0118 nA and pulse duration 50 ms, so that the current $I_{\mathrm{L,R}}$ is as follows:

$$I_{\mathrm{L,R}}(t) = \sum_i \Pi(t - t_i). \tag{13}$$

These currents, corresponding to left and right clicks, are fed onto the corresponding selective populations (left or right) of the PPC, as well as to the accumulator, which we now define. The accumulator is an implementation of a drift-diffusion model (Ratcliff, 1978), with parameters for drift, noise, and input stimulus (Brunton et al., 2013). Time evolution of the accumulator value a at time t is given by the following:

$$\frac{da}{dt} = \lambda a + I_{\mathrm{L}}(t) - I_{\mathrm{R}}(t) + \xi(t), \tag{14}$$

where λ is the drift and $\xi(t)$ is Gaussian white noise with a mean of 0 and a standard deviation of 0.1. We set $\lambda = 0$, corresponding to leak-free integration, for the simulations in Figure 6, but small values of λ did not significantly alter these plots.

To obtain the relationship between an accumulator and PPC/PFC firing rates, we selected four time points relative to response onset in the PPC/PFC ($t = 200, 250, 300, 350$ ms) and obtained the distribution of firing rates and accumulator values for each of those time points (Hanks et al., 2015). For each of the time points, we binned the accumulator values from −7 to 7, with a bin size of 2 and calculated the mean firing rate for each bin. Applying this procedure to the four selected time points, we obtain Figure 6*C*, where the firing rate is shown as a function of time and is color coded as a function of accumulator value. Finally, to obtain the relationship between firing rate and accumulator value for the PPC and PFC, we averaged the firing rate over time for each of the accumulator values $i = -7, \ldots 7$ and scaled the corresponding firing rates r_i to account for differences in dynamic range (Hanks et al., 2015) as follows:

$$r_i^{\mathrm{norm}} = \frac{r_i - \min\{r_i\}}{\max\{r_i\} - \min\{r_i\}},$$

to finally obtain Figure 6*B*.

Target selection in a visual search task. In Figures 7 and 8, we simulated a perceptual DM task analogous to target selection in visual search (Sato et al., 2001) with the two-module circuit model. Each module contains two populations that are selective to a target and a distractor, respectively. As in the WM paradigm, external stimuli enter as currents into the PPC module. These applied currents reflect the external stimulus as follows:

$$I_{\mathrm{app,i}} = C(A_{\mathrm{target}} - I_{\mathrm{motion}}) \cdot \left[\exp\left(\frac{-(t - t_{\mathrm{target}})}{\tau_{\mathrm{decay}}}\right) - \exp\left(\frac{-(t - t_{\mathrm{target}})}{\tau_{\mathrm{rise}}}\right)\right] + I_{\mathrm{motion}}, \tag{15}$$

where $I_{\mathrm{motion}} = I_e \cdot \left(1 \pm \frac{c'}{100\%}\right)$, $I_e = 0.0118$ nA scales the overall strength of the input and c', referred to as the contrast, and sets the bias of the input for one population over the other, which is equivalent to the coherence in the study by Wong and Wang (2006); A_{target} and t_{target} determine the amplitude and the onset of the target, respectively; the time constants τ_{decay} and τ_{rise} determine the approximate decay and rise of the target-induced transient response; and C is a normalization factor. A zero-contrast stimulus applies equal input I_e to each population in Module 1 (PPC, also denoted as selection cells). In all of the simulations and with $c' > 0$, the target-selective population receives the greater biased input. Due to noise, however, this does not guarantee that the target population will win, especially for low-contrast values.

Because our model provides instantaneous firing rates for a population, we can define measures of neural activity directly for individual trials. For DM simulations (Figs. 7, 8), we define and calculate the reaction time as the time at which the firing rate of a population in Module 2 (PFC, also denoted as action cells) crosses a threshold (Hanes and Schall, 1996). We measure the discrimination time for selection cells in Module 1 (PPC) through a threshold on the absolute difference between firing rates between the two populations. We measure the onset time for action cells in Module 2 through a threshold on firing rate for the winning population. The thresholds for reaction time, discrimination time, and

onset time are 40, 12, and 7 Hz, respectively. Reaction times are divided into quintiles for each of the contrast conditions. This is analogous to the division into short, intermediate, and long reaction time groups (Sato et al., 2001).

Disruption of pathway-specific excitation–inhibition balance. In Figure 8, we examined the effects of pathway-specific excitation–inhibition balance in intermodule projections. To this end, we systematically decreased the inhibitory weights of the projections from Module 1 to Module 2 to alter the E/I ratio. The altered inhibitory weights $J^{*}_{(\mathrm{diff})}$ are as follows:

$$J^{*}_{(\mathrm{diff})} = c_{\mathrm{inh}} \cdot J^{1\to 2}_{(\mathrm{diff})}, \quad (16)$$

where $J^{1\to 2}_{(\mathrm{diff})}$ is the original (i.e., unaltered) inhibitory synaptic weight between population A (B) in Module 1 and population B (A) in Module 2 and $c_{\mathrm{inh}} \in \{0.2, 0.4, 0.8, 1\}$. According to Equation 16, $c_{\mathrm{inh}} = 1$ corresponds to balanced excitation and inhibition onto both populations in Module 2, while $c_{\mathrm{inh}} < 1$ corresponds to imbalance (i.e., a net positive weight onto the populations; see Eqs. 6 and 7).

We fit the reaction time versus the contrast relation in Figures 7*C*, top, and 8*A* using the following exponential function:

$$R(c') = a_1 + a_2 \cdot \exp(-a_3 \cdot (c' - a_4)),$$

where R is reaction time, c' is the contrast and a_1, a_2, a_3, a_4 are free parameters of the fit. Similarly, we fit the fraction of correct trials versus contrast relation in Figures 8*A* and 7*C*, bottom, with a sigmoid, as follows:

$$F(c') = a_1 + \frac{a_2}{1 + \exp(-a_3 \cdot (c' - a_4))},$$

where F is the fraction of correct trials, and a_1, a_2, a_3, and a_4 are parameters of the fit.

Tradeoff amelioration and parameter exploration in a two-module circuit. In Figure 9, we systematically varied the local structures $J_S^{1\to 1}$ and $J_S^{2\to 2}$ to explore the behavior of the two-module circuit in response to sustained and transient inputs for DM and WM computations, respectively. We varied both $J_S^{1\to 1}$ and $J_S^{2\to 2}$ from 0.34 to 0.42 nA in steps of 0.01 nA, while keeping other parameters fixed, and calculated the discrimination threshold and robust stimulus range as follows. The discrimination threshold for DM was calculated as in Equation 11 for each local structure pair $\{J_S^{1\to 1}, J_S^{2\to 2}\}$. To calculate the robust stimulus range for WM, we used a paradigm similar to the one used for Figure 5*A* in which a target is presented at $t = 0$ ms and is held in WM. The minimum current required to induce a transition to the target-related high memory state is given by $I_{\mathrm{target}} = I_{\mathrm{min}}$. A distractor subsequently intervenes during the delay period at $t = 1300$ ms. For each local structure pair $\{J_S^{1\to 1}, J_S^{2\to 2}\}$, we varied the distractor current $I_{\mathrm{distractor}}$ from 0 to 2 nA in steps of 0.002 nA. The maximum distractor current below which the target-related memory state is robust (i.e., at $t = 3000$ ms the distractor does not override the target representation) is $I_{\mathrm{distractor}} = I_{\mathrm{max.}}$ The robust stimulus range is calculated as $I_{\mathrm{max}} - I_{\mathrm{min}}$ for each local structure pair $\{J_S^{1\to 1}, J_S^{2\to 2}\}$. For an adequate comparison between the distractor-filtering abilities in the control and "no PFC feedback" cases, we scaled the robust stimulus range and calculated the relative stimulus range I_{rel} as follows:

$$I_{\mathrm{rel}} = \frac{I_{\mathrm{max}} - I_{\mathrm{min}}}{I_{\mathrm{min}}}, \quad (17)$$

for each local structure pair $\{J_S^{1\to 1}, J_S^{2\to 2}\}$. In Figures 9, *C* and *F*, we plotted the discrimination threshold and relative stimulus range as a function of the difference in structure along the line $J_1 + J_2 = 0.76$ nA, thus preserving the total structure in the two-module system.

Results

We have designed and characterized a distributed circuit model that supports persistent activity for WM and slow integration over time and winner-take-all competition for DM. The model is comprised of two reciprocally connected modules that model the PPC and PFC. Each module consists of two populations of excitatory neurons, and each population is selective to one of two spatial, directional, or object stimuli (Fig. 1; see also Materials and Methods). Local connectivity (i.e., connectivity within a module) is specified by recurrent excitation and cross-inhibition. Long-range connectivity (i.e., connectivity across modules) is specified by feedforward and feedback projections that are net excitatory between same-selectivity populations and net inhibitory between different-selectivity populations. Model parameters were chosen so that the same architecture and parameter set could capture several important neurophysiological dynamics observed in both WM and DM.

Tradeoffs in working memory and decision-making for a local circuit

We first characterized WM and DM function in a local, one-module network. Prior studies have shown that a local attractor network can perform both WM and DM (Wang, 2002; Wong and Wang, 2006). However, how well the same network can perform both functions or what tradeoffs exist between the optimization of local circuit properties for WM versus DM has been studied less (Standage and Paré, 2011). Of particular interest is the role of the local recurrent structure, parameterized by J_S, here defined as the total recurrent strength including self-excitation within a population and cross-inhibition between populations (see Eq. 5 in Materials and Methods).

Working memory and robustness against distractors

In the attractor network framework, the key requirement for WM function is multistability [i.e., the coexistence of multiple stable fixed points (attractor states) in the absence of a stimulus; Amit and Brunel, 1997; Wang, 1999; Brunel and Wang, 2001]. In the absence of a stimulus, the simplified two-population network studied here supports the following three stable states: one symmetric baseline state and two asymmetric memory states (Wong and Wang, 2006). Before stimulus onset, the network is in the symmetric baseline state with both populations, *A* and *B*, at low activity (Fig. 2*A*,*B*). After one population is sufficiently activated, it is able to remain persistently in a stimulus-selective, high-activity state in the absence of a stimulus.

In addition to maintenance over time, robust WM requires shielding internal representations from interference by both internal noise and external distraction. There is evidence that PPC and PFC have different susceptibilities for distractors to disrupt persistent activity. In general, PFC exhibits persistent activity that is robust against distractors, whereas posterior association areas in the temporal and parietal lobes exhibit persistent activity that is disrupted by distractors (Miller et al., 1993, 1996; Constantinidis and Wang, 2004; Qi et al., 2010; Suzuki and Gottlieb, 2013). Motivated by these findings, we explored the mechanisms of robustness against distraction and how they depend on network structure. In our model, distractors can be modeled as an intervening input during the WM delay to a different population than the one activated by the WM target. For instance, if population *A* is active in the memory state, a distractor is modeled as a subsequent input to population *B*.

We found that recurrent structure J_S (i.e., the total self-excitation and cross-inhibition within the one-module network) plays a crucial role in determining the robustness against distractors in the local network (Fig. 2*A*–*D*). With moderate structure, the network generates WM-related persistent activity, but it is distractible. As shown in Figure 2*A*, the distractor input switches the network to representing the distractor (red) instead of the target (blue). With high structure, in contrast, the distractor is filtered out and the target representation is robustly maintained

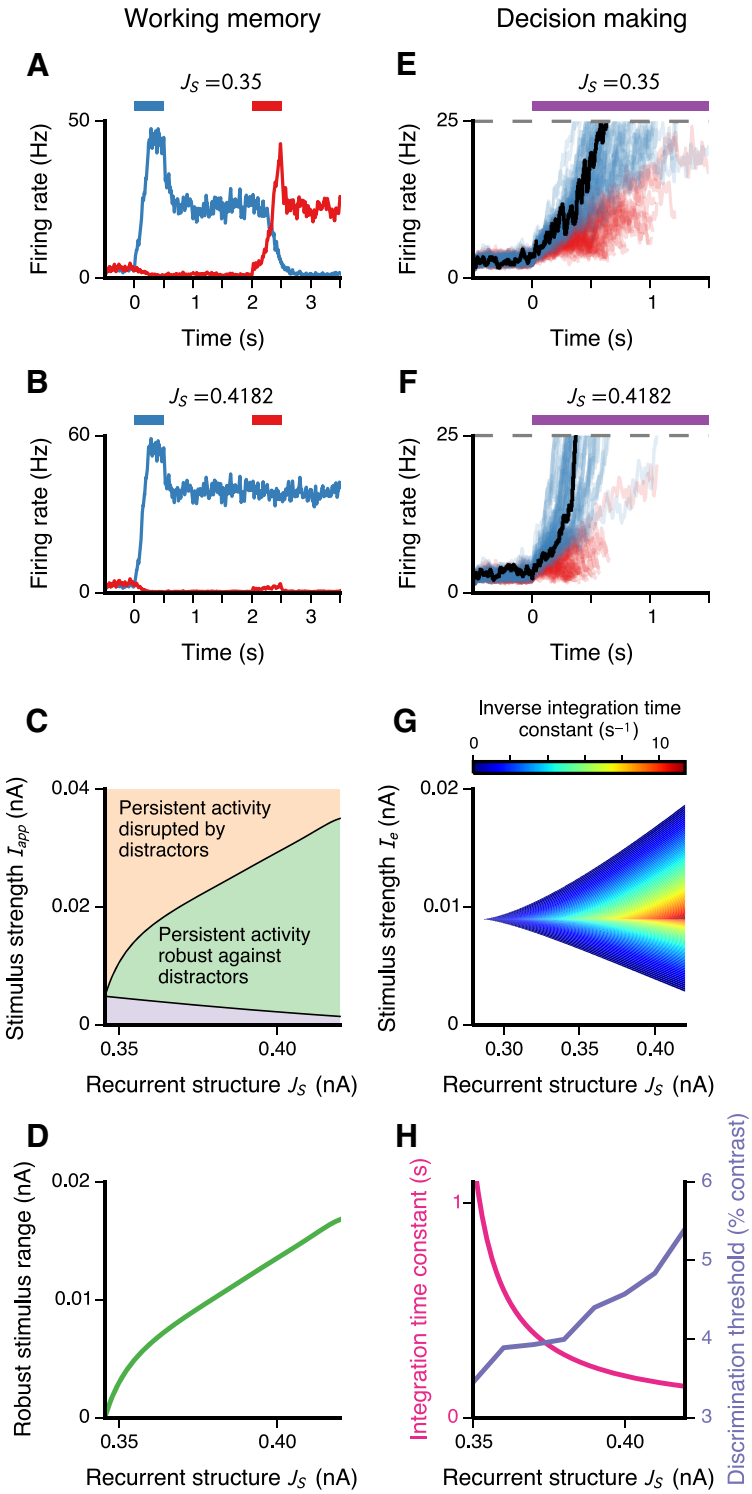

Figure 2. Tradeoffs between WM and DM function in a local attractor network model. ***A***, Neural activity for a single WM trial. The colored bars mark presentation of input current to the population A (blue) and B (red), with strength $I_{app} = 0.0295$ nA. For a recurrent structure $J_S = 0.35$ nA, the circuit generates a stimulus-selective persistent memory state, but it is vulnerable to intervening distractors. ***B***, At increased recurrent structure J_S, WM activity in the circuit is robust against distractors. ***C***, Robustness of WM as a function of recurrent structure J_S and stimulus strength I_{app}. In the purple lower region, the stimulus is too weak for the target to induce a transition from the (stable) baseline state to the memory state. In the green middle region, the network can perform WM that is robust against intervening distractors. In the orange upper region, the stimulus current is strong enough for a distractor to disrupt target-related memory. ***D***, WM robustness increases with increasing recurrent structure J_S. The robust stimulus range is defined as the range of stimulus strength I_{app} in which persistent activity is robust against distractors (i.e., by the height of the green region in *C*). ***E***, Neural activity during DM for zero-contrast stimulus (i.e., equal strength input to both populations), for 100 trials in which the blue population first reached threshold. The colored bar marks stimulus presentation, with strength $I_e = 0.0118$ nA. The black trace marks the firing rate of the winning population for the trial with median reaction time. Note that the time range shown here is shorter than for the WM simulations in *A* and *B*. ***F***, At increased recurrent structure J_S, integration is shorter, limiting evidence accumulation, and ramping to threshold occurs sooner. ***G***, Integration time constant as a function of recurrent structure J_S and stimulus strength I_e for a zero-contrast signal. The integration time constant is defined as the absolute value of the inverse eigenvalue of the unstable mode of the saddle point in the system (Wong and Wang, 2006). The inverse of the integration time constant is plotted. The white region marks where the symmetric state is stable, and therefore the network is not in a winner-take-all regime. ***H***, DM performance degrades with increasing recurrent structure. For a fixed stimulus strength (here with $I_e = 0.0118$ nA), the integration time constant decreases with J_S. Correspondingly, the discrimination threshold increases, indicating degraded performance. Note that the two single modules shown in ***A***, ***B***, ***E***, and ***F*** have the same circuit parameters as the two modules of the distributed circuit in subsequent figures.

(Fig. 2*B*). The network mechanism for robustness results from the combined effects of lateral inhibition and recurrent excitation set by J_S (Brunel and Wang, 2001). When population *A* is in the high-activity state, distraction requires both activation of population *B* and deactivation of population *A*. Activation of population *B* by the distractor input is counteracted by lateral inhibition. Recurrent excitation within population *A* counteracts the ability of inhibition from population *B* to deactivate it.

A dynamic systems analysis can formally characterize the robustness of WM to distractors through the bifurcation diagram defining how the fixed points of the system change as a function of input current to one population I_{app}. The bifurcation diagram gives the range of input current strengths in which WM is robust against distractors. We can examine how this range varies as a function of recurrent structure J_S. Figure 2, *C* and *D*, shows how this robustness range of I_{app} changes with J_S. We found that the distractibility threshold (boundary between green and orange regions) increases with J_S. The memory induction threshold (boundary between purple and green regions) decreases as J_S increases, widening the robustness range further. The net effect is that the robust stimulus range increases with higher recurrent structure J_S (Fig. 2*D*). This analysis suggests that increased robustness against distractors in PFC, compared with PPC, may be due to higher network structure (Brunel and Wang, 2001).

Decision-making and slow integration of evidence

We now consider the ability of the attractor network to perform perceptual DM functions. In the attractor network framework,

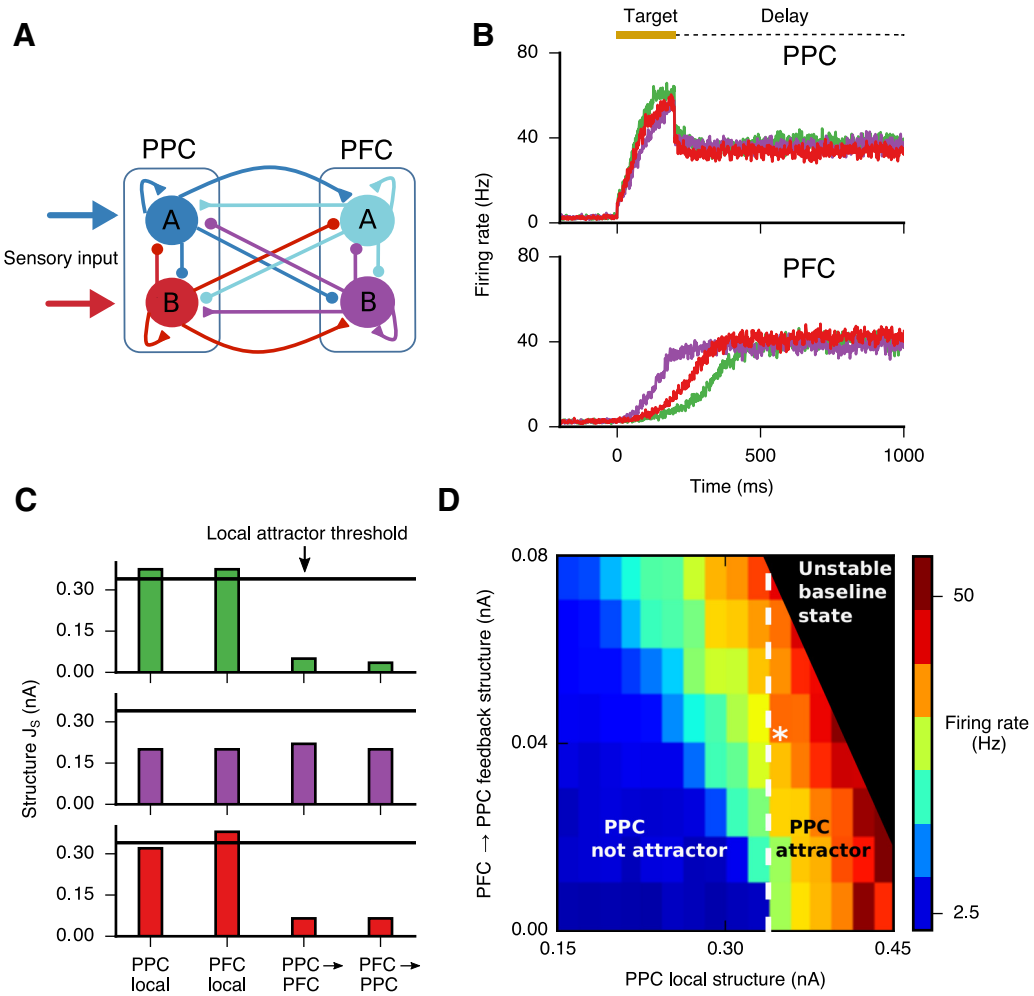

Figure 3. A distributed cortical model reproduces spatially selective persistent activity through local and long-range connections. ***A***, The circuit is composed of two reciprocally connected modules, PPC and PFC, and each module consists of two excitatory neural populations selective to a stimulus A and B, respectively. The circuit model is endowed with self-excitation and cross-inhibition. Neurons in the PPC receive the sensory stimulus and convey the information to the PFC via long-range net-excitatory and net-inhibitory projections, as in Figure 1. ***B–D***, Local and long-range structures jointly contribute to persistent activity. ***B***, PPC and PFC populations reach the same level of persistent activity in the steady state across the three scenarios depicted in ***C***, demonstrating the joint contributions of long-range and local connectivity. Gold bar denotes stimulus presentation. ***C***, Structure values reflecting local (within-module) and long-range (across-module) connectivity for three scenarios are shown: (1) PPC and PFC both independently support persistent activity (green, top); (2) neither PPC nor PFC is capable of persistent activity independently (purple, middle); and (3) only PFC independently supports persistent activity (bottom, red). Black horizontal line denotes the threshold for a local module to support persistent activity independently (i.e., multiple stimulus-selective attractor states). ***D***, Steady-state firing rate for the activated population of the PPC module in the memory state, as a function of PPC local structure and PFC ⟶ PPC feedback. The PFC local and feedforward PPC ⟶ PFC structures are fixed. In the region in the upper right corner, the baseline state is unstable. In the region to the right of the white dashed line, the PPC is an independent attractor. The white asterisk marks the parameter values used for the WM and DM simulations in Figures 4, 5, 6, 7, and 8.

two key requirements for DM are slow accumulation of evidence over time and winner-take-all competition (Wang, 2002; Wong and Wang, 2006). Inputs to the network are characterized by the strength of the evidence or contrast (see Eq. 10 in Materials and Methods). For a zero-contrast stimulus, the two populations receive equal input, differing only through the noise term. The network nevertheless performs categorical selection through one population going to a high-activity state and the other to a low-activity state.

To subserve perceptual DM, the network should be able to integrate evidence over time when the signal-to-noise ratio is low (Gold and Shadlen, 2007). In the model, a decision is made when the corresponding neural population reaches a threshold firing rate. As shown in Figure 2, *E* and *F*, the recurrent structure J_S plays a crucial role in determining the timescale of integration as reflected in ramping neural activity, here with a zero-contrast stimulus. With moderate structure, the network ramps relatively slowly (median decision time, ~560 ms), indicating that the network implements slow integration of sensory evidence. With a higher recurrent structure, the network ramps substantially faster (median decision time ~430 ms), indicating a more limited duration over which the network integrates sensory evidence.

A dynamic systems analysis can formally characterize the integration timescale. In response to a zero-contrast stimulus, the symmetric state of the network is a saddle point in the (S_1, S_2) phase plane (Wong and Wang, 2006). The timescales associated with this saddle point, along with the strength of noise, largely determine the timescales of integration. The antisymmetric mode, whose positive eigenvalue indicates that it is the unstable mode, is the direction in which integration occurs and that leads to categorical choice. The integration timescale τ_{int} can be defined as the inverse of this eigenvalue. Figure 2*G* shows the dependence of τ_{int} on network structure J_S and stimulus current strength I_e, for zero-contrast stimulus. As the system approaches the bifurcations that form the boundaries of the winner-take-all regime, the integration timescale increases toward infinity. For a fixed I_e in a winner-take-all DM regime (Fig. 2*G*), increasing J_S

decreases the integration timescale, which limits the duration over which perceptual integration can occur. If less evidence is integrated, we expect to have more errors for a fixed contrast (coherence in Wong et al. (2007); Gold and Shadlen (2007)) value. This degradation in performance is measured by the increase of the discrimination threshold, defined as the contrast necessary to achieve a predefined performance level, here 81.6% correct (see Eq. 11 in Materials and Methods). Indeed, the shortened integration timescale degrades DM performance, as reflected in the discrimination threshold from the psychometric function (Fig. 2*H*). Therefore, although recurrent structure must be above a threshold value in order for the network to perform winner-take-all selection, further increasing recurrent structure limits the gradual accumulation of sensory evidence.

These findings illustrate a tension between WM and DM function in a local attractor network. There is a tradeoff between the two as recurrent structure J_S is varied: higher structure increases robustness against distractors for WM (Fig. 2*D*), but at the expense of shortening the integration timescale that degrades DM performance (Fig. 2*H*). As we show and discuss below, this performance tradeoff may be ameliorated in a distributed circuit in which local modules have different strengths of recurrent structure.

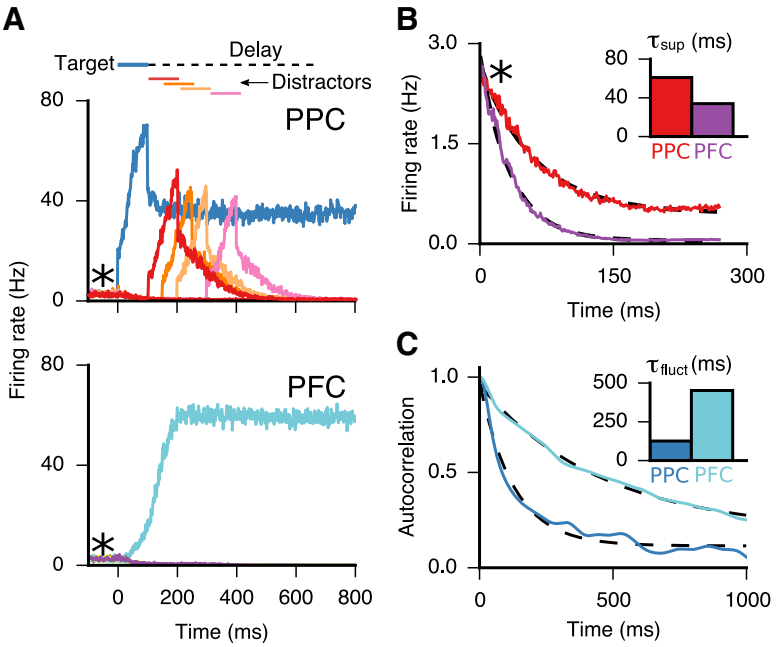

Figure 4. Firing-rate dynamics in the PPC–PFC circuit during a WM task. ***A***, Top, the blue trace shows the response of the target-selective neural population in the PPC in response to a target presented at $t = 0$ ms, with no distractors. Red, orange, light orange, and pink traces show the responses of the distractor-selective population to distractors presented at $t = 100$, 150, 200, and 300 ms, respectively. Bottom, The cyan trace shows activity of the target-selective population in the PFC that receives the stimulus indirectly through the long-range projections from the PPC, in the no-distractor condition. The other traces show the distractor-selective population in response to distractors, as for PPC. However, these responses are not visible due to the strong filtering by surround suppression within PFC. ***B***, Temporal dynamics of the two suppressed populations in PPC and PFC around the time of stimulus presentation (asterisks in ***A***). ***C***, Autocorrelation of the spontaneous firing rate shows the difference in fluctuation timescales τ_{fluct} between the PPC (127 ms, blue) and the PFC (438 ms, cyan). Firing rate traces were smoothed with a Gaussian window of 20 ms width before calculating the autocorrelation, and dashed lines are exponential fits.

Persistent activity in a circuit model with local and long-range connections

Visual WM recruits persistent activations that are distributed across PPC and PFC (Chafee and Goldman-Rakic, 1998), which are mediated by PPC–PFC interactions (Chafee and Goldman-Rakic, 2000; Ferraina et al., 2002; Salazar et al., 2012; Dotson et al., 2014). The distributed circuit model we developed is composed of two reciprocally connected modules that can support persistent activity independently (Fig. 2*A*,*B*). What are the roles of local (within-module) versus long-range (across-module) connections in supporting persistent WM states in this distributed circuit model? We model visual stimuli as inputs to PPC (Fig. 3*A*, Module 1) following the dorsal visual pathway (Felleman and Van Essen, 1991) and the ordering of activations during bottom-up visual processing as well as during target selection (Buschman and Miller, 2007; Ibos et al., 2013; Siegel et al., 2015), although some experiments suggest that input can also rapidly reach PFC through other pathways (Katsuki and Constantinidis, 2012a). Figure 3*B* shows the model circuit response to a stimulus input to one PPC population for different values of the recurrent structure J_S (Fig. 3*C*). PPC responds vigorously to the stimulus and propagates this signal to PFC. Following the offset of the stimulus (i.e., during the WM delay), both PPC and PFC encode the stimulus through selective persistent activity. On the basis of WM delay activity alone, PPC and PFC therefore have similar WM activity, as observed experimentally (Chafee and Goldman-Rakic, 1998; Qi et al., 2010; Suzuki and Gottlieb, 2013).

We examined the roles of local and long-range structure in determining persistent activity. Figure 3, *B* and *C*, shows how different combinations of local and long-range parameters can give rise to similar delay activity in the PPC and PFC modules. A transient stimulus input can switch the state of the circuit from a low-activity baseline state to a persistent and selective high-activity memory state. The stable and persistent delay activity is a reflection of the attractor dynamics in the combined PPC–PFC circuit. Interestingly, these attractor dynamics are present in the combined circuit even if the individual modules are not attractor networks independently (Fig. 3*C*). This demonstrates that in a distributed network, the observation of persistent activity in a local area does not necessarily imply that the area is independently capable of multistability. We then examined the relationship between the firing rate, memory states, and structure in the distributed circuit. In Figure 3*D*, the tradeoff in local and long-range structure is made explicit for the PPC firing rate in a high memory state. In conclusion, persistent activity in the PPC–PFC circuit is supported by both local and long-range structure, with the persistent activity states determined by the total (i.e., combined local and long-range) structure.

Working memory with distractors in the parietal-prefrontal circuit

The ability for a WM circuit to encode and maintain information robustly while filtering out distractors is crucial for WM function (Sakai et al., 2002; Suzuki and Gottlieb, 2013). Our distributed cortical model is able to selectively encode a target stimulus in WM in the presence of distractors (Figs. 4, 5). Within the PPC and PFC modules, one neural population is selective to the target,

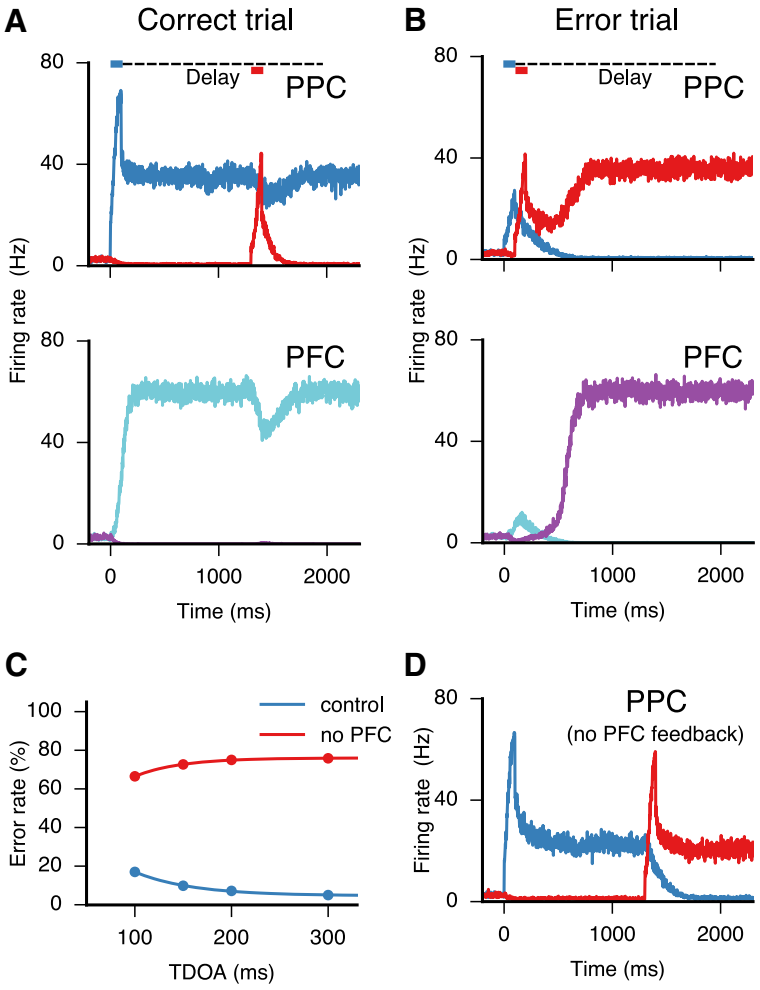

Figure 5. Relationship between neural dynamics and behavioral performance during a WM task with distractors (colors correspond to the schematic in Fig. 3*A*). ***A***, Example of a correct, not distracted, trial. The target-selective population in PPC encodes the target in WM following stimulus onset at $t = 0$ ms (top, blue), while the distractor-selective PPC population transiently but strongly encodes the distractor following its presentation at $t = 1300$ ms (top, red). After distractor offset, feedback from the PFC switches the PPC back to encoding the target, enabling a correct response at the end of the trial. The PFC (bottom) is activated by the response of the PPC to the target, which is maintained in WM by the target-selective population in the PFC as well. Distractor presentation causes a transient suppression of the delay activity in the neurons encoding the target (cyan), but the distractor is not represented strongly (magenta) as it is in the PPC. ***B***, Example of an error, distracted, trial. If the target precedes the distractor by a short interval (100 ms in this example), there is an increased probability of the distractor representation overriding the target representation, so that the distractor is encoded in persistent activity in both PPC and PFC (top, red; and bottom, purple). ***C***, Simulated behavioral performance as a function of TDOA. Distractibility decreases with longer TDOA (blue). Simulated lesion of PFC greatly increases distractibility (red). ***D***, Effects of the removal of PFC→PPC feedback. Absence of PFC feedback onto the PPC forces the PPC to encode the last presented stimulus, leaving it vulnerable to distractors.

and the other to the distractor. Figure 4*A* shows how the distributed circuit responds to target and distractor stimuli during WM. Presentation of the target stimulus activates the selective population in PPC, which transmits this information to the PFC module via the feedforward long-range projections. Following stimulus withdrawal, the target is encoded in PPC and PFC in persistent activity. An intervening distractor, presented during the delay, competes with the mnemonic target representation in the network activity (Fig. 4*A*). If WM is robust, the response of the distractor is transient, and target-coding persistent activity is maintained.

Several notable observations of distractor processing in the model are in line with single-neuron recordings from PPC and PFC during visuospatial WM. First, distractor responses are weaker than the target response (Falkner et al., 2010; Suzuki and Gottlieb, 2013; Zhang et al., 2017). This surround suppression is mainly due to the cross-inhibition from the active target population to the distractor population. Second, the peak distractor amplitude also decreases as the time of distractor presentation relative to target increases (Suzuki and Gottlieb, 2013). This results from the network dynamics of the synaptic gating variables. At shorter distractor onset times, the synaptic gating variables of the target and distractor populations have not reached their steady-state levels (i.e., a high value for the target population and a low value for the distractor population). Therefore, the suppressive effect from the target population on the distractor responses will slightly increase over time, a consequence of the local network having slow synaptic-gating variables. Finally, and most strikingly, distractor responses are markedly different in PFC compared with PPC; PPC represents the distractor strongly during its presentation (Powell and Goldberg, 2000; Falkner et al., 2010; Suzuki and Gottlieb, 2013), while PFC strongly filters distractors (Everling et al., 2002; Suzuki and Gottlieb, 2013). In addition to being indirectly activated by the stimulus, the PFC module has a higher local structure J_S and thus stronger recurrent dynamics than the PPC. The stronger recurrence makes target encoding in PFC more robust against transient distractor inputs, which are effectively filtered. From the dynamic systems perspective, the PFC module has a wider basin of attraction than the PPC (Fig. 2; Brunel and Wang, 2001; Wong and Wang, 2006). To summarize, the transient encoding of distractors is weaker than target encoding, and weaker in the PFC than in the PPC.

Single-neuron recordings have found that WM activity in PPC and PFC generates surround suppression even on baseline activity, in the absence of distractors (Funahashi et al., 1989; Falkner et al., 2010; Suzuki and Gottlieb, 2013). We analyzed differences between the PPC and PFC modules in terms of the dynamics of this surround suppression during target encoding (Fig. 4*B*). Relative to PPC, the surround suppression of the distractor population in PFC is (1) stronger (i.e., toward a lower baseline activity) and (2) more rapid (i.e., with a lower time constant). These features are nontrivial, given that the stimulus directly drives PPC, whereas PFC is activated via projections from PPC. They result from the higher structure in PFC relative to PPC. Both of these features are in line with single-neuron features of distractor suppression (Suzuki and Gottlieb, 2013).

The timescale of surround suppression can be contrasted to the intrinsic timescale of activity fluctuations in the baseline state. Primate cortex shows a hierarchical organization of this dynamic feature, with intrinsic timescales increasing along the cortical hi-

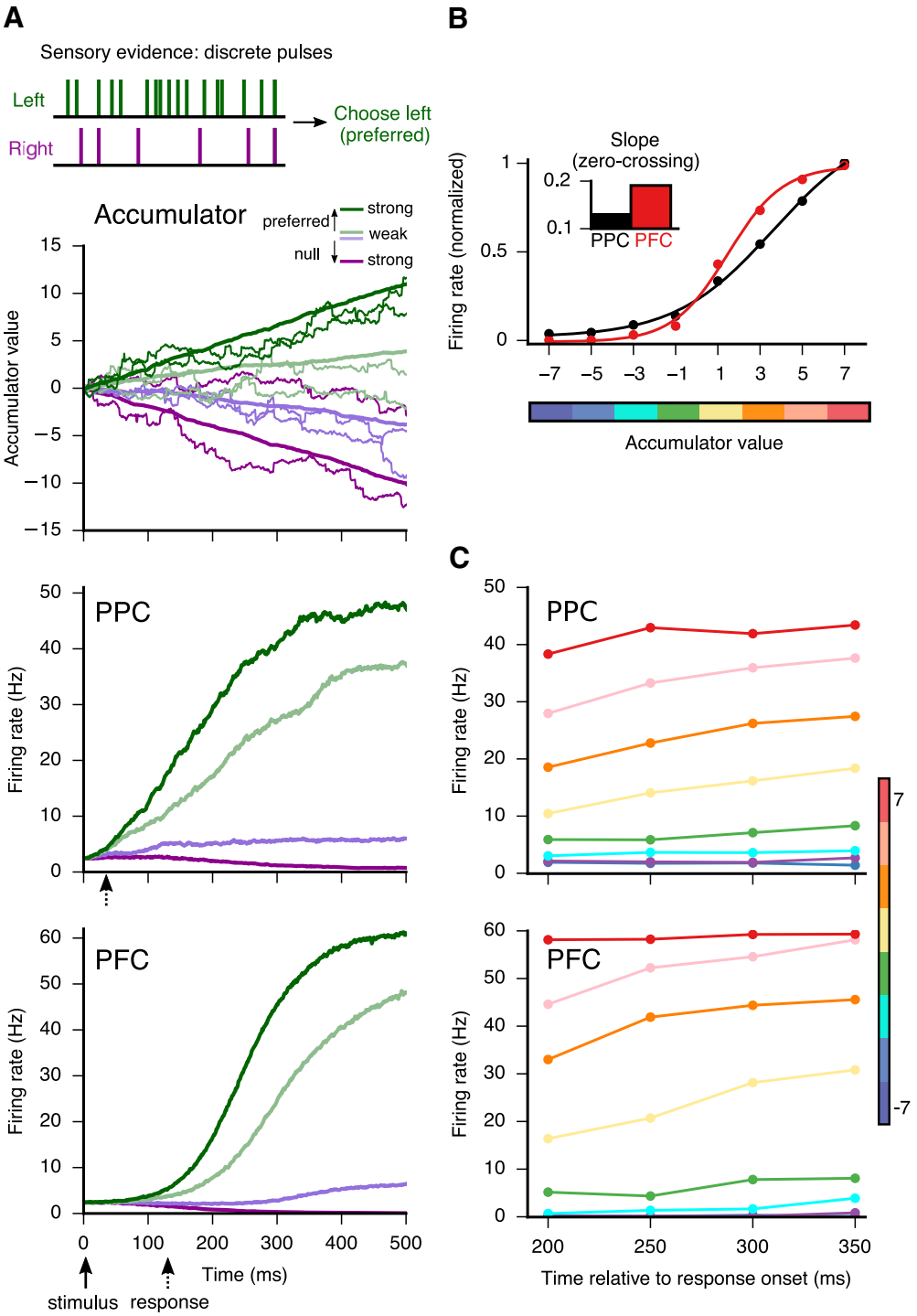

Figure 6. PPC and PFC differentially encode accumulated evidence during perceptual DM. ***A***, The theoretical accumulator (top), PPC (middle), and PFC (bottom) integrate sensory evidence as a function of time and trial difficulty. Thick traces show an average over 60 trials for each difficulty condition, while thin traces in the accumulator show single trials. Traces corresponding to PPC and PFC include both correct and error trials. ***B***, The firing rate vs accumulator relationship is more categorical, with a steeper slope at zero accumulator value, in the PFC than in the PPC, which has a more graded coding. The slopes at zero crossing are 0.13 and 0.19 for PPC and PFC, respectively (see also Hanks et al. (2015)). ***C***, The relationship between firing rate in PPC and PFC and accumulator value as a function of time is stable. The eight accumulator values (from purple to red) correspond to the horizontal axis in ***B***.

erarchy, with dorsolateral PFC exhibiting a longer timescale than posterior parietal area LIP (lateral intraparietal area; Murray et al., 2014b). We examined whether this hierarchical organization was consistent with the distributed circuit model we developed. To this end, we calculated the autocorrelation of spontaneous fluctuations in the PPC and PFC (Fig. 4C). We found that for the given parameter set, the fluctuations timescale of PFC is 3.44 times slower than that of PPC, analogous to the findings of Murray et al. (2014b), who found a change by a factor of 2. Notably, this ordering of intrinsic timescale (longer in PFC than PPC) is opposite to the ordering of the surround-suppression timescale (shorter in PFC than PPC). These contrasting time-scale orderings suggest that the difference between PPC and PFC cannot be attributed to a fixed property of the neural response such as one

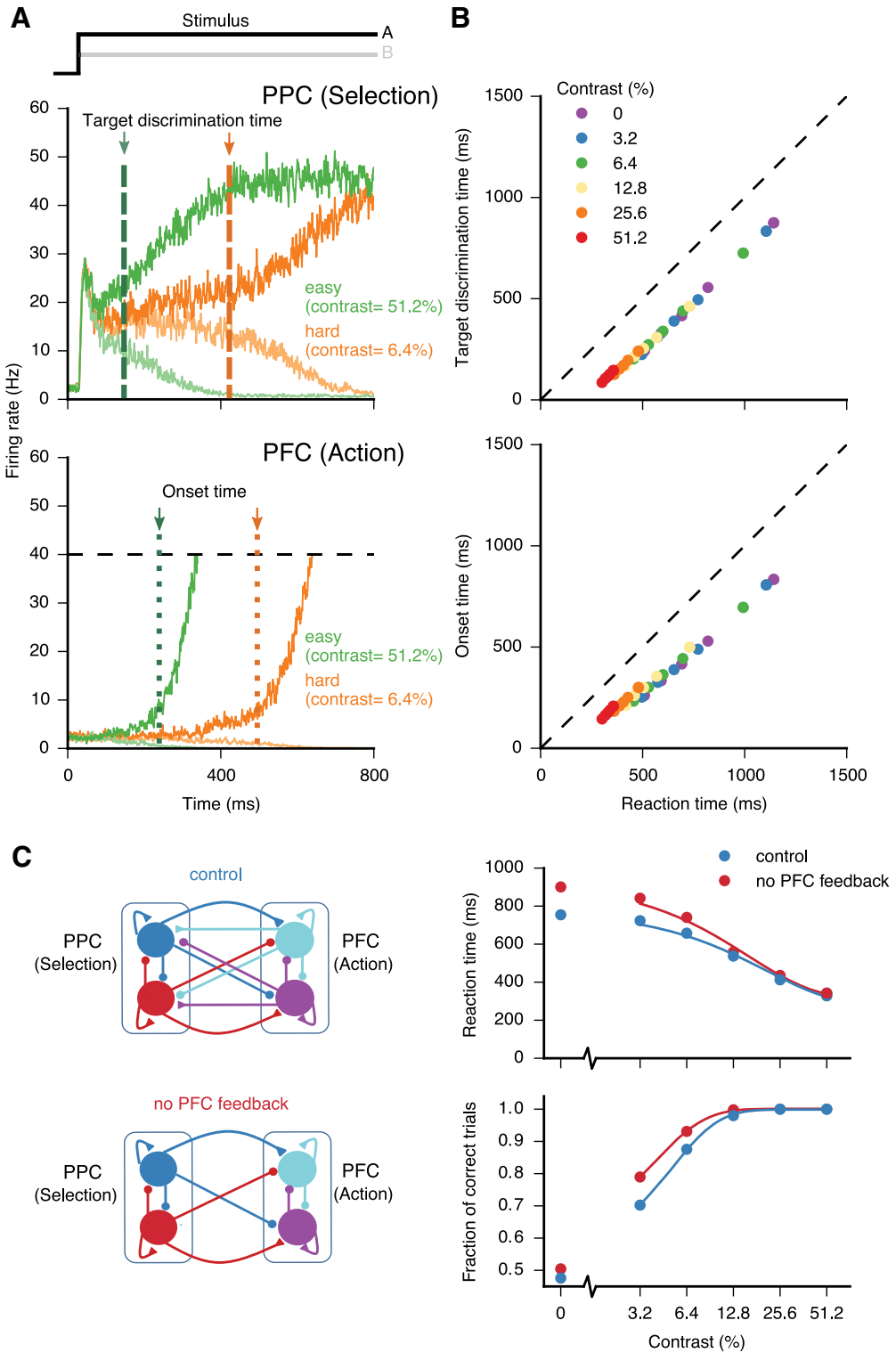

Figure 7. Dynamics of the distributed circuit model during a visuospatial DM task. ***A***, Top, Target and distractor cells in the selection module receive stimulus inputs, integrate perceptual evidence, and discriminate the inputs (marked by target discrimination time). Bottom, Following target discrimination in the selection module, the corresponding population of action cells is activated and begins ramping (marked by onset time). When one of the action populations reaches a threshold of 40 Hz (black dashed lines), an overt response is triggered and a reaction time is registered. Green (orange) traces correspond to easy (hard) trials. Target and distractor cells are shown in thick and thin lines, respectively. Dashed lines mark the target discrimination time in selection cells defined as the time when the difference in firing rate of the two populations has reached 12 Hz. Dotted lines mark the onset time in action cells defined as the time when the firing rate of one of the populations has reached 7 Hz. ***B***, Target discrimination times in the selection module (top) and onset times in the action module (bottom) correlate with reaction times, both across and within contrast conditions. Reaction times were split into quintiles for each contrast level. Only correct trials are shown. ***C***, Psychometric and chronometric curves as a function of contrast for control and no PFC feedback (right).

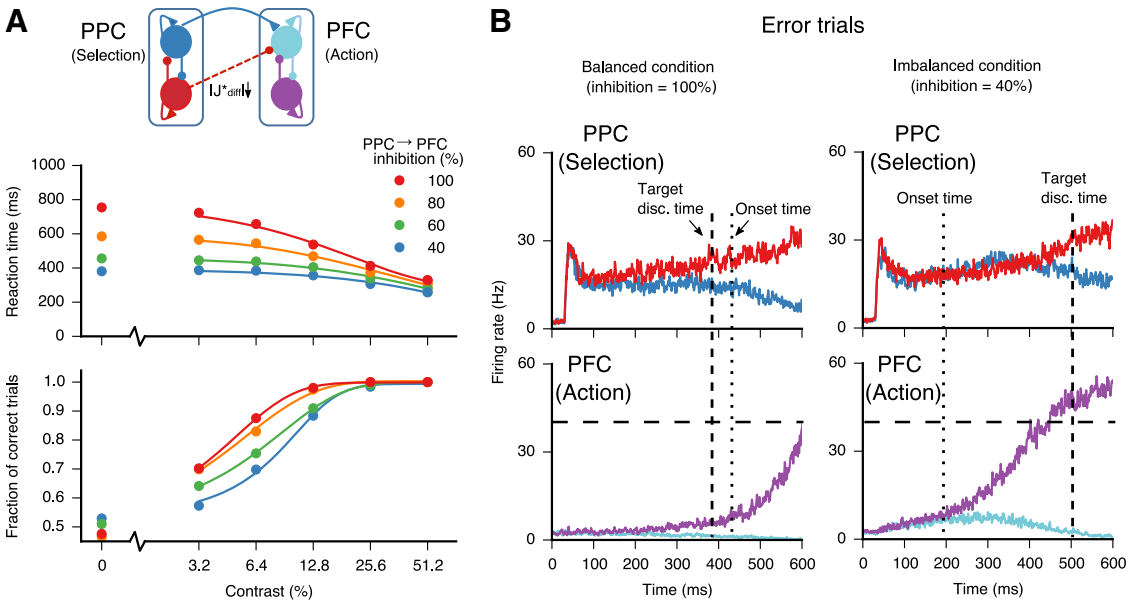

Figure 8. Pathway-specific excitation-inhibition balance disruption and speed–accuracy tradeoff. The degree of balance disruption is quantified by the percentage of the total inhibition required to balance excitation to a cell (100% corresponds to full balance). ***A***, Speed–accuracy tradeoff. Reaction times decrease both as a function of contrast and the degree of balance disruption (top). The fraction of correct trials increases as a function of contrast but decreases as a function of the degree of balance disruption (bottom). ***B***, Error trials with (left) and without (right) pathway-specific balance. For balanced trials (100% inhibition), errors are due to mis-selection from cells in the selection module and subsequent ramping of a population of cells in the action module (top). For imbalanced trials (i.e., with excitatory bias; 40% inhibition), errors can be due to early ramping in cells in the action module before any divergence has begun in the selection module (bottom). Dashed black lines mark the target discrimination time in the selection module, and dotted black lines mark the onset time in the action module.

area being more sluggish than the other. The suppression timescale reflects winner-take-all competition, which is stronger (i.e., faster) in the PFC because of its higher recurrent structure. On the other hand, the intrinsic timescale reflects the timescale of decay of the autocorrelation function of an area given a noisy input, which for linear systems was shown to increase with larger recurrent feedback (Chaudhuri et al., 2015).

Distractor-induced errors in working memory

We now examine the circuit's dynamics in relation to behavioral performance on a WM task with distractors (Fig. 5). An example of a correct trial is shown in Figure 5*A*. The PPC first encodes the target presented at $t = 0$ ms, which is subsequently encoded in persistent activity in both PPC and PFC, and after 1300 ms a distractor is presented to another population in the PPC. During distractor presentation, the distractor-selective population in PPC encodes the distractor strongly in its firing rate, while persistent activity in the target-selective population is transiently and mildly perturbed. The strong distractor response of the PPC is due to two features of the model: stimulus input into PPC is strong and PPC is a weak attractor network, so that the target-selective population, while in the high memory state, does not strongly suppress distractor cells. When the distractor population has higher activity than target cells in PPC, the locus of attention is transiently shifted to the distractor location (Bisley and Goldberg, 2003, 2006). Representation of the distractor in PPC is thought to be functionally desirable, as it allows the PPC to flexibly function as a saliency map (Bisley and Goldberg, 2010).

After the distractor stimulus is withdrawn, an interesting dynamic occurs in PPC: feedback from the PFC switches the PPC back to encoding the target. This switch back to encoding the target in PPC is a feature of distributed processing in the frontoparietal circuit. In an isolated local attractor network, a strong distractor response would switch the state of the network from encoding the target to encoding the distractor even after withdrawal of the distractor stimulus (Fig. 2*A*; Compte et al., 2000; Brunel and Wang, 2001; Murray et al., 2014a). In the distributed circuit model, this switch back to encoding the target in PPC is accomplished by feedback projections from PFC. The target cells in PFC send excitation to the target population in PPC and inhibition to the distractor population in PPC. Because PPC is a weak attractor network, this combination of same-selectivity excitation and cross-selectivity inhibition from PFC can effectively switch PPC back to the target memory state. A single module such as PPC can achieve distractor suppression (Fig. 2*B*), but feedback from PFC makes the system resistant to distractors while simultaneously allowing for the distractor response to transiently surpass the mnemonic target response in PPC (Fig. 5*A*; Suzuki and Gottlieb, 2013). Single-neuron recordings in LIP have shown that PPC networks can switch back to encoding the target in memory after transiently but strongly encoding the distractor (Bisley and Goldberg, 2003; Falkner et al., 2010; Suzuki and Gottlieb, 2013). Furthermore, and consistent with the model, there is experimental evidence that long-range projections between PFC and PPC produce both enhancement and suppression of activity of the recipient cells (Chafee and Goldman-Rakic, 2000).

The PFC exhibits markedly different activity than the PPC in response to distractors. The PFC is activated indirectly, via the response of PPC to the stimulus. As with the PPC, information about the target is maintained in the PFC by persistent activity in the target-selective population. Subsequent distractor presentation causes a transient suppression of delay activity in the PFC. The transient suppression of the target population in PFC is primarily attributable to feedforward different-selectivity inhibition from PPC, rather than local lateral inhibition from the distractor population in PFC. This is consistent with the finding of Suzuki and Gottlieb (2013) that target-selective neurons in the PFC can show suppression by the distractor stimulus without distractor-selective PFC neurons being strongly activated. In contrast to the PPC, the distractor is not represented strongly in the PFC (Fig.

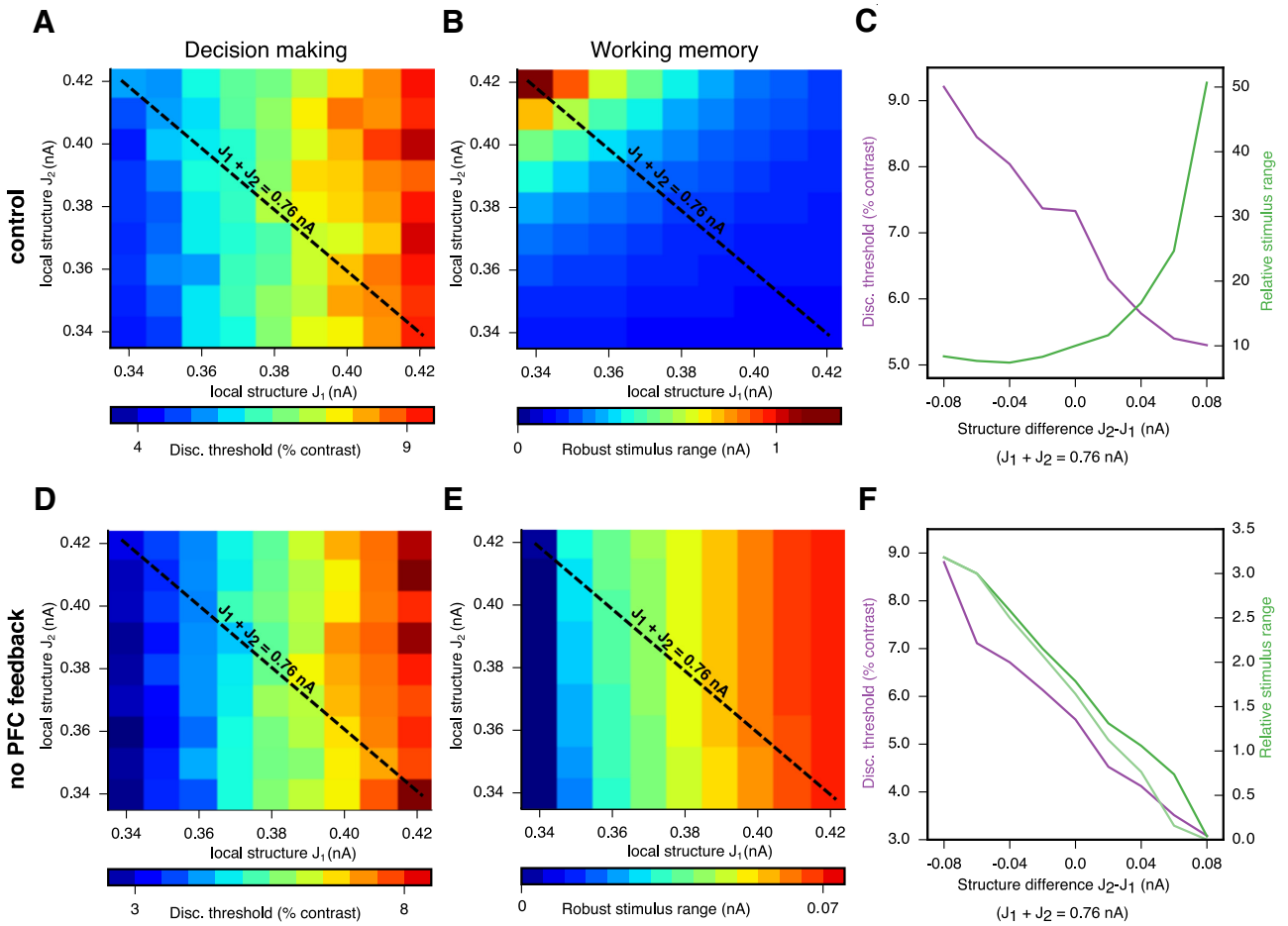

Figure 9. Amelioration of the tradeoff between DM and WM function in a distributed circuit with interareal differences in local recurrent structure. The local structures for Modules 1 and 2, $J_S^{1\to1} = J_1$ and $J_S^{2\to2} = J_2$, respectively, are varied systematically to obtain the performance measures of "discrimination threshold" for DM and "robust stimulus range" for WM. ***A–F***, In ***A–C***, the distributed circuit is endowed with both feedfoward and feedback connections, while in ***D–F*** the feedback connection from PFC is absent (i.e., $J_S^{2\to1} = J_T^{2\to1} = 0$). ***A***, The discrimination threshold increases with increasing J_1, while largely insensitive to J_2. ***B***, The robust stimulus range increases with increasing J_2 and with decreasing J_1. ***C***, The discrimination threshold decreases and the relative stimulus range increases by varying the structure difference J_2–J_1 while keeping the total structure $J_1 + J_2 = 0.76$ nA fixed. Performance in both measures improves with $J_2 > J_1$ (i.e., an increase in local recurrent strength from Module 1 to Module 2). ***D***, Without PFC feedback, the dependence of the discrimination threshold for DM on structure is similar to the control case. ***E***, Without PFC feedback, the robust stimulus range for WM is largely independent of J_2 and the magnitude of the currents that render the system distractible is an order of magnitude below that of the control case. ***F***, Both the discrimination threshold and relative stimulus range (Eq. 17) decrease in the case without PFC feedback as the structure difference J_2–J_1 is varied while keeping the total structure $J_1 + J_2 = 0.76$ nA fixed. Readout of the response from the PFC or PPC module is marked in dark or light green, respectively.

5*A*, bottom). The difference in distractor filtering across modules arises because, first, the external input activates the PFC module only indirectly and, second, the PFC module is modeled as a strong attractor network, with increased local structure relative to the PPC. By combining self-excitation and cross-inhibition, this local structure allows the target population in the memory state to maintain higher activity than the distractor population throughout the delay. This is the mechanism for the robustness of WM against distraction in the circuit.

We investigated the behavior of the PPC–PFC circuit during error trials. An error trial occurs when the PFC module fails to represent the target at the time of decision or readout, following the delay period (at $t > 3000$ ms). During an error trial, the distractor population fires at a higher rate than the target population (Fig. 5*B*). This also corresponds to the locus of attention being shifted to the distractor location (Bisley and Goldberg, 2003, 2006). Figure 5*C* shows that the proportion of errors, and thus distractibility, decreases with greater time separation between target and distractor onset, a behavioral feature that is in line with experimental findings (Suzuki and Gottlieb, 2013). As with the timing dependence of the distractor amplitude in the PPC (Fig. 4*A*), this time course of distractibility is due to the transient dynamics of synaptic gating variables as they evolve toward a steady-state attractor state encoding the target.

Prefrontal inactivation impairs robustness against distractors

In our model, we simulated inactivation or lesion of the PFC by removing the PFC → PPC feedback inputs to the PPC and characterized its effects on neural activity and robustness of WM against distractors (Fig. 5*C*,*D*). As shown in Figure 5*D*, PFC inactivation renders the system vulnerable to distractors. Since the PPC exhibits attractor dynamics as an independent local network, the PPC can still encode the target into the delay period through persistent activity, providing some capacity for WM with only PPC engagement. However, without feedback from PFC, the distractor stimulus switches the network to encoding the distractor, which it continues to encode through the subsequent delay. This is in line with experiments finding that PFC inactivation induces error responses to the distractor location (Suzuki and Gottlieb, 2013). This demonstrates the key role of PFC → PPC feedback projections, in the intact circuit, in switching the PPC network back to encoding the target following

distractor withdrawal (Fig. 5*A*). The resulting behavioral vulnerability to distractors is exemplified by an increase in the error rate with respect to control (Fig. 5*C*). That PFC is not essential for simple WM maintenance, but plays a key role in robustness of WM and filtering of distractors, is in line with conclusions from a range of experimental findings, both in monkeys (Suzuki and Gottlieb, 2013) and in humans (Sakai et al., 2002; Feredoes et al., 2011).

The model findings suggest two modes of operation during WM. When PFC is engaged, the network can operate in a "remember first" regime, storing the initial stimulus and filtering subsequent stimuli. When PFC is not engaged, PPC on its own operates in a "remember last" regime, storing the location of the most recent stimulus, which may also be functionally desirable for a saliency map (Bisley and Goldberg, 2010). As previously mentioned, the amplitude of the transient response to the distractor in PPC is lower compared with that of the target. When the PFC is engaged, the peak distractor response is suppressed (Fig. 5*A*, top) compared with when the PFC is inactivated (Fig. 5*D*).

Evidence accumulation during perceptual decision-making

Having characterized differential roles in WM for the two modules of our distributed circuit, we now examine how it performs in perceptual DM tasks. As shown above in Figure 2, the recurrent structure J_S of a local network shapes how it accumulates perceptual evidence over time to guide a decision, which suggests the PPC and PFC modules will show differential responses to accumulated evidence during perceptual DM. To probe this issue, we adapted a task paradigm used to study perceptual integration in the mouse (Brunton et al., 2013; Hanks et al., 2015). Hanks et al. (2015) found that in the mouse cortex, frontal and parietal areas differ in their representations of accumulated evidence: parietal neurons encoded the accumulator value in a graded fashion, whereas prefrontal neurons encoded the accumulator value more categorically.

In the two-alternative forced choice paradigm of Hanks et al. (2015), the subject receives auditory input consisting of a sequence of Poisson-distributed clicks emitted from the left and from the right side, and the subject is rewarded for reporting which side (left vs right) had the higher-frequency signal (Brunton et al., 2013; Hanks et al., 2015). We model clicks as current pulses whose onset is represented by a set of Poisson-distributed times for each trial, where each trial is characterized by a rate pair sorted according to difficulty. For instance, a 18 clicks/s (left):16 clicks/s (right) trial is "hard" whereas a 30:4 clicks/s trial is "easy." The click-triggered current pulses define the inputs to a theoretical accumulator as well as to the circuit model via the first module, the PPC (see Materials and Methods for details).

Figure 6*A*, top, shows the accumulator value as a function of time, for example trials of varying difficulty. Positive accumulator values correspond to the preferred tuning direction of a neuron, and negative values correspond to the nonpreferred direction. The difficulty of the trials is reflected in the slope of the accumulator versus time plot, where a higher or lower slope in absolute value corresponds to an easy or hard trial, respectively. The trial-averaged firing rates of the PPC and PFC as a function of time and difficulty are shown in Figure 6*A*, middle and bottom, respectively. The difficulty of the trials is also reflected in the instantaneous slopes of the firing rate, but, due to the attractor dynamics and the coupling between the modules, the firing rates in the PPC and PFC are not as linear as a perfect accumulator.

To examine how the theoretical accumulator value is represented in the firing activity of a neural population, we obtained an explicit relationship between accumulator value and the firing rate, following the approach of Hanks et al. (2015) (see Materials and Methods). The relationship between firing rate and accumulator value as a function of time is relatively stable for both the PPC and PFC (Fig. 6*C*, top and bottom), which is similar to empirical findings (Hanks et al., 2015). Importantly, the spacing between the firing rates for different accumulator values is more uniform in the PPC, reflecting a more quasi-linear encoding of the accumulator value compared with the PFC. We obtained the global relationship between firing rate in the PPC and PFC and accumulator by averaging the plots in Figure 6*C* with respect to time and scaling the ranges from 0 to 1 (Fig. 6*B*). We found that, although the accumulator value is encoded in the firing rate for both the PPC and PFC, there is a difference in slope between PPC and PFC firing rates at zero accumulator value that reflects the qualitative difference observed by Hanks et al. (2015): a more categorical accumulator encoding in PFC than in PPC. Indeed, if we map the dynamics of a local circuit to a phase (decision) space where each axis represents one of the two decision variables (Wong and Wang, 2006), increasing the recurrency J_S of the circuit results in a shorter trajectory from the spontaneous state to the threshold. This dynamic difference between PFC and PPC manifests itself as a steeper slope for the firing rate of PFC as a function of both time and the hypothetical accumulator, which scales linearly with time. These results suggest that the differences in recurrent structure between parietal and prefrontal circuits may contribute to the differences in accumulator encoding between these cortical regions.

Perceptual decision-making across functional cell types

In the primate, the PPC and PFC are key cortical areas engaged in perceptual DM tasks such as visual search and target selection (Schall and Thompson, 1999; Thomas and Paré, 2007; Purcell et al., 2010). Within frontoparietal circuits, saccadic target selection involves at least two stages of processing: selection or discrimination of the relevant target among potential distractors; and preparation of an action or response following that selection (Schall and Thompson, 1999; Woodman et al., 2008). Single-neuron recordings have found that signals related to these two stages are represented heterogeneously across different functional cell types, which are distributed across PFC and PPC circuits. Within the frontal eye field (FEF) and other areas, two broad types of neurons—visual cells and movement cells—show distinct dynamics during visual target selection tasks.

The dynamics of visual and movement cells appears to reflect the processes in perceptual selection and action preparation, respectively (Schall, 2015). Visual cells respond strongly at stimulus onset. After the initial visual transient, they distinguish target from distractor through higher firing rates. Visual selection cells have been characterized in LIP (Ipata et al., 2006; Thomas and Paré, 2007), FEF (Thompson et al., 1996; Sato et al., 2001; Sato and Schall, 2003), and the subcortical superior colliculus (McPeek and Keller, 2002; White and Munoz, 2011). In contrast, movement cells are not activated by stimulus onset, but their response is tied to saccade onset. Their activity ramps in motor preparation, with an apparent firing-rate threshold that drives a saccade to their associated movement field (Hanes and Schall, 1996; Woodman et al., 2008). Movement cells are also found in FEF (Hanes and Schall, 1996; Hanes et al., 1998; Woodman et al., 2008) and superior colliculus (McPeek and Keller, 2002), but appear to be much more sparse in LIP (Ferraina et al., 2002).

We sought to test whether our distributed circuit model could capture these differences between visual and movement functional cells in frontoparietal circuits. In the context of saccadic target selection, we associate Module 1 with selection (visual) cells and Module 2 with action (movement) cells, and these functional cell types may be differentially distributed across PPC and PFC. Sensory input enters into selection cells (Module 1), as proposed for a sensorimotor cascade (Purcell et al., 2010; Schall, 2013). Choice and reaction time are set by a firing-rate threshold on action cells (Module 2; Hanes and Schall, 1996; Woodman et al., 2008). The goal of the visual target selection task is to make a saccade toward a target in the presence of distractors. The difficulty of the task is dictated by the contrast, which reflects the salience of the target with respect to that of distractors (Sato et al., 2001; Thomas and Paré, 2007). High or low contrast corresponds to low or high target-distractor similarity, respectively.

In the model, stimulus input activates selection cells (Fig. 7*A*). When the stimulus appears, there is a pronounced visual transient in the firing rates of both target and distractor populations in the selection module. The target-selective population receives more input than the distractor population (the amount depending on the contrast; see Eq. 15), and the activity of the two populations diverges due to competitive dynamics (Wong and Wang, 2006), resulting in a discrimination of the target from the distractor. Due to the feedforward projection from the selection module to the action module, action cells receive stimulus signals indirectly. In a correct trial, the target population reaches the response threshold and a reaction time is recorded. In a high-contrast condition, the reaction time is lower because there is a larger difference in input current to the target and distractor populations of the selection module (Wong and Wang, 2006). Thus, the contrast-dependent differences in input current to the populations in the selection module along with the amplification of those differences due to the recurrent dynamics of both modules eventually leads to a categorical choice in the action cells.

The high-activity population, corresponding to the choice, is consistent in both selection and action modules, but there are two important differences in the dynamics between the modules. First, the competitive dynamics in the action module, compared with the selection module, result in a steeper firing-rate ramping as a function of time (Fig. 7*A*). This is because the action module has a higher recurrent structure J_S, and stronger attractor dynamics, than the selection module. Weaker structure in the selection module enables better integration of perceptual evidence for target selection, as shown in Figure 2. Second, the pronounced transient activation of selection cells, due to the appearance of the target, is not represented in the action module (Fig. 7*A*). This is due to the pathway-specific E/I balance in the feedforward selection → action projection (Eq. 7). Therefore, net inputs to the action module reflect the difference of activity between the populations in the selection module. Thus, an action-cell response above baseline will be observed only when the activities in the selection module have diverged, which is in line with theories and evidence of "discrete flow" between selection- and action-related stages in perceptual DM (Woodman et al., 2008). These results suggest that the functional distinction between selection and action cells arises from a difference in structure in the respective modules and the existence of pathway-specific E/I balance onto action cells.

Single-neuron recordings during visual search have characterized how dynamics of functional cell types relate to reaction times, through the target-distractor discrimination time in visual cells and the onset time in movement cells (see Materials and Methods). For visual cells, the target-distractor discrimination time correlates with reaction time, across and within search difficulty conditions (Sato et al., 2001; McPeek and Keller, 2002; Ipata et al., 2006; Thomas and Paré, 2007; White and Munoz, 2011). For movement cells, the onset time correlates with reaction time (McPeek and Keller, 2002; Woodman et al., 2008; White and Munoz, 2011). We computed analogous measures for selection and action cells in our model. In line with experimental findings, we found that reaction time correlates with discrimination time in selection cells and with onset time in action cells, across different contrast conditions as well as across reaction time variability within a contrast condition (Fig. 7*B*). This implies that although selection and action are distinct stages in perceptual DM, they are consistent and reflect the feedforward nature of the two-module circuit architecture.

We investigated the performance and speed of the distributed DM circuit responses in Figure 7*C*. We found a monotonic dependence on contrast for both measures (Roitman and Shadlen, 2002). Furthermore, we studied the role of PFC (action) feedback on PPC (selection) cells. Interestingly, the absence of feedback leads to more accurate and slower trials—a form of speed–accuracy tradeoff (Heitz and Schall, 2012; Wimmer et al., 2015)—because less total structure improves the slow integration of evidence (compare to Fig. 2*H*).

Pathway-specific E/I balance and speed–accuracy tradeoff

We have shown that action cells are activated only after the selection cells have diverged to select an option, because the projection from selection to action cells exhibits pathway-specific E/I balance. To further characterize the role of pathway-specific E/I balance in the feedforward projection from selection to action cells, we parametrically reduced the strength of feedforward inhibition while holding constant the strength of feedforward excitation (Eq. 16). For each level of inhibition, we computed a chronometric and a psychometric plot: reaction time as a function of contrast and accuracy as a function of contrast (Fig. 8*A*). As feedforward inhibition decreases, reaction times decrease, but so does accuracy. Both of these effects are more pronounced at lower contrast values. Therefore, perturbing pathway-specific E/I balance implements a speed–accuracy tradeoff during target selection.

We then examined how errors arise in the model under control balanced and reduced inhibition conditions. Figure 8*B* shows representative single trials from the two conditions. In the control balanced condition, all errors arise due to mis-selection of the distractor instead of the target in selection cells (Fig. 8*B*, left), which is in line with single-neuron recordings finding mis-selection in visual cells during search tasks (Thompson et al., 2005; Shen and Paré, 2007; Trageser et al., 2008). In contrast, under reduced inhibition a new type of error trial can occur (Fig. 8*B*, right). Action cells prematurely select a response before target selection has completed within the selection cells, causing errors because the action module makes a decision for target or distractor in a quasi-random manner. Under reduced inhibition, when the target and distractor selection cells are activated, but not yet diverged, the action module receives a nonspecific net-excitatory input. This net-excitatory input can induce quasi-random winner-take-all DM (Wong and Wang, 2006), giving rise to an imbalance-dependent type of error. Our findings suggest that in a healthy physiological state, the projection from Module 1 cells to Module 2 cells should be in a state near E/I balance, because this configuration produces errors consistent with electrophysiolog-

ical recordings (Thompson et al., 2005; Shen and Paré, 2007; Trageser et al., 2008).

Tradeoff amelioration in a distributed circuit

Having described how a two-module distributed circuit can capture key features of multiple electrophysiological findings related to WM and DM, we return to the issue of the computational advantages of such a distributed circuit compared with a single-module circuit. In particular, we showed in Figure 2 that there is a tradeoff in WM and DM performance in that by varying the recurrent structure J_S the one-module network exhibits increased performance for one measure (e.g., robustness) and a decreased performance for the other (e.g., discrimination). In Figure 9*A–C*, we show that a distributed circuit composed of two modules with local recurrent structures J_1 and J_2 can ameliorate the tradeoff in the single-module circuit. As the difference in structures J_2–J_1 is increased while preserving the total structure in the circuit, there are increases in both performance measures, robustness in WM, and discrimination in DM (Fig. 9*C*; see also Materials and Methods). This amelioration occurs when there is a hierarchical increase in local recurrent structure from Module 1 (PPC, selection) to Module 2 (PFC, action). We suggest that this feature constitutes a desirable design principle for multi-regional networks.

We also evaluated the importance of the feedback projection from the second module to the first in Figure 9*D–F*. Without top-down feedback, the tradeoff between WM and DM function is not ameliorated: as the difference in structures increases, DM discrimination performance increases (Fig. 7*C*) but robustness in WM decreases. Furthermore, the stimulus range for which the system is robust to distractors is substantially smaller than the control case, which is consistent with Figure 5*D*. These results suggest that distractor filtering in WM, and the tradeoff amelioration in general, can be strongly regulated by top-down feedback.

Discussion

In this study, we propose a parsimonious circuit model for distributed computations subserving WM and DM, two core cognitive functions that recruit overlapping frontoparietal circuits. We highlight a tradeoff that exists when optimizing the recurrent structure for robustness against distractors in WM versus slow integration of evidence in DM; the model developed here ameliorates this tradeoff by extension to two modules with a hierarchical increase in recurrent strengths. We found that across both WM and DM paradigms, the circuit model captures a wide range of salient, empirically observed features of neural activity in frontoparietal circuits. We summarize the findings of the model with respect to the circuit architecture and its relationship to WM and DM computations.

First, Module 1 (PPC or selection cells), which receives sensory input (Felleman and Van Essen, 1991; Buschman and Miller, 2007; Ibos et al., 2013; Siegel et al., 2015), is a weak attractor network. This property is beneficial so that PPC can transiently encode distractors and function as a saliency map, and so that its memory state can be effectively controlled by weak feedback projections from PFC (e.g., to switch the state back to encoding the target following distractor presentation). In the context of perceptual DM, Module 1 (selection cells) should be a weak attractor network to support integration of perceptual evidence with a long timescale to improve accuracy.

Second, Module 2 (PFC or action cells) is a strong attractor network. In the context of WM, this property is functionally beneficial because PFC can thereby provide robustness against distractors, filtering out the effects of strong distractor responses in PPC. In the context of DM, it is functionally beneficial for action cells to be a strong attractor network because this enables them to ramp quickly to threshold to drive choice following the upstream target selection. This difference in recurrent structure also predicts a difference in the representation of accumulated value, which is more graded in Module 1 and more categorical in Module 2.

Third, the modules are interconnected via reciprocal projections that are structured: net-excitatory between same-selectivity populations and net-inhibitory between different-selectivity populations. The feedforward Module 1 → Module 2 projection should be structured to propagate signals for both WM and DM. The feedback Module 2 → Module 1 projection is especially important in the context of WM. This projection being structured allows PFC to switch PPC back to encoding the target following distractor presentation.

Differential roles for PPC and PFC in working memory

Although PPC and PFC are both involved in active WM maintenance, converging evidence from a range of methodologies suggests differential roles, with PPC associated with attentional saliency and selection (Wardak et al., 2012, 2002) and PFC associated with robustness of WM and filtering of distractors (Sakai et al., 2002; Feredoes et al., 2011; Suzuki and Gottlieb, 2013). Our distributed WM circuit model captures multiple key features from these studies and suggests that the different roles of PPC and PFC may be due in part to distinct dynamic behaviors arising from their recurrent microcircuitry (Chaudhuri et al., 2015), as suggested by electrophysiological recordings (Katsuki et al., 2014; Murray et al., 2014b).

The model also proposes a key role for feedback from PFC to PPC during distractor processing, namely to strengthen target representations via feedback excitation and filter distractor representations via feedback inhibition (Figs. 4, 5). In line with this proposal, Feredoes et al. (2011) performed combined stimulation-imaging experiments in humans and found a key role for feedback from PFC to posterior areas during distractor filtering to enhance target representations (for PFC-mediated enhancement of WM capacity in PPC, see also Edin et al. (2009)). Along these lines, we suggest that the effective strength of the lateral inhibition in PPC is not a purely local property and can be flexibly controlled by top-down (i.e., feedback) prefrontal engagement (Falkner et al., 2010, 2013). Recording in LIP during visuospatial WM, Falkner et al. (2010, 2013) found that surround suppression of distractors, and target representation, are strengthened by top-down cognitive modulation (e.g., by motivation). Our model predicts that PFC inactivation should disrupt this modulation of surround suppression in PPC. Furthermore, if PFC is inactivated or the PFC–PPC feedback projection is very weak, the system can operate in a 'remember-last' regime (Fig. 5*D*), whereas PFC engagement can switch the system into a 'remember-first' regime (Fig. 5*A*).

In modeling responses to intervening distractors (Figs. 4, 5) we considered input to the PPC module only, which resulted in highly attenuated responses in the PFC—to a greater extent than the findings by Suzuki and Gottlieb (2013). Thus, in our model we effectively investigated a subcircuit in the PFC that strongly filters distractors and whose response is much smaller than the more sensory-driven PPC circuit. The effects of simultaneous external inputs to both modules and their interplay with the hi-

erarchical recurrent interactions remains a topic for future computational studies.

Evidence accumulation in the frontoparietal network

Gradual accumulation of perceptual evidence is reflected in DM-related neuronal activity in frontoparietal circuits (Gold and Shadlen, 2007; Schall, 2013; Brody and Hanks, 2016). Drift-diffusion models (Usher and McClelland, 2001; Bogacz et al., 2006; Purcell et al., 2010) as well as recurrent circuit models (Wang, 2002; Wong and Wang, 2006) can account for the accumulation process and subsequent ramping behavior in the neural dynamics, including more discrete "jumping" modes (Miller and Katz, 2010; Wang, 2012; Latimer et al., 2015; Lo et al., 2015). Our model predicts that, due to the interareal differences in local recurrent structure, representations of accumulated evidence are more graded in PPC and more categorical in PFC, as supported by single-neuron recordings (Hanks et al., 2015). Recent inactivation studies found that even though PPC shows DM-related signals, it does not play a causal role in perceptual DM (Erlich et al., 2015; Katz et al., 2016). The random-dot motion and auditory accumulation tasks where such deficits were observed are different from the visuospatial attention task we modeled in Figures 7 and 8. Normal performance in such visual or visuospatial attention tasks has been shown to require LIP (Wardak et al., 2002, 2004; Goard et al., 2016; Licata et al., 2017). In rats, inactivation of PPC disrupts the accumulation of visual but not auditory evidence (Raposo et al., 2014). Furthermore, sensory signals could reach areas such as the FEF (primate) or FOF (rodent) through other pathways (Buschman and Miller, 2007; Katsuki and Constantinidis, 2012a; Hanks et al., 2015) that are in principle capable of evidence integration independent from LIP (Fig. 2*G*). Furthermore, it is unclear whether these results can be accounted for via compensatory mechanisms and/or distributed processing within a broader region in PPC that includes LIP. The issue of LIP involvement in decision-making tasks should be addressed by future experiments that carefully consider behavioral performance, task design, species-specific differences, and the presence of evidence accumulation in the neural activity of multiple simultaneously recorded areas (Pisupati et al., 2016; Hanks and Summerfield, 2017).

Dynamics and localization of visual and movement cells

In the context of perceptual DM, the selection (Module 1) and action (Module 2) cells in our model capture key dynamic features of visual and movement cells, respectively, which have been characterized in single-neuron recordings from LIP, FEF, and superior colliculus (SC) (Thompson et al., 1996; Hanes et al., 1998; Sato et al., 2001; McPeek and Keller, 2002; Ipata et al., 2006; Thomas and Paré, 2007; Woodman et al., 2008; Schall, 2015). Visual and movement cells in cortex may have distinct biophysical properties (Cohen et al., 2009b), laminar distributions (Pouget et al., 2009), and long-range projections (Gregoriou et al., 2012). Our results suggest that both modules could represent neural populations that are distributed across multiple areas (e.g., visual cells in LIP/FEF/SC, movement cells in FEF/SC).

An important feature of our distributed circuit model is the pathway-specific E/I balance in the feedforward projection from selection to action cells. This E/I balance regulates a cascade of activations across functional cell types and can be characterized as discrete flow, since response preparation in action cells begins only after target selection is completed in selection cells (Woodman et al., 2008). Pathway-specific balance thereby provides a mechanism for discrete flow without requiring direct gating of neuronal responses (Wang et al., 2004; Purcell et al., 2010; Yang et al., 2016). The sensitivity of action cells to differences in upstream selection cells is related to a geometrical characterization of neural dynamics, whereby some directions in neural state space elicit responses, while those in the null space do not (Vogels and Abbott, 2009; Kaufman et al., 2014; Li et al., 2016).

We disrupted pathway-specific E/I balance by systematically decreasing the strength of feedforward inhibition onto the action cells. A strongly imbalanced circuit produced a distinct type of error: Action cells started ramping before the populations of selection cells had diverged in activity (i.e., before selection was accomplished). The imbalanced condition may capture dynamics recorded from visual and movement cells in FEF under speed-demanding compelled response paradigms (Stanford et al., 2010). The modulation of feedforward inhibition in our model resulted in a smooth tradeoff between speed and accuracy, a plausible mechanism among others (Bogacz et al., 2010; Stanford et al., 2010; Heitz and Schall, 2012; Standage et al., 2014; Hanks et al., 2015).

Limitations and future directions

Future studies can build upon and extend the present model to address a number of important questions. One direction is to extend the two-population discrete network studied here to a quasi-continuous network in which neurons exhibit smoothly varying tuning of a parametric stimulus variable (Compte et al., 2000; Furman and Wang, 2008) to explore effects that depend on the similarity and distance between distractors and targets held in WM (Suzuki and Gottlieb, 2013; Murray et al., 2014a). An important question for future research is how to include sufficient heterogeneity in the model so as to account for the dynamic responses of individual cells but stable representations at the population level during WM tasks (Murray et al., 2017; Spaak et al., 2017). Generalization to more than two populations would enable modeling of set-size effects in visual search tasks, whereby the number of stimuli affects behavior and visual cell and movement cell activity (Balan et al., 2008; Woodman et al., 2008; Cohen et al., 2009a). Extension to a spiking circuit model would enable modeling the neural signatures of synchronization between PPC and PFC during cognitive processing (Pesaran et al., 2008; Ardid et al., 2010; Salazar et al., 2012; Dotson et al., 2014). Another extension of the model is the inclusion of a remapping module for inhibitory control (Lo and Wang, 2016) that transforms the visual representation of a target to a saccade goal away from the target (i.e., an antisaccade; Gottlieb and Goldberg, 1999; Munoz and Everling, 2004).

Our study suggests important design principles for constructing multiregional neural circuit models of distributed cognitive function, such as the interplay between long-range and local connectivity in recurrent dynamics and computation, the roles of specialized microcircuit properties across the cortical hierarchy, and the implications of balanced excitation and inhibition in long-range interactions. The parsimonious model studied here may therefore instantiate features of a canonical cognitive circuit useful for studying distributed computation in the brain.

References

Abbott LF, Chance FS (2005) Drivers and modulators from push-pull and balanced synaptic input. Prog Brain Res 149:147–155. CrossRef Medline

Amit DJ, Brunel N (1997) Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. Cereb Cortex 7:237–252. CrossRef Medline

Ardid S, Wang XJ, Gomez-Cabrero D, Compte A (2010) Reconciling coherent oscillation with modulation of irregular spiking activity in selective

attention: gamma-range synchronization between sensory and executive cortical areas. J Neurosci 30:2856–2870. CrossRef Medline
Balan PF, Oristaglio J, Schneider DM, Gottlieb J (2008) Neuronal correlates of the set-size effect in monkey lateral intraparietal area. PLoS Biol 6:e158. CrossRef Medline
Bisley JW, Goldberg ME (2003) Neuronal activity in the lateral intraparietal area and spatial attention. Science 299:81–86. CrossRef Medline
Bisley JW, Goldberg ME (2006) Neural correlates of attention and distractibility in the lateral intraparietal area. J Neurophysiol 95:1696–1717. CrossRef Medline
Bisley JW, Goldberg ME (2010) Attention, intention, and priority in the parietal lobe. Annu Rev Neurosci 33:1–21. CrossRef Medline
Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD (2006) The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychol Rev 113:700–765. CrossRef Medline
Bogacz R, Wagenmakers EJ, Forstmann BU, Nieuwenhuis S (2010) The neural basis of the speed-accuracy tradeoff. Trends Neurosci 33:10–16. CrossRef Medline
Brody CD, Hanks TD (2016) Neural underpinnings of the evidence accumulator. Curr Opin Neurobiol 37:149–157. CrossRef Medline
Brunel N, Wang XJ (2001) Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. J Comput Neurosci 11:63–85. CrossRef Medline
Brunton BW, Botvinick MM, Brody CD (2013) Rats and humans can optimally accumulate evidence for decision-making. Science 340:95–98. CrossRef Medline
Buschman TJ, Miller EK (2007) Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. Science 315:1860–1862. CrossRef Medline
Chafee MV, Goldman-Rakic PS (1998) Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task. J Neurophysiol 79:2919–2940. Medline
Chafee MV, Goldman-Rakic PS (2000) Inactivation of parietal and prefrontal cortex reveals interdependence of neural activity during memory-guided saccades. J Neurophysiol 83:1550–1566. Medline
Chaudhuri R, Knoblauch K, Gariel MA, Kennedy H, Wang XJ (2015) A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. Neuron 88:419–431. CrossRef Medline
Christophel TB, Klink PC, Spitzer B, Roelfsema PR, Haynes JD (2017) The distributed nature of working memory. Trends Cogn Sci 21:111–124. CrossRef Medline
Clewley R (2012) Hybrid models and biological model reduction with PyDSTool. PLoS Comput Biol 8:e1002628. CrossRef Medline
Cohen JY, Heitz RP, Woodman GF, Schall JD (2009a) Neural basis of the set-size effect in frontal eye field: timing of attention during visual search. J Neurophysiol 101:1699–1704. CrossRef Medline
Cohen JY, Pouget P, Heitz RP, Woodman GF, Schall JD (2009b) Biophysical support for functionally distinct cell types in the frontal eye field. J Neurophysiol 101:912–916. CrossRef Medline
Compte A, Brunel N, Goldman-Rakic PS, Wang XJ (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. Cereb Cortex 10:910–923. CrossRef Medline
Constantinidis C, Procyk E (2004) The primate working memory networks. Cogn Affect Behav Neurosci 4:444–465. CrossRef Medline
Constantinidis C, Wang XJ (2004) A neural circuit basis for spatial working memory. Neuroscientist 10:553–565. CrossRef Medline
Domenech P, Redoute J, Koechlin E, Dreher JC (2017) The neuro-computational architecture of value-based selection in the human brain. Cereb Cortex. Advance online publication. Retrieved November 6, 2017. CrossRef Medline
Dotson NM, Salazar RF, Gray CM (2014) Frontoparietal correlation dynamics reveal interplay between integration and segregation during visual working memory. J Neurosci 34:13600–13613. CrossRef Medline
Duncan J (2010) The multiple-demand (md) system of the primate brain: mental programs for intelligent behaviour. Trends Cogn Sci 14:172–179. CrossRef Medline
Edin F, Klingberg T, Johansson P, McNab F, Tegnér J, Compte A (2009) Mechanism for top-down control of working memory capacity. Proc Natl Acad Sci U S A 106:6802–6807. CrossRef Medline
Erlich JC, Brunton BW, Duan CA, Hanks TD, Brody CD (2015) Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. Elife 4:e05457. CrossRef Medline
Everling S, Tinsley CJ, Gaffan D, Duncan J (2002) Filtering of neural signals by focused attention in the monkey prefrontal cortex. Nat Neurosci 5:671–676. CrossRef Medline
Falkner AL, Krishna BS, Goldberg ME (2010) Surround suppression sharpens the priority map in the lateral intraparietal area. J Neurosci 30:12787–12797. CrossRef Medline
Falkner AL, Goldberg ME, Krishna BS (2013) Spatial representation and cognitive modulation of response variability in the lateral intraparietal area priority map. J Neurosci 33:16117–16130. CrossRef Medline
Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex 1:1–47. Medline
Feredoes E, Heinen K, Weiskopf N, Ruff C, Driver J (2011) Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual working memory. Proc Natl Acad Sci U S A 108:17510–17515. CrossRef Medline
Ferraina S, Paré M, Wurtz RH (2002) Comparison of cortico-cortical and cortico-collicular signals for the generation of saccadic eye movements. J Neurophysiol 87:845–858. CrossRef Medline
Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J Neurophysiol 61:331–349. Medline
Furman M, Wang XJ (2008) Similarity effect and optimal control of multiple-choice decision making. Neuron 60:1153–1168. CrossRef Medline
Goard MJ, Pho GN, Woodson J, Sur M (2016) Distinct roles of visual, parietal, and frontal motor cortices in memory-guided sensorimotor decisions. Elife 5:e13764. CrossRef Medline
Gold JI, Shadlen MN (2007) The neural basis of decision making. Annu Rev Neurosci 30:535–574. CrossRef Medline
Goldman-Rakic PS (1995) Cellular basis of working memory. Neuron 14:477–485. CrossRef Medline
Gottlieb J, Goldberg ME (1999) Activity of neurons in the lateral intraparietal area of the monkey during an antisaccade task. Nat Neurosci 2:906–912. CrossRef Medline
Gregoriou GG, Gotts SJ, Desimone R (2012) Cell-type-specific synchronization of neural activity in fef with v4 during attention. Neuron 73:581–594. CrossRef Medline
Hanes DP, Schall JD (1996) Neural control of voluntary movement initiation. Science 274:427–430. CrossRef Medline
Hanes DP, Patterson WF 2nd, Schall JD (1998) Role of frontal eye fields in countermanding saccades: visual, movement, and fixation activity. J Neurophysiol 79:817–834. Medline
Hanks TD, Summerfield C (2017) Perceptual decision making in rodents, monkeys, and humans. Neuron 93:15–31. CrossRef Medline
Hanks TD, Kopec CD, Brunton BW, Duan CA, Erlich JC, Brody CD (2015) Distinct relationships of parietal and prefrontal cortices to evidence accumulation. Nature 520:220–223. CrossRef Medline
Heitz RP, Schall JD (2012) Neural mechanisms of speed-accuracy tradeoff. Neuron 76:616–628. CrossRef Medline
Ibos G, Duhamel JR, Ben Hamed S (2013) A functional hierarchy within the parietofrontal network in stimulus selection and attention control. J Neurosci 33:8359–8369. CrossRef Medline
Ipata AE, Gee AL, Goldberg ME, Bisley JW (2006) Activity in the lateral intraparietal area predicts the goal and latency of saccades in a free-viewing visual search task. J Neurosci 26:3656–3661. CrossRef Medline
Katsuki F, Constantinidis C (2012a) Early involvement of prefrontal cortex in visual bottom-up attention. Nat Neurosci 15:1160–1166. CrossRef Medline
Katsuki F, Constantinidis C (2012b) Unique and shared roles of the posterior parietal and dorsolateral prefrontal cortex in cognitive functions. Front Integr Neurosci 6:17. CrossRef Medline
Katsuki F, Qi XL, Meyer T, Kostelic PM, Salinas E, Constantinidis C (2014) Differences in intrinsic functional organization between dorsolateral prefrontal and posterior parietal cortex. Cereb Cortex 24:2334–2349. CrossRef Medline
Katz LN, Yates JL, Pillow JW, Huk AC (2016) Dissociated functional significance of decision-related activity in the primate dorsal stream. Nature 535:285–288. CrossRef Medline
Kaufman MT, Churchland MM, Ryu SI, Shenoy KV (2014) Cortical activity in the null space: permitting preparation without movement. Nat Neurosci 17:440–448. CrossRef Medline

Latimer KW, Yates JL, Meister ML, Huk AC, Pillow JW (2015) Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. Science 349:184–187. CrossRef Medline

Lawrence BM, White RL 3rd, Snyder LH (2005) Delay-period activity in visual, visuomovement, and movement neurons in the frontal eye field. J Neurophysiol 94:1498–1508. CrossRef Medline

Li N, Daie K, Svoboda K, Druckmann S (2016) Robust neuronal dynamics in premotor cortex during motor planning. Nature 532:459–464. CrossRef Medline

Licata AM, Kaufman MT, Raposo D, Ryan MB, Sheppard JP, Churchland AK (2017) Posterior parietal cortex guides visual decisions in rats. J Neurosci 37:4954–4966. CrossRef Medline

Lo CC, Wang XJ (2016) Conflict resolution as near-threshold decision-making: a spiking neural circuit model with two-stage competition for antisaccadic task. PLoS Comput Biol 12:e1005081. CrossRef Medline

Lo CC, Wang CT, Wang XJ (2015) Speed-accuracy tradeoff by a control signal with balanced excitation and inhibition. J Neurophysiol 114:650–661. CrossRef Medline

Machens CK, Romo R, Brody CD (2005) Flexible control of mutual inhibition: a neural model of two-interval discrimination. Science 307:1121–1124. CrossRef Medline

McPeek RM, Keller EL (2002) Saccade target selection in the superior colliculus during a visual search task. J Neurophysiol 88:2019–2034. Medline

Meister ML, Hennig JA, Huk AC (2013) Signal multiplexing and single-neuron computations in lateral intraparietal area during decision-making. J Neurosci 33:2254–2267. CrossRef Medline

Miller EK, Li L, Desimone R (1993) Activity of neurons in anterior inferior temporal cortex during a short-term memory task. J Neurosci 13:1460–1478. Medline

Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. J Neurosci 16:5154–5167. Medline

Miller P, Katz DB (2010) Stochastic transitions between neural states in taste processing and decision-making. J Neurosci 30:2559–2570. CrossRef Medline

Mitchell DJ, Bell AH, Buckley MJ, Mitchell AS, Sallet J, Duncan J (2016) A putative multiple-demand system in the macaque brain. J Neurosci 36:8574–8585. CrossRef Medline

Munoz DP, Everling S (2004) Look away: the anti-saccade task and the voluntary control of eye movement. Nat Rev Neurosci 5:218–228. CrossRef Medline

Murray JD, Anticevic A, Gancsos M, Ichinose M, Corlett PR, Krystal JH, Wang XJ (2014a) Linking microcircuit dysfunction to cognitive impairment: effects of disinhibition associated with schizophrenia in a cortical working memory model. Cereb Cortex 24:859–872. CrossRef Medline

Murray JD, Bernacchia A, Freedman DJ, Romo R, Wallis JD, Cai X, Padoa-Schioppa C, Pasternak T, Seo H, Lee D, Wang XJ (2014b) A hierarchy of intrinsic timescales across primate cortex. Nat Neurosci 17:1661–1663. CrossRef Medline

Murray JD, Bernacchia A, Roy NA, Constantinidis C, Romo R, Wang XJ (2017) Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. Proc Natl Acad Sci U S A 114:394–399. CrossRef Medline

Pesaran B, Nelson MJ, Andersen RA (2008) Free choice activates a decision circuit between frontal and parietal cortex. Nature 453:406–409. CrossRef Medline

Pisupati S, Chartarifsky L, Churchland AK (2016) Decision activity in parietal cortex—leader or follower? Trends Cogn Sci 20:788–789. CrossRef Medline

Pouget P, Stepniewska I, Crowder EA, Leslie MW, Emeric EE, Nelson MJ, Schall JD (2009) Visual and motor connectivity and the distribution of calcium-binding proteins in macaque frontal eye field: implications for saccade target selection. Front Neuroanat 3:2. CrossRef Medline

Powell KD, Goldberg ME (2000) Response of neurons in the lateral intraparietal area to a distractor flashed during the delay period of a memory-guided saccade. J Neurophysiol 84:301–310. Medline

Purcell BA, Heitz RP, Cohen JY, Schall JD, Logan GD, Palmeri TJ (2010) Neurally constrained modeling of perceptual decision making. Psychol Rev 117:1113–1143. CrossRef Medline

Qi XL, Katsuki F, Meyer T, Rawley JB, Zhou X, Douglas KL, Constantinidis C (2010) Comparison of neural activity related to working memory in primate dorsolateral prefrontal and posterior parietal cortex. Front Syst Neurosci 4:12. CrossRef Medline

Raposo D, Kaufman MT, Churchland AK (2014) A category-free neural population supports evolving demands during decision-making. Nat Neurosci 17:1784–1792. CrossRef Medline

Ratcliff R (1978) A theory of memory retrieval. Psychol Rev 85:59–108. CrossRef

Roitman JD, Shadlen MN (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. J Neurosci 22:9475–9489. Medline

Sakai K, Rowe JB, Passingham RE (2002) Active maintenance in prefrontal area 46 creates distractor-resistant memory. Nat Neurosci 5:479–484. Medline

Salazar RF, Dotson NM, Bressler SL, Gray CM (2012) Content-specific fronto-parietal synchronization during visual working memory. Science 338:1097–1100. CrossRef Medline

Sato TR, Schall JD (2003) Effects of stimulus-response compatibility on neural selection in frontal eye field. Neuron 38:637–648. CrossRef Medline

Sato T, Murthy A, Thompson KG, Schall JD (2001) Search efficiency but not response interference affects visual selection in frontal eye field. Neuron 30:583–591. CrossRef Medline

Schall JD (2001) Neural basis of deciding, choosing and acting. Nat Rev Neurosci 2:33–42. CrossRef Medline

Schall JD (2013) Macrocircuits: decision networks. Curr Opin Neurobiol 23:269–274. CrossRef Medline

Schall JD (2015) Visuomotor functions in the frontal lobe. Annu Rev Vis Sci 1:469–498. CrossRef Medline

Schall JD, Thompson KG (1999) Neural selection and control of visually guided eye movements. Annu Rev Neurosci 22:241–259. CrossRef Medline

Shen K, Paré M (2007) Neuronal activity in superior colliculus signals both stimulus identity and saccade goals during visual conjunction search. J Vis 7(5):15 1–13. CrossRef Medline

Siegel M, Buschman TJ, Miller EK (2015) Cortical information flow during flexible sensorimotor decisions. Science 348:1352–1355. CrossRef Medline

Spaak E, Watanabe K, Funahashi S, Stokes MG (2017) Stable and dynamic coding for working memory in primate prefrontal cortex. J Neurosci 37:6503–6516. CrossRef Medline

Standage D, Paré M (2011) Persistent storage capability impairs decision making in a biophysical network model. Neural Netw 24:1062–1073. CrossRef Medline

Standage D, Blohm G, Dorris MC (2014) On the neural implementation of the speed-accuracy trade-off. Front Neurosci 8:236. CrossRef Medline

Stanford TR, Shankar S, Massoglia DP, Costello MG, Salinas E (2010) Perceptual decision making in less than 30 milliseconds. Nat Neurosci 13:379–385. CrossRef Medline

Suzuki M, Gottlieb J (2013) Distinct neural mechanisms of distractor suppression in the frontal and parietal lobe. Nat Neurosci 16:98–104. CrossRef Medline

Thomas NW, Paré M (2007) Temporal processing of saccade targets in parietal cortex area LIP during visual search. J Neurophysiol 97:942–947. CrossRef Medline

Thompson KG, Hanes DP, Bichot NP, Schall JD (1996) Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. J Neurophysiol 76:4040–4055. Medline

Thompson KG, Bichot NP, Sato TR (2005) Frontal eye field activity before visual search errors reveals the integration of bottom-up and top-down salience. J Neurophysiol 93:337–351. CrossRef Medline

Trageser JC, Monosov IE, Zhou Y, Thompson KG (2008) A perceptual representation in the frontal eye field during covert visual search that is more reliable than the behavioral report. Eur J Neurosci 28:2542–2549. CrossRef Medline

Usher M, McClelland JL (2001) The time course of perceptual choice: the leaky, competing accumulator model. Psychol Rev 108:550–592. CrossRef Medline

Vogels TP, Abbott LF (2009) Gating multiple signals through detailed balance of excitation and inhibition in spiking networks. Nat Neurosci 12:483–491. CrossRef Medline

Wang XJ (1999) Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. J Neurosci 19:9587–9603. Medline

Wang XJ (2001) Synaptic reverberation underlying mnemonic persistent activity. Trends Neurosci 24:455–463. CrossRef Medline

Wang XJ (2002) Probabilistic decision making by slow reverberation in cortical circuits. Neuron 36:955–968. CrossRef Medline

Wang XJ (2008) Decision making in recurrent neuronal circuits. Neuron 60:215–234. CrossRef Medline

Wang XJ (2012) Neural dynamics and circuit mechanisms of decision-making. Curr Opin Neurobiol 22:1039–1046. CrossRef Medline

Wang XJ (2013) The prefrontal cortex as a quintessential "cognitive-type" neural circuit: working memory and decision making. In: Principles of frontal lobe function, Ed 2 (Stuss DT, Knight RT, eds), pp 226–248. New York: Oxford UP.

Wang XJ, Tegnér J, Constantinidis C, Goldman-Rakic PS (2004) Division of labor among distinct subtypes of inhibitory neurons in a cortical microcircuit of working memory. Proc Natl Acad Sci U S A 101:1368–1373. CrossRef Medline

Wardak C, Olivier E, Duhamel JR (2002) Saccadic target selection deficits after lateral intraparietal area inactivation in monkeys. J Neurosci 22: 9877–9884. Medline

Wardak C, Olivier E, Duhamel JR (2004) A deficit in covert attention after parietal cortex inactivation in the monkey. Neuron 42:501–508. CrossRef Medline

Wardak C, Ben Hamed S, Olivier E, Duhamel JR (2012) Differential effects of parietal and frontal inactivations on reaction times distributions in a visual search task. Front Integr Neurosci 6:39. CrossRef Medline

White BJ, Munoz DP (2011) Separate visual signals for saccade initiation during target selection in the primate superior colliculus. J Neurosci 31: 1570–1578. CrossRef Medline

Wimmer K, Compte A, Roxin A, Peixoto D, Renart A, de la Rocha J (2015) Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area MT. Nat Commun 6:6177. CrossRef Medline

Wong KF, Wang XJ (2006) A recurrent network mechanism of time integration in perceptual decisions. J Neurosci 26:1314–1328. CrossRef Medline

Wong KF, Huk AC, Shadlen MN, Wang XJ (2007) Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision making. Front Comput Neurosci 1:6. CrossRef Medline

Woodman GF, Kang MS, Thompson K, Schall JD (2008) The effect of visual search efficiency on response preparation: neurophysiological evidence for discrete flow. Psychol Sci 19:128–136. CrossRef Medline

Yang GR, Murray JD, Wang XJ (2016) A dendritic disinhibitory circuit mechanism for pathway-specific gating. Nat Commun 7:12815. CrossRef Medline

Zhang W, Falkner AL, Krishna BS, Goldberg ME, Miller KD (2017) Coupling between one-dimensional networks reconciles conflicting dynamics in LIP and reveals its recurrent circuitry. Neuron 93:221–234. CrossRef Medline