Vision Research 48 (2008) 1345–1373

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Temporal dynamics of decision-making during motion perception in the visual cortex

Stephen Grossberg *, Praveen K. Pilly [1]

Department of Cognitive and Neural Systems, Center for Adaptive Systems, Center of Excellence for Learning in Education, Science, and Technology, Boston University, 677 Beacon Street, Boston, MA 02215, USA

ARTICLE INFO

Article history:
Received 23 July 2007
Received in revised form 19 February 2008

Keywords:
Motion perception
Direction discrimination
Decision-making
Visual cortex
Aperture problem
Motion capture
Noise-saturation dilemma
Recurrent competitive field
Bayesian inference
Stochastic decision models
Diffusion models
Speed–accuracy trade-off
Shunting model
On-center off-surround network
MT
MST
LIP
Basal ganglia
Psychometric function

ABSTRACT

How does the brain make decisions? Speed and accuracy of perceptual decisions covary with certainty in the input, and correlate with the rate of evidence accumulation in parietal and frontal cortical "decision neurons". A biophysically realistic model of interactions within and between Retina/LGN and cortical areas V1, MT, MST, and LIP, gated by basal ganglia, simulates dynamic properties of decision-making in response to ambiguous visual motion stimuli used by Newsome, Shadlen, and colleagues in their neurophysiological experiments. The model clarifies how brain circuits that solve the aperture problem interact with a recurrent competitive network with self-normalizing choice properties to carry out probabilistic decisions in real time. Some scientists claim that perception and decision-making can be described using Bayesian inference or related general statistical ideas, that estimate the optimal interpretation of the stimulus given priors and likelihoods. However, such concepts do not propose the neocortical mechanisms that enable perception, and make decisions. The present model explains behavioral and neurophysiological decision-making data without an appeal to Bayesian concepts and, unlike other existing models of these data, generates perceptual representations and choice dynamics in response to the experimental visual stimuli. Quantitative model simulations include the time course of LIP neuronal dynamics, as well as behavioral accuracy and reaction time properties, during both correct and error trials at different levels of input ambiguity in both fixed duration and reaction time tasks. Model MT/MST interactions compute the global direction of random dot motion stimuli, while model LIP computes the stochastic perceptual decision that leads to a saccadic eye movement.

1. Introduction

The brain's ability to carry out context-appropriate perceptually based decisions in response to ambiguous and probabilistic situations plays an essential role in ensuring animal and human survival. How the speed and accuracy of decisions varies with the ambiguity of environmental information is of particular interest (Gold & Shadlen, 2007; Luce, 1986).

A valuable paradigm for studying decision-making, which links psychophysics and neurophysiology, has been developed by Newsome, Shadlen, and colleagues (Roitman & Shadlen, 2002; Shadlen & Newsome, 2001). This research studies how brain dynamics in lateral intraparietal (LIP) area relate to saccadic behavior of monkeys (% accuracy, reaction time), that are based upon discriminating the motion direction of random dot motion stimuli at various degrees of coherence.

In these experiments, two kinds of tasks were employed: *fixed duration* (FD) and *reaction time* (RT) tasks. Macaques were trained to discriminate net motion direction and report it via a saccade. Random dot motion displays, covering a 5° diameter aperture centered at the fixation point on a computer monitor, were used to control motion coherence; namely, the fraction of dots moving non-randomly in a particular direction from one frame to the next in each of the three interleaved sequences (see Appendix A.1 for details about the motion algorithm).

* Corresponding author. Fax: +1 617 353 7755.
E-mail addresses: steve@bu.edu (S. Grossberg), advait@cns.bu.edu (P.K. Pilly).

[1] Authorship in alphabetical order. S.G. was supported in part by the National Science Foundation (NSF SBE-0354378), and the Office of Naval Research (ONR N00014-01-1-0624). P.K.P. was supported in part by the National Institutes of Health (NIH R01-DC-02852), the National Science Foundation (NSF IIS-02-05271 and NSF SBE-0354378), and the Office of Naval Research (ONR N00014-01-1-0624).

0042-6989/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2008.02.019

Varying the motion coherence provides a quantitative way to manipulate the ambiguity of directional information that the monkey uses to make a saccadic eye movement to a peripheral choice target in the judged motion direction, and thus the task difficulty. More coherence resulted in better accuracy and faster responses.

In the FD task (Roitman & Shadlen, 2002; Shadlen & Newsome, 2001), monkeys viewed the moving dots for a fixed duration of 1 s, and then made a saccade to the target in the judged direction after a variable delay. In the RT task (Roitman & Shadlen, 2002), monkeys had theoretically unlimited viewing time, and were trained to report their decision as soon as the motion direction was determined. The RT task allowed measurement of how long it took the monkey to make a decision, which was defined as the time from the onset of the motion until when the monkey initiated a saccade. The two monkeys in the Roitman and Shadlen (2002, p. 9476) were shaped during RT task training to initiate the choice saccade within "~1 s" after the dots turn on. In each RT task trial, the monkeys had to wait for a minimum of about 1 s (one monkey: 800 ms, and the other: 1200 ms) after motion onset to receive a reward however rapidly they responded. Human subjects in a similar RT task usually respond around 1 s from motion onset for the weakest coherence without any speed instruction (Palmer, Huk, & Shadlen, 2005, p. 385).

Neurophysiological recordings were done in LIP while the monkeys performed these tasks. The recorded neurons had receptive fields (RF) that encompassed just one target, and did not include the circular aperture in which the moving dots were displayed. Also, they were among those that showed sustained activity during the delay period in a memory-guided saccade task. It was found that the speed and accuracy of perceptual decisions covaried strongly with the rate of evidence accumulation in LIP cells.

Several authors (Gold & Shadlen, 2001, 2007; Jazayeri & Movshon, 2006; Ma, Beck, Latham, & Pouget, 2006; Rao, 2004) have suggested that these data illustrate Bayesian inference in the brain. Indeed, Gold and Shadlen (2001) have suggested that the "logarithm of the likelihood ratio (log LR) provides a natural currency for trading off sensory information, prior probability and expected value to form a perceptual decision" (p. 10). See Section 4.2 for their proposal. Despite their intuitive appeal, Bayesian models heretofore have not processed the perceptual stimuli that were used in the experiments, and have not disclosed novel brain mechanisms of decision-making.

Non-Bayesian models for the above dataset also exist (Ditterich, 2006a, 2006b; Mazurek, Roitman, Ditterich, & Shadlen, 2003; Wang, 2002), but none of them clarifies how the perceptual ambiguity that is created by the randomly moving dots is gradually transformed by the brain into a perceptual decision in response to the non-randomly moving dots. In particular, previous models have missed important brain principles and mechanisms that are at play in the dots task by ignoring the motion processing that extracts a dynamic neural representation of the directional uncertainty inherent in the random dot motion stimulus. They model decision-making properties only after assuming that the neural code of sensory uncertainty is provided. Section 4.3 details similarities and differences between our model and previous work in the field.

We here show how the data may be quantitatively simulated by a biophysically realistic model of interactions within and between Retina/LGN and cortical areas V1, MT, MST, and LIP, gated by basal ganglia (BG). The model achieves these results when it is directly activated by the visual stimuli that were used in the experiments (Fig. 1). Our results have been briefly reported in Pilly and Grossberg (2005, 2006). These model circuits solve two general design problems that are faced by the brain.

The *aperture problem* (Wallach, 1935) occurs whenever objects move with respect to spatially limited receptive fields of neurons: how can an unambiguous direction and speed of global object motion be determined from local motion signals that are ambiguous at most locations on the object?

The *noise-saturation problem* (Grossberg, 1973, 1980) occurs within all neurons because their activations fluctuate within a small interval of possible values: how does a network of cells, or cell populations, remain sensitive to the spatially distributed pattern of their inputs as they vary greatly in size through time? A special case of such networks shows how the most highly activated cell, or cell population, is selected to make a decision.

Our model's successful simulations of perceptual decision-making data support the hypothesis that the brain designs which solve the aperture problem and noise-saturation problem also underlie perceptual decision-making during random dot motion direction discrimination tasks. Individual moving dots do not experience the aperture problem. However, we claim that principles and mechanisms that have evolved in the brain to tackle the aperture problem can also help us intuitively understand the data at hand. The aperture problem is faced by any localized neural motion sensor, such as a neuron in the early visual pathway, which responds to a local contour moving through its aperture-like receptive field. Only when the contour within the aperture contains features, such as line terminators, object corners, junctions, high contrast blobs, or dots, can the local motion detector accurately signal the direction and speed of motion.

For example, the direction of motion of a featureless straight line seen behind an occluding aperture is thus ambiguous. When the aperture is circular, the line seems to move perpendicular to its orientation. When the aperture is rectangular, as during the barberpole illusion (Wallach, 1935), moving lines may appear to move in the direction parallel to the longer edges of the rectangle within which they move, even if their actual motion direction is not parallel to these edges. The brain must solve this problem in order to detect the correct motion directions of important moving objects in the world.

To overcome aperture ambiguities, the visual cortex embodies two complementary processes of *motion integration* and *motion segmentation*. The former process joins nearby motion signals into a single object, while the latter keeps them separate as belonging to different objects. The visual cortex uses the relatively few unambiguous motion signals arising from image features, called *feature tracking* signals, to inhibit the more numerous ambiguous signals from contour interiors. For example, during the barberpole illusion, feature tracking signals from the moving line ends along the longer edges of the bounding rectangle of the barberpole compute an unambiguous motion direction. These feature tracking signals gradually propagate into the interior of the rectangle. This *motion capture* process selects the feature tracking motion direction from the possible ambiguous directions along the lines within the rectangle, and suppresses the ambiguous motion signals corresponding to other directions that are generated by the moving lines (Ben-Av & Shiffrar, 1995; Bowns, 1996, 2001; Castet, Lorenceau, Shiffrar, & Bonnet, 1993; Chey, Grossberg, & Mingolla, 1997, 1998; Grossberg, Mingolla, & Viswanathan, 2001; Lorenceau & Gorea, 1989; Mingolla, Todd, & Norman, 1992). When a scene does not contain any unambiguous motion signals, the ambiguous motion signals cooperate to compute a consistent object motion direction and speed (Grossberg et al., 2001; Lorenceau & Shiffrar, 1992).

The brain thus needs to ensure that a sparse set of unambiguous feature tracking motion signals can gradually capture a greater number of ambiguous motion signals to determine the global direction and speed of object motion. In the case of random dot motion discrimination tasks, the signal dots at any coherence level produce locally unambiguous, though short-lived, motion signals.

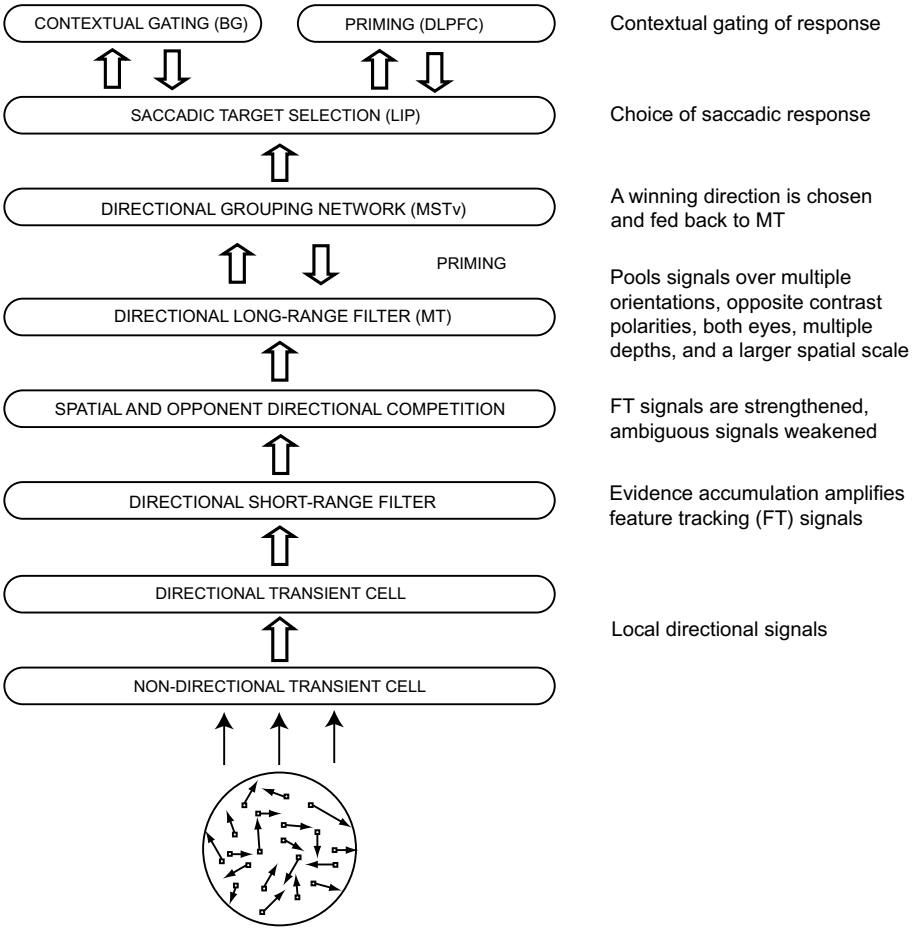

Fig. 1. Retina/LGN–V1–MT–MST–LIP–BG model processing stages. See text and Appendix for details.

The model shows how the same brain circuits that help resolve the aperture problem can also enable a small number of coherently moving dots to capture, as much as possible, the random motion directions caused by a large number of unambiguous, but incoherently moving, dots.

2. Model

Model processing stages are summarized in Fig. 1. Model equations and parameters are provided in Appendix A. The motion processing stages of the model were adapted from the Motion Boundary Contour System (mBCS) model, which clarifies how cortical areas V1, MT, and MST interact together to solve the aperture problem and create the brain's best representation of object motion direction and speed (Berzhanskaya, Grossberg, & Mingolla, 2007; Chey et al., 1997; Grossberg et al., 2001). The output from MST in the mBCS model is thus a *perceptual* representation that is not in the correct coordinates to command eye or arm movements towards a goal object.

The Motion Decision BCS (mdBCS) model, also known for short as the MOtion DEcision (MODE) model, that is developed herein adds an LIP decision-making circuit that discriminates motion direction based on inputs from the distributed motion representation of the model MST processing stage, and is gated by a simplified basal ganglia circuit. This enhanced model converts random dot motion stimuli into stochastic directional movement commands that are sensitive to the amount of directional coherence in the stimuli. The model processing stages are as follows.

2.1. Motion processing by V1, MT, and MST

The model change-sensitive receptors, non-directional transient cells and directional transient cells, compute local directional signals in response to image random dot motion. In particular, whenever a dot shows up at a spatial location, after being either randomly relocated or moved by a fixed displacement in the signal direction, a non-directional transient pulse is elicited at that location (see Appendix Eqs. (9)–(11)). This feeds into a directional transient cell network (Appendix Eqs. (12)–(14)) where local directional signals are computed. For example, suppose a dot arrives at location $[i + 1]$ from the leftward location $[i]$. Then the rightward inhibitory interneuron at location $[i]$ inhibits the leftward inhibitory interneuron and transient cell at location $[i + 1]$ well enough that they cannot recover to above-baseline firing when the dot does arrive at location $[i + 1]$ a little later (cf., Barlow & Levick, 1965). As a result, the leftward transient response is not obtained (see Fig. 2). Other directional transient cells at location $[i + 1]$ may be activated if they are not similarly inhibited in advance by their corresponding null direction interneuron at the location which is one unit from $[i + 1]$ in the preferred direction (see Appendix A.3.2 and Fig. 3a and b). In addition, directional inhibitory interneurons preserve direction sensitivity at a wide range of motion speeds. Appendix A.3.2 notes that these interneurons may be compared with starburst amacrine cells (Appendix Eq. (12)) and thus predicts that starburst cells may help transient cells to retain their directional sensitivity in response to motion at variable speeds.

The directional transient cell output signals feed into the directional short-range filter in V1, which accumulates monocular and

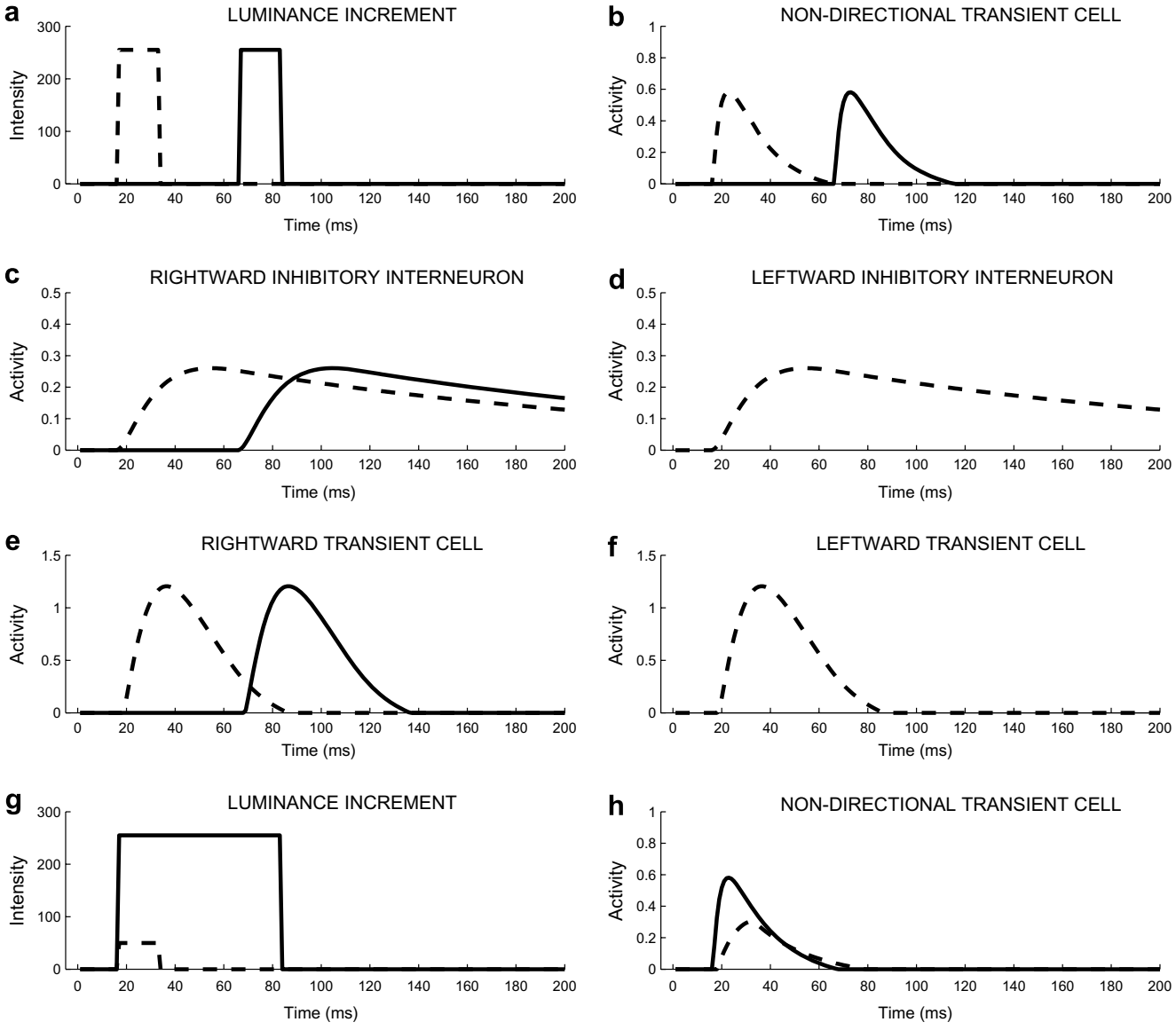

Fig. 2. Change-sensitive receptors, non-directional transient cells, and directional transient cell network: suppose a bright dot (intensity = 255) appears at location [i] on a dark background (intensity = 0) in one frame, and then again at location [$i + 1$] three frames later, with the frame rate 60 Hz, leading to a percept of rightward motion. Consider only two directions, Right and Left, for simplicity. Dashed and solid curves in (b–f) show rectified dynamics of cells in the model front end corresponding to locations [i] and [$i + 1$], respectively. (a) The luminance increment, (b) the non-directional transient cell response, (c) and (d) refer to activities of Right and Left inhibitory interneurons, and (e) and (f) correspond to Right and Left transient cell responses at the corresponding locations. Observe that for a luminance increment of size 255, the activities of non-directional transient cells and directional transient cells last about 50 ms and 70 ms, respectively. Note also that inhibitory interneurons remain active for a longer period, which ensures direction selectivity at a wide range of speeds. (e) and (f) reveal how directional selectivity manifests in the model. In (g) and (h), the effects of decreasing the luminance increment caused by dot onset and increasing the dot on-duration on the non-directional transient cell response is shown. Comparison of (h) and (b) reveals that the transient cells are only sensitive to the onset and size of the luminance increment (contrast) and not to how long it stays on. Luminance change duration, beyond a critical value, does not matter and lower contrast results in a weaker pulse and a delayed peak.

contrast polarity-sensitive evidence for motion in a given direction at each position, and thereby amplifies the feature tracking (FT) signals (Appendix Eqs. (15)–(17)). The output of the short-range filter is thresholded so that when a dot moves in a particular direction, the short-range filter output corresponding only to that direction survives. Thus, the model shows how individual dot motion creates unambiguous local motion signals, and hence does not suffer from the aperture problem (see Fig. 3a–c).

Such local directional processes can be fooled when there are multiple dots in each frame, some dots move incoherently, or independent random dot motion sequences are interleaved through time. The directional transient cells generate local directional signals between *any* two dots that occur with an appropriate spatio-temporal displacement, and the directional short-range filter integrates directional evidence from any active directional transient cells that occur within its directionally selective receptive field. Typically some directions will be amplified more than others by the short-range filter. However, lower motion coherence, higher dot density, and more interleaving of stimulus frames increase the probability that incorrect directional signals will be generated in the short-range filter, and thereby reduce the impact of correct local groupings (see Fig. 3d–f and g–i) in determining a clear motion directional percept. Apart from being few in number, these correct directional signals also have a short life span because a new set of signal dots are chosen every frame. Therefore, the motion stream must somehow enable a relatively sparse set of short-lived and correct feature tracking signals to gradually discount the more numerous incorrect local directional groupings.

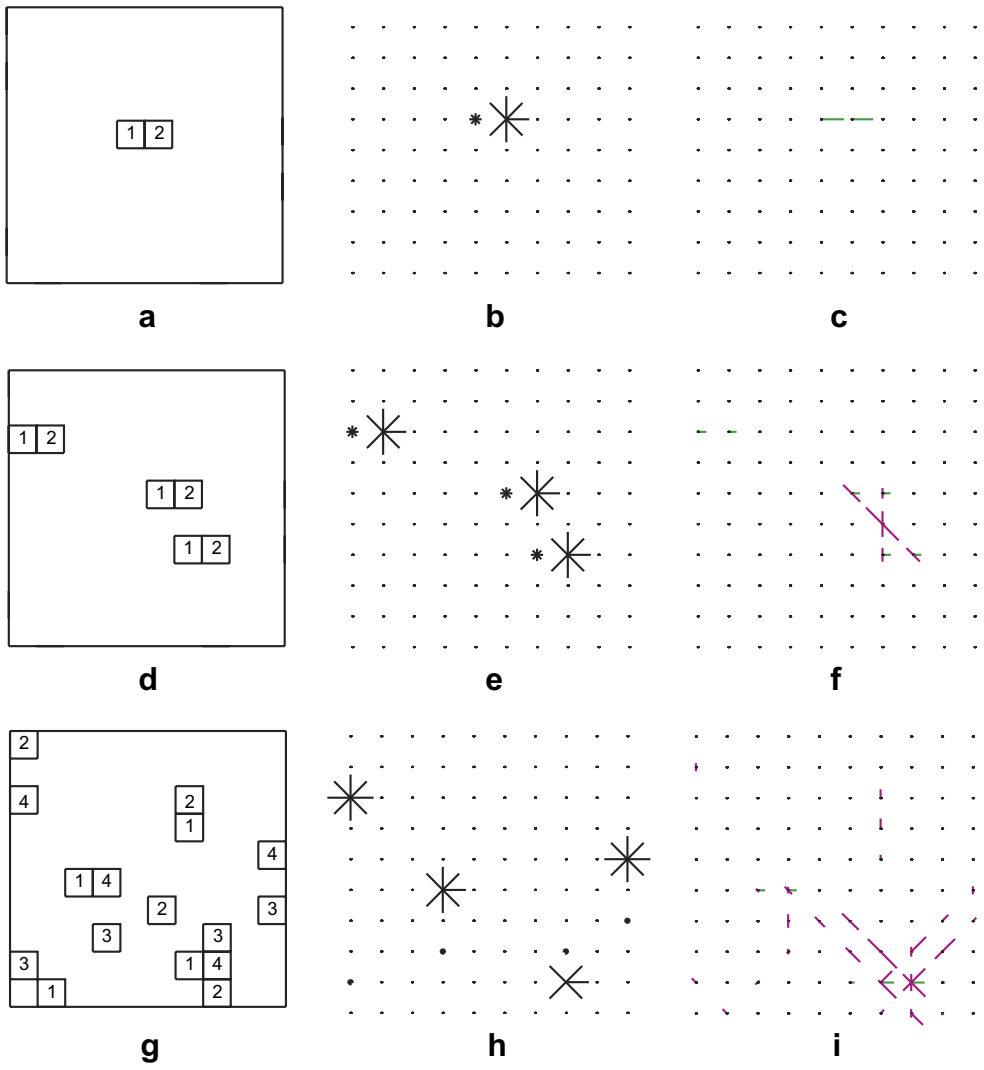

Fig. 3. Single moving dots do not suffer from aperture problem, but a random dot motion stimulus, which consists of multiple dots, does create local directional ambiguity. A vector representation is used to show sampled activities of rectified directional transient cells (b,e, and h) and thresholded directional short-range filter cells (c,f, and i). Since the model is sensitive to eight equally spaced directions at each spatial location, each point in the plot can potentially have eight vectors emanating from it with magnitudes that are proportional to the real-time rectified responses of the corresponding model cells. In (a), (d), and (g), the number on a dot refers to the frame (#) in which it appears. (a) A dot moves right by 1 pixel in one frame at 60 Hz, say from location $[i,j]$ to $[i+1,j]$. In (b), the temporal signature of the directional transient cells (waning at $[i,j]$ and waxing at $[i+1,j]$), and also inhibition of the null direction transient cell at location $[i+1,j]$ can be observed at $t = 60$ ms from the onset of the first frame. (c) The output threshold of the short-range filter stage is enough so that only the rightward directional signals (green) survive. In (d), three dots all move to the right from different locations at 60 Hz, two of which are relatively closer. Even at 100% coherence, (f) the short-range filters have, by $t = 60$ ms, detected inconsistent local groupings (magenta). In (g), four dots move rightward at 50% coherence with three frames interleaving at 60 Hz. The simulated transient cell activities sampled at $t = 100$ ms illustrate how inconsistent groupings can also arise because of random correlations due to interleaving of multiple sequences. These examples illustrate why, in order to perform random dot motion direction discrimination, the brain may need to use principles and mechanisms that are similar to those that help solve the aperture problem.

Spatial and opponent directional competitions (Appendix Eqs. (18)–(21)) further selectively strengthen FT signals, weaken ambiguous motion signals, and create speed-sensitive receptive fields. These cells are predicted to be a Where cortical stream analog of the simple cells for form perception that are well known to occur in the What cortical stream. A long-range filter then gives rise to true directional cells by pooling output signals with the same, or nearly the same, directional preference that survive the competitive stage (Appendix Eqs. (22)–(27)). Within each direction, these signals are pooled over multiple orientations, opposite contrast polarities, both eyes, and a larger spatial scale. The cells that receive inputs from the directional long-range filter are also predicted to be depth-selective. They are predicted to occur in MT, and to be analogous to complex cells in the form pathway. Neither of the properties of sensitivity to speed, and pooling across multiple orientations, opposite contrast polarities, both eyes, and a range of depths, is needed in our current simulations of direction discrimination. Also, form–motion interactions between V2 and MT (Berzhanskaya et al., 2007) are not invoked because motion processing is rate-limiting in response to random dot motion stimuli.

The directional cells activate a directional grouping network, proposed to occur within cortical area MST (more specifically, MSTv), within which all directions compete to determine a winner at each position (Appendix Eq. (28)). Enhanced feature tracking signals typically win this competition over ambiguous motion signals at their positions. Motion capture begins when model MST cells that encode the winning directions feed back to model MT cells via a top-down spatial filter (Appendix Eq. (23)), where they indirectly boost directionally consistent cell activities by suppressing inconsistent directional cells over the spatial region to which they project. Our simulations were performed with and without

internal cellular noise in both MT and MST in order to assess the effects of such noise on the model's stochastic decision-making.

This shift in the spatial locus of unambiguous feature tracking signals continues to propagate across space as the MST-to-MT feedback process cycles through time. The action of this feedback loop was predicted to solve the aperture problem, and to generate a representation of global object direction and speed (Berzhanskaya et al., 2007; Chey et al., 1997; Grossberg et al., 2001). Pack and Born (2001) have reported neurophysiological data which support this prediction.

The rate and strength of motion capture by the MT/MST feedback loop is reflected in the decision-making properties of model LIP, as it receives its inputs from MST. Model LIP hereby simulates the temporal dynamics of decision-making in the neurophysiological and behavioral data. The intuitive idea is that the feedback loop needs more time to capture the incoherent motion signals, and cannot achieve as high a level of asymptotic response magnitude, when there are more of them competing with the emerging winning direction. A key point of this article is thus that the effectiveness of the motion capture process depends on the input coherence and exposure duration. LIP converts the spatially distributed directional motion signals from MST into an eye movement command, in the manner noted below, and thereby enables the monkey to report its decision via a saccade.

2.2. Probabilistic decision-making by LIP with BG gating

As noted above, the motion information from the model MST stage (Fig. 1) is spatially distributed and needs to be converted into a form where it can command an eye movement to the choice target that corresponds to the judged net motion direction. These targets are present at a specific eccentricity in the respective directions of the motion during both the training and the recording phases of the experiments. In these experiments, LIP cells do exhibit decision-related activity that correlates with a saccadic eye movement to one of the two or more possible choice targets. The model LIP circuit (see Appendix Eqs. (29)–(34)) converts the spatially distributed directional motion signals into the activation of cells that code for saccadic eye movements in specific directions.

This LIP circuit is modeled using a kind of decision circuit that has become classical in the neural modeling literature; namely a network of cells that obey membrane equation, or shunting, dynamics and interact via a recurrent on-center off-surround network (Grossberg, 1973, 1980). Such a network is often called a recurrent competitive field (RCF), and its variants have been used by a number of authors to model the dynamics of perceptual or motor decisions in both deterministic models (e.g., Brown, Bullock, & Grossberg, 2004; Chey et al., 1997; Francis, Grossberg, & Mingolla, 1994; Francis & Grossberg, 1996a, 1996b) and stochastic models (Boardman, Grossberg, Myers, & Cohen, 1999; Cisek, 2006; Grossberg, Boardman, & Cohen, 1997; Grossberg & Myers, 2000; Usher & McClelland, 2001). It should be noted that even a deterministic RCF typically describes the temporal evolution of population mean activities and cell firing frequencies. It is only when properties that depend upon the variance of cell firing become rate-limiting that explicit noise terms add explanatory power.

There is another sense in which RCFs, and indeed all shunting on-center off-surround models, embody probabilistic properties. The shunting dynamics of a RCF lead to automatic gain control and self-normalizing properties, so that the total activity of a processing channel in such a model is often approximately conserved. This total activity plays the role of a real-time probability distribution. Self-normalization enables such a network to maintain its sensitivity in response to distributed inputs whose total number and size can vary wildly through time. This property is useful when estimating the directional coherence of inputs whose total number and distribution can vary randomly through time, as occurs in the experiments that are simulated in this article. This robustness is illustrated by the fact that model properties remain qualitatively unchanged when cellular noise is added to MT and MST, as illustrated in Fig. 15.

The RCF that models LIP herein has the following properties (see Fig. 4 and Appendix Eqs. (29) and (34)):

(1) Each cell activity (y_d) corresponding to movement direction d in the LIP RCF is associated with the peripheral choice target within its response field.
(2) Each LIP cell gets excited by the spatially distributed population activity of the foveal MST pool of neurons coding its preferred motion direction (see term S_d in Appendix Eqs. (29) and (34)). It is assumed that these pro-saccade connections are gradually strengthened as a result of extensive operant conditioning on the task. As a result, each cell's activity predicts that the global motion of the random dots is in the direction from the fovea to its choice target.
(3) Each cell also receives a bottom-up excitatory input due to the presence of a choice target within its response field (see term T_C in Appendix Eqs. (29) and (34)). This input produces the above-baseline activity observed before the onset of the dots.
(4) Each cell receives recurrent self-excitation ($f(y_d)$) via a sigmoidal signal function, and recurrent inhibition ($h(y_D)$) from other cells via another sigmoidal signal function (Grossberg, 1973, 1980).
(5) Each LIP cell is subject to an individual internal noise process that influences stochastic choice dynamics (see term W in Appendix Eqs. (29) and (34)). A similar hypothesis has been used to quantitatively simulate the temporal dynamics of speech categorization data (Boardman et al., 1999; Grossberg, Boardman, et al., 1997; Grossberg & Myers, 2000). This noise process contributes to the known variability in the read-out from sensory to motor areas. Model simulations show that the LIP cellular noise, in combination with the randomness of the moving dot stimuli, can explain the variability in saccadic decision-making, whether or not there is noise at the MT and MST processing stages.
(6) During the RT task, when one of the competing LIP cells first reaches a fixed decision threshold (Γ_1), a directional decision is initiated by opening a basal ganglia (BG) movement gate that increases the gain of the corresponding cell's self-excitation (see g_d in Appendix Eq. (29)) to a high value (see g_{BG} in Appendix A.3.6). This event triggers the final stage of LIP firing, namely pre-saccadic enhancement that is independent of motion coherence. LIP hereby transitions from the sensory mode to the motor mode. All LIP cells continue to integrate their sensory inputs, but the selected cell population does so at a higher self-excitatory gain. A saccade to the associated target is initiated when the winning cell's firing rate reaches a criterion level (Γ_2). The reaction time (RT) is the time from stimulus onset to when the choice saccade is made, not the time when the decision is initiated. Both decision time and RT are emergent properties, and covary with each other.
(7) As noted in Section 1, Roitman and Shadlen (2002) delayed reward until ~1 s after motion onset during RT task training. During the RT task, mostly during weaker coherence trials, cells in the RCF may take longer than 1 s to reach the decision threshold. If none of the activities sampled at a random time between 1000 and 1100 ms has crossed the decision threshold (see term G_d in Appendix Eq. (29)), then a volitional top-down signal boosts the LIP cell with the largest activity. A choice made in this manner is termed a "forced

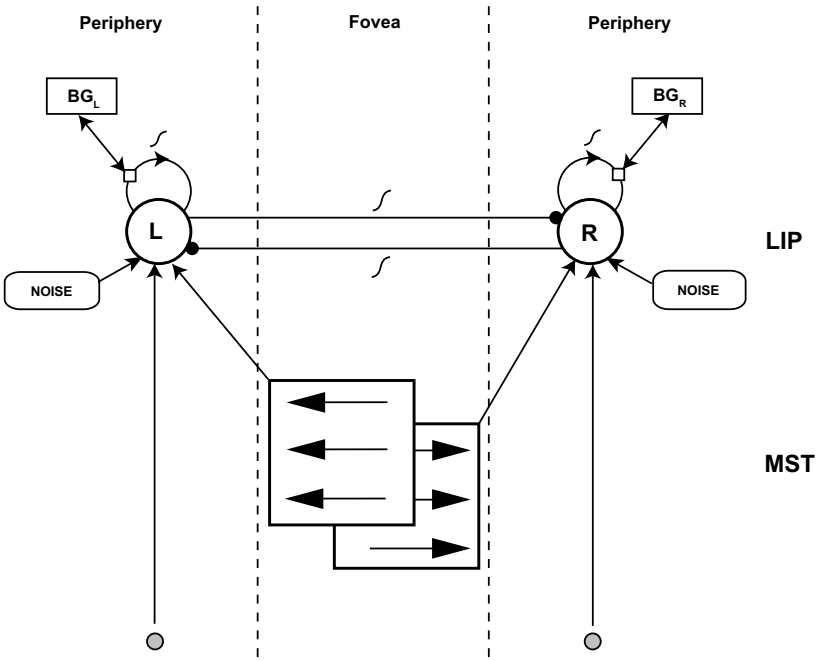

Fig. 4. Components of the LIP decision circuit. The activities of the cells in the recurrent competitive field predict the perceptual decision regarding the direction of the randomly moving dots. Each cell receives two bottom-up inputs: one from the form stream due to the presence of the peripheral choice target (indicated by a small circle) in the response field, and the other from the foveal MST pool of the motion stream tuned to its preferred direction. The latter sensory input comes from outside the response field. This non-classical neural connection is thought to result from training on the stereotypical task. Also, each cell self-excites itself and is inhibited by other competing cells in the field via different sigmoidal signal functions, which are indicated on the respective connections. The gain of the recurrent self-excitation is regulated by the basal ganglia. Two time-appropriate top-down inputs, an inhibitory one that simulates post-saccadic suppression, and an excitatory one that helps the decision circuit make a forced choice in trials of extended uncertainty, are not depicted. Internal noise processes to each cell help to simulate the probabilistic nature of perceptual decisions.

choice". Section 3.2 cites evidence that support this hypothesis. The model monkey is also simulated without this volitional mechanism, and can still make coherence-sensitive choices, but some longer RTs are generated as a result.

(8) Cell activities decay faster in the FD task, in keeping with the experimental observation that the gain of the LIP response is smaller in the FD task than in the RT task (Roitman & Shadlen, 2002, p. 9485), and also that LIP responses tend to saturate during the viewing duration in the FD task.

(9) During the FD task, at the end of the viewing duration of 1 s, the model chooses the motion direction corresponding to the maximally active LIP cell, by increasing the self-excitatory gain of this cell (see term g_{delay} in Appendix A.3.6). This event causes a persistent and slow buildup activity during the variable delay period after the input from motion areas shuts off. The basal ganglia increase this gain further (term g_{d} in Appendix Eq. (34) jumps to g_{BG}), resulting in coherence-independent pre-saccadic enhancement, as in the RT task.

(10) After a saccade begins, all LIP cells receive a strong inhibitory signal. In vivo, the source of this signal is possibly from frontal eye field (FEF) post-saccadic cells (Brown et al., 2004; Bruce, Goldberg, Bushnell, & Stanton, 1985) (see term G_{I} in Appendix Eqs. (29) and (34)).

Section 3 discusses in detail how these LIP RCF properties help to explain the simulated data.

3. Results

3.1. Data summary

The recorded LIP neurons show visuo-motor responses. They have properties of both buildup and burst cells (Munoz & Wurtz, 1995a) that are found in superior colliculus (SC). First, visual targets present in the receptive fields contribute to above-baseline LIP firing before the dots turn on. Even though the motion stimulus is not presented within their classical receptive fields, these neurons still respond with direction selectivity, probably because of extensive training on the tasks during which new stimulus–response associations are learned (Bichot & Schall, 1996). This property has also been observed for SC neurons in monkeys trained to perform an FD task (Horwitz, Batista, & Newsome, 2004a; Horwitz & Newsome, 2001a).

On correct trials during the decision-making period, more coherence in the preferred direction causes faster LIP cell activation, on average, in both the tasks (Fig. 5), and also higher maximal cell activation in the FD task (Fig. 5c–f). More coherence in the opposite direction causes faster cell inhibition in both the tasks, and also lower minimal cell activation in the FD task. Thus on correct trials, the instantaneous difference in average LIP activity between judgments of net motion being towards the receptive field (T_{in} choices) and judgments of motion being away from the receptive field (T_{out} choices) increases with coherence. In other words, the correct trial predictiveness of LIP cell responses is proportional to % coherence.

The temporal dynamics of LIP decision neurons also correlate with behavioral properties of perceptual decision-making (Fig. 6). In both FD and RT tasks, more coherence in the motion translates into more accurate decisions (Fig. 6a and b). Also, RT task accuracy at weaker coherence levels is slightly better than FD task accuracy. The psychometric function ranges from about chance level to 100% accuracy as the dot coherence varies from 0% to 100%, respectively. In addition, a longer viewing duration in the FD task tends to improve performance at all motion strengths (Fig. 2A and B in Gold & Shadlen, 2003), revealing a speed–accuracy trade-off (Fig. 7) wherein performance asymptotes at shorter durations for higher coherences. More coherence also results in faster reaction times

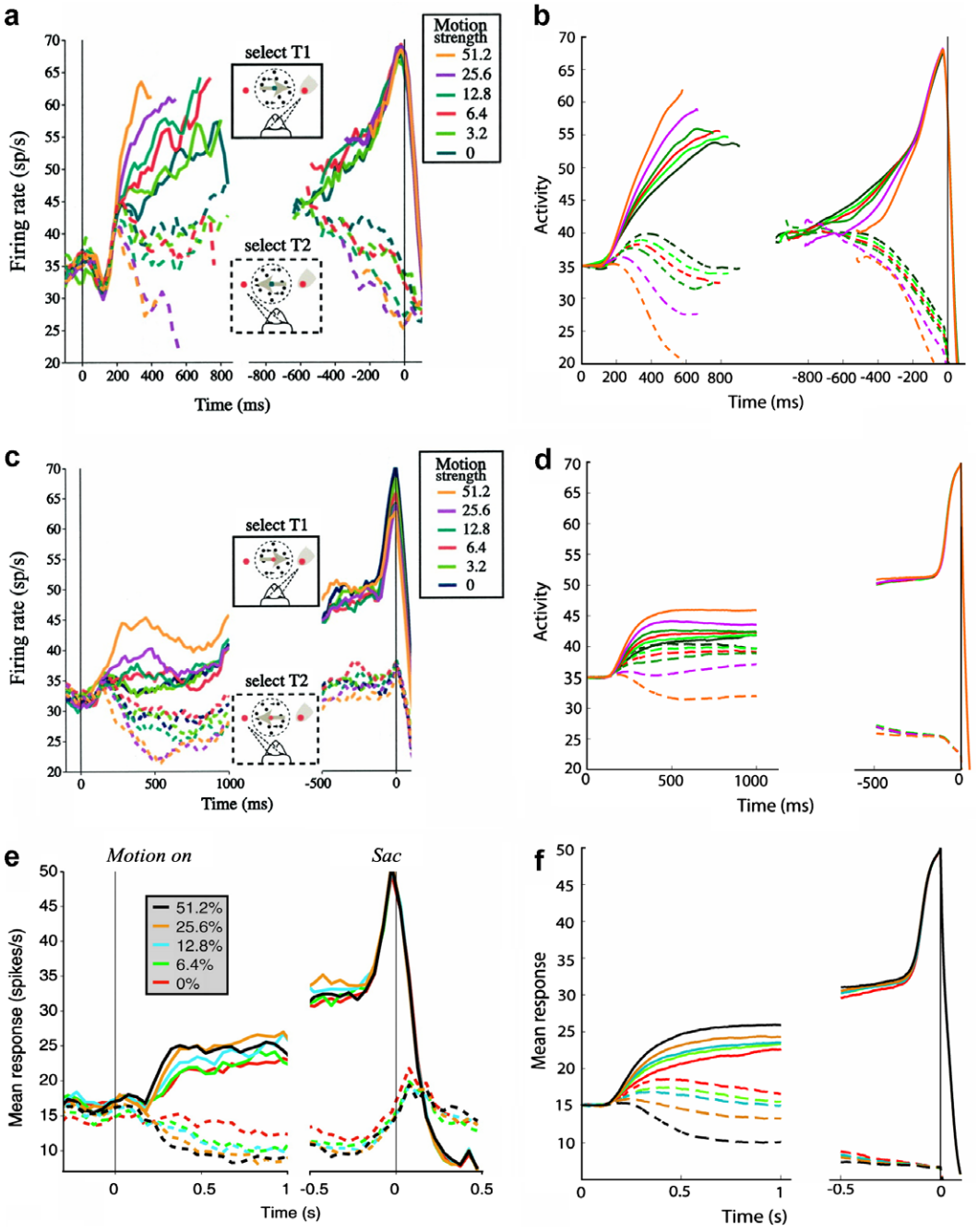

Fig. 5. Temporal dynamics of LIP neuronal responses during the fixed duration (FD) and reaction time (RT) tasks. (a) Average responses of a population of 54 LIP neurons among correct trials during the RT task (Roitman & Shadlen, 2002). The left part of the plot is time-aligned to the motion onset, and includes activity only up to the median RT, and excludes any activity within 100 ms backward from saccade initiation (which roughly corresponds to pre-saccadic enhancement in firing). The right part of the plot is time-aligned to the saccade initiation, and excludes any activity within 200 ms forward from motion onset (which corresponds to initial transient pause in firing). (b) Model simulations replicate LIP cell recordings during the RT task. In both data and simulations for the RT task, the average responses were smoothed with a 60 ms running mean. (c) Average responses of a population of 38 LIP neurons among correct trials during the 2002 FD task (Roitman & Shadlen, 2002), during both the motion viewing period (1 s) and a part (0.5 s) of the delay period before the saccade is made. (d) Model simulations mimic LIP cell recordings during the 2002 FD task. (e) Average responses of a population of 104 LIP neurons among correct trials during the 2001 FD task (Shadlen & Newsome, 2001), during both the motion viewing period (1 s) and a part (0.5 s) of the delay period before the saccade is made. (f) Model simulations emulate LIP cell recordings during the 2001 FD task. In (a–f), solid and dashed curves correspond to trials in which the monkey correctly chose the right target (T_{in}) and the left target (T_{out}), respectively. Cell dynamics (rate of rise or decline, and response magnitude) reflect the incoming sensory ambiguity (note the different colors; the color code for the various coherence levels is shown in the corresponding data panels), and the perceptual decision (solid: T_{in} choices, dashed: T_{out} choices). Note that for 0% coherence, even though there is no correct choice per se, the average LIP response rose or declined depending on whether the monkey chose T_{in} or T_{out}, respectively. [Data in (a,c) and (e) is reprinted with permission from Roitman and Shadlen (2002) and Shadlen and Newsome (2001), respectively.]

in the RT task (Fig. 6c and d). Moreover, the monkey responds with relatively slower reaction times on error trials when compared to correct trials. Reaction time standard errors of mean (SEM) decrease with coherence on correct trials, and increase with coherence on error trials. Error trial SEMs are greater than those for correct trials (Table 2 in Roitman & Shadlen, 2002).

Cell responses on error trials and 0% coherence trials during both FD and RT tasks reveal that LIP firing reflects the perceptual

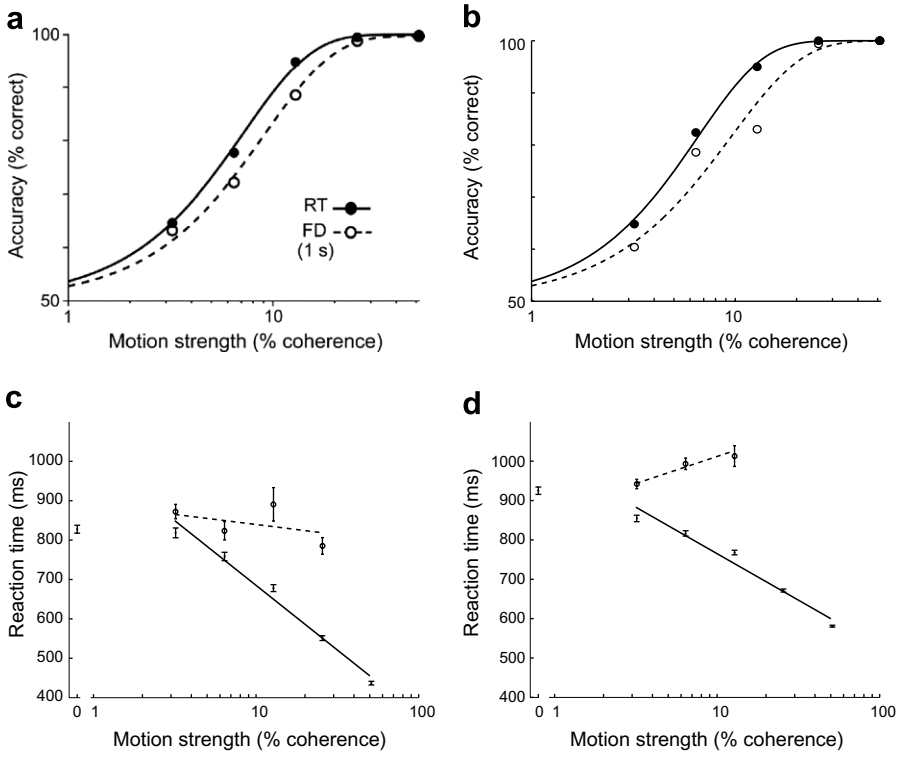

Fig. 6. Psychometric and chronometric data during the FD and RT tasks (Roitman & Shadlen, 2002). (a) Accuracy data (% correct) as a function of motion coherence (% certainty) is fit using a cumulative Weibull distribution function for both FD and RT tasks (Fig. 6B in Mazurek et al., 2003). The ability to discriminate motion direction depends on the stimulus strength. Accuracy in the RT task is slightly better than that in the FD task for lower coherence levels. (b) Model simulations emulate these data. Solid curve corresponds to the RT task, and dashed curve to the 1s FD task. Number of trials is 500. The two parameters, α (threshold) and β (steepness), obtained for each psychometric function were: $\alpha_{RT} = 6.5746$, $\beta_{RT} = 1.3466$, $\alpha_{FD} = 9.5544$, and $\beta_{FD} = 1.2379$. Note that $\frac{\alpha_{RT}}{\alpha_{FD}} < 1$. (c) Reaction time data (ms) as a function of motion coherence (% certainty) is linear fit using a weighted, least-squares estimate (as per the convention in Fig. 3B of Roitman and Shadlen, 2002). The plot is prepared from the data for T_{out} (left target) choices in Table 2 of Roitman and Shadlen (2002). Data for T_{in} (right target) choices gives a similar plot. Less ambiguity implies a faster decision. Solid line corresponds to correct trials, and dashed line to error trials. Error bars shown are standard errors of mean (SEM). SEMs decrease with coherence on correct trials, but increase with coherence on error trials. Moreover, error trials have relatively higher SEMs. (d) Model simulations emulate the RT data on both correct and error trials. Note in particular that the model is able to produce slower error trial RTs, unlike the alternative model in Mazurek et al. (2003). Also, the behavior of SEMs with respect to coherence and correctness of trials is captured in the simulations. The number of trials is 500. In (a–d), the abscissa is in the $\log_{10}$ scale. [Data in (a) is reprinted with permission from Mazurek et al. (2003).]

decision, regardless of the true direction and strength of the random dot motion stimulus (Figs. 8 and 9). This becomes particularly apparent on weaker motion strength trials when the monkey is prone to making wrong decisions. It is this feature that distinguishes the so-called "decision" LIP responses from "sensory" MT and MST responses. However, the error trial predictiveness of LIP cells decreases as coherence increases. That is, on error trials, the difference in average LIP activity between T_{in} and T_{out} choices is inversely proportional to percent coherence (see Fig. 9 in particular). Also, the gain, or rate of growth, of average LIP responses for erroneous T_{in} choices is reduced when compared to that on correct trials, and is further reduced with coherence. This is also true for the rate of decline in average LIP activity for T_{out} choices. To observe these interesting properties, please compare solid color (T_{in} correct) with dashed gray (T_{in} error) curves in Fig. 8 (RT task), and solid black (T_{in} correct) with dashed black (T_{in} error) curves in Fig. 9 (FD task).

Moreover, on trials resulting in correct T_{in} choices, coherence does not differentiate the final stages of LIP firing, ~100 ms before the saccade begins, when motor signals dominate the LIP response (see Figs. 5, 8, and 9; top right portion of the plots). On the other hand, for correct T_{out} choices in the RT task, coherence has a systematic influence on the LIP cell response throughout the trial (see Fig. 5a and b, bottom right portion).

An analysis of the relationship between LIP response and RT reveals that LIP encodes not only "where, but also when, to move the eyes" (Roitman & Shadlen, 2002, p. 9485), since RT correlates with the rate of buildup in LIP response for correct T_{in} choices (Fig. 10) at each motion strength.

Another interesting characteristic of LIP physiology is the coherence-independent dip and then rise in activity, lasting about 100 ms, that begins approximately 90 ms after motion onset in the RT task (Fig. 5a). Interestingly, this stimulus-insensitive transient pause in LIP firing is not that prominent in the FD task; see Fig. 5c and e.

When both FD and RT tasks are conducted on the same set of monkeys in alternating block of trials, LIP neuronal recordings reveal that the gain of the LIP response is smaller in the FD task than in the RT task (Roitman & Shadlen, 2002), and also that LIP responses tend to saturate during the fixed viewing duration in the FD task. Compare Fig. 5a and b with c and d.

Mazurek, Hanks, Yang, and Shadlen (2005) manipulated prior probability in the RT task. As the odds of one direction being correct in a block of trials was increased, the monkey responded with relatively more accuracy and faster RTs to motion in the more probable direction, and less accuracy and slower RTs to motion in the other direction at all coherence levels. The rate of growth or decay in LIP activity modulated as if there was some extra coherent motion in the biased direction. This bias also caused a slight positive offset in the activity of the corresponding LIP population before motion onset (reported in Shadlen & Newsome, 2001 too). The effect of varying the number of choices has also been studied (Churchland, Tam, Plamer, Kiani, & Shadlen, 2005), with more choices resulting in relatively slower reaction times, lesser

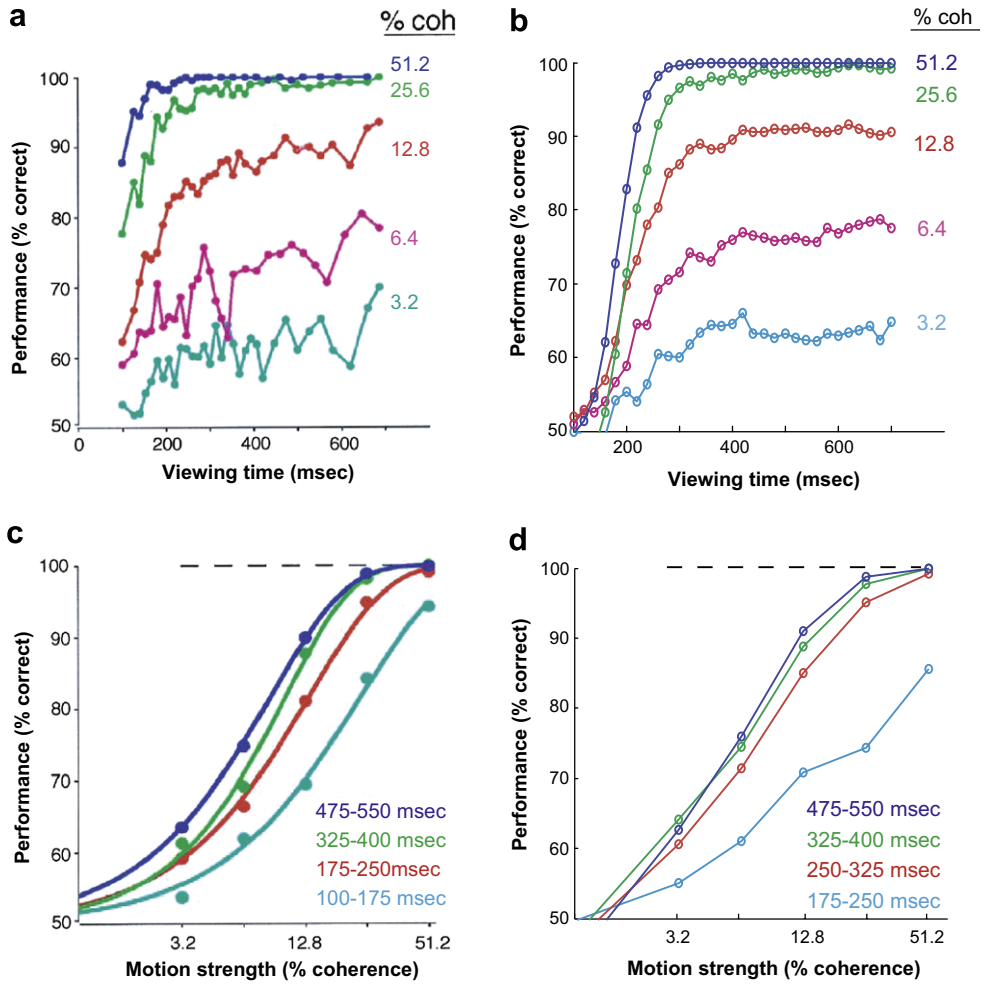

Fig. 7. Influence of viewing duration on performance at various coherences in the FD task paradigm. (a) Data from Gold and Shadlen (2003) shows that, the more time the dots are observed, the better is the performance. This effect saturates at every coherence level. (b) Model simulations (2001 FD task) reproduce this influence of viewing time. (c) The psychometric function as a function of duration ranges. More viewing time tends to shift the psychometric function to the left, thus reducing the discrimination threshold. (d) The simulated psychometric functions capture these data trends. [Data in (a,c) is reprinted with permission from Gold and Shadlen (2003).]

accuracy, and lower firing rates at the beginning of the decision-making epoch.

Recordings from MT and MST neurons with random dot motion presented within their receptive fields revealed roughly linear relationships with positive and negative slopes between response magnitude and coherence of motion in the preferred and null directions, respectively (Britten, Shadlen, Newsome, & Movshon, 1993; Celebrini & Newsome, 1994); see Fig. 11.

3.2. Data explanations

This section explains the data properties that have just been summarized. Section 2 discusses how the random dot motion stimulus causes local directional ambiguity, similar in many respects to the aperture problem, in the short-range filter stage (V1), and how recurrent processing in the MT–MST circuit allows the locations with aperture ambiguities to be captured by FT signals. The main new explanatory concept in analyzing the current dataset is that, in the dots task, the effectiveness of this motion capture process is limited by the coherence of the moving dots, and also by the viewing duration. This effectiveness is reflected in a bigger contrast or difference in activity among MT/MST directional cell populations (Fig. 11). This section explains how better performance and faster reaction time may proportionally be derived from this dynamic difference. The time course of the activities in model MT and MST are consistent with those recorded in MT (Britten et al., 1993). One interesting point in Fig. 11 from both data and simulations is that the motion-sensitive neurons are active well above the baseline for a 0% coherence stimulus. This can be explained by noting that, if there is no stimulus-induced directional bias, as is the case at 0% coherence, then random local groupings are formed equally in all directions.

Although MT and MST provide the trial-to-trial neural basis of directional ambiguity on which decisions are made, they exhibit low choice probabilities (MT: Britten, Newsome, Shadlen, Celebrini, & Movshon, 1996; MST: Celebrini & Newsome, 1994); that is, they "covary only weakly with what the animal decides" (p. 1930 in Shadlen & Newsome, 2001) especially at weaker coherences. When monkeys are trained to report the perceived direction through a saccadic eye movement, the recorded dynamics of appropriate LIP cells correlate strongly with decision-making behavior. It is on most error trials that MT/MST ("sensory") and LIP ("decision") cells differ in their winning direction. The current consensus in neuroscience is that some sort of a noisy accumulation of sensory signals is the basic mechanism underlying perceptual decision-making in the brain (Smith & Ratcliff, 2004; see Section 4.3 for more references).

In our model, LIP is modeled as a recurrent competitive field in which individual cells are selective for motion directions of foveal stimuli presented outside their response fields. Distinct targets that

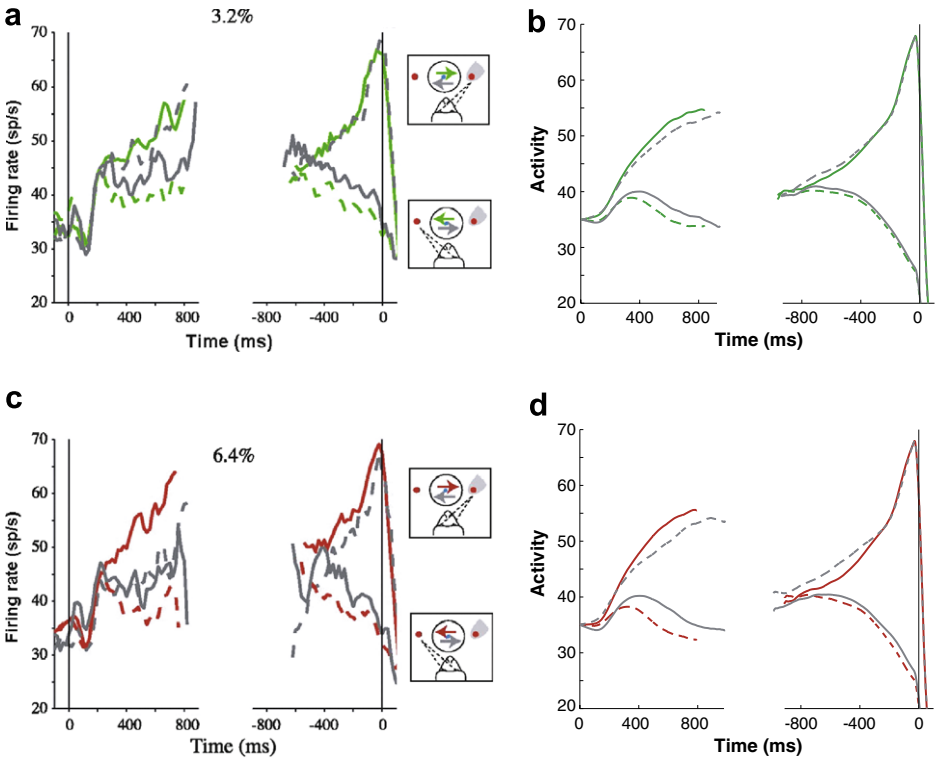

Fig. 8. Temporal dynamics of LIP responses during error trials in the RT task (Roitman & Shadlen, 2002) for two low coherence levels. Same conventions as in Fig. 5a are followed. (a) Average responses from 54 LIP neurons on correct and error trials during the RT task for 3.2% coherence. (b) Model simulations replicate LIP cell recordings on error trials of the RT task for 3.2% coherence. (c) Average responses from the same population of LIP neurons on correct and error trials during the RT task for 6.4% coherence. (d) Model simulations again capture the data. In (a–d), the colored curves represent correct trials, and the gray curves represent error trials. Solid and dashed curves correspond to input stimuli whose motion is towards or away from the right target (T_{in})/receptive field, respectively. LIP responses on error trials show that LIP reflects both the choice the monkey makes, and also the true direction and strength of the dots. Gray curves in the left portion of the plots show that the rates of buildup and decline in average LIP activity are relatively lower on error trials. Also, note that the median RT is relatively more on the error trials. [Data in (a,c) is reprinted with permission from Roitman and Shadlen (2002).]

are visible at a specific eccentricity around the foveal locations, where the dots are displayed, guide the possible alternatives of motion direction the monkey *has* to choose from. Each LIP cell population receives convergent excitatory directional inputs from spatially distributed cells in foveal MST. This distributed activation provides evidence for the perceptual decision that the global motion is in the direction from the fovea towards its response field. Operant conditioning from foveal MST neurons to these non-foveal LIP neurons is assumed to strengthen the corresponding fovea-to-response field connections. This learning process is not simulated here.

In the model, a choice target creates an excitatory input whose magnitude is enough for the cells to achieve above-baseline activity before the dots turn on (see Fig. 5 and Appendix Eq. (31)). While each LIP cell gets excited, it is also inhibited by other active LIP cells in the recurrent competitive field. Since a higher percent coherence in a particular direction, on average, causes faster and more widespread captured motion signals in MST cells, it also causes in LIP cells faster activation via the recurrent on-center and faster inhibition via the recurrent off-surround for the preferred and non-preferred LIP cells, respectively (Fig. 5).

The model's LIP-BG loop (Figs. 1 and 4) gates the release of choice saccades during both the FD and RT tasks. It is well known that such a BG gate controls the release of saccades in vivo; e.g., Hikosaka and Wurtz (1983, 1989). In the RT task, a decision is made after either directional LIP activation exceeds a threshold. In particular, stronger motion strength results in faster rise of activity for one of the LIP cells to the threshold, causing faster decisions, and thereby faster reaction times (Fig. 6c and d). In the FD task, monkeys are trained to wait to move until the fixation point is extinguished. In the model, fixation point extinction triggers a GO signal. Thus, for a saccade to be initiated from the fixation point to the chosen target in either task, the basal ganglia need to first open the gate that releases the final decision stage in the target LIP cell population. This is computationally achieved in the model by switching the gain of self-excitation of the selected LIP cell to a higher value. Strong recurrent self-excitation then gets activated which manifests as steep pre-saccadic enhancement, or burst, in LIP firing just before the eyes move. In the model, this gain switches to a high enough value such that the recurrent signal outweighs any differential sensory excitation. This property is what makes the LIP cell firing for correct T_{in} choices independent of percent coherence in the post-threshold-crossing (RT task, Fig. 5a and b) or post-GO signal (FD task, Fig. 5c–f) stage. This is when LIP switches, as it were, from the sensory driven mode to the motor decision mode. For correct T_{out} choices, coherence plays a role even in the final stages of LIP firing in the RT task because the model BG do not similarly increase the gain of recurrent inhibition from the selected cell (Fig. 5a and b).

This simple BG mechanism is consistent with detailed models of how basal ganglia gates may be dynamically controlled when monkeys learn to do a range of saccadic eye movement tasks (Brown et al., 2004). The LIP burst response is transformed into the execution of the appropriate saccadic eye movement further downstream in the brain. Several modeling articles about saccadic and smooth pursuit eye movements detail these subsequent eye movement stages (Gancarz & Grossberg, 1998, 1999; Grossberg, Srihasam, & Bullock, submitted for publication; Srihasam, Bullock, & Grossberg, 2008).

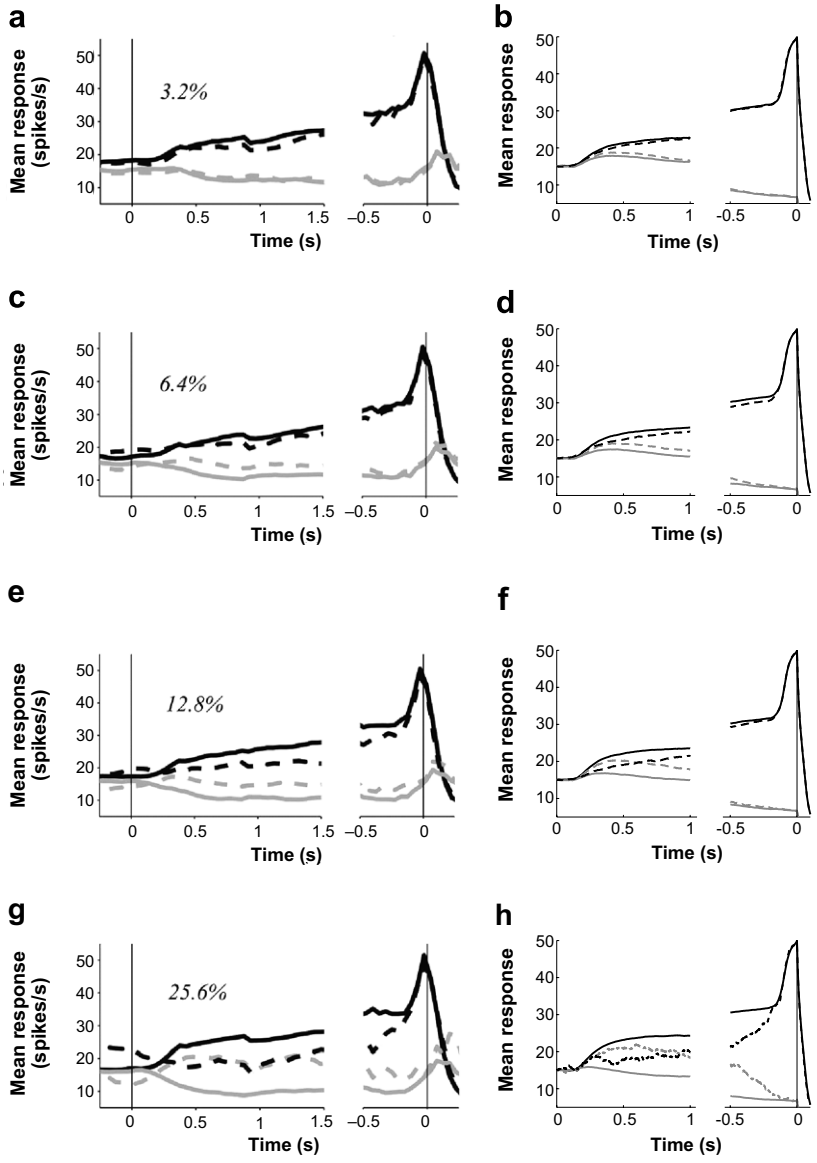

Fig. 9. Temporal dynamics of LIP responses during error trials in the 2001 FD task (Shadlen & Newsome, 2001). (a, c, e, and g) Average responses of LIP neurons on correct and error trials during the FD task for the coherence levels: 3.2%, 6.4%, 12.8%, 25.6%, respectively. Data is from trials in which the viewing duration was 2 s. (b, d, f, and h) Model simulations of 2001 FD task (1 s viewing duration) capture the data features. In (a–h), the left portion of the plots correspond to mean responses during the motion viewing period, and the right portion of the plots correspond to those during the delay period, 500 ms backward from saccade initiation. Solid curves represent correct trials, and dashed curves error trials. Black and gray curves correspond to trials in which the motion was decided to be towards and away from the right target/receptive field, respectively. Cell dynamics on error trials show that LIP codes both the perceptual decision and the stimulus. The rate of buildup in response and the response magnitude (black dashed curves) are relatively lower on error trials, and so is the rate of decline in activity (gray dashed curves). On error trials, especially at higher coherences, LIP activity takes relatively more time to become "predictive" of the decision. [Data in (a,c,e,g) is reprinted with permission from Shadlen and Newsome (2001).]

High choice probability in LIP is possible because the internal noise in each LIP cell, modeled as a Brownian motion process (Appendix Eqs. (29) and (34)), enables the LIP cells to proportionately make erroneous choices at lower coherence levels. Stronger motion strength implies a relatively stronger bottom-up excitatory input from MST to the preferred LIP cell, which dilutes the effect of noise fluctuations on its dynamics at each time instant. In such a case, the noise does not hinder the LIP cell from firing faster and reaching higher levels of firing, on average. The noise has relatively more influence on the moment-by-moment dynamics of LIP response at weaker coherences.

An error trial occurs when a non-preferred LIP cell manages to be the first to reach the decision threshold in the RT task, or to have the highest activity by the end of the viewing duration in the FD task. This can occur because of two reasons or a combination of both: (a) On an error trial, the net sensory evidence in MT/MST itself may be biased in favor of a non-preferred direction because of randomness inherent in the random dot motion stimulus. (b) Even though MT and MST fire the most for the signal direction, the moment-by-moment noise fluctuations that drive the evolving LIP responses may eventually result in a non-preferred LIP cell being chosen. Low choice probabilities of MT and MST are typically determined by (b). The possibility of both (a) and (b) decreases with increase in percent coherence. Hence performance in both the tasks improves with coherence (Fig. 6a and b).

In the FD tasks, the decision is made at the end of motion viewing. The monkeys are trained to remember the choice during a variable delay period before the choice saccade can commence. During this period, the selected LIP cell response is sustained, despite the loss of motion excitation, and also builds up slowly. This is another

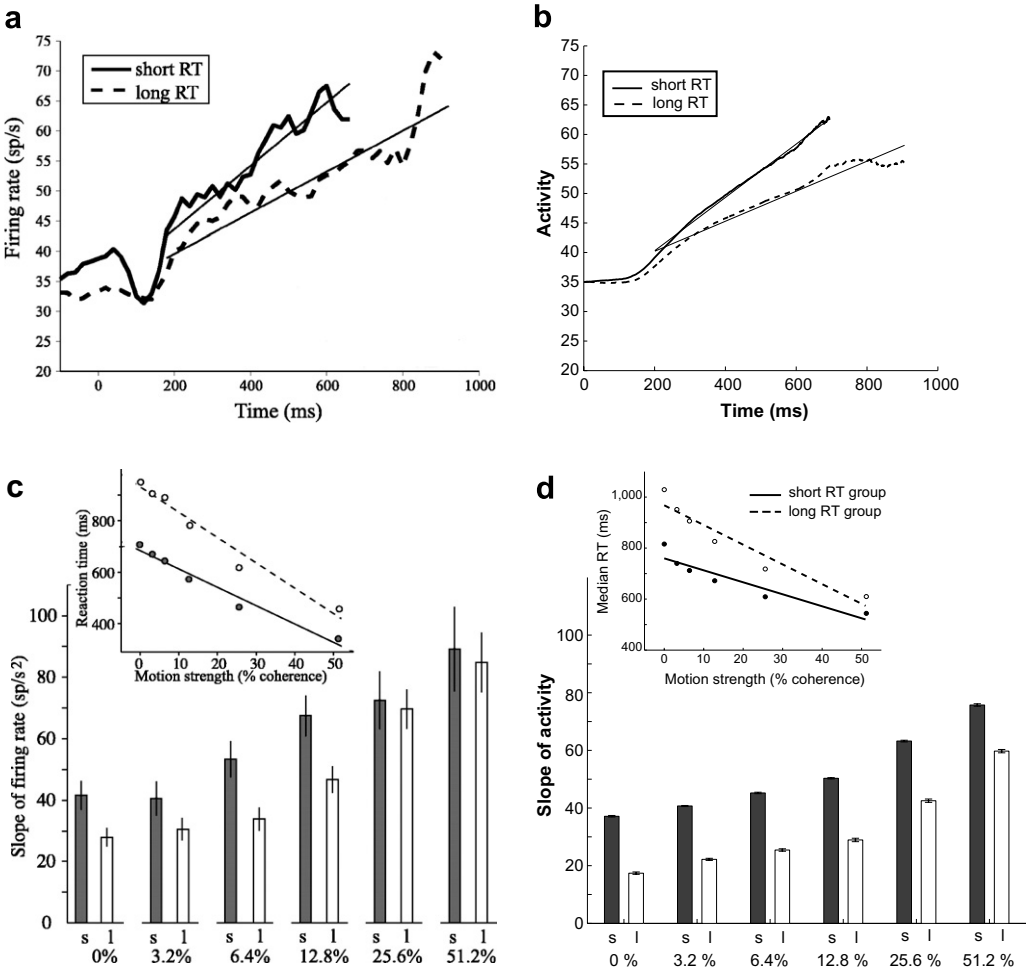

Fig. 10. Relationship between LIP response and reaction time. At each non-zero coherence, correct trials for the right target (T_{in}) were sorted into two groups: short RT and long RT, based on whether the RT is smaller or larger than the median RT. For 0% coherence, trials resulting in T_{in} choices were considered. (a) Average LIP responses for both groups at 6.4% coherence, which are linear fit from 200 ms after motion onset to the median RT of the group, while excluding any activity 100 ms backwards from saccade initiation. (b) Model simulation reflects this relationship. The solid curve corresponds to short RT group, and the dashed curve to long RT group. (c) The histogram shows for each coherence level and group the slope of the linear fit to the average LIP response starting at 200 ms from motion onset, ending at the median RT of the group, and excluding any spikes within 100 ms before the eyes move. The inset shows at each coherence the median RT for both groups formed from trials of correct T_{in} choices. (d) Model simulations reproduce these data trends. Dark bars correspond to short RT group, and light bars to long RT group. In (c) and (d), the error bars represent 95% confidence intervals (CI). [Data in (a,c) is reprinted with permission from Roitman and Shadlen (2002).]

difference with sensory neurons. In the model, the gain of recurrent excitation of the winning LIP cell switches to a value which is just high enough to not only compensate for the input from motion areas getting shut off, but also to help its activity to grow slowly during the delay period for all motion strengths. The recurrent self-excitatory interactions within the LIP recurrent competitive field thus enable persistent activity for the chosen LIP cell after the motion offset. This gain increase hypothesis is consistent with data showing that an intention to saccade to a particular location results in the deployment of some attention to that location (Bisley & Goldberg, 2003). Similar anticipatory activity is seen in superior colliculus (SC) buildup neurons in delayed saccade tasks (Grossberg, Boardman, et al., 1997; Grossberg, Roberts, Aguilar, & Bullock, 1997; Munoz & Wurtz, 1995a, 1995b). The fact that the LIP responses in the FD task have a slower rate of growth and a tendency to saturate when compared to the RT task (Fig. 5) is explained in the model by using a slightly higher passive decay rate parameter for the FD task. This parameter change may reflect a task-sensitive change in the monkey's LIP responsiveness given the difference in the two experimental conditions (see p. 1930 in Shadlen & Newsome, 2001). The tendency of LIP responses to saturate in the FD task may, in turn, explain the tendency of performance as a function of viewing duration to saturate, especially at higher motion strengths (see Fig. 7).

A "forced choice" is made in the RT task in the manner noted in Section 2; namely, if no choice is made within a random time between 1000 and 1100 ms, then a volitional top-down signal is activated at the LIP cell with the most activity (see term G_d in Appendix Eq. (29)). Fig. 12 shows the model's RT distributions with this top-down signal at work. The fraction of simulation trials in which a choice needs to be forced decreases with coherence (see Fig. 12d). Fig. 13 shows that the model also works without this top-down mechanism, but allows for some longer RTs to occur in the 1500–2000 ms range.

Our proposed mechanism underlying "forced choices" is compatible with a different mechanism for the same purpose described in Ditterich (2006b), in which the gain of MT signals into LIP cells monotonically increases with time, and with physiological data recently reported by Churchland, Kiani, and Shadlen (2007) of a stimulus-independent time-varying signal, which increases with time in each LIP population.

The basic idea behind these mechanisms is to trigger a decision despite only a partial resolution among LIP cells as time passes. Ditterich (2006b) argued that this may be important for a monkey

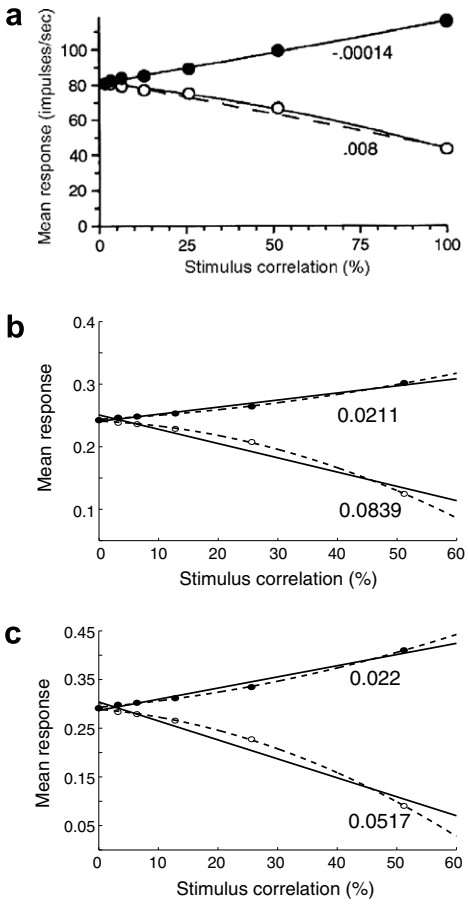

Fig. 11. Average responses in motion-sensitive areas as a function of stimulus correlation. (a) Mean responses of an MT neuron at different coherence levels across several trials for motion in both its preferred and anti-preferred directions (Britten et al., 1993). (b) Average activity in a model MT cell pool across 500 trials at different coherence levels for motion in both its preferred and anti-preferred directions. (c) Average activity in a model MST cell pool across 500 trials at different coherence levels for motion in both its preferred and anti-preferred directions. Thus (b) and (c) demonstrate the analog sensitivity of the model to the direction and strength of random dot motion stimuli. In (a–c), closed and open circles show responses to preferred and anti-preferred random dot motion stimuli, respectively. Solid and dashed lines correspond to least-squares estimated linear and quadratic fits, respectively. The quadratic function describes the data better than the linear function. For each quadratic fit, the ratio of the quadratic coefficient to the linear coefficient is computed, and is shown near the fit. The value of this ratio gives a feel for the degree of non-linearity in the function relating the coherence of the dots to the mean response of the motion-sensitive area. In (b) and (c), the standard errors of mean (SEM) were too small to report in the plots. [Data in (a) is reprinted with permission from Britten et al. (1993).]

who tries to obtain a reward in each trial. On a weak coherence trial, extended motion viewing may be necessary before the direction of the dots can be correctly determined, but this requires maintained fixation in order for the trial not to be aborted. Given that the possibility of breaking fixation increases with time for both monkeys and humans alike, there is a trade-off between continuing to look at the dots to make a correct choice leading to reward, and forcing a choice to prevent loss of reward due to a break in fixation.

As in the data, the simulations also show that the RT task performance is better than that of the 1 s FD task at weaker motion strengths (Fig. 6a and b). This may happen because the monkey has as much access as it needs to sensory information in ambiguous trials in the RT task, unlike in the FD task (Roitman & Shadlen, 2002). Simulated reaction time (RT) histograms in fact reveal that a greater number of correct trial reaction times are in excess of 1 s at lower coherence levels (see Fig. 12a and c). It is noteworthy that the matches between data and model of RT distributions (Fig. 12), the differences between correct and error trial LIP responses (Figs. 8 and 9), RT properties (Fig. 6), and the relationship between LIP response and reaction time (Fig. 10) are emergent properties of the model that were not specifically sought in the model design and parameter choices.

We have already seen how noise influences LIP model cells to make erroneous choices. This happens because the variance of a Brownian motion process increases with time as individual white noise samples are integrated by cell dynamics. The system hereby becomes more noisy as time elapses, thereby increasing the likelihood of errors with time. This is why the simulated average error trial RTs are longer than those of correct trials (Fig. 6). Also, observations of error trial LIP cell responses in both tasks (Figs. 8 and 9) indicates the rate of change in activity is lower on error trials when compared to correct trials, which correlates well with longer RTs, as the analysis in Fig. 10 shows.

Fig. 14 provides some samples of the entire time course of simulated LIP responses during both correct and error trials in the FD and RT tasks. Note in particular the distinct visual and saccadic components present in the LIP response dynamics, and also the sustained component in the variable delay period of FD tasks.

The LIP recurrent competitive field (RCF) helps to solve the noise-saturation problem because its cells self-normalize the network's total activity, even if their inputs vary greatly in size, without losing their sensitivity to the relative sizes of each input. In vision, this property is often called *contrast normalization*. This self-normalization property plays a role akin to computing real-time probabilities. Thus, an RCF shows how distributed pattern processing by networks of cells, or cell populations, can behave like a real-time probability calculus. When the recurrent interactions between the cells are controlled by faster-than-linear signal functions (e.g., power laws), or sigmoid signal functions, they can cause contrast enhancement and noise suppression (Grossberg, 1973, 1988) of their input pattern, including the limiting case of selecting a single winning cell (winner-take-all). RCFs, and simplifications thereof, have been used to model many different types of decision-making over the years since their discovery, including perceptual and cognitive decision-making (Chey et al., 1997; Francis et al., 1994; Usher & McClelland, 2001) and reach decisions in dorsal premotor cortex (Model: Cisek, 2006; Data: Cisek & Kalaska, 2005).

Self-normalizing competition, present in other stages too (although in different forms), helps to amplify coherent motion signals while suppressing incoherent motion signals. For example, lower dot coherence renders population activities of neurons in MT and MST that are tuned to opponent directions more indistinguishable, thus generating a distributed neural representation of high input motion ambiguity. Strengthening one directional pool either by bottom-up input or by microstimulation (Ditterich, Mazurek, & Shadlen, 2003) results in weakening the pools coding the other directions, the null direction in particular (see Fig. 11). When motion coherence is high, there is less ambiguity, so fast feedforward processing (cf., Thorpe, Fize, & Marlot, 1996) can occur in the MT/MST circuit. Low coherence slows down processing because activities of all conflicting groupings are reduced by self-normalizing competition, and it takes longer for any one of them to exceed output thresholds. These normalized spatial activity distributions hereby behave like real-time probabilities that reflect system uncertainty. Feedback is automatically engaged from MST to MT that helps to choose among these ambiguous alternatives, and to contrast enhance the strongest grouping. At all model V1, MT, MST, and LIP stages, selected cells respond to more coherence in the input by increasing their activities more rapidly to higher levels of firing, thereby generating faster and more accurate decisions as described earlier in this section. This self-organizing system thus trades accuracy against speed, and illustrates how cortical dynam-

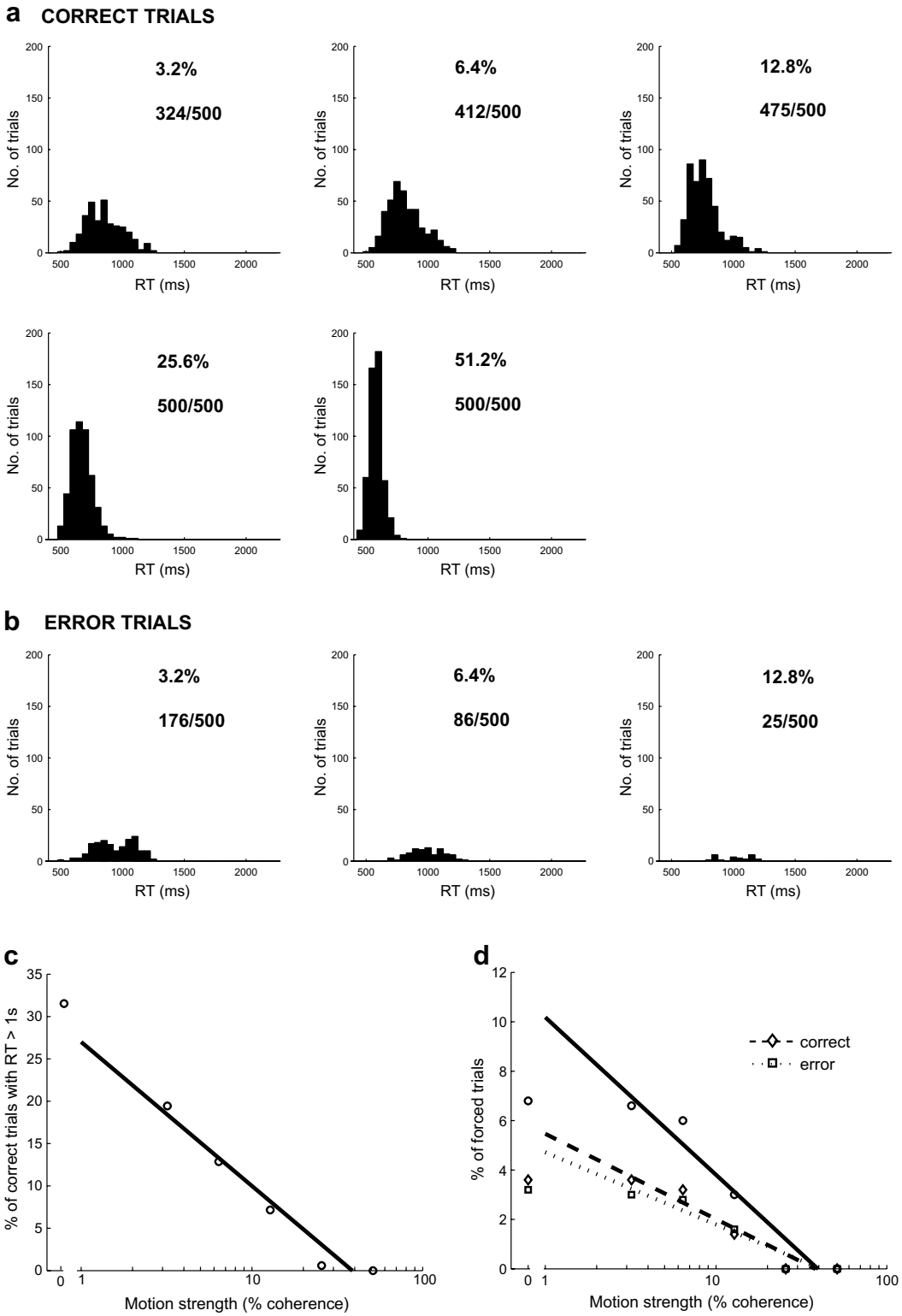

Fig. 12. Reaction time histograms from simulations of the RT task. In (a), RT statistics from the correct trials are shown. We can observe that as the coherence increases, the RT histogram becomes narrower and shifts to the left. Also, as the coherence decreases, quite a number of trials have RTs in excess of 1000 ms. This may explain why RT task accuracy is slightly better than that of 1s FD task at weaker coherence levels. In (b), RT histograms from the error trials are shown. The model monkey does not make any errors at stronger coherences: 25.6% and 51.2%. Note the fraction of trials in each panel for the corresponding condition. The bin-width used in generating the histograms is 50 ms. The RT distribution data from Roitman and Shadlen (2002) experiments is available in Ditterich (2006a). Panel (c) plots the proportion of correct trials with reaction time greater than 1000 ms in the simulated RT task at each motion strength. (d) Forced choice task design requires the model to make a choice in each RT task trial however ambiguous the stimulus is. Sometimes, at low coherences in particular, increased viewing duration does not sufficiently clarify the direction of the dots. In such cases, the model is forced to make a choice via term G_d in Appendix Eq. (29). In this plot, open circles indicate the proportion of trials (out of 500), when a forced choice was invoked, as a function of motion strength. This fraction falls off linearly with respect to $\log_{10}$ scale of percent coherence. Open diamonds and squares refer to correct and error decisions, respectively. The forced choices turn out to be correct at slightly better than chance, implying they are not "pure" guesses.

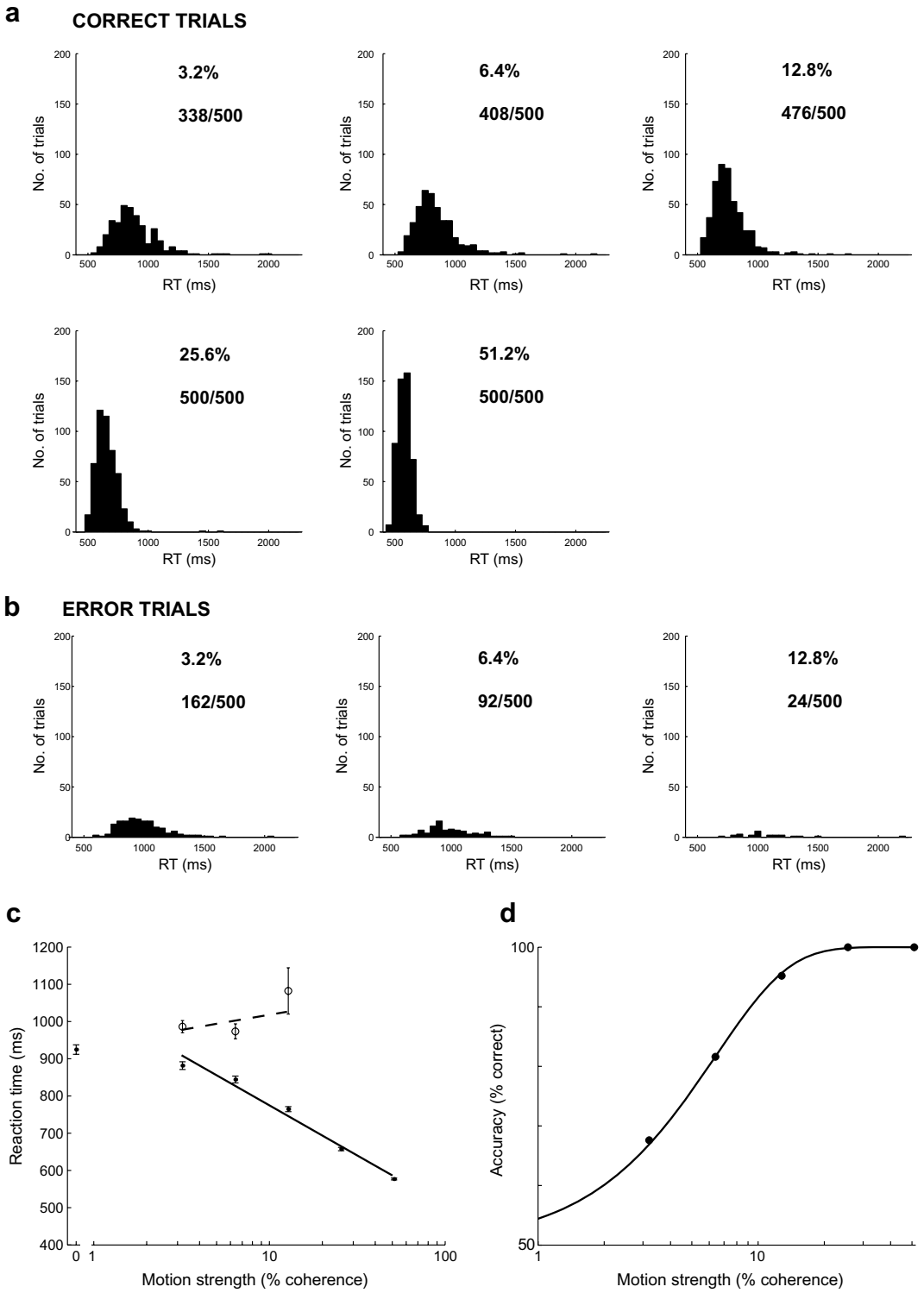

Fig. 13. Simulations of the RT task without the mechanism to force a decision within about 1 s; that is, with term G_d in Appendix Eq. (29) set equal to zero through the trial. (a) Correct trial RT distributions for each non-zero coherence level. Given there is no restriction on the amount of time to make a response, a few RTs fall in the 1500–2000 ms range, unlike in the earlier simulations shown in Fig. 12. (b) Error trial RT distributions for the three lower non-zero coherences. The model did not make errors at the two higher motion strengths. (c) Correct (solid) and error (dashed) trial RT statistics. (d) Psychometric function. Other results for the RT task remained intact (not shown here).

ics go beyond Bayesian concepts, while clarifying why probability theory ideas are initially so appealing (see Section 4.2).

Experiments in which prior probability is manipulated (Mazurek et al., 2005) can be explained by top-down priming from prefrontal cortex (PFC), which can leaky integrate the number of occurrences of a particular choice, and feed it back to LIP as a modulatory attentional prime (Grossberg, 2003; Miller, Erickson, & Desimone, 1996; Srihasam et al., 2008), both before and after motion onset. This hypothesis predicts that prior probabilities will play a lesser role if PFC working memory storage is reset or otherwise weakened between consecutive trials in a block. The effects of MT microstimulation (Ditterich et al., 2003) are similar to those of

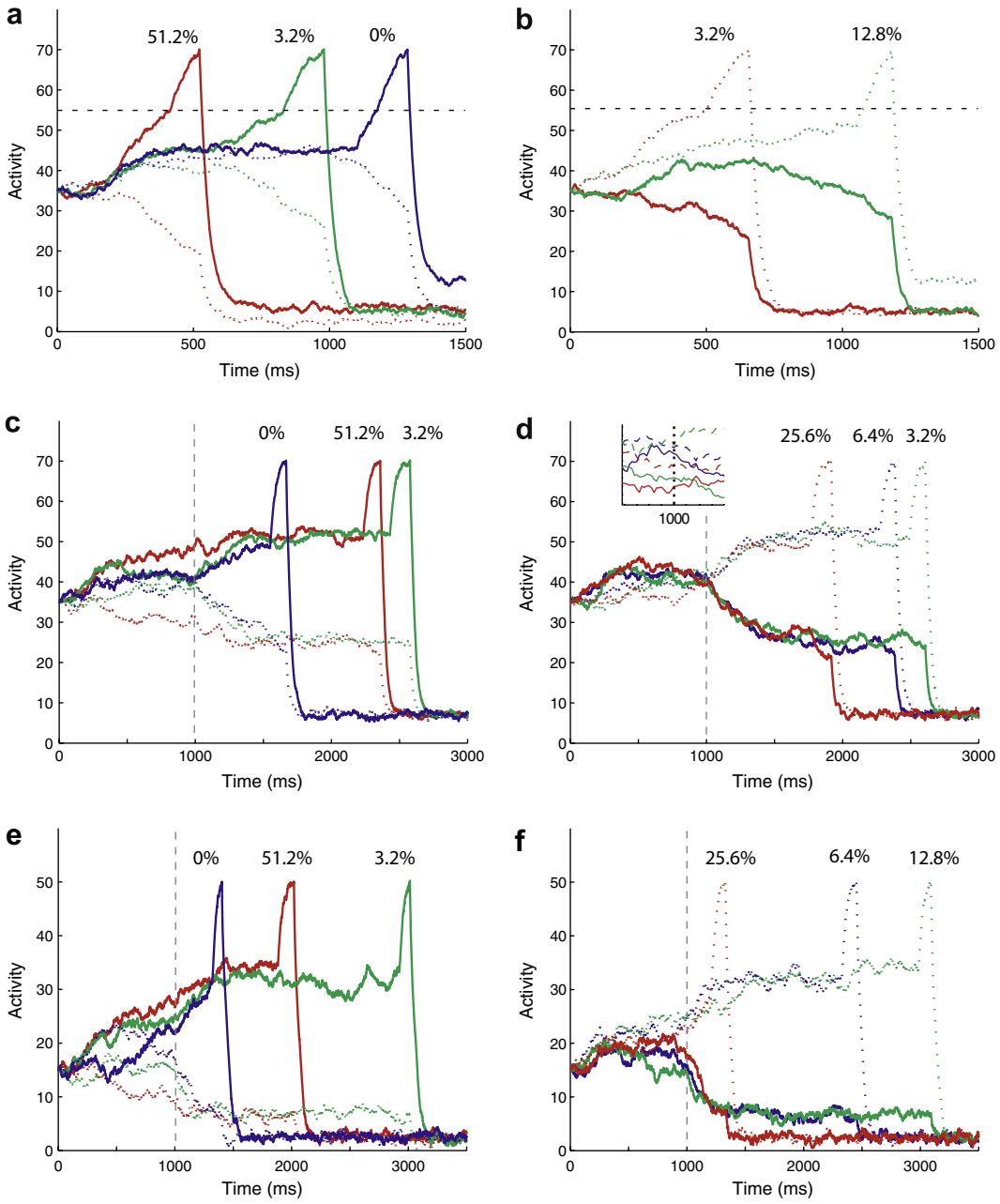

Fig. 14. Simulated LIP responses during some representative correct and error trials in the RT and the FD tasks. (a) Entire time course of the two LIP cells coding the preferred direction (solid) and the null direction (dashed), respectively, during some example correct trials of the RT task at three different coherence levels: 51.2%, 3.2%, and 0%. The choice in the 0% coherence case shown here is forced because even after ~1 s, neither competing cell has approached the threshold. (b) Temporal dynamics of the two LIP cells during some illustrative error trials at two coherence levels: 3.2% and 12.8%. (c) and (d), and (e) and (f), are for the 2002 and 2001 FD tasks, respectively. Since the fixation point can be turned off at variable times after motion offset, the time of the actual saccade does not depend on motion strength in the FD tasks. The inset in (d) shows how erroneous decisions are made at the end of motion viewing (1 s).

prior probability on the accuracy and reaction time of perceptual decisions. They can be modeled by adding a steady excitatory input to each cell in a subset of model MT for the duration that the stimulation was applied. Within the model, microstimulating a particular MT or MST directional pool clearly has the same effect as increasing coherent motion in the corresponding direction (Celebrini & Newsome, 1995; Salzman, Murasugi, Britten, & Newsome, 1992).

The influence of more potential choices on LIP responses, and hence behavior (Churchland et al., 2005), can also be explained by self-normalization in the RCF. More choice targets create more recurrent inhibition on an LIP cell from other LIP cells, which implies a smaller initial activity for the selected LIP response. Increasing the number of alternatives is thus similar to decreasing the motion strength in a two-alternative, forced choice (2AFC) task. As a result, more choices, on average, predict slower reaction times and more errors in decision-making.

Additional considerations are needed to understand the first 200 ms of the LIP cell responses after motion onset. The coherence-insensitive initial dip and rise in activity (Fig. 5a) is explained as follows: before motion onset, attention is divided among the two peripheral choice targets (covert) and the foveal fixation point (overt). The abrupt appearance of the motion dots in the foveal region may automatically attract all attention (Egeth & Yantis, 1997),

thereby rapidly creating a strong foveal transient signal (Bisley, Krishna, & Goldberg, 2004) that inhibits both peripheral LIP populations. The foveal transient is independent of the strength of the moving dots. The coherence-independent rise in activity results from the offset of the initial burst, and the time when LIP activities start fanning-out is when the discriminative motion signals start reaching the LIP choice network. Such lateral inhibition is also consistent with recent experimental data and modeling of transient pauses in SC neuronal firing induced in part by shape-changing cues occurring in a region well outside the movement field (Li, Kim, & Basso, 2006). Unlike in the RT task, the interval between the onset of choice targets and that of dots is fixed in the FD task (see p. 9476 in Roitman & Shadlen, 2002), thus effectively eliminating the 'surprise' element. The lack of transient pause in LIP firing during the FD task (Fig. 5c and e) may be a result of this difference.

4. Discussion

4.1. Noise in perceptual decision-making

Cellular noise in LIP influences temporal dynamics and thus the variability in the accuracy and reaction time of perceptual decisions that LIP controls. Model simulations to this point are deterministic until the MST stage. In this form, the model approximates all internal neuronal noise from Retina to LIP by noise that drives just the LIP output stage. Trial-to-trial variability in the model until MST is elicited only by the randomness inherent in the random dot motion stimuli. However, both MT and MST have been shown to exhibit response variance even if the same sequence of random dots is input again and again (Britten et al., 1993; Celebrini & Newsome, 1994).

The above simulations support the hypothesis that noise in LIP, plus the randomness in the stimulus itself, which generates random cell responses throughout the motion processing stages, is sufficient to quantitatively simulate key data properties. In order to test this directly, we also ran our simulations introducing similar noise processes in both MT and MST (see Appendix Eqs. (23) and (28)), in addition to the noise in LIP. We tried two different standard deviations (σ_0) for the internal noise in MT and MST: one comparable to that used in LIP, and the other five times greater. In both cases, all our results remained qualitatively the same (see illustrative Fig. 15). For the higher MT and MST noise condition, the psychometric function seems to shift slightly to the right, as one would expect by virtue of more overall noise in the system (threshold ratio $\alpha_{\mathrm{RT},\sigma_0=0.1}/\alpha_{\mathrm{RT},\sigma_0=0.5} = 0.893 < 1$; see Appendix A.3.7 for how threshold α is defined). Our data explanations still remain the same. These simulations show that the absence of cellular noise in the motion processing system is not a rate-limiting factor as long as there is cellular noise in the model's LIP output stage.

4.2. Bayesian inference in the brain?

The probabilistic nature of decision-making has led to the proposal that classical statistical concepts, such as the Bayes rule, may apply to decision-making in the brain (Gold & Shadlen, 2001, 2007; Knill & Pouget, 2004; Pouget, Dayan, & Zemel, 2003). It is typically argued that variability in neuronal responses and the fact that prior knowledge or information influences what we see and how we behave implies explicit Bayesian inference in the brain. Bayesian and other general statistical concepts formalize aspects of the *unconscious inference* hypothesis of Helmholtz (2000); namely, that visual system works by making inferences about the world based upon probable interpretations of incomplete or ambiguous data. Kanizsa (1979) and his disciples have, in contrast, described many visual percepts that cannot be explained by unconscious inference. The Helmholtz vs. Kanizsa perspectives exemplify top-down vs. bottom-up approaches to visual perception, as two extremes in a continuum. The neural model proposed herein characterizes and joins together both bottom-up and top-down processes to explain challenging behavioral and brain data.

The Bayes rule, and related statistical concepts, are so general that they may be applied to problems in any science. In particular, the Bayes rule follows from writing the probability of any two events I and S, namely $p(I,S)$, in two different ways and then dividing both sides of this identity by a term on one side of the equation; see Eqs. (1) and (2) below. This generality is part of its broad appeal, but is also its weakness in not providing enough constraints to discover new models of any particular science. In particular, it does not in itself help to discover the brain design principles and mechanisms that can effectively process ambiguous motion stimuli. The present article proposes how the brain may make perceptual decisions in response to motion stimuli by using organizational principles and detailed neuronal mechanisms that go beyond a general Bayesian formulation, and clarifies why a Bayesian approach may initially seem so appealing.

In the Bayesian framework, given a stimulus I which falls on the Retina, the Bayesian model, or ideal observer, asks what percept S is most likely to have caused it. The basic idea is to find the S that maximizes the posterior probability $p(S/I)$ via the Bayes rule:

$$p(S/I) = \frac{p(I/S)p(S)}{p(I)}. \tag{1}$$

To derive Eq. (1), the probability $p(I,S)$ of the events I and S is written in two different ways:

$$p(I,S) = p(I/S)p(S) = p(S/I)p(I), \tag{2}$$

and then divided by $p(I)$. The specific usage depends on which of I and S is observable. I could be a spatial input (static image), or a spatio-temporal input (video). S could be either a distributed pattern of neuronal firing in some low or mid-level brain region (contour, motion, or depth perception, etc.), or a symbolic percept in some high-level brain region (face, house, or object recognition, etc.). Typical psychophysical experiments require subjects to detect or discriminate a perceptual feature and/or estimate its strength. The likelihood $p(I/S)$ assigns a probability to each percept S depending on the amount of consistency with the image I, the prior $p(S)$ is the probability of the scene description S, and $p(I)$ acts as a normalizing factor (Kersten, Mamassian, & Yuille, 2004; Kersten & Yuille, 2003). For example, Bayesian models of speed perception assume a prior that slower speeds are more frequently perceived (Stocker & Simoncelli, 2006; Weiss, Simoncelli, & Adelson, 2002).

Bayesian and other general statistical approaches do not address how the *units* of stimulus I and percept S are generated by the brain, such as illusory contours, which are emergent properties not explicit in the luminance levels of an inducing image. Bayesian models also do not explain percepts of environments whose statistics change through time, where there is no obvious likelihood $p(I/S)$, or novel rare percepts, where there is no obvious prior $p(S)$. Most importantly, it does not explain the brain processes whereby we perceive.

How might Eq. (1) be applied to motion direction discrimination? Gold and Shadlen (2001) propose that the brain may accumulate the logarithm of the posterior ratio (log PR) in a sensory-motor association area such as cortical area LIP. In this situation, the brain has to discriminate between opponent motion directions, say S_1 and S_2, given the spatio-temporal input $\vec{I}$. The posterior ratio is defined as:

$$\mathrm{PR} = \frac{p(\vec{I}/S_1)p(S_1)}{p(\vec{I}/S_2)p(S_2)}. \tag{3}$$

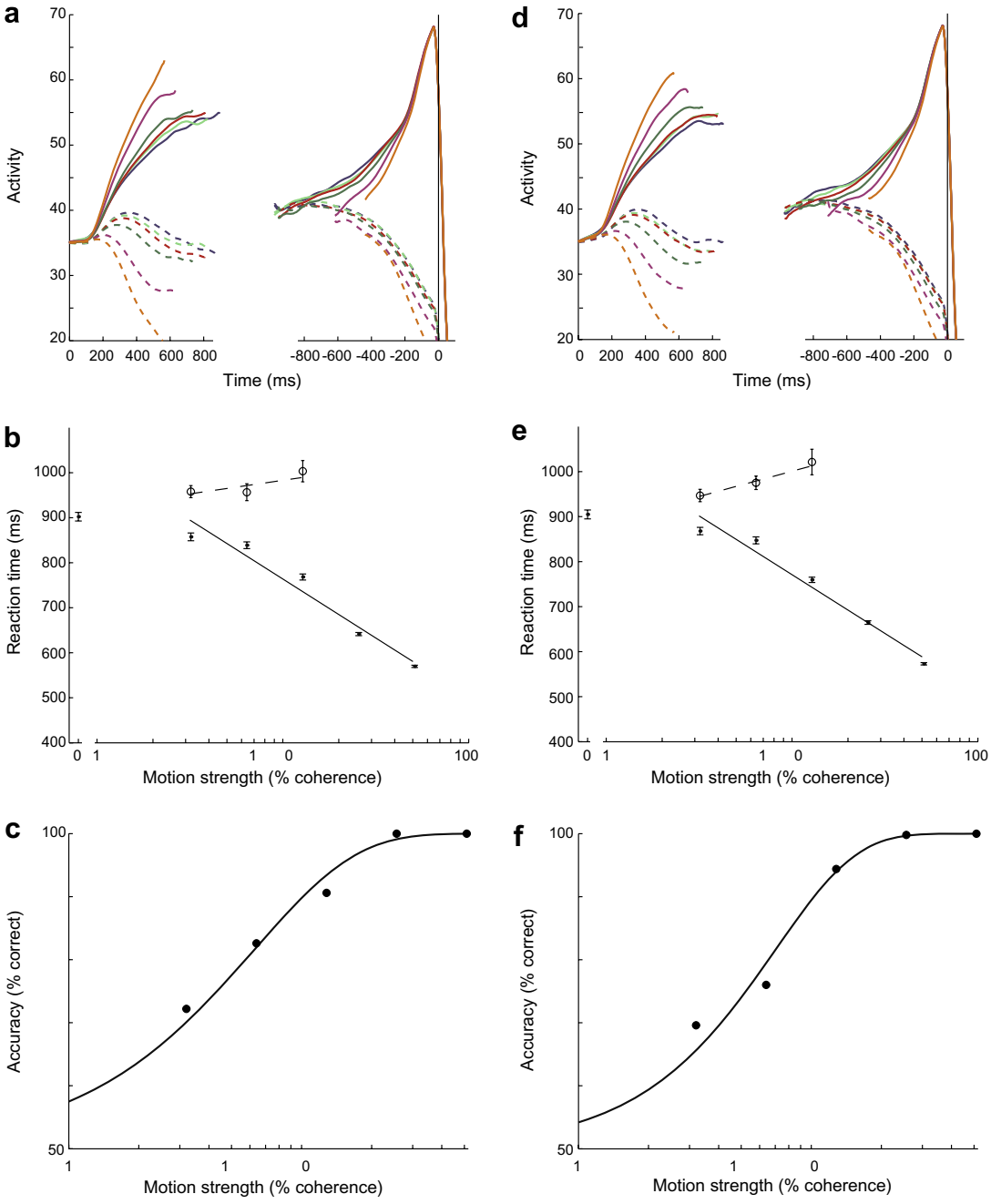

Fig. 15. Simulations of the RT task with noise in both MT and MST. See Appendix Eqs. (23) and (28). The level of noise was about the same as that used in LIP (see Appendix). (a) Correct trial LIP cell dynamics for all motion strengths in either motion direction. Same conventions as in Fig. 5a are used. (b) Correct (solid) and error (dashed) trial RTs at various coherences. (c) Psychometric function. Noise in MT and MST also did not affect other RT task results (not shown here). (d–f) Simulations of the RT task with more noise in MT and MST. The level of noise was five times greater than that used in the simulations shown in (a)–(c).

Taking a logarithm on both sides yields a difference of two logs:

$$\log \mathrm{PR} = \log p(\vec{I}\,/S_1)p(S_1) - \log p(\vec{I}\,/S_2)p(S_2). \quad (4)$$

When both directions are equally probable, $p(S_1) = p(S_2)$, log PR reduces to log-likelihood ratio (log LR):

$$\log \mathrm{LR} = \log p(\vec{I}\,/S_1) - \log p(\vec{I}\,/S_2). \quad (5)$$

Gold and Shadlen (2001, 2007) suggest that the difference in activities of two populations in middle temporal area (MT) that are tuned to opponent directions approximates a scaled version of log LR. Motion opponency is known to occur in MT (Albright, Desimone, & Gross, 1984; Heeger, Boynton, Demb, Seidemann, & Newsome, 1999; Snowden, Treue, Erickson, & Andersen, 1991), and plays a role in our model as well (Figs. 1 and 4), but its proposed theoretical equivalence with log LR does not add to our understanding of the neuronal dynamics during perceptual decision-making in response to a wide variety of perceptual stimuli. In particular, it is not clear how the concept of log LR extends to multiple-choice tasks, and any log approximation to neuronal activations and firing is biologically implausible because the log grows without bound, unlike cellular activations and firing, as its argument increases.

Given these concerns, it is natural to ask if such general statistical formulations provide the best currently available approach to

modeling perception and decision-making? As illustrated herein and in a growing number of other studies, biophysically detailed models of visual cortex can quantitatively simulate neurophysiologically recorded dynamics of identified cells in anatomically supported circuits, and the perceptual and behavioral data that they generate as emergent properties. Moreover, these models embody new design principles, mechanisms, local circuits, and global architectures that are proposed to solve fundamental problems facing the brain. The present model shows how brain circuits may carry out hypothesis-testing and decision-making in response to both statistically changing and novel rare events.

It should also be noted that any filtering operation, such as the short-range and long-range filters (see Appendix Eqs. (15) and (22)), may be interpreted as a prior (namely, the current neural signal) multiplied by a conditional probability or likelihood (namely, the filter connection strength to the target neurons). Likewise, a contrast-enhancing operation, such as the LIP recurrent on-center off-surround network (see Appendix Eq. (29)), that responds to such a filter may be viewed as maximizing the posterior. These insights have been discussed in the neural modeling literature for a long time; e.g., Grossberg (1978). However, as Fig. 1 and the total set of model equations illustrate, such local processes do not, in themselves, embody the design principles and computational intelligence of a behaviorally competent neural system.

In order to contribute to the characterization of how the brain carries out motion processing and decision-making, the present work further develops and integrates models that have been gradually elaborated to explain and simulate increasingly large data bases about motion-based percepts and behaviors through the years. This emerging motion model, which is often called the 3D FORMOTION model, has simulated a large number of motion percepts (see Fig. 16 for two examples). These percepts include aspects of barberpole illusion; motion capture; spotted barberpole illusion; triple barberpole illusion; occluded translating square illusion; motion transparency; chopsticks illusion; short-range and long-range apparent motion, including beta, gamma, reverse-contrast gamma, delta, split, ternus, and reverse-contrast ternus motion; visual inertia; Korté's laws; second-order motion; form-motion interactions, including the line motion illusion, motion induction, and transformational apparent motion; how component angles, contrasts, and durations influence barberpole and (type 1 and type 2) plaid coherence and incoherence; the effects on perceived speed of stimulus contrast, duration, dot density, orientation, length, and spatial frequency; and the effects of parvocellular and magnocellular lateral geniculate nucleus lesions on motion perception (Baloch & Grossberg, 1997; Baloch, Grossberg, Mingolla, & Nogueira, 1999; Berzhanskaya et al., 2007; Chey et al., 1997, Chey, Grossberg, & Mingolla, 1998; Francis & Grossberg, 1996b; Grossberg et al., 2001; Grossberg & Rudd, 1989, 1992). The parameters in this model are chosen within a robust parameter range to enable all of the processing stages to contribute their share to the disambiguation of locally ambiguous motion direction.

The fact that the current explanations and simulations of the temporal dynamics of decision-making in response to probabilistically defined displays of moving dots may be added to this compendium of results, without explicitly involving Bayesian concepts, provides additional evidence that specific brain designs to solve the aperture problem and the noise-saturation problem are indeed at work.

4.3. Stochastic decision models

Various sorts of stochastic reaction time models have long been used in cognitive psychology to account for both psychometric and chronometric functions of human subjects in a number of perceptual and cognitive tasks (e.g., Busemeyer & Townsend, 1993; Luce, 1986; Ratcliff, 1978, 1980; Ratcliff, van Zandt, & McKoon, 1999; Ratcliff & Smith, 2004; Smith, 1990, 2000; Smith & Vickers, 1988). Most of these models are variants of a diffusion model, or a random-walk model in discrete-time. In the race or accumulator model for multiple-choice RT tasks, an accumulator cell y_d for each possible alternative d gradually integrates its corresponding noisy evidence until one of the cell activities first rises to a decision threshold, thereby deciding both the choice and the decision time. The reaction time is computed by adding to the decision time an independent, random non-decision time to account for sensory and motor latencies. All diffusion models make assumptions regarding how perceptual stimuli map into internal sensory representations, which in turn are converted into decision variables (y_d) as follows:

$$\frac{dy_d}{dt} = S_d + \sigma \frac{dW}{dt}, \tag{6}$$

where S_d is called the drift rate of the diffusion process for alternative d, which is an increasing function of its internal signal, and the stimulus-independent parameter σ scales the standard deviation of the Brownian motion process W. The noise term $\sigma \frac{dW}{dt}$ controls choice and decision time variability. In an FD task, the choice is typically governed by the decision cell with the greater activity at the end of the fixed duration (Ratcliff, 1978).

This simple model can fit the psychometric function, RTs on correct trials, and effects of speed–accuracy trade-off by adjusting the decision threshold based on priority instructions (Reddi & Carpenter, 2000), but needs extensions to tackle relative error trial RTs (Palmer et al., 2005). Some of the variations include dynamic drift rate (Ratcliff, 1980), trial-to-trial variability in drift rate, which can simulate longer error trial RTs (Ratcliff & Rouder, 1998), and trial-to-trial variability in the starting point of integration, which can simulate faster error trial RTs in the urgency condition (Ratcliff et al., 1999). Other variants include leaky integration (Smith, 2000; Usher & McClelland, 2001), differences between alternative internal signals in the sensory read-out stage (Ditterich, 2006a, 2006b; Mazurek et al., 2003), and lateral inhibition among accumulator cells (Usher & McClelland, 2001).

Neural correlates of decision-making have only recently been investigated in both primates (Glimcher, 2003; Gold & Shadlen, 2007; Platt, 2002; Romo & Salinas, 2001; Schall, 2003) and humans (Heekeren, Marrett, Bandettini, & Ungerleider, 2004). Accordingly, general diffusion models have begun to model both behavioral and neurophysiological data in, for example, LIP (Mazurek et al., 2003) and SC (Ratcliff, Cherian, & Segraves, 2003). Studies such as Ratcliff et al. (2003) provide good fits to neurophysiological data from buildup cells in the deeper layers of the SC during a two-choice decision task. Although such a diffusion model is more dynamical than a general statistical model, it is a formal model that does not embody the brain mechanisms that characterize neural decision processes.

Neural models currently exist that quantitatively simulate the temporal dynamics of both burst and buildup cells in the SC (Gancarz & Grossberg, 1999; Grossberg, Roberts, et al., 1997). Unlike the diffusion model, the SC neural model clarifies the different functional roles of both cell types, notably how they help to adaptively recode the several types of SC inputs (retinotopic visual inputs, head-centered auditory inputs, and planned prefrontal inputs) into the same coordinate representation wherein all the inputs can competitively choose the next saccadic decision. The current MODE model attempts to provide a similar functional and mechanistic neural understanding of motion-based decision-making.

Four other models (Ditterich, 2006a, 2006b; Mazurek et al., 2003; Wang, 2002) have discussed the Shadlen and Newsome data. We now highlight similarities and differences between our model and these alternative models.

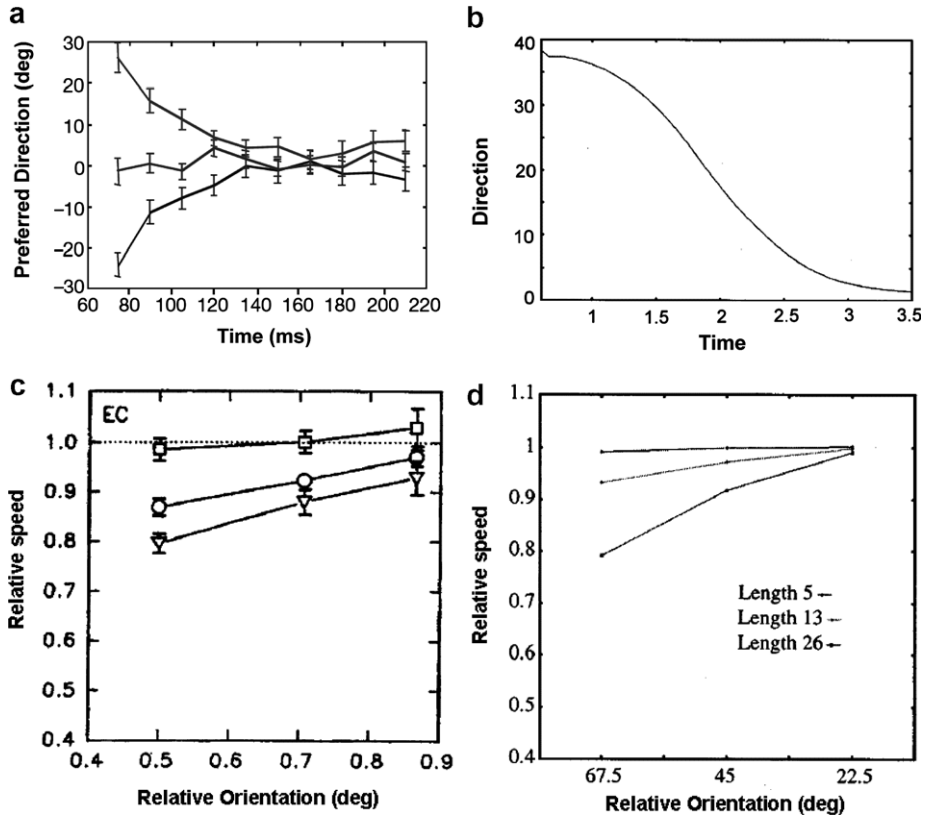

Fig. 16. The 3D FORMOTION model (Berzhanskaya et al., 2007; Chey et al., 1997, 1998; Francis & Grossberg, 1996a, 1996b; Grossberg et al., 2001) helps to explain many temporally evolving motion percepts in addition to those derived from random dot stimuli. (a) Data (adapted from Fig. 2c in Pack and Born, 2001) show that MT gradually solves the aperture problem with time. When monkeys are shown stimuli which comprised a field of white bars moving coherently in a fixed direction on a black background, MT cells are tuned initially to the direction perpendicular to the orientation of the bars, but eventually reflect the true motion direction. (b) Simulations from Chey et al. (1997) predicted these observations. Here the stimulus is a horizontally moving line tilted at 45° from vertical. (d) Simulations from Chey et al. (1997) also account for data from Castet et al. (1993) shown in (c). Here a line moves in a fixed direction at a fixed speed for a fixed duration (167 ms). The longer the line, and the more its orientation is tilted from the orientation perpendicular to movement direction, the slower the line seems to be moving. These data can be understood as the effect of feature tracking signals at the line ends propagating into the interior of the line, and gradually selecting the consistent feature motion direction there, while being averaged with the directions and speeds of ambiguous motion signals along the length of the line. The feature tracking signals gradually select the consistent motion direction and speed through time, much as they do in (a). [Data in (a) is reprinted with permission from Pack and Born (2001), and panels (b,c,d) from Chey et al. (1997).]

In Ditterich (2006a, 2006b), several variations of a diffusion model are presented that attempt to fit the psychometric function, the correct trial RTs, the relatively longer error trial RTs, the roughly symmetric RT distributions, and some aspects of the average LIP responses on correct trials in the RT task. In addition to stationary versions of the diffusion model, different kinds of non-stationarities were also tested: trial-to-trial variability in the drift rate, time-variant decision threshold, time-variant additive signal, and time-variant sensory gain. Based on the overall quality of the best possible fits for each variation (see Table 1 in Ditterich, 2006a), it was proposed that a diffusion model with time-variant sensory gain is most likely the underlying mechanism in LIP that makes decisions in the dots task. Model structure has a difference read-out from two opponent MT pools to two LIP cells, with uncorrelated Gaussian noise in each MT directional pool.

This model describes mean MT responses (Britten et al., 1993; see Fig. 11) by linear functions of the coherence and direction of random dot motion. It also assumes a logistic function for the time-variant sensory gain to explain slower error RTs without altering the signal-to-noise ratio, and a normally distributed non-decision time (see Fig. 6a in Ditterich, 2006a). To fit the more constraining LIP neural data, Ditterich further assumes that the two sensory difference signals that feed LIP decision cells are not perfectly anti-correlated, that an LIP choice cell stops integrating after reaching the decision threshold while the other LIP cells continue accumulating, and also that the activities of the LIP cells do not increase beyond a limited span by clipping them.

Analysis of the estimated residual times from the various considered models led to the conclusion that integration in LIP could be quite leaky (p. 998 in Ditterich, 2006a). An interesting point of this model, as noted in Section 3.2, regards how the time-variant sensory gain hypothesis helps the monkey to maximize its reward rate (Ditterich, 2006b). Section 3.2 details a different mechanism in our model that realizes the same property.

Mazurek et al. (2003) presented a stationary diffusion-type model (also considered in Ditterich, 2006a, 2006b) with most of above assumptions, and applied it to both the FD and RT tasks. In order to account for data from the FD task, they proposed that the LIP cells freeze their activities when one of them reaches the decision threshold, and that there is a delay in the detection of a threshold-crossing before the activities can freeze (p. 1263). Even with these assumptions, model simulations do not show the observed shift in the psychometric functions of 1 s FD and RT tasks (their Fig. 6B). Their model also exhibits faster error trial RTs, contrary to data (their Fig. 4B).

Wang (2002) and Wong and Wang (2006) developed an attractor neural network model that was used to simulate the current dataset. It has properties of decay, recurrent self-excitation and inhibition like our model LIP. Assumptions are made similar to those of Ditterich (2006a, 2006b) and Mazurek et al. (2003) to ob-

tain MT responses. Unlike Mazurek et al. (2003), Wang generates persistent activity in LIP cells in the FD task using strong recurrent excitation within the same network that integrates noisy evidence in the initial period of the trial. As noted previously, our model LIP RCF employs self-excitation too, but also in the RT task to explain the steep, coherence-independent motor preparatory activity, and for quantitative simulations of LIP cell dynamics in the FD task through the entire trial, the delay period in particular (compare our Fig. 5d and f to Fig. 9A in Wong & Wang, 2006).

Both Mazurek et al. (2003) and Wang (2002) employ the same decision rule in both FD and RT tasks, namely the same decision threshold-crossing. In Wang's FD task simulations, if none of the attractor cells wins (that is, reaches the threshold) within the given duration, a random choice is made. It is not clear in Mazurek et al. (2003) about what happens if no threshold-crossing occurs within the fixed duration.

As we have seen in the RT task, the average LIP responses for T_{in} choices, when aligned to saccade onset and grouped by coherence level, reveal a firing level, or decision threshold, above which coherence does not differentiate the dynamics (see Fig. 5a). Even if it is assumed that FD task choices are also based on a threshold-cross, as in the RT talk, the disparate response ranges during the decision formation epoch in the two tasks (compare Fig. 5a (RT task) and 5c (FD task): data from the same monkeys) at least rule out a common threshold value. As a result, the Mazurek et al. (2003) model does not match the recorded LIP response range during motion viewing in the FD task (their Fig. 7C), despite its specialized hypotheses for the FD task. Wang (2002) avoided this issue by emphasizing that model simulations are only qualitative. In contrast, our model makes FD task choices based on the maximally activated LIP cell at the end of the viewing duration (as proposed in Gold & Shadlen, 2003, p. 635), and also simulates the task-dependent changes in LIP responses.

Our MODE model shares with these alternative models the hypotheses that decisions are made in LIP by accumulating noisy motion signals from MT/MST, and that noise influences the erroneous choices made by LIP. Our model does not, however, need to make many of the special hypotheses of the other models because it explicitly represents the spatially distributed neural dynamics of motion and decision processing in response to the stimuli used in the experiments.

Leakage and lateral inhibition are modeled in the *leaky competing accumulator (LCA) model* (Usher & McClelland, 2001) that simulates human behavioral data from other perceptual tasks in both FD and RT paradigms. Its simulations show how leakage alleviates the need for trial-to-trial variability in drift rate (Ratcliff, 1978; Ratcliff & Rouder, 1998) to explain how improvement in performance due to more exposure duration stops at longer durations. Usher and McClelland use a *Grossberg Additive Model* with threshold-linear competitive signals that have been proved capable of making decisions (cf., Grossberg, 1978, 1988). The model also incorporates a noise term to simulate stochastic decision data:

$$\frac{dy_d}{dt} = -AY_d + \lambda S_d - g\sum_{D\neq d} Y_D + \sigma\frac{dW}{dt}, \text{ and } Y_d = [y_d]^+. \quad (7)$$

Usher and McClelland (2001, p. 553) acknowledged these antecedents to their use of the Additive Model, and also (p. 555 and 557) that they do not incorporate recurrent excitation in Eq. (7), which plays an important role of selecting winners in response to inputs that can vary widely in size (Grossberg, 1973, 1978). Our model LIP (Appendix Eqs. (29) and (34)) uses a recurrent competitive field, or RCF, also called a *Grossberg Shunting Model*, that has properties of automatic gain control, self-normalization, and self-maintenance of activity between maximum and minimum levels, unlike the Additive Model (Grossberg, 1973, 1978, 1988). As noted above, the Shunting Model embodies these properties in a network with recurrent excitation and inhibition. They enable our LIP RCF to respond without a loss of sensitivity to spatially distributed inputs from a brain region like MST whose cell activations can vary widely in number and size through time, and can thereby control effective choice behavior.

The biggest departure of our model from previous ones is that it describes the spatially distributed dynamics of multiple brain areas (Retina/LGN, V1, MT, MST) in the motion pathway leading to LIP, and interactions of the basal ganglia with LIP to make decisions. In particular, we show how principles and mechanisms that have evolved in the brain to tackle the aperture problem and the noise-saturation problem may also explain data about the dots task. Our model also meets the challenge of directly processing the random dot motion experimental stimuli as real-time inputs at 60 Hz, and its emergent properties fit the widest range of decision-making data in this paradigm of any available model.

Model simulations are consistent with the interpretation of Newsome, Shadlen, and colleagues regarding the different roles of "sensory" MT/MST and "decision" LIP cells in the motion tasks (see Discussion in Shadlen & Newsome, 2001). In particular, model MT/MST cells are involved in motion capture and resolve the veridical direction of the kinetic dots as much as possible, irrespective of the choice made in each trial by LIP cell dynamics. Accordingly, we propose a lower choice probability in LIP if, for example, the monkey were to shows its decision using a hand movement (Fig. 7 in Nichols & Newsome, 2002), not an eye movement, or if the task prevents saccade planning during the motion viewing period (Gold & Shadlen, 2003; Horwitz, Batista, & Newsome, 2004b).

The model can also explain how microstimulation to MT and MST influences the accuracy of perceptual decisions (Celebrini & Newsome, 1995; Salzman et al., 1992), how MT microstimulation influences correct and error trial RTs (Ditterich et al., 2003), and the effects of prior probability (Mazurek et al., 2005) and multiple choices (Churchland et al., 2005). The model also simulates the influence of viewing duration on performance at each motion strength, and thereby the psychometric function (Gold & Shadlen, 2003).

In addition, our model goes beyond alternate models by: incorporating the difference in LIP responsiveness to the two task conditions, in particular that the "gain of the LIP response is greater in the RT version of the task" (p. 9485 in Roitman & Shadlen, 2002); considering the visual contribution to LIP response due to the presence of the choice target in the receptive field (p. 1931 in Shadlen & Newsome, 2001); simulating the entire time course of LIP responses during both FD and RT tasks on both correct and error trials for all experimented coherences and directions, in particular the pre-motion onset period activity, the pre-saccadic enhancement (for T_{in} choices) and post-saccadic inhibition in activity (for both tasks), and the variable delay period activity (for FD task); and highlighting the important role of BG in contextually gating the saccadic response, depending on whether the task is an FD or an RT type.

4.4. Some model predictions

The MODE model explains and simulates the major data properties derived from the popular random dot motion discrimination experiments in terms of how the brain solves the aperture problem by using feature tracking signals to provide the best estimates possible of object motion directions. The motion processing stages of the model are capable of handling arbitrary motion stimuli. We predict that several stimulus manipulations should shift the psychometric function to the right, increase reaction times, and have influences on MT, MST, and LIP responses similar to those that occur due to lowering the motion strength. These manipulations include using higher dot density, more interleaved sequences, briefer

signal dots, and more spatial displacement of signal dots between frames.

Another prediction regards the lack of coherence-independent initial transient pause in LIP activity in the FD tasks that can be seen in Fig. 5c and e, unlike in the RT task (see Fig. 5a), due to the predictability of motion onset after the choice targets turn on. Self-normalization of the recurrent competitive field predicts lower LIP pre-motion onset period activity in multiple-choice experiments that use more choice targets (cf., Basso & Wurtz, 1998).

The model hypothesis that a volitionally activated top-down signal can help to make a "forced choice" in difficult RT task trials is a prediction that already has some support from recent data (Churchland et al., 2007). Also testable is the model explanation for how the average LIP responses in the RT task for correct T_{out} choices continue to be influenced by motion strength, even in the final stages before the saccade is initiated.

4.5. Outstanding issues

All oculomotor areas have been shown to be "choice-predictive" in the direction discrimination experiments discussed in this article. This is usually understood by the fact that the monkeys were trained to indicate their perceptual decision using a saccadic eye movement, and not some other response. In particular, cells in other saccadic areas (SC, FEF, Area 46), whose response fields contain just one choice target, exhibit direction-selective and decision-related activity (Horwitz & Newsome, 2001b; Kim & Shadlen, 1999) similar to those in LIP. Given the extensive reciprocal connectivity among these areas, it remains to analyze in each particular experimental paradigm which saccade-related area selects the target and relays it to other areas. The locus of the "decision-maker" is of particular interest in "abstract" versions of the basic direction discrimination task, which dissociate the decision-making process from programming of the movement that leads to reward (Gold & Shadlen, 2003; Horwitz et al., 2004b). In these experiments, the targets are not present during motion viewing, and appear only after motion offset and at unpredictable locations. Hence, a choice target cannot be selected and the saccade vector cannot be computed until after the kinetic dots shut off, unlike in simpler tasks in which the saccade can be planned while the motion cues are still on. Involvement of the prefrontal cortex is likely in such memory and rule-based planned movements.

Another issue concerns how the monkey learns the task strategy within an operant conditioning framework (Born, 2004; Uka & DeAngelis, 2004)? In the experiments simulated herein, monkeys had to choose between two peripheral choice targets based on motion information extracted from the foveal display. It is likely that only those pools of foveal neurons in MT and MST that compute either right or left motion direction gradually get associated with appropriate populations in several oculomotor structures because of extensive training on the task. As a result, LIP cells are able to show direction selectivity to motion stimuli presented well outside their classical receptive fields (Shadlen & Newsome, 2001). A monkey whose brain is well-shaped for this particular task may not be able to immediately perform an almost identical task requiring, instead, an up vs. down direction discrimination, because the relay links of the pertinent sensory representation to the applicable choice networks, wherever they are present, are not yet learned (Freedman & Assad, 2006). If, moreover, a monkey has been trained to do both the above tasks, interesting issues of task strategy-switching that can be triggered by contextual cues come to the fore.

This is in contrast to human subjects who can perform any version of the task, even the abstract ones, just by a mere instruction, probably because of facilitation by language. Explaining the neural mechanisms whereby language achieves such flexibility is a large theoretical challenge.

The model hypotheses about how the basal ganglia may gate movement decisions are highly simplified. A more thorough analysis of how laminar prefrontal cortical circuits interact with the basal ganglia, sensory cortices, and other subcortical structures to learn and perform movement tasks is provided by the TELOS model of Brown et al. (2004). Here, after the model incrementally learned five different movement tasks that monkeys have been trained to perform, it quantitatively simulated the neurophysiologically recorded dynamics of seventeen different cell types in these brain regions when those tasks were performed. Future modeling work will incorporate a more complete representation of such prefrontal and basal ganglia interactions with sensory cortices.

Appendix A. Model equations and parameters

A.1. Inputs

Model inputs were random dot motion sequences at varying levels of coherence. They were generated in MATLAB 7.1 as described in the neurophysiological experiments being simulated; see http://cns.bu.edu/~advait/RDMstimuli.html. Three independent two-dimensional movie sequences were first produced, and then interleaved to obtain the actual input, which was presented to the model in real time at the frame rate of 60 Hz. Each frame was a 60×60 grid with 100 white dots (intensity = 255) of size 1×1 on a black background (intensity = 0). The number of dots per frame may appear to be more than 100 because of visual persistence. In the first frame in each of the three sequences, the dots were randomly placed. For the subsequent frames, a fixed fraction of dots (% coherence) from the immediately previous frame were randomly chosen, and displaced by 1 pixel in the signal direction (either left or right), and the remaining dots were relocated randomly. Thus in every frame, a new fraction of dots was chosen to move in the signal direction. Varying this fraction hence controls the motion strength of the dot display. Given the three frames interleaving and 60 Hz refresh rate, the effective speed of the signal dots was 20 pixels/s, and there was systematic correlation among the temporary signal dots with a phase of three frames. If any signal dot moved out of the grid, then it was wrapped around to appear from the opposite side and thus be within the grid.

The phenomenology of the inputs thus created was confirmed by testing informally with many human subjects. The simulation experiment comprised 500 trials at each of six coherence levels, which were as follows: 0%, 3.2%, 6.4%, 12.8%, 25.6%, and 51.2%. Zero-percent coherence corresponds to no correct global motion direction. For all non-zero coherences, the correct direction was always "Right" for the sake of simplicity. In the monkey experiments, trials were randomly and equally distributed between the "Left" and the "Right" directions both to prevent monkeys from developing a bias towards either direction and to study how the non-selected target cells would respond (see Section 3). This is not an issue in the simulations because we explicitly model two LIP populations which prefer the two opponent motion directions (see Appendix A.3.6). The model was found to work also for the input frame rate of 75 Hz (Results not included). Note that the model integration time-step $\Delta t = 0.001$ s is much smaller than $1/60 = 0.0167$ s or $1/75 = 0.0133$ s (inter-frame duration) so that information embedded in inter-frame transitions is registered.

A.2. Shunting and habituative dynamics

Each cell in the network is modeled using a membrane equation (Grossberg, 1973; Hodgkin & Huxley, 1952), which can be written as:

$$\frac{dV}{dt} = \alpha[-\beta V + (\gamma - V)I_{ex} - (V + \delta)I_{in}]. \quad (8)$$

In Eq. (8), V represents the cell membrane potential; $\frac{dV}{dt}$ is the rate at which V changes; parameter α scales cell response speed; parameter β controls the passive decay rate; parameters γ and $-\delta$ are the saturation, or maximum and minimum, potentials, respectively; and I_{ex} and I_{in} are total excitatory and inhibitory inputs, respectively. Eq. (8) thus represents a leaky interaction between time-varying excitatory and inhibitory signals. Membrane or shunting equations additionally have properties of automatic gain control and self-normalization. Some model cells exhibit additive dynamics, which is a special case of the shunting equations where gain control effects are not significant, and one type has habituative dynamics. The model network obeys a large system of simultaneous first-order, nonlinear differential equations, so analytic solutions are not feasible. Instead, cell dynamics were numerically integrated in MATLAB 7.1 using the forward Euler method with a fixed time-step of 0.001s. This time-step was small enough to guarantee numerical stability. In particular, the activity of each cell V in the model was monitored through time to check if it remained within its firing range, namely between $-\delta$ and γ, in order to detect any potential divergences, and we never found any. Model stages until MST (see Fig. 1) are organized in a columnar architecture, with each column processing eight equally spaced directions at a spatial location. A total of about 180,000 cells were engaged for the simulation of each trial, and a total of 3000 trials were simulated (500 trials/coherence, 6 coherences).

A.3. Model processing stages

The following cell types and interactions model the processing stages from Retina to cortical area LIP that are summarized in Fig. 1. The design of the cells in the motion processing stages of the model (i.e., model stages until MST) are adapted from Berzhanskaya et al. (2007) and Grossberg et al. (2001). The activities of all cells at rest are initialized to zero unless otherwise noted. The notation $[\ ]^+$ is used to denote half-wave rectification.

A.3.1. Change-sensitive receptors: Non-directional transient cells

Model non-directional transient cells (Grossberg et al., 2001) simplify how magnocellular cells in Retina and lateral geniculate nucleus (LGN) transiently respond to temporal luminance changes in the visual input. Magnocellular ON and OFF cells give transient responses to bright stimulus onset and offset or dark stimulus offset and onset, respectively (Baloch et al., 1999; Schiller, 1992). Non-directional ON transient cell activities b_{ij} at each spatial location (i,j) are modeled as follows:

$$b_{ij} = [x_{ij}z_{ij} - \theta_b]^+. \tag{9}$$

In Eq. (9), cell activities x_{ij} perform shunted, and thus bounded, leaky integration of input luminance increments I_{ij}:

$$\frac{dx_{ij}}{dt} = A_1(-B_1x_{ij} + (1 - x_{ij})I_{ij}). \tag{10}$$

The output from cell activity x_{ij} is gated by a habituative transmitter z_{ij} which is initially fully accumulated at 1. Non-zero activation x_{ij} results in habituation, or depression, of the transmitter gate z_{ij} according to the following equation (Grossberg, 1972, 1980):

$$\frac{dz_{ij}}{dt} = A_2(1 - z_{ij} - K_2x_{ij}z_{ij}). \tag{11}$$

When a non-zero input I_{ij} turns on, x_{ij} in Eq. (10) approaches 1 with a rate proportional to $(1 - x_{ij})I_{ij}$, balanced by passive decay to 0 at a rate proportional to $-B_1x_{ij}$. As x_{ij} grows, the transmitter z_{ij} begins to habituate, or depress. When I_{ij} shuts off, x_{ij} returns to 0, whereas z_{ij} recovers to 1. These inverse changes in x_{ij} and z_{ij} create a transient pulse of activation in their product $x_{ij}z_{ij}$ in Eq. (9). The output threshold θ_b in Eq. (9) ensures that the duration of a typical non-directional transient cell signal is roughly about 50 ms (Kaplan & Benardete, 2001). Parameters are $A_1 = 1$, $B_1 = 10$, $A_2 = 1$, $K_2 = 50$, and $\theta_b = 0.1$.

These transient responses are sensitive to the contrast of the moving dots; i.e., the magnitude of I_{ij}, but are insensitive to the duration for which stimulus turns on or off beyond a critical duration (see Fig. 2). ON and OFF transient cells code the leading and trailing edges, respectively, of a moving bright object (Schiller, 1992). Here, only ON cells sensitive to luminance increments are simulated for simplicity. Modeling OFF cells and their cross talk with ON cells becomes crucial if the model primate were to judge the direction of second-order motion, such as contrast-reversing or reverse-phi motion stimuli (Anstis, 1970; Anstis & Rogers, 1975; Baloch et al., 1999; Chubb & Sperling, 1989).

A.3.2. Directional transient cells

Two cell types interact to realize directional selectivity at a wide range of speeds, directional transient cells, which generate output signals, and directional inhibitory interneurons which influence these directional output signals (Grossberg et al., 2001). This predicted interaction is consistent with rabbit retinal data concerning how bipolar cells interact with inhibitory starburst amacrine cells and direction-selective ganglion cells, and how starburst cells interact with each other and with ganglion cells (Fried, Münch, & Werblin, 2002). The predicted role of starburst cells in ensuring directional selectivity at a wide range of speeds has not yet been tested. The motion circuits in the model are sensitive to 8 equally spaced directions (Right, Top Right, Top, Top Left, Left, Down Left, Down, Down Right).

Directional inhibitory interneuron (cf., starburst amacrine cell) activities c^{d}_{ij} integrate transient cell inputs b_{ij} as follows:

$$\frac{dc^{\mathrm{d}}_{ij}}{dt} = A_3(-c^{\mathrm{d}}_{ij} + C_3b_{ij} - K_3[c^{\mathrm{D}}_{XY}]^+). \tag{12}$$

In all equations starting now, superscript d denotes the directional preference of the cell. These interneurons receive excitatory inputs from transient non-directional cells b_{ij} at the same position, and inhibition from directional interneurons $[c^{\mathrm{D}}_{XY}]^+$ of opposite directional preference D at the position (X,Y) which is offset by 1 unit from (i,j) in the direction d. For example, for the direction of motion $\phi = \frac{\pi}{4}$, $X = i + 1$, $Y = j + 1$, and D refers to $\phi = \frac{5\pi}{4}$. Inhibition is a lot stronger than excitation ($K_3 > C_3$) so that a directional signal can be vetoed or at least greatly suppressed if the stimulus arrives from the null direction (Barlow & Levick, 1965). This arrangement results in excitation that lags behind inhibition temporally. Parameters are $A_3 = 5$, $C_3 = 5$, and $K_3 = 20$.

Directional transient cell (cf., ganglion cell) activities e^{d}_{ij} in the next layer combine transient inputs b_{ij} with inhibitory interneuronal signals $[c^{\mathrm{D}}_{XY}]^+$.

$$\frac{de^{\mathrm{d}}_{ij}}{dt} = A_4\left(-e^{\mathrm{d}}_{ij} + C_4b_{ij} - K_4[c^{\mathrm{D}}_{XY}]^+\right), \tag{13}$$

and they generate rectified output signals:

$$E^{\mathrm{d}}_{ij} = [e^{\mathrm{d}}_{ij} - \theta_e]^+. \tag{14}$$

This circuit is depicted in Fig. 8 of Grossberg et al. (2001). Parameters are $A_4 = 50$, $C_4 = 5$, $K_4 = 20$, and $\theta_e = 0.2$. Output threshold θ_e in Eq. (14) ensures that the duration of directional transient cell outputs is about 70 ms. Directional inhibitory interneurons operate on a slower time-scale ($A_3 < A_4$) so that directional selectivity is extracted despite slower speeds.

A.3.2.1. Directional short-range filter. Short-range filter cell activities f^{d}_{ij} in V1 spatio-temporally accumulate motion in each direction d as follows:

$$\frac{df_{ij}^{d}}{dt} = A_5\left(-f_{ij}^{d} + \sum_{XY} E_{XY}^{d} G_{XYij}^{d}\right). \quad (15)$$

In Eq. (15), E_{ij}^{d} is the rectified directional transient cell output (Eq. (14)), and G_{XYij}^{d} is an anisotropic Gaussian kernel whose major axis is aligned with the preferred direction of motion d, and centered at (i,j):

$$G_{XYij}^{d} = \frac{G}{2\pi\sigma_x\sigma_y}\exp\left(-0.5\left(\left(\frac{[X-i]\cos(\phi)+[Y-j]\sin(\phi)}{\sigma_x}\right)^2 + \left(\frac{-[X-i]\sin(\phi)+[Y-j]\cos(\phi)}{\sigma_y}\right)^2\right)\right), \quad (16)$$

where the kernel has values $G = 100$, $\sigma_x = 2$, $\sigma_y = 0.5$, and $\phi = \left(\frac{d-1}{4}\right)\pi$. Short-range filter cells are thresholded and rectified:

$$F_{ij}^{d} = [f_{ij}^{d} - \theta_f]^{+} \quad (17)$$

with output threshold $\theta_f = 0.5$. This threshold value ensures that, when a dot moves, only the true direction of the dot survives. The short-range filter can thus accumulate directional evidence from a succession of individually activated directional transient cells. Parameter $A_5 = 0.5$.

A.3.3. Spatial and opponent directional competition

Directional signals are influenced by two types of competition, which help to amplify feature tracking signals especially at motion discontinuities: spatial competition within each direction contrast enhances more active cells within a spatial range that is defined by an on-center anisotropic Gaussian kernel J_{XYij}^{d}, and an off-surround, offset isotropic Gaussian kernel K_{XYij}. Directionally opponent cells at each location compete as well, generating a push–pull effect. Both types of competition act through shunting terms that divisively normalize cell activities. These competing cell activities h_{ij}^{d} obey:

$$\frac{dh_{ij}^{d}}{dt} = A_6\left(-h_{ij}^{d} + (1-h_{ij}^{d})\sum_{XY} F_{XY}^{d} J_{XYij}^{d} - (h_{ij}^{d}+0.5)\left[\sum_{XY} F_{XY}^{d} K_{XYij} + K_6 F_{ij}^{D}\right]\right). \quad (18)$$

The anisotropic excitatory kernel J_{XYij}^{d} obeys:

$$J_{XYij}^{d} = \frac{J}{2\pi\sigma_x\sigma_y}\exp\left(-0.5\left(\left(\frac{X-i}{\sigma_x}\right)^2 + \left(\frac{Y-j}{\sigma_y}\right)^2\right)\right). \quad (19)$$

For horizontal motion directions $\phi = 0$ and $\phi = \pi$, the kernel has values $\sigma_x = 3$ and $\sigma_y = 1$. Kernels for other directions are obtained by rotation as in Eq. (16). The isotropic inhibitory kernel K_{XYij} is:

$$K_{XYij} = \frac{K}{2\pi\sigma^2}\exp\left(-0.5\left(\frac{(X-i_0)^2+(Y-j_0)^2}{\sigma^2}\right)\right) \quad (20)$$

with $\sigma = 4$. The center $(i_0 j_0)$ of the off-surround kernel K_{XYij} is offset from (i,j) by 1 U in the direction D which is opposite to the cell's preferred direction d. This arrangement results in inhibition that trails excitation spatially. The activities h_{ij}^{d} are half-wave rectified to generate the output signals:

$$H_{ij}^{d} = [h_{ij}^{d}]^{+}. \quad (21)$$

Parameters are $A_6 = 5$, $K_6 = 1$, $J = 75$, and $K = 75$.

A.3.4. Directional long-range filter

Middle temporal area (MT) cells m_{ij}^{d} receive a combination of excitatory bottom-up inputs N_{ij}^{d} that filter the output signals from the competition stage H_{ij}^{d} through a long-range anisotropic Gaussian filter L_{XYij}^{d}:

$$N_{ij}^{d} = \sum_{XY} H_{XY}^{d} L_{XYij}^{d}, \quad (22)$$

and top-down matching signals $[T_{XY}^{e}]^{+}$ from medial superior temporal area (MST):

$$\frac{dm_{ij}^{d}}{dt} = A_7(-m_{ij}^{d} + (1-m_{ij}^{d})N_{ij}^{d} - (m_{ij}^{d}+0.5) \times \sum_{e}\sum_{XY} w^{de}[T_{XY}^{e}]^{+} P_{XYij}) + \sigma_0\frac{dW_0}{dt}. \quad (23)$$

In Eq. (23), W_0 is a Brownian motion process that injects noise into MT dynamics, and parameter σ_0 scales the standard deviation of the Brownian process. A similar noise process influences MST responses too (Eq. (28)). The simulations in Figs. 5–14 did not activate noise in MT/MST, and hence σ_0 was set to zero. In the simulations shown in Fig. 15, MT/MST had their noise terms activated. Two different values of σ_0 were used: 0.1 and 0.5. The first value is about the same as that used in LIP ($\sigma = 5$), but scaled for the different activity ranges $\left(5 \times \frac{(80-0)}{(1-(-0.5))} = 0.0938 \approx 0.1\right)$.

The long-range filter L_{ijXY}^{d} for horizontal motion directions $\phi = 0$ and $\phi = \pi$ obeys:

$$L_{XYij}^{\phi=0,\pi} = \frac{L}{2\pi\sigma_x\sigma_y}\exp\left(-0.5\left(\left(\frac{X-i}{\sigma_x}\right)^2 + \left(\frac{Y-j}{\sigma_y}\right)^2\right)\right) \quad (24)$$

with $\sigma_x = 10$ and $\sigma_y = 3$. The major axis of the kernel is along the preferred direction d. Kernels for other directions are obtained by rotation as in Eq. (16). Note that the long-range filter L_{XYij}^{d} is larger in spatial scale than the short-range filter G_{XYij}^{d}. The spatial extent of top-down MST feedback is determined by an inhibitory Gaussian filter P_{XYij}:

$$P_{XYij} = \frac{P}{2\pi\sigma^2}\exp\left(-0.5\left(\frac{(X-i)^2+(Y-j)^2}{\sigma^2}\right)\right) \quad (25)$$

with $\sigma = 8$. Biased competition among all motion directions, which is important for aperture problem resolution, happens within MT and MST. Accordingly, inhibition is received from all directions except d. The inhibitory weight w^{de} in Eq. (23) between direction d and another direction e is given by:

$$w^{de} = \begin{cases} 0, & e = d \\ 1, & e \neq d, D \\ 2, & e = D \end{cases}, \quad (26)$$

where D is the direction opposite to d. Since the excitatory input N_{ij}^{d} is from the preferred direction, this directionally asymmetric inhibition tends to amplify d and suppress other motion directions, including the opponent direction D. MST-to-MT feedback is thus modeled by a modulatory on-center, off-surround network, a type of top-down attentive matching signal that can also explain how directional selectivity is learned (Grossberg, 2003). As a result of this process, the degree to which a direction at a given location is directly suppressed or indirectly boosted depends on the spatial distance from a feature tracking signal, and its consistency with the feature tracking direction. Surviving directional signals from MST try to select the consistent direction in MT and propagate that direction across space via the inhibitory kernel P_{XYij}. In the dots task, the effectiveness of this selection and propagation process depends on percent coherence or motion strength. The output signals from the long-range filter stage obey:

$$M_{ij}^{d} = [m_{ij}^{d}]^{+}. \quad (27)$$

Parameters are $A_7 = 10$, $L = 10$, and $P = 1$.

A.3.5. Directional grouping

MST activity T_{ij}^{d} is described by

$$\frac{dT_{ij}^{d}}{dt} = A_8\left(-T_{ij}^{d} + (1 - T_{ij}^{d})C_8 M_{ij}^{d} - (T_{ij}^{d} + 0.5)\sum_{e}\sum_{XY} w^{de}[T_{XY}^{e}]^{+} P_{XYij}\right) + \sigma_0 \frac{dW_0}{dt}. \quad (28)$$

In Eq. (28), the bottom-up input M_{ij}^{d} is the rectified MT output. Recurrent MST inhibitory connections help to select a winning direction to be fed back to MT. The spatial extent and directional asymmetry of this inhibitory interaction are again defined by kernel P_{XYij} and weighting coefficients w^{de} in Eqs. (25) and (26), respectively. Parameters are $A_8 = 10$ and $C_8 = 10$. As in Eq. (23), MST was subjected to cellular noise via term W_0 in the simulations reported in Fig. 15.

The self-normalizing directional activities in MT and MST are sensitive to the ambiguity inherent in the random dot motion stimulus, as shown by model simulations (see Fig. 11). They are spatially distributed, as are the moving dots that excite them. LIP processing receives these spatially distributed directional signals and converts them into probabilistic directional eye movement commands, as follows.

A.3.6. Decision cells and decision gating

Lateral intraparietal area (LIP) firing rate dynamics are related to the speed and accuracy of perceptual decisions in random dot motion direction discrimination tasks. LIP shows decision-related activity because monkeys have been trained to communicate their decision about the motion direction with an eye movement to the appropriate choice target. LIP is known to be an association area, where visual, cognitive, and saccadic signals all play a role. We model two LIP populations representing the two choice targets using a recurrent competitive field (RCF). The LIP equation for the RT task (Roitman & Shadlen, 2002) is as follows:

$$\frac{dy_d}{dt} = -A_9 y_d + (B_9 - y_d)[\lambda S_d + T_C + g_d f(y_d) + G_d] - y_d[g_2 h(y_D) + G_I] + \sigma \frac{dW}{dt}, \quad (29)$$

where y_d is the activity of the LIP population that codes direction d, S_d represents the population activity of the foveal MST pool tuned to direction d, namely:

$$S_d = \sum_{XY}[T_{XY}^{d}]^{+}, \quad (30)$$

T_C is the bottom-up input due to the presence of the visual target within the receptive field of each LIP population, $f(y_d)$ is the recurrent on-center feedback signal, G_d is an excitatory, higher level input which is only invoked to make a forced choice, especially after a reasonable time has elapsed without a decision, y_D is the activity of the LIP population with the opponent motion preference, $h(y_D)$ is the recurrent off-surround feedback signal, G_I is an inhibitory, higher level signal which causes inhibition after the eyes start moving, W is a Brownian motion process that simulates the stochastic components of LIP activity, σ^2 scales the variance of the Brownian motion process, and g_d and g_2 are gains of the recurrent on-center and off-surround terms, respectively. The value of the visual signal T_C is chosen so that the two LIP population activities equilibrate at an activity θ before the onset of the dots, as in the data:

$$T_C = \frac{\theta[A_9 + g_2 h(\theta)]}{B_9 - \theta} - g_d f(\theta). \quad (31)$$

The above-baseline activity θ is attributed to the presence of the choice target within the receptive field of the LIP population. The excitatory recurrent signal is defined by:

$$f(x) = \frac{x^n}{\mu^n + x^n}. \quad (32)$$

Basal ganglia (BG) gate both the competing saccade plans from realization. When the activity of either population (say direction d) reaches the decision threshold Γ_1, the BG are assumed to open a movement gate that switches the gain g_d to g_{BG}. As a result, the activity of that population is rapidly amplified, thereby simulating the pre-saccadic enhancement in activity seen for about 100 ms before the eyes start moving. The inhibitory recurrent signal is defined by:

$$h(x) = \frac{x^m}{\mu^m + x^m}. \quad (33)$$

Parameters are $A_9 = 1$, $B_9 = 80$, $\lambda = \frac{5}{60\times 60} = 1.39 \times 10^{-3}$, $g_d = 2$, $g_2 = 5$, $g_{BG} = 10$, $\mu = 50$, $n = 5$, $m = 10$, $\Gamma_1 = 55$, $\sigma = 5$, and $\theta = 35$ (estimated from data).

Inputs G_d and G_I in Eq. (29) are initially equal to zero. When either activity (y_d or y_D) reaches $\Gamma_2 = 70$ (see Figs. 7, 9, and 11 in Roitman & Shadlen, 2002), the eyes begin to move, as can be seen in the data. Post-saccadic suppression from FEF post-saccadic cells now comes into play by turning G_I on from 0 to 30. Reaction time (RT) in each trial is recorded as the time from motion onset until when either activity y_d or y_D crosses Γ_2; that is, when the choice saccade is actually initiated. The random dot motion display turns off upon saccade onset. We assume post-saccadic suppression, and decay due to loss of sensory inputs, reset LIP activities to zero before a new trial begins. The value for the parameter B_9 is chosen as $\Gamma_2 + 10$, so that the LIP cell activity y_d can exceed criterion level Γ_2. The value of Γ_1 is also chosen from data by identifying the level of firing in the RT task for correct T_{in} choices, after which differences in input motion strength do not seem to play a role.

Sometimes, neither activity can reach threshold Γ_1 in a reasonable time (about 1 s). The model is then forced to make a choice, as follows: if for some random time between 1000 ms and 1100 ms, the model has not taken a decision, then the direction d corresponding to the greater LIP activity is chosen by turning on G_d from 0 to $2g_d$. Simulations done without this volitional mechanism at work are reported in Fig. 13.

Since the Brownian motion process $W(t)$ in Eq. (29) is an independent increments process, the increments $[W(t) - W(s)]$ are zero-mean, Gaussian random variables with variance $t - s$ for $t > s$. Thus at each time step ($\Delta t = 0.001s$) of numerical integration, a zero-mean, Gaussian random variable of variance $\sigma^2 \Delta t$ is added to the evolving LIP activity. The same is true in Eq. (34) below.

The FD task simulations use the same equation:

$$\frac{dy_d}{dt} = -A_9 y_d + (B_9 - y_d)[\lambda S_d + T_C + g_d f(y_d)] - y_d[g_2 h(y_D) + G_I] + \sigma \frac{dW}{dt}. \quad (34)$$

The various parameters here have the same meaning as those described for the RT task. Experimental data from two FD tasks (Roitman & Shadlen, 2002; Shadlen & Newsome, 2001) are simulated. First, we model the 2002 FD task which was done on the same macaques in alternating blocks of trials with the RT task. In this task, the dots turn off after 1 s of viewing. The monkey needs to make a choice based on all evidence accumulated in the 1 s. Thus, the model monkey basically chooses that direction with the greater LIP activity at the end of the motion viewing period. The passive decay rate A_9 is given a slightly higher value (4.5) for the FD task; see Section 3.2. In the 2002 FD task, the monkeys have to remember their choice during an ensuing delay period, whose duration is varied randomly between 500 ms and 1500 ms, before the fixation point is extinguished, which cues the monkey to go ahead with the chosen saccade. This is enabled by a jump in the recurrent on-center gain g_d of the judged direction d to g_{delay} at the end of the motion stimulus viewing period. As a result, activity y_d is pre-

vented from being attenuated during the delay period, and also builds up slowly, as observed in the data. When the fixation point turns off, the basal ganglia are assumed to open the gate for the appropriate saccade plan, which results in the typical pre-saccadic enhancement in activity. This is realized by a further, instant gain boost for the on-center gain g_d to g_{BG}. Post-saccadic inhibition is triggered when y_d crosses $\Gamma_2 = 70$ by turning on G_I from 0 to 30 just as in the RT task. Parameters g_{delay} and g_{BG} are 9 and 35.

The 2001 FD task recorded from different monkeys and hence different LIP populations. The LIP responses during the task were therefore somewhat different, for example in the dynamic range of firing ($0 - \Gamma_2$ Hz), and in the above-baseline firing due to the presence of the target before the onset of dots (θ). Accordingly, the simulated LIP decision circuit was the same as that of the 2002 FD task, but some parameters were not: $A_9 = 5$, $B_9 = 60$, $\lambda = \frac{3}{60\times 60} = 8.33 \times 10^{-4}$, $g_d = 2$, $g_2 = 7.5$, $\mu = 25$, $n = 5$, $m = 10$, $\sigma = 5$, $\theta = 15$, $\Gamma_2 = 50$. The values for θ and Γ_2 were again chosen from data (see Figs. 8 and 12 in Shadlen & Newsome, 2001), and $B_9 = \Gamma_2 + 10$. In this FD task, the delay period duration is randomly chosen between 200 ms and 2000 ms. The recurrent on-center gain g_d associated with the chosen direction d jumps from 2 to 5.5 (g_{delay}) at the start of the delay period, and then to 25 (g_{BG}) when the FP is turned off. As in the other task simulations, post-saccadic inhibition begins when y_d crosses Γ_2, which turns on G_I from 0 to 30.

A.4. Psychometric function

For each task, the simulated proportion correct data, $p(C)$, defined to be the fraction out of 500 trials at each percent coherence C, is fit with a cumulative Weibull distribution function for two-alternative, forced choice tasks:

$$p(C) = \frac{1}{2} + \frac{1}{2}\left(1 - \exp\left[-\left(\frac{C}{\alpha}\right)^{\beta}\right]\right) \tag{35}$$

using a maximum likelihood estimation procedure presented in Myung (2003). Parameter α is the threshold of the psychometric function, and corresponds to the coherence level that elicits ~82% accuracy. Parameter β controls the steepness of the psychometric function. Since there are just two alternatives, the Weibull function yields a chance performance of 0.5, or 50%, at 0% coherence $\left(p(0) = \frac{1}{2}\right)$.

References

Albright, T. D., Desimone, R., & Gross, C. G. (1984). Columnar organization of directionally sensitive cells in visual area MT of the macaque. *Journal of Neurophysiology, 51*, 16–31.

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research, 10*, 1411–1430.

Anstis, S. M., & Rogers, B. J. (1975). Illusory reversals of visual depth and movement during changes in contrast. *Vision Research, 15*, 957–961.

Baloch, A. A., & Grossberg, S. (1997). A neural model of high-level motion processing: Line motion and formotion dynamics. *Vision Research, 37*, 3037–3059.

Baloch, A. A., Grossberg, S., Mingolla, E., & Nogueira, C. A. M. (1999). Neural model of first-order and second-order motion perception and magnocellular dynamics. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 16*, 953–978.

Barlow, H. B., & Levick, W. R. (1965). The mechanism of directionally selective units in the rabbit's retina. *Journal of Physiology, 178*, 477–504.

Basso, M. A., & Wurtz, R. H. (1998). Modulation of neuronal activity in superior colliculus by changes in target probability. *Journal of Neuroscience, 18*, 7519–7534.

Ben-Av, M. B., & Shiffrar, M. (1995). Disambiguating velocity estimates across image space. *Vision Research, 35*, 2889–2895.

Berzhanskaya, J., Grossberg, S., & Mingolla, E. (2007). Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. *Spatial Vision, 20*, 337–395.

Bichot, N. P., & Schall, J. D. (1996). Visual feature selectivity in frontal eye fields induced by experience in mature macaques. *Nature, 381*, 697–699.

Bisley, J. W., & Goldberg, M. E. (2003). Neuronal activity in the lateral intraparietal area and spatial attention. *Science, 299*, 81–86.

Bisley, J. W., Krishna, B. S., & Goldberg, M. E. (2004). A rapid and precise on-response in posterior parietal cortex. *Journal of Neuroscience, 24*, 1833–1838.

Boardman, I., Grossberg, S., Myers, C., & Cohen, M. (1999). Neural dynamics of perceptual order and context effects for variable-rate speech syllables. *Perception & Psychophysics, 6*, 1477–1500.

Born, R. T. (2004). Taking strategies to task. *Neuron, 42*, 185–187.

Bowns, L. (1996). Evidence for a feature tracking explanation of why type II plaids move in the vector sum direction at short durations. *Vision Research, 36*, 3685–3694.

Bowns, L. (2001). IOC, vector sum, and squaring: Three different motion effects or one? *Vision Research, 41*, 965–972.

Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience, 13*, 87–100.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Visual Neuroscience, 10*, 1157–1169.

Brown, J. W., Bullock, D., & Grossberg, S. (2004). How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Networks, 17*, 471–510.

Bruce, C. J., Goldberg, M. E., Bushnell, M. C., & Stanton, G. B. (1985). Primate frontal eye fields: II. Physiological and anatomical correlates of electrically-evoked eye movements. *Journal of Neurophysiology, 54*, 714–734.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory. *Psychological Review, 100*, 432–459.

Castet, E., Lorenceau, J., Shiffrar, M., & Bonnet, C. (1993). Perceived speed of moving lines depends on orientation, length, speed and luminance. *Vision Research, 33*, 1921–1936.

Celebrini, S., & Newsome, W. T. (1994). Neuronal and psychophysical sensitivity to motion signals in extrastriate MST of the macaque monkey. *Journal of Neuroscience, 14*, 4109–4124.

Celebrini, S., & Newsome, W. T. (1995). Microstimulation of extrastriate area MST influences performance on a direction discrimination task. *Journal of Neurophysiology, 73*, 437–448.

Chey, J., Grossberg, S., & Mingolla, E. (1997). Neural dynamics of motion grouping: From aperture ambiguity to object speed and direction. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 14*, 2570–2594.

Chey, J., Grossberg, S., & Mingolla, E. (1998). Neural dynamics of motion processing and speed discrimination. *Vision Research, 38*, 2769–2786.

Chubb, C., & Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Sciences of the United States of America, 86*, 2985–2989.

Churchland, A. K., Kiani, R., & Shadlen, M. N. (2007). LIP neurons combine accumulated evidence with a representation of elapsed time—urgency—to mediate a perceptual decision. *Society of Neuroscience Abstracts, 507.3*.

Churchland, A. K., Tam, M., Plamer, J., Kiani, R., & Shadlen, M. N. (2005). Responses of LIP neurons reflect accumulation of evidence in a multiple choice decision task. *Society of Neuroscience Abstracts, 16.8*.

Cisek, P. (2006). Integrated neural processes for defining potential actions and deciding between them: A computational model. *Journal of Neuroscience, 26*, 9761–9770.

Cisek, P., & Kalaska, J. F. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action. *Neuron, 45*, 801–814.

Ditterich, J. (2006a). Stochastic models of decisions about motion direction: Behavior and physiology. *Neural Networks, 19*, 981–1012.

Ditterich, J. (2006b). Evidence for time-variant decision making. *European Journal of Neuroscience, 24*, 3628–3641.

Ditterich, J., Mazurek, M., & Shadlen, M. N. (2003). Microstimulation of visual cortex affects the speed of perceptual decisions. *Nature Neuroscience, 6*, 891–898.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation and time course. *Annual Review of Psychology, 48*, 269–297.

Francis, G., & Grossberg, S. (1996a). Cortical dynamics of boundary segmentation and reset: Persistence, afterimages and residual traces. *Perception, 35*, 543–567.

Francis, G., & Grossberg, S. (1996b). Cortical dynamics of form and motion integration: Persistence, apparent motion, and illusory contours. *Vision Research, 36*, 149–173.

Francis, G., Grossberg, S., & Mingolla, E. (1994). Cortical dynamics of feature binding and reset: Control of visual persistence. *Vision Research, 34*, 1089–1104.

Freedman, D. J., & Assad, J. A. (2006). Experience-dependent representation of visual categories in parietal cortex. *Nature, 443*, 85–88.

Fried, S. I., Münch, T. A., & Werblin, F. S. (2002). Mechanisms and circuitry underlying directional selectivity in the retina. *Nature, 420*, 411–414.

Gancarz, G., & Grossberg, G. (1998). A neural model of the saccade generator in the reticular formation. *Neural Networks, 11*, 1159–1174.

Gancarz, G., & Grossberg, S. (1999). A neural model of saccadic eye movement control explains task-specific adaptation. *Vision Research, 39*, 3123–3143.

Glimcher, P. W. (2003). The neurobiology of visual-saccadic decision making. *Annual Review of Neuroscience, 26*, 133–179.

Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences, 5*, 10–16.

Gold, J. I., & Shadlen, M. N. (2003). The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *Journal of Neuroscience, 23*, 632–651.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision-making. *Annual Review of Neuroscience, 30*, 535–574.

Grossberg, S. (1972). A neural theory of punishment and avoidance: II. Quantitative theory. *Mathematical Biosciences, 15*, 253–285.
Grossberg, S. (1973). Contour enhancement, short term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics, 52*, 213–257.
Grossberg, S. (1978). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen & F. Snell (Eds.). *Progress in theoretical biology* (Vol. 5, pp. 233–374). New York, NY: Academic.
Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review, 87*, 1–51.
Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks, 1*, 17–61.
Grossberg, S. (2003). How does the cerebral cortex work? Development, learning, attention, and 3D vision by laminar circuits of visual cortex. *Behavioral and Cognitive Neuroscience Reviews, 2*, 47–76.
Grossberg, S., & Myers, C. W. (2000). The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review, 107*, 735–767.
Grossberg, S., Boardman, I., & Cohen, C. (1997). Neural dynamics of variable-rate speech categorization. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 418–503.
Grossberg, S., Mingolla, E., & Viswanathan, L. (2001). Neural dynamics of motion integration and segmentation within and across apertures. *Vision Research, 41*, 2521–2553.
Grossberg, S., Roberts, K., Aguilar, M., & Bullock, D. (1997). A neural model of multimodal adaptive saccadic eye movement control by superior colliculus. *Journal of Neuroscience, 17*, 9706–9725.
Grossberg, S., & Rudd, M. E. (1989). A neural architecture for visual motion perception: Group and element apparent motion. *Neural Networks, 2*, 421–450.
Grossberg, S., & Rudd, M. E. (1992). Cortical dynamics of visual motion perception: Short-range and long-range apparent motion. *Psychological Review, 99*, 78–121.
Grossberg, S., Srihasam, K., & Bullock, D. (submitted for publication). Neural dynamics of saccadic and smooth pursuit eye movement coordination during visual tracking of unpredictably moving targets.
Heeger, D. G., Boynton, G. M., Demb, J. B., Seidemann, E., & Newsome, W. T. (1999). Motion opponency in visual cortex. *Journal of Neuroscience, 19*, 7162–7174.
Heekeren, H. R., Marrett, S., Bandettini, P. A., & Ungerleider, L. G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature, 431*, 859–862.
Helmholtz, H. (2000). *Treatise on physiological optics*. Bristol, UK: Thoemmes.
Hikosaka, O., & Wurtz, R. H. (1983). Visual and oculomotor functions of monkey substantia nigra pars reticulata. IV. Relation of substantia nigra to superior colliculus. *Journal of Neurophysiology, 49*, 780–798.
Hikosaka, O., & Wurtz, R. H. (1989). The basal ganglia. In R. Wurtz & M. Goldberg (Eds.), *The neurobiology of saccadic eye movements* (pp. 257–281). Amsterdam: Elsevier.
Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology, 117*, 500–544.
Horwitz, G. D., Batista, A. P., & Newsome, W. T. (2004a). Direction-selective visual responses in macaque superior colliculus induced by behavioral training. *Neuroscience Letters, 366*, 315–319.
Horwitz, G. D., Batista, A. P., & Newsome, W. T. (2004b). Representation of an abstract perceptual decision in macaque superior colliculus. *Journal of Neurophysiology, 91*, 2281–2296.
Horwitz, G. D., & Newsome, W. T. (2001a). Target selection for saccadic eye movements: Direction-selective visual responses in the superior colliculus. *Journal of Neurophysiology, 86*, 2527–2542.
Horwitz, G. D., & Newsome, W. T. (2001b). Target selection for saccadic eye movements: Prelude activity in the superior colliculus during a direction-discrimination task. *Journal of Neurophysiology, 86*, 2543–2558.
Jazayeri, M., & Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience, 9*, 690–696.
Kanizsa, G. (1979). *Organization in vision: Essays on gestalt perception*. New York, NY: Praeger.
Kaplan, E., & Benardete, E. (2001). The dynamics of primate retinal ganglion cells. *Progress in Brain Research, 134*, 17–34.
Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology, 55*, 271–304.
Kersten, D., & Yuille, A. (2003). Bayesian models of object perception. *Current Opinion in Neurobiology, 13*, 1–9.
Kim, J.-N., & Shadlen, M. N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neuroscience, 2*, 176–185.
Knill, D., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neuroscience, 27*, 712–719.
Li, X., Kim, B., & Basso, M. A. (2006). Transient pauses in delay-period activity of superior colliculus neurons. *Journal of Neurophysiology, 95*, 2252–2264.
Lorenceau, J., & Gorea, A. (1989). 'Blobs' are critical in perceiving the direction of moving plaids. *Perception, 18*, 539.
Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research, 32*, 263–273.
Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York, NY: Oxford University.
Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience, 9*, 1432–1438.
Mazurek, M. E., Hanks, T., Yang, T., & Shadlen, M. N. (2005). Prior probability changes the rate of evidence accumulation in a motion discrimination task: Behavior and LIP physiology. *Society of Neuroscience Abstracts, 621.3*.
Mazurek, M. E., Roitman, J. D., Ditterich, J., & Shadlen, M. N. (2003). A role for neural integrators in perceptual decision making. *Cerebral Cortex, 13*, 1257–1269.
Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience, 16*, 5154–5167.
Mingolla, E., Todd, J. T., & Norman, J. F. (1992). The perception of globally coherent motion. *Vision Research, 32*, 1015–1031.
Munoz, D. P., & Wurtz, R. H. (1995a). Saccade-related activity in monkey superior colliculus. I. Characteristics of burst and buildup cells. *Journal of Neurophysiology, 73*, 2313–2333.
Munoz, D. P., & Wurtz, R. H. (1995b). Saccade-related activity in monkey superior colliculus. II. Spread of activity during saccades. *Journal of Neurophysiology, 73*, 2334–2348.
Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology, 47*, 90–100.
Nichols, M. J., & Newsome, W. T. (2002). Middle temporal visual area microstimulation influences veridical judgments of motion direction. *Journal of Neuroscience, 22*, 9530–9540.
Pack, C. C., & Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in area MT of macaque brain. *Nature, 409*, 1040–1042.
Palmer, J., Huk, A., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision, 5*, 376–404.
Pilly, P. K., & Grossberg, S. (2005). Brain without Bayes: Temporal dynamics of decision-making in the laminar circuits of visual cortex. *Society of Neuroscience Abstracts, 591.1*.
Pilly, P. K., & Grossberg, S. (2006). Brain without Bayes: Temporal dynamics of decision-making during form and motion perception by the laminar circuits of visual cortex. *Journal of Vision, 6*, 886.
Platt, M. L. (2002). Neural correlates of decisions. *Current Opinion in Neurobiology, 12*, 141–148.
Pouget, A., Dayan, P., & Zemel, R. S. (2003). Computation and inference with population codes. *Annual Review of Neuroscience, 26*, 381–410.
Rao, R. (2004). Bayesian computation in recurrent neural circuits. *Neural Computation, 16*, 1–38.
Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59–108.
Ratcliff, R. (1980). A note on modeling accumulation of information when the rate of accumulation changes over time. *Journal of Mathematical Psychology, 21*, 178–184.
Ratcliff, R., Cherian, A., & Segraves, M. (2003). A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of simple two-choice decisions. *Journal of Neurophysiology, 90*, 1392–1407.
Ratcliff, R., van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review, 106*, 261–300.
Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9*, 347–356.
Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111*, 333–367.
Reddi, B. A. J., & Carpenter, R. H. S. (2000). The influence of urgency on decision time. *Nature Neuroscience, 3*, 827–830.
Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience, 22*, 9475–9489.
Romo, R., & Salinas, E. (2001). Touch and go: Decision-making mechanisms in somatosensation. *Annual Review of Neuroscience, 24*, 107–137.
Salzman, C. D., Murasugi, C. M., Britten, K. H., & Newsome, W. T. (1992). Microstimulation in visual area MT: Effects on direction discrimination performance. *Journal of Neuroscience, 2*, 2331–2355.
Schall, J. D. (2003). Neural correlates of decision processes: Neural and mental chronometry. *Current Opinion in Neurobiology, 13*, 182–186.
Schiller, P. H. (1992). The ON and OFF channels of the visual system. *Trends in Neuroscience, 15*, 86–92.
Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology, 86*, 1916–1936.
Smith, P. L. (1990). A note on the distribution of response times for a random walk with Gaussian increments. *Journal of Mathematical Psychology, 34*, 445–459.
Smith, P. L. (2000). Stochastic dynamic models of response time and accuracy: A foundational primer. *Journal of Mathematical Psychology, 44*, 408–463.
Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences, 27*, 161–168.
Smith, P. L., & Vickers, D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology, 32*, 135–168.
Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The response of area MT and V1 neurons to transparent motion. *Journal of Neuroscience, 11*, 2768–2785.
Srihasam, K., Bullock, D., & Grossberg, S. (2008). Target selection by frontal cortex during coordinated saccadic and smooth pursuit eye movements. *Journal of Cognitive Neuroscience*, in press.
Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience, 9*, 578–585.
Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*, 520–522.

Uka, T., & DeAngelis, G. C. (2004). Contribution of area MT to stereoscopic depth perception: Choice-related response modulations reflect task strategy. *Neuron, 42*, 297–310.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108*, 550–592.

Wallach, H. (1935). On the visually perceived direction of motion. *Psychologische Forschung, 20*, 325–380.

Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron, 36*, 955–968.

Wong, K.-F., & Wang, X.-J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience, 26*, 1314–1328.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience, 5*, 598–604.