

# 1 Text S1

2 In text S1, we provide details for the stochastic optimal control framework used to model  
3 reach and eye movements.

## 4 1 Hand modeling.

5 To preserve interpretability, it is important to use a model no more complex than absolutely nec-  
6 essary. In the current study, we modeled the dynamics of the hand as a “point of mass” ( $m =$   
7 1 Kg), in Cartesian coordinates, with 2-dimensional position  $\mathbf{p}_t^h = [x_h(t), y_h(t)]^T$  and velocity  
8  $\dot{\mathbf{p}}_t^h = [\dot{x}_h(t), \dot{y}_h(t)]^T$ <sup>1</sup>. The combined actions of all muscles is represented by the force vector  
9  $\mathbf{f}_t = [f_x(t), f_y(t)]^T$ . The motor command  $\mathbf{u}_t$  is transformed into these forces  $\mathbf{f}_t$  through second-  
10 order muscle-like low-pass filters with constants  $\tau_1$  and  $\tau_2$ , and by adding control-dependent mul-  
11 tiplicative noise<sup>1</sup>.

$$\tau_1\tau_2\ddot{\mathbf{f}}_t + (\tau_1 + \tau_2)\dot{\mathbf{f}}_t + \mathbf{f}_t = \mathbf{u}_t \quad (1)$$

12 The second-order low pass filters can be written as a pair of coupled first-order filters with outputs  
13  $\mathbf{g}$  and  $\mathbf{f}$ , as:

$$\tau_1\dot{\mathbf{g}}_t + \mathbf{g}_t = \mathbf{u}_t, \quad \tau_2\dot{\mathbf{f}}_t + \mathbf{f}_t = \mathbf{g}_t \quad (2)$$

14 Given that  $\mathbf{p}^G = [x^G, y^G]$  is the position that the hand is planned to arrive, at the end of the  
15 movement (i.e., the current goal of the task), the discrete-time state is described by the following  
16 10<sup>th</sup>-dimensional vector  $\mathbf{x}_t$ .

$$\mathbf{x}_t = [x_h(t), y_h(t), \dot{x}_h(t), \dot{y}_h(t), f_x(t), f_y(t), g_x(t), g_y(t), x_G, y_G]^T \quad (3)$$

17

The discrete-dynamics of the hand are given by Eq. (4):

$$\begin{aligned} \mathbf{p}_{t+\delta t}^h &= \mathbf{p}_t^h + \dot{\mathbf{p}}_t^h \delta t \\ \dot{\mathbf{p}}_{t+\delta t}^h &= \dot{\mathbf{p}}_t^h + \mathbf{f}_t \frac{\delta t}{m} \\ \mathbf{f}_{t+\delta t} &= \mathbf{f}_t \left(1 - \frac{\delta t}{\tau_2}\right) + \mathbf{g}_t \frac{\delta t}{\tau_2} \\ \mathbf{g}_{t+\delta t} &= \mathbf{g}_t \left(1 - \frac{\delta t}{\tau_1}\right) + \mathbf{u}_t (1 + \sigma_c \epsilon_t) \frac{\delta t}{\tau_1} \end{aligned} \quad (4)$$

18

where  $\delta t = 0.01s$  is the sampling period of discretization. In the current framework, we as-

19

sume that the motor control signal  $\mathbf{u}_t$  is corrupted by multiplicative noise described by the product

20

term  $\sigma_c \epsilon_t \mathbf{u}_t$ , with  $\sigma_c = 1$ , which is a unitless variable defined as the noise magnitude related to the

21

control signal magnitude <sup>1</sup>.

22

## 2 Reaching to a single goal.

23

In the current study, each controller is linked with a neuron  $i$  from the ‘‘motor plan formation field’’.

24

The purpose of the controller is to generate an optimal policy that drives the hand to a distance  $r$

25

from the current location, towards the preferred direction  $\phi$  of the neuron  $i$ . Within the stochastic

26 optimal control framework, control policies originate as the solution of an optimization problem.  
 27 The basic idea is to find a sequence of motor commands that acquire as much reward as possible,  
 28 while spending as little effort as possible. For reaching tasks, given the kinematics of the hand  
 29 and the sensory and motor noise in estimating and controlling the state of the hand, the stochastic  
 30 optimal control framework finds an optimal policy  $\mathbf{u}^*$ , for time instances  $t = [t_1, t_2, \dots, t_T]$  that  
 31 optimizes the cost function  $J$  described in Eq. (5), for each state of the hand and the environment  
 32  $\mathbf{x}_t$ .

$$J(\mathbf{x}_t, \pi) = \|\mathbf{p}_T^h - \mathbf{p}^G\|^2 + \|\dot{\mathbf{p}}_T^h\|^2 + \|\mathbf{f}_T\|^2 + \sum_{t=1}^T \pi(\mathbf{x}_t)^T R \pi(\mathbf{x}_t) \quad (5)$$

33 The first 3 terms define that current task of the controller - i.e., drive the hand to the position  
 34  $\mathbf{p}^G = [r \cos(\phi) \ r \sin(\phi)]$  (first term) and stop there (second and third term). The last term is the  
 35 motor command cost (i.e., action cost) that penalizes the effort required to arrive at the position  $\mathbf{p}^G$   
 36 after time  $T$  starting from the current state. The matrix  $R = \frac{1}{T} \begin{bmatrix} r_x & 0 \\ 0 & r_y \end{bmatrix}$  is the control-dependent  
 37 cost of the hand motion in the x and y dimension.

38 We can write Eq. (5) in the general form of optimal control cost function as follows:

$$J = (\mathbf{x}_T - S\mathbf{p}^G)^T Q_T (\mathbf{x}_T - S\mathbf{p}^G) + \sum_{t=1}^{T-1} \pi(\mathbf{x}_t)^T R \pi(\mathbf{x}_t) \quad (6)$$

39 where  $S$  matrix picks out the actual position of the hand and the current goal-position  $\mathbf{p}^G$  at  
40 the end of the movement, from the state-vector. The time-varying matrix  $Q_T$  describes the state-  
41 dependent cost and is the zero matrix for any time  $t < T$  and is equal to the Hessian matrix of the  
42 cost function evaluated at the end of the movement  $T$ .

43 To minimize the cost function in Eq. (6), a model of the system dynamics and sensory feed-  
44 back must be incorporated. A plethora of experimental studies provide evidence that the sensory  
45 system uses an internal forward model to predict the next state of the system at time  $t + 1$ ,  $\hat{\mathbf{x}}_{t+1|t}$ ,  
46 based on the sensory feedback  $\mathbf{y}_t$ , the current state estimate  $\hat{\mathbf{x}}_t$  and the control commands  $\mathbf{u}_t$ <sup>2</sup>. This  
47 prediction is necessary to overcome control instabilities due to noisy sensors and temporal delays.  
48 In the current study, we modeled the hand and the state space using linear dynamics and measure-  
49 ment as a discrete linear system, Eq. (7), considering that the motor commands are corrupted by  
50 multiplicative noise, normally distributed with zero mean and standard deviation proportional to  
51 the magnitude of the control commands and the state variables<sup>1</sup>.

$$\begin{aligned}\mathbf{x}_{t+1} &= A\mathbf{x}_t + B\mathbf{u}_t + \xi_t + C(\mathbf{u}_t)\epsilon_t \\ \mathbf{y}_t &= H\mathbf{x}_t + \omega_t\end{aligned}\tag{7}$$

52 where  $A$ ,  $B$  and  $H$  are the actual system dynamics and observation matrices, respectively.

$$A = \begin{bmatrix} 1 & 0 & \delta t & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \delta t & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \frac{\delta t}{m} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & \frac{\delta t}{m} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 - \frac{\delta t}{\tau_2} & 0 & \frac{\delta t}{\tau_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 - \frac{\delta t}{\tau_2} & 0 & \frac{\delta t}{\tau_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 - \frac{\delta t}{\tau_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 - \frac{\delta t}{\tau_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

53

$$B = \begin{bmatrix} \mathbf{0}_{6 \times 2} \\ \frac{\delta t}{\tau_1} & 0 \\ 0 & \frac{\delta t}{\tau_1} \\ \mathbf{0}_{2 \times 2} \end{bmatrix}, H = \begin{bmatrix} \mathbf{I}_{6 \times 6} & | & \mathbf{0}_{6 \times 4} \end{bmatrix} \quad (8)$$

54

55 The noise variables  $\xi_t$ ,  $\omega_t$  and  $\epsilon_t$  are normally distributed variables with zero mean and co-  
56 variance  $\Omega^\xi \geq 0$ ,  $\Omega^\omega \geq 0$  and  $\Omega^\epsilon = I$ , respectively.  $C(\mathbf{u}_t)$  is a scaling matrix for control-dependent

57 system noise, such as  $C(\mathbf{u})\epsilon = \sum_i C_i \mathbf{u} \epsilon^i$ , where  $\epsilon^i$  is the  $i^{th}$  component of the random variable  $\epsilon$ .

58 Given the belief about the state at time  $t$  and the current task of the controller, the controller  
59 will suggest an optimal policy  $\pi^*(\mathbf{x}_t) = \mathbf{u}^*$  that minimizes the expected cost function in Eq. (5).  
60 This form of optimal control is a modified version of the Linear Quadratic Gaussian (LQG) regu-  
61 lator, since the dynamics of the system are linear, the expected cost function is quadratic and the  
62 noise is Gaussian, but with signal-dependent noise <sup>1</sup>.

63 The optimal policy is incorporated into the system model to generate a feedback controller  
64 that uses its forward model to make predictions  $\hat{\mathbf{y}}_t$  from knowledge of controls, dynamics and  
65 sensory measurements. These predictions are combined with actual sensory feedback  $\mathbf{y}_t$  using  
66 Eq. (9) to update the belief about the state in time  $t + 1$ .

$$\hat{\mathbf{x}}_{t+1|t+1} = (A - BL_t)\hat{\mathbf{x}}_{t+1|t} + K_t (\mathbf{y}_t - H\hat{\mathbf{x}}_{t+1|t}) \quad (9)$$

67 where  $K_t$  is the Kalman gain at time  $t$ .

### 68 **3 Eye modeling**

69 In the current study, we modeled the globe and the surrounding tissues of the eye by an elastic  
70 element with stiffness  $k$  that pulls the eye away from the equilibrium point ( $x = 0$ ), and a viscous  
71 element  $\mu$  that resists in this motion (for more details about this eye model see <sup>3</sup>). The one-  
72 dimensional dynamics of the eye is given by Eq. (10).

$$m\ddot{x} = -kx - \mu\dot{x} + u \quad (10)$$

73 where  $m$  is the inertia of the eye and  $u$  are the motor commands.

74 We used a simple model to translate motor commands  $u$  into forces:  $\alpha_1 \dot{f} + \alpha_2 f = u$ , where  
 75  $f$  describes the instantaneous force generated by the extra-ocular muscles to move the eye.

76 Let's assume that the current goal is to move the eye at the position  $\mathbf{p}^G = [x_G, y_G]^T$ . The  
 77 discrete-time state is given by the  $8^{th}$ -dimensional vector  $\mathbf{x}_t$ , Eq. (11).

$$\mathbf{x}_t = [x_e(t), \dot{x}_e(t), y_h(t), \dot{y}_e(t), f_x, f_y, x_G, y_G]^T \quad (11)$$

78 We extended this one-dimensional model of the eye to a more realistic two-dimensional model, in  
 79 which the discrete-time dynamics of the eye are given by Eq. (12):

$$\begin{aligned} \mathbf{p}_{t+\delta t}^e &= \mathbf{p}_t^e + \dot{\mathbf{p}}_t^e \delta t \\ \dot{\mathbf{p}}_{t+\delta t}^e &= -\delta t \frac{k}{m} \mathbf{p}_t^e + \left(1 - \delta t \frac{\mu}{m}\right) \dot{\mathbf{p}}_t^e + \frac{\delta t}{m} \mathbf{f}_t \\ \mathbf{f}_{t+\delta t} &= \left(1 - \delta t \frac{\alpha_2}{\alpha_1}\right) \mathbf{f}_t + \mathbf{u}_t (1 + \sigma_c \epsilon_t) \frac{\delta t}{\alpha_1} \end{aligned} \quad (12)$$

80 where  $\mathbf{p}_t^e = [x_e(t), y_e(t)]^T$  and  $\dot{\mathbf{p}}_t^e = [\dot{x}_e(t), \dot{y}_e(t)]^T$  is the 2-dimensional position and velocity  
 81 of the eye, respectively. Similarly, with the hand model, the product term  $\sigma_c \epsilon_t \mathbf{u}_t$  describes the  
 82 multiplicative noise added to the control signal  $\mathbf{u}_t$ , with  $\sigma_c = 1$  and  $\epsilon_t$  a vector of zero-mean  
 83 random variables with covariance  $\Omega^e = \mathbf{I}$ . Additionally,  $\delta t = 0.001\text{s}$  is the sampling period of  
 84 discretization.

## 85 **4 Saccade to a single goal**

86 Even though saccades are brief in time and sensory information seems to have no role in the control  
87 of eye movements, recent findings suggest that saccades are not open-loop movements. Instead,  
88 the motor commands that generate saccades benefit from a forward internal model that monitors  
89 the motor commands and predicts their consequences <sup>4</sup>.

90 We modeled eye movements within the stochastic optimal control framework using a similar  
91 approach with the hand movements presented in a previous section. However, the difference with  
92 eye movements is that the sensory feedback does not affect the control of the eye movements. Each  
93 controller is linked to a neuron  $i$  from the motor plan formation field, and generates a policy that  
94 drives the eye towards the preferred direction  $\phi$  of the neuron  $i$ . The cost function is the same as  
95 the one used in reaching movements (see Eq. (5)).

96 The dynamics of the eye can be transformed into the form of Eq. (7), with the following  
97 matrices:



$$\begin{aligned}
A = & \begin{bmatrix}
1 & \delta t & 0 & 0 & 0 & 0 & 0 & 0 \\
-\frac{k\delta t}{m} & 1 - \frac{\mu\delta t}{m} & 0 & 0 & \frac{1}{m}\delta t & 0 & 0 & 0 \\
0 & 0 & 1 & \delta t & 0 & 0 & 0 & 0 \\
0 & 0 & -\frac{k\delta t}{m} & 1 - \frac{\mu\delta t}{m} & 0 & \frac{1}{m\delta t} & 0 & 0 \\
0 & 0 & 0 & 0 & 1 - \frac{\alpha_1\delta t}{\alpha_2} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 - \frac{\alpha_1\delta t}{\alpha_2} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix} \\
B = & \begin{bmatrix}
\mathbf{0}_{4 \times 2} \\
\frac{\delta t}{\alpha_1} & 0 \\
0 & \frac{\delta t}{\alpha_1} \\
\mathbf{0}_{2 \times 2}
\end{bmatrix}
\end{aligned} \tag{13}$$

98

99 For a monkey eye, we used time constants  $\tau_1 = 0.260$  s and  $\tau_2 = 0.012$  s as proposed by  
100 Keller<sup>5</sup> and Robinson<sup>6</sup>. These time constants are related with the constants of the system dynamic  
101 matrices  $A$  and  $B$  as follows:  $\mu = \tau_1 + \tau_2$ ,  $m = \tau_1\tau_2$ . Based on these studies<sup>5,6</sup>, we set  $k = 1$ ,  
102  $\alpha_1 = 0.004$  and  $\alpha_2 = 1$ .

## References

104  
103

- 105 1. Todorov, E. *Stochastic optimal control and estimation methods adapted to the noise charac-*  
106 *teristics of the sensorimotor system*, Neural Comput., 17, 1084-1108 (2005)
- 107 2. Diedrichsen, J., Shadmehr, R., & Ivry, R. *The coordination of movement. Optimal feedback*  
108 *control and beyond*, Trends Cogn Sci.,14, 31-39 (2010)
- 109 3. Haith, A., Reppert, R., & Shadmehr, R. *Evidence for hyperbolic temporal discounting of re-*  
110 *ward in control of movements.*, J Neurosci. 32, 11727-11736 (2012)
- 111 4. Ethier, V., Zee, D., & Shadmehrm R. *Changes in control of saccades during gain adaptation,*  
112 *J Neurosci.* 28, 13929-13937 (2010).
- 113 5. Keller, E. *Accommodative vergence in the alert monkey. Motor unit analysis*, Vision Res. 13,  
114 1565-1575 (1973)
- 115 6. Robinson, D., Gordon, J., & S. Gordon *A model of the smooth pursuit eye movement system,*  
116 *Biol Cybern.* 55, 43-57 (1986).