A Single, Continuously Applied Control Policy for Modeling Reaching Movements with and without Perturbation

Zhe Li
lizhe07@mails.tsinghua.edu.cn
Department of Biomedical Engineering and Center for Brain-Inspired Computing Research, Tsinghua University, Beijing 100084, China

Pietro Mazzoni
mazzonip@wustl.edu
Department of Neurology, Washington University in St. Louis, St. Louis, MO 63110, U.S.A.

Sen Song
songsen@tsinghua.edu.cn
Department of Biological Engineering and Center for Brain-Inspired Computing Research, Tsinghua University, Beijing 100084, China

Ning Qian
nq6@columbia.edu
Department of Neuroscience and Department of Physiology and Cellular Biophysics, Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY 10032, U.S.A.

It has been debated whether kinematic features, such as the number of peaks or decomposed submovements in a velocity profile, indicate the number of discrete motor impulses or result from a continuous control process. The debate is particularly relevant for tasks involving target perturbation, which can alter movement kinematics. To simulate such tasks, finite-horizon models require two preset movement durations to compute two control policies before and after the perturbation. Another model employs infinite- and finite-horizon formulations to determine, respectively, movement durations and control policies, which are updated every time step. We adopted an infinite-horizon optimal feedback control model that, unlike previous approaches, does not preset movement durations or use multiple control policies. It contains both control-dependent and independent noises in system dynamics, state-dependent and independent noises in sensory feedbacks, and different delays and noise levels for visual and proprioceptive feedbacks. We analytically derived an optimal solution that can be applied continuously to move an effector toward a target regardless of whether, when, or where the

target jumps. This single policy produces different numbers of peaks and "submovements" in velocity profiles for different conditions and trials. Movements that are slower or perturbed later appear to have more submovements. The model is also consistent with the observation that subjects can perform the perturbation task even without detecting the target jump or seeing their hands during reaching. Finally, because the model incorporates Weber's law via a state representation relative to the target, it explains why initial and terminal visual feedback are, respectively, less and more effective in improving end-point accuracy. Our work suggests that the number of peaks or submovements in a velocity profile does not necessarily reflect the number of motor impulses and that the difference between initial and terminal feedback does not necessarily imply a transition between open- and closed-loop strategies.

1 Introduction

Submovement decomposition is a major class of methods in motor research that has been applied to analyze and compare various experimental conditions and subject groups (Flash & Henis, 1991; Milner, 1992; Lee, Port, & Georgopoulos, 1997; Krebs, Aisen, Volpe, & Hogan, 1999; Novak, Miller, & Houk, 2002). It assumes that a reaching movement is a superposition of primitives, each produced by a motor impulse or command. Typically a velocity profile is decomposed into a sum of shifted and scaled versions of a standard profile, and the number of fitted components (submovements) is interpreted as indicating the number of underlying discrete motor impulses or commands. An alternative view is that a continuously applied control policy can explain movement kinematics, including submovements. Indeed, stochastic optimal control models have been shown to produce movement kinematics similar to observed ones using continuous control signals without explicitly summing discrete submovements (Harris & Wolpert, 1998; Todorov & Jordan, 2002; Scott, 2004; Tanaka, Krakauer, & Qian, 2006; Shadmehr & Krakauer, 2008; Qian, Jiang, Jiang, & Mazzoni, 2013).

The target perturbation paradigm (Georgopoulos, Kalaska, & Massey, 1981; Soechting & Lacquaniti, 1983; Flash & Henis, 1991; Heath, Hodges, Chua, & Elliott, 1998) provides a particularly interesting case to contrast these two views. In this paradigm, subjects are instructed to reach with their hand to a target. At some point before or after movement onset, the target jumps to a new position, and subjects have to correct their movements to reach the new position. It is known that depending on the jump time, the perturbation can greatly change movement kinematics compared with the corresponding no-perturbation condition. In particular, the velocity profile varies with jump time and may show a plateau or a new peak. The submovement view contends that such changes in velocity profile indicate

changes in the number of discrete motor commands or impulses for controlling the movements (Fishbach, Roy, Bastianen, Miller, & Houk, 2005, 2007). In contrast, the continuous-control view argues that there is no fundamental difference between the conditions with and without target perturbation (Pélisson, Prablanc, Goodale, & Jeannerod, 1986; Desmurget & Grafton, 2000); in both cases, the system estimates its state (using sensory feedback and internal prediction) to correct errors and move the effector toward the target in a continuous process. Although perturbation introduces a large error, it is corrected in the same way as errors in the no-perturbation case.

Optimal feedback control models have already been used to explain target-perturbation experiments. Some (Hoff & Arbib, 1993; Liu & Todorov, 2007) used finite-horizon formulations, which require a predetermined movement duration to define the cost function and solve the optimization problem (see the discussions in Tanaka et al., 2006 and Qian et al., 2013). To simulate target perturbation, they have to be given two movement durations (taken from experimental data) to compute two solutions for use before and after the target jump, respectively. Consequently, the models are forced to use two control policies to simulate target perturbation experiments. Although each policy is run continuously, the requirement of two policies and a switch between them weaken the argument for the continuous-control view. Moreover, requiring the models to "know" two durations seems biologically implausible, particularly when the target jumps to an unpredictable position. Another model (Rigoux & Guigon, 2012) employed both infinite- and finite-horizon formulations. The infinite-horizon component specifies a movement duration, which is then used to determine the corresponding finite-horizon control policy, and both the duration and control policy are updated at each time step. The model has the advantage of treating movements with and without target perturbation uniformly. However, it has to compute a large number of durations and policies equal to the number of time steps for any movement. Moreover, if target perturbation increases movement duration (as it often does), the model requires more control policies.

We therefore investigated whether a single, continuously applied control policy could simulate movements both with and without target perturbation. We and others recently proposed infinite-horizon stochastic optimal feedback control models that do not preset or precompute movement durations or rely on finite-horizon control policies (Huh, Todorov, & Sejnowski, 2010; Jiang, Jiang, & Qian, 2011; Qian et al., 2013). In this letter, we extend and apply this framework to target perturbation. We also model the observation that subjects can perform the perturbation task even when their vision is suppressed during the target jump and they do not see their hands during reaching (Goodale, Pélisson, & Prablanc, 1986; Pélisson et al., 1986). A recent study found that visual and proprioceptive feedbacks have not only different noise levels but also different delays (Crevecoeur, Munoz, &

Scott, 2016). Although our study does not directly concern this interesting finding, we incorporated it into our model to ensure that our conclusions are valid under more realistic feedback conditions.

A related debate is the relative importance of feedforward versus feedback control of reaching. There is little doubt that reliable sensory feedback can improve performance (Bernstein, 1967; Todorov & Jordan, 2002). However, sensory feedback in the initial phase of reaching does not seem to be as important as that in the terminal phase (Beaubaton & Hay, 1986). This has led to the proposal that motor control uses a crude or even open-loop mechanism initially and switches to a refined, closed-loop mechanism later, implying a two-stage process. However, few models have considered the alternative that a single control policy may appear to assign different importance to initial and terminal feedback. We therefore extended the previous infinite-horizon model (Phillis, 1985; Jiang et al., 2011; Qian et al., 2013) to implement Weber's law of sensory perception. This law states that uncertainty in estimating a value is proportional to the value estimated. We therefore assume, for example, that the uncertainty in judging the distance between the hand and target is proportional to this distance (Baird, 1970). We were able to implement Weber's law via state-dependent noise in the sensory feedback equation because we represented the effector state relative to the target. Surprisingly, this well-known sensory law has not been incorporated into optimal feedback control models. We examined whether the law naturally rendered initial and terminal sensory feedback less and more effective, respectively, without assuming a switch between feedforward and feedback strategies.

Some preliminary results were reported in abstract form (Fangwen, Zhaoping, & Qian, 2012).

2 Results

A stochastic optimal feedback control model is defined by equations specifying the system dynamics, sensory feedback (observation), state estimation, control law, and cost function (e.g., Phillis, 1985; Todorov, 2005; Qian et al., 2013). Similar to previous work, we used the following system dynamics:

$$\mathrm{d}x = (Ax + Bu)\mathrm{d}t + Fx\mathrm{d}\beta + Yu\mathrm{d}\gamma + G\mathrm{d}\omega, \tag{2.1}$$

where x is the state vector and u is the control signal. The A and B matrices are determined by the motor plant and Newtonian dynamics (see the appendix for examples). The F, Y, and G terms represent state-dependent, control-dependent, and control-independent noise in the dynamics. β and γ are scalar Wiener processes and ω a vector of Wiener processes; their formal differentials (random-walk steps in time $\mathrm{d}t$) are gaussian white noises

with zero mean and variance dt. The target state is defined to be $x = 0$, so that x really represents the effector state relative to the target.

We modified commonly used observation equations to better represent sensory feedback. First, we considered multiple feedback types or modalities with different parameters,

$$dy_s = C_s x_s dt + Z_s x_s d\delta_s + D_s d\xi_s, \tag{2.2}$$

where $s = 1, 2, \ldots, f$ indicates different types or modalities of sensory feedback with an observation y_s vector for each feedback. For this study, we limited $f = 2$ with $s = 1, 2$ for vision and proprioception, respectively, but our formulation and analytical derivation can be readily extended to larger f. C_s are the observation matrices. Second, the observation vectors y_s are based on the delayed state vectors x_s, which are related to the actual state vector x according to the following low-pass filters (Jiang et al., 2011),

$$\tau_s dx_s = (-x_s + x) dt \tag{2.3}$$

where τ_s are the time constants that determine the delays of the different feedback types (Crevecoeur et al., 2016). (Additionally, for each feedback type, different time constants can be applied to different components of the state vector, but this is omitted here for simplicity.) Using low-pass filters instead of hard time shifts (Todorov, 2005) to model delays is analytically advantageous. It is also physiologically plausible because delays caused by neural processing are more like low-pass filtering than hard time shifts (Reichardt, 1987; Egelhaaf & Reichardt, 1987; Qian & Andersen, 1997). Third, there are both state-dependent noise (the Z_s term) and state-independent noise (the D_s term) in each observation equation. Because x represents the effector state relative to the target, the state-dependent noise naturally implements Weber's law of sensory processing (see section 2.1). The Z_s term was not included previously (Jiang et al., 2011; Qian et al., 2013; Phillis, 1985). δ_s and ξ_s are scalar and vector Wiener processes, respectively. All the random processes in equations 2.1 and 2.2 are independent of each other. Importantly, equations 2.2 and 2.3 allow different delays and noise levels for different sensory feedbacks.

Since our analysis follows Phillis (1985), which does not include delay equations, we combined the delay equations with the system dynamics by defining (Jiang et al., 2011)

$$x^* \equiv \begin{bmatrix} x \\ x_1 \\ x_2 \end{bmatrix}, \quad dy^* \equiv \begin{bmatrix} dy_1 \\ dy_2 \end{bmatrix}, \tag{2.4}$$

and rewriting equations 2.1 to 2.3 as

$$dx^* = (A^*x^* + B^*u)dt + F^*x^*d\beta + Y^*ud\gamma + G^*d\omega^*, \tag{2.5}$$

$$dy^* = C^*x^*dt + Z^*x^*d\delta + D^*d\xi^*, \tag{2.6}$$

where the new, starred symbols are defined by the original ones as

$$A^* \equiv \begin{bmatrix} A & \\ 1/\tau_1 & -1/\tau_1 \\ 1/\tau_2 & & -1/\tau_2 \end{bmatrix}, \; B^* \equiv \begin{bmatrix} B \\ 0 \\ 0 \end{bmatrix}, \; C^* \equiv \begin{bmatrix} 0 & C_1 & 0 \\ 0 & 0 & C_2 \end{bmatrix},$$

$$F^* \equiv \begin{bmatrix} F \\ 0 \\ 0 \end{bmatrix}, \; Y^* \equiv \begin{bmatrix} Y \\ 0 \\ 0 \end{bmatrix}, \; G^* \equiv \begin{bmatrix} G \\ 0 \\ 0 \end{bmatrix},$$

$$Z^* \equiv \begin{bmatrix} 0 & Z_1 & 0 \\ 0 & 0 & Z_2 \end{bmatrix}, \; D^* \equiv \begin{bmatrix} D_1 \\ & D_2 \end{bmatrix}, \; d\omega^* \equiv d\omega, \; d\xi^* \equiv \begin{bmatrix} d\xi_1 \\ d\xi_2 \end{bmatrix},$$

and the 0's and 1's represent zero and identity matrices of proper dimensions.

Our state-estimation and control-law equations are similar to those in previous stochastic optimal feedback control models:

$$d\hat{x}^* = (A^*\hat{x}^* + B^*u)dt + K(dy^* - C^*\hat{x}^*dt), \tag{2.7}$$

$$u = -L\hat{x}^*, \tag{2.8}$$

where $\hat{x}^*$ is the estimation of x^*. The two terms in equation 2.7 represent internal prediction (using an efference copy of u and a forward model of the plant) and corrections from the sensory feedbacks, respectively. (Some studies view internal prediction as a form of feedback—see Desmurget & Grafton, 2000—but we use *feedback* to denote sensory inputs only.)

The goal of optimal feedback control is to determine the Kalman-estimator matrix K and control-law matrix L by minimizing an appropriate cost function. To avoid explicit movement duration required by finite horizon formulations, we adopted an infinite-horizon approach (Phillis, 1985; Jiang et al., 2011; Qian et al., 2013) with the following steady-state cost function:

$$J = \lim_{t \to \infty} \langle \tilde{x}^T U \tilde{x} + x^T Q x + u^T R u \rangle, \tag{2.9}$$

where $\tilde{x} = x - \hat{x}$ is the state estimation error and the brackets indicate expectation over all relevant noises. The three terms represent expected costs (per unit time) of estimation error, reaching error, and energetic effort, respectively. We need to find K and L that minimize J under the constraints

of equations 2.5 to 2.8. We analytically derived the optimal solution as (see the appendix):

$$K = P_{22}C^{\mathrm{T}}(D^*D^{*\mathrm{T}} + Z^*P_{11}Z^{*\mathrm{T}})^{-1}, \tag{2.10}$$

$$L = (R + Y^{*\mathrm{T}}(S_{11} + S_{12})Y^*)^{-1}B^{*\mathrm{T}}S_{11}, \tag{2.11}$$

$$\bar{A}^{\mathrm{T}}S + S\bar{A} + \bar{F}^{\mathrm{T}}S\bar{F} + \bar{Y}^{\mathrm{T}}S\bar{Y} + \bar{Z}^{\mathrm{T}}S\bar{Z} + V = 0, \tag{2.12}$$

$$\bar{A}P + P\bar{A}^{\mathrm{T}} + \bar{F}P\bar{F}^{\mathrm{T}} + \bar{Y}P\bar{Y}^{\mathrm{T}} + \bar{Z}P\bar{Z}^{\mathrm{T}} + \bar{G}\bar{G}^{\mathrm{T}} = 0, \tag{2.13}$$

where P is the correlation matrix of the extended state vector that concatenates x^* and $\tilde{x}^* \equiv x^* - \hat{x}^*$, S is a matrix of Lagrange multipliers, V is a matrix related to the cost terms, subscripts 11 and 12 refer to the upper left and upper right quarter blocks of a square matrix, and the barred matrices are defined with the corresponding unbarred ones (see the Appendix). Phillis's original solution can be recovered by setting $Z^* = 0$. For a given motor plant, the optimal K and L can be found quickly by iterating equations 2.10 to 2.13 to convergence (Jiang et al., 2011; Qian et al., 2013). We define $K \equiv [K_1 \, K_2]$ where K_1 and K_2 are the first and last half columns of K and correspond to the Kalman estimators for visual and proprioceptive feedbacks, respectively. By setting K_1 or K_2 or both to 0 during simulations, we can model different feedback conditions.

2.1 Weber's Law and Different Effects of Initial and Terminal Visual Feedback.

Weber's law of sensory perception has the form

$$\Delta p = kp, \tag{2.14}$$

where Δp is the change in a stimulus property that produces a just noticeable difference in the perception of the property and p is the current value of the property. The dimensionless proportionality constant k is called the Weber fraction. For Δp to be perceptually noticeable, it has to overcome the noise in neural representations of p. Therefore, equation 2.14 implies that noise in p representation is proportional to p. Because in our model, the state vector x represents the end effector relative to the target, the state-dependent observation noise (the Z_s term of equation 2.2) implements Weber's law: as the hand state gets closer to the target state, x becomes smaller in magnitude, and the uncertainty in observation of x becomes smaller accordingly. (State-dependent observation noise does not implement Weber's law for models that use separate, absolute representations of the effector and target.)

The Z_s matrix thus determines Weber fractions in our model. For the diagonal Z_s we used, the diagonal terms are proportional to Weber fractions for the components of x, which include position, velocity, and acceleration (see the Appendix). We determined the ratios among the diagonal

terms of Z_1 using the Weber fractions of visual judgments of distance (Baird, 1970; McKee, Levi, & Bowne, 1990; Levi & Klein, 1992), velocity (McKee, Silverman, & Nakayama, 1986), and acceleration (Mueller & Timney, 2016). In other words, the properties measured in visual psychophysical experiments were assumed to be applicable to the visual guidance of the hand's reaching movements to targets. Following Crevecoeur et al. (2016), we further assumed that proprioception is faster but noisier than vision and let $\tau_1 = 100$ ms, $\tau_2 = 50$ ms, $D_2 = 2.5D_1$, and $Z_2 = 2.5Z_1$. A complication is that proprioception can sense the hand but not the target. Therefore, proprioceptive feedback of the state should really be viewed as resulting from the integration of proprioception of the hand and visual perception or memory of the target, and we assume that Weber's law applies to this integration. Our conclusions in this letter are not sensitive to these or other parameters. For example, we obtained qualitatively similar results when the factor relating proprioceptive and visual noises varied from 2 to 5, or when both delay constants were set to 100 ms or a value from 50 to 150 ms, or when Weber's law for proprioception was eliminated (by setting Z_2 to 0 and doubling D_2).

We used the optimal solution for a single-joint motor plant (see the appendix for details) to simulate the four visual feedback conditions for reaching movements in Beaubaton and Hay (1986). That study compared subjects' movements under full vision, no vision, vision in the initial half and vision in the terminal half of the hand-position trajectory. For all conditions, the target was always visible and stationary. When the hand was invisible, subjects could not use vision to estimate the hand state, so we set $K_1 = 0$. Figure 1 shows the simulated position trajectories (panel a) and the end-point variabilities (panel b) in the four conditions. Consistent with experimental findings, we found that the initial-vision condition produced a more variable end-point position than did the terminal-vision condition, and that with the end-point variability as the measure, the initial- and terminal-vision conditions were, respectively, similar to no-vision and full-vision conditions.

Figure 1b is based on simulation results at 1 s but our conclusion does not depend on this choice of end point. Figure 2 shows the SDs of the hand position as a function of time for the four conditions. The conclusion is the same for any time after 0.5 s as the end time (Figure 1a shows that it is reasonable to assume that movements ended after 0.5 s). Thus, the model explains different effects of initial and terminal visual feedback on end-point accuracy without assuming two mechanisms. Intuitively, because of Weber's law, visual feedback toward the end of a movement is highly accurate due to small distances between the hand and target states, and consequently, its presence versus absence largely determines end-point variability. On the other hand, our simulations suggest that initial feedback had a significant impact on the positional variability during the movement (see Figures 1a and 2). Indeed, in the early phase of movements, the terminal-vision and no-vision conditions were similar
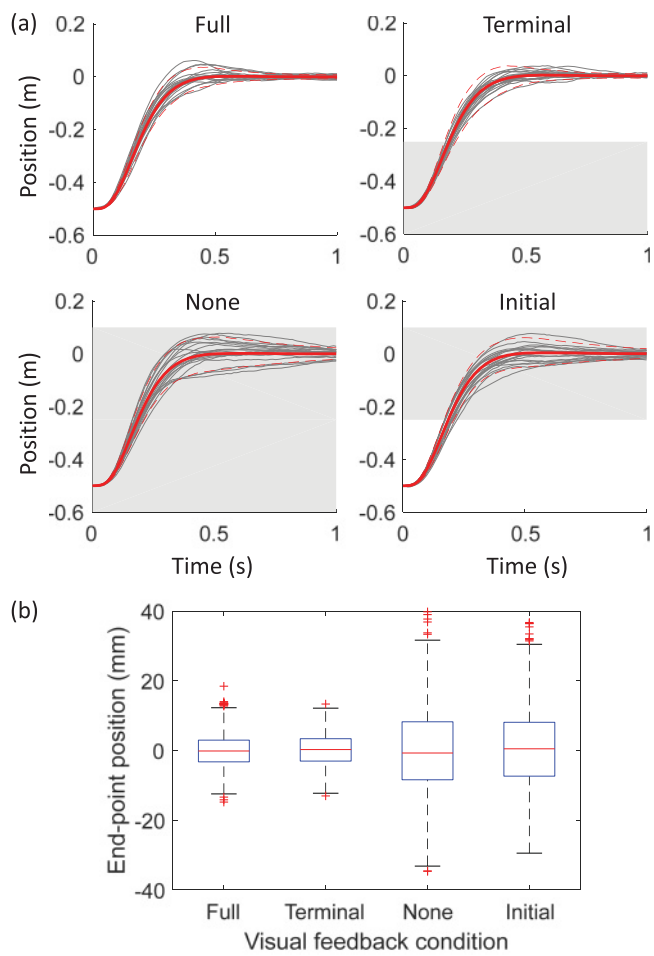
Figure 1: Simulations of the four vision conditions (Beaubaton & Hay, 1986). (a) The four conditions are full vision (Full), no vision (None), vision in the initial half (Initial), and vision in the terminal half (Terminal), of the hand-position trajectory. The white and gray backgrounds indicate visible and invisible hands, respectively. In each panel, the solid red curve is the average hand position trajectory and the two dashed red curves are the 5 and 95 percentiles estimated from 1000 simulations. The dark gray curves are 20 sample trials. (b) The box plot for each condition shows the median end-point position (long red line), 25 and 75 percentiles (blue box), and outliers (red crosses), from 1000 simulations at 1 sec. Consistent with Beaubaton and Hay (1986), the terminal-vision condition was more accurate than the initial-vision condition. In addition, the terminal- and initial-vision conditions were similar to the full- and no-vision conditions, respectively.
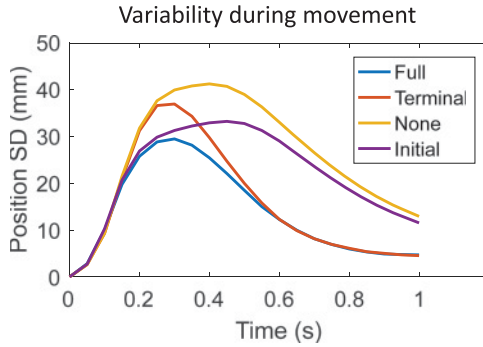
Figure 2: SDs of the hand position as a function of time for the four vision conditions. The blue, orange, yellow, and purple curves are for the full-vision, terminal-vision, no-vision, and initial-vision conditions, respectively. At the beginning, all four conditions had similar SDs because even when visual feedback was available, it was crude (because of Weber's law) and delayed.

to each other. This is not surprising because these two conditions were identical up to the midpoint of each trajectory, and from that point on, the continuous Newtonian dynamics kept the two conditions similar until accurate feedback in the terminal-vision condition gradually made a difference on convergence. Likewise, at the beginning, the initial-vision and full-vision condition were similar to each other. We note that the experiments also showed larger end-point biases in the initial- and no-vision conditions than in the terminal and full-vision conditions (Beaubaton & Hay, 1986). Since this could be trivially explained by assuming that visual perception of the hand state is less biased than proprioception, we did not model this result explicitly.

Finally, Beaubaton and Hay (1986) found that the difference in end-point accuracy between the initial/no-vision conditions and the terminal/full vision conditions became smaller for slower movements. To explain this result, we first defined the movement duration of a trial as the first time when the velocity had been less than 0.05 m/s for 40 ms, and then used this criterion to sort the simulated trials from fast to slow. We then plotted the end-point accuracies of the four vision conditions for the fast trials (durations less than 0.45 s) and slow trials (durations more than 0.6 s) in Figures 3a and 3b, respectively. (Each of the two groups of trials constituted about a quarter of the total number of trials.) Consistent with the experimental finding (Beaubaton and Hay, 1986), the differences in end-point accuracy between the initial/no-vision conditions and the terminal/full vision conditions were smaller for the slow trials.

2.2 Target perturbation at various times. We next used the same model to simulate target-perturbation experiments. In Figure 4, the target was
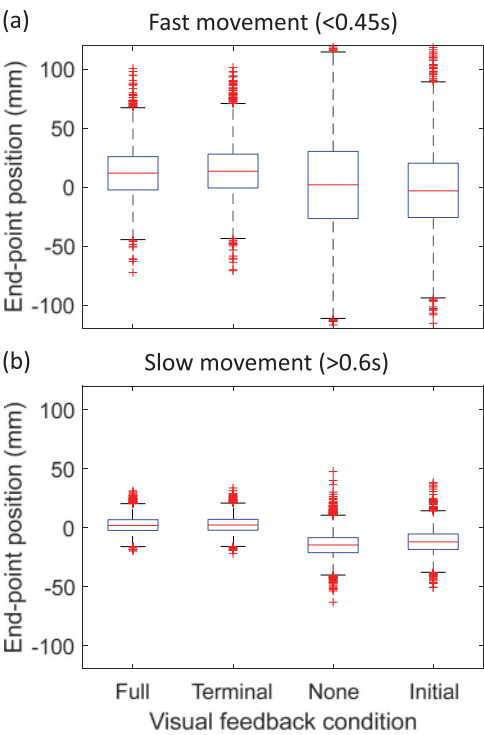
Figure 3: End-point-accuracy difference among the four vision conditions for fast and slow trials. Fast (<0.45 s) and slow (>0.6 s) trials of Figure 1 are analyzed separately. The differences in end-point accuracy between the initial/no-vision conditions and the terminal/full vision conditions were smaller for the slow trials (bottom panel) than for the fast trials (top panel), consistent with the experimental finding of Beaubaton and Hay (1986).

initially 0.5 m away from the hand but jumped farther away from the hand by 0.1 m at various times to make a total distance of 0.6 m. The two columns of the figure show the position and velocity profiles, respectively. Different rows show results of different jump times (no jump, 0.1, 0.2, 0.3, and 0.4 s after reaching starts). Figure 5 shows the corresponding results for the target jumped closer to the hand by 0.1 m to make a total distance of 0.4 m.

The simulations in Figures 4 and 5 reproduced some key features of the target-perturbation experiments (Georgopoulos et al., 1981; Soechting & Lacquaniti, 1983; Flash & Henis, 1991; Heath et al., 1998; Goodale et al., 1986; Pélisson et al., 1986; Fishbach et al., 2005, 2007). When the target jumped early during reaching, the kinematics were similar to those of the no-jump condition, with smooth and single-peaked mean velocity profiles. When the target jumped late, however, the mean velocity profiles became less
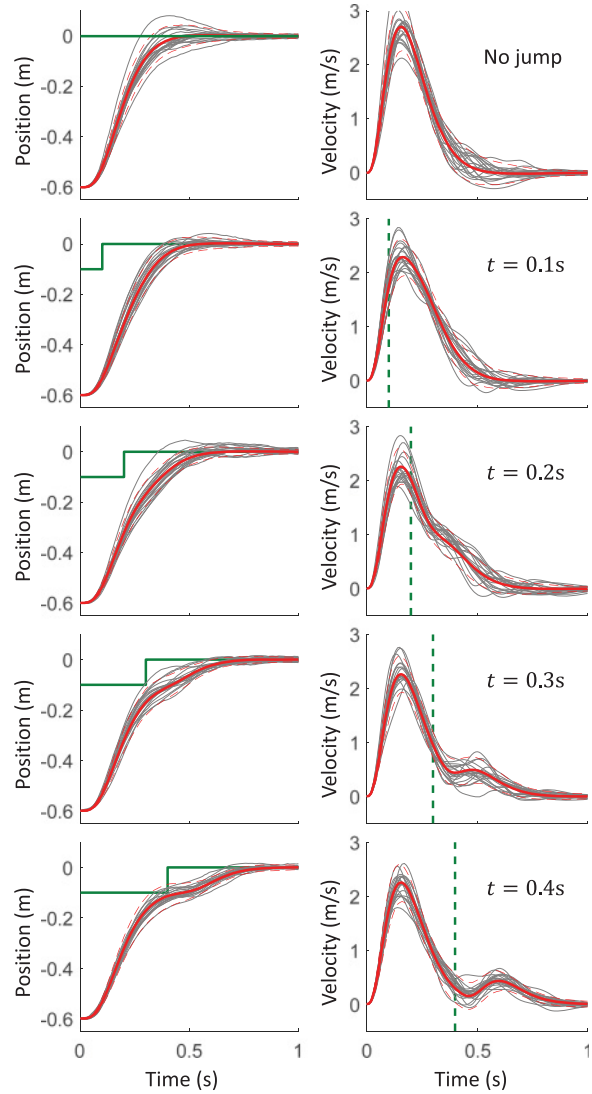
Figure 4: Simulations of target jump to a far position. The target was initially 0.5 m away from the hand, but jumped farther away from the hand by 0.1 m at various times to make a total distance of 0.6 m. The top row shows results for the no-jump condition with a distance of 0.6 m. The remaining rows show results for jumps at $t = 0.1, 0.2, 0.3, 0.4$ s after the start of reaching. The two columns show the position trajectories and velocity profiles, respectively. For the position plots, the solid green lines indicate the target positions with the steps at the jump times. For the velocity plots, the dashed vertical green lines indicate the same jump times.
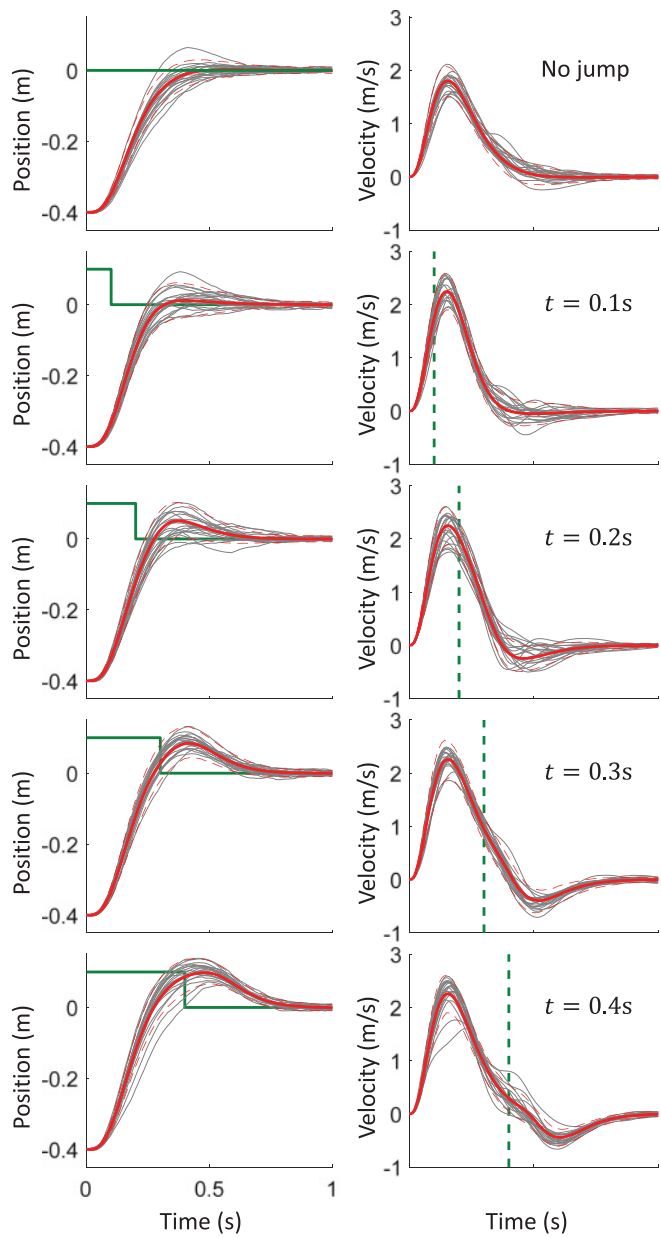
Figure 5: Simulations of target jump to a near position. The target was initially 0.5 m away from the hand but jumped closer to the hand by 0.1 m at various times to make a total distance of 0.4 m. The format was identical to that of Figure 4.

regular and showed a plateau or second peak. Importantly, all simulations in Figures 4 and 5 were produced by a single policy with fixed estimator K and control law L. In all cases, the model was run continuously, without any prefixed or precomputed duration, to move the hand toward the target regardless of whether, when, and where the target jumped. At any time, the estimator combines sensory feedback with internal prediction to produce estimated state vector for guiding control. To our knowledge, this is the first computational demonstration that a single feedback control policy with fixed parameters simulates different kinematic features of target-perturbation experiments.

In a variation of the target-perturbation paradigm (Goodale et al., 1986), subjects did not see their hands during the movement. In addition, the target jumped when the subjects' saccadic eye movement to the target was at peak speed (around the start of reaching). Consequently, the subjects were unaware of the jump because of transient saccadic suppression of vision. It was found that subjects corrected for the jump just fine even though they were unaware of it and that the movement kinematics with and without the jump were similar. We modeled the invisible hand by setting $K_1 = 0$. Since our model represents the hand state relative to the target whose visual feedback was also suppressed around peak saccadic speed, proprioception alone could not determine the state during saccadic suppression. We therefore modeled the unaware condition by also setting $K_2 = 0$ for 50 ms and letting the target jump at the midpoint of this time window. When both K_1 and K_2 were 0, the first term of equation 2.7, which represents internal prediction, still estimated the state, effectively providing a memory mechanism. The results of our simulations for movements with and without target jump are shown in Figure 6. The model produced similar position and velocity profiles for the two conditions and corrected for the jump occurred during the 50 ms blind window, reproducing the experimental findings. Two factors contributed to the results. First, although the target jumped when sensory feedbacks were suppressed, the new target location was observed *after* the saccade, which guided subsequent control. Second, as noted above, earlier sensory feedback was not as important as later feedback because of Weber's law. The similarity between the jump and no-jump conditions was due to the fact that the jump occurred at the start of the hand movement (Goodale et al., 1986), regardless of the hand visibility for the rest of the reaching. Indeed, the jump and no-jump conditions were still similar when we simulated the visible hand condition (results not shown).

2.3 Submovement Decomposition. We discussed in section 1 the debate on whether the number of decomposed submovements could determine the number of discrete motor impulses or commands for movement control. We investigated this issue by applying a submovement decomposition method (Lee et al., 1997) to the velocity profiles of sample trials produced by our model for target-perturbation experiments (see Figures 4
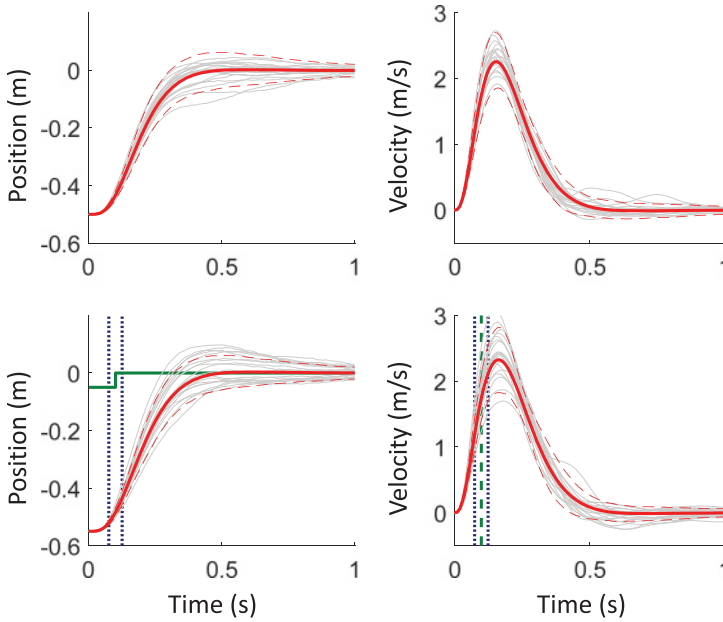
Figure 6: Simulations of target jump during saccadic suppression of vision. The target was initially 0.5 m away from the hand. The top row shows the position and velocity traces for the no-jump condition. The bottom row shows the results for the case where the target jumped away from the hand by 0.05 m in a 50 ms time window (marked by dotted vertical lines) during which the vision was suppressed. The jump and no-jump conditions produced similar results (in agreement with the experiment) because the jump time was early.

and 5). Three examples of decomposition for each of various jump conditions are shown in Figure 7. By applying the method to 1000 velocity profiles from each condition, we obtained the histograms in Figure 8. Similar to the results of decomposing observed velocity profiles (Fishbach et al., 2005, 2007), we found with our simulations that no-jump and early-jump conditions produced fewer numbers of decomposed submovements than those of the late-jump conditions. For a given jump condition, different trials could produce different numbers of submovements because the sampled noises (from the same distributions) fluctuated. For a given trial, the number of submovements depends on the shape of the standard profile and the criterion of fit. With symmetric standard profile (Lee et al., 1997) and a criterion of accounting for 99% variance, we found that all velocity profiles we tried, including those from the no-jump condition, produced multiple submovements. Importantly, the velocity profiles were all simulated with exactly the same single control policy run continuously in time. Therefore,
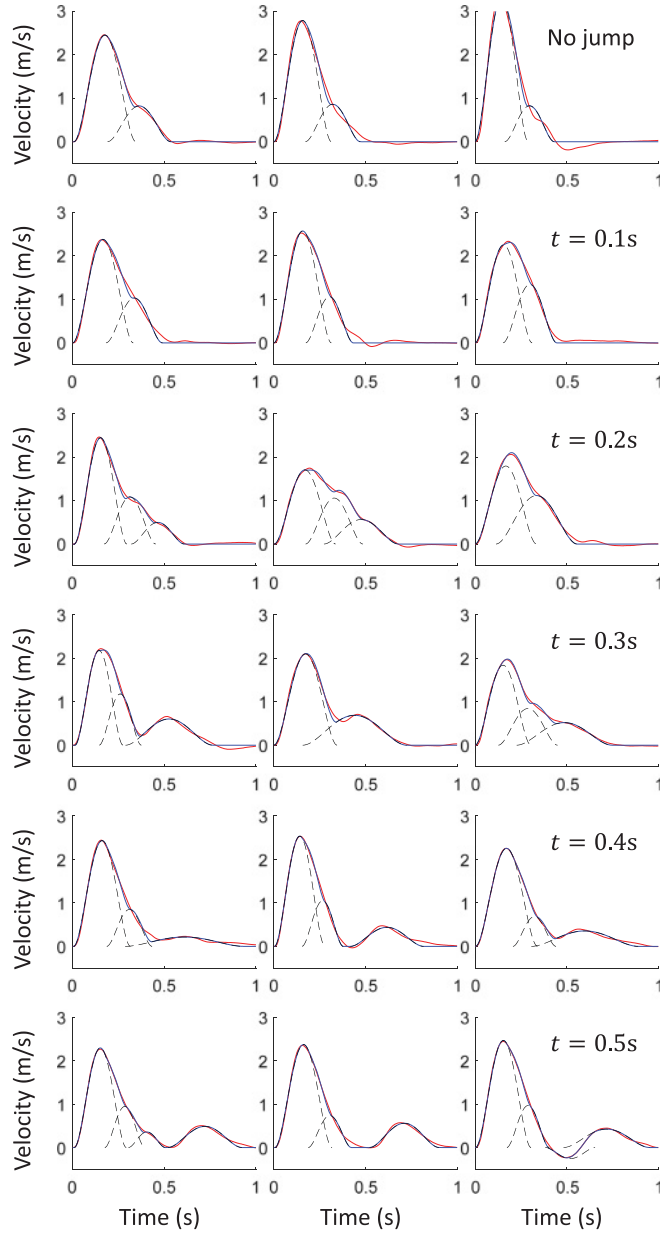
Figure 7: Submovement decomposition of simulated velocity profiles. Three trials from each jump condition in Figure 4 are decomposed using the methods of Lee et al. (1997) with the criterion of accounting for 99% variance. Dashed blue curves indicate submovements. Solid blue and red curves indicate the fitted and originally simulated velocity profiles, respectively.
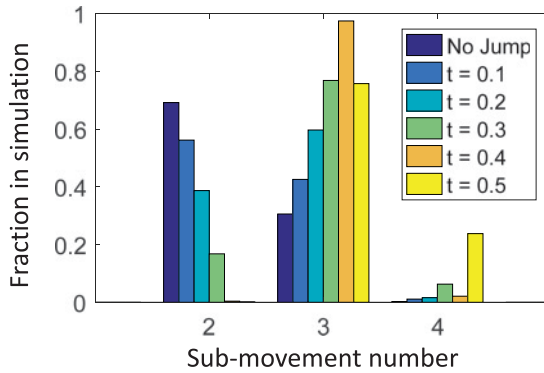
Figure 8: Distributions of numbers of decomposed submovements. The frequency histograms of the numbers of decomposed submovements for the various jump conditions in Figure 4 are shown. The frequency of each condition was estimated from 1000 trials of simulations. Profiles from later-jump conditions required more submovements to fit than those from no-jump or early-jump conditions.

our results suggest that the number of decomposed submovements does not necessarily reflect the number of discrete motor impulses, commands, or policies.

A related experimental finding is that velocity profiles of slower movements require more submovements to fit (Morasso, Ivaldi, & Ruggiero, 1983; Darling, Cole, & Abbs, 1988; van der Wel, Sternad, & Rosenbaum, 2009; Levy-Tzedek, Krebs, Song, Hogan, & Poizner, 2010; Shmuelof, Krakauer, & Mazzoni, 2012). To explain this finding, we used the same criterion for Figure 9 to define the duration of a trial and sort the simulated trials of the no-jump condition from fast to slow. We then divided the trials into fast, medium, and slow groups according to whether the duration is less than 0.5 s, between 0.5 and 0.6 s, and greater than 0.6 s, and determined the submovement decomposition for each group. The results in Figure 9 show that our single-policy model reproduced the experimental finding that slower trials required more submovements to fit. Again, the conclusion is that the number of submovements does not necessarily reflect the number of underlying motor impulses, commands, or control policies.

2.4 Two-Dimensional Target Perturbation. In the above, we modeled one-dimensional target perturbation in which the target jumped parallel to the movement path to become either closer to or farther away from the hand. We also used a two-dimensional motor plant described in Liu and Todorov (2007) to simulate target jumps perpendicular to the initial movement path. Figure 10 shows example simulations in which the hand started
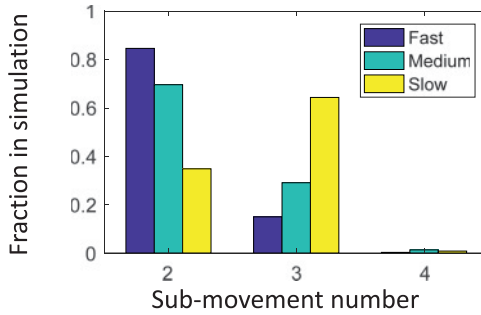
Figure 9: The dependence of submovement number on movement duration. The simulated trials of the no-jump condition of Figure 4 were divided into fast, medium, and slow groups according to whether the movement duration was less than 0.5 s, between 0.5 and 0.6 s, or greater than 0.6 s. The submovement decomposition was done for each group separately to produce the frequency histogram.

moving horizontally to a target 0.3 m to the right, which then jumped perpendicular to the initial movement path by 0.05 m in either direction. The jump time was 100, 200, or 300 ms after the movement onset. Since in the experiment of Liu and Todorov (2007), subjects did not see their hands during reaching, we set the vision part of the Kalman matrix $K_1 = 0$ as before. We averaged movement trajectories over 4000 trials in order to compare them with the reported, average results (Liu & Todorov, 2007). Consistent with experimental observations, the simulated average hand trajectories turned gradually toward the new target locations, with sharper turns for later jumps (see Figure 10a), the average speed profiles became more irregular for later jumps (see Figure 10b), and the movement duration was longer for later jumps (see Figure 10c). Following Liu and Todorov (2007), the terminations of the curves in Figure 10c indicate movement durations, and the termination criterion was the time when average hand velocity was smaller than 0.5 cm/s for 40 ms. (Note that this criterion was for the average velocity across trials and different from the one for individual trials used above.) As we mentioned in section 1, Liu and Todorov (2007) simulated these perturbations with a finite-horizon model and had to preset two movement durations to compute two control policies for each jump time. We used the same motor plant, but because of the infinite-horizon formulation, we were able to apply a single policy for all the simulations in Figure 10 without presetting any movement duration.

3 Discussion

In this study, we first extended an infinite-horizon stochastic optimal feedback control model (Phillis, 1985; Jiang et al., 2011; Qian et al., 2013) by
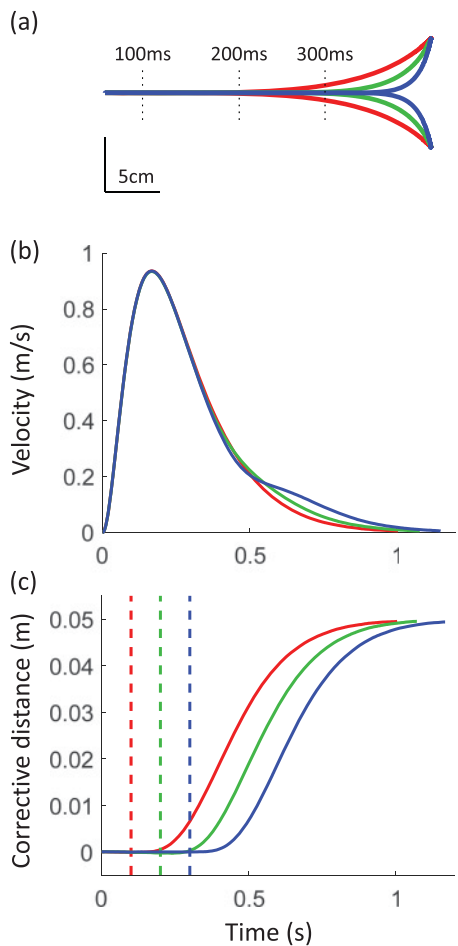
Figure 10: Simulations of two-dimensional target perturbation. The hand started moving horizontally to a target 0.3 m to the right, which then jumped perpendicular to the initial movement path by 0.05 m in either direction. The jump time was 100, 200, or 300 ms after the movement onset. (a) The average movement trajectories. The red, green, and blue curves represent jumps at 100, 200, and 300 ms, respectively. (b) The average velocity profiles for the three jump times. (c) The average corrective distance along the perturbation axis (perpendicular to the initial movement direction) for the three jump times.

incorporating Weber's law of sensory perception. We were able to achieve this by adding state-dependent noise to the observation equation because the model represents the effector state relative to the target. Models that use separate, absolute representations of effector and target in the state vector

would need to compare the effector and target states to realize Weber's law. We also considered visual and proprioceptive feedback with different delays and noise levels. We analytically derived an optimal solution for the extended model. We then applied the model to show that terminal vision was more effective in reducing end-point variability than initial vision, and these two partial vision conditions produced end-point accuracies similar to those of the full-vision and no-vision conditions, respectively, and that the differences among the conditions became smaller for slower movements. These simulation results explain Beaubaton and Hay's findings coherently without assuming a switch between crude and fine or between open- and closed-loop strategies. We also applied the same model (and parameters) to simulate target-perturbation experiments for various times of target jump. Consistent with experimental findings (Georgopoulos et al., 1981; Soechting & Lacquaniti, 1983; Flash & Henis, 1991; Heath et al., 1998; Goodale et al., 1986; Pélisson et al., 1986; Fishbach et al., 2005, 2007), the model produced smooth, single-peaked mean velocity profiles when the target jumped early and irregular mean profiles with a plateau or a second peak when the target jumped late. Additionally, the model explained the observations that subjects can perform the task even when they do not see the jump and when their hands are invisible during reaching (Goodale et al., 1986). Moreover, we found that the simulated velocity profiles, including those from the no-jump condition, appeared to have multiple submovements, and the number of submovements was larger for trials that were slower or perturbed later. Importantly, all of these simulations were produced by a single, fixed control policy. The model did not assume different mechanisms at different periods of movements or for different jump times and locations nor require different numbers of policies for movements of different durations. Finally, we simulated two-dimensional target perturbations in which the target jumped perpendicular to the initial movement direction. Here again we used a single policy to simulate different jump times and directions.

Our study demonstrates for the first time that a single control policy can explain a broad range of kinematic features observed in different experimental conditions and trials. This demonstration offers a more parsimonious interpretation of the observations. Thus, ineffective initial visual feedback does not necessarily imply open-loop control at the start of movements. Indeed, in the extreme case of no sensory information about the initial state vector at all, it would be impossible to decide in which direction to move the hand. Likewise, different kinematic features for different perturbation times or different movement durations do not necessarily imply different control strategies, and the number of peaks or decomposed submovements in a velocity profile does not necessarily reflect the number of motor impulses or commands or policies. Instead, our work suggests that variations in experimental conditions might naturally lead to variations in kinematics without changes in underlying control mechanisms or policies. For example, the jump time determines when the large perturbation error

occurs during a movement, and that could be sufficient to account for the jump-time dependence of kinematics. Similarly, different trials have different samples of noise sequences, and this fluctuation could be sufficient to explain trial-by-trial variation of kinematics.

Most optimization models for biological motor control use a finite-horizon formulation (Flash & Hogan, 1985; Uno, Kawato, & Suzuki, 1989; Harris & Wolpert, 1998; Todorov & Jordan, 2002; Scott, 2004; Todorov, 2004; Diedrichsen, Shadmehr, & Ivry, 2010). Because they require movement duration to define the cost functions, these models have to be given the duration before solving the optimization problems and initiating movements. To avoid this problem, we and others have recently advocated infinite-horizon formulations (Huh et al., 2010; Jiang et al., 2011; Qian et al., 2013), which can predict movement duration instead of prefixing it. The contrast between finite- and infinite-horizon formulations is particularly striking for modeling target-perturbation experiments. Because target jump changes movement duration, finite-horizon models have to be given two preset durations—one for the original target location and the other for the final target location. Moreover, the second duration depends on when and where the target jumps and the hand state at the time of jump, and finite-horizon models have to be given different second durations for different situations. There is also a model that uses both infinite- and finite-horizon approaches, with the former determining the movement duration and the latter specifying a control policy for that duration (Rigoux & Guigon, 2012). Because the model updates the duration and the corresponding policy at each time step, its number of required finite-horizon policies equals the number of time steps and changes with different movement durations for different conditions. We showed in this letter that our purely infinite-horizon approach provides a simpler explanation of target-perturbation experiments. Our model does not need to be given or compute movement durations explicitly, and it uses a single policy to move the hand toward the target regardless of whether, when, or where the target jumps or whether the jump is visible.

Computational modeling shows only theoretical possibilities. Whether stochastic optimal feedback control models or submovement decomposition models better account for biological motor control is an empirical question. Similarly, only experiments can help differentiate the validity of infinite-horizon models presented here and elsewhere (Huh et al., 2010; Jiang et al., 2011; Qian et al., 2013), the finite-horizon models (Harris & Wolpert, 1998; Todorov & Jordan, 2002; Scott, 2004; Tanaka, Krakauer, & Qian, 2006), and the mixed formulation (Rigoux & Guigon, 2012). We previously discussed the similarities and differences, and the strengths and weaknesses, of finite- and infinite-horizon models and speculated that they both might be employed by biological systems (Qian et al., 2013). Specifically, parsimonious, infinite-horizon policies might provide default control mechanisms, whereas finite-horizon policies might be called for when

additional optimization and the associated computational cost are justified. Our study is broadly consistent with this speculation: it shows that an infinite-horizon model with a single policy is able to perform a variety of tasks, but a finite-horizon model with multiple policies might help refine these tasks (e.g., by forcing their completion in a prefixed time frame with less trial-by-trial variation of movement duration).

We already noted that because our model represents the effector state relative to the target, we can implement Weber's law easily by introducing state-dependent noise in the observation equations. Another consequence of this relative, hand-centered representation is that the control signal and the corresponding force vanish as the effector approaches the target even for constant control-law matrix L (see equation 2.8). (If force is needed to maintain the effector, one can simply add a component to the state vector or a term in equation 2.8.) In contrast, other models often use a state vector that represents the effector and target separately with respect to a reference frame not centered on the hand (e.g., Todorov, 2005). With such absolute representation, the state vector is not zero even when the effector reaches the target. Consequently, L has to vary with time or state and drop to zero to eliminate the control signal (and the force) when the hand reaches the target. Obviously, different choices of the reference frame produce different absolute representations of the effector and target and, consequently, different L dependence on time or state. Experimentally measured force gains do change with time or state (Liu & Todorov, 2007; Dimitriou, Wolpert, & Franklin, 2013), although it is unclear which absolute representation they support. These and other findings suggest that the brain may use both absolute and relative state representations (Graziano, 2001; Dimitriou et al., 2013). How to determine the reference frame for the absolute-representation component is an open question for future investigation.

Our explanation of Beaubaton and Hay's (1986) experiments relies on Weber's law and is not identical to that based on the difference between foveola and peripheral vision. The fovea spans only about 1 degree of visual angle, roughly the size of the thumbnail at arm's length. In contrast, hand movements in many experiments and in daily life can often span visual angles of 10 degrees or more so that during most parts of such movements, the eyes can foveate only the hand or the target but not both. Therefore, foveal vision of both hand and target is available only when the hand has almost reached the target.

Our model makes testable predictions. First, in Figures 1 and 2, although toward the end of the movements, the initial- and no-vision conditions were similar and the terminal- and full-vision conditions were similar, the model predicts the opposite before the final convergence on the target: there is greater positional variance in the terminal- and no-vision conditions than in the initial- and full-vision conditions. Second, a closely related prediction is that the time to reach the midpoint of the position trajectory has much larger variance in the terminal- and no-vision conditions than in the initial- and

full-vision conditions. Confirmation of these predictions will be further evidence for the effects of visual feedback even at the beginning of movements. Third, although initial feedback is ineffective in reducing final positional variance, it should still be useful for crude aspects of movements such as general direction. The experiment of Bard, Hay, and Fleury (1985) seems to support this prediction. Finally, in a previous paper (Qian et al., 2013) we compared different predictions from the finite- and infinite-horizon control models (e.g., the absence versus presence of trials in which the hand overshoots the target). This study suggests another test that could potentially distinguish between the models. As we mentioned in section 1, the finite-horizon models require two movement durations to compute two control policies for use before and after the target jump. If the target jumps at an unpredictable time or to an unpredictable location, it is impossible to know the second movement duration before the jump. If, in addition, the jump is undetected because of saccadic suppression of vision during the jump (Goodale et al., 1986; Pélisson et al., 1986), then the finite-horizon models would not know when to reset the movement duration or switch the control policy. In contrast, according to our infinite-horizon model, movement duration is not preset but an emergent property that depends on when and where the target jumps regardless of whether the jump moment is detected. Therefore, for an unpredictable and undetected target jump, our infinite-horizon model predicts a change of duration while the finite-horizon models do not.

In summary, we have proposed an infinite-horizon stochastic optimal feedback control model for understanding target-perturbation experiments. Because we incorporated Weber's law, the model also explains differences between initial and terminal visual feedback. Our model can reproduce different kinematic features under various conditions with a single control policy and suggests a parsimonious new interpretation of relevant experimental findings.

Appendix: Model Details and Derivation

As noted in the text, we extended the formulation of Phillis (1985) and Qian et al. (2013) in two ways. First, we added the Z term to each feedback equation to implement Weber's law of sensory perception. Second, we included multiple sensory feedbacks with different delays and noise levels to represent, for example, differences between vision and proprioception. In equations 2.1 and 2.2, x is the state n-vector, u is the control m-vector, and each y_s is a sensory-feedback k-vector. A, B, C_s, D_s, F, G, Y, and Z_s are matrices of proper sizes. A and B are determined by the motor plant and Newtonian mechanics (see the examples that follow). More explanations can be found in Qian et al. (2013). We transformed these equations and the delay equation 2.3 into equations 2.5 and 2.6 by defining the new state x^* and observation y^* (see equation 2.4), which, for two sensory feedbacks, are $3n$- and

2k-vectors, respectively. The procedure can be extended to more than two feedbacks. Equations 2.7 and 2.8 are standard estimation and control-law equations, where $\hat{x}^*$ is also 3n-vector.

The cost function defined in equation 2.9 contains three terms, at the steady state. $\langle \tilde{x}^T U \tilde{x} \rangle$ represents estimation error, $\langle x^T Q x \rangle$ measures the movement performance since x is the effector state relative to the target, and $\langle u^T R u \rangle$ is the effort or energetic cost. U, Q, and R are constant matrices that are positive definite or semidefinite (Qian et al., 2013). Equation 2.9 can be expressed in x^* and $\tilde{x}^* \equiv x^* - \hat{x}^*$ as

$$J = \lim_{t \to \infty} \langle \tilde{x}^{*T} U^* \tilde{x}^* + x^{*T} Q^* x^* + u^T R u \rangle, \tag{A.1}$$

where

$$U^* \equiv \begin{bmatrix} U \\ & 0 \\ & & 0 \end{bmatrix}, \quad Q^* \equiv \begin{bmatrix} Q \\ & 0 \\ & & 0 \end{bmatrix}.$$

The problem is to find Kalman-estimator matrix K and control-law matrix L that jointly optimize the cost J under the constraints of equations 2.5 to 2.8. To solve the problem, we adapted Phillis's approach by first defining

$$X \equiv \begin{bmatrix} x^* \\ \tilde{x}^* \end{bmatrix}, \quad d\bar{\omega} \equiv \begin{bmatrix} d\omega^* \\ d\xi^* \end{bmatrix}, \quad \bar{G} \equiv \begin{bmatrix} G^* & 0 \\ G^* & -KD^* \end{bmatrix},$$

$$\bar{A} \equiv \begin{bmatrix} A^* - B^* L & B^* L \\ 0 & A^* - KC^* \end{bmatrix}, \quad \bar{F} \equiv \begin{bmatrix} F^* & 0 \\ F^* & 0 \end{bmatrix},$$

$$\bar{Y} \equiv \begin{bmatrix} -Y^* L & Y^* L \\ -Y^* L & Y^* L \end{bmatrix}, \quad \bar{Z} \equiv \begin{bmatrix} 0 & 0 \\ -KZ^* & 0 \end{bmatrix},$$

where X is a 6n-vector termed extended-state vector, and then combining equations 2.5 to 2.8 as

$$dX = \bar{A} X dt + \bar{F} X d\beta + \bar{Y} X d\gamma + \bar{Z} X d\delta + \bar{G} d\bar{\omega}. \tag{A.2}$$

The $\bar{Z}$ term stems from the state-dependent noise terms Z_s not present in the original formulation (Phillis, 1985; Qian et al., 2013).

With further definitions,

$$P \equiv \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \equiv \langle XX^T \rangle, \tag{A.3}$$

$$V \equiv \begin{bmatrix} Q^* + L^T R L & -L^T R L \\ -L^T R L & L^T R L + U \end{bmatrix}, \tag{A.4}$$

where P is the correlation matrix of the extended state X and the subscripts refer to the four block matrices of $3n \times 3n$ each, the cost, equation A.1 becomes

$$J = \lim_{t \to \infty} \langle X^\mathrm{T} V X \rangle = \lim_{t \to \infty} \langle \mathrm{tr}(X^\mathrm{T} V X) \rangle = \lim_{t \to \infty} \langle \mathrm{tr}(V X X^\mathrm{T}) \rangle$$

$$= \lim_{t \to \infty} \mathrm{tr}(V P). \tag{A.5}$$

We next converted the stochastic differential, equation A.2, into a deterministic one. First, equation A.2 has the formal solution (Tanaka et al., 2006)

$$X(t) = e^{\int_0^t \bar{A}(t_e)\mathrm{d}t_e} X(0) + \int_0^t e^{\int_{t'}^t \bar{A}(t_e)\mathrm{d}t_e} \left[\left(\bar{F} X \mathrm{d}\beta \right)\big|_{t'} \right.$$

$$\left. + \left(\bar{Y} X \mathrm{d}\gamma \right)\big|_{t'} + \left(\bar{Z} X \mathrm{d}\delta \right)\big|_{t'} + \left(\bar{G} \mathrm{d}\bar{\omega} \right)\big|_{t'} \right], \tag{A.6}$$

for the general case that $\bar{A}$ may be a function of t. Then, using the definition of P, we have

$$P(t) = \langle e^{\int_0^t \bar{A}(t_e)\mathrm{d}t_e} X(0) X(0)^\mathrm{T} e^{\int_0^t \bar{A}^\mathrm{T}(t_e)\mathrm{d}t_e} \rangle$$

$$+ \int_0^t \int_0^t e^{\int_{t'}^t \bar{A}(t_e)\mathrm{d}t_e} \langle \left(\bar{F} X \mathrm{d}\beta \right)\big|_{t'} \left(\mathrm{d}\beta X^\mathrm{T} \bar{F}^\mathrm{T} \right)\big|_{t''}$$

$$+ \left(\bar{Y} X \mathrm{d}\gamma \right)\big|_{t'} \left(\mathrm{d}\gamma X^\mathrm{T} \bar{Y}^\mathrm{T} \right)\big|_{t''} + \left(\bar{Z} X \mathrm{d}\delta \right)\big|_{t'} \left(\mathrm{d}\delta X^\mathrm{T} \bar{Z}^\mathrm{T} \right)\big|_{t''}$$

$$+ \left(\bar{G} \mathrm{d}\bar{\omega} \right)\big|_{t'} \left(\mathrm{d}\bar{\omega}^\mathrm{T} \bar{G}^\mathrm{T} \right)\big|_{t''} \rangle e^{\int_{t''}^t \bar{A}^\mathrm{T}(t_e)\mathrm{d}t_e}. \tag{A.7}$$

The double stochastic integral can be divided into three parts: $t' > t''$, $t' < t''$, and $t' = t''$. Because all the random processes are assumed to be independent, the first two parts are zero. For example, consider $t' > t''$ and the first expectation term $\langle \left(\bar{F} X \mathrm{d}\beta \right)\big|_{t'} \left(\mathrm{d}\beta X^\mathrm{T} \bar{F}^\mathrm{T} \right)\big|_{t''} \rangle$. Because $\mathrm{d}\beta|_{t'}$ cannot affect X at either the same time t' or the previous time t'', it can be factored out, and its expectation is zero. For $t' < t''$, $\mathrm{d}\beta|_{t''}$ can be factored out and has zero expectation. Similar arguments apply to the other noise terms of equation A.7. The $t' = t''$ part of the integral is along the diagonal line of the (t', t'') plane and is therefore a single integral. Since for standardized Wiener process β, $\mathrm{d}\beta$ is gaussian white noise with zero mean and variance $\mathrm{d}t$, we have $\langle \mathrm{d}\beta^2 \rangle = \mathrm{d}t$. Equation A.7 then becomes

$$P(t) = e^{\int_0^t \bar{A}(t_e)\mathrm{d}t_e} P(0) e^{\int_0^t \bar{A}^\mathrm{T}(t_e)\mathrm{d}t_e}$$

$$+ \int_0^t e^{\int_{t'}^t \bar{A}(t_e)\mathrm{d}t_e} (\bar{F} P \bar{F}^\mathrm{T} + \bar{Y} P \bar{Y}^\mathrm{T} + \bar{Z} P \bar{Z}^\mathrm{T} + \bar{G} \bar{G}^\mathrm{T}) e^{\int_{t'}^t \bar{A}^\mathrm{T}(t_e)\mathrm{d}t_e} \mathrm{d}t, \tag{A.8}$$

which is a deterministic equation. Finally, differentiating this equation produces

$$\dot{P} = \bar{A}P + P\bar{A}^{\mathrm{T}} + \bar{F}P\bar{F}^{\mathrm{T}} + \tilde{Y}P\tilde{Y}^{\mathrm{T}} + \tilde{Z}P\tilde{Z}^{\mathrm{T}} + \bar{G}\bar{G}^{\mathrm{T}}. \tag{A.9}$$

At steady state, we have $\dot{P} = 0$, which is equation 2.13 in the text.

The problem now becomes optimizing the cost, equation A.5, under the constraint of $\dot{P} = 0$. Similar to Phillis (1985), we found the solution, equation 2.10 to 2.13, via the Lagrange multiplier method. The augmented cost function is

$$J_a = J + \sum_{ij} \dot{P}_{ij}S_{ij} = J + \mathrm{tr}(\dot{P}S), \tag{A.10}$$

where S is the matrix of Lagrange multipliers. Since P is symmetric, S is also symmetric, so that each unique constraint has a unique multiplier. Additionally, standard Kalman filter has the property that the estimate $\hat{x}^*$ is orthogonal to the estimation error $\tilde{x}^*$; this implies that $P_{12} = P_{22}$ and the P_{12} constraints are not new. Therefore, S has the form

$$S \equiv \begin{bmatrix} S_{11} & 0 \\ 0 & S_{22} \end{bmatrix}. \tag{A.11}$$

By setting $\partial J_a/\partial K = 0$, $\partial J_a/\partial L = 0$, and $\partial J_a/\partial P = 0$ and using the identity $\partial(AXBX^T)/\partial X = A^T XB^T + AXB$ for any A, B, X matrices of proper dimensions, we obtained the solution equations 2.10 to 2.12. For a given motor plant and parameter set, optimal K and L can found numerically by iterating equations 2.10 to 2.13 to convergence (Jiang et al., 2011; Qian et al., 2013).

We used two motor plants for the simulations in this letter. The first was a single-joint arm described in our previous studies (Tanaka et al., 2006; Qian et al., 2013). The state vector x contained angular position, velocity, acceleration, and jerk of the hand as components:

$$x = \begin{bmatrix} \theta \\ \dot{\theta} \\ \ddot{\theta} \\ \dddot{\theta} \end{bmatrix}. \tag{A.12}$$

The system followed second-order Newtonian dynamics and a second-order equation relating neural control signal u to muscle force (Winters & Stark, 1985), leading to

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -\alpha_1 & -\alpha_2 & -\alpha_3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ b_u \end{bmatrix}, \tag{A.13}$$

in equation 2.1, where

$$\alpha_1 = \frac{b}{t_a t_e I}, \ \alpha_2 = \frac{1}{t_a t_e} + \left(\frac{1}{t_a} + \frac{1}{t_e}\right)\frac{b}{I},$$

$$\alpha_3 = \frac{b}{I} + \frac{1}{t_a} + \frac{1}{t_e}, \ b_u = \frac{1}{t_a t_e I}, \tag{A.14}$$

with the arm's moment of inertia $I = 0.25$ kg m^2, viscosity $b = 0.2$ kg m^2/s, and muscle activation and excitation time constants $t_a = 0.03$ s and $t_e = 0.04$ s. For the figures, we converted elbow angle θ to hand position along the movement path by multiplying θ with the forearm length ($L_0 = 0.35$ m). The matrices for the noise terms in equation 2.1 were $F = 0$, $Y = 0.02B$, and $G = 0.03I_4$, where I_4 is the 4×4 identity matrix. The observation matrices were

$$C_s = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

for $s = 1, 2$. The fourth component of the state vector was assumed to be unobservable (Todorov, 2004; Qian et al., 2013) as there is no evidence that the sensory systems can perceive the third-order temporal derivative of position. Following Crevecoeur et al. (2016), we assumed that visual feedback was more delayed but less noisy than proprioceptive feedback and set $\tau_1 = 100$ ms, $\tau_2 = 50$ ms, $Z_2 = 2.5Z_1$, and $D_2 = 2.5D_1$. As noted in the text, we set diagonal entries of Z_1 as proportional to Weber fractions of visual perception of position, velocity, and acceleration reported in the literature (Baird, 1970; McKee et al., 1990; Levi & Klein, 1992; Mueller & Timney, 2016):

$$Z_1 = \frac{1}{2.5} \begin{bmatrix} 0.03 & 0 & 0 & 0 \\ 0 & 0.06 & 0 & 0 \\ 0 & 0 & 0.3 & 0 \end{bmatrix}.$$

D_1 can be any matrix of small numbers, and we let

$$D_1 = \frac{1}{2.5} \begin{bmatrix} 0.0005 & 0 & 0 \\ 0 & 0.005 & 0 \\ 0 & 0 & 0.025 \end{bmatrix}.$$

The matrices in the cost function were $Q = \mathrm{diag}(1, 0.01, 0, 0)$, $R = 0.0001$, and $U = \mathrm{diag}(1, 0.1, 0.01, 0)$.

The second motor plant, taken from Liu and Todorov (2007), was a point-mass effector on a plane driven by a two-dimensional neural control signal. The system followed second-order Newtonian dynamics and a first-order equation relating the control signal to muscle force. The state vector x

contains the horizontal and vertical components of the position, velocity, and force:

$$
x = \begin{bmatrix} p_x \\ p_y \\ v_x \\ v_y \\ f_x \\ f_y \end{bmatrix}.
\tag{A.15}
$$

Importantly, we represented these quantities relative to the target. The system dynamics was characterized by

$$
A = \begin{bmatrix} 0 & 1 & & & & \\ & 0 & & 1 & & \\ & & -b/m & & 1/m & \\ & & & -b/m & & 1/m \\ & & & & -1/\tau & \\ & & & & & -1/\tau \end{bmatrix},\quad
B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1/\tau & 0 \\ 0 & 1/\tau \end{bmatrix},
\tag{A.16}
$$

where the mass $m = 1$ kg, viscosity $b = 10$ Ns/m, and muscle time constant $\tau = 50$ ms (Liu & Todorov, 2007). Noise terms in equation 2.1 were $F = 0$, $G = \mathrm{diag}(0.1, 0.1, 1, 1, 0, 0)$. The control-dependent noise had both horizontal and vertical components $Y_x u d\gamma_x$ and $Y_y u d\gamma_y$ to model noise in two-dimensional control signal (Todorov & Jordan, 2002).

$$
Y_x = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 5 & 0 \\ 0 & 5 \end{bmatrix},\quad
Y_y = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 5 \\ -5 & 0 \end{bmatrix}.
$$

In equation 2.6, we let $C = I_6$, $2.5Z_1 = Z_2 = \mathrm{diag}(0.03, 0.03, 0.06, 0.06, 0.3, 0.3)$, $2.5D_1 = D_2 = \mathrm{diag}(0.001, 0.001, 0.01, 0.01, 0.1, 0.1)$. Matrices in cost function were $Q = \mathrm{diag}(0.02, 0.02, 0.2, 0.2, 0, 0)$, $R = 0.0002I_2$ and $U = \mathrm{diag}(0.25, 0.25, 0.025, 0.025, 0.0025, 0.0025)$.

To model target jump to a new position, we modified the relevant position component of x while keeping other components unchanged. Movement is simulated using the Euler method with time step $dt = 1$ ms.

Acknowledgments

References

Baird, J. C. (1970). *Psychophysical analysis of visual space*. New York: Pergamon Press.

Bard, C., Hay, L., & Fleury, M. (1985). Role of peripheral vision in the directional control of rapid aiming movements. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *39*(1), 151–161.

Beaubaton, D., & Hay, L. (1986). Contribution of visual information to feedforward and feedback processes in rapid pointing movements. *Human Movement Science*, *5*(1), 19–34.

Bernstein, N. (1967). *The co-ordination and Regulation of Movements*. Oxford: Pergamon.

Crevecoeur, F., Munoz, D. P., & Scott, S. H. (2016). Dynamic multisensory integration: Somatosensory speed trumps visual accuracy during feedback control. *Journal of Neuroscience*, *36*(33), 8598–8611.

Darling, W. G., Cole, K. J., & Abbs, J. H. (1988). Kinematic variability of grasp movements as a function of practice and movement speed. *Experimental Brain Research*, *73*(2), 225–235.

Desmurget, M., & Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Sciences*, *4*(11), 423–431.

Diedrichsen, J., Shadmehr, R., & Ivry, R. B. (2010). The coordination of movement: Optimal feedback control and beyond. *Trends in Cognitive Sciences*, *14*(1), 31–39.

Dimitriou, M., Wolpert, D. M., & Franklin, D. W. (2013). The temporal evolution of feedback gains rapidly update to task demands. *Journal of Neuroscience*, *33*(26), 10898–10909.

Egelhaaf, M., & Reichardt, W. (1987). Dynamic response properties of movement detectors: Theoretical analysis and electrophysiological investigation in the visual system of the fly. *Biological Cybernetics*, *56*(2), 69–87.

Fangwen, Z., Zhaoping, L., & Qian, N. (2012). Infinite-horizon optimal feedback control models for biological systems: Application to target jump. Cosyne abstracts, Salt Lake City.

Fishbach, A., Roy, S. A., Bastianen, C., Miller, L. E., & Houk, J. C. (2005). Kinematic properties of on-line error corrections in the monkey. *Experimental Brain Research*, *164*(4), 442–457.

Fishbach, A., Roy, S. A., Bastianen, C., Miller, L. E., & Houk, J. C. (2007). Deciding when and how to correct a movement: Discrete submovements as a decision making process. *Experimental Brain Research*, *177*(1), 45–63.

Flash, T., & Henis, E. (1991). Arm trajectory modifications during reaching towards visual targets. *Journal of Cognitive Neuroscience*, *3*(3), 220–230.

Flash, T., & Hogan, N. (1985). The coordination of arm movements: An experimentally confirmed mathematical model. *Journal of Neuroscience*, *5*(7), 1688–1703.

Georgopoulos, A. P., Kalaska, J. F., & Massey, J. T. (1981). Spatial trajectories and reaction times of aimed movements: Effects of practice, uncertainty, and change in target location. *Journal of Neurophysiology*, *46*(4), 725–743.

Goodale, M. A., Pélisson, D., & Prablanc, C. (1986). Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature*, *320*(6064), 748–750.

Graziano, M. (2001). Is reaching eye-centered, body-centered, hand-centered, or a combination? *Reviews in the Neurosciences*, *12*(2), 175–185.

Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, *394*(6695), 780–784.

Heath, M., Hodges, N. J., Chua, R., & Elliott, D. (1998). On-line control of rapid aiming movements: Unexpected target perturbations and movement kinematics. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *52*(4), 163–173.

Hoff, B., & Arbib, M. A. (1993). Models of trajectory formation and temporal interaction of reach and grasp. *Journal of Motor Behavior*, *25*(3), 175–192.

Huh, D., Todorov, E., & Sejnowski, T. (2010). Infinite horizon optimal control framework for goal directed movements. In *Proceedings of the 9th Annual Symposium on Advances in Computational Motor Control* (vol. 12).

Jiang, Y., Jiang, Z. P., & Qian, N. (2011). Optimal control mechanisms in human arm reaching movements. In *Proceedings of the 2011 30th Chinese Control Conference* (pp. 1377–1382). Piscataway, NJ: IEEE.

Krebs, H. I., Aisen, M. L., Volpe, B. T., & Hogan, N. (1999). Quantization of continuous arm movements in humans with brain injury. *Proceedings of the National Academy of Sciences*, *96*(8), 4645–4649.

Lee, D., Port, N. L., & Georgopoulos, A. P. (1997). Manual interception of moving targets II. On-line control of overlapping submovements. *Experimental Brain Research*, *116*(3), 421–433.

Levi, D. M., & Klein, S. A. (1992). "Weber's law" for position: The role of spatial frequency and contrast. *Vision Research*, *32*(12), 2235–2250.

Levy-Tzedek, S., Krebs, H. I., Song, D., Hogan, N., & Poizner, H. (2010). Non-monotonicity on a spatio-temporally defined cyclic task: Evidence of two movement types? *Experimental Brain Research*, *202*(4), 733–746.

Liu, D., & Todorov, E. (2007). Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. *Journal of Neuroscience*, *27*(35), 9354–9368.

McKee, S. P., Levi, D. M., & Bowne, S. F. (1990). The imprecision of stereopsis. *Vision Research*, *30*(11), 1763–1779.

McKee, S. P., Silverman, G. H., & Nakayama, K. (1986). Precise velocity discrimination despite random variations in temporal frequency and contrast. *Vision Research*, *26*(4), 609–619.

Milner, T. E. (1992). A model for the generation of movements requiring endpoint precision. *Neuroscience*, *49*(2), 487–496.

Morasso, P., Ivaldi, F., & Ruggiero, C. (1983). How a discontinuous mechanism can produce continuous patterns in trajectory formation and handwriting. *Acta Psychologica*, *54*(1–3), 83–98.

Mueller, A. S., & Timney, B. (2016). Visual acceleration perception for simple and complex motion patterns. *PLoS One*, *11*(2), 1–10.

Novak, K. E., Miller, L. E., & Houk, J. C. (2002). The use of overlapping submovements in the control of rapid hand movements. *Experimental Brain Research*, *144*(3), 351–364.

Pélisson, D., Prablanc, C., Goodale, M. A., & Jeannerod, M. (1986). Visual control of reaching movements without vision of the limb. *Experimental Brain Research*, *62*(2), 303–311.

Phillis, Y. A. (1985). Controller design of systems with multiplicative noise. *IEEE Transactions on Automatic Control*, *30*(10), 1017–1019.

Qian, N., & Andersen, R. A. (1997). A physiological model for motion-stereo integration and a unified explanation of Pulfrich-like phenomena. *Vision Research*, *37*(12), 1683–1698.

Qian, N., Jiang, Y., Jiang, Z.-P., & Mazzoni, P. (2013). Movement duration, fitts's law, and an infinite-horizon optimal feedback control model for biological motor systems. *Neural Computation*, *25*(3), 697–724.

Reichardt, W. (1987). Evaluation of optical motion information by movement detectors. *Journal of Comparative Physiology A*, *161*(4), 533–547.

Rigoux, L., & Guigon, E. (2012). A model of reward- and effort-based optimal decision making and motor control. *PLoS Computational Biology*, *8*(10), 1–13.

Scott, S. H. (2004). Optimal feedback control and the neural basis of volitional motor control. *Nat. Rev. Neurosci.*, *5*(7), 532–546.

Shadmehr, R., & Krakauer, J. W. (2008). A computational neuroanatomy for motor control. *Experimental Brain Research*, *185*(3), 359–381.

Shmuelof, L., Krakauer, J. W., & Mazzoni, P. (2012). How is a motor skill learned? Change and invariance at the levels of task success and trajectory control. *Journal of Neurophysiology*, *108*(2), 578–594.

Soechting, J. F., & Lacquaniti, F. (1983). Modification of trajectory of a pointing movement in response to a change in target location. *Journal of Neurophysiology*, *49*(2), 548–564.

Tanaka, H., Krakauer, J. W., & Qian, N. (2006). An optimization principle for determining movement duration. *Journal of Neurophysiology*, *95*(6), 3875–3886.

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nat. Neurosci.*, *7*(9), 907–915.

Todorov, E. (2005). Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. *Neural Computation*, *17*(5), 1084–1108.

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nat. Neurosci.*, *5*(11), 1226–1235.

Uno, Y., Kawato, M., & Suzuki, R. (1989). Formation and control of optimal trajectory in human multijoint arm movement. *Biological Cybernetics*, *61*(2), 89–101.

van der Wel, R. P. R. D., Sternad, D., & Rosenbaum, D. A. (2009). Moving the arm at different rates: Slow movements are avoided. *Journal of Motor Behavior*, *42*(1), 29–36.

Winters, J. M., & Stark, L. (1985). Analysis of fundamental human movement patterns through the use of in-depth antagonistic muscle models. *IEEE Transactions on Biomedical Engineering*, *32*(10), 826–839.