

**Effects of changing noise in dynamics of reaching on changes  
in control of reaching: An optimal control perspective**

by

Tushar D. Rane

A thesis submitted to The Johns Hopkins University in conformity with the  
requirements for the degree of Master of Science in Engineering.

Baltimore, Maryland

May, 2007

© Tushar D. Rane 2007

All rights reserved

# Abstract

Traditionally, the objective of the motor controller in the brain is thought to be to guide the limb to a desired trajectory while making a reaching movement. The controller in traditional models, thus, adapts internal models of the body and the environment to guide the limb to a 'desired trajectory'. However, there is significant evidence suggesting that even after extensive training in a novel environment, the trajectories followed by people do not converge to the so called 'desired trajectory'. Recently, optimal feedback control framework has been proposed as a model of motor control. According to this model, the aim of any action made by the controller is to achieve behavioral goals (i.e. earn rewards) while minimizing energy (i.e. costs). The minimization of these costs is constrained by the forward dynamics of the task, which makes the motor commands issued by the controller highly dependent on task dynamics. We hypothesized that the dependence of the motor commands issued by the controller on the task dynamics is not limited to the expected value of the parameters defining the task dynamics but also the variance in these parameters i.e. the learner's uncertainty about these parameters. This hypothesis in conjunction

with stochastic optimal feedback control framework has quite interesting predictions about movements made by people in environments with varying levels of uncertainties in the parameters defining dynamics of movement in the environment. We have been able to demonstrate through a series of experiments that these predictions indeed match behavioral data and thus, the control policy followed by the motor controller is dependent on its uncertainty about the parameters defining the task dynamics.

Thesis Advisor: Dr. Reza Shadmehr

Thesis Reader: Dr. Kechen Zhang

Thesis Reader: Dr. Mark Shelhamer

# Acknowledgements

I would like to thank my advisor Dr. Reza Shadmehr to provide me with constant guidance and support to complete this project. I would also like to thank Dr. Jun Izawa for being my mentor throughout this project. Thanks to thesis readers Dr. Kechen Zhang and Dr. Mark Shelhamer for their helpful comments to improve this thesis. And finally, thanks to my family, friends and colleagues for providing me with constant support and affection, making my experience as a graduate student gratifying.

# Contents

<b>Abstract</b>	<b>ii</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>List of Tables</b>	<b>vii</b>
<b>List of Figures</b>	<b>viii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Kinematic costs . . . . .	2
1.2 Dynamic costs . . . . .	3
1.3 Minimum Endpoint Variance Model . . . . .	5
1.4 Stochastic Optimal Feedback Control . . . . .	6
1.4.1 Minimal intervention principle . . . . .	7
<b>2 Problem formulation and solution</b>	<b>8</b>
2.1 Optimal feedback control model . . . . .	8
2.2 Optimal Control . . . . .	15
2.3 Optimal Estimation . . . . .	17
2.4 Introducing Model Parameter Uncertainty . . . . .	18
<b>3 Model Predictions</b>	<b>23</b>
3.1 Simulation Parameters . . . . .	23
3.2 Predictions for a point-to-point reaching task . . . . .	27
3.2.1 Biased viscous curl force field with off-diagonal noise . . . . .	27
3.2.2 Biased viscous curl force field with diagonal noise . . . . .	34
3.3 Predictions for a via-point reaching task . . . . .	34
<b>4 Experiments</b>	<b>39</b>
4.1 Apparatus . . . . .	39
4.2 Reach experiment with biased certain force field . . . . .	40

4.3	Reach experiment with biased variable force field . . . . .	45
4.3.1	Off-Diagonal Noise . . . . .	45
4.3.2	Diagonal Noise . . . . .	49
4.4	Via-point experiment with unbiased uncertain force field . . . . .	52
<b>5</b>	<b>Discussion and Conclusion</b>	<b>60</b>
5.1	Future work . . . . .	63
	<b>Appendices</b>	<b>65</b>
<b>A</b>	<b>Solving for the Optimal Controller and the Optimal Estimator with model parameter uncertainty</b>	<b>66</b>
A.1	Optimal Control Policy . . . . .	66
A.2	Optimal Estimator . . . . .	72
<b>B</b>	<b>Simulation parameters</b>	<b>78</b>
B.1	Common Parameters for all simulations . . . . .	78
B.1.1	Noises . . . . .	79
B.1.2	Cost Function . . . . .	80
B.2	Simulation specific parameters . . . . .	81
B.2.1	Biased viscous curl force field with off-diagonal and diagonal noise . . . . .	81
B.2.2	Via-point reaching task . . . . .	82
	<b>Vita</b>	<b>88</b>

# List of Tables

2.1	List of Notation . . . . .	11
4.1	Reach Experiment Protocol (Biased Force Field with zero variance) .	42
4.2	Reach Experiment Protocol (Biased Force Field with High Variance Off Diagonal noise) . . . . .	47
4.3	Experimental protocol for via-point task . . . . .	56

# List of Figures

2.1	The Stochastic Optimal Feedback Control System . . . . .	9
3.1	Simulation results for biased clockwise viscous curl force field with off-diagonal noise . . . . .	30
3.2	Simulation results for biased counterclockwise viscous curl force field with off-diagonal noise . . . . .	31
3.3	Simulation results for incomplete learning of a biased clockwise viscous curl force field with off-diagonal noise . . . . .	33
3.4	Simulation results for biased clockwise viscous curl force field with diagonal noise . . . . .	35
3.5	Simulation results for the via-point task with unbiased force field and off-diagonal noise . . . . .	37
4.1	Average trajectory in a zero variance viscous curl clockwise force field	44
4.2	Average trajectory in a zero variance viscous curl counterclockwise force field . . . . .	46
4.3	Average trajectory and speed profile comparison for movements in high off diagonal noise force field vs those in zero noise force field . . . . .	50
4.4	Peak speed comparison for movements in high off diagonal noise force field vs those in zero noise force field . . . . .	51
4.5	Average trajectory and speed profile comparison for movements in high diagonal noise force field vs those in zero noise force field . . . . .	53
4.6	Peak speed comparison for movements in high diagonal noise force field vs those in zero noise force field . . . . .	54
4.7	Average trajectory and speed profile comparison for the via-point task in zero variance unbiased force field vs that in high variance unbiased force field . . . . .	58
4.8	Via-point speed comparison for via-point task in high variance unbiased force field vs that in zero noise unbiased force field . . . . .	59

# Chapter 1

## Introduction

Human beings are capable of making highly skilled movements fairly easily. We usually take the controlling mechanism behind these movements for granted. However, the human motor control system has to keep making complex calculations to continuously adapt to and learn the properties of the environment in which the movements are being made. For many years motor control scientists have been trying to model the motor control system, leading to important improvements in our understanding of the working of the motor control system. Here I present a brief summary of some of the important models of the working of motor control system that have emerged over past few years.

Reaching movements form an important part of our daily lives. The movements we perform to grasp an object like a coffee cup or a ball form an example of reaching movements. The properties of the environment in which reaching movements are

performed can be changed to probe the adaptation and learning mechanisms of the motor system. The changes in properties of the environment can relate to changes in the kinematic map relating target position to arms configuration (e.g. use of prism glasses) or the change in dynamics of the arm (e.g. use of a curl force field) [18]

The human motor control system constantly keeps adapting and learning to improve the behavioral performance. The motor system has to play with multiple parameters to formulate the best strategy to achieve a behavioral goal. So the motor system is thought to be working towards optimizing a cost function which is reflective of the success in achieving the behavioral goal and movement constraints. Over the years multiple different cost functions have been proposed to explain the data observed in behavioral experiments.

## **1.1 Kinematic costs**

The cost function in models based on kinematics contain only geometrical and time-based properties of motion and the variables of interest are the positions and their corresponding velocities, accelerations and higher derivatives. Flash and Hogan [9][11] hypothesized that a major goal of the motor control system is to produce smooth movements. They proposed that the objective of the motor control system is the optimization of a cost function which is reflective of the smoothness of the movement. The cost function proposed was the square of the magnitude of jerk or rate of change of acceleration integrated over the whole movement.

$$Cost = \frac{1}{2} \int_0^{t_f} \left( \left( \frac{d^3x}{dt^3} \right)^2 + \left( \frac{d^3y}{dt^3} \right)^2 \right) dt \quad (1.1)$$

Where  $x$  and  $y$  are time varying hand position co-ordinates in the laboratory fixed Cartesian coordinate system and  $t_f$  is the time at which the arm should reach the final position from the initial position. The dynamic optimization of this cost function was able to explain straight-line Cartesian hand paths with bell shaped velocity profiles as observed in the empirical data for rapid movements without accuracy constraints. [13]

## 1.2 Dynamic costs

The cost function in models based on dynamics depends on the dynamics of the arm, and the variables of interest include joint torques, forces acting on the hand and muscle commands. Considering the observation that people produce curvilinear movements when asked to move over a large range, a phenomenon unexplainable by the minimum jerk model, Uno et al [23] suggested an alternative cost function called the minimum torque change cost function.

$$Cost = \frac{1}{2} \int_0^T \left( \left( \frac{d\tau_1}{dt} \right)^2 + \left( \frac{d\tau_2}{dt} \right)^2 \right) dt \quad (1.2)$$

where  $\tau_1$  and  $\tau_2$  are the torques at the two arm joint (i.e. shoulder and elbow) and  $T$  is the time allowed for the arm to reach the final position from the starting position.

An important difference between minimum jerk model and the minimum torque change model is the separability of movement planning and execution. The trajectory obtained by minimizing 'jerk' is in terms of positions and velocities of arm as functions of time and a separate mechanism is required to achieve this trajectory. On the other hand, the solution of the problem of minimization of torque change specifies a sequence of motor commands necessary to achieve the movement, making the planning and execution steps the same. Uno et al were able to show that the minimum torque change model was able to predict the trajectories they observed in their empirical data.

Minimum jerk and minimum torque change model have been important in understanding different aspects of motor control. They, however, have certain features making them unsatisfying. These models do not provide any satisfactory explanation for why the Central Nervous System (CNS) would choose to optimize quantities like jerk and torque change. These models make the assumption that the objective of the CNS is to produce smooth movements, however, no particular advantage of smoothness is specified. Also, even if the objective of the CNS was to minimize jerk or torque change, it is hard to imagine how the CNS might be computing integrals of complex quantities like jerk and torque change over the course of a movement. [10][13]

### 1.3 Minimum Endpoint Variance Model

To overcome the problems associated with these models, Harris and Wolpert proposed that while making reaching movements, the objective of the CNS is to minimize the variance of the final limb position, in the presence of biological noise [10]. It is hypothesized that the noise in the neuronal control signal (i.e. motor neuron firing rate) will cause the trajectories to deviate from the desired path. These deviations would accumulate over the duration of the movement making the final position variable.

If the noise in the neural control signal were independent of the control signal, then the accumulated error can be minimized by making the movement as fast as possible. However, one of the key assumptions in the minimum variance model proposed by Harris and Wolpert is that the noise in the neural control signal increases with the mean level of the signal. This assumption is supported by a variety of physiological [3][15] and psychophysical data [8]. As a result of this assumption, making fast movements would actually lead to higher variability in final position as moving fast for low pass systems like the arm would require control signals which lead to higher noise. On the other hand, moving slow would lead to increase in the amount of time taken to achieve the target. Hence, the movement trajectory obtained by minimizing the end point variance with the assumption of signal dependent noise is a trade-off between movement duration and accuracy at the endpoint [10][13]

The minimum endpoint variance model provided a biologically plausible theory for arm movements. To operate according to this model, the CNS would be able to

use a measure of movement cost (i.e. endpoint variance), which is directly available to the CNS as against the minimum jerk and torque change models which require the CNS to construct highly complicated functions of the quantities available to it for estimation of movement cost. The smoothness of the movement comes naturally since non-smooth movements would require higher control signals which in turn lead to higher noise, increasing the endpoint variability of the movement. [10]

A recurring theme in all the models listed above is the separation between “trajectory planning” and “trajectory execution”. As a result of this separation, the behavioral goal is replaced by a so called ‘desired trajectory’. Accurate execution of this desired trajectory guarantees the achievement of the behavioral goal. However, there is an extensive body of evidence [21] which suggests that the trial to trial fluctuations are much larger in task irrelevant movement-parameters as compared to the task relevant movement-parameters or the parameters crucial for achievement of the behavioral goal. This implies that the motor control system tries to restrict the motor variability to a task irrelevant subspace rather than suppressing variability over the whole space of movement-parameters. This would be impossible if the behavioral goal is replaced by a desired trajectory.

## 1.4 Stochastic Optimal Feedback Control

Todorov and Jordan proposed Stochastic optimal feedback control as a theory of motor control [21]. An important feature of this model is the fact that whenever

the task allows redundant solutions, rather than resolving the redundancy all at once before starting the movement, the controller achieves optimal behavior by resolving the redundancy moment by moment making use of a feedback control law. As the resolution of redundancy is delayed till the last moment, the control law can make use of additional task completion opportunities created by fluctuations from the average trajectory. Such exploitation of redundancy can explain several phenomena related to motor coordination like task-constrained variability, goal-directed corrections and motor synergies.

#### **1.4.1 Minimal intervention principle**

A distinct feature of Stochastic Optimal Feedback control is the ‘Minimal intervention principle’. This principle states that the controller lets motor variability accumulate in the task irrelevant directions rather than trying to eliminate it. This is an optimal thing to do for the controller since the task of the controller is to achieve the goal and by definition making any attempts to reduce motor variability in the task irrelevant dimensions wouldn’t increase the probability of achieving the goal. These attempts can actually increase the probability of missing the goal since both noise and energy spent are control-dependent and can potentially increase. Allowing uncertainty to accumulate can improve the chances of achieving the goal by allowing the controller to further suppress uncertainty in task relevant directions.

# Chapter 2

## Problem formulation and solution

### 2.1 Optimal feedback control model

Under the optimal control framework, the motor control system is modeled as a feedback controller as shown in Fig. 2.1. In this system, the optimal controller computes the “optimal policy” to achieve a goal given its model of the environment. The “optimal policy” is computed so that it minimizes a measure of cost incurred by the biological plant in achieving the goal. The motor commands issued by the controller act on the biological plant or the hand for reaching movements, producing a change in the state of the biological plant. The sensory system i.e. visual system, proprioceptive system etc. sense a noisy function of the change in state of the biological plant. The optimal estimator receives a copy of motor commands issued by the controller i.e. efference copy and the noisy observations from the sensory system. There

is considerable behavioral evidence suggesting that the motor controller should have a copy of the commands it generates so that it can predict the consequences of its own actions [12][1][17] [26][4]. This becomes highly important both in online control of movements as well as co-ordination of movements of different body parts. The optimal estimator makes a prediction about the state of the biological plant based on its model of the world and compares it with the observations from the sensory system. Based on the error in its estimate, the optimal estimator updates its estimate of the state of the biological plant and feeds this information to the controller. Making use of this state estimate, the optimal feedback controller generates motor commands following its optimal policy.

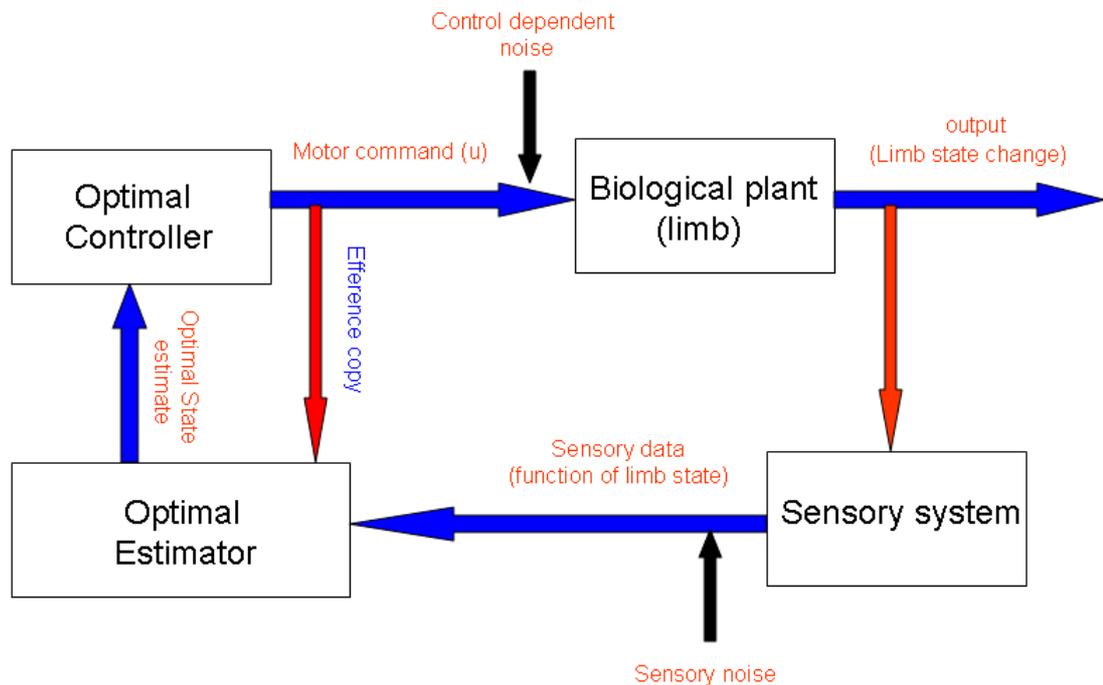


Figure 2.1 The Stochastic Optimal Feedback Control System

Testing optimality principles in biological movement empirically requires us to be able to solve the stochastic optimal control and estimation problem. Due to the intractability of the solution of the optimal feedback control problem with a biologically correct, control-dependent noise, open loop optimization of control signals with a control-dependent noise has been proposed in the field of sensorimotor control [10]. However, for the stochastic partially observable plants like the musculoskeletal system, the control sequence obtained by open-loop optimization has been shown to be suboptimal [21].

The stochastic optimal feedback control problem with additive gaussian noise can be solved efficiently in the Linear-Quadratic-Gaussian(LQG) framework. However, many robust phenomena observed in human motor control like trajectory smoothness, speed-accuracy trade-offs and structured motor variability are linked to signal-dependent nature of sensorimotor noise [10][21]. So, it necessary to extend the LQG framework to include biologically plausible noises. Recently, Todorov [20] proposed a co-ordinate descent algorithm to solve the stochastic optimal control and estimation problem for a partially observable plant with biologically plausible noise in the LQG framework. The optimization problem formulation incorporates state-dependent, control-dependent and internal noise as explained below.

For simplicity, we shall be following the notations similar to those used by Todorov for all subsequent derivations[20]. The notations being used are listed in Table 2.1.

Consider a discrete linear dynamic system with state  $\mathbf{x}_t \in \mathfrak{R}^m$ , control  $\mathbf{u}_t \in \mathfrak{R}^p$

and feedback  $\mathbf{y}_t \in \mathfrak{R}^k$  in discrete time  $t$ .

$$\text{Dynamics} \quad \mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t + \xi_t + \sum_{i=1}^c \varepsilon_t^i \bar{C}_i \mathbf{u}_t \quad (2.1)$$

$$\text{Feedback} \quad \mathbf{y}_t = H\mathbf{x}_t + \omega_t + \sum_{i=1}^d \epsilon_t^i D_i \mathbf{x}_t \quad (2.2)$$

$$\text{Cost per step} \quad \text{cps}_t = \mathbf{x}_t^T Q_t \mathbf{x}_t + \mathbf{u}_t^T R \mathbf{u}_t \quad (2.3)$$

Table 2.1 List of Notation

---

$\mathbf{x}_t \in \mathfrak{R}^m$	state vector at time step $t$
$\mathbf{u}_t \in \mathfrak{R}^p$	control signal
$\mathbf{y}_t \in \mathfrak{R}^k$	sensory observation
$n$	total number of time steps
$A, B, H$	system dynamics and observation matrices
$\xi_t, \omega_t, \varepsilon_t, \epsilon_t, \eta_t$	zero mean noise terms
$\Omega^\xi, \Omega^\omega, \Omega^\varepsilon, \Omega^\epsilon, \Omega^\eta$	covariances of noise terms
$C_1, \dots, C_d$	scaling matrices for state-dependent system noise
$\bar{C}_1, \dots, \bar{C}_d$	scaling matrices for control-dependent system noise
$D_1, \dots, D_d$	scaling matrices for state-dependent observation noise
$Q_t, R$	matrices defining state and control-dependent costs
$\hat{\mathbf{x}}_t$	state estimate
$\mathbf{e}_t$	estimation error
$\Sigma_t$	conditional estimation error covariance
$\Sigma_t^e, \Sigma_t^{\hat{\mathbf{x}}}, \Sigma_t^{\hat{\mathbf{x}}e}$	unconditional covariances
$v_t$	optimal cost-to-go function
$S_t^{\hat{\mathbf{x}}}, S_t^e, s_t$	parameters of the optimal cost-to-go function
$K_t$	filter gain matrices
$L_t$	control gain matrices

---

The system is initialized with a known state mean  $\hat{\mathbf{x}}_1$  and state covariance  $\Sigma_1$ . The matrices  $A$  and  $B$  define the dynamics of the system. The next state of the system depends on its previous state scaled by matrix  $A$  and the control signal  $\mathbf{u}_t$  generated during the previous time step  $t$  scaled by matrix  $B$ . The term  $\sum_{i=1}^c \varepsilon_t^i \bar{C}_i \mathbf{u}_t$  defines

the way control signal dependent noise has been implemented in this formulation. It has been shown in previous research that the standard deviation in the muscle force is a linear function of the mean force in both static [21] and isometric force tasks [16]. If  $\mathbf{u}$  is a vector of control signals and  $\varepsilon$  is a vector of zero mean random numbers, multiplicative noise can be expressed as  $\bar{C}(\mathbf{u})\varepsilon$ , where  $\bar{C}(\mathbf{u})$  is a matrix whose elements depend linearly on  $\mathbf{u}$ . The linear relationship between the matrix  $\bar{C}$  and the vector  $\mathbf{u}$  can be expressed by making the  $i^{th}$  column of  $\bar{C}$  equal to  $\bar{C}_i\mathbf{u}$ , where  $\bar{C}_i$  are constant scaling matrices. Given this structure of matrix  $\bar{C}$ , it is easy to see that  $\bar{C}(\mathbf{u})\varepsilon = \sum_i \bar{C}_i\mathbf{u}\varepsilon^i$ , where  $\varepsilon^i$  is the  $i^{th}$  component of random vector  $\varepsilon$ .

We constantly receive sensory feedback from various sensory modalities to help us control our movements online. Vision and Proprioception form two of the most important sensory modalities for online movement control. The sensory feedback is incorporated in the model through the feedback equation. On every time step, the system receives feedback  $\mathbf{y}_t$ , which is a linear function of the current state of the system. The observation matrix  $H$  defines the linear relationship between state  $\mathbf{x}_t$  and feedback  $\mathbf{y}_t$ . Analogous to the multiplicative nature of motor noise, the accuracy of visual position estimates can be well modeled with multiplicative noise, whose standard deviation is proportional to the eccentricity[2][24]. This scaling law has been confirmed in the visuomotor setting [7]. To incorporate this effect in the model, the feedback equation has multiplicative observation noise in the form  $D(\mathbf{x})\epsilon = \sum_i D_i\mathbf{x}\epsilon_i$ , where  $\mathbf{x}$  is the state of the plant and the environment.

For quantitative modeling of sensorimotor control, we require a scalar cost function as a measure of success. The step by step cost function proposed in [7] has two components to it. The first term  $\mathbf{x}_t^T Q_t \mathbf{x}_t$  is the ‘state cost’. For simple point to point reaching movements, this cost can be converted to a ‘tracking cost’ i.e. a cost which penalizes the controller for not being at the desired position at any time point during the movement. The matrices  $Q_1, \dots, Q_n$  are symmetric positive semi-definite matrices which scale the cost over the course of the movement. Note that if you set all the  $Q$  matrices to a high value for the ‘tracking cost’ case, you are essentially forcing the controller to follow a pre-specified ‘desired trajectory’. However, if the objective of a movement is just to be at a spatial location at the end of the movement, you would penalize the controller only for not being at the final desired position at the end of the movement while setting the cost to zero during the movement. This provides the optimization algorithm with maximum possible flexibility to search for a control policy which can be used to achieve the goal with minimum effort. The second component of the cost function  $\mathbf{u}_t^T R \mathbf{u}_t$  is the ‘control cost’. The matrix  $R$  is a symmetric positive definite matrix which scales the ‘control cost’. This component penalizes the controller for producing large motor commands. This component can be thought of as being indicative of the energy required for producing the motor commands. The larger the motor commands you produce, the larger is the energy spent in doing so. Hence, it would make sense for the controller to try and minimize the energy required to complete a motor task.

Each movement is modeled as being made of  $n$  time steps. The independent random variables  $\xi_t \in \mathfrak{R}^m, \omega_t \in \mathfrak{R}^k, \varepsilon_t \in \mathfrak{R}^c$  and  $\epsilon_t \in \mathfrak{R}^d$  define and scale various noise terms incorporated in the model. The state  $x$  is not directly available to the system. So, the system continuously needs to make an estimate of its state making use of the feedback it gets on every time step. For linear systems the optimal way to update the state estimate is to use the Kalman filter. It is easy to see that when the control-dependent and the state-dependent noise terms are zero (i.e.  $\bar{C}_1, \dots, \bar{C}_d = 0$  and  $D_1, \dots, D_d = 0$ ), the system reduces to the classic LQG system with additive noise only. This classic LQG problem has been solved long ago. [14]

<i>Linear-Quadratic Regulator</i>	<i>Kalman Filter</i>
$\mathbf{u}_t = -L_t \hat{\mathbf{x}}_t$	$\hat{\mathbf{x}}_{t+1} = A \hat{\mathbf{x}}_t + B \mathbf{u}_t + K_t (\mathbf{y}_t - H \hat{\mathbf{x}}_t)$
$L_t = (R + B^T S_{t+1} B)^{-1} B^T S_{t+1} A$	$K_t = A \Sigma_t H^T (H \Sigma_t H^T + \Omega^\omega)^{-1}$
$S_t = Q_t + A^T S_{t+1} (A - B L_t)$	$\Sigma_{t+1} = \Omega^\varepsilon + (A - K_t H) \Sigma_t A^T$

In this case, the matrices  $L$ , defining the optimal control law, do not depend on any of the noise covariances or the Kalman gain matrices  $K$ . Similarly the Kalman gain matrices  $K$  do not depend on the cost or optimal control law. The dependence of optimal motor commands  $\mathbf{u}_t$  on the history of control and feedback signals is only through the state estimate  $\hat{\mathbf{x}}_t$ , updated recursively by the Kalman filter. This independence of optimal control law formulation and state estimation does not hold if we add state-dependent and control-dependent noise to the model. So, for the purpose of analytical tractability, the state estimate is assumed to be updated according to a

linear recursive filter.

$$\hat{\mathbf{x}}_{t+1} = A\hat{\mathbf{x}}_t + B\mathbf{u}_t + K_t(\mathbf{y}_t - H\hat{\mathbf{x}}_t) + \eta_t \quad (2.4)$$

here, the term  $\eta_t \in \mathfrak{R}^m$  is the so called ‘Internal noise’. Generally it is assumed that the neural representations work as flawlessly as a digital computer which is unrealistic. So, to incorporate the fluctuations in neural representations of the model, a term for ‘Internal noise’ is added to the model. The Kalman filter gains are assumed to be non-adaptive and are determined in advance. A non-adaptive filter for this model is sub-optimal. However, the difference between the performance with adaptive vs non-adaptive filtering was explored numerically in [20] and not found to be substantially different for models similar to musculoskeletal plants.

## 2.2 Optimal Control

It can be shown that given the system dynamics model described above, the cost expected to accumulate after a particular time step  $t$ , if the controller follows an optimal control policy for the remaining time steps,  $(\mathbf{u}_t = \pi(\hat{\mathbf{x}}_t))$  has a quadratic form [20]

$$v_t(\mathbf{x}_t, \hat{\mathbf{x}}_t) = \mathbf{x}_t^T S_t^x \mathbf{x}_t + \mathbf{e}_t^T S_t^e \mathbf{e}_t + s_t \quad (2.5)$$

where  $\mathbf{e}_t = \mathbf{x}_t - \hat{\mathbf{x}}_t$  is a quantity defined for simplicity in the derivation. This quantity which is the difference between the actual state and the state estimate can be called the estimation error.

Minimizing this cost function with respect to the control law  $\pi(\hat{\mathbf{x}}_t)$  gives us the optimal control law. However, the cost function as shown above is a function of the actual state  $\mathbf{x}_t$ , which is not available to the controller. The best estimate of the state available to the controller at any time point  $t$  is  $\hat{\mathbf{x}}_t$ . So, the cost function actually minimized to obtain the optimal control law is the expected value of the cost function (Equation 2.5) given the state estimate  $\hat{\mathbf{x}}_t$ . Solving this problem gives us a system of equations which lets us calculate the optimal control law recursively backward in time (Equation 2.6).

The cost matrices are initialized so that  $S_n^{\mathbf{x}} = Q_n, S_n^{\mathbf{e}} = s_n = 0$ . As stated before, the optimal filter gain matrices  $K$  are assumed to be fixed to pre-known values. The iterative relations for computing the optimal control law show that the feedback gain matrices  $L$  do not directly depend on the additive gaussian noise terms  $\eta, \xi, \omega$ . However, as shown in the next section, the computation of the optimal filter gains  $K$  depends on all the noises, which in turn affect the calculation of feedback gain matrices  $L$ .

$$\begin{aligned}
\text{Controller } \mathbf{u}_t &= -L_t \hat{\mathbf{x}}_t \\
L_t &= \left( R + B^T S_{t+1}^{\mathbf{x}} B + \sum_i \bar{C}_i^T (S_{t+1}^{\mathbf{x}} + S_{t+1}^{\mathbf{e}}) \bar{C}_i \right)^{-1} B^T S_{t+1}^{\mathbf{x}} A \\
S_t^{\mathbf{x}} &= Q_t + A^T S_{t+1}^{\mathbf{x}} (A - BL_t) + \sum_i D_i^T K_t^T S_{t+1}^{\mathbf{e}} K_t D_i; \quad S_n^{\mathbf{x}} = Q_n \\
S_t^{\mathbf{e}} &= A^T S_{t+1}^{\mathbf{x}} BL_t + (A - K_t H)^T S_{t+1}^{\mathbf{e}} (A - K_t H); \quad S_n^{\mathbf{e}} = 0 \\
s_t &= Tr(S_{t+1}^{\mathbf{x}} \Omega^\xi + S_{t+1}^{\mathbf{e}} (\Omega^\xi + \Omega^\eta + K_t \Omega^\omega K_t^T)) + s_{t+1}; \quad s_n = 0
\end{aligned}$$

## 2.3 Optimal Estimation

In the previous section, the feedback gain matrices for the optimal control law were calculated for a fixed set of optimal filter gain matrices. Now our objective is to calculate the updated sequence of optimal filter gain matrices such that they minimize the cost function for the optimal feedback gain matrices  $L$  obtained in the previous section. Thus, the filter gain matrices  $K$ , rather than being calculated to minimize the estimation error are calculated to minimize the cost function for a given set of feedback gain matrices. As shown in [20] new optimal filter gain matrices can be calculated in a forward pass through time. The optimal filter gain matrix  $K_t$  at any time point  $t$  can be found analytically as long as  $K_{t+1}, \dots, K_{n-1}$  still have values for which the feedback gain matrices  $L_{t+1}, \dots, L_{n-1}$  are still optimal.

$$\begin{aligned}
 \text{Estimator } \hat{\mathbf{x}}_{t+1} &= (A - BL_t) \hat{\mathbf{x}}_t + K_t (\mathbf{y}_t - H\hat{\mathbf{x}}_t) + \eta_t \\
 K_t &= A \Sigma_t^e H^T \left( H \Sigma_t^e H^T + \Omega^\omega + \sum_i D_i (\Sigma_t^e + \Sigma_t^{\hat{\mathbf{x}}} + \Sigma_t^{\hat{\mathbf{x}}e} + \Sigma_t^{e\hat{\mathbf{x}}}) D_i^T \right)^{-1} \\
 \Sigma_{t+1}^e &= \Omega^\xi + \Omega^\eta + (A - K_t H) \Sigma_t^e A^T + \sum_i C_i L_t \Sigma_t^{\hat{\mathbf{x}}} L_t^T C_i^T; \quad \Sigma_1^e = \Sigma_1 \\
 \Sigma_{t+1}^{\hat{\mathbf{x}}} &= \Omega^\eta + K_t H \Sigma_t^e A^T + (A - BL_t) \Sigma_t^{\hat{\mathbf{x}}} (A - BL_t)^T \\
 &\quad + (A - BL_t) \Sigma_t^{\hat{\mathbf{x}}e} H^T K_t^T + K_t H \Sigma_t^{e\hat{\mathbf{x}}} (A - BL_t)^T; \quad \Sigma_1^{\hat{\mathbf{x}}} = \hat{\mathbf{x}}_1 \hat{\mathbf{x}}_1^T \\
 \Sigma_{t+1}^{\hat{\mathbf{x}}e} &= (A - BL_t) \Sigma_t^{\hat{\mathbf{x}}e} (A - BL_t)^T - \Omega^\eta; \quad \Sigma_1^{\hat{\mathbf{x}}e} = 0
 \end{aligned}$$

Given these sets of equations, we can estimate the optimal feedback gain matrices  $L$  and the optimal filter gain matrices  $K$ . The iterations are started by assuming

a starting value for the sequence  $K_1, \dots, K_{n-1}$ . Then, we keep on iterating between the system of equations for optimal controller and that for optimal estimator until convergence.

## 2.4 Introducing Model Parameter Uncertainty

A motor controller needs to have a model of functioning of the surrounding environment as well as the body to be able to execute desired movements. Such models are called as ‘Internal models’ which can be viewed as the motor controllers view of the functioning of the surrounding world and the body[13][25]. The internal models help the controller predict in advance, the change in state of the body given that it has issued a certain motor command. It is very essential for the controller to be able to predict the consequences of its motor commands while performing essentially no feedback movements like saccades since the movement times are too fast for the controller to receive any feedback from the sensory system. This ability to predict is also essential in improving the performance of the controller in online control of the movement since the feedback is received from the sensors with considerable delay. The predictions from the internal models can then be compared with sensory feedback to obtain an actual estimate of state of the body to correct the internal model if there are any discrepancies.

Human beings have to constantly deal with changing properties of the surrounding environment as well as the body itself. One can think of multiple instances involving

such changes of properties which force the motor control system to adapt its internal models about the functioning of the surrounding world as well as the body in order to execute desired movements properly. The changes in the properties, however, are not always deterministic. One simple example would be mass of an object we are trying to lift. We are generally very good at judging the mass of an object by making use of various cues like the color of the object, surface texture of the object etc. If a person is going to lift an object he/she is used to lifting, the person doesn't need to think much about it. On the other hand, if the person is about to lift an unknown object made of an unknown material, he/she would approach much carefully. In other words, there is a difference between the way a person executes a motor task when the person is fairly certain about the properties of the entities involved as against when the person has fairly high uncertainty about the properties of the entities.

A natural conclusion from the above observations would be that the motor control system must have some way of representing uncertainty regarding the parameters of the Internal model so that these uncertainties can be considered while formulating the optimal control policy to execute a motor task. To incorporate the uncertainty of the motor controller regarding the parameters of the model, we made the following simple modification to the system dynamics equation 2.1.

$$\begin{aligned}
 \mathbf{x}_{t+1} &= (A + \eta\varepsilon_t) \mathbf{x}_t + B\mathbf{u}_t + \xi_t + \sum_{i=1}^d \gamma_t^i \bar{C}_i \mathbf{u}_t \\
 \mathbf{x}_{t+1} &= A\mathbf{x}_t + B\mathbf{u}_t + \xi_t + \eta\varepsilon_t \mathbf{x}_t + \sum_{i=1}^d \gamma_t^i \bar{C}_i \mathbf{u}_t
 \end{aligned} \tag{2.6}$$

where  $\eta$  is a parameter which modulates the variance in the model parameter  $A$  and  $\varepsilon_t$  is a random variable drawn from the distribution  $N(0, 1)$ . The uncertainty term  $\eta\varepsilon_t\mathbf{x}_t$  can be expressed as a state-dependent ( $x$ ) noise in the state update equation as shown in Equation 2.7. Modulating the matrices  $C$  can be interpreted as changing the variance in parameter  $A$  i.e the uncertainty of the controller regarding internal model parameter  $A$ .

$$\text{Dynamics } \mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t + \xi_t + \sum_{i=1}^d \varepsilon_t^i C_i \mathbf{x}_t + \sum_{i=1}^d \gamma_t^i \bar{C}_i \mathbf{u}_t \quad (2.7)$$

$$\text{Feedback } \mathbf{y}_t = H\mathbf{x}_t + \omega_t + \sum_{i=1}^d \epsilon_t^i D_i \mathbf{x}_t \quad (2.8)$$

$$\text{Cost per step } \mathbf{cps}_t = \mathbf{x}_t^T Q_t \mathbf{x}_t + \mathbf{u}_t^T R \mathbf{u}_t \quad (2.9)$$

This extended version of this stochastic optimal feedback control problem is simple to solve following the steps followed in [20]. The optimal control policy obtained by solving this extended problem is shown in the following set of equations. All the notations used in the above equations have meanings as described in table 2.1. Please

refer to the appendix for the detailed derivation.

$$\begin{aligned}
\text{Controller } \mathbf{u}_t &= -L_t \hat{\mathbf{x}}_t \\
L_t &= (R + B^T S_{t+1}^x B + \bar{C}_t)^{-1} B^T S_{t+1}^x A \\
S_t^x &= Q_t + A^T S_{t+1}^x (A - BL_t) + N_t \\
S_t^e &= (A - K_t H)^T S_{t+1}^e (A - K_t H) + A^T S_{t+1}^x B L_t \\
s_t &= s_{t+1} + \text{Tr}(M_t) \\
N_t &= \sum_{i=1}^d (C_i^T (S_{t+1}^x + S_{t+1}^e) C_i + D_i^T K_t^T S_{t+1}^e K_t D_i) \\
\bar{C}_t &= \sum_{i=1}^d (\bar{C}_i^T (S_{t+1}^x + S_{t+1}^e) \bar{C}_i) \\
M_t &= S_{t+1}^x \Omega^\xi + S_{t+1}^e (\Omega^\xi + K_t \Omega^\omega K_t^T + \Omega^\eta)
\end{aligned}$$

Following the methodology similar to [20], we can also obtain the changed set of equations for the optimal estimator as shown below

$$\begin{aligned}
\text{Estimator } \hat{\mathbf{x}}_{t+1} &= (A - BL_t) \hat{\mathbf{x}}_t + K_t (\mathbf{y}_t - H \hat{\mathbf{x}}_t) + \eta_t \\
K_t &= A \Sigma_t^e H^T \left( H \Sigma_t^e H^T + \Omega^\omega + \sum_{i=1}^d D_i \Sigma_t^x D_i^T \right)^{-1}
\end{aligned}$$

$$\begin{aligned}
\Sigma_{t+1}^{\hat{\mathbf{x}}} &= (A - BL_t) \Sigma_t^{\hat{\mathbf{x}}} (A - BL_t)^T + K_t \left( \sum_{i=1}^d D_i (\Sigma_t^{\hat{\mathbf{x}}} + \Sigma_t^{\mathbf{e}} + \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}}) D_i^T \right) K_t^T + \Omega^\eta + \\
&\quad + K_t (H \Sigma_t^{\mathbf{e}} H^T + \Omega^\omega) K_t^T + (A - BL_t) \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} H^T K_t^T + K_t H \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} (A - BL_t)^T \\
\Sigma_{t+1}^{\mathbf{e}} &= (A - K_t H) \Sigma_t^{\mathbf{e}} (A - K_t H)^T + \sum_{i=1}^d \left( C_i (\Sigma_t^{\hat{\mathbf{x}}} + \Sigma_t^{\mathbf{e}} + \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} + \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}^T}) C_i^T \right) + \\
&\quad K_t \left( \sum_{i=1}^d \left( D_i (\Sigma_t^{\hat{\mathbf{x}}} + \Sigma_t^{\mathbf{e}} + \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} + \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}^T}) D_i^T \right) + \Omega^\omega \right) + \Omega^\xi + \Omega^\eta + \\
&\quad \sum_{i=1}^d (\bar{C}_i L_t \Sigma_t^{\hat{\mathbf{x}}} L_t^T \bar{C}_i^T) \\
\Sigma_{t+1}^{\hat{\mathbf{x}}\mathbf{e}} &= K_t H \Sigma_t^{\mathbf{e}} (A - K_t H)^T - K_t \left( \sum_{i=1}^d D_i (\Sigma_t^{\hat{\mathbf{x}}} + \Sigma_t^{\mathbf{e}} + \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} + \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}^T}) D_i^T + \Omega^\omega \right) K_t^T - \\
&\quad \Omega^\eta + (A - BL_t) \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} (A - K_t H)^T
\end{aligned}$$

# Chapter 3

## Model Predictions

### 3.1 Simulation Parameters

With the framework we developed in the previous chapter, we have a way to estimate the optimal behavior an ideal optimal feedback control system should exhibit in various different kinds of motor environments. These predictions can then be compared with human behavior to answer questions like does human motor control system actually exhibit behavior in accordance with optimal feedback control theory?, Does the optimal policy followed by the motor controller depend on its uncertainty about the world? and if yes, can these changes be explained by the optimal feedback control theory?

We focused our attention to point-to-point reaching movements in a novel dynamic environment called a viscous curl force field [18]. The effect of a viscous curl force

field is similar to altering the viscosity of the surrounding environment. The force produced by this force field is given by

$$f = Z\dot{x} \quad (3.1)$$

where for a simple 2-D case,  $f = [f_x, f_y]$ ,  $\dot{x} = [V_x, V_y]$  and  $f_x, V_x$  and  $f_y, V_y$  are the force and velocity in the  $x$  and  $y$  direction respectively. The elements of the matrix  $Z$  are the parameters controlling the effect of the force field. We can introduce uncertainty in these parameters as shown below.

$$f = [Z + \beta\eta] \dot{x} \quad (3.2)$$

where  $\beta$  is the noise gain modulation parameter and  $\eta \sim N(0, 1)$ .

Since we are limited to a linear dynamical system, we modeled the hand as a point mass system. The dynamics of a point mass in a viscous curl force field can be easily represented as a linear dynamical system. For a simple point-to-point reaching task, the state was a  $8 \times 1$  vector with the elements of the state being  $\mathbf{x} = [P_x, V_x, P_y, V_y, F_x, F_y, T_x, T_y]$  where  $P_x$  and  $P_y$  are the hand co-ordinates,  $V_x$  and  $V_y$  are the hand velocities,  $F_x$  and  $F_y$  are the forces produced by the hand and  $T_x$  and  $T_y$  are the target positions in  $x$  and  $y$  directions respectively. The motor command  $\mathbf{u}$  is a  $2 \times 1$  vector  $\mathbf{u} = [u_x, u_y]$  where  $u_x$  and  $u_y$  are the  $x$  and  $y$  components of the motor commands. The translation of motor commands  $u_x$  and  $u_y$  into forces  $F_x$  and  $F_y$  was modeled as a first order dynamical system

$$\tau \frac{\partial F_i}{\partial t} = u_i - F_i; \quad (3.3)$$

the parameter  $\tau$  lets us control the latency between issue of motor command  $u_i$  from the brain and the expression of force by the muscles  $F_i$ . Given the above assumptions, we can generate the matrices  $A$  and  $B$  defining system dynamics (Eqn 2.6).

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & ZM_{11} & 0 & ZM_{12} & M_{11}^{-1} & M_{12}^{-1} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & ZM_{21} & 0 & ZM_{22} & M_{21}^{-1} & M_{22}^{-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{-1}{\tau} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{-1}{\tau} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} dt + I; \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \frac{1}{\tau} & 0 \\ 0 & \frac{1}{\tau} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} dt \quad (3.4)$$

where matrix  $M$  is the matrix defining mass properties of the hand, the matrix  $ZM$  is the matrix formed by the product of matrices  $Z$  and  $M^{-1}$ ,  $dt$  is the time step used in the discrete dynamical system and  $I$  is an identity matrix of size  $8 \times 8$ .

It is easy to see the working of the dynamics system we just defined. The 1st and 3rd elements of the state are hand co-ordinates  $P_x$  and  $P_y$  respectively. They are just being updated on every time step by adding  $V_x dt$  and  $V_y dt$  to their values on previous time step respectively i.e. by adding the area below the velocity curve for the respective direction on every time step. The velocities  $V_x$  and  $V_y$  are updated on every time step by adding the area below the acceleration curve in the respective direction on every time step to their values on previous time step. The viscous curl

force field generates forces on the point mass depending on the velocity of the mass. Hence, the acceleration at any time step is computed by considering the effect of forces  $F_x$  and  $F_y$  produced on the last time step as well as the forces produced by the curl force field on the hand. The update rule for the forces  $F_x$  and  $F_y$  basically implements the dynamics shown in the Eqn 3.3 and the target position  $T_x$  and  $T_y$  remains unchanged over the course of the movement.

The matrix  $Q$  defining the ‘tracking cost’ (Eqn 2.9) is such that the controller is penalized for not being at the target at the pre-specified movement completion time  $t_f$ . The matrix  $Q$  is a zero matrix at all other time points. The matrix  $R$  defining the ‘control cost’ is set to a constant value throughout the movement time.

The control signal only affects the forces produced by the plant directly. The change in velocity and position, however, doesn’t depend directly on the motor commands. So, the control-dependent noise matrix  $\bar{C}$  (Refer to Eqn 2.7) was defined as a  $8 \times 2$  matrix with all the elements zero except the elements  $\bar{C}_{51}$  and  $\bar{C}_{62}$ . Please note that this structure of the control-dependent noise matrix ensures that any control-dependent noise affects only the forces produced by the plant directly.

Now we are interested in varying the uncertainty of the learner/controller regarding the model parameters and find out the difference in the optimal control policies for different levels of uncertainty. In the case of viscous curl force fields, the matrix  $Z$  is the parameter we can control and introduce uncertainty in, to test the predictions of the model. So we define the matrix  $C$  which controls variance of the noise in model

parameters (Refer to Eqn 2.7) as shown below

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \eta_{11} & 0 & \eta_{12} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \eta_{21} & 0 & \eta_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (3.5)$$

The parameter  $\eta$  can be changed to change the uncertainties in the model parameters  $Z$ . We simulated the average behavior of the optimal controller in all different tasks described below for three levels of variance viz. High Variance (HV), Medium Variance (MV) and Zero Variance (ZV).

Please refer to the appendix for the specific values of the parameters used for simulations.

## 3.2 Predictions for a point-to-point reaching task

### 3.2.1 Biased viscous curl force field with off-diagonal noise

We calculated the optimal policy predicted by the stochastic optimal control model for three different levels of uncertainties and simulated the average behavior of the sys-

tem using this optimal policy. These simulations are done for a clockwise biased viscous force field i.e. the elements of the matrix  $Z$  used were  $Z = \begin{bmatrix} 0 & 13 \\ -13 & 0 \end{bmatrix} Ns/m$ . The force field is such that if the hand velocity is in the positive  $y$  direction, the force field tends to push the hand towards positive  $x$  direction. The noise is called ‘off-diagonal noise’ since we introduced noise only in the non-zero parameters of the field matrix  $Z$  i.e. the off-diagonal elements. Thus, for these simulations, parameters  $\eta_{11}$  and  $\eta_{22}$  from Eqn 3.5 were set to zero for all simulations in this section while the parameters  $\eta_{12}$  and  $\eta_{21}$  were changed for different levels of uncertainties. The idea was to introduce noise only in the new parameters being learnt by the system.

Now, if we have a system which doesn’t care about the uncertainty in the model parameters, we would predict that the average behavior of the system for different levels of uncertainties is the same given that the expected value of the model parameters is the same for all cases. However, the stochastic optimal feedback control model, predicts different average behavior of the controller for different levels of uncertainties in the model parameters as shown in Fig 3.1. The simulations are done for a reach length of 9 cm to match the reach length in the experiments reported in next chapter.

A noteworthy prediction of the model is that the average trajectory for the certain clockwise force field (ZV) is curved towards the negative  $x$  direction. Hence, the average trajectory followed by the system is curved in direction opposite to the direction of the force field i.e. the system shows overcompensation on an average. This prediction is contradictory to the conventional idea of a straight line being the desired

trajectory of the motor system for making any point-to-point reaching movements. As shown in the following chapter, we find that this prediction indeed matches the average human behavior in a biased clockwise viscous curl force field.

It is a well accepted fact that human beings show symmetric bell shaped speed profiles when making point-to-point reaching movements [9][23]. From the simulations, it is clear that the model predicts a symmetric speed profile for the zero noise case as expected. However, as the level of uncertainty increases, the peak speed increases and the speed profile becomes more and more skewed. Note that in this particular case, the state-dependent component of the noise in system dynamics is directly proportional to the velocities  $V_x$  and  $V_y$ . Thus, lower the velocities, lower the value of the absolute noise in system dynamics. Now, the objective of the controller is to be at the target at the movement completion time  $t_f$ . Now since the absolute value of the noise is proportional to the velocities  $V_x$  and  $V_y$ , it is easier to achieve the goal if you increase the speed in the initial phase of the movement and cover major portion of the distance without caring much about how you cover the distance and then slow down so that the perturbation is lower in magnitude as well as more certain towards the end of the movement. This gives you better control over your movement towards the end which is essential since you have the spatial constraint of being present in the target to be able to successfully complete the task.

Another important difference is the separation of the average trajectories for different levels of uncertainties. As the level of the uncertainty increases, the average

trajectory shown by the system becomes more and more straighter.

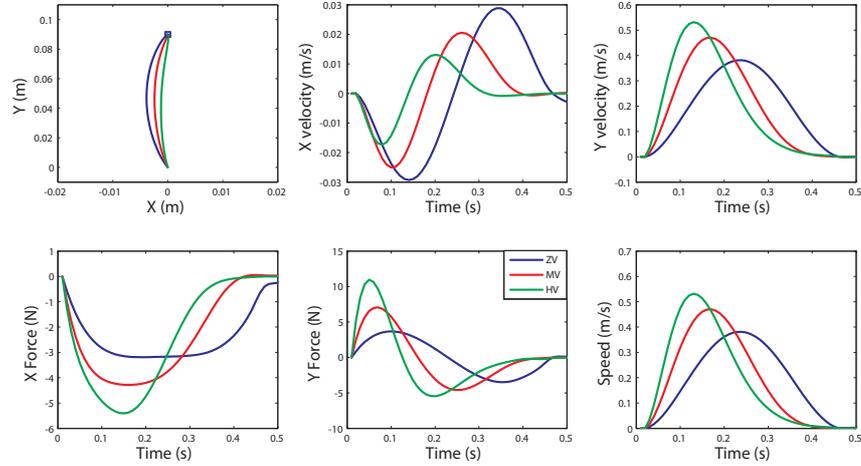


Figure 3.1 Simulation results for biased clockwise viscous curl force field with off-diagonal noise. The plots show the progression of different components of state  $x$  over the course of an average movement. The Blue, Red and Green curves correspond to the Zero Variance, Medium Variance and High Variance cases respectively.

We did the same simulation for counterclockwise field viscous curl force field as well. Here the elements of the matrix  $Z$  are such that if hand velocity is in the positive  $y$  direction, the force field tends to push the hand towards negative  $x$  direction. The actual elements of matrix  $Z$  used were  $Z = \begin{bmatrix} 0 & -13 \\ 13 & 0 \end{bmatrix} Ns/m$ . The simulation results are indicated in Fig 3.2

The predictions are essentially the same as those for the clockwise force field, except the curvature of the average trajectories is in the opposite direction. Since,

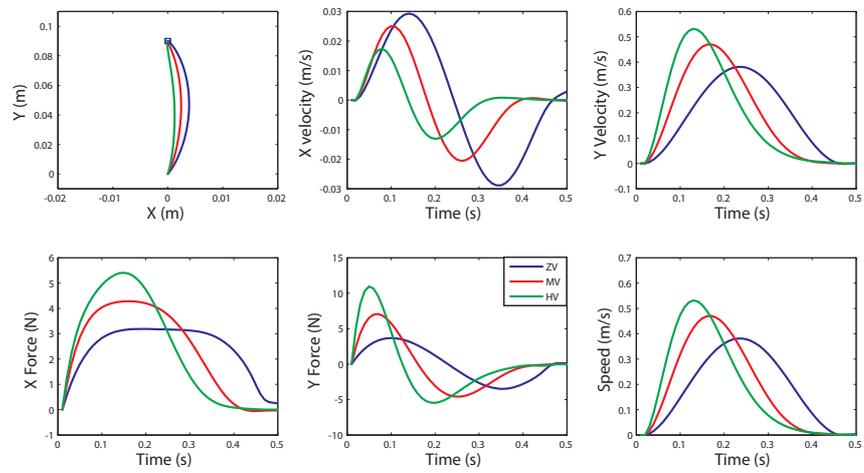


Figure 3.2 Simulation results for biased counterclockwise viscous curl force field with off-diagonal noise. The plots show the progression of different components of state  $x$  over the course of an average movement. The Blue, Red and Green curves correspond to the Zero Variance, Medium Variance and High Variance cases respectively.

our model is an approximation of the limb used for making reaching movements, we decided to make sure that the overcompensation observed in human behavior for the clockwise field is not just a result of the nonlinear dynamics of the human arm.

The simulations in Fig 3.1 are based on the assumption that the controller has a perfectly adapted internal model of its and the surrounding environment's working. This, however, need not always be true. In the case of the biased force fields, presumably, the internal model starts from an estimate of zero bias and slowly increases its estimate of bias in the force field based on its observations. Fig 3.3 shows the effect on the average behavior of the system if the optimal policy followed by the controller is based on incompletely adapted internal model i.e. the force field bias estimate of the control system is a fraction of the actual bias of the force field. These simulations are done for a fixed level of high uncertainty (HV) from Fig 3.1. The average trajectory predicted by the model is close to a straight line if the controller's estimate regarding the bias of the curl force field is correct, the same as that in Fig 3.1. However, as the learned fraction of the bias becomes smaller, the average trajectory shows a peculiar overcompensation during early part of the movement and an undercompensation towards the later part of the movement. The significance of these simulations will become clear when we analyze human behavior in a biased curl force field with high uncertainty.

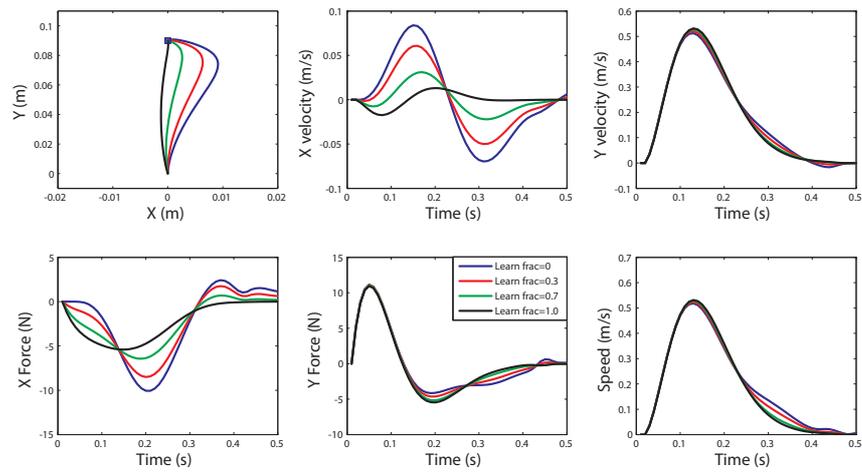


Figure 3.3 Simulation results for incomplete learning of a biased clockwise viscous curl force field with off-diagonal noise. The plots show the progression of different components of state  $x$  over the course of an average movement. The Blue, Red and Green curves correspond to the Zero Variance, Medium Variance and High Variance cases respectively.

### 3.2.2 Biased viscous curl force field with diagonal noise

All the simulation conditions were exactly the same for this section as in the previous section except that now we introduced noise in the diagonal components of the matrix  $Z$  i.e. the components of the matrix  $Z$  which were at the default zero value and the controller didn't need to learn them. The idea was to see if there are any differences in the optimal policy followed by the learner if the uncertainty is increased in the parameters whose bias is already learnt by the learner vs the case in which the learner is learning a parameter which is noisy to begin with. Fig 3.4 shows the simulation results for this case. The predictions are quite similar to the predictions for the off-diagonal noise case. Even in this case, the average trajectory tends to become straighter and the speed profile tends to become more skewed with higher peak speed as the variance level increases.

## 3.3 Predictions for a via-point reaching task

To substantiate our hypothesis about the role of learner's uncertainty in the optimal policy followed by the learner to do a motor task, we decided to test the predictions of the model on a different motor task i.e. the via-point task. In this task, the learner is supposed to make a reaching movement such that he/she has to cross a via-point on the path to the target at a pre-specified via-point time and then reach the target before certain pre-specified movement time. The uncertainty introduced

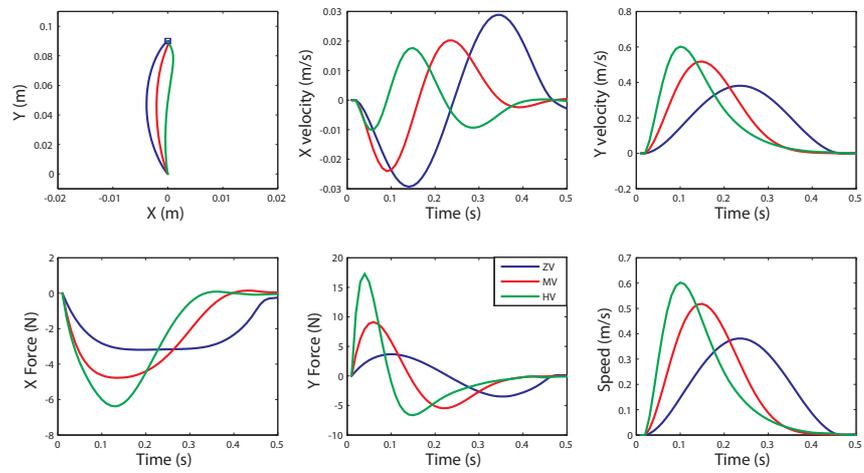


Figure 3.4 Simulation results for biased clockwise viscous curl force field with diagonal noise. The plots show the progression of different components of state  $x$  over the course of an average movement. The Blue, Red and Green curves correspond to the Zero Variance, Medium Variance and High Variance cases respectively.

in the force field in this task is exactly the same as that used for reaching task in a biased force field with off-diagonal noise. The force field used, however, was an unbiased force field i.e. the expected value of the elements of matrix  $Z$  defining the force field was zero. For the simulations corresponding to the via-point task, the state vector needs to have an estimate of the via-point. So, the state vector is changed to

$$\mathbf{x} = [P_x, V_x, P_y, V_y, F_x, F_y, via_x, via_y, T_x, T_y] \quad (3.6)$$

where all other symbols have the same meaning as specified in the previous section and  $via_x$  and  $via_y$  are the  $x$  and  $y$  co-ordinates of the via-point. Since the position of the via-point doesn't change over the duration of the movement, the changes to be made to the matrices  $A$  and  $B$  are simple. The matrices are set so that the position of via-point remains constant throughout the movement duration. The system knows the correct location of the via-point through proper initialization of the state vector  $\mathbf{x}$ .

The matrix  $Q$  defining the 'tracking cost' (Eqn 2.9) is such that the controller is penalized for not being at the via-point at the pre-specified via-point time  $t_v$ , and for not being at the target at the pre-specified movement completion time  $t_f$ . The matrix  $Q$  is a zero matrix at all other time points. The matrix  $R$  is the same as that used for the simulations corresponding to the point-to-point reaching task.

Fig 3.5 shows the simulation results. The simulation results show that the average trajectory followed by the system is the same for all levels of uncertainties. The main difference in the control policy, however, shows up in the speed profiles for different

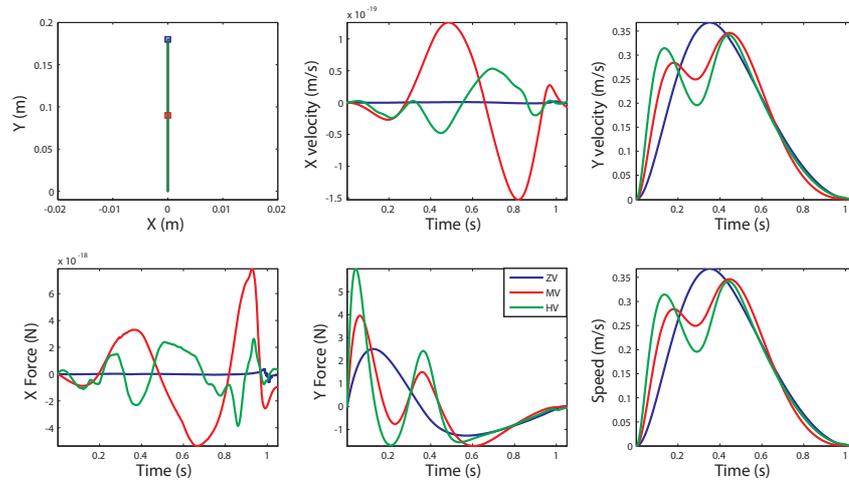


Figure 3.5 Simulation results for the via-point task with unbiased force field and off-diagonal noise. The plots show the progression of different components of state  $x$  over the course of an average movement. The Blue, Red and Green curves correspond to the Zero Variance, Medium Variance and High Variance cases respectively.

uncertainties. The objective of the task is to pass through the via-point at a specific time point  $t_v$  and then reach the target before time point  $t_f$ . In the viscous curl force field as discussed in previous section, the uncertainty in the perturbing force is velocity dependent. So, it is in the best interest of the system to reduce its speed when coming close to the via-point to be able to satisfy the spatial constraint of being present within the via-point at the required time. One possible solution is to reduce the speed to zero at the via-point crossing time and then increasing speed again to reach the target. This, however, cannot always be feasible since there is an added constraint of minimizing the ‘control cost’ which goes higher if a lot of high amplitude motor commands are produced. Since reducing the speed to complete zero would require the system to suddenly accelerate to a high speed and then decelerate to meet the time constraint for the target, the optimal solution obtained is a tradeoff between the cost of not being at the via-point at the via-point crossing time  $t_v$  and the control cost.

# Chapter 4

## Experiments

### 4.1 Apparatus

For all experiments described in this thesis, right handed subjects sat on a chair in front of a 2-dimensional robotic arm, holding the handle of the robotic arm. The robotic arm could be controlled to produce forces on the subjects hands in a plane. The robot position and speed were sampled at the rate of 100 Hz. The targets were projected onto a white screen above the robotic arm using a DLP projector (Optima, EP739, refresh rate: 70Hz) mounted overhead. White screen covered the robotic arm and subjects hand holding the arm so that the subject couldn't view their arms. The chair could be adjusted in height and position so that subjects could comfortably move the robotic arm and comfortably view the target projected on the screen. All the experimental protocols were approved by the Johns Hopkins Medicine IRBs. All

the subjects recruited for the experiments were healthy right handed subjects.

## **4.2 Reach experiment with biased certain force field**

In traditional motor control literature straight line point-to-point movements are considered to be the “desired” way of the motor controller to make point-to-point reaching movements. There has been evidence from recent literature that this is not the case. A popular way to manipulate the motor environment of a person is to make the person do reaching movements holding a two degree of freedom robotic manipulandum. Using the robot, different force perturbations can be applied to the subject’s hand which essentially means temporarily changing the dynamics of the surrounding motor environment for the subject. A common perturbation used is the viscous curl force field[18]. To probe the extent of learning of the new environment, the experiments usually have “catch” trials dispersed randomly between training trials [6][19][18]. No force perturbation is applied during the “catch” trials. A recent article shows that even after training people for a long time in a novel environment, their trajectories didn’t return to straight line trajectories or the so called “desired trajectory”, if no catch trials were used during training[5]. In this particular study [5], subjects were trained over three days but in multiple directions. Also, each day, training began with a null block of movements i.e. a set of movements without any

force perturbations, which would have the effect of washout on the learning of making movement in the force field. So one might argue that the training might not be long enough for any particular direction for the controller to be able to learn to achieve straight line trajectory in the novel environment.

To address this issue, we decided to do a simple study to check if this was indeed true. We trained subjects for three consecutive days for making reaching movements in a single direction. The subjects were trained for making 9 cm reaching movements over three days in a zero variance clockwise viscous curl force field. The matrix  $Z$  (Eqn 3.1) defining the force field was  $Z = \begin{bmatrix} 0 & 13 \\ -13 & 0 \end{bmatrix} Ns/m$ . The table 4.1 shows the sequence of blocks over three days we followed for this experiment. In the table, ‘no field’ block refers to a block of training trials without any perturbation while ‘ZV FF’ refers to a block of trials with a zero variance clockwise viscous curl force field. On each trial, the subjects started from a square indicating starting position and made a reaching movement to a target indicated by another square on the screen. The subjects were told to imagine that they were playing a video game and the objective of the game was to reach the target before a specific time limit. If the subjects were able to reach the target in time, they were rewarded with a visual explosion of the target, while if they were slow, they were given feedback in the form a blue colored target. To motivate subjects to perform the task better, their score i.e. the number of explosions of target in a block, was constantly displayed on the screen. 8 subjects participated in this study.

Table 4.1 Reach Experiment Protocol (Biased Force Field with zero variance)

	Block 1	Block 2	Block 3	Block 4	Block 5
Day 1	<i>No Field</i>	<i>ZV FF</i>	<i>ZV FF</i>	<i>ZV FF</i>	<i>ZV FF</i>
Day 2	<i>ZV FF</i>	<i>ZV FF</i>	<i>ZV FF</i>	<i>ZV FF</i>	
Day 3	<i>ZV FF</i>	<i>ZV FF</i>	<i>ZV FF</i>	<i>ZV FF</i>	

On the first day, people were familiarized with the robotic manipulandum through a block of 150 movements with no force perturbations. This block was followed by 4 blocks of force field training. On the second day and third days, the subject had four blocks of force field training on each day. We hypothesized that if the objective of the controller is to converge to a ‘desired trajectory’, this long term training for three consecutive days in a single direction should be enough for the controller to learn the dynamics of the new environment and converge onto its desired trajectory by the end of training.

Fig. 4.1 shows the gist of the results of the experiment. In plot A of the figure, the black curve shows the average trajectory over subjects for the no force field condition. This curve was obtained by averaging the trajectories for each subject over the last 50 trials of Block 1 of Day 1 and then by averaging the average trajectory over subjects. The green and the blue curves were derived in a similar way from first 50 trials of the first force field training block (i.e. Day 1 Block 2) and last 50 trials of the last force field training block on the last day (i.e. Day 3 Block 4) respectively. The average trajectories were obtained after aligning all the trajectories so that their starting x-co-ordinate was the same. From the results, we can clearly see the dissociation

between the trajectory in the no force field condition (black curve) vs the trajectory after completion of training (blue curve). The average trajectory after completion of training shows peculiar overcompensation as predicted by the simulations for the reaching task in a zero variance clockwise viscous curl force field shown in Fig 3.1. The difference between average trajectories for the first and last training blocks is indicative of how the average trajectory changed after optimization to minimize the cost associated with making a movement in the changed dynamic environment was completed. The plot B in the Fig. 4.1 shows the overcompensation averaged over subjects for the no force field training block followed by that for last force field training block on all three days. Overcompensation was calculated as the maximum deviation of a trajectory from the straight line joining the starting point and the target in the direction opposite to the direction of the force field. We can clearly see a sustained overcompensation higher than the average overcompensation for the no force field training block on all three days.

The model of the arm we have is a simplistic linear model. So, one might suspect that the overcompensation for this particular direction of force field might be due to nonlinear dynamics of the arm rather than the result of formation of a new optimal control policy. So, to strengthen our belief in the reason for the curved trajectory not being a result of nonlinear dynamics of the arm, we tested behavior in an oppositely directed force field (counterclockwise force field). The sequence of blocks followed was exactly the same as that for clockwise force field with the only difference being

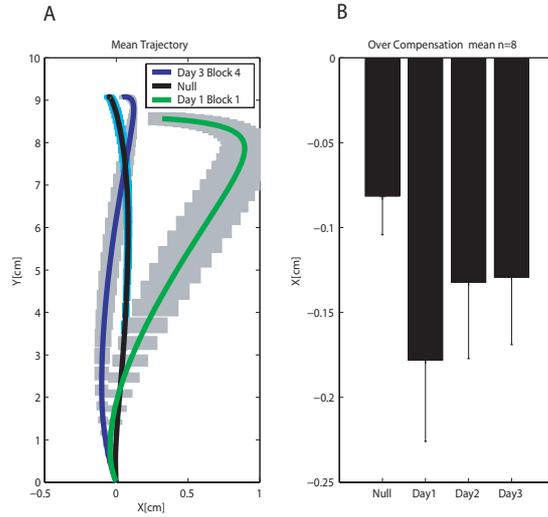


Figure 4.1 Average trajectory in a zero variance viscous curl clockwise force field. Plot A: Black curve shows the average trajectory in the no force field condition (averaged over last 50 trials of (Day 1 Block 1) and then averaged over subjects). The green and blue curves show similarly calculated averages for first 50 trials of (Day 1 Block 2) and last 50 trials of (Day 3 Block 4) respectively. The gray shading indicates standard error over subjects ( $n=8$ ). Plot B: Average overcompensation over subjects for (Day 1 Block 1) followed by that for last force field training block on each day. Error bars indicate standard error over subjects ( $n=8$ ).

the changed direction of the force field. The matrix  $B$  used for this case was  $B = \begin{bmatrix} 0 & -13 \\ 13 & 0 \end{bmatrix} Ns/m$ . Fig. 4.2 shows the results for this experiment. All the colors and curves have the same meaning as in Fig. 4.1. Thus, we have a strong indication that indeed the average trajectory in the force fields never converges to the average trajectory in the null field, making the possibility of the motor controller trying to approach a desired trajectory highly unlikely.

## 4.3 Reach experiment with biased variable force field

### 4.3.1 Off-Diagonal Noise

Now, to test the effects of increasing the learner’s uncertainty regarding the model parameters, we trained a group of subjects to make reaching movements in a high variance biased force field. The bias of the force field parameters was the same as that for the clockwise force field in the ‘Reach Experiment with biased certain force field’. However, the force field parameters were changed from trial to trial by adding gaussian noise to the constant bias of the force field parameters. The variance of the gaussian noise added was the same as that for the high variance case (HV) in the simulations of reaching movements in a biased viscous curl force field with off-diagonal noise (Fig 3.1). The sequence of blocks followed for this experiment is shown

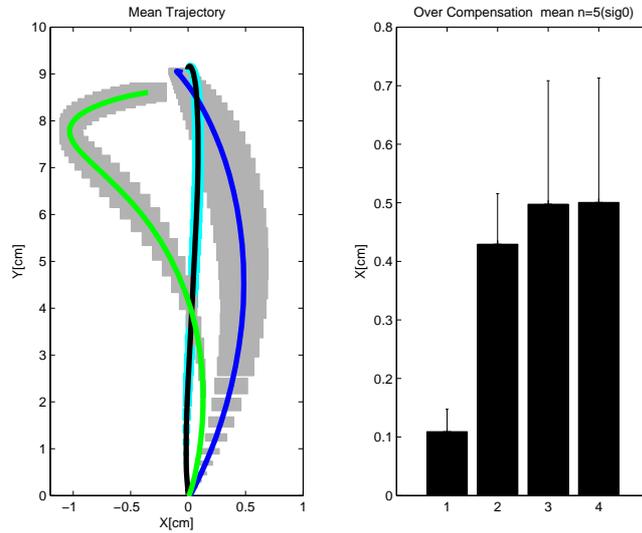


Figure 4.2 Average trajectory in a zero variance viscous curl clockwise force field. Plot A: Black curve shows the average trajectory in the no force field condition (averaged over last 50 trials of Block 1 on Day 1 and then averaged over subjects). The green and blue curves show similarly calculated averages for first 20 trials of (Day 1 Block 2) and last 50 trials of (Day 3 Block 4) respectively. The gray shading indicates standard error over subjects ( $n=5$ ). Plot B: Average overcompensation over subjects for no force field block(Day 1 Block 1) followed by that for last force field training block on each day. Error bars indicate standard error over subjects ( $n=5$ ).

in Table 4.2. The additional block of 50 trials on Day 3 (Day 3 Block 5) tested the behavior of the subjects in a zero variance force field. This block helped us compare the average behavior of the motor controller when trained in different conditions (high variance and zero variance force fields) in the same environment i.e. the zero variance force field. 8 subjects participated in this study.

Table 4.2 Reach Experiment Protocol (Biased Force Field with High Variance Off Diagonal noise)

	Block 1	Block 2	Block 3	Block 4	Block 5
Day 1	<i>No Field</i>	<i>HV FF</i>	<i>HV FF</i>	<i>HV FF</i>	<i>HV FF</i>
Day 2	<i>HV FF</i>	<i>HV FF</i>	<i>HV FF</i>	<i>HV FF</i>	
Day 3	<i>HV FF</i>	<i>HV FF</i>	<i>HV FF</i>	<i>HV FF</i>	<i>ZV FF</i>

Fig. 4.3 shows the comparison of average trajectories over subjects for different conditions. In plot A of the figure, the black curve shows the average trajectory over subjects for the no force field condition. This curve was obtained by averaging the trajectories for each subject over the last 50 trials of Block 1 of Day 1 and then by averaging the average trajectory over subjects. The red and the blue curves were derived in a similar way from 50 trials of Block 5 on Day 3 for the subjects who were trained in the high variance force field and last 50 trials of Block 4 Day 3 for the subjects trained in the certain biased force field(last force field training block on the last day) respectively. One can clearly see the separation of the average trajectories resulting due to training of the motor controller in the high variance vs the zero variance conditions. The average trajectory corresponding to training in high variance

condition, however, doesn't match the model prediction shown in Fig 3.1 exactly. We observe a peculiar undercompensation during the later part of the movement. In high variance training, the motor controller needs to learn two aspects of the force field, the bias as well as the variance of the force field parameters. Since there is no direct way for the controller to make observations of these parameters, it is not unlikely for the controller to not have been able to estimate the bias and variance correctly over the training that was provided. The simulations shown in Fig 3.3 can explain the average trajectory observed after training in the high variance force field. As the learned fraction of the bias in the force field reduces, we see more and more undercompensation during the later part of the trajectory in Fig 3.3. Plot B of Fig 4.3 shows the validation of another important prediction of the model. The model predicts that with increasing variance in the force field parameters, the speed profile should become more skewed and the peak speed should increase as compared to the speed profile for the zero variance force field condition. We find that this prediction is indeed true from the average speed profiles shown in Fig. 4.3. The speed profiles were obtained from the same data that was used for obtaining the average trajectories of the corresponding color. The gray shading indicates the standard error over subjects. One might question that why should the average trajectories match the predictions from based on incomplete learning of the model while the speed profiles match the predictions from optimization based on a completely learned model. However, if we look at the speed profile predictions based on different levels of incomplete learning,

we can see that they are more or less similar or overlapping during the earlier part of the movement which can explain the difference in the peaks of the speed profiles for the HV group vs the ZV group. The data, however, doesn't show significant reduction towards the end of the movement for the HV group as compared to the ZV group. If we look at the model predictions for incomplete learning, we see that the speed towards the end of the movement increases as the fraction of learning reduces. Thus, we suspect that the effect might be too small due to incomplete learning to be able to observe any statistically significant results.

Fig 4.4 shows a bar plot comparing the average peak speed over subjects for the zero variance condition (ZV)(calculated for the last 50 trials of Block 4 on Day 3) vs the high variance condition (HV)(calculated for the last 50 trials of Block 5 on Day 3). The error bars indicate the standard error over subjects.

### 4.3.2 Diagonal Noise

We wanted to study the difference in behavior when subjects were exposed to an environment where they had to learn new noisy parameters (i.e. the off-diagonal noise case) vs when they had to learn only the noise in the parameters (i.e. the diagonal noise case) as explained in the chapter on model predictions. The sequence of blocks followed was exactly the same as that for the Off-diagonal noise case. The only difference was that the noise parameters used were the same as those used for simulations of the high variance case (HV) in the simulations for the Clockwise biased

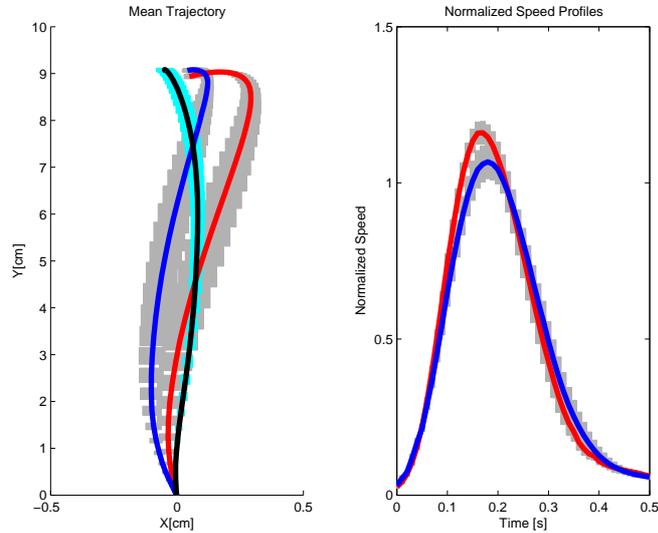


Figure 4.3 Average trajectory and speed profile comparison for movements in high off diagonal noise force field vs those in zero noise force field. Plot A: Black curve shows the average trajectory in the no force field condition (averaged over last 50 trials of (Day 1 Block 1) and then averaged over all subjects). The red curve is similarly calculated average from last 50 trials of the ZV FF block on Day 3 for subjects from the HV group and the blue curve is similarly calculated average from the last 50 trials of the last training block on Day 3 for subjects from the ZV group. The gray shading indicates the standard error over subjects.(n=16 for black curve and n=8 for blue and red curves) Plot B: Average speed profiles for the corresponding curves in plot A. The speed profiles were normalized with the average peak speed in the no force field condition before averaging over subjects.

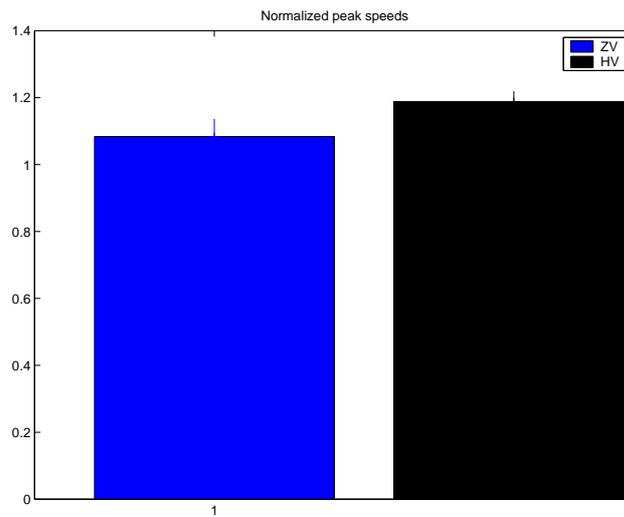


Figure 4.4 Peak speed comparison for movements in high off diagonal noise force field vs those in zero noise force field. Blue bar is the average peak speed over last 50 trials of (Day 3 Block 4) for all subjects from ZV group while black bar is the average peak speed over the 50 trials of (Day 3 Block 5) for all subjects from the HV group. The error bars indicate the standard error over subjects( $n=8$ )

viscous curl force field with diagonal noise (Fig 3.4). 6 subjects participated in this study.

Fig 4.5 shows the results of the experiment. Both Plots A and B in the Fig 4.5 correspond to the plots A and B in Fig 4.3. We see that the separation in the trajectories for the high variance (HV-Diag) case vs the zero variance case (ZV) is very similar to the off-diagonal noise case. However, the results for the change in the speed profile are much more prominent in this case. Here, we can see clear correspondence to both predictions of the model for the speed profile i.e. increase in peak speed during the earlier part of the movement and reduction in the speed towards the later part of the movement for the HV-Diag group as compared to the ZV group.

Fig 4.6 shows a bar plot comparing the average peak speed over subjects for the zero variance condition (ZV)(calculated for the last 50 trials of Block 4 on Day 3) vs the high variance condition (HV-Diag)(calculated for the last 50 trials of Block 5 on Day 3). The error bars indicate the standard error over subjects.

## **4.4 Via-point experiment with unbiased uncertain force field**

To substantiate our hypothesis about the role of learner’s uncertainty in the optimal policy followed by the learner to do a motor task, we did another experiment

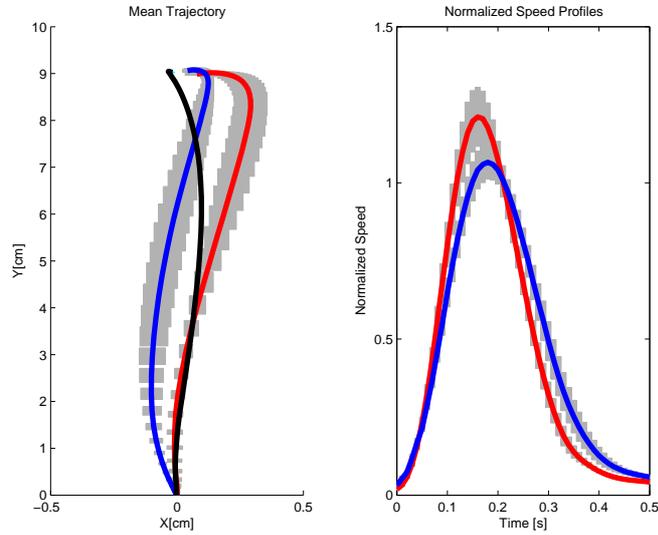


Figure 4.5 Average trajectory and speed profile comparison for movements in high diagonal noise force field vs those in zero noise force field. Plot A: Black curve shows the average trajectory in the no force field condition (averaged over last 50 trials of (Day 1 Block 1) and then averaged over all subjects). The red curve is similarly calculated average from last 50 trials of the ZV FF block on Day 3 for subjects from the HV-Diag group and the blue curve is similarly calculated average from the last 50 trials of the last training block on Day 3 for subjects from the ZV group. The gray shading indicates the standard error over subjects.(n=14 for black curve and n=8 for blue curve and n=6 for red curve) Plot B: Average speed profiles for the corresponding curves in plot A. The speed profiles were normalized with the average peak speed in the no force field condition before averaging over subjects.

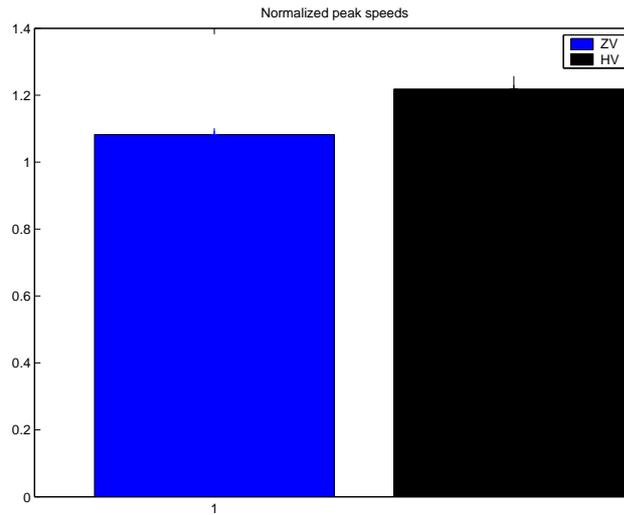


Figure 4.6 Peak speed comparison for movements in high diagonal noise force field vs those in zero noise force field. Blue bar is the average peak speed over last 50 trials of (Day 3 Block 4) for all subjects from ZV group while black bar is the average peak speed over the 50 trials of (Day 3 Block 5) for all subjects from the HV-Diag group. The error bars indicate the standard error over subjects (n=8 for blue bar and n=6 for black bar)

trying to probe the role of the learner’s uncertainty about the world in the control policy followed by the learner.

The task was a via-point task. The subject’s task was to pass through a via-point at a specific ‘via-point time’ and then reach the target before a certain ‘movement time’. The task constraints used i.e. via-point time, movement time, target and via-point distance were exactly the same as those used for the simulations related to the via-point task in the ‘Model Predictions’ section. After each trial the subject was given feedback about his/her actual timing by showing them arrows indicating difference between their actual time of passing the via point and the required time of passing through via point. The subjects were shown a blue filled square at the target if they were slower than required to reach the target. Correct timings were rewarded with visual explosions of the via point and the target. A score of the number of times subjects were able to get the timings right during each block was constantly displayed on the screen. Subjects were motivated to perform good on the task by offering them remuneration proportional to their score on the task.

The sequence of blocks for the experiment is shown in Table 4.3. The force field used for this experiment was an unbiased force field i.e. the expected value of the elements of matrix  $B$  (Eqn 3.1 defining the force field was zero. Subjects were trained to do the task for 150 trials during the first block in a zero variance force field (ZV force field). Now, since the field used was unbiased, ZV force field essentially means making movements in a null environment with no external forces acting on the hand.

Then for the following 4 blocks, subjects were trained to do the task in an unbiased high variance force field (HV force field). The variance of this force field was the same as that used for the HV case in the via-point task simulations shown in Fig ???. Then subjects were trained to make movement in the ZV force field again for subsequent three blocks.

Table 4.3 Experimental protocol for via-point task

Block 1	Block 2	Block 3	Block 4
<i>ZV force field</i>	<i>HV force field</i>	<i>HV force field</i>	<i>HV force field</i>
Block 5	Block 6	Block 7	Block 8
<i>HV force field</i>	<i>ZV force field</i>	<i>ZV force field</i>	<i>ZV force field</i>

Fig. 4.7 shows the results of the experiment. In plot A of the figure, the black curve shows the average trajectory over subjects for the no force field condition in the first block. This curve was obtained by averaging the trajectories for each subject over the last 50 trials of Block 1 and then by averaging the average trajectory over subjects. The red curve was derived in a similar way from last 50 trials of the last force field training block (i.e. Block 5). Plot B from Fig. 4.7 shows the average speed profiles for the same blocks as those for the average trajectories in plot A. The gray shading in both the plots indicates the standard error over subjects. The average trajectories are more or less similar for the high variance (HV) and zero variance(ZV) conditions as predicted by the model. But the important thing is dramatic change in the average speed profile from the ZV condition to the HV condition. As predicted by the model (Fig. 3.5), to reduce the variability of position at the via-point, the learner

should start reducing his/her speed when close to the via-point and then increase the speed again after crossing the via-point to reach the target in time. This leads to generation of a characteristic speed profile with two peaks, as can be clearly seen in the results of the experiment for the HV condition.

Fig. 4.8 shows a bar plot comparing the average speeds at the via-point crossing time over subjects for the ZV condition (calculated for the last 50 trials of Block 1) vs the HV condition (calculated for the last 50 trials of Block 5). The error bars indicate the standard error over subjects.

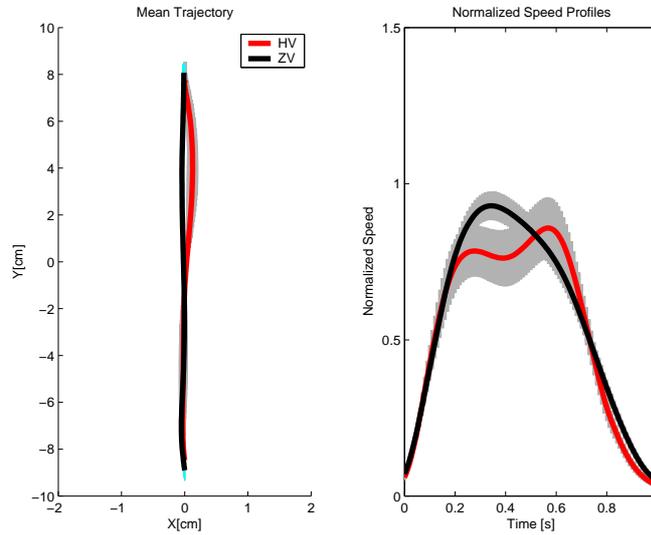


Figure 4.7 Average trajectory and speed profile comparison for the via-point task in zero variance unbiased force field vs that in high variance unbiased force field. Plot A: The black curve shows the average trajectory for the zero variance force field case (averaged over last 50 trials of Block 1 and then averaged over subjects). The red curve shows similarly calculated average trajectory for last high variance force field training block (Block 5). Plot B: The speed profiles for corresponding average trajectories in plot A. The gray shading indicates standard error over subjects( $n=11$ ).The speed profiles were normalized with the average peak speed in the zero variance unbiased force field condition before averaging over subjects.

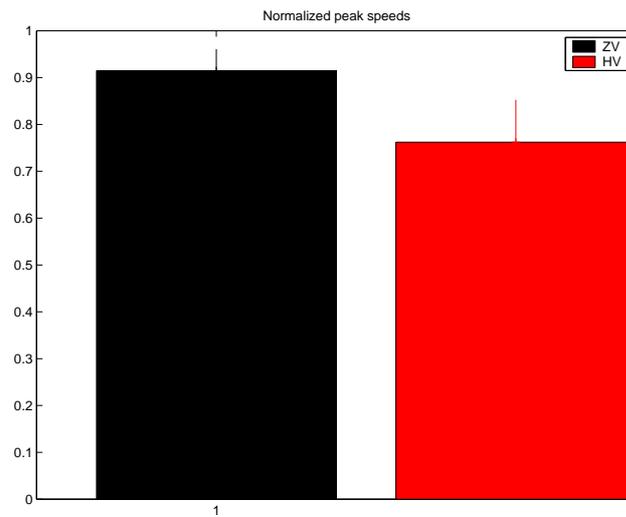


Figure 4.8 Via-point speed comparison for via-point task in high variance unbiased force field vs that in zero noise unbiased force field. Black bar is the average via-point speed over last 50 trials of Block 1(ZV) for all subjects while red bar is the average via-point speed over the 50 trials of Block 5(HV) for all subjects. The error bars indicate the standard error over subjects(n=11)

# Chapter 5

## Discussion and Conclusion

Traditionally motor control theories have focused on trying to explain the observed human behavior in perturbed motor environments by proposing several different criteria for optimization. [10][23][9]. However, all these theories proposed that the redundancy in the possible ways of successful achievement of a behavioral goal is removed by calculation of a 'desired trajectory' during a 'trajectory planning' phase. The job of the motor controller is to try to execute this 'desired trajectory' as faithfully as possible for guaranteed achievement of the goal. However, based on the observation that the trial to trial variability is higher in task irrelevant movement parameters as compared to task relevant movement parameters, Todorov and Jordan [21] proposed optimal feedback control as a theory of motor control. According to this theory the controller instead of calculating a 'desired trajectory' calculates an 'optimal policy' for successful achievement of a behavioral goal in a novel motor environment. This

optimal policy helps the controller to delay taking the optimal action till the last moment to take advantage of every possible task completion opportunity created due to deviation from the average trajectory.

Human beings are believed to have 'Internal Models' of the working of their own body and the surrounding environment. Having such predictive models is necessary for predicting the consequences of one's own actions, which is extremely useful for online control of a movement when the only feedback available comes from noisy delayed sensors. Changes in these internal models should presumably lead to re-optimization of the policy followed by the motor controller to achieve a behavioral goal. We hypothesized that there is no reason for the internal models to only have a representation of the expected value of it's parameters. The internal models should infact have a representation of uncertainty in it's parameters to be able to better accomplish motor tasks in a novel motor environment.

Solution of the 'stochastic optimal feedback control' optimization problem by Todorov [20] gave us a framework to think about the effects of biologically plausible noises in the internal model on the optimal motor policies followed by the motor controller to successfully accomplish a behavioral goal. Through a small change in Todorov's model formulation, we were able to obtain a representation of model parameter uncertainty in the internal model. This stochastic optimal control problem with an enhanced internal model could be done by simply following the methodology in [20].

Now that we had a framework to understand the consequences of various kinds of uncertainties in the internal model, we could predict the optimal policies that the controller should ideally follow in various situations. We focused our attention on the viscous curl force field paradigm [18]. We predicted the average behavior of the controller for optimal policies formulated with internal models of different forms of uncertainties and validated them experimentally.

One of the important outcomes of our experiments has been substantial evidence contrary to the idea of the objective of the controller being to converge on to a desired straight line trajectory even when making reaching movements in a viscous curl force field. Traditionally due to the presence of catch trials or the no force field trials during the training phase for force fields, the curvature in the movements was probably reduced and went unnoticed. Theoretically, adding catch trials is like increasing the learner's uncertainty regarding the force field parameters. Now, according to our model predictions for reaching task in a biased viscous curl force field (Fig 3.1), the average trajectory becomes more and more straighter as the learner's uncertainty regarding the model parameters increases. This could explain the straighter trajectories observed in training with catch trials [5]. Our experiment with constant viscous curl force field training for making movements in a single direction for three days showed strong evidence for the change in the optimal policy followed by the controller for the changed motor environment (i.e. the force field), in accordance with the model predictions, and no tendency to converge on to the straight line trajectory. This pre-

diction of curved trajectory was observed in human behavior for both clockwise and counterclockwise force fields.

The most important result of the experiments was the demonstration of the effect of change in mere uncertainty (with constant expected value) of model parameters on change in optimal policy followed by the motor controller. Both the tasks that we trained people on (i.e. the reaching task and the via-point task) demanded presence at particular spatial locations at a specific times during the movement for successful completion of the task. The viscous curl force field is a velocity dependent force field by its very definition. So introducing uncertainty in the force field parameters implies introducing an uncertainty in the internal model proportional to the velocities. So, intuitively one would expect the controller to try to minimize movement speed just before arriving at constrained spatial locations in order to minimize uncertainty at these constrained spatial locations, which is exactly what the model predicted and we qualitatively validated these predictions through experiments.

## 5.1 Future work

The dynamical system we used for our simulations was a linear dynamical problem since the solution to the stochastic optimal control problem was only available for linear dynamical systems. The model based on a linear dynamical system although not having an accurate representation of the dynamics of the arm, serves as an excellent qualitative predictor for human reaching behavior in novel environments. However,

with recent approximations to the solution of the stochastic optimal control problem available for non-linear dynamical systems[22], we can have better model predictions for behavior in novel environments which might even be able to numerically match observed behavior.

Our analysis of the experimental results has focused on the average trajectories observed after completion of the optimization process. However, we still lack an understanding of how exactly this optimization might be progressing over training time. More experiments specifically designed to address this issue need to be done to improve our understanding of the formulation of optimal motor policies in novel environments.

One more interesting aspect of the movements in environments with different levels of uncertainties in dynamical model parameters is the variation of the ‘variance’ of hand path over the course of the movements. For the simple point-to-point reaching task, one would expect the movement variance to be small at the starting position and the target and to be higher in the middle part of the movement since the exact trajectory followed to move from the starting position to the target doesn’t interfere with task constraint i.e. being at the target before certain time after starting the movement [21]. Similarly, the variance of the trajectory for the parts of the movement between the starting position and the via-point and the via-point and the target is expected to be greater than that at the starting position, via-point and the target. However, we still don’t know how exactly the variance of the trajectory varies with

change in the dynamical model parameter uncertainties.

# Appendix A

## Solving for the Optimal Controller and the Optimal Estimator with model parameter uncertainty

### A.1 Optimal Control Policy

The system dynamics are expressed as following

$$\text{Dynamics } \mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t + \xi_t + \sum_{i=1}^d \epsilon_t^i C_i \mathbf{x}_t + \sum_{i=1}^d \gamma_t^i \bar{C}_i \mathbf{u}_t \quad (\text{A.1})$$

$$\text{Feedback } \mathbf{y}_t = H\mathbf{x}_t + \omega_t + \sum_{i=1}^d \epsilon_t^i D_i \mathbf{x}_t \quad (\text{A.2})$$

$$\text{Cost per step } \text{cps}_t = \mathbf{x}_t^T Q_t \mathbf{x}_t + \mathbf{u}_t^T R \mathbf{u}_t \quad (\text{A.3})$$

All the notations have the same meaning as described in Table 2.1. The state estimate of the dynamic system is assumed to be updated according a linear recursive filter for analytical tractability

$$\hat{\mathbf{x}}_{t+1} = A\hat{\mathbf{x}}_t + B\mathbf{u}_t + K_t(\mathbf{y}_t - H\hat{\mathbf{x}}_t) + \eta_t \quad (\text{A.4})$$

We define the estimation error as  $\mathbf{e}_t = \mathbf{x}_t - \hat{\mathbf{x}}_t$ . Now, we can show through induction that the optimal cost-to-go function or the cost expected to accumulate under the optimal control law after a time step  $t$  has the quadratic form

$$v_t(\mathbf{x}_t, \hat{\mathbf{x}}_t) = \mathbf{x}_t^T S_t^x \mathbf{x}_t + (\mathbf{x}_t - \hat{\mathbf{x}}_t)^T S_t^e (\mathbf{x}_t - \hat{\mathbf{x}}_t) + s_t \quad (\text{A.5})$$

$$= \mathbf{x}_t^T S_t^x \mathbf{x}_t + \mathbf{e}_t^T S_t^e \mathbf{e}_t + s_t \quad (\text{A.6})$$

Now, consider an optimal control law denoted by  $\mathbf{u}_t = \pi(\hat{\mathbf{x}}_t)$ . Let  $v_t^\pi(\mathbf{x}_t, \hat{\mathbf{x}}_t)$  be the cost-to-go function corresponding to the optimal control law. Since the control law  $\pi(\hat{\mathbf{x}}_t)$  is optimal for all time points  $t+1, \dots, n$ , the cost-to-go function satisfies  $v_{t+1}^\pi = v_{t+1}$ . Given this condition, the cost-to-go function satisfies the Bellman equation: [20]

$$v_t^\pi(\mathbf{x}_t, \hat{\mathbf{x}}_t) = \mathbf{x}_t^T Q_t \mathbf{x}_t + \pi(\hat{\mathbf{x}}_t)^T R \pi(\hat{\mathbf{x}}_t) + E[v_{t+1}(\mathbf{x}_{t+1}, \hat{\mathbf{x}}_{t+1}) | \mathbf{x}_t, \hat{\mathbf{x}}_t, \pi]$$

From Eqn A.4 and Eqn A.2

$$\hat{\mathbf{x}}_{t+1} = A\hat{\mathbf{x}}_t + B\mathbf{u}_t + K_t \left( H\mathbf{x}_t + \omega_t + \sum_{i=1}^d \epsilon_t^i D_i \mathbf{x}_t - H\hat{\mathbf{x}}_t \right) + \eta_t \quad (\text{A.7})$$

Hence, using Eqn A.1 and the definition of  $\mathbf{e}_t$ , we get

$$\mathbf{e}_{t+1} = (A - K_t H) \mathbf{e}_t + \xi_t - K_t \omega_t - \eta_t + \sum_{i=1}^d (\epsilon_t^i C_i \mathbf{x}_t + \gamma_t^i \bar{C}_i \pi(\hat{\mathbf{x}}_t) - \epsilon_t^i K_t D_i \mathbf{x}_t) \quad (\text{A.8})$$

Now we will calculate the expected values and the covariance matrices of the random variables of our interest  $\mathbf{x}_{t+1}$  and  $\mathbf{e}_{t+1}$ , which shall be useful in calculating the cost-to-go function.

$$E[\mathbf{x}_{t+1}|\mathbf{x}_t, \hat{\mathbf{x}}_t, \pi] = A\mathbf{x}_t + B\pi(\hat{\mathbf{x}}_t) \quad (\text{A.9})$$

$$E[\mathbf{e}_{t+1}|\mathbf{x}_t, \hat{\mathbf{x}}_t, \pi] = (A - K_t H)\mathbf{e}_t \quad (\text{A.10})$$

$$Cov[\mathbf{x}_{t+1}|\mathbf{x}_t, \hat{\mathbf{x}}_t, \pi] = \Omega^\xi + \sum_{i=1}^d \left( C_i \mathbf{x}_t \mathbf{x}_t^T C_i^T + \bar{C}_i \pi(\hat{\mathbf{x}}_t) \pi(\hat{\mathbf{x}}_t)^T \bar{C}_i^T \right) \quad (\text{A.11})$$

$$Cov[\mathbf{e}_{t+1}|\mathbf{x}_t, \hat{\mathbf{x}}_t, \pi] = \Omega^\xi + K_t \Omega^\omega K_t^T + \Omega^\eta + \sum_{i=1}^d \left( C_i \mathbf{x}_t \mathbf{x}_t^T C_i^T + \bar{C}_i \pi(\hat{\mathbf{x}}_t) \pi(\hat{\mathbf{x}}_t)^T \bar{C}_i^T + K_t D_i \mathbf{x}_t \mathbf{x}_t^T D_i^T K_t^T \right) \quad (\text{A.12})$$

Using the relations derived in Eqn A.5, Eqn A.1 and the expected values and covariances we just calculated, we get

$$v_t^\pi(\mathbf{x}_t, \hat{\mathbf{x}}_t) = \mathbf{x}_t^T Q_t \mathbf{x}_t + \pi(\hat{\mathbf{x}}_t)^T R \pi(\hat{\mathbf{x}}_t) + E[\mathbf{x}_{t+1}^T S_{t+1}^x \mathbf{x}_{t+1} + \mathbf{e}_{t+1}^T S_{t+1}^e \mathbf{e}_{t+1} + s_{t+1}] \quad (\text{A.13})$$

Now using the relation

$$E[\mathbf{x}^T A \mathbf{x}] = Tr(A\Sigma) + c^T A c \quad (\text{A.14})$$

where  $\Sigma = Var[x]$  and  $c = E[x]$ , we get

$$\begin{aligned}
v_t^\pi(\mathbf{x}_t, \hat{\mathbf{x}}_t) = & \mathbf{x}_t^T Q_t \mathbf{x}_t + \pi(\hat{\mathbf{x}}_t)^T R \pi(\hat{\mathbf{x}}_t) + (A\mathbf{x}_t + B\pi(\hat{\mathbf{x}}_t))^T S_{t+1}^\mathbf{x} (A\mathbf{x}_t + B\pi(\hat{\mathbf{x}}_t)) + \\
& ((A - K_t H) \mathbf{e}_t)^T S_{t+1}^\mathbf{e} ((A - K_t H) \mathbf{e}_t) + s_{t+1} + \\
& Tr \left( S_{t+1}^\mathbf{x} \left( \Omega^\xi + \sum_{i=1}^d (C_i \mathbf{x}_t \mathbf{x}_t^T C_i^T + \bar{C}_i \pi(\hat{\mathbf{x}}_t) \pi(\hat{\mathbf{x}}_t)^T \bar{C}_i^T) \right) \right) + \\
& Tr \left( S_{t+1}^\mathbf{e} \left( \Omega^\xi + K_t \Omega^\omega K_t^T + \Omega^\eta + \sum_{i=1}^d \bar{C}_i \pi(\hat{\mathbf{x}}_t) \pi(\hat{\mathbf{x}}_t)^T \bar{C}_i^T \right) \right) + \\
& Tr \left( S_{t+1}^\mathbf{e} \left( \sum_{i=1}^d (C_i \mathbf{x}_t \mathbf{x}_t^T C_i^T + K_t D_i \mathbf{x}_t \mathbf{x}_t^T D_i^T K_t^T) \right) \right) \tag{A.15}
\end{aligned}$$

Using the property

$$Tr(xy^T) = y^T x$$

where  $x$  and  $y$  are vectors, we get

$$\begin{aligned}
v_t^\pi(\mathbf{x}_t, \hat{\mathbf{x}}_t) = & \mathbf{x}_t^T \left( Q_t + A^T S_{t+1}^\mathbf{x} A + \sum_{i=1}^d (C_i^T (S_{t+1}^\mathbf{x} + S_{t+1}^\mathbf{e}) C_i + D_i^T K_t^T S_{t+1}^\mathbf{e} K_t D_i) \right) \mathbf{x}_t + \\
& \mathbf{e}_t^T (A - K_t H)^T S_{t+1}^\mathbf{e} (A - K_t H) \mathbf{e}_t + s_{t+1} + \\
& \pi(\hat{\mathbf{x}}_t)^T \left( R + B^T S_{t+1}^\mathbf{x} B + \sum_{i=1}^d \bar{C}_i^T (S_{t+1}^\mathbf{x} + S_{t+1}^\mathbf{e}) \bar{C}_i \right) \pi(\hat{\mathbf{x}}_t) + \\
& 2\pi(\hat{\mathbf{x}}_t)^T B^T S_{t+1}^\mathbf{x} A \mathbf{x}_t + Tr \left( S_{t+1}^\mathbf{x} \Omega^\xi + S_{t+1}^\mathbf{e} (\Omega^\xi + K_t \Omega^\omega K_t^T + \Omega^\eta) \right) \tag{A.16}
\end{aligned}$$

Let us define some symbols to make our equations compact

$$\begin{aligned}
N_t &= \sum_{i=1}^d (C_i^T (S_{t+1}^{\mathbf{x}} + S_{t+1}^{\mathbf{e}}) C_i + D_i^T K_t^T S_{t+1}^{\mathbf{e}} K_t D_i) \\
\bar{C}_t &= \sum_{i=1}^d (\bar{C}_i^T (S_{t+1}^{\mathbf{x}} + S_{t+1}^{\mathbf{e}}) \bar{C}_i) \\
M_t &= S_{t+1}^{\mathbf{x}} \Omega^\xi + S_{t+1}^{\mathbf{e}} (\Omega^\xi + K_t \Omega^\omega K_t^T + \Omega^\eta)
\end{aligned}$$

Using these symbols, the cost-to-go function becomes

$$\begin{aligned}
v_t^\pi(\mathbf{x}_t, \hat{\mathbf{x}}_t) &= \mathbf{x}_t^T (Q_t + A^T S_{t+1}^{\mathbf{x}} A + N_t) \mathbf{x}_t + \mathbf{e}_t^T (A - K_t H)^T S_{t+1}^{\mathbf{e}} (A - K_t H) \mathbf{e}_t + \\
& s_{t+1} + \pi(\hat{\mathbf{x}}_t)^T (R + B^T S_{t+1}^{\mathbf{x}} B + \bar{C}_t) \pi(\hat{\mathbf{x}}_t) + 2\pi(\hat{\mathbf{x}}_t)^T B^T S_{t+1}^{\mathbf{x}} A \mathbf{x}_t + \\
& + Tr(M_t)
\end{aligned} \tag{A.17}$$

The cost-to-go function is, however, a function of the real state  $\mathbf{x}_t$ , which is not available to the controller. The only thing available to the controller is the state estimate  $\hat{\mathbf{x}}_t$ . So, we take the expected value of the cost-to-go function over the true state and minimize it w.r.t. the control policy  $\pi$

$$\begin{aligned}
E[v_t^\pi(\mathbf{x}_t, \hat{\mathbf{x}}_t) | \hat{\mathbf{x}}_t] &= constant + \pi(\hat{\mathbf{x}}_t)^T (R + B^T S_{t+1}^{\mathbf{x}} B + \bar{C}_t) \pi(\hat{\mathbf{x}}_t) + \\
& 2\pi(\hat{\mathbf{x}}_t)^T B^T S_{t+1}^{\mathbf{x}} A \hat{\mathbf{x}}_t
\end{aligned} \tag{A.18}$$

Minimizing this expectation w.r.t. the control policy  $\pi$ , we get the optimal control policy

$$\begin{aligned}
\mathbf{u}_t &= \pi(\hat{\mathbf{x}}_t) = -L_t \hat{\mathbf{x}}_t; \\
L_t &= (R + B^T S_{t+1}^{\mathbf{x}} B + \bar{C}_t)^{-1} B^T S_{t+1}^{\mathbf{x}} A
\end{aligned} \tag{A.19}$$

We found the optimal control policy but we still need to figure out how to obtain the matrices  $s_t, S_t^{\mathbf{x}}$  and  $S_t^{\mathbf{e}}$ . Substituting  $\mathbf{u}_t = -L_t \hat{\mathbf{x}}_t$  in the Eqn A.17

$$\begin{aligned} v_t^\pi(\mathbf{x}_t, \hat{\mathbf{x}}_t) &= \mathbf{x}_t^T (Q_t + A^T S_{t+1}^{\mathbf{x}} A + N_t) \mathbf{x}_t + \mathbf{e}_t^T (A - K_t H)^T S_{t+1}^{\mathbf{e}} (A - K_t H) \mathbf{e}_t + \\ &\quad s_{t+1} + (-L_t \hat{\mathbf{x}}_t)^T (R + B^T S_{t+1}^{\mathbf{x}} B + \bar{C}_t) (-L_t \hat{\mathbf{x}}_t) + \\ &\quad 2(-L_t \hat{\mathbf{x}}_t)^T B^T S_{t+1}^{\mathbf{x}} A \mathbf{x}_t + Tr(M_t) \end{aligned} \quad (\text{A.20})$$

Now using  $L_t^T (R + B^T S_{t+1}^{\mathbf{x}} B + \bar{C}_t) L_t = L_t^T B^T S_{t+1}^{\mathbf{x}} A = A^T S_{t+1}^{\mathbf{x}} B L_t$

$$\begin{aligned} v_t^\pi(\mathbf{x}_t, \hat{\mathbf{x}}_t) &= \mathbf{x}_t^T (Q_t + A^T S_{t+1}^{\mathbf{x}} A + N_t) \mathbf{x}_t + \mathbf{e}_t^T (A - K_t H)^T S_{t+1}^{\mathbf{e}} (A - K_t H) \mathbf{e}_t + \\ &\quad s_{t+1} + Tr(M_t) + \hat{\mathbf{x}}_t^T L_t^T B^T S_{t+1}^{\mathbf{x}} A \hat{\mathbf{x}}_t - 2\hat{\mathbf{x}}_t^T L_t^T B^T S_{t+1}^{\mathbf{x}} A \mathbf{x}_t + \\ &\quad \mathbf{x}_t^T L_t^T B^T S_{t+1}^{\mathbf{x}} A \mathbf{x}_t - \mathbf{x}_t^T L_t^T B^T S_{t+1}^{\mathbf{x}} A \mathbf{x}_t \end{aligned}$$

$$\begin{aligned} v_t^\pi(\mathbf{x}_t, \hat{\mathbf{x}}_t) &= s_{t+1} + Tr(M_t) + \mathbf{x}_t^T (Q_t + A^T S_{t+1}^{\mathbf{x}} A + N_t - A^T S_{t+1}^{\mathbf{x}} B L_t) \mathbf{x}_t \\ &\quad + \mathbf{e}_t^T (A - K_t H)^T S_{t+1}^{\mathbf{e}} (A - K_t H) \mathbf{e}_t + (\mathbf{x}_t - \hat{\mathbf{x}}_t)^T L_t^T B^T S_{t+1}^{\mathbf{x}} A (\mathbf{x}_t - \hat{\mathbf{x}}_t) \end{aligned}$$

$$\begin{aligned} v_t^\pi(\mathbf{x}_t, \hat{\mathbf{x}}_t) &= s_{t+1} + Tr(M_t) + \mathbf{x}_t^T (Q_t + A^T S_{t+1}^{\mathbf{x}} (A - B L_t) + N_t) \mathbf{x}_t + \\ &\quad \mathbf{e}_t^T \left( (A - K_t H)^T S_{t+1}^{\mathbf{e}} (A - K_t H) + A^T S_{t+1}^{\mathbf{x}} B L_t \right) \mathbf{e}_t \end{aligned}$$

Comparing with  $v_t^\pi(\mathbf{x}_t, \hat{\mathbf{x}}_t) = \mathbf{x}_t^T S_t^{\mathbf{x}} \mathbf{x}_t + \mathbf{e}_t^T S_t^{\mathbf{e}} \mathbf{e}_t + s_t$ , we get

$$s_t = s_{t+1} + Tr(M_t) \quad (\text{A.21})$$

$$S_t^{\mathbf{x}} = Q_t + A^T S_{t+1}^{\mathbf{x}} (A - B L_t) + N_t \quad (\text{A.22})$$

$$S_t^{\mathbf{e}} = (A - K_t H)^T S_{t+1}^{\mathbf{e}} (A - K_t H) + A^T S_{t+1}^{\mathbf{x}} B L_t \quad (\text{A.23})$$

Thus, we showed that the cost-to-go function remains in the assumed quadratic form shown in Eqn A.5 for any time step  $t$  given that it is true for the time step  $t + 1$ ,

completing the induction proof.

## A.2 Optimal Estimator

According to the assumption in the previous section of Kalman gains not being functions of  $\mathbf{x}$  and  $\hat{\mathbf{x}}$ , we need to minimize the unconditional expectation of the cost-to-go function  $v_{t+1}$  with respect to  $K_t$

$$E[v_{t+1}(\mathbf{x}_{t+1}, \hat{\mathbf{x}}_{t+1})] = E_{\mathbf{x}_t, \hat{\mathbf{x}}_t} [E[v_{t+1}(\mathbf{x}_{t+1}, \hat{\mathbf{x}}_{t+1}) | \mathbf{x}_t, \hat{\mathbf{x}}_t, L_t]] \quad (\text{A.24})$$

taking only the terms involving  $K_t$  in the expression for  $E[v_{t+1}(\mathbf{x}_{t+1}, \hat{\mathbf{x}}_{t+1}) | \mathbf{x}_t, \hat{\mathbf{x}}_t, L_t]$

$$\begin{aligned} f(K_t) &= ((A - K_t H) \mathbf{e}_t)^T S_{t+1}^e ((A - K_t H) \mathbf{e}_t) + \\ &\quad Tr(S_{t+1}^e (K_t \Omega^\omega K_t^T + K_t D_i \mathbf{x}_t \mathbf{x}_t^T D_i^T K_t^T)) \\ f(K_t) &= \mathbf{e}_t^T (A - K_t H)^T S_{t+1}^e (A - K_t H) \mathbf{e}_t + \\ &\quad Tr(S_{t+1}^e K_t (\Omega^\omega + D_i \mathbf{x}_t \mathbf{x}_t^T D_i^T) K_t^T) \end{aligned}$$

Using  $Tr(AB) = Tr(BA)$

$$\begin{aligned} f(K_t) &= \mathbf{e}_t^T (A - K_t H)^T S_{t+1}^e (A - K_t H) \mathbf{e}_t + \\ &\quad Tr\left(K_t \left(\Omega^\omega + \sum_{i=1}^d D_i \mathbf{x}_t \mathbf{x}_t^T D_i^T\right) K_t^T S_{t+1}^e\right) \end{aligned}$$

Using the Eqn A.14 and defining

$$a(K_t) = E_{\mathbf{x}_t, \hat{\mathbf{x}}_t} [f(K_t)]; \quad \Sigma_t^{\mathbf{x}} = E[\mathbf{x}_t \mathbf{x}_t^T]; \quad \Sigma_t^e = E[\mathbf{e}_t \mathbf{e}_t^T]$$

we get the unconditional expectation of the  $K_t$  dependent expression in the cost-to-go function

$$\begin{aligned}
a(K_t) &= \text{Tr} \left( (A - K_t H)^T S_{t+1}^e (A - K_t H) \Sigma_t^e \right) + \\
&\quad \text{Tr} \left( K_t \left( \Omega^\omega + \sum_{i=1}^d D_i \Sigma_t^x D_i^T \right) K_t^T S_{t+1}^e \right) \\
a(K_t) &= \text{Tr} \left( \text{constant} + H^T K_t^T S_{t+1}^e K_t H \Sigma_t^e - H^T K_t^T S_{t+1}^e A \Sigma_t^e - A^T S_{t+1}^e K_t H \Sigma_t^e \right) + \\
&\quad \text{Tr} \left( K_t \left( \Omega^\omega + \sum_{i=1}^d D_i \Sigma_t^x D_i^T \right) K_t^T S_{t+1}^e \right)
\end{aligned}$$

using  $\text{Tr}(ABC) = \text{Tr}(BCA) = \text{Tr}(CAB)$

$$\begin{aligned}
a(K_t) &= \text{Tr} \left( \text{constant} + K_t^T S_{t+1}^e K_t H \Sigma_t^e H^T - K_t^T S_{t+1}^e A \Sigma_t^e H^T - K_t H \Sigma_t^e A^T S_{t+1}^e \right) + \\
&\quad \text{Tr} \left( K_t \left( \Omega^\omega + \sum_{i=1}^d D_i \Sigma_t^x D_i^T \right) K_t^T S_{t+1}^e \right) \tag{A.25}
\end{aligned}$$

using the properties

$$\begin{aligned}
\frac{\partial \text{Tr}(XA)}{\partial X} &= A^T, \quad \frac{\partial \text{Tr}(X^T A)}{\partial X} = A \quad \text{and} \\
\frac{\partial \text{Tr}(X^T BXC)}{\partial X} &= \frac{\partial \text{Tr}(XCX^T B)}{\partial X} = BXC + B^T X C^T
\end{aligned}$$

we can minimize the unconditional expectation of the  $K_t$  dependent expression in the cost-to-go function w.r.t  $K_t$

$$\begin{aligned}
\frac{\partial \text{Tr}(a(K_t))}{\partial K_t} &= S_{t+1}^e K_t H \Sigma_t^e H^T + S_{t+1}^e K_t (H \Sigma_t^e H^T)^T - S_{t+1}^e A \Sigma_t^e H^T - \\
&\quad (H \Sigma_t^e A^T S_{t+1}^e)^T + 2S_{t+1}^e K_t \left( \Omega^\omega + \sum_{i=1}^d D_i \Sigma_t^x D_i^T \right)
\end{aligned}$$

$$\begin{aligned} \frac{\partial Tr(a(K_t))}{\partial K_t} = & 2S_{t+1}^e K_t \left( H \Sigma_t^e H^T + \Omega^\omega + \sum_{i=1}^d D_i \Sigma_t^x D_i^T \right) \\ & - 2S_{t+1}^e A \Sigma_t^e H^T \end{aligned}$$

setting the derivative to zero and solving for  $K_t$ , we get

$$K_t = A \Sigma_t^e H^T \left( H \Sigma_t^e H^T + \Omega^\omega + \sum_{i=1}^d D_i \Sigma_t^x D_i^T \right)^{-1} \quad (\text{A.26})$$

Thus, we obtained a relation for calculating the optimal kalman gains over the course of a movement. But, we still need to find out a way to calculate the covariance matrices  $\Sigma^{\hat{\mathbf{x}}}$ ,  $\Sigma^{\mathbf{e}}$  and  $\Sigma^{\hat{\mathbf{x}}\mathbf{e}}$ . We calculate these matrices in a forward pass through time. Since the variables  $\mathbf{x}$ ,  $\hat{\mathbf{x}}$  and  $\mathbf{e}$  are deterministically related, we can calculate the covariance for the third variable given we know the covariance for two variables. We choose to calculate the covariance matrices for  $\hat{\mathbf{x}}$  and  $\mathbf{e}$  since the equations are most compact for these variables [20]. Given the covariances of  $\hat{\mathbf{x}}$  and  $\mathbf{e}$ , the covariance of variable  $\mathbf{x}$  is given by

$$\Sigma_t^x = E \left[ (\mathbf{e} + \hat{\mathbf{x}}) (\mathbf{e} + \hat{\mathbf{x}})^T \right] = \Sigma_t^e + \Sigma_t^{\hat{\mathbf{x}}} + \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} + \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}^T} \quad (\text{A.27})$$

Now, from Eqn A.7

$$\begin{aligned} \hat{\mathbf{x}}_{t+1} = & A \hat{\mathbf{x}}_t + B \mathbf{u}_t + K_t \left( H \mathbf{x}_t + \omega_t + \sum_{i=1}^d \epsilon_t^i D_i \mathbf{x}_t - H \hat{\mathbf{x}}_t \right) + \eta_t \\ \hat{\mathbf{x}}_{t+1} = & A \hat{\mathbf{x}}_t + B \mathbf{u}_t + K_t \left( \left( H + \sum_{i=1}^d \epsilon_t^i D_i \right) (\mathbf{e}_t + \hat{\mathbf{x}}_t) + \omega_t - H \hat{\mathbf{x}}_t \right) + \eta_t \\ \hat{\mathbf{x}}_{t+1} = & \left( A - B L_t + K_t \sum_{i=1}^d \epsilon_t^i D_i \right) \hat{\mathbf{x}}_t + K_t \left( H + \sum_{i=1}^d \epsilon_t^i D_i \right) \mathbf{e}_t + \\ & K_t \omega_t + \eta_t \end{aligned} \quad (\text{A.28})$$

Similarly, from Eqn A.8

$$\begin{aligned}
\mathbf{e}_{t+1} &= (A - K_t H) \mathbf{e}_t + \xi_t - K_t \omega_t - \eta_t + \\
&\quad \sum_{i=1}^d (\varepsilon_t^i C_i \mathbf{x}_t + \gamma_t^i \bar{C}_i \pi(\hat{\mathbf{x}}_t) - \epsilon_t^i K_t D_i \mathbf{x}_t) \\
\mathbf{e}_{t+1} &= (A - K_t H) \mathbf{e}_t + \xi_t - K_t \omega_t - \eta_t + \\
&\quad \sum_{i=1}^d ((\varepsilon_t^i C_i - \epsilon_t^i K_t D_i) (\mathbf{e}_t + \hat{\mathbf{x}}_t) - \gamma_t^i \bar{C}_i L_t \hat{\mathbf{x}}_t) \\
\mathbf{e}_{t+1} &= \left( A - K_t H + \sum_{i=1}^d (\varepsilon_t^i C_i - \epsilon_t^i K_t D_i) \right) \mathbf{e}_t + \xi_t - K_t \omega_t - \eta_t + \\
&\quad \left( \sum_{i=1}^d (\varepsilon_t^i C_i - \epsilon_t^i K_t D_i - \gamma_t^i \bar{C}_i L_t) \right) \hat{\mathbf{x}}_t \tag{A.29}
\end{aligned}$$

Now, we know that

$$\Sigma_t^{\hat{\mathbf{x}}} = E [\hat{\mathbf{x}}_t \hat{\mathbf{x}}_t^T]; \Sigma_t^{\mathbf{e}} = E [\mathbf{e}_t \mathbf{e}_t^T]; \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} = E [\hat{\mathbf{x}}_t \mathbf{e}_t^T] \tag{A.30}$$

Using these definitions,

$$\begin{aligned}
\Sigma_{t+1}^{\hat{\mathbf{x}}} &= (A - BL_t) \Sigma_t^{\hat{\mathbf{x}}} (A - BL_t)^T + K_t \left( \sum_{i=1}^d D_i \Sigma_t^{\hat{\mathbf{x}}} D_i^T \right) K_t^T + K_t H \Sigma_t^{\mathbf{e}} H^T K_t^T + \\
&\quad K_t \left( \sum_{i=1}^d D_i \Sigma_t^{\mathbf{e}} D_i^T \right) K_t^T + K_t \Omega^\omega K_t^T + \Omega^\eta + (A - BL_t) \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} H^T K_t^T + \\
&\quad 2K_t \left( \sum_{i=1}^d D_i \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} D_i^T \right) K_t^T + K_t H \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} (A - BL_t)^T
\end{aligned}$$

Simplifying

$$\begin{aligned}
\Sigma_{t+1}^{\hat{\mathbf{x}}} &= (A - BL_t) \Sigma_t^{\hat{\mathbf{x}}} (A - BL_t)^T + K_t \left( \sum_{i=1}^d D_i (\Sigma_t^{\hat{\mathbf{x}}} + \Sigma_t^{\mathbf{e}} + \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}}) D_i^T \right) K_t^T + \Omega^\eta + \\
&\quad + K_t (H \Sigma_t^{\mathbf{e}} H^T + \Omega^\omega) K_t^T + (A - BL_t) \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} H^T K_t^T + K_t H \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} (A - BL_t)^T
\end{aligned} \tag{A.31}$$

Now calculating the covariance of variable  $\mathbf{e}$ ,

$$\begin{aligned}
\Sigma_{t+1}^{\mathbf{e}} &= \left( A - K_t H + \sum_{i=1}^d (\varepsilon_t^i C_i - \epsilon_t^i K_t D_i) \right) \Sigma_t^{\mathbf{e}} \left( A - K_t H + \sum_{i=1}^d (\varepsilon_t^i C_i - \epsilon_t^i K_t D_i) \right)^T + \\
&\quad \left( \sum_{i=1}^d (\varepsilon_t^i C_i - \epsilon_t^i K_t D_i - \gamma_t^i \bar{C}_i L_t) \right) \Sigma_t^{\hat{\mathbf{x}}} \left( \sum_{i=1}^d (\varepsilon_t^i C_i - \epsilon_t^i K_t D_i - \gamma_t^i \bar{C}_i L_t) \right)^T + \\
&\quad \left( A - K_t H + \sum_{i=1}^d (\varepsilon_t^i C_i - \epsilon_t^i K_t D_i) \right) \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}^T} \left( \sum_{i=1}^d (\varepsilon_t^i C_i - \epsilon_t^i K_t D_i - \gamma_t^i \bar{C}_i L_t) \right)^T + \\
&\quad \left( \sum_{i=1}^d (\varepsilon_t^i C_i - \epsilon_t^i K_t D_i - \gamma_t^i \bar{C}_i L_t) \right) \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} \left( A - K_t H + \sum_{i=1}^d (\varepsilon_t^i C_i - \epsilon_t^i K_t D_i) \right)^T + \\
&\quad \Omega^\xi + K_t \Omega^\omega K_t^T + \Omega^\eta
\end{aligned}$$

$$\begin{aligned}
\Sigma_{t+1}^{\mathbf{e}} &= (A - K_t H) \Sigma_t^{\mathbf{e}} (A - K_t H)^T + \sum_{i=1}^d (C_i \Sigma_t^{\mathbf{e}} C_i^T + K_t D_i \Sigma_t^{\mathbf{e}} D_i^T K_t^T) + \\
&\quad \sum_{i=1}^d (C_i \Sigma_t^{\hat{\mathbf{x}}} C_i^T + K_t D_i \Sigma_t^{\hat{\mathbf{x}}} D_i^T K_t^T + \bar{C}_i L_t \Sigma_t^{\hat{\mathbf{x}}} L_t^T \bar{C}_i^T) + \Omega^\xi + K_t \Omega^\omega K_t^T + \Omega^\eta + \\
&\quad \sum_{i=1}^d \left( C_i \left( \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} + \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}^T} \right) C_i^T + K_t D_i \left( \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}} + \Sigma_t^{\hat{\mathbf{x}}\mathbf{e}^T} \right) D_i^T K_t^T \right)
\end{aligned}$$

Simplifying further

$$\begin{aligned}
\Sigma_{t+1}^e &= (A - K_t H) \Sigma_t^e (A - K_t H)^T + \sum_{i=1}^d \left( C_i \left( \Sigma_t^{\hat{\mathbf{x}}} + \Sigma_t^e + \Sigma_t^{\hat{\mathbf{x}}e} + \Sigma_t^{\hat{\mathbf{x}}e^T} \right) C_i^T \right) + \\
&K_t \left( \sum_{i=1}^d \left( D_i \left( \Sigma_t^{\hat{\mathbf{x}}} + \Sigma_t^e + \Sigma_t^{\hat{\mathbf{x}}e} + \Sigma_t^{\hat{\mathbf{x}}e^T} \right) D_i^T \right) + \Omega^\omega \right) + \Omega^\xi + \Omega^\eta + \\
&\sum_{i=1}^d \left( \bar{C}_i L_t \Sigma_t^{\hat{\mathbf{x}}} L_t^T \bar{C}_i^T \right)
\end{aligned} \tag{A.32}$$

Now calculating the covariance between variables  $\hat{\mathbf{x}}$  and  $\mathbf{e}$

$$\begin{aligned}
\Sigma_{t+1}^{\hat{\mathbf{x}}e} &= -K_t \sum_{i=1}^d D_i \Sigma_t^{\hat{\mathbf{x}}} D_i^T K_t^T + K_t H \Sigma_t^e (A - K_t H)^T - K_t \sum_{i=1}^d D_i \Sigma_t^{\hat{\mathbf{x}}} D_i^T K_t^T - \\
&K_t \Omega^\omega K_t^T - \Omega^\eta + (A - B L_t) \Sigma_t^{\hat{\mathbf{x}}e} (A - K_t H)^T - 2K_t \sum_{i=1}^d D_i \Sigma_t^{\hat{\mathbf{x}}e} D_i^T K_t^T
\end{aligned}$$

Simplifying

$$\begin{aligned}
\Sigma_{t+1}^{\hat{\mathbf{x}}e} &= K_t H \Sigma_t^e (A - K_t H)^T - K_t \left( \sum_{i=1}^d D_i \left( \Sigma_t^{\hat{\mathbf{x}}} + \Sigma_t^e + 2\Sigma_t^{\hat{\mathbf{x}}e} \right) D_i^T + \Omega^\omega \right) K_t^T - \Omega^\eta + \\
&(A - B L_t) \Sigma_t^{\hat{\mathbf{x}}e} (A - K_t H)^T
\end{aligned} \tag{A.33}$$

# Appendix B

## Simulation parameters

This section lists the values of simulation parameters used for the simulations shown in the 'Model Predictions' section.

### B.1 Common Parameters for all simulations

Mass property of the point mass being controlled by the controller in the model was set to

$$M = \begin{bmatrix} 4.0 & 0 \\ 0 & 1.5 \end{bmatrix}$$

The time constant for conversion of motor commands  $u$  to forces  $f$  (Eqn 3.3) was set to  $\tau = 0.12$  sec.

The observation matrix  $H$  (Eqn 2.7) was set so that the feedback vector  $y$  was

$$y = [P_x, P_y, V_x, V_y, T_x, T_y]$$

$$y = [P_x, P_y, V_x, V_y, via_x, via_y, T_x, T_y]$$

for the reach and via-point simulations respectively. All the symbols have the same meaning as described in the 'Model Predictions' section.

### B.1.1 Noises

The variance of gaussian noise ( $\Omega^\xi$ ) in system dynamics equation (Eqn 2.7) was set to a diagonal matrix of the size of the state  $x$  with value of all diagonal elements being 1.

The variance of the gaussian noise ( $\Omega^\omega$ ) in feedback equation (Eqn 2.8) was set to a diagonal vector with the elements along the diagonal being

$$[0.08, 0.08, 0.8, 0.8, 0.08, 0.08]$$

$$[0.08, 0.08, 0.8, 0.8, 0.08, 0.08, 0.08, 0.08]$$

for the reach and via-point experiment respectively. Please note that the noise variances are varied in accordance with elements of the feedback vector  $y$  and have units as the units for the feedback vector elements squared.

The variance of the gaussian noise ( $\Omega^\eta$ ) in the state update equation (Eqn 2.4) was set to a diagonal matrix of the size of the state estimate vector  $\hat{\mathbf{x}}$  with all the

diagonal elements being 0.3

Please note that although the values for the variances of these noises were chosen randomly, the effects of uncertainty on the control policy seen in the simulations were fairly robust with respect to change in the variances of these noises.

The matrix  $\bar{C}$  defining the control dependent noise in the system dynamics equation (Eqn 2.7) was set to a zero matrix of the size of the matrix  $B$  in the system dynamics equation with the only non-zero elements being  $\bar{C}_{51} = \bar{C}_{62} = 0.01$

### **B.1.2 Cost Function**

The matrix  $Q$  defining the 'state cost'(Eqn 2.9) was set so that the controller was penalized for not being at target or the via-point at the required movement time or via-point time. The penalty was the squared distance of the current position in the state vector from the target or the via-point scaled by a factor of 20.

The matrix  $R$  defining the control cost was set to a diagonal matrix of size 2X2 with the diagonal elements being 0.000000001.

Please note that the size of the effect of uncertainty observed in simulations is dependent on the ratio of the state cost to the control cost but the effect even if small can be observed for a wide range of ratios of the cost.

## B.2 Simulation specific parameters

### B.2.1 Biased viscous curl force field with off-diagonal and diagonal noise

For clockwise force field,  $Z = \begin{bmatrix} 0 & 13 \\ -13 & 0 \end{bmatrix} Ns/m.$

For counterclockwise force field,  $Z = \begin{bmatrix} 0 & -13 \\ 13 & 0 \end{bmatrix} Ns/m.$

Movement time was set to 0.45 sec. The target distance from the starting position was 9 cm.

#### Off-diagonal noise

The matrix  $C$  defining the state dependent noise in the system dynamics equation (Eqn 2.7) was set to a zero matrix with the only nonzero elements being  $C_{24} = 0.0975$  and  $C_{42} = -0.26$  for the high variance ('HV') case. The same elements were halved for the medium variance ('MV') case.

#### Diagonal noise

The matrix  $C$  defining the state dependent noise in the system dynamics equation (Eqn 2.7) was set to a zero matrix with the only nonzero elements being  $C_{22} = 0.0975$  and  $C_{44} = -0.26$  for the high variance ('HV') case. The same elements were halved for the medium variance ('MV') case.

## B.2.2 Via-point reaching task

The movement time was set to 1.0 sec. The via-point time was 0.4 sec. The distance of the via-point and target from the starting position was 9cm and 18cm respectively.

The matrix  $C$  defining the state dependent noise in the system dynamics equation (Eqn 2.7) was set to a zero matrix with the only nonzero elements being  $C_{24} = 0.0617$  and  $C_{42} = -0.1644$  for the high variance ('HV') case. The same elements were halved for the medium variance ('MV') case.

All the values of elements of the matrix  $C$  correspond to the actual noise levels tested in the 'Experiments' section.

# Bibliography

- [1] Sarah J. Blakemore, Susan J. Goodbody, and Daniel M. Wolpert. Predicting the Consequences of Our Own Actions: The Role of Sensorimotor Context Estimation. *J. Neurosci.*, 18(18):7511–7518, 1998.
- [2] Chistina A. Burbeck and Yen Lee Yap. Two mechanisms for localization? evidence for separation-dependent and separation-independent processing of position information. *Vision Research*, 30(5):739–750, 1990.
- [3] H. P. Clamann. Statistical analysis of motor unit firing patterns in a human skeletal muscle. *Biophysical journal*, 9(10):1233–1251, Oct 1969. LR: 20041117; PUBM: Print; JID: 0370626; ppublish.
- [4] Jrn Diedrichsen, Timothy Verstynen, Andrew Hon, Steven L. Lehman, and Richard B. Ivry. Anticipatory adjustments in the unloading task: Is an efference copy necessary for learning? *Experimental Brain Research*, 148(2):272–276, 2003.
- [5] O. Donchin and R. Shadmehr. Change of desired trajectory caused by training in

- a novel motor task. *Conference proceedings : ...Annual International Conference of the IEEE Engineering in Medicine and Biology Society.IEEE Engineering in Medicine and Biology Society.Conference*, 6:4495–4498, 2004. PUBM: Print; JID: 101243413; ppublish.
- [6] Opher Donchin, Joseph T. Francis, and Reza Shadmehr. Quantifying Generalization from Trial-by-Trial Behavior of Adaptive Systems that Learn with Basis Functions: Theory and Experiments in Human Motor Control. *J. Neurosci.*, 23(27):9032–9045, 2003.
- [7] Todorov E. *Studies of goal-directed movements*. PhD thesis, Massachusetts Institute of Technology, 1998.
- [8] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. 1954. *Journal of experimental psychology.General*, 121(3):262–269, Sep 1992. LR: 20061115; PUBM: Print; JID: 7502587; ppublish.
- [9] T. Flash and N. Hogan. The coordination of arm movements: an experimentally confirmed mathematical model. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 5(7):1688–1703, Jul 1985. LR: 20061115; PUBM: Print; JID: 8102140; ppublish.
- [10] C. M. Harris and D. M. Wolpert. Signal-dependent noise determines motor planning. *Nature*, 394(6695):780–784, Aug 20 1998. LR: 20061115; PUBM:

- Print; JID: 0410462; CIN: Nature. 1998 Aug 20;394(6695):725-6. PMID: 9723613; ppublish.
- [11] N Hogan. An organizing principle for a class of voluntary movements. *J. Neurosci.*, 4(11):2745–2754, 1984.
- [12] R. S. Johansson and G. Westling. Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects. *Experimental Brain Research*, 56(3):550–564, 1984.
- [13] M. I. Jordan and D. M. Wolpert. *The Cognitive Neurosciences (M. Gazzaniga, ed.)*, chapter Computational motor control. 1999.
- [14] Davis M. and Vinter R. *Stochastic modelling and control*. London: Chapman and Hall, 1985.
- [15] P. B. Matthews. Relationship of firing intervals of human motor units to the trajectory of post-spike after-hyperpolarization and synaptic noise. *The Journal of physiology*, 492 ( Pt 2)(Pt 2):597–628, Apr 15 1996. LR: 20061115; PUBM: Print; JID: 0266262; ppublish.
- [16] Schmidt R., Zelaznik H., Hawkins B., Frank J., and Quinn J. Motor-output variability: A theory for the accuracy of rapid motor acts. *Psychol Rev.*, 86(5):415–451, 1979.

- [17] Miall R.C. Forward models for physiological motor control. *Neural Networks*, 9:1265–1279(15), November 1996.
- [18] R. Shadmehr and F. A. Mussa-Ivaldi. Adaptive representation of dynamics during learning of a motor task. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 14(5 Pt 2):3208–3224, May 1994. LR: 20061115; PUBM: Print; JID: 8102140; ppublish.
- [19] K. A. Thoroughman and R. Shadmehr. Learning of action through adaptive combination of motor primitives. *Nature*, 407:742–747, October 2000.
- [20] E. Todorov. Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. *Neural computation*, 17(5):1084–1108, May 2005. PUBM: Print; JID: 9426182; ppublish.
- [21] E. Todorov and M. I. Jordan. Optimal feedback control as a theory of motor coordination. *Nature neuroscience*, 5(11):1226–1235, Nov 2002. LR: 20041117; PUBM: Print; JID: 9809671; CIN: Nat Neurosci. 2002 Nov;5(11):1110-1. PMID: 12404002; ppublish.
- [22] E. Todorov and W. Li. A generalized iterative lqg method for locally-optimal feedback control of constrained nonlinear stochastic system. *Conference proceedings : ...Proc. of American Control Conference.Conference*, pages 300–306, 2005.

- [23] Y. Uno, M. Kawato, and R. Suzuki. Formation and control of optimal trajectory in human multijoint arm movement. *Biological Cybernetics*, 61:89–101, 1989/06/01.
- [24] DAVID WHITAKER and KEZIAH LATHAM. Disentangling the role of spatial scale, separation and eccentricity in weber’s law for position. *Vision Research*, 37(5):515–524, 1997/3.
- [25] Steven P. Wise and R. Shadmehr. *Encyclopedia of the Human Brain*, chapter Motor Control. 2002.
- [26] D. M. Wolpert and M. Kawato. Multiple paired forward and inverse models for motor control. *Neural Networks*, 11(7-8):1317–1329, 1998/10/11.