UCLouvain

École polytechnique de Louvain

Adaptive human reaching movements

Modelling of sub-movement motor learning : sensory delays and curl field perturbations

Author: **Sarah GLUME**
Supervisor: **Frédéric CREVECOEUR**
Readers: **Philippe LEFÈVRE, Renaud RONSSE**
Academic year 2018–2019
Master [120] in Biomedical Engineering

Abstract

When we need to perform a goal-directed hand movement, our brain must compute the adequate command signals to send to the muscles of the arm in order to accurately reach the target. For this purpose, our brain uses an internal representation of the biomechanics of the upper limb, as well as of the external environment. However, in the event of an early exposure to an unexpected perturbation during the execution of the movement, the brain must quickly react by determining the nature of the perturbation and adapting the control signals accordingly. Motor learning is the ability of the human brain to learn new representations of the external environment when performing a motor task. Computationally, motor learning has been studied over movement repetitions, using the framework of optimal feedback control theory and in particular the Linear-Quadratic-Gaussian controller. In this work, we perform numerical simulations of such perturbed reaching movements with the hypothesis of "sub-movement" motor learning, namely that the central nervous system is able to compensate for the unexpected disturbance during the course of the movement. We start by reproducing the results of the recent work of Crevecoeur et al. [1] that uses a Least Squares identification technique in order to estimate the intensity of the disturbance during the execution of the movement. Afterwards, motivated by the biological likelihood, we introduce a temporal delay in the sensory feedback loop of the model. Our results show that the numerical implementation of the model used in this work can successfully reproduce results in agreement with the hypothesis of online motor learning, even when sensory motor noise and temporal feedback delays are taken into account. Furthermore, the model can also be flexibly adapted to more general force fields, such as a curl force field. Altogether, these advances will help to develop more realistic mathematical models of the motor system, allowing at the same time a better understanding of the motor function.

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Professor Frédéric Crevecoeur, who drove and motivated me through my learning path in the fascinating field of motor control. Indeed, he was always there to help, advise and motivate me throughout the realization of my master thesis.

Many thanks also to Professor Lefèvre and Professor Ronsse for agreeing to be part of the thesis jury as readers.

I am very grateful for the unfailing support and understanding of my parents and my brother Geoffroy throughout my years of engineering studies and particularly these few final week. They taught that willpower and hard work are the most important things whatever the project that I undertake and even if some people try to discourage me.

Moreover, I sincerely thank my other siblings for their continuous encouragements, especially my engineer godfather. A special thought goes to my guardian angels up there, and I would also like to thank my second family for their support.

Thanks a lot to my friends for supporting me in any occasion, for their invaluable help, and for all the funny moments that we spent together.
A special thank to Céline and Antoine for their suggestions during this final year, and to Andine for the proofreading of this manuscript.

Finally, words are certainly not sufficient to express my thankfulness to the most patient and supportive one, who is always there when I need help and self-confidence. Thank you so much Benjamin, my love.

Contents

List of Figures

Chapter 1

Introduction

Monday morning, 7 am, your alarm clock rings and a new day begins.

One of the first movements of the day consists in reaching the alarm clock in order to turn it off. This everyday action, seemingly effortless, is actually the result of complex and swift interactions between the sensory inputs and the hand, all monitored through the central nervous system.

In fact, this so-called "reaching movement" is the focus of active research in neuroscience since it allows us to get remarkable insights into the functioning of the human brain. In particular, how the brain is able to quickly integrate multiple information in order to compute the appropriate control command to send to the muscles is the central question addressed by the field of motor control. Interestingly, it turns out that mathematical methods from the fields of dynamical systems and control theory are well-suited for the study of reaching movements. This enables the use of computer simulations to analyze different control paradigms in order to compare them to empirical data and to determine which model most accurately describes the control laws employed by the brain.

Now, imagine that your dog (which is of course looking forward to greeting you) hears the alarm clock, runs to your room and bumps into your arms before you reach the off button. If this happens regularly, your brain will likely adapt its representation of the environment to take this perturbation into account when it occurs, in order to counteract it and to reach the off button anyway. Motor learning is this aptitude of the brain to learn new representations of the external environment in order to improve smoothness and accuracy when performing a motor task.

The present work focuses on the sub-movement motor adaptation (or online motor learning), i.e. the ability of the central nervous system to update its

internal representation of the environment during the movement, as opposed to over repetitions of the movement. This highlights a very fast time scale for the motor adaptation and opens challenges for the computational approaches required for the numerical simulation of reaching movements. We address some of these challenges throughout this document.

1.1 Motivations

Computational approaches to motor control and learning paved the way for a better understanding of the motor function. Indeed, the use of numerical simulations allows us to investigate many facets of the motor abilities in humans through the use of powerful algorithms. However, the mathematical formulation of such problems always implies several assumptions and simplifications that have consequences on the accuracy of the resulting simulations. This is not a problem as long as the researcher is well aware of those assumptions and simplifications. Still, investigating the relaxation of some assumptions by building more complex models can also provide important insights.

In this work, we improve an existing model of online motor learning by incorporating a temporal delay in the computations. This is motivated by the biological fact that information cannot travel instantaneously across the central nervous system. Moreover, we consider more general types of disturbances occuring during the course of the movement, and we briefly introduce key elements for the extension of the existing model to a more realistic geometry. Altogether, these improvements aim at producing more biologically grounded simulations of the motor function.

Overall, the long-term goals of the fields of motor control and learning are, among others, to improve the condition of motor-impaired patients, as well as to help developments in the field of robotics. In this context, we hope that our contributions will help to produce more precise models, thereby partially filling the gap between theory and reality.

1.2 State of the Art

The early research regarding motor behaviour can be traced back prior to the 20th century [3]. In parallel, the fields of neurophysiology and psychology studied the fundamental basis of voluntary movements. Whereas neurophysiological approaches provided insights into the neural processes that are involved in the production of a movement, without taking into account the characteristics of the movement itself,

psychological studies were more interested in the description of the high-level skills (e.g. motor "performance") required in a stimulus-response setup [4]. At that time, little attention was paid to the actual neural mechanisms that are responsible for the generation of an adequate motor response to a goal-directed visual stimulus. It is only later, in the late 1970's, that both fields merged together to offer a more exhaustive view of the human motor behaviour.

More recently, a focus has emerged on the mathematical description of the motor function, thereby opening opportunities for the use of powerful methods from the field of control theory [5, 6, 7]. This exciting approach offers several advantages such as the possibility of performing numerical simulations, and constitutes a critical step towards a deeper insight into the neural basis of motor control.

1.2.1 Motor control

Everyday life movements look effortless and become almost unconscious, even though they reflect complex interactions between limbs and the central neural system (CNS) [8]. The field of motor control investigates the ability of the human CNS to generate commands that are sent to muscles in order to achieve this kind of goal-directed movements accurately [9], while dealing with a mechanically uncertain environment [2]. In this context, we can distinguish three kinds of responses to external disturbances. First, long-latency reflexes play an important role in motor control since they consist in fast ($\sim 50 - 100$ms) corrective responses to counter mechanical perturbations thanks to sensory information and ensure task success. They differ from the short-latency reflexes, also called spinal reflexes, that happen in a shorter time interval ($\sim 20 - 50$ms). Finally, voluntary reactions, whose latency is higher than 100ms, differ from short and long-latency reflexes since they involve making a decision consciously in order to correct the movement [1, 2, 10, 8].

In his review [8], Stephen Scott argues that the three fundamental levels of analysis of the motor function, namely the motor behaviour, the limb mechanics and the neural control, can be connected through the theory of optimal feedback control in order to study the strategy employed by the brain to plan and control movement during long-latency reflexes [5, 8, 2, 11, 12, 13]. In particular, the Linear-Quadratic-Gaussian (LQG) control framework is promoted in several independent studies [14, 2, 10, 6]. This comes mostly from the fact that a great amount of results and algorithms have been derived in the context of LQG control.

Arm reaching movements are the most studied human movements and studies have demonstrated that the motor commands are based on an internal model of the involved dynamics [15]. If the environmental forces do not correspond to the

expected dynamics, the predicted model is inaccurate and the trajectory of the hand is deviated. The internal representation is then learned and adapted with practice across repetitions [16]. This was correlated with electromyography (EMG) measures by Thoroughman and Shadmehr in [17]. This is made possible by the sensory feedback we collect from what happens actually and the comparison to the belief the brain estimated with motor command and previous prediction about the arm state [9, 18].

1.2.2 Motor learning

The oldest report on visuomotor adaptation experiment was written in 1867 by Hermann von Helmholtz [9]. The study of visuomotor adaptation consists in experiments in which visual feedback of the motor commands is distorted, by means of lenses for example, during arm movements. This constitutes one of the well studied types of the motor adaptation framework [19, 20, 9, 21].

The other well studied version of motor adaptation is the adaptation to a force-field applied to the subject's hand during the arm motion [9]. Typically, this perturbation varies during the hand movement. Shadmehr and colleagues led one of the first studies of the modelling of motor adaptation to force field perturbations in [22]. They studied the appearance of aftereffects after the unexpected removal of a force field applied to the arm during a reaching movement. These are characterized by deviations in unperturbed movements consecutive to the exposition to the perturbation. With the same idea, other types of experiments can be performed with a similar force field effect, such as adding weight onto the subject's hand in [19], execution of arm reaching movements in presence of Coriolis forces (in a rotating room) in [23] and [24] or in microgravity conditions in [25] and [26].

In the past 20 years, a lot of studies were led in the field of motor learning [27], which is the short-term motor adaptation of the perturbed movement through practice. The field can also involve motor learning with long-term practice, i.e. learning of new high-level motor skills such as playing tennis, playing guitar or dancing [27, 21].

Various modelling approaches can explain different tested aspects of the short-term motor learning, on a trial-by-trial basis. Some authors, among which Burdet in [28], suggested a stiffening mechanism of the muscle that is engaged after the maximum deviation in order to smoothly correct the trajectory. Others researchers, such as Izawa in [29], modelled internal representation errors as stochastic state-dependent Gaussian errors, which can be accurately regulated by a stochastic optimal control technique, such as the LQG controller. Finally, the motor adapta-

tion was also addressed by switching between different controllers, depending on the exerted forces [1]. These are supposed to be learned with previously experienced situations.
Authors also highlighted "savings" phenomenons, which correspond to a faster relearning of a task previously performed by the subject. The opposite effect is the anterograde interference, meaning slower learning when a perturbation is applied to the subject, after previous adaptation to the opposite load during a reaching task [9, 21].

The above mentioned approaches have allowed the modelling of most of the features observed during reaching laboratory experiments. However, Crevecoeur and colleagues in [1] were the first ones to investigate online motor learning. This consists in the adaptation of the internal model during the course of the reaching movement, as opposed to across repetitions. This approach therefore involves a very fast time scale for motor adaptation. Online motor learning opens perspectives for the characterization of previously unmodelled phenomena, such as aftereffetcs during the movement when there is removal of the force field applied perpendicularly to the arm.

1.3 Organization of the thesis and contributions

In Chapter 2, we present the mathematical methods used throughout this work. The Linear-Quadratic-Gaussian (LQG) control framework is introduced, as well as the Recursive Least Squares identification technique used for the online estimation of the perturbation parameter.

In Chapter 3, we follow the recent work of Crevecoeur and coworkers [1] and reproduce the numerical experiments related to sub-movement motor adaptation. The results show that the least squares identification method allows us to successfully estimate the force field parameter, thereby enabling the simulation of online learning. This chapter forms the core of the present work, as following chapters discuss extensions of these simulations.

Motivated by neurophysiological faithfulness, our contribution in Chapter 4 addresses the addition of a sensory feedback delay in the feedback loop of the model introduced in Chapter 3.

Chapter 5 further extends the model of Chapter 4 by considering a more general kind of external disturbance, namely a curl force field. As opposed to previous chapters, this force field results from the combination of two orthogonal force

field, in the x- and y-direction respectively. Our results show that the model, as well as the implementation, are able to flexibly take such perturbations into account.

Finally, Chapter 6 introduces some key elements for future investigations that will take into account a more complex geometric representation of the arm.

Chapter 2

Linear modelling of motor control and motor learning

As explained in the introduction, the optimal feedback control theory offers great opportunities for the study of motor control using a linear dynamical model. Indeed, the motion of a human arm can be modelled as a sensorimotor system (Figure 2.1). This model is constructed by means of a linear control system, in which the plant S represents the arm and the controller C illustrates neural control [2]. The arm is characterized by a state vector, containing several state variables such as its position and velocity. Meanwhile, the controller is an entity that sends motor commands, u, in order to control the evolution of these state variables over time, based on the collected sensory feedback, y. This evolution is expressed by the arm trajectory [2].

In this framework, the Linear-Quadratic-Gaussian controller is a linear system control approach well known in the motor control field due to its consistent behaviour with experimental results [14]. Furthermore, an identification method for parameters of the system controlled by a Linear-Quadratic-Gaussian controller, called the Recursive Least Square method, was proposed in [30]. Therefore, both methods can be combined in order to model the online learning of an unknown disturbance, in our case an external force field. In this chapter, the theoretical foundations of models studied in the following chapters are laid.

2.1 Linear-Quadratic-Gaussian control

In this section, the Linear-Quadratic-Gaussian control is explained. This piece of this thesis is largely inspired by lecture notes from the course "LGBIO2060 - Modelling of Biological Systems" [31], written by Professor Crevecoeur, and by lecture

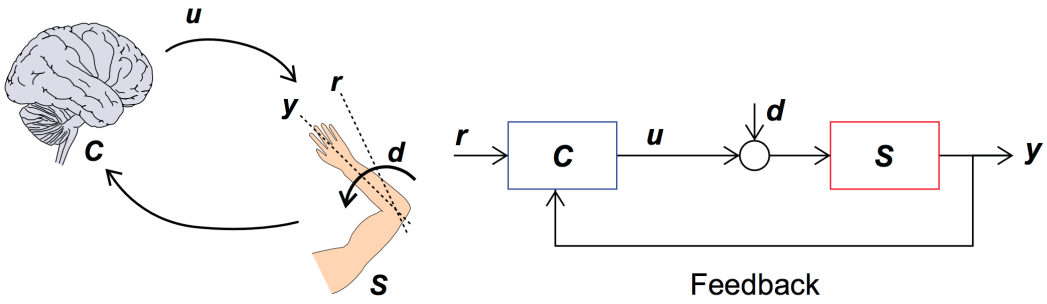

Figure 2.1: **Schematic illustration of the interaction between the central nervous system and the human arm during a movement (sensorimotor system).** **C** is the controller and **S** is the plant. **r** is the goal-target of the movement, **y** is the feedback/output state vector describing the trajectory of the arm, **u** is the motor command signals sent to the arm and **d** is a potential disturbance applied on the arm during the movement. The figures are adapted from the work of Crevecoeur et al. [2].

notes of P.M. Mälikä in [32]. Extensive mathematical details and explanations are also available in the work of Emanuel Todorov [14] and Joseph Nashed and colleagues [10].

The Linear-Quadratic-Gaussian control framework is defined for a system with **L**inear dynamics, whose control law minimizes a **Q**uadratic cost function, and whose additive noisy disturbance is modeled as a random variable following a **G**aussian distribution.

Let us consider a linear system described by the following discrete-time state space representation:

$$\begin{aligned} \mathbf{x}_{t+1} &= A\mathbf{x}_t + B\mathbf{u}_t + \varepsilon_t \\ \mathbf{y}_t &= H\mathbf{x}_t + \omega_t \end{aligned}$$

where $\mathbf{x}_t \in \mathbb{R}^n$ is the state vector containing the state variables, $\mathbf{u}_t \in \mathbb{R}^m$ is the command vector composed of control variables and $\mathbf{y}_t \in \mathbb{R}^p$ is the feedback vector. The noise parameters $\varepsilon_t \sim \mathcal{N}(0, \Omega_\varepsilon)$ and $\omega_t \sim \mathcal{N}(0, \Omega_\omega)$ are centered Gaussian disturbances with variances equal to Ω_ε and Ω_ω respectively. We assume that the initial state vector $\mathbf{x}_1$ is provided.

The cost function, that is considered at each time $t = 1, 2, ..., N-1$ step, is equal to the sum of a quadratic cost on the current state $\mathbf{x}_t$ and a quadratic cost on the command $\mathbf{u}_t$. At $t = N$, the cost function is reduced to a quadratic cost related

to the final state $\mathbf{x}_N$ since the command is undefined. The following expressions depict the used cost function :

$$\begin{aligned} J_t(\mathbf{x}_t, \mathbf{u}_t) &= \mathbf{x}_t^T Q_t \mathbf{x}_t + \mathbf{u}_t^T R \mathbf{u}_t \\ J_N(\mathbf{x}_N) &= \mathbf{x}_N^T Q_N \mathbf{x}_N \end{aligned}$$

where Q_N, Q_t and R are symmetric and positive definite (or semidefinite) matrices of appropriate dimensions.

Optimal control problem

The problem encounters in this work, called the LQG problem, is an optimal control problem which consists in finding a sequence of commands $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{N-1}$ so that the scalar quadratic cost function is minimized :

$$J \equiv E\left[J_N + \sum_{t=1}^{N-1} J_t(\mathbf{x}_t, \mathbf{u}_t)\right]$$

Solution of the LQG problem : complete state information

For the fully observable case (i.e. when the state vector is known perfectly), the optimal solution of the optimal control problem is a linear state feedback controller and a cost-to-go at each time step t given by :

$$V_t(\mathbf{x}_t, \mathbf{u}_t) = \mathbf{x}_t^T S_t \mathbf{x}_t + s_t$$

where S_t is a symmetric, positive definite (or semidefinite) matrix and s_t are non-negative scalar quantities. Indeed, this is the form of solution obtained by solving the functional Bellman equation :

$$\begin{aligned} V_t &= \min_{\mathbf{u}_t}\left[J_t(\mathbf{x}_t, \mathbf{u}_t) + E(V_{t+1}|\mathbf{x}_t, \mathbf{u}_t)\right] \\ &= \min_{\mathbf{u}_t}\left[\mathbf{x}_t^T Q_t \mathbf{x}_t + \mathbf{u}_t^T R \mathbf{u}_t + E(V_{t+1}|\mathbf{x}_t, \mathbf{u}_t)\right] \end{aligned}$$

This statement can be proved by induction. First of all, the previous claim is assumed as true for $t = N$, with $S_N = Q_N$ ans $s_N = 0$. Then, if we manage to show that 2.1 is true for $t + 1$, then the claim will also be true for $1 \leq t < N - 1$. Expanding the conditional expected value of V_{t+1}, given $\mathbf{x}_t$ and $\mathbf{u}_t$ from the induction

assumption gives :

$$\begin{aligned} V_{t+1} &= \min_{\mathbf{u}_t}\Bigg[\mathbf{x}_t^T\Big(Q_t + A^T S_{t+1} A\Big)\mathbf{x}_t + \mathbf{u}_t^T\Big(R + B^T S_{t+1} B\Big)\mathbf{u}_t \\ &\quad + 2\mathbf{x}_t^T A^T S_{t+1} B\mathbf{u}_t + tr(S_{t+1}\Omega_\varepsilon) + s_{t+1}\Bigg] \end{aligned}$$

This equation is a quadratic form in $\mathbf{u}_t$, which is minimized when $\mathbf{u}_t$ satisfies :

$$\begin{aligned} \mathbf{u}_t &= -C_t\mathbf{x}_t & (2.1) \\ &= -\Big(R + B^T S_{t+1} B\Big)^{-1} B^T S_{t+1} A\mathbf{x}_t & (2.2) \end{aligned}$$

Therefore, it can be observed that the optimal control policy is a linear function of the state vector. By plugging the expression of the optimal control variable into V_t, we obtain :

$$\begin{aligned} V_t &= \mathbf{x}_t^T S_t \mathbf{x}_t + s_t \\ &= \mathbf{x}_t^T\Big(Q_t + A^T S_{t+1}(A - BC_t)\Big)\mathbf{x}_t + s_t \end{aligned}$$

where $s_t = s_{t+1} + tr(S_{t+1}\Omega_\varepsilon) > 0$. The required expression for V_t is found and we can see that $S_t \geq 0$ to complete the proof.

In practice, the optimal feedback gains are given by the following backward recursion :

$$\begin{aligned} S_N &= Q_N \\ s_N &= 0 \\ C_t &= \Big(R + B^T S_{t+1} B\Big)^{-1} B^T S_{t+1} A \\ S_t &= Q_t + A^T S_{t+1}\Big(A - BC_t\Big) \\ s_t &= s_{t+1} + tr(S_{t+1}\Omega_\varepsilon) \end{aligned}$$

Solution for the LQG problem : incomplete state information

The fully observable case assumes perfect knowledge of the state vector. In many cases in the real world as for the control of a sensorimotor system, it happens to

be an unrealistic assumption. The measurement of the state at a certain time step (sensory feedback) can be modelled by a noisy mixture of state variables.

The estimate of the state vector is assumed to be a convex combination of prior and feedback :

$$\begin{aligned}\hat{\mathbf{x}}_{t+1} &= (1 - K_t) \times \text{prior} + K_t \times \text{feedback} \\ \hat{\mathbf{x}}_{t+1} &= A\hat{\mathbf{x}}_t + B\mathbf{u}_t + K_t(\mathbf{y}_t - H\hat{\mathbf{x}}_t)\end{aligned}$$

In the previous expressions, K_t denotes the gain of the Kalman filter used to estimate the state [33], which weights the contribution of feedback in the update rule. The estimation error induced by the imperfect knowledge of the state variable is given by :

$$\mathbf{e}_{t+1} = (A - K_t H)\mathbf{e}_t + \varepsilon_t - K\omega_t$$

The optimal Kalman gain K_t is computed such that it minimizes the estimation error of the state variabble :

$$\begin{aligned}K_t &= \arg\min_K ||\mathbf{e}_t||^2 \\ &= \arg\min_K \left[\text{tr}\left(E(\mathbf{e}_{t+1}\mathbf{e}_{t+1}^T\right)\right]\end{aligned}$$

From the error expression above, the terms in the error covariance matrix that depend on K_t are given by :

$$a(K_t) = \text{tr}\left(- 2K_t H\Sigma_t + K_t(H\Sigma_t H^T + \Omega_\omega)K_t^T\right)$$

which by minimizing over K_t give:

$$\nabla a(K_t) = 0 \implies K_t = A\Sigma_t H^T(H\Sigma_t H^T + \Omega_\omega)^{-1}$$

However, in practice, optimal state estimates and Kalman filter gains are obtained using the following forward recursion (with Σ_1 known) :

$$\begin{aligned}\hat{\mathbf{x}}_{t+1} &= A\hat{\mathbf{x}}_t + B\mathbf{u}_t + K_t(\mathbf{y}_t - H\hat{\mathbf{x}}_t) \\ K_t &= A\Sigma_t H^T(H\Sigma_t H^T + \Omega_\omega)^{-1} \\ \Sigma_{t+1} &= \Omega_\varepsilon + (A - K_t H)\Sigma_t A^T\end{aligned}$$

Finally, the whole closed loop system can be written by means of the following equations :

$$\begin{bmatrix}\mathbf{x}_{t+1} \\ \mathbf{e}_{t+1}\end{bmatrix} = \begin{bmatrix}A - BC_t & BC_t \\ 0 & A - K_t H\end{bmatrix}\begin{bmatrix}\mathbf{x}_t \\ \mathbf{e}_t\end{bmatrix} + \begin{bmatrix}\varepsilon_t \\ \varepsilon_t - K_t\varepsilon\omega_t\end{bmatrix}$$

Its stability depends on the eigenvalues of $A - BC_t$ and $A - K_tH$ separately. This principle is known as the separation principle of LQG control. In fact, the estimation and control steps in the LQG control problem are separated as a result of the independence between the error covariance and the command inputs $\mathbf{u}_t$.

The solution of the previous Bellman equation is the same when the true state $\mathbf{x}_t$ is replaced by the estimated state $\hat{\mathbf{x}}_t$. Therefore, under partial state observation, the optimal control policy from Equation 2.1 can be applied to the estimated state. This is known as the certainty equivalence principle.

2.2 Recursive Least Squares parameter identifier

In their book [30], Robert R. Bitmead, Michel Gevers and Vincent Wertz present a candidate adaptive optimal controller, in which system parameters can be identified inside the LQG closed loop. It consists in providing to the controller design estimates of these parameters in order to compute the sequence of controller gains by using Equation 2.1. The Recursive Least Squares identification approach provides the estimates of the system parameters.

This method consists in the iterative update of a vector containing the parameters estimates, by means of the following adaptive rule :

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \gamma R_t^{-1} \psi_t(\hat{\theta}_{t-1}) \mathbf{e}_t(\hat{\theta}_{t-1}) \tag{2.3}$$

In Equation 2.3, the following elements can be noticed :

- $\hat{\theta}_t$ is the vector of the current parameters estimates.
- ψ_t is the regressor and is expressed as the partial derivative of the output state vector with respect to the parameters vector :

$$\psi_t(\hat{\theta}_{t-1}) \quad = \quad \frac{\partial \hat{\mathbf{y}}_{t|t-1}(\hat{\theta}_{t-1})}{\partial \theta}$$

- $\mathbf{e}_t(\hat{\theta}_t)$ is the prediction error, computed from the observed output $\mathbf{y}_t$ and the estimated output $\hat{\mathbf{y}}_{t|t-1}$ provided with the previous estimate :

$$\mathbf{e}_t(\hat{\theta}_{t-1}) \quad = \quad \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}$$

- γ is a positive scalar that determines the step size of the adaptation, chosen to be constant in this case, called the learning rate. Convergence is achieved depending on the value of γ. It has to be small but not infinitesimal.

- R_t is a matrix that can take different forms and which is updated at each time step t. In our modelling, we use a time constant matrix, such that $R_t = I$, the identity matrix (with appropriate dimensions). The updating algorithm is then called a gradient algorithm.

It can be demonstrated that, under certain assumptions described in [30], this iterative process helps to stabilize the closed loop of the controller optimal gains computation [1].

Finally, we noticed that this method is simply called iterative Least Squares identification when the vector θ is composed of only one element.

Chapter 3

Adaptive control : point mass model

The human arm can be modelled as a biomechanical linear control system. The model parameters and the state vector of this system are defined using on the one hand equations describing the dynamics of the system, and on the other hand a model of the muscle that quantifies its reaction to a neural command, in terms of forces.

There exist different planar representations of the arm. They mostly differ from each other by the number of degrees of freedom considered, as well as by the biomechanical likelihood with the actual human arm geometry. Furthermore, in order to take advantage of the power of linear control techniques such as the Linear-Quadratic-Gaussian controller (Section 2.1), we restrict ourselves in this work to models of the arm described by linear equations. Therefore, the point mass model is the first model we chose to employ as a widespread model in the motor control theory field.

Here, we would like to model the online adaptation of a hand movement to an external force field applied during the entire movement, when reaching a fixed target. When subjects are not aware of the perturbation, the trajectory of the hand is dramatically deviated. Indeed, the true dynamics of the system is not perfectly known by the system controller. It has been physiologically studied by Shadmehr and coauthors in [22] that, with practice, the trajectory tends to become as close as possible to the one without perturbation.

With the aim of modelling this motor learning, our hypothesis is that there is a very fast timescale of motor learning, namely a sub-movement adaptation to the perturbation, whose potential presence has been demonstrated by Crevecoeur and

colleagues in [1].

3.1 Models and methods

Our goal is to implement a method to model the sub-movement motor adaptation to an unexpected perturbation. Therefore, the model proposed in this chapter combines a Linear-Quadratic-Gaussian controller and an iterative Least Square algorithm (Chapter 2), in order to simulate the online learning of a force field scaling factor, also called the perturbation parameter.
The model used in this chapter is mostly inspired by the adaptive control model developed by Crevecoeur and colleagues in [1] and is based on the hypothesis that the system knows, exactly and at each moment, the position and the velocity of the point mass, as well as the forces involved.

Point mass model

The point mass model is based on the translation of a point mass in the horizontal plane, as depicted in Figure 3.1. This mass is set close to the one of a human arm (around 2.5 kg).

The motion equations are derived using Newton's second law (sum of forces exerted on a mass equals the mass multiplied by its acceleration). Three forces act on the mass m in this model. The first one is a viscous force, which is proportional to the velocity, with a dissipative proportionality constant G (0.1 Nsm^{-1}) and opposite sign. The other two forces involved are the controlled force, whose components are F_x and F_y, and the environmental force field perturbation F_{ENV}, which is equal to $\theta\dot{y}$. Only a purely lateral disturbance is considered here. Depending on the direction of this force field (clockwise or counterclockwise), the force field intensity θ is equal to $\pm L$ (Nsm^{-1}), or to zero if there is no perturbation.

$$m\ddot{x} = -G\dot{x} + F_x + F_{ENV} \tag{3.1}$$
$$m\ddot{y} = -G\dot{y} + F_y \tag{3.2}$$

In Equations 3.1 and 3.2, $\dot{x}$ and $\dot{y}$ are the velocity components in the two directions of the plane.
Moreover, in order to approximate the muscle dynamics, a first-order linear filter between command variables u_x and u_y and controlled forces is assumed, with a time constant τ fixed to 0.08s. This value has been chosen because it constitutes a trade-off between physiological values [34] and the value used in [1].

$$\tau\dot{F}_x = u_x - F_x \tag{3.3}$$
$$\tau\dot{F}_y = u_y - F_y \tag{3.4}$$

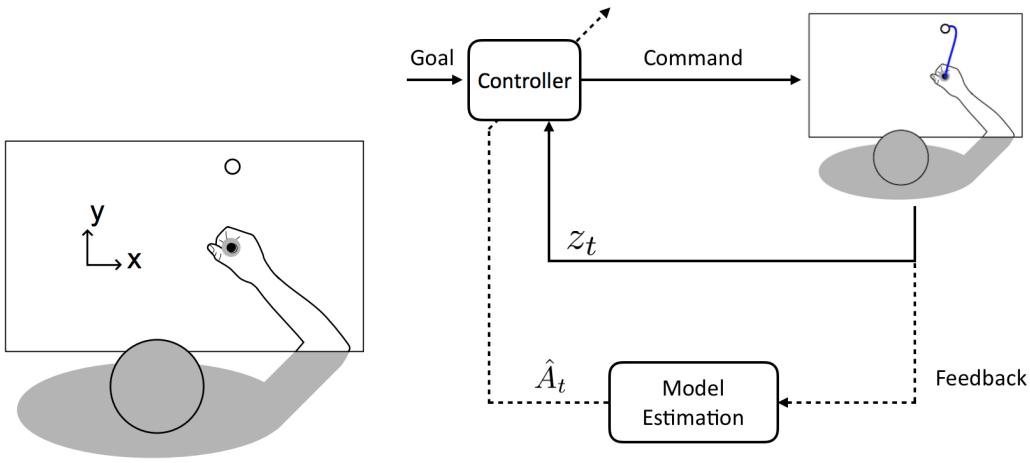

Figure 3.1: **Point mass model, with complete state information. Left :** Schematic illustration of the simulated experiment, for which the subject has to perform a reaching movement in the horizontal plane. The hand is modelled as a point mass. **Right :** Schematic representation of an adaptive controller. The inner loop (solid line) accounts for the correction of state deviations. The outer loop (dashed line) accounts for the model corrections. These figures are adapted from the work of Crevecoeur et al. [1].

Considering vectors $\mathbf{z} = [x,\ y,\ \dot{x},\ \dot{y},\ F_x,\ F_y,\ x^*,\ y^*]^T$ and $\mathbf{u} = [u_x\ u_y]^T$ and the constant goal target position $(x^*,\ y^*)$, the differential equations of the true dynamics can then be gathered into an linear algebraic form $\dot{\mathbf{z}} = A_c\mathbf{z} + B_c\mathbf{u}$ (c index stands for "continuous-time"). A_c and B_c matrices are defined as follows

$$A_c = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -G/m & \theta/m & 1/m & 0 & 0 & 0 \\ 0 & 0 & 0 & -G/m & 0 & 1/m & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1/\tau & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1/\tau \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad B_c = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1/\tau & 0 \\ 0 & 1/\tau \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Adaptive control system

Now that we have a continuous representation of the system, we would like to discretize the time axis with a chosen time step of $\delta t = 0.01$s, in order to add stochastic disturbances and to obtain a discrete-time LQG optimal control problem [6, 31]. Let $\mathbf{z}_t$ and $\mathbf{u}_t$ denote respectively the state vector and the control vector at

each time t, such that $\mathbf{z}_t = [x_t,\ y_t,\ \dot{x}_t,\ \dot{y}_t,\ F_{x,t},\ F_{y,t},\ x^*,\ y^*\ ]^T$ and $\mathbf{u}_t = [u_{x,t},\ u_{y,t}]^T$. The discrete-time state space representation of this system is

$$\mathbf{z}_{t+1} \quad = \quad A\mathbf{z}_t + B\mathbf{u}_t + \varepsilon_t \tag{3.5}$$

where A and B are the discretized matrices. They can be obtained by using explicit Euler integration technique, as done in [10] and [34]. This is based on a first order Taylor expansion over one time step δt [1, 31]. System matrices become $A := I + \delta t A_c$ and $B := \delta t B_c$.
In Equation 3.5, the motor noise ε_t is a zero-mean Gaussian disturbance with covariance matrix $\Omega_w = BB^T$, as explained in [1] and for which variance is close to 10^{-2}, such as in [10].

Nevertheless, in the case where the central nervous system is not aware of the external force field applied on the hand, the expected dynamics is described, at each time t, by a matrix $\hat{A}_t$. The latter contains a time varying estimate of the force field scaling factor $\hat{L}_t$ and is given by

$$\hat{A}_t = \begin{bmatrix} 1 & 0 & \delta t & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \delta t & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 - \delta t G/m & \boxed{\hat{L}_t/m} & \delta t/m & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 - \delta t G/m & 0 & \delta t/m & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -\delta t/\tau & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -\delta t/\tau \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

However, the Linear-Quadratic-Gaussian controller is not acquainted with the true dynamics, even if it is assumed that it knows which parameter is unknown. As a consequence, the time series of the feedback control gains $-C_t(\hat{A}_t)$ has to be computed with the dynamics estimate $\hat{A}_t$. As explained in Section 2.1, the optimal control policy is $\mathbf{u}_t = -C_t(\hat{A}_t)\mathbf{z}_t$, determined by the application of the controller gains to the current state vector, which is perfectly known by the controller.

At each time t, it is possible to predict an estimate $\hat{\mathbf{z}}_t$ of the state vector that should be obtained with the expected dynamics matrix $\hat{A}_t$ and the control signals $\mathbf{u}_t$:

$$\hat{\mathbf{z}}_{t+1} \quad = \quad \hat{A}_t\hat{\mathbf{z}}_t + B\mathbf{u}_t \tag{3.6}$$

When the true and the expected dynamics are not identical, meaning that $\hat{A}_t \neq A$, and that the point mass is subject to an unknown environmental force

field, there is a prediction error $\mathbf{e}_t$ defined as

$$\begin{aligned} \mathbf{e}_t &= \mathbf{y}_t - \hat{\mathbf{y}}_t & (3.7) \\ &= \mathbf{z}_t - \hat{\mathbf{z}}_t & (3.8) \end{aligned}$$

where Equation 3.8 stems from the fact that we assume that the information about the current state vector in complete for the controller. As a consequence, the feedback vector is exactly the current state vector and there is no temporal delay.

This error can then be used in an iterative Least Square parameter identification method, which is explained in Section 2.2 and can be used in this case for a vector containing only one element $\hat{L}_t$. It allows us to calculate an updated estimate of the force field factor with the following adaptive rule :

$$\begin{aligned} \hat{L}_{t+1} &= \hat{L}_t + \gamma \frac{\partial \hat{\mathbf{y}}_{t+1|t}}{\partial L} \mathbf{e}_t & (3.9) \\ &= \hat{L}_t + \gamma \frac{\partial \hat{\mathbf{z}}_{t+1|t}}{\partial L} \mathbf{e}_t & (3.10) \end{aligned}$$

for which the initial value is $\hat{L}_1 = 0$ and where the partial derivative of the state vector with respect to the unknown parameter is given by

$$\frac{\partial \hat{\mathbf{z}}_{t+1|t}}{\partial L} = [0,\ 0,\ \frac{1}{m}\hat{\dot{y}}_t,\ 0,\ 0,\ 0,\ 0,\ 0\ ] \quad (3.11)$$

In Equations 3.9 and 3.10, the learning rate γ can take a range of values that will be discussed in Section 3.2, where the values 0.05, 0.1 ans 0.25 and 0.5 were specifically studied and compared with the case where there is no sub-movement motor learning ($\gamma = 0$). Indeed, as explained in Section 2.2, learning rate has to be small but not infinitesimal [30].
The updated value of the perturbation parameter $\hat{L}_{t+1}$ can then be used to update the expected dynamics matrix $\hat{A}_{t+1}$ and to compute another time series of controller gains $-C_t(\hat{A}_{t+1})$.

Cost function

In order to compute the LQG controller gains in closed loop, we define the following cost function :

$$J(\mathbf{z}_N, \mathbf{u}_t) = w_1[(x_N - x^*)^2 + (y_N - y^*)^2] + w_2[\dot{x}_N^2 + \dot{y}_N^2] + \sum_{t=1}^{N-1} w_R ||\mathbf{u}_t||^2 \quad (3.12)$$

where w_1, w_2 and w_R are initially set to 1000, 20 and 10^{-5}, respectively, with the aim of getting bell-shaped velocity profiles. It did not appear essential to add

intermediate state penalties, because this cost function allows a qualitative fitting with the experimentally observed lateral force profiles.

3.2 Results

In this section, we present the results obtained for the adaptive control of a point mass moving in the horizontal plane, as defined in Section 3.1, with a complete state information. Simulations were repeated over 15 times, in order to obtain smooth averaged positions, velocities and forces profiles. The basic duration of hand reaching movement was set to 610 ms, the starting position was set at the origin of the Cartesian coordinate system and the target position was set 15 cm in the y-direction. The force field scaling factor was set to $\theta = 13$ Nsm^{-1} when it was in the clockwise direction and to $\theta = -13$ Nsm^{-1} in the counterclockwise direction.

First of all, only results for the clockwise perturbation force field are exposed. We comment the results related to the counterclockwise perturbation at the end of this section.

Different learning rates γ used in the parameter estimation with the Least Square technique (Section 2.2) were tested and compared to the case without motor learning, in order to find critical and best values for a relatively smooth deviated reaching movement. Simulated trajectories of the hand in the horizontal plane are shown in Figure 3.2.A, for these different γ values.

Estimation of the force field scaling factor

The estimation of the scaling factor of the force field $\hat{L}_t$ evolves throughout time and reaches different values, according to the chosen learning rate. Given its definition in Equation 3.10, the slope of the curve in Figure 3.2.B, depends on the partial derivative of the state vector with respect to the parameter (Equation 5.4), calculated by means of the current velocity in the y direction, $\dot{y}_t$. Moreover, the prediction error $\mathbf{e}_t$ is needed. These relationships appear correctly in Figure 3.2, with different learning rates.

We noticed that increasing the duration of the reaching movement does not mean that the parameter is better or faster estimated. In fact, the inverse situation happens because the time span is larger but with an unchanged traveled distance, the amplitude of the velocity $\dot{y}$ is then lower. Same learning rates were tested and gave the same behaviours that the one observed when the simulated movement

lasts 610 ms.

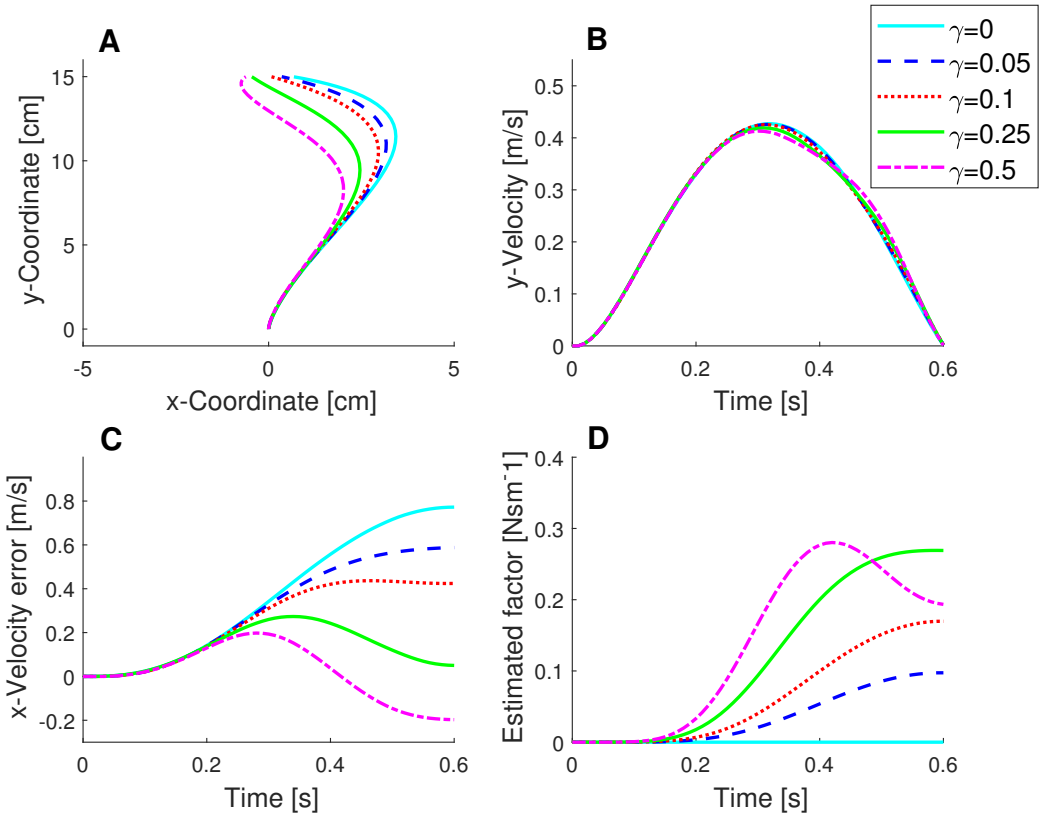

Figure 3.2: **Numerical experiments, complete state information, for different learning rates. A :** Mean trajectories in the horizontal plane. **B :** Mean velocity profiles in the y-direction. **C :** Evolution of the mean prediction error for the velocity in the x-direction. **D :** Evolution of the mean estimate of the force field scaling factor.

Critical and best learning rates

The prediction error (Figure 3.2.C) is stabilized along around the last hundred milliseconds, which is also the case for the parameter estimate near its final value, and tends to zero for what we define as the best values of the learning rate.

Nevertheless, for certain values of the learning rate, such as $\gamma = 0.5$, we observe that the error becomes negative and keeps growing in the negative area, leading to a poor convergence of the estimation of the force field factor at the end of the

reaching movement in Figure 3.2.D. This is also correlated with the trajectory in Figure 3.2.A, where, for $\gamma = 0.5$, there is a second correction needed at the end of the reaching movement. This means that the first correction is too large compared to the true perturbation force field applied on the point mass, because of an early overestimation of the perturbation parameter.

Worst, for large learning rates values, over $\gamma = 0.85$, the force field factor becomes negative and does not converge. As explained in [30] and in Section 2.2, γ has to be small.

Therefore, it seems that there is a certain critical value of the learning rate γ that should not be exceeded if the hand reaching movement is to remain smooth and without second correction. This is moreover confirmed by Figure 3.3.A, for a learning rate $\gamma = 0.5$, where we can guess the beginning of a third peak of lateral force at the end of the movement, when this critical learning rate is overtaken. This is also the case, to a lesser extent, for the lateral velocity profile in Figure 3.3.C. However, these observations do not qualitatively represent the results of the laboratory experiments for similar reaching movements, such as studied by Crevecoeur and coauthors in [1].

The maximum value that can take the learning rate for these simulations is $\gamma = 0.27$, with a precision in the order of 10^{-2}. This leads to an end estimate of the parameter equal to 0.273. For the sake of convenience, the value $\gamma = 0.25$ is used in the following analyses. Furthermore, the difference between the profiles obtained for these two values is very small. For this value of the learning rate, the force field factor reached an estimate of 0.271 at the end of the movement duration.

Influence of the learning rate value

Since the prediction error is lower when the learning rate increases, until the critical rate is reached, the maximal possible learning rate should also constitute the optimal one. Indeed, this the one for which the perturbation applied on the hand during the movement is "learned the best".

First of all, the second lateral force peak (peak end force in the x direction) presents a strong diminution, related to the learning rate, as depicted in Figures 3.3.A and 3.4.B. From the case of the absence of motor learning ($\gamma = 0$), the end peak force is decreased by ~ 2N with $\gamma = 0.05$. Compared to the latter, another decrease by ~ 1.8N is observed with $\gamma = 0.1$. Nonetheless, this maximum value of the lateral force occurs exactly at the same time for these learning rates. On the other hand, the peak end force appears 40ms sooner for $\gamma = 0.25$ and is anew

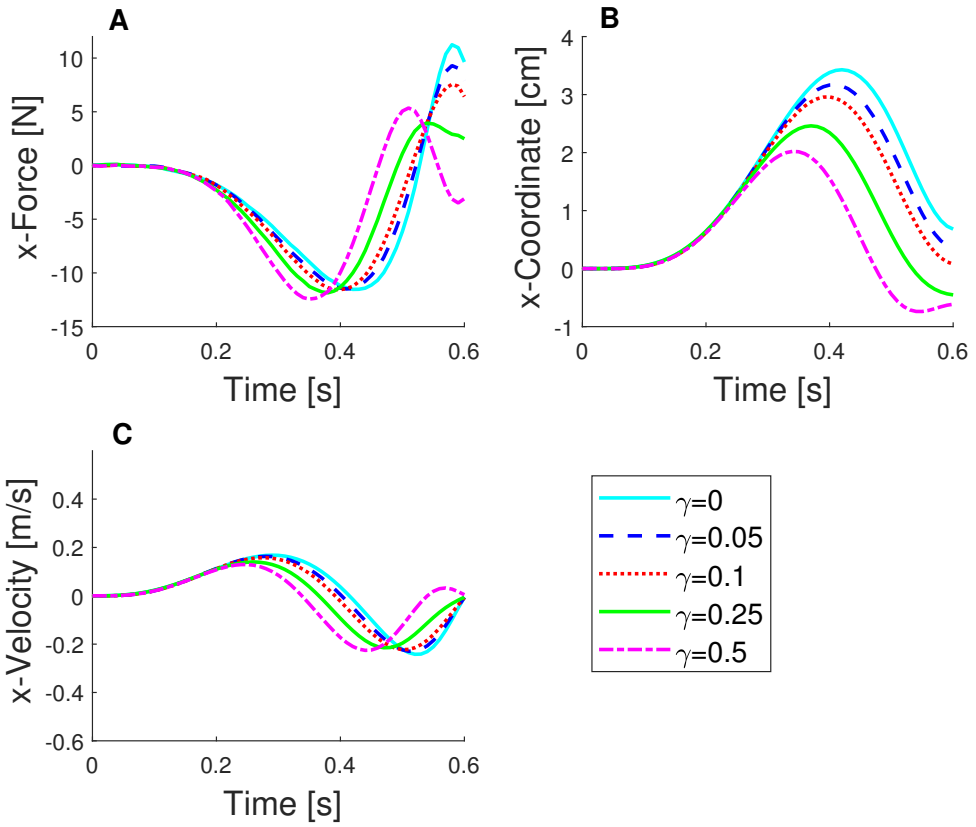

Figure 3.3: **Numerical experiments, complete state information, for different learning rates. A :** Mean lateral force profiles. **B :** Evolution of the mean x-position of the point mass. **C :** Mean velocity profiles in the x-direction.

reduced by 5N and 3.5N, compared to the two smallest learning rates.
These observations are confirmed by the computation of the correlation between the controlled force in the x direction and the external force field, evaluated for a range of possible learning rates and averaged over 15 trial simulations in Figure 3.4.A. The correlation significantly follows a linear trend, with respect to the learning rate.

The first peak of lateral force can also be analyzed (Figure 3.3.A), even though differences between curves are quite light. Actually, the peak occurs 20ms sooner for each increase of the learning rate (from 0.05 to 0.25). This is also fairly larger, increasing by almost 0.4N between $\gamma = 0.05$ and $\gamma = 0.1$.

From another point of view, the maximum lateral deviation (visible in Figure 3.3.B) occurs 10ms sooner for the learning rate of 0.1 and 40ms sooner for the

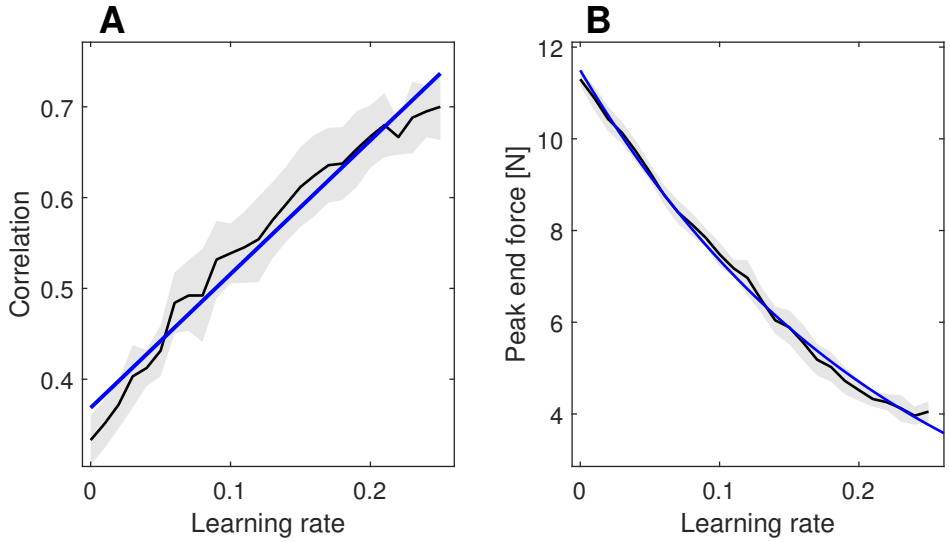

Figure 3.4: **Numerical experiments, complete state information. A :** Evolution of the correlation between the lateral force and the applied force field, with respect to the learning rate. Correlation was computed during a time interval from 100ms to 550ms and averaged across experiment repetitions. The shaded area represents one SEM. The blue curve is a significant linear regression. **B :** Evolution of the lateral peak end force, with respect to the learning rate. The shaded area represents one SEM. The blue curve is a decreasing exponential fit.

learning rate of 0.25, compared to the deviation with the learning rate equal to 0.05, which is a little bit earlier than the maximum deviation when there is no motor learning (10ms).
However, it can be seen that the end lateral position, meaning the x-coordinate at the end of the reaching movement, is not exactly equal to the target position, for $\gamma = 0.25$. Besides, a small overshoot appears (less than an half centimeter).

Finally, the position in the y direction and tangential force profiles are shown in Figure 3.5. Learning rate values do not influence these, except that, between 0.4s and 0.5s, the slope is less sharp compared to the case without online learning.

Counterclockwise force field

Similar results are observed when the force field perturbation is set to $-13\dot{y}_t$, as depicted in Figure 3.6. There is a strong reduction of the lateral end force peak, which which decreases with increasing learning rates, until a critical learning rate is reached. This critical learning rate is identical to the critical value of γ for a clockwise force field. The relationship between the lateral peak end force and the

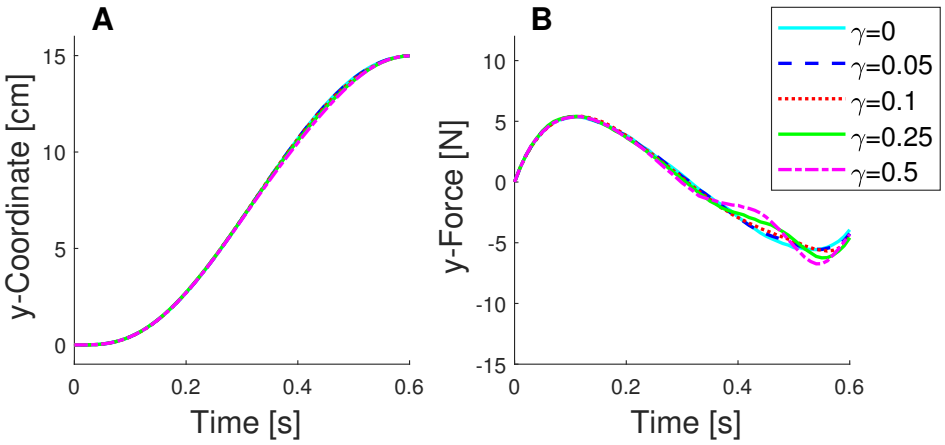

Figure 3.5: **Numerical experiments, complete state information, for different learning rates. A :** Evolution of the mean y-position of the point mass. **B :** Mean tangential force profiles.

learning rate also follows an exponential curve, with opposite sign, compared to the clockwise perturbation. Finally, the correlation between lateral force and external force field follows a linear trend with respect to the learning rate, which is identical to Figure 3.4.A and is consequently not repeated in this part.

3.3 Discussion

The objective of this chapter was to build an adaptive control model similar to the one presented in [1], assuming that the current state is exactly known by the controller. It aims to model the fact that human beings can execute reaching movements towards a goal target, even when they are subjected to an unknown mechanical disturbance all along the movement. The main hypothesis of this model is that the central nervous system is able to learn online and adapt the internal representation of the system dynamics at a very fast time scale, within the movement.

This model qualitatively fits the experimental laboratory results obtained by Crevecoeur and colleagues in [1]. Indeed, lateral force profiles correspond to those observed with subjects for which the target has to be reached in a time range from 60 to 80ms. However the amplitudes are not exactly identical, we observe the same tendency to reduce the end peak force when there is practice, meaning an increasing learning rate in our simulations. Likewise, the correlation between the lateral force and the applied force field is linearly related to the level of motor

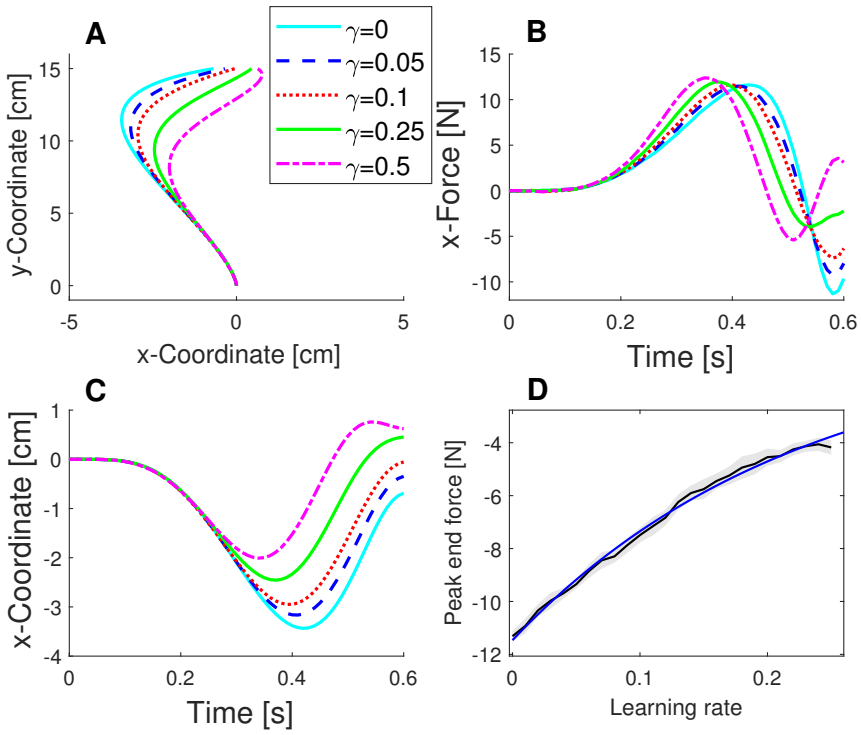

Figure 3.6: **Numerical experiments, complete state information, for different learning rates and counterclockwise force field. A :** Mean trajectories in the horizontal plane. **B :** Mean lateral force profiles. **C :** Evolution of the mean x-position of the point mass. **D :** Evolution of the lateral peak end force, with respect to the learning rate. The shaded area represents one SEM. The blue curve is an increasing exponential fit.

learning, associated in Figure 3.4.A with an increasing learning rate. Therefore, this model is a good candidate for linking online adaptation and practice throughout trials and motor learning on a trial-by-trial basis.

Most of previous studies about motor learning has been focused on the trial-by-trial learning, considering that there is no change in the internal representation during the ongoing movement. For example, in [28], authors modelled the correction of a deviated movement by introducing a stiffening mechanism allowing to correct the movement smoothly. Others have hypothesized that the central nervous system contains savings of previously used internal models and retrieves the appropriate controller, depending on the perturbation, as explained in Nonetheless, these models do not exactly match with the experiments studied in [1] and do not explain all their results.

In fact, in this paper, the internal representation is assumed to be the same at the beginning of the movement, with or without perturbation. Then the switching controllers cannot be a consistent model. Another proof is the absence of practice before the experiments and the fact that the subjects showed variability in their behaviour during altered trials even though the force field has a unique amplitude. In our model, using previously learned models is not necessary because we assume that the parameter related to the disturbance can potentially be tracked continuously during the movement, even in real time because the system is fully observable. Moreover, this model is consistent with all the experimental results in [1].

From a physiological point of view, we can wonder which value of the learning rate is representative of the instantaneous motor adaptation during an unexpectedly altered movement. We can guess that this value depends on the subject, it is quite small but is nonzero for healthy humans. Another question that emerges is, if the online motor learning is improved across trials, how fast is the change from one learning rate to another, with the aim of reaching the target accurately and with less effort.

From a modelling point of view, some limitations can be highlighted. Indeed, considering the translation of a simple point mass in the case of complete state information is not a realistic model of the arm and of functioning of the nervous system. This should lean on a more complex representation of the human arm, involving nonlinear dynamics equations. Furthermore, it has to account of the sensorimotor noise and sensory temporal delays in the feedback loop.

Chapter 4

Adaptive control : sensory feedback delays

Our interest is now directed to the addition of sensory temporal delay in the feedback loop along with Gaussian disturbances. Indeed, the adaptive control model used in Chapter 3 is relatively simple compared to the complexity of the neuromusuloskeletal human system, and neglects some biological factors such as latencies in the transfer of information.

Therefore, we keep the translation of a point mass in the horizontal place but we add sensorimotor noise and temporal delays between the execution of a command and the sensory feedback of a copy of the state vector.

4.1 Models and methods

The point mass model, depicted in Section 3.1, is again used with the same model parameters. The difference between this part of the work and Chapter 3 lies in the adaptive control system.

Adaptive control system with sensory feedback delays

Let us consider the same state and control vectors at each time t as used in the case without sensory delay (Section 3.1) : $\mathbf{z}_t = [x_t,\ y_t,\ \dot{x}_t,\ \dot{y}_t,\ F_{x,t},\ F_{y,t},\ x^*,\ y^*\ ]^T$ and $\mathbf{u}_t = [u_{x,t},\ u_{y,t}]^T$. However, when the current state vector is not fully observable and a feedback delay of h time steps is added, the sensory feedback vector becomes

$$\mathbf{y}_t \quad = \quad \mathbf{z}_{t-h} + \omega_t \tag{4.1}$$

The temporal delay for long-latency reflexes is larger or equal to 50ms [1, 10, 2]. We therefore chose the value of 50ms. In terms of time steps of 10ms, this corresponds

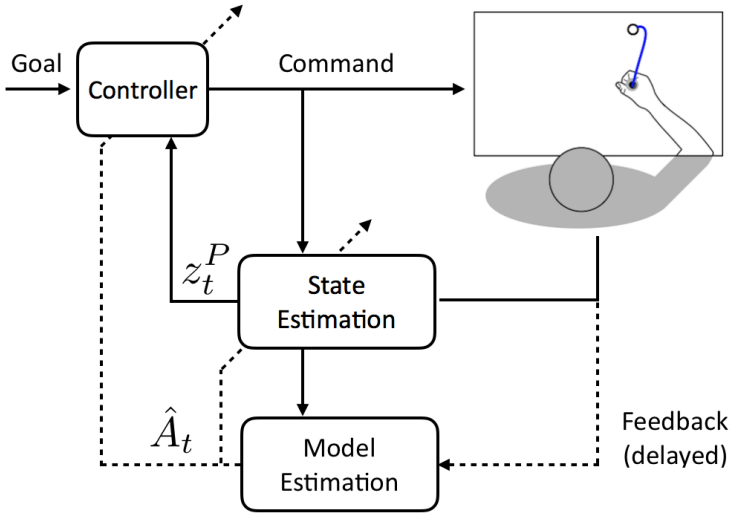

Figure 4.1: **Point mass model, with incomplete state information.** Schematic representation of an adaptive controller. The inner loop (solid line) accounts for the correction of state deviations. The outer loop (dashed line) accounts for the model corrections. This figure are adapted from the work of Crevecoeur et al. [1].

to a delay of $h = 5$ time steps. Therefore, the feedback vector can be equivalently defined as

$$\mathbf{y}_t \quad = \quad H[\mathbf{z}_t,\ \mathbf{z}_{t-1},\ \mathbf{z}_{t-2},\ \mathbf{z}_{t-3},\ \mathbf{z}_{t-4},\ \mathbf{z}_{t-5}]^T + \omega_t \tag{4.2}$$

where H is named the observability matrix, such that

$$H = [0_{8\times8}, 0_{8\times8}, 0_{8\times8}, 0_{8\times8}, 0_{8\times8}, H_h] \tag{4.3}$$

with $0_{8\times8}$ is the null matrix with the same dimension as the state vector and $H_h = I_{8\times8}$, the identity matrix, as used in [10].

In Equations 4.1 and 4.2, $\omega_t \sim \mathcal{N}(\mathbf{0}, \Omega_\omega)$ accounts for the sensory noise, which is modeled as a Gaussian disturbance with variance set to 10^{-6} [10].

In order to consider the h previous state vectors needed for the definition of the feedback vector, we must build an extended state vector, equal to $\mathbf{z}_t^{ext} = [\mathbf{z}_t,\ \mathbf{z}_{t-1},\ \mathbf{z}_{t-2},\ \mathbf{z}_{t-3},\ \mathbf{z}_{t-4},\ \mathbf{z}_{t-5}]^T$ along with an extended discrete-time state space representation with extended matrices $A^{ext} \in \mathbb{R}^{48\times48}$ and $B^{ext} \in \mathbb{R}^{48\times2}$. These matrices can be computed from the expressions of A and B used in the discrete-time

state space representation in Section 3.1, as well as the motor noise vector ε_t^{ext}, in order to obtain expressions that are consistent with the equations of the dynamics [11].

$$A^{ext} = \begin{bmatrix} A & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} \\ I_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} \\ 0_{8\times8} & I_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} \\ 0_{8\times8} & 0_{8\times8} & I_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} \\ 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & I_{8\times8} & 0_{8\times8} & 0_{8\times8} \\ 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & I_{8\times8} & 0_{8\times8} \end{bmatrix} \quad B^{ext} = \begin{bmatrix} B \\ 0_{8\times2} \\ 0_{8\times2} \\ 0_{8\times2} \\ 0_{8\times2} \\ 0_{8\times2} \end{bmatrix} \quad \varepsilon_t^{ext} = \begin{bmatrix} \varepsilon_t \\ 0_{8\times1} \\ 0_{8\times1} \\ 0_{8\times1} \\ 0_{8\times1} \\ 0_{8\times1} \end{bmatrix}$$

The expected dynamics is described, at each time step t, by a matrix $\hat{A}_t$, described in Section 3.1, and containing a current (time varying) estimate of the force field scaling factor $\hat{L}_t$. Using the same reasoning as previously to extend the state space matrices, this matrix must also be extended to match the new state vector dimension. Hence, we obtain the following expression for $\hat{A}_t^{ext}$:

$$\hat{A}_t^{ext} = \begin{bmatrix} \hat{A}_t & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} \\ I_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} \\ 0_{8\times8} & I_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} \\ 0_{8\times8} & 0_{8\times8} & I_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} \\ 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & I_{8\times8} & 0_{8\times8} & 0_{8\times8} \\ 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & 0_{8\times8} & I_{8\times8} & 0_{8\times8} \end{bmatrix}$$

However, the LQG controller does not know the true dynamics. Indeed, although we assume that the controller knows the nature of the parameter to learn, it has no knowledge of its true value. As a consequence, the sequence of feedback control gains $-C_t^{ext}(\hat{A}_t^{ext})$ have to be computed with the extended estimate $\hat{A}_t^{ext}$. Let us notice that the matrices of control gains are also extended in order to satisfy the new dimensions of the system of equations. Thus, the optimal control policy, $\mathbf{u}_t = -C_t^{ext}(\hat{A}_t^{ext})\mathbf{z}_t^P$, is determined in this case by the application of the controller gain to the current estimated state vector $\mathbf{z}_t^P$ obtained by using the Kalman filter gain. This vector approximates the real state vector $\mathbf{z}_t^{ext}$ based on the delayed feedback vectors and system knowledge. Since the true dynamics matrix A^{ext} is not known by the central nervous system, the estimated state vector is updated using the following expression :

$$\mathbf{z}_{t+1}^P = \hat{A}_t^{ext}\mathbf{z}_t^P + B^{ext}\mathbf{u}_t + K_t(\mathbf{y}_t - H\mathbf{z}_t^P) + \xi_t \tag{4.4}$$

where K_t is the current Kalman gain matrix, computed using the expected dynamics matrix $\hat{A}_t$, following the steps explained in Section 2.1, and $\xi_t \sim \mathcal{N}(\mathbf{0}, \Omega_\xi)$ is

the prediction noise, with variance set to $\Omega_\xi = 10^{-6}$.

Furthermore, the estimate of the state vector $\hat{\mathbf{z}}_t^{ext}$, obtained from the expected dynamics matrix $\hat{A}_t^{ext}$ and the current control signals $\mathbf{u}_t$, is predicted as follows:

$$\hat{\mathbf{z}}_{t+1}^{ext} = \hat{A}_t^{ext}\hat{\mathbf{z}}_t^{ext} + B^{ext}\mathbf{u}_t \tag{4.5}$$

When the point mass is subjected to an unknown environmental force field and that the expected dynamics does not match exactly the true one ($\hat{A}_t^{ext} \neq A^{ext}$), the prediction error $\mathbf{e}_t$ becomes

$$\mathbf{e}_t = \mathbf{y}_t - \hat{\mathbf{y}}_t \tag{4.6}$$
$$= \mathbf{z}_{t-h} - \hat{\mathbf{z}}_{t-h} \tag{4.7}$$

The adaptive rule for the estimation of the perturbation parameter, using the iterative Least Square identification technique (Section 2.2), is then :

$$\hat{L}_{t+1} = \hat{L}_t + \gamma\frac{\partial\hat{\mathbf{y}}_{t+1|t}}{\partial L}\mathbf{e}_t \tag{4.8}$$
$$= \hat{L}_t + \gamma\frac{\partial\hat{\mathbf{z}}_{t-h+1|t-h}}{\partial L}\mathbf{e}_t \tag{4.9}$$

for which the initial value is $\hat{L}_1 = 0$ and where the partial derivative of the feedback vector with respect to the unknown parameter is given by

$$\frac{\partial\hat{\mathbf{z}}_{t-h+1|t-h}}{\partial L} = [0,\ 0,\ \frac{1}{m}\hat{\dot{y}}_{t-h},\ 0,\ 0,\ 0,\ 0,\ 0\ ] \tag{4.10}$$

In Equations 4.8 and 4.9, the learning rate γ can take a range of values that will be discussed in section 4.2, as well as compared with the values tested in Section 3.2. The updated value of the perturbation parameter $\hat{L}_{t+1}$ can then be used to update the expected dynamics matrix $\hat{A}_{t+1}$ and to compute another sequence of controller gains $-C_t^{ext}(\hat{A}_{t+1}^{ext})$.

Cost function

The cost function used to implement the LQG controller closed loop is the same as the one applied in Section 3.1.

Observability matrix

In the field of motor control, the observability matrix is often defined as in Equation 4.3. However, different observability matrices can be used in the definition of the feedback vector $\mathbf{y}_t$, depending on the number of states variables that are observed in

the feedback vector. In the case of online motor learning during reaching, with the adaptive rule given by Equation 4.9, if the feedback loop can not access every state variable, there is a risk of a lack of information for the computation of the prediction error (equation 3.82) or of the derivative of the feedback vector in equation 4.8. The role of the observability matrix is therefore crucial.

4.2 Results

In this section, we present the results obtained for the adaptive control of a point mass moving in the horizontal plane, as defined in Section 4.1, with an incomplete state information. Indeed, the controller knows the state vector but only with a sensory delay of 50ms. Repeated simulations were performed with the same parameters as the ones used in Section 3.2.

First of all, results related to the clockwise perturbation force field are exposed and compared with the complete state information case of Section 3.2. We comment the results related to the counterclockwise perturbation at the end of this section.

Different learning rates γ, involved in the parameter estimation within the Least Squares algorithm (Section 2.2), were tested and compared to the case without motor learning and the learning rates of Section 3.2.

Estimation of the force field scaling factor

Given the adaptive rule of the force field factor $\hat{L}_t$ in Equation 4.9, its estimate converges to different values, depending on the value of the learning rate γ. For an identical learning rate value, it can be seen in Figures 4.2.D and 4.3.A that the final estimate, at the end of the reaching movement, is larger for the case with sensory delays compared to the case with a complete sate information. This can be explained by the fact that the tangential velocity (Figure 4.2.B) and the prediction error used for the estimation of the parameter (Figure 4.2.C) at time t are actually measures of what happened at time $t - h$, when they are higher. Indeed, we can notice that the maximum of the prediction error in Figure 4.3.B occurs 0.05s later when we account for sensory delay. As a consequence, with temporal feedback delays, a larger learning rate of $\gamma = 0.4$ is needed to reach a final value of the prediction error that is as small as the value obtained with $\gamma = 0.25$, for the fully observable case studied in the previous chapter.

Simulated trajectories, in Figure 4.2.A, are considerably deviated, and precision with respect to the target at the end of the movement is not optimal. In fact, the

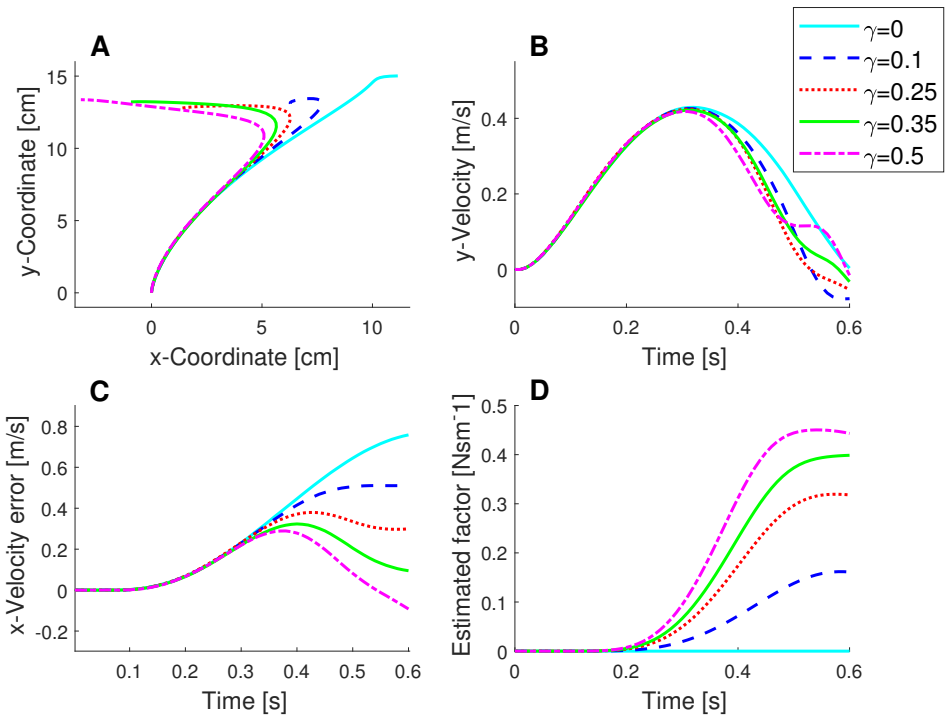

Figure 4.2: **Numerical experiments, incomplete state information, for different learning rates. A :** Mean trajectories in the horizontal plane. **B :** Mean velocity profiles in the y-direction. **C :** Evolution of the mean prediction error for the velocity in the x-direction. **D :** Evolution of the mean estimate of the force field scaling factor.

accuracy for the x-coordinate at the end of the movement is reasonable, at least for $\gamma = 0.35$ in which case there is a small overshoot of 0.6cm. However, the error in the final y-coordinate is larger, approaching 2cm. This happens even with a larger cost on the y coordinate or with a larger duration of the movement. This observation is confirmed by the y-velocity profile in Figure 4.2.B and tangential force profile in Figures 4.3.D and 4.4.B. We observe that the y-force and velocity drop down faster than in the case without delay and try to catch up during the last 100ms. This is consistent with the fact that the y-target is not exactly reached.

Nonetheless, we observe that in the case of no learning (i.e. $\gamma = 0$), the deviation attains more than 10cm. This confirms the fact that online learning is essential for the point mass to accurately reach the target. The difference at this point was smaller for the case without sensory delays.

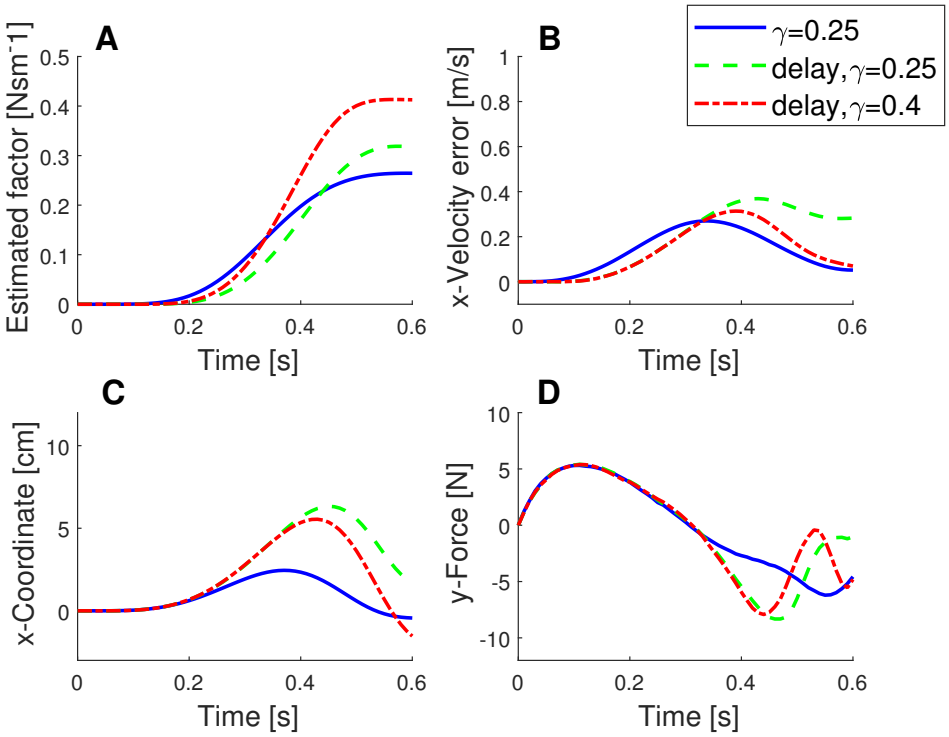

Figure 4.3: **Comparison between numerical experiments with complete and incomplete state information (instantaneous or delayed feedback). A :** Evolution of the mean estimate of the force field scaling factor. **B :** Evolution of the mean prediction error for the velocity in the x-direction. **C :** Evolution of the x-position of the point mass. **D :** Mean tangential force profiles.

Critical learning rates

During the numerical experiments, we noticed that some learning rates are more effective than others in order to model a reaching movement in presence of delays. Interstingly, they are not equal to values observed for the fully observable case. The critical value of γ for the representation of consistent trajectories (even if it is deviated) and lateral force profile is $\gamma = 0.43$, although the lateral force values are relatively large for such a movement. For larger learning rate values, the prediction error becomes negative and grows in the negative area (Figure 4.2.C) and the estimate of the perturbation parameter does not stabilize at the end of the movement duration (Figure 4.2.D). Moreover, above $\gamma = 0.43$, the tangential force profile (Figure 4.4.B) presents more than 2 oscillations at the end of the movement, meaning instabilities. Moreover, the lateral error with respect to the target is still growing (Figure 4.5.B).

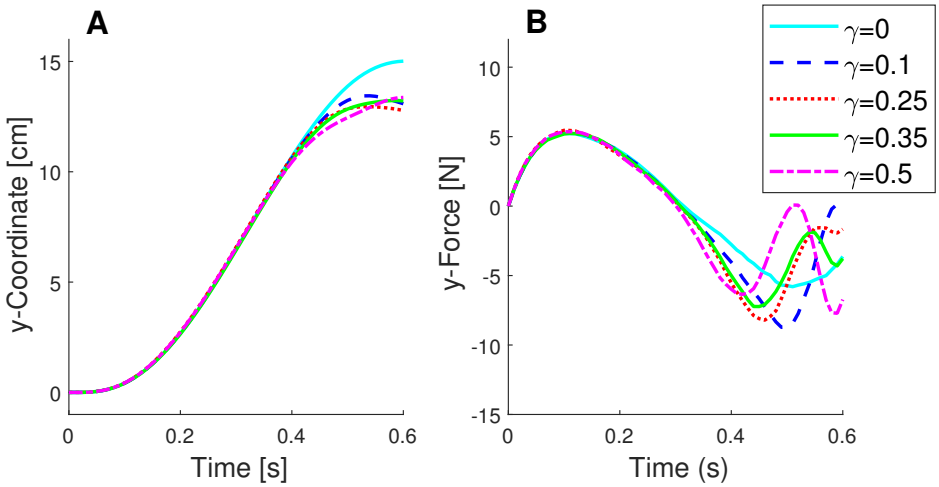

Figure 4.4: **Numerical experiments, incomplete state information, for different learning rates. A :** Evolution of the mean y-position of the point mass. **B :** Mean tangential force profiles.

Influence of the learning rate value

In terms of lateral force, in Figure 4.5.A, an increase of the learning rate leads to a decrease of the end peak, at least between $\gamma = 0.25$ and $\gamma = 0.35$ where the peak diminishes by ~ 4N. This can also be seen in Figure 4.6.B. Also, the correlation between the lateral force and the applied force field follows a linear increase with growing learning rate value, as shown in Figure 4.6.A. Furthermore, the first lateral force peak appears 20ms sooner for $\gamma = 0.35$ and $\gamma = 0.25$, compared to $\gamma = 0.1$, and its amplitude is more than 5N bigger.

On the other hand, among the available learning rates below $\gamma = 0.43$, a bigger learning rate does not especially mean an improved accuracy of the movement arounf the target, in particular in the y-direction (Figure 4.4.A). Indeed, even if the target is 15cm away in front of the starting position, the maximum reached value is around 13.2cm, whatever the learning rate value, except when there is no online learning.

Counterclockwise force field

Same behaviours as described in this section are observed for the case of a perturbation in the opposite direction and some of them are shown in Figure 4.7. In particular, we can see that the deviations of the trajectories are of the same order of magnitude in Figure 4.7.A, and the amplitude of the peak end lateral force decreases with the increase learning rate in Figure 4.7.D.

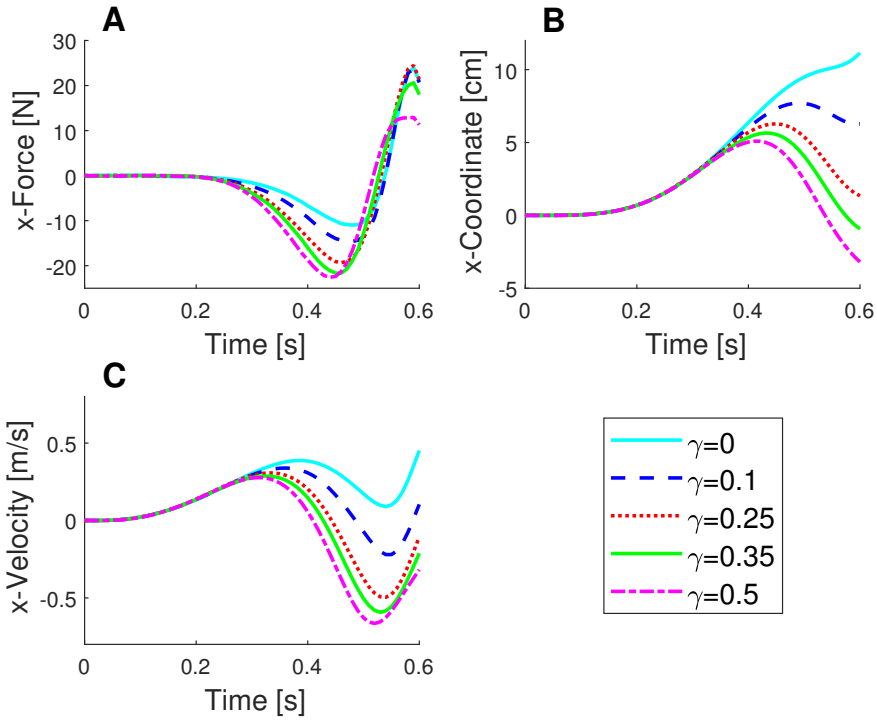

Figure 4.5: **Numerical experiments, incomplete state information, for different learning rates. A :** Mean lateral force profiles. **B :** Evolution of the mean x-position of the point mass. **C :** Mean velocity profiles in the x-direction.

Initialization of the parameter estimate

The initialization of the parameter estimate was set to $\hat{L} = 0$ at the beginning of each reaching movement, whatever the value of the learning rate. In Figure 4.8, we can observe the effect of the initialization of $\hat{L}_t$ to a nonzero value for the adaptive rule (Equation 4.9), and considering $\gamma = 0.35$. The initialization values were set to the end value obtained for the perturbation parameter for $\gamma = 0.1$ and $\gamma = 0.25$. The precision of the end position, compared to the target, seems better and the maximum deviation in the x direction is lower.

4.3 Discussion

The goal of this chapter was to improve the online learning model presented in Chapter 3 by adding a 50ms sensory delay along with sensorimotor noise to create a more realistic model.

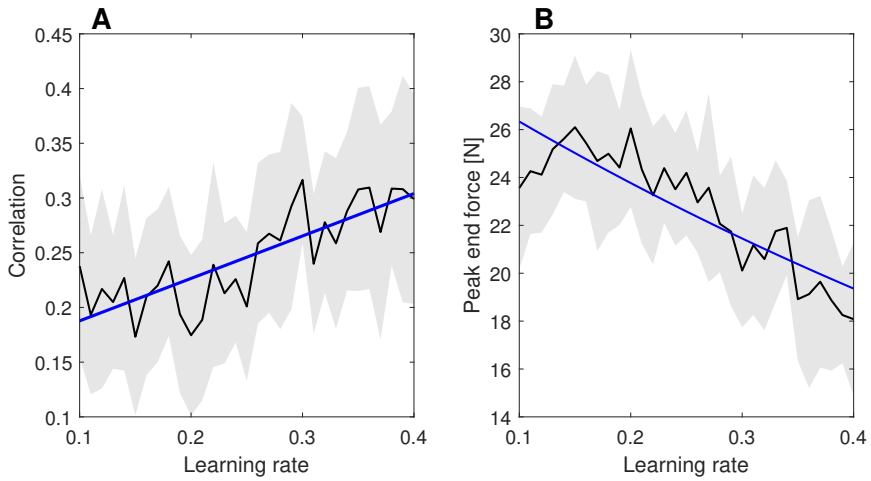

Figure 4.6: **Numerical experiments, incomplete state information. A :** Evolution of the correlation between the lateral force and the applied force field 50ms sooner, with respect to the learning rate. Correlation was computed during a time interval from 100ms to 550ms and averaged across experiment repetitions. The shaded area represents one SEM. The blue curve is a significant linear regression. **B :** Evolution of the lateral peak end force, with respect to the learning rate. The shaded area represents one SEM. The blue curve is a decreasing exponential fit.

As seen in Section 4.2, big differences compared to the previous model are the oscillations in the tangential force and velocity profiles, as well as the amplitude of the maximum lateral deviation and the poor correction of the trajectories and precision at the end of the movement. Moreover, the amplitude of the lateral force seems to be higher than the one observed with experimental results in [1] and does not seem physiologically consistent. Even though, its profile appears to be qualitatively similar. It seems that there is a trade-off between precision in the y-coordinate and the correction of the lateral deviation during a short time span.

These effects can be explained firstly by the presence of sensory delay and mostly by the fact that the current state is not fully observable. Therefore, the control signals have to be computed with an estimation of the state vector at time t, calculated by means of a Kalman filter gain and current knowledge of the system (Section 4.1). This knowledge depends essentially on the expected dynamics, that does not correspond to the true one because of the unknown perturbation, and the state of the system measured 50ms formerly.

However, we can wonder if the assumption to begin the movement with a zero initialization of the scaling factor force field estimate is appropriate. Indeed, in Section

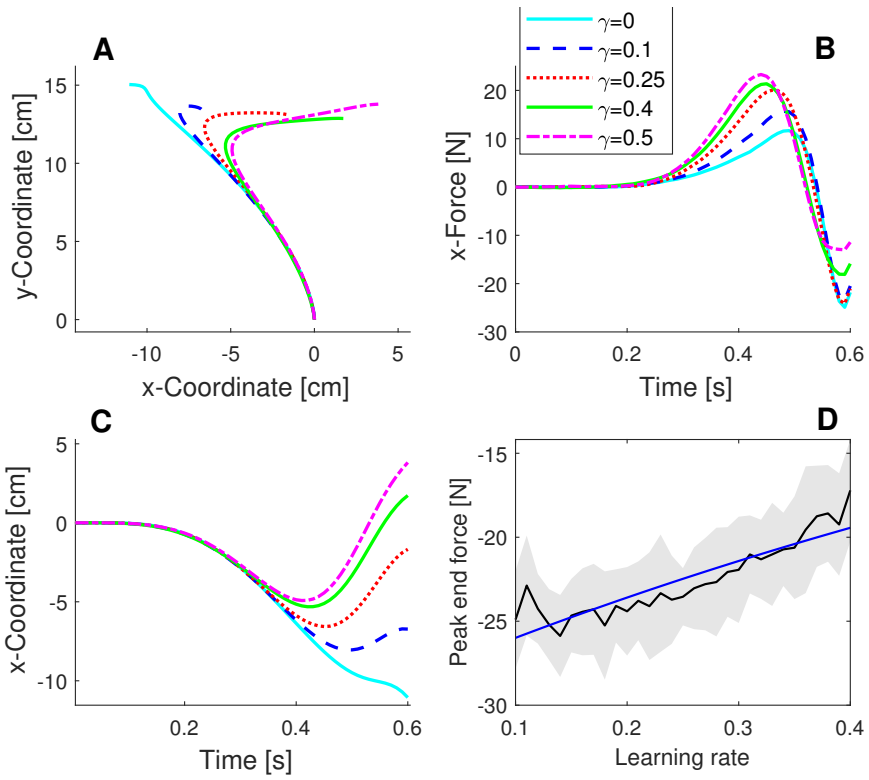

Figure 4.7: **Numerical experiments, incomplete state information, for different learning rates and counterclockwise force field. A :** Mean trajectories in the horizontal plane. **B :** Mean lateral force profiles. **C :** Evolution of the mean x-position of the point mass. **D :** Evolution of the lateral peak end force, with respect to the learning rate. The shaded area represents one SEM. The blue curve is an increasing exponential fit.

3.3, we drew a parallel between the increase of the learning rate and the trial-by-trial learning with practice, by means of savings with practice. Our hypothesis would be that, besides an increase of the online learning, the short-term motor memory could record the value on the previously calculated parameter, with a lower learning rate. It should explain why, when a human is subjected to the perturbation for the first time, the trajectory is dramatically deviated, to the point of not being able to reach the goal target, as seen in Figure 4.2.A for no online learning. Indeed the results, obtained with such an assumption at the end of Section 4.2, seem promising.

On the other hand, another factor to be determined is the approximated duration of the model update (outer loop of the system). Indeed, it is agreed that long-latency reflexes have a latency of more than 50ms, meaning that motor correc-

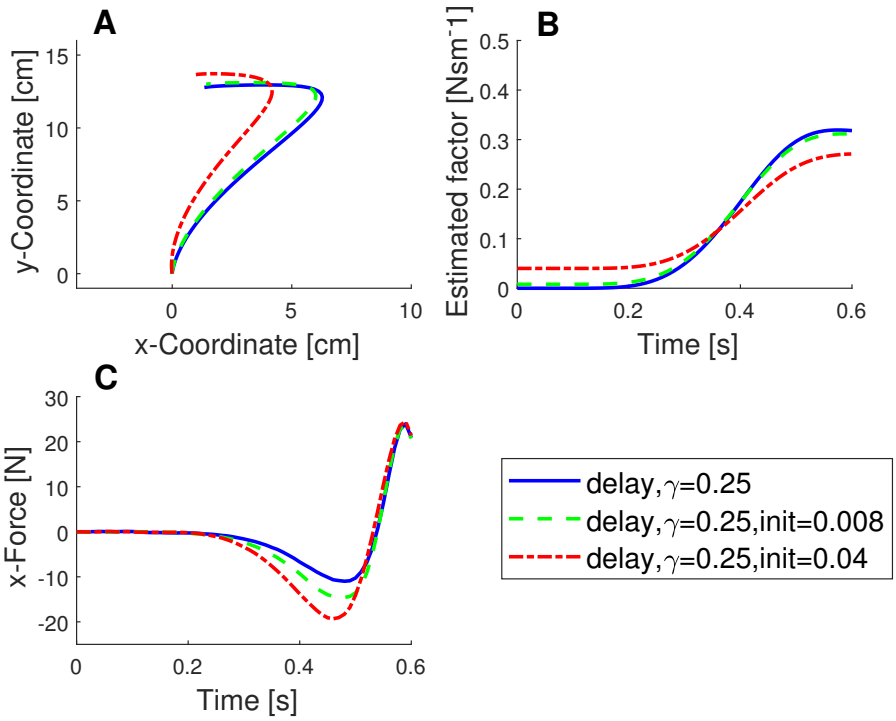

Figure 4.8: **Comparison between numerical experiments, incomplete state information (delayed feedback), for different estimate initialization values. A :** Mean trajectories in the horizontal plane. **B :** Evolution of the mean estimate of the force field scaling factor. **C :** Mean lateral force profiles.

tion are engaged in this range of time but we have to find the latency of the model correction. This correction is perhaps slower than the correction of the movement.

Finally, the model presented in this chapter still contains non-realistic features. Indeed, the point mass model does not take into account the complex geometry of a human arm and the associated nonlinear dynamics. This constitutes the next step for the modelling of online motor learning.

Chapter 5

Adaptive control : point mass in a curl field

In Chapter 4, the adaptive control model developed in Chapter 3 was modified accordingly in order to take sensory delays into account. Now, we are interested in the case of a multiple components perturbation applied to the arm during a reaching movement. In this chapter, we use an example of curl force field [35], which is defined as the force field resulting from the combination of two orthogonal force fields, in the x- and y-direction respectively.

5.1 Models and methods

The translation of a point mass model depicted in Section 3.1 is used again. The adaptive control framework developed in Section 4.1 with sensory delays is also employed.

However, the force field applied to the point during the reaching movement is consists in two components, one for each direction of the horizontal plane and of opposite sign.

Adaptive control system with sensory feedback delays

Since we take the sensory feedback delay into account, the control system described in Section 4.1 is putting into practice, as depicted in Figure 4.1. The same state variables, command variables and parameters are used.

The applied force field is a curl field, defined with two non-zero orthogonal components in the horizontal plane, such as the environmental disturbance is

$$\begin{bmatrix} F_{ENV,x} \\ F_{ENV,y} \end{bmatrix} = \begin{bmatrix} 0 & L \\ -L & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \tag{5.1}$$

where L is set to 9.19 Nsm^{-1}, in order to obtain a resultant intensity equal to 13 Nsm^{-1}. $L_1 = L$ and $L_2 = -L$ are the scaling factors of the force field.
As a consequence, the true dynamics matrix A and the expected dynamics matrix $\hat{A}_t$ are impacted. Indeed, in Equation 3.2 of the point motion for the y-coordinate, the force field component $F_{ENV,y}$ is added. The linear algebraic form of the system is composed of the continuous-time A_c and B_c matrices as follows

$$A_c = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -G/m & L_1/m & 1/m & 0 & 0 & 0 \\ 0 & 0 & L_2/m & -G/m & 0 & 1/m & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1/\tau & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1/\tau \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad B_c = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1/\tau & 0 \\ 0 & 1/\tau \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

It is noticed that B_c remains unchanged. The system has then to be transformed in a discrete-time state space representation, applying the Euler explicit technique, as explained in Section 3.1.

On the other hand, the time-varying expected dynamics matrix $\hat{A}_t$ contains the two estimates $\hat{L}_{1,t}$ and $\hat{L}_{2,t}$ of the force field scaling factors L_1 and L_2, respectively, at each time t :

$$\hat{A}_t = \begin{bmatrix} 1 & 0 & \delta t & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \delta t & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 - \delta t G/m & \boxed{\hat{L}_{1,t}/m} & \delta t/m & 0 & 0 & 0 \\ 0 & 0 & \boxed{\hat{L}_{2,t}/m} & 1 - \delta t G/m & 0 & \delta t/m & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -\delta t/\tau & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -\delta t/\tau \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

In order to take the sensory feedback delay into account, the state vector and matrices of the discrete-time state space representation have to be extended, following the process in Section 4.1. From now, this is then these extended matrices that are considered.

The extended matrix $\hat{A}_t$ in used to compute the sequence of Linear-Quadratic-Gaussian controller gains $-C_t(\hat{A}_t^{ext})$ that are applied to the estimate of the current state vector computed by Equation 4.4 in order to calculate the current motor

command signals : $\mathbf{u}_t = -C_t(\hat{A}_t^{ext})\mathbf{x}_t^P$. More details can be retrieved in Section 4.1.

Afterwards, we compute the prediction error $\mathbf{e}_t$, as in Equation 4.7, that appears when the true and the expected dynamics are not consistent with each other.
Let $\hat{\mathbf{L}}_t$ denote a vector of the scaling factors estimates $[\hat{L}_{1,t}, \hat{L}_{2,t}]^T$. The adaptive rule for the estimation of the perturbation parameters, using the Recursive Least Square identification technique (Section 2.2), is then :

$$\hat{\mathbf{L}}_{t+1} = \hat{\mathbf{L}}_t + \gamma I_2 \frac{\partial \hat{\mathbf{y}}_{t+1|t}}{\partial \mathbf{L}} \mathbf{e}_t \tag{5.2}$$

$$= \hat{\mathbf{L}}_t + \gamma I_2 \frac{\partial \hat{\mathbf{z}}_{t-h+1|t-h}}{\partial \mathbf{L}} \mathbf{e}_t \tag{5.3}$$

for which the initial vector is $\hat{\mathbf{L}}_1 = [0, 0]^T$ and where I_2 is the identity matrix $\in \mathbb{R}^2$. The partial derivatives of the feedback vector with respect to the unknown parameters are given by

$$\frac{\partial \hat{\mathbf{z}}_{t-h+1|t-h}}{\partial L_1} = [0,\ 0,\ \frac{1}{m}\hat{\dot{y}}_{t-h},\ 0,\ 0,\ 0,\ 0,\ 0\ ] \tag{5.4}$$

$$\frac{\partial \hat{\mathbf{z}}_{t-h+1|t-h}}{\partial L_2} = [0,\ 0,\ 0,\ \frac{1}{m}\hat{\dot{x}}_{t-h},\ 0,\ 0,\ 0,\ 0\ ] \tag{5.5}$$

In Equations 5.2 and 5.3, the learning rate γ can take a range of values that will be discussed in section 5.2. Moreover, it could be interesting to turn one of the true scaling factors to 0 Nsm^{-}1 and observe how the results of the adaptive rule react.

The updated values of the perturbation parameters $\hat{L}_{1,t+1}$ and $\hat{L}_{2,t+1}$ can then be used to update the expected dynamics matrix $\hat{A}_{t+1}$ and to compute another sequence of controller gains $-C_t^{ext}(\hat{A}_{t+1}^{ext})$.

Cost function

The cost function used to implement the LQG controller closed loop is the same as the one applied in Section 3.1.

5.2 Results

In this section, results obtained with the model explained in Section 5.1 are analyzed, for the application of a curl force field perturbation during the translation of a point mass in the horizontal plane. The goal of these results is to show how the adaptive controller reacts when it has to estimate two force field scaling factors,

with identical learning rates γ. The consequences on the learning rate value are also commented.

In order to validate this implementation, we compare the results obtained for the curl field described FF_A in Section 5.1 with the following field FF_B :

$$FF_A : \begin{bmatrix} F_{ENV,x} \\ F_{ENV,y} \end{bmatrix} = \begin{bmatrix} 0 & 9.19 \\ -9.19 & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \quad \text{vs.} \quad FF_B : \begin{bmatrix} F_{ENV,x} \\ F_{ENV,y} \end{bmatrix} = \begin{bmatrix} 0 & 13 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix}$$

Indeed, the second force field FF_B is the same as the one tested in the precedent chapter, and it must then produce similar curves compared to the ones observed in Section 4.2. Moreover, it allows us to verify that a zero scaling factor leads indeed to a zero parameter estimate, thereby validating our implementation.

Trajectories in presence of the perturbation FF_B correctly fit with trajectories observed in Section 4.2, as well as the estimate of the nonzero force field factor and forces profiles. These results are confirmed in Figures 5.1.B and .D as well as in Figures 5.3.C and .D.

Concerning the curl force field FF_A, trajectories in Figure 5.1.A display a good correction in the x-direction for learning rates $\gamma = 0.25$ and $\gamma = 0.35$. However, the target is not exactly reached, especially in the y-coordinate. Indeed, we can guess a certain trade-off in the accuracy between the two coordinates, as highlighted in Section 4.2, certainly due to the short time span of the movement and the application of the perturbation. In addition to that, it seems logical because of a force field component in the y-direction.

On the other hand, the evolution of the estimate $\hat{L}_{1,t}$, in Figure 5.1.C, is consistent with the previous analyses and the smaller value to correct (9.19 Nsm^{-1}). The only difference is the critical value of the learning rate that still enables the convergence of the parameter estimate. In fact, this value seems to depend not only on the presence of delay or not, but also on the type of force field applied on the point mass.

More surprisingly, even if the same learning rate is used for the estimate of the second factor $\hat{L}_{2,t}$ in Figure 5.2.A, its end value is far from the value reached by the estimate of $\hat{L}_{1,t}$, even if their respective intensities are the same. Furthermore, the estimate grows in the positive area even if the true value is negative, and vice et versa. Nevertheless, as mentioned earlier, the trajectory is consistent with what was expected when applying such a force field FF_A. Moreover, the evolution of the estimate $\hat{L}_{2,t}$ is consistent with the evolution of the component of the prediction

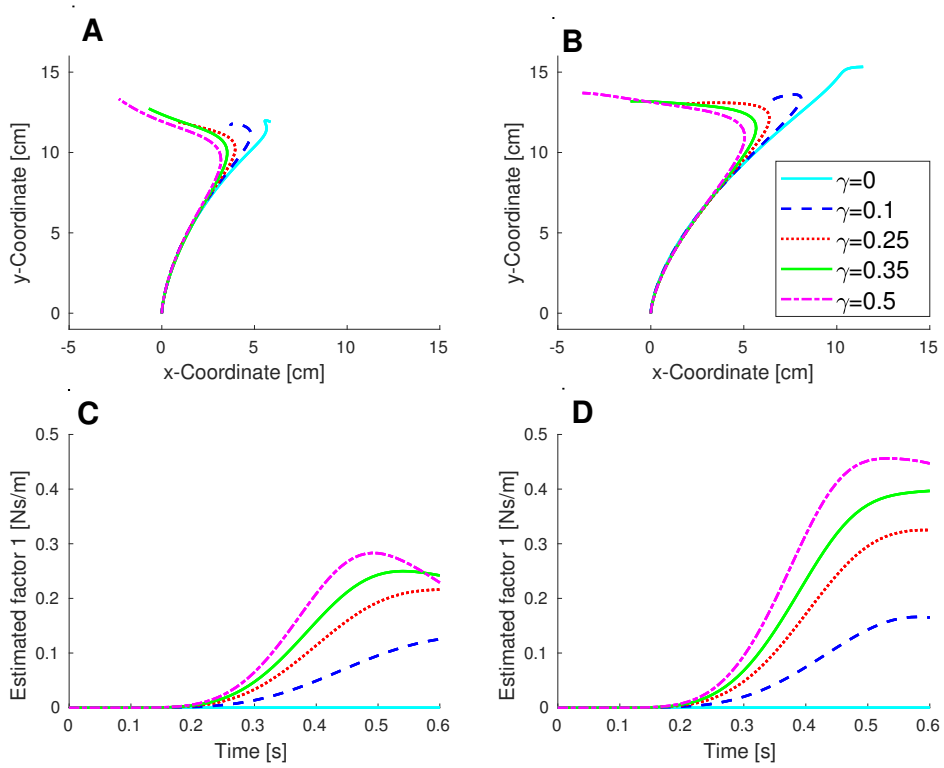

Figure 5.1: **Numerical experiments, incomplete state information, for different learning rates. A :** Mean trajectories for the curl force field. **B :** Mean trajectories for the x-direction force field. **C :** Evolution of the mean estimate of the x-component scaling factor, for the curl force field. **D :** Evolution of the mean estimate of the force field factor for the x-direction force field.

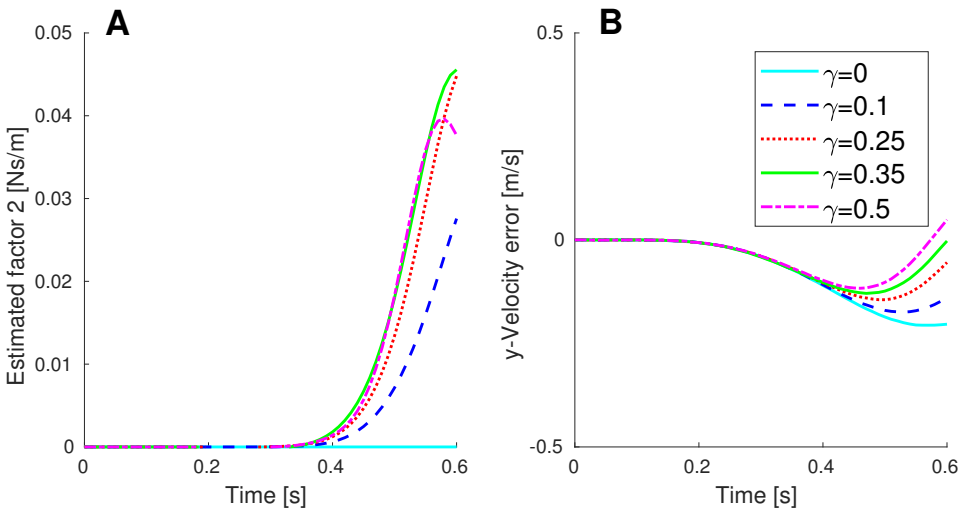

Figure 5.2: **Numerical experiments, incomplete state information, for different learning rates. A :** Evolution of the mean estimate of the y-component scaling factor, for the curl force field. **B :** Evolution of the mean prediction error for the velocity in the y-direction.

error used for its calculation with Equation 5.5, represented in Figure 5.2.B.

Finally, the lateral force profiles with FF_A in Figure 5.3.B present similar behaviours than the profiles observed with FF_B. However, because of the presence of a y-component force field, the tangential force profiles in Figure 5.3.B are affected at the end of the movement, when the prediction error is the largest.

5.3 Discussion

Te goal of this chapter was to demonstrate the flexibility of the adaptive control model presented in Chapter 4, considering sensory delays. Therefore, a more general kind of force field, named the curl force field, was tested.

First of all, similar results compared to Section 4.2 are observed when the force field corresponds to a force purely applied in the x-direction. Secondly, the trajectory correction and tangential force profile are consistent for the curl force field. However, the estimate of the second scaling factor (for the y-component of the force field) is not computed as accurately as the first one. This is due to fact that the synergistic effect of the curl force field is hard to unravel for the controller. Our hypothesis to explain this phenomenon is that, since the adaptive controller

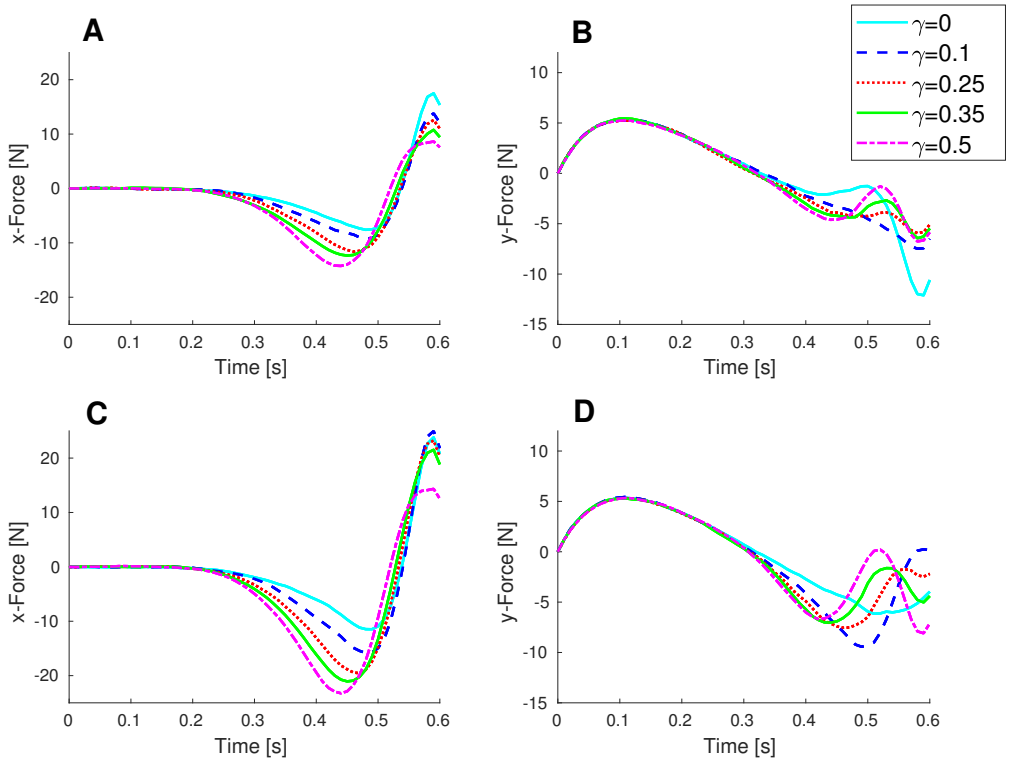

Figure 5.3: **Numerical experiments, incomplete state information, for different learning rates. A :** Mean lateral force profiles for the curl force field. **B :** Mean tangential force profiles for the curl force field. **C :** Mean lateral force profiles for the x-direction force field. **D :** Mean tangential force profiles for the x-direction force field.

does not know in advance that the perturbation is the result of the combination of two forces applied along orthogonal directions, it prioritizes the parameter to estimate, according more importance to the components along which the prediction error is larger.

Afterwards, as highlighted in Chapters 3 and 4, these findings have to be further investigated in future work with physiologically determined values for the learning rates, as well as the latency of the outer and the inner loop in Figure 4.1.

Finally, the major limitation of this model lies in the linear modelling of the dynamics. We should keep in mind that the true complex biomechanics of the arm has to be described by non linear equations in order to better reflect the reality.

Chapter 6

Adaptive control : 2-joint arm model

The main limitation of the previous adaptive models in Chapters 3, 4 and 5 is the representation of the arm by means of a point mass translation motion in the horizontal plane. Nonetheless, a non-linear arm representation leads to a non-linear biomechanical system dynamics and then, to the inability to use linear control theory explained in Chapter 2. This can be partly bypassed by employing local approximation and by neglecting non-linear non-linear terms.

In this chapter, we bring such a linearized model to light, as one of the potential adaptive control model of a human arm. A basic force field factor estimation is tested by way of preliminary results and foundations are built for future work in the adaptive control framework applied to the arm reaching movements.

6.1 Models and methods

2-link arm model

The 2-link human arm model consists in a planar representation of the arm composed of two segments with two joints, at the shoulder and at the elbow. As depicted in Figure 6.1, the origin of the system is set to the shoulder joint and the angles q_s and q_e describe the position of the hand in the horizontal plane. Each segment is represented by an index : $_1$ reflects characteristics of the upper part of the arm and $_2$ refers to the forearm/hand segment.

Moreover, the dynamics of multi-joint movements cannot be reduced to simple Newton's equations, as it was the case for the point mass model, because a torque at one of the two joints can lead to the motion of the other one. Therefore, we

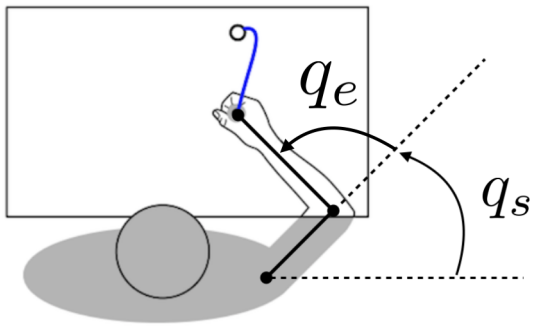

Figure 6.1: **2-link joint arm model.** Schematic illustration of the simulated experiment, for which the subject has to perform a reaching movement in the horizontal plane.

use equations proposed by Li and Todorov in [34], by neglecting the non-linear centripetal and Coriolis forces and using local approximation when needed.

Considering the joint angles vector $\mathbf{q} = [q_s, q_e]^T$, the dynamics equations of such a human arm representation can be gathered as follows :

$$\mathcal{M}\ddot{\mathbf{q}} = \mathbf{T} - \mathcal{B}\dot{\mathbf{q}} \tag{6.1}$$

In Equation 6.1, $\mathcal{M} \in \mathbb{R}^{2\times2}$ is the inertia matrix (positive semidefinite), $\mathbf{T} = [T_s, T_e]^T$ is the joint torque and $\mathcal{B} \in \mathbb{R}^{2\times2}$ is the joint friction matrix.

The intertia and the joint friction matrices are defined as

$$\mathcal{M} = \begin{bmatrix} a_1 + 2a_2cos(q_e) & a_3 + a_2cos(q_e) \\ a_3 + a_2cos(q_e) & a_3 \end{bmatrix} \quad \mathcal{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

where the elements of the joint friction matrix are $b_{11} = b_{22} = 0.05$ and $b_{21} = b_{12} = 0.025$.

The inertia matrix elements are expressed as

$$a_1 = I_1 + I_2 + m_2 l_1^2 \tag{6.2}$$
$$a_2 = m_2 l_1 s_2 \tag{6.3}$$
$$a_3 = I_2 \tag{6.4}$$
$$cos(q_e) \approx cos(q_{e,0}) \tag{6.5}$$

where m_2 is the mass of the forearm link (1kg), I_1 and I_2 are the moments of inertia of the two links (0.025kgm^2 and 0.0.045kgm^2), l_1 and l_2 are the length of each link (30cm and 33cm), s_2 is the distance from the centre of mass of the forearm link and the joint centre (16cm) and $q_{e,0}$ is the starting angle at the elbow joint.

Finally, the muscle dynamics is approximated with a first order linear filter between the command variables $\mathbf{u} = [u_e, u_y]^T$ and the controlled joint torques $\mathbf{T}$, with a time constant τ set to 0.06s, as in [31] :

$$\tau \dot{\mathbf{T}} = \mathbf{u} - \mathbf{T} \tag{6.6}$$

Considering vector $\mathbf{z} = [q_s,\ q_e,\ \dot{q}_s,\ \dot{q}_e,\ T_s,\ T_e,\ q_s^*,\ q_e^*]^T$ and the constant goal target angles position $(q_s^*,\ q_e^*)$, the differential equations of the dynamics can then be gathered into an linear algebraic form $\dot{\mathbf{z}} = A_c\mathbf{z} + B_c\mathbf{u}$ (c index stands for "continuous-time").

Furthermore, we consider a force field applied in order to disturb the arm motion and defined as a dissipative velocity-dependent elbow joint load [36] :

$$\begin{bmatrix} T_{ENV,s} \\ T_{ENV,e} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & L \end{bmatrix} \begin{bmatrix} \dot{q}_s \\ \dot{q}_e \end{bmatrix} \tag{6.7}$$

where L is set to -1.5Nsm^{-1}. This equation has to be added in the continuous representation of the system, in order to get the true dynamics equations.

Adaptive control model

We would like to obtain a discrete-time state space representation of the system, in order to add stochastic disturbances and to solve a discrete-time LQG optimal feedback problem, as described in Section 3.1 and Figure 3.1(right). Let $\mathbf{z}_t$ and $\mathbf{u}_t$ denote respectively the state vector and the control vector at each time t, such that $\mathbf{z}_t = [q_{s,t},\ q_{e,t},\ \dot{q}_{s,t},\ \dot{q}_{e,t},\ T_{s,t},\ T_{e,t},\ q_s^*,\ q_e^*\ ]^T$ and $\mathbf{u}_t = [u_{x,t},\ u_{y,t}]^T$. The discrete-time state space representation of this system is

$$\mathbf{z}_{t+1} = A\mathbf{z}_t + B\mathbf{u}_t + \varepsilon_t \tag{6.8}$$

where A and B are the discretized matrices. They can be obtained by using explicit Euler integration technique, as done in Section 3.1. In Equation 6.8, the motor noise ε_t is a zero-mean Gaussian disturbance with covariance matrix $\Omega_w = BB^T$.

Nevertheless, the case where the force field in Equation 6.7 is unexpectedly applied to the hand, the current expected dynamics is described by a matrix $\hat{A}_t$ at each time t. The latter contains an estimate of the time varying unknown scaling factor $\hat{L}_t$. This expected matrix is then used to compute the gains sequence of the LQG controller, applied to the current state vector in order to determine the control vector : $\mathbf{u}_t = -C_t(\hat{A}_t)\mathbf{z}_t$. Indeed, we start with the fully observable case.

At each time t, it then possible to calculate a prediction error, as defined in Equation 3.8, between the observed state vector $\mathbf{z}_t$ and the expected state vector $\hat{\mathbf{z}}_t$.

The adaptive rule for the estimate of the force field factor is the same as the rule depicted in Equations 3.9 and 3.10 :

$$\begin{aligned} \hat{L}_{t+1} &= \hat{L}_t + \gamma \frac{\partial \hat{\mathbf{y}}_{t+1|t}}{\partial L} \mathbf{e}_t \\ &= \hat{L}_t + \gamma \frac{\partial \hat{\mathbf{z}}_{t+1|t}}{\partial L} \mathbf{e}_t \end{aligned}$$

for which the initial value is $\hat{L}_1 = 0$, γ is the learning rate and where the partial derivative of the state vector with respect to the unknown parameter is given by

$$\frac{\partial \hat{\mathbf{z}}_{t+1|t}}{\partial L} = [0,\ 0,\ \frac{-1}{\det\mathcal{M}}(a_3 + a_2 cos(q_{e,0})\hat{\dot{q}}_{e,t},\ \frac{1}{\det\mathcal{M}}(a_1 + 2a_2 cos(q_{e,0})\hat{\dot{q}}_{e,t},\ 0,0,0,0[\quad (6.9)$$

The updated value of the perturbation parameter $\hat{L}_{t+1}$ can then be used to update the expected dynamics matrix $\hat{A}_{t+1}$ and to compute another sequence of controller gains $-C_t(\hat{A}_{t+1})$.

Cost function

The cost function used to implement the LQG controller closed loop is the same as the one applied in Section 3.1, except that the control cost weight w_R is set to 10^{-3}.

Planar transformations

In order to switch from cartesian coordinates (x, y) of the hand to angular coordinates of the joint, we can use the formulae derived by Hollerbach and Flash in [37]:

$$q_e = \cos^{-1}\left(\frac{x^2 + y^2 - l_1^2 - l_2^2}{2 l_1 l_2}\right) \quad (6.10)$$

$$q_s = \tan^{-1}\left(\frac{y}{x}\right) - \tan^{-1}\left(\frac{l_2 \sin q_e}{l_1 + l_2 \cos q_e}\right) \quad (6.11)$$

In Equations 6.10 and 6.11, l_1 and l_2 stand for the length of the upper arm and the forearm respectively. It is also possible to convert torques $\mathbf{T}$ to forces $\mathbf{f}$ using the following expression [22]:

$$\mathbf{f} = (J(\mathbf{q})^T)^{-1}\mathbf{T} \quad (6.12)$$

In Equation 6.12, $J(\mathbf{q})^T$ is the transpose of the Jacobian of the mapping from (q_s, q_e) to (x, y) and is computed as follows:

$$J(\mathbf{q})^T = \begin{bmatrix} \frac{\partial x}{\partial q_s} & \frac{\partial x}{\partial q_e} \\ \frac{\partial y}{\partial q_s} & \frac{\partial y}{\partial q_e} \end{bmatrix} \tag{6.13}$$

6.2 Future perspectives

Our preliminary results show that the convergence of the parameter estimate update (Equation 3.9) is possible only with very small learning rate values γ, in the order of 0.00001. As explained in Section 2.2, the Recursive Least Square technique, associated with a LQG optimal controller, needs a small but not infinitesimal learning rate γ. Therefore, these results do not seem consistent with this statement.

Furthermore, in order to simulate any kind of reaching movements, more general equations have to be developed for the planar transformations, with the aim of covering all the quadrant sectors of the horizontal plane. Actually, the equations described in Section 6.1 correspond to an arm representation in the first quadrant, also called the right work space in [22]. Without these general equations, some problems could occur with the computation of the Jacobian matrix in Equation 6.12, when the hand is located in the left quadrant for example.

Moreover, another track to follow could be the conversion of the state feedback vector into Cartesian coordinates before the use of the adaptive rule, following the equations in the end of the Section 6.1. Nevertheless, the Jacobian matrix can also lead to convergence problems for the adaptive rule if it is difficult to obtain its inverse.

Finally, future work would be focused on the modelling of more complex force fields, with four nonzero components, except of inly one in Equation 6.7.

To conclude, we can qualified the 2-link joint arm model as a good approximation of the true human arm. In fact, good results are obtained in the framewotk of LQG control problem, when the dynamics is fully known by the controller and by using the linear approximation of the nonlinear dynamics equations explained in Section 6.1.

Chapter 7

Conclusion

The motor function is probably one of the most basic, yet essential abilities of human beings. Over the last decades, a great body of work, across several disciplines, has been devoted to the understanding of this function. In particular, mathematical approaches have offered important insights into the neural mechanisms underlying the processes involved in motor control as well as motor learning, through the application of control theory methods.

In this thesis, we investigated the capacity of the central nervous system to perform an ongoing adaptation of its internal representation of the environment, in the presence of an unexpected disturbance. Linear-Quadratic-Gaussian control and Least Squares identification methods have proved useful to accurately model such online motor learning. Across numerical experiments, we have shown that it was feasible to take into account sensory feedback delays, as well as general perturbation force fields. This constitutes an important step towards more biologically realistic computational models of the motor behaviour during reaching movements.

7.1 Limitations

Even though the main contribution of this work was to incorporate some biological faithfulness into an existing model of online motor learning, there remains several major limitations to the current approach.

In our opinion, future work should definitely try to physiologically quantify the learning rate parameter γ used in our model. Indeed, as shown in our numerical simulations, this parameter has a great impact on the general look of the results since it controls the rate at which the central nervous system is able to learn a representation of an ongoing perturbation. Nevertheless, we still have no clue of

what is the biological value of γ.

Furthermore, we did not investigate the neuroanatomical foundations of online motor learning. Although we have highlighted the need to take this fast time scale into account, we did not try to locate the computational center of this process inside the cerebral cortex. Collaboration with neuroanatomists should provide a deeper appreciation of this topic.

Finally, a major constraint to the current approach lies in the choice of a linear modelling. Of course, non linear models of motor control and learning should bring important advances in our understanding of the motor function in the future. However, this also comes with many theoretical and practical challenges to be still addressed.

Bibliography

[1] Frederic Crevecoeur, Jean-Louis Thonanrd, and Philippe Lefevre. A sub-movement time scale of human motor adaptation. *BioRxiv*, page 269134, 2018.

[2] Frederic Crevecoeur and Isaac Kurtzer. Long-latency reflexes for inter-effector coordination reflect a continuous state feedback controller. *Journal of neurophysiology*, 120(5):2466–2483, 2018.

[3] Richard A Schmidt, Timothy D Lee, Carolee Winstein, Gabriele Wulf, and Howard N Zelaznik. *Motor control and learning: A behavioral emphasis*. Human kinetics, 2018.

[4] JA Scott Kelso. *Human motor behavior: An introduction*. Psychology Press, 2014.

[5] Emanuel Todorov and Michael I Jordan. Optimal feedback control as a theory of motor coordination. *Nature neuroscience*, 5(11):1226, 2002.

[6] Dan Liu and Emanuel Todorov. Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. *Journal of Neuroscience*, 27(35):9354–9368, 2007.

[7] Daniel M Wolpert and Mitsuo Kawato. Multiple paired forward and inverse models for motor control. *Neural networks*, 11(7-8):1317–1329, 1998.

[8] Stephen H Scott. Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews Neuroscience*, 5(7):532, 2004.

[9] Reza Shadmehr, Maurice A Smith, and John W Krakauer. Error correction, sensory prediction, and adaptation in motor control. *Annual review of neuroscience*, 33:89–108, 2010.

[10] Joseph Y Nashed, Frédéric Crevecoeur, and Stephen H Scott. Influence of the behavioral goal and environmental obstacles on rapid feedback responses. *Journal of neurophysiology*, 108(4):999–1009, 2012.

[11] Jun Izawa and Reza Shadmehr. On-line processing of uncertain information in visuomotor control. *Journal of Neuroscience*, 28(44):11360–11368, 2008.

[12] Jörn Diedrichsen. Optimal task-dependent changes of bimanual feedback control and adaptation. *Current Biology*, 17(19):1675–1679, 2007.

[13] Jörn Diedrichsen, Reza Shadmehr, and Richard B Ivry. The coordination of movement: optimal feedback control and beyond. *Trends in cognitive sciences*, 14(1):31–39, 2010.

[14] Emanuel Todorov. Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. *Neural computation*, 17(5):1084–1108, 2005.

[15] Mitsuo Kawato. Internal models for motor control and trajectory planning. *Current opinion in neurobiology*, 9(6):718–727, 1999.

[16] Kurt A Thoroughman and Reza Shadmehr. Learning of action through adaptive combination of motor primitives. *Nature*, 407(6805):742, 2000.

[17] Kurt A Thoroughman and Reza Shadmehr. Electromyographic correlates of learning an internal model of reaching movements. *Journal of Neuroscience*, 19(19):8573–8588, 1999.

[18] Jörn Diedrichsen, Olivier White, Darren Newman, and Níall Lally. Use-dependent and error-based learning of motor behaviors. *Journal of Neuroscience*, 30(15):5159–5166, 2010.

[19] John W Krakauer, Maria-Felice Ghilardi, and Claude Ghez. Independent learning of internal models for kinematic and dynamic control of reaching. *Nature neuroscience*, 2(11):1026, 1999.

[20] John W Krakauer, Zachary M Pine, Maria-Felice Ghilardi, and Claude Ghez. Learning of visuomotor transformations for vectorial planning of reaching trajectories. *Journal of Neuroscience*, 20(23):8916–8924, 2000.

[21] Eran Dayan and Leonardo G Cohen. Neuroplasticity subserving motor skill learning. *Neuron*, 72(3):443–454, 2011.

[22] Reza Shadmehr and Ferdinando A Mussa-Ivaldi. Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience*, 14(5):3208–3224, 1994.

[23] James R Lackner and Paul Dizio. Rapid adaptation to coriolis force perturbations of arm trajectory. *Journal of neurophysiology*, 72(1):299–313, 1994.

[24] John W Krakauer. Motor learning and consolidation: the case of visuomotor rotation. In *Progress in motor control*, pages 405–421. Springer, 2009.

[25] James R Lackner and Paul DiZio. Motor function in microgravity: movement in weightlessness. *Current Opinion in Neurobiology*, 6(6):744–750, 1996.

[26] Frédéric Crevecoeur, Joseph McIntyre, J-L Thonnard, and Philippe Lefèvre. Gravity-dependent estimates of object mass underlie the generation of motor commands for horizontal limb movements. *Journal of neurophysiology*, 112(2):384–392, 2014.

[27] Daniel M Wolpert, Jörn Diedrichsen, and J Randall Flanagan. Principles of sensorimotor learning. *Nature Reviews Neuroscience*, 12(12):739, 2011.

[28] Etienne Burdet, Rieko Osu, David W Franklin, Theodore E Milner, and Mitsuo Kawato. The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature*, 414(6862):446, 2001.

[29] Jun Izawa, Tushar Rane, Opher Donchin, and Reza Shadmehr. Motor adaptation as a process of reoptimization. *Journal of Neuroscience*, 28(11):2883–2891, 2008.

[30] Robert R. Bitmead, Michel Gevers, and Vincent Wertz. *Adaptive Optimal Control. The Thinking Man's GPC*. Prentice Hall International Series in Systems and Control Engineering, 1990.

[31] Frédéric Crevecoeur and Philippe Lefèvre. Lecture notes of the course lgbio2060 - modelling of biological systems, 2018.

[32] PM Mäkilä. Kalman filtering and linear quadratic gaussian control. *Lecture notes for course*, 7604120, 2004.

[33] Greg Welch, Gary Bishop, et al. An introduction to the kalman filter. 1995.

[34] Weiwei Li and Emanuel Todorov. Iterative linearization methods for approximately optimal control and estimation of non-linear stochastic system. *International Journal of Control*, 80(9):1439–1453, 2007.

[35] Stephen H Scott. The computational and neural basis of voluntary motor control and planning. *Trends in cognitive sciences*, 16(11):541–549, 2012.

[36] Tyler Cluff and Stephen H Scott. Rapid feedback responses correlate with reach adaptation and properties of novel upper limb loads. *Journal of Neuroscience*, 33(40):15903–15914, 2013.

[37] John M Hollerbach and Tamar Flash. Dynamic interactions between limb segments during planar arm movement. *Biological cybernetics*, 44(1):67–77, 1982.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain
Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/epl**