



Task representation in human motor control: sequential goals and receding horizon allow for concise modelling of motor adaptation

Étienne Moullet

► To cite this version:

Étienne Moullet. Task representation in human motor control: sequential goals and receding horizon allow for concise modelling of motor adaptation. Neuroscience. Sorbonne Université (France), 2022. English. NNT: . tel-03869527v1

HAL Id: tel-03869527

<https://hal.science/tel-03869527v1>

Submitted on 26 Oct 2022 (v1), last revised 24 Nov 2022 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

présentée pour obtenir

le titre de docteur délivré par Sorbonne Université

École doctorale: Science Mécanique, Acoustique, Électronique et Robotique de Paris

par

Etienne MOULLET

**Task representation in human motor control :
sequential goals and receding horizon allow for
concise modelling of motor adaptation**

Véronique PERDEREAU	Professeure, Sorbonne Université	Présidente du Jury
Bastien BERRET	Professeur, Université Paris-saclay	Rapporteur
Aymar DE RUGY	Directeur de recherche CNRS	Rapporteur
Amandine AFTALION	Directrice de recherche CNRS	Examinatrice
Pauline MAURICE	Chargée de recherche CNRS	Examinatrice
Emmanuel GUIGON	Chargé de Recherche CNRS	Co-directeur de thèse
Agnès ROBY-BRAMI	Directrice de recherche INSERM, émérite	Directrice de thèse

Institut des Systèmes Intelligents et de Robotique (ISIR)
Pyramide Tour 55, 4 place Jussieu
UMR CNRS 7222, Paris, France

Abstract

Human motor control is a complex process ruling over a vast variety of voluntary tasks (posture, reaching, writing, walking...) while presenting persistent characteristics (coordination, structure of variability, movement segmentation, speed/accuracy trade-off...). Several theories brought better understanding of the mechanisms underlying the richness of motor behavior, yet failing to provide a unified framework for the production of human movement. Such a general computational theory was formulated by E. Guigon and relies on three modelling principles: a universal optimal feedback control policy, control with a receding time horizon, and task representation by a series of via-points updated at fixed frequency. In this thesis, we show that contrary to their appearing constraining nature, the combination of these principles offers powerful predictions, providing insights on the functioning of motor control and its neural bases. First, under the light of experimental data, we propose a novel view on motor adaptation to dynamic perturbations, casting it at the task representation level (goal selection) rather than at the control level (action selection). Second, we show that the receding horizon fundamental principle extends the stationary property of optimal solutions and allows for a simple neural representation of the universal controller: a small neural network (a Multi Layer Perceptron with two hidden layers) trained on simulated data accounted for multiple properties of cortical motoneurons described in the literature. Additionally, we briefly explore within our framework the differences between force control and position control tasks as well as co-contraction and its temporal evolution during movements.

Keywords: Human, Motor Control, Adaptation, Neural Network, Task Representation, Cocontraction

Remerciements

Une thèse est une longue aventure, difficile mais aussi passionnante. Elle est aussi l'occasion de rencontres et un catalyseur pour les relations humaines préexistantes. En conséquence, la liste des personnes que je souhaite remercier est un peu longue...

Je tiens tout d'abord à remercier le jury d'avoir examiné ce travail de thèse. Les remarques sur le manuscrit et les échanges lors de la soutenance furent très enrichissants.

Mes encadrants de thèse ensuite.

Agnès, merci pour ta bienveillance, ta disponibilité. Merci pour ta sagesse.

Emmanuel, merci de m'avoir transmis une (petite) part de ton immense savoir. Merci pour ton exigence et ta rigueur scientifique, qui m'accompagneront pour la suite. Merci également pour ton humanité et l'aide considérable que tu m'as apporté au cours de la rédaction de ce manuscrit.

A Guillaume. Professeur Morel. Merci de m'avoir accordé ta confiance en m'accueillant à l'ISIR. Merci pour nos longues discussions autour de la recherche autour du handicap et celles à venir.

A Lucas. Nous n'avons passé que quelques mois dans le même bureau, mais nous avons tant partagé dans ce court laps de temps, et soudé un lien indéfectible. Je ne saurai que te remercier d'être toi, et te donne rendez-vous aussi souvent que possible pour les 50 prochaines années.

A Marc. Je ne croyais guère aux « bromances » et pourtant... Merci pour ta motivation permanente, ta curiosité insatiable, ta créativité sans borne et l'enthousiasme que tu injectes dans tes relations humaines et dans ton travail. J'ai hâte que nous collaborions à changer le monde. Autour d'une bière, évidemment.

A ceux qui ne sont plus des collègues depuis déjà bien longtemps, mes frères et sœurs d'armes : Eléonore, Mégane, Josh, Clémence, Jésus, Jimmy, Mario, Omar, Félix, Charlotte. Merci d'avoir été présents pour tous les moments d'enthousiasme et de découragements, de réflexion intense, de détente autour de mots-fléchés ou

de verres au Sciences. Merci d'avoir supporté mes retards (bien que souvent involontaires), mes questions à rallonge et mes blagues douteuses. Merci, aussi, pour l'aide matérielle dont j'ai souvent eu besoin (diplôme de découpe d'aliments au CROUS par validation des acquis pour tous !).

Eléonore, merci pour ton sourire perpétuel, ta candeur si rafraichissante, ton écoute, ta bienveillance et ta noblesse d'âme.

Mégane, merci pour la force tranquille que tu es, pour la malice qui se cache parfois derrière un petit air blasé. Merci pour ta rigueur horaire (#team11h30), merci pour nos discussions sérieuses et décalées.

Josh, merci pour les pôzs (hum), nos instants de déprime commune, ou au contraire de remotivation réciproque. Merci pour ta disponibilité et ton entrain pour m'aider. Merci aussi pour les moments d'échanges et d'ouverture culturelle. Merci, enfin, pour le bonhomme adorable et incroyable d'intelligence qui se cache derrière ta timidité.

Clémence. Merci en premier lieu pour Slipknot, symptômes de ton ouverture au monde et ton appétit insatiable de découverte. Merci pour les pôzs avec Josh, merci pour ton rire éclatant et entraînant. Merci pour la réparation de mon fauteuil et pour tes nobles intentions envers le monde du handicap.

A Jésus. Ooooh Jézoush. Merci pour le soleil que tu es, pour ta positivité en toutes circonstances. Merci pour ton incroyable sens scientifique et ta capacité à te projeter dans des domaines qui s'écartent du tien. Merci pour ton incroyable gentillesse et ton ouverture au monde.

A Jimmy. Merci pour les rôtis, pour les teqs pafs. Merci pour ton alternance de sérieux et de folie. Merci, Monsieur Jimmy.

A Mario. Merci pour les lasagnes, les tiramisus. Merci pour les leçons de français, pour ton franc parler et pour investissement sans faille dans la vie professionnelle et amicale.

Omar, merci de m'avoir démontré que l'on pouvait suivre ses rêves.

Félix, merci d'avoir pris un temps le relais en termes de blagues douteuses. Merci pour ta patate perpétuelle et la sensibilité qui se camoufle en dessous.

Charlotte. Merci pour ta bonne humeur systématique et ton entrain à tracter des luges.

Aux autres nouveaux doctorants, que je n'ai pas (encore) eu la chance de connaître, merci pour la bonne humeur et l'animation que vous avez su prolonger alors que la vague de doctorant précédente n'était plus tout à fait à même de le faire.

Aux permanents de l'équipe Agathe, Ludovic, Nathanaël, Fabien, Jérôme, Brahim, Marie-Aude, Wael, Philippe, merci pour votre bienveillance, vos conseils et l'ambiance d'entraide et de franche camaraderie qui règne dans l'équipe.

A Gaël, merci pour nos travaux pendant ton stage, qui n'ont certes pas donné lieu à publication, mais ont nourri de très intéressantes réflexions.

Professeur Denys. Merci pour votre bienveillance d'abord, puis nos discussions, initialement dans un contexte étonnant, qui ont mené à ce que je considère comme une forme de mentorat, sans lequel cette thèse n'aurait jamais commencé. Je souhaite de tout cœur que nous puissions collaborer dans le futur.

Professeur Mozer. Merci pour vos conseils avisés, qui m'ont mené à me lancer dans l'aventure doctorale puis, plus de 3 ans après à la poursuivre dans la recherche académique. J'espère sincèrement pouvoir vous présenter mes remerciements « en chair et en os » et, qui sait, avoir ensemble des discussions scientifiques qui pourraient mener à collaboration.

A Marion. Bientôt 30 ans d'amitié. Une constante, une présence rassurante et motivante. Merci d'être là. Toujours.

A mes Shifus : Doudou, Annabelle, Bastien, Brice, Hassan, Davide, Willou, Benou... Merci pour nos folies, les burgers, les bières, le soutien inconditionnel et les discussions parfois profondes. Je préfère ne pas compter les années dernières nous, mais j'ai hâte des prochaines !

A ma mémé Gisèle. Merci de m'avoir aidé à grandir et de continuer à le faire. Merci pour l'exemple de courage, d'abnégation que tu es. Merci pour les bugnes, les crêpes et pour le cocon plein de douceur que tu as créé pour tes petits enfants.

A mes grands parents qui ne sont plus là : Pépé Pierrot, Papy Georges, Mamy Henriette. Merci pour ces belles années d'enfance avec vous, qui résonnent encore en l'homme que je suis.

A mes oncles et tantes, Myriam, Eric, Jacky, Brigitte et Christine. Merci pour d'avoir été cette si belle extension familiale. Merci pour l'éclectisme de vos personnalités, histoires et sagesse de vie, qui m'ont fourni autant de briques pour bâtir l'homme que je suis aujourd'hui.

A mes cousins, Mathilde, Maxime, Thomas, Aurore. Merci pour la complicité, le soutien, les aventures. Merci pour votre présence, même lointaine, et pour les moments délicieux de nos retrouvailles.

A ma petite sœur de cœur, Victoria. Merci pour les longues soirées à refaire le monde, à parler à cœurs ouverts. Merci pour les sushis, les ageofs, les parties d'échecs. Merci pour la lumière de ton sourire et de ton âme.

A ma moitié, mon amoureuse. A Eva. Merci pour ton soutien indéfectible, d'avoir supporté les couchers tardifs et les réveils matinaux. Merci pour les fous rires, les accents bizarres, les sessions cuisine décadente. Merci de mettre ce doux fou artistique dans notre quotidien, qui le rend exceptionnel. Merci d'être l'étoile de ma vie.

A mon frère, Simon. Merci pour les centaines d'aventures étant petits, pour nos inventions farfelues, nos jeux dans les forêts, nos bastons. Merci pour l'exemple que tu as toujours été par ce mélange de compétition bienveillante et de paroles piquantes mais motivantes. Merci pour ces actes qui parlent plus que mille mots d'amour. Merci d'être là, toujours. Et enfin merci, avec Anaïs, d'avoir amené dans ma vie ce petit soleil d'Iris. Comptez sur moi pour être un tonton à la fois gâteau et instructeur, qui lui parlera peut-être un peu de contrôle moteur !

A mon infirmière, ma secrétaire, mon intendante, ma confidente, ma coiffeuse, ma jardinière. A mère. A mon coach, mon kiné, mon menuisier, mon technicien en tout, mon co-inventeur. A mon père. A mes parents, merci d'avoir été présent à mes côtés à tout instant. Merci de m'avoir aidé, sans faillir, à traverser les épreuves qui se sont présentées. Merci de m'avoir inculqué le courage, l'abnégation, la persévérance, la bienveillance. Merci de m'avoir accompagné, soutenu et encouragé dans tous mes projets fous. Un jour j'écirai, j'en fais le serment, l'aventure de nos vies, afin de clamer les héros que vous êtes à mes yeux.

Contents

Abstract	i
Acknowledgements	ii
List of Figures	vii
List of Tables	xi
1 Introduction and context	1
1.1 Introduction and context	1
1.1.1 What is motor control?	2
1.1.2 Theoretical frameworks	2
1.1.3 Neural bases of motor control	2
1.1.4 Our approach	3
1.2 Structure of this manuscript	3
2 Modelling human motor control	5
2.1 Introduction	6
2.2 Sensorimotor system	8
2.2.1 Central Nervous System	8
2.2.2 Spinal level	9
2.2.3 Neurons	9
2.2.4 Joints	10
2.2.5 Muscles and tendons	10
2.3 Challenging properties	11
2.3.1 Redundancy	11
2.3.2 Noises	12
2.3.3 Delays	13
2.3.4 Variability	14
2.3.5 Nonlinearity	15
2.3.6 Smoothness vs. segmentation	15

2.4	Modelling	17
2.4.1	State and dynamics	17
2.4.2	Observation	20
2.4.3	Noises	20
2.4.4	Delays	21
2.4.5	Overall process	21
2.5	Computational frameworks	22
2.5.1	Equilibrium point	23
2.5.2	Optimal control	25
2.5.2.1	Linear Quadratic Regulator	27
2.5.2.2	Linear Quadratic Gaussian controller	27
2.5.2.3	Stochastic optimal feedback control	29
2.5.2.4	iLQR and iLQG	30
2.5.2.5	Task representation	30
2.5.3	Achievements of the SOFC framework	31
2.5.4	Limitations of the SOFC framework	32
2.6	Optimal control with sequential terminal constraints and receding horizon	33
2.6.1	Modelling principles	33
2.6.1.1	Universal Control policy	34
2.6.1.2	Receding horizon	35
2.6.1.3	Sequential goals	36
2.6.2	Implementation	37
2.6.2.1	Task representation	38
2.6.2.2	Optimal computation	38
2.6.2.3	State estimation	41
2.7	The common content of motor signals	41
2.8	Discussion	47
2.8.1	Everything is "movement"	48
2.8.1.1	Relationship with postural models	48
2.8.2	Everything is task representation	49
2.8.3	Stationary control	51
3	Adaptation to motor perturbations as a redirection of goals	53
3.1	Introduction	53
3.2	Simulation methods	55
3.2.1	Computational modelling	55
3.2.2	Compensation model	56
3.2.3	Reoptimization model	56

3.2.4	Redirection model	59
3.2.5	Numerical solution	60
3.3	Experiment	60
3.3.1	Ethics statement	60
3.3.2	Participants	60
3.3.3	Apparatus	60
3.3.4	Experimental procedure	61
3.3.5	Data processing and analysis	62
3.4	Results	63
3.4.1	Predictions	63
3.4.2	Two representative participants	66
3.4.3	All participants	68
3.4.4	Redirection model	69
3.4.5	Parametric study of the reoptimization model	70
3.5	Discussion	72
4	Neural implementation of optimal control	77
4.1	Introduction	78
4.2	Materials and methods	79
4.2.1	Neural network model	79
4.2.2	Dataset and training	80
4.2.3	Simulation and data analysis	80
4.2.4	Parametric analysis	81
4.2.5	Dimensionality reduction	82
4.3	Results	82
4.3.1	Network training and performance	82
4.3.2	Directional tuning	85
4.3.3	Relationship between neurons and EMG	88
4.3.4	Influence of arm posture	89
4.3.5	Dimensionality reduction	89
4.4	Discussion	96
4.4.1	Disclaimer	96
4.4.2	Distributed representation of the universal controller	96
4.4.3	Relation to previous models	97
4.4.4	Role of the motor cortex	97
4.4.5	Dimensionality reduction	98
4.4.6	Limitations	99
4.4.7	Future developments	100
4.4.8	Extension	100

5	Discussion and Perspectives	101
5.1	Motor control involves a universal controller	101
5.2	Optimality and the content of action	102
5.3	Perspectives	103
	Appendices	105
A	Optimal Control	107
A.1	Analytic optimization for linear problems	107
A.1.1	Statement of the problem	107
A.1.2	Adjoint problem	108
A.1.3	Conditions for optimality	108
A.1.4	Application to a simple case	109
A.1.4.1	Dynamics	109
A.1.4.2	Optimization problem	110
A.1.5	Solving Euler-Lagrange equations	111
A.2	Stationarity with receding horizon	113
B	State Estimation	115
B.1	State estimation and motion	115
B.2	The extended Kalman filter	116
B.3	Discretized continuous formulation	118
B.4	Delays	118
C	Supplementary Materials to Chapter 3	121
D	Supplementary Materials to Chapter 4	133
D.1	Influence of network architecture and size	133
D.2	More on the performance of the network	133
E	Position vs. force control	141
E.1	Position control	142
E.1.1	Issue with receding horizon optimal control	142
E.1.2	Adapted task representation for receding horizon optimal control	145
E.1.2.1	Goal position adjustment	146
E.1.2.2	Goal activation/excitation adjustment	147
E.2	Force control	148
E.3	Modelling predictions: comparison	149
F	Cocontraction	153

F.1 Cocontraction during postural task	154
F.2 Cocontraction during movement	155

Bibliography	175
---------------------	------------

List of Figures

2.1	Hill-type model of muscle and tendons	11
2.2	Redundancies in a motor task	12
2.3	The multiple levels of occurrence of noises in the sensorimotor system	13
2.4	Illustration of the Uncontrolled Manifold in the task space	14
2.5	Impact of filtering level on velocity profiles	16
2.6	Segmentation of slow movement	16
2.7	Simplified model of an arm	18
2.8	Feedforward Control	22
2.9	Feedback Control	22
2.10	Equilibrium Point theory	24
2.11	Perturbed reaching of a line	25
2.12	Estimation/Control	28
2.13	Velocity scaling	33
2.14	Optimal Control with sequential terminal constraints and receding horizon	37
2.15	Task representation (1/2)	39
2.16	Task representation (2/2)	40
2.17	Acceleration and jerk profiles for diverse motor tasks	43
2.18	Frequency contents of experimental and synthetic data	45
2.19	Frequency content of data recorded with various sensors	46
2.20	Two trials with different velocity fluctuations	47
3.1	Description of the experiment	61
3.2	Model adjustment based on data of participant P7	64
3.3	Model predictions	67
3.4	Data of participant P7 (1/2)	68
3.5	Data of participant P7 (2/2)	69
3.6	Data of participant P5	70
3.7	All participants	71
3.8	Simulations of the redirection model	72
4.1	Neural network and training set	81
4.2	Schema of the principal component analysis	83
4.3	Training curves	83
4.4	Neural network simulations	84
4.5	First layer analysis	85

4.6	Muscles directional tuning	86
4.7	Distribution of neuron's preferred direction - First layer	87
4.8	Error in the reconstruction of movement direction	88
4.9	Correlation between neuronal activity in the first hidden layer and EMG	89
4.10	Matching peaks between neuronal activity in the first hidden layer and EMG	90
4.11	Neuron's preferred direction relatively to shoulder muscle	91
4.12	Neuron's preferred direction relatively to elbow muscle	92
4.13	Influence of arm posture	93
4.14	Neural trajectories of two first Principal Components in the first hidden layer	94
4.15	Shoulder vs elbow EMG for the movements of the DIRECTION experiment.	95
C.1	Predictions of the compensation model	121
C.2	Time at which the angle derivative of the before-effect trajectory becomes positive	122
C.3	Participants whose behavior is incompatible with the reoptimization model (1/3)	123
C.4	Participants whose behavior is incompatible with the reoptimization model (2/3)	124
C.5	Participants whose behavior is incompatible with the reoptimization model (3/3)	125
C.6	Participants whose behavior is partially compatible with the reoptimization model (1/2)	126
C.7	Participants whose behavior is partially compatible with the reoptimization model (2/2)	127
C.8	Two participants that failed to improve their behavior with training	128
C.9	Parametric study of the model: influence of feedback delay.	129
C.10	Parametric study of the model: influence of noise ratio	130
C.11	Parametric study of the model: influence of muscle gain ratio	131
C.12	Parametric study of the model: influence of boundary conditions	132
C.13	Power spectrum analysis	132
D.1	Influence of network architecture and size	134
D.2	Performances of the main neural network (two hidden layers - 64/268 neurons)	135
D.3	Distribution of neuron's preferred direction - Second layer	136
D.4	Performances of a smaller two hidden layers neural network (64/128 neurons)	137
D.5	Performances of a single hidden layer neural network (1024 neurons)	138
D.6	Temporal activation of the neural network	139
D.7	Neural trajectories of two first Principal Components in the second hidden layer	140

E.3	Optimal control facing constant force	145
E.4	Control as a function of position	146
E.5	Adapted task representation for receding horizon optimal control . .	147
E.6	Position control with goal activation/excitation adjustment	148
E.7	Optimal control producing force	149
E.8	Force control with goal activation/excitation adjustment	150
E.9	Position control vs force control	151
F.1	Linear mass actuated by two muscles (rectangles) and their tendons (springs).	153
F.2	Reaching task and without cocontraction instructions	157

List of Tables

2.1	Parameters of the simulated arm	20
3.1	Simulation and optimization parameters	59

Chapter 1

Introduction and context

Science is what we understand well
enough to explain to a computer
(Art is everything else we do)

Donald Knuth

What I cannot create, I do not
understand

Richard Feynman

1.1 Introduction and context

Action is the mean by which living beings interact with their environment. Diminution or loss in the capacity to produce actions ("movement disorders") are often observed following injuries, diseases or ageing, affect the normal performance of activities of daily living, and are a central issue in modern healthcare systems. Progresses in the understanding of how actions are produced and in the definition and diagnosis of pathological actions are mandatory to improve handling and rehabilitation of movement disorders. A proper understanding of how actions are produced faces difficult challenges. The first challenge is to properly define the problem of motor control. The second challenge is to choose a theoretical framework in which the problem and the very many observations on properties of action can be handled. The third challenge is to address the neural bases of motor control. One can recognize here the levels of analysis of [Marr \[1982\]](#).

1.1.1 What is motor control?

Motor control is a complex problem with multiple levels of redundancy (task-space, body-space, muscle-space, neural-space), nonlinearities, uncertainty, noise and time delays [Bernstein, 1967; Glencross, 1980; Franklin and Wolpert, 2011]. Despite this complexity, living beings produce a large repertoire of actions (walking, running, flying, reaching, grasping, speaking, singing, writing, drawing, looking, smiling, swimming, standing) with an apparent ease. The resulting actions are geared towards behavioral goals rather than towards the reproduction of stereotyped patterns [Bernstein, 1967; Todorov and Jordan, 2002]. They are flexible in time and space, and are effortlessly modified to take into account changes in the body and the environment [Shadmehr and Mussa-Ivaldi, 1994; Liu and Todorov, 2007; Nashed et al., 2012]. A fundamental point is that the complexity of motor control cannot be bypassed by considering so-called simple actions (e.g. single joint movements): multijoint movements are not scaled-up versions of single-joint movements due the presence of intersegmental dynamics.

1.1.2 Theoretical frameworks

Following the claims of Donald Knuth and Richard Feynman, motor control cannot be understood outside a consistent theoretical framework. On the one hand, reasoning on neurons, muscles, reflexes or brain circuits is clearly an inadequate method. On the other hand, drawing conclusions from observations of movements could give rise to speculations (e.g. power laws; Huh and Sejnowski 2015). Several frameworks coexist: the dynamical approach [Kelso, 1995; Warren, 2006], the equilibrium point theory [Feldman and Levin, 1995], the Passive Motion Paradigm [Mohan and Morasso, 2011], the Active Inference theory [Friston, 2011], the stochastic optimal feedback control approach [Todorov and Jordan, 2002]. This is not the place for a thorough presentation and discussion of these frameworks (for elements, see Mohan et al. 2019; Latash 2021; Guigon 2022 and related comments). Here the point is to insist on the necessity of reasoning on motor control within a theoretical framework to avoid the pitfalls of purely descriptive approaches.

1.1.3 Neural bases of motor control

A valuable outcome of motor control studies is to contribute to a computational neuroanatomy/neurology of movement [Shadmehr and Krakauer, 2008; Haar and Donchin, 2020; Parr et al., 2021], i.e. to identify where in the nervous system and possibly how critical operations of motor control are produced. Although we are still far from an accurate picture of the "motor brain", undeniable progress has been made to clarify the respective role of key structures such as the motor cortex, the basal ganglia and the cerebellum. Further insights could be gained by building neural network controllers [Lillicrap and Scott, 2013; Kalidindi et al., 2021].

1.1.4 Our approach

This work is an attempt to address the three above-mentioned challenges in a unified, parameter-free theoretical framework [Guigon, 2022]. This framework which inherits from the stochastic optimal feedback control framework of Todorov and Jordan [2002] claims that each and every motor task is produced by one and the same *universal* controller working on a specific task representation. It explains the production of discrete, continuous, rhythmic and temporally-constrained movements, and their parametric and statistical properties (scaling laws, power laws, speed/accuracy tradeoffs). On the one hand, we thoroughly explored two main novel issues. First, we asked whether adaptation to dynamic perturbations (force fields) can be explained in the proposed framework, i.e. whether adaptation can occur with changes at the task level, without changes at the control level. Second, leveraging on recent advances in deep learning methods [LeCun et al., 2015], we asked whether an artificial neural network can be trained to approximate the universal controller. We answered positively to this question and we explored the distributed representations of motor control processes in this network and compared them to the properties of motor cortical neurons. On the other hand, we only scratched the surface of two important topics: (1) the role of cocontraction; (2) the contrast between position control and force control.

1.2 Structure of this manuscript

This thesis manuscript is composed of 3 main chapters :

- **Chapitre 2** - We describe stochastic optimal control and summarise both its capabilities and restrictions. Then, we present optimal control with sequential terminal constraints and receding horizon : its founding modelling principles, its implementation particularities regarding task representation and a few of its predictions.
- **Chapitre 3** - We run an experiment of adaptation to a velocity-force field exerted by a robotic arm. We propose a novel view of motor adaptation through task representation, that account for some key kinematic properties of adapted movements.
- **Chapitre 4** - We show that the stationary properties of optimal control with receding horizon simplifies the necessary computations to the point of enabling a feasible neural implementation. We train a feedforward neural network on data simulated by our model and generate trajectories that account for observed tuning characteristics of cortical motoneurons described in the literature.

Chapter 2

Modelling human motor control

2.1	Introduction	6
2.2	Sensorimotor system	8
2.2.1	Central Nervous System	8
2.2.2	Spinal level	9
2.2.3	Neurons	9
2.2.4	Joints	10
2.2.5	Muscles and tendons	10
2.3	Challenging properties	11
2.3.1	Redundancy	11
2.3.2	Noises	12
2.3.3	Delays	13
2.3.4	Variability	14
2.3.5	Nonlinearity	15
2.3.6	Smoothness vs. segmentation	15
2.4	Modelling	17
2.4.1	State and dynamics	17
2.4.2	Observation	20
2.4.3	Noises	20
2.4.4	Delays	21
2.4.5	Overall process	21
2.5	Computational frameworks	22
2.5.1	Equilibrium point	23
2.5.2	Optimal control	25
2.5.3	Achievements of the SOFC framework	31
2.5.4	Limitations of the SOFC framework	32
2.6	Optimal control with sequential terminal constraints and receding horizon	33

2.6.1	Modelling principles	33
2.6.2	Implementation	37
2.7	The common content of motor signals	41
2.8	Discussion	47
2.8.1	Everything is "movement"	48
2.8.2	Everything is task representation	49
2.8.3	Stationary control	51

2.1 Introduction

Human motor control is one of the most introspective endeavour humans can delve into [Bernstein, 1967]. It may be summarized in a general question: How does our brain controls our muscles in order to perform movements? From this apparently simple question stems a large amount of fundamental questions about our very body and mind functioning: How are we able to repeatedly perform a motor task without doing twice the same movement? How can we resist to external and unexpected perturbation? How do we maintain abilities while submitted to fatigue, injuries, or aging? How are we able to switch from one motor task to the other, or even perform several tasks at once? What is motor expertise and how is it acquired? All these questions (and many others) are at the crossroads of biomechanics, neurophysiology and neurosciences.

Our body is in perpetual interaction with its environment to either induce motion or maintain a given position or posture. First and foremost, as Newton's fundamental laws of dynamics state, movements are direct consequences of actions - torques or forces. In our own referential, a first fundamental step is to understand the kinematics of our own body. Not only is our skeleton composed of 206 bones, but they mainly articulate with each other in very complex joints involving cartilaginous contacts, ligaments, and tendons, far from the commonly human-engineered prismatic or revolute joints. These joints are actuated by 639 muscles (acting through their tendons), themselves composed of a variety of muscle fibers. These muscle fibers are innervated by alpha moto-neurons (and afferent pathways) emerging from the spinal cord, itself connected to the brain. Even if extremely simplified, this description already allows to glimpse the complexity of the system at hands while trying to understand human movements.

To face this challenge, scientists have adopted a constructivist methodology: from behavioral observations, make assumptions on a given functionality or mechanism, "implement them into a model", make predictions, confront them to additional observations and iterate over the multiple facets of human movement. Two crucial words were employed here: **observation** and **model**. Studying pathologies allows to attribute functionalities to anatomical structures by reverse reasoning: if several people presenting the same damaged part of the brain may not perform a given task, it can be speculated that the part in question plays a crucial role in this task for healthy humans (*e.g.* deafferentation, Parkinson's disease..) [Shadmehr and Krakauer, 2008; Haar and Donchin, 2020]. Additionally, experimental studies set up in precise conditions can bring to light meaningful properties of human movement by fine analysis of electromyograms (EMG), joint or end-effector trajectories or any relevant body variable.

A model often takes the form of a set of equations and parameters describing a given phenomenon and its evolution through time by its *dynamics*. It does not have to be perfect for it to be useful: newtonian theory of gravity is not exact but gives satisfying predictions for a wide range of non extreme gravitational phenomenon. Similarly, approximations may be accepted while modelling parts of the system that constitute the human body and the Central Nervous System (CNS) to accommodate, for example, for computational or theoretical capacities. Moreover, if the studied motor behavior involves only parts of the body it may be relevant to ignore the rest of the body: for upper limbs movements, for instance, we may consider the trunk to be an inertial reference frame and focus on modelling the arm.

Once a model of the system of interest is chosen rises the core problematic of human motor control: designing a schema describing how an intent leads to a movement. Depending on the theory adopted to study motor control, such a schema may take the form of a *physical law* (*e.g.* in the equilibrium point hypothesis) or of control policy (in control theory). In its general conception, a control policy may be seen as an algorithm producing a sequence of control signals that, when transmitted to the muscles will produce a movement achieving a given motor task. It must be noted, though, that such a computational representation is merely an attempt to replicate in simulations the outcome of the neurophysiological process occurring in the brain. Moreover, although an accurate enough model is necessary for the scientist to perform simulations, its knowledge by the CNS is not *a priori* a necessary condition. For instance one can perfectly drive a car without knowing precisely the thermochemistry of gas combustion, the mechanics of gear transmission or the surface properties of their tires.

If a set of model and control policy fails to account for certain characteristics of a motor behavior, two possibilities arise: either the model is at fault or the control policy is. Should the model be changed to be more realistic or the control policy modified it is of uttermost importance to verify that the updated model is still able to account for previously demonstrated results. In this regard, an accumulation

of *ad-hoc* models and policy have no explanatory power, as their predictions are bounded to the specific tasks or conditions for which they were built. Throughout the years, many theories have proposed their own control policies in attempt to account for experimental data, but few managed to be consistent with the broad range of motor tasks and their emerging decisive characteristics.

In this chapter, we first provide a brief description of the different body components playing a part in motor control [Kandel et al., 2013], and discuss on how they should be integrated in the modeling conditioning the definition of a control policy. Then, we present the successive control framework milestones that led to the formulation of the control theory adopted in this thesis: Optimal feedback control with receding horizon.

2.2 Sensorimotor system

2.2.1 Central Nervous System

The central nervous system is composed of the brain and the spinal cord. The brain itself is composed of several structures organized on a hierarchical basis. The brain stem (medulla, pons and midbrain) is situated above the spinal cord with cranial sensory and motor nerves. The cerebellum is connected backward to the brain stem. The forebrain includes medial structures (thalamus) and the cerebral hemispheres. The cortical neurons are situated on the surface of the brain (gray matter) interconnected with short- or long-range fiber tracts (white matter). The basal ganglia are groups of deep gray matter structures tightly interconnected and with the cerebral cortex. The cerebral cortex is divided in lobes. The occipital lobe includes the primary and secondary visual areas and the temporal lobe the auditory areas. The parietal lobe includes the primary somatosensory area and secondary integrative areas integrating vision and proprioception. The primary motor cortex area is situated in the posterior part of the frontal lobe, close to the parietal somatosensory area. The premotor area and supplementary motor areas lie forward to the primary motor area. The prefrontal areas are more related to planning and cognition. The brain treats sensory information coming up from the afferent pathways of the spinal cord (+ optical nerve) to generate motor commands in order to perform a motor task. These inputs are in turn sent to the efferent pathways of the spinal cords which distribute them to the muscles through the nerves. The planification of motor control involve large cortical networks including the parietal associative areas and multiple frontal areas [Battaglia-Mayer and Caminiti, 2019]. The supraspinal descending command is mainly ensured by the pyramidal tract issued from the primary motor area [Chouinard and Paus, 2006]. Motor execution also involves regulatory loops involving the cerebellum and basal ganglia working in parallel. The cerebellar system contribute to timing and sensorimotor coordination and learning [Manto et al., 2012]. The basal ganglia

contribute to action selection, inhibit unwanted motor responses and is involved in skill learning [Turner and Desmurget, 2010].

2.2.2 Spinal level

At spinal level, alpha motoneurons are prolonged by axons that innervate a set of muscular fibers (motor unit) [Pierrot-Deseilligny and Burke, 2005]. There are two kinds of receptors (sensors) in the muscle: i) muscle spindles include an intrafusal muscular fiber (organized in parallel with the main extrafusal fibers) and Ia sensory endings (connected to large and fast afferent sensory fibers) and slower type II fibers. Spindles are sensitive to stretch and stretch velocity of the muscle. Intrafusal muscular fibers are excited by gamma motoneurons which regulate the sensitivity of the spindles. Alpha-gamma coactivation during movement avoids the silencing of Ia afferents during muscular contraction; ii) the Golgi tendon organ, are sensors organized in series with the muscular fibers, connected with fast Ib afferent fibers, they are sensitive to the force of the contraction. Ia and Ib afferents are connected to the spinal segment of the corresponding muscle and regulate the reflex activity of the motoneurons via segmental interneurons (e.g. reciprocal inhibition due to Ia afferents, Ib inhibition, Renshaw inhibition). Slower fibers (spindle type II and cutaneous afferent fibers) project to several spinal segments, also connected via propriospinal interneurons. Spinal segmental and intersegmental reflexes and central pattern generators which contribute to the spinal control of movement are under the control of multiple descending spinal pathways [Weiler et al., 2021].

2.2.3 Neurons

Neurons are single cells of very peculiar form: far from the common representation of an ellipsoidal membrane containing a nucleus surrounded by cytoplasm, neurons have axons that can go through the whole body and measure up to 2 meters long. When a neuron discharges, an action potential propagates along its axons until it synapses. Here the electric potential induces the emission of a neurotransmitter (either excitatory or inhibitory) that crosses the synaptic cleft up to the dendrite of the connected neuron, creating a post-synaptic potential. Once the latter surpasses a given threshold, the second neuron in turn discharges along all its axons. Motoneurons innervate a set of muscle fibers: extrafusal muscle fibers for *alpha-motoneurons*, intrafusal muscle fibers of muscle spindles for *beta-motoneurons* and intrafusal muscle fibers found within the muscle spindle for *gamma-motoneurons*.

2.2.4 Joints

Joints are at the conjunction of bones. Their mechanics depend not only on the shape of the bones involved but on the ligaments restricting their relative movements and the muscles actuating it. Their kinematics may thus be rich and complex. For instance, movements of the single shoulder joint often summarized by 3 degrees of freedom (DoF) ball joint (flexion/extension, abduction/adduction, axial rotation) also includes 6 extra DoF due to clavicle and scapula (sterno-clavicular and acromioclavicular joints) but the mobility of the scapula within the shoulder complex is limited since the kinematic chain is closed due to the sliding of the scapula on the thorax (scapulo-thoracic joint).

2.2.5 Muscles and tendons

A muscle is made of muscle fibers and is innervated by motoneurons. A motor unit corresponds a motoneuron and the group of muscle fibers it innervates. Several fiber types classifications coexist but all concord in stating that motor fiber properties range from fast to slow, weak to strong and fatiguable to fatigue-resistant [Burke, 2011; Scott et al., 2001b]. Moreover, the size of a motoneuron is proportional to the size and number of its fibers and the excitability of a motoneuron is inversely proportional to its size [Henneman et al., 1965]. These properties give rise to the Henneman's size principle [Henneman et al., 1974]: motor units are recruited according to their size - smaller, weaker motor units are recruited first. Additionally, muscle force can be modulated by changing the motoneurons' firing rate. This gives two means of force modulation: an orderly change in the number of motor-units recruited and changes in the frequency of action potentials of the already recruited motor units

Taking a step back to a more macroscopic, empiric view, muscle force generation have been shown to follow a highly complex non-linear dynamic, depending on excitation, muscle length and velocity, and maximum force. It was first formalized in the form of the famous Hill equation [Hill, 1953]:

$$(F + a)v = b(F - F_0) \quad (2.1)$$

where F is the load in the muscle, v is the velocity of contraction, F_0 is the maximum isometric load generated in the muscle and a and b are constant parameters. A commonly used representation of force generation is the Hill-type muscle, describing it as a simplified mechanical chain (see Fig. 2.1): muscle fibers are represented as force generators acting in parallel to a damping element and a spring-like element. Conversely, tendons transmitting muscle force to their insertions exhibit a force-strain relationship subdivided in non-linear and linear phases. Yet, they are often modeled as spring-like elements (K serial) connected in series to the muscle fibers. Finally, the mentioned insertions are often spatially distributed

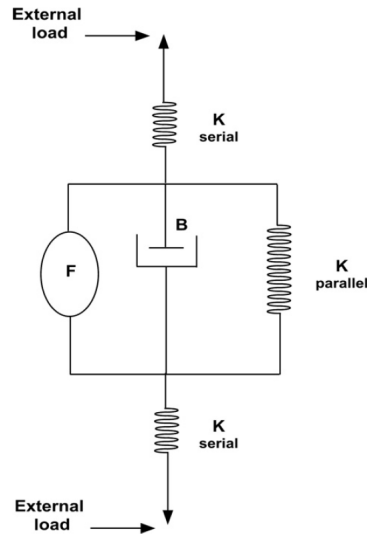


FIGURE 2.1: Hill-type model of muscle and tendons. Muscle fibers are represented as force generators (F) acting in parallel to a damping element (B) and a spring-like element (K parallel). Tendons can be modeled as spring-like elements (K serial) connected in series to the muscle fibers.

one the bones, and present a pennation angle, making the characterization of the lever effects on the actuated joint hard to describe analytically.

2.3 Challenging properties

2.3.1 Redundancy

Having briefly listed the principal components of the sensory-motor system, one of its key characteristics emerges: *redundancy* or *abundance* [Bernstein, 1967; Gera et al., 2010; Latash, 2012]. These terms summarize the fact that there are always more variables potentially contributing to a motor tasks than the constraints defining it.

Let us glimpse at the different levels of redundancy by considering one of the most common tasks: reaching a point with our hand.

The position in space of the hand is described by its three Cartesian coordinates x , y , and z (setting aside its orientation, that may be described by its Euler angles). One can easily observe that there are an infinity of body joint configurations corresponding to a single hand position (Fig. 2.2A): grab a door handle and move your body. One can bend the knees, extend the elbow, elevate the shoulder, lean forward, the many joints of the body will find ways to keeping his hand in the same position.

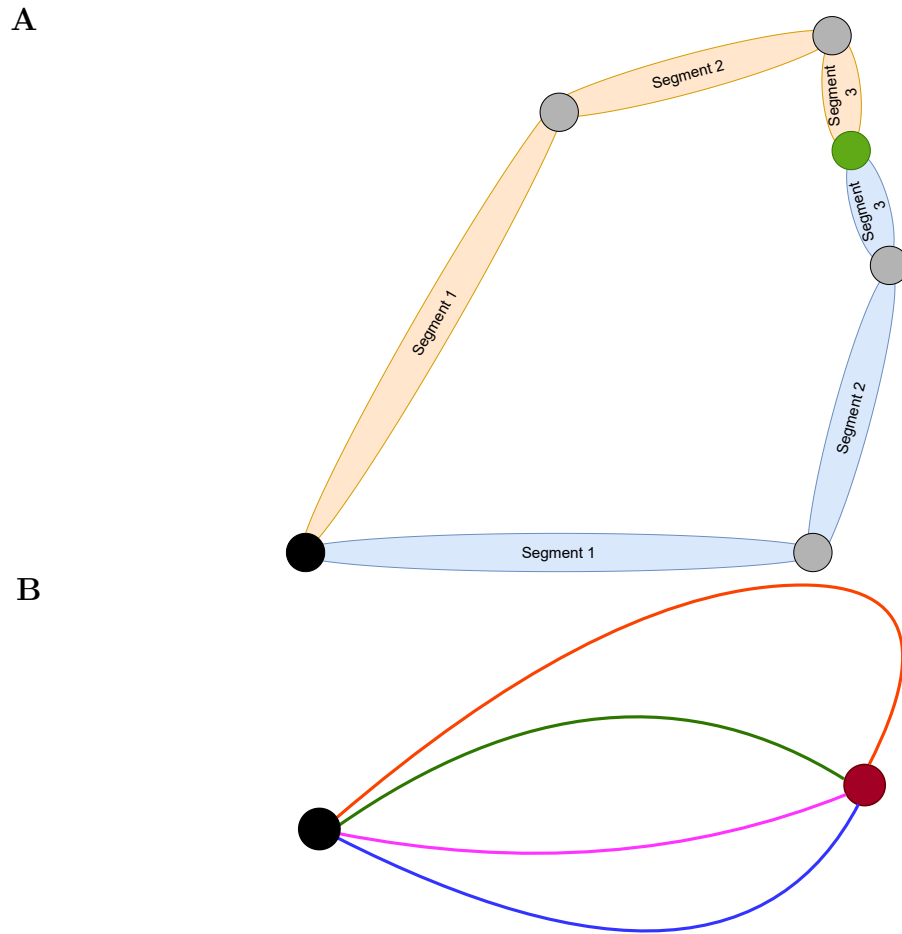


FIGURE 2.2: **Redundancies in a motor task.** **A.** Geometrical redundancies: a limb composed of three segments can adopt multiple joint configurations (light brown and light blue) to place its end effector at a given cartesian position (green circle). **B.** A reaching task (from black circle to red circle) may be achieved following several paths (blue, violet, green and orange lines).

Then, to go from a given point to another may be done through infinite number of hand paths (Fig. 2.2B), that can be followed at an infinity of paces. The corresponding movements are the consequences of torques generated at each joint by muscles contractions. Joints are often crossed by multiple muscles: elbow joint is for instance actuated by three flexor muscles (biceps, brachialis, and brachioradialis) and three extensor muscles (the heads of the triceps). In turn, each muscle consists in hundreds of motor units whose generated force depends on their firing rate.

2.3.2 Noises

Human perception and motor actions are not perfect. Both sensory feedback and motor commands are corrupted by noises, building up from the sensors themselves

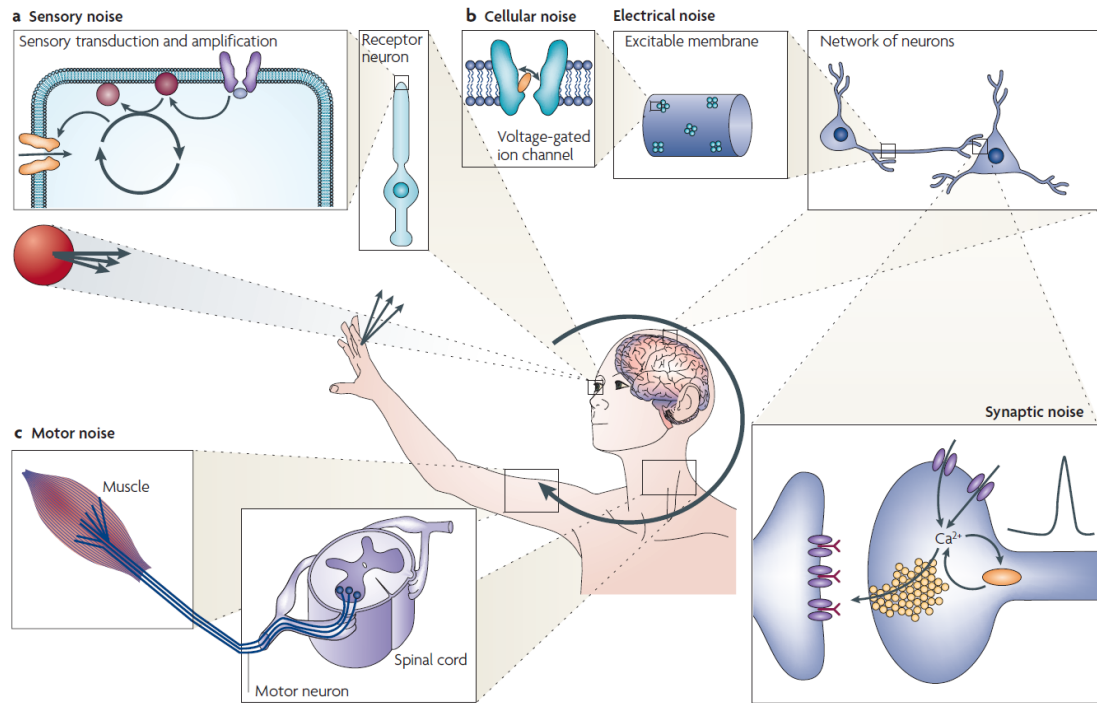


FIGURE 2.3: **The multiple levels of occurrence of noises in the sensorimotor system.** Reproduced from [Faisal et al. \[2008\]](#).

(spindles, GTO, skin stretches, the eye, ...), to the transmission and data processing [[Faisal et al., 2008](#)] (see Fig 2.3). Sensory noise degrades the observation of the state of the system at all scales and as well as of the environment state, and motor noise induce discrepancy between intended and effective actions. Additionally, motor noise is signal-dependent: its variance increases with mean signal magnitude [[Harris and Wolpert, 1998](#)].

2.3.3 Delays

Similarly to the noises plaguing its perception and actions, the sensorimotor system has to deal with delays at both ends of the control process [[Franklin and Wolpert, 2011](#)]: afferent sensory information is received with delays (~ 100 ms) and efferent control signals take time to be transmitted (~ 10 -40 ms) and to consequently induce force increase in muscles (~ 25 ms). Additionally, the exact value of these delays depend on the sensory modality and muscles involved. Moreover, the different existing reflexes (stretch reflex, spinal reflexes, long-loop reflexes, ...) have their own time constants, rooting the sensorimotor system in plural pasts.

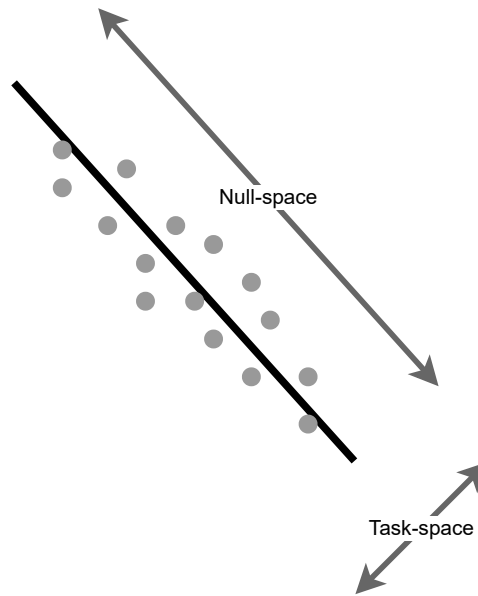


FIGURE 2.4: **Illustration of the Uncontrolled Manifold in the task space.** For a task consisting in pointing to a line (black line), the variance across trials of the final points (grey dots) is widely spread on the null-space (along the line) and limited in the task space (perpendicularly to the line).

2.3.4 Variability

All scales of redundancy may not be of equal importance but coupled with the presence of noises in the sensorimotor system they give rise to an simple yet astonishing property of human movement summarized by [Bernstein \[1967\]](#) as "*repetition without repetition*". This expression underlines the fact that while repeatedly and successfully performing a given motor task, humans will never produce twice the same exact trajectory.

Interestingly, the variance across trials of the endpoint trajectory is of lesser magnitude than the one affecting the evolution of the angles of engaged joint. Additionally, skilled motor behavior is not only able to manage all available degrees of freedom to perform a task, it actually exploits them to do so, justifying the use of term *abundance* [[Latash, 2012](#)]. This phenomenon is in particular observed in motor experts: mastering a task does not consist in reducing overall variance across trials but rather in shaping their variability to improve task performance [[Nisky et al., 2014](#)]. The Uncontrolled Manifold (UCM) describes how the *variability* inherent to irrepressible noises (in action or perception) is spread across a null-space, on which variance does not affect performance [[Scholz and Schöner, 1999](#)] (see Fig. 2.4). This method, although giving great insights on the structure of variability and synergies lacks an explanatory and predictive power. This was solved by the introduction of the *minimum intervention principle* by [Todorov and Jordan \[2002\]](#), which will be further described in section 2.5.2.3.

2.3.5 Nonlinearity

In linear systems the addition of two inputs will result in the addition of their corresponding individual outputs. This makes them easy to predict, and thus to control. The human sensorimotor system, unfortunately, doesn't fall in this category: the kinematic and inertial dynamic of the multiple and complex joints of the body, the relationship between muscle force and velocity, the activation of motor units by descending commands all add up to a highly non-linear system [Zajac and Gordon, 1989].

2.3.6 Smoothness vs. segmentation

In the laboratory or hospital service, correlates of action can be measured as time series of data at multiple levels in neural, bodily and environmental space. Processing and analysis of these data lead to qualitative and quantitative descriptions from which interpretations are drawn. As an example, consider the seminal study of Flash and Hogan [1985] on voluntary planar unconstrained movements of the upper limb. The movements were recorded with precision potentiometers, digitized at a rate of 100 Hz, and the digitized signals were low-pass filtered at 5.2 Hz. The authors observed that the movements had bell-shaped velocity profiles that were well described as the smoothest possible profiles (in the mathematical sense of involving minimum changes in acceleration or minimum jerk; jerk is the derivative of acceleration). Over the years, the minimum jerk model has become a central model in the field of motor control and robotics, and the idea of smoothness, although in different flavors, has played a cardinal role in many models of motor control [Uno et al., 1989; Harris and Wolpert, 1998; Todorov and Jordan, 2002].

Yet, as illustrated in Fig. 2.6 (reaching movement; data from van der Kooij et al. [2015]), a different conclusion could have been drawn from the same data depending on the actual level of “smoothing” of the data. Except the unfiltered velocity profile which probably contains artifacts (Fig. 2.5, black), there is no simple way to say which of the more irregular (less filtered) profile (red), the smoothest (more filtered) profile (purple) or other profiles (green, blue) best represents the “true” velocity. This trivial observation has far reaching implications all the more so that pathological movements are often characterized by their lack of smoothness [Cirstea and Levin, 2000; Rohrer et al., 2002].

Moreover, slower movements exhibit several velocity peaks and consequently numerous acceleration and jerk peaks. Guigon et al. [2019] showed that slow movements were in fact *segmented*: when instructed to draw a line at constant - and limited - speed, participants produced fluctuating velocity profiles (see Fig. 2.6). Moreover, the frequency of velocity peaks - as well as acceleration and jerk peaks - was constant, irrespective of the overall movement speed.

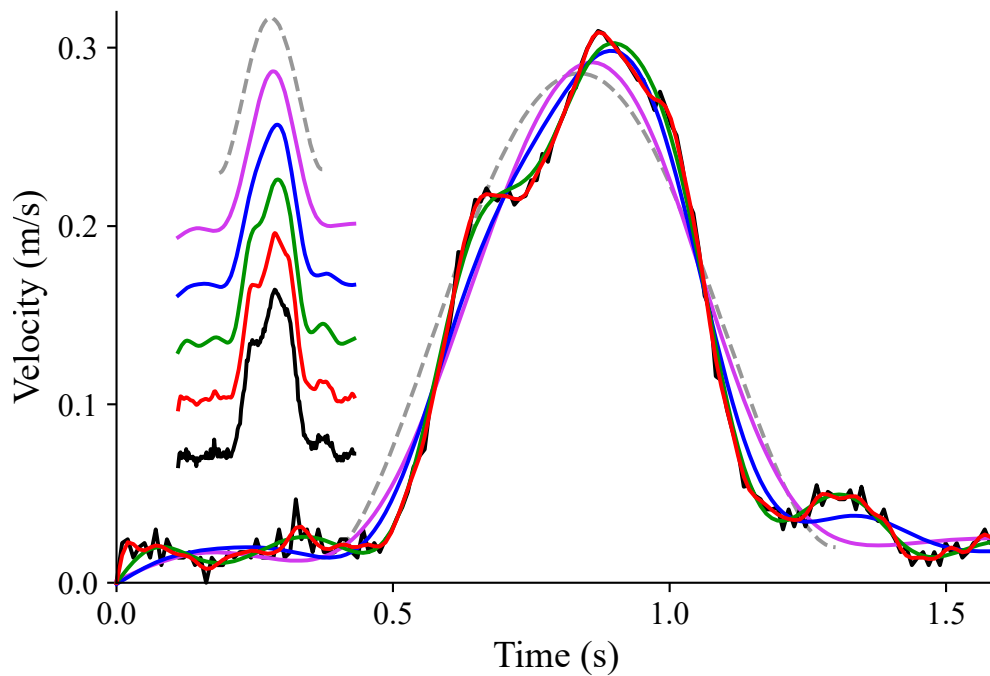


FIGURE 2.5: **Impact of filtering level on velocity profiles.** Velocity profiles of a single trial of arm displacement from van der Kooij et al. [2015] corresponding to different levels of low-pass filtering: (black) unfiltered data; (red) filtered at 20 Hz; (green) filtered at 5 Hz; (blue) filtered at 3 Hz; (purple) filtered at 2 Hz. The dotted gray line is a minimum jerk velocity profile adjusted by hand to the purple profile. Filter: 4th order Butterworth. The velocity profiles are reproduced in inset for legibility. van der Kooij et al. [2015]: <https://doi.org/10.5061/dryad.sm925>

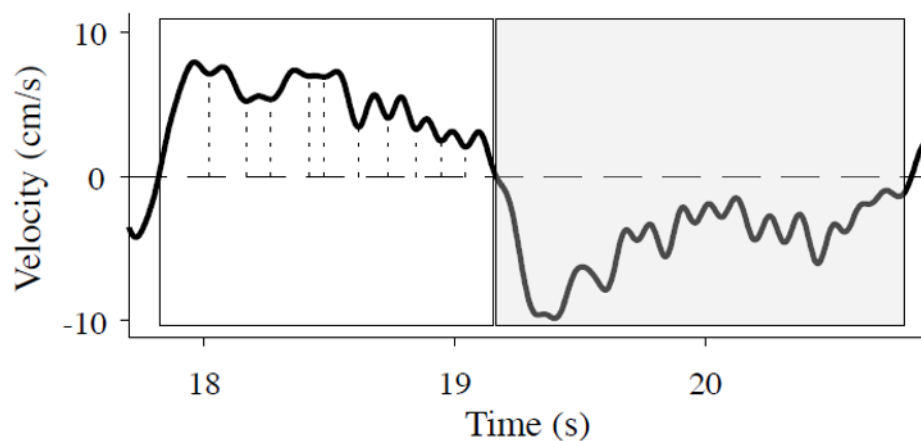


FIGURE 2.6: **Segmentation of slow movement.** Velocity profile for a single back and forth slow movement. From Guigon et al. [2019].

2.4 Modelling

The different features and properties of the human sensorimotor system converge to a tautological statement: human motor control is an extremely complex process to study and model. Yet, it would be vain to try to build an overarching model: for instance, delving deep into the propagation of action potential along the axon of a neuron or the formation of cross-bridges between actin and myosin would not be relevant to the comprehension of motor coordination. Conversely, modelling the hand by a linearly actuated punctual mass would not bring much insight to the same endeavour. Moreover, the level of details of a model needs to be put in perspective with the computational burden of corresponding simulations.

In the end, modelling the dynamics of human movement needs to find a compromise between oversimplification and overcomplexification. The approach followed in this thesis consists in restricting realism to the following model of the movement of a human hand. Our principle is to investigate to what extent complex motor behavior can stem from simple and synthetic modelling concepts.

2.4.1 State and dynamics

A most general representation of the dynamics of a system described by a state $\mathbf{x}(t)$ is a vectorial differential equation

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t), \quad (2.2)$$

where $\mathbf{u}(t)$ is a control signal. Here, we make an assumption of *stationarity* upon the dynamic:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \quad (2.3)$$

Qualitatively, it ensures repeatability: setting aside stochastic effects, a movement (or simulation) starting from the same initial state and performed with a given control sequence will induce the same state evolution regardless of the time at which it is accomplished. This assumption incarnates the choice to not take into account evolving phenomena such as muscle fatigue, injuries or aging.

Additionally, we make the following modelling choices:

- **Geometry:** the trunk is considered as the inertial frame. The arm is composed of 2 segments representing the upper and fore-arm. The former is linked to the trunk by a simple rotational joint - the shoulder - and is in turn attached to the forearm by a joint of same type figuring the elbow. The hand is then assimilated to the tip of the forearm. The wrist joint is neglected for its action is mostly restricted to maintaining a given posture and grip for the tasks considered in this work which are almost exclusively

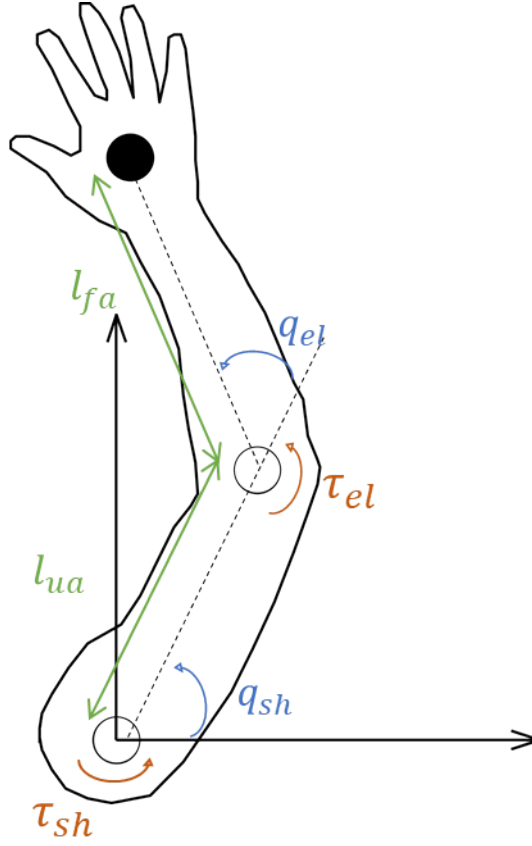


FIGURE 2.7: **Simplified model of an arm.** All geometrical and inertial parameters are given in Tab. 2.1

reaching tasks. Therefore, our model, shown in Fig. 2.7, is described by the shoulder and elbow angles:

$$\mathbf{q} = \begin{bmatrix} q_{sh} \\ q_{el} \end{bmatrix} \quad (2.4)$$

following a dynamic of the form:

$$\ddot{\mathbf{q}} = \mathbf{M}(\mathbf{q})^{-1} (\boldsymbol{\tau}_u + \boldsymbol{\tau}_e - \mathbf{c}(\mathbf{q}, \dot{\mathbf{q}})) \quad (2.5)$$

where $\mathbf{M}(\mathbf{q})$ is the inertia matrix, $\mathbf{c}(\mathbf{q}, \dot{\mathbf{q}})$ the vector of velocity-dependent torques, $\boldsymbol{\tau}_u$ the control torque produced by actuators and $\boldsymbol{\tau}_e$ the torque due to external forces applied on the arm. We define

$$\mathbf{M}(\mathbf{q}) = \begin{bmatrix} d_1 + 2d_2 \cos q_{el} & d_3 + 2d_2 \cos q_{el} \\ d_3 + d_2 \cos q_{el} & d_3 \end{bmatrix} \quad (2.6)$$

and

$$\mathbf{c}(\mathbf{q}, \dot{\mathbf{q}}) = d_2 \begin{bmatrix} 2\dot{q}_{sh}\dot{q}_{el} + \dot{q}_{el}^2 \\ -\dot{q}_{sh}^2 \end{bmatrix} \sin q_{el} \quad (2.7)$$

where m_{ua} and m_{fa} are the link masses (upper-arm and forearm), l_{ua} and l_{fa} the link lengths, I_{ua}^{sh} and I_{fa}^{el} their moments of inertia expressed at their

proximal joint, s_{sh} and s_{el} the distances from their center of mass to their proximal joint, $d_1 = I_{ua}^{sh} + I_{fa}^{el} + m_{fa}l_{sh}^2$, $d_2 = m_{fa}l_{sh}s_{el}$ and $d_3 = I_{fa}^{el}$.

In this work, the external torque is either null or applied to the tip of the forearm and is expressed as:

$$\boldsymbol{\tau}_e = \boldsymbol{\tau}_e^\phi = \mathbf{J}(\mathbf{q})^T \mathbf{f}_{ext} \quad (2.8)$$

where $\mathbf{J}(\mathbf{q})$ is the Jacobian matrix of kinematics

$$\mathbf{J}(\mathbf{q}) = \begin{bmatrix} -l_{ua} \sin q_{sh} - l_{fa} \sin (q_{sh} + q_{el}) & -l_{fa} \sin (q_{sh} + q_{el}) \\ l_{ua} \cos q_{sh} + l_{fa} \cos (q_{sh} + q_{el}) & l_{fa} \cos (q_{sh} + q_{el}) \end{bmatrix} \quad (2.9)$$

- **Actuation:** avoiding the insertion and dynamic complexities of muscle insertions in section 2.2.5, we chose actuation to be done by torque generators at each joint $j \in \{sh, el\}$. But as a muscle and its innervating motoneurons can be globally modelled as a second-order low-pass filter between the control input and the force output [van der Helm and Rozendaal, 2000], we grant them the corresponding state

$$\mathbf{x}_j = \begin{bmatrix} \alpha_j \\ \varepsilon_j \end{bmatrix} \quad (2.10)$$

and dynamics

$$\begin{cases} \dot{\alpha}_j = \frac{\varepsilon_j - \tau_j}{\nu} \\ \dot{\varepsilon}_j = \frac{u_j - \varepsilon_j}{\nu} \end{cases} \quad (2.11)$$

where α_j muscle activation, ε_j muscle excitation, u_j input control and ν is a time constant. We connect actuation dynamics to inertial dynamics by defining joint torques from actuator activations:

$$\tau_j = g_j \alpha_j \quad (2.12)$$

We write the whole state of our system as $\mathbf{x} = \begin{bmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \\ \mathbf{x}_{sh} \\ \mathbf{x}_{el} \end{bmatrix}$ and the control

$\mathbf{u} = \begin{bmatrix} u_{sh} \\ u_{el} \end{bmatrix}$ in order to define its state dynamics in the form:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \quad (2.13)$$

The parameters used in the simulations throughout this work are given in Tab. 2.1.

Parameter	Value	Unit
m_{ua}	1.4	kg
m_{fa}	1.1	kg
l_{ua}	0.3	m
l_{fa}	0.33	m
s_{sh}	0.11	m
s_{sh}	0.16	m
I_{ua}^{sh}	0.025	kg.m ²
I_{fa}^{sh}	0.045	kg.m ²
ν	0.05	s
g_{sh}	2	-
g_{el}	1	-

TABLE 2.1: Parameters of the simulated arm

2.4.2 Observation

We consider perception to be direct:

$$\mathbf{y} = \mathbf{H}\mathbf{x} \quad (2.14)$$

where \mathbf{y} is the observed state and \mathbf{H} is the observation matrix.

2.4.3 Noises

Identifying where noises occur is a challenging question in itself: the imperfect observed position of the hand, for instance, is likely coming from of measurement errors coming from eyesight, proprioception as well as information transmission and treatment. A precise definition would require a thorough modelling of the sensorimotor system, which is out of range for the moment, and out of the scope of this thesis. Hence, we limit noise modelling to the following [Todorov and Jordan, 2002; Guigon et al., 2008]:

- we consider noises to be independent white Gaussian noise *i.e.* each noise has a normal distribution with zero mean;
- we consider noises to be either additive or proportional to the considered signal;
- noise variances are scaled across the coordinates of the state (noise over joint angles is lower than noise over joint velocity, ...).

2.4.4 Delays

Similarly to noises, delays appear at all stages of human motor control on diverse timescales. We decide to focus on observation delay, which is considered to be homogeneous across the coordinates of the state:

$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t - \Delta), \quad (2.15)$$

where Δ is the delay.

2.4.5 Overall process

All things considered, the motor process is described by

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) + \mathbf{n}_{dyn}(t) \\ \mathbf{n}_{dyn}(t) &= \boldsymbol{\omega}_x(t) + \boldsymbol{\epsilon}_x(t) \odot \mathbf{x}(t) + \mathbf{f}(t)(\mathbf{x}(t), \boldsymbol{\omega}_u(t) + \boldsymbol{\epsilon}_u(t) \odot \mathbf{u}(t)) \end{aligned} \quad (2.16)$$

where \odot denotes the element-wise multiplication operator. The observation process is:

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{H}\mathbf{x}(t - \Delta) + \mathbf{n}_{obs}(t - \Delta) \\ \mathbf{n}_{obs}(t) &= \boldsymbol{\xi}(t) + \mathbf{H}\boldsymbol{\kappa}(t) \odot \mathbf{x}(t) \end{aligned} \quad (2.17)$$

where:

- $\boldsymbol{\omega}(t) = \begin{bmatrix} \boldsymbol{\omega}_x(t) \\ \boldsymbol{\omega}_u(t) \end{bmatrix}$ is a $(8 + 2)$ -dimensional zero-mean Gaussian random vector with power spectral density matrix $\boldsymbol{\Omega}^\xi(t)$ representing signal independent motor noise (SINm),
- $\boldsymbol{\epsilon}(t) = \begin{bmatrix} \boldsymbol{\epsilon}_x(t) \\ \boldsymbol{\epsilon}_u(t) \end{bmatrix}$ is a $(8 + 2)$ -dimensional zero-mean Gaussian random vector with power spectral density matrix $\boldsymbol{\Omega}^\epsilon(t)$ representing signal dependent motor noise (SDNm),
- $\boldsymbol{\xi}(t)$ is a p -dimensional ($p \in \{4, 8\}$) zero-mean Gaussian random vector with power spectral density matrix $\boldsymbol{\Omega}^\omega(t)$ representing signal independent sensory noise (SINs),
- $\boldsymbol{\kappa}(t)$ is a 8-dimensional zero-mean Gaussian random vector with power spectral density matrix $\boldsymbol{\Omega}^\kappa(t)$ representing signal dependent sensory noise (SPo).

It should be noted that, in this thesis, noises will only be used as statistics to bootstrap the estimator (Kalman filter, see section 2.6.2.3 and B).

2.5 Computational frameworks

Human motor control modeling has revolved for a long time around the dichotomy between *feedforward* and *feedback* control. *Feedforward* control relies on an inverse model of the dynamics of the system it is designed to control in order to compute the commands necessary to reach a goal (Fig. 2.8). It allows for precise task definition, but fails if the modeled dynamic of the system is slightly different from the real one. On the other hand, a *feedback* controller builds its commands from the difference between a given reference and the feedback it gets from the controlled system (Fig. 2.9). A famous feedback controller in the robotic community is the *Proportional Integral Derivative* (PID) controller which is praised for its robustness - for any external perturbation of the system will be observed and incorporated in the commands. *Equilibrium point theory* produced interesting results by coupling a feedback gain to a virtual minimum jerk (optimal) trajectory, effectively using feedforward control in a preparation phase feedback control during the execution phase [Hogan, 1984b]. It is important to note that even if the reference trajectory is optimal in the sense that it minimizes the integral jerk, the injected controls (here, taking the form of torques) are not the product of an optimization process.

In the optimal control field, *Linear Quadratic Gaussian* (LQG) control figures as a both stable and flexible framework, fusing *feedforward* and *feedback* control into an optimal feedback controller, the *Linear Quadratic Regulator* (LQR) and

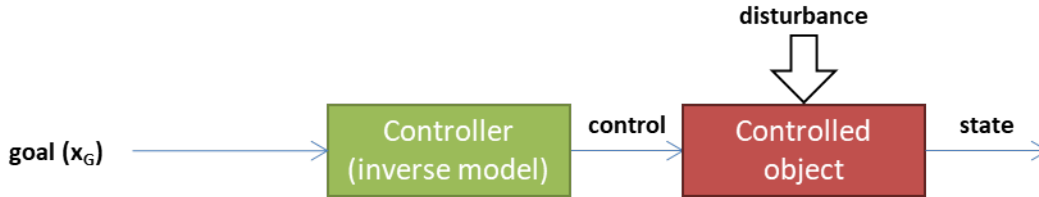


FIGURE 2.8: Feedforward Control: control is computed on the basis of an inverse model of the dynamics of the system, without knowing its actual evolution.

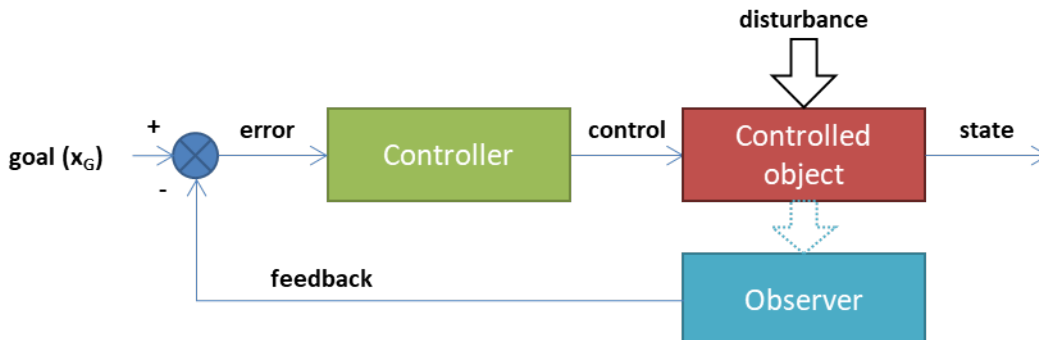


FIGURE 2.9: Feedback Control: the control is built on a measured error to be gradually decreased, without knowing the dynamics of the controlled system.

incorporating an optimal state estimator, the *Linear Quadratic Estimator*, commonly referred to as the Kalman filter (see Fig. 2.12). The LQG proved to be a founding brick in the edifice of optimal control: it allowed to implement complex tasks (reaching to points or lines, velocity constraints, ...) along with compromises between task completion and effort all while being exposed to observation and motor noises.

It was the *Stochastic Optimal Feedback Control* (SOFC), introduced by Todorov and Jordan [2002], that elegantly solved the apparent paradox of motor variability (see section 2.3.4) by extending the LQG. SOFC fuses together trajectory planning and execution into a single optimization process. Moreover, in attempting to account for the speed-accuracy trade-off ruling every precision task - the well-known *Fitts law* [Fitts, 1954] - it takes into account not only observation noise but also multiplicative noise, proportional to the control signal. Despite the major breakthrough SOFC brought, it still presents some drawbacks. First, a specific controller needs to be built for each and every new task. Plus, control laws computed by SOFC are time invariant: an increase or decrease of task time mainly results in a scaling of velocity (and other derivative) profiles, failing to account for increase time after a perturbation and for movement segmentation.

In the following sections, we will see how the need for a prolongation of the constructive approach that led to SOFC was answered to by optimal control with sequential terminal constraints and receding horizon by edicting new principles to tackle these issues.

2.5.1 Equilibrium point

Feedforward and feedback control can be combined inside a paradigm subdividing motor control into two phases of *motion planning* and *trajectory execution*: a - potentially optimal - feedforward trajectory is prepared and then followed, acting as an attracting reference to both induce free movement and corrective responses to external perturbations. Hogan [1984b] presented such a control paradigm starting from the following statement: for a given dynamical system (*e.g.* an arm) described by a joint state \mathbf{x} , to any vector of muscle activations \mathbf{A} should correspond an equilibrium state - or a posture - \mathbf{x}_0 . As a consequence, he postulates that movements are planned by setting a sequence of postures constituting a temporal *virtual trajectory* $\mathbf{x}_0(\{\mathbf{A}(t)\}) = \mathbf{x}_0(t)$. Movement is then induced by a torque defined by a gain - or angular stiffness - that drives the body along this reference trajectory, that was generated during the planning phase (see Fig. 2.10). Using the same notations as in Eq. 2.5 this model may be written as:

$$\mathbf{M}\dot{\mathbf{x}}(t) + \mathbf{c} - \boldsymbol{\tau}_e = \mathbf{K}(\mathbf{x}_0(t) - \mathbf{x}(t)) \quad (2.18)$$

Note that the actuation control and dynamic to generate the wished torque is omitted in this formulation.

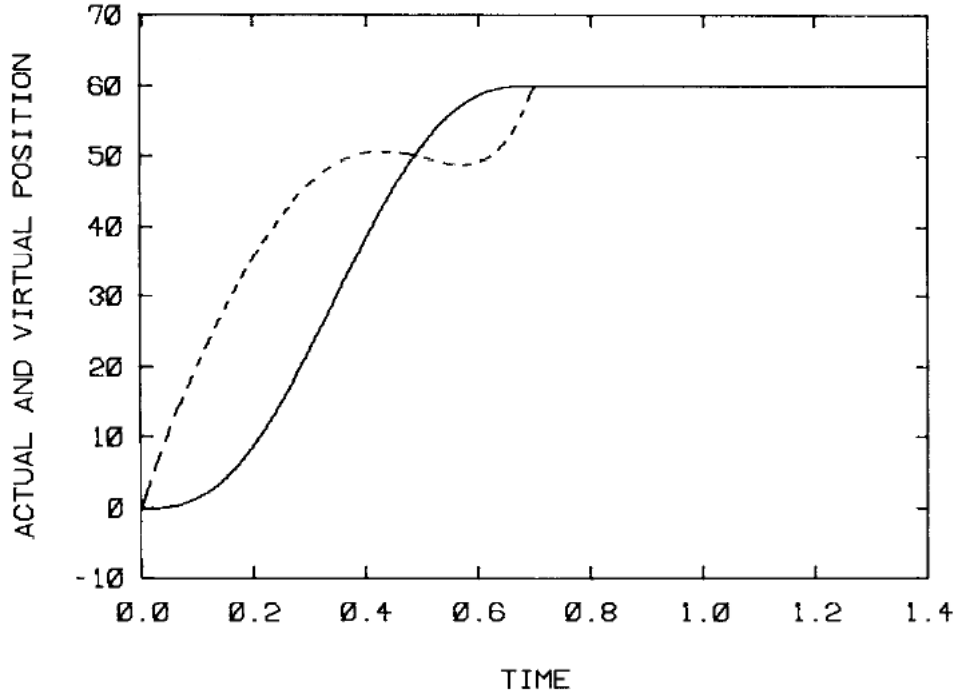


FIGURE 2.10: One dimensional trajectory generated by the equilibrium point theory. Dashed line is the virtual trajectory and plane line is the actual trajectory. Extracted from [Hogan, 1984b].

Interestingly, the virtual trajectory $\mathbf{x}_0(t)$ followed a principle of optimality: in order to reproduce the smoothness of actual movements, trajectories were to begin and end with zero velocity and to minimize the following cost-function:

$$J = \int_{t_0}^{t_f} \|\gamma(t)\| dt \quad (2.19)$$

where t_0 and t_f are the starting and end time of the movement, and $\gamma = \frac{d^3 \mathbf{p}(t)}{dt^3}$ is the **jerk**: the third derivative of the position, or the rate of change of acceleration. These so-called *minimum jerk* trajectories have been widely used as they are the product of an analytically solved optimal problem but virtually any trajectory could serve as a reference trajectory. Similarly, the concept of reference trajectory was often called for in attempt of modeling answers to perturbations. It was even extended to account for motor adaptation by Shadmehr and Mussa-Ivaldi [1994] (see Chapter 3).

A major downfall of control policies based on reference trajectories is the lack of versatility. Let us consider a motor task consisting in the reaching of a line. In this framework, the preparation phase leads to the choice of a point and its corresponding trajectory, minimizing a given cost-function. If unperturbed, the simulated trajectory would match the smooth, apparently optimal actual movements. In the presence of perturbations (*e.g.* internal noises or unexpected external forces), though, this control policy would persevere in getting back to the

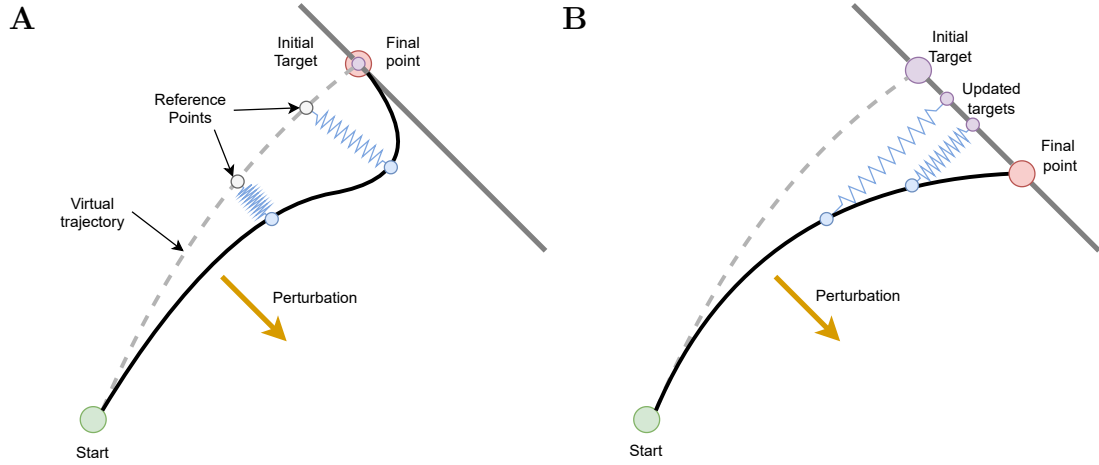


FIGURE 2.11: **Perturbed reaching of a line.** **A.** Equilibrium point - The virtual trajectory (grey dashed line) is computed in a preparation phase as the optimal trajectory to reach the objective line (grey line). It is composed of a series reference points the controller tries to reach successively. Consequently, when exposed to a perturbation the planned target (pale red) is reached regardless of line-defined task. **B.** Optimal and experimental behavior - As the task is to reach the line, the lateral perturbation is not corrected and the controller virtually tries to reach updated targets, sliding along the task line.

planned trajectory, while the observed (and optimal) strategy is to not correct the deviation in the direction of the task line (see Fig. 2.11 and Todorov and Jordan 2002). We will see hereafter that only (certain forms) of optimal control account for such strategies.

2.5.2 Optimal control

In its most general formulation [Kirk, 2004; Stengel, 1994; Bryson, 1999], an optimal control problem is defined on a dynamic system of the form:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) \quad (2.20)$$

subjected for $t \in [t_0, t_f]$ to the constraints:

$$\mathbf{x}(t_0) = \mathbf{x}_0 \quad (2.21)$$

$$\boldsymbol{\psi}(\mathbf{x}(t_f), t_f) = 0 \quad (2.22)$$

$$\mathbf{g}(\mathbf{x}(t), \mathbf{u}(t), t) \leq 0 \quad (2.23)$$

The aim is to find an admissible control policy \mathbf{u}^* that causes the defined system to follow a state evolution \mathbf{x}^* admissible regarding Eq 2.21 to Eq. 2.23 and minimizing a cost-functional J :

$$\mathbf{u}^* = \underset{\mathbf{u}'}{\operatorname{argmin}}(J(\mathbf{u}', t_0, t_f)) \quad (2.24)$$

$$J(\mathbf{u}', t_0, t_f) = \phi(\mathbf{x}(t_f), t_f) + \int_{t_0}^{t_f} L(\mathbf{x}(t), \mathbf{u}'(t), t) dt \quad (2.25)$$

\mathbf{u}' is then called an *optimal control* and \mathbf{x}^* the corresponding *optimal trajectory*.

This definition describes an optimal control problem with a given initial condition (\mathbf{x}_0), a free endpoint subject to an algebraic equality constraint (ψ), and control and state trajectories subject to algebraic inequality constraints (g). The cost functional includes an endpoint cost term (ϕ) and a trajectory cost term (L).

In the field of human motor control, the considered system will generally be all or parts of the human body whose dynamics is composed of a passive, inertial part (*e.g.* limb dynamics) as well as an actuation part (*e.g.* muscle activation dynamics). A motor task is defined as the evolution of the state (\mathbf{x}) describing the body (or part of it) from an initial state (\mathbf{x}_0) to a terminal state answering to a final cost objective (ϕ) and/or constraint (ψ) relative to a goal (\mathbf{x}_f) while respecting anatomical and physiological constraints (g) - *e.g.* joint amplitude or maximum muscle force. To accomplish such a task, motor units are excited by neural inputs (\mathbf{u}) which evolution is considered to be optimal regarding a defines performance index - or cost-function (J).

Note that the terms presented in this general formulation are not necessarily present in all optimization problems: in this work, continuous constraints (g) won't be used and final cost objectives will only appear in LQR (section 2.5.2.1), LQG (section 2.5.2.2) and SOFC (section 2.5.2.3).

The Euler-Lagrange equations stem from the calculus of variations [Kirk, 2004] and set necessary analytical conditions for solutions to an optimal problem. A derivation of these equations and a simple application are provided in Appendix A. The Hamilton-Jacobi-Bellman (HJB) equation define a necessary and sufficient condition for optimality when the problem has a fixed end time and doesn't include path constraints. It emerges from the principle of optimality: a trajectory on a given interval $[t_0, t_f]$ is optimal if and only if it is also optimal on all sub-intervals of the form $[t, t_f]$. This equation is solved backward in time, providing the core of numerical solving of optimal problems. Multiple methods of general (direct or indirect) non-linear programming have been developed to solve optimal problems.

The successive theories that built optimal motor control toolbox have been guided by the necessary compromise between realism and computational capabilities. Across the different frameworks, the adopted dynamic might be simplified and the goal and cost-function definition may vary. Hereafter, we give an overview of the funding theories that applied optimal control theory in attempt of studying human motor control.

2.5.2.1 Linear Quadratic Regulator

The LQR figures as an exception in the optimal control field, as it is an analytical solution to the well known Hamilton-Jacobi-Bellman (HJB) equation which generally calls for numerical solving. For this framework to be applied, the optimization problem needs to be built on a linear dynamic and a quadratic cost function. Thus, considering a system described by a state $\mathbf{x}_t \in \mathbb{R}^n$ and controlled by a signal $\mathbf{u}_t \in \mathbb{R}^m$ following a linear dynamic of the discrete form:

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t \quad (2.26)$$

we define a quadratic cost-function over timesteps $t = [0 \cdots N]$:

$$J = \mathbf{x}_N^T \mathbf{Q}_N \mathbf{x}_N + \sum_{t=0}^{N-1} (\mathbf{x}_t^T \mathbf{Q}_t \mathbf{x}_t + \mathbf{u}_t^T \mathbf{R}_t \mathbf{u}_t) \quad (2.27)$$

where \mathbf{Q}_t and \mathbf{R}_t are sequences of (positive definite) weighting matrices, representing respectively *tracking cost* and *effort cost*. Qualitatively, they set priorities between the state's and control's components to be minimized. The optimal feedback controller is then a Linear Quadratic Regulator:

$$\mathbf{u}_t = -\mathbf{L}_t \mathbf{x}_t \quad (2.28)$$

where the feedback gains \mathbf{L}_t are computed backward in time following the recursive equation:

$$\mathbf{L}_t = (\mathbf{R}_t + \mathbf{B}^T \mathbf{S}_{t+1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{S}_{t+1} \mathbf{A} \quad (2.29)$$

$$\mathbf{S}_t = \mathbf{Q}_t + \mathbf{A} \mathbf{S}_{t+1} (\mathbf{A} - \mathbf{B} \mathbf{L}_t) \quad (2.30)$$

Task representation is embedded in \mathbf{Q}_t , as explained in 2.5.2.5, corresponding to a cost-objective similar to ϕ in section 2.5.2.

2.5.2.2 Linear Quadratic Gaussian controller

The aforementioned LQR provides an optimal control law to be applied on the current state of a system whose dynamic is perfectly known. In reality, not only may the dynamic be noisy but the exact state of the system might be partially observed and this measurement plagued with noise. We thus consider the following discrete noisy dynamic, and observation process:

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \boldsymbol{\omega}_t \quad (2.31)$$

$$\mathbf{y}_t = \mathbf{H}\mathbf{x}_t + \boldsymbol{\xi}_t \quad (2.32)$$

where $\boldsymbol{\omega}_t$ and $\boldsymbol{\xi}_t$ are independent multivariate normal random variables with mean 0 and covariance matrices $\boldsymbol{\Omega}^\omega$ and $\boldsymbol{\Omega}^\xi$ (the discrete equivalent of power spectral

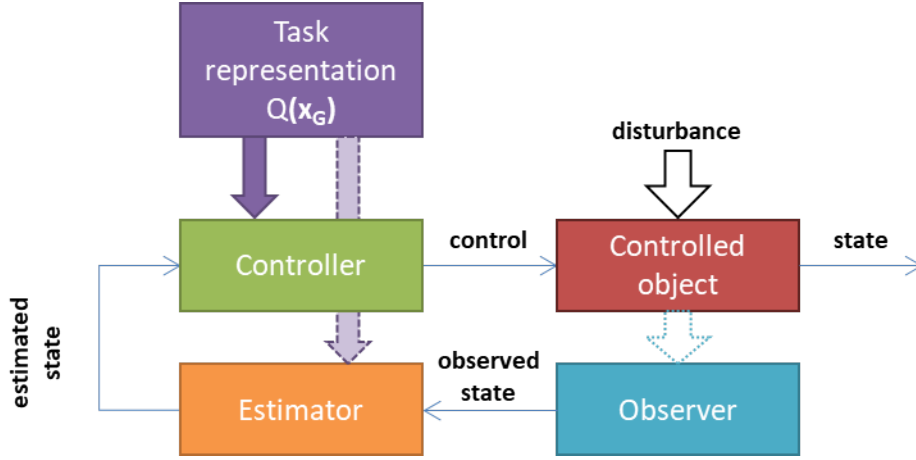


FIGURE 2.12: Estimation/Control architecture: in LQG (2.5.2.2), SOFC (2.5.2.3) and iLQR (2.5.2.4), control and estimation are built to be optimal regarding different criterion. In SOFC and iLQR, both the controller and the estimator embed task representation, while only the controller does in LQG.

density matrices in section 2.4.3). The cost-function from Eq. 2.27 is modified:

$$J(\mathbf{u}) = \langle \mathbf{x}_N^T \mathbf{Q}_N \mathbf{x}_N + \sum_{t=0}^{N-1} (\mathbf{x}_t^T \mathbf{Q}_t \mathbf{x}_t + \mathbf{u}_t^T \mathbf{R} \mathbf{u}_t) \rangle_{\omega, \xi} \quad (2.33)$$

where $\langle \cdot \rangle_{\omega, \xi}$ is the expectation operator over noises ω and ξ . Then the LQR needs to be coupled with a Linear Quadratic Estimator - also known as a Kalman filter - to compute the control upon estimated state (Fig. 2.12):

$$\mathbf{u}_t = -\mathbf{L}_t \hat{\mathbf{x}}_t \quad (2.34)$$

where the estimated state is:

$$\hat{\mathbf{x}}_{t+1} = \mathbf{A} \hat{\mathbf{x}}_t + \mathbf{B} \mathbf{u}_t + \mathbf{K}_t (\mathbf{y}_t - \mathbf{B} (\mathbf{A} \hat{\mathbf{x}}_t + \mathbf{B} \mathbf{u}_t)) \quad (2.35)$$

and the Kalman gain is computed forward in time following (a more complete description of state estimation is given in Appendix B):

$$\mathbf{K}_t = (\mathbf{A} \mathbf{P}_t \mathbf{A}^T + \Omega^\omega) \mathbf{H}^T (\mathbf{H} (\mathbf{A} \mathbf{P}_t \mathbf{A}^T + \Omega^\omega) \mathbf{H}^T + \Omega^\xi)^{-1} \quad (2.36)$$

$$\mathbf{P}_{t+1} = (\mathbf{I}_n - \mathbf{K}_t \mathbf{H}) \mathbf{P}_t \quad (2.37)$$

As in LQR (section 2.5.2.1), task representation is embedded in \mathbf{Q}_t , as explained in 2.5.2.5. Note that consequently, state estimation is independent of the task at hand.

2.5.2.3 Stochastic optimal feedback control

SOFC is an extension of the LQG (section 2.5.2.2) embedding the mathematical formulation of a powerful view of motor control based on the observation of motor variability as a consequence of redundancy and noise: trajectory planning and execution should not be viewed as two successive and separated phases but as a unified, continuous optimization process. This view allows for the concise yet powerful *Minimum Intervention Principle* [Todorov and Jordan, 2002]:

"Deviations from the average trajectory are corrected only when they interfere with task performance."

Additionally, to account for speed-accuracy trade-off [Fitts, 1954; Harris and Wolpert, 1998], multiplicative noise was incorporated into Eq. 2.31:

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \sum_{i=1}^m \mathbf{C}_i \mathbf{u}_t \epsilon_{i,t} \quad (2.38)$$

where m is the number of controls (see section 2.5.2.1) and \mathbf{C}_i are constant $n \times m$ matrices. Those matrices were constructed in Todorov and Jordan [2002] in a way that propagates the noise from the control of one actuator to the other. Arguably this reproduces the fact that during effective actuation of human limb geometry (*e.g.* an arm), noise in the torque generated at a given joint would be projected onto all Cartesian coordinates, depending on the joint configuration.

The observation process is of the same type as Equation 2.32:

$$\mathbf{y}_t = \mathbf{H}\mathbf{x}_t + \boldsymbol{\xi}_t \quad (2.39)$$

$\boldsymbol{\epsilon}_t = \begin{bmatrix} \epsilon_{1,t} \\ \vdots \\ \epsilon_{m,t} \end{bmatrix}$ and $\boldsymbol{\xi}_t$ are independent multivariate normal random variables with

mean 0 and covariance matrices $\boldsymbol{\Omega}^\epsilon$ and $\boldsymbol{\Omega}^\xi$. Equation 2.33 was then amended to take into account proportional noise:

$$J(\mathbf{u}) = \langle \mathbf{x}_N^T \mathbf{Q}_N \mathbf{x}_N + \sum_{t=0}^{N-1} (\mathbf{x}_t^T \mathbf{Q}_t \mathbf{x}_t + \mathbf{u}_t^T \mathbf{R} \mathbf{u}_t) \rangle_{\epsilon, \xi} \quad (2.40)$$

where $\langle \rangle_{\epsilon, \xi}$ is the expectation operator over noises ϵ_i and $\boldsymbol{\xi}$ and \mathbf{Q} embed task definition (see 2.5.2.5). The optimal control is

$$\mathbf{u}_t = -\mathbf{L}_t \hat{\mathbf{x}}_t \quad (2.41)$$

with the usual state estimation:

$$\hat{\mathbf{x}}_{t+1} = \mathbf{A}\hat{\mathbf{x}}_t + \mathbf{B}\hat{\mathbf{u}}_t + \mathbf{K}_t(\mathbf{y}_t - \mathbf{H}\hat{\mathbf{x}}_t) \quad (2.42)$$

where \mathbf{K}_t and \mathbf{L}_t are recursively defined as

$$\mathbf{K}_t = \mathbf{A}\Sigma_t^e \mathbf{H}^T (\mathbf{H}\Sigma_t^e \mathbf{H} + \mathbf{\Omega}^\xi)^{-1} \quad (2.43)$$

$$\Sigma_{t+1}^e = (\mathbf{A} - \mathbf{K}_t \mathbf{H}) \Sigma_t^e \mathbf{H}^T + \sum_i^n \mathbf{C}_i \mathbf{L}_t \Sigma_t^{\hat{x}} \mathbf{L}_t^T \mathbf{C}_i^T \quad (2.44)$$

$$\Sigma_{t+1}^{\hat{x}} = \mathbf{K}_t \mathbf{H} \Sigma_t^e \mathbf{A}^T + (\mathbf{A} - \mathbf{B} \mathbf{L}_t) \Sigma_{t+1}^{\hat{x}} (\mathbf{A} - \mathbf{B} \mathbf{L}_t)^T \quad (2.45)$$

$$\mathbf{L}_t = \left(\mathbf{B} \mathbf{S}_{t+1}^x \mathbf{B}^T + \mathbf{R} + \sum_{i=1}^n \mathbf{C}_i^T (\mathbf{S}_{t+1}^x + \mathbf{S}_{t+1}^e) \mathbf{C}_i \right)^{-1} \mathbf{B}^T \mathbf{S}_{t+1}^x \mathbf{A} \quad (2.46)$$

$$\mathbf{S}_t^x = \mathbf{Q}_t + \mathbf{A}^T \mathbf{S}_{t+1}^x (\mathbf{A} - \mathbf{B} \mathbf{L}_t) \quad (2.47)$$

$$\mathbf{S}_t^e = \mathbf{A}^T \mathbf{S}_{t+1}^x \mathbf{B} \mathbf{L}_t + (\mathbf{A} - \mathbf{K}_t \mathbf{H})^T \mathbf{S}_{t+1}^e (\mathbf{A} - \mathbf{K}_t \mathbf{H}) \quad (2.48)$$

2.5.2.4 iLQR and iLQG

The optimal control methods presented until now only apply to linear dynamics. However, as seen in section 2.3.5 the sensorimotor system is highly non-linear. LQR and LQG were then extended to iLQR [Li and Todorov, 2004] and iLQG [Todorov and Li, 2005] to handle non-linear dynamics. Briefly, these methods linearize the system's dynamics around a nominal trajectory on which a control law is computed and applied to improve iteratively the nominal trajectory.

2.5.2.5 Task representation

A crucial characteristics of methods presented from section 2.5.2.1 to 2.5.2.4, all relying on the Linear-Quadratic-Gaussian formalism, is that task definition is done within the cost-function. More precisely, it is embedded in the matrices \mathbf{Q}_t in Eqs. 2.27, 2.33 and 2.40. Consequently, control laws in Eqs 2.29 and 2.46 are built on this task representation. Moreover, in SOFC, as the Kalman gain depends on the control law (in a recursive way - see section 2.43), state estimation too depends on the task.

Additionally, Todorov and Jordan give two possible implementations for a task:

- embedding the task's goal coordinates in the tracking cost, *i.e.* $\mathbf{Q}_t = \mathbf{Q}_t(\mathbf{w}_t, \mathbf{x}_{G,t})$
- augment the state with the task's goal coordinates and setting only weights

in the tracking cost *i.e.* the state is $\tilde{\mathbf{x}}_t = \begin{bmatrix} \mathbf{x}_t^T \\ \mathbf{x}_{G,1}^T \\ \vdots \\ \mathbf{x}_{G,N}^T \end{bmatrix}$ and $\mathbf{Q}_t = \mathbf{Q}_t(\mathbf{w}_t)$

where \mathbf{w}_t are weights and $\mathbf{x}_{G,t}$ are goals to reach (or approach) during the task.

With the first approach, two different tasks automatically imply two different sets of \mathbf{Q}_t and two different control laws. The second is a little more versatile: a set of \mathbf{Q}_t and its corresponding control law contain only the timing and weights of the while the actual goals can be chosen *via* the augmented state $\tilde{\mathbf{x}}_t$.

2.5.3 Achievements of the SOFC framework

The Linear-Quadratic-Gaussian formalism laid a solid ground for the motor control community, putting an end to the feedforward/feedback debate by introducing a more comprehensive control/estimation architecture. LQR (section 2.5.2.1) defined an optimal control law (or feedback gain) that is independent of the initial state of the system, making these control paradigms efficient for repeating and achieving a given task with variable initial states. It also allowed to set priorities between subgoals and effort, even though this feature presents some drawbacks regarding task definition (see section 2.5.4).

LQG (section 2.5.2.2) broadened optimal control to stochastic systems, introducing the Kalman gain that provides an optimal estimation of the state of the system, built on the balance of confidence that can be put into state observation and state propagation prediction.

SOFC (section 2.5.2.3, [Todorov and Jordan 2002](#)) introduced signal-dependent noise in the dynamic and the minimum intervention principle. From emerges a key feature of this theory: trajectory planning and execution are no longer sequential but simultaneous throughout the task. A direct consequence of this principle is a declination of the Uncontrolled Manifold [[Scholz and Schöner, 1999](#)], illustrated in Fig. 2.4 (see also [Todorov and Jordan 2002](#), fig. 2): when reaching for a line, perturbations in the direction of the line are not corrected. In addition to this spatial structure of variability, temporal organisation is also accounted for: serving the purpose of maximizing final performance, variance is allowed to be high in the early stages of the task (see [Todorov and Jordan 2002](#), fig. 4). Moreover, since corrective actions are optimal regarding the task at hand, redundancy is not only solved in a pre-computed planned trajectory, but dynamically, spreading optimally variance and the burden of corrections and on available actuators (see [Todorov and Jordan 2002](#), fig. 6), as observed for instance in bimanual pointing tasks, examined through the UCM framework [[Domkin et al., 2002](#)].

iLQR and iLQG (section 2.5.2.4) later prolong this framework to non-linear dynamics.

SOFC framework has since been widely employed and extended in various studies extending but not limited to: introducing gaussian trajectory goals to produce graffitis [[Berio et al., 2017](#)], widening its capabilities to obstacle avoidance [[Singh et al., 2018](#)] or to deal with goals uncertainty [[Haith et al., 2015](#)].

2.5.4 Limitations of the SOFC framework

Despite being a major breakthrough in the computational approach of human movement, a few limitations can be opposed to the SOFC framework.

First, the task is mostly defined in the cost-function (section 2.5.2.5). Even in the augmented state representation, in which the state objectives are not in the weighting matrices themselves, the weights actually have to be adapted in magnitude and timing to the task. This leads to the major downfall that a new controller (and estimator) needs to be computed for each task encountered, and potentially stored. The computational burden of such a process and the neural implementation of the storage of several controller is not realistic. Furthermore, the very setting of the weights themselves seems like a ill-defined problem to solve beforehand: what could be the rationale deciding to favor the completion of a velocity goal over a excitation one, or over the overall effort throughout the task? [Müller et al., 2017]

Second, movement time is not predicted by the model but rather a parameter defining the task (Eq. 2.33). As a consequence, the first order impact of modulating the task-time (setting apart the increased variability due to bigger controls) will be a scaling of the control outputs and, as a result, a scaling of each derivative of the kinematics of the generated trajectory (Fig. 2.13). Yet, in reality movements are segmented: long duration movements (typically longer than 0.5 s) will exhibit velocity fluctuations (Fig. 2.6). As a matter of fact, as exposed in section 2.3.6, human movements show a very consistent frequency content (*e.g.* the number of peaks per seconds) over a large variety of tasks (reaching, drawing, scribbling, ...) through the successive derivatives of the position (velocity, acceleration and jerk). Conversely, modulating the task distance (*e.g.* for a reaching task) while keeping the same time results in the same type of scaling mentioned before, giving rise to isochrony not only for fast but for all movements [Guigon et al., 2019].

Finally, as the time remaining to accomplish a given task decreases as it is done, a perturbation in the last moments would either result in tremendous efforts to reach the goal afterwards or in the failing of the task. Attempts have been made to adapt feedback gains to handle task perturbations [Liu and Todorov, 2007; Dimitriou et al., 2013] but do not tackle the issue of the computational and storage burden expose hereabove.

The aforementioned limitations should not cast a sententious shadow upon SOFC and its breakthrough, but they call to be handled. The next section expose how it was achieved comprehensively in a optimal control model inheriting SOFC capabilities and built on elementary modelling principles.

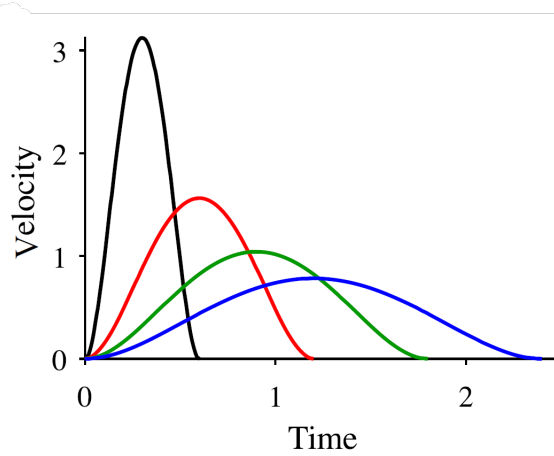


FIGURE 2.13: **Velocity scaling.** Simulations of trajectories with the same start and ending point with varying task time from the shortest (black) to the longest (blue). Extracted from [Guigon et al., 2019]

2.6 Optimal control with sequential terminal constraints and receding horizon

SOFC came on top of tens of years of fruitful research navigating between modeling predictions and experimental observations challenging and enriching each other. It is the same constructive process that led Emmanuel Guigon to propose a model [Guigon, 2022] that would build up a broader, comprehensive computational theory on the solid ground laid down by Todorov and Jordan [2002]. In order to reconcile optimal control with movements segmentation and overcome the aforementioned limitations, three principles were stated.

2.6.1 Modelling principles

The principles exposed hereafter were defined in the framework of control theory [Guigon, 2022]:

Motor control is considered as a control problem in which the behavior of a controlled object is governed by a controller through a control policy and a set of goals to achieve.

They are thoroughly explained in the referenced paper, but we will nonetheless give a brief summary of their motivation and core essence.

2.6.1.1 Universal Control policy

As exposed in 2.5.4, SOFC defines a controller upon the task it is called to achieve. This very property is at the core of one of the arguments often made against optimal control: the feasibility of the neural implementation of such a control framework. Indeed, the computational burden of building the control law for a given task makes the construction "on the fly" of a controller adapted to the task at hand hard to conceive. On the other hand, human motor control is extremely reactive in successively perform different tasks. Moreover, as will be exposed in Chapter 3, we humans have the incredible ability to quickly adapt to external perturbations never encountered before, and the even more amazing ability to "remember" these perturbations to re-adapt even quicker in any re-exposition to the learned perturbations. These converging arguments suggest that a tremendous number of control policies should, in a way, be stored in the brain. Such an assertion is not reasonable, but does not rule out optimal control in itself. We show hereafter that it is not only feasible to build a flexible controller from a universal control policy, but that, coupled with the two subsequent principles, it presents a large power of prediction and a plausible neural implementation.

The universal control policy we consider is inspired from model predictive control (MPC): at each moment, a new control is computed to optimally reach a given goal from the current estimated state. In contrast, in their original formulation LQR (section 2.5.2.1), SOFC (section 2.5.2.3) and iLQR (section 2.5.2.4) derive a control law for the whole fixed-time trajectory in a *preparation phase* which is then applied during the *execution phase* to determine the current optimal control based on the best estimation available of the current state. Yet, they can be used in our chosen framework by computing a whole control law at each timestep and only apply its first control.

General equations 2.20 to 2.25 need to be amended according to the MPC framework and to account for stochasticity. The dynamics at hand becomes:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) + \mathbf{n}_{dyn}(\mathbf{x}(t), \mathbf{u}(t), t) \quad (2.49)$$

At a time t , we seek to find an optimal control $\mathbf{u}(t)$ to take the system from its current state $\mathbf{x}(t)$ to a state satisfying a terminal condition

$$\psi(\mathbf{x}(T(t)), \mathbf{x}_G(t)) = 0 \quad (2.50)$$

where $\mathbf{x}_G(t)$ is a potentially time-varying goal, within a potentially time-varying time frame $T(t)$. The control $\mathbf{u}(t)$ is the evaluation at time t of the function \mathbf{u}^* , defined on $[t, T(t)]$ which is optimal regarding a cost functional

$$\mathbf{u}^* = \underset{\mathbf{u}'}{argmin} \left\{ \int_t^{T(t)} L(\mathbf{x}(\tau), \mathbf{u}'(\tau), \tau) d\tau \right\} \quad (2.51)$$

To summarise, the optimal control at time t depends on the current state of the system, the defined terminal condition, the predictable, deterministic part of the dynamics, the chosen cost-function, and the time to reach the current goal:

$$\mathbf{u}(t) = \boldsymbol{\pi}(\mathbf{x}(t), \boldsymbol{\psi}, \mathbf{x}_G(t), \mathbf{f}, T(t)) \quad (2.52)$$

To define our control policy to be *universal*, we assume that the dynamic upon which it is built is constant, irrespective of time and tasks, and corresponds to the body's own dynamic so that:

$$\mathbf{u}(t) = \boldsymbol{\pi}_u(\mathbf{x}(t), \boldsymbol{\psi}, \mathbf{x}_G(t), T(t)) \quad (2.53)$$

A strong implication of this choice is that the dynamic used to compute the control cannot be adapted to any task or interaction (see section 2.8.2 and Chapter 3).

2.6.1.2 Receding horizon

One of the most challenging concept in modelling human motor control is time. Before even trying to account for temporal characteristics of human motor behavior, one faces a fundamental modelling issue: defining the *beginning* and the *end* of a movement. In experimental setups, defining movement onset and offset is tricky: even in studying simple motor tasks in controlled environment, absolute immobility is never observed on unconstrained limbs. To overcome this difficulty, one often considers velocity thresholds to detect the beginning or ending of movement, but other methods were developed [Botzer and Karniel, 2009]. Yet, this difficulty expresses an intuitive observation: the human body is a system in ever-going motion, even if of small amplitude.

Modelling does not face such embarrassment, for initial and final velocities (and generally states) can be freely set in a controller at times - or *horizon* - chosen following three paradigms:

- **fixed horizon:** at time t of the simulation, the control is computed to reach the current goal \mathbf{x}_G within a decreasing time $T(t) = T_0 - t$, where T_0 is a time fixed in advance;
- **infinite horizon:** at any time of the simulation, the control is computed to reach the current goal \mathbf{x}_G within an infinite time $T(t) = \infty$, and the actual task time is consequential to the optimization process [Rigoux and Guigon, 2012];
- **receding horizon:** at any time of the simulation, the control is computed to reach the current goal \mathbf{x}_G within a fixed time $T(t) = T_H$.

A control policy with fixed horizon cannot account for the temporal flexibility motor control naturally exhibits: the task is performed within the predefined time T_0 , even when exposed to noises or external perturbations while in reality movements are prolonged to compensate for them. A fixed horizon paradigm is thus not satisfactory.

Furthermore, it is observed that only fast movements are smooth. Optimal control methods exposed in Section 2.5.2 with fixed horizon are time-invariants: setting different task times T_0 leads to a simple scaling of velocities (and its subsequent derivatives), allowing smooth movements of any duration. Likewise, infinite horizon generate smooth movements in all configurations.

Receding horizon is highly flexible: if perturbations drive the controlled system away from its goal, the effective completion time will be automatically extended without any update on the controlled policy. Moreover, while isochronous behavior can be produced by setting up diverse fixed goals, the segmentation of movements [Guigon et al., 2019] can also be accounted for. Finally, with this setting the control policy becomes stationary (independent of time), contributing to the feasibility of a neural implementation of optimal control (see Chapter 4).

Hence, receding horizon was employed, with a *fixed* - over time, space and tasks - value of $T_H = 0.28$ s, which was determined in a study of slow movements [Guigon et al., 2019]. The universal control policy described in Eq. 2.53 becomes stationary:

$$\mathbf{u}(t) = \boldsymbol{\pi}_u^{T_H}(\mathbf{x}(t), \boldsymbol{\psi}, \mathbf{x}_G(t)) \quad (2.54)$$

In the particular case of a full-state terminal constraint, it simplifies further as:

$$\mathbf{u}(t) = \boldsymbol{\pi}_u^{T_H}(\mathbf{x}(t), \mathbf{x}_G(t)) \quad (2.55)$$

This choice imposes constraints on the representation of the timing of tasks (see 2.5.2.5) but we will see in Chapter 4 that the stationarity of the control has interesting implications regarding neural implementation.

2.6.1.3 Sequential goals

In the LQG and SOFC framework trajectories may be defined by a sequence of goals, embedded in the cost-function (see 2.5.2.5). Two issues come out of this feature: (1) the reaching time between two successive goals is fixed in advance, without any flexibility regarding the impact of potential perturbations; (2) Movements have marked start and ending time and place, which seems a bit odd considering that one does not think one movement after another but rather as a natural flow of actions, continuously and seamlessly achieving motor tasks.

An alternative view is proposed: rather than seeing it as a strict succession of motor tasks, motor control should be pictured as a perpetual reaching task

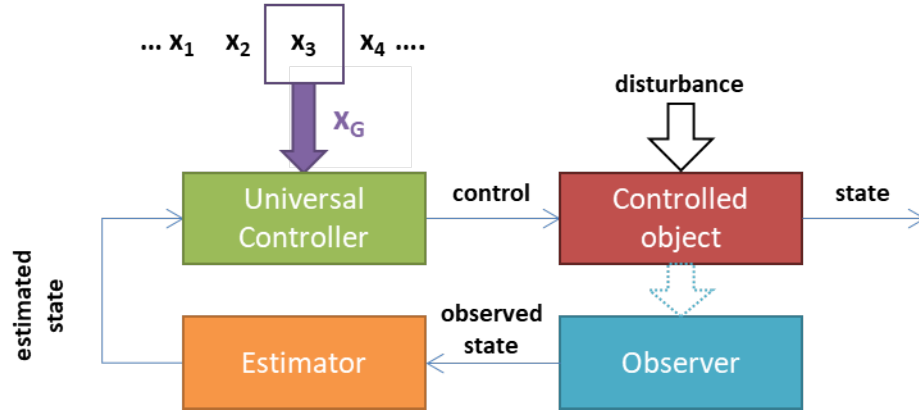


FIGURE 2.14: **Optimal Control with sequential terminal constraints and receding horizon.** The controller is universal and compute control toward a goal \mathbf{x}_G that is set at a fixed frequency from a sequence of states. State is estimated to counter motor and observation noises.

toward a current goal \mathbf{x}_G that can be updated in order to fulfill complex motor tasks. Importantly, even complex, continuous motor tasks are to be decomposed as a series of discrete goals to be targeted successively or overlapping. Additionally, if the task at hand does not present temporal constraints, \mathbf{x}_G is updated at fixed and unique frequency ($1/T_G$, with $T_G = 0.13$ s). This value was identified in a study of slow movements to account for velocity fluctuations at ~ 8 Hz [Guigon et al., 2019] and, as T_H , should be seen as a constant. In the same paper are given rules on the application and most importantly definition of sequential goals: they can be set as partial constraints on the state (*e.g.* only position and velocity are set) as well as functions (to constrain end-effector in a redundant system for instance).

Interestingly, this value coupled with the one of T_H implies that a new goal generally won't be reached in a single time-frame of T_G , allowing the representation of continuous tasks in this framework (see section 2.6.2.1).

2.6.2 Implementation

The schematic adopted representation of human motor control is given in Fig. 2.14: in order to control the state \mathbf{x}_t of a potentially noisy controlled system, a continuous and potentially noisy observation \mathbf{y}_t is performed, while a state estimation $\hat{\mathbf{x}}_t$ is used as an initial state for the determination of an optimal control \mathbf{u}_t to reach a current goal \mathbf{x}_G .

2.6.2.1 Task representation

As previously exposed, we define the control policy to be universal. This first principle demands all task representation to be embedded in the goals provided to the controller. Formally, Eq. 2.50 defines them as terminal boundary conditions (allowing for instance to define full state objectives as well as endpoint cartesian objectives) but we will hereafter refer to them as *goals*, written \mathbf{x}_G . In this framework:

- **reaching** is achieved by setting \mathbf{x}_G to \mathbf{x}_{reach} and maintaining it long enough for the system to stabilize. For unperturbed movements and with the determined horizon time 0.28 s, it takes roughly 0.4 s, regardless of the objective state (or distance) giving rise to isochrony for fast movements (Fig. 2.15A). In order to perform slower reaching movements are done by setting intermediary via-points before the actual goal (Fig. 2.15B)
- **posture** is achieved by maintaining \mathbf{x}_G to $\mathbf{x}_{posture}$ for as long as the task goes on. With perfect control and observation, this results in trying to get the system to a state he is already in. In the presence of noise and/or perturbations, deviations from the objective state induce corrective actions, potentially generating oscillations (Fig. 2.16C).
- **tracking** is achieved by extracting a sequence goals from the trajectory to follow and successively feeding them to the controller as its current goal (Fig. 2.16D).

Guigon [2022] provides a more detailed and comprehensive list of task representation: drawing, scribbling, rhythmic movements, ...

Resisting static force is a task that has been thought to be handled by a static controller, while movements were generated by a dynamic controller according to the *separation principle* [Guigon et al., 2007a]. In our framework, though, no such distinction is allowed. Appendix E proposes diverse task representation to position control opposing static forces but also to force control. In essence, the goal state \mathbf{x}_G needs to include an overshoot of the actual aimed position, an overshoot of the activation and excitation corresponding to the force or a combination of those. Although proper study need to be ran to confront them to experimental data, such a task representation show an interesting "trick": the goals set to the controller may be different from the actual objective state.

2.6.2.2 Optimal computation

Section 2.6.1.1 stated a first principle that defines the element of the optimization process: the dynamic considered is bound to be the one of the system to control

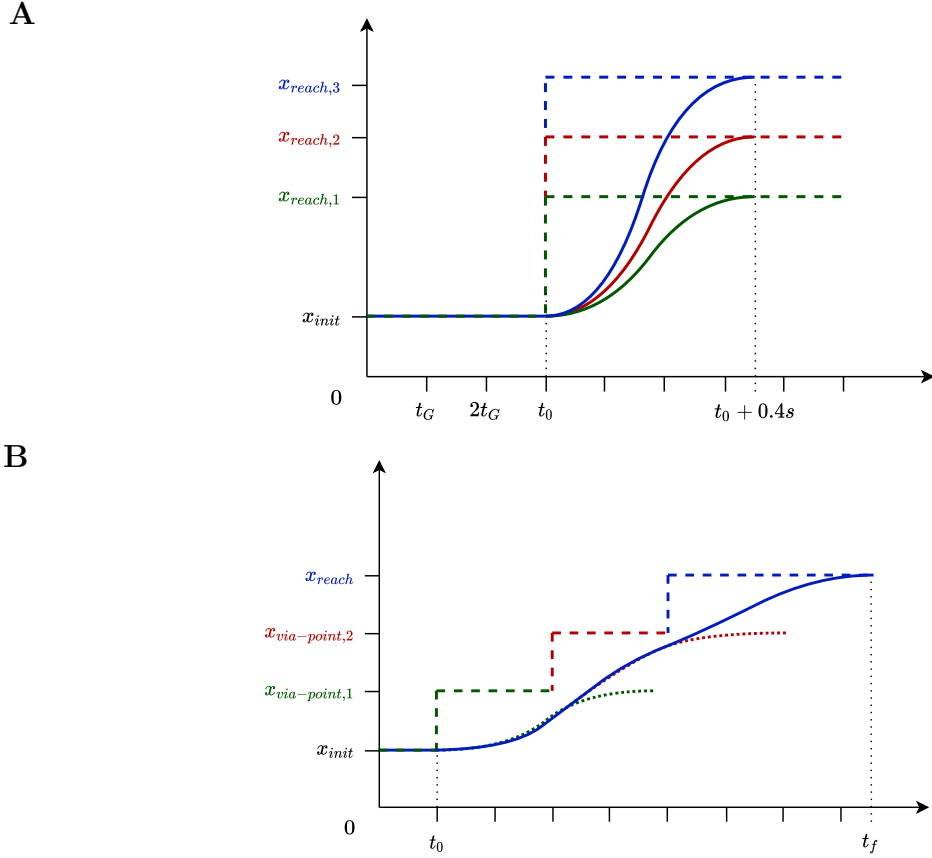
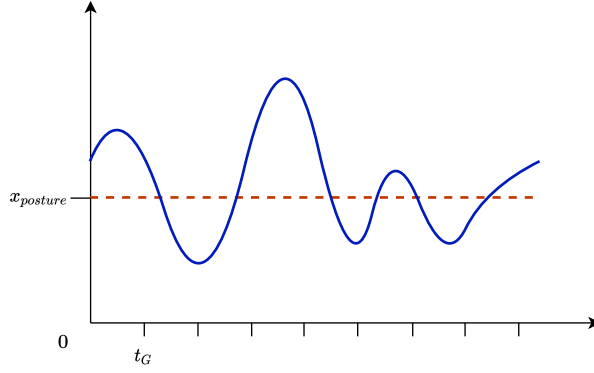


FIGURE 2.15: **Task representation (1/2)**. All figures are illustrations: y-axis represent general state, with arbitrary unit. Dashed red lines represent the current goal the controller is trying to reach. Plain line represent actual trajectories. **A.** Reaching task with three different goals, all attained within the same time (roughly 0.4 s), showing *isochrony*. **B.** Slower reaching task: two intermediary via-points are successively fed to the controller before the actual goal. Dotted lines show the trajectories followed if those via-points were pursued.

(*i.e.* the body or used limbs), regardless of any interaction with the environment (see Chapter 3). The second principle (section 2.6.1.2) identified the optimization time to be receding. Finally, task representation is defined by the third principle as a series of state to be fed at a fixed frequency to the controller as its current terminal objective. Two elements remain to be determined:

- **cost-function:** many cost-function definitions have been introduced in attempt of reproducing diverse human movement characteristics, and attempts have been made to combine them in hybrid cost-functions and evaluate their respective contributions [Berret et al., 2011]. Arguments could be made on the likelihood of cost-functions, but given the approximations made on the modelling of the phenomena at hand (inertial dynamics, muscles insertions and activation, ...), we postpone the endeavour of finding the best cost-function to the point where it will appear to be a critical element in accounting for motor control characteristics. Hence, we stick to the *effort*

C



D

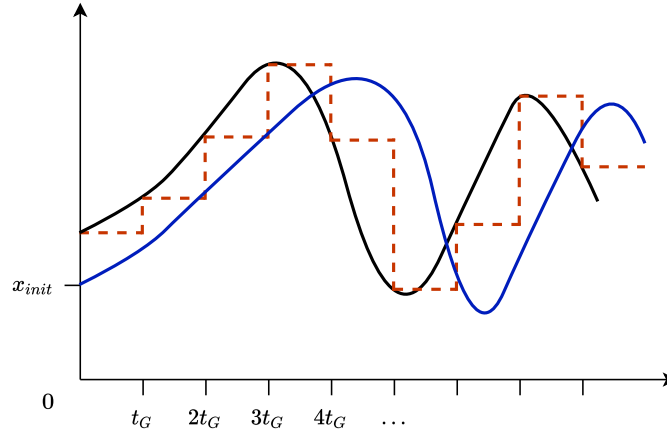


FIGURE 2.16: **Task representation (2/2)**. All figures are illustrations: y-axis represent general state, with arbitrary unit. Dashed red lines represent the current goal the controller is trying to reach. Plain blue line represent actual trajectory. **C.** Postural task: the controller always aims at the same state, correcting for inner or external perturbations. **D.** Tracking task: via-points (dashed red lines) are extracted from the followed trajectory (plain black line).

cost-function adopted by [Todorov and Jordan \[2002\]](#):

$$J = \int_t^{t+t_H} \mathbf{u}_\tau^T \mathbf{u}_\tau d\tau \quad (2.56)$$

- **controller current goal:** as stated in sections 2.6.1.3 and 2.6.2.1, tasks are to be translated into series of goals, defining successively the current goal of the controller. The general formulation expresses goals as an absolute constraint (ψ in Equation 2.50) but they may be implemented as an objective within the cost-function, alongside the effort:

$$J = \int_t^{t+t_H} (\mathbf{u}^T \mathbf{u} + (\mathbf{x} - \mathbf{x}_G)^T \mathbf{Q}(t)(\mathbf{x} - \mathbf{x}_G) d\tau \quad (2.57)$$

with $\mathbf{Q}(t) = \text{diag}(\mathbf{w}_g)$ where under the condition that the weights in \mathbf{w}_g be high enough for the controller not to make a compromise between task

completion and effort. This requirement is fundamental in order to use iLQR as an optimal solver in this framework.

2.6.2.3 State estimation

The sensorimotor system is plagued with noise (section 2.3.2) and delays (section 2.3.3) and the state of the system is known only through a noisy observation process:

$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t - \Delta) + \mathbf{n}_{obs}(\mathbf{x}(t - \Delta), t) \quad (2.58)$$

To incarnate the ability of human motor control to overcome this difficulty, state estimation was introduced in the LQG (section 2.5.2.2) by the means of a *Kalman filter*. In SOFC (section 2.5.2.3) and iLQR (section 2.5.2.4), the control law was recurrently defined from the Kalman gain. Yet, it was shown that such a close coupling between control and estimation is not necessary for the *minimum intervention principle* to apply and structured variability to be accounted for [Guigon et al., 2008]: a deterministic controller and a stochastic estimator perform equally well. Thus, the control computation described in section 2.6.1.1 ought to be based on the best estimate of the current state: $\hat{\mathbf{x}}(t)$.

State estimation was performed through an extended Kalman filter - which is the general, non-linear formulation (see Appendix B). In order to approach the continuous nature of human movement, a discretized version of the continuous formulation of the Kalman filter was adopted (although with small enough simulation time-steps results are similar with the discrete Kalman filter):

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{f}(\hat{\mathbf{x}}(t), \mathbf{u}(t)) + \mathbf{K}(t)(\mathbf{y}(t) - \mathbf{H}\hat{\mathbf{x}}(t)) \quad (2.59)$$

2.7 The common content of motor signals

One way to circumvent the smoothness issue raised in section 2.3.6 is to choose a smoothing level once and for all, and use it for any and every dataset we intend to analyse. In this way, the data will be more or less smooth depending on the chosen level, but there will be no opportunity to adjust the smoothness at will. Furthermore, analysis of acceleration and jerk profiles can be interesting to reveal hidden irregularities in velocity profiles. These points are illustrated in Fig. 2.17 where acceleration and jerk profiles for samples of tongue movement during speech (Fig. 2.17A), drawing (Fig. 2.17B,C) and displacement of the center of pressure during posture (Fig. 2.17D) are shown. Two observations can be made. First, there is an apparent common temporal structure for the three tasks despite differences in task nature, displacement amplitude, acceleration and jerk (Fig. 2.17A,B,D). Second, there is no apparent differences of temporal structure between healthy

and pathological drawing movements (healthy elderly vs patient with Parkinson's disease; Fig. 2.17B,C).

The present work is cast in the framework of computational motor control [Todorov and Jordan, 2002]. In this framework, an action is the outcome of a control process that elaborates a signal to drive an object (e.g. a set of body segments and a muscular system) toward a goal. A simple and widely used case consists in a mass point (m) driven by a second-order linear muscle (time constant ν). Its dynamics writes

$$\begin{aligned}\dot{p}(t) &= v \\ \dot{v}(t) &= \frac{\alpha(t)}{m} \\ \dot{\alpha}(t) &= \frac{\varepsilon(t) - \alpha(t)}{\nu} \\ \dot{\varepsilon}(t) &= \frac{u(t) - \varepsilon(t)}{\nu}\end{aligned}\tag{2.60}$$

where p , v , α , ε , u are position, velocity, muscle activation, muscle excitation, and control signal, respectively, and the purpose of control is to find an appropriate control policy $u(t)$ to obtain some intended behavior [Harris and Wolpert, 1998; Todorov and Jordan, 2002]. Irrespective of the exact control policy, the model indicates that characteristics of the produced behavior (e.g. $p(t)$) are a by-product of the filtering and integration of the control signal by the muscular system and the object dynamics. According to this model, if we record a position signal during a behavioral motor task, it is expected that its successive derivatives (velocity, acceleration, jerk, ...) should become an increasingly closer approximation of the putative control signal as the order of differentiation increases. Consider the experimental data shown in Fig. 2.18A (a single trial of slow back-and-forth arm movements; Guigon et al. 2019). A position signal was recorded, successive derivatives were obtained numerically, all the signals were low-pass filtered at 10 Hz and the peak frequency (number of min+number of max/duration/2) in each signal was calculated. Note that the peak frequency in the position signal itself (black; Fig. 2.18A,B) is not relevant since it corresponds to task requirements (in fact, a position signal can be transformed in a monotonic signal without changing the frequency content of its derivatives). We observed by eye that the frequency of fluctuations increased with the order of differentiation (from top to bottom; Fig. 2.18A). This property was found over the entire dataset together with a decrease in the variability in peak frequency (Fig. 2.18B). This latter observation indicates that widely different velocity profiles are produced by a comparatively restricted set of temporal patterns of acceleration. These observations suggest, according to the proposed computational framework, that the putative (unobservable) control signal for this dataset should contain fluctuations in a ~ 6 -10 Hz range. Three sets of synthetic signals are provided to weight this conclusion and show that it is

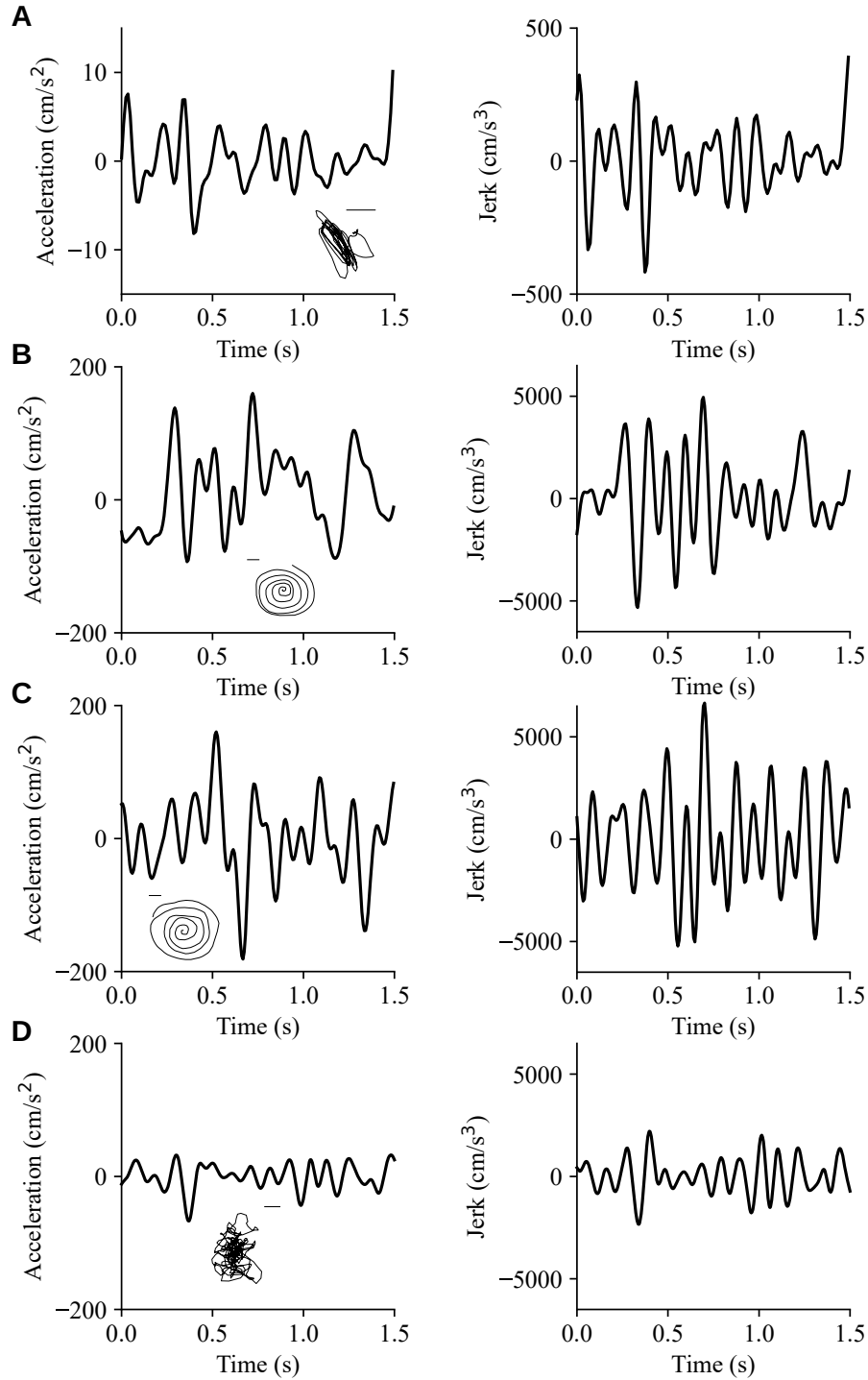


FIGURE 2.17: **Acceleration and jerk profiles for diverse motor tasks.** **A.** 1.5-s extract of acceleration and jerk profiles for tongue tip displacement during speech (inset; calibration 5 mm). Data from Kuberski and Gafos [2019]. **B.** Same as **A** for a 1.5 s extract of a single spiral drawing trial (inset; calibration 5 cm). Data from a control participant from Drotár et al. [2016]. **C.** Same as **B** for a participant with Parkinson's disease. **D.** Same as **A** for a 1.5 s extract of the centre of pressure trajectory during quiet stance (inset; calibration 1 cm). Data from Apthorp et al. [2014]. All the time series are low-pass filtered at 10 Hz. Kuberski and Gafos [2019]: <https://doi.org/10.5281/zenodo.2273898> / Drotár et al. [2016]: <http://bdalab.utko.feec.vutbr.cz/> / Apthorp et al. [2014]: <https://doi.org/10.6084/m9.figshare.1126648.v1>

neither an epiphenomenon nor an artifact (Fig. 2.18C,E,G). First, a model with fluctuations of the control signal at a unique, fixed frequency (~ 8 Hz; see section 2.6 and Guigon et al. 2019) can generate synthetic signals that are consistent with the experimental data (no fitting was attempted; Fig. 2.18C,D). The same model with fluctuations of the control signal at variable frequencies fails to account for the experimental observations (not shown). Second, not all signals (e.g. a purely sinusoidal signal) display a scaling of peak frequencies (Fig. 2.18E,F). Third, random signals (here a shuffled version of the data in Fig. 2.18 A) display the scaling property but without a decrease in the variability in peak frequency (Fig. 2.18G,H).

To further assess the content of motor signals, we selected a dataset that contains data from several independent recording systems. In Wang and Majewicz Fey [2018], participants performed rhythmic arm movements in a Fitts' paradigm (14 participants, 36 trials/participant). Arm position was recorded in 3D with an haptic device, and IMUs collected velocity and acceleration signals. A sample trial is shown in Fig. 2.19A and its successive derivatives in Fig. 4B. For this trial, the peak frequency increased with the order of differentiation (Fig. 2.19B, top to bottom). This property was also observed across the trials although with a substantial variability (Fig. 2.19E, top to bottom; Fig. 2.19H, left). Note that we are not interested in the temporal structure of the position signal itself since it mostly reflects task requirements. We noticed that the peak frequencies varied weakly with the mean velocity calculated on each trial (Fig. 2.19E), although a Bayesian test systematically rejected an absence of variation. Similar observations were made for velocity recordings (Fig. 2.19C,F and H, center) and acceleration recordings (Fig. 2.19D,G and H, right). Note that, in the latter case, peak frequency was displayed as a function of mean acceleration (Fig. 2.19G). Mean peak frequency consistently increased with the order of differentiation for the three types of signal (Fig. 2.19H) which supports the idea that it is not merely an artifact of data processing (numerical differentiation, filtering). Yet differences were observed, e.g. between acceleration obtained by an accelerometer and acceleration obtained by differentiation (Fig. 2.19H, green) which indicates that data processing can introduce artifacts. The mean peak frequency covaried with the "true" frequency content of the signals (estimated by the mean centroid frequency of the power spectrum calculated on the unfiltered signals; Fig. 2.19H, horizontal bars) which further confirms the suitable nature of the peak frequency measure.

To explain the variability in peak frequency, we considered two trials with widely different kinematics: large and long velocity fluctuations (Fig. 2.20A) vs small and short fluctuations (Fig. 2.20B). They were characterized by large differences in velocity and acceleration peak frequencies (Fig. 2.20A,B; red and green), but minimal difference in jerk peak frequencies (Fig. 2.20B; blue). This means that a common temporal structure in higher derivatives can correspond to dramatically different structures in the lower derivatives. It is unsurprising in mathematical terms, but interesting as a strategy for motor control.

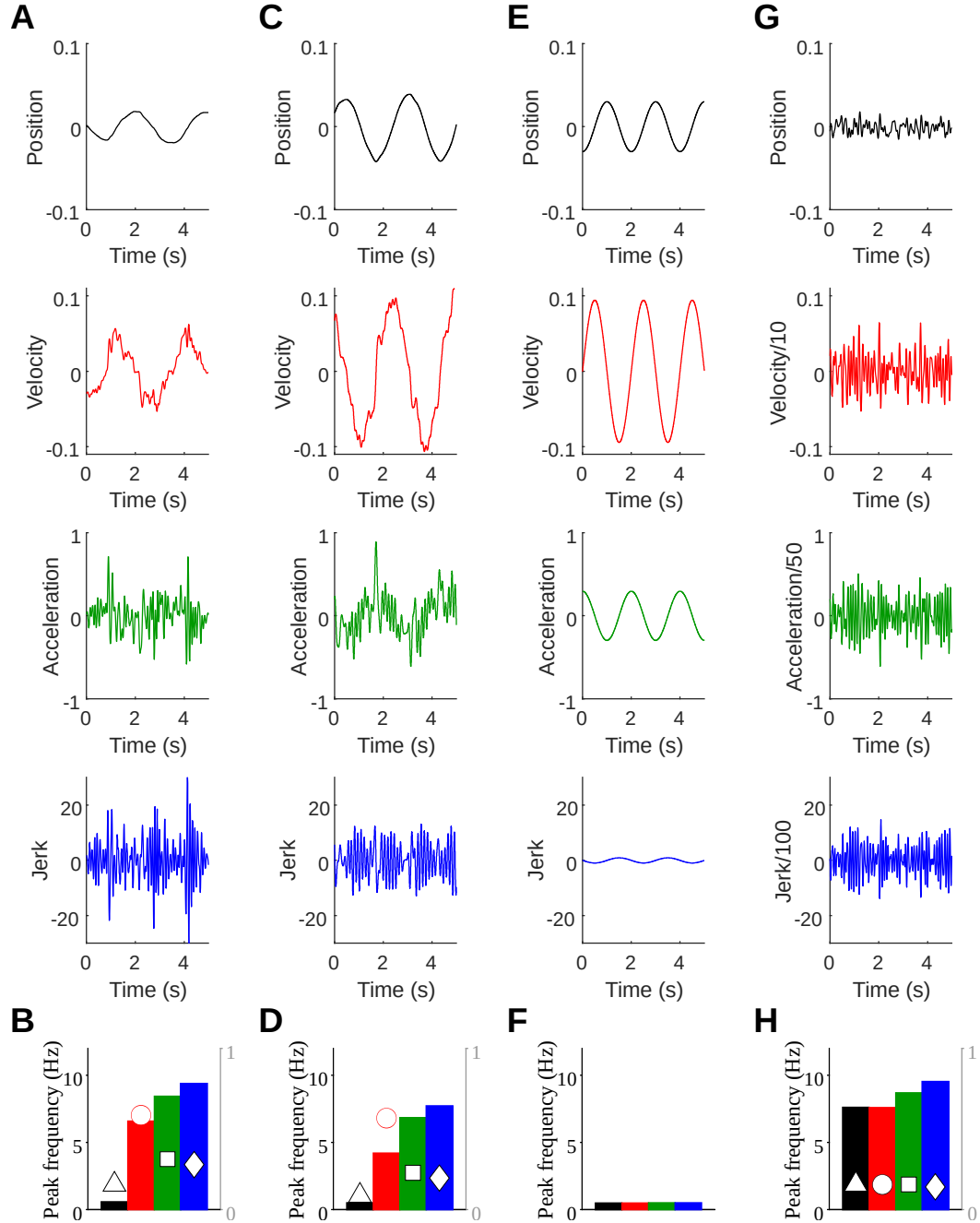


FIGURE 2.18: **Frequency contents of experimental and synthetic data.**

A. A 5-s extract of position, velocity, acceleration and jerk profiles for one trial of slow arm displacement. Data from Guigon et al. [2019]. **B.** Histogram of mean peak frequency across the complete dataset for data in **A**. The symbols give the standard deviation (right axis). **C.** Synthetic data generated using the proposed model. **D.** Histogram for data in **C**. **E.** Synthetic sinusoidal trajectory. **F.** Histogram for data in **E**. **G.** Shuffled signal from **A**. **H.** Histogram for data in **G**. Guigon et al. [2019]: <https://doi.org/10.6084/m9.figshare.5977894>

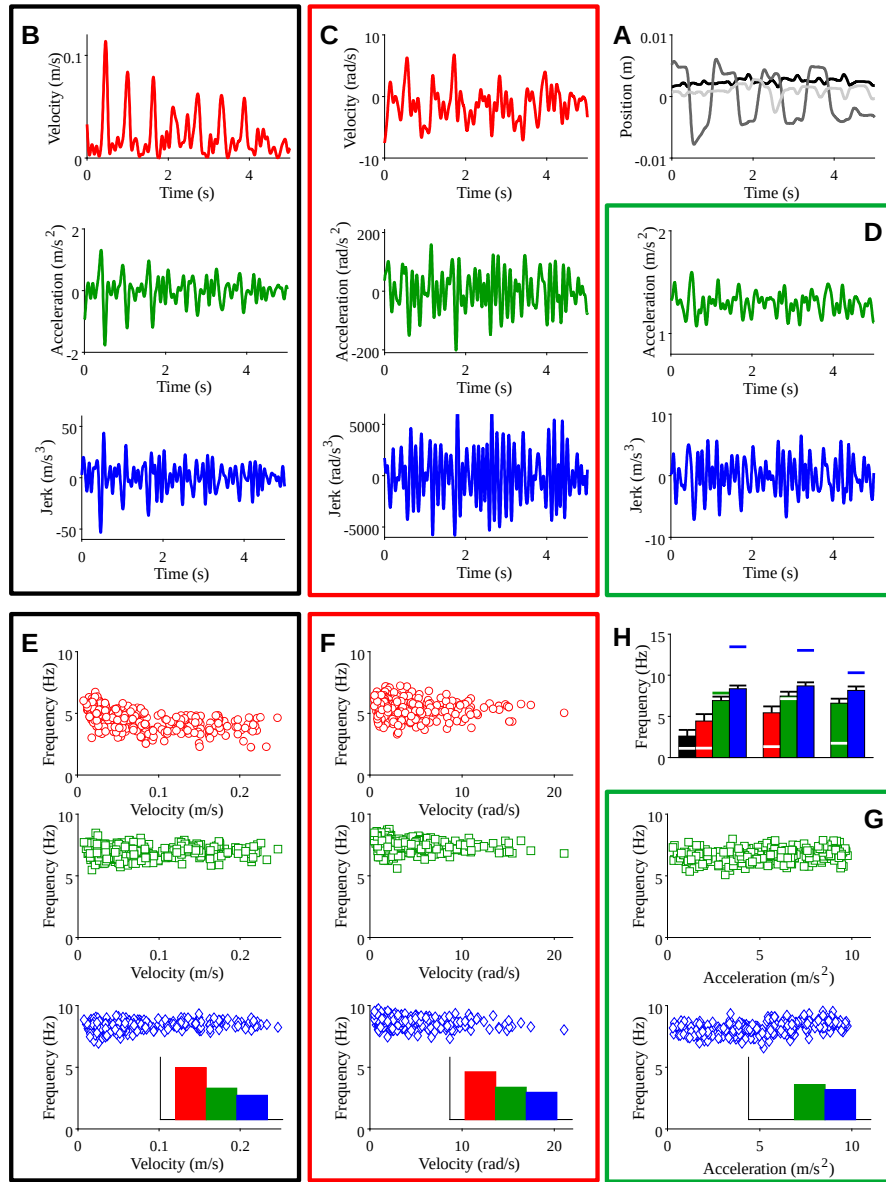


FIGURE 2.19: **Frequency content of data recorded with various sensors** (data from Wang and Majewicz Fey [2018]). **A.** A 5-s extract of position recorded with an haptic device (3D coordinates: black, dark gray, light gray) for one trial. **B.** Cartesian velocity, acceleration and jerk for data in **A**. **C.** One-axis velocity signal recorded in an IMU for the trial in **A** and its derivatives. **D.** One-axis acceleration signal recorded in with an IMU for the trial in **A** and its derivative. **E.** Velocity, acceleration and jerk peak frequency for the position signal as a function of mean velocity for all the trials (only 1 over 5 trials shown). One trial is shown in **A**, **B**. Inset: standard deviation. **F.** Velocity, acceleration and jerk peak frequency for the recorded velocity signal as a function of mean velocity for all the trials (only 1 over 5 trials shown). One trial is shown in **C**. **G.** Acceleration and jerk peak frequency for the recorded acceleration signal as a function of mean acceleration for all the trials (only 1 over 5 trials shown). One trial is shown in **D**. **H.** Histogram of peak frequency for **E** (left), **F** (center), **G** (right). Horizontal lines (white or colored) indicate the centroid frequency of the power spectrum calculated on unfiltered data. Wang and Majewicz Fey [2018]: <https://doi.org/10.7910/DVN/YN89X6>

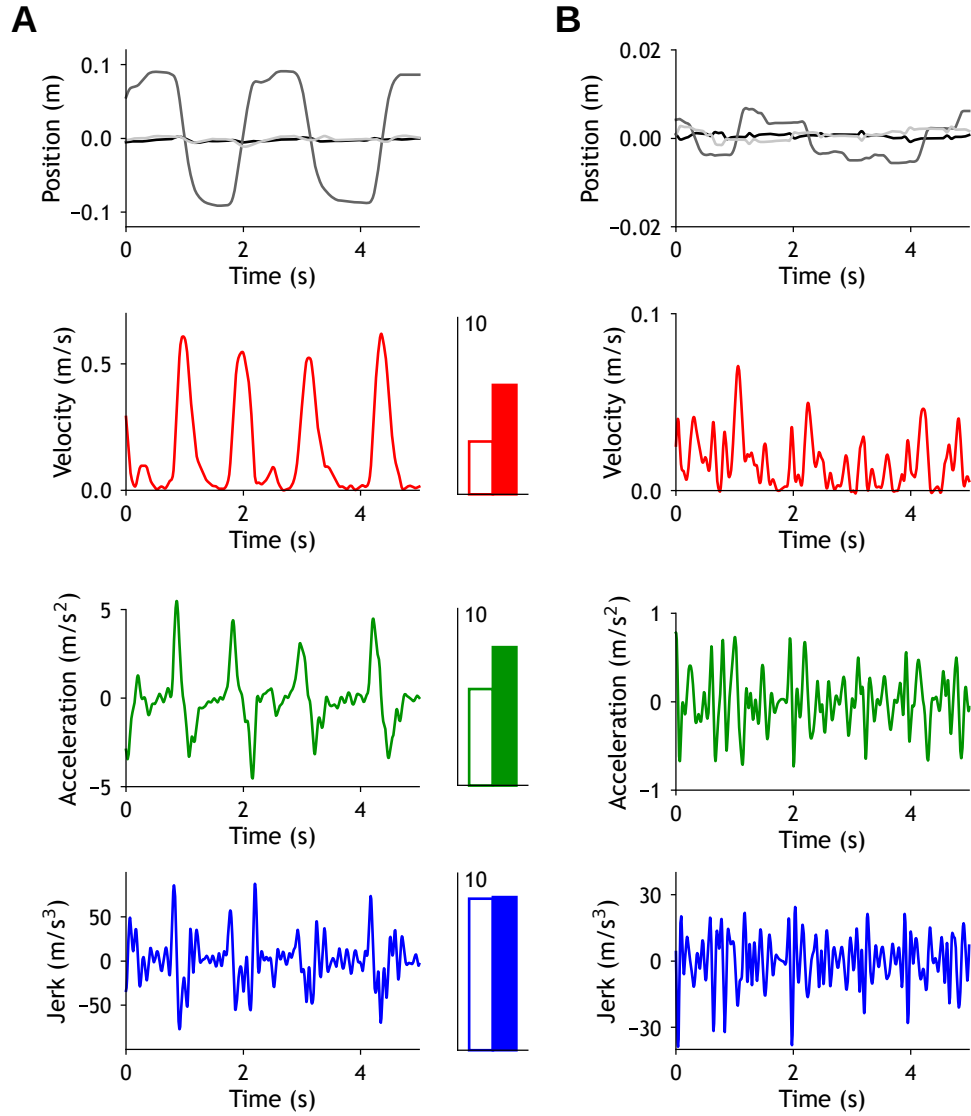


FIGURE 2.20: Data from Wang and Majewicz Fey [2018]. Two trials with different velocity fluctuations. The central histograms indicate the peak frequency for the traces in **A** (open) and **B** (closed).

These observations are very general and can be reproduced on many different datasets for different tasks (arm movement, gait, posture, ...) and different species (human, monkey, rat, ...) [Guigon et al., 2022]. They put strong constraints on models of motor control. In fact, most models, in particular those based on optimal control, fail to account properly for the content of motor signals (see section 2.5.4).

2.8 Discussion

The study of human motor control has come a long way since the first crucial observations that laid a solid ground for attempts in modelling how the CNS produces voluntary movements. Optimal control offered an efficient computing tool

to reproduce the general apparent smoothness and gracefulness of rapid reaching movements. It was further used to generate reference trajectories to which the body would be bound to by the intermediary of gains both provoking motion and ensuring stability in presence of perturbations. Yet, the limitation of correction to deviations that impact performance (described by the uncontrolled manifold) could not be accounted for. The LQG proposed a most pertinent structure fusing optimal control and optimal estimation, allowing to drive linear stochastic systems according to trajectory goals embedded in a cost function.

SOFC extended this framework by introducing signal proportional noise and minimizing a cost function expressed as an expectation over noises, thus formulating the *minimum intervention principle* [Todorov and Jordan, 2002]. This elegant solution to the redundancy problem accounted for structured variability in both task and joint space. yet showed some limitations: (1) a strict definition of task time prevented timing adjustments to perturbations and did not account for movements segmentation; (2) the task definition was embedded in the control law, making unrealistic a direct neural implementation.

Optimal control with receding horizon and sequential goals prolonged these works proposing three modelling principles [Guigon, 2022]: (1) the control policy is universal; (2) the time to reach a goal is constant; (3) goals are updated at a fixed frequency. This model managed to reproduce a wide range of task (reaching, writing, scribbling, rhythmic tasks, ...) and account for most properties of human movement (scaling laws, isochrony, segmentation, power laws, speed/accuracy tradeoff, structured variability). Most importantly it proposed a decisive view of motor control that have three strong implications: (1) the CNS is in a perpetual reaching task; (2) each and every motor task is translated into a sequence of goals to reach with a unique, constant and universal controller; (3) at all times, control is unequivocally defined by the current estimated state of the body and the current goal to reach.

2.8.1 Everything is "movement"

2.8.1.1 Relationship with postural models

Sections 2.6.1 and 2.6.2 modelled the control process as a stationary computation depending solely on the current estimated state of the system and the current goal to reach:

$$\mathbf{u}(t) = \boldsymbol{\pi}(\hat{\mathbf{x}}(t), \mathbf{x}_G(t)) \quad (2.61)$$

The motor action - or torques - may thus be expressed as:

$$\boldsymbol{\tau}(t) = \boldsymbol{\Gamma}(\hat{\mathbf{x}}(t), \mathbf{x}_G(t)) \quad (2.62)$$

where $\mathbf{\Gamma}$ is the solution in $\boldsymbol{\tau}$ of the differential system defined by Eq. 2.10. The whole dynamic of the system may thus be written as:

$$\mathbf{M}\dot{\mathbf{x}}(t) + \mathbf{c} - \boldsymbol{\tau}_e = \mathbf{\Gamma}(\mathbf{x}_G(t), \hat{\mathbf{x}}(t)) \quad (2.63)$$

As a reminder, the model proposed by the equilibrium point theory was (Eq. 2.18)

$$\mathbf{M}\dot{\mathbf{x}}(t) + \mathbf{c} - \boldsymbol{\tau}_e = \mathbf{K}(\mathbf{x}_0(t) - \mathbf{x}(t)) \quad (2.64)$$

Here we see a striking resemblance between a theory portraying motor control as perpetual posture and ours, viewing it as perpetual movements. The main advantage of our proposed model is that it embeds control dynamics, optimality, redundancy solving and minimum intervention principle. Hopefully this conceptual similarity/symmetry will contribute to reconcile the remaining divergence between these two theories.

2.8.2 Everything is task representation

The proposed model is built upon three simple, elementary principles. The first one states that all motor tasks are managed by a single, unique and universal controller. This bold claim has already been validated on numerous tasks by Guigon [2022] (see 2.6.2.1 for the schematic representation of some of them). In all those tasks the human body was isolated from its surrounding environment, evolving freely. In reality though, we are in perpetual interaction with our environment. In certain cases, those interactions can be omitted while modelling for they are either negligible by nature (*e.g.* friction in scribbling tasks) or orthogonal to the task by experimental design (*e.g.* planar manipulandum) making the *free hypothesis* valid. But in other conditions not only are interactions non negligible but their astonishing good management by the CNS is of great interest for the study of human motor control. It is in particular the case for perturbations, which may be divided in three categories:

- **stochastic perturbations:** be the nature of their cause internal (*e.g.* inner noises - see section 2.3.2) or external (*e.g.* intermittent force pikes, vibrations), their intensity and time of occurrence cannot be anticipated nor modelled. At best can their statistics - in the form of spectral density matrix - be evaluated. Optimal state estimation is a powerful approach that exploits the perturbations' statistics to minimise their effect on the computation of voluntary controls, but observation and action delays (see section 2.3.3) may in certain circumstances (perturbations of high frequency or intensity) make descending commands insufficient to perform the motor task at hand. Fortunately, the human body is equipped with tools allowing to produce corrective actions without involving the brain. The first to come

into play is of purely mechanical nature and lies in the impedance of muscle-tendons units (see section 2.2.5). Its efficiency can be adapted by the means of the simultaneous activation of agonist and antagonist muscles: *cocontraction* [Hogan, 1984a]. The second to intervene is the set of reflexes pathways that generate automatic, peripheral commands to modify muscle activity. Several mechanisms allow to modulate the sensitivity of those reflexes.

- **continuous perturbations:** interactions with our environment is often continuous (*e.g.* when moving in water) and may lead to the application of unknown forces upon our end effector, which follow their own dynamics (*e.g.* the fluid dynamics). Despite these forces, motor tasks may be achieved: at first, movements may be clumsy, but after a surprisingly short time task performance rises up, and trajectories come close to those under unperturbed conditions, approaching back a form of optimality. This phenomenon is called **motor adaptation**, and has been mostly studied using force fields exerted via robotic manipulanda [Shadmehr and Mussa-Ivaldi, 1994].
- **tool manipulation:** during the manipulation of complex objects (*e.g.* surgical tools, active prosthesis), one is submitted to aforementioned continuous perturbations. Yet, motor tasks are not to be achieved despite the interacting object, but *with* the object, redefining not only the global dynamics but the end effector.

After having been confronted to and validated on free tasks, it is thus essential to examine how interaction tasks may be cast in our proposed framework.

Similarly to the separation principle [Guigon et al., 2007a], cocontraction is often modelled as an additive command, injected alongside movement commands following a reference trajectory [Gribble et al., 1998; Latash and Gottlieb, 1991]. In our proposed model, no reference trajectory definition is allowed, but a resembling additive command may be included. This was the object of an exploratory theoretical study presented in Appendix F, in which a candidate task representation and preliminary modelling predictions are proposed.

As it is the case for the human body, continuous perturbations may be described and predicted by a set of differential equations constituting a dynamics. It is therefore tempting to describe motor adaptation as either the addition of force commands countering the predicted action of the perturbation [Shadmehr and Mussa-Ivaldi, 1994] or as the inclusion of the perturbation's dynamics into a new controller, built on a new global dynamic [Izawa et al., 2008]. In our framework, the first modelling principle does not allow any modification of the controller, and motor adaptation can only be cast in the form of task representation. In Chapter 3, we propose a novel view on motor adaptation and compare it to competing motor adaptation theories in the light of an experimental study.

Manipulating a tool requires *a minima* to have knowledge of its geometry and of the nature of its mechanical link with our body in order to define a new

end effector. With this relatively realistic hypothesis, defining a goal for the end effector of the tool can be translated into a goal for the linking point on the body. But is this information enough to be able to handle any tool? Or does one need to be able to estimate the velocity of the newly formed joint? To evaluate its inertial properties? Or ultimately to know the proper dynamics of the tool? These considerations could bring valuable information on how the CNS handles tools and prosthesis and ultimately to improve control laws.

2.8.3 Stationary control

The first two modelling principles we adopted - universal control policy (section 2.6.1.1) and receding horizon (section 2.6.1.2) - make the computation of optimal control stationary (Eq. 2.54). In the particular case of a linear representation of the controlled system, control is even given by the product of a constant matrix and a vector combining the current state of the system (or its best available estimation in the case of a stochastic system) and the current goal to reach. In a more realistic non-linear representation, stationarity holds, and control is found by applying a non-linear, time independent function to the current state and goal.

This property paves the way for putting an end to a recurrent objection often made to optimal control as a description of human motor control: it would see the CNS as an intelligent controller, acting as a computer rather than a system answering to the laws of physics. It appears from Latash [2010] that this view stems from the sequential representation made of processes in control theory: the *controller* receives information and then "decides" a control appropriate to the task it is trying to achieve. Feldman [2015] state that motor control should be described as a *physical law* which is defined by a set of equations bonding *law-constrained variables* (e.g. the angle of a pendulum) to parameters that can be divided into two subsets: *invariant parameters* (e.g. the gravitational constant) and *controllable parameters* (e.g. the length of the pendulum). It is argued that motor control is driven by voluntary changes of controllable parameters.

The proposed model can be perfectly cast in this view: control is in our framework directly, continuously and unequivocally linked to sensory information and the current motor task by a function with constant parameters. Control thus meets the criteria to be defined as a *law-constrained variables*, the goal of the controller as a *controllable parameters*, and the control law as a *physical law*. Universal optimal control with receding horizon is then a perfectly valid motor control theory according to the own requirements of the community of equilibrium point theory (also known as the lambda theory and referent control theory).

The next step to go further in the validation of our view of motor control is to verify the feasibility of its neural implementation. Chapter 4 addresses this issue by assessing the possibility for our optimal control law to be approximated by the composition of elementary activation functions inside a simple neural network.

Moreover, the activation properties of our artificial neurons are compared to those of the primary motor cortex.

Chapter 3

Adaptation to motor perturbations as a redirection of goals

3.1	Introduction	53
3.2	Simulation methods	55
3.2.1	Computational modelling	55
3.2.2	Compensation model	56
3.2.3	Reoptimization model	56
3.2.4	Redirection model	59
3.2.5	Numerical solution	60
3.3	Experiment	60
3.3.1	Ethics statement	60
3.3.2	Participants	60
3.3.3	Apparatus	60
3.3.4	Experimental procedure	61
3.3.5	Data processing and analysis	62
3.4	Results	63
3.4.1	Predictions	63
3.4.2	Two representative participants	66
3.4.3	All participants	68
3.4.4	Redirection model	69
3.4.5	Parametric study of the reoptimization model	70
3.5	Discussion	72

3.1 Introduction

Motor behavior is at the same time highly stable and widely flexible [Bernstein, 1967; Glencross, 1980; Krakauer et al., 2019]. On the one hand, a large repertoire

of skilled, efficient behaviors (e.g. speech production, handwriting, gait, ...) is maintained for decades and can possibly resist to injury, aging, disease or brain damage. On the other hand, a few movements performed in a novel sensorimotor environment (e.g. wearing prismatic glasses, holding a visco-elastic manipulum, ...) or in some altered physiological state (e.g. muscular fatigue, pain, ...) can induce lasting changes in motor performance [Shadmehr and Mussa-Ivaldi, 1994; Martin et al., 1996; Takahashi et al., 2006; Bouffard et al., 2014]. A proper cooperation between stability and flexibility is mandatory so that:

- ingrained skills remain sensitive to steady and persistent changes in the environment, the body and the nervous system but are not outrageously influenced by temporary, incidental events;
- new skills can develop at any time upon request.

What kind of organization of skills is compatible with such a requirement?

There are two views on motor learning and skill acquisition [Krakauer et al., 2019]. The first view holds that learning occurs at the action selection (control) level and modifies the mapping between intended goals and actions appropriate to achieve these goals. For instance, in the typical laboratory example of adaptation to a velocity-dependent force field (dynamic perturbation; Bernstein 1967; Shadmehr and Mussa-Ivaldi 1994), learning has been described either as a compensation process, i.e. a mapping is learned between states and compensatory forces that oppose to the applied forces [Shadmehr and Mussa-Ivaldi, 1994], or as a reoptimization process, i.e. a mapping is learned between goals and optimal forces to achieve the goals in the presence of the applied forces [Izawa et al., 2008]. According to the second view, learning occurs at the goal selection level and modifies the mapping between intended and actual goals irrespective of how to achieve these goals. For instance, adaptation to a visuomotor rotation of the visual display (kinematic perturbation) results from a redirection process, i.e. a remapping between target and movement vectors [Wang and Sainburg, 2005]. Although the latter learning process appears more flexible and frugal than the former, it is unclear whether it can account for adaptation to dynamic perturbations, i.e. when new patterns of force need to be learned.

Models based on compensation or reoptimization are well formulated models that can be used to make predictions on adaptation to dynamic perturbations (velocity-dependent force field; Shadmehr and Mussa-Ivaldi 1994; Izawa et al. 2008). In particular, the shape of after-effect trajectories, i.e. late trajectories in the absence of the force field after adaptation, should incorporate a "negative image" of the forces induced by the applied force field and thus resemble to mirror-transformed before-effect trajectories, i.e. early trajectories in the presence of the force field before adaptation (this is exactly the case for the compensation model; Shadmehr and Mussa-Ivaldi 1994). The shape of before-effect trajectories has

been thoroughly documented. They are initially curved "away" from the baseline (unperturbed) trajectory with a late ensuing correction toward the target [Shadmehr and Mussa-Ivaldi, 1994]. We have not found any study that quantitatively documents the shape of after-effect trajectories. Yet qualitative observations on published figures suggest that after-effect trajectories do not obey to the predicted mirror organization (fig. 5 in Thoroughman and Taylor 2005; fig. 4 in Hwang et al. 2006; fig. 1 in Nozaki et al. 2006; fig. 2 in Huang and Shadmehr 2007; fig. 2 in Darainy et al. 2009; fig. 1b in Sun et al. 2022). In fact after-effect trajectories seem to resemble "kinematic" trajectories, i.e. trajectories observed during visuomotor rotation or target jump tasks rather than to "dynamic" trajectories observed during force field tasks (examples of contrast between kinematic and dynamic trajectories in fig. 3 in Diedrichsen et al. 2005; fig. 6 in Torrecillos et al. 2015). Thus they could be compatible with a redirection process as if adaptation corresponded to aiming toward spatially remapped targets.

The goal of this study is to clarify the nature of before-effect and after-effect trajectories during a force field adaptation task and gather evidence for or against between compensation/reoptimization and redirection processes as a basis for motor adaptation.

3.2 Simulation methods

3.2.1 Computational modelling

We simulated displacements of a planar two-link arm whose dynamics and parameters were described in Chapter 2, section 2.4.1. As a reminder, the dynamics is written :

$$\ddot{\mathbf{q}} = \mathbf{M}(\mathbf{q})^{-1} (\boldsymbol{\tau}_u + \boldsymbol{\tau}_e - \mathbf{c}(\mathbf{q}, \dot{\mathbf{q}})) \quad (3.1)$$

where $\mathbf{q} = \begin{bmatrix} q_{sh} \\ q_{el} \end{bmatrix}$ are the shoulder and elbow angles, $\mathbf{M}(\mathbf{q})$ the inertia matrix, $\mathbf{c}(\mathbf{q}, \dot{\mathbf{q}})$ the vector of velocity-dependent torques, $\boldsymbol{\tau}_u$ the control torque produced by actuators and $\boldsymbol{\tau}_e$ the torque due to external forces applied on the arm.

Displacements were perturbed by a velocity-dependent force field producing a force proportional to v_{ψ}^{hand} , the hand cartesian velocity projected onto the task direction vector $\boldsymbol{\delta}_{\psi}$ (angle direction ψ is measured from initial hand position to target position and 0 is rightward) and perpendicular to this direction (colinear to $\boldsymbol{\delta}_{\psi}^{\perp}$, the normal vector to $\boldsymbol{\delta}_{\psi}$). The force applied is thus $\mathbf{F}_{\psi} = \phi v_{\psi}^{hand} \boldsymbol{\delta}_{\psi}^{\perp}$. Described in the joint space, the force field is

$$\mathbf{D} = \mathbf{R}(\psi)^{-1} \begin{bmatrix} 0 & 0 \\ \phi & 0 \end{bmatrix} \mathbf{R}(\psi) \quad (3.2)$$

where ϕ is the force level ($\phi > 0$ for a counterclockwise perturbation, $\phi < 0$ for a clockwise perturbation) and $\mathbf{R}(\psi)$ the rotation matrix of angle ψ . The perturbation torque is

$$\boldsymbol{\tau}_e = \boldsymbol{\tau}_e^\phi = \mathbf{J}(\mathbf{q})^T \mathbf{D} \mathbf{J}(\mathbf{q}) \dot{\mathbf{q}} \quad (3.3)$$

where $\mathbf{J}(\mathbf{q})$ is the Jacobian matrix defined in 2.9

In all the simulations, the initial arm configuration was $[45^\circ, 90^\circ]$, movement amplitude was 0.1 m, movement direction was $\psi = 90^\circ$, and force (field) level was $\phi = 2$ Ns/m. Four conditions were considered: *baseline*, in the absence of the force field; *before-effect*, in the presence of the force field before adaptation; *adapted*, in the presence of the force field after adaptation; *after-effect*, in the absence of the force field after adaptation.

3.2.2 Compensation model

The compensation model is taken from Shadmehr and Mussa-Ivaldi [1994]. The principle is the following. First we derive a desired 1-s spatial trajectory for a 0.1 m forward displacement based on a 0.5 s 0.1-m long minimum-jerk trajectory [Flash and Hogan, 1985] followed by a 0.5-s stationary posture. Second we use the arm inverse kinematics to obtain the desired angular trajectory $\mathbf{q}^*(t)$, and the arm inverse dynamics (Eq. 3.1) to calculate the joint torques $\boldsymbol{\tau}_u^*(t)$ which produce the desired angular trajectory. Third we obtain actual angular trajectories using

$$\ddot{\mathbf{q}} = \mathbf{M}(\mathbf{q})^{-1} (\boldsymbol{\tau}_u^* + \boldsymbol{\tau}_e + \boldsymbol{\tau}_c - \mathbf{B}(\mathbf{q} - \mathbf{q}^*) - \mathbf{c}(\mathbf{q}, \dot{\mathbf{q}})) \quad (3.4)$$

where $\boldsymbol{\tau}_c$ is a compensation torque built by adaptation, and \mathbf{B} a feedback gain along the desired trajectory ($\mathbf{B} = 20\mathbf{I}_2$ Nm/rad, where \mathbf{I}_2 is the 2×2 identity matrix). The four conditions are: baseline, $\boldsymbol{\tau}_e = \boldsymbol{\tau}_e^0$ and $\boldsymbol{\tau}_c = \mathbf{0}_2$; before-effect, $\boldsymbol{\tau}_e = \boldsymbol{\tau}_e^\phi$ and $\boldsymbol{\tau}_c = \mathbf{0}_2$; adapted, $\boldsymbol{\tau}_e = \boldsymbol{\tau}_e^\phi$ and $\boldsymbol{\tau}_c = \boldsymbol{\tau}_e^{-\phi} = -\boldsymbol{\tau}_e^\phi$; after-effect, $\boldsymbol{\tau}_e = \boldsymbol{\tau}_e^0 = \mathbf{0}_2$ and $\boldsymbol{\tau}_c = \boldsymbol{\tau}_e^{-\phi} = -\boldsymbol{\tau}_e^\phi$.

3.2.3 Reoptimization model

The reoptimization model is an extension of the model described in Izawa et al. [2008]. The control torque $\boldsymbol{\tau}_u = \begin{bmatrix} \tau_u^{sh} \\ \tau_u^{el} \end{bmatrix}$ is derived from a control input $\mathbf{u} = \begin{bmatrix} u^{sh} \\ u^{el} \end{bmatrix}$ according to Eq. 2.10.

We rewrite the dynamics (Eq. 2.16) as $\dot{\mathbf{x}} = \mathbf{F}_0(\mathbf{x}) + \mathbf{n}_{dyn}$ for the unperturbed dynamics or $\dot{\mathbf{x}} = \mathbf{F}_\phi(\mathbf{x}) + \mathbf{n}_{dyn}$ for the perturbed dynamics, where \mathbf{n}_{dyn} is additive noise on the dynamics. We formulate an optimal feedback control

problem for this dynamics as a search for a control policy $u(t)$ to reach a goal $\mathbf{x}^\# = [q_{sh}^\#, q_{el}^\#, \dot{q}_{sh}^\#, \dot{q}_{sh}^\#, \alpha_{sh}^\#, \alpha_{el}^\#, \epsilon_{sh}^\#, \epsilon_{sh}^\#]$ while minimizing the cost

$$\mathcal{J}_F = \int_t^{t+T_H} (\mathbf{u}^T \mathbf{u} + (\mathbf{x} - \mathbf{x}^\#)^T \mathbf{Q}(t)(\mathbf{x} - \mathbf{x}^\#)) dt \quad (3.5)$$

where

$$\begin{aligned} \mathbf{Q}(t) &= \mathbf{0}_{8,8} && \text{for } t < t_H - \delta \\ \mathbf{Q}(t) &= \mathbf{diag}([w_q^\#, w_q^\#, w_{\dot{q}}^\#, w_{\dot{q}}^\#, w_\alpha^\#, w_\alpha^\#, w_\epsilon^\#, w_\epsilon^\#]) && \text{for } t \geq t_H - \delta \end{aligned} \quad (3.6)$$

where $w_q^\#, w_{\dot{q}}^\#, w_\alpha^\#, w_\epsilon^\#$ are cost parameters (exponent $\#$ indicates the parameter depends on the goal), \mathbf{F} is either \mathbf{F}_0 or \mathbf{F}_ϕ to indicate whether optimization applies to the unperturbed or the perturbed dynamics, T_H is the planning horizon [Guigon, 2022], and $\mathbf{diag}[\]$ indicates the diagonal matrix with listed values on the diagonal. Note that the goal formulation as an objective here is but an artifact from optimization computation. Weights were set with high values to ensure the goal as a constraint, non subjected to compromises (see 2.6.1).

In Izawa et al. [2008], optimization runs on a fixed duration (0.5 s) and thus cannot be used to simulate before-effect and after-effect conditions which require flexible time to produce online movement corrections. Control with a planning horizon offers an efficient solution to time flexibility as at any time and in any changing situation due to a perturbation there always remains the duration of a planning horizon to reach designated goals [Guigon, 2022]. The initial boundary condition for the optimization process is given by $\mathbf{x}(t) = \hat{\mathbf{x}}(t)$, where $\hat{\mathbf{x}}(t)$ is the estimated value of \mathbf{x} provided by an optimal state estimator using forward modeling and delayed sensory feedback with delay Δ [Guigon et al., 2008; Guigon, 2022]. The state estimation is given by

$$\hat{\mathbf{x}}(t) = \hat{\mathbf{F}}(\hat{\mathbf{x}}(t)\mathbf{u}(t)) + \mathbf{K}(t)(\mathbf{y}(t - \Delta) - \mathbf{H}\hat{\mathbf{x}}(t)) \quad (3.7)$$

where $\hat{\mathbf{F}}$ is the dynamic considered for estimation which is either \mathbf{F}_0 or \mathbf{F}_ϕ (see below), $\mathbf{H} = [\mathbf{I}_4 \mathbf{0}_4]$ is a 4×8 observation matrix (\mathbf{I}_4 is the 4×4 identity and $\mathbf{0}_4$ the 4×4 null matrix), indicating that only the position and velocity are observed, $\mathbf{K}(t)$ the Kalman gain and

$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{n}_{obs} \quad (3.8)$$

where \mathbf{n}_{obs} is additive observation noise. The Kalman gain is given by

$$\mathbf{K}(t) = \mathbf{A}(t)\mathbf{P}(t)\mathbf{H}^T (\mathbf{H}\mathbf{P}(t)\mathbf{H}^T + \mathbf{\Omega}^\omega)^{-1} \quad (3.9)$$

where

$$\mathbf{A}(t) = \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \quad (3.10)$$

and

$$\mathbf{P}(t + \delta) = \mathbf{\Omega}^\xi + (\mathbf{A}(t) - \mathbf{K}(t)\mathbf{H})\mathbf{P}(t)\mathbf{A}(t)^T \quad (3.11)$$

where δ is the integration timestep, $\mathbf{\Omega}^\omega$ the covariance matrix of observation (sensory) noise \mathbf{n}_{obs} (4-dimensional, zero-mean, Gaussian random vector) and $\mathbf{\Omega}^\xi$ the covariance matrix of dynamic (motor) noise \mathbf{n}_{dyn} (8-dimensional, zero-mean, Gaussian random vector). We take

$$\mathbf{\Omega}^\omega = \sigma_\omega \times \mathbf{diag}[1, 1, 10, 10] \quad (3.12)$$

and

$$\mathbf{\Omega}^\xi = \sigma_\xi \times \mathbf{diag}[1, 1, 10, 10, 100, 100, 1000, 1000] \quad (3.13)$$

where σ_ω and σ_ξ are the variance of sensory and motor noise, respectively. The state estimator is formulated to be optimal taking into account the feedback delay as explained in Appendix B.

To control movement duration, the goal \mathbf{x}_G is updated every T_G within a series of successive intermediate goals (via-points) $S = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$ with $\mathbf{x}_n = \mathbf{x}^*$, i.e.:

- $\mathbf{x}^\# = \mathbf{x}_0$ at $t = 0$,
- $\mathbf{x}^\# = \mathbf{x}_1$ at time $t = T_G$,
- ...,
- $\mathbf{x}^\# = \mathbf{x}_n = \mathbf{x}_*$ at time $t = nT_G$,

where \mathbf{x}^* is the final goal of the movement [Guigon et al., 2019; Guigon, 2022]. Note that movements are not bounded to end at a preset time. For anytime $t \geq nT_G$, the goal to reach will remain \mathbf{x}^* , allowing for endtime adjustments for perturbation compensation. The four conditions are:

- **baseline:** $\tau_e = \tau_e^0 = \mathbf{0}$, $\mathcal{J} = \mathcal{J}_{F_0}$, $\hat{\mathbf{F}} = \mathbf{F}_0$;
- **before-effect:** $\tau_e = \tau_e^\phi$, $\mathcal{J} = \mathcal{J}_{F_0}$ (the trajectory is planned based on the unperturbed dynamics but executed against a perturbation), $\hat{\mathbf{F}} = \mathbf{F}_0$ (the estimator is unaware of the perturbation);
- **adapted:** $\tau_e = \tau_e^\phi$, $\mathcal{J} = \mathcal{J}_{F_\phi}$ (the trajectory is planned based on the perturbed dynamics and executed against a perturbation), $\hat{\mathbf{F}} = \mathbf{F}_\phi$ (the estimator is tuned to the perturbed dynamics);
- **after-effect:** $\tau_e = \tau_e^0 = \mathbf{0}$, $\mathcal{J} = \mathcal{J}_{F_\phi}$ (the trajectory is planned based on the perturbed dynamics but executed in the absence of the perturbation), $\hat{\mathbf{F}} = \mathbf{F}_\phi$ (the estimator remains tuned to the perturbed dynamics).

Parameter	Value	Unit
Δ	0.12	s
δ	0.01	s
σ_ω	1	-
σ_ξ	1	-
T_H	0.28	s
T_G	0.13	s
w_q^*	10	-
$w_{\dot{q}}^*$	0.1	-
w_α^*	0.01	-
w_ϵ^*	0.01	-
$w_q^\#$	w_q^*	-
$w_{\dot{q}}^\#$	$w_{\dot{q}}^*$	-
$w_\alpha^\#$	0	-
$w_\epsilon^\#$	0	-

TABLE 3.1: Simulation and optimization parameters

The same series of via-points S is used in all the conditions. The fact that the estimator becomes adapted to the perturbed dynamics is consistent with experimental observations [Flanagan et al., 2003; Davidson and Wolpert, 2005].

Parameters are given in Tab. 3.1. # marks intermediate goals ($\mathbf{x}^\# \neq \mathbf{x}^*$), for which only the position is constrained. The final goal state is

$$\mathbf{x}^* = [60.7^\circ, 60^\circ, 0, 0, 0, 0, 0, 0],$$

i.e. the final shoulder and elbow angles corresponding to a 0.1 m forward displacement, zero final velocity, activation and excitation.

3.2.4 Redirection model

The redirection model is taken from Guigon [2022] and customized to the current formulation. The baseline and before-effect conditions are the same as for the reoptimization model. In the adapted and after-effect conditions, the cost function is \mathcal{J}_{F_0} , i.e. the controller is unaware of the perturbation, but the series of via-points S used in the baseline and before-effect conditions is replaced by a new series of via-points S' which defines adaptation.

As adaptation in this model does not rely on the knowledge of the "current" dynamic, the estimator here remains unaware of the perturbation: $\hat{\mathbf{F}} = \mathbf{F}_0$.

3.2.5 Numerical solution

The reoptimization and redirection models were simulated numerically using the iLQR method proposed by [Li and Todorov \[2004\]](#).

3.3 Experiment

3.3.1 Ethics statement

The experiment was approved by Comité d’Ethique de La Recherche at Sorbonne Université (CER-2021-112). Participants signed a consent form prior to participating in the experiment and in accordance with the ethical guidelines of Sorbonne Université and in accordance with the Declaration of Helsinki.

3.3.2 Participants

Twenty-two volunteers (20–30 yr old, 8 female) participated in the behavioral experiment. According to the Edinburgh Protocol of handedness [[Oldfield, 1971](#)], 18 were right-handed, 2 left-handed and 2 ambidextrous. They had no known neurological disorders and normal or corrected to normal vision and they were uninformed of the methodological details of the experiment.

3.3.3 Apparatus

Participants were seated on a chair and used their dominant hand (their most comfortable hand for ambidextrous participants) to move the handle of a robotic arm programmed to constrain the displacement of the hand in an horizontal plane and apply force perturbations along this plane. The robot used was a Virtuose 6D¹ HAPTIONTM and was chosen for its high back-drivability [[Perret and Vercruysse, 2014](#)]. Task instructions, feedback information, and visual feedback of hand displacement were provided on a monitor placed vertically in front of the participant. The flow of the task was controlled by a personal computer running Windows 7 (Microsoft Corporation, USA). The 3D position of the robot was recorded at 1000 Hz, and stored on the computer for offline processing and analysis using custom written Matlab scripts (Mathworks, Natick, MA, USA).

¹<https://www.haption.com/en/products-en/virtuose-6d-en.html%7D%7D>

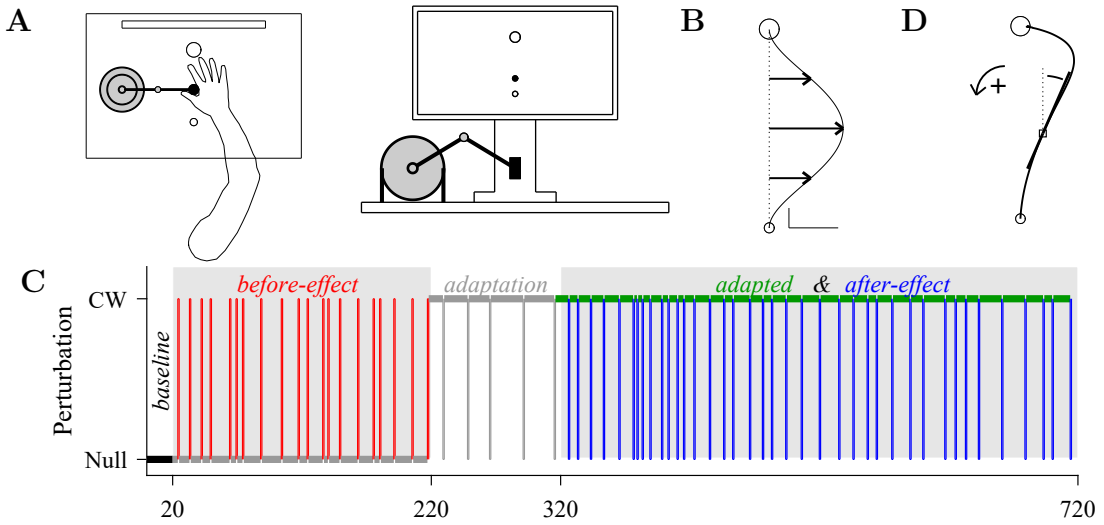


FIGURE 3.1: **Description of the experiment.** **A.** Experimental setup. (left) Top view. The small open circle is the start position and the large open circle the target position. The black circle is the robot handle. The elongated open rectangle is a top view of a monitor. (right) Front view. The start position, target position, and visual feedback of hand position (black circle) are shown on the monitor. The black rectangle is the robot handle. The scales are not respected. **B.** Simulated velocity-dependent force field. A minimum-jerk velocity profile with a 0.3 m/s peak was multiplied by a 5 N/m force field. Vertical scale: 1 cm. Horizontal scale: 1 N. **C.** Experimental protocol. The force field level (null or CW) is indicated by the horizontal black (baseline block), green (adaptation block) or gray (before-effect and after-effect blocks) thick line segments. The vertical line segments indicate catch trials: unexpected CW force field in the before-effect block (red); unexpected null force field in the adaptation block (gray) and in the after-effect block (blue). Only the colored trials (black: baseline; red: before-effect; green: adapted; blue: after-effect) were analyzed. **D.** Graphical definition of the trajectory angle. At one point along the trajectory (open square), the trajectory angle is the angle between start position/target position direction (dashed line) and the tangent to the trajectory (thick line).

3.3.4 Experimental procedure

The participants were asked to make forward and return arm reaching movements with a constant moderate tempo (peak velocity between 0.25 and 0.35 m/s) (see Chapter 3.2). The movement was performed from a start position to a target position located 0.1 m away using visual information displayed on the monitor (start position: 0.6-cm diameter blue circle; target position: 1-cm diameter black circle; moving cursor: 0.3-cm diameter red dot - see Fig. 3.1A). To start a trial, the participants placed the cursor at the start position and began to move when ready. The peak velocity was computed on line in order to inform the participant about his/her velocity. Once the cursor stopped inside the target circle (Cartesian velocity < 0.01 m/s), the circle colored in blue for too slow movements (peak velocity along task direction < 0.25 m/s), in red for too fast movements (peak velocity along task direction > 0.35 m/s) or in green for movements of proper

speed. There was no constraint of accuracy except to stop into the target circle. The return movement was unconstrained except the need to stop inside the start circle to start the next trial (velocity < 0.01 m/s).

On some trials, a velocity-dependent force field (3.1B) was applied during the forward displacement as defined by Eq. 3.2 with $\psi = \frac{\pi}{2}$ and $\phi = \pm 10$ Ns/m. The force field was CCW ($\phi > 0$) for half of the participants. The participants performed four blocks of trials (Fig. 3.1C): block 1 (20 trials, 100% vs 0% of null field vs force field), block 2 (200 trials, 90% vs 10%), block 3 (100 trials, 5% vs 95%), block 4 (min 150 trials, max 400 trials, 10% vs 90%). The last block involved many trials to maximize the number of recordings of after-effect trajectories. Yet the participants were offered the possibility to stop the experiment after 150 trials if they felt exhausted or bored. A pause was proposed between each block. The instructions given to the participants were the following:

"Perform forward reaching movements to the target at a proper speed as indicated by the color code (blue, green, red) and return to the starting position at your own pace. Make a brief pause in the target and at the starting position and avoid rhythmic back and forth movements. Sometimes the robot may perturb your movement. Whenever it happens, continue to obey to the task instructions."

At the start of recording, the participants were already familiar with the robot as they performed unrelated preliminary trials of force and position measurements. The robot was transparent and easy to manipulate.

3.3.5 Data processing and analysis

Raw data were used to obtain the planar trajectory of the hand for each trial. A symmetry relative to the start position/target position axis was applied to the trajectories of participants receiving a CCW perturbation. Velocity and acceleration were calculated numerically from the two-sample difference of the position and velocity signals, respectively. Position, velocity and acceleration were filtered with a fourth-order Butterworth low-pass filter with a cutoff at 10 Hz. Valid trials were detected by a peak velocity between 0.25 and 0.35 m/s in the forward part of the movement. For each valid trial, the forward trajectory was extracted by detection of movement onset and offset with a velocity threshold of 0.01 m/s and two time-varying quantities were calculated: (1) the angle (counted positive in the CCW direction) of the tangent to the trajectory relative to the line between the start position and the target position; (2) the time derivative of this angle which is closely related to the curvature of the trajectory. The valid trials were divided into four categories: baseline (trials of block 1), before-effect (perturbed trials of block 2), adapted (perturbed trials of block 4), after-effect (unperturbed trials of block 4). For each category, mean trajectory, mean angle and mean angle derivative were calculated over the trials.

The rationale for the choice of the filter cutoff frequency is the following. A power spectrum analysis was performed on the unfiltered timeseries using a specific method for short-duration timeseries [de Grosbois and Tremblay, 2016]. The results are shown in Fig. C.13 for velocity, acceleration and jerk pooled across trials and participants, separately for each category (baseline, before-effect, adapted, after-effect). Much of the power was below 10 Hz.

A classical Student's t-test was used to assess the sign of the trajectory angle derivative ($H_0 : = 0$ vs $H_1 : \neq 0$). A p -value < 0.05 was taken to support H_1 . A p -value > 0.05 indicated that we could not reject H_0 . To assess the status of H_0 vs H_1 in the latter case, we calculated the Bayes factor bf_{10} which is the ratio between the likelihood of the data under H_1 and H_0 [Rouder et al., 2009]. Bayes factors were interpreted according to the following table: $1 < bf_{10} < 3$: anecdotal; $bf_{10} > 3$: substantial. The Bayes factors were calculated with the Matlab toolbox FieldTrip (<https://www.fieldtriptoolbox.org/>; Oostenveld et al. 2011).

3.4 Results

We designed a force field adaptation experiment with a large number of trials and a small fraction of catch trials (unexpected addition or removal of the force field) to obtain "pure" before-effect and after-effect trajectories uncontaminated by ongoing learning processes [Thoroughman and Shadmehr, 2000]. Twenty two participants were asked to make fast, planar, forward arm reaching movements from a start position to a target position located 0.1 m away in the presence of a null field or a perpendicular CW or CCW velocity-dependent force field (Fig. 3.1A,B). The participants performed four blocks of trials (Fig. 3.1C) and we identified baseline, before-effect, adapted, and after-effect trajectories (see Material and methods, section 3.2). For data analysis, all trajectories were displayed with a CW deviation, i.e. for a CCW perturbation, a vertical symmetry was applied to the trajectories. A trajectory was described by (1) the angle (counted positive in the CCW direction) of its tangent relative to the hand-target direction (Fig. 3.1D); (2) the time derivative of the trajectory angle (see Material and Methods for details - section 3.2).

3.4.1 Predictions

The compensation model makes immediate predictions on the shape of before-effect and after-effect trajectories and corresponding velocity profiles (Fig. C.1). These prediction will not be further considered: they are robust but lack pertinence as the compensation model is not a general model of motor control (see Discussion - section 3.5).

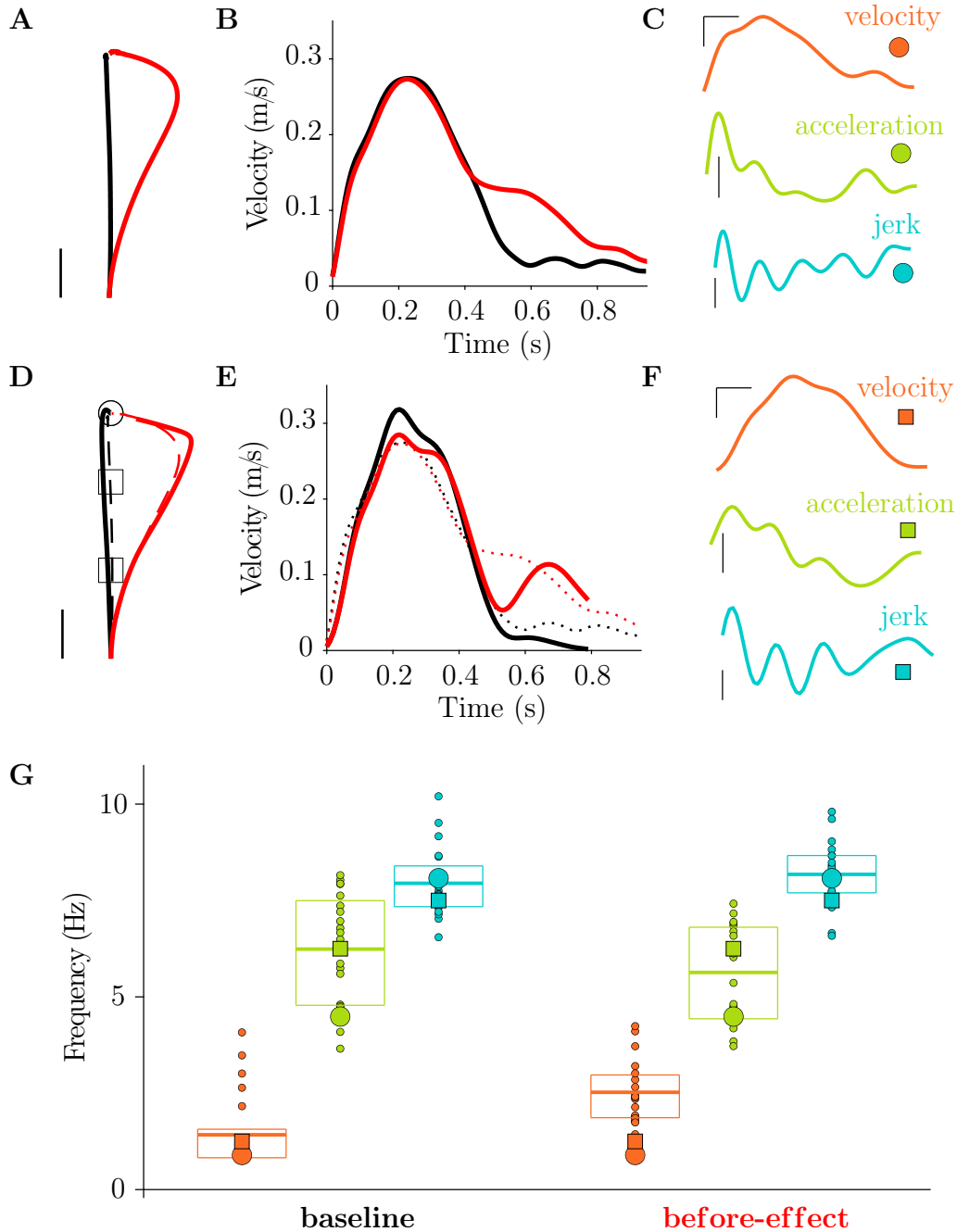


FIGURE 3.2: Model adjustment based on data of participant P7.

A. Mean baseline (black; 17 trials) and before-effect (red; 19 trials) trajectories for P7. Scale: 2 cm. **B.** Mean velocity profiles of baseline and before-effect trajectories for P7. **C.** Velocity (scale: 0.1 m/s), acceleration (scale: 2 m/s²) and jerk (scale: 30 m/s³) profiles for a single baseline trial (P7). Time scale: 0.1 s. **D.** Simulated baseline (plain black) and before-effect (plain red) trajectories compared to experimental trajectories (dashed; data from A). Squares are via-points for the simulated trajectories. Same scale as in A. **E.** Simulated velocity profiles. **F.** Velocity (same as black in E), acceleration and jerk profiles for the simulated baseline trajectory. The profiles have been truncated to match the duration of the trial in C. Same scales as in C. **G.** Peak frequency for velocity (orange), acceleration (light green) and jerk (light blue) profiles for individual trials (small dots). Large circles correspond to the trial in C. Large squares correspond to the simulated trial in F. Thick lines are mean values and boxes indicate 25-75 percentiles.

In order to build precise predictions for the reoptimization model, we proceeded in the following way. We considered one participant (P7) and analyzed detailed characteristics of her motor behavior (Fig. 3.2). We calculated the mean baseline and before-effect trajectories (Fig. 3.2A) and velocity profiles (Fig. 3.2B). For each single trial (e.g. a baseline trial; Fig. 3.2C), we calculated a discrete measure of the frequency content (peak frequency; number of minima+number of maxima/duration/2) of velocity, acceleration and jerk traces. We plotted the peak frequency of all trials for the two types of trial (Fig. 3.2G). These results show that the smoothness of mean trajectories and velocity profiles (Fig. 3.2A,B) is an artifact of averaging widely nonsmooth and variable single trials (Fig. 3.2C,G). Although these observations are not surprising [Vallbo and Wessberg, 1993; Guigon et al., 2019], they cannot be explained by models that produce temporally invariant smooth movements [Flash and Hogan, 1985; Harris and Wolpert, 1998; Todorov and Jordan, 2002]. To circumvent this difficulty, we considered a model which explains the frequency content of movements by the pursuit of intermediate goals (via-points) updated at 8 Hz (see Material and methods, section 3.2; Guigon et al. 2019; Guigon 2022). We searched for a series of via-points S and model parameters that account for experimental paths and velocity profiles of baseline and before-effect trajectories (Fig. 3.2D,E; for a parametric study of the model, see section 3.4.5). The series S contained two intermediate via-points (squares; Fig. 3.2D) at 36% and 72% of the distance to the target in the direction of the target, and the target itself (circle; Fig. 3.2D). Note that we did not search for the "best fit", as all single trials were different (Fig. 3.2C,G). Note also that the intensity of the modeled force field (ϕ) was lower than that of the experimental field (see Discussion). The amplitude and frequency contents of the resulting movement were consistent with the experimental data (Fig. 3.2F,G).

At this stage, the proposed model is appropriate for trajectory formation and online motor control during perturbations and further accounts for many characteristics of motor behavior [Guigon, 2022]. We can now obtain proper predictions for the reoptimization model (Fig. 3.3). The adapted trajectory was not a straight path but an overcompensation (green; Fig. 3.3A) which is consistent with Izawa et al. [2008]. Its velocity profile was close to the baseline velocity (green vs black; Fig. 3.3B). The after-effect trajectory had the expected mirror organization relative to the before-effect trajectory (blue vs red; Fig. 3.3A) and a velocity profile which resembled the before-effect profile (blue vs red; Fig. 3.3B). The mirror effect is quantitatively described in Fig. 3.3C,D. The trajectory angles had opposite monotonic trends for before-effect and after-effect trajectories over the first ~ 0.6 s (blue vs red; Fig. 3.3C) with corresponding changes in the sign of the derivatives (blue vs red; Fig. 3.3D). In the following, we will focus on the early part of the trajectories (0.4 s; dotted boxes in Fig. 3.3C,D; Fig. 3.3E,F) since trajectory averaging for experimental data may produce unreliable results for the late part of the trajectory. Two quantitative observations are relevant: (1) the angle derivative of the before-effect trajectory became positive at 0.29 s (vertical red dashed line; Fig. 3.3F). This result is consistent with experimental data in P7 and across all

the participants (Fig. C.2); (2) the angle derivative of the after-effect trajectory became negative at 0.31 s (vertical blue dashed line; Fig. 3.3F) which means that the derivative is negative 7.5% of the time during the first 0.4 s. For comparison with experimental data, we will use this number rather than the time of change in sign which might not be well defined in the data (e.g. due to multiple changes in sign).

On the one hand, the expected positive sign of the derivative of the after-effect trajectory angle would add support to the reoptimization model. On the other hand, a null or negative derivative would suggest to reject the reoptimization model.

3.4.2 Two representative participants

Results for participant P7 shown in Fig. 3.4 and 3.5 (same format as in Fig. 3.3) followed the typical pattern observed in force-field adaptation experiments [Shadmehr and Mussa-Ivaldi, 1994; Izawa et al., 2008]: 1. The mean baseline trajectory was straight (black; Fig. 3.4A); 2. The mean before-effect trajectory deviated in the direction of the perturbation with a late hook-like correction (red; Fig. 3.4A); 3. The mean adapted trajectory was straighter than the mean before-effect trajectory but not as straight as the baseline trajectory (green; Fig. 3.4A); 4. The mean after-effect trajectory was deviated in the direction opposite to the perturbation (blue; Fig. 3.4A); 5. The velocity profiles had a large initial peak followed by one or more smaller peaks (Fig. 3.4B); 6. Single trial trajectories were variable but consistent with the mean trajectory (Fig. 3.4C); 7. Lateral deviation decayed exponentially across trials ($R^2 = 0.69$; Fig. 3.4D).

To test the reoptimization model, we analyzed the time course of the mean trajectory angle (Fig. 3.5E) and trajectory angle derivative (Fig. 3.5F). As expected, the mean angle of the before-effect trajectory decreased until 0.3 s and then increased (red; Fig. 3.5E,F and inset). The mean angle of the after-effect trajectory was initially approximately constant and then decreased (blue; Fig. 3.5E,F and inset). To assess the statistical significance of this observation, we performed a t-test on the sign of the angle derivative ($H_0 : = 0$ vs $H_1 : \neq 0$; $N=34$ trials) at each timestep. The corresponding p -value was > 0.05 for the first 0.1 s (Fig. 3.5G), which indicates that we cannot reject the hypothesis that the angle derivative is zero. The p -value was < 0.05 after 0.1 s (Fig. 3.5G), which mean that the angle derivative was significantly different from zero and negative. We calculated the Bayes factor bf_{10} for H_1 vs H_0 which indicated that the data were 1 to 5 times more likely under H_0 than under H_1 when $p > 0.05$ (Fig. 3.5H).

A different behavior was observed for participant P5 (Fig. 3.6). The mean before-effect and after-effect trajectories were symmetrically organized as shown in Fig. 3.6A,C,D. A statistical analysis indicated that the angle derivative of the after-effect trajectory was non-zero and positive between 0.1 and 0.3 s following

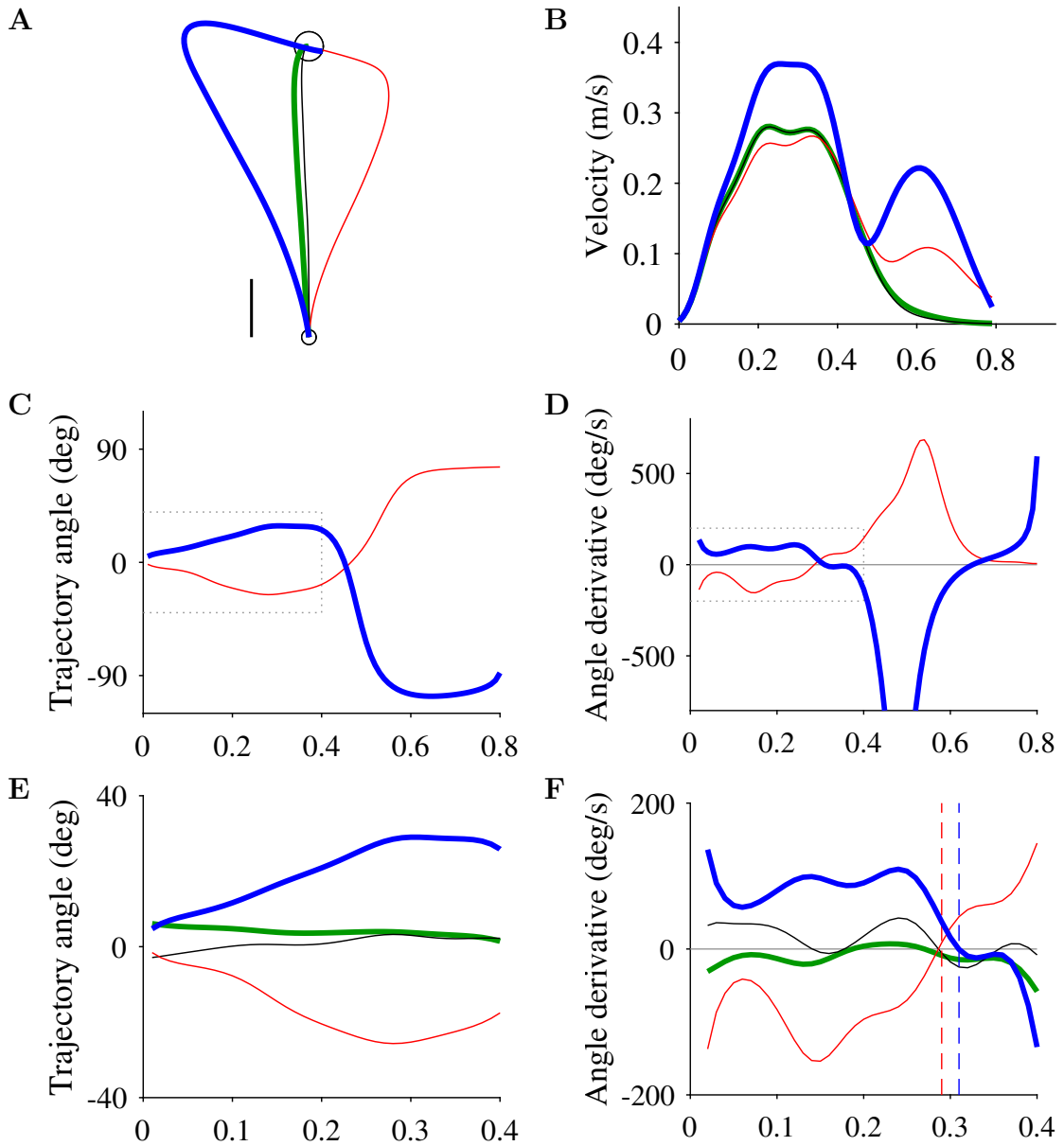


FIGURE 3.3: **Model predictions.** **A.** Simulated adapted (green) and after-effect (blue) trajectories corresponding to simulated baseline (black) and before-effect (red) trajectories shown in Fig. 3.2D and reproduced here with thin lines. Scale: 2 cm. **B.** Simulated velocity profiles. **C.** Trajectory angle for **A**. **D.** Trajectory angle derivative for **A**. **E.** Zoom on trajectory angle (dotted box in **C**). **F.** Zoom on trajectory angle derivative (dotted box in **D**). Vertical dashed lines indicate the time of change in derivative sign.

movement onset (p -value < 0.05 , Fig. 3.6E; $bf_{10} > 3$, Fig. 3.6F). Although the behavior of this participant matches some predictions of the reoptimization model, the experimental path and velocity profile of the after-effect trajectory were different from the predicted path and velocity profile (Fig. 3.6A,B vs Fig. 3.3A,B).

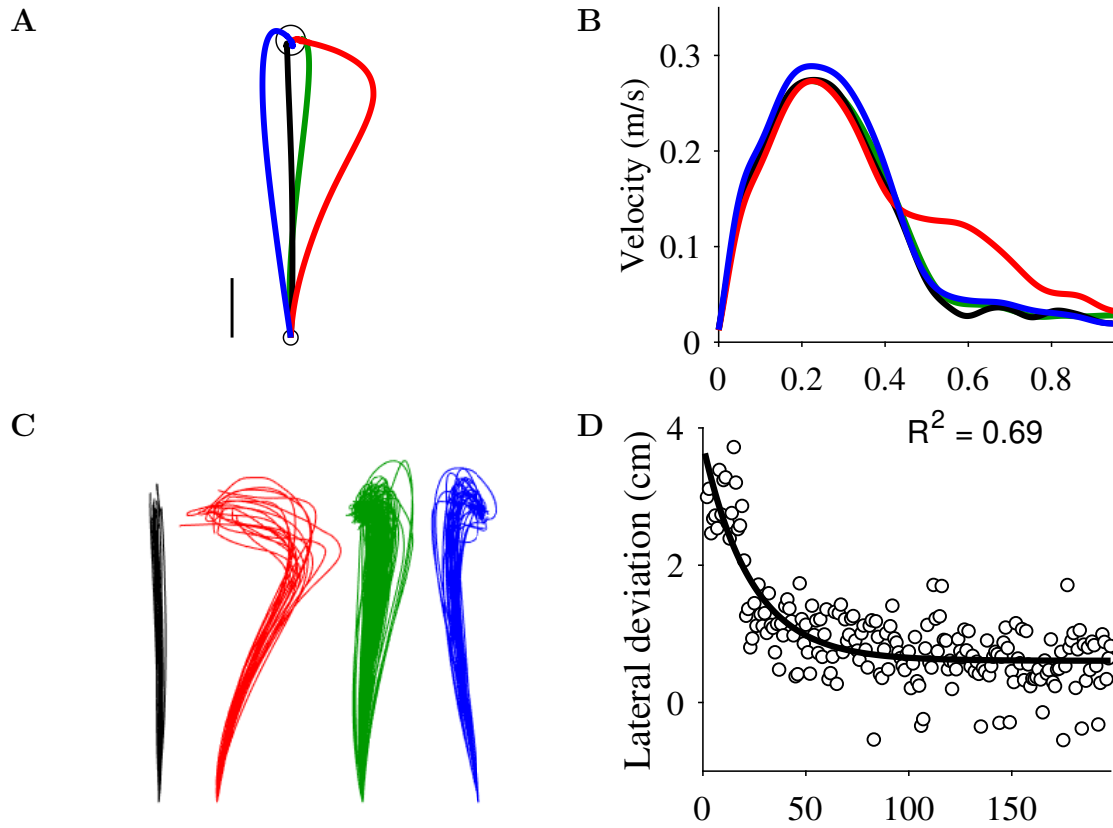


FIGURE 3.4: **Data of participant P7 (1/2).** **A.** Mean trajectories. Scale: 2 cm. **B.** Mean velocity profiles. **C.** Single trials (17, 19, 107, 34 trials, from left to right). Same scale as in **A**. **D.** Changes in lateral deviation (maximum of trajectory deviation from the start position/target position line) with training. The data were taken from **C** (red and green). Fitting an exponential decay is shown.

3.4.3 All participants

For each participant, we calculated the percentage of time over the first 0.4 s during which the angle derivative of the after-effect trajectory was statistically null or negative using both p -values and Bayes factors (see Fig. 3.5G,H). The reoptimization model predicts that this percentage should be around 7.5 (Fig. 3.3F). The experimental percentage was far from the predicted percentage for all the participants (Fig. 3.7A,B). The behavior of ten participants (blue bars; Fig. 3.7) was similar to the behavior of P7 (see Figs. 3.5 and 3.5; Figs. C.3, C.4 and C.5). The behavior of eight participants (dark blue bars; Fig. 3.7) was similar to the behavior of P5 (see Fig. 3.6). The quantitative results for these participants are shown exhaustively in Fig. C.6. It can be observed that the behavior of these participants is rather homogeneous and differs qualitatively and quantitatively from the predicted behavior in terms of path and velocity profile.

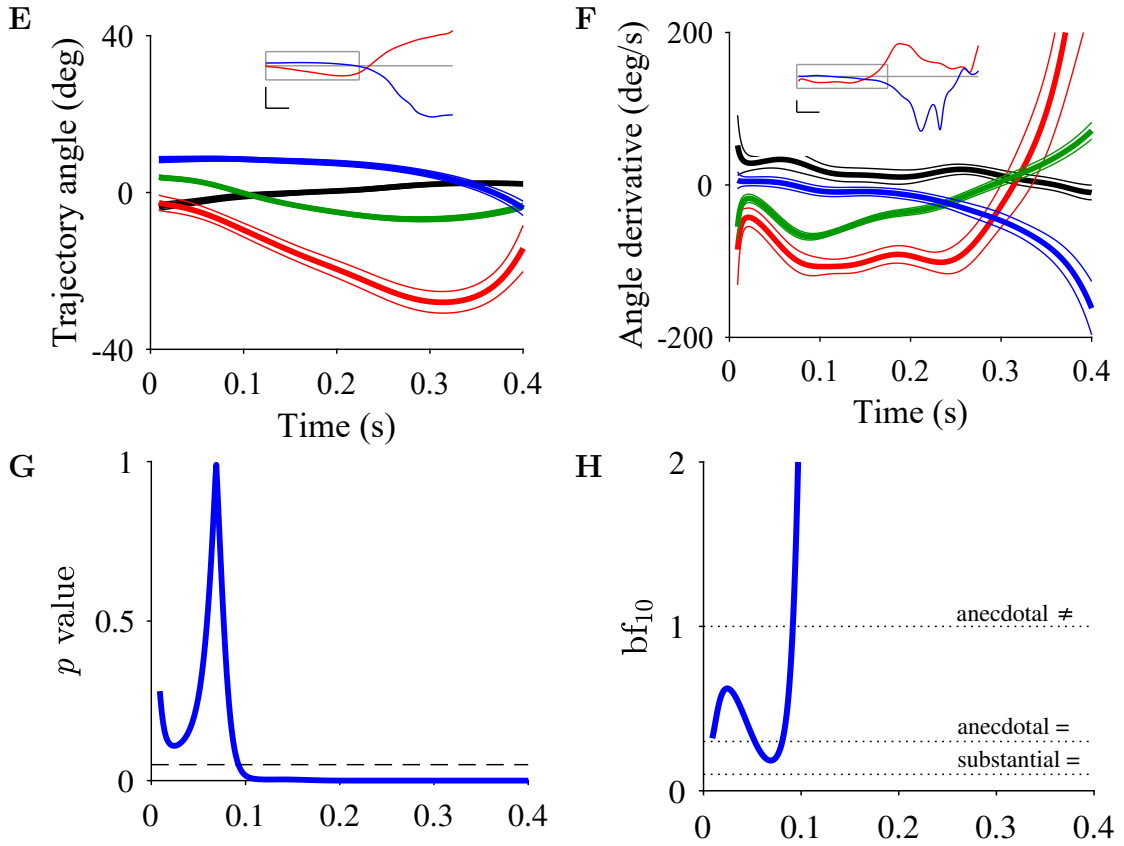


FIGURE 3.5: **Data of participant P7 (2/2).** **E.** Mean trajectory angle over the first 0.4 s with the 95% confidence interval. Inset: mean trajectory angle over the first 0.8 s; the box indicates the 0-0.4 s window. Scale: 0.1 s, 60 deg. **F.** Same as **E** for mean trajectory angle derivative. Scale: 0.1 s, 200 deg/s. **G.** p -value of a test $\neq 0$ vs $= 0$ for trajectory angle derivative in **F**. The dotted line indicates 0.05. **H.** Bayes factor for the test $\neq 0$ vs $= 0$. The dotted lines delimitate regions of interpretation of Bayes factors.

The two remaining participants (light blue bars; Fig. 3.7) failed to improve their behavior with training (Fig. C.8).

3.4.4 Redirection model

We simulated adaptation through redirection using an ad-hoc series of via-points S' to obtain an adapted trajectory (green; Fig. 3.8A, left) which resembles a real adapted trajectory (green; Fig. 3.4C). We generated small variations around these via-points to obtain an ensemble of adapted trajectories (green; Fig. 3.8A, right). The corresponding after-effect trajectories (blue; Fig. 3.8A) resembled real after-effect trajectories (blue; Fig. 3.4C). The velocity profiles, the trajectory angles and the trajectory angle derivatives were consistent (Fig. 3.8B,C,D). As expected, we observed an absence of mirror effect, the angle derivative of after-effect trajectories being positive $< 10\%$ of the time in the first 0.4 s (Fig. 3.8D).

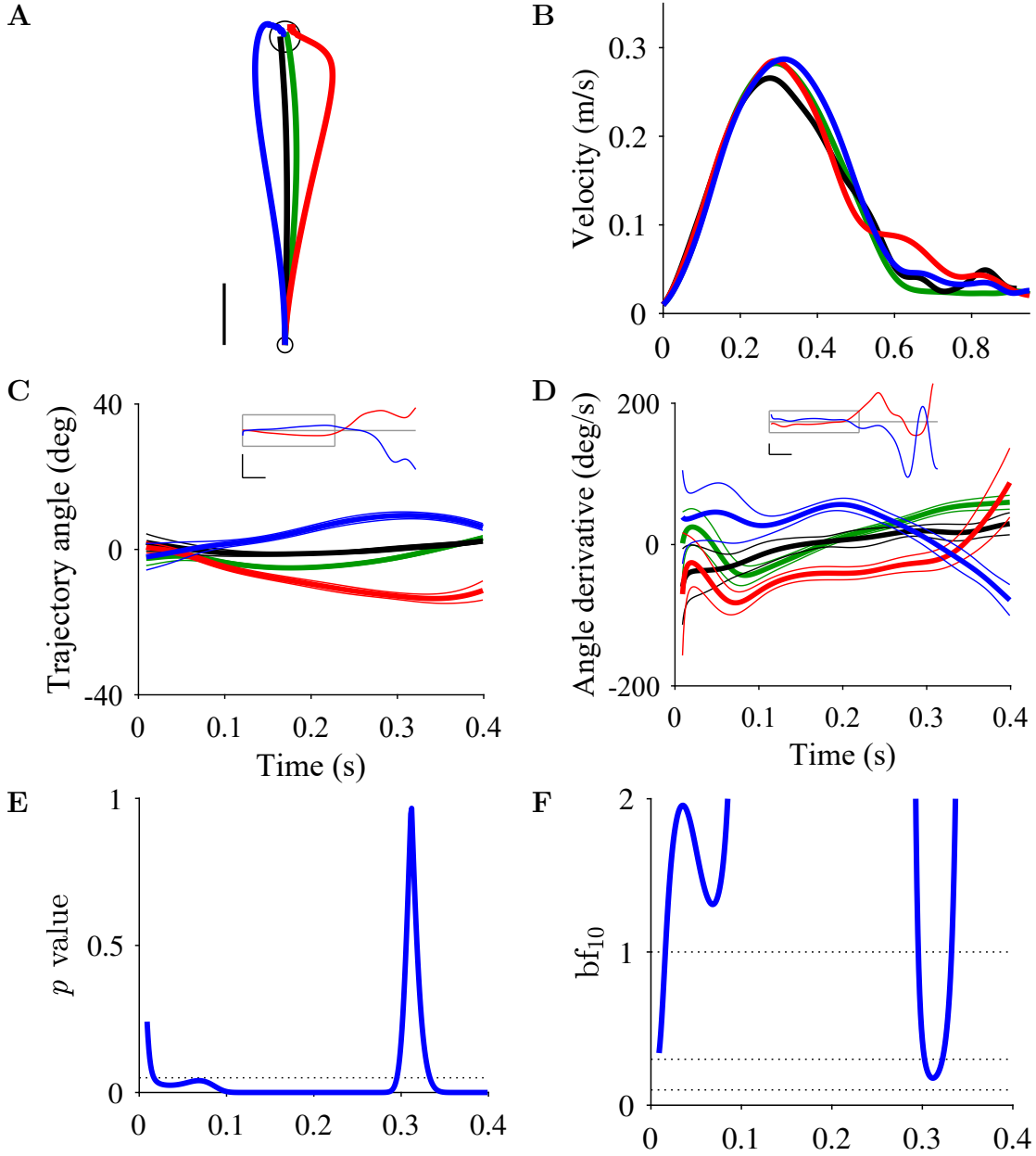


FIGURE 3.6: **Data of participant P5.** Same organization as Fig. 3.4 and 3.5 with C, D, E, F corresponding to E, F, G, H.

3.4.5 Parametric study of the reoptimization model

The conclusions of this study are highly dependent on the predictions of the reoptimization model. As the model contains parameters, it is important to understand the influence of these parameters on the proposed predictions. We explored the role of 3 parameters: the feedback delay Δ , the noise ratio σ^ξ/σ^ω of motor to sensory noise variance used in the estimator, and the muscle gain ratio between shoulder and elbow g_{sh}/g_{el} . The first two parameters modulate how sensory information participates to state estimation (see Appendix B). The third parameter calibrates the contribution of shoulder and elbow torques to coordination. The trajectories,

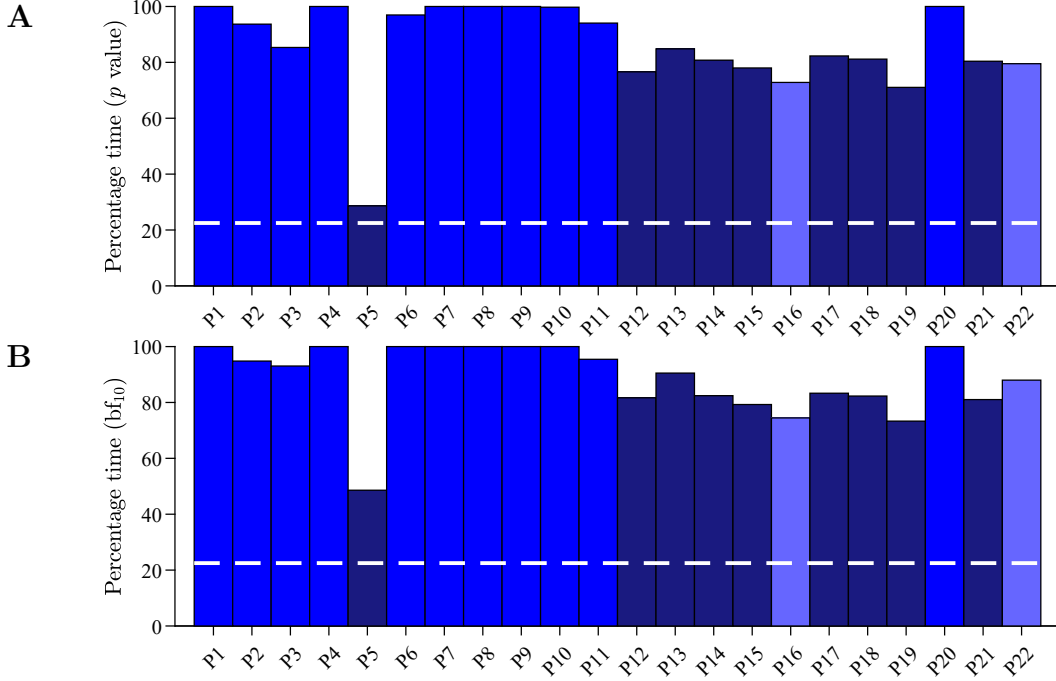


FIGURE 3.7: **All participants.** **A.** Percentage of time over the first 0.4 s during which the angle derivative of the after-effect trajectory is statistically null or negative using t-test p -values. Dashed white line is model prediction. Color code for the participants: blue, behavior incompatible with the reoptimization model; dark blue, behavior partially compatible with the reoptimization model; light blue, behavior with no effect of training. **B.** Same as **A** using Bayes factor.

velocity profiles, trajectory angles and trajectory angle derivatives were consistent across variations of these parameters (Figs. C.9, C.10, C.11). The mirror organization between before-effect and after-effect trajectories was robustly observed. We note that the time at which the angle derivative of the before-effect trajectory becomes negative (Fig. 3.3F) varied with the feedback delay (Fig. C.9D) and the torque ratio (Fig. C.11D).

A set of parameters ($w_q^\#$, $w_{\dot{q}}^\#$, $w_\alpha^\#$, $w_\epsilon^\#$) specify the boundary conditions at the via-points, i.e. whether position, velocity, activation and excitation are forced to take specified values. The predictions were built with $w_q^\# \neq 0$ and $w_{\dot{q}}^\# = w_\alpha^\# = w_\epsilon^\# = 0$ (constraints only on position). The role of these parameters is illustrated in Fig. C.12. Although they have little influence on acceleration (Fig. C.12A), they have a clear effect on jerk, with a higher level of jerk whenever the constraints are not exclusively on position (Fig. C.12B). Only the lower level of jerk (constraints on position) is consistent with experimental data (Fig. 3.3C,D).

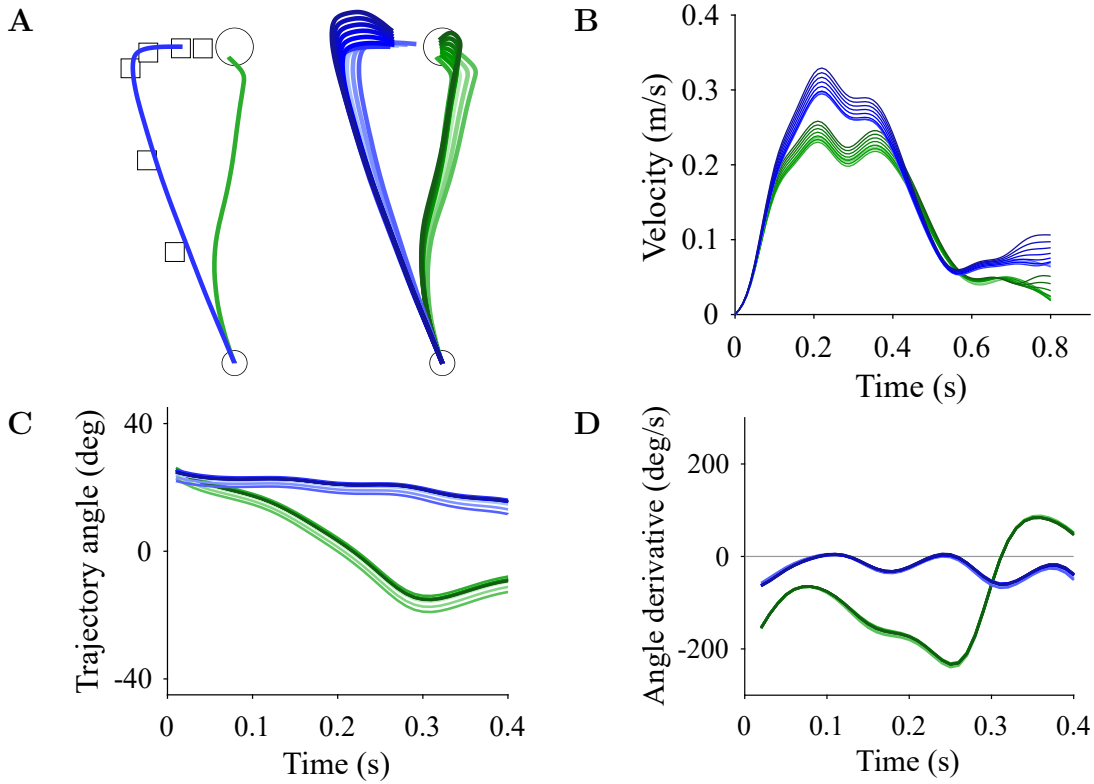


FIGURE 3.8: **Simulations of the redirection model.** **A.** (left) Adapted (green) and after-effect (blue) trajectories corresponding to a series of via-points (squares). (right) Multiple adapted and after-effect trajectories corresponding to variations of the series of via-points. **B.** Corresponding velocity profiles. **C.** Corresponding trajectory angles. **D.** Corresponding trajectory angle derivatives.

3.5 Discussion

Classical computational models of motor adaptation assume that learning occurs at the action selection level [Shadmehr and Mussa-Ivaldi, 1994; Izawa et al., 2008]. We derived predictions for these models which show that after-effect trajectories following adaptation to a velocity-dependent force field are close to a mirror of before-effect trajectories (Figs. 3.6, C.1). Experimental data collected in twenty-two participants did not follow these predictions (Figs. 3.4, 3.4, 3.6, 3.7, Figs. C.3, C.4, C.5 and C.6). We discuss implications and limitations of these observations.

Hundreds of force field adaptation studies have been performed since the seminal study of Shadmehr and Mussa-Ivaldi [1994], but none of them have quantitatively documented properties of after-effect trajectories. Although many published figures could informally be used to gain qualitative information on before-effect and after-effect trajectories and their differences (to mention a few: figs. 3 and 7 in Lackner and DiZio 1994; fig. 13 in Shadmehr and Mussa-Ivaldi 1994; see Introduction for other references), they are not sufficient to draw a firm conclusion. The lack of a specific interest for after-effect trajectories might be related to the prevalent view in computational motor control that adaptation results from

changes at the control level and properties of after-effect trajectories are a direct by-product of these changes [Shadmehr and Mussa-Ivaldi, 1994; Izawa et al., 2008]. In this framework, an after-effect trajectory reflects a kind of compensation that attempts to negate the forces induced by the applied force field and thus inherits from properties of the force field. As velocity along the trajectory increases, the compensation force increases and the after-effect trajectory curves away from the baseline trajectory in mirror with the before-effect trajectory. Both the compensation and the reoptimization models [Shadmehr and Mussa-Ivaldi, 1994; Izawa et al., 2008] obey to this scenario (Figs. 3.3, C.1). The data collected in this study are incompatible with these models. The expected mirror organization was completely absent in 11/22 participants, and present but far smaller than expected in the remaining participants (Fig. 3.7). Yet, we did not average data across the participants, provided single participant analyses (Figs. C.3, C.4, C.5, C.6), and made the raw data available to give a chance to any possible interpretation.

At this stage, it is interesting to consider the implication of adaptation at the action selection level. It would mean that each new adaptation requires the building of a dedicated control policy which inherits from general motor abilities (e.g. when we hold a manipulandum in a force field task, we do not need to relearn motor coordination from scratch), but remains insulated from general and specific skills (e.g. it does not interfere with our ability to walk or to play the piano). The corresponding motor architecture comes with a heavy computational burden to build, maintain, update, share and exploit each learned ability in each proper context. A solution based on the storage of multiple controllers has been proposed [Wolpert and Kawato, 1998], but does not address the associated computational burden. Furthermore, its proposed implementation through the huge computational power of the cerebellum is probably incompatible with the predominant sensory nature of cerebellar processing [Gao et al., 1996; Nixon, 2003]. Besides these computational issues, interlimb transfer of force field adaptation [Criscimagna-Hemminger et al., 2003; Malfait and Ostry, 2004] and adaptation by mere observation [Mattar and Gribble, 2005] are also inconsistent with adaptation at the action selection level.

An alternative view is that adaptation to a novel motor environment relies on changes at the goal selection level (redirection model), i.e. aiming toward appropriately chosen successive spatial goals (e.g. via-points) would mimic adaptation and after-effects in a force field (Fig. 3.8). In the proposed scenario, the "memory" of the perturbation is not a continuous mapping between state and force but a discrete set of via-points. This scenario can be reproduced in a simulation by a three-step target jump protocol involving two intermediate via-points and the actual target where the via-points have been chosen by hand to obtain an adapted trajectory which is close to the baseline trajectory (Fig. 3.8). Our data are not incompatible with this view. Yet, they cannot be said to support it since the proposed adaptation mechanism remains uncompletely specified: there is no principled recipe to select or learn to select proper via-points. Interestingly, the redirection model

is versatile enough to account for the whole dataset. The absence of mirror effect in some participants and its presence in others can be explained by a specific configuration of via-points (Fig. 3.8A,B).

The proposed mechanism should not be linked to the kind of explicit, cognitive strategy that can be used to compensate for a visuomotor rotation simply by changing the aiming direction [Mazzoni and Krakauer, 2006]. Here, the proper choice of via-points is conceived as the outcome of a learning process. What drives the learning process is not specified, but could possibly be cast in a cost/benefit framework (e.g. effort vs accuracy). We have no information on the explicit or implicit nature of the mechanism. Yet, in post-experiment interviews, we noted that several participants believed that the after-effect deviations were due to a force field and not to their own behavior, which suggests that they probably had little conscious control on their behavior once adapted to the perturbation.

A possible concern is the seemingly adhoc nature of the proposed scenario. However this scenario is derived from a consistent theoretical construct which accounts for the production of fast and slow movements, the distinction between discrete and rhythmic movements, the ubiquity of isochronous behaviors, the existence of scaling laws, power laws and speed-accuracy tradeoffs [Guigon, 2022]. Thus motor control would involve a unique, general-purpose, task-independent action selection mechanism (controller) and each task would have its own representation defined as a series of successive intermediate goals updated at a fixed frequency and pursued at a fixed horizon. In this framework, a skilled movement is not defined by the operation of a dedicated, "skilled" controller, but the use of a dedicated, "skilled" task representation. Consider the following example. It is probable that none of the readers of this article have the tennis skills of a top ranked tennis player. Yet they do not mostly look clumsy in activities of daily living, probably have their own motor skills, and should be able to sometimes produce a magnificent backhand worthy of a good tennis player. Accordingly, the main difference between a novice and an expert would not be found at a control level, e.g. a difference in mastering coordination, but at the level of task representation, i.e. how successive goals are consistently set to properly elicit and guide actions. Our view of motor adaptation can effortlessly be cast in this framework. Interestingly, the computational burden associated with the storage of multiple controllers is significantly alleviated with the storage of multiple task representations. Task representations are discrete sets which are much more frugal in neural resources than continuous mappings. Furthermore, they can be scaled spatially and shared between effectors, accounting for motor equivalence. Issues related to the stability and flexibility of skills appear much less enigmatic when skills are conceived as task representations rather than controllers.

The main limitation of this study is its strong reliance on computational modeling. Our conclusions are based on the divergence between experimental data and predictions of the compensation/reoptimization models. So it is fundamental to check that the proposed predictions are both robust and realistic. As far as

robustness is concerned, there is no difficulty with the compensation model which is well-formulated and easy to simulate. However, this model has little general relevance for motor control as it does not provide solutions to central problems such as trajectory formation and coordination [Todorov and Jordan, 2002]. For this reason, we have not pursued comparisons with this model. The reoptimization model is based on optimal feedback control [Todorov and Jordan, 2002] and has been updated here to account for proper online feedback control [Guigon, 2022]. It generates movements with realistic trajectories, velocity profiles, and amplitude and frequency contents (Fig. 3.2). We have shown that its predictions are robust to parameter changes.

An unsettled and interesting issue is related to the intensity of the applied force field. The predictions were obtained with a 2-Ns/m force field as compared to the 10-Ns/m field of the experiment. It should be noted that these force field are applied to slightly different systems: in the simulations, the perturbation is a pure force acting on the modeled hand while the experiment relied on a robot to exert this force, injecting its own inertial effects into the effective dynamics. Though, the used robot was chosen for its high back-drivability so that this effect, if not negligible, would not explain the force field discrepancy between simulation and experiment. In the model, for a given perturbation intensity, the size of the lateral deviation of the before-effect trajectory is determined by the interplay between the operation of the state estimator and the dynamics of the arm. Two observations can be made. First, changes in parameters of the state estimator (feedback delay, noise ratio) can reduce the impact of the force field. Yet even a fine tuning of these parameters would not lead to a realistic trajectory deviation for a 10-Ns/m field. Second, parameters of the dynamics have also an influence on the response to perturbations. For instance, we have assessed the influence of the muscle gain ratio (Fig. C.11). This parameter reflects the relative efficiency of shoulder and elbow muscles but its value is not easy to set as it depends on the muscle maximum force, its tendon properties, the physiological cross-sectional area, the innervation ratio, the moment arm and the modulation of force production by firing rate and recruitment in pools of motoneurons of each muscle. Furthermore we cannot play freely with this parameter as it has a strong impact on the timing of the movement (Fig. C.11D). Other parameters of the dynamics cannot be modified as they pertain to intrinsic characteristics of the arm.

We propose two ideas to obtain quantitatively more realistic deviations with respect to the intensity of the perturbation. The first idea is to use a more realistic dynamics for the modeled arm. For simplicity, we considered the control of a planar two-link arm. Yet the participants were free to use all available degrees of freedom from the trunk to the wrist. The corresponding kinematic chain would likely offer a larger inertial resistance to perturbations. The second idea is in fact an extension of the first one and invokes impedance to account for resistance to perturbations, i.e. not only inertia, but also joint viscosity and stiffness, could contribute to the resistance [Hogan, 1985; Burdet et al., 2001]. In the simulations,

we used a long feedback delay (0.12 s) to clearly indicate that any kind of instantaneous, short-latency and medium-latency visco-elastic contributions of muscles and tendons remained unmodeled. A model of these contributions is feasible for perturbations about a static posture [Crevecoeur and Scott, 2014], but remains elusive for perturbations during ongoing movements. Note that the very efficient elastic feedback along the desired trajectory used in the compensation model cannot be included in the reoptimization model or the redirection model due to the absence of a desired trajectory.

Chapter 4

Neural implementation of optimal control

4.1	Introduction	78
4.2	Materials and methods	79
4.2.1	Neural network model	79
4.2.2	Dataset and training	80
4.2.3	Simulation and data analysis	80
4.2.4	Parametric analysis	81
4.2.5	Dimensionality reduction	82
4.3	Results	82
4.3.1	Network training and performance	82
4.3.2	Directional tuning	85
4.3.3	Relationship between neurons and EMG	88
4.3.4	Influence of arm posture	89
4.3.5	Dimensionality reduction	89
4.4	Discussion	96
4.4.1	Disclaimer	96
4.4.2	Distributed representation of the universal controller	96
4.4.3	Relation to previous models	97
4.4.4	Role of the motor cortex	97
4.4.5	Dimensionality reduction	98
4.4.6	Limitations	99
4.4.7	Future developments	100
4.4.8	Extension	100

4.1 Introduction

If, on the one hand, we claim that we have understood, through a model, how the human produces movements, on the other hand, it is natural to reflect upon the way the human nervous system participates to this motor function. The better our understanding, the more we could hope to bring a useful contribution to a computational neuroanatomy of movement, i.e. the non-trivial assignment of operations to specific brain regions [Shadmehr and Krakauer, 2008; Haar and Donchin, 2020]. It is an attempt to sketch a neurological scheme in which elements of the model are univocally associated to some nervous structures.

Key issues of this approach are important since they concern not only the normal functioning of the brain but also its pathologies through “virtual lesions” of the model. Hence motor deficits found in different pathologies (stroke, Parkinson’s disease, cerebellar syndromes, ...) could be cast in a theoretical framework that would guide rehabilitation processes (e.g. Han et al. 2008).

A link between the model and the nervous system is made possible by techniques of artificial neural networks which allow virtually any mathematical function to be learned and reproduced in a distributed fashion by a set of interconnected elementary processing units which resemble neurons [Yuste, 2015; Kriegeskorte and Golan, 2019]. Once properly trained, an artificial neural network, as any real neural network, can be submitted to an “electrophysiological” study to decode the way neurons contribute to the production of the desired function and to proceed to comparisons with experimental data [Sussillo et al., 2015; Saxena et al., 2021]. In fact there is nothing really new in this approach [Rumelhart et al., 1986; Zipser and Andersen, 1988] except the great renewal of these techniques due to increased computing capacities [LeCun et al., 2015].

Our problem is thus to build a neural network which approximates the function of our universal controller (section 2.6.1.1). In fact this is neither a new nor an intractable problem, but it has some interesting specificities. There is quite a long tradition of studying neural networks controllers [Nguyen and Widrow, 1989]. The goal is in general to discover a control law by learning, either by simple error minimization as in the truck backer-upper example [Nguyen and Widrow, 1989], or by reinforcement learning [Levine, 2013; Lillicrap et al., 2015]. We are not directly interested by these methods since we assume that the proper controller is known. Our approach, based on training a deep neural network with input/output pairs generated by the controller, is close to that of Berniker and Körding [2015] and Sánchez-Sánchez et al. [2016]. Yet there is a main difference in the nature of the network outputs. In the latter studies, temporal patterns of control and/or full trajectories are spatially represented on the network output layer. What is actually learned is not the controller itself but some patterns of control.

Once we have built a neural network which properly approximates a motor controller, we become involved in the century-old debate on the nature of processing and representations in motor regions of the brain [Todorov, 2000; Scott, 2008; Reimer and Hatsopoulos, 2009; Vyas et al., 2020]. Here we propose to settle the debate by accepting the idea that any observed neural property related to motor control is a by-product of the motor system being a motor controller.

4.2 Materials and methods

4.2.1 Neural network model

We have designed a control policy that accounts for a wide range of observations on motor control [Guigon et al., 2019; Guigon, 2022; Moullet et al., 2022]. This policy is a stationary, optimal feedback control policy that, under the assumption of a full-state terminal constraint can be written (section 2.6.1.2, Eq. 2.55):

$$\mathbf{u}(t) = \boldsymbol{\pi}(\mathbf{x}(t), \mathbf{x}_G(t)) \quad (4.1)$$

where \mathbf{x} is the current state of the body, \mathbf{x}^* its desired state, and \mathbf{u} the instantaneous control vector that induces the state evolution of the body over one timestep when it transits optimally between \mathbf{x} and \mathbf{x}_G in a planning horizon. A full movement can be simulated by coupling the control policy with the body dynamics

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \quad (4.2)$$

On the one hand, the control policy is a function which is a solution of a complex, nonlinear optimization problem. On the other, it looks like a stationary mapping between vectors. It is easy to see that in the linear case (linear dynamics \mathbf{F}), the control policy is linear and can be represented by a linear mapping. In the nonlinear case, the policy is nonlinear and could possibly be replaced by a nonlinear mapping, e.g. a multilayer neural network which can be written

$$\mathbf{u} = \mathbf{N}_{w,f}(\mathbf{x}, \mathbf{x}_G) \quad (4.3)$$

where $\{w, f\}$ indicates that the network is defined by weights and transfer functions [Hinton, 2007; Kriegeskorte and Golan, 2019].

We built a multi layer perceptron with two hidden layers (Fig. 4.1A) for the control of a planar two-link arm (Fig. 4.1B). The input layer contained 16 neurons, 8 for \mathbf{x} and 8 for \mathbf{x}_G , and the output layer contained two neurons (shoulder and elbow controls). The hidden layers had 64 and 256 neurons, respectively. The layers were fully connected. The transfer function was a RELU function for the hidden layers and a linear function for the output layer.

4.2.2 Dataset and training

We trained the neural network with a dataset built in the following way. We generated 200 0.5-s trajectories (timestep 0.001 s) starting from the same position ($[40^\circ, 110^\circ]$) toward twenty targets uniformly spaced on ten uniformly spaced circles (radius from 0.1 to 0.2 m) (Fig. 4.1C) with null terminal velocities, activations and excitations. Each trajectory provided 500 input (current state, target state) / output (controls) pairs which gives a dataset with 100000 examples. This dataset was randomly split between a training base and a testing base (respectively 75% and 25% of the examples) to monitor the learning process.

The network was implemented in Keras 2.5.0 with a tensorflow 1.12.0 backend. Training was performed using MSE (Mean Square Error) as a cost function, a batch size of 25 and the gradient descent was realized with Adam [Kingma and Ba, 2014], with a 0.0001 learning rate. Models were trained up to 60 epochs in order to select *a posteriori* the optimal epoch. At each period, performance error was first measured by two quantities: (1) the training error is the MSE calculated over the batches of the training set; it is the internal error used to train the network; (2) the test error is the MSE calculated over the test set; (3) the reconstruction error is an error based on the difference between the optimal trajectories (produced using the actual optimal control policy) and the corresponding trajectories produced by the network:

$$R_{loss}^e = \sqrt{\sum_{i=1}^8 \sum_{t=0}^{500} \left\| \mathbf{p}_{i,t}^O - \mathbf{p}_{i,t}^{NN,e} \right\|^2} \quad (4.4)$$

where $\mathbf{p}_{i,t}$ is the hand Cartesian position at time t for a trajectory of index i and O and NN denote trajectories generated respectively *via* optimal control and the neural network trained until epoch e . Eight 0.5 s trajectories were considered, starting from the same position ($[40^\circ, 110^\circ]$) toward eight targets uniformly spaced on a 0.15 m radius circle. Note that only the four trajectories corresponding to the cardinal points were included in the training dataset.

4.2.3 Simulation and data analysis

To produce a movement, the neural network was coupled with our simplified model of an arm (Fig. 4.1D). At each time, current goals (task representation) and sensory feedback from the arm are entered in the input layer of the network. The network produces motor commands in its output layer which are sent to the arm which moves according to its dynamics (Eq. 2.13) and generates new perfect sensory feedback. Note that we do not consider either noises or delays, for we wish to emulate only the controller of the computational schema in Fig. 2.14, which has no knowledge of any noise or delay. We thus focus here on the ability of the neural network to produce the right output (control) from a given input.

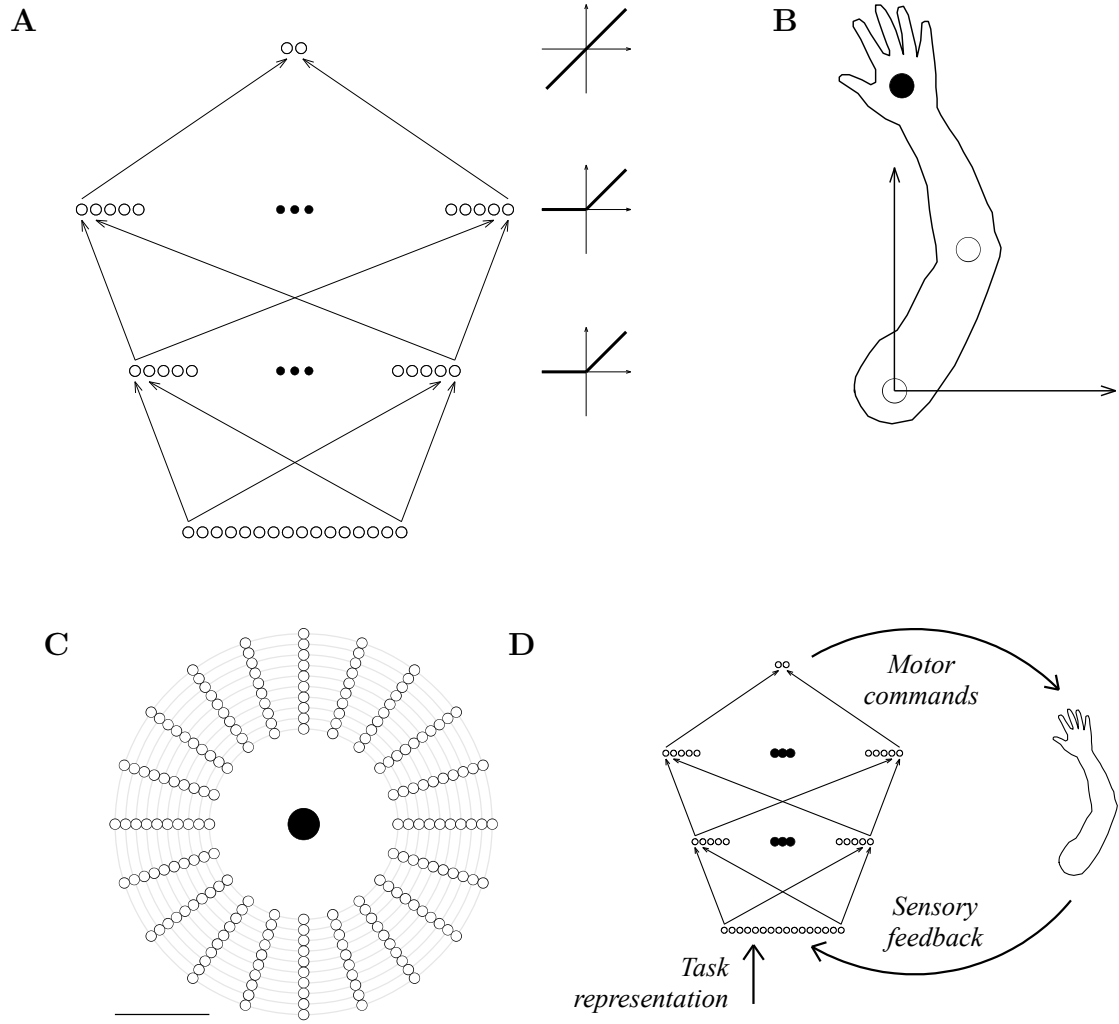


FIGURE 4.1: **Neural network and training set.** **A.** Network architecture and transfer function. **B.** Schema of the arm controlled by the neural network. **C.** Training set of the network. Scale: 0.1 m. **D.** Closed-loop architecture of the model.

Movements of different directions, amplitudes and velocities were simulated using Eqs. 4.2 and 4.3. EMG was defined as the excitation state. Output of neurons in hidden layers 1 and 2 was defined as neuronal activity.

4.2.4 Parametric analysis

We simulated a DIRECTION experiment in which movement direction was varied [Georgopoulos et al., 1982]. The directional tuning of neuronal activity and EMG was analyzed [Georgopoulos et al., 1982]. The quantity of interest d (neuronal activity, EMG) was regressed against target direction θ according to

$$d = b_0 + b_1 \cos \theta + b_2 \sin \theta \quad (4.5)$$

which was used to defined the preferred direction θ according to

$$d = b_0 + c_1 \cos(\theta - \theta_0) \quad (4.6)$$

The p -value of the regression was taken as a measure of directional tuning. As d is a function of time, a preferred direction can be calculated at every timestep [Sergio and Kalaska, 1998]. For a population of N neurons, we note $\mathbf{PD}_i(t)$ the preferred direction vector of neuron i which is the unitary vector along the preferred direction of the neuron. Distributions of preferred directions were built and were tested for bimodality (Hartigan’s dip test; Hartigan and Hartigan 1985). A population vector was calculated as

$$\mathbf{P}(t) = \sum_{i=1}^N w_i(t) \mathbf{PD}_i(t) \quad (4.7)$$

where $w_i(t) = d_i(t) - a_i$, with $a_i = \text{mean}(d_i(t))$ [Georgopoulos et al., 1988].

4.2.5 Dimensionality reduction

For each movement direction, the DIRECTION experiment produced a $Q \times N$ array of neuronal activity, where Q is the number of timesteps ($Q = 500$ for a 0.5-s movement) and N the number of neurons ($N = 64$ for the first hidden layer, $N = 256$ for the second hidden layer). A principal component analysis (PCA) was applied to this array: the rows (timesteps) were observations and the columns (neurons) were variables. The PCA transformed the array into new variables (principal components, PCs; N vectors of size Q) and coefficients (N vectors of size N) which can be used to reconstruct the variables by linear combinations of the PCs (Fig. 4.2). We considered the first two PCs which account in general for $> 90\%$ of the variance of the data, and reflect the dominant signals in the population of neuron. The same method could be applied to EMG. Yet, as there are only two muscles in the model, this method is no relevant here.

4.3 Results

4.3.1 Network training and performance

As expected, the training and test errors calculated at each epoch decreased during the course of training (Fig. 4.3A). Yet, they proved not to be a proper criterion for training convergence, for models with reasonable losses proved to generate non-realistic trajectories (see Fig. 4.4B). The reconstruction error defined in Eq. 4.4 also decreased, but with a slower time course (Fig. 4.3B). The model trained up

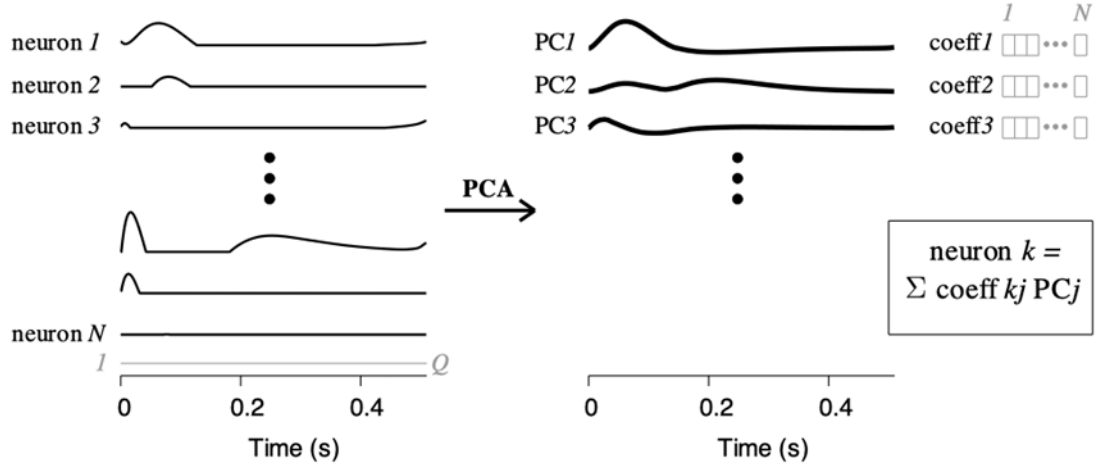


FIGURE 4.2: **Schema of the principal component analysis.** N is the number of neurons. Q is the number of timesteps.

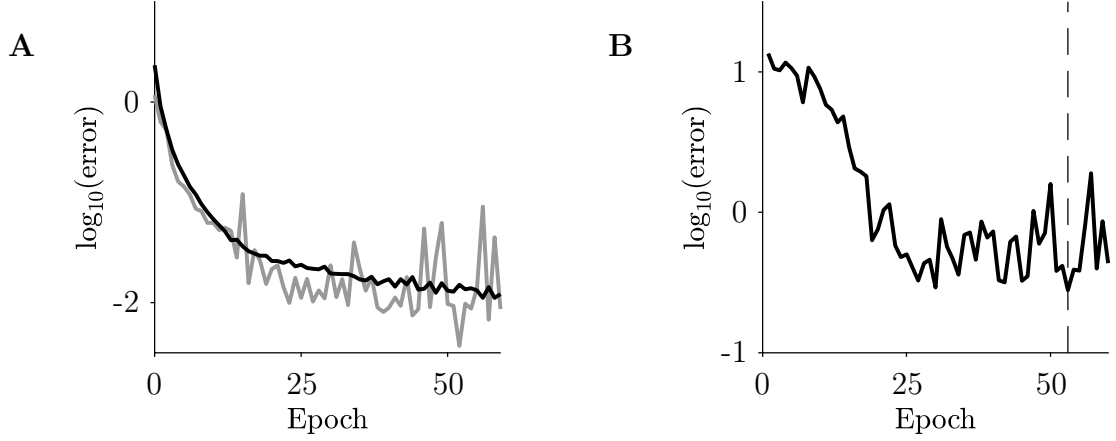


FIGURE 4.3: **Training curves.** **A.** Training error (black) and test error (gray). **B.** Reconstruction error. Vertical dashed line: epoch of minimum reconstruction error.

to 53 epochs was chosen, corresponding to the minimum of reconstruction error (vertical dashed line in Fig. 4.3B), and showed to satisfyingly reproduce optimal control (see Fig. 4.4).

We simulated 15-cm, 0.5-s trajectories in eight directions (at 45° intervals) from the starting position used in the training set (Fig. 4.4A, inset). The trajectories were more or less straight with a bell-shaped velocity profile (Fig. 4.4A,C). For comparison, trajectories produced by a partially trained network are shown in Fig. 4.4B. Shoulder and elbow muscle displayed an agonist-antagonist pattern - their EMG took positive and negative values, as would the difference between agonist and antagonist activities ($EMG \sim EMG_{ag} - EMG_{ant}$) - and their activities were tuned to movement direction (Fig. 4.4D,E). We provide a more detailed description of the network performance in Appendix C.

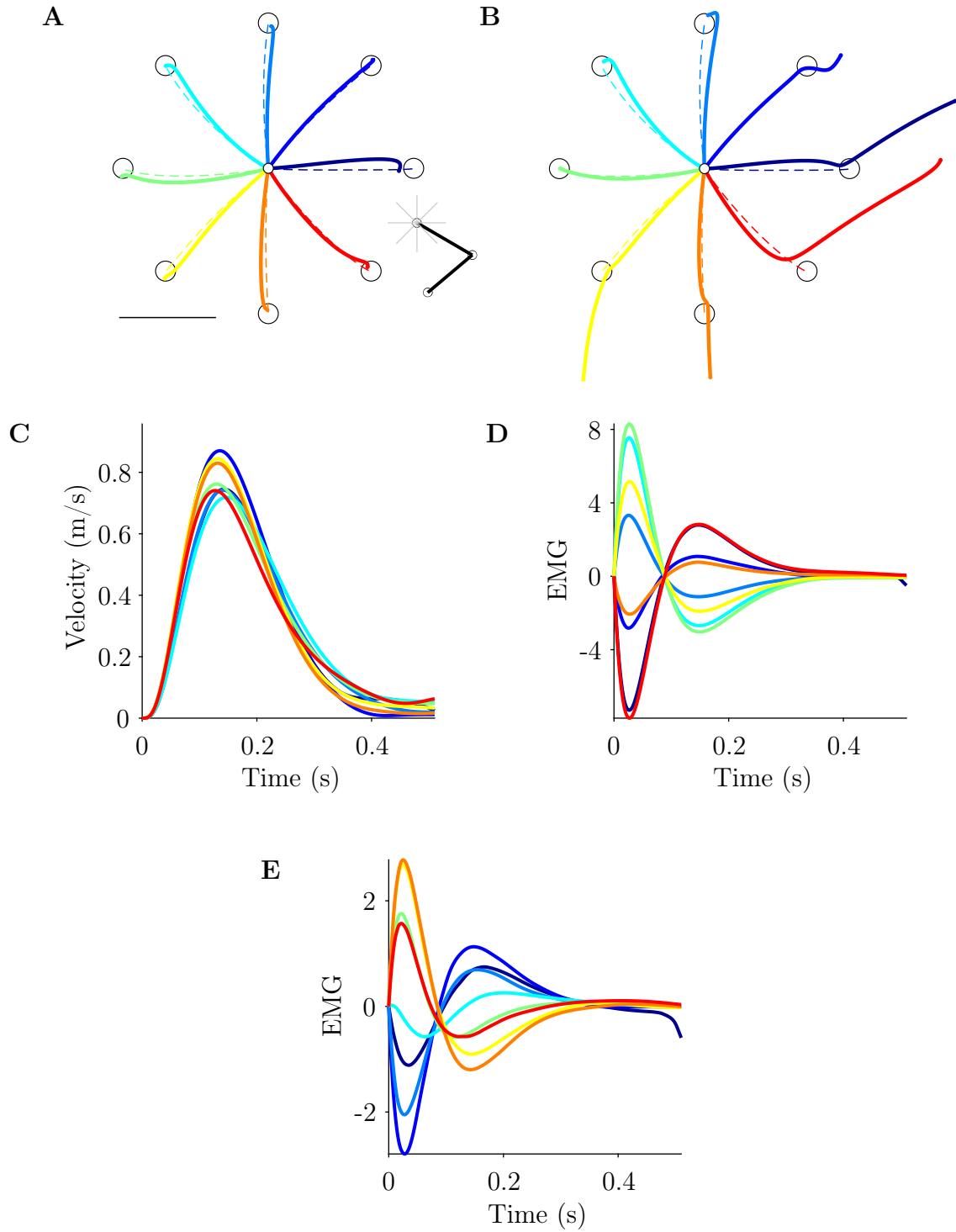


FIGURE 4.4: **Neural network simulations.** **A.** Trajectories toward eight targets generated with the neural network (plane thick lines) and optimal control (dashed thin lines). Scaling: 0.1 m. Inset: initial arm posture. **B.** Same as **A.** with a partially trained neural network (until epoch 17 vs 53 for trajectories in **A**). **C.** Velocity profiles. Color code as in **A**. **D.** Shoulder EMG. **E.** Elbow EMG.

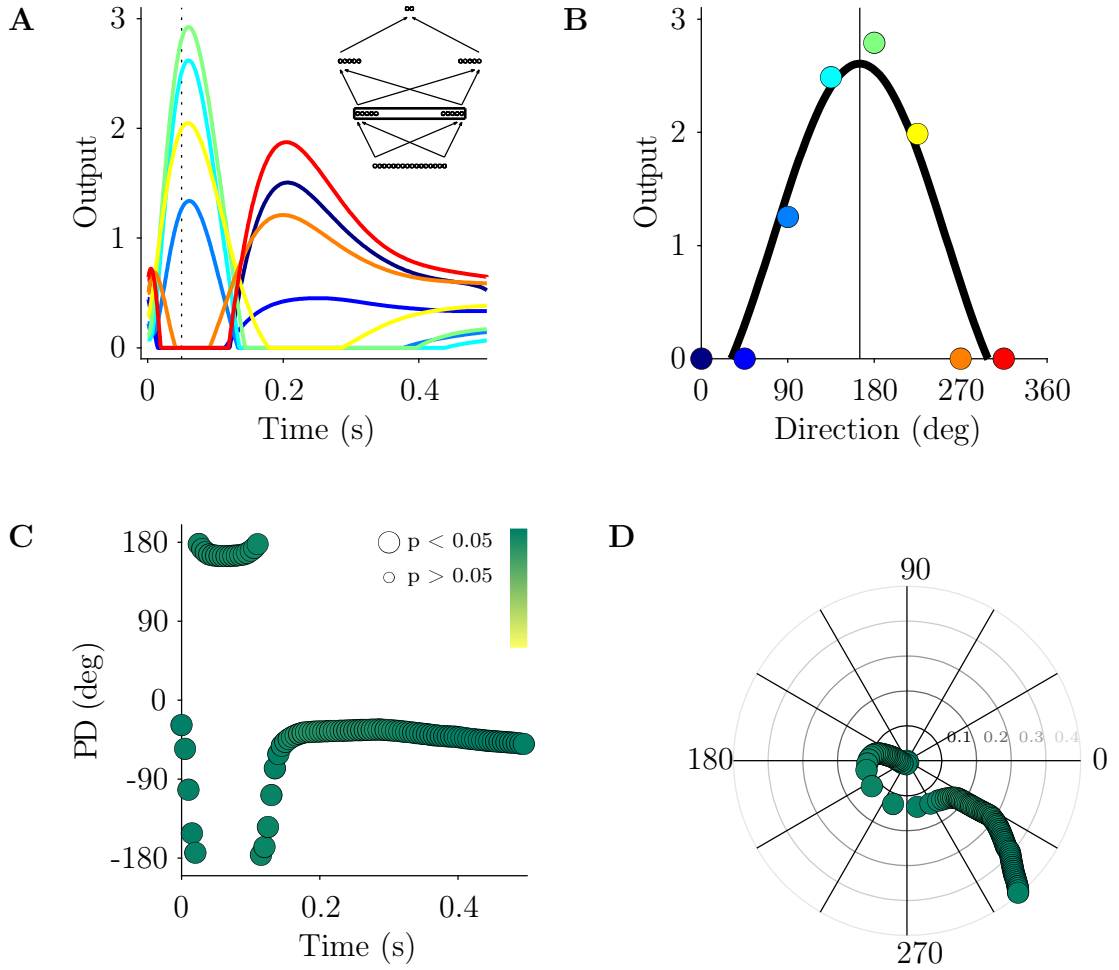


FIGURE 4.5: **First layer analysis.** **A.** Output of neuron 18 of the first hidden layer for 8 movement directions. Color code as in Fig. 4.4 **B.** Tuning curve of output at 0.05 s (dotted line in **A**). The preferred direction is indicated by the vertical line. **C.** Neuron's preferred direction every 5 ms. Color code: regression R^2 . **D.** Same data as **C** in polar representation. Each point of **C** is represented by the polar coordinates $r, \theta = (time, PD)$. Radial circles indicate time.

4.3.2 Directional tuning

We considered a single neuron in the first hidden layer (Fig. 4.5). Its output varied with movement direction (Fig. 4.5A). The neuron was directionally tuned at 0.05 s after movement onset (preferred direction $PD = 164.9^\circ$, p -value = 0.0023; Fig. 4.5B). In fact, the neuron was directionally tuned at every timestep along the movement and its preferred direction changed in time (Fig. 4.5C,D; see fig. 6 in Sergio et al. 2005).

In the same way, both the shoulder and elbow muscles were directionally tuned and changed their tuning in time (Fig. 4.6; fig. 3 in Sergio and Kalaska 1998).

In the first hidden layer, all the neurons had a nonzero output at least at one point in time for one of the 8 movements. Ninety-eight percent of the neurons

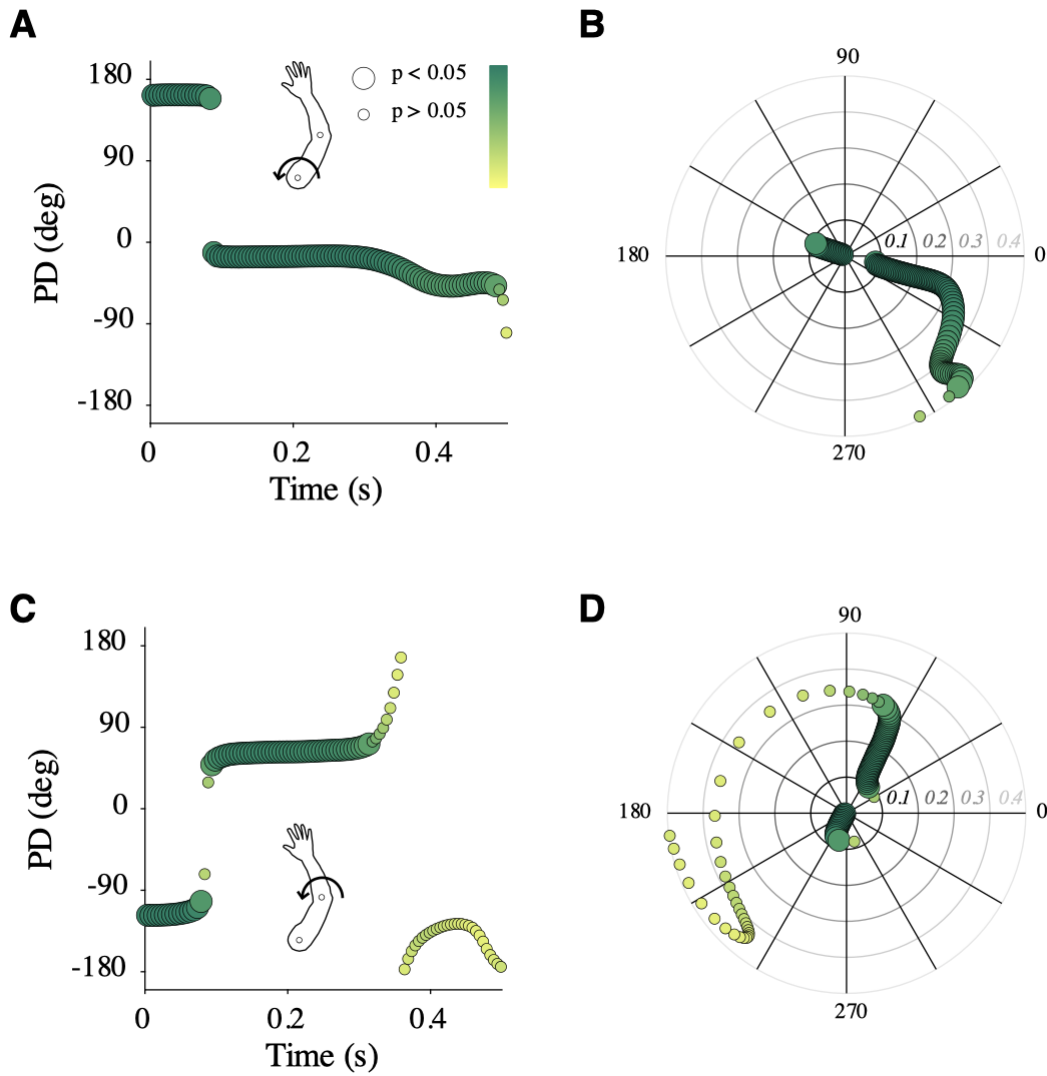


FIGURE 4.6: **Muscles directional tuning.** **A.** Shoulder muscle's preferred direction every 5 ms. **B.** Same data as **A** in polar representation. **C.** Same as **A** for elbow muscle. **D.** Same as **B** for elbow muscle.

were directionally tuned at least at one point in time (mean 87.2% of the time points). The distributions of neuron's preferred directions were calculated every 0.05 s (Fig. 4.7). The distributions were anisotropic with a main axis (in the second quadrant) at $148.3 \pm 8.7^\circ$ (fig. 3 in Scott et al. 2001a, fig. 4 in Suminski et al. 2015)

In the second hidden layer, the neurons were much more sparsely active (Fig. D.6) and much less directionally tuned (Fig. D.3). Forty percent of the neurons had zero output for the 8 movements. Of the remaining neurons, 55% were directionally tuned at least at one point in time (mean 5.4% of the time points). As expected, reconstruction of movement direction from neuronal activities in the first hidden layer was biased due to the anisotropic distribution of preferred directions (Fig. 4.8; fig. 2 in Scott et al. 2001a).

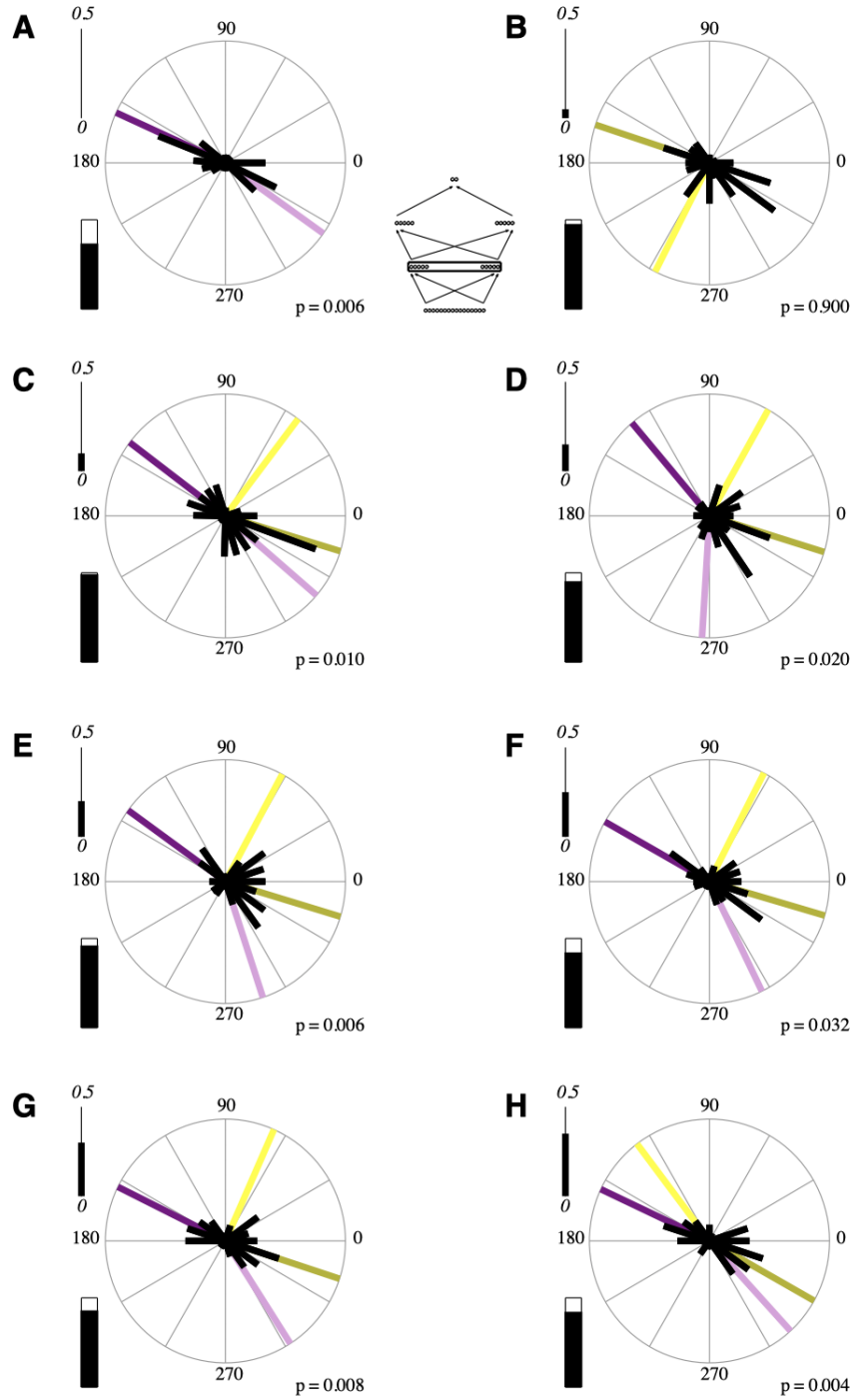


FIGURE 4.7: Distribution of neuron's preferred direction (circular histogram, 20 bins, black bars) for the first hidden layer calculated every 0.05 s (**A** to **H**). On each plot, the upper gauge indicates time (between 0 and 0.5 s) and the lower gauge, the percent of tuned neurons (between 0 and 100). The circle corresponds to 15 neurons. The purple lines indicate peaks of a bimodal distribution (when significant, Hartigan's dip test; p -value in the lower right). The yellow lines indicate PD of muscles (shoulder: dark; elbow: light).

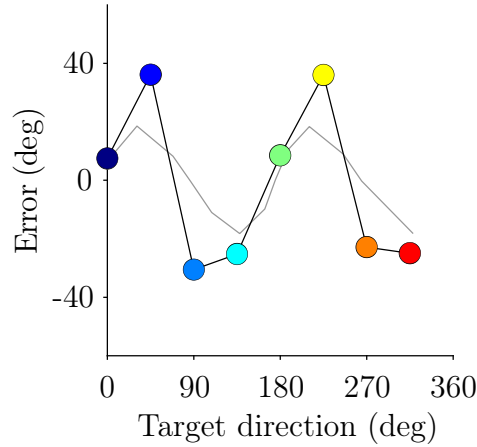


FIGURE 4.8: Error in the reconstruction of movement direction using the population vector from the first hidden layer at movement onset. The gray line corresponds to data from [Scott et al. \[2001a\]](#) shifted leftward to account for the differences in the major axis of the PD distribution.

Third we analyzed the relationship between the preferred directions of neurons and the preferred directions of the muscles. We calculated the distribution of differences between neuron's preferred directions and shoulder and elbow muscle's preferred direction (Figs. 4.11, 4.12). These results are consistent with experimental observations: in [Griffin et al. \[2015\]](#), 30% of corticomotoneuronal cells recorded during wrist movements had a PD within $0-45^\circ$ of their target muscle (almost parallel) and 30% within $135-180^\circ$ (almost opposite).

4.3.3 Relationship between neurons and EMG

We assessed the relationship between neuronal activity and EMG in three ways. First, we calculated the mean squared correlation coefficient of neuronal activity and EMG across movement directions for the shoulder and elbow muscles (Fig. 4.9) which is consistent with the observed low neuron-EMG similarity [[Schieber and Rivlis, 2007](#)].

Second, we analyzed the extent to which peaks of neuronal activity were accompanied by peaks in EMG activity [[Griffin et al., 2008](#)]. For each neuron, we searched for peaks of activity and calculated the percentage of movement directions in which a peak in EMG was present in the same time window. The calculus was made for the primary peaks of neuronal activity and EMG (Fig. 4.10A,B) and for the primary peaks of neuronal activity and any peaks of EMG (Fig. 4.10C,D). Again we see a strong mismatch between neuronal activity and EMG [[Griffin et al., 2008](#)].

Third we analyzed the relationship between the preferred directions of neurons and the preferred directions of the muscles. We calculated the distribution of neuron's preferred directions relative to shoulder and elbow muscle's preferred

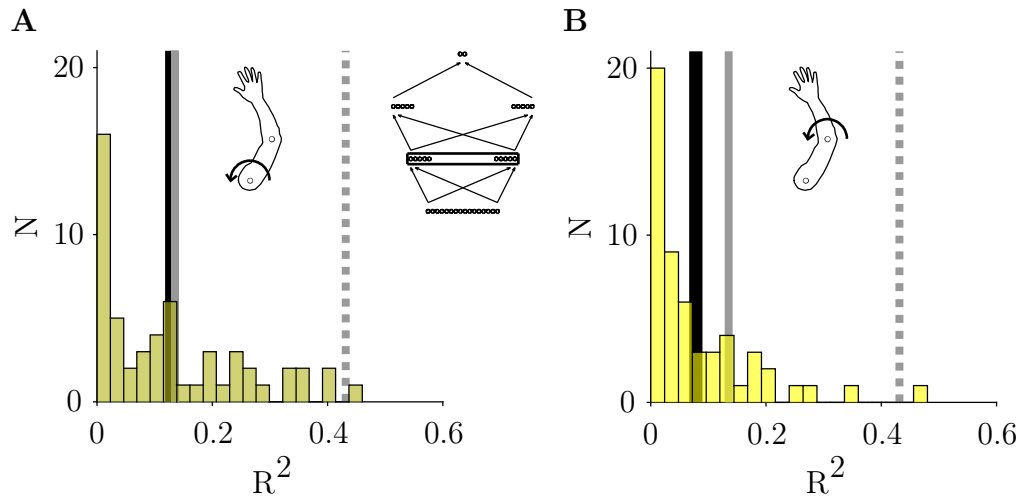


FIGURE 4.9: Mean squared correlation coefficient between neuronal activity in the first hidden layer and EMG. **A.** Shoulder muscle. **B.** Elbow muscle. Black vertical solid line: mean R^2 . Gray vertical line: data from [Schieber and Rivlis \[2007\]](#), solid is mean, dotted is max.

direction (Figs. 4.11, 4.12). These results are consistent with experimental observations: in [Griffin et al. \[2015\]](#), 30% of corticomotoneuronal cells recorded during wrist movements had a PD within 0-45° of their target muscle (almost parallel) and 30% within 135-180° (almost opposite).

4.3.4 Influence of arm posture

We explored the influence of initial arm posture on the primary axis of distributions of preferred directions. We trained four neural networks with varying elbow angles (shoulder/elbow: 30/80°, 30/90°, 30/100°, 30/110°). The training set contained 0.5-s trajectories toward twenty targets uniformly spaced on ten uniformly spaced circles (radius from 0.05 to 0.15 m). We simulated 0.1-m, 0.5-s trajectories in eight directions (at 45° intervals) from each starting position. We observed that the preferred directions of neurons (Fig. 4.13A,C) and elbow muscle (Fig. 4.13B,C) rotated with change in elbow angle. The rotation for neurons was in the same direction as that the elbow muscle but with a smaller magnitude (Fig. 4.13C). [[Yanai et al., 2008](#); [Oby et al., 2013](#)].

4.3.5 Dimensionality reduction

We analyzed the behavior of population of neurons in the DIRECTION experiment (Fig. 4.14A, inset) using a dimensionality reduction method (PCA; see Materials and Methods; [Russo et al. 2018](#)). For each movement direction, we extracted the first two principal components (PCs) from the output of the 64 neurons of the first

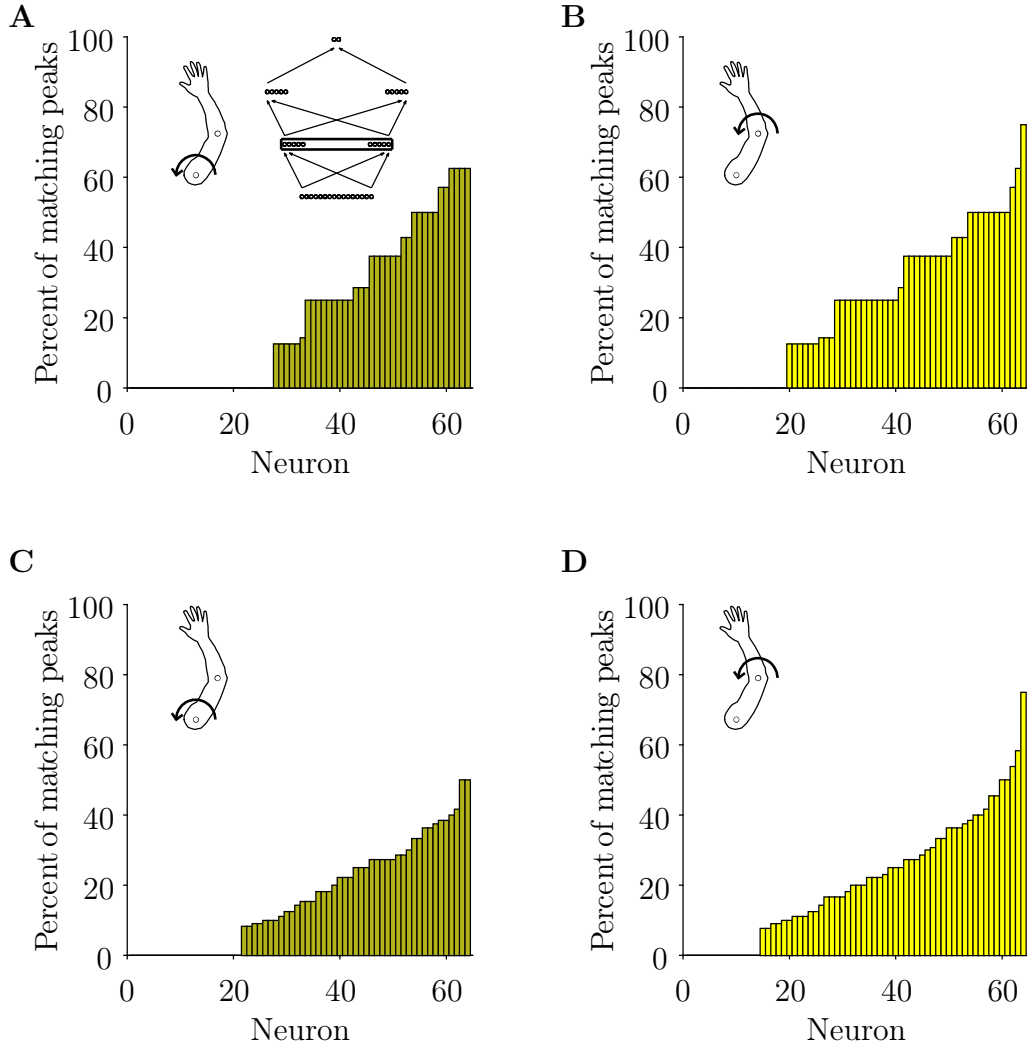


FIGURE 4.10: Matching peaks between neuronal activity in the first hidden layer and EMG. **A.** Primary peaks of neuronal activity and EMG for the shoulder muscle. **B.** Same as **A** for the elbow muscle. **C.** Primary peaks of neuronal activity and any peaks of EMG for the shoulder muscle. **D.** Same as **C** for the elbow muscle.

hidden layer. These PCs (e.g. Fig. 4.14B, inset) correspond to two typical temporal patterns from which the output of all the neurons can be linearly reconstructed. The first two PCs accounted for 90.5% of the variance of neuronal outputs (mean across movement directions). The neural trajectories were in general elliptical and rotated all in the same direction [Russo et al., 2018; Saxena et al., 2021]. Muscle trajectories (Fig. 4.15) and neural trajectories in the second hidden layer (Fig. D.7) were less elliptical and rotated in different directions [Russo et al., 2018; Saxena et al., 2021]. For comparison, neural trajectories in the first hidden rotated in the counterclockwise (CCW) direction for all movement directions (test with 8, 16 and 32 movement directions); neural trajectories in the second hidden layer rotated CCW in 12%, 19% and 19% of the movement directions, for 8, 16 and 32 directions, respectively; muscle trajectories rotated CCW in 25%, 31% and 38% of the movement directions, for 8, 16 and 32 directions, respectively. The dominant

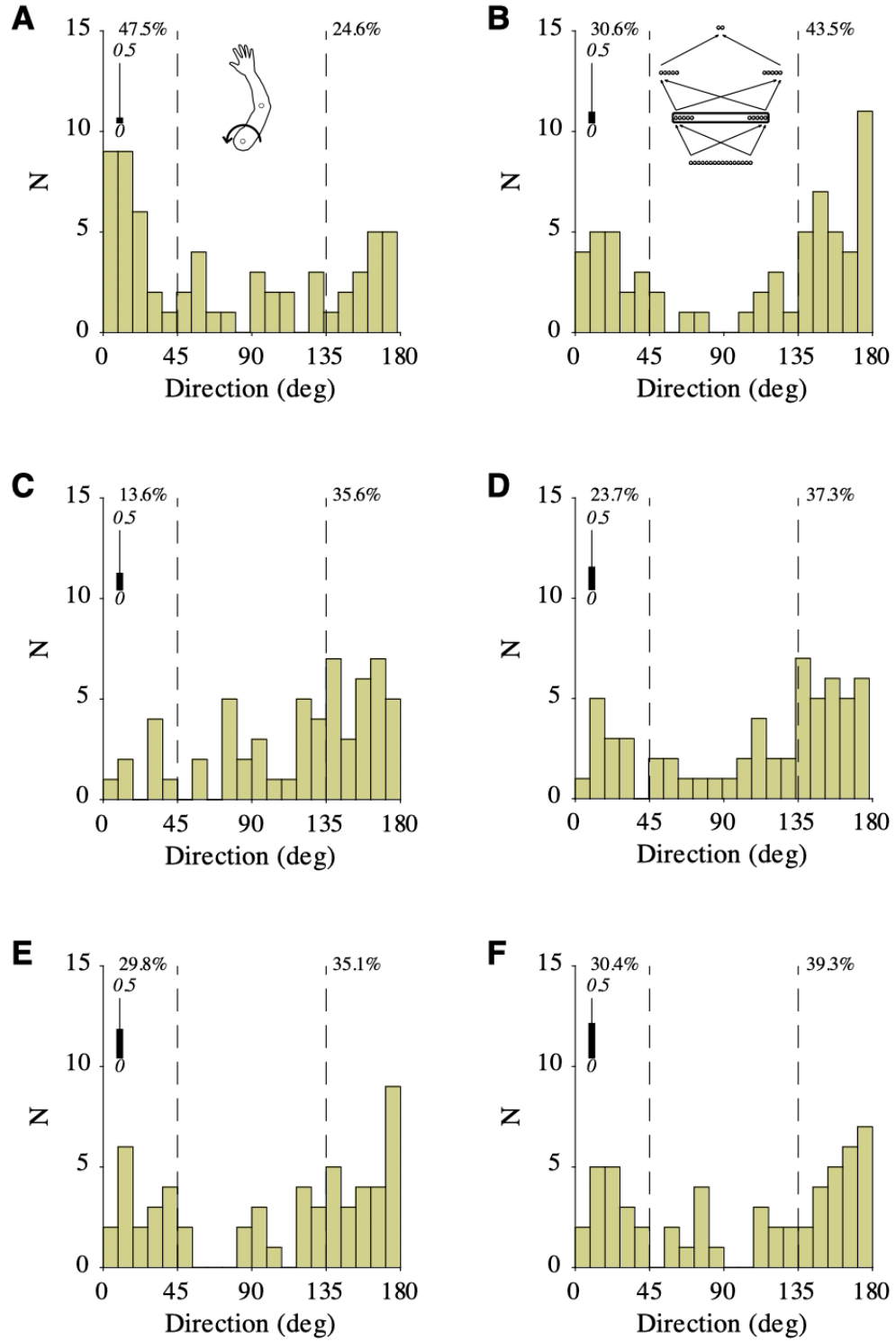


FIGURE 4.11: Distribution of the neuron's preferred direction in the first hidden layer relative to shoulder muscle preferred direction every 0.05 s (A to H). 0° on the abscissa corresponds to the shoulder muscle's preferred direction. On each plot, the gauge indicates time (between 0 and 0.5 s).

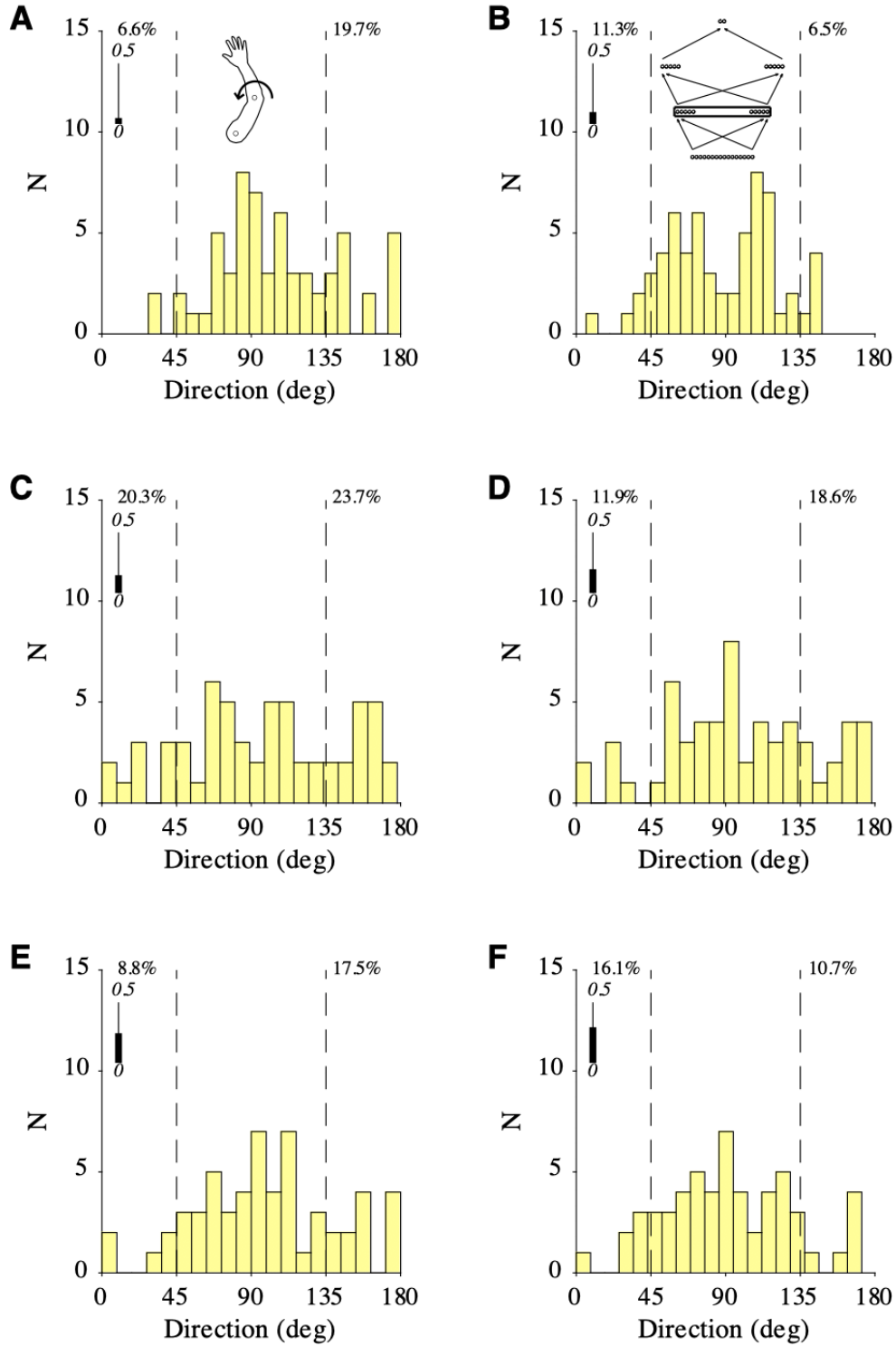


FIGURE 4.12: Distribution of the neuron's preferred direction in the first hidden layer relative to elbow muscle preferred direction every 0.05 s (A to H). 0° on the abscissa corresponds to the elbow muscle's preferred direction. On each plot, the gauge indicates time (between 0 and 0.5 s).

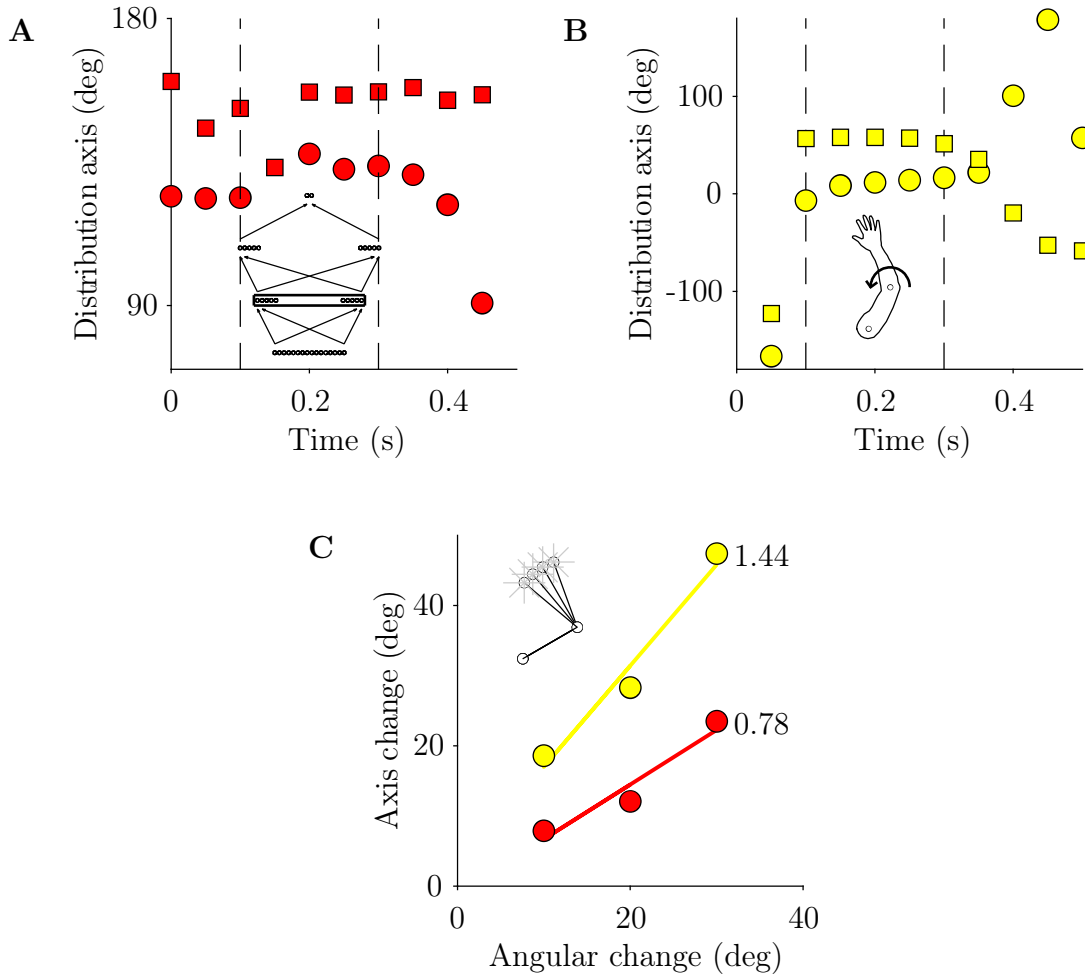


FIGURE 4.13: **Influence of arm posture.** **A.** Temporal evolution of the primary axis of the distribution of preferred directions for neurons in the first hidden layer. Two initial postures: shoulder/elbow $30/80^\circ$ (circle), $30/110^\circ$ (square). Missing points correspond to an absence of bimodal distribution. **B.** Same as **A** for the elbow muscle. **C.** Mean change in axis orientation as a function of change in elbow angle ($80-90$, $80-100$, $80-110^\circ$) for neurons (red) and elbow muscle (yellow) calculated over $0.1-0.3$ s (vertical dashed line in **A** and **B**). Regression lines and slopes are shown.

signals in the first hidden layer were clearly different from those in the muscles and in the second hidden layer. The exact meaning of what is observed using a dimensionality reduction method is unclear. The only significant issue is the fact that the observed properties emerge from neural network processing and are consistent with experimental observations [Russo et al., 2018; Saxena et al., 2021].

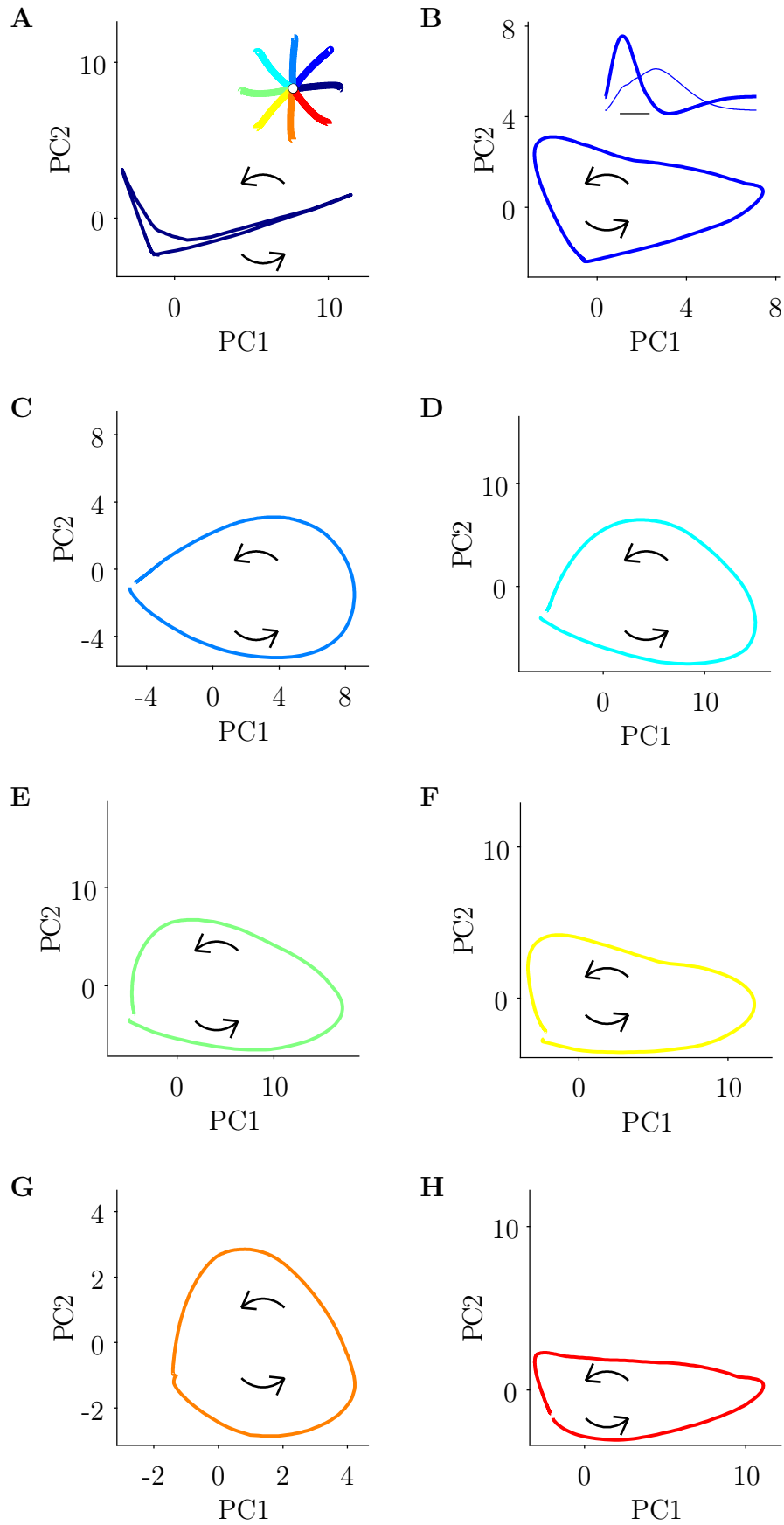


FIGURE 4.14: Neural trajectories in the first hidden layer corresponding to the first two PCs extracted from neural outputs for each movement direction of the DIRECTION experiment. Inset in **A**: movement trajectories. Inset in **B**: the two PCs for the corresponding direction (thick line: PC1; thin line: PC2). Arrows indicate the direction of rotation.

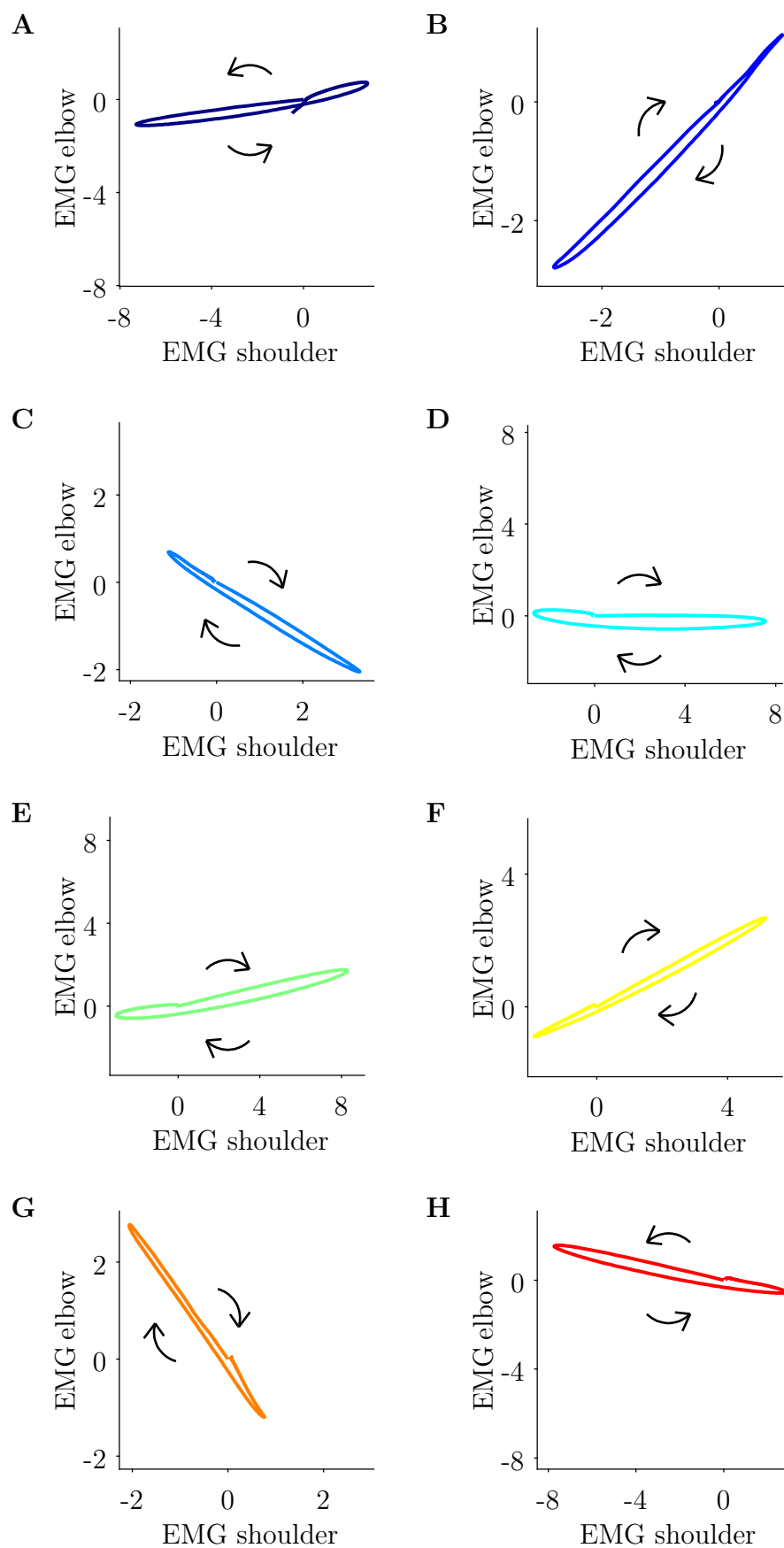


FIGURE 4.15: Shoulder vs elbow EMG for the movements of the DIRECTION experiment.

4.4 Discussion

We have shown that a neural network trained to approximate the universal controller of a planar two-link arm exhibits many characteristic properties of operations in the primary motor cortex. As these properties concern both the representational view and the dynamical view of motor cortical functions, our results suggest that these views are not necessary and can be replaced by the view that the motor cortex implements the proposed universal controller (see section 2.6). Premises of this idea are present in the literature [Todorov, 2000; Guigon et al., 2007a; Lillicrap and Scott, 2013; Kalidindi et al., 2021].

4.4.1 Disclaimer

This work is not about neural networks *per se*. We do not claim that we have found the best network architecture, the best training set or the best training procedure. We have not systematically quantified the performances of the neural network (interpolation, extrapolation). These issues would call for a specific treatment which is well beyond the scope of our work. Our goal was to show the feasibility of learning a distributed representation of the universal controller of a planar two-link arm and to carry out a comparison between the resulting neural network and the motor cortex.

4.4.2 Distributed representation of the universal controller

The observed features of neurons are emergent properties of a neural network trained to control a planar arm. Interestingly, similar features were observed in networks with different architectures (sequential or recurrent networks, number of layers, number of neurons per layer), different training sets and different training procedures (batch size, number of epochs). This observation suggests that these features correspond to a robust, well-defined, and invariant “neuronal” solution to the proposed control problem. The only really surprising feature is directional tuning. There is no specific reason why neurons should become directionally tuned and their activity well described by a cosine function. Other features such as the anisotropy of distributions of preferred directions and the variations in distribution axis with arm posture are less unexpected as they probably result from the geometrical and inertial properties of the arm [Scott et al., 2001a; Lillicrap and Scott, 2013]. The weak relationship between neuronal activity and EMG is not unexpected as the neurons are only indirectly involved in the production of the EMG. The presence of rotational dynamics is consistent with previous neural network models trained to produce optimal trajectories [Berniker and Penny, 2019; Kalidindi et al., 2021].

Our network architecture with two hidden layers raises two questions. First, why do we need two hidden layers? We have no theoretical answer to this question. We observed that networks with a single hidden layer had poorer performance than networks with two hidden layers even when they have more parameters (Fig. D.1, Fig. D.5). Second, what is the specific role of each layer? Again, we have no clearcut answer. We observed that most neurons in the first hidden layer were active and directionally tuned (Fig. D.6A ; Fig. 4.7) while neurons in the second hidden layer were sparsely active and weakly tuned (Fig. D.3; Fig. D.6B). We tried to reduce the number of neurons in the second hidden layer (128 instead of 256) but we obtained a poorer performance (Fig. D.4), which suggests that inactive neurons were not useless.

4.4.3 Relation to previous models

The main novelty of our approach is to train a neural network to produce a control policy. The knowledge embedded in the network is not a set of optimal trajectories discovered by learning [Lillicrap and Scott, 2013; Kalidindi et al., 2021], but an approximation of a well-defined function which is the universal controller of a planar arm. This means that the network need not be trained on the specific set of trajectories that we want to produce, but can be trained on any kind of training set, provided the states encountered in the training set have a sufficient overlap with those needed to produce the desired trajectories. Another interesting point is related to the stationary nature of the universal controller. A simple multilayered feedforward network trained with backpropagation is sufficient to learn the controller. There is no need of a recurrent network and backpropagation through time to train the network [Lillicrap and Scott, 2013; Sussillo et al., 2015; Kalidindi et al., 2021; Saxena et al., 2021]. We do not claim that our approach is more biologically plausible but only that the burden of training recurrent networks can be avoided. A proper dynamics can emerge from a feedforward network that controls a dynamical system (planar arm) and receives sensory feedback [Kalidindi et al., 2021].

4.4.4 Role of the motor cortex

Our results suggest that properties observed in the activity of motor cortical neurons are a by-product of the motor cortex being a controller. This view concurs with the conclusion of previous studies [Todorov, 2000; Guigon et al., 2007a; Lillicrap and Scott, 2013; Kalidindi et al., 2021]. The apparent contradictions in the debate between the representational view and the dynamical view of motor cortical functions [Reimer and Hatsopoulos, 2009; Shenoy et al., 2013; Schwartz, 2016] might well be settled by saying that these views are neither necessary nor sufficient to account for processing in the motor cortex. Yet, behind this seemingly

idle debate and the proposed conclusion that the motor cortex is a controller, there remains several unsolved issues on the role of the motor cortex. First, we considered data from nonhuman primates which fail to acknowledge the input/output organization of the motor cortex, i.e. the fact, properly studied in rodents, that the motor cortex is a “multilayered” network with multiple outputs (e.g. toward the spinal cord and the basal ganglia) [Weiler et al., 2008; Isomura et al., 2009; Kiritani et al., 2012; Currie et al., 2020; Park et al., 2022]. It would be interesting to combine “primate” tasks (e.g. reaching in multiple directions; Bollu et al. 2019; Becker et al. 2020) and the precise anatomo-physiological resolution available in rodents. It could put additional constraints on a neural network model of the motor cortex. Second, even in primates, there is no simple causal role of the motor cortex in the production of movements [Travis and Woolsey, 1956; Carlson and Devinsky, 2009; Murata et al., 2015]. The situation is still more complex in cats and rodents in which dexterity is lost after motor cortex or corticospinal tract lesions, but most motor functions remain intact [Grillner, 1985; Kawai et al., 2015; Serradj et al., 2021]. It is unclear how to conciliate these observations and a role of the motor cortex as a controller. A possible bypass could be to consider that the corticospinal tract is necessary for skilled but not unskilled behaviors [Serradj et al., 2021].

4.4.5 Dimensionality reduction

As in many previous studies, we have used a method of dimensionality reduction [Churchland et al., 2012; Russo et al., 2018; Saxena et al., 2021]. The basic method is principal component analysis applied on recorded neuronal activity with neurons as variables and time course of activity as observations. Taking the first two or three PCs reveal dominant low-dimensional features in the neuronal population activity. The same method can also be applied to EMG activity when multiple muscles are simultaneously recorded. A typical observation is the difference in low-dimensional activity between motor cortical neurons and muscles [Russo et al., 2018; Saxena et al., 2021]. This difference was observed in a recurrent network trained to produce observed patterns of muscular activity [Saxena et al., 2021] and in our non-recurrent network trained to be a controller. The specificity of this observation is unclear (but see below).

Churchland et al. [2012] developed a new reduction method (jPCA) which specifically reveals rhythmic oscillatory tendencies in the population data. They observed that an oscillatory pattern was present even in nonrhythmic discrete reaching movements. Since its inception, this method has been used not only to analyze experimentally recorded neuronal activities but also to disclose similar properties in artificial neural networks [Hennequin et al., 2014; Sussillo et al., 2015; Michaels et al., 2016; Kalidindi et al., 2021]. We have also applied the jPCA

method¹ to our network and obtained the expected rotational patterns (results not shown). Recently, Lebedev et al. [2019] showed, using Churchland’s data and simulations, that “rotational patterns occurred in neuronal populations when there was a temporal sequence in peak firing rates exhibited by individual neurons”. Accordingly, they casted some doubt upon the specificity and the potential explanatory power of low-dimensional features of neuronal activities. We observed a typical temporal sequence in peak firing rates in our neural network (Fig. D.6).

4.4.6 Limitations

There are several important limitations to this work. First, we have not thoroughly explored the influence of the network architecture, the training procedure (batch size, number of epochs) and the training set (number and nature of samples) on the performance of the network (how efficiently it learns the control policy) and emergent properties of neurons. Preliminary observations indicate that networks with two hidden layers are more efficient than networks with a single hidden layer but that the usual training loss (MSE) proved to be a poor criterion for assessing the convergence to the optimal neural network. This might be explained by the very nature of the training dataset and the dynamics of the modelled system. As the trajectories composing the dataset exhibit well known bell-shaped velocity profiles, a majority of the inputs present small velocities, while high velocity inputs are scarce. It is then most probable that the rapid decrease of both training and test loss in the early epochs is due to an efficient fitting of the numerous low velocity examples. Conversely, the fitting of the few high velocity ones only slightly decreases the loss giving the appearance of a quick convergence of the learning process. A brute-force data generation method could be to homogeneously draw values from maximum to minimum values of each of the 16 inputs. But not only would this dramatically increase the size of the dataset (the combination of 10 draws of these 16 coordinates would make up to 10^{16} inputs), but the generated inputs could be highly incoherent (high velocities at positions close to the starting point, uncorrelated activation and excitation...) potentially leading to a failure of the optimization process. The proposed reconstruction error is a more meaningful metric for our use than the traditional mean square error loss but was used *a posteriori* to assess the models performance. An improvement could be found in using it as a training loss as well as in enriching it with velocity criterion.

Second, some interesting questions are related to role of excitation and inhibition in motor cortex [Isomura et al., 2009]. To explore these questions, it could be possible to build and train a network with neurons which are either excitatory or inhibitory. Third, a spatial organization seems to exist in the motor cortex in terms of functional clusters of neurons which might be due to local recurrent circuit dynamics [Hira et al., 2013]. Although fully connected recurrent networks

¹Code can be found here: <https://churchland.zuckermaninstitute.columbia.edu/content/code>

are probably physiologically unrealistic, local recurrent circuits exist in the brain and could be added to a feedforward architecture. Fourth, intracortical microstimulations in motor cortex can elicit complex, behaviorally relevant movements [Brown and Teskey, 2014; Graziano, 2016]. There is no counterpart of this observation in the network. Fifth, we have used pull-push linear muscles to avoid bounds in the numerical optimization process. It should not be too difficult to use instead pairs of agonist-antagonist nonlinear pull-only muscles. The influence of this change on properties of the network is unknown. For instance, Suminski et al. [2015] have shown nonlinear characteristics of muscle such as the force/length and force/velocity relationships are important to account for the temporal evolution of the primary axis of the distribution of preferred directions (their fig. 6 and fig. 7).

4.4.7 Future developments

The proposed neural network took direct and perfect observations of the state of controlled arm as an input. Yet, it is well established that noises and delays plague the CNS (sections 2.3.2 and 2.3.3). In the adopted control theory, a Kalman filter was used to build the best possible estimation of the state of the system in spite of those inner imperfections (sections 2.6.2.3 and B). Such a control/estimation is supported by neuroanatomical evidence [Scott, 2004]. It would be most relevant to extend the approach applied in this work to the estimation process. A first step would be to assess the feasibility of its neural implementation as a neural network: state estimation is a complex process combining observation, dynamics propagation of the estimated state and covariance prediction that would likely require different network architecture and training method. Second should be studied intrinsic properties of the network similarly - but not limited - to directional tuning. A comparison to actual recordings of the neural activity of neuroanatomical structures suspected to be involved in state estimation would then contribute to clarify the neural bases of motor control. neural network model of the estimator

4.4.8 Extension

An interesting extension of this work would be to train a neural network to control a more complex dynamics (e.g. a 7-dof arm in the 3D space; Guigon et al. [2007b]). The network would alleviate the huge computational burden of computing online optimal control solutions. It would open new avenues for the efficient and realistic control of complex systems for physic-based character animation [Wang et al., 2012] or real robots [Taïx et al., 2013].

Chapter 5

Discussion and Perspectives

In this thesis we have obtained two important results. First, adaptation to a force field perturbation is not properly described by a model which maximizes performance in the presence of the perturbation (chapter 3; reoptimization model; Izawa et al. 2008). Second, a feedforward multilayered neural network can be trained to approximate an optimal controller and the resulting distributed computation is reminiscent of corresponding operations in the motor cortex (chapter 4). We have also provided preliminary results on the problem of position vs force control, and the role of cocontraction (Appendix E, F). We have already discussed at length these results in the respective chapters. Here we step back and assess their far-reaching implications.

5.1 Motor control involves a universal controller

The starting point of our work is the statement that motor control is governed by a task-independent, universal controller dedicated to mastering body dynamics. This means that, as we encounter a new environment with its specific dynamics, we consider it as a perturbation and interact with it through feedback control. The only flexibility is offered by the task representation, i.e. the way we set boundary conditions on body states to reach intended goals. Consider the case of stick balancing. A first solution would be to control the ensemble body+stick. In this case, goals can be set directly on the state of the stick (e.g. its angle relative to the vertical, its velocity, ...) and control can achieve these goals. A second solution is to control only the body. Goals can only be set on the body, e.g. how to displace the hand to keep the stick near the vertical. Here, the task representation is the indirect definition of body goals from the state of the stick. We claim that the latter solution is closer to a human strategy than the former. This claim could be tested in a stick balancing experiment, either with a real stick [Milton et al., 2016] or in a virtual environment [Yoshikawa et al., 2016]. As with the adaptation experiment

(chapter 3), we would compare experimental data with model's simulations and try to account for specific properties of stick balancing (e.g. non-Gaussian fluctuations in the acceleration distribution of the stick, power-law distributions of corrective movements, ...). This would constitute another strong challenge to our claim about a universal controller. A central argument for the universal controller is the striking contrast between the long road of motor development (years; [Thelen 1995](#)) and the brief timescale of adaptation processes (minutes to hours; [Shadmehr and Mussa-Ivaldi 1994](#)). Although it is difficult to have some certainty on this issue, we can question the fact that the brain could elaborate a full-blown new controller for each task at hand. Interestingly, the study of [de Rugy et al. \[2012\]](#) which is often taken to argue against optimal motor control models, for once, would be consistent with our view. They showed that human participants failed to reoptimize their muscle recruitment patterns following (virtual) changes in muscle actions. They interpreted their results by the existence of "habitual" coordination patterns that are unaffected by selective modifications of the peripheral apparatus. Our model suggests that these habitual patterns are learned optimal patterns generated by the universal controller.

5.2 Optimality and the content of action

A central claim in computational motor control is optimality [[Scott, 2004](#); [Todorov, 2004](#); [Diedrichsen et al., 2010](#)]. Accordingly, our movements should be optimal for some cost function, although the exact nature of this function remains debated [[Berret et al., 2011](#)]. Yet this view gives only an incomplete account of movement production. In fact, most movements are unlikely to be optimal for any of the cost functions proposed in the literature. Although strong signal filtering might give an impression that smoothness is the very essence of movement [[Flash and Hogan, 1985](#)], velocity and acceleration fluctuations on a fast timescale (e.g. 8 Hz) are a true feature of a large repertoire of movements [[Vallbo and Wessberg, 1993](#); [Guigon et al., 2019, 2022](#)]. Even fast movements such as those that we have recorded in the force field adaptation task contain such fluctuations (chapter 3, 3.2, C.13). The failure of an optimality principle for movement is no surprise: the ubiquitous encounter of isochronous behaviors [[Denier van der Gon and Thuring, 1965](#); [Lacquaniti et al., 1983](#); [Kodl et al., 2011](#)] is already an argument against optimality per se. These observations call into question the appropriateness of optimal control models applied to the understanding of human movements. The problem is probably benign in general but might become serious when dealing with detailed characteristics of movement kinematics or precision. Yet it says nothing about optimal planning and optimal energy expenditure [[Aftalion and Trélat, 2021](#); [Brown et al., 2021](#)].

Neural bases of motor control An ultimate goal of motor control modeling is to improve our understanding of the neural bases of motor control. A frequent argument against the optimal control approach is its computational burden and the associated difficulty to imagine how a neuronal

tissue could produce such complex operations. We have made a decisive step towards addressing this issue: (1) a simple feedforward neural network can be trained to produce a nonlinear optimal feedback control policy; (2) properties of the network match those of the primate motor cortex. Although much work remains to be done at the machine learning level (structure and size of the network, training set, robustness, ...), these observations are very satisfactory. In fact, part of the success is probably due to the nature of the control policy: a stationary, parameter-free, universal control policy. Thus, it gives additional support to our modeling hypotheses [Guigon et al., 2019; Guigon, 2022; Moullet et al., 2022].

5.3 Perspectives

This work offers two broad perspectives. The first perspective is to extend the proposed framework to a broader range of tasks and to more complex dynamics. Interesting tasks are those in which an unmodeled dynamics is present (e.g. stick balancing; see above). We expect that these tasks can be handled properly through an adequate choice of a task representation, without the need to integrate the unmodeled dynamics in the controller. For instance, working with an upper limb prosthesis controlled through compensatory body motions [Legrand et al., 2021] has been successfully modeled in our framework (results not shown). The second perspective is to deepen the neural bases of motor control. The proposed computational neuroanatomy of motor control involves not only the motor cortex but also some other parts of the cerebral cortex, the basal ganglia and the cerebellum [Shadmehr and Krakauer, 2008; Haar and Donchin, 2020]. In particular, the role of the cerebellum is within reach through a neural network trained to perform state estimation.

Appendices

Appendix A

Optimal Control

A.1 Analytic optimization for linear problems

A.1.1 Statement of the problem

We consider a system described by a state \mathbf{x} driven by a dynamic \mathbf{f} on which we have a control \mathbf{u} :

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t], \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (\text{A.1})$$

We seek to find an optimal control $\mathbf{u}(t)$ minimizing a cost-function over a given time lapse, taking the form:

$$\mathcal{J} = \phi[\mathbf{x}(t_f), t_f] + \int_{t_0}^{t_f} \mathcal{L}[\mathbf{x}(t), \mathbf{u}(t), t] dt \quad (\text{A.2})$$

Under the terminal constraints:

$$\psi[\mathbf{x}(t_f), t_f] = 0 \quad (\text{A.3})$$

$$\phi[\mathbf{x}(t + t_H), t_H] = 0 \quad (\text{A.4})$$

Note: as we focus on a linear problem, no inequality constraints are taken into account here.

A.1.2 Adjoint problem

The system dynamics may be written as the time-varying constraint, :

$$\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] - \dot{\mathbf{x}}(t) = \mathbf{0} \quad (\text{A.5})$$

which is to be satisfied over the entire time interval $[t_0, t_f]$ and may therefore be adjoined to the integrand of \mathcal{J} , using the time-varying Lagrange coefficient $\boldsymbol{\lambda}(t)$, as well as the terminal constraints, using $\boldsymbol{\nu}(t)$:

$$\begin{aligned} \mathcal{J}_A = & \phi[\mathbf{x}(t_f), t_f] + \boldsymbol{\nu}^T(t) \psi[\mathbf{x}(t_f), t_f] \\ & + \int_{t_0}^{t_f} \mathcal{L}[\mathbf{x}(t), \mathbf{u}(t), t] + \boldsymbol{\lambda}^T(t) (\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] - \dot{\mathbf{x}}(t)) dt \end{aligned} \quad (\text{A.6})$$

Defining the Hamiltonian \mathcal{H} :

$$\mathcal{H}[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] = \mathcal{L}[\mathbf{x}(t), \mathbf{u}(t), t] + \boldsymbol{\lambda}^T(t) \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] \quad (\text{A.7})$$

we write

$$\begin{aligned} \mathcal{J}_A = & \phi[\mathbf{x}(t_f), t_f] + \boldsymbol{\nu}^T(t) \psi[\mathbf{x}(t_f), t_f] \\ & + \int_{t_0}^{t_f} \mathcal{H}[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] - \boldsymbol{\lambda}^T(t) \dot{\mathbf{x}}(t) dt \end{aligned} \quad (\text{A.8})$$

and integrating by part:

$$\begin{aligned} \mathcal{J}_A = & \phi[\mathbf{x}(t_f), t_f] + \boldsymbol{\nu}^T(t) \psi[\mathbf{x}(t_f), t_f] + [\boldsymbol{\lambda}^T(t_0) \mathbf{x}(t_0) - \boldsymbol{\lambda}^T(t_f) \mathbf{x}(t_f)] \\ & + \int_{t_0}^{t_f} \mathcal{H}[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t] + \dot{\boldsymbol{\lambda}}^T(t) \mathbf{x}(t) dt \end{aligned} \quad (\text{A.9})$$

A.1.3 Conditions for optimality

A necessary condition for optimality is that the first-order of control variations on the cost function be zero throughout the time interval. First variations of \mathcal{J}_A can be expressed as:

$$\begin{aligned} \Delta \mathcal{J}_A &= \left\{ \boldsymbol{\lambda}^T \Delta \mathbf{x}(\Delta \mathbf{u}) \right\} \Big|_{t=t_0} + \int_{t_0}^{t_f} \left\{ \frac{\partial \mathcal{H}}{\partial \mathbf{u}} \Delta \mathbf{u} + \left[\partial \mathcal{H} \mathbf{x} - \dot{\boldsymbol{\lambda}}^T \right] \Delta \mathbf{x}(\Delta \mathbf{u}) \right\} dt \\ &+ \left\{ \left[\frac{\partial \phi}{\partial \mathbf{x}} + \boldsymbol{\nu}^T \frac{\partial \psi}{\partial \mathbf{x}} - \boldsymbol{\lambda}^T \right] \Delta \mathbf{x}(\Delta \mathbf{u}) \right\} \Big|_{t=t_f} \\ &\triangleq \Delta \mathcal{J}_A(t_0) + \Delta \mathcal{J}_A(t_0, t_f) + \Delta \mathcal{J}_A(t_f) \end{aligned} \quad (\text{A.10})$$

where $\Delta \mathbf{u}(t)$, $t_0 \leq t \leq t_f$ is an arbitrary (presumably small) function, $\Delta \mathbf{x}(\Delta \mathbf{u})$ is a functional that denotes state perturbations arising from control perturbations and Δt_f is a change in final time. As stated, in the vicinity of the optimal trajectory, the first variations on the cost function must equal zero, the four parts of $\Delta \mathcal{J}_A$ must equal zero too:

- It can be assumed that the initial state is unaffected by the initial control, which means $\Delta\mathcal{J}_A(t_0) = 0$ regardless of the adjoint vector.
- We wish to choose the adjoint vector time history such that $\Delta\mathcal{J}_A$ is insensitive to arbitrary (nonzero) values of $\Delta\mathbf{u}$ and $\Delta\mathbf{x}(\Delta\mathbf{u})$ in the time interval:

$$\dot{\boldsymbol{\lambda}}^T(t) = -\frac{\partial\mathcal{H}[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t]}{\partial\mathbf{x}}, \quad t_0 \leq t \leq t_f \quad (\text{A.11})$$

subject to the terminal conditions:

$$\boldsymbol{\lambda}^T(t_f) = \left. \frac{\partial\phi[\mathbf{x}(t), t]}{\partial\mathbf{x}} \right|_{t=t_f} + \boldsymbol{\nu}^T \left. \frac{\partial\psi[\mathbf{x}(t), t]}{\partial\mathbf{x}} \right|_{t=t_f} \quad (\text{A.12})$$

- $\Delta\mathcal{J}_A(t_f)$ is zero by the previous equation.
- For $\Delta\mathcal{J}_A(t_0, t_f)$, to be zero, in addition to equation (A.11), the following must be satisfied:

$$\frac{\partial\mathcal{H}[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t]}{\partial\mathbf{u}} = 0 \quad (\text{A.13})$$

Equations (A.11) (A.12) and (A.13) are the well-known Euler-Lagrange equations.

Denoting $\frac{\partial\mathbf{f}}{\partial\mathbf{x}}$ and $\frac{\partial\mathbf{f}}{\partial\mathbf{u}}$ by \mathbf{F} and \mathbf{G} , equations (A.11) and (A.13) become:

$$\dot{\boldsymbol{\lambda}}(t) = -\mathbf{F}^T(t)\boldsymbol{\lambda}(t) - \left[\frac{\partial\mathcal{L}}{\partial\mathbf{x}} \right]^T \quad (\text{A.14})$$

$$\left[\frac{\partial\mathcal{L}}{\partial\mathbf{u}} \right]^T + \mathbf{G}^T(t)\boldsymbol{\lambda}(t) = 0 \quad (\text{A.15})$$

A.1.4 Application to a simple case

A.1.4.1 Dynamics

For both the sake of illustration and simplification we study a simple system consisting in a 1D-mass m described by a state \mathbf{x} driven by a muscle-like actuation dynamic \mathbf{f} on which we have a control \mathbf{u} where:

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} = \begin{bmatrix} \text{position} \\ \text{velocity} \\ \text{activation} \\ \text{excitation} \end{bmatrix} \quad (\text{A.16})$$

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), t] = \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dot{x}_4(t) \end{bmatrix} = \begin{bmatrix} x_2(t) \\ \frac{x_3(t)}{m} \\ \frac{x_4(t) - x_3(t)}{\tau} \\ \frac{u(t) - x_4(t)}{\tau} \end{bmatrix} \quad (\text{A.17})$$

or

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), t] = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) \quad (\text{A.18})$$

with

$$\mathbf{F} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{m} & 0 \\ 0 & 0 & -\frac{1}{\tau} & \frac{1}{\tau} \\ 0 & 0 & 0 & -\frac{1}{\tau} \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\tau} \end{bmatrix} \quad (\text{A.19})$$

A.1.4.2 Optimization problem

We seek to find a control $\mathbf{u}(t)$ minimizing the cost-function:

$$\mathcal{J} = \int_{t_0}^{t_f} \mathcal{L}[\mathbf{u}(t), t] d\tau \quad (\text{A.20})$$

Which means solving the problem

$$\mathcal{U}_{[t_0, t_f]} = \arg \min_{\mathbf{u}} \int_{t_0}^{t_f} \mathcal{L}[\mathbf{x}(\tau), \mathbf{u}(\tau), \tau] d\tau \quad (\text{A.21})$$

We define an "ecological" cost-function as the square sum of the control coordinates:

$$\mathcal{L}[\mathbf{x}(\tau), \mathbf{u}(\tau), \tau] = \mathbf{u}^T(\tau) \mathbf{u}(\tau) \quad (\text{A.22})$$

We set the terminal constraints:

$$\phi[\mathbf{x}(t_f), t_H] = \mathbf{c} \odot (\mathbf{x}(t_f) - \mathbf{x}_f) = \mathbf{0} \quad (\text{A.23})$$

that can be written as:

$$\begin{bmatrix} c_1(x_1(t_f) - x_{1f}) \\ c_2(x_2(t_f) - x_{2f}) \\ c_3(x_3(t_f) - x_{3f}) \\ c_4(x_4(t_f) - x_{4f}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (\text{A.24})$$

where the coefficients $c_i = \{0, 1\}$, 1 denoting that the i th coordinate of the state is constrained to the value x_{if} and 0 denoting that it is set free to be optimized. We hereby define \mathbf{c}_f^c as the vectorised concatenation of the terminally constrained coordinates.

A.1.5 Solving Euler-Lagrange equations

Applying (A.12) to our case, defined in (A.22), we find:

$$\boldsymbol{\lambda}(t_f) = \begin{bmatrix} \lambda_1(t_f) \\ \lambda_2(t_f) \\ \lambda_3(t_f) \\ \lambda_4(t_f) \end{bmatrix} = \begin{bmatrix} c_1\nu_1 \\ c_2\nu_2 \\ c_3\nu_3 \\ c_4\nu_4 \end{bmatrix} \quad (\text{A.25})$$

Applying (A.14) to our case, defined in (A.22), we find:

$$\dot{\boldsymbol{\lambda}}(t) = -\mathbf{F}^T(t)\boldsymbol{\lambda}(t) \quad (\text{A.26})$$

Applying (A.15) to our case, defined in (A.22), we find:

$$\mathbf{u}(t) = -\frac{\mathbf{G}^T\boldsymbol{\lambda}(t)}{2} = -\frac{\lambda_4(t)}{2\tau} \quad (\text{A.27})$$

Which leads to express (A.17) as:

$$\dot{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(t) - \frac{\mathbf{G}\mathbf{G}^T}{2}\boldsymbol{\lambda}(t) \quad (\text{A.28})$$

We now merge (A.28) and (A.26):

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\boldsymbol{\lambda}} \end{bmatrix} = \begin{bmatrix} \mathbf{F} & -\frac{\mathbf{G}\mathbf{G}^T}{2} \\ 0 & -\mathbf{F}^T \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} \quad (\text{A.29})$$

Which evaluates as:

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\boldsymbol{\lambda}} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{m} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{\tau} & \frac{1}{\tau} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{\tau} & 0 & 0 & 0 & -\frac{1}{2\tau^2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{m} & \frac{1}{\tau} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau} & \frac{1}{\tau} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} \quad (\text{A.30})$$

This differential system finds a solution that we represent as a transition matrix $\Phi(t)$, defined by four sub-matrices:

$$\begin{bmatrix} \mathbf{x}(t) \\ \boldsymbol{\lambda}(t) \end{bmatrix} = \Phi(t) \begin{bmatrix} \mathbf{x}(t_0) \\ \boldsymbol{\lambda}(t_0) \end{bmatrix} = \begin{bmatrix} \Phi_{11}(t) & \Phi_{12}(t) \\ \Phi_{21}(t) & \Phi_{22}(t) \end{bmatrix} \begin{bmatrix} \mathbf{x}(t_0) \\ \boldsymbol{\lambda}(t_0) \end{bmatrix} \quad (\text{A.31})$$

As we know the values of $\mathbf{x}(t_0)$, we now wish to express $\boldsymbol{\lambda}(t_0)$ as a function of the initial state \mathbf{x}_0 and the constrained final state \mathbf{x}_f .

Evaluating (A.31) at $t = t_0$, we find:

$$\mathbf{x}(t_f) = \begin{bmatrix} \Phi_{11}(t_f) & \Phi_{12}(t_f) \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \boldsymbol{\lambda}(t_0) \end{bmatrix} = \Phi_{11}(t_f)\mathbf{x}_0 + \Phi_{12}(t_f)\boldsymbol{\lambda}(t_0) \quad (\text{A.32})$$

Thus:

$$\boldsymbol{\lambda}(t_0) = \Phi_{12}^{-1}(t_f) (\mathbf{x}(t_f) - \Phi_{11}(t_f)\mathbf{x}_0) \quad (\text{A.33})$$

The coordinates of $\mathbf{x}(t_f)$ are known to be equal to those of \mathbf{x}_f for indexes i for which $c_i = 1$. Conversely, we thus need to find $x_j(t_f)$ for the m indexes j for which $c_j = 0$. Fortunately, for each unknown $x_j(t_f)$, (A.25) gives an equation constraining the corresponding $\lambda_j(t_f)$.

Evaluating (A.31) at $t = t_f$ once again, we find:

$$\boldsymbol{\lambda}(t_f) = \begin{bmatrix} \Phi_{21}(t_f) & \Phi_{22}(t_f) \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \boldsymbol{\lambda}(t_0) \end{bmatrix} = \Phi_{22}(t_f)\boldsymbol{\lambda}(t_0) \quad (\text{A.34})$$

Thus:

$$\boldsymbol{\lambda}(t_0) = \Phi_{22}^{-1}(t_f)\boldsymbol{\lambda}(t_f) \quad (\text{A.35})$$

Combining (A.33) and (A.35), we find:

$$\Phi_{12}^{-1}(t_f) (\mathbf{x}(t_f) - \Phi_{11}(t_f)\mathbf{x}_0) = \Phi_{22}^{-1}(t_f)\boldsymbol{\lambda}(t_f) \quad (\text{A.36})$$

Which, in this 1D-mass example is a 4-dimension system, with $2 * m$ unknowns $x_j(t_f)$ and $\lambda_i(t_f)$ (where i are the indexes for which $c_i = 1$ and j are the indexes for which $c_j = 0$). This systems can be solved in the form:

$$\begin{bmatrix} \mathbf{x}_j(t_f) \\ \boldsymbol{\lambda}_i(t_f) \end{bmatrix} = \Gamma(t_f) \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_f^c \end{bmatrix} \quad (\text{A.37})$$

Knowing now perfectly $\mathbf{x}(t_f)$, we can evaluate λ_0 with (A.33) and the optimization problem is solved with (A.31).

A.2 Stationarity with receding horizon

An important feature of the optimal control derived in A.27 is that it depends only on x_0 and x_f^c :

$$\begin{aligned} \mathbf{u}(t) &= -\frac{\mathbf{G}^T \boldsymbol{\lambda}(t)}{2} = -\frac{\mathbf{G}^T (\Phi_{21}(t)\mathbf{x}_0) + \Phi_{21}(t)\boldsymbol{\lambda}_0(t_0))}{2} \\ &= -\frac{\mathbf{G}^T (\Phi_{21}(t)\mathbf{x}_0 + \Phi_{21}(t)\Phi_{12}^{-1}(t_f) (\mathbf{x}(t_f) - \Phi_{11}(t_f)\mathbf{x}_0))}{2} \end{aligned} \quad (\text{A.38})$$

Which can be written as:

$$\mathbf{u}(t) = \boldsymbol{\psi}(\mathbf{x}_0, \mathbf{x}_f^c, t, t_H) \quad (\text{A.39})$$

$$\mathbf{u}(t) = \boldsymbol{\Psi}(t, t_H) \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_f^c \end{bmatrix} \quad (\text{A.40})$$

And there lies the power of optimal control with receding horizon: at each time step the control is computed with the same time limit $t_f = 0.28$ s, and only the initial control is used. In this situation, switching to a discrete representation of task time (as opposed to the previous time t that refers to the optimizing time), each control is in the linear case given by

$$\mathbf{u}_t = \boldsymbol{\Psi}_0(t_H) \begin{bmatrix} \mathbf{x}_t \\ \mathbf{x}_f^c \end{bmatrix} \quad (\text{A.41})$$

and

$$\mathbf{u}_t = \boldsymbol{\psi}(\mathbf{x}_t, \mathbf{x}_f^c, t_H) \quad (\text{A.42})$$

Appendix B

State Estimation

B.1 State estimation and motion

Human perception is neither instantaneous nor perfect. In addition, both motor controls effectively delivered to motor units, as well as the consequential motor outputs may differ from those intended during the action selection process. But regardless of those imperfections the motor control process remains robust, and motor tasks can generally be achieved. Interestingly, perception and action imperfections, combined with the many redundancies characterising the human body (as well as, in many cases, the motor task itself) give rise to what is called structured variability. This phenomenon was first described by Bernstein in his founding study of skilled motor behavior: he observed that blacksmiths were able to repeatedly perform a hammering task while never repeating twice the same movement. The randomness of the different errors affecting the motor process is tamed by the control policy to achieve the task at hand, displaying a organisation of kinematic properties over repetition of a given task. The very existence of skilled behavior in spite of perception and action imperfections calls for a necessary functionality in the motor control process: state estimation. The most used implementation of state estimation is the Kalman filter for its capacity to take into account both proportional and additive noises on perception and actions as well as process delay. Although the exact nature and characteristics of perception and action imperfections is yet to be precisely determined, a handy and common description of their probabilistic behavior is white Gaussian noise.

B.2 The extended Kalman filter

As stated, we seek a way of building an estimation of the current state of our system that would be more accurate than its measurements alone. Aside the observations our body can make of itself, the only insightful information our brain has at its disposal is a representation of the dynamic of the body it controls. An efficient tool for fusing different measurements of a system and its ruling dynamic is the Stratonovich–Kalman–Bucy filter (which will be referred to as the Kalman filter, for the sake of conciseness). Although it cannot be claimed that the Kalman is the exact state estimation process effectively neurally implemented in our brain, it is nonetheless a concise and powerful method for simulations, accounting for the global process of estimation.

Qualitatively, the Kalman produces an estimation of the state of system by computing an average between the prediction of the state by the propagation of the dynamic from the previous state and the observations of the state at its disposal. For a noisy process, there will be a noticeable discrepancy between measured and predicted states, coming either from a perturbation of the dynamic of the system or from inaccurate measurements. The Kalman filter makes use of the representation we may have of the "intensity" of the different noises affecting our process in order to assign corresponding "weights" to available measurements and state prediction: if the modeled dynamic is highly noisy and the observations trustworthy, the Kalman filter will mainly rely on measurements ; conversely, if the dynamic is well known and the sensors of poor accuracy, the Kalman filter will produce an almost feedforward estimate.

While the Kalman is renown for its use over linear dynamics (used to solve LQG problems - at the foundation of control theory), an extended version can be formulated for more general non-linear dynamics:

$$\dot{\mathbf{x}}(t) = \overline{\mathbf{f}}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t)) \quad (\text{B.1})$$

associated with a general observation:

$$\mathbf{y}(t) = \mathbf{H}(\mathbf{x}(t), \mathbf{v}(t)) \quad (\text{B.2})$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ is the state of the system, $\mathbf{u}(t) \in \mathbb{R}^m$ the control signal driving the system and $\mathbf{y}(t) \in \mathbb{R}^p$ the observation of the system and $\mathbf{w}(t)$ and $\mathbf{v}(t)$ represent respectively dynamic and observation noises which are both assumed to be zero mean multivariate Gaussian noises with covariance matrices $\mathbf{\Omega}^w$ and $\mathbf{\Omega}^v$. These equations may be discretized as state transition using first order forward Euler method at time t with timestep δ :

$$\mathbf{x}_{t+1} = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t) \quad (\text{B.3})$$

$$\mathbf{y}_t = \mathbf{H}(\mathbf{x}_t, \mathbf{v}_t) \quad (\text{B.4})$$

Where

$$\mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t) = \mathbf{x}_t + \delta \bar{\mathbf{f}}(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t) \quad (\text{B.5})$$

$$\mathbf{y}_t = \mathbf{H}(\mathbf{x}_t, \mathbf{v}_t) \quad (\text{B.6})$$

As previously stated, the state estimate $\hat{\mathbf{x}}_t$ is built upon prediction through the (potentially wrong) representation of the system's dynamic and (potentially erroneous) observation of the system, fused by the intermediate of the Kalman filter gain \mathbf{K}_t . Its effective calculation is done by forward iteration throughout the whole considered process, keeping in "memory" the occurrences of noises in order to minimize the overall error: the expected value of the square of the magnitude of between the actual state and its estimate $E[\|\mathbf{x} - \hat{\mathbf{x}}\|^2]$. This computation can be decomposed in 4 steps:

$$\begin{aligned} \text{Predicted state estimate} &: \hat{\mathbf{x}}_{t+1|t} &= & \mathbf{f}(\hat{\mathbf{x}}_{t|t}, \mathbf{u}_t) \\ \text{Predicted covariance estimate} &: \mathbf{P}_{t+1|t} &= & \mathbf{A}_t \mathbf{P}_{t|t} \mathbf{A}_t^T + \mathbf{L}_t \boldsymbol{\Omega}^w \mathbf{L}_t^T \\ \text{Innovation} &: \bar{\mathbf{y}}_t &= & \mathbf{y}_t - \mathbf{H}(\hat{\mathbf{x}}_{t+1|t}) \\ \text{Innovation covariance} &: \mathbf{S}_t &= & \mathbf{H}_t \mathbf{P}_{t+1|t} \mathbf{H}_t^T + \mathbf{M}_t \boldsymbol{\Omega}^v \mathbf{M}_t^T \\ \text{Near optimal Kalman gain} &: \mathbf{K}_t &= & \mathbf{P}_{t+1|t} \mathbf{H}_t^T \mathbf{S}_t^{-1} \\ \text{Updated state estimate} &: \hat{\mathbf{x}}_{t+1|t+1} &= & \hat{\mathbf{x}}_{t+1|t} + \mathbf{K}_t \bar{\mathbf{y}}_t \\ \text{Updated covariance estimate} &: \mathbf{P}_{t+1|t+1} &= & (\mathbf{I}_n - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t+1|t} \end{aligned} \quad (\text{B.7})$$

with

$$\mathbf{A}_t = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \quad (\text{B.8})$$

$$\mathbf{B}_t = \frac{\partial \mathbf{H}}{\partial \mathbf{u}} \quad (\text{B.9})$$

$$\mathbf{L}_t = \frac{\partial \mathbf{f}}{\partial \mathbf{w}} \quad (\text{B.10})$$

$$\mathbf{M}_t = \frac{\partial \mathbf{H}}{\partial \mathbf{v}} \quad (\text{B.11})$$

In short, the state estimate can be written as:

$$\hat{\mathbf{x}}_{t+1} = \mathbf{f}(\hat{\mathbf{x}}_t, \mathbf{u}_t) + \mathbf{K}_t (\mathbf{y}_t - \mathbf{H}(\mathbf{f}(\hat{\mathbf{x}}_t, \mathbf{u}_t))) \quad (\text{B.12})$$

or

$$\hat{\mathbf{x}}_{t+1} = \mathbf{f}(\hat{\mathbf{x}}_t, \mathbf{u}_t) - \mathbf{K}_t \mathbf{H}(\mathbf{f}(\hat{\mathbf{x}}_t, \mathbf{u}_t)) + \mathbf{K}_t \mathbf{y}_t \quad (\text{B.13})$$

We can recognise here a form of interpolation,, which becomes even more obvious assuming that the observation \mathbf{H} is linear - *i.e.* $\mathbf{H}(\mathbf{x}_t) = \mathbf{H}_t \mathbf{x}_t$:

$$\hat{\mathbf{x}}_{t+1} = (\mathbf{I}_n - \mathbf{K}_t \mathbf{H}_t) \mathbf{f}(\hat{\mathbf{x}}_t, \mathbf{u}_t) + \mathbf{K}_t \mathbf{y}_t \quad (\text{B.14})$$

B.3 Discretized continuous formulation

When drawing the timestep in Eq. B.7 to infinitely small values, a continuous Kalman gain may be defined upon continuous dynamics (Eq. B.1):

$$\mathbf{K}(t) = \mathbf{P}(t)\mathbf{H}^T[\mathbf{M}(t)\boldsymbol{\Omega}^v\mathbf{M}(t)^T]^{-1} \quad (\text{B.15})$$

where the covariance matrix follows the Riccati Equation :

$$\dot{\mathbf{P}}(t) = \mathbf{A}(t)\mathbf{P}(t) + \mathbf{A}(t)\mathbf{P}(t)^T + \mathbf{L}(t)\boldsymbol{\Omega}^w\mathbf{L}(t)^T - \mathbf{K}(t)\mathbf{H}\mathbf{P}(t) \quad (\text{B.16})$$

and the estimated state propagation is governed by :

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{f}(\hat{\mathbf{x}}(t), \mathbf{u}(t)) + \mathbf{K}(t)[\mathbf{y}(t) - \mathbf{H}\hat{\mathbf{x}}(t)] \quad (\text{B.17})$$

Then, a discretized version may be defined :

$$\mathbf{K}_t = \mathbf{A}_t\mathbf{P}_t\mathbf{H}^T[\mathbf{H}\mathbf{P}_t\mathbf{H}^T + \mathbf{M}_t\boldsymbol{\Omega}^v\mathbf{M}_t^T]^{-1} \quad (\text{B.18})$$

where the covariance matrix follows the Riccati Equation :

$$\mathbf{P}_{t+1} = \mathbf{A}_t\mathbf{P}_t\mathbf{A}_t^T + \mathbf{L}_t\boldsymbol{\Omega}^w\mathbf{L}_t^T - \mathbf{K}_t[\mathbf{H}\mathbf{P}_t\mathbf{H}^T + \mathbf{M}_t\boldsymbol{\Omega}^v\mathbf{M}_t^T]\mathbf{P}_t \quad (\text{B.19})$$

B.4 Delays

To take delay into account, the discrete Kalman filter has to be built around an augmented state storing timesteps until they are processed. Given a state \mathbf{x}_t describing our system at a given time and denoting d as the number of timesteps between an observation and its processing, the feedback is written as:

$$\mathbf{y}_t = \tilde{\mathbf{H}}\tilde{\mathbf{x}}_t + \mathbf{n}_t^o \quad (\text{B.20})$$

where $\tilde{\mathbf{x}}_t = [\mathbf{x}_t; \mathbf{x}_{t-1}; \dots; \mathbf{x}_{t-d}]^T$ is the augmented state and $\tilde{\mathbf{H}} = [\mathbf{H}, \dots, \mathbf{0}]$ is the delayed feedback process. The motor process is then expressed as

$$\tilde{\mathbf{x}}_{t+1} = \tilde{\mathbf{A}}\tilde{\mathbf{x}}_t + \tilde{\mathbf{B}}\tilde{\mathbf{u}}_t + \tilde{\mathbf{n}}_t^m \quad (\text{B.21})$$

with

$$\tilde{\mathbf{A}} = \left[\begin{array}{c|c} \mathbf{A} & \mathbf{0} \\ \hline \mathbf{I}_{n*d} & \mathbf{0} \end{array} \right] \quad (\text{B.22})$$

$$\tilde{\mathbf{B}} = [\mathbf{B}, \dots, \mathbf{0}]^T \quad (\text{B.23})$$

$$\tilde{\mathbf{n}}_t^m = [\mathbf{n}_t^m, \dots, \mathbf{0}]^T \quad (\text{B.24})$$

The augmented estimated state is expressed as:

$$\hat{\mathbf{x}}_{t+1} = \tilde{\mathbf{A}}_t \hat{\mathbf{x}}_t + \tilde{\mathbf{B}} \mathbf{u}_t + \tilde{\mathbf{K}}_t (\mathbf{y}_t - \tilde{\mathbf{B}} \hat{\mathbf{x}}_t) \quad (\text{B.25})$$

where the augmented Kalman gain $\tilde{\mathbf{K}}_t$ is then computed following Eq. B.7 using the augmented matrices:

$$\tilde{\boldsymbol{\Omega}}_t^\xi = \boldsymbol{\Omega}_t^\xi \quad (\text{B.26})$$

$$\tilde{\boldsymbol{\Omega}}_t^\omega = \begin{bmatrix} \boldsymbol{\Omega}_t^\omega \\ \mathbf{0}_{n \times d, n+m} \end{bmatrix} \quad (\text{B.27})$$

Finally, with $\hat{\mathbf{x}}_t = [\hat{\mathbf{x}}_t; \hat{\mathbf{x}}_{t-1}; \dots; \hat{\mathbf{x}}_{t-d}]^T$ the last component $\hat{\mathbf{x}}_{t-d}$ is used to compute the next control command.

Appendix C

Supplementary Materials to Chapter 3

The following figures constitute supplementary materials for Chapter 3

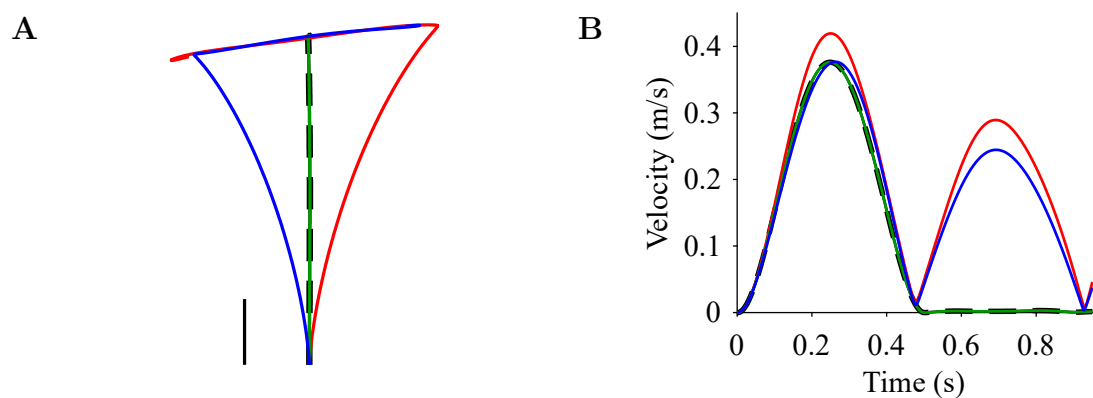


FIGURE C.1: **Predictions of the compensation model.** **A.** Simulated trajectories. **B.** Simulated velocity profiles.

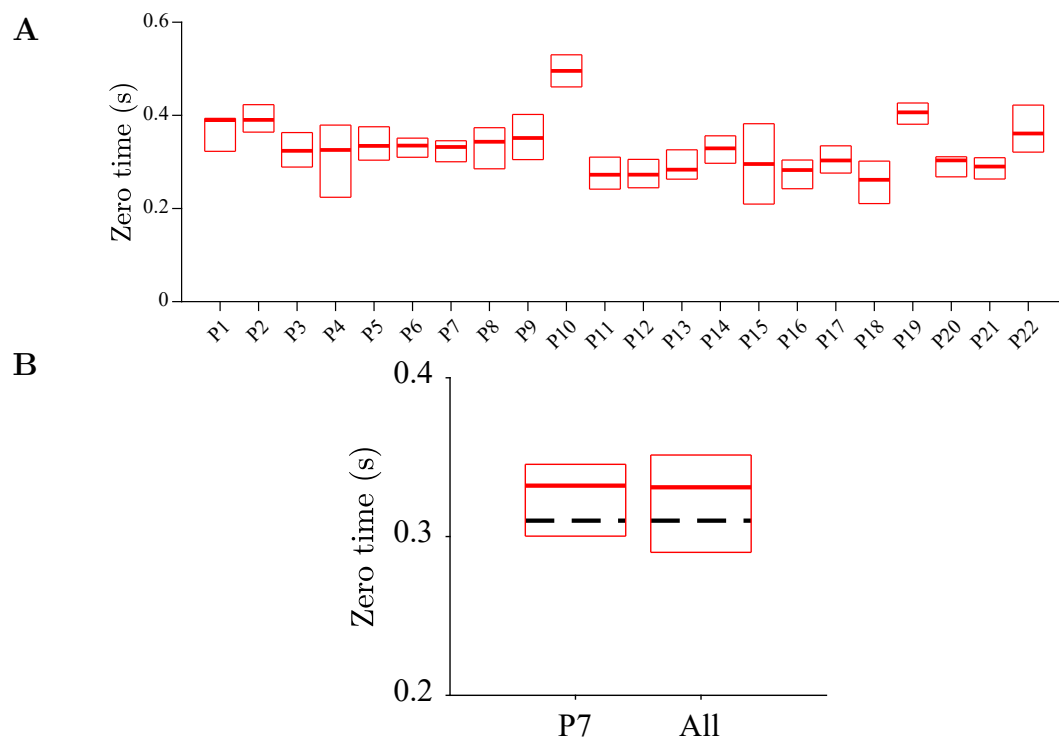


FIGURE C.2: **Time at which the angle derivative of the before-effect trajectory becomes positive.** **A.** All participants with mean value (thick line) and 25-75 percentiles (box). **B.** Data of participant P7 and mean of all the participants. The black dashed line is the model prediction.

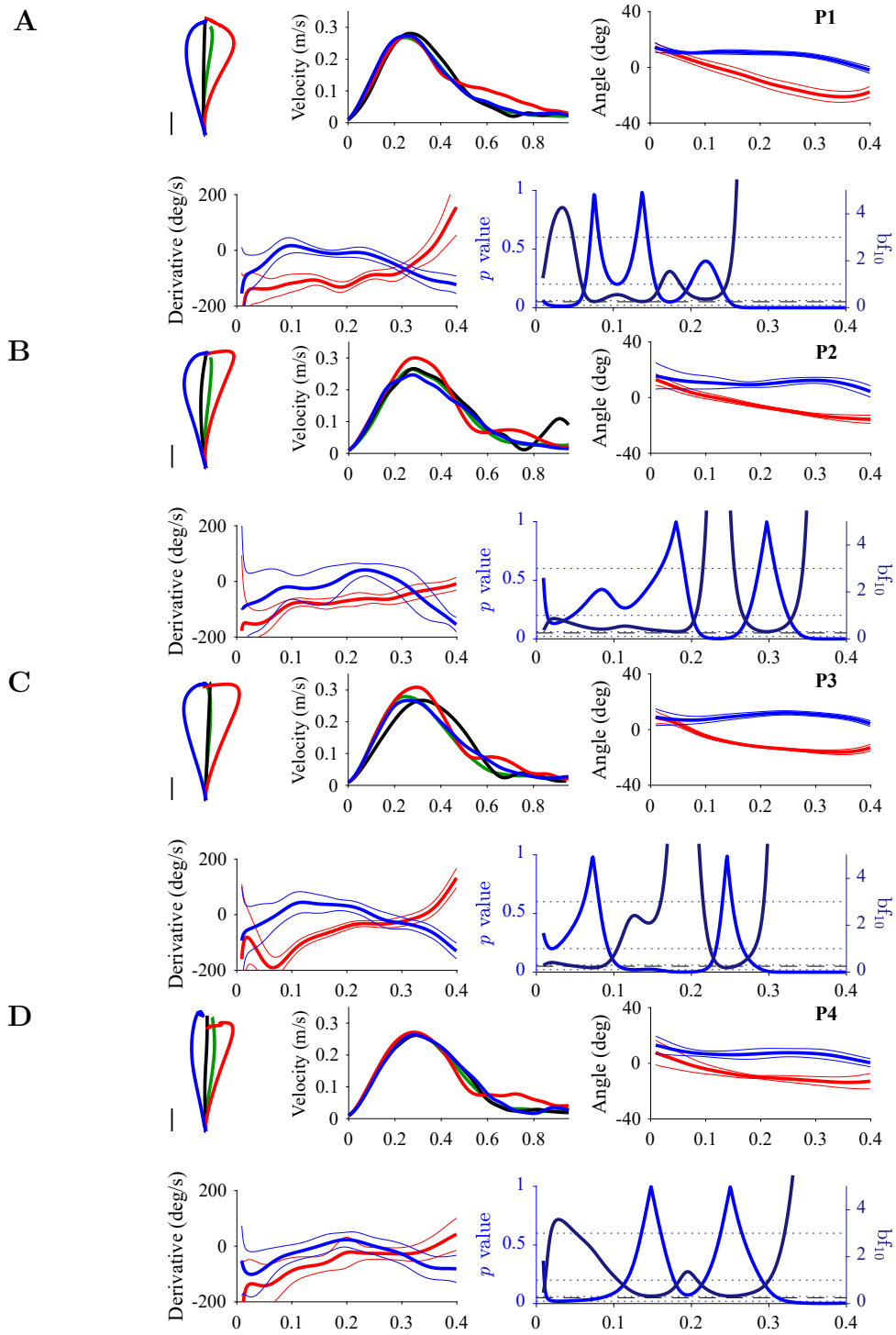


FIGURE C.3: **Participants whose behavior is incompatible with the reoptimization model (1/3).** Same format as in Fig. 3.4. For b_{f10} , the dotted lines correspond, from bottom to top, to substantial =, anecdotal =, anecdotal \neq , and substantial \neq .

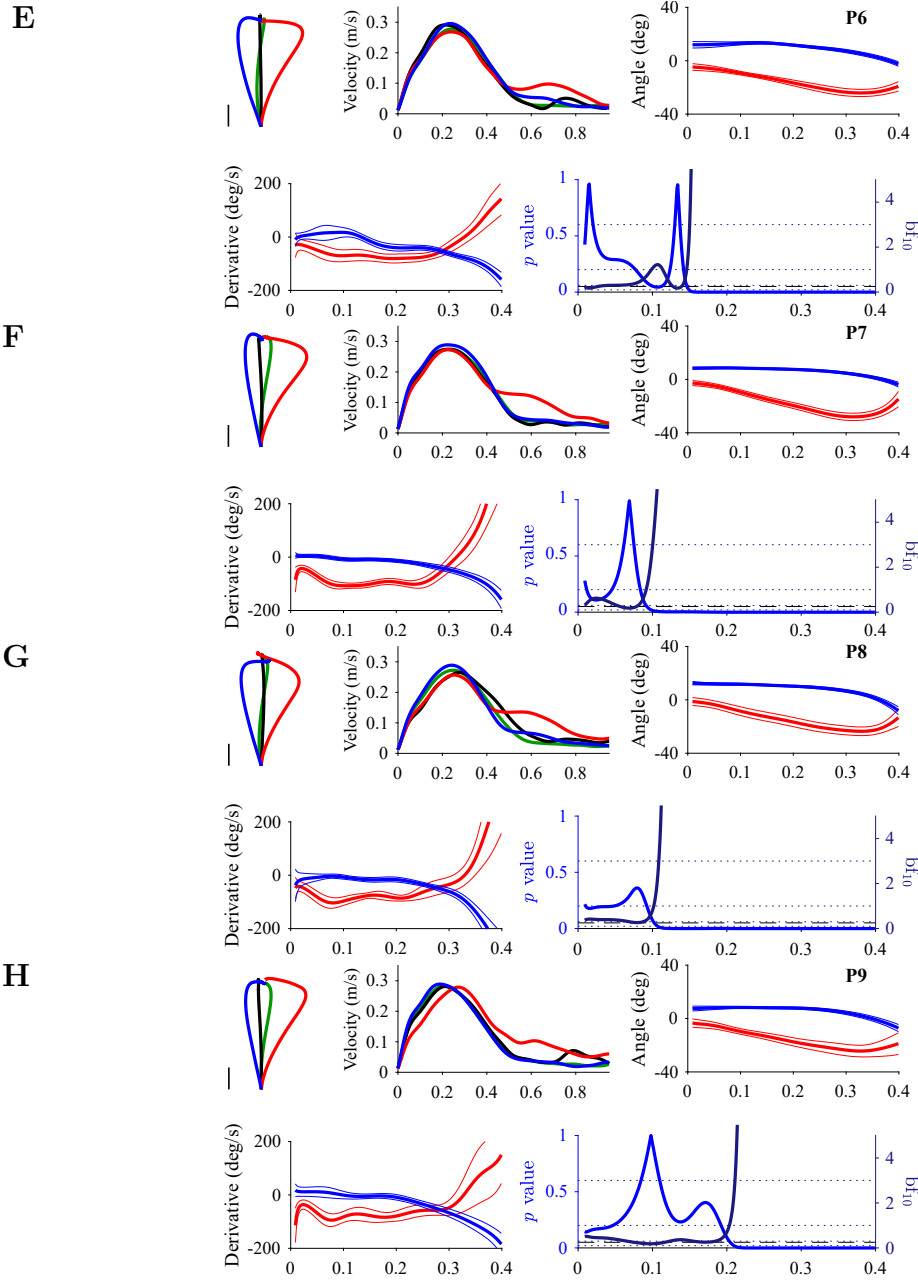


FIGURE C.4: **Participants whose behavior is incompatible with the reoptimization model (2/3).** Same format as in Fig. 3.4. For bf_{10} , the dotted lines correspond, from bottom to top, to substantial =, anecdotal =, anecdotal \neq , and substantial \neq .

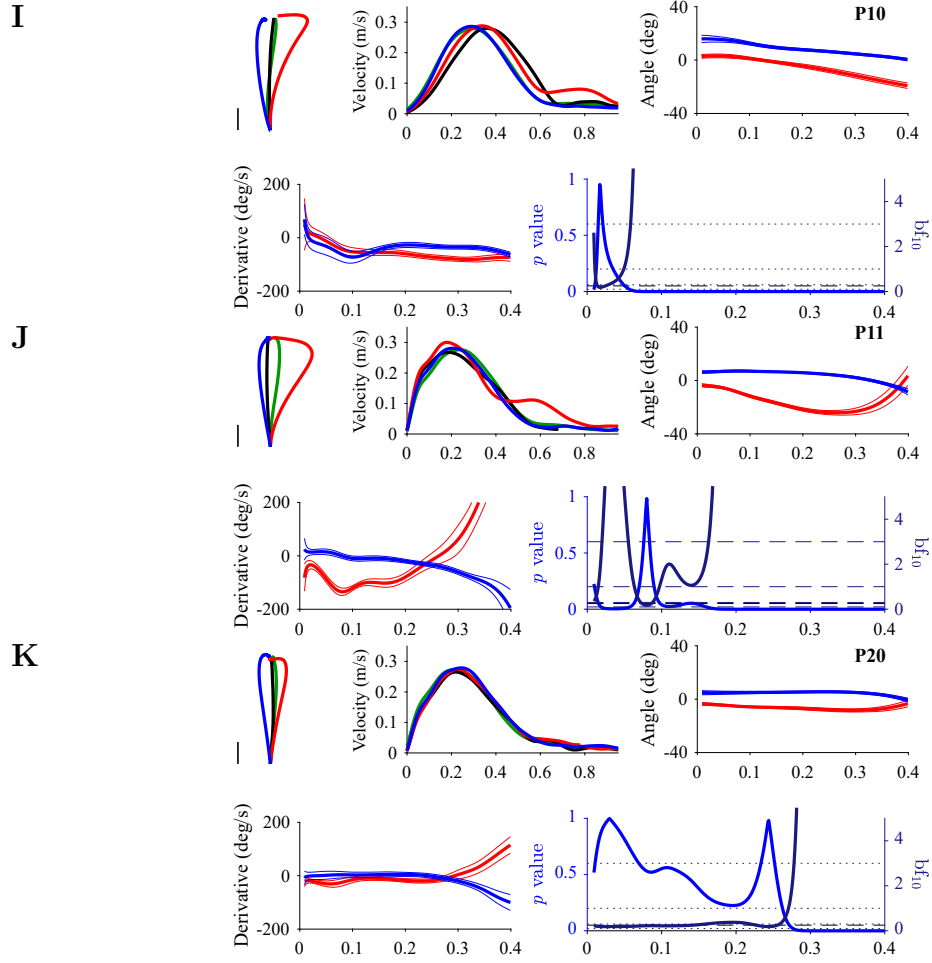


FIGURE C.5: **Participants whose behavior is incompatible with the reoptimization model (3/3).** Same format as in Fig. 3.4. For bf_{10} , the dotted lines correspond, from bottom to top, to substantial =, anecdotal =, anecdotal \neq , and substantial \neq .

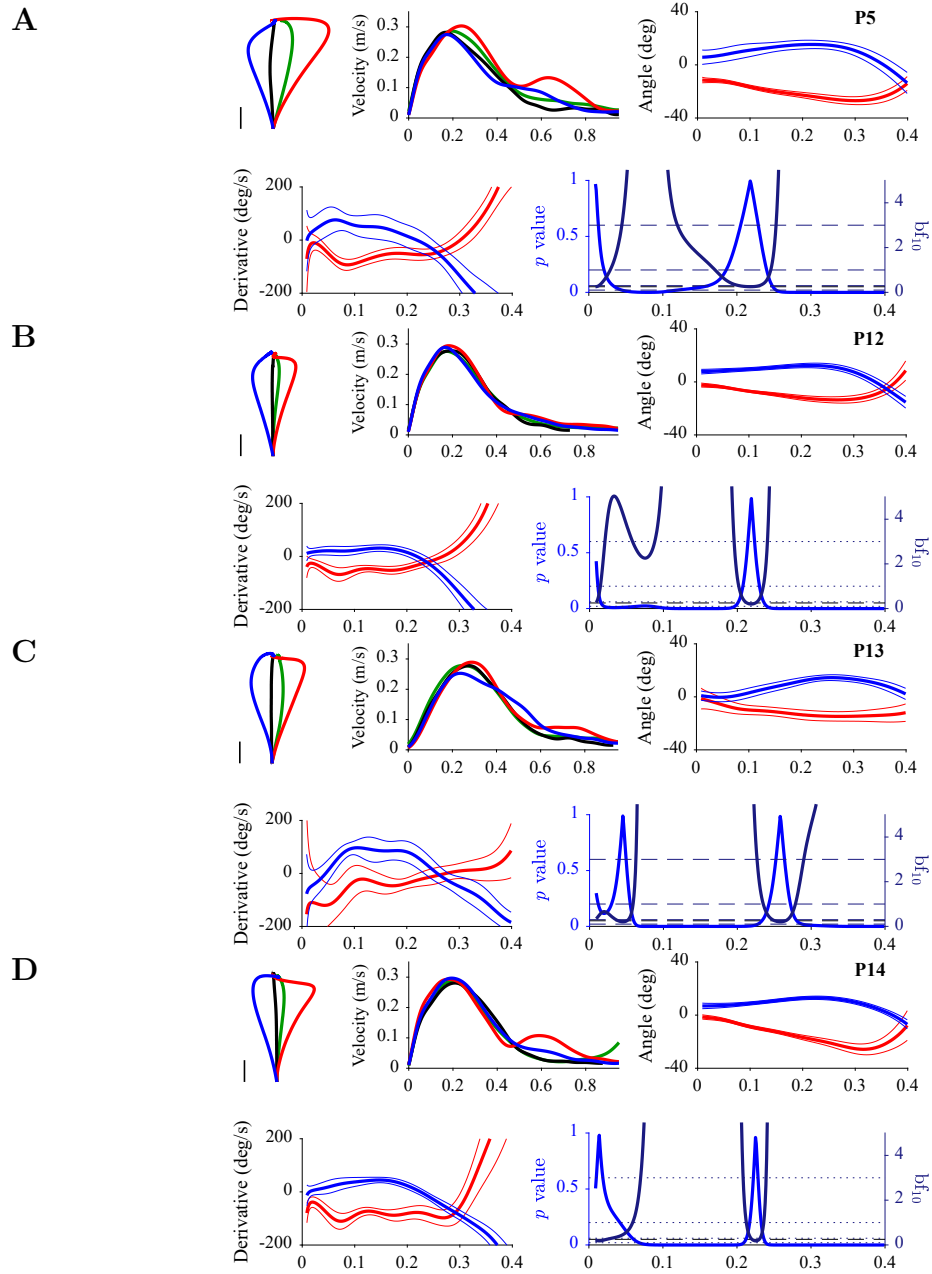


FIGURE C.6: Participants whose behavior is partially compatible with the reoptimization model (1/2). Same format as Fig. C.3

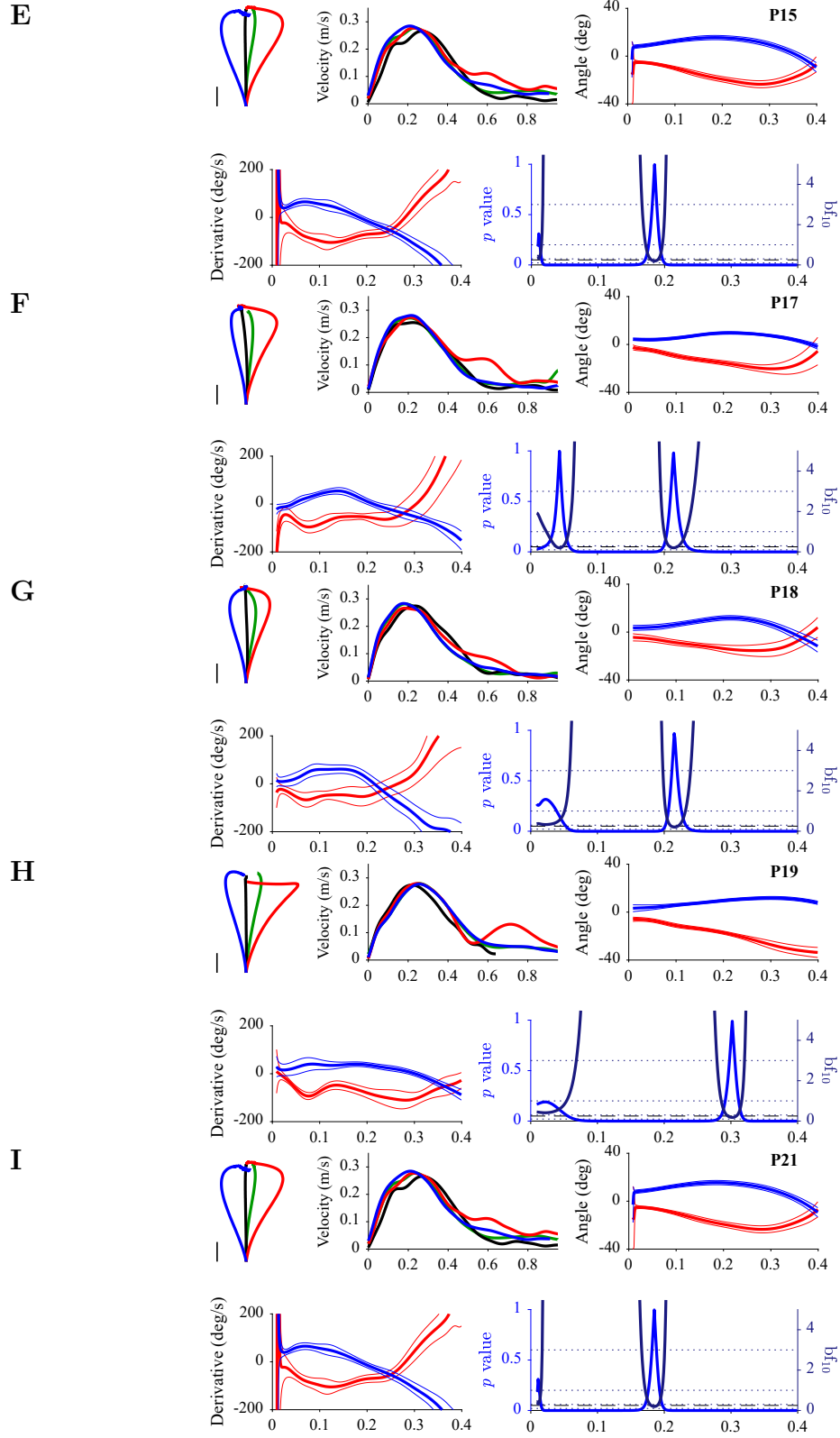


FIGURE C.7: Participants whose behavior is partially compatible with the reoptimization model (2/2). Same format as Fig. C.3

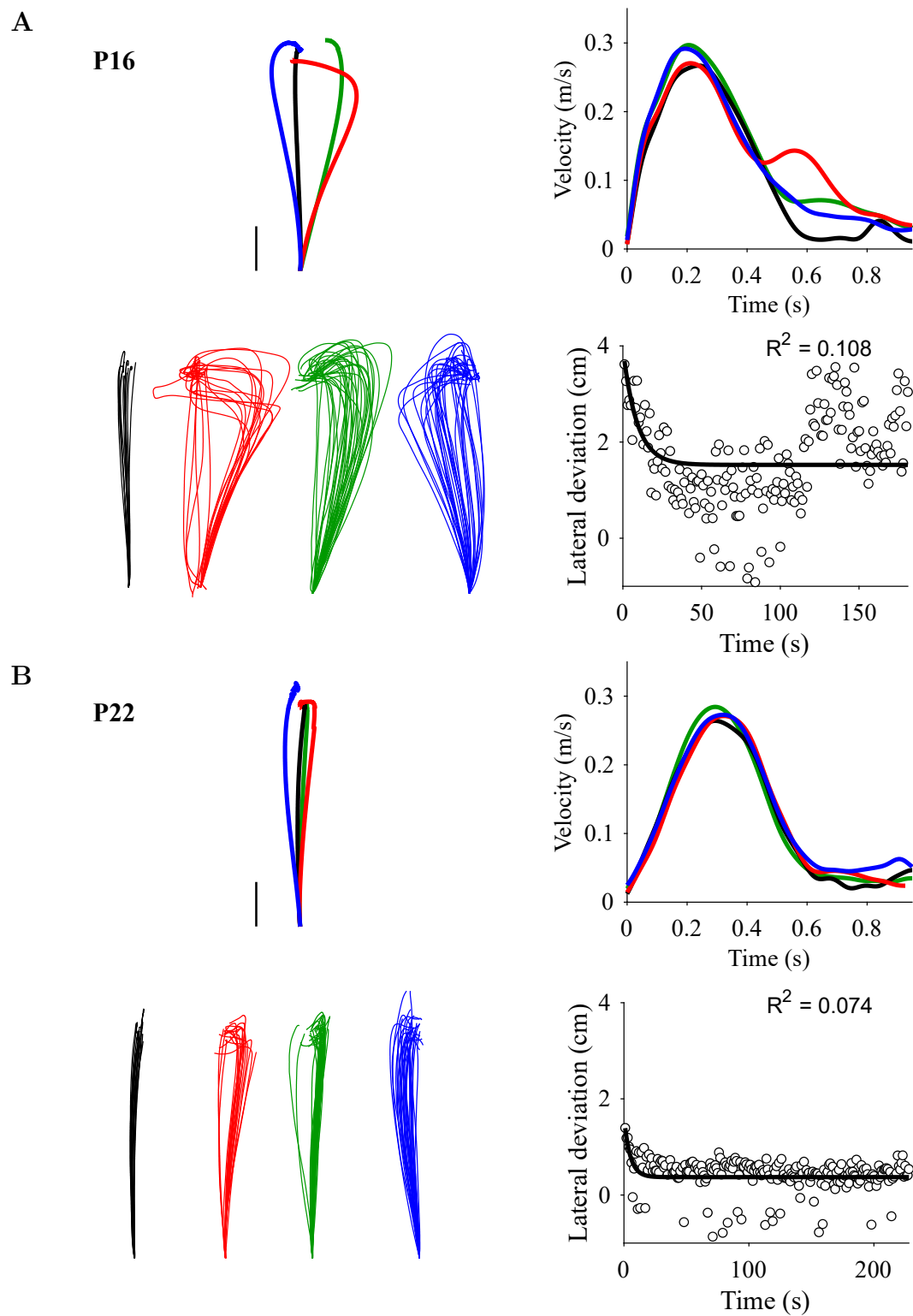


FIGURE C.8: Two participants that failed to improve their behavior with training. Same format as Fig. 3.4.

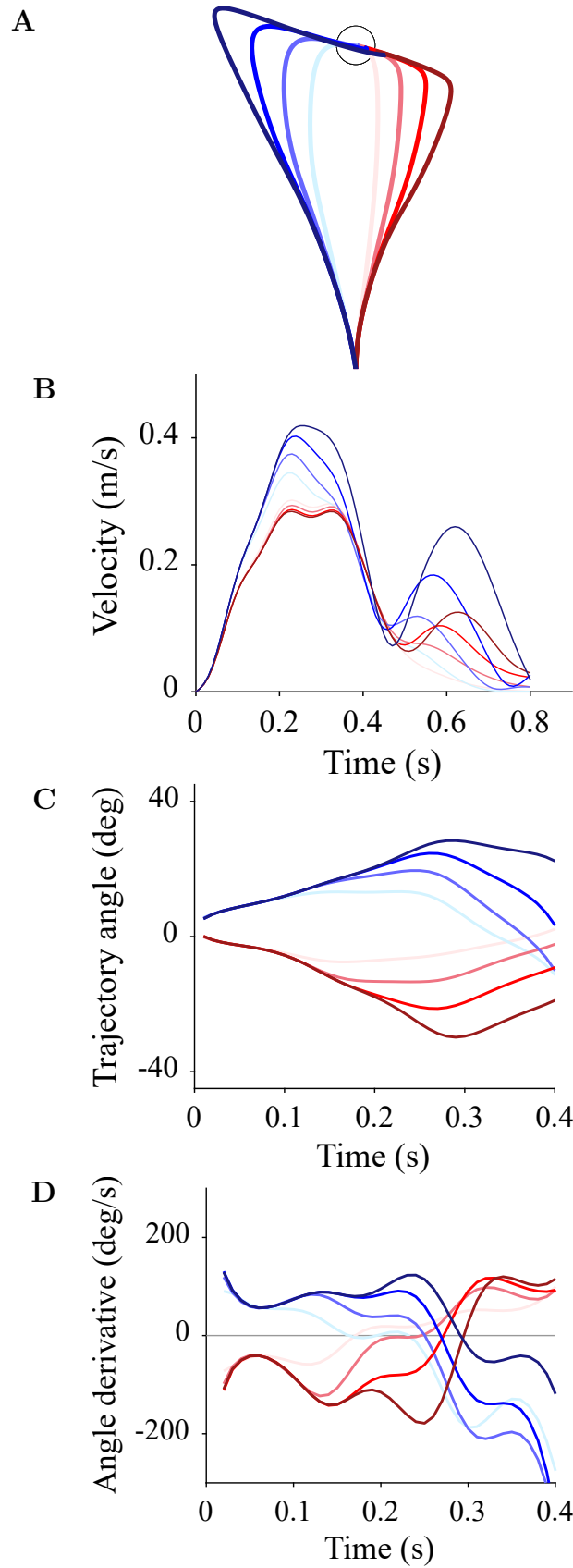


FIGURE C.9: **Parametric study of the model: influence of feedback delay.** **A.** Before-effect (red) and after-effect (blue) trajectories. Feedback delay: 0, 0.05, 0.1, 0.15 s; light to dark color. **B.** Velocity profile. **C.** Trajectory angle. **D.** Angle derivative.

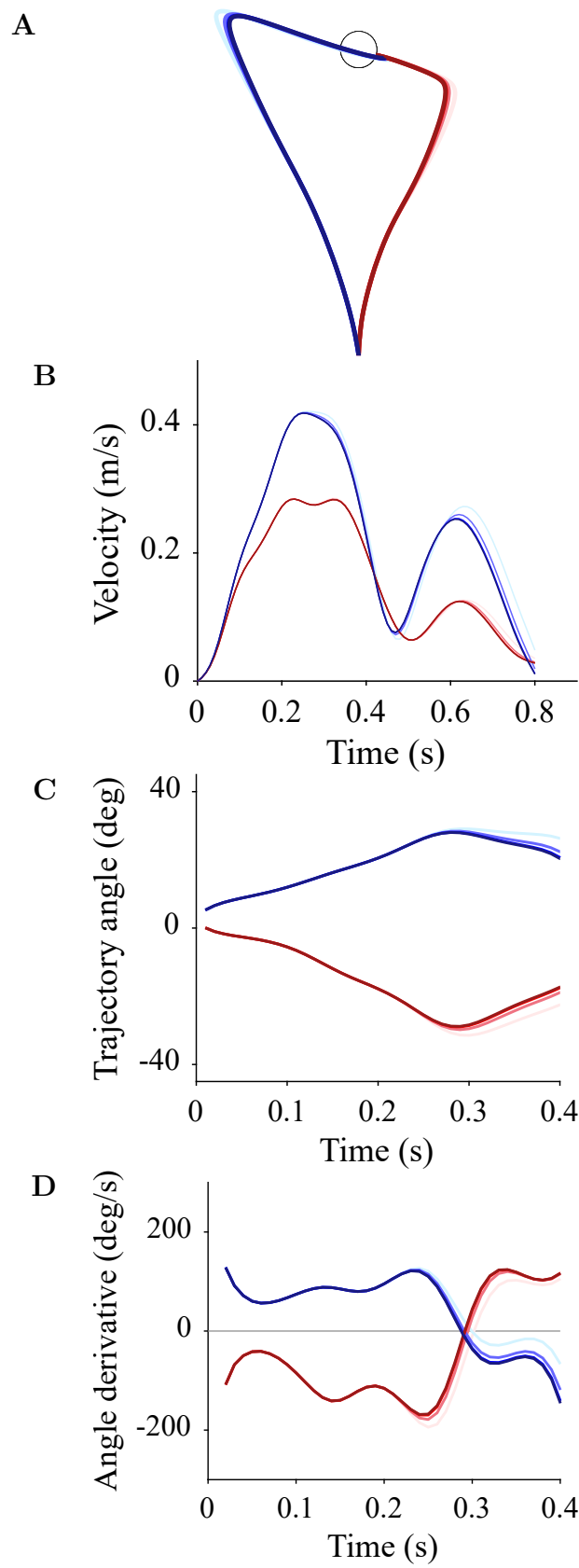


FIGURE C.10: **Parametric study of the model: influence of noise ratio** σ^ξ/σ^ω (motor/sensory), used for estimation. Same format as Fig. C.9. Noise ratio: 0.1, 1, 10, 100; light to dark color.

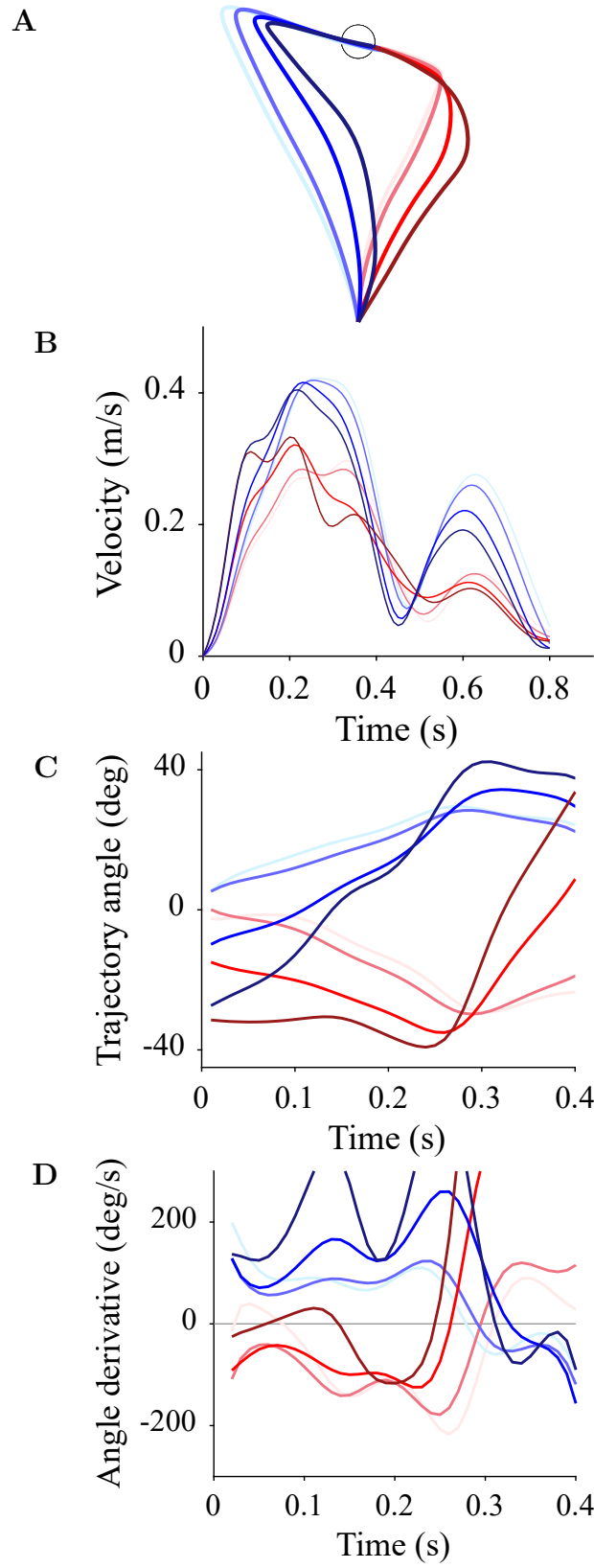


FIGURE C.11: **Parametric study of the model: influence of muscle ratio g_{sho}/g_{el} (shoulder/elbow).** Same format as Fig. C.9. Muscle gain ratio: 1, 2, 5, 10; light to dark color.

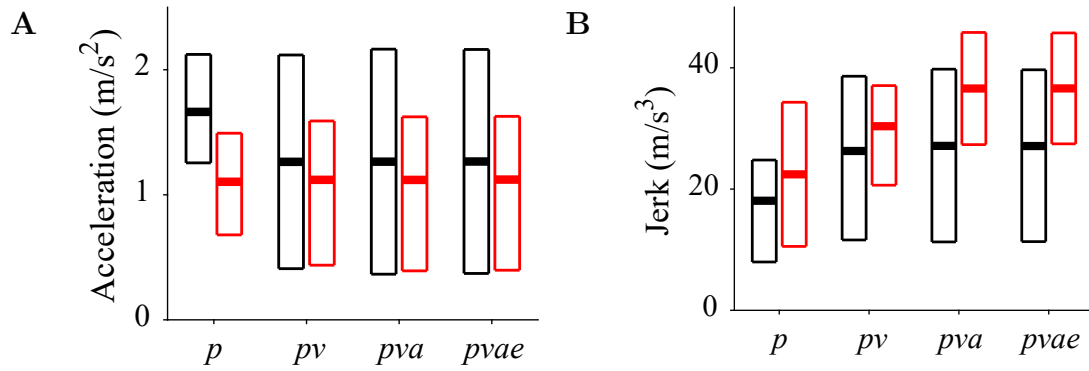


FIGURE C.12: **Parametric study of the model: influence of boundary conditions.** **A.** Mean and 25-75 percentiles of positive acceleration peaks for baseline (black) and before-effect (red) trajectories for different boundary conditions at via-points: *p*: only position; *pv*: position and velocity; *pva*: position, velocity and activation; *pvae*: position, velocity, activation and excitation. **B.** Same as **A** for jerk.

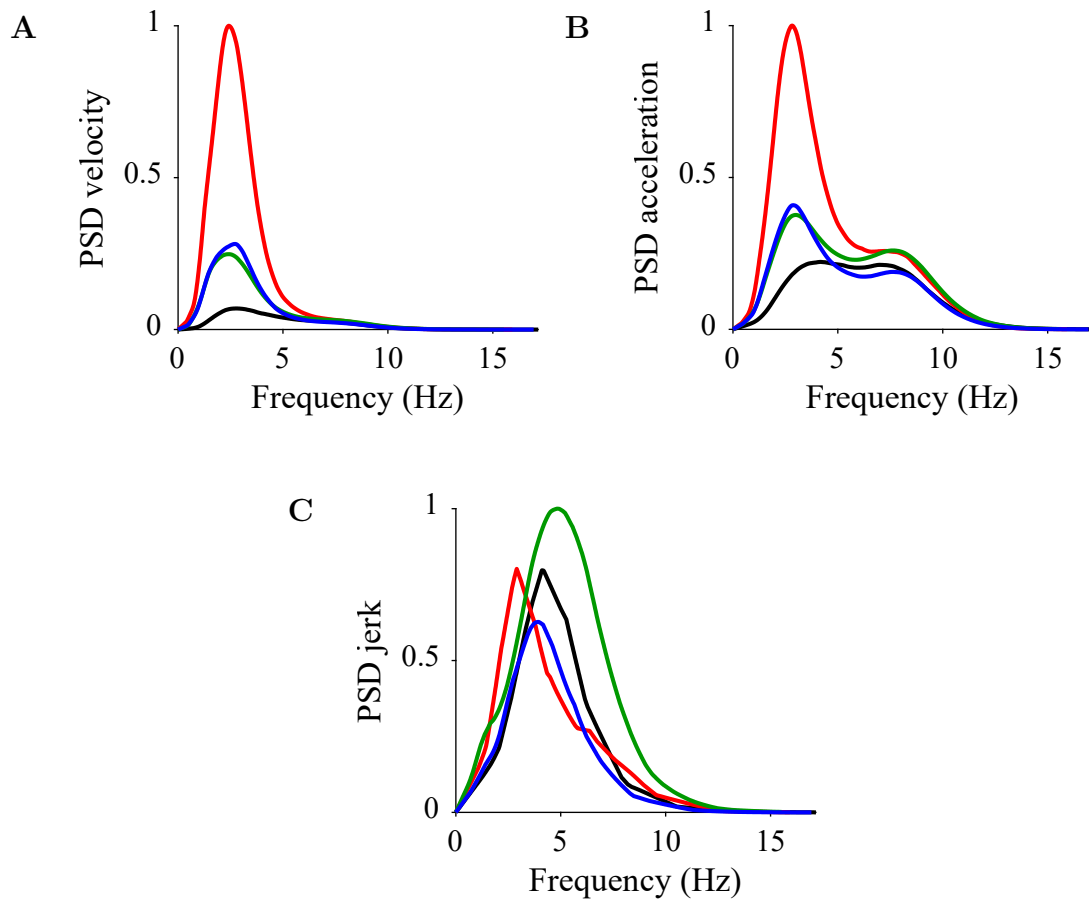


FIGURE C.13: **Power spectrum analysis.** **A.** Power spectrum density (arbitrary unit) of velocity average across trials and participants, for baseline (black), before-effect (red), adapted (green) and after-effect (green) trials. **B.** Same as **A** for acceleration. **C.** Same as **A** for jerk.

Appendix D

Supplementary Materials to Chapter [4](#)

The following figures constitute supplementary materials for Chapter [4](#)

D.1 Influence of network architecture and size

We trained a number of networks with either one or two hidden layers and with layers of various size (256, 512 and 1024 for the network with one hidden layer; 64-128, 64-192, 64-256 for the network with two hidden layers). The training, test and reconstruction error decreased with the number of weights in the network and were lower for two vs one hidden layer (Fig. [D.1](#)).

D.2 More on the performance of the network

We tested the network on 0.12-m, 0.17-m and 0.22-m trajectories in 32 evenly distributed directions (Fig. [D.2](#)). None of these trajectories belongs to the training set. Although the trajectories are not perfect, the network displays correct trends of interpolation and extrapolation.

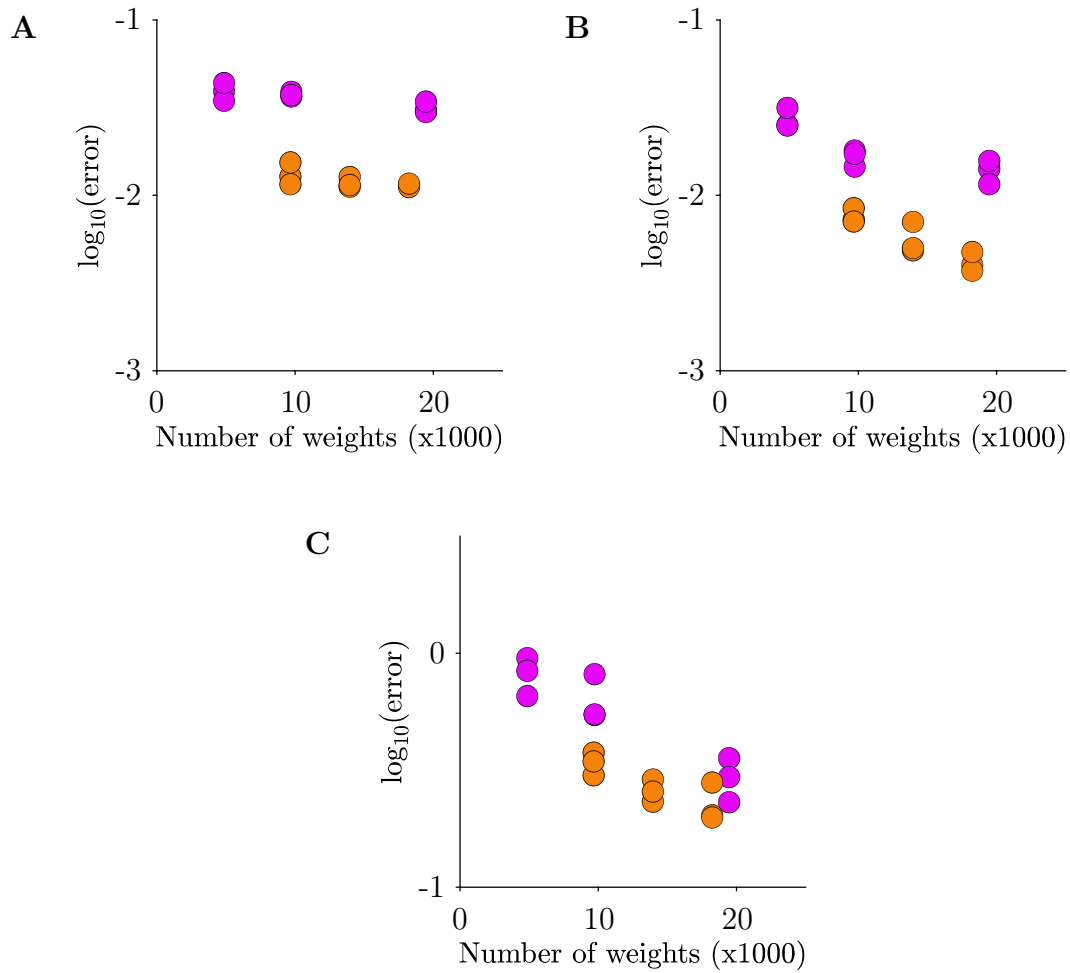


FIGURE D.1: **Influence of network architecture and size.** **A.** Training error as a function of the number of weights in the network. Purple: network with one hidden layer (256, 512, 1024 neurons); orange: networks with two hidden layers (64/128, 64/192, 64/256 neurons). Multiple points with the same number of weights correspond to the same network for different initial postures (shoulder/elbow: 40/100 deg, 40/110 deg, 40/120 deg). **B.** Test error. **C.** Reconstruction error.

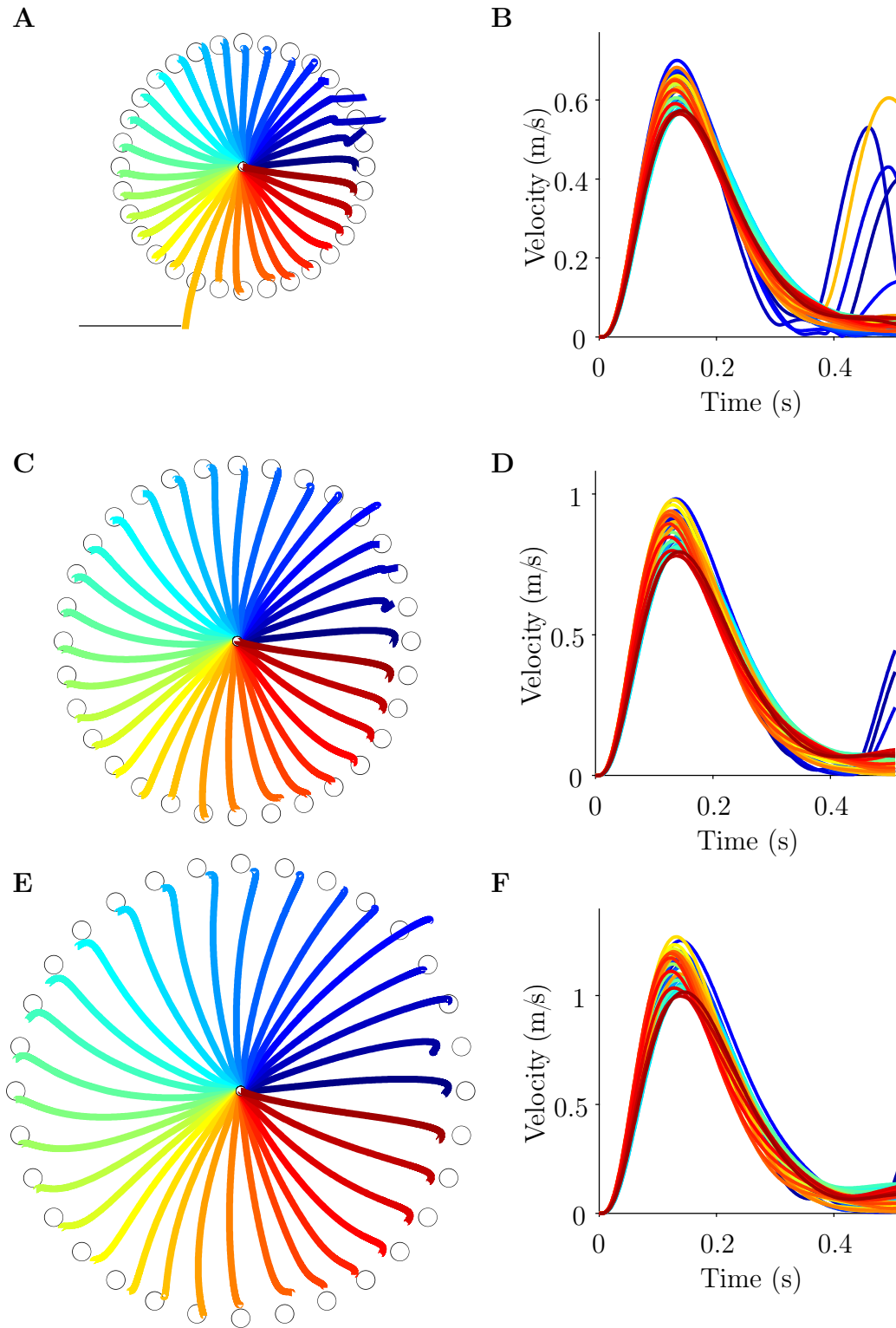


FIGURE D.2: Performances of the main neural network (two hidden layers - 64/268 neurons). **A.** 32 12-cm trajectories. Scale: 0.1 m. **B.** Velocity profiles for A. **C.** 32 17-cm trajectories. **D.** Velocity profiles for C. **E.** 32 22-cm trajectories. **F.** Velocity profiles for E.

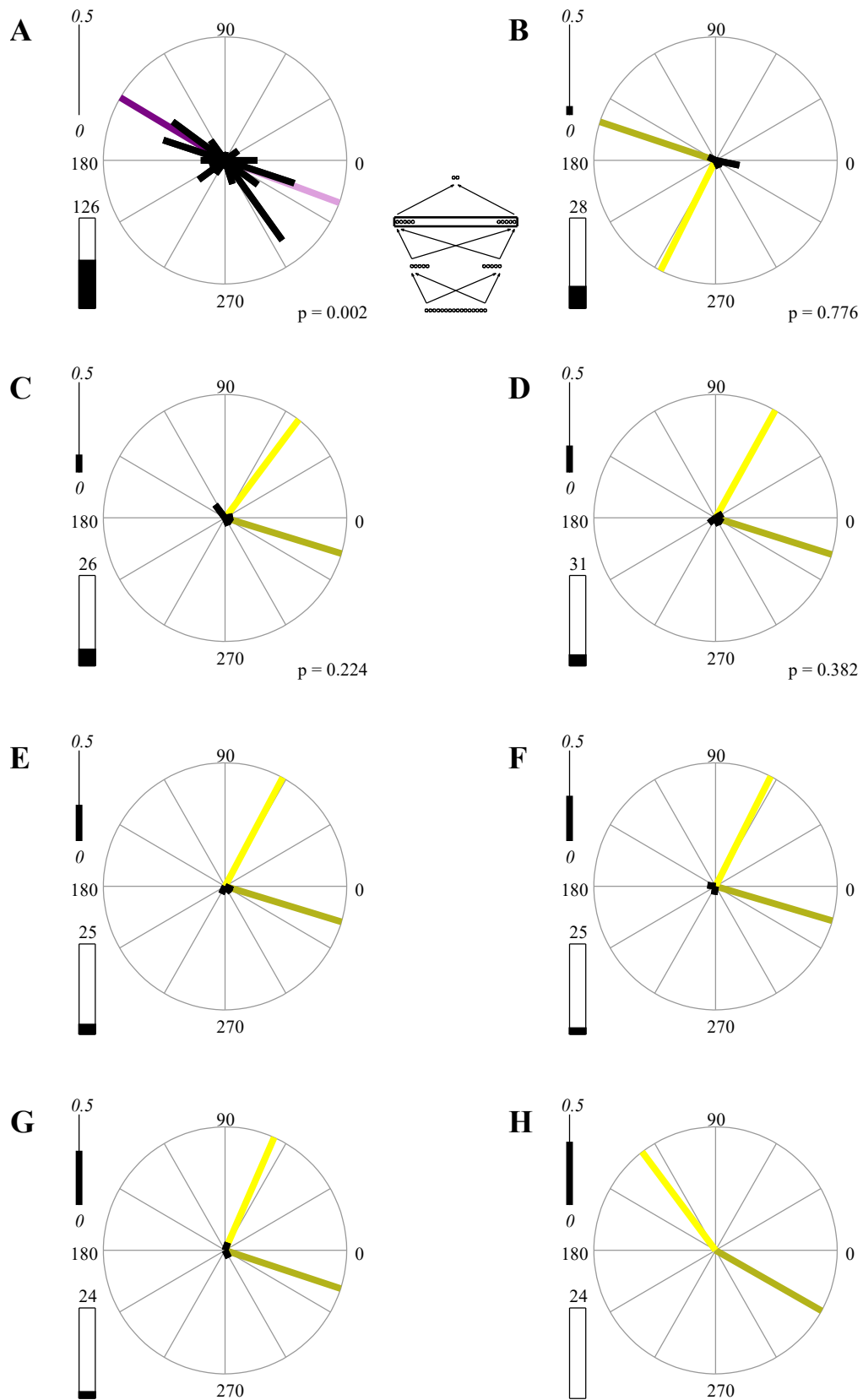


FIGURE D.3: Distribution of neuron's preferred direction (circular histogram, 20 bins, black bars) for the second hidden layer calculated every 0.05 s (A to H) Same format as Fig. 4.7 for the second hidden layer.

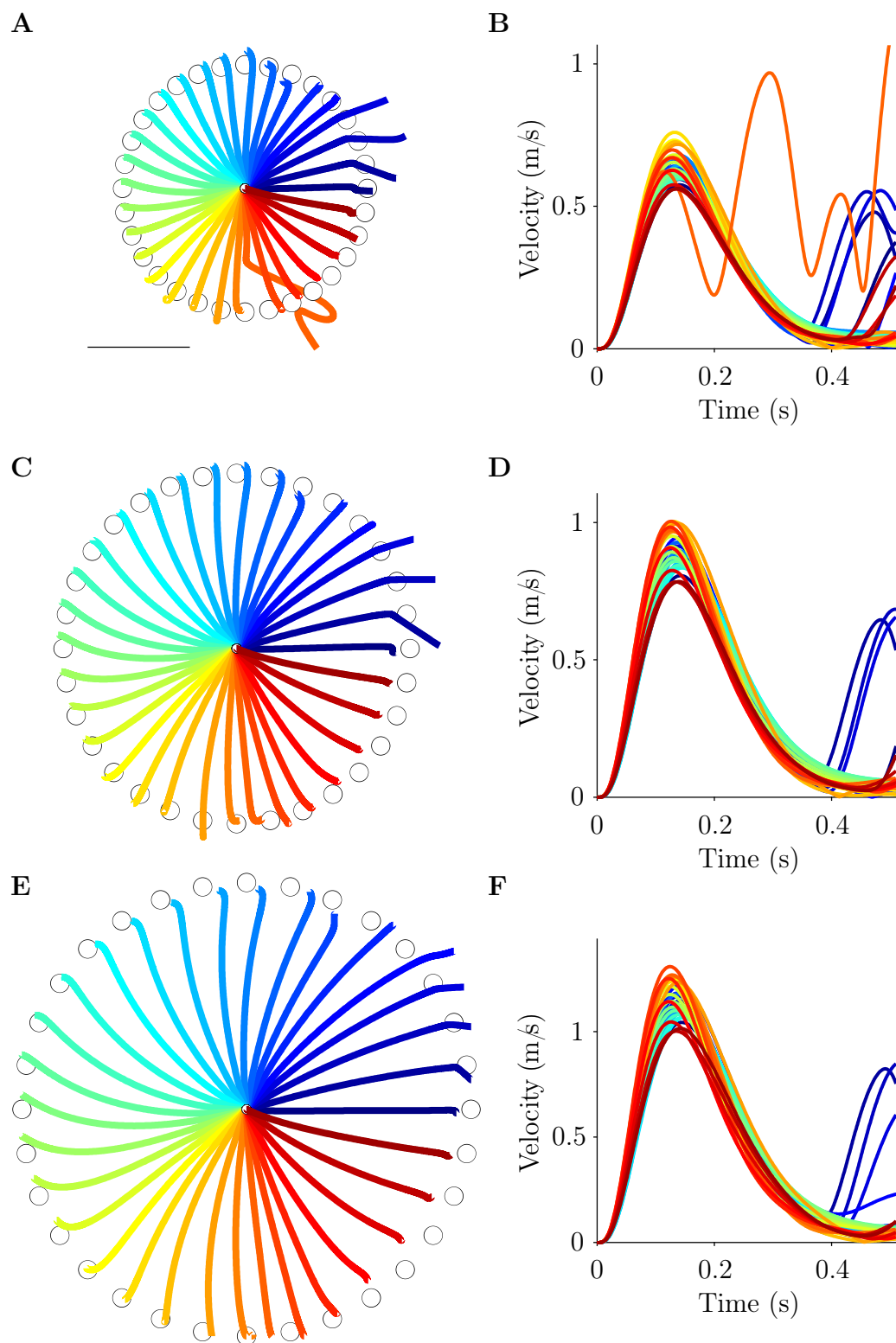


FIGURE D.4: Performances of a smaller two hidden layers neural network (64/128 neurons). Same format as Fig. D.2.

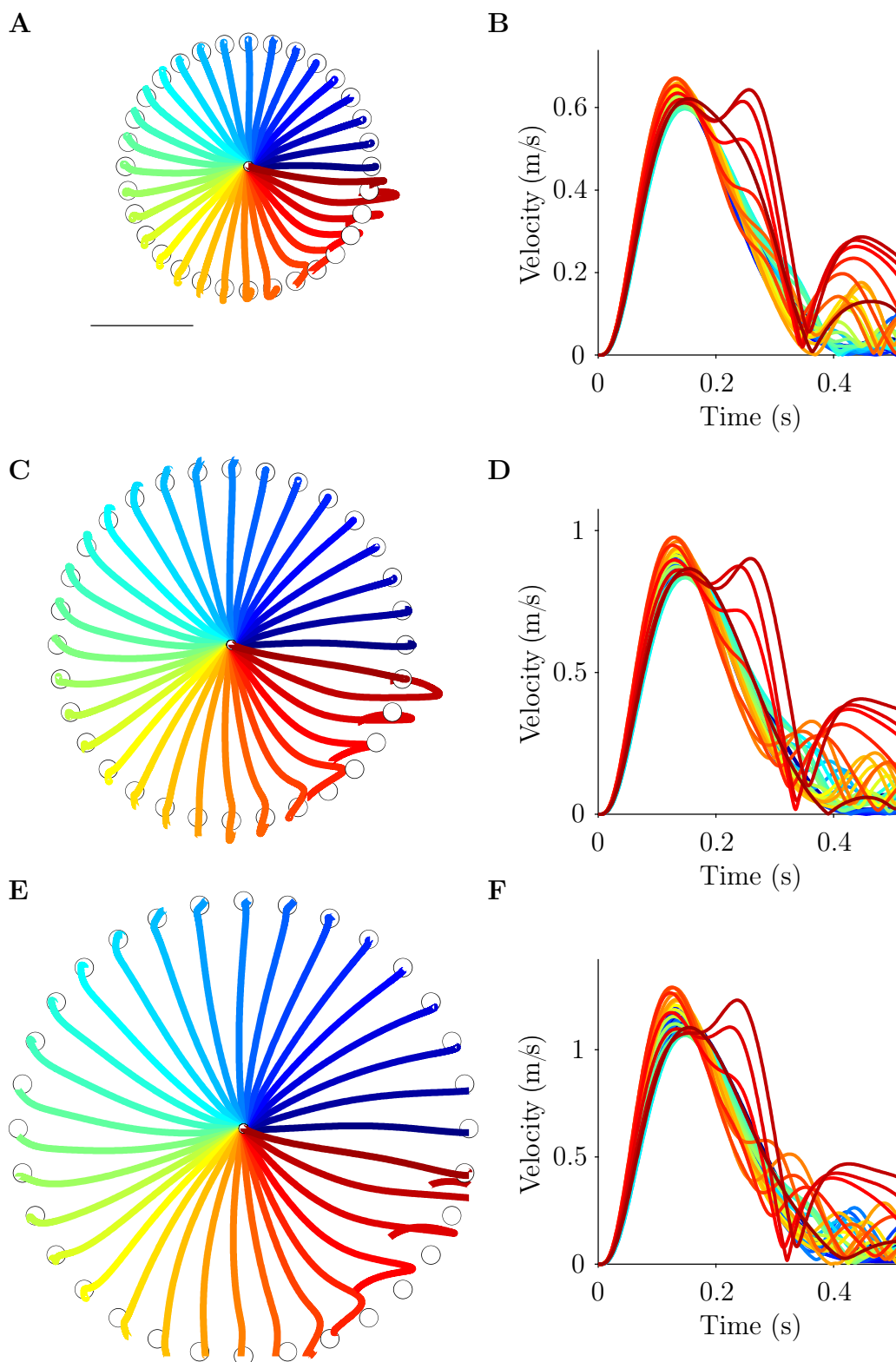


FIGURE D.5: Performances of a single hidden layer neural network (1024 neurons). Same format as Fig. D.2.

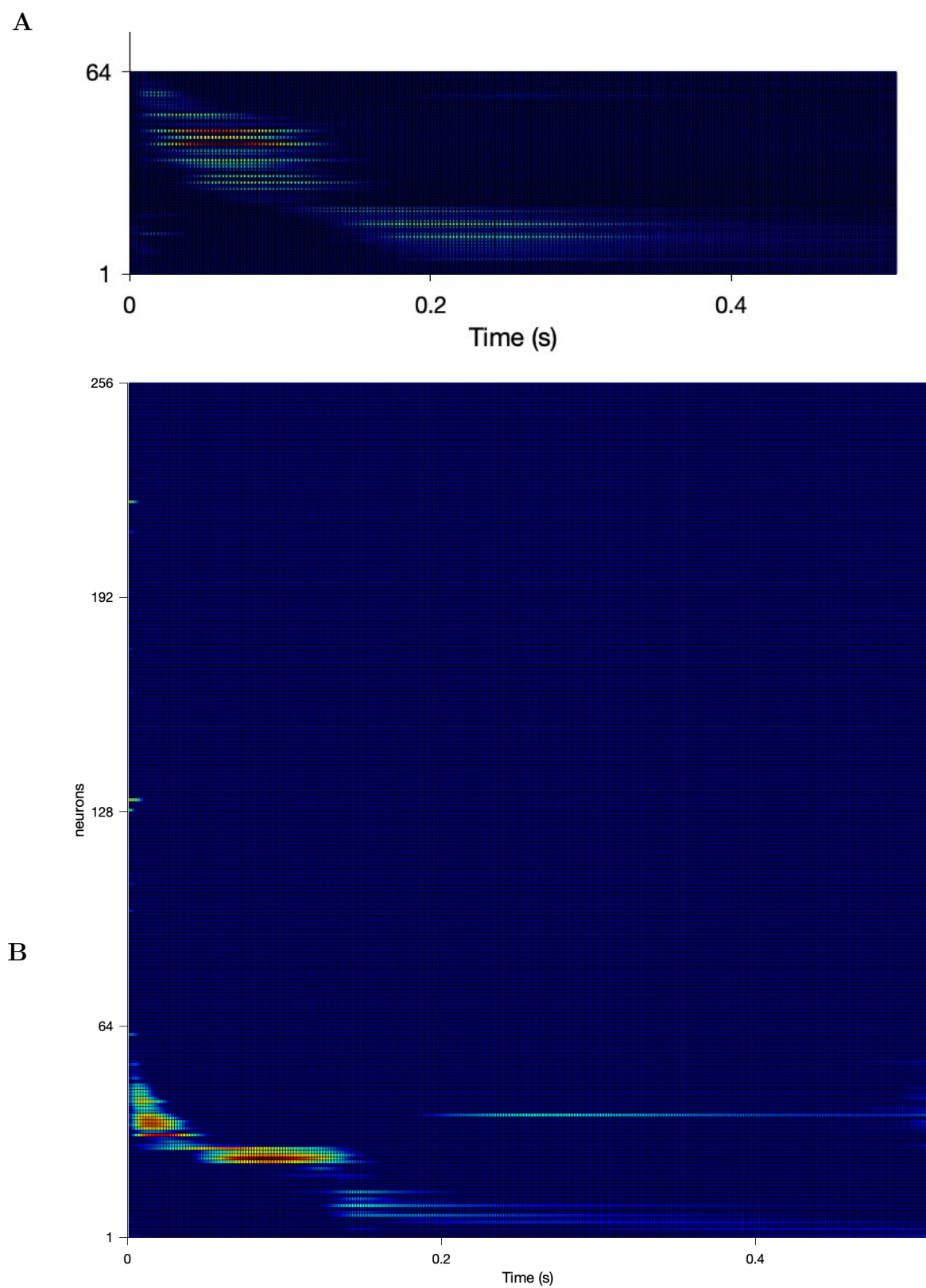


FIGURE D.6: **Temporal activation of the neural network.** **A.** Activity of all neurons in the first hidden layer of single movement (0.15 m , 0°). The neurons are sorted by time-to-peak activity. **B.** Same as **A** for the second hidden layer.

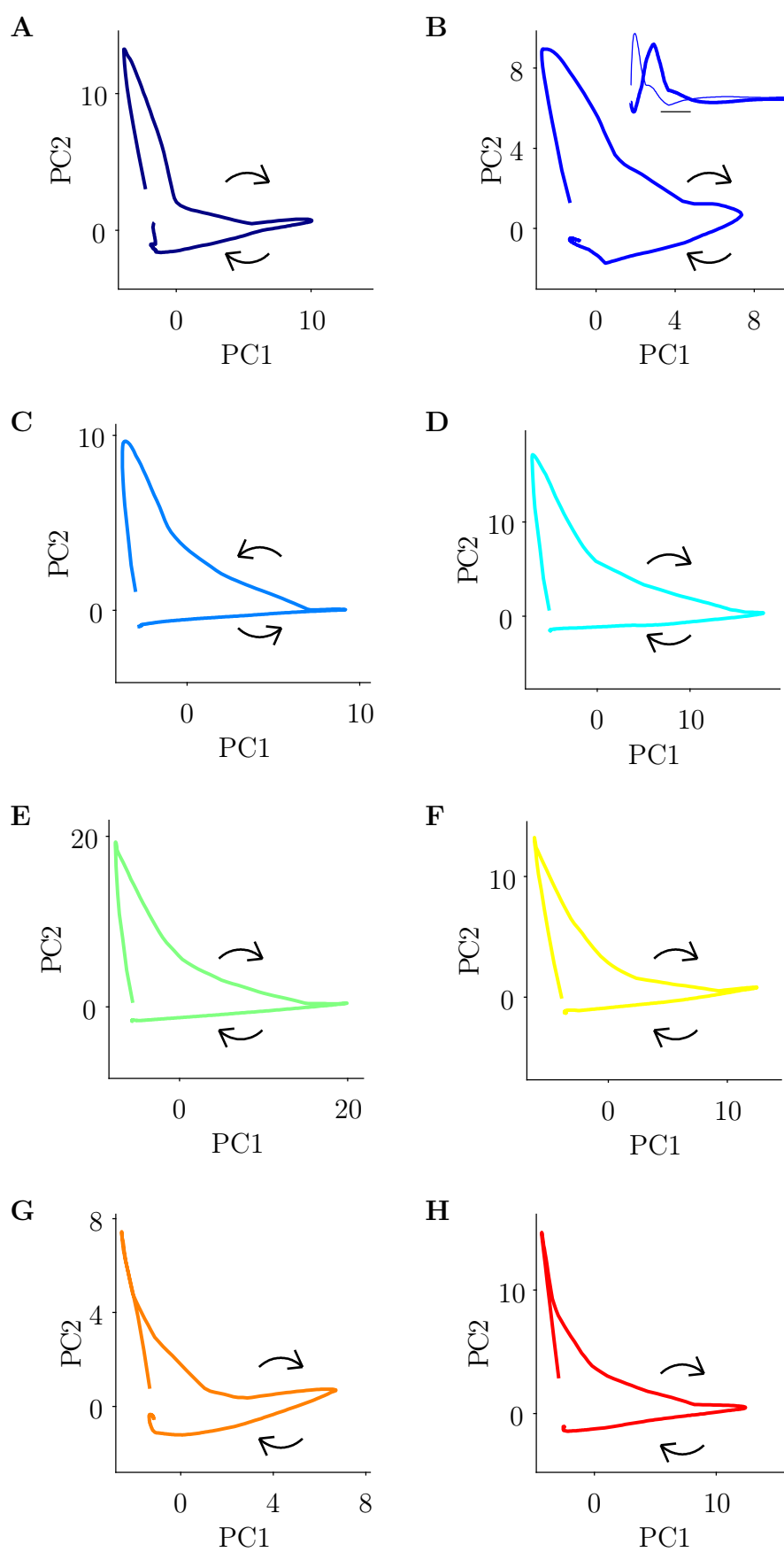


FIGURE D.7: Same as Fig. 4.14 for the second hidden layer.

Appendix E

Position vs. force control

With a receding horizon motor control is defined to be a perpetual movement towards goals. In this framework, posture is merely a movement toward a point - or joint state - already attained. Oscillations about the target are but direct consequences of noises: postural observations, control inputs or goal definition may not be perfect, generating positional errors leading to corrective actions that in turn may induce slight offsets, give rise to ever-going adjustments. It also grants our model the ability to naturally answer to external intermittent perturbations without requiring any adaptation of the control policy or modification of task time. Furthermore, interesting predictions arise when considering one the most simple interaction with our environment: constant forces.

Position control and force control tasks have been widely used to study muscle fatigue [Enoka and Duchateau, 2008]. In several studies, participants were asked to perform two tasks with diverse body parts (elbow [Hunter et al., 2002; Mottram et al., 2005; Rudroff et al., 2007], finger [Maluf et al., 2005], ankle [Magalhaes et al., 2019]): (1) maintain a fixed position while exposed to a constant force F (position control) ; (2) maintain the same force level F (force control). Intuitively, these two tasks look alike, as they require the same force production. Yet, surprisingly, time to failure was shorter for position control than for force control, consistently across all body parts and muscles involved. Explanations regarding the neurophysiological phenomena at play may involve the recruitment of motor units, amplitude of EMG bursts or reflexes through afferent pathways [Enoka and Duchateau, 2008].

Yet, the candidate mechanisms contemplated remain at the peripheral level, and a reason for such a difference may be found in the CNS at the control level. The representation of these tasks in our framework can provide insight in this regard.

E.1 Position control

E.1.1 Issue with receding horizon optimal control

Let us first consider a very simple one-dimensional mass, which state is described by

$$\mathbf{x}(t) = \begin{bmatrix} p(t) \\ v(t) \\ \alpha(t) \\ \varepsilon(t) \end{bmatrix} \quad (\text{E.1})$$

where p is the position of the mass, v its velocity, a the activation - or force - of its actuator and e its excitation, and is driven by a control u through a dynamic:

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} v(t) \\ \alpha(t) \\ \frac{m}{\varepsilon(t) - \alpha(t)} \\ \frac{\nu}{u(t) - \varepsilon(t)} \end{bmatrix} \quad (\text{E.2})$$

where m is the mass and $\nu = 0.05$ s is a time constant. Finally, let us define a

$$\text{simple reaching task starting from state } \mathbf{x}_0 = \begin{bmatrix} p_0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ to } \mathbf{x}_G = \mathbf{x}_f = \begin{bmatrix} p_f \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Fixed time optimal controllers (*e.g.* LQR - section 2.5.2.1 - or SOFC - section 2.5.2.3) producing symmetrical bell-shape velocity profiles. Section 2.5.2.5 showed how such a task is to be implemented in our framework, reproducing similar trajectory velocity profiles [Guigon et al., 2019; Guigon, 2022], with a slight difference in peak velocity timing and amplitude. The reason for this difference is exposed in section 2.5.2.5: the goal of the controller \mathbf{x}_G is to be set equal to \mathbf{x}_f but time to reach it is at all times set to $T_H = 0.28$ s (see section 2.6.1.2). Hence, the control signal decreases as the mass closes in on the target, and so do excitation and activation. This means that the target \mathbf{x}_f is actually never exactly reached but rather asymptotically approached. In reality, the remaining distance becomes negligible after approximately 0.4 s.

Now, let us add a constant opposing force F :

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} v(t) \\ \frac{\alpha(t) - F}{m} \\ \frac{\nu}{u(t) - \varepsilon(t)} \end{bmatrix} \quad (\text{E.3})$$

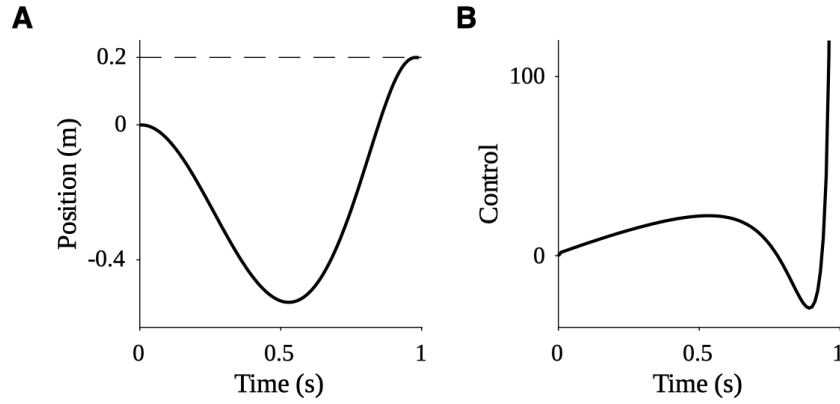


FIGURE E.1: **Simulation of position control against a constant force with a fixed time horizon.** **A.** Time course of position. **B.** Time course of control. Parameters: $m = 1$ kg, $\nu = 0.05$ s.

We performed two simulations for a displacement of 0.2 m in 1 s against a 10 N force (not integrated in the controller), using a fixed task time optimal controller and a receding horizon controller. The fixed time controller succeeded in reaching the target (Fig. E.1) while the receding horizon controller plateaued to a position undershooting the target (Fig E.2B,C). Let us give an explanation for these different behaviors.

With a fixed time controller, the time remaining to reach the current goal decreases as the movement unfolds (see 2.6.1.2). Consequently, the controls computed in the presence of an opposing force will be automatically (potentially exponentially) increased during the latest moments of the task and compensate for the perturbation (see Fig. E.1,B), and the target is successfully reached (see Fig. E.3A).

For optimal control with receding horizon, there is no predefined task end time (see 2.6.1.2). Rather, task may be considered to be accomplished when the state objective is attained (within observation margin error) or, if impossible, when equilibrium is reached. Fig. E.2A illustrates how, in the receding horizon case, an equilibrium is reached.

State equilibrium is defined by $\dot{\mathbf{x}}_{eq} = \mathbf{0}$ which yields

$$\mathbf{x}_{eq} = \begin{bmatrix} p_{eq} \\ 0 \\ F \\ F \end{bmatrix} \quad (\text{E.4})$$

where p_{eq} is the position reached at equilibrium, illustrated by the plateau reached in Fig. E.2. Its exact value is easier to find through simulations than calculations, but a interesting reasoning can be made on state equilibrium, that will prove to be useful to circumvent this undershooting issue.

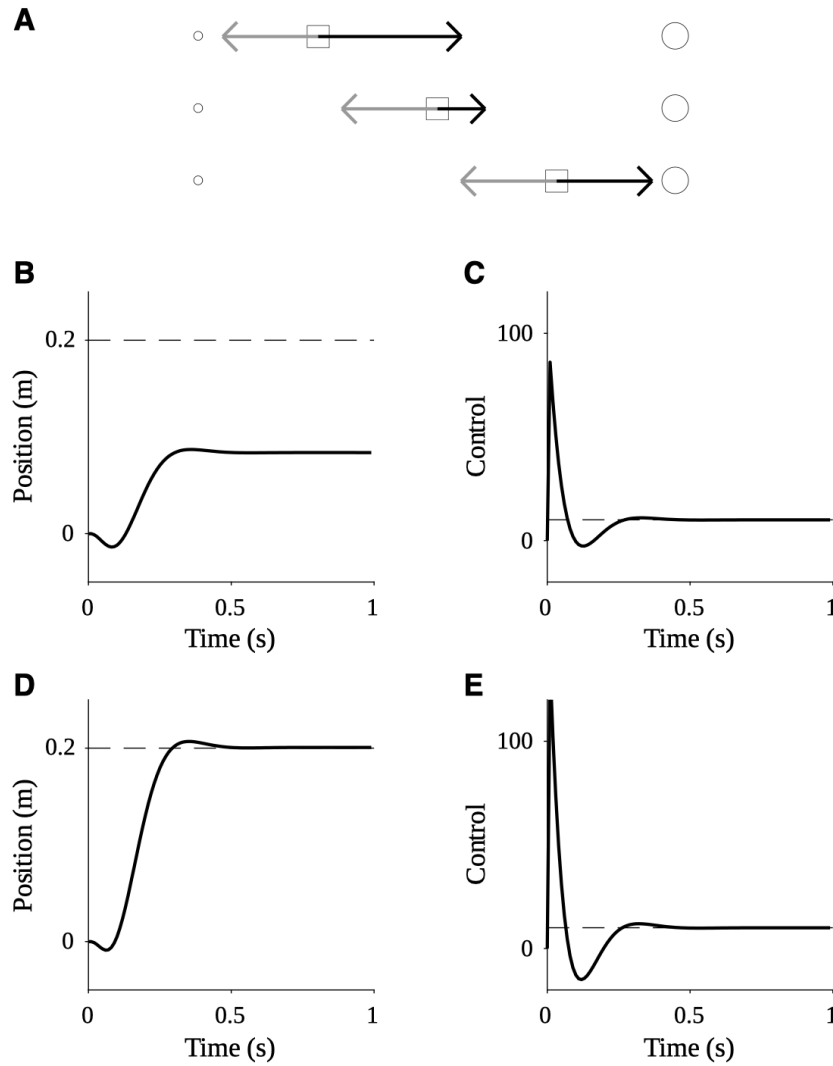


FIGURE E.2: **Simulation of position control against a constant force with a receding time horizon.** **A.** Schema to explain the failure of optimal control with a receding horizon in a position control task. The start position is the small circle on the left. The target position is the large circle on the right. The mass is depicted by a square and moves from the left to the right due to two forces: a constant force (gray arrow) and a control force (black arrow). Three successive position of the mass are shown (from top to bottom). Early in the movement, the controlled force is large (top). Later, the controlled force is small (middle). Then an equilibrium between the forces is reached, the mass stops moving and undershoots the target position (bottom). **B.** Time course of position corresponding to the scenario in **A**. The goal state is $[0.2 \ 0 \ 0 \ 0]$. **C.** Time course of control for **B**. **D.** Time course of position corresponding to a scenario in which the goal state is $[0.2 \ 0 \ 22.3 \ 22.3]$. **E.** Time course of control for **D**.

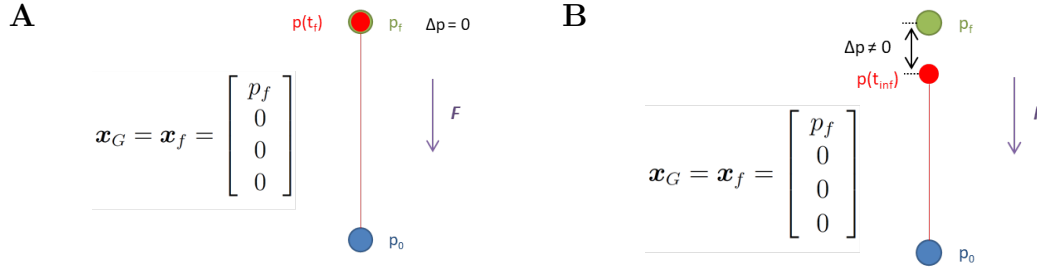


FIGURE E.3: **Optimal control facing constant force with the goal of the controller set to the actual goal.** **A.** Optimal control with fixed end time: the target is reached. **B.** Optimal control with receding horizon: target is undershot.

It is straightforward to see from Eq. E.3 that control at equilibrium is $u_{eq} = F$. Additionally, we showed in Chapter 2 (see Eq. 2.55) that the control in a given configuration is a function of the current state and the current goal to reach:

$$u(t) = \pi(\mathbf{x}(t), \mathbf{x}_G(t)) \quad (\text{E.5})$$

The system will thus be at rest at \mathbf{x}_{eq} such that

$$\pi(\mathbf{x}_{eq}, \mathbf{x}_f) = F \quad (\text{E.6})$$

A method to find the equilibrium position \mathbf{x}_{eq} may thus be the following:

1. compute control $\pi(\mathbf{x}_{search,i}, \mathbf{x}_f)$ with a series of state $\mathbf{x}_{search,i}$

$$\mathbf{x}_{search,i} = \begin{bmatrix} p_{search,i} \\ 0 \\ F \\ F \end{bmatrix} \quad (\text{E.7})$$

where $p_{search,i} < p_f$.

2. draw the curve of the control as a function of p_{search} and interpolate to find p_{eq} , the value satisfying Eq. E.4.

Fig. E.4 was built following this method and shows that, consistently with Fig. E.2, the system previously defined reach a point of equilibrium $p_{eq} \sim 8.2$ cm.

E.1.2 Adapted task representation for receding horizon optimal control

Keeping in mind that the dynamic embed in the controller may not be changed, for our system to reach the target position p_f despite the opposing force F , only

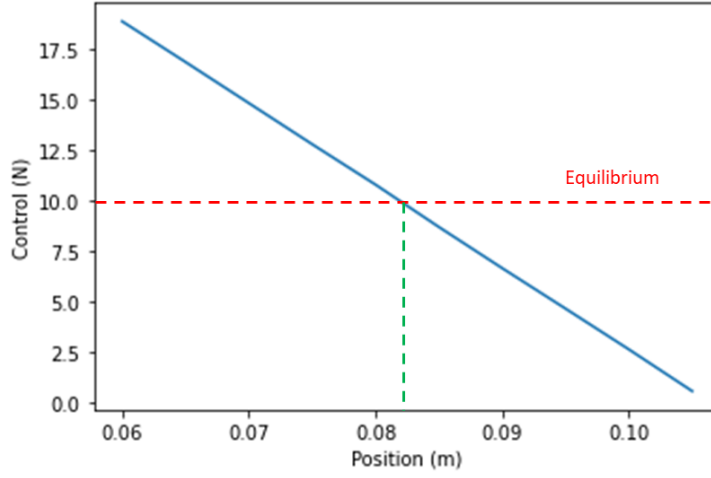


FIGURE E.4: **Control as a function of position.** Control as a function of position for a reaching task toward a target placed at 0.2 m with an opposing force of 10 N (plain blue line). Equilibrium is reached when control is equal to 10 N (dashed red line), corresponding to a position $p_{eq} \sim 11.8$ cm

the goal fed to the controller \mathbf{x}_G may be adjusted. The problem to solve is thus to find a controller goal \mathbf{x}_G that satisfies:

$$\pi(\mathbf{x}_{eq}^f, \mathbf{x}_G) = F \quad (\text{E.8})$$

with

$$\mathbf{x}_{eq}^f = \begin{bmatrix} p_f \\ 0 \\ F \\ F \end{bmatrix} \quad (\text{E.9})$$

Such a task representation \mathbf{x}_G can be constructed in two ways, presented hereafter.

E.1.2.1 Goal position adjustment

A first option is to displace the objective fed to the controller further away so that the equilibrium point matches the position actually intended (see Fig. E.5A) :

$$\mathbf{x}_G = \begin{bmatrix} p_f + \Delta p \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (\text{E.10})$$

As the chosen dynamics is linear, the undershoot as observed in Fig. E.4 is not dependent of the objective position p_f . This means that by setting the controller's goal as noted above to $p_f + \Delta p$ with the "overshoot" $\Delta p = p_f - p_{eq}$ allows to compensate for a constant force. The application to the previous simulation thus yields with an of $\mathbf{x}_G = [0.318, 0, 0, 0]$ to compensate for the force of 10 N.

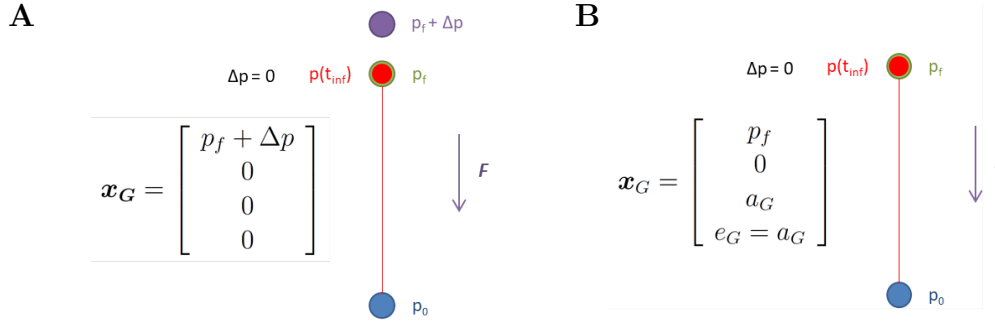


FIGURE E.5: **Adapted task representation for receding horizon optimal control.** **A.** Goal position adjustment. The goal of the controller is different from the actual goal: objective position shifted onward by Δp . **B.** Goal activation/excitation adjustment. The goal of the controller is different from the actual goal: activation and excitation are set to a non-zero, adapted value.

It should be noted, though, that this task representation would not be adapted to redundant and/or non-linear systems, for a different end effector position implies a different joint configuration, and different effect of applied forces. The task representation presented hereafter is not plagued with this drawback.

E.1.2.2 Goal activation/excitation adjustment

A second possibility is to set up the controller state goal to the actual objective position p_f , but with non-null activation and excitation (see Fig. E.5B):

$$\mathbf{x}_G = \begin{bmatrix} p_f \\ 0 \\ \alpha_G \\ \varepsilon_G = \alpha_G \end{bmatrix} \quad (\text{E.11})$$

Similarly to the method presented in section E.1.1, we computed multiple controls $\pi(\mathbf{x}_f, \mathbf{x}_{G,i})$ with

$$\mathbf{x}_{G,i} = \begin{bmatrix} p_f \\ 0 \\ \alpha_{G,i} \\ \varepsilon_{G,i} = \alpha_{G,i} \end{bmatrix} \quad (\text{E.12})$$

where $\alpha_{G,i}$ were a range of activations. Fig. E.6 shows the relationship between the control u and the activation and excitation goals ($\alpha_{G,i}$ and $\varepsilon_{G,i}$) with an opposing force of 10 N. For the same position $p_f = 0.2$ m to be the equilibrium point, control is to be equal to 10 N, which requires $\alpha_G = \varepsilon_G = 22.3$ N.

This task representation is more versatile than the previous: in this example dynamics goals are set as activation which is equivalent to a force, but in the general case they would be expressed as forces generated at the end-effector, which translates additively as joint torques (see Eq. 2.8).

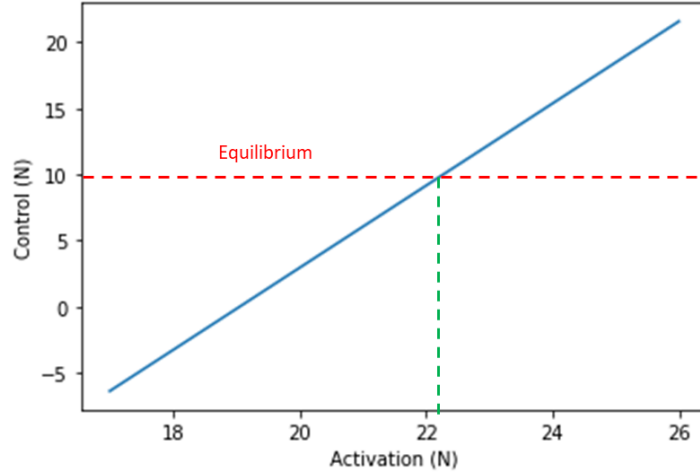


FIGURE E.6: **Position control with goal activation/excitation adjustment.** Control as a function of the goal activation with a force of 10 N (plain blue line). Equilibrium is reached when control is equal to 10 N (dashed red line), corresponding to an activation goal of ~ 22.3 N. Slope is 3.10.

E.2 Force control

In our framework, a task of pure force control can be modelled for our system (1D mass of 1 kg) as:

$$\mathbf{x}_f = \begin{bmatrix} * \\ * \\ \alpha_f \\ e_f = \alpha_f \end{bmatrix} \quad (\text{E.13})$$

where $*$ indicates that the coordinate is unconstrained. Similarly to the "positional" reaching task defined in E.1, we simulated an "activation" reaching task, starting from zero activation and excitation up to 10 N for a fixed time controller and a receding horizon controller. Unsurprisingly, the fixed time controller achieved the task (Fig. E.7A) while the receding horizon controller plateaued to an undershoot and ask representation need to be adapted for optimal control with receding horizon (Fig. E.7A).

The problem to solve is here to find a controller goal \mathbf{x}_G that satisfies:

$$\pi(\mathbf{x}_f, \mathbf{x}_G) = F \quad (\text{E.14})$$

with

$$\mathbf{x}_f = \begin{bmatrix} * \\ * \\ F \\ F \end{bmatrix} \quad (\text{E.15})$$

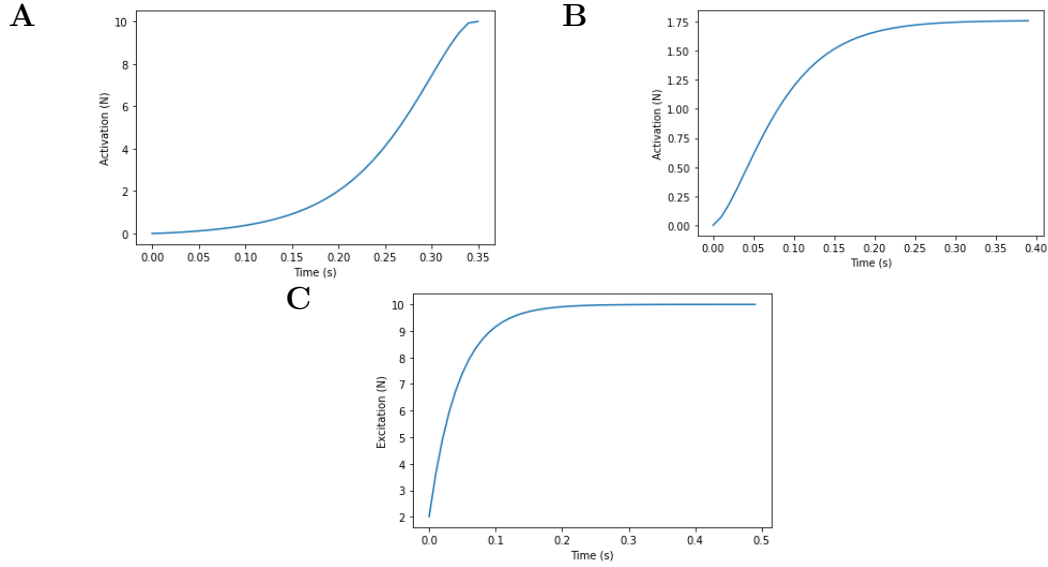


FIGURE E.7: **Optimal control producing force.** **A.** Optimal control with fixed end time, setting the goal of the controller to the actual goal: the target force of 10 N is reached. **B.** Optimal control with receding horizon, setting the goal of the controller to the actual goal: the target force of 10 N is highly undershot. **C.** Optimal control with receding horizon, setting the goal of the controller to $\mathbf{x}_G = [*, *, 56.8, 56.8]$: the target force of 10 N is reached.

The same method as in E.1.2.2 was adopted to find the activation α_G defining

$$\mathbf{x}_G = \begin{bmatrix} * \\ * \\ \alpha_G \\ \varepsilon_G = \alpha_G \end{bmatrix} \quad (\text{E.16})$$

satisfying Eq. E.14. Fig. E.8 shows the relationship between the control u between the control u and the activation and excitation goals to produce a force of 10 N, evaluating $\alpha_G = \varepsilon_G = 56.8$ N. Note that higher values of final activation and excitation need to be set as a goal than for position control.

E.3 Modelling predictions: comparison

To make predictions on the difference between these two tasks, noise must be taken into account. Out of the noises listed in 2.3.2, we decided to analyse the impact of signal dependent motor noise ϵ , updating the dynamic of Eq. E.3 to:

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} v(t) \\ \frac{\alpha(t) - F}{m} \\ \frac{\varepsilon(t) - \alpha(t)}{\nu} \\ \frac{u(t)(1 + \epsilon) - \varepsilon(t)}{\nu} \end{bmatrix} \quad (\text{E.17})$$

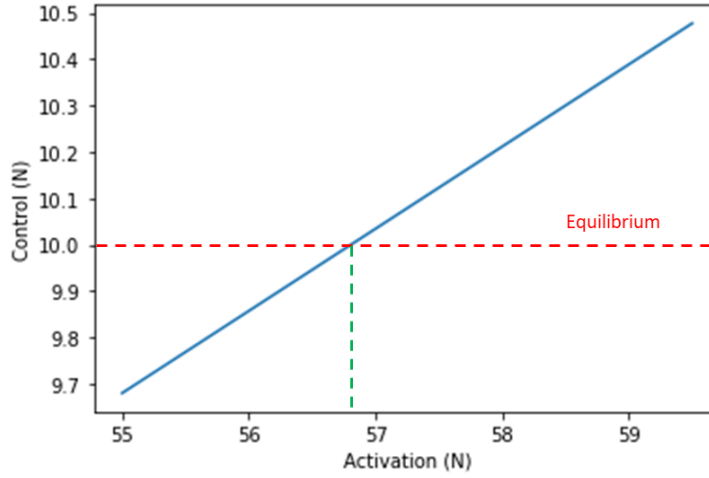


FIGURE E.8: **Force control with goal activation/excitation adjustment.** Control as a function of the goal activation with a force of 10 N (plain blue line). Equilibrium is reached when control is equal to 10 N (dashed red line), corresponding to an activation goal of ~ 56.8 N. Slope is 0.18.

We simulated a position control task around a position at 0.2 m, facing a force of 10 N and with signal dependent noise of mean zero and standard deviation $\sigma_\epsilon = 0.4$. Similarly, we simulated a force control task to produce a force of 10 N, with signal dependent noise of mean zero and standard deviation $\sigma_\epsilon = 0.4$. Both tasks lasted 4 s. Note that no estimation was taken into account in these simulations.

Time series of each coordinate of the state involved are showed in Fig. E.9 for both tasks. Remarkably, while subjected to the same noise intensity, control fluctuations were noticeably stronger in the position control task than in force control task (Fig. E.9E,H). This discrepancy, even if reduced, logically impacts excitation (Fig. E.9D,G) and activation (Fig. E.9) profiles: even if they are of comparable amplitude, they don't have the same frequency content. Interestingly, for position control, velocity (Fig. E.9B) and position (Fig. E.9A) fluctuations were limited. Even if associated with fluctuating controls, position control appears to be more stable than force control.

An explanation could be found in the significantly different slopes of the curves of Figs E.6 and E.8: with a smaller slope, a given noise occurrence in the control implies a lower change in the aimed activation.

One of the observed differences between position control and force control tasks lied in the content of EMGs. Here, simply out of the two task representations in our framework emerged discrepancies in the frequency content of excitation, which is commonly associated with EMG signals. Even if qualitative and very partial, these results suggest that the observed difference between those tasks could find an explanation in the task representation cast in our optimal control with receding horizon framework.

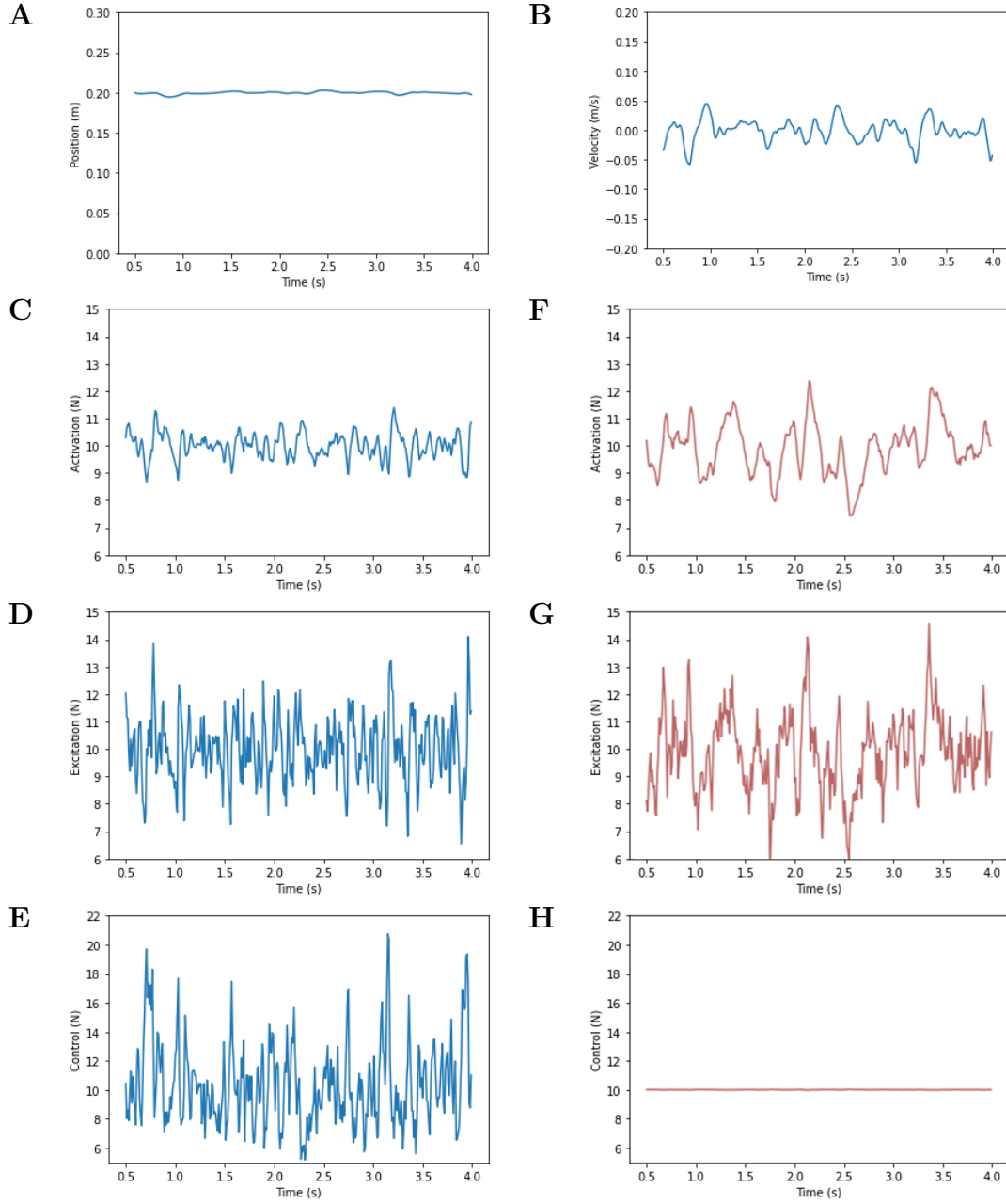


FIGURE E.9: **Position control vs force control** **A.** Position time series for the position control task. **B.** Velocity fluctuations for the position control task. **C.** Activation fluctuations for the position control task. **D.** Excitation fluctuations for the position control task. **E.** Control fluctuations for the position control task. **F.** Activation fluctuations for the force control task. **G.** Excitation fluctuations for the force control task. **H.** Control fluctuations for the force control task. Time series were cut before 0.5 s to put aside transient phases.

Appendix F

Cocontraction

Cocontraction - or coactivation - is defined by the simultaneous activation of antagonist muscles acting on a given joint. Its primary shown effect was that it increases the *stiffness* of the joint in question around a given posture: any perturbation leading to a deviation from the desired position will induce a immediate passive correcting force by stretching one of the tendons involved [Hogan, 1984a]. It also appears to improve movement accuracy [Gribble et al., 2003]. Moreover, recent studies showed that cocontraction also plays a role in the short term reflex loops, as well as in the perception of perturbation and induce active responses [Dideriksen et al., 2015; Saliba et al., 2020; Koelewijn and van den Bogert, 2022]. Cocontraction is often measure with tonic EMG [Gribble et al., 2003], but is also associated with joint stiffness [Gomi and Kawato, 1996], although the complex dynamics linking muscle activity and velocity to generated forces and joint stiffness calls for cautiousness in drawing conclusions. Nevertheless, cocontraction, and more generally muscle-tendon impedance thus appears to be a key factor in motor control and should in future developments be incorporated in the proposed model. Although its integration in simulations was out of the scope of this work, it is of interest to glimpse at the way it can be cast in our framework.

For the sake of simplicity, let us consider a one-dimensional mass m actuated by a set of agonist/antagonist muscle-tendon units (Fig. F.1), described by the state



FIGURE F.1: Linear mass actuated by two muscles (rectangles) and their tendons (springs).

$$\mathbf{x}(t) = \begin{bmatrix} p(t) \\ v(t) \\ \alpha^+(t) \\ \varepsilon^+(t) \\ \alpha^-(t) \\ \varepsilon^-(t) \end{bmatrix} \quad (\text{F.1})$$

where p is the position of the mass, v its velocity and $\alpha^\#$ and $\varepsilon^\#$ are the activation and excitation of the muscle $\# \in \{+, -\}$ (+ is the agonist and $-$ the antagonist), and is driven by a control $\mathbf{u} = [u^+, u^-]$ through a dynamic:

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} v(t) \\ \frac{f_m^+(t) - f_m^-(t)}{m} \\ \frac{\varepsilon^+(t) - \alpha^+(t)}{\nu} \\ \frac{u^+(t) - \varepsilon^+(t)}{\nu} \\ \frac{\varepsilon^-(t) - \alpha^-(t)}{\nu} \\ \frac{u^-(t) - \varepsilon^-(t)}{\nu} \end{bmatrix} \quad (\text{F.2})$$

where m is the mass and $\nu = 0.05$ s is a time constant, and $f_m^\#$ is the muscle force, considered to be in equilibrium with tendon force $f_t^\#$

$$f_m^\# = f_t^\# \quad (\text{F.3})$$

where

$$\begin{cases} f_m^\# &= a_m^\# k_m^\# (L_m^\# - L_m^0) & \text{if } L_m^\# - L_m^0 > 0 \\ f_m^\# &= 0 & \text{if } L_m^\# - L_m^0 < 0 \end{cases} \quad (\text{F.4})$$

$$\begin{cases} f_t^\# &= a_t^\# k_t^\# (L_t^\# - L_t^0) & \text{if } L_t^\# - L_t^0 > 0 \\ f_t^\# &= 0 & \text{if } L_t^\# - L_t^0 < 0 \end{cases} \quad (\text{F.5})$$

where $k_m^\#$ represents muscle efficiency, $k_t^\#$ represent the stretching gain of the tendons and L_m^0 and L_t^0 are muscles' and tendons' slack lengths.

F.1 Cocontraction during postural task

In appendix E, section E.1.2.2 presented how a linear mass can be controlled to oppose a constant force. A postural task around a position p_f with a cocontraction level defined as $F_{co} = \min(f_m^+, f_m^-)$ can be implemented in a similar way in setting a goal with non-zero activations and excitations:

$$\mathbf{x}_{G,co} = \begin{bmatrix} p_f \\ 0 \\ F_{co} \\ F_{co} \\ F_{co} \\ F_{co} \end{bmatrix} \quad (\text{F.6})$$

A control \mathbf{u}_{co} can then be found to reach equilibrium between the two muscles, depending on the desired position and on parameters of each muscle and tendon. Remember, though, that due to the receding horizon, the goal's activation and excitation values cannot be directly derived for the cocontraction level to obtain, but need to be regressed following the same process as described in appendix E, section E.1.2.2 :

$$\mathbf{x}_{G,co} = \begin{bmatrix} p_G \\ 0 \\ \alpha_{G,co}^+ \\ \varepsilon_{G,co}^+ = \alpha_{G,co}^+ \\ \alpha_{G,co}^- \\ \varepsilon_{G,co}^- = \alpha_{G,co}^- \end{bmatrix} \quad (\text{F.7})$$

Then, following the common mechanism of impedance, a perturbation provoking a deviation from the maintained posture will immediately induce a change in muscles and tendons lengths, and consequently in tendons strains and forces, generating a corrective action to reduce deviation. Note that, after diverse delays, reflexes and descending commands will come into play.

F.2 Cocontraction during movement

Cocontraction is often cast in the equilibrium point theory (section 2.5.1): a reference trajectory is defined in a preparation phase and then executed by the computation of the necessary forces or torques to follow it. One can thus simply create cocontraction during movement by increasing the current forces by the desired level of cocontraction (potentially taking into account the moment arms in non-linear cases) [Latash and Gottlieb, 1991; Gribble et al., 1998]. Our framework, though, does not allow continuous addition of commands, for task representation is restricted to the definition of sequential goals (section 2.6.1.3).

Section 2.6.2.1 described a reaching task as a two phase process: (1) maintain the controller's goal \mathbf{x}_G^i for long enough to reach equilibrium corresponding to an initial state \mathbf{x}_i ; (2) at a time t_0 switch and maintain the controller's goal to a value \mathbf{x}_G^f corresponding (after convergence) to the final state \mathbf{x}_f .

We showed in section F.1 that setting a level of cocontraction around a given position is done by defining adapted activation and excitation in both muscles.

We may thus embed initial cocontraction in \mathbf{x}_G^i and terminal cocontraction in a goal \mathbf{x}_G^f .

But the evolution of cocontraction during movement (between the t_0 and equilibrium) itself cannot be controlled, and is a direct product of the optimal control policy and dynamics of the system. We ran two fixed horizon simulations of 0.28 s for a reaching task from position $p_i = 0.1$ cm to a position $p_f = 0.15$ cm: one without any cocontraction instruction and the other with an initial cocontraction level $F_{co,i} = 1$ N and a final cocontraction level $F_{co,f} = 2$ N.

In both cases, target was reached (Fig. F.2A,D) and velocity profile was slightly different from usual, with a delayed velocity peak (Fig. F.2B,F). Without cocontraction instructions, cocontraction still appeared at peak velocity (Fig. F.2E), coherently with the deactivation of the agonist muscle and activation of the antagonist muscle (Fig. F.2C) required to induce deceleration. Note that Without cocontraction instructions, there is a slight delay between the cocontraction peak (~ 0.18 s) and the crossing of activations (~ 0.22 s) (F.2E), most likely due to muscle-tendons dynamics. Finally, with cocontraction instructions, cocontraction still evolved in a non-monotonic fashion during movement (F.2H): starting from the initial fixed 1 N value, it first decreased, then increased about the middle of the movement, reaching a peak at peak velocity and plateauing at movement end.

These preliminary predictions are globally consistent with the non-monotonic stiffness reported by Gomi and Kawato [1996] for reaching movements, even though the modelled system here is extremely simplified, and stiffness is not entirely determined by cocontraction.

Further development should consist in running the simulations with receding horizon (which was not possible due to computational complications), adopting more realistic arm and muscle models, possibly integrate reflexes. Moreover efforts should be made in clearing the blurred lines subsisting in cocontraction experimental assessment through stiffness or muscle activity.

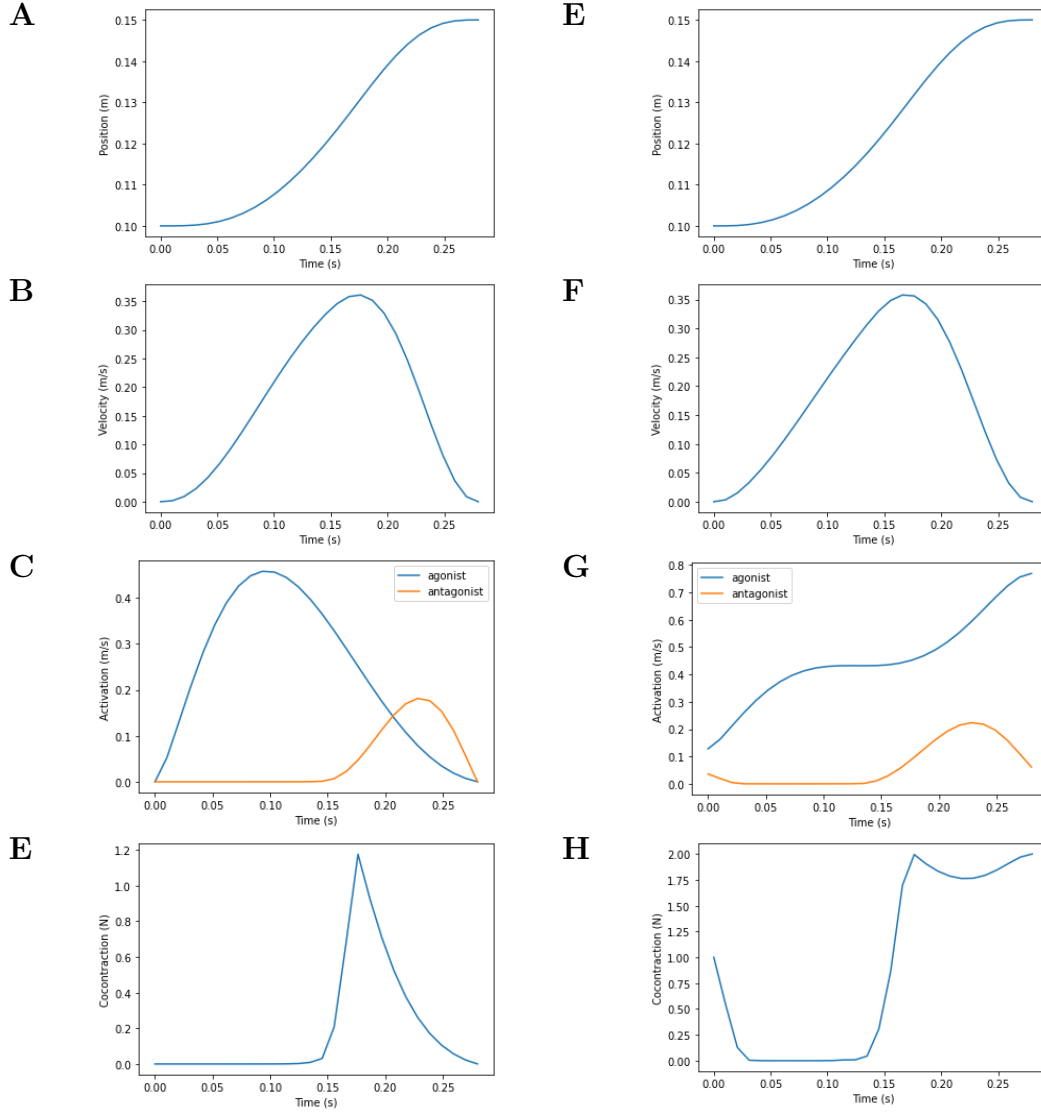


FIGURE F.2: **Reaching task with and without cocontraction instructions** **A.** Position over time without cocontraction instructions. **B.** Velocity over time without cocontraction instructions. **C.** Cocontraction over time without cocontraction instructions. **D.** Muscle activations over time without cocontraction instructions. **E.** Position over time with cocontraction instructions. **F.** Velocity over time with cocontraction instructions. **G.** Muscle activations over time with cocontraction instructions. **H.** Cocontraction over time with cocontraction instructions. Parameters are $m = 1 \text{ kg}$, agonist and antagonist muscle-tendon units having the same properties $L_m^0 = 0.07 \text{ m}$, $L_t^0 = 0.05 \text{ m}$, $\nu = 0.05 \text{ s}$, $k_m = 100$, $k_t = 500$.

Bibliography

- Aftalion, A. and Trélat, E. (2021). Pace and motor control optimization for a runner. *J Math Biol*, 83(1):9. *Cited page 102*
- Apthorp, D., Nagle, F., and Palmisano, S. (2014). Chaos in balance: Non-linear measures of postural control predict individual variations in visual illusions of motion. *PLoS One*, 9(12):e113897. *Cited page 43*
- Battaglia-Mayer, A. and Caminiti, R. (2019). Corticocortical systems underlying high-order motor control. *J Neurosci*, 39(23):4404–4421. *Cited page 8*
- Becker, M., Calame, D., Wrobel, J., and Person, A. (2020). Online control of reach accuracy in mice. *J Neurophysiol*, 124(6):1637–1655. *Cited page 98*
- Berio, D., Calinon, S., and Fol Leymarie, F. (2017). Dynamic graffiti stylisation with stochastic optimal control. In *Proc 4th International Conference on Movement Computing, Article No. 18*. *Cited page 31*
- Berniker, M. and Körding, K. (2015). Deep networks for motor control functions. *Front Comput Neurosci*, 9:32. *Cited page 78*
- Berniker, M. and Penny, S. (2019). A normative approach to neuromotor control. *Biol Cybern*, 113(1-2):83–92. *Cited page 96*
- Bernstein, N. (1967). *The Co-ordination and Regulation of Movements*. Pergamon Press, Oxford, UK. *6 citations page 2, 6, 11, 14, 53, and 54*
- Berret, B., Chiovetto, E., Nori, F., and Pozzo, T. (2011). Evidence for composite cost functions in arm movement planning: An inverse optimal control approach. *PLoS Comput Biol*, 7(10):e1002183. *2 citations page 39 and 102*
- Bollu, T., Whitehead, S., Prasad, N., Walker, J., Shyamkumar, N., Subramaniam, R., Kardon, B., Cohen, I., and Goldberg, J. (2019). Automated homepage training of mice in a hold-still center-out reach task. *J Neurophysiol*, 121(2):500–512. *Cited page 98*
- Botzer, L. and Karniel, A. (2009). A simple and accurate onset detection method for a measured bell-shaped speed profile. *Front Neurosci*, 3:61. *Cited page 35*

- Bouffard, J., Bouyer, L., Roy, J., and Mercier, C. (2014). Tonic pain experienced during locomotor training impairs retention despite normal performance during acquisition. *J Neurosci*, 34(28):9190–9195. Cited page [54](#)
- Brown, A. and Teskey, G. (2014). Motor cortex is functionally organized as a set of spatially distinct representations for complex movements. *J Neurosci*, 34(41):13574–13585. Cited page [100](#)
- Brown, G., Seethapathi, N., and Srinivasan, M. (2021). A unified energy-optimality criterion predicts human navigation paths and speeds. *Proc Natl Acad Sci USA*, 118(29):e2020327118. Cited page [102](#)
- Bryson, A. (1999). *Dynamic Optimization*. Prentice-Hall, Englewood Cliffs, NJ. Cited page [25](#)
- Burdet, E., Osu, R., Franklin, D., Milner, T., and Kawato, M. (2001). The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature*, 414(6862):446–449. Cited page [75](#)
- Burke, R. (2011). *Motor Units: Anatomy, Physiology, and Functional Organization*, pages 345–422. John Wiley and Sons, Ltd. Cited page [10](#)
- Carlson, C. and Devinsky, O. (2009). The excitable cerebral cortex fritsch g, hitzig e. uber die elektrische erregbarkeit des grosshirns. *arch anat physiol wissen* 1870;37:300-32. *Epilepsy Behav*, 15(2):131–132. Cited page [98](#)
- Chouinard, P. and Paus, T. (2006). The primary motor and premotor areas of the human cerebral cortex. *Neuroscientist*, 12(2):143–152. Cited page [8](#)
- Churchland, M., Cunningham, J., Kaufman, M., Foster, J., Nuyujukian, P., Ryu, S., and Shenoy, K. (2012). Neural population dynamics during reaching. *Nature*, 487(7405):51–56. Cited page [98](#)
- Cirstea, M. and Levin, M. (2000). Compensatory strategies for reaching in stroke. *Brain*, 123(5):940–953. Cited page [15](#)
- Crevecoeur, F. and Scott, S. (2014). Beyond muscles stiffness: Importance of state-estimation to account for very fast motor corrections. *PLoS Comput Biol*, 10(10):e1003869. Cited page [76](#)
- Criscimagna-Hemminger, S., Donchin, O., Gazzaniga, M., and Shadmehr, R. (2003). Learned dynamics of reaching movements generalize from dominant to nondominant arm. *J Neurophysiol*, 89(1):168–176. Cited page [73](#)
- Currie, S., Ammer, J., Premchand, B., Wu, Y., Eleftheriou, C., Faisal, A., Hennig, M., and Duguid, I. (2020). Dense movement with embedded sparse action-type representations in the output layer of motor cortex. Preprint (bioRxiv 2020.10.27.357087v1). Cited page [98](#)

- Darainy, M., Mattar, A., and Ostry, D. (2009). Effects of human arm impedance on dynamics learning and generalization. *J Neurophysiol*, 101(6):3158–3168. Cited page [55](#)
- Davidson, P. and Wolpert, D. (2005). Widespread access to predictive models in the motor system: A short review. *J Neural Eng*, 2(3):S313–S319. Cited page [59](#)
- de Grosbois, J. and Tremblay, L. (2016). Quantifying online visuomotor feedback utilization in the frequency domain. *Behav Res Meths*, 48(4):1653–1666. Cited page [63](#)
- de Rugy, A., Loeb, G., and Carroll, T. (2012). Muscle coordination is habitual rather than optimal. *J Neurosci*, 32(21):7384–7391. Cited page [102](#)
- Denier van der Gon, J. and Thuring, J. (1965). The guiding of human writing movements. *Kybernetik*, 2(4):145–148. Cited page [102](#)
- Dideriksen, J., Negro, F., and Farina, D. (2015). The optimal neural strategy for a stable motor task requires a compromise between level of muscle cocontraction and synaptic gain of afferent feedback. *J Neurophysiol*, 114(3):1895–1911. Cited page [153](#)
- Diedrichsen, J., Hashambhoy, Y., Rane, T., and Shadmehr, R. (2005). Neural correlates of reach errors. *J Neurosci*, 25(43):9919–9931. Cited page [55](#)
- Diedrichsen, J., Shadmehr, R., and Ivry, R. (2010). The coordination of movement: Optimal feedback control and beyond. *Trends Cogn Sci*, 14(1):31–39. Cited page [102](#)
- Dimitriou, M., Wolpert, D., and Franklin, D. (2013). The temporal evolution of feedback gains rapidly update to task demands. *J Neurosci*, 33(26):10898–108909. Cited page [32](#)
- Domkin, D., Laczko, J., Jaric, S., Johansson, H., and Latash, M. (2002). Structure of joint variability in bimanual pointing tasks. *Exp Brain Res*, 143(1):11–23. Cited page [31](#)
- Drotár, P., Mekyska, J., Rektorová, I., Masarová, L., Smékal, Z., and Faundez-Zanuy, M. (2016). Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson’s disease. *Artif Intell Med*, 67:39–46. Cited page [43](#)
- Enoka, R. and Duchateau, J. (2008). Muscle fatigue: What, why and how it influences muscle function. *J Physiol (Lond)*, 586(1):11–23. Cited page [141](#)
- Faisal, A., Selen, L., and Wolpert, D. (2008). Noise in the nervous system. *Nat Rev Neurosci*, 9(4):292–303. Cited page [13](#)
- Feldman, A. (2015). *Referent Control of Action and Perception*. Springer, Montreal. Cited page [51](#)

- Feldman, A. and Levin, M. (1995). The origin and use of positional frames of reference in motor control. *Behav Brain Sci*, 18(4):723–744. Cited page [2](#)
- Fitts, P. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J Exp Psychol*, 47(6):381–391. 2 citations page [23](#) and [29](#)
- Flanagan, J., Vetter, P., Johansson, R., and Wolpert, D. (2003). Prediction precedes control in motor learning. *Curr Biol*, 13(2):146–150. Cited page [59](#)
- Flash, T. and Hogan, N. (1985). The coordination of arm movements: An experimentally confirmed mathematical model. *J Neurosci*, 5(7):1688–1703. 4 citations page [15](#), [56](#), [65](#), and [102](#)
- Franklin, D. and Wolpert, D. (2011). Computational mechanisms of sensorimotor control. *Neuron*, 72(3):425–442. 2 citations page [2](#) and [13](#)
- Friston, K. (2011). What is optimal about motor control? *Neuron*, 72(3):488–498. Cited page [2](#)
- Gao, J., Parsons, L., Bower, J., Xiong, J., Li, J., and Fox, P. (1996). Cerebellum implicated in sensory acquisition and discrimination rather than motor control. *Science*, 272(5261):545–547. Cited page [73](#)
- Georgopoulos, A., Kalaska, J., Caminiti, R., and Massey, J. (1982). On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *J Neurosci*, 2(11):1527–1537. Cited page [81](#)
- Georgopoulos, A., Kettner, R., and Schwartz, A. (1988). Primate motor cortex and free arm movements to visual targets in 3-dimensional space. II. Coding of the direction of movement by a neuronal population. *J Neurosci*, 8(8):2928–2937. Cited page [82](#)
- Gera, G., Freitas, S., Latash, M., Monahan, K., Schöner, G., and Scholz, J. (2010). Motor abundance contributes to resolving multiple kinematic task constraints. *Motor Control*, 14(1):83–115. Cited page [11](#)
- Glencross, D. (1980). Levels and strategies of response organization. In Stelmach, G. and Requin, J., editors, *Tutorials in Motor Behavior, Advances in Psychology, Vol 1*, pages 551–566. North-Holland, Amsterdam. 2 citations page [2](#) and [53](#)
- Gomi, H. and Kawato, M. (1996). Equilibrium-point control hypothesis examined by measured arm stiffness during multijoint movement. *Science*, 272(5258):117–120. 2 citations page [153](#) and [156](#)
- Graziano, M. (2016). Ethological action maps: A paradigm shift for the motor cortex. *Trends Cogn Sci*, 20(2):121–132. Cited page [100](#)

- Gribble, P., Mullin, L., Cothros, N., and Mattar, A. (2003). Role of cocontraction in arm movement accuracy. *J Neurophysiol*, 89(5):2396–2405. Cited page [153](#)
- Gribble, P., Ostry, D., Sanguineti, V., and Laboissière, R. (1998). Are complex control signals required for human arm movement? *J Neurophysiol*, 79(3):1409–1424. 2 citations page [50](#) and [155](#)
- Griffin, D., Hoffman, D., and Strick, P. (2015). Corticomotoneuronal cells are “functionally tuned”. *Science*, 350(6261):667–670. 2 citations page [88](#) and [89](#)
- Griffin, D., Hudson, H., Belhaj-Saïf, A., McKiernan, B., and Cheney, P. (2008). Do corticomotoneuronal cells predict target muscle EMG activity? *J Neurophysiol*, 99(3):1169–1986. Cited page [88](#)
- Grillner, S. (1985). Neurobiological bases of rhythmic motor acts in vertebrates. *Science*, 228(4696):143–149. Cited page [98](#)
- Guigon, E. (2022). A computational theory for the production of limbs movements. *Psychol Rev*. 15 citations page [2](#), [3](#), [33](#), [38](#), [48](#), [49](#), [57](#), [58](#), [59](#), [65](#), [74](#), [75](#), [79](#), [103](#), and [142](#)
- Guigon, E., Baraduc, P., and Desmurget, M. (2007a). Coding of movement- and force-related information in primate primary motor cortex: A computational approach. *Eur J Neurosci*, 26(1):250–260. 4 citations page [38](#), [50](#), [96](#), and [97](#)
- Guigon, E., Baraduc, P., and Desmurget, M. (2007b). Computational motor control: Redundancy and invariance. *J Neurophysiol*, 97(1):331–347. Cited page [100](#)
- Guigon, E., Baraduc, P., and Desmurget, M. (2008). Optimality, stochasticity, and variability in motor behavior. *J Comput Neurosci*, 24(1):57–68. 3 citations page [20](#), [41](#), and [57](#)
- Guigon, E., Chafik, O., Jarrassé, N., and Roby-Brami, A. (2019). Experimental and theoretical study of velocity fluctuations during slow movements in humans. *J Neurophysiol*, 121(2):715–727. 15 citations page [15](#), [16](#), [32](#), [33](#), [36](#), [37](#), [42](#), [44](#), [45](#), [58](#), [65](#), [79](#), [102](#), [103](#), and [142](#)
- Guigon, E., Moullet, E., and Roby-Brami, A. (2022). The common temporal structure of kinematic and dynamic motor signals. Preprint. 2 citations page [47](#) and [102](#)
- Haar, S. and Donchin, O. (2020). A revised computational neuroanatomy for motor control. *J Cogn Neurosci*, 32(10):1823–1836. 4 citations page [2](#), [7](#), [78](#), and [103](#)
- Haith, A., Huberdeau, D., and Krakauer, J. (2015). Hedging your bets: Intermediate movements as optimal behavior in the context of an incomplete decision. *PLoS Comput Biol*, 11(3):e1004171. Cited page [31](#)

- Han, C., Arbib, M., and Schweighofer, N. (2008). Stroke rehabilitation reaches a threshold. *PLoS Comput Biol*, 4(8):e1000133. *Cited page 78*
- Harris, C. and Wolpert, D. (1998). Signal-dependent noise determines motor planning. *Nature*, 394(6695):780–784. *5 citations page 13, 15, 29, 42, and 65*
- Hartigan, J. and Hartigan, P. (1985). The Dip test of unimodality. *Ann Stat*, 13(1):70–84. *Cited page 82*
- Henneman, E., Clamann, H., Gillies, J., and Skinner, R. (1974). Rank order of motoneurons within a pool: Law of combination. *J Neurophysiol*, 37(6):1338–1349. *Cited page 10*
- Henneman, E., Somjen, G., and Carpenter, D. (1965). Excitability and inhibitability of motoneurons of different sizes. *J Neurophysiol*, 28(3):599–620. *Cited page 10*
- Hennequin, G., Vogels, T., and Gerstner, W. (2014). Optimal control of transient dynamics in balanced networks supports generation of complex movements. *Neuron*, 82(6):1394–1406. *Cited page 98*
- Hill, A. (1953). The mechanics of active muscle. *Proc R Soc Lond B Biol Sci*, 141(902):104–117. *Cited page 10*
- Hinton, G. (2007). Learning multiple layers of representation. *Trends Cogn Sci*, 11(10):428–434. *Cited page 79*
- Hira, R., Ohkubo, F., Ozawa, K., Isomura, Y., Kitamura, K., Kano, M., Kasai, H., and Matsuzaki, M. (2013). Spatiotemporal dynamics of functional clusters of neurons in the mouse motor cortex during a voluntary movement. *J Neurosci*, 33(4):1377–1390. *Cited page 99*
- Hogan, N. (1984a). Adaptive control of mechanical impedance by coactivation of antagonist muscles. *IEEE Trans Automat Control*, AC-29(8):681–690. *2 citations page 50 and 153*
- Hogan, N. (1984b). An organizing principle for a class of voluntary movements. *J Neurosci*, 4(11):2745–2754. *3 citations page 22, 23, and 24*
- Hogan, N. (1985). The mechanics of multi-joint posture and movement control. *Biol Cybern*, 52(5):315–331. *Cited page 75*
- Huang, V. and Shadmehr, R. (2007). Evolution of motor memory during the seconds after observation of motor error. *J Neurophysiol*, 97(6):3976–3985. *Cited page 55*
- Huh, D. and Sejnowski, T. (2015). Spectrum of power laws for curved hand movements. *Proc Natl Acad Sci USA*, 112(29):E3950–E3958. *Cited page 2*

- Hunter, S., Ryan, D., Ortega, J., and Enoka, R. (2002). Task differences with the same load torque alter the endurance time of submaximal fatiguing contractions in humans. *J Neurophysiol*, 88(6):3087–3096. *Cited page 141*
- Hwang, E., Smith, M., and Shadmehr, R. (2006). Adaptation and generalization in acceleration-dependent force fields. *Exp Brain Res*, 169(4):496–506. *Cited page 55*
- Isomura, Y., Harukuni, R., Takekawa, T., Aizawa, H., and Fukai, T. (2009). Microcircuitry coordination of cortical motor information in self-initiation of voluntary movements. *Nat Neurosci*, 12(12):1586–1593. *2 citations page 98 and 99*
- Izawa, J., Rane, T., Donchin, O., and Shadmehr, R. (2008). Motor adaptation as a process of reoptimization. *J Neurosci*, 28(11):2883–2891. *9 citations page 50, 54, 56, 57, 65, 66, 72, 73, and 101*
- Kalidindi, H., Cross, K., Lillicrap, T., Omrani, M., Falotico, E., Sabes, P., and Scott, S. (2021). Rotational dynamics in motor cortex are consistent with a feedback controller. *Elife*, 10:e67256. *4 citations page 2, 96, 97, and 98*
- Kandel, E., Schwartz, J., Jessell, T., Siegelbaum, S., and Hudspeth, A., editors (2013). *Principles of Neural Science, 5th ed.* McGraw-Hill Professional, New York, NY. *Cited page 8*
- Kawai, R., Markman, T., Poddar, R., Ko, R., Fantana, A., Dhawale, A., Kampff, A., and Olveczky, B. (2015). Motor cortex is required for learning but not for executing a motor skill. *Neuron*, 86(3):800–812. *Cited page 98*
- Kelso, J. (1995). *Dynamic Patterns*. MIT Press, Cambridge, MA. *Cited page 2*
- Kingma, D. and Ba, J. (2014). Adam: A method for stochastic optimization. Preprint (arXiv 1412.6980). *Cited page 80*
- Kiritani, T., Wickersham, I., Seung, H., and Shepherd, G. (2012). Hierarchical connectivity and connection-specific dynamics in the corticospinal-corticostriatal microcircuit in mouse motor cortex. *J Neurosci*, 32(14):4992–5001. *Cited page 98*
- Kirk, D. (2004). *Optimal Control Theory: An Introduction*. Dover, Mineola, NY. *2 citations page 25 and 26*
- Kodl, J., Ganesh, G., and Burdet, E. (2011). The CNS stochastically selects motor plan utilizing extrinsic and intrinsic representations. *PLoS One*, 6(9):e24229. *Cited page 102*
- Koelewijn, A. and van den Bogert, A. (2022). Antagonistic co-contraction can minimize muscular effort in systems with uncertainty. *PeerJ*, 10:e13085. *Cited page 153*

- Krakauer, J., Hadjiosif, A., Xu, J., Wong, A., and Haith, A. (2019). Motor learning. *Compr Physiol*, 9(2):613–663. *2 citations page 53 and 54*
- Kriegeskorte, N. and Golan, T. (2019). Neural network models and deep learning. *Curr Biol*, 29(7):R231–R236. *2 citations page 78 and 79*
- Kuberski, S. and Gafos, A. (2019). The speed-curvature power law in tongue movements of repetitive speech. *PLoS One*, 14(3):e0213851. *Cited page 43*
- Lackner, J. and DiZio, P. (1994). Rapid adaptation to Coriolis force perturbations of arm trajectory. *J Neurophysiol*, 72(1):299–313. *Cited page 72*
- Lacquaniti, F., Terzuolo, C., and Viviani, P. (1983). The law relating the kinematic and figural aspects of drawing movements. *Acta Psychol (Amst)*, 54(1-3):115–130. *Cited page 102*
- Latash, M. (2012). The bliss (not the problem) of motor abundance (not redundancy). *Exp Brain Res*, 217(1):1–5. *2 citations page 11 and 14*
- Latash, M. (2021). Laws of nature that define biological action and perception. *Phys Life Rev*, 36:47–67. *Cited page 2*
- Latash, M. and Gottlieb, G. (1991). An equilibrium-point model for fast, single-joint movement: II. Similarity of single-joint isometric and isotonic descending commands. *J Mot Behav*, 23(3):179–191. *2 citations page 50 and 155*
- Latash, M. L. (2010). Two archetypes of motor control research. *Motor control*, 14(3):e41–e53. *Cited page 51*
- Lebedev, M., Ossadtchi, A., Mill, N., Urpí, N., Cervera, M., and Nicolelis, M. (2019). Analysis of neuronal ensemble activity reveals the pitfalls and shortcomings of rotation dynamics. *Sci Rep*, 9(1):18978. *Cited page 99*
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444. *2 citations page 3 and 78*
- Legrand, M., Jarrassé, N., de Montalivet, E., Richer, F., and Morel, G. (2021). Closing the loop between body compensations and upper-limb prosthetic movements: A feasibility study. *IEEE Trans Med Robot Bionics*, 3(1):230–240. *Cited page 103*
- Levine, S. (2013). Exploring deep and recurrent architectures for optimal control. Preprint (arXiv 1311.1761). *Cited page 78*
- Li, W. and Todorov, E. (2004). Iterative linear-quadratic regulator design for nonlinear biological movement systems. In *Proc First International Conference on Informatics in Control, Automation, and Robotics*, pp 222–229. *2 citations page 30 and 60*

- Lillicrap, T., Hunt, J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., and Wierstra, D. (2015). Continuous control with deep reinforcement learning. Preprint (arXiv 1509.02971v5). *Cited page 78*
- Lillicrap, T. and Scott, S. (2013). Preference distributions of primary motor cortex neurons reflect control solutions optimized for limb biomechanics. *Neuron*, 77(1):168–179. *3 citations page 2, 96, and 97*
- Liu, D. and Todorov, E. (2007). Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. *J Neurosci*, 27(35):9354–9368. *2 citations page 2 and 32*
- Magalhaes, F., Mello, E., and Kohn, A. (2019). Association between plantarflexion torque variability in quiet stance and during force and position tasks. *Somatosen Mot Res*, 36(4):241–248. *Cited page 141*
- Malfait, N. and Ostry, D. (2004). Is interlimb transfer of force-field adaptation a cognitive response to the sudden introduction of load? *J Neurosci*, 24(37):8084–8089. *Cited page 73*
- Maluf, K., Shinohara, M., Stephenson, J., and Enoka, R. (2005). Muscle activation and time to task failure differ with load type and contraction intensity for a human hand muscle. *Exp Brain Res*, 167(2):165–177. *Cited page 141*
- Manto, M., Bower, J., Conforto, A., Delgado-García, J., da Guarda, S., Gerwig, M., Habas, C., Hagura, N., Ivry, R., Mariën, P., Molinari, M., Naito, E., Nowak, D., Oulad Ben Taib, N., Pélisson, D., Tesche, C., Tilikete, C., and Timmann, D. (2012). Consensus paper: Roles of the cerebellum in motor control—the diversity of ideas on cerebellar involvement in movement. *Cerebellum*, 11(2):457–487. *Cited page 8*
- Marr, D. (1982). *Vision*. Freeman, New York, NY. *Cited page 1*
- Martin, T., Keating, J., Goodkin, H., Bastian, A., and Thach, W. (1996). Throwing while looking through prisms. II. Specificity and storage of multiple gaze-throw calibrations. *Brain*, 119(4):1199–1211. *Cited page 54*
- Mattar, A. and Gribble, P. (2005). Motor learning by observing. *Neuron*, 46(1):153–160. *Cited page 73*
- Mazzoni, P. and Krakauer, J. (2006). An implicit plan overrides an explicit strategy during visuomotor adaptation. *J Neurosci*, 26(14):3642–3645. *Cited page 74*
- Michaels, J., Dann, B., and Scherberger, H. (2016). Neural population dynamics during reaching are better explained by a dynamical system than representational tuning. *PLoS Comput Biol*, 12(11):e1005175. *Cited page 98*
- Milton, J., Meyer, R., Zhvanetsky, M., Ridge, S., and Insperger, T. (2016). Control at stability’s edge minimizes energetic costs: Expert stick balancing. *J R Soc Interface*, 13(119):20160212. *Cited page 101*

- Mohan, V., Bhat, A., and Morasso, P. (2019). Muscleless motor synergies and actions without movements: From motor neuroscience to cognitive robotics. *Phys Life Rev*, 30:89–111. *Cited page 2*
- Mohan, V. and Morasso, P. (2011). Passive motion paradigm: An alternative to optimal control. *Front Neurorobotics*, 5:4. *Cited page 2*
- Mottram, C., Christou, E., Meyer, F., and Enoka, R. (2005). Frequency modulation of motor unit discharge has task-dependent effects on fluctuations in motor output. *J Neurophysiol*, 94(4):2878–2887. *Cited page 141*
- Moullet, E., Roby-Brami, A., and Guigon, E. (2022). What is the nature of motor adaptation to dynamic perturbations? *PLoS Comput Biol*, 18(8):1–23. *2 citations page 79 and 103*
- Müller, J., Oulasvirta, A., and Murray-Smith, R. (2017). Control theoretic models of pointing. *ACM Trans Comput Hum Interact*, 24(4):27. *Cited page 32*
- Murata, Y., Higo, N., Hayashi, T., Nishimura, Y., Sugiyama, Y., Oishi, T., Tsukada, H., Isa, T., and Onoe, H. (2015). Temporal plasticity involved in recovery from manual dexterity deficit after motor cortex lesion in macaque monkeys. *J Neurosci*, 35(1):84–95. *Cited page 98*
- Nashed, J., Crevecoeur, F., and Scott, S. (2012). Influence of the behavioral goal and environmental obstacles on rapid feedback responses. *J Neurophysiol*, 108(4):999–1009. *Cited page 2*
- Nguyen, D. and Widrow, B. (1989). The truck backer-upper: An example of self-learning in neural networks. In *Proc International Joint Conference on Neural Networks, Vol 2*, pp 357–363. *Cited page 78*
- Nisky, I., Hsieh, M., and Okamura, A. (2014). Uncontrolled manifold analysis of arm joint angle variability during robotic teleoperation and freehand movement of surgeons and novices. *IEEE Trans Biomed Eng*, 61(12):2869–2881. *Cited page 14*
- Nixon, P. (2003). The role of the cerebellum in preparing responses to predictable sensory events. *Cerebellum*, 2(2):114–122. *Cited page 73*
- Nozaki, D., Kurtzer, I., and Scott, S. (2006). Limited transfer of learning between unimanual and bimanual skills within the same limb. *Nat Neurosci*, 9(11):1364–1366. *Cited page 55*
- Oby, E., Ethier, C., and Miller, L. (2013). Movement representation in the primary motor cortex and its contribution to generalizable EMG predictions. *J Neurophysiol*, 109(3):666–678. *Cited page 89*
- Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1):97–113. *Cited page 60*

- Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci*, 2011:156869. *Cited page 63*
- Park, J., Phillips, J., Guo, J., Martin, K., Hantman, A., and Dudman, J. (2022). Motor cortical output for skilled forelimb movement is selectively distributed across projection neuron classes. *Sci Adv*, 8(10):eabj5167. *Cited page 98*
- Parr, T., Limanowski, J., Rawji, V., and Friston, K. (2021). The computational neurology of movement under active inference. *Brain*, 144(6):1799–1818. *Cited page 2*
- Perret, J. and Vercruysse, P. (2014). Advantages of mechanical backdrivability for medical applications of force control. In *Proc Conference on Computer/Robot Assisted Surgery (CRAS)*, pages 84–86. *Cited page 60*
- Pierrot-Deseilligny, E. and Burke, D. (2005). *The Circuitry of the Human Spinal Cord: Its Role in Motor Control and Movement Disorders*. Cambridge University Press, Cambridge, UK. *Cited page 9*
- Reimer, J. and Hatsopoulos, N. (2009). The problem of parametric neural coding in the motor system. In Sternad, D., editor, *Progress in Motor Control. A Multidisciplinary Perspective*, pages 243–259. Springer, Berlin. *2 citations page 79 and 97*
- Rigoux, L. and Guigon, E. (2012). A model of reward- and effort-based optimal decision making and motor control. *PLoS Comput Biol*, 8(10):e1002716. *Cited page 35*
- Rohrer, B., Fasoli, S., Krebs, H., Hughes, R., Volpe, B., Frontera, W., Stein, J., and Hogan, N. (2002). Movement smoothness changes during stroke recovery. *J Neurosci*, 22(18):8297–8304. *Cited page 15*
- Rouder, J., Speckman, P., Sun, D., Morey, R., and Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon Bull Rev*, 16(2):225–237. *Cited page 63*
- Rudroff, T., Christou, E., Poston, B., Bojsen-Moller, J., and Enoka, R. (2007). Time to failure of a sustained contraction is predicted by target torque and initial electromyographic bursts in elbow flexor muscles. *Muscle Nerve*, 35(5):657–666. *Cited page 141*
- Rumelhart, D., Hinton, G., and Williams, R. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088):533–536. *Cited page 78*
- Russo, A., Bittner, S., Perkins, S., Seely, J., London, B., Lara, A., Miri, A., Marshall, N., Kohn, A., Jessell, T., Abbott, L., Cunningham, J., and Churchland, M. (2018). Motor cortex embeds muscle-like commands in an untangled population response. *Neuron*, 97(4):953–966. *4 citations page 89, 90, 93, and 98*

- Saliba, C., Rainbow, M., Selbie, W., Deluzio, K., and Scott, S. (2020). Co-contraction uses dual control of agonist-antagonist muscles to improve motor performance. Preprint (bioRxiv 2020.03.16.993527v1). *Cited page 153*
- Sánchez-Sánchez, C., Izzo, D., and Hennes, D. (2016). Learning the optimal state-feedback using deep networks. In *Proc IEEE Symposium Series on Computational Intelligence (SSCI)*, pp 1-8. *Cited page 78*
- Saxena, S., Russo, A., Cunningham, J., and Churchland, M. (2021). Motor cortex activity across movement speeds is predicted by network-level strategies for generating muscle activity. Preprint (bioRxiv 2021.02.01.429168). *5 citations page 78, 90, 93, 97, and 98*
- Schieber, M. and Rivlis, G. (2007). Partial reconstruction of muscle activity from a pruned network of diverse motor cortex neurons. *J Neurophysiol*, 97(1):70–82. *2 citations page 88 and 89*
- Scholz, J. and Schöner, G. (1999). The uncontrolled manifold concept: Identifying control variables for a functional task. *Exp Brain Res*, 126(3):289–306. *2 citations page 14 and 31*
- Schwartz, A. (2016). Movement: How the brain communicates with the world. *Cell*, 164(6):1122–1135. *Cited page 97*
- Scott, S. (2004). Optimal feedback control and the neural basis of volitional motor control. *Nat Rev Neurosci*, 5(7):532–546. *2 citations page 100 and 102*
- Scott, S. (2008). Inconvenient truths about neural processing in primary motor cortex. *J Physiol (Lond)*, 586(5):1217–1224. *Cited page 79*
- Scott, S., Gribble, P., Graham, K., and Cabel, W. (2001a). Dissociation between hand motion and population vectors from neural activity in motor cortex. *Nature*, 413:161–165. *3 citations page 86, 88, and 96*
- Scott, W., Stevens, J., and Binder-Macleod, S. (2001b). Human skeletal muscle fiber type classifications. *Phys Ther*, 81(11):1810–1816. *Cited page 10*
- Sergio, L., Hamel-Paquet, C., and Kalaska, J. (2005). Motor cortex neural correlates of output kinematics and kinetics during isometric-force and arm-reaching tasks. *J Neurophysiol*, 94(4):2353–2378. *Cited page 85*
- Sergio, L. and Kalaska, J. (1998). Changes in the temporal pattern of primary motor cortex activity in a directional isometric force versus limb movement task. *J Neurophysiol*, 80(3):1577–1583. *2 citations page 82 and 85*
- Serradj, N., Marino, F., Moreno-Lopez, Y., Agger, S., Sloan, A., and Hollis, E. (2021). Refinement of corticospinal neuron activity during skilled motor learning. Preprint (bioRxiv 2021.03.22.436415). *Cited page 98*

- Shadmehr, R. and Krakauer, J. (2008). A computational neuroanatomy for motor control. *Exp Brain Res*, 185(3):359–381. 4 citations page 2, 7, 78, and 103
- Shadmehr, R. and Mussa-Ivaldi, F. (1994). Adaptive representation of dynamics during learning a motor task. *J Neurosci*, 14(5Pt2):3208–3224. 10 citations page 2, 24, 50, 54, 55, 56, 66, 72, 73, and 102
- Shenoy, K., Sahani, M., and Churchland, M. (2013). Cortical control of arm movements: A dynamical systems perspective. *Annu Rev Neurosci*, 36:337–359. Cited page 97
- Singh, A., Berman, S., and Nisky, I. (2018). Stochastic optimal control for modeling reaching movements in the presence of obstacles: Theory and simulation. In *Proc 7th IEEE International Conference on Biomedical Robotics and Biomechanics (Biorob)*, pages 997–1004. Cited page 31
- Stengel, R. (1994). *Optimal Control and Estimation*. Dover, New York, NY. Cited page 25
- Suminski, A., Mardoun, P., Lillicrap, T., and Hatsopoulos, N. (2015). Temporal evolution of both premotor and motor cortical tuning properties reflect changes in limb biomechanics. *J Neurophysiol*, 113(7):2812–2823. 2 citations page 86 and 100
- Sun, X., O’Shea, D., Golub, M., Trautmann, E., Vyas, S., Ryu, S., and Shenoy, K. (2022). Cortical preparatory activity indexes learned motor memories. *Nature*, 602(7896):274–279. Cited page 55
- Sussillo, D., Churchland, M., Kaufman, M., and Shenoy, K. (2015). A neural network that finds a naturalistic solution for the production of muscle activity. *Nat Neurosci*, 18(7):1025–1033. 3 citations page 78, 97, and 98
- Taïx, M., Tran, M., Souères, P., and Guigon, E. (2013). Generating human-like reaching movements with a humanoid robot: A computational approach. *J Comput Sci*, 4(4):269–284. Cited page 100
- Takahashi, C., Nemet, D., Rose-Gottron, C., Larson, J., Cooper, D., and Reinkensmeyer, D. (2006). Effect of muscle fatigue on internal model formation and retention during reaching with the arm. *J Appl Physiol*, 100(2):695–706. Cited page 54
- Thelen, E. (1995). Motor development. A new synthesis. *Am Psychol*, 50(2):79–95. Cited page 102
- Thoroughman, K. and Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. *Nature*, 407(6805):742–747. Cited page 63
- Thoroughman, K. and Taylor, J. (2005). Rapid reshaping of human motor generalization. *J Neurosci*, 25(39):8948–8953. Cited page 55

- Todorov, E. (2000). Direct cortical control of muscle activation in voluntary arm movements: A model. *Nat Neurosci*, 3(4):391–398. *3 citations page 79, 96, and 97*
- Todorov, E. (2004). Optimality principles in sensorimotor control. *Nat Neurosci*, 7(9):907–915. *Cited page 102*
- Todorov, E. and Jordan, M. (2002). Optimal feedback control as a theory of motor coordination. *Nat Neurosci*, 5(11):1226–1235. *16 citations page 2, 3, 14, 15, 20, 23, 25, 29, 30, 31, 33, 40, 42, 48, 65, and 75*
- Todorov, E. and Li, W. (2005). A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems. In *Proc American Control Conference, Vol 1, pp 300-306*. *Cited page 30*
- Torrecllos, F., Alayrangues, J., Kilavik, B., and Malfait, N. (2015). Distinct modulations in sensorimotor postmovement and foreperiod beta-band activities related to error salience processing and sensorimotor adaptation. *J Neurosci*, 35(37):12753–12765. *Cited page 55*
- Travis, A. and Woolsey, C. (1956). Motor performance of monkeys after bilateral partial and total cerebral decortications. *Am J Phys Med*, 35(5):273–310. *Cited page 98*
- Turner, R. and Desmurget, M. (2010). Basal ganglia contributions to motor control: A vigorous tutor. *Curr Opin Neurobiol*, 20(6):704–716. *Cited page 9*
- Uno, Y., Kawato, M., and Suzuki, R. (1989). Formation and control of optimal trajectory in human multijoint arm movement - minimum torque change model. *Biol Cybern*, 61(2):89–101. *Cited page 15*
- Vallbo, A. and Wessberg, J. (1993). Organization of motor output in slow finger movements. *J Physiol (Lond)*, 469:673–691. *2 citations page 65 and 102*
- van der Helm, F. and Rozendaal, L. (2000). Musculoskeletal systems with intrinsic and proprioceptive feedback. In Winters, J. and Crago, P., editors, *Biomechanics and Neural Control of Posture and Movement, Chp 11*, pages 164–174. Springer, New York, NY. *Cited page 19*
- van der Kooij, K., Brenner, E., van Beers, R., and Smeets, J. (2015). Visuomotor adaptation: How forgetting keeps us conservative. *PLoS One*, 10(2):e0117901. *2 citations page 15 and 16*
- Vyas, S., Golub, M., Sussillo, D., and Shenoy, K. (2020). Computation through neural population dynamics. *Annu Rev Neurosci*, 43:249–275. *Cited page 79*
- Wang, J., Hamner, S., Delp, S., and Koltun, V. (2012). Optimizing locomotion controllers using biologically-based actuators and objectives. *ACM Trans Graph*, 31(4):25. *Cited page 100*

- Wang, J. and Sainburg, R. (2005). Adaptation to visuomotor rotations remaps movement vectors, not final positions. *J Neurosci*, 25(16):4024–4030. Cited page [54](#)
- Wang, Z. and Majewicz Fey, A. (2018). Human-centric predictive model of task difficulty for human-in-the-loop control tasks. *PLoS One*, 13(4):e0195053. 3 citations page [44](#), [46](#), and [47](#)
- Warren, W. (2006). The dynamics of perception and action. *Psychol Rev*, 113(2):358–389. Cited page [2](#)
- Weiler, J., Gribble, P., and Pruszynski, J. (2021). Spinal stretch reflexes support efficient control of reaching. *J Neurophysiol*, 125(4):1339–1347. Cited page [9](#)
- Weiler, N., Wood, L., Yu, J., Solla, S., and Shepherd, G. (2008). Top-down laminar organization of the excitatory network in motor cortex. *Nat Neurosci*, 11(3):360–366. Cited page [98](#)
- Wolpert, D. and Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Netw*, 11(7-8):1305–1316. Cited page [73](#)
- Yanai, Y., Adamit, N., Israel, Z., Harel, R., and Prut, Y. (2008). Coordinate transformation is first completed downstream of primary motor cortex. *J Neurosci*, 28(7):1728–1732. Cited page [89](#)
- Yoshikawa, N., Suzuki, Y., Kiyono, K., and Nomura, T. (2016). Intermittent feedback-control strategy for stabilizing inverted pendulum on manually controlled cart as analogy to human stick balancing. *Front Comput Neurosci*, 10:34. Cited page [101](#)
- Yuste, R. (2015). From the neuron doctrine to neural networks. *Nat Rev Neurosci*, 16(8):487–497. Cited page [78](#)
- Zajac, F. and Gordon, M. (1989). Determining muscle’s force and action in multi-articular movement. *Exerc Sports Sci Rev*, 17:187–230. Cited page [15](#)
- Zipser, D. and Andersen, R. (1988). A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature*, 331(6158):679–684. Cited page [78](#)

