الجـــــمـــهــــورية الجـزائــريــــة الديمـقــراطــيـــة الــشــعــبــيــــة
REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
وزارة التـــعـــــلـيــــــم العـالي و البـحث العــلــمـي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
جـــامعة أبي بـكــــــــر بــلــــقـــــايـد – تـــــلمســان –
Université Aboubakr Belkaïd – Tlemcen –
Faculté de TECHNOLOGIE

THESE

Présentée pour l'obtention du **grade** de **DOCTORAT 3ème Cycle**

En : Automatique

Spécialité : Automatique

Par : MEZIANI Yeser

Sujet

A Kinematic Framework for Upper Extremity Rehabilitation Assessment: Expectation Maximization as a Motor Learning Model

Soutenue publiquement, le 29 / 06 / 2022 , devant le jury composé de :

Mme CHOUKCHOU BRARAM Amal	Professeur	Université de Tlemcen	Président
M. Guy BOURHIS	Professeur des universités	Université de Lorraine	Directeur de thèse
M. Yann MORERE	Maitre de conférences HDR	Université de Lorraine	Co- Directeur de thèse
M. HADJ-ABDELKADER M. Amine	Professeur	Université de Tlemcen	Directeur de thèse
Mme ZENATI Nadia	Directrice de Recherche	CDTA - Alger	Examinateur
M. Éric MONACELLI	Professeur des universités	Université de Versailles-saint Quentin	Examinateur
M. BENMANSOUR Mohammed	Professeur – chef de service	CHU Tlemcen	Invité

A Kinematic Framework for Upper Extremity Rehabilitation Assessment : Expectation-Maximization as a Motor Learning Model

Yeser Meziani

March 21, 2020
Version: v1.5.4

Unviersité de Tlemcen & Université de Lorraine

Doctorate Dissertation

A Kinematic Framework for Upper Extremity Rehabilitation Assessment : Expectation-Maximization as a Motor Learning Model

Yeser Meziani

Submitted in partial fulfillment of the requirements for a double Doctorate Degrees
in Control Engineering

1. Reviewer Nadia Zenati
Equipe IRVA
CDTA Alger

2. Reviewer Eric Monacelli
LISV, UVSQ
Université de Versailles-saint Quentin

Supervisors Mohammed A. Hadj-Abdelkader and
Yann Morère and
Guy Bourhis

March 21, 2020

A Kinematic Framework for Upper Extremity Rehabilitation Assessment : Expectation-Maximization as a Motor Learning Model
Doctorate Dissertation, March 21, 2020
Reviewers: Nadia Zenati , Eric Monacelli
Supervisors: Mohammed A. Hadj-Abdelkader , Yann Morère and Guy Bourhis

Unviersité de Tlemcen
Laboratoire d'Automatique de Tlemcen
Chetouan, 13000, Tlemcen

Université de Lorraine
Laboratoire de Conception, Optimisation et Modélisation des Systèmes
Metz, 57070, Metz

Abstract

Motor learning as a recovery mechanism is assumed to be a framework that drived and guided physical therapy and now since the advent of robotics doing the same to the rehabilitation devices. The rehabilitation process presents the intersection of many different interconnected facets that co-interact to produce recovered movements. The use of the technology introduces many benefits while contributing to the complexity of the phenomena at hand.

We interest our research to the passive exoskeleton training of the upper limb. We propose an adaptive intra patient assessment scale that is capable of detecting changes in intra-patient performance during robotic training. Motor learning, the process of our brain's acquiring newer motor skills or relearning those he lost due to neurological or traumatic incident is our portal to investigating this phenomenon. The interaction of the system that is composed of the device, the incentive in form of exercise games and the patients with all its level of existence, physiological, psychological, and cognitive is the system of study. The components present heterogeneous qualities and dynamically driven changes. The system output in the form of the executed trajectories is our gauging instrument to investigate the interactions within the system.

We formulate the trajectory model as a Markov Chain and use the Kalman Filter to estimate the smoothed states. While dynamics are time-variant we model the assumptions about the movement into a dynamical formulation and estimate its parameters from data. To account for the time variability we introduce parallel noise source to the dynamics and estimate it using an Expectation-Maximization algorithm. The temporal nature being another facet of the kinematic phenomena, we assume a variable temporal alignment and estimate it using Expectation-Maximization iteration to increase the likelihood of the estimated model compared to the observed trajectories.

Once learned the model dependent and extracted parameters are used to compare between differences in performance. The properties of the clinical assessment tools are investigated and results are formulated to answer the commonly reported needs. Stemming from the same fundamentals of motor learning, we aimed to define a new visual assessment instrument that is intended to fulfill the need of patient-first easily communicated feedback form. We present and assess clinical properties of the tools while providing validating results on clinical data attesting the longitudinal sensitivity of the tool. The underlying assumption of the visualization was then assessed using an objective measure of maximum probability value derived using a probabilistic model of the trajectories and expected on a highly likely trajectory model learned using a Kernel-Near-Neighbors Regressor.

Abstract (French)

L'apprentissage moteur en tant que mécanisme de récupération est supposé être un cadre qui a guidé les principes de la thérapie physique et qui, depuis l'avènement de la robotique, fait de même pour les dispositifs de réadaptation. Le processus de réadaptation représente l'intersection de nombreuses facettes interconnectées qui co-interagissent pour produire des mouvements récupérés. L'utilisation de la technologie présente de nombreux avantages tout en contribuant à la complexité du phénomène en question.

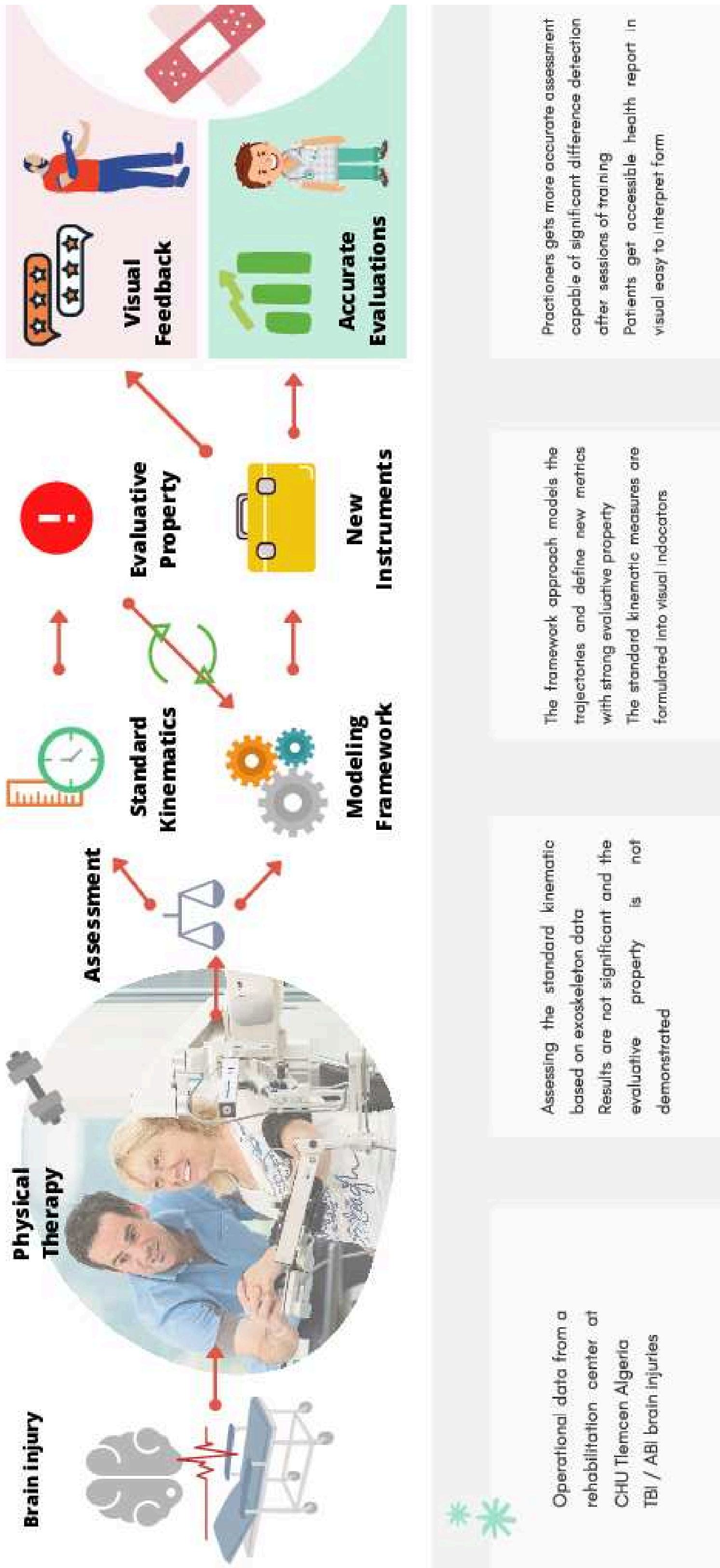
Nous intéressons notre recherche à l'entraînement passif par exosquelette du membre supérieur. Nous proposons une échelle d'évaluation adaptative intra-patient qui est capable de détecter les changements de performance intra-patient pendant l'entraînement robotique. L'apprentissage moteur, le processus par lequel notre cerveau acquiert de nouvelles capacités motrices ou réapprend celles qu'il a perdues suite à un incident neurologique ou traumatique, est notre entrée pour étudier ce phénomène. L'interaction du système composé de l'appareil, de l'incitation sous forme de jeux sérieux et des patients avec tous leurs niveaux d'existence, physiologiques, psychologiques et cognitifs est le système d'étude. Les composants présentent des qualités hétérogènes et des changements dynamiques.

La sortie du système sous la forme de trajectoires exécutées est notre instrument de mesure pour étudier les interactions au sein du système. Nous formulons le modèle de trajectoire comme une chaîne de Markov et utilisons le filtre de Kalman pour estimer les états lissés. Alors que la dynamique varie dans le temps, nous modélisons les hypothèses sur le mouvement dans une formulation dynamique et estimons ses paramètres à partir des données. Pour tenir compte de la variabilité temporelle, nous introduisons une source de bruit parallèle à la dynamique et l'estimons à l'aide d'un algorithme d'espérance-maximisation.

La nature temporelle n'étant qu'une seule facette du phénomène cinématique, nous supposons un alignement temporel variable et l'estimons en utilisant l'itération d'espérance-maximisation pour augmenter la vraisemblance du modèle estimé par rapport aux trajectoires observées. Une fois appris, les paramètres dépendants du modèle et extraits sont utilisés pour comparer les différences de performance. Les propriétés des outils d'évaluation clinique sont étudiées et les résultats sont formulés pour répondre aux besoins communément signalés.

En partant des mêmes principes fondamentaux de l'apprentissage moteur, nous avons cherché à définir un nouvel instrument d'évaluation visuelle destiné à répondre au besoin d'un formulaire de retour d'information facile à communiquer pour le patient. Nous présentons et évaluons les propriétés cliniques de l'outil tout en fournissant des résultats de validation sur des données cliniques attestant de la sensibilité longitudinale de l'outil. L'hypothèse sous-jacente de la visualisation a ensuite été évaluée à l'aide d'une mesure objective de la valeur de probabilité maximale dérivée d'un modèle probabiliste des trajectoires et appliquée sur un modèle de trajectoire hautement probable appris à l'aide d'un régresseur Kernel K-Nearest Neighbors (KNN).

Graphical Abstract



Acknowledgement

Thanks and praise is Allah's for guiding me to survive and thrive, if rightfully done it is through his guidance, all inadequacies are dutifully mine.

To my family, this work is dedicated.

To the Rehabilitation center at Tlemcen Hospital (HUC) both staff and patients whose contributions you are about to read -all I did is put them into chunky words and lofty symbols- A heartful Thank you.

To everyone who help in part or in whole during these past couples years.

To open source community for the wealth of tools that made our work too possibly efficient, Thanks are stacked to the Stackexchange gurus, keep on questioning guys!

...

To the giants on whom shoulders we stand, I humbly present an attempt.

Contents

1 General Introduction

> *The key to staying sharp in old age, is in your fingers. From your fingers to your brain and back again. If you keep your fingers busy, you'll live to see one hundred.*
>
> **— A Japanese Centenarian**

1.1 Topic & Context

Research in fields of rehabilitation, neuroscience, biomechanics, robotics and artificial intelligence has brought major progress to our understanding of how the brain processes happen, what changes happen at what anatomical levels. We have a firm grasp on many pieces of the puzzle, yet for some reason we are slightly at a loss to achieve a major breakthrough to seemingly broader horizons bringing those pieces together to reach more holistic understanding. The informed reader would instantly see the answer: further research you would say. This is true of what we have done so far in this field -like any other- yet the pluridisciplinarity of the application calls upon much more interactions from many different knowledge backgrounds and domain expertise and not to mention that the study involves the "human brain".

The introduction of much of the technology into the physical rehabilitation care delivery cycle goes a long way into enhancing the efficiency and the quality of the care, it also did augment our understanding of the phenomenon underlying the 'microscopic' root causes to 'macroscopic' changes therapists assess and observe during practice. By this far, the critical reader would have stricken the idea: why are we having this discussion in the first place? We have the questions -too many to start with- and we have more technologies to approach answers, what is missing then to reach this breakthrough? The immediate answer would be extending our toolset with more adapted tools. Since the application is a multi-disciplinary field, instead of adopting the "fork strategy" to research -in control engineers' jargon- we would benefit from "joining" existing research to expand our tools and methods of investigation and to bring together the compiled fragments of knowledge hopefully closer together.

1.2 Motivation & Problem Statement

It is striking to see that compared to the number of methodologies and techniques available in the adjacent literature few are benefiting research in the field of rehabilitation. It's also strikingly concerning that the major understandings we have of the functions of the human brain came from understanding the source of disfunction in literature's famous pathological case studies. The rehabilitation context is a major source we believe to further this understanding while assisting many patients struggling to regain their motor abilities after major neural incident. From a much pragmatic point of view, the numbers are increasingly concerning to the worldwide populations, practitioners and policy makers and payers. Stroke remains the most recurrent disability provoking incident, as USA and UE statistics puts it on top of their list [132]. In the US 795K cases/per year of stroke or recurrent attacks are reported, while EU's Italy suffers 200K cases per year [126].

On the other side of the globe the largest world population also presents a staggering rate of stroke as China reports 2M cases per year [124] 66% of which struggles with consequent motor deficits. Locally, Tlemcen CHU admits 5K cases per year for post stroke physical rehabilitation. In terms of incurred socioeconomical burden due to stroke, in western countries in general, stroke is the third largest cause of DALYs [1], meanwhile it is the second largest in developing countries [161].

In the US for example, the public expenses covering medical care for stroke are estimate at 25.2B USD [132]. Hospitalization costs for the stroke victims depending on diagnosis, sex, age and type of attack are reported at 20K USD. Therapy care takes up a fair share for per patient costs where intensive therapy costs $7382 USD, a figure that can be reduced to $5152 if therapy is robot assisted [161].

While neurological disorders holds the lion's share in motor deficiencies, orthopedic surgery and injuries such as wrist fractures is among the first causes of physical therapy visits in the US. These accidents result in disfunction requiring 67 days up to 20 weeks for recovery, and resulted in self-reported problems with productivity

[1] Disability-adjusted life years (DALYs) represent the number of lost years because of early death or years lived with a form of disability compared to population average. These can be consequences of accidents such as Traumatic Brain Injury, Spinal Cord Injury and motor neurons lesions. Besides to common neurological pathologies: Multiple Sclerosis, Cerebral Palsy, Guillain Barre Syndrome, Essential Tremor and Parkinson [105]

(50%) and self-care (40%) [123] similarly to what is reported for stroke victims in this sector.

A number of techniques has been developed in laboratories setups that proved to be highly informative, yet the most compelling of these technologies is the deployment of exoskeleton devices. These devices are relied upon to deliver constant repetitive amount of training for both upper and lower limbs and are candidate to provide deeper insights on the recovery course in objective systematic way.

1.3 Research Focus

Geographically, the area of Tlemcen agglomeration with a population of 378K has, according to the CHU of Tlemcen (Algeria), an average 5K cases of stroke yearly requiring physical rehabilitation at the CHU's Physical Rehabilitation Center (PRC).

In addition to stroke victims many different patients approach the center due to other pathologies. Objective assessments, feedback and intercation tools can augment clinical decision. The objectivity in these rehabilitation tools can boost efficiency and effectiveness of the care provided to the patients. The tools used at Tlemcen's PRC feature an upper limb exercise orthotic exoskeleton device, the Armeo Spring depicted in Fig. 1.1. The device is passive and relies on springs to compensate for the weight of the upper limb during trainings.

The device is augmented with a PC based proprietary software permitting the following functionalities:

- Virtual Reality serious exercises simulating different Activities of Daily Life (ADLs);
- Logging task parameters and events during training sessions;
- Assessment tasks providing standardized exercises to test patient's performance;
- Logging assessment results and producing feedback reports;
- Collecting and recording patients joint movements trajectories during assessment sessions;

Fig. 1.1.: Armeo Spring device, a passive orthotic exoskeleton for upper limb rehabilitation training. The reference frame attached to the shoulder sensor defines all the positional measurements recorded in the dataset used for this study. Image courtesy of Hocoma Inc. Switzerland.

The patients use this device to target specific upper limb deficiencies while following traditional physical therapy.

The interactions with the virtual environments are done by moving the device. Meanwhile, end effector movements are reproduced by the cursor on the screen. In grasping exercises a force sensor captures the grip at the robot's handle (end effector) and permits catching targets to move and place following the game's scenario. The Assessment task presented to the patients focuses on the $X-Y$ vertical plane of the reaching task. The patient is presented with continuously appearing visual targets on the screen to be pointed using the cursor before the end of a time slot. Difficulty of the session increases the number of targets appearing per session. Targets appear in random succession motivating nonstop reaching movements in different directions. The raw CHU-PRC dataset description is defined in annex A.

We formulate the definition of an elementary trajectory (ET) as the end effector trajectory performed to point a target. The start of the trajectory is either initiated by the appearance of the first target in a session, the previous target position in case of successful task or the previous position at the end of the dedicated time slot. A drawing representing the definition of ETs is depicted in Fig. 1.4.

Outcome measures are defined based on the kinematically evaluated trajectories. Inspecting the kinematic nature of the movement through a modeling technique is

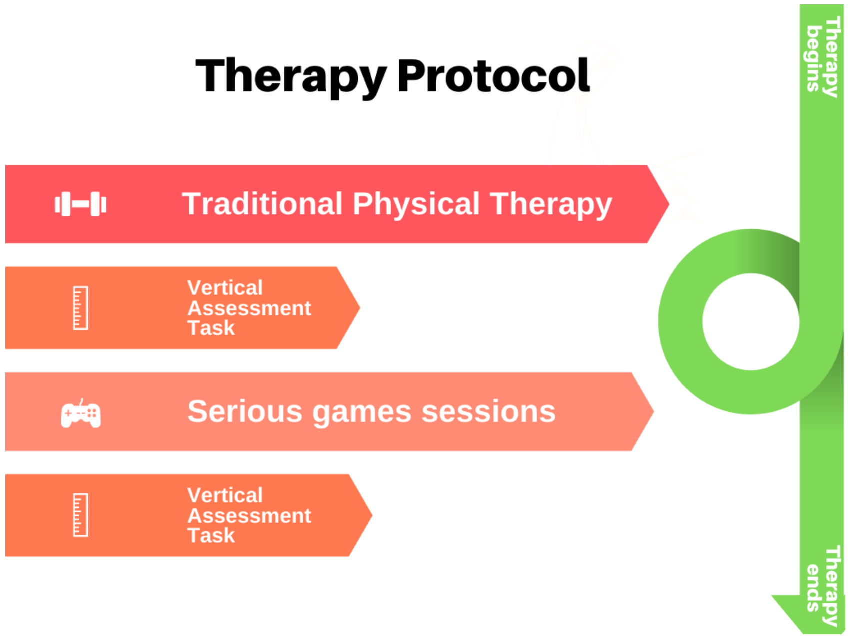


Fig. 1.2.: A description of the training protocol followed at the University Hospital of Tlemcen.

adopted as recovery process potential parameterization technique. By developing the trajectory model, this study aims to propose a methodology that is capable of adaptively monitoring patient's progress and at the same time relate further understanding on the underlying dynamics of the impaired movements. This can be achieved through accurately parameterizing the process and providing sensitive scales to the progress of the rehabilitation. The methodology aims to establish a valid assessment tool and a mean of parameterization of the process that can help reply to commonly reported challenges.

1.4 Relevence of the Research

Getting finer grip on the recovery mechanisms and recovery course is at the core of the major current research directions.

The research on objective assessment using kinematic measures is reaching its fruitful harvest summer as we saw the tool being used progressively by trial studies involving robotic systems. Standardization and validation endeavors are increasingly adopted and published. Meanwhile, the compiled evidence suggests that many challenges are still to be addressed.

Fig. 1.3.: The pointing task generates the data. Patients are using the exoskeleton to drive the cursor and point out targets. The trajectories of the end effector during these sequences of movements constitute the elementary study units noted ET. The Armeo Spring picture is courtesy of Hocoma Inc. Switzerland.

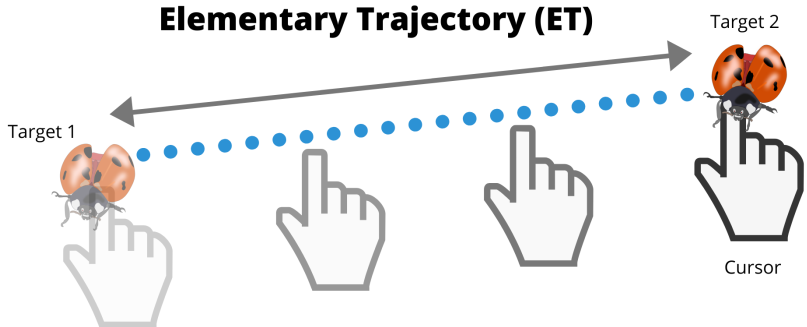


Fig. 1.4.: A description of the vertical assessment exercise task where patients are required to point at targets appearing randomly on the screen by moving a cursor using the Armeo device. An elementary trajectory is thus the trajectory followed by the end effector to achieve this elementary task.

From a virtual reality perspective, Virtual Reality games exercises are attention consuming care delivery medium that would benefit from tools for task difficulty adaptation. As much as that would benefit automatic task setting, a constant per trial or a per session monitoring and assessment scale would be beneficial to the therapist and patients alike.

Modeling recovery using computational approaches is a promising research orientation that holds the future of technology assisted rehabilitation. By providing extensible modeling tools, researchers can take advantage of the theoretical framework to systematically address questions and test interactions or drive conclusions. This can be highly enlightening to the findings related to motor learning theory and our understanding of human body functions.

Assessment and feedback tools are at the heart of the needs of key stakeholders -practitioners and patients- and can be immediately fulfilled by the introduction of these tools into the physical therapy setting.

Modeling robotic trajectories in scenarios involving the interactions of humans and robots is the topic of collaborative robotics, and shared control strategies for example. A tool for predicting human-aware movement trajectories can be extremely useful in establishing user friendly and more secure path planning routines. The additional modeling approach would be a candidate for researchers in these adjacent application fields to help compliant control strategies based on single human users movement profiles for example.

Overall the effectiveness and quality of care through adaptive training, high resolution assessment, monitoring, and outcome predictions can be beneficial from an organizational perspective as it would increase their patients satisfaction along their recovery, practitioners effectiveness and efficacity, while heightening payers money value.

1.5 Aims & Objectives

The aims of the current study can be related as follows:

- For the practitioners to provide an easily retrieved objective scale of rehabilitation status;
- For the patients to provide feedback on their status and development throughout the training course;

- For the rehabilitation device to permit the possibility to adaptively setup tasks according to the patients instantaneous performance;
- For the researcher, expanding their toolset with newer methods and techniques to help frame and investigate different research questions;
- For the healthcare system, to eventually help maintain high care efficiency and efficacy and thus maximizing the return on the expenditure.

The objectives are listed as follows:

- Investigating the state of the art kinematic scales using operational dataset from the CHU-PRC;
- Modeling the cartesian trajectories during Armeo assessment task using adaptive learning technique;
- Defining assessments instruments to quantify and study performance changes accurately;
- Investigate the clinical properties of the proposed instruments using standard clinical guidelines;
- Develop feedback forms that can be intuitively communicated to patients and can objectively map recovery progress.

1.6 Thesis Structure

The next part of the dissertation handles the state of the art perspective on the problem. We detail the literature covering the issue to define major trends and methodologies. The state of the art tools are then evaluated in a pre-study and results are presented to motivate the current issues and challenges that need to be addressed to further extend our knowledge on the topic.

Chapter 2

The literature review aims at developing a basic understanding of the phenomenon, the underlying causes and mechanisms that play a critical role in the process of recovery, lists common needs by the stakeholders of the rehabilitation ecosystem, and introduces the different tools used in practice. In the chapter the information are presented gradually in a comprehensible way hoping to provide a reference form for anyone interested in tackling this research as we will discuss in the general conclusion.

Chapter 3

In this chapter we conduct early data analysis and cross check state of the art tools (kinematics). By using measurements reported in the trials literature in robotic therapy, we proceed to investigate metrics performance and properties on a rehabilitation operational dataset. We highlight description of the dataset and early results on quantifying recovery from kinematic data. We also present the difficulties and the challenges and motivate for further research on the topic.

***

We then introduce the theoretical contribution. We start by providing the theoretical background and its relation to the topic. We state the rational of the methodology choice and then relate the methodology and the algorithms proposed to model the executions during therapy sessions. The assessment tools are then presented and results are detailed. Finally, conclusions are drawn in light of the current state of the art and potential extensions.

Chapter 4

In this chapter, an introduction of the proposed modeling technique is presented. The framework used is formulated within the proper context and referring to the underlying theoretical formulations. Once the framework with its algorithms are presented, a layout of the defined metrics and visuals is presented. A detailed plan of the statistical testing and methodology employed in the study is also presented to facilitate reconstruction of results.

Chapter 5

The results of each phase of the framework are listed and commented. Afterwards, a discussion details the intricacies and potential shortcomings of the work while also highlighting the major results and consequences of the presented results. To tie the results and findings to the actual motivation of the work a conclusion links these findings to the stated motivations and draws parallels to the needs of each of the stakeholders.

Chapter 6

To further the results from the feasibility study and cross check the elementary results. The chapter tackles a clinimetrics evaluation study. The study investigates the effects and properties listed in the feasibility study and further details the properties in clinical jargon for application specific needs.

***

Afterwards we focuse on another parallel aspect of rehabilitation that is of interest to the stakeholders in the rehabilitation context, Feedback. The rehabilitation of upper limb is presented in a visual form using a preprocessing technique, visualization propositions are made to help communicate performance and discuss findings.

Chapter 7

In this chapter we present the algorithmic formulation and detailed implementation of visualization technique. The data preprocessing strategy as well as the plotting functions used are presented.

Chapter 8

The chapter presents the results and the discussions of the visual feedback tools. To establish their feasibility, results of different visual techniques are listed and compared between patients from the rehabilitation center dataset.

Chapter 9

Finally, the conclusion chapter synthesises and lists the research contributions and presents the executive summary. We also introduce the viable research directions to extend and capitalize on the findings of this research.

2 Upper Limb Rehabilitation Between Clinical Requirements and Technological Promises

"Use it and improve it, or lose it"

— **Principle of neurorehabilitation**

In this chapter, we provide a comprehensive overview of the literature. We start by giving ample background covering injuries inducing motor deficit. We present the perspectives of each rehabilitation stakeholder, and we try to highlight their needs. We discuss the therapy tools and introduce the assessment of physical capacities. We present the state-of-the-art kinematic scale along with its current limitations and challenges. We finalize the chapter by synthesizing the literature on advanced rehabilitation methods and assessment techniques. We try to give considerable interest to kinematic only methodologies.

2.1 Main Disability Causes

Stroke is leading the statistics in terms of population affected still, many pathologies cause similar motor deficits upon the upper limb. These causes provoking the upper limb motor disability can be distinguished into the following families:

- Acquired Brain Injuries (ABI): nonprogressive, non-hereditary brain injuries transpiring after stroke, trauma, hypoxia, and infection [16] as well as certain neurological diseases such as Multiple Sclerosis, Guillain Barre Syndrome, Essential Tremor, Cerebral Palsy, and Parkinson [105];
- Traumatic Brain Injury (TBI): Classifies pathologies such as Spinal Cord Injury, and injuries inflicted by motor neurons [105].

Moreover, there exist workplace injuries that lead to disability and substantial time loss. This type of accident is handled in occupational rehabilitation [63]. Among these injuries, fractures to the elbow hinder the forearm's role in supporting and stabilizing the hand [123].

2.1.1 Post Stroke

Following the onset, stroke victims experience cognitive and motor deficiencies instantly. After the acute care, patients are heavily motor-impaired, and the post-stroke therapy takes off to try to minimize the effects of the attack. It also aims to recover as much as possible of the remaining motor abilities in a race that clinicians refer to as *time is brain*.

Terminology Stroke attacks lead the statistics in terms of disability provoking neurological conditions. In the related literature, heavily pathology related terminology is used, for instance:

- The continuum of stroke phase classification: acute, sub-acute, and chronic stages post-stroke (see Fig. 2.1);
- The acute phases span from onset to 7 days. The subacute continues until six months post-stroke and onwards into the chronic phase;
- Ipsilateral/ipsilesional is an interchangeable term referring to the unaffected Upper Limb (UL) side in hemiplegia. This side suffers weaknesses nevertheless when matched to healthy subjects;
- Contralateral/contralesional refers to the affected side;
- The upper limb is segmented compared to the trunk. The segments are the shoulder and arm that are considered proximal. The forearm and hand are distal;
- Activities involving the use of a single hand are unimanual. Likewise, bimanual tasks involve both hands.

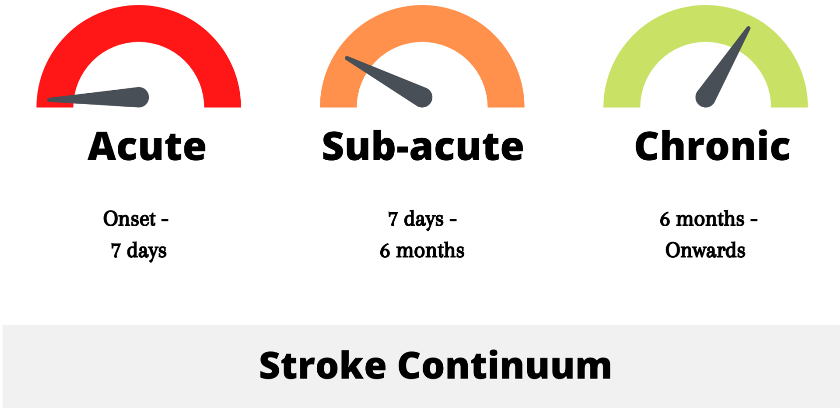


Fig. 2.1.: The post stroke continuum, the stages stroke victims go through after stroke onset.

Stroke Sequelae Post-stroke impairment engenders a variety of sensory-motor, cognitive, and psychological symptoms. Stroke patients suffer several impairments that decrease the movement quality [13, 14, 14, 57, 120], for example:

- Th paretic side exhibits weaknesses caused by the joint torque impairments;
- A reduction in active and passive range of motion;
- Spasticity defined as the intrinsic muscle resistance to the passive movement;
- Inability to fractionate movements;
- The presence of abnormal synergies;
- Deteriorated trajectories with the presence of higher deviation;
- Hyper muscle tonicity causing jagged moves;
- Inhibition of somatosensation;
- Poor limb coordination;
- Increased antagonistic muscles reflexes;
- Paresis.

Proprioception A fundamental stroke sequela that highly affects the recovery potential is proprioception. Research suggests that 33% to 50% of stroke survivors suffer impaired proprioception. Proprioception refers to the perception of joint position, force, sense of movements, and effort [43, 48]. Proprioception is a crucial component affecting recovery post-stroke. At the neuromotor level, it is a critical feedback signal enabling neural plasticity. These roles qualify it as a fundamental component for motor control and learning. This feedback source permits the motor control phenomena to generate smooth and well-coordinated movements. The loss of proprioception results in difficulties in learning movements. The loss also impedes improving the quality of the remaining motor abilities over time with functional repetitions. This form of feedback is paramount for many basic motor functionalities. Namely, joint angle control that is responsible for posture and coordination. Any deficit to this sense emerges in the form of inadequate control of movement and posture. This deficit also increases the reliance on vision to correct the movements [40].

Spasticity Spasticity is an inherent stroke sequela as it affects 50% of victims post-stroke. Other neuro conditions such as Multiple Sclerosis and Spinal Cord Injury exhibit this sequela in 85% and 71% of cases, respectively. It often worsens after stroke, reportedly increasing from 25% of patients at day three to 46% after 12 months. The European SPASM Network describes spasticity as a disorder of sensory-motor control following a motor neuron lesion syndrome. This deficit is present in many forms, sustained or intermittent involuntary muscle activations, a muscle tone dependent on velocity increasing the stretch reflex by exaggerating tendon jerks or contractures. It highly hurts the rehabilitation, limiting function and independence while causing pain and additional impairments [133].

Reaching movement disorders During typical daily activities (ADLs), the reaching component is predominantly part of the movement sequence and is thus essential to the achievement of the desired tasks. Although relatively simple to the observer, the stereotypical healthy motor control features well-coordinated movements which are dependent on hand-eye coordination, inter and intra- limb coordination [149],higher Central Nervous System (CNS) planning and relies on the high redundancy of the human musculoskeletal system, including high interactions with the outer environment [122]. At the control level, synergies are coactivations of motor units to achieve a certain functional goal. The collected set of synergies make up the motor control building blocks that allow the coordination of complicated movements [149]. For instance, reaching outwards to grab a cup of tea is achieved by executing a set of synergies. Synergies are the motor components overlapped to produce a specific motor output. Weakness of the upper limb is considered a motor control deficit. At its core, a decrease in recruiting motor units[1], the absence of synchronization, and a lower firing rate at the synaptic terminals causing an incongruous motor output [30]. Consequently, the impaired motor control and synergies, besides the weaknesses of the upper limb, affect the movements, essentially reaching movements. Fractionated movements with near or full stops are stereotypical symptoms. The velocity of the hand has a fluctuating profile. This translates into a lower mean velocity compared to higher local peaks [84, 120]. The reaching movement seems to be drastically increased by the presence of weight support for the hemiparetic side. The increase appears to be unrelated to muscle weakness [57].

Strikingly, upper limb torques at elbow and shoulder levels are impaired after a hemiparesis [89]. Stroke survivors lose independent joint control. The contralateral motor pathway, the cortico-reticulospinal, recruited to substitute the lesion damaged pathways induce disturbances in the form of coactivations which alter the limb dynamics. This effect is known as the flexion synergy affecting the elbow extension. It involves the association of shoulder abductors and elbow flexors, where higher shoulder torque causes the issue [57].

A key distinction to make is the effects of the plane of the movement and the gravity support of arm weight. Horizontal reaching, while reflective of the loss of independent joint control, does not relate to the dynamic component of this loss. By actively elevating and maintaining the arm above support, the patient puts increased demand on the shoulder abduction torque. This engagement emphasizes the source of the impairment. This component presents an inter-subject variability and is related to stroke severity [57]. This makes it clinically important to evaluate and motivate the use of 3D or vertical assessment tasks.

2.1.2 Post TBI

Traumatic Brain Injuries (TBI) are not neurological conditions at source but have similar chronic effects and complications. TBIs are induced by accidents such as military blasts or ballistic injuries and the different forms of accidents affecting the nervous system or the brain. The diffuse of the impact source affects the deficit. The resulting changes are believed to depend on many factors, including medical care, genetics, and prior medical conditions in addition to psychological influences

[1] Fibers of the muscle tissue.

[59]. The accident launches a sequence of complications spanning inflammations, excitotoxicity, and ischemia. These complications induce what is similar to the degenerative nature in Acquired Brain Injuries that influences neuroplasticity during recovery [59].

In addition to the physiological complications, psychological drain in form of depression, stress and the flood of cortisol affects cognition and memory that are essential for motor skill relearning during the rehabilitation. TBIs reduce the hippocampus' volume, thus impeding learning and memory processing [59]

2.2 Recovery Mechanisms

2.2.1 A Neuroscientific Perspective

Studies into neuroscience have informed us about the processes that drive recovery after a neurological accident. The recovery begins for instance at the penumbral areas located right outside the lesioned area. This is followed by the reduction of fluid (edema) around the lesion with restitution from diaschisis [13].

The Central Nervous System (CNS) can restitute the damage and initiate recovery through the neuroplasticity of the brain. This property allows for reorganization of the anatomical and functional parts of the CNS [13, 132]. This plasticity is excited through the sensory input experience and motor learning [132].

The neural adaptation mechanisms are listed in [132] as follows:

- Function transfer from infarcted to intact areas by recruiting concurrent motor neuron regions;
- Modification of the cortical representations as a result of this transfer;
- Reinforcing redundant parallel synapses;
- An increased neuronal myelination;
- Increased dendrite sprouting;
- Forming new synapses.

Anatomical neuroimaging through functional Magnetic Resonance Imaging (fMRI) provided evidence for reorganization at partial and complete recovery [30].

To achieve motor function recovery, multiple brain regions collaborate. The activation of both ipsi- and contra-lateral sensorimotor cortex with the premotor areas and supplementary motor area in addition to the parietal cortex are shown to play a role [30].

Modulation of cortical motor output and motor learning tasks are found to involve the cerebellum as well. The combination of these neurological processes helps mitigate the neural deficits, not necessarily through damage replacement but rather by unmasking redundant neural pathways that replace the damaged main systems [132]. To incite the neuronal recovery mechanisms, rehabilitation exercise is known to boost optimal neuroplastic changes and enhance function through the use of high intensity repetitive multisensory task-specific training [56, 132]. The effects also benefit the proprioceptive acuity within the trained workspace [40].

The recent brain imaging studies suggest that cortical representations diminish and reexpand with the introduction of activity [152].

Proprioceptive feedback plays a critical role in reorganization and affects the overall neuromotor recovery through the use-dependent plasticity [40]. Thus, motivating investigation of ways to augment feedback modalities in the interest of boosting the use-dependent plasticity.

2.2.2 A Motor Learning Perspective

In humans, repetitive task-based training and experiences that involve cognitive components such as attention and memory result in permanent neuroplasticity. This process is referred to as Motor Learning (ML) [39]. Although goal orientation is common to the tasks referred to in motor learning, the developmental-ists argue not. The early neonate developmental stage of fetal movements begins with no specific intentions and is based purely on reflexes. Nevertheless, these early building blocks reflect late on the motor skills acquisition. The ML process spans the entire lifetime of the human and is apparent in gait, environmental interactions, and motor skills. ML also contributes to the decline in motor function for the elderly [119]. This effect is also seldom explicitly mentioned, yet it often gets factored in when researchers consider age-matched groups. Prehensile, ballistic, and manipulative Object-interaction skills are considered fundamental motor skill [119]. Consequently, modeling these skills would permit us to characterize the progress of the skill acquisition. Additionally, it enables composing fundamental components to parameterize larger, more complicated sequences. Further, by subdividing a throwing task using a components model, researchers showed that the subunits evolved at different rates within individuals and across individuals. These effects obscure the motor skills competencies assessments [119].

The challenge facing human motor learning and control mechanisms is to find a movement pattern to execute a certain task. The challenge is highlighted as the motor problem which was defined by Bernstein in 1967. The problem is invigorated by task dependency and degenerate relationships. The problem revolves around the presence of multiple configurations to reach the same outcome. The general belief holds that efficiency is a constraint on candidate solutions. The immediate indicator of motor learning apparent to the observer is a performance increase, typically viewed as speedy, smooth, and stable movements. This effect is what we colloquially term skilled motor performance [81, 156]. Skillful motor execution also relates to the individual's ability to output a predictable, temporally, and energetically efficient movement. In addition to these general optimizations, impairment-related deficits constrain the pool of possible solutions to the motor problem [119].

ML is referred to as the progressive motor skill acquisition characterized by a static error as skill reaches its stable phase [120]. The trial and error strategy favors optimizing the speed-accuracy trade-off which is related to the Fitts law. This optimization is task-dependent in individuals [81]. Consider learning new complex movements such as piano playing. The task involves detecting errors in executions and adapting iteratively by correcting to eliminate the errors [39]. Researchers investigated the effects of errors by applying exaggerated force fields perturbation to the movement. They concluded that the movement afterward resembled more the normal patterns [81]. Stroke victims move inefficiently, due to the effects of impaired motor memory and feedforward feedback. These impairments limit their motor relearning capacity [39]. Neurorehabilitation is similar and obeys the same governing principles of motor learning and is believed to drive motor recovery [81, 120]. Post-stroke rehabilitation programs increase motor recovery whereas, its time course differs based on measures and individuals. The final level achieved remains unaffected [81]. Researchers tie the motor learning process to the cortical reorganization suggesting that the repetitive motor activity alone in animals did not necessarily elicit neuroplasticity [25]. This brings back the findings relating to proprioceptive feedback and somatosensation besides proper corrective feedback and initiation and highlights their contribution to reorganization. Relatably, primary cortex activations and increased synapses recruitment visualized using fMRI were shown to follow active training compared to passive [39]. Similarly, [18] reported that active engagement induces plasticity and increases motor function after therapy whereas, passive training did not.

The exercise dosage is a critical variable in rehabilitation as studies reported increases to synaptic density in the motor cortex following 400 trials and not after 60 [81]. The fact remains, these are a single variable in the recovery equation. Intermixing the training contributed to higher retention of the acquired skills after stroke. From a motor control perspective, this can be because the brain solves a new problem for each trial instead of relying on memorized movements [81]. Consider the UL as an open kinematic chain with three segments (upper arm, forearm, and hand), this is a highly redundant structure. The redundancy obviated at the kinematic level stretches to the neuromuscular level. The CNS can choose different activation patterns resulting in the same posture [112].

On the motor control level, rehabilitation induces selective coordination of muscle contractions, the synergies combine forces to generate a newer kinematic and dynamic pattern [39, 81]. The real-time motor controller of the human has been shown to have a memory as such, patients' force output exponentially decays over the course of movement allowing momentum to carry over into the later phase of the movement [81].

Research showed that the cerebellum handles the adaptation driven by error. It allows for the calculation of proper control policies with the anatomical regions possibly facilitating their execution. Consequently, skill learning requires more computations than only adaptation [81].

Motor control is dependent on sensory feedback for motor output corrections. Control functions are thought to share pathways with voluntary movements which means they get impaired after an accident [103].

2.3 Recurrent Needs in the Context of Robotic Rehabilitation

2.3.1 Patients' Requirements

From the patients' perspective, introducing technological devices, such as robots, into therapy diminishes the interaction with the human therapist. As a result, motivation decreases amidst low levels of encouragement and entertainment [63].

As we have explored previously, motor learning processes rely heavily on the patient's active involvement during training. Motivation is an unbridgeable requirement to reach and sustain the levels of involvement. So much so, it is reported to strongly predict the therapy outcome. As a dependent analysis variable, motivation is known to be affected by difficulty levels during training, awareness of subjective performance as well as frequency and quality of feedback [42]. The reader is referred to [84] for a detailed presentation of the challenges and difficulties leading to low transfer of functional ability. These findings invite interest in developing feedback and performance metrics to help sustain and increase the patients' motivation. When queried in [115], more than 60% of patients exhibited high appraisal towards the standardized tests throughout therapy rating it as important or very important. Nevertheless, only a quarter reported that during the care episode, the tests were only conducted at admission and discharge. The patients exhibited good levels of awareness of tests' frequency and their importance during care. This further highlights the recurrent problem of longitudinal follow-up tools.

2.3.2 Practitioners' Requirements

Therapists constitute the golden ring in the rehabilitation provision cycle, and answering their needs should permit them to efficiently and effectively provide care for their patients. Meanwhile, it will help them best exploit the developed technologies.

In acute stroke and clinical practice in general, there is a time restraint precluding lengthy evaluations. Furthermore, it is a major selection criterion for the assessment method [57, 150]. Consequently, the clinical viability of the metrics is tied to their expedition and absence of flooring or ceiling effects [57]. The choice of the assessment method was found to be mainly decided by the patient's condition [150].

The necessity of intensive and prolonged rehabilitation with the possibility to assess human performance objectively through the provision of therapeutic outcomes are among the most recurrent needs [67, 84]. They stress the need for longitudinal assessments believed to hold change-related information in patients' recovery during care episodes. Eventually, this information helps decision-making concerning intervention adaptation [115]. Outcome measurements need to be intended for communication to patients and significant others during neurologic therapy and to share the decision-making of care plan. Additionally, it positively affects the communication between providers and patients. Measures are needed to be generic (e.g., not condition-specific) in addition to having a strong clinical utility: free, takes less than 20 min to administer with no special equipment [115].

Therapists also require the possibility to assess abnormal movements patterns related to individual impairments [14]. The hands-on therapy allows the therapist to feel the challenging movement and witness these patterns on patients' movements, while the introduction of robots obfuscate these details. An assessment technique should build on this requirement as the abnormal movements are at the source of declined function and holds information concerning their origin. Therapists invigorate the fact that tasks, rhythmic or discrete, rely on different neuromechanisms [67].

These assessment tools would help provide crucial feedback for any semi or autonomous training system. The feedback -they believe- is considered an extrinsic motivator [84]. While acknowledging that predetermined tasks' difficulty seems to improve patients' performance, they reprimand not having enough exercise variability to suit individual targets [132]. The need highlights possibilities for adaptation of tasks and training. This also introduces a need for individualized assessment techniques to track these effects closely.

Therapist acceptance of these tools would lead to indirectly satisfy the patient [14]. When surveyed in the same study, 91% of therapists claimed that requiring hands-on therapy is a barrier to using the technologies. All participants rated feedback from the device as important or very important. Specifically, biofeedback on muscle activation (71%) and joint position (54%) is believed to be a useful tool [14]. Therapists essentially require that outcome measures are related to the International Classification of Functioning, Disability, and Health (ICF) activity domain [115] (refer to annex B.1 for a definition of the ICF classification model.).

Figure 2.2 indentifies the different components of the ICF model. Authors in [149] restated the important aspects affecting assessment use in practice:

- Knowledge, education, and perceived value of the outcome measurement (psychometric quality);
- Organizational support or priority for outcome measurement's use;
- Practical feasibility (e.g., time, cost);
- Clinical usefulness and impact on treatment.

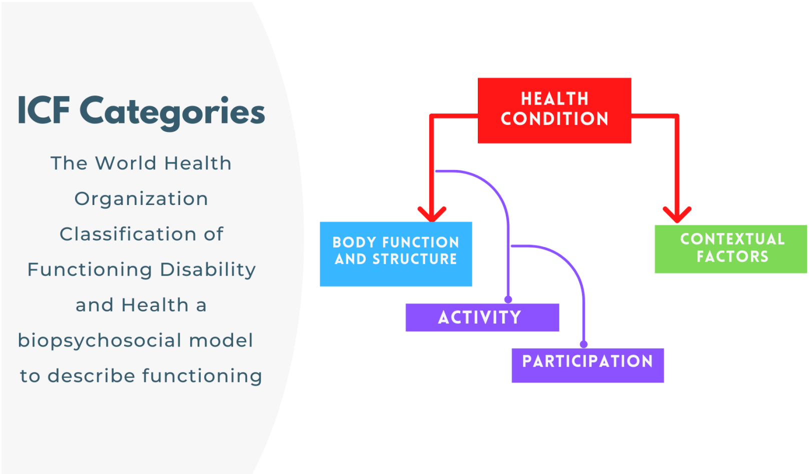


Fig. 2.2.: The ICF categorization.

2.3.3 Stakeholders' Requirements

Organizational perspective is present in the form of numerous guidelines and recommendations governing the design of rehabilitation exercises [24]. These recommendations aim to maximize the efficiency of the treatment while compensating for the shortage of qualified personnel and therapists delivering these services. This can be achieved by opening doors for automation. This shortage becomes noticeable given the growing demographics of the elderly and the motor-impaired patients victims of some pathologies such as stroke. European Robot Road-map in 2010 emphasized the urgency to provide intelligent robotic solutions for healthcare assistance that is capable of automating a large portion of the caregiving procedure [3]. Although these guidelines are recommendations, the use of ubiquitous technologies for rehabilitation training stresses the need to automate the process of assessment, diagnosis, and monitoring. This consists of developing a more inclusive patient-robot-therapist loop, where continuous feedback is present between all elements. Achieving this would enable the full exploitation of these tools while freeing the therapist to care for more patients at once.

From a payer's perspective, health insurers tend to limit or deny rehabilitation for post-stroke patients claiming that recovery plateaus several months after onset and into the chronic phase. This idea of a restricted rehabilitation window after stroke has long vanished. Surfacing neuroscience evidence shows that cortical plasticity is a process spanning through months after the cerebral accident [30].

As healthcare costs continue to grow, the economical concerns stress the need to provide treatment through devices useful at home and clinic in a cost-effective way to reach greater populations and extend the therapy period.

2.4 Rehabilitation Care Tools

2.4.1 Classic Physiotherapy

Physical rehabilitation is an active motor skill relearning starting after the disability provoking accident and aims to enable people with a form of disability to regain and maintain optimal function within their environment [13, 18, 149]. Therapy is a neurodevelopment treatment aiming at the normalization of tone and movements patterns through therapist intervention. This concept is referred to as bobath [13]. When training is used in conjunction with cognitive occupational and pharmaceutical interventions it helps catalyze recovery. Timing the start of the exercise and patient's participation and acceptance are crucial for effectiveness [59].

Therapy activity is inter individually variable and involves different sub-processes. For instance, one of the variables that can result in additional motor recovery is compensation strategies. Compensations manifest as the use of alternative, inefficient movement patterns to achieve the desired move precluding actual recovery [13].

Upper limb functional recovery appears to depend on paresis. It is achieved by 79% of acute stroke patients with mild paresis otherwise, it is limited to 18% if paresis is severe [30]. The Copenhagen Stroke Study reported data showing that recovery in activities of daily living (ADLs) function occurs in most patients within 13 weeks from stroke onset [161]. The ADLs often include a dexterous phase to accomplish the task. This phase relies on interlimb coordination and hand function critical to the upper limb recovery for post stroke [149]. In practice, rehabilitation therapists follow two distinct approaches to fulfill patients' needs, namely [149]:

- Top-down approach: focusing on activities and participation levels according to ICF classifications;
- Bottom-up approach focuses on the body structure and function of the ICF.

Using the top-down approach, the focus of the therapists is on involvement during training centering on ADLs in living environments. This is called the client-centered practice. It involves therapist evaluation of ADLs' relevancy and favors their practice in therapy. By so doing, the therapy aims to alleviate the real-life functional difficulties [149].

In the bottom-up approach, therapists evaluate different skills components and underlying neurological impairments. The approach gets criticized because it separates the assessment from the ADLs. The focus it gives to the impairment level over function is believed to undermine the generalizability into performances in real ADLs [149]. We will revisit this concept of functional transfer when we deal with robotic devices further on. It is important to note that this is relative to the assessment strategy level regardless of the technology used.

Therapy aims to maximize the attention and effort to guarantee motor relearning and retention while also repeating exercises with enough dosage. Therapy monotony limits both processes. Overall, therapy training has positive effects on mood and cognitive function due to the release of BDNF [2]. Adaptive therapy and assistance help to better the recovery results by eliminating monotony and stagnation [105].

Good things are not indefinitely so, therapy can harm the recovery if slacking is present due to excess assistance driving the patient to minimize the effort and attention. Learning is supposed to be error-driven and thus faster improvement might be achieved when error increases [105]. Literature showed that early introduction of exercise training for instance makes it a stressor with detrimental effects on the recovery [59].

2.4.2 Robotically mediated Training

The early devices date back to the 1990s and were called haptic interfaces [132]. In 1988, a double link planar arm provided continuous passive rehabilitation mobilization. Clinical studies on MIT MANUS followed in 1992, advocating it as a device for upper limb post-stroke rehabilitation [56] focusing on shoulder and elbow therapy. Rehabilitation robots were endorsed for use by the American Heart Association in 2012 [56].

With these devices showing promising results, emerging large trials reported higher gains and retention using robotic therapy and proving cortical reorganization [22, 108, 109, 133] which provide ample evidence to establish their effects.

In contrast to other fields of application, rehabilitative robots are intended as instrumented assistive tools for patients to increase their range of motion and boost

[2]Neurotrophic Support, the BDNF neurotrophin is an agent for synaptic plasticity, neurogenesis and neuronal survival post-TBI [59]

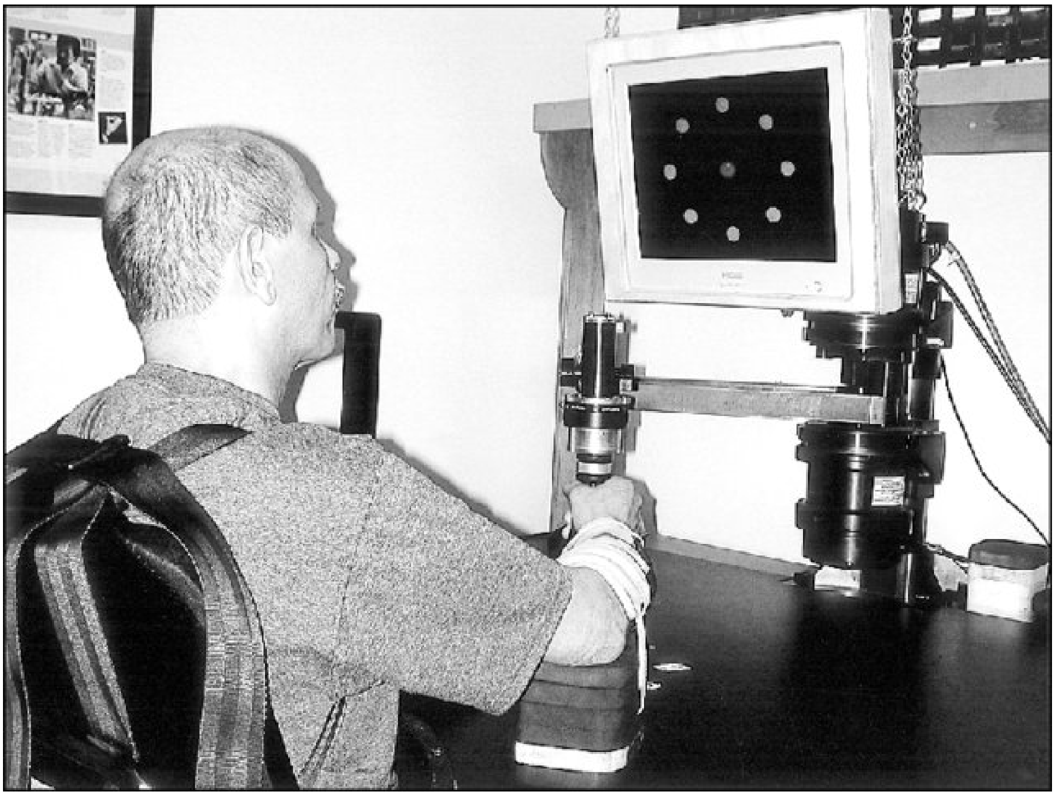


Fig. 2.3.: The MIT-Manus rehabilitation end effector. Copyright Michael Reding, used with permission.

their levels of motivation and interest. It allows therapists to extend their capabilities in delivering therapy and care. Needless to say that the robots are unlikely to gain the skills of a therapist [63, 102] making them far from replacing their roles.

While a great motivation towards developing rehabilitative robots would be saving costs, a tangible benefit is replying to therapist needs by providing an objective sensitive assessment of motor gains. Since the clinical evaluations are cumbersome and often neglected during practice, robot therapy can bridge this lack in the therapy provision cycle [17, 102]. Different robot rehabilitation systems targeting other areas of the body and relying on different therapy modalities soon emerged in the 2000s [63]. For instance, devices targeting the wrist, hand, and fingers as well as ADLs. Other systems targeting the lower limb, gait, and ankle training also appeared. Mostly, rehabilitation robotics focuses on neurological rehabilitation [63]. The main difference between neurological and orthopedic rehabilitation is the clinical goal setting of the therapy. Patients suffering a form of Central Nervous System injury usually focus on achieving cortical reorganization restoring their motor function. Whereas orthopedic cases usually suffer musculoskeletal lesions without cognitive degradation. Hence, the goal for the therapy becomes the restoration of the functional range of motion, muscular strength recovery, pain elimination [123].

Robotic neurorehabilitation favors limb mobilization through complex multi-joint movements, favoring neuroplasticity through stimulation with cognitive components of the tasks. In orthopedic rehabilitation, the focus is rather on incremental single joint mobilization [123].

Rehabilitation robotics techniques are based on motor learning where the standard stated is use it and improve it, or lose it which is key to the success of rehabilitation [39, 85].

The devices can be classified into two categories namely: Assistive robots and therapeutic robots according to [51, 105]. Therapeutic devices are intended to supply rehabilitation therapy training while assistive devices help compensate for lost skills. Our interest revolved around rehabilitation robotics providing therapy training. Different rehabilitation robotic systems were proposed for clinical or at-home remote settings [40], 120 systems are listed in [105] which classifies them per application as support for ADLs or therapy and physical training devices.

These systems are bundled with serious games in virtual reality environments that are rewarding and provide high engagement [39]. These rehabilitation robots are proving to be highly beneficial rehabilitation tools [13, 6, 42, 81, 93] as they are capable of delivering high-intensity training both in sub-acute and chronic stroke [40]. This intensity can facilitate motor skill relearning and retention in the long term. The training intensity is linked to significant cortical improvement in the disabled population after a year with an additional reduction of pain at shoulder level [134]. Particularly, the continuous challenging and assistance boost the motor learning and coordination during robot therapy [42].

In terms of assistance provided by the device, the authors [105] distinguish the active, passive, haptic, and coaching devices. In terms of exercise, they span two categories: active and passive modalities. Active training involves robot-assisted movements while passive exercise relies on the patient's carrying out the movement without force assistance.

The Transfer Dilemma The fundamental question that faced robotic rehabilitation concerned whether patients adapt or acquire motor control of the upper limb. Investigation of the evidence suggests beneficial effects of robotic therapy on the upper limb impairments and an increase in motor control in form of increased speed and more natural activation synergies. On the other hand, the transfer of the performance gains into the much-needed ADLs lags reportedly [13, 81, 105].

Researchers suggest a potential argument that the patient has learned to associate the changes related to training with the robot rather than their body. The literature treating the subject lacks definitive evidence in comparisons between robot control schemes, impairment levels, differences between the three and two-dimensional training, all of which are factors that might influence the task acquisition during training.

When investigated further the patients seemed to attribute the errors in executions to themselves and modify their movements execution. Suggestions that generalization might be the source of the transfer issue show that similar tasks differed by direction. These results suggest the necessity to train across wider workspace in different directions to ensure full skill learning [81]. Subsequently, the multi-planarity of the exercises seems to induce motor cortex excitation [132].

Another potential issue is that transfer is related to the integration of both distal and proximal arm members when most early trials focused on proximal training. In addition to these effects, the large disparity in robotic platforms and protocols used in early clinical trials might be a potent contributor [13].

An important aspect of the functional environment is the fact that ADLs are executed solely in a natural setting, whereas most therapy mediums resorted to some form of gravity compensation [152]. The fact that motor skill training was in a gravity-free environment might also contribute to the generalizability of the motor control strategy and thus to the functional transfer of the learned skills.

Many of the robotic devices presented assistive forces to help patients during the execution of the intended movements and to provide corrections while training. This mode of therapy is called passive training and might encourage slacking, which is detrimental to motor learning and recovery. This is particularly true when we recognize motor learning literature does not consider passive movement as actions since they lack intention [119].

The active training modalities were shown to be more effective as they require patients' movement initiation. Resistive training is challenging movements progressively using active robots and was found to increase the function as well [13]. In conclusion, the studies favor active-assisted rather than passive training. They also favor the graduation in amount and typology of the feedback provided to the patient as a function of effort and participation.

Structural Issues From the shoulder to the wrist joint, the human arm is approximated by five degrees of freedom: shoulder horizontal abduction/adduction, shoulder elevation, shoulder internal/external rotation, elbow flexion/extension, forearm pronation/supination [85]. For better training and usability, the device does not have to interfere or impede the freedom of any joint.

A review of the robotic devices for upper limb rehabilitation suggests that the robotic end effectors limit the control of joint level torque and can only produce combinations of elbow and shoulder joints instead of targeting specific joints. Exoskeleton structures are also presenting a challenge with singularity configurations that eliminate a single degree of freedom in the structure [99]. In exploring the different control approaches for the robotic devices, the research suggested that at the joint level, torque is mainly compensating for gravitational effects compared to centrifugal and inertial loads [99].

Actuated robots use electric motors with stiff gearboxes with high torque output that might be concerning for patients with spasticity. Wearable robots highlighted this issue and introduced the actuators impedance control to solve this issue while adding increased complexity and costs [152]. This affects patients' perception of the device as they are intimidating and constraining if we consider the variability in the human limb dimensions and properties [5]. The fact remains that the existing robotic systems may only provide training towards some degrees of freedom and not the full range of freedom [132].

Assessment The increase of knowledge diffusion and the advent of technology help bridge the gap between the providers and the rehabilitation professionals [132], thus permitting us to train and inform the users about the potential benefits of these new rehabilitative technologies. The recurrent issues against the classical assessment and scoring methods are that they are subjective, qualitative and, time consuming [51].

Rehabilitation robotics are often geared towards providing supplementary assessments and intelligence enabling tools for the therapist instead of replacing their roles [63]. Robotics also provide the possibility to assess movements in a repeatable objective way which is critical to the effectiveness of therapy and for promoting evidence-based decision making [39].

The therapist in practice can adjust assistance, resistance, as well as the different assessment and therapy tasks that are to be carried out in a robot-mediated session. This permits the adaptation between the user and the robotic therapy which is an important aspect of the robot rehabilitation compared to classic therapy [63]. A critical misconception is that robots are not dehumanizing the therapy, since the device is unlikely to replace the multilevel interaction of the therapist, nor his experience and dexterity in manual operations during exercise [132].

Examplary Rehabilitation Robots Authors [47] propose a robotic platform designed to provide both kinetic and kinematic measurements of the upper limb during the autonomous training session. The hand movement was parameterized and visualized using measurements collected through the platform.

The IntelliArm platform [139] is a robotic platform (see Fig. 2.4) designed with the full spectrum of rehabilitation processes in mind. The robot provides the ability to run pre-evaluation and diagnosis. Besides, it provides strenuous and safe stretching passively, assistive active reaching training, and quantitative outcome evaluation at the joint level. The robot provides the full measurements of kinetic and kinematic variables of the movements.

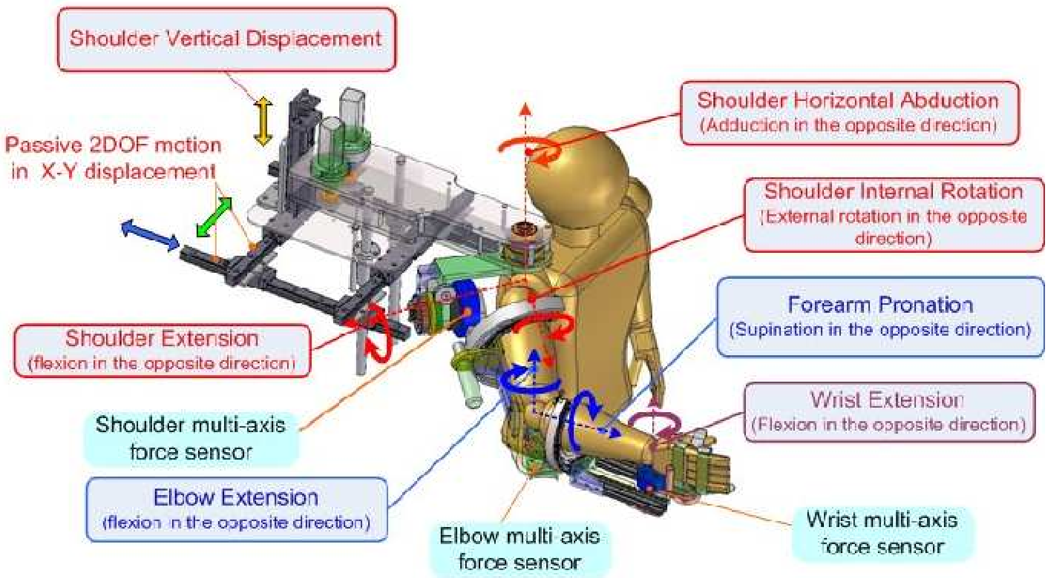


Fig. 2.4.: The IntelliARM exoskeleton [139].

The platform focuses on the ability to track abnormal joint level couplings during movements and relate them to the patient's level of motor control. The platform offers the ability to quantify muscle properties and active control ability that is at the source of the impairments. These functionalities are proposed through the same platform and thus obviate the need to transfer patients between devices.

The ARMIN I device (see Fig. 2.5) was tested for its feasibility on three chronic stroke cases and showed sustained FMA performance after 8 weeks follow up [118]. A next-generation ARMIN III [70] was introduced as an ADL supporting exoskeleton. The device is equipped with a virtual environment to support task motivation and initiation. The control strategy is shown to effectively account for the variability of the human movements around the desired task trajectory.

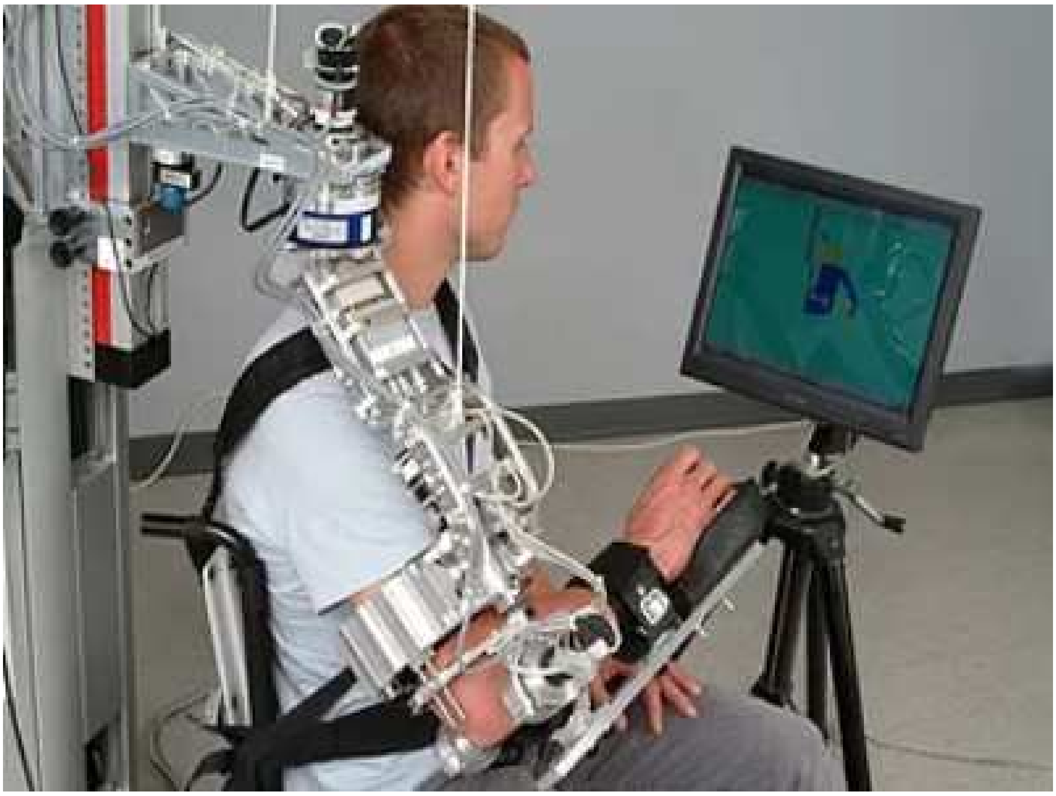


Fig. 2.5.: The ARMin I exoskeleton [114].

In [5] a low-cost cable-driven robot prototype was proposed (see Fig. 2.6). The task assessment is based on prior knowledge of the path trajectory of the circular task. The robot presents bigger workspaces, transportability, flexibility, and configurability. Additionally, it has low inertia compared to bulky exoskeletons. The task evaluation is done through circular trajectory modeling. The model is dependent on the load of the robot and the length of the trajectory. The movement period is used to assess the patient's performance (it should have at worst close value to the predicted period). The solution and the actual performance are presented visually to assess the session's performance by plotting the parameters color-coded plane and visualizing the optimization solution and the actual performances.

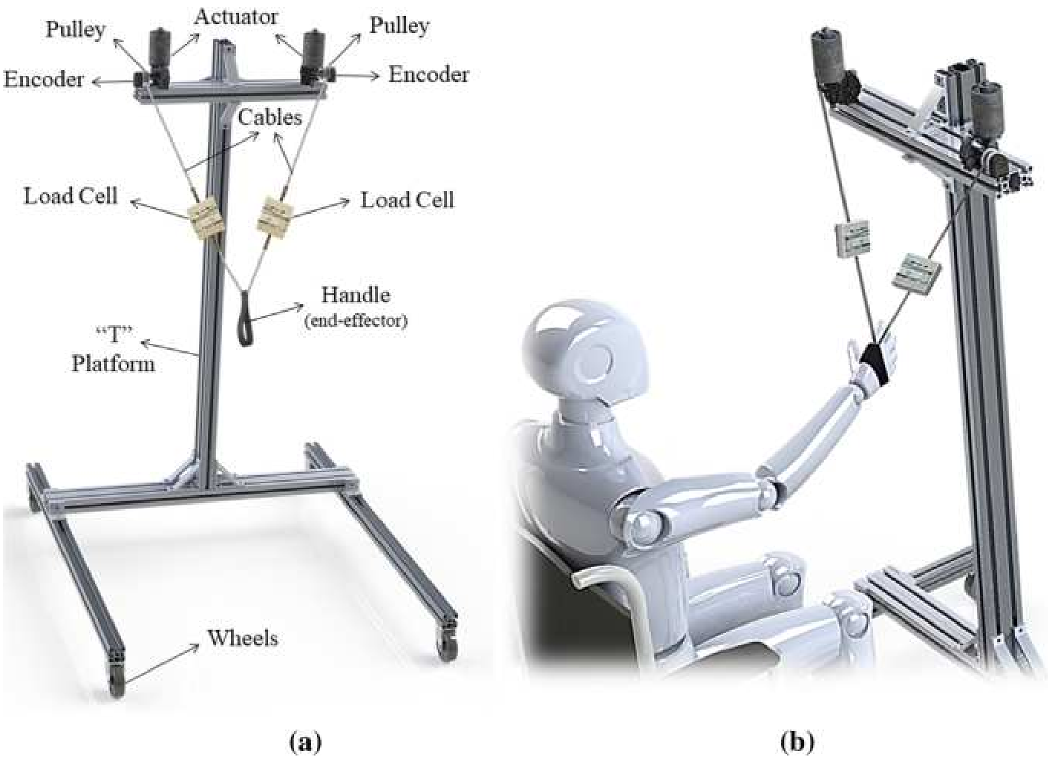


Fig. 2.6.: The Cable Driven Rehabilitative Robot [5].

2.4.3 Telerehabilitation

Telerehabilitation is a parallel development field interested in developing remote therapy or teleoperated systems to help deliver therapy close to the patients. For the case of teleoperated systems [14], the therapist interacts haptically through an arm robot controlling another system over distance. The biofeedback from the information collected by the patient's robot helps the therapist conduct therapy remotely. The authors investigated the modality of feedback presented to the users.

They also implement haptic rectification of the patient's impaired movements to the therapist's arm by programming the devices to mimic each other. The therapist can thus guide the patient's hand through the right movement and provide conscious assistance to the patient. Technologically, this functionality can be achieved using disturbance observer [51]. A different aspect of telerehabilitation is the use of social robots [76] where ZORA, a NAO-based rehabilitation companion provides instructions, motivation, therapist guidance, and rewards (see Fig. 2.7).

The robot was conceived as a friendly buddy, as much as an intermediate between the therapist and the patients. It also provides demonstrations of the exercises and

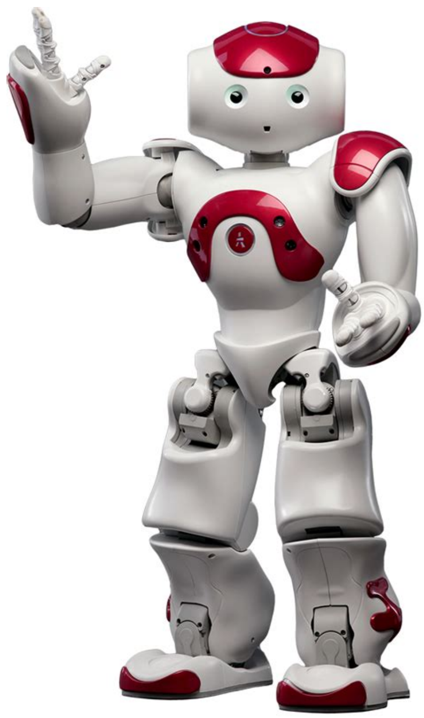


Fig. 2.7.: The ZORA telerehabilitation system. Courtesy of zorarobotics.

supports psychologically the patient during training. Assessment of its effects on the session was carried out by professionals who found it to be a great motivator for the patients.

2.5 Physical Recovery Assessment

According to [115] rehabilitation assessments come at the intersection between innovative research, stakeholders' requirements, needs, and expectations. In the healthcare community, CPGs or Clinical Practice Guidelines are evidence-based recommendations guiding clinical decisions. They also help identify gaps in the literature and potent research avenues . These CPGs often suggest new assessment techniques and appropriate care associated with the different diagnoses.

Steering communities publish CPGs as part of the process of knowledge translation (KT) at the heart of both clinical research and practice. The knowledge translation is the process of systematic synthesis and exchange of knowledge and ethically applying it to better health services and products and to strengthen the healthcare system.

For instance, the components of the physical therapy examinations are according to the Guide to Physical Therapist Practice: history, systems review, tests, and measures. Of interest to our research, the highlight on the outcome measure (OM) refers to standardized measures used to monitor a defined construct during the care episode. Terms such as standardized assessments, instruments, and tools often refer to OM [115]. We use these terms interchangeably in our work.

These standardized tests have a preset of properties such as reliability, validity and are uniform and consistent. The difference between OMs and a standardized test is that the latter is relevant to an intervention [150]. Using standardized measures is recommended, and their selection is decided based on their psychometric properties and clinical utility. They are used to diagnose, discriminate, assess change over time and predict. Their use throughout the care episode is considered good clinical practice [115]. In the case of TBI, the measures are used to quantify severity, monitor spontaneous recovery, and establish baseline function scoring [59].

Typically these assessments target the following areas: mobility, shower/bath assessment, home visit assessment, dressing assessment, upper limb assessment, cognitive assessment, leisure interests and activities, perceptual assessment, ADLs assessment, cooking assessment, work visit, driving assessment [150]. Furthermore, the therapist reported they assess their care through goal achievement, self-reported function, patient satisfaction, clinical intuition, and standardized measures. Remarkably, these tools to a large extent remain highly subjective measures [150]. Moreover, when querying about the goal achievement method, researchers found that only a few reported having a goal-setting session at the start of care. This is a stark contrast to the proportion of therapists using it as an assessment tool in the first place [150]. From the therapist's perspective, standardized measures are the least used tool. Meanwhile, the majority of therapists felt their incorporation was important [150].

2.5.1 Why Do We Assess Patients?

The functional assessment stands as a fundamental component of the activity and functional rehabilitation therapy [63]. The focus of the clinical scales is on providing an understanding of the source of the sensorimotor problem while tracking the evolution of the recovery over time and predicting the therapy outcomes. The three functions are diagnosis, monitoring, and prediction [149]. The clinical assessment permits comparing patients and organizational outcomes, gauging the effectiveness of the therapy, and helps expand our knowledge about the phenomena [115].

2.5.2 Benefits of Using Clinical Scales & Outcome Measures

Using the Outcome Measures OMs in practice is believed to increase the efficacy of care delivery among many other benefits [115, 150]. From the patient's perspective, OMs help monitor changes in patient's health over time and enable quantification of these observations and the scored function. Achieving detailed examinations enables the care facility to provide more patient-centered care and tailored therapy planning. It enables the care providers to identify at-risk patients that necessitate prioritized care.

Therapists can move towards evidence-based therapy practice based on OMs, with the added benefit of easing the communication between the care stakeholders. Standardized OMs enable comparisons of outcomes associated with different interventions, providers, and patients or diagnostic groups. This allows for the curation of a larger dataset that can be used to analyze differences and effects. This data enables researchers to answer clinical questions of relevance. They can also serve as a care facility performance metric by comparing outcomes amongst clinicians and settings. For the payers, the inclusion of OMs help measure healthcare costs and enhance the reimbursement plans offered accordingly.

2.5.3 FMA The Standard Clinical Scale

The utility of the clinical assessment is in its effectiveness, efficiency, comparability, and predictability [128]. The Fugl Meyer Assessment (FMA) is the most widely reported quantitative measure of motor impairment for stroke with proven validity and reliability including resistance to compensatory effects [81, 100, 149]. Brain studies have provided evidence of the correlation between functional scales and brain lesions [16].

The scale designates five domains of assessment: motor function, sensory function, balance, joint range of motion, and pain. A trained rater spends 30 minutes assessing FMA items using a three-point Likert scale. A subscale termed FMA-UE is of relevance to our research. It is dedicated to the upper extremity ranging from 0 to 36 [122].

The increased score on the FMA scale can be interpreted as increased force output and range of motion and is not necessarily valuable for the real-life ADLs [81]. The scale does not deal with the functional activities domain of the ICF and is considered as body function domain scale [100]. This is particularly noteworthy, as a large proportion of the studies dealing with robotic therapy are using it. This limits the ability to quantify transfer through cross-study meta-analysis.

Clinical Scale Limitations & Challenges The clinical scale although established in practice as a valid tool has its limitations. The scale faces several challenges specifically when we consider the new technological advancements introduced into the field of rehabilitation.

A major issue for the clinical scale is the sensitivity to the subtle change in motor performance during rehabilitation therapy [100]. Iteratively this issue can be related to the composition of the scale. The FMA reflects strength, compensatory movements, and motor control aspects at once. Thus, an increase in the motor control ability as seen through the performance measures might not be visible on the score [89].

The patients present large inter-subject variability in their capacity to recover [81], thus detecting enough recovery to judge changes as significant might not be accessible using cross-population statistics. On the other hand, demographic variables served to explain 89% of the FMA variability when excluding severely impaired patients from the regression [81]. These remarks tie back to what the clinicians define as the minimal clinically important difference (MCID) [161]: " the smallest difference in the score of a clinical outcome perceived by patients as beneficial and which would mandate, in the absence of troublesome side effects and excessive costs, a change in the patient's management".

These binding theories introduce an interesting trade-off between what is deemed clinically useful and what is a significant increase in motor learning and control. Specifically, when we consider the effects of inter-individual subtleties on these scales. Moreover, with the advent of measurement technologies and instrumentations, moving away from the subjectivity of the current standards towards more objective quantifications of effects appears tangible. In broader terms, when we consider limitations in design and methods, scholars often reported challenges summarized as follows:

- Routine assessment during therapist practice is very limited [149];
- Subjectivity of the scale due to idiosyncratic rater qualities [63];
- Focuses on behavior and physical or visual inspection [103];

- Limit the precision and reliability of the assessment [63];
- Coarse measures ensuring reliability while hindering functional level changes [120, 164];
- Rate criteria instead of a normal standard [164];
- Present high floor and ceiling effects [103];
- Undermine the effectiveness of the therapy by obscuring effects [128];
- Cannot provide assessment during therapy [128];
- FMA assess the motor function instead of activity [120];
- Absence of strategy of relaying mechanism level to motor control impairment level [122];
- The composite score as a standardized test was controversial and is relaying often on the outdated theory of general motor ability [119];
- According to the ICF the OMs can be classified into one of the categories it enlists. The OMs aim to enhance the quality of care as well as standardize practice by eliminating variations. The latter aim seems unlikely to have been reached [115]. This suggests an inherent problem with the current standardized OMs.

When questioning the practitioners, the reasons reported for not using clinical assessment measures are [115, 150]:

- Time consuming for a therapist with a busy setting, large caseload, and long waiting list;
- The need is for standardized assessments that are expedient and easy to administer;
- Absence of financial compensation;
- Equipment availability;
- Not specific enough to measure therapy outcome;
- Not holistic enough or not patient-centered enough;
- Not sensitive enough to measure the change during intervention;
- The perception of tests as an additional burden on patients;

- Lack of knowledge of the assessment for their practice;
- Lack of knowledge of administration and scoring of these measures;
- Large proportion reported not having any training related to the measures, training therapists on the use of the outcome measures, for instance, was found to increase their usage;
- The caseloads were formed of mixed diagnostic groupings within daily caseloads. Mostly for the interviewed therapists, the largest proportion of patients were older people[3];
- The assessments are not usually suitable for other diagnostic groups.

2.5.4 Kinematic Assessment Scales

The fact that the standard in clinical assessment is still subjective highlights the importance of objective measurements as a tool to enhance recognition of patients' improvement with finer granularity [120]. Studies started to emerge stressing the need to improve assessment methods for rehabilitation [164]. The robot therapy mainly consists of goal-directed planar reaching tasks, which are weight-supported. Most kinematic parameters describe temporal and spatial properties of endpoint or joint trajectories (e.g. position, velocity, and movement time), classified into the Body Functions and Structure ICF domain [161].

In the literature, the most frequently reported kinematic measures are: movement speed (37 studies), movement accuracy (31), number of peak speed (23), movement distance (25), movement duration (11), and peak speed (13) [161].

Many reviews focused on the kinematic measures obtained by robots during training to provide objective assessment and outcome measures [140, 87, 106, 120, 147, 161].

Kinematic Scale Benefits The new robotic devices provide scale with many interesting benefits for the clinical use such as:

- Assessed based on behavioral tasks [103];
- Quantify sensory, motor and cognitive abilities [103];

[3]The diagnostic groups being: stroke, arthritis, orthopedic conditions, progressive neurological conditions, dementia, falls, cancer, brain injury, spinal injury, respiratory conditions, hand injuries

- Inter rater reliability that ranges from good to excellent [103];
- Have higher resolution [128];
- Provide objective continuous assessment scales [122];
- Ceiling and floor are less common compared to ordinal scale [122];
- Can quantify physical properties kinematic or dynamic [128];
- Distinguish the effects of the therapy parameters on motor learning [128, 33, 164];
- Quantify proprioceptive, postural control,coordination, visuomotor reactions and reaching performance [164];
- Kinetic information permits evaluation of force, muscle tone and paresis [120];
- Kinematically assess the range and coordination of UL at joint level [120];
- It permits to discriminate the compensatory movements from the recovery [120];
- Reach kinematic analysis is a sensitive scale [30];
- Permit analyzing UL interjoint coordination and control strategy [30];
- Assessment measures with improved objectivity, repeatability, precision and ease of application [161];
- Detect changes for ipsilesional limb compared to contralesional [164].

The Assessment Tasks Used in Conjunction to Kinematic Scales To calculate these kinematic measures, researchers have devised several standardized tasks to quantify specific functions and aptitudes. Authors [120] provided a taxonomy of the assessment task and the movement aspects they evaluate by reviewing the literature, they report:

Point to point reaching: In the transverse plane, the task evaluates the feedforward, and feedback control, perturbation compensation relating to temporal efficacy, ease, smoothness, planning, and movement efficiency. In the sagittal plane, it evaluates the range of motion, feed-forward feedback control, and gravity compensation relating to planning, temporal efficiency, and smoothness.

Free reaching: In the sagittal frontal plane, it evaluates the range of motion, perturbation compensation, feed-forward, and feedback control, gravity compensation relating to planning, temporal efficiency, range, smoothness, and movement efficiency.

Activities of daily living: In the form of virtual games, they evaluate functional ability in all planes relating to the range, interlimb coordination, and temporal efficiency.

The Validity of the Kinematic Scale Clinical acceptance of the new technology-based assessments is contingent upon the provision of consistent data with valid outcome measures such as the FMA scale [4] .

In addition, the evidence from anatomical measures using brain imaging techniques like the Cortico Spinal Tract (CST) integrity will help establish these scales [108]. Robotic measures are shown to correlate with clinical scales assessing adults' and children's impairments [16].

Many kinematic parameters showed a significant correlation value with the FMA, for instance, the movement accuracy (r=0.674), movement speed (r=0.53), and the jerk metric (r=-.578). Authors listed other studies' results reporting significant correlations of movement speed with the FMA scale (r=0.53), and with the Box and Block Test related to dexterity (r=-0.44), the jerk metric with the FMA (r=-0.578). The movement time correlated with the FMA (between 0.4 to 0.65) [161]. Ultimately, the correlation of findings from clinical scales, robotic measures, and anatomical measures may provide a deeper understanding of the motor deficits causes and the recovery mechanisms [108].

Kinematics and Sequelae Kinematics have been shown to quantify spasticity after robot therapy and Botulin Toxin injection [133]. Peak velocity, final angle, creep (or release) angles, between arm peak velocity difference are kinematic covariants of spasticity [27].

Proprioception during neurorehabilitation was identified from bimanual task's kinematics post-stroke when compared to controls [48]. During these tasks, patients try to recreate reference positions in the absence of vision. Proprioceptive acuity is estimated by matching errors.

[4] Often termed the golden standard in clinical assessment.

The kinematics also assess joint control and movements [57] including their relation to the coordination and motor aptitudes of post-stroke patients.

Kinematics Classification The robotic measures are commonly classified into different taxonomies: kinematic, kinetic, and neuromechanical measures. Another classification takes into account motor control theory and stratifies measures into: movement initiation, feedback control, corrective responses, and total movement metrics [122].

Authors in [120] provided detailed taxonomy. We synthesize the major elements and those of direct impact on our study in the following:

Movement Planning Comprises measures that relate to the feed-forward sensorimotor control of movement measured through the changes in directions, time to initiation, and the starting speed. These measures correlated to FMA and reflect the improved synergies, volition, feed-forward, and somatosensory loss.

The speed profile indicates details about the control strategy. For example, peak velocity at 33% to 50% of the velocity profile, with a left shift in the peak speed is indicative of the visually guided reaching movement [30].

Temporal Efficiency Comprises measures that relate to the time required to complete the task. This measure is found to be lower for movements in the center or ipsilateral compared to contralateral sides. This is believed to differ between patients with different affected areas of the brain and healthy controls. It reflects both losses of somatosensation, feedback, and paresis and is associated with the loss of fractioned movement.

Movement Accuracy Comprises measures that relate to the distance between the target and the prehension organs. The loss of accuracy can be related to the reduction of somatosensation [100], feedback and paresis.

Movement Efficacy Relates to the redundancy of the possible configurations that permit the access of the same position in space. Path length for chronic patients improved significantly while also significant impairments are noticed for those with right hemiparesis.

The measure correlated with the module of the gravity compensation. The normalized trajectory only captures effects for sub-acute patients and does not appear to correlate with FMA. It reflects losses of somatosensation feedback and paresis.

Intra Limb Coordination It relates to the coupling between the elbow and shoulder using the joint angle correlations for the paretic side. Decoupling happens at discharge and consists of restoring normal synergy and reducing the excessive elbow elevation. These reflect measures of feed-forward, somatosensory loss and are associated with the loss of fractioned movement.

The Range of Motion Comprises measures that relate the extent of the reach during movements. It is known to be influenced by gravitational compensation compared to the planar or transverse plane. When bimanual tasks are assessed the effect favors the planar tasks. This suggests the gravitational effect on unimanual tasks. These measures are associated with jagged movements due to abnormal muscle tone.

Ease of Movement Comprises measures that relate the ability to execute the movement effortlessly measured as mean speed and peak speed. Higher mean speed reflects decreased abnormal effort. The normalized measure showed improvements for all subjects across populations. The measure can distinguish the chronic and sub-acute as well as directional and rotational differences in movements.

The decrease of velocity can be due to paresis [100]. Peak velocity is linearly related to the force that is generated during the movement process [30] which makes it a proxy to deduce kinetic information from simple kinematics. This is particularly important as we seek to minimize the costs of the devices and increase their utility, potentially through computational methods.

Movement Smoothness Comprises measures that relate to the movement trajectory profile with a bell-shaped velocity curve. These trajectories minimize the jerk measured as the third derivative of the movement. Arrest time considered at 10% of the mean time, normalized jerk measure, and speed ratios were also significant in detecting smoothness. The jerkiness can be related to abnormal muscle tone [100]. Findings suggest that movement smoothness is a result of learned inter-joint coordination and it increases with motor recovery [30].

Kinematic Assessment While not exhaustive in nature, this section lists exemplary kinematic assessment studies noting the different conclusions the researchers were able to make based on the measured metrics. The main aim is to illustrate the theoretical aspects and the reported benefits through applied research results. Reaching task was assessed using, time, peak velocity, percentage of time to peak velocity and normalized jerk were all significantly better at the end of training but not sustained at retention test [30]. Within the subject, effects were significant for all the metrics.

In [107] authors aimed to identify performance changes for subacute and chronic patients using the InMotion device (see Fig. 2.8). The task used was horizontal reach with an assist as needed controller. Measures reported were mean velocity, number of speed peaks, and speed and acceleration ratios following different directions. Clinical FMA and defined measures showed decreased impairments after treatment for both groups with higher gains achieved by subacute subjects specifically in movement quality. Effects were noticed on both groups after the administration of 20 sessions.

Using different tasks featuring horizontal reach the authors [67] defined metrics as the ten repetitions sum and called it a variation coefficient. Measures were assessed for learning effects using 10 consecutive trials with 15 stroke patients without finding effects. The task was determinant of the suitability of the measures. The measures showed reliability for test-retest evaluation.

The kinARM robot (see Fig. 2.9) was used with chronic patients performing the reaching task in a study of clinical validity and usefulness for the metrics [122]. Twelve measures were calculated and tested for correlation with the FMA-UE with ten of them showing correlations. The lateral differences between the paretic and non-paretic sides were also performed with the paretic side significantly different. The unassisted task revealed that feedforward control is impaired for the hemiparetic arm, with errors in reaching the target and weak peak velocity during the initial phases.

A gravity compensated reach in eight directions was used to assess the mean velocity, the mean deviation from the straight line, and the FMA UE score. Authors [56] found that velocity was recovered before the accuracy which supposedly comes in line with the effects of a speed-accuracy trade-off mentioned in motor learning. Exploring the relationship between motor control deficits and the Kinematic measures, in [100] authors retrospectively assessed 44 stroke patients using kinematic measurements. Comparing measures from patients to the healthy controls, permitted the detection of significant differences in between-group statistics for accuracy,

Fig. 2.8.: The InMotion End effector rehabilitation robot. Courtesy of Bioniklabs.

velocity, and smoothness for patients within the same clinical score. In another study [38], researchers followed 3 weeks of conventional therapy augmented by the use of different robots for therapy. Robot training consisted of horizontal plane reach and was assessed using FMA, Modified Ashworth Scale, and kinematic measures. The results showed that there were no differences in between-group effects for the different robots used. Measure's responsiveness effects were Large for the FMA score and the speed. The study concluded that gains achieved during therapy programs were not dependent on the type of robot.

The Armeo Spring device was used to assess precision velocity and smoothness using kinematic measures. The results suggested that motor outcome evolution happened at different temporal scales highlighted differences between groups with ABI and Cerebral Palsy [16]. For the sake of our own research a more detailed

Fig. 2.9.: The KINARM rehabilitation exoskeleton. From Wikimedia.

presentation of the Armeo device was provided in 1.3. Robot-based speed measures correlated best with FMA. The construct validity of the measures was established for all the robot measurements. The FMA is a timed test therefore, the speed is likely to play a role [108].

2.5.5 Basic Properties and Assumptions For the Kinematic Scale

The kinematic scale is built assuming a set of assumptions and properties that consisted of the motivating hypotheses. We present these fundamentals and relate their effects on the scale and any potential challenges associated with them.

Gravity Compensation Gravity compensation is omnipresent for almost every single rehabilitative robot device. Understanding the implications and the effects of this design is critical to understanding the weaknesses and strengths of the current scale.

The colloquial terminology of strength is in physics the interaction between mass, force, and moments. Thus, the physical composition of the body has a great influence on these forces. Human movements are built through the fusion of retractions, extensions, or isometric stabilizing forces and moments [160]. A major force applied to the upper limb during activities of daily living is the gravitational moment which is considered a source of movement errors. The moment is often modeled as a trigonometric function that either substituted for or substracted if against the movement [160]. This definition is basic, and separation of the gravitation moment effects is challenging in reality due to rotational moves and overlapping signals which makes them inseparable [74].

In a related study [74], authors went on to suggest removing the gravitational component for accelerations. These components are often used as assessment measures in ADLs. They concluded that the choice of metric explains the different degrees of variance in daily physical activity. Gravity compensations ease the difficulty in normal reaching, extend the distance and increase the speed [161]. Gravity balancing robots and impedance assistance are useful also in decreasing the jagged movements of stroke victims [120] since the compensation helps with movement ease, permitting the patient to continuously exert effort and complete the task. This includes his ability to reduce the interaction torques and maintain normal coactivation of antagonist's muscle during reach [120].

A key recommendation for defying gravity influence is to train with targets in different levels instead of planar tasks [120]. Relatably, gravity compensation effects on a 3D reach task were investigated in [136], and EMG readings suggested reduced muscle activity of -25% to -50% with similar activations for either compensated or free trials. The kinematic measures also remained similar for the different trials. They concluded that the gravity compensation acts as a separate intervention stimulating the function by affecting the muscle activity. Motor control during the facilitated task relies on the intact capacities to generate and coordinate movements with higher intensity [136].

Smoothness is considered to be influenced by this compensation. This is due to the effects of compensation on speed peaks in the velocity profile which is used to quantify this property [120]. While compensation for these moments can be beneficial, research shows [85] that this can lead to over or under-compensation of the arm weight. The characteristics of the spring used in the passive exoskeleton contribute to disturbances that might affect the movements. Interestingly, the adaptive reduction of the compensation leads to increased active ROM and permits task execution within an increased workspace.

Properties of Impaired and Healthy Trajectories As we have seen the reaching task is the fundamental component in many ADLs and is used as the basic assessment task in the majority of the robotic solutions available today. The move requires inter-joint coordination and is extensively studied [120]. In attempting to define the desired trajectories during rehabilitation therapy, the early proposition dates back to the 1980s where movements were defined in visual space as a stereotypically straight path with smooth velocities [39, 81, 117]. Seminal research by Hogan and Flash [62] found that smoothness can be mathematically parameterized as minimization of the third time derivative of the position knowing the length and the duration in what is afterward known as the minimum jerk trajectory. Meanwhile, the ADL trajectories presented a major challenge to this model as they are far more complex and were shown to consistently differ from this model [81, 117]. The model is considered when no accuracy constraint is applied since complex and precise reaching movements present asymmetric multi-peaked velocity profiles [82, 117]. Researchers are considering questions as to whether assuming an invariant optimal trajectory would better affect the recovery compared to invariant cost function [81].

The sub-movement theory supposes that the reaching movement is composed of discrete units that are combined by the motor controller of the CNS into complex movements [82]. From a motor control standpoint, the claim is reasonably paralleled by the multiple sub motor and neuronal units that collaborate to produce the desired motion through an unknown cost function. These sub-movements are supposed to follow the Minimum Jerk Model differing in length, amplitude, and initiation time. The motor problem defined earlier is expressed as finding a combination of different overlapping sub-movements that achieve the desired task [117]. Parameters of the sub-movements, namely duration and rate are shown to be intrinsic to patients and independent of the pathological level [120]. According to the kinematic measures, the optimization of the Minimum Jerk Model strategy can be noticed on the velocity profile where time shift of peak speed towards the middle of the motion is observed in chronic patients introducing the famous bell-shaped profile [120].

By inspecting the normalized speed profile, a consistent pattern emerged characterized as kinematic properties of blended segments. Researchers in [141] simulated the sub movement by blending two-component curves and compared the kinematics of the resulting trajectories. Thus, smoother movements are believed to be the result of the blending of sub-movements and their decrease over therapy which can be noticed on the normalized frequency spectrum of velocity signal [120].

For instance, the 'quiet standing' protocol representing the control of posture was criticized as an invalid generic solution [119]. According to [77], the movement primitives at the basis of human motions are sub-movements, oscillations, and mechanical impedance combined either sequentially or in parallel for movement and forces. This is believed to be the strategy to approach learning, performance, and retention of complex skills. The focus here is on the biomechanical aspects of the movement rather than goal-driven functional tasks [119].

2.5.6 Kinematic Scale Limitations and Challenges

Considering the evidence and the literature presented above, the kinematic scales offer a great advantage compared to the standard clinical assessment scale. First off, the objectivity of the measures permitted the researchers to draw many conclusions and quantify the effects between different mechanisms driving the rehabilitation and the observed patient performance.

Nevertheless, the kinematic scale has its limitations that we have slightly referenced in previous sections. In this paragraph, we try to detail these aspects exhaustively.

Challenges related to rehabilitation devices The usability and accessibility of any tool designed for rehabilitation care is a real challenge [84]. The kinematic measures are, to an extent, hardly dependent on the device and the environment which makes it hard for them to substitute the classical scales [132]. This effect is further stressed as robotic platforms come in a wide range of structures, assessment tasks, sensors, and control approaches [13, 38, 68, 99].

For the most part, exoskeleton devices require a transformation enabling the parameterization of the assessment in the anatomical reference frame [120]. This transformation has to factor in the effects of the constraints and fastening between the upper limb and the device on the biomechanical joints during slow movement in rehabilitation training [44]. It is important to ensure a constant coupling between the robot and its user during training [120] which are hard to ascertain during real care practice.

Another concurrent challenge is the need to account for the effect of the intrinsic dynamics of the device. Studies have reported these constraints in the form of planar movement surfaces [120] and the proposed solution uses the transparent mode to rely on the controller to dissociate the patient from the robot during the assessment.

The most prevalent device in rehabilitation delivery is the end effector structure. These devices are known, for instance, to not allow the exploitation of the full arm workspace [84, 161]. This is of utmost importance as the optimal assessment is conditioned on allowing free movement. The end effector structure is also considered to lack the involvement of proximal limb coordination while relying on synergistic tasks to assess patients. Thus, obscuring a major upper limb degree of freedom [120]. The fact that these devices do not constrain elbow rotation is considered dangerous for unsupported tasks [102].

Therefore, the devices are confined to horizontal tasks only. The research has been clear, the robotic therapy does not exceed the classical therapy effects under similar intensity and frequency [120]. This can be true to the nature of the robot devices that have been in use which often offer active assistance for training. This also was stated clearly, passive motion using robots does not have advantages over classic therapy [56]. The fact has been iteratively related to the controllers used in conjunction with rehabilitation robots which tend to provide excessive assistance encouraging patients slacking [56].

This aspect of rehabilitation robotics was inherited from early rehabilitation protocols based on delaying active movement to reduce spasticity. This theory is outdated, and research showed it does not hold [134]. For once, even if the training intensity has increased, the effects of the robot assistance on the kinematics measuring performance and active participation of the patient is not fully characterized [18, 56, 71]. For all the robotic devices, the lurking effects of external forces, including gravity, inertia, and passive or rotational mechanical forces, directly interfere with the muscle-induced activity [120]. Whereas when using end effector measures, interferences effects get hidden.

Challenges related to theoretical conception When trying to formulize the theory driving the kinematic assessment scale, the first question that comes to mind is the nuances of importance to the clinical practice. The measures are related to task-specific improvements, therefore dissociating the impairments from the function is relevant [81, 90]. The first involves the physical aptitude of the patient whereas the last relies on that aptitude to execute and fulfill a relevant function.

The upper limb functional tasks which we seek to assess are generally handled flexibly with a single limb taking multiple potential functional roles. Thus, quantifying a stereotypical model of the phenomena becomes a nonobvious task [149]. Scholars attempted to parameterize these dynamics through the modeling of the assessment measures during time, yet the proposed modeling techniques are rather descriptive and lack the integration of fundamental functions such as motor learning during recovery [11]. The kinematic scale, while reflective of the performance level apparent as faster, smoother, and stable movements, still overlooks the interferences of cofactors such as a decrease in spasticity, muscle impedance, and increase activation [11]. Trying to assess the efficacy using endpoint movements covers the influence of compensatory strategies and proximal control [100, 120]. The compensation appears as a motor pattern adapting the remaining function to the incurred loss by substituting with trunk displacement or rotation, scapular elevation, shoulder abduction, and internal rotations. These effects leave us with a partial view on true performances [120] as they positively affect the kinematic measured.

Challenges related to clinical utility When it comes to clinical adoption, the assessment metrics are challenged by the inter-subject variability and multiple confounding factors in the stroke population, these are the factor that hinders predictability [151]. Additionally, diversity in age and time to onset proved critical because spontaneous recovery and muscle strength highly affect recovery potentials [41].

These aspects affect the measure time course and thus the conclusions that we can induce. Moreover, scholars showed concerns about the kinematics' sensitivity and whether they are measuring dimensions of interest, such as voluntary control, reflexes, and synergies [102]. The kinematic scale is thought to lack the characterization of the longitudinal sensitivity of its measures [100]. For instance, the minimally meaningful confidence interval of change in task time was (-2.4,-8.4) [116] which can be a coarse measure if we consider timed tasks and chronic population with only recovery effects apparent.

The clinical meaningfulness of these measures is considered lacking. Nevertheless, good clinical evidence is starting to emerge [109]. The kinematic measures are reported to have a weak to moderate correlation to the FMA scale [161]. The correlations were termed "controversial" as differences emerged in effects, significance, and signs [161]. Consequently, weak functional transfer of gains to life independence is often reported [102, 161]. The kinematic scale suffers a lack of clinical use standardization [120]. The evaluation of their clinimetrics and concurrent validity are reported to be lacking [100]. This reflects on the clinicians' trust in these tools. We note a lack of interjection to therapeutic intervention planning as 20% of studies in [145] used robotics for both therapy and clinical assessments in stroke populations.

Kinematic and the ICF According to [161], of the 49 kinematic measures clustered into ICF domains, they found 44 belonged to the body functions and structure, five into activities, and none in participation nor environmental factors. When further investigating the cluster referring to body function, none of the metrics was found to assess individual components of body function coordination, force, and impedance modulation.

These factors hinders causality conclusions in clinical practice [149]. As most kinematic measures are related to the ICF body function level, incongruently to clinically meaningful gains which are at the activity level or beyond [161]. Besides, clinical scales (like FMA) consider arm and hand movements at once, including measures from both arm and hand, which would better predict this scale [78].

Research Uses and Challenges Researchers' struggles seem to be amplified as costly robotic rehabilitation, specifically during pilot studies, limit the progress of this research towards large-scale conclusive studies [102]. Researchers are trying to process and extract meaningful and useful outcome measures from the robotic data [102, 128]. One of the main research interests is boosting the motivation and the confidence of patients through data processing and presentation [102] which might be attained by providing patient-specific feedback for example. The direction towards autonomous rehabilitation care requires supervision and sustaining patients' motivation during therapy [84]. This can be achieved through monitoring and adaptability and would require a suitable assessment scale to furnish these functionalities.

The current kinematic measures seem to be interesting for neuroscience studies as they tackle simple abstracted tasks simplified to allow isolation of fundamental aspects and their underlying anatomical root [149]. This can help build upon our understanding of the underlying mechanisms of recovery and the associated patterns observed using the kinematic scale.

Scholars are also interested in common questions that would find use in assessment measures. For instance, standard assessment of interlimb coordination is relevant as it is key to understand motor control solving motor redundancy for optimality [149]. Interlimb coordination is a challenging aspect to measure as it often uses bimanual tasks which require complicated devices.

Nuances between experimental setups, theoretical formulations, and modalities can benefit from sensitive scales. The classical passive training focuses on position error control and trajectory planning, whereas the physical state of training the impaired limb is dynamically changing due to uncertain internal or external factors [124]. Therefore, accounting for and detecting these subtle differences might allow the construction of scales that can dissociate the different effects and hence, robustly parameterize the rehabilitation process.

A late study [103] showed the current kinematics derived from healthy subjects presented large disparity with a skewed distribution reflecting strategical choices difference. These findings are double-edged. Firstly, the kinematic scale is supposed to have low intra-patient variability, considering the need for normative data from control subjects. On the other hand, this is a promising finding, in that, it can offer strategic level insight on movement planning and coordination at the expense of higher variability.

2.6 Advanced Rehabilitation Tools

As researchers continue to gauge the fundamental questions in neurorehabilitation, addressing the challenges linked to the current kinematic scale motivates further investigation into additional modeling techniques and assessment instruments. This opens up the door for further advanced tools conceived to reply to remaining research questions.

Newer Perspectives As aggregate evidence shows us [98], robotic mediated therapy does not, alone at least, provide any additional gains in recovery compared to conventional therapy. This is not a strong justification to relate to funding boards while attempting to justify a huge upfront investment in a robotic platform. To weigh over this detriment, the instantaneous inclination would be to gear the technology towards enhancing Clinical Intelligence and supporting Evidence Based Practice in physical and occupational therapy which is highly lagging [73, 137].

Moreover, this cannot be fully enabled using the traditional scales whose frequency of use lies within the average of 3 months [55]. The hypothesis is that researchers should strategize providing assessment techniques accompanying the robot to establish a counterbalancing added benefit and reply to the need for full rehabilitation monitoring [153]. This orientation should move us closer to personalized therapy as iteratively recommended in clinical guidelines [72].

Of interest to our research, is exploring the possibilities of extending these tools to account for cofounding variables obscuring the rehabilitation process [60]. These extensions would furnish more accurate objective assessments and deeper phenomenological understanding. Issues remain open, for instance, what drives the recovery process and whether it is possible to predict the long-term outcome of a rehabilitation treatment [25]. Surprisingly, the physical assistance mechanisms promoting motor learning or relearning are still poorly understood [25]. Interestingly, authors in [11] explored two types of models adopted by researchers for recovery parameterization for instance:

- *Phenomenological Models:* These models use power law to define the time-course evolution of recovery during robotic rehabilitation.
- *State-Space Models:* These approaches try modeling neurological recovery as linear or nonlinear dynamical systems.

They concluded by recommending exploring models of adaptation to characterize the recovery course. Besides, they suggested that models of the subacute population must account for spontaneous recovery as well as therapy-induced motor learning, in contradiction to those in the chronic phase, for whom only the latter plays a role in recovery. The understanding of these mechanisms has often been reported as lacking [25] and as we shall see, not many studies attempted to bridge this gap.

Technical Considerations While we are considering robot-mediated therapy the motion analysis approach can be based on different acquisition hardware. In the clinically oriented literature, we find electro goniometers, inertial or magnetoinertial sensors, and optoelectronic systems [126].

A critical complication in their regards is misalignment which worsens with dynamic activities. Rotation planes are obscured by the placement of the sensors which results in lost information. Sensors placed on the body are also prone to alter the natural motion and obstruct the motion, specifically in sensor network-based settings needed to capture the entire upper limb motion.

Camera systems are the golden standard when seeking accurate reconstruction but are complicated to deploy due to the high costs as well as high structured acquisition environments with a time-consuming operational procedure (Marker installation, Calibrations, and Familiarization).

No system is all around perfect, and weighing the benefits and availability of each technology is critically decided based on the objectives of the study.

Different advanced rehabilitation modalities are conditioned on the use of one or another of this technological equipment. For our purpose, we present techniques that are relevant either in the research problem or in theory to those we are investigating.

2.6.1 Movement Quality

Advanced approaches investigating movement quality are reported in the literature. Authors [46] intended to explicitly model the joint velocities using a linear correlation between joints. The model used principal component analysis to determine constraints in pointing task redundancy. These constraints are supposed to define a subspace with the least variance of speed and quantify the difference between the natural constraint, therapist imposed and robot imposed constraints. As a result, the kinematic coupling was detected using three components explaining 96% of the variation in data. They also concluded that the therapist constraints did not alter the end-effector trajectory while decreasing the elbow angle.

The use of functional Principal Component Analysis to model motor control recovery was assessed on a 3D task with gravity assistance using the end effector robot in [89]. The methods resulted in 3 components explaining 99% of the trajectories variance. Mahalanobis distance was proposed as a measure to differentiate between the patients and elderly control group. While no effects were noticed on FMA or Action Research Arm Test.

The functional Principal Component Analysis showed a significant decrease using the average mean difference measure which defines the distance between the groups. The clinical scales were concluded to involve more components in addition to MC which translates in their constant values. The authors go on to motivate a question about whether the robot is aiding task skill or in MC recovery? This was speculatively answered as the skill increase is not a surprise since hemiparesis is considered an MC impairment and not a learning impairment. The conclusion stated that generalization and adaptation from training to assessment are causes for noticeable improvements and the MC was rather a subsequent and not major effect as seen on FMA and The Action Research Arm Test (ARAT)[5].

Researchers in [8] presented another approach to applying the functional Principal Component Analysis framework. The study involved 33 subjects in a reaching and

[5]A functional test for upper extremity focusing on performance corrdination dexterity and functionning most for stroke ABI and Multiple Sclerosis. It comprises 19 component and is observational scale.

manipulation task. The reaching trajectories were captured using a vision system. Then, they analytically estimated the joint level configuration. They used three components to model trajectories and were able to express 80% of variation. Further, using a single component, they were able to generate human-like free movements. Adding more components was necessary to deviate from obstacles.

The synergy between the elbow and shoulder joints in reaching is investigated in [112] through drawing an eight-like graph of the variation of elbow and shoulder velocities. The graph defines a metric named, kinematic coefficient C. The metric represents the synergy between the two angles and showed sparse results for the patient compared to healthy controls. The task involved the whole functional workspace[6] without time constraints. Bell-shaped velocities in joint level result in the eight-like figure. Hemiparetic patients had variable velocities, especially for the elbow joint. The metric permitted the distinction of the two groups as well as subject level differences.

The results while not able to identify motor control strategy, obtained useful information about its mechanism. The results confirmed the recovery of strength happens from proximal to distal. The variability noted in hemiparetic performances was related to the tendency of subjects from the neurological pathology to introduce altered and non-homogeneous motor patterns with a content function representing the repeatability of the same movement.

The belief is that movement during therapy is undertaken within a manifold of the space representing a plane. This hypothesis was tested in [168]. Reaching movements were used in the form of the peg hole VR exercise (see Fig.2.10) and movements were captured using the phantom device with Multiple Sclerosis patients. The results showed that planar tendency was present and its orientation angles follow a binomial law.

Using Fitts law, the controls complied while patient's data showed higher times proportional to distance. Comparison to Minimum Jerk Model showed that patients significantly deviated, whereas controls seem to have better fits. The conclusion is that this boundedness to the sub-plane is favorable for cases of control of robot devices. The reference was made to the literature showing the same tendency using wrist motion planar piece-wise approximations.

Using robotic devices to capture the movements introduced a further need to express viable and computationally efficient Inverse Kinematics (IK). Particularly, the IK would permit researchers to link back to joint motions more relevant to

[6] Ipsilateral, frontal, contralateral, proximal, medial, and distal parts of the workspace.

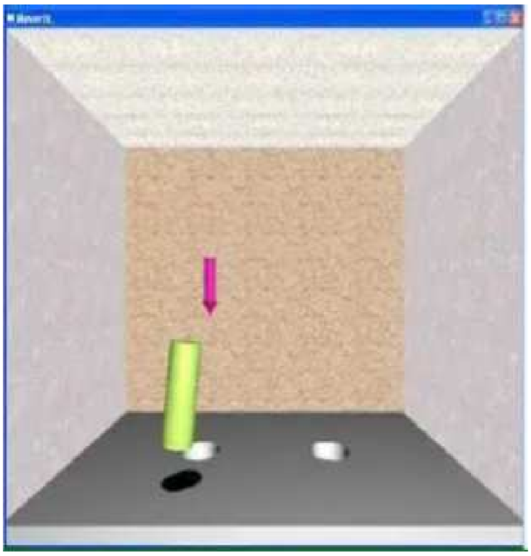


Fig. 2.10.: An example of the peg hole virtual reality assessment task for upper limb.

rehabilitation therapy. In [114] authors approached the inverse kinematic analytical solution for redundancy problems assuming the slow rehabilitation's movements.

They solved the redundancy by introducing an additional Inertial Measurement Unit capturing the elbow configuration. Similarly, authors [15] used two IMUs to deterministically solve the problem and calculate joint trajectories. In a similar fashion [34] attempted to solve for reachable workspace using kinematic modeling of the upper extremity for exoskeleton therapy. The usable workspace or joint configuration would characterize movement qualities at joint and visual frames to assess the therapy progress.

The authors in [48] investigated the proprioception acuity and proposed an algorithm based on Neural Networks and Radial Basis Functions to process coordinates during the reached task. The study resulted in detecting statistical between-group differences.

2.6.2 Behavior Modeling

Another parallel granted researchers diligence was in conceiving models and frameworks to approach the behaviors during rehabilitation training and exercises. The recommendations of using adaptation models to fit the recovery course and the introduction of variables factoring in the effects of instantaneous recovery for the acute population were advocated [11]. The evolution of the learning performances was modeled by the use of a logarithmic equation fitted to data in [39]. The model-related parameters served to differentiate post-treatment differences and engagement measures correlated to the functional outcome.

Authors [25] suggested that modeling recovery as a dynamic process might explain the driving mechanisms of recovery and potentially predict the outcome. The approach was to model the per trial dynamics of recovery during robot therapy. The task consisted of arm extension in a non assisted reaching movement. The task was carried out with or without vision and featured nine stroke patients.

They assume the model to be a linear, discrete, time-invariant dynamical system to account for different evolution time scales. The model expresses the relationship between voluntary control as the sum of learning, driving, and assistance. The planar horizontal robot Braccio-de-Ferro was used. In addition, they introduced a Boolean representing the use of vision. The results suggest large positive learning rates. The vision was found beneficial to the mildly impaired contrary to the severe cases where it did not contribute.

The adaptation model used in the study presented less Variance Accounted For and this can be due to the absence of the exploratory component to learning. Other causes advocate the intervention of psychological effects on performance during task training. The sensorimotor adaptation representing the MC fast response in correcting errors is better substituted with ML which can tolerate errors while slowly increasing performances and have stronger retention comparatively.

To measure the FMA at higher correlation rates the authors [92] proposed a nonlinear combination of kinematic measurements. The composite measures results showed $.9$ correlation factors and increased the effects significantly between 7 to 90 days.

Learning the kinematic model of a task appears to be a promising endeavor since trajectories encode for both performance and strategies, while also being easily portable between systems. To this end, we can find the process of learning the kinematic models of the patient during a specific task. These methods are often conceived for path planning for active exoskeleton control. We refer to [171] for a detailed overview of these methods and techniques for rehabilitation applications.

Particularly, of interest in our context is the emphasis the authors put on the fact that the potential final model of the patient's performance remains a hidden outcome. They point out that the proposed modeling techniques need to account for motion limitations due to fatigue and stiffness among other factors influencing the performance of the patient. The fact remains that an active research endeavor is modeling the recovery process whereas the few approaches proposed within the rehabilitation context still lag in terms of clinical validity and general utility.

End effectors limitation to Cartesian space was approached in [126]. The benefits of using the joint movement analysis were proposed as quantifying per-joint residual motor capability, assessing strategies, and their dependency on pathology. It can also benefit safety issues where joint limits are set for each patient to prevent overstretching and injuries. The swivel angle of the elbow is estimated using a single IMU which was valid assuming slow movements. The results suggest that motion variability decrease after recovery in both joint and task spaces. The increase in recovery was followed by increased excursion of the elbow and shoulder. They concluded that FMA differences between patients are related to the task-involved DoF, which are not necessarily assessed DoFs during FMA.

Another group sought to enable patient rehabilitation planning by helping them learn their movements [94]. As we saw in an early section, the Minimum Jerk Model has its limitations and cannot reproduce the patient style of motion. The study proposes the use of spline decomposition jointly with optimization costs to model the human trajectories during human-robot interactions. The authors presented a Dynamic Movement Primitives based framework where seven functions were used to model joint trajectory. Using Locally Weighted Regression the parameters were estimated from healthy motion data. The test was conducted on a Kuka LWR anthropomorphic robot and results showed a reconstruction error lower than 5% allowing 0.002m on a Cartesian movement of length 0.4m and 0.08 rad error for a 1.57rad in joint rotation. The motion style index was defined and quaternions parameters were also found to differ using the Wilcoxon test.

In [58] the authors suggest a learning algorithm for modeling the time and space alignment from the movements as a distribution over the parameters. The model was then used to assess the correctness of the movements and provide haptic assistance to the patients. They defined a standardized score based on the expert model and user distributions. The method was extended to define a Reinforcement Learning policy-based optimization algorithm that was modeling performance constraints of the desired trajectory and sample parameters for expert-like trajectory generation.

In [117] ADL trajectories were learned for optimal control of assistance during therapy. They modeled the elbow angular trajectories. They used overlapped sub-movement to predict the joint trajectories in ADL tasks. The parameters characterizing the sub-movements were learned from data using three Artificial Neural Networks for the duration, start, and amplitude separately. All of the predictions successfully attained the target. The Variance Accounted For of the trajectories was

98.9% with a root mean squared error of 6.51° degrees. The method did not find any task or subject level differences for the Artificial Neural Networks reconstruction ability.

In [138], a musculoskeletal model of the patient is built by approaching the human body with a multi-joint kinematic model and estimating the parameters from the data afterward. An extended review paper in [96] provides more details on studies implementing similar models. These kinematic models define the joint structure of the body during a rehabilitation exercise from a morphological point of view to assess the movement conducted during training or estimate the interventions' effects.

Another work proposed in [110] presents a learning technique for the kinematic model in the context of gait trajectories identifying centroids in trajectories from multiple healthy users' demonstrations and interpolating them to model a step. Learning the Kinematic model as a discriminant representation of the user without considering the variability in his performance would consist of an undermining factor to the generalizability of similar approaches. Meanwhile, in [93] the authors used the Dynamic Movement Primitives to encode for the trajectories accomplished during the Activities of Daily Living (ADLs).

The challenge that the Dynamic Movement Primitives framework presents, is the assumption of availability and knowledge of the model of the system which is indeed the case with robotic arms. The Dynamic Movement Primitives-generated trajectories are also smooth splines that would fail to capture the imperfections that persist in the patient's behavior.

A Locally weighted regression has been used by [7] to learn robot arm control by approximating the local model at each position. This approach while time-efficient for online learning does not provide enough generalization in the output model, it also may present a challenge for continuous monitoring applications since it is a memory-based approach and may present overflow challenges.

In [166], subject robot interactions were studied by building a model of daily life activities using the standard minimum jerk model. The model was insufficient to capture all the settings that the subjects demonstrated. The authors also noted some significant differences in task execution and strategies. They claim these differences would be more apparent in pathological subjects. A serious note was put on establishing a more sensitive model to capture the curvatures demonstrated during the reach movements for all the panel members.

To tackle these limitations, evolutive or iterative approaches were proposed. For instance, in [32] the authors presented an evaluation of a hybrid of both model-based and model-free reinforcement learning approaches and the respective algorithms used for training in the context of trajectory learning. A persistent challenge with these propositions is that they still assume fully known Markov Decision Processes besides some system model knowledge.

For trajectory tracking, Iterative Learning Control [2] for motor learning has been proposed as a theoretical framework to approach the evolution of the learning. These methods rely on the iterative approximation of the injected controls to ensure convergence towards the final model. However, this can only match perfectly repetitive tasks such as robotic manipulation. Besides, the assumption of knowledge of the patient model trajectory is unavailable in the context of rehabilitation training. The methodology assumes invariance in the dynamics governing the system evolution, whereas in rehabilitation, the dynamics are a hidden model to estimate.

Moreover, the dynamics have the potential to evolve frequently considering the recovery process taking place. By using the Kalman iterations and a clipping filter to handle major errors, scholars [165] proposed to predict the movement of the patients. They used the kinematic joint trajectory to predict the intention of the patient during robot assisted movements. As a result, they were able to closely follow the evolution of the trajectory. This technique allowed to overcome lag and abrupt changes that usually destabilize the filter predictions.

In [83] the authors presented an approach to learning an optimal task demonstrated by a user. The resulting model aims to control robot interactions through the reproduction of the movement.

2.6.3 Adaptive Rehabilitation

Ensuring patients' engagement during the robot therapy training was approached either by employing assist as needed controls, detection of intent, or virtual reality games. Adaptive training protocols and intent detection are believed to promote recovery and thus improve the outcomes of therapy [18]. For example, the Unitherapy software supports multiple medium hardware used for rehabilitation training and assessment. It offers the potential for remote monitoring and adaptability of the task. The assessment is proposed through an extensive list of kinematics from task trajectories [84].

The work supposes that adapting the assistance using an exponential feature of reaching kinematics can be useful. The task parameters are changed for each session. Using this algorithm the difficulty changed five sessions earlier than the therapist suggested. This helps match task difficulty to the patient's skill and accounts for local variations due to direction changes [42].

Automatic evaluation of outcomes can be highly useful in tracking patients during semi or fully autonomous therapy sessions. Particularly, for the current robotic solutions that lack customization and always require therapists monitoring during training. They propose a state-space model to study the hidden dynamics of the motor improvement variable. The model of a random walk observation equation was expressed as a log-linear probability model. The performance was assessed based on the measures of abilities. Nevertheless, the challenge is that performances on the task level might encompass many effects that can influence the recovery effects [125].

Active end effector robot was used in [33] for testing metrics related to 3D reaching task. The authors suggest an adaptive setup of the duration and the reach according to subject performances. In [128], children using the Armeo were assessed using a newly defined P Parameter depending on the task level performances. The Median increased significantly between the pre and post-sessions. The oscillations remarked on the sessions' performances were attributed to the motivation efforts and fatigue as a potential factor affecting the patient's limb during tests. The P measure was used as a per session evolution measure.

Authors [123] proposed a method to predict performance scores as a linear combination of kinesiological observation. The method implemented variable difficulty training parameters to individualize the treatment. The study involves orthopedic patients and results from robotic therapy for these populations are limited. The game parameters were predicted using linear regression of kinematic measures and the results showed a monotonic fit between the predicted performance and the real measures. The adaptation of the training is assumed to be beneficial to neuroplasticity and motor relearning.

Many studies presented an intelligent agent to handle game difficulty for rehabilitation exercises based on virtual reality [4, 19, 20, 162]. Both [86, 155] are examples of approaches involving an assistive robot instead of a virtual environment. Their theoretical frameworks range between the Bayesian framework, fuzzy inference, and Markov Decision Processes with Reinforcement learning. To retrieve the parameters of the exercise they often rely on multi-modal sensing using RGB-D cameras. Otherwise recovering sensors' readings from the assistive robot were adopted for robot-based systems. Applications of telerehabilitation focus on building such assessment and monitoring models for online feedback for therapists and users as seen in [19, 129].

The authors of [54] employed a Hidden Markov Model (HMM) to follow and decide dynamically on the rehabilitation exercise parameters concerning the actual performance of the patient. In [22], HMMs were used to follow on the execution of a rehabilitation exercise captured by an RGB-D camera and to provide feedback on the correctness of the execution. A shared limitation to these approaches is that, while accounting for the environment's specific actions set (i.e. game controls), they fail to generalize easily across rehabilitation systems and remain thus system-dependent.

In an assessment-oriented application, rule-based techniques rely on a predefined set of rules to score and give feedback to the patient and therapists. As an example, a kinematic rule-based modeling technique was proposed by [170] which defines an encoding of the exercise execution rules. Exercise executions are then compared against the established base truth to give feedback and assessment. Fuzzy inference is used to decide on quality assessment. This approach lacks a dynamic update of the rule definition. The major benefit of the method is the ability to incorporate useful advice into the feedback given to the patient by specifying the execution error committed during the exercise.

In [23] the authors used an HMM to detect a set of features of a patient's movements during the exercise using an RGB-D camera. They then provide an assessment based on a set of predefined rules that were concerted with a therapist. The definition of these rules is a non-trivial task. It involves a considerable amount of work to handwrite these rules and use them later for assessment. The defined rules are also specific to the task and are not generalizable.

In attempting to investigate the underlying motor skill learning, authors [49] presented a Human Machine Interface capturing movements through an instrumented glove and controlling the cursor on the screen. The cursor movements were modeled as a 2DoF robot arm. The users were kept oblivious to these choices. The mapping of hand posture and the cursor position was approached using different Machine learning techniques. ML group learned the mapping earlier than controls who used a single function for the entire session. They concluded that the CNS corrected for the task-related errors hence the variability is confined to non-task-related dimensions. The reduction of control signals dimensionality to achieve the reaching skill is a form of motor learning.

2.7 Literature Synthesis

In this section we aim to properly define our topic and position our research within the body of knowledge. The literature investigating similar aspects is synthesized systematically in the following sections. Namely, in our synthesis, we are interested in robotic-mediated therapy research defining advanced kinematic modeling and upper limb rehabilitation assessment.

We considered protocols with gravitational compensation strategy to alleviate strain and enable early phase rehabilitation. The selected studies and the major findings are presented in Fig. 2.11. In the following sections, we provide the synthesis of the key aspects namely: research questions and motivating frame, devices used, assessment tasks, population, modeling frameworks, methodologies, highlights, and findings.

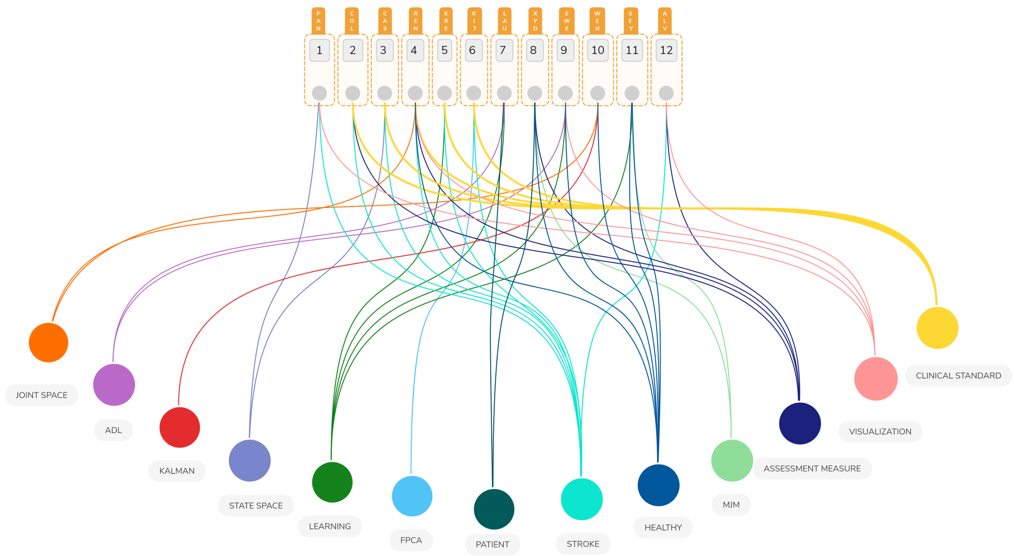


Fig. 2.11.: The thematic map of the studies reported in the literature synthesis. The authors names are coded as follows, PAN [125], COL [42], CAS [25], REN [139], KRE [92], KIT [89], XYD [168], EWE [58], WEN [165], SEY [117] and ALV [5].

2.7.1 Motivations and Research Questions

The addressed research questions reported in these studies stem from an inherent need for an automated (or semi) task/robot adaptation and trial-by-trial rehabilitation training monitoring tools [25, 125].

Adaptation was closely linked to the motivation and the psychological state of the patient. It is deemed critical to active involvement and consequently highly affects therapy outcomes.

The performance metrics used so far seem to happen at different time scales, as such, accounting for the temporal and spatial disparity was also investigated [25, 42, 58].

Some of the limitations of kinematics that were mentioned are lower side effects in terms of statistical changes reported on clinical data [92] as well as the assumption of apriori knowledge of the spatial range of trajectories at recovery end [89].

Kinematics often assume some characteristic model of movement such as Minimum Jerk Model which supposes start and end with zero velocity [168] and are reflective of very specific aspects of recovery as an investigated whole [89]. As such, exploring the recovery hypothesis and its underlying mechanisms seemed a core motivation as well [25, 42].

Prediction of performance through different methodologies has also been hypothesized and investigated in kinematics at joint level [5, 94, 117, 165], and an end effector with clinical scales [92].

Efforts were made to extend the functionality of the device by incorporating newer applications like stretching and diagnosis [139] or exploring newer types of platforms in the case of cable robots which are less intimidating compared to other devices and come with higher workspace [5].

2.7.2 Devices

When accounting for the used devices, we can note the following three types:

Robots Robots of different structure were used to record and assess movement in the studies namely MEMOS [42, 125], BdF [25, 42], MIT-MANUS [92], Reogo [89] representing variants of the end effector structure. These devices permit the user to interact with the system, often a computer-based exergame or simulated task, by moving the effector around the horizontal plane. A four DOF exoskeleton NESM with a 5DoF grasping extension was used in [94].

Moreover, in [139] the IntelliArm exoskeleton presents extended structure and functionalities. Exoskeleton devices are worn and attached to the entire upper limb and permit better weight support while also permitting the interaction and recording of movements at the joint level. A cable robot prototype was reported [5], where a moving platform linked to cables is attached to the wrist-hand during training.

Haptic devices Haptic joysticks were reported as an acquisition tool in the case of Phantom HF Haptic [168] (see Fig. 2.12) and Haption Virtuose 6D [58] (see Fig. 2.13). The haptic devices act similarly to end effector devices, where the patient guides the joystick to execute a task in a virtual environment. This permits the assessment of the pointing and tracking tasks. It also presents a haptic force feedback source as well as an assistance device to correct trajectory tracking.

Fig. 2.12.: Phantom Premium System. Courtesy of 3dsystems.

Recording prosthesis Laboratory research finds the use of instrumented prototype orthosis for joint movement recording as an exoskeleton [165] or a one uniaxial electro goniometer [117]. These devices represent laboratory setups for measurements and trajectories acquisition during training.

2.7.3 Tasks

In terms of evaluation tasks, the mostly reported task was the reach to target in [25, 42, 89, 92, 117, 125, 139, 168]. The reach to grab has been investigated in [58, 168] as a pick and place task, a Nine Hole Pegboard Test [168], and as ADLs

Fig. 2.13.: The Haption Virtuose 6D. Courtesy of Haption SA.

[94, 117]. The trajectory tracking tasks were reported, although using different patterns between studies (circle, etc.) [5, 58, 89, 92]. Joint level mobilization in the form of isolated joint movements was reported in [139, 165].

2.7.4 Populations samples and pathologies

The populations involved in the studies were predominantly post-stroke at chronic phase (N=61) [5, 25, 42, 89, 125, 139] or acute (N=208) [92]. Other investigated pathologies were Limb Girdle Muscular Dystrophy (N=4) [94] and MS (N=7) [168]. Healthy user(N=19) were reported in [58, 117, 139, 165, 168].

2.7.5 Theoretical model

Overall the reported theoretical frameworks are State-space (linear or nonlinear), Kalman filtering, Artificial Neural Networks, Dynamic Movement Primitives, functional Principal Component Analysis, modified Minimum Jerk Model, DTW, HMMs, Gaussian model, and Fitts's law. The methodologies can be categorized into:

State Space Modeling Authors are approaching the recovery as a non-linear system [25, 125]. In this approach a measure of interest, often a kinematic metric or composed score, is modeled following a dynamic Non-Linear System in time. The parameterization of the course evolution permits hypothesizing, identifying, and quantifying the different variables of interest affecting the evolution of recovery as seen through this measure.

Time series modeling (Cartesian/joint space coordinates) These methodologies are attempting to infer a model based on the trajectories recorded from either able-bodied or pathological populations. This was attempted through functional Principal Component Analysis [89] supposing that the movements can be divided into a set of components that are functionally related to generating the observed trajectories. Artificial Neural Networks [58, 117] are used to model parameters' variation based on task characteristics. The approach permitted the generation of proper trajectories to achieve the task.

Dynamic Movement Primitives [94] is another attempt to parametrically model the trajectories as a set of components and learn the parameters through the observations. HMM, and RL policy [94] approach to modeling trajectories permits the quantification of temporal and spatial disparity in the observed trajectories while also building a model based on the observations to assess the patient executions during training or guide the training by generating proper ADL trajectories.

Kalman filtering [165] uses the autoregressive filtering technique to estimate the trajectories under different conditions. Once the filter is properly tuned, it permits the prediction of movements online which can be crucial for robot control and patient security. This permits the identification of the proper dynamics that are driving the trajectories and test hypotheses.

Score derivation Normalized root-mean-square deviation [139], likelihood from a Gaussian model [94], model parameters [5, 42, 168], Neural Networks output [92] were all reported as a standalone scores of the evolution of the recovery during upper limb rehabilitation.

2.7.6 Methodologies

In terms of adopted methodologies, the shared drawback is the reported small sample sizes. Only a single trial was found [92]. Comparisons to clinical scores standards were reported in [25, 42, 89, 92, 139]. Only two studies reported comparative studies with controls [139, 168].

Testing the models' parameters using repeated-measures statistical tests [58, 94, 125]. Correlations between observed and model results was used in [42, 92, 117, 139, 168]. Mahalanobis distance [89] and Akaike Information Criterion measure of the derived models [25] constituted standardized measures for between models performance comparison.

Commonly reported quantities like Root Mean Square Error, Normalized Root Mean Squared Error were used in [117], Variance Accounted For in [25, 117, 125]. The visual comparison of results [165] were used for trajectories fitting as well. In the case of robot assistance, only a single study mentioned the use of transparent mode during tests [94]. A single study reported comparative results between affected and unaffected sides [89].

2.7.7 Highlights & Findings

State Space Modeling Using the state-space model revealed a difference between directions of trained movements [125]. The model permitted the adaptation of task difficulty five session earlier compared to therapist [42, 125]. The task adaptation was reported to improve patients' post-treatment FMA score [42].

Suggestions were made about the recovery believed to be driven by performance, specifically for those with spasticity. Using a state model with a variable refering to Vision presence, found it correlated to FMA and negatively to Ashworth scores. Meanwhile the learning parameter assumed as voluntary control component correlated to FMA after three months [25].

Time series modeling Adaptive Kalman Filtering, functional Principal Component Analysis, Dynamic Movement Primitives and Artificial Neural Networks, Hidden Markov Models/Reinforcement Learning proved capable of generating movement trajectories and successfully attain targets during simulated reaching tasks [58, 89, 94, 117, 165].

This was achieved through the use of different trajectory models. Nevertheless, parameterization was always estimated using patient or subjects' data. Thus, it permits the quantification of subject-specific trajectories parameterization.

The resulting parameterizations were used to generate patient-aware trajectories, provide a personalized assessment during live executions, detect differences and investigate hypotheses about the underlying mechanisms of the recovery evolution. It permitted investigating subject or pathology specific characteristics as well.

Assessment results The anthropomorphic quality of a trajectory was assessed using a measure called distance from the physiological joint limits [94]. Another dimension of interest was the loss of joint individuation. This dimension was evaluated using a normalized root mean distance squared and was correlated to motor status scores. This permitted identification of the specific joint with abnormal couplings [139].

Artificial neural networks were used to learn composed kinematic metrics which permitted the researchers to achieve a gain of 83% effect size at validation step [92].

When investigating topological features of the movements, the trajectories of Multiple Sclerosis patients seemed to deviate specifically on the horizontal plane projections compared to the healthy users. While significant temporal differences were validated with Fitts law, significant between-groups differences in parameters were found [168].

Modeling the desired pattern trajectory was proposed to predict the right movement period with 0.99 correlation and 1.5% error [5]. Through using functional principal component analysis model parameters did detect differences while clinical measures remained unchanged [89].

2.8 Chapter Summary

This chapter served for scoping and synthesizing research in the advanced kinematic movement assessment during robot-mediated rehabilitation. As kinematics are increasingly reported in trials evaluating the different effects of therapy strategies to help objectively quantify the recovery, there have been nevertheless challenges that motivated further investigations.

The identified literature showed promising techniques and frameworks that were used to extend the kinematic assessment. The extensions help to reply to many of the listed challenges in the literature review. The recovery modeling permit parameterizing the recovery process in a deterministic way, identifying and quantifying variables of interest influencing the variation of the recovery. Further, the definition of multiple scores served to investigate obscured aspects of the recovery, deliver instant feedback, and serve to adapt therapy sessions according to patients' performance.

Pathological movements were often generated using the Minimum Jerk Model approach, this also served as a major source of control input for assistive and rehabilitative robots. Regardless, the framework has its limitations. Gaining an understanding of the pathological trajectories enabled distinctions between groups, identifying intentions during arm movements, and generating proper patient-centered ADL trajectories. This was achieved by incorporating some further understanding of recovery mechanisms. For instance, the sub-movements theory is augmented by nonlinear learning techniques to estimate the proper parameterization.

The studies presented promising findings in reply to many of the challenges referred to earlier. Proper modeling for instance of the recovery might enable the investigations of recovery mechanisms in a deterministic way using objective theoretical models. Testing hypotheses and confirming them using these techniques would enable these conclusions across a wide range of rehabilitation settings assuming that sufficient kinematic data can be gathered. Thus, enabling further advantages to use robots and extending their utility beyond intensifying training dose.

When assessing the listed literature we can note that study sample sizes were small, which can be due to the feasibility nature of most of this research. Control groups were reported in two studies, and only one study investigated trial data. Thus, from a clinical strength of evidence perspective, the results are weak. Psychometric qualification of the techniques is also lacking and validation or follow-up studies were reported only for the state-space model proposition. The fact that only kinematic data is used permits the data to be gathered from multiple acquisition tools and thus also helps with the portability of these techniques.

The methodologies presented in this synthesis served to address many of the existing challenges that are reported for kinematics. It permits as well answering many of the remaining questions, as we gain further understanding in the form of models. Modeling phenomena is a strong research tool that helps us hypothesize, test, verify and ultimately extend our knowledge about the topic.

As the recovery presents a multifaceted interconnected process, proper modeling would enable us to separate and quantify different influence sources and their effect magnitude. Thus, deepening our understanding of the phenomena. With the ease of portability, this can also be enabled via many rehabilitation systems to extend robotic rehabilitation benefits.

2.9 The Take Home Message

Current kinematics has several challenges that would necessitate further developments. Specifically, the consideration of the future trends in the rehabilitation engineering field at large and those relative to the rehabilitation robots and devices invites research into more user dependent methodologies.

For instance, the research showed that many of the underlying theoretical frameworks that were used to establish the current kinematic scale have narrow applications and come with many simplifications. The research synthesis suggests that advanced methodologies incorporating the movement assessment during rehabilitation exercises are very promising.

Based on further understanding of the neuro composition of the brain after a deficit-provoking incident and advancement in adjacent fields such as autonomous computing and intelligence, the avenue for establishing more advanced tools that can give us a firm grasp on the theoretical drivers of the recovery is possible.

Kinematic analysis of the movement is shown to come at the intersection of many research directions we presented. Translational value from such modeling the technique can be highly beneficial and is thus, prioritized for investigation. The clinical validation of the tools is also of importance, as clinical research mandates the use of statistical analysis tools to establish the feasibility and the validity of any proof of concept tool for the domain.

Many requirements and needs are also stated reportedly by all of the stakeholders. These needs require researchers to take them into account to produce relevant tools for their target users. Investigating the communication tools such as feedback and data processing techniques is lacking despite the abundance of the data. The introduction of the new devices into healthcare motivates the investigation of artificial intelligence techniques to automate knowledge extraction.

2.10 Conclusion

By pinpointing the research gap and the potentially relevant literature, we characterized the state-of-the-art. In the following chapter, we first begin by cross-validating the established kinematic assessment tools on the Tlemcen CHU-PRC dataset.

We check the results in comparison to the major highlighted issues and the current trends. By providing clinical evidence we start off designing and presenting our theoretical proposition and the results of our study.

3 Investigating Kinematic Measures' Clinimetrics: A Case Series Report

> *The main problem of motor control and biology in general: the lack of an adequate language.*
>
> — **M. Latash & M. Levin**

As we have seen in the literature review chapter, the clinical practice started to slowly adopt the kinematic scale as a complement to clinical assessment. We have also seen that these kinematics came short, according to the researchers studying them, to reply to many challenges specifically longitudinal follow-up.

On a research basis, this property in assessment instruments can help many recent endeavors such as adaptive and patient-centered care as well as increase the understanding and efficiency of recovery. Technology-based functional training particularly can find multiple uses of these instruments in adaptation and monitoring functions.

In this pre-study, we aim to cross-check the kinematics quality as evolutive instruments. This also serves as a validation study to test the reported challenges and limitations as well as functionality. The study serves as a benchmark to compare our methodology to the state-of-the-art validated tool in practice and research today.

3.1 Study Sample Constitution

Firstly, six patients with discharge FMA measurements are included as a convenience sample. The patients' pathologies are post-stroke, post-traumatic, and neurode generative pathology. The group composition is heterogeneous in terms of age and stroke patients are older compared to the other two pathologies. This caseload for instance is representative of the daily care practice at the CHU-PRC.

Tab. 3.1.: Population description.

Patient	Age	Gender	Affected Arm	FMA-UE	Pathology
S1	45	Female	Right	54	Post-stroke
S2	15	Male	Left	44	Post-stroke
S3	54	Female	Right	47	Post-stroke
S4	30	Female	Right	48	Chronic inflammatory demyelinating polyneuropathy
S5	28	Female	Right	34	Chronic inflammatory demyelinating polyneuropathy
S6	32	Male	Right	44	Post Traumatic tetra paresis

A between-group analysis is limited due to the large disparity in subjects. From a motor learning perspective, the pool also differs significantly in terms of: recovery mechanisms employed, type and location of neuro damage incurred, and inhibitory complications that follow the specific accident. The subjects included are listed in Table 3.1.

3.2 Study Design

This work presents a longitudinal observational study on rehabilitation exercise data. The study was conducted, on records of exoskeleton upper limb rehabilitation exercises retrieved from the CHU-PRC of Tlemcen, Algeria.

The rehabilitation procedure is presented in section 1.3. The study aims to parameterize the recovery process through repeated measures in time. Outcome measures are defined based on the Cartesian end-effector trajectories.

3.2.1 Kinematic Metrics

The Cartesian exercise trajectories are used to define the elementary trajectories (ET). The ETs are then used to evaluate the kinematics. The kinematics of interest to us in this study are defined as follows:

- Distance traveled: the distance the effector travels during the task;
- Mean velocity: the average movement velocity during the task;
- Movement path ratio: the ratio of the distance traveled to the ideal Euclidean distance to the target;
- Movement time: the time elapsed between the appearance of the target and its capture or disappearance.

3.2.2 Clinimetric Properties

The clinimetrics of the kinematic measures are then assessed. The relevant properties as defined in [147] are:

- Clinical divergent validity: the relationship of the measure to a similar scale or measure with its ability to identify different impairments' levels. This measure is evaluated by comparing against the Fugl-Meyer Assessment's scores;
- Responsiveness: The ability of a given kinematic measure to detect changes in the evaluated impairment. To assess this property we undertake a statistical significance study of intra-patient kinematic measures' evolution. The property is defined as internal responsivity.

3.2.3 The hypotheses

According to the previous studies [147] and our assumptions about the metrics, we seek to validate a set of hypotheses. We postulate that for each measure during this evaluation, as the patient's recovery level advances, the metrics should behave as follows:

- Distance traveled: the distance would decrease as the patient becomes more accurate in his moves;
- Mean velocity: better motor control would result in higher velocities;
- Movement path ratio: the ratio would decrease towards one with increased movement accuracy;
- Movement time: time spent to complete the task would diminish.

3.3 Results & Discussions

In the sequel, the statistical significance is reported at a significance level $\alpha = .05$. Results were evaluated using Python scripts using Statsmodels and NumPy modules to perform the regression calculations.

Tab. 3.2.: Kinematics' correlations with the FMA-UE scores.

Measure	$r - value$	$p - value$
Distance	.615	.193
Velocity	.874	.022
Path-Ratio	-.067	.89
Time	-.545	.262

3.3.1 Clinical Convergent Validity

The clinical convergent validity of the measures proposed in section 3.2.1 is represented in Table 3.2 as evaluated in form of correlation coefficients with the FMA-UE scale and their respective $p - values$. According to the hypotheses listed in section 3.2.3 the only metric that increased significantly with the score of the FMA-UE is the velocity matching the a priori set hypothesis. Meanwhile, the correlation coefficient's signs for the Time and the Path ratio match the hypotheses. Whereas, surprisingly enough, the Distance presented a positive correlation coefficient instead of a negative one suggesting that further outlier results persisted in the dataset.

Performing the test on averaged windows of data points seemed to affect significance and corrected for outlier effects. Although a basic outlier elimination procedure was applied the systematic presence of errors necessitated domain knowledge expertise to set thresholds instead of relying on the infected data to estimate them from a pure data-driven perspective. The significance of the results can be related to sample size in this case since consistent results were found when applying tests to averaged windows.

3.3.2 Responsiveness of Measure

To investigate the responsiveness of the studied measures, we list regression analysis results as a characteristic of the evolution of each measure during the course of training.

Results for linear regression We start with analyzing first-order regression results as an indication of the responsiveness of the measure and cross-check the hypotheses that we set forth earlier. The reported parameters are the slope and intercept of the regression line, the $r - value$ as the correlation coefficient, r^2 as the quality of fit measure with the associated $p - value$ describing the percentage of variation in the

Subject		Distance		Velocity		PathRatio		Time
S1	-	.0, 0.014	+	0.006, .07	-	0.002, .001	-	0.0005, .16
S2	-	1.7, .25	-	0.4, .25	-	0.06, .17	-	0.02, .15
S3	-	0.006, .009	-	0.003, .0004	-	0.0001, .02	-	0.0005, .005
S4	+	0.02, .05	+	0.01, .004	+	0.003, .024	+	0.004, .006
S5	-	0.08, .25	+	0.07, .01	-	0.005, .2	-	0.02, .3
S6	-	0.002, .21	-	0.06, .005	+	0.01, .002	+	0.02, .0

Fig. 3.1.: 1^{st} order regression recap table. Key: Red non significant p, Green significant p, - negative correlation, + positive correlation. Numerics: SE standard error, R squared

metric represented by the model. SE represents the standard error of the estimate: the average expected error with the linear model.

Responsiveness of the distance traveled Figure 3.1 represents the different regression coefficients for each patient during the course of their training. When examining the evolution of the parameters we can see that overall the distance has a very weak 1^{st} order regression significance for most patients. Nevertheless, a significant result is shown for patient S2, with an $r^2 = .25$ which is the highest value in the population.

We can argue that distance traveled, although non significant, negatively related to the evolution of the patient, might be because of movements boundedness at the ET level by the standard work-space span and the varying positioning of the targets on the screen during the exercise. Diminishing the traveled distance is possible to a certain extent, that is the average ET straight-line length. This might suggest that the distance metric presents a ceiling effect and would not be responsive enough for consistent follow-up on the recovery.

The responsiveness of the velocity metric Examining the regression results in Figure 3.1 points to significant statistics for two patients with no clear tendency since we have varying slope signs. The $r^2 = .25$ is the highest value for patient S2. This might lead to concluding that the velocity can be used as a discriminant metric rather than an evolution metric as it can distinguish between the tendencies which would potentially contrast differences on the functional level for the monitored patient.

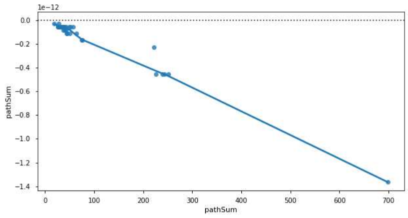
At a practical level, this can be useful as increasing performance means continuing therapy while consistent challenges and decrease of performance should invite more analysis on the part of the therapist to understand what difficulties are causing this decline in performance.

Responsiveness of the path ratio measure Reported in Figure 3.1 are the regression results for the patients in the study. Examining the results, we can note that the path ratio has three significant results with negative slopes for patients S2, S3 and S5. We also note significant r-squared with $r^2 = .20$ and $r^2 = .17$. This suggests that the metric is sensible, at least weakly, to the evolution of the state of the patient following the set hypothesis.

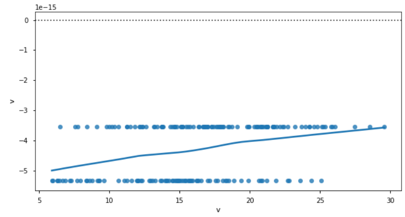
This is interesting as the smoothness of the movement represents a Motor Learning and control aspect of the recovery. The finding also suggests that although sometimes significantly decreasing the amount of variance that is expressed by a first-order was rather weak. That is the recovery, at least from our cases' perspective, did not respond proportionally to the training amount.

Responsiveness of the time metric The results in Figure 3.1 show three significant results for the time with negative slopes suggesting that the time decreases as the evolution advances. The three results are weakly significant with an $r^2 = .16$, $r^2 = .15$ and $r^2 = .3$ for patients S1, S2, and S5. This consistency in signs and significance suggests that time is a viable monitoring metric that is expected to decrease as the recovery is taking place. This also is compatible with the aforementioned hypothesis.

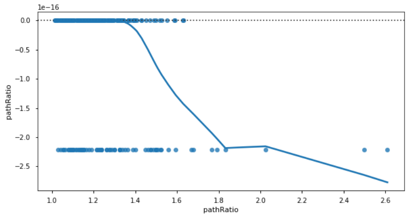
To examine the model's representativeness of the underlying data, we plot the residuals depicting the differences $\delta_t = y_{predicted} - y_{real}$. The reported results suffer from a weak representative power as shown on the residual plots of Fig. 3.2. We can see a clear "staircase" pattern on the residuals plots suggesting that our model does not have enough expressive power to represent the data and thus resulting in biased statistical tests. This undermines to an extent any conclusions we attempt to draw from the reported results using linear relationships.



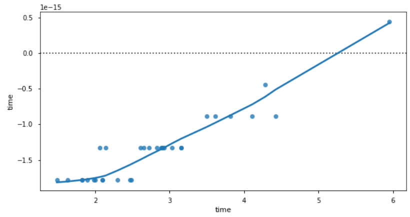
(a) S5 distance residuals



(b) S3 speed residuals



(c) S4 pathRatio residuals



(d) S2 time resuduals

Fig. 3.2.: Examplary first order regression residuals plots.

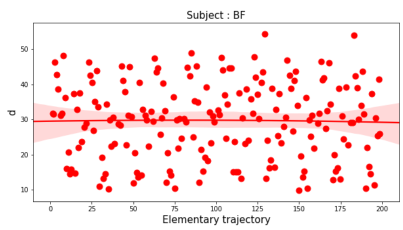
Results for second order regression Results are communicated by listing the regressors coefficients $\beta_0, \beta_1, \beta_2$ with their respective $p-values$ $\beta_0 p, \beta_1 p, \beta_2 p$. r^2 is the quality of fit and F is the f statistic for the model given the three regressors and the patient's metric vector.

The responsiveness of the distance The distance regression presented two weak results for patients S3 and S2 for whom we note a decreasing tendency in the total traveled distance. For patient S2 the $p - values$ are insignificant except for the intercept. The coefficients were all significant for patient S3. Observing the corresponding plot in Fig. 3.4 (f) shows that this sign might be due to the outlier values recorded at the beginning of the training which involved extreme cases that were not repeated after more training. On the contrary, Fig. 3.4 (e) presents a more likely case where we see, in concert with our hypothesis, a decrease in the total distance as training advances. Insignificance might be due, but not limited to, the small-sized sample.

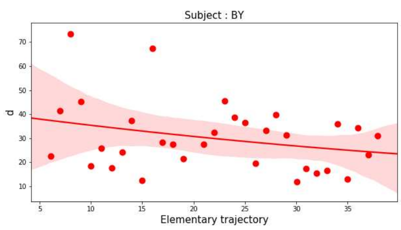
The responsiveness of the velocity Only one patient had a significant r^2 result, patient S3 with a significant $p - value$ for all the coefficients. Although comparing with Fig. 3.5 we can see that the higher velocities at the beginning of the training, that corresponded to the largest traveled distances might have caused this notable significance. On the most part, examining the regression on Fig. 3.5 we see an

Subject	Distance	Velocity	PathRatio	Time
S1	0.03, .0	9.19, .09	0.13, .001	21.69, .18
S2	9.6, .37	7.8, .32	3.58, .18	2.92, .15
S3	1.62, .013	1.01, .008	3.24, .03	2.53, .02
S4	3.2, .07	2.15, .05	5.2, .1	5.61, .12
S5	1.62, .1	0.63, .04	3.77, .2	7.42, .34
S6	0.17, .01	0.44, .03	0.38, .02	0.03, .002

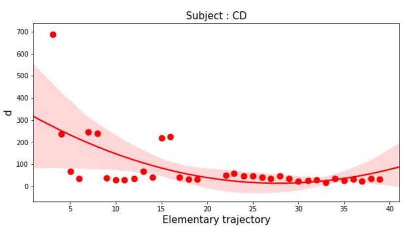
Fig. 3.3.: 2^{nd} order regression recap table. Key: coef. 1^{st} order, coef. 2^{nd} order: Red i non-significant p, Green significant p. Numerics: F statistic, R squared



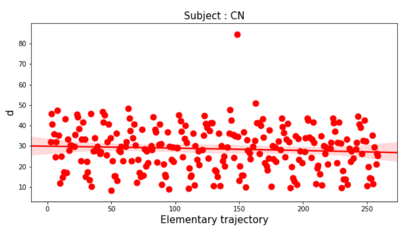
(a) Distance evolution for S1



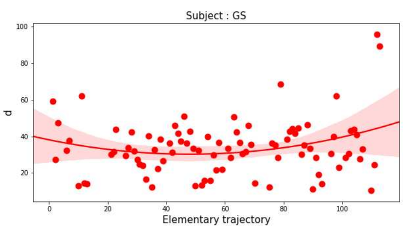
(b) Distance evolution for S2



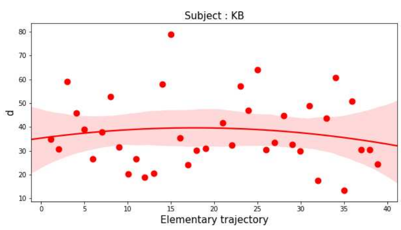
(c) Distance evolution for S3



(d) Distance evolution for S4

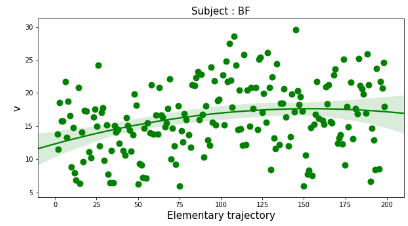


(e) Distance evolution for S5

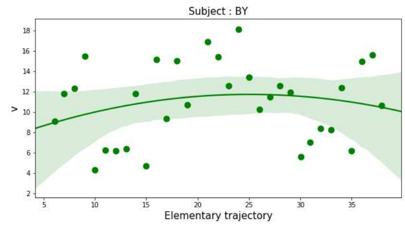


(f) Distance evolution for S6

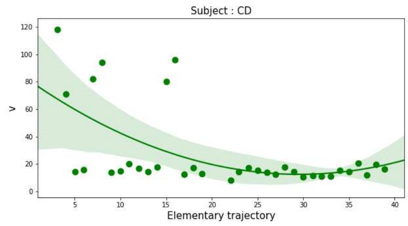
Fig. 3.4.: Second order regression plots for distance traveled measure.



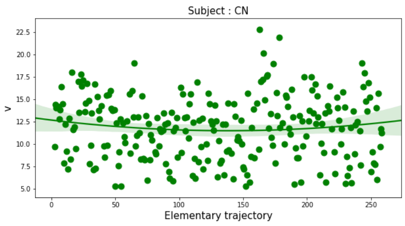
(a) Speed evolution for S1



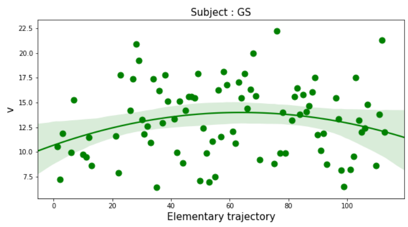
(b) Speed evolution for S2



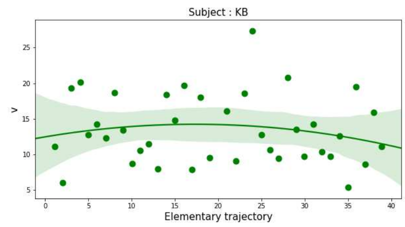
(c) Speed evolution for S3



(d) Speed evolution for S4



(e) Speed evolution for S5



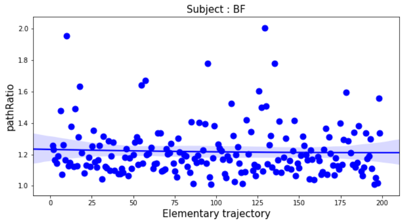
(f) Speed evolution for S6

Fig. 3.5.: Second order regression plots for speed measure.

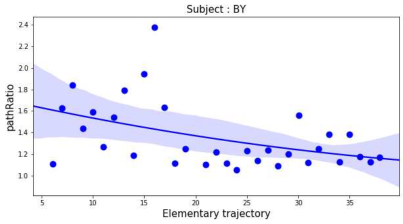
increasing tendency for the velocity metric as training progresses. This result defends our proposed hypothesis.

The responsiveness of the path-ratio Examining results in Figure 3.3 we note that patients S2, S5, and S3 had weakly significant r^2 results, while only S5 had significant $p - values$ for the regression coefficients. Fig. 3.6 illustrates how this might be due to the model overfitting the large values at the end of training. In the meanwhile, for both S2 and S3 the decreasing tendency in the path ratio values is consistent.

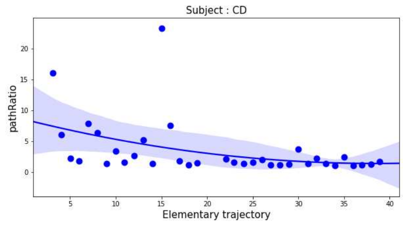
The responsiveness of the time For time evolution, patients S2, S1, S5, and S3 presented weakly significant r^2 values, while only S1 and S5 presented significant $p - values$ for the regression coefficients. Observing the regressions on Fig. 3.7 the time spent on the task has a decreasing tendency for most.



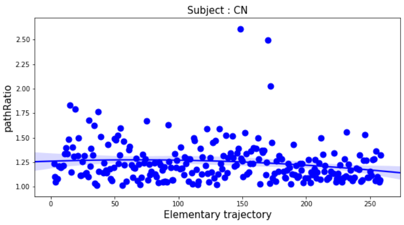
(a) pathRatio evolution for S1



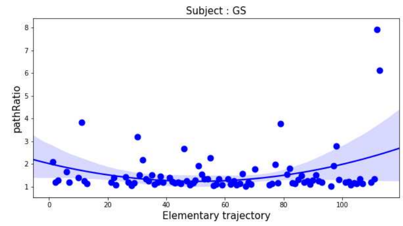
(b) pathRatio evolution for S2



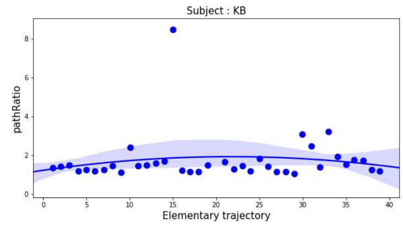
(c) pathRatio evolution for S3



(d) pathRatio evolution for S4

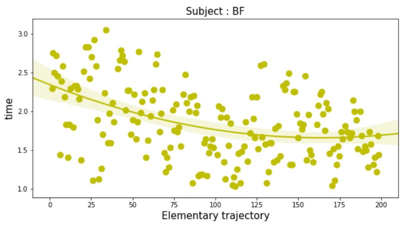


(e) pathRatio evolution for S5

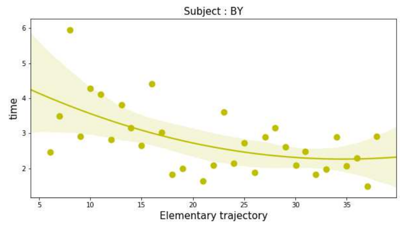


(f) pathRatio evolution for S6

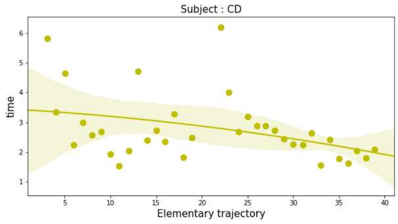
Fig. 3.6.: Second order regression plots for pathRatio measure.



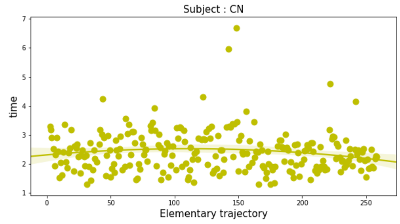
(a) Time evolution for S1



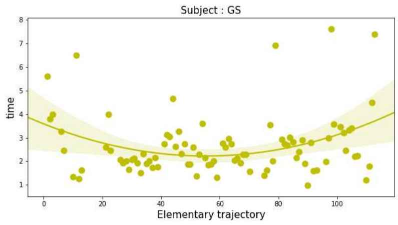
(b) Time evolution for S2



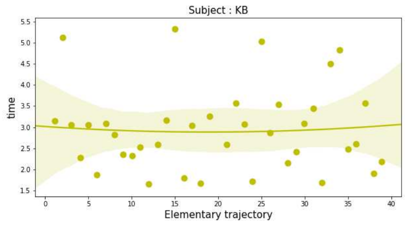
(c) Time evolution for S3



(d) Time evolution for S4



(e) Time evolution for S5

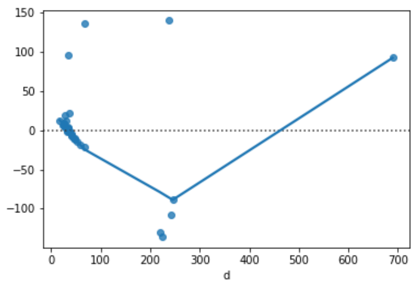


(f) Time evolution for S6

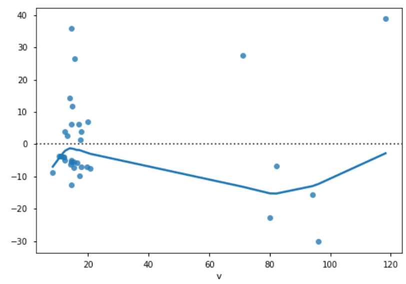
Fig. 3.7.: Second order regression plots for time measure.

By remarking the Fig. 3.8 we note that overall the residuals resulting from the second-order models were sparser. The fact reassures the pre-supposition that the second-order model captured more of the tendencies of the metrics and better fit their time course evolution. From a statistical standpoint, the regression analysis is indeterminant as it reports high conditioning numbers due to regressors collinearity.

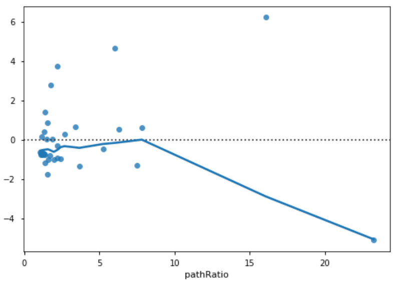
Notably, this affects our coefficients' $p - values$ and thus the significance of the results. Per contra, the ill-conditioning does not have a severe consequence on our attempt to determine the responsiveness of the measure to the change in the recovery state as measured by the $r - squared$, since it is only dependent on the squared sum of errors [31].



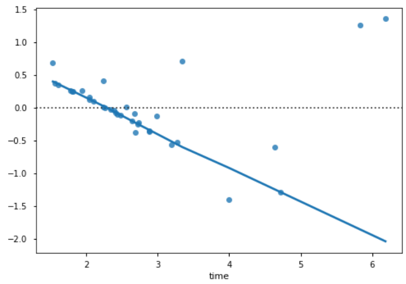
(a) Residuals plot for distance



(b) Residuals plot for velocity



(c) Residuals plot for path-ratio



(d) Residuals plot for time

Fig. 3.8.: The evolution of the second order regression residuals for the four metrics. Example case is of subject S4.

3.4 Conclusion

From the results discussed earlier we can come up with many conclusions. First, that the clinical validity of the metric studied matches the suggested hypotheses for all metrics. Metrics such as path ratio and distance might not be sensible enough to small state changes for intrapatient monitoring. Time and velocity showed consistent tendencies suggesting their sensibility to changes and thus for monitoring applications.

The regression study, albeit presenting weak evidence for most patients showed an overall consistency with the hypotheses. We noted that this type of characterization of the metrics can help us understand their strengths and shortcomings and thus enables their efficient use.

From a research perspective, the monitoring and longitudinal evidence derived from the study are rather weak. We showed that the kinematics showed significant results for a few cases. The results for velocity and time measures are rather expected, in that the task is a timed exercise and the setup pushes the patients to exhibit an increased velocity and thus reduce travel time considering the average distances are consistent between sessions.

The path-ratio presented more tendency to proportionally follow the number of exercises accomplished. Speed and time seemed to be a learning variable presenting more fits with the second-order model. Overall the significance is not qualitative for several reasons.

The SE values reported for the significant fits are higher and thus suggest the higher variability in the measures compared to the fit. Additionally, the statistical results exhibited highly condition numbers and suggest, regardless of more residuals sparsity, that the fits are not conclusive although the tendency is approved.

Likewise, the same challenge reported by the therapists concerning the utility of the measures across different pathologies for the stroke patients' performance measures (velocity and time) seemed relevant, while the results are not consistent in the neurodegenerate or post-traumatic cases.

Finally, the challenges reported in the literature review and reiterated in this prestudy merit further interest. When we consider the tendencies in rehabilitation engineering research compared to these limitations, we can see the importance of further understanding the recovery of motor function during recovery. From a methodological standpoint, the pure data science analysis proved limited in accounting for prior understanding and knowledge and might beneficially be augmented by expert input. In the next chapters, we present a new theoretical formulation and assessment tools to monitor training and bridge some of the gaps presented in the first chapters.

4 Towards Adaptive & Finer Rehabilitation Assessment Measures: A Computational Learning Framework for Trajectories Modeling

> *It is not science. Instinct is the opposite of science: research tells you what others have learned, instinct tells you what you have learned. Science studies other people.*
>
> — **Tim Grover**

In this chapter, we introduce our first contribution. The modeling approach that we are proposing came to address many challenges that we have faced and read in the literature concerning the kinematic evaluation of the rehabilitation process. We detailled the exploratory data analysis as means of addressing the literature review and as an assessment of the state of the art in the field. We then commented and highlighted the challenges that are left unaddressed and then laid the foundation to introduce our proposition as a potential answer to fulfill the needs of the end-users for both patients and practitioners.

We are adopting a statistical trajectory modeling approach to establish a mean of assessing the patient's performance. We seek to find the underlying model of the patient's hand movement during the exercise on an orthotic exoskeleton by capturing the motion of the end-effector using an HMM and imitation learning technique. Further analysis of the imitation learning technique in the context of robotic trajectory modeling is referenced in [121].

Our main contribution is to provide an instrument to assess the longitudinal evolution of patients during the instrumented rehabilitation training sessions. Hence, a framework would be developed to model the trajectories recorded during patient movements to permit the portability of this methodology to other exoskeleton devices, motion capture, or telerehabilitation systems. The framework relies on a data-estimated dynamic model. To measure the evolution of the rehabilitation new metrics will be proposed and tested statistically. We hypothesize that the resulting model will detect significant changes early on in the course of rehabilitation. We also hypothesize the significance of the findings will be pathology-agnostic. Finally, we aim to evaluate the framework on an operational dataset to ensure these captured changes are detectable in the deployment environment. Our starting point is the work that has been conducted by Coates et al. [37] which proposes an HMM trajectory representation and an algorithm to infer the ideal trajectory of a given task. Our reasoning to opt for this methodology could be articulated in the following manner:

- Firstly, the HMMs representation permits us to capture the fine details of the trajectories considering that we use an upsampled chain to model the optimal execution. Meanwhile, in robotics the trajectory modeling techniques often outcome a smooth trajectory. The fact of smoothing the resulting trajectories inadvertently contradicts our attempt to capture non-optimal and errored executions demonstrated by the patient;

- Secondly, the statistical representation would likely fit the stochasticity of the human Patient Controller and the Orthotic Exoskeleton system that we are studying, hence referred to as PCOE;

- Thirdly, the iterative procedure used permits constant update of the outcome model;

- Fourthly, the assumption of noisy demonstrations and the smoothing procedure would permit us to alleviate the rigor constraint on the data quality. This is important as quality is often imposed by the equipment; for instance, the sensors used on the Armeo Spring exoskeleton;

- Lastly, the portability of this methodology, learning state trajectories would permit the extrapolation towards a control task for any active-assist intervention using motorized exoskeletons as well as being applicable in the virtual environment systems.

To be able to capture the patient model, we first start by encoding the system dynamics. To this end, we implement a learning algorithm as in [1] which permits us to approximate the state transition model. We present a newer iteration and cost function that can perform as well with better consistency.

We show results presenting the ability of the model to keep satisfactory error rates while predicting state changes with both the original and the proposed implementations. To model the patient ideal trajectory, we implemented a variation of [37] where we used the Expectation-Maximization (EM) algorithm to infer the underlying model.

The resulting HMM model serves as the basis for defining a set of six new metrics that are studied statistically to investigate their properties. The reader is invited for a reminder on basic definitions and concepts related to the formulations introduced in this chapter in annex D. In sections 4.1, 4.2, 4.3 we start by presenting the framework and detailing its components. We end with detailing the definition of the proposed metrics.

4.1 The Framework Presentation

We propose to model the trajectories executed by the patient during the exercises. The outcome model serves as a tool to define a finer measure of the level of recovery. The algorithm is an EM aiming to infer an ideal model from a task demonstration presented at the input. The framework composition is presented in Fig. 4.1 and involves two main components:

- The dynamic model approximates the transition between states at times t and $t+1$. The model follows the exhibited system properties by learning from the data. The system is the complex entity we note PCOE given the fact that the controller and actuator, i.e., the patient is changing dynamically during training. From a robotics standpoint, the only known parameters are the geometric configuration of the 6DoF robot. The patient acts as a controller that drives the muscles representing the actuators. The contractions of the muscles result in the final movement of the device;

- The model trajectory learning algorithm permits the inference of the hidden model trajectory of the patient for a given task from the demonstrated exercise.

We are thus attempting to encode the dynamics and the kinematic trajectory model of the patient during this elementary task.

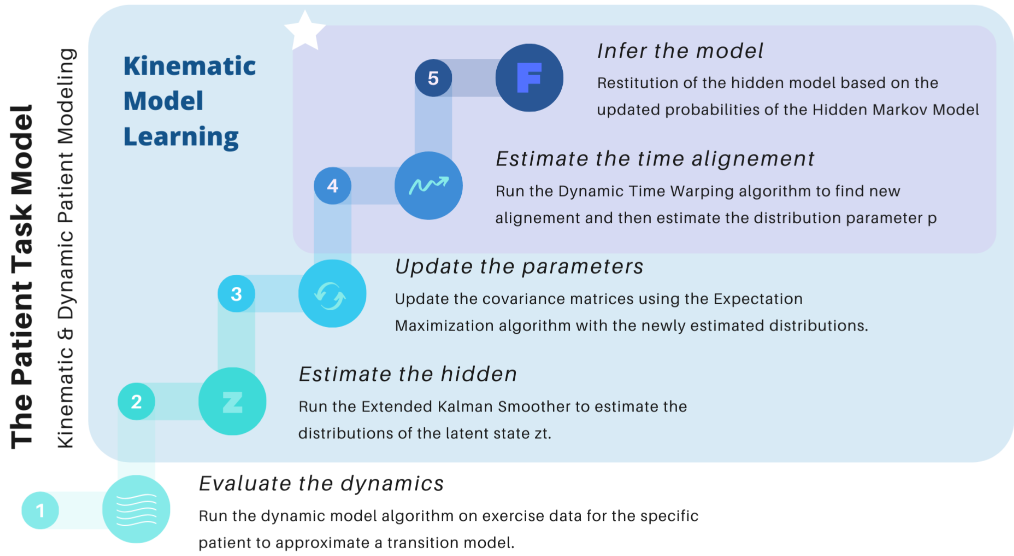


Fig. 4.1.: The overall framework permitting the modeling of the patient behavior from exercise demonstrations by learning from data.

4.2 State Transition Model Learning Algorithm

4.2.1 Piecewise Function Approximation Definition

We consider that the PCOE system is observed over discrete periods $t \in \Phi.\Phi = t \in (1, \ldots, T)$ where T is the length of the particular task we are studying in time stamps. The system state is defined as $z = [z_1, z_2]^\intercal$ coordinates of the end effector in the reference frame attached to the tip of the gantry located at the device's shoulder sensor as appears in Fig. 1.1.

Assuming the evolution of the state is in the Cartesian space and varies smoothly locally, we approach these variations using a third-order polynomial. The approximation can alternatively be seen as assuming the 3^{rd} derivative of the position, namely the jerk is constant which entails a smooth movement. The optimization problem is determining the transition polynomial model's coefficients to minimize the prediction error over an H-wide prediction window.

In our case, the ET's maximum time length does not exceed 2s for tasks completed with a successful catch. Besides, a resampled frequency $F = ET/MedianETlength$, we can consider the transition between state z_t and z_{t+1} to be governed by the following equations:

$$
\begin{aligned}
\hat{z}_1(t+1) &= \alpha_{z_1} z_1^3(t) + \beta_{z_1} z_1^2(t) + \gamma_{z_1} z_1(t) + z_1(t), && (4.1)\\
\hat{z}_2(t+1) &= \alpha_{z_2} z_2^3(t) + \beta_{z_2} z_2^2(t) + \gamma_{z_2} z_2(t) + z_2(t), && (4.2)\\
(1)\ \&\ (2) &\implies \hat{z}_{t+1} = A z_t^3 + B z_t^2 + C z_t + z_t && (4.3)\\
\text{with} \quad & z = [z_1, z_2]^\intercal, && (4.4)\\
A &= diag([\alpha_{z_1}, \alpha_{z_2}]), \text{and} && (4.5)\\
B &= diag([\beta_{z_1}, \beta_{z_2}]), \text{and} && (4.6)\\
C &= diag([\gamma_{z_1}, \gamma_{z_2}]) && (4.7)
\end{aligned}
$$

We employ the optimization lagged error criterion algorithm [1] to minimize the cost of successive predictions of state trajectories. The assumed local transition model is thus the approximation of the form:

$$\hat{z}_{t+1} = A z_t^3 + B z_t^2 + C z_t + z_t \tag{4.8}$$

Where A , B and C are diagonal coefficient matrices in $\mathbb{R}^{2\times 2}$ as defined in Eq. (4.8). To run the optimization algorithm, we use exercise files with measurement logs of the position of the end effector during the predefined assessment task.

4.2.2 The Algorithm

We evaluate our coefficient matrices directly on the Cartesian coordinates namely $z = [z_1, z_2]^\intercal, z \in \mathbb{R}^2$. We attempt to approximate the parameters of the cost value function by conducting estimates improvement using a black-box optimization technique on the exercise data. For this minimization, we define a cost value function as the cumulative incurred error during a prediction window of width H. The prediction errors are calculated starting from each timestamp t and swiping the entire dataset. Predictions are calculated using the model in Eq. (4.8).

In practice, we run the algorithms on empirically determined hyperparameters, a learning rate α_l set at 0.3 and a stopping criteria defined by a tolerance threshold ϵ of 10^{-5} resulted in satisfactory accuracy at 68 iterations. The learning rate is set

Algorithm 1 Compass : State Transition Model Learning

Initialize the A_0, B_0, C_0 matrices by minimizing the one step prediction error:
OSPE = $z_{i+1} - z_i - A_0 z_i^3 - B_0 z_i^2 - C_0 z_i$
repeat
  for $i = 0, \ldots, T-1$ **do**
    Run $\hat{z}_{i+1} = A_j \hat{z}_i^3 + B_j \hat{z}_i^2 + C_j \hat{z}_i + \hat{z}_i$
  end for
  for $i = 0, \ldots, T-1$ **do**
    for $h = 0, \ldots, H-1$ **do**
      for $t = 0, \ldots, h-1$ **do**
        $S = S + A_j \hat{z}_{i+t|i}^3 + B_j \hat{z}_{i+t|i}^2 + C_j \hat{z}_{i+t|i}$
      end for
      $cost = cost + \|z_{i+h} - z_i - S\|_2^2$
    end for
  end for
  $(\hat{A}, \hat{B}, \hat{C}) = \arg\min_{A,B,C} cost$
  $A_{j+1} = (1 - \alpha_l) A_j + \alpha_l \hat{A}$
  $B_{j+1} = (1 - \alpha_l) B_j + \alpha_l \hat{B}$
  $C_{j+1} = (1 - \alpha_l) C_j + \alpha_l \hat{C}$
until $\|A_{j+1} - A_j\| + \|B_{j+1} - B_j\| + \|C_{j+1} - C_j\| < \epsilon$

such that we change the estimated parameters slowly. By imposing this constraint, we ensure our learned coefficients are not majorly influenced given that the training data-set contains sharp turns and closed loops.

The optimization was conducted using a Powell solver [135]. This optimization problem is defined as a multidimensional ill-conditioned, due to the cost function that is nonlinear in the parameters. The approximation that we are taking following the referenced work is that the coefficient matrices used in evaluating the sliding sub-sum S in Algorithm 1 do not depend on previous estimates, which permits us to have a tractable problem formulation.

The function is defined as the integral squared norm of the error incurred on successive predictions with complexity $O(N_c.T.H^2)$, where N_c is the number of calls of the minimization solver, T is the model length and H is the window width. The Powell method does not depend on the differentiation of the cost function. This method makes comparatively fewer cost function calls during optimization iterations that is making it a robust and efficient choice.

4.3 Patient Model Learning Algorithm

4.3.1 Hidden Markov Model Definition

We consider each ET as a Hidden Markov Model where state z represent a stochastic process taking a countable set of values $z_i \in [-ROM_i, +ROM_i], i = 1, 2$ where ROM_i represents the maximum range of motion for the i^{th} coordinate. This value is commonly fixed for the assessment exercise to permit comparability between patients. The stochasticity of the state z is our means to encode for the variability of the PCOE's states [34, 12]. Following [37], we note z_t as the hidden variable representing the ideal state at time instant t. The hidden state vector z is defined as $z = z_t, z_t \in \mathbb{R}^2, t \in \Phi$. A naive transition model is assumed to be governed by the equation:

$$z_{t+1} \sim \mathcal{N}(f_1(z_t), Q) \tag{4.9}$$

The function f_1 represents the state transition model characterizing the PCOE system learned with Algorithm 1. The model noise is a zero-mean Gaussian distribution parameterized by covariance Q. Each $k^{th} ET$ we observe represents an emission of the hidden vector $\mathbf{z}$ and is noted $y_k = y_j, j = 1, \ldots, N^k, y_j \in \mathbb{R}^2$, where the measurement state y_j is evaluated through the observation model :

$$y_{t+1} \sim \mathcal{N}(h(z_t), R) \tag{4.10}$$

The length of the HMM is fixed to twice the median of the observations' lengths. The resulting model is obtained at a higher frequency. The higher frequency HMM permits mainly to accommodate the acceleration effects that can be noted as increased spacing or "jumps" in the time series of each coordinate. This effect is notable considering the fixed sampling rate of the device. The naive predictor in Eq (4.9) does not natively accommodate this change and we can benefit from additional steps to compensate for the error. The function h represents the observation model and is assumed to follow:

$$y_{t+1} \sim \mathcal{N}(z_{\tau_j}, R) I_\Gamma(t), \tag{4.11}$$

Where R represents the measurement noise covariance and I_Γ is an indicator function for $t \in \Gamma$, where $\Gamma = t, t = \tau_0, \ldots, \tau_{N-1}$. The time alignment τ_j follows the probability law:

$$P(\tau_{j+1} | \tau_j, p) = p_j, \sum_{j=1}^{3} p_j = 1 \tag{4.12}$$

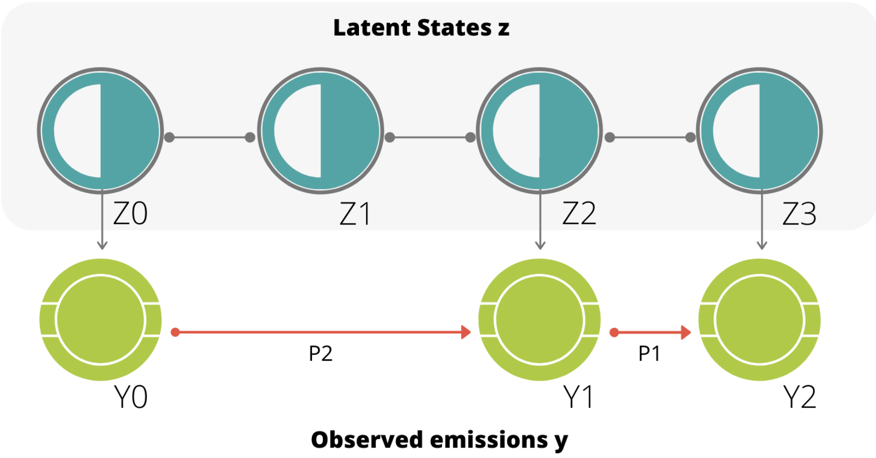


Fig. 4.2.: The trellis diagram representing the hidden Markov model with the aligned observations vector. The alignment exemplifies the probabilities defined in Eq.(4.12).

with

$$p_j = \begin{cases} p_1 & \text{if } \tau_{j+1} - \tau_j = 1 \\ p_2 & \text{if } \tau_{j+1} - \tau_j = 2 \\ p_3 & \text{if } \tau_{j+1} - \tau_j = 3 \end{cases}$$

The relationship of the hidden state vector $\mathbf{z}$ with the observations vector $\mathbf{y}$ is illustrated in the Fig. 4.2. Following the trellis diagram, the dynamic model returns, at timestamp t, the naive prediction of the z_t using the equation (4.9). On the other hand, the observation model h returns, at timestamp $t+1$, the last aligned smoothed hidden state z_{τ_j} with an added Gaussian noise defined by the covariance matrix R.

To increase the accuracy of the model we augment the dynamic model f of the HMM to be:

$$z_{t+1} \sim \mathcal{N}(f_1(z_t) + \zeta_t, Q) \tag{4.13}$$

while ζ is defined as: $\zeta_j = \xi(y_{\tau_j} - f_1(y_{\tau_{j-1}}))$ and $\zeta_0 \equiv 0$. The extended ζ_t vector is defined as follows:

$$\zeta_t = \begin{cases} \zeta_{\tau_j} & \text{if } I_\Gamma(t) = 1 \\ 0 & \text{if } otherwise \end{cases}$$

ξ is a dampening coefficient determined empirically[1]. The bias ζ permits to correct for the prediction errors of the model evaluated on the observations $\mathbf{y}$. We re-inject the bias into the model to compensate for the "observed" error in the model while predicting future states.

To estimate the parameters of these models namely, the distributions of the HMM state vector $\mathbf{z}$, the error covariance matrices Q, R and the alignment distribution τ an optimization problem is formulated as:

$$max_{\tau,Q,R,\mathbf{p}} logP(\mathbf{y}, \mathbf{z}, \tau; Q, R, \mathbf{p}) \tag{4.14}$$

An Expectation-Maximization Algorithm 2 is used to estimate these parameters. Firstly, an Extended Kalman Smoother (EKS) is used to infer the $\mathbf{z}$. Note that while τ is a learned parameter the problem cannot be properly defined without fixing it since the EKS iterations depend on deterministic alignment. We use the proposed uniform multinomial as an initialization value for the probability distribution A. We also assume that the state's components vary independently. Therefore we defined the covariance matrices as diagonal initialized at $Q, R = diag(\epsilon),\ Q, R \in \mathbb{R}^{2\times 2}$.

The smoothed distributions at the end of the Kalman iterations are used to update the covariance matrices, following the procedure:

$$\delta\bar{z}_t = \bar{z}_{t+1|T-1} - f(\bar{z}_{t|T-1}) \tag{4.15}$$

$$A_t = \mathcal{F}(\bar{z}_{t|T-1}) \tag{4.16}$$

$$L_t = \Sigma_{t|t} A_t^\intercal \Sigma_{t+1|t}^{-1} \tag{4.17}$$

$$P_t = \Sigma_{t+1|T-1} - \Sigma_{t+1|T-1} L_t^\intercal A_t^\intercal - A_t L_t \Sigma_{t+1|T-1} \tag{4.18}$$

$$Q = \frac{1}{T} \sum_{t=0}^{T-1} \delta\bar{z}_t \delta\bar{z}_t^\intercal + A_t \Sigma_{t|T-1} A_t^\intercal + P_t \tag{4.19}$$

With $\bar{z}$ as the center of the Gaussian at position t in the HMM chain. $\mathcal{F}$ represents the Jacobian matrix of the f function[2] defined in Eq. (4.13).

$$\delta y_t = y_t - h(\bar{z}_{t|T-1}) \tag{4.20}$$

$$R = \frac{1}{T} \sum_{t=0}^{T-1} \delta y_t \delta y_t^\intercal + \quad \mathcal{H}(\bar{z}_{t|T-1}) \Sigma_{t|T-1} \mathcal{H}(\bar{z}_{t|T-1})^\intercal \tag{4.21}$$

[1]The ξ was set to $2/3$ in our experiment.

[2]The Jacobian matrix of the measurement model is evaluated to: $\mathcal{F} = diag([(1),(2)]), (1) = 3\alpha_{z_1} z_1^2 + 2\beta_{z_1} z_1 + \gamma_{z_1}, (2) = 3\alpha_{z_2} z_2^2 + 2\beta_{z_2} z_2 + \gamma_{z_2}$

$\mathcal{H}$ represents the Jacobian matrix of the h function[3] defined in Eq. (4.11).

We then proceed to use the estimated model to infer τ such that, given the distribution in Eq. (4.12) and fixing the covariance to the previously updated matrices, we have the following:

$$max_{\tau}\mathcal{L}(Q, R, \mathbf{p}, \tau|\mathbf{z}, \mathbf{y}) \tag{4.22}$$

$$max_{\tau}logP(\mathbf{y}, \mathbf{z}; Q, R, \mathbf{p}, \tau) \tag{4.23}$$

$$max_{\tau}logP(\mathbf{y}; \mathbf{z}, \tau)P(\tau) \tag{4.24}$$

$$max_{\tau}[\ell(\mathbf{z}, \tau|\mathbf{y}) + \ell(\tau)] \tag{4.25}$$

for each observation vector $\mathbf{y}^k$ we determine the vector τ that best aligns the observed to the hidden. This process is performed using the subsequent formula:

$$L_{0,t} = logP(y_0; z_{\tau_0}, p)P(\tau_0 = t) \tag{4.26}$$

$$L_{i,t} = logP(y_i; z_{\tau_i}, \tau_i = t) \tag{4.27}$$

$$\max_{t'}[logP(\tau_i; \tau_{i-1} = t')L_{i-1,t'}]$$

$$\text{with } t' \in \{t-3, t-2, t-1\} \tag{4.28}$$

$$\text{and } t' \in [2t-3, 2t+3] \tag{4.29}$$

The Algorithm

The implementation follows the suggestion in [37]. The pseudo-code is listed in Algorithm 2 and we provide implementation specifications in what follows.

Algorithm 2 Patient Trajectory Model Learning

```
for y in observations: do
  Initialize μ_0 = y_0 , P_0 = (10^{-9})I
  Initialize the parameters p_prior = 1/3 and Q, R = εI
  repeat
    Run the Extended Kalman Smoother.
    Update the covariance matrices Q and R for process and measurement noises.
    Update once for each Observation.
    Run the Dynamic Time Warping (DTW) algorithm to find the best τ.
    Estimate p_posterior using maximum likelihood for multinomial distribution.
  until D_KL(p_prior|p_posterior) < ε_τ
end for
```

[3]The Jacobian matrix of the observation model is evaluated to: $\mathcal{H} = I(2)$ an identity matrix in $\mathbb{R}^{2\times 2}$

We start the iteration assuming a uniform multinomial distribution for τ values and an $\epsilon = 10^{-5}$ for the covariance matrices. We start by running the standard EKS Filter using observations $\mathbf{y}$ as measurements of the hidden. After reaching step T we run backward pass that permits the estimation of the smoothed distributions' parameters, $\mathcal{N}(\mu_{t|T-1}, \Sigma_{t|T-1})$ [61].

We execute the Dynamic Time Warping algorithm [143] by maximizing the likelihood function as defined in Eqs. (4.26)-(4.29). The inferred solution is the time alignment vector τ that maximizes the likelihood of the observation given the alignment and the updated distribution.

We update τ to the vector with the higher likelihood values and uses the Kullback-Leiber divergence as a termination test for τ convergence with a threshold $\epsilon_\tau = 10^{-2}$. The convergence can be safely assumed although not optimality, given the stability of the problem statement of maximizing the likelihood function which acts as a Lyapunov function for the EM iterations [142].

4.3.2 Assessment Metrics

Once the HMM model is evaluated for each ET during the session, we use it to define a set of metrics. The detailed definitions and formulation of the metrics are listed in Table 4.1.

These metrics are derived from information theory and bayesian statistics pertaining to fundamental properties of HMM models. For the sake of our research continuous discrete state hidden markov models. We can observe error measures in the Euclidean sense, measures of probability density and variance.

Additionally measures of likelihood are often used in Bayesian statistics and were also investigated in our study. These measures derive mostly off the Expectation-Maximization technique where likelihood estimation was done for spatial and temporal aspects of the trajectories.

4.4 Conclusion

This chapter presented the formulation of the theoretical modeling technique that constitutes the main contribution of the current work. A dynamic model estimated through exercise data is used to provide a process equation of the state space modeling framework.

Metric	Symbol	Definition	Equation	Hypothesis
Probability	cdf	The total probability defined as integral probability at each state of the hidden chain model.	$cdf(y_j) = \int_{-\inf}^{y_j} \mathcal{N}(y_j : \mathbf{z}_{A(j)}, P_{A(j)})dj \quad A : \Phi \rightarrow \Gamma$	Increase
Total Variance	var	The sum of the covariance matrices along the hidden chain.	$V_T = \sum_t P_t, t \in \Phi$	Decrease
M Distance	$\mathcal{M}$	M distance is defined as the sum of the squared root of the standardized differences.	$\mathcal{M} = \sum_j \sqrt{\frac{e_j}{P_{A(j)}}}, e_j = y_j - \mathbf{z}_{A(j)}, j \in (0, \dots, N)$	Decrease
Akaike Information Criterion	$\mathcal{AIC}$	The scaled log of the squared errors.	$L.log(\sum_j e_j^2), j \in (0, \dots, N)$	Decrease
Total Likelihood	$\mathcal{L}_T$	The total likelihood of the hidden model chain.	$\sum_j -\frac{1}{2}.log(\lvert P_{A(j)} \rvert) + e_j^\intercal.P_{A(j)}^{-1}.e_j + 2\pi\ ,\ j \in (0, \dots, N)$	Increase
Alignment Likelihood	$\mathcal{L}_\tau$	The total likelihood of the alignment vector τ.	$\sum_j \ell(j : j-1) + \ell(y_j : \mathbf{z}_{A(j)}, P_{A(j)}, j)$	Decrease
Entropy	$\mathcal{E}$	The total sum of the scaled log determinant of the covariance matrix.	$\sum_t \frac{1}{2} log(2\pi e \lvert P_t \rvert)\ , t \in \Phi$	Decrease

Tab. 4.1.: The Metrics defined based on the HMM model of the ET. The metrics are measures of standardized distance and of probability.

The Hidden Markov Model is presented as the patient trajectory model. The Kalman smoothing technique used an extended generative model, based on the dynamic model estimated in the first step, and is used to fit the trajectory data.

Once fitted the estimation of the time alignment is then computed using the dynamic time warping technique. The maximization step uses the standard update equation for maximum likelihood for both noise covariance and time alignment multinomial distributions. This framework is then used to define a set of measurements that are useful to account for the longitudinal evolution of patients' performance as objective estimates from the data.

Implementation details are also presented alongside the algorithms and their formulations. In the next chapter, we investigate the feasibility and the validity of the measures proposed in this chapter on the operational dataset.

5 The Learning Framework Feasibility as a Valid Modeling & Assessment Instrument

> *Any noticeable discrepancy in the recorded movements of the exoskeleton would have a physiological source at the functional or cognitive level. Thus, we are capturing a meaningful event.*
>
> — **The Chapter Highlight**

The previous chapter laid the theoretical background to our methodology and presented the assessment instruments that we advocate to assess trial-by-trial performance changes. This chapter tackles the validation study results.

Beforehand, the statistical methodology employed, the data preprocessing and the underlying definitions are detailed when applicable. We first introduce the dynamic model and evaluate it on the raw data. The prediction model is then used within the framework formulations to model trajectories and assess per session differences.

The results are then discussed in light of their potential utility. Finally, we state the limitations of the current results and highlight the motivation for the following study.

5.1 Learning The Dynamic Model

5.1.1 Experiment Setup & Methodology

We start by retrieving an exercise file from our CHU-PRC dataset for the task defined in Fig. 1.3 We train the learning Algorithm 1 on the entire normalized exercise dataset (approximately 4000 timestamps) with different combinations of local models. The data was standard normalized before calculations. We then chose the best three models to detail. Firstly, for the $cubmodel$ predictions are performed based on the transitions defined in Eq. (4.8) while for the $sqmodel$ a second-order variation of Eq. (4.8) was used as the prediction equation. Finally, for the $linmodel$ predictions were made using similar transition functions to the one defined in [1]. Algorithm 1 iterations were then updated accordingly to define the appropriate cost functions before optimizing. We then proceeded to validate using the held-out method on an unseen exercise file to evaluate the accuracy of the predictions using error mean and standard deviations for the 20 timestamps starting at each time step t in the data. According to the definition in Eq. 4.13, the most possible number for blind predictions, i.e. without corrections by measurements y_t , during EKS iterations is three timestamps. Accordingly, we detail the performance of the first three steps predictions for the trained models. In terms of execution time, the dynamic model is learned in 15 min up to 1 hour depending on the imposed tolerance.

5.1.2 Experiment Results

Fig. 5.1 shows that the cubmodel defined as local cubic spline centered in z_t achieves better accuracy results. The sqmodel which implemented a parabolic model resulted in the second-best accuracy, while the original linmodel did not present any further accuracy improvements. A detailed evolution of the mean error rate for the first three predictions is shown in Fig. 5.1, where we can note that both linmodel and sqmodel achieve comparable accuracy levels while reaching the third step with slightly better accuracy and consistency for the latter with mean values of $1.53 \times 10^{-3}(2.25 \times 10^{-3}$) and 1.13×10^{-3} (1.85×10^{-3}) respectively. We note that cubmodel achieves lower standard deviation overall which entails better prediction accuracy with better consistency, justifying the proposition.

5.2 Task Trajectory Learning Experiment

5.2.1 Data Preparation

For each session record, we extract the global trajectory on the plane $X - Y$. We then process the global trajectory to extract single elementary trajectories ETs. After segmentation, the ETs are selected based on task completion success. The obtained ETs are then resampled to the same rate to have a similar length per ET.

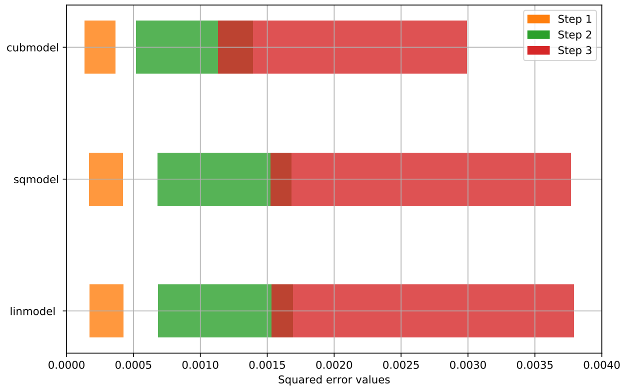


Fig. 5.1.: The broken bars plot represents the mean squared error evolution through the first three prediction steps. The minimum edge of the rectangle coincides with the mean error value of the model, while the length represents the variation of the measure as defined by the standard deviation σ. The graphs show that in addition to better accuracy the cubic model had lower standard deviations.

The common length is set to the median of ETs lengths in a single session. The common length constraint is imposed by the alignment phase where observed trajectories are aligned to a double lengthed HMM. The resulting data are represented in Fig. 5.2.

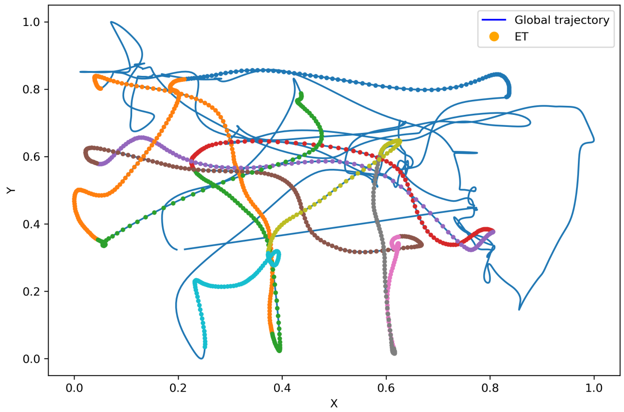


Fig. 5.2.: The training data was a sub-sample of the original exercise recording where we selected for the successful catch trajectories among the trajectories recorded represented in dotted lines against the whole session trajectory.

The resulting trajectories are then translated to the origin and rotated such that each ends on the x-axis. This will help us unify the directions for all the trajectories and provide a baseline for comparison. An example of transformed trajectories is shown in Fig. 5.3 and thus permits us to transform the learning into encoding absolute displacement instead of absolute position.

The software has been developed using Python v3.7 with basic Anaconda modules. Data have been normalized using the Min-Max normalization to have the same scale for the position readings.

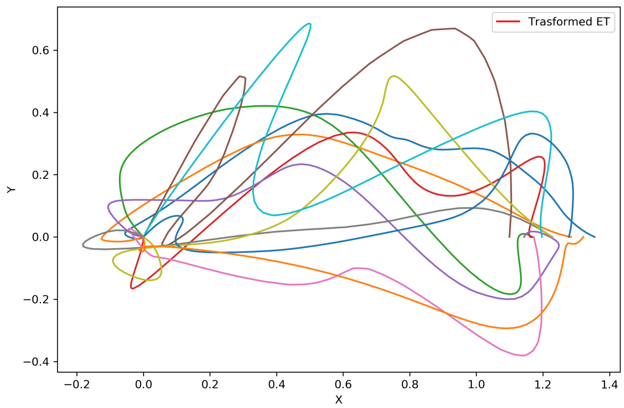


Fig. 5.3.: The selected ETs are then resampled to the same rate to have the same number of timestamps per observation. The data are also unit normalized to ensure the same scale for all observations. ETs are then translated to the origin and rotated to define the movement as absolute displacements in the same direction.

5.2.2 Study Sample

To test the basic functions of the framework, we used data from three subjects. Subject (S1), a 54 years old female, a stroke survivor who appeared to restore a large proportion of her motor function by the end of the therapy. Subject (S2), a 9 years old girl, suffering from infantile hemiplegia and struggled to regain her abilities and sustain them. To compare performance, a healthy subject C1, a 27 years old male, attempted the same care protocol and provided control data used for discriminate tests below.

5.2.3 Properties Definitions & Statistical Analysis

Following the clinical research methodologies, a set of statistical tests are used to assess basic metrics properties. First, the metrics did not prove normally distributed as a result of the Shapiro Wilks test. Consequently, we employed non-parametric tests instead of their parametric counterparts. The median and the Interquartile Range (IQR) were used as descriptive statistics. The statistical significance thresholds α are set to $.05$.

Evolution property Firstly, we study the utility of each metric as an evolution metric. Hence, its ability to measure the evolution of the recovery process in a significant way. To this end, we used a within-subject dependent test for metrics samples at the beginning and the end of the treatment noted *pre* and *post* respectively. The test used is the two-tailed Wilcoxon Signed Rank test which is equivalent to the one-sample t-test. It assumes that the samples are dependent and with similar shapes to their distributions. The dependent variable is assumed to be continuous.

Results of $p - values$ are reported to eleven degrees of freedom $df = 11$. The effect size was defined as $ES = W/n$ where W is the Wilcoxon Signed Rank statistic and n is the maximum sum of ranks. The W statistic is the sum of the positive ranks and is reported as zero whenever there is none. A zero effect size follows a zero statistic by definition. Considering the fixed variability of the values of each metric and the continuous scale, the assumptions of the Wilcoxon Signed Ranks test are satisfied.

Discriminative property Besides, we study the characteristic of a metric as a discriminative measure which refers to its ability to differentiate between different subjects at different stages of recovery. To this end, we used a between-subjects sample test for significance comparisons. The Kruskal Wallis test is used to check the differences between three or more samples. The same assumptions about the shapes of the distributions apply. Results of $p - values$ are reported to $df = 2$. Mutual differences between subjects were done via a post-hoc analysis. An independent sample two-tailed Mann Whitney Ranks Sum test was used to assert the mutual differences. Results of $p - values$ are reported to $df = 11$ and a Bonferroni rule of $(\alpha/3)$% ought to be used when considering for significance. The effect size was defined as $ES = U/n_1 n_2$ where U is the Mann-Whitney test statistic representing the score for the second sample and n_1 and n_2 are the sample sizes.

Predictive property Afterward, we assess the possibility of using the metric as predictive measure, representing the possibility of detecting trends of recovery status as a correlation relationship. To this end, we evaluated Spearman's ρ. The coefficient represents the strength of the monotonic relationship between the number of ETs conducted, as a form of training repetition, and the given metric. Results of $p - values$ were reported to $df = 21$.

Strength of evidence To measure the likelihood of the alternative hypothesis of the non-parametric test we employ the Bayes factor bound (BFB) metric which defines the maximum odds of the alternative being true given the $p - value$ of the test. This measure helps us ascertain the robustness of the conclusion and provide further evidence for the sake of reproducibility. It is defined as follows: $BFB = -1/eplnp$. All statistical tests were calculated using custom Python scripts using implementations of Scipy v1.3.2.

5.2.4 Task Model Learning Results

We proceed to establish a task model using the modified dynamic model defined in Eq. 4.13 augmented with a bias term ζ. In Fig. 5.4(a) we note the beginning of the iterations where the inferred model is too errored and the alignment is not properly maintained with a respectively higher accumulation of error along the trajectory. The iterative process gives us a proper estimation of the noise covariances and better alignment sequences which are observed in Fig. 5.4(b) where alignment is achieved by the end of the iterations. The maximum likelihood alignment vector τ is chosen, this is matched with a decrease in the errors accumulated in the model.

A closer examination of the aligned demonstrations and the resulting model is shown in zoomed Fig. 5.4(a) where we see the beginning of the iterations with higher errors at the observations y_t . The final iteration in Fig. 5.4(b) shows us how the converged τ better fits the demonstrations and the model all while keeping lower error values compared to early iterations. The 95% confidence interval around the smoothed trajectories does not show any particular accumulation of error, neither does the blind predictions in the trajectory. The smoothing procedure helps towards better and consistent modeling and thus prevents convergence to biased models at the end of learning iterations. The bias vector shifts in position with the same amplitude in coherence with its definition and has a maximum amplitude of 4% of the amplitude of the normalized measurement. Generation of a single report file takes 1 minute 33 seconds which is fairly reasonable for the end of session report. The very nature of the metrics definition involving measurements from a full episode (i.e. an ET), limits to an extent a fully online use of the modeling framework without further modifications and experimentations.

5.2.5 The Statistical Analysis Results

Modeling the patient's behavior as a means of extracting the underlying quality and performance of the rehabilitation exercise was presented in different forms and many studies proposed methodologies to proceed to that end.

Subject	Metric		*pre* statistics median(iqr)	*post* statistics median(iqr)	W statistic	$p-value$	Effect size	BFB
S1	AIC	z_1	-1445(73.1)	-1083(59.4)	0	.002	.0	27
		z_2	-1478(57.4)	-1116(51.4)	0	.002	.0	27
	L_τ		-4139(9138)	-8039(2576)	25	.27	.32	1
	L_T		-88.3(2.39)	-64.9(1.41)	0	.002	.0	27
	M		130.5(8.75)	96.2(7.23)	0	.002	.0	27
	E		-2056(33.8)	-1499(17)	0	.002	.0	27
S2	AIC	z_1	-1271(77.6)	-1238(60.5)	19	.11	.24	1
		z_2	-1310(51.8)	-1266(87.1)	15	.059	.19	2
	L_τ		-8291(4914)	-11807(6827)	22	.18	.28	1
	L_T		-78.7(2.18)	-74.9(2.5)	0	.002	.0	27
	M		118 (8.28)	112(9.72)	23	.21	.29	1
	E		-1813(18.7)	-1729(31.2)	2	.004	.026	27

Tab. 5.1.: A comparison between descriptive statistics for *pre* and *post*-treatment data shows similar trends between subjects. A Wilcoxon Ranks Sum Tests for *pre* and *post*-treatment differences in metrics' values detected many significant results. Nonsignificance is reported for both subjects for likelihood alignment L_τ, this is due to the high value of the IQR reported for all its samples. High odds ratio always favors the alternative and low effect sizes are still reported for the non-significant differences.

Our objective in this study was to characterize the evolution of the patient during the rehabilitation exercises through the study of the evolution of the model of their trajectories. The modeling approach listed above evaluates a Bayesian model of each trajectory execution and thus permits us to define a set of new measures that we studied throughout the training.

The properties tested, namely the evolutive and predictive properties of the metrics, are for testing the evolution of the recovery over time and this is not the case of a healthy user whom we used as a comparative ideal for attained recovery. As such, the tests for both properties did not include results for subject C1.

5.2.6 Wilcoxon Ranks Sum Test

Observing the results of Table 5.1 shows that increase in median values for metrics AIC, E, L_T and the decrease in the median of L_τ, M were consistent between the two subjects. IQR values were of the same order of magnitude for the same metric between *pre* and *post* samples and between samples of the two subjects.

To assess the hypothesis that the metrics measurements evaluated based on ETs from *pre* and *post*-treatment samples evolve significantly, and are thus capable of detecting differences between the performance levels of the patients, we used the Wilcoxon Signed Rank test. Differences between values of the entropy E and the total likelihood L_T were significant for subject S1.

An odds ratio of 27:1 in favor of the alternative reinforced this result. Despite having non-significant statistics, the other metrics reported a moderate effect size of $.2$.

For subject S2 all metrics showed statistically significant differences between values at *pre* and *post*-treatment assessments. Significant differences had a 27:1 odds ratio in favor of the alternative hypothesis. The difference between the values of alignment likelihood L_τ was insignificant, with a moderate effect size of $.32$.

The results suggest that the metrics did not reach significant results for subject S2 simply because the process of recovery is still underway for that particular patient.

This can be further justified if we examine the magnitude of differences between the median values *pre* and *post*-treatment for subject S1 that are twice the magnitude of the measure for subject S2. A further argument for this statement is the evaluation of the effect size that shows a small effect size for the metrics despite not reaching the statistically significant threshold.

The results for the alignment likelihood were consistently insignificant, a result that can be related to the high values of IQR which suggests the high dispersion in values. This dispersion makes the detection of a clear difference unfeasible and thus results in insignificant statistics.

5.2.7 Kruskal Wallis test

Evaluation of the between patients' differences can help distinguish notable differences between their performances. We studied the differences as observed through the Kruskal Wallis test and reported its results in Table 5.2. The test shows

		C1/S1/S2 *pre* data			C1/S1/S2 *post* data	
Metric	H statistic	$p-value$	BFB	H statistic	$p-value$	BFB
AIC z_1	23.4	8×10^{-6}	3866	28.1	8×10^{-7}	32495
AIC z_2	27	1×10^{-6}	19480	30.3	3×10^{-7}	91402
L_τ	7.23	.027	135	17	.0002	222
L_T	25.1	4×10^{-6}	8220	31.1	2×10^{-7}	136268
M	13.7	.001	51	22.8	2×10^{-5}	1801
E	31.1	2×10^{-7}	136268	31.1	2×10^{-7}	136268

Tab. 5.2.: Investigating the differences between the three subjects at the start and end of treatment was conducted using the Kruskal Wallis test. The significant results are shown for all metrics and are reported with high odd ratios as evidence for the alternative. The results show that the metrics are significantly different between at least two of the subjects.

		Patients											
		S1/S2				C1/S2				C1/S1			
Metric		U statistic	$p-value$	Effect size	BFB	U statistic	$p-value$	Effect size	BFB	U statistic	$p-value$	Effect size	BFB
AIC	*pre* z_1	142	$6{\times}10^{-5}$	.99	630	105	.06	.73	2	3	$8{\times}10^{-5}$	.02	507
	post z_1	0	$4{\times}10^{-5}$	.0	984	128	.001	.89	41	144	$4{\times}10^{-5}$	1.0	984
	pre z_2	144	$4{\times}10^{-5}$	1.0	984	121	.005	.84	14	0	$4{\times}10^{-5}$	.0	984
	post z_2	3	$8{\times}10^{-5}$	.02	507	143	$5{\times}10^{-5}$	.99	786	144	$4{\times}10^{-5}$	1.0	984
L_τ	*pre*	46	.14	.32	1	19	.002	.13	25	74	.93	.51	6
	post	43	.1	.3	2	9	.0003	.063	148	18	.002	.13	29
L_T	*pre*	144	$4{\times}10^{-5}$	1.0	984	106	.05	.74	2	0	$4{\times}10^{-5}$	.0	984
	post	0	$4{\times}10^{-5}$	.0	984	144	$4{\times}10^{-5}$	1.0	984	144	$4{\times}10^{-5}$	1.0	984
M	*pre*	17	.002	.12	35	64	.67	.44	1	127	.001	.88	35
	post	132	.0006	.92	84	29	.014	.20	6	4	$1{\times}10^{-4}$	.03	409
E	*pre*	144	$4{\times}10^{-5}$	.0	984	144	$4{\times}10^{-5}$	1.0	984	0	$4{\times}10^{-5}$	.0	984
	post	0	$4{\times}10^{-5}$	.0	984	144	$4{\times}10^{-5}$	1.0	984	144	$4{\times}10^{-5}$	1.0	984

Tab. 5.3.: Mann Whitney Test results assert that the metrics defined are capable of discriminating against the different subjects both at the onset of training as well as at the end. This can be seen in the significant results reported for comparisons between the three subjects. Large effect sizes and odd ratios also accompanied the results testifying the strength of the evidence reported.

that the differences between the metrics for *pre* and *post*-treatment samples between the subjects were significant. This was reinforced with extreme odds reported in favor of the alternative.

5.2.8 Mann Whitney Ranks Sum test

Further analysis of these results using mutual tests or Mann Whitney is presented in Table 5.3. Comparing results for subject S2 to healthy control C1 gives significant results at the Bonferroni corrected threshold for all metrics except for L_T, M, and AIC_{z_1} all being insignificant for *pre*-treatment samples comparisons. Despite the important $p-values$, moderate to large effect sizes were reported $.74$, $.44$ and $.73$ respectively.

Similarly comparing subject S1 to the healthy control showed significant results for all metrics except the alignment likelihood L_τ. Nevertheless, the metric presented a large effect size of $.5$.

The results show that the proposed metrics can differentiate between the patients at the onset and the end of the treatment. Besides, it proves that the metrics are capable of characterizing the differences between the subjects at all stages of recovery.

A result that can be very promising as it can be linked to the differences in treatment response that each patient will exhibit and that are otherwise undetectable. Whilst some metrics were reported to have insignificant statistics, the same conclusions may still be pronounced considering the reported high to moderate effect size.

5.2.9 Spearman's ρ correlation

Further examining the evolution of each metric between the start and the end of the training can be seen in Table 5.4 which detailed the correlation coefficients. These coefficients relay the strength of the monotonic relationship between the metric and the exercise repetitions, represented by the number of ETs accomplished.

Examining the results we can note that for the metrics proposed the coefficients reported for patient S1 are large effect sizes relating to the clear trend between the values of the metric at the beginning and the end of the training. These results are in line with the findings of the Wilcoxon tests. This also provides further details to the nature of the incremental relationship between the metric and the evolution of training repetitions. The examination of the signs of the coefficient shows that similar trends are present for all the metrics and both subjects despite not reaching significance for patient S2.

In addition to examining the trends of the metrics, the underlying structure of the model is represented by the different parameters that we estimate through the optimization techniques of the trajectory modeling Algorithm 2. The observations that we can note on the value of the model parameters are a significant decrease in the total probability of the model and a decrease in the total variance.

The model noise variance is also increasing with a medium effect size. Thus, We can detect changes at the level of the model that is characteristic of the changes in the structure of the trajectories that the patient exhibits throughout his training. This is a very compelling result as the intuition of the proposition of the model is the ability to capture the changes in the executions of the patient at different stages of recovery.

		Patients					
		S2			S1		
Metric		ρ statistic	$p-value$	BFB	ρ statistic	$p-value$	BFB
AIC	z_1	.13	.48	1	.77	2×10^{-8}	859964
	z_2	.46	.01	7	.75	1×10^{-7}	207282
L_τ		-.5	.006	12	-.32	.05	2
L_T		.66	9×10^{-5}	458	.72	6×10^{-7}	44415
M		-.35	.06	2	-.7	1×10^{-6}	21528
E		.64	.0002	231	.79	5×10^{-9}	3754432
cdf_{z_1}		-.15	.44	1	-.7	2×10^{-6}	15531
cdf_{z_2}		.13	.5	1	-.73	3×10^{-7}	76180
var_{z_1}		.24	.22	1	-.37	.024	4
var_{z_2}		.01	.94	7	-.46	.004	16
Q_1		.44	.016	6	.53	.0007	72

Tab. 5.4.: The evolution of the metrics is tested using Spearman's rank. The test quantifies the existence and the strength of the monotonic variation of the metric with additional training. The significance of results was not consistent between the two subjects, whereas, coherent signs of correlations were reported. The model parameter Q_1 showed a significant increase for subjects and with moderate effect sizes asserting that the model adapts to the changes in the demonstrated dynamics of the movement while encoding for the trajectories.

Detecting a clear trend on the parameters of the model indicates the ability of the proposed framework to encode for those differences in these parameters through the optimization in Algorithm 2 and the estimatation of their values.

To assess the strength in the monotonic evolution of the metric between *pre* and *post*-phases we use Spearman's rank. Results are reported in Table 5.4. For subject S2 we note three metrics AIC_{z_2}, L_T, E that significantly increased with moderate to large effect sizes and with moderate odds for the alternative. L_τ showed a significant decreasing trend with a large effect size and low odds ratio. A single model parameter Q_1 showed a significant increasing trend with moderate effect size and low odds ratio.

For subject S1, metricsAIC, L_T, E showed a significant increasing trend with large effect sizes and extremely likely odds in favor of the alternative. Metric M showed an increasing tendency with large effect sizes and high odds as well. Model parameters all decreased significantly at the exception of Q_1 that showed instead an increasing tendency. Results were associated with a large effect size and extreme to likely odds in favor of the alternative.

5.2.10 Responsivity of the metrics

To further support this claim we carried a longitudinal analysis of the evolution of the metric values through the course of training for subject S1. We plotted the results of the $p - value$ of the Wilcoxon test between the metric values at the beginning of the training and the next session completed by the patient that resulted in Fig. 5.6.

From these results, we can say that the metrics proposed can characterize differences in the status of recovery as early as after three sessions. These results can be highly beneficial in applications where the need for close follow-up is strictly imposed, for instance in the telerehabilitation applications. On the other hand, the discriminative property is also present although the fluctuations may still leave several open questions to research as we can see in Fig. 5.7.

The measure of the entropy of the model presented the most interesting results between the metrics as it is capable of significantly comparing the differences between the samples from different subjects regardless of the stage of recovery. The AIC presents another interesting behavior where, by the end of the training, it can detect no differences between the healthy user and the recovered patient. In the meanwhile, it still showed differences present between the recovered and the non-recovered patient. At the beginning of the training though it shows differences between all three patients.

Examining the results of Fig. 5.6, we can note that for all the metrics the $p-values$ have a decreasing tendency towards the significance threshold. The metrics on the exception of the alignment likelihood are detecting significant differences between the samples as soon as the third session since the start of the rehabilitation training.

Closely following the evolution of the $p - values$ for the Mann Whitney test as depicted in Fig. 5.7 demonstrates that the entropy measure can distinguish subjects throughout the course of the training. The likelihood alignment does not have significant differences between the subject and the healthy user, while it presents some significant results at the onset of the rehabilitation between the two patients.

The AIC measures are fluctuating throughout the training. The results by the end of treatment show that it is unable to detect significant differences between the patient and the healthy user by the end of the training. At the same phase differences with the other patient are significant which denotes the recovery has progressed significantly for the subject S1 compared to S2.

The M distance also presents fluctuations and is switching between significant and nonsignificant results throughout the training. The same remarks can be observed on the plots of the total likelihood values which are significant at the start and the end while fluctuating throughout the training.

5.3 Experiments' Results Discussion

5.3.1 The utility

The rehabilitation exercises can be more effectively and accurately assessed using our proposed framework and the defined metrics. The new resulting metrics present strong evolutive, discriminative, and predictive properties that are key to assessment adoption for clinical validation and practice.

We modeled the trajectories of the end effector during rehabilitation exercises using an Armeo Spring exoskeleton using an EM procedure. The resulting model permitted the definition of several metrics. The new metrics proposed to serve as a finer assessment measure of the patient's recovery. The new methodology permits detecting changes after 3 sessions with significant results for a number of the proposed metrics.

A major concern for any assessment technique is its sensitivity to the measured construct i.e. the recovery level. By showing a finer responsiveness property, the proposed framework promises to enable researchers to establish a much finer objectively evaluated kinematic-based scale that correlates with the recovery process.

A strong discriminative nature for many of the resulting metrics would permit a proper classification of the patients into performance groups. By capitalizing on the correlation properties of the metrics, a narrow horizon future prediction of the recovery outcome might be feasible.

These findings promise to extend the kinematic approach that has been commonly used by researchers to evaluate upper limb recovery, specifically when using exoskeleton devices. Some studies have stretched similar approaches further by proposing training to therapists using such models to ensure safe execution and reproducibility of correct exercises [144].

Another interesting use case has been offered in the context of modeling for diagnosis of Parkinson [169]. A challenge that researchers often face is the inability to detect significance in recovery as measured by the traditional kinematic properties as we have shown in our clinimetric study on the same dataset (refer to chapter 3).

Similar issues can be due to the small sample size and the heterogeneity of the population among others. The detection of longitudinal evolution by providing a finer assessment scale would enable researchers to study more precisely the morphological or neurological processes associated with the recovery while enabling the possibility of parameterizing many subtle differences. Thus, enabling the possibility to accurately associate the phenomenon happening at the different levels and draw conclusions about causality or the underlying mechanisms of recovery.

In considering the underlying experimental unit as defined by the elementary trajectory, potential portability is open to any other application where the analysis of trajectories is of interest, for classification purposes for example. Meanwhile, exploring other trajectory capturing devices might prove useful. This is especially true When we know the multitude of the devices reported in the kinematic rehabilitation assessment.

Comparative studies that investigate the different approaches to physical rehabilitation interventions can benefit from the use of quantitative measures that are capable of detecting the small effects more reliably.

In subsequent research, we would aim to conduct the study of the clinimetric properties of the metrics that we derived. This study will provide us with the properties and characteristics of each metric and its best use cases. Clinimetric characterization of the metrics should also help move them forward towards clinical adoption by asserting their validity.

5.3.2 The shortcomings

No experiment is perfect and we can acknowledge some limitations for the results that we presented. A case study of three subjects aimed to explore the feasibility of the tools proposed is far from providing enough evidence to establish the proposed framework. Nevertheless, the random nature of the framework and the heterogeneity of the data used can stand arguments in favor of the results provided.

The updated transition function while presenting relatively better accuracy has a higher iteration execution time. In the case of online deployment for a step-by-step prediction the higher accuracy model is encouraged since this training is to be done apriori.

As for any learning technique, the model derived can only be as good as the data used to train it, and thus a good quality sensory reading from the exoskeleton is our baseline for all the models built and we are bound to the accuracy that it provides while detailing results of this study.

The data derived represent exactly the exoskeleton motions and does not necessarily represent the actual arm joint movements. Although there is not a direct relationship between the arm and exoskeleton angular displacement, the resultant movement still mirrors some of the degrees of freedom of the arm joints and the end effector is submitted to the full control of the human controller and driver.

Arguably, any noticeable discrepancy in the recorded movements of the exoskeleton would have a physiological source at the functional or cognitive level and we are thus capturing a meaningful event.

The re-sampling of the data into similar lengths results in unwanted data changes and thus providing a more independent way to update the model would be necessary to guarantee reliability without the need to introduce artificial measures into the data.

The normalization of the data while contributing to the stability of the algorithm might present the challenge when testing on newer data where coordinates maximum values are not known beforehand and thus might affect the final inferred model.

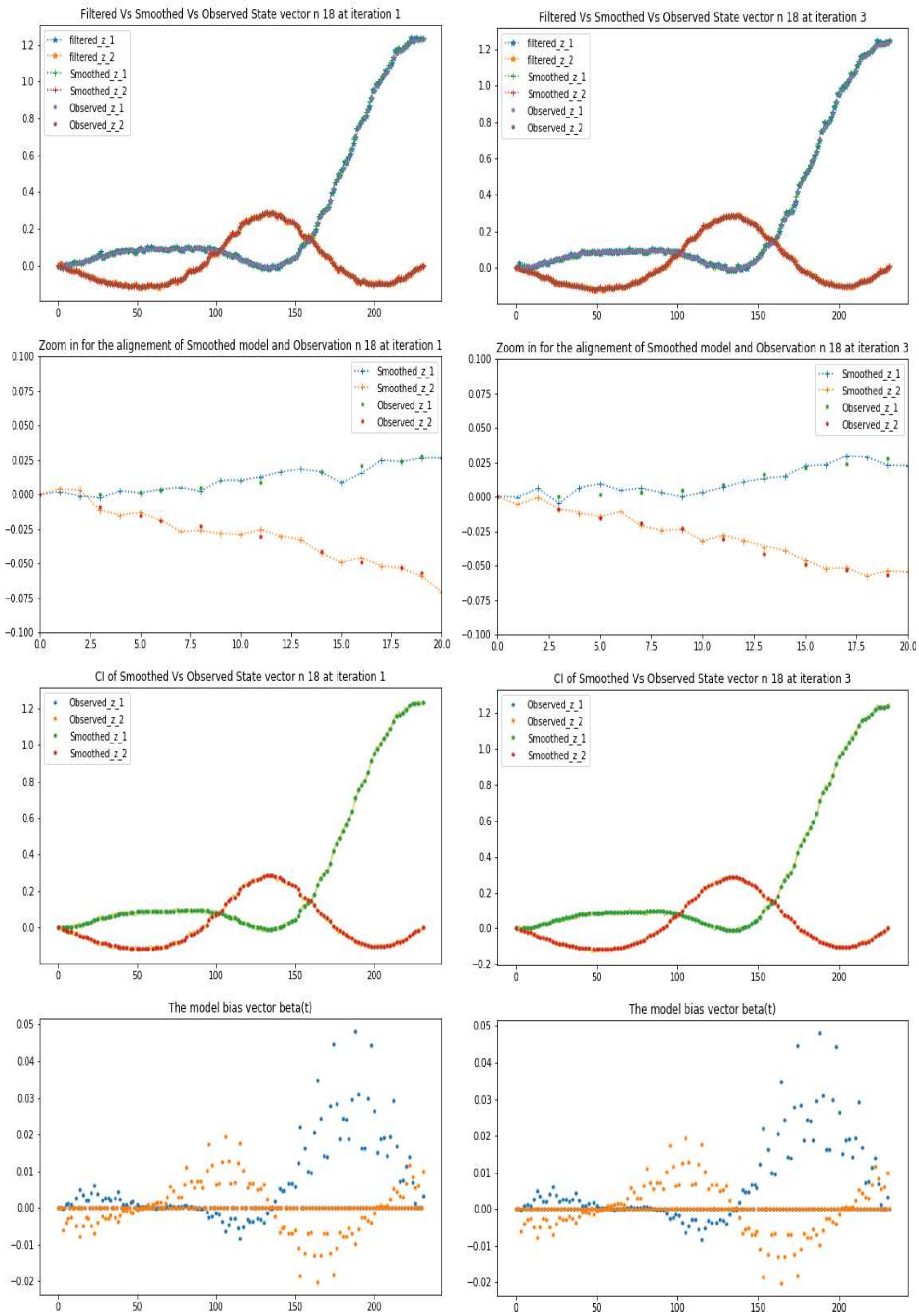
5.4 Conclusion

The reliability and clinical conformity of an individualized objective assessment toolset were presented in this chapter detailing the result of two experiments. First, we explored the feasibility of the proposed framework and the derived metrics as a tool of assessment. Evolutive, discriminative, and predictive properties were assessed for each metric.

The results also presented a high responsivity and longitudinal sensitivity for within-subject between-session evaluation of the motor relearning and the physical recovery of the patients. Measures are calculated automatically with computer programs and are based on robot assessment with zero patient or rater intervention making it an objective reliable score.

We showed significant results between successive sessions on patients' data and as late as three sessions into training. We are aiming through this work to further advance this framework for clinical use by providing more evidence from clinical data.

Follow-up and close monitoring of rehabilitation can serve as a crucial tool in the toolbox of physical therapy practitioners as well as for movement assessment and modeling researchers.



(a) Beginning of the iterations on the 18^{th} ET (b) Convergence of the iterations for the 18^{th} ET

Fig. 5.4.: An example iterations over a single ET where we note that the algorithm starts with the standard alignment and error rates. The algorithm then reiterates over the observation and estimates a better alignment vector. The new alignment can be noted on the zoomed plots where observations represented in dots shifted positions between the start and end of iterations. The error on the resulting smoothed trajectories becomes less apparent as an effect of better matching of the observations to the hidden model.

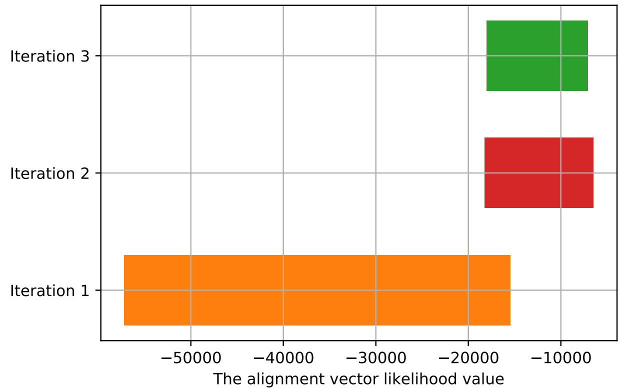


Fig. 5.5.: The evolution of the total likelihood as estimated by the alignment EM iteration, the two extremities of each box represents the min and max values for all ETs. The values evolution between different iteration steps is depicted. We can note the clear increase in likelihood values for the observations given the alignment after the first iteration.

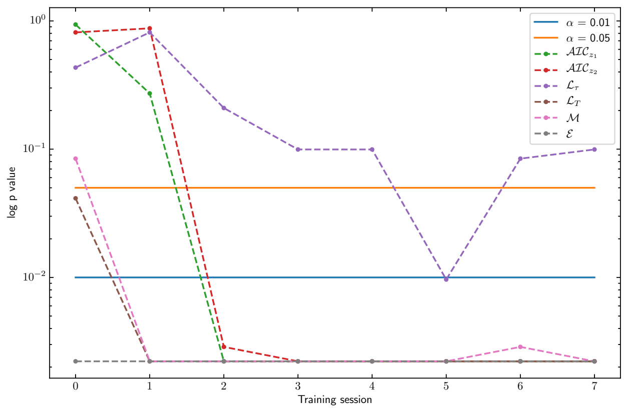
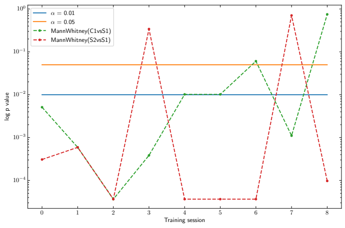
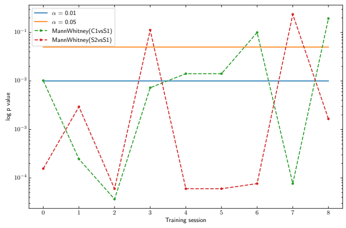


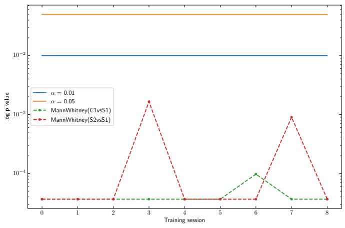
Fig. 5.6.: The evolution of the $p-values$ for the statistical Wilcoxon test for baseline metric measurements and those after each additional session of training for patient S1. α represent statistical significance thresholds. The results show that except the L_τ all the metrics are capable of detecting significant differences between the performance for the post-stroke subject after one or two additional sessions. The results are moderately to strongly in favor of the alternative considering the BFB associated with the $p-values$ below 10^{-3}.



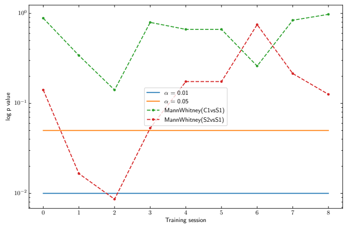
(a) Results for metric AIC_{z_1}



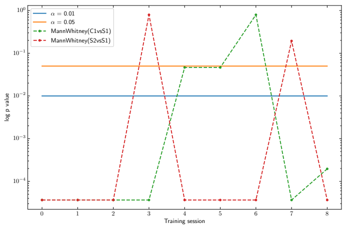
(b) Results for metric AIC_{z_2}



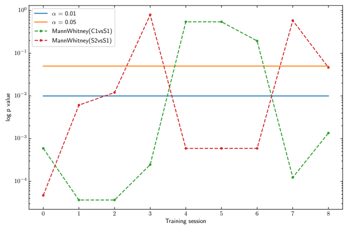
(c) Results for metric E



(d) Results for metric L_τ



(e) Results for metric L_T



(f) Results for metric M

Fig. 5.7.: Plots of the log $p - value$ of the Mann Whitney tests for mutual differences between subjects C1, S1, and S2. α represent statistical significance thresholds. The discriminative power of each metric is viewed in terms of the evolution of the $p - value$ of the between-subject Mann Whitney test. The results show that the entropy measure reports significant differences consistently. The $p - value$ fluctuates near the significance threshold for the other metrics between the start and end of the rehabilitation. AIC metric seems to reach insignificance of differences between subject S1 and healthy control, suggesting a good metric for end of rehabilitation application.

6 Clinimetrics of the Assessment Measures: A Retrospective Case Series Report

> *If we note that the metric Q is the contribution of the dynamic model and is learned throught the Expectation Maximization loop, we can be highly confident that the learning framework is providing a dynamic score of the physical rehabilitation of the upper limb.*
>
> — **The Chapter Highlight**

In the previous chapter, we presented the feasibility of the modeling technique and its utility to detect session-level performance differences. This chapter presents a validation experiment of these findings by extending our cases to a cohort of patients.

We discuss the results in light of the clinical utility of these tools. The assessment of the clinimetrics of these instruments is performed to qualify the measures and assess the evidence on held-out data. Moreover, the Critical Appraisal Skills Program's (CASP) checklist for a diagnostic study is used to discuss the findings from a clinical perspective.

6.1 Methodology

6.1.1 Patient Cohort

The observational study was conducted based on our Tlemcen CHU-PRC dataset. A wide range of pathologies has been served using the Armeo device. To establish this study, the subject inclusion criteria were: the availability of assessment task history records, and the availability of the associated discharge Fugl-Meyer Upper Extremity (FMA-UE) score [65]. Subjects who have met our criteria in addition to a healthy control were included and are listed in Table 6.1.

Tab. 6.1.: The cohort of subjects included in the study.

Subject	Age	Gender	Affected Arm	FMA-UE	Pathology
C1	27	Male	Right	N/A	Healthy
C2	9	Female	Right	N/A	Infantile hemiplegia
S1	45	Female	Right	54	Post stroke
S2	15	Male	Left	44	Post stroke
S3	28	Female	Right	47	Chronic inflammatory demyelinating polyneuropathy
S4	30	Female	Right	48	Chronic inflammatory demyelinating polyneuropathy
S5	32	Male	Right	34	Post Traumatic tetraparesis
S6	54	Female	Right	44	Post stroke

6.1.2 Statistical analysis

Exploratory statistical tests due to non-conformity to the normality assumptions, based on the Shapiro Wilks test, non-parametric tests are used. All tests were examined for their prior assumptions before their use. Firstly, evolutive property, the discriminative property, predictive property and Bayesian statistics are defined similarly to the previous chapter (refer to section 5.2.3).

The motor skill performance curve It is known that the performance of the motor skills is reflective to some degree of the motor learning phenomena. To draw the performance curve a second-order regression model was established to capture the dynamic of the evolution of each metric throughout the course of rehabilitation. The Ordinary Least Squares technique derived the fitting model. A visualization "reg-chart" was introduced to condense the regression results into an accessible visual form relevant to the skill performance curve.

Clinical convergent validity A linear regression model was established based on the mean metric value evaluated at the last assessment session in relation to the Fugl-Meyer Upper Extremity (FMA-UE) score [65].

Factor analysis A Principle Component Analysis (PCA) was used to determine the number of potential factors in the defined set of assessment metrics. An exploratory factor analysis was then evaluated followed by an Oblimin rotation of resulting factors.

6.1.3 Clinical evidence quality

Critical Appraisals Skills Program CASP [26] provides a detailed checklist to assess a diagnostic study. We are evaluating our study according to the items listed in the checklist and results are detailed in the discussion as a preamble to the conclusions derived based on the results. This consists of a step towards communicating engineering results with clinical researchers in the expectation of easier adoption into clinical trial research. The checklist items are provided in Annex C.

6.2 Evaluation Results From Operational Data

6.2.1 Are the metrics clinically valid ?

Firstly, the construct validity is investigated using correlation assessment between the instruments proposed, i.e. the numerical metrics, and the clinical golden standard FMA-UE. The results of the correlation analysis are reported in Table 6.2 Meanwhile, none of the results reached statistical significance, which is logically justified given the limited sample number (six FMA administrations).

The correlations coefficients results showed that metrics AIC, M, E, L_T and cdf_{z_1} correlated moderately with the FMA score, while cdf_{z_2} and var showed near moderate effect size. The linear model is representative of no more than 20% of the variation of the relationship which is considered low. The most intriguing finding is the fact that Q which represents the contribution of the proposed learning technique, is shown to have a large correlation and a moderate $r - squared$.

Tab. 6.2.: Table of correlation with the FMA-UE

Metrics	ρ	$r - squared$ (%)	$p - value$
AIC_{z_1}	.45	20.44	.37
AIC_{z_2}	.47	22.1	.35
L_τ	.29	8.0	.58
L_T	.43	18.6	.39
M	-.42	17.68	.41
E	.44	19.45	.38
cdf_{z_1}	-.44	19.19	.39
cdf_{z_2}	-.38	14.26	.46
var_{z_1}	-.39	15.00	.45
var_{z_2}	-.34	11.25	.52
Q	.61	36.66	.23

Despite the insignificant $p - values$ the observation of the graph of the results displayed in Fig. 6.1 shows a clear tendency on the exception of L_τ.

6.2.2 Content Validity

To interrogate the content validity of the measures, a PCA was conducted based on the calculations of the correlations matrix R since we are looking only for components based on the correlation structure in the variables. Besides, the interpretations of the covariance-based components might be biased considering the disparity in the

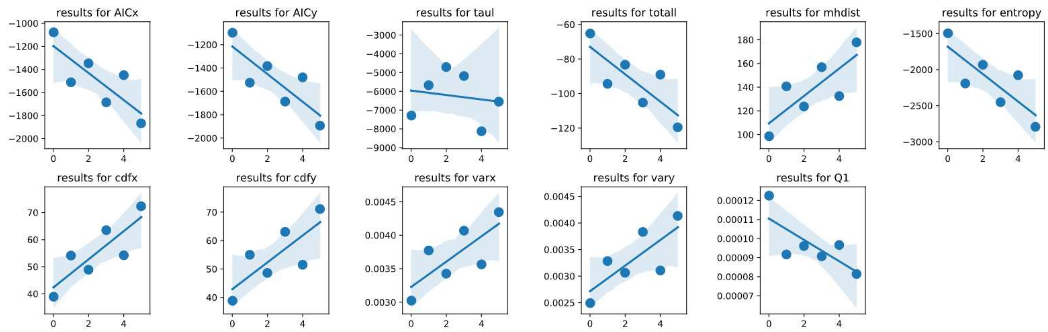


Fig. 6.1.: The regression plots of each metric shows that the variation of the FMA-UE score is represented by the regression plot closely for most of the variables. The observation comes as an illustration of the moderate effect size. The sample size can safely be said to be the only detriment to reaching significance.

variances of the columns. The PCA analysis resulted in two major components after exclusion of the L_T which showed no correlation with other metrics. Noting the results from the PCA analysis, we consider a two factors model might be a legitimate choice for our Exploratory Factor Analysis.

Tab. 6.3.: Table representing the Oblimin rotated loadings for the metrics and according to the determined two factors.

Metric	Factor 0	Factor 1
$AICx_{z_1}$	**0.9597**	0.1846
$AICx_{z_2}$	**0.9217**	0.2608
L_T	**0.9901**	-0.1333
M	**-0.9285**	0.2167
E	**0.9993**	0.0603
cdf_{z_1}	**-0.8870**	0.1054
cdf_{z_2}	**-0.8742**	0.0741
var_{z_1}	-0.1188	**0.9983**
var_{z_2}	-0.0311	**0.9357**
Q	0.1537	**0.9640**

The factor analysis reveals two main factors among the defined metrics. Results showed in Table 6.3 clusters the metrics var and Q together into the second factor.

Upon closely examining the properties and definitions of the metrics it can be noted that these three values represent variance measures, thus, they are grouped into a single category that encodes for the confidence in the evaluation of the position.

The metrics AIC, M, E, L_T and cdf_{z_2} represent, according to their definitions, measures of error either squared, scalar or standardized. These metrics encode for the error in the distance between the model and the observed state.

The analysis highlights this categorization and points it through the high loadings that were determined using the Oblimin rotation. The factors coincide with the World Health Organization International Classification of Functioning, Disability, and Health (ICF), Body and Function domain of assessment of the movement where variability represents the smoothness of the movement and errors metrics represent the accuracy of the achieved movement. Considering convergent validity and content validity results, we stipulate that the validity of the metrics is moderately demonstrated although not conclusive.

6.2.3 Are the new instruments responsive ?

Evaluating the responsiveness of the proposed instruments, a Spearman correlation rank between metric values at *pre* and *post*-rehabilitation provided significant results and moderate to large correlation for most cases except the metric L_τ and subjects C2 and S4. These results show a strong monotonic relationship between the metric and rehabilitation progress. This denotes a good predictive property for the metrics. To further quantify the responsiveness, a second-order regression model was fitted to the longitudinal data for each patient.

Results of the regression analysis are condensed into Fig. 6.2. The results of L_τ were not significant for any subject so it was excluded from the analysis. Subjects C2, S4, S2 and S1 had significant fits for metrics AIC, M, E and L_T cdf_{z_2} while no decisive result can be noticed for the concavity or convexity of the parabola since coefficients signs are not consistent. The associated $r - squared$ ranged from large to moderate for subjects S1, S2, S3, and S5 for the same metrics.

Model metrics cdf_{z_1} and var did not seem to follow a 2^{nd} order model and showed insignificant fits for all subjects except S4. R^2 for the resulting model was low to moderate for subjects S1 and S2. Metric Q presented two significant fits for subjects S1 and S4 while having low $r - squared$ values. Responsiveness is suggested for the metrics AIC, M, E and L_T to be following a second-order model. Meanwhile, for the rest of the metrics, the results suggest that a linear relationship would best describe their evolution.

6.2.4 Can we detect session-level differences in performances?

The property of a clinical evaluation instrument is assessed using the Wilcoxon Signed-Rank test where results indicate the significant differences in the values between the first session of training and the following ones. The results presented in Fig. 6.3 show that the metrics present significant capability of detecting changes over the course of training on the exception of L_τ which had only occasional significance. The results also present the instruments' high sensitivity and thus their strong evaluative property.

6.2.5 Metrics as a discriminative instrument

After carrying the Kruskal Wallis test on packed metric vectors for each of the subjects and those of C1 and C2 from *pre* and *post*-treatment sessions, results were significant in favor of the alternative that at least two of the vectors differed. We, therefore, proceeded to evaluate the Mann-Whitney results for mutual differences. A *pre-post* comparison for all six subjects was significant overall and had strong evidence of the alternative. We then followed by a longitudinal study of the evolution of the metrics for four subjects S1, S3, S4, and C2, who have completed more than 2 assessment sessions during their training.

Results are reported in Figures Figures 6.4-6.7. The L_τ metric presented some significant discrimination specifically for subjects S3 and S4 with no clear tendency. The AIC results showed symmetric tendency for discrimination between subject-healthy and subject-non-recovered tests where we can note at the middle of the training a significant test is reported for both comparisons. The E metric seems to discriminate healthy-subject successfully while failing to do so for the non-recovered-S4 tests.

On the other hand, M and L_T discriminate C1 and failed C2 comparisons for a single occasion for subjects S5 and C2. Results of the same metrics on S1 showed a symmetric behavior where insignificance is exchanged between the two cases. The results summarize a strong sensitivity and discriminative power for the established metrics, and although no clear tendency is noted, the property is highly present in the reported results.

6.3 Discussion

6.3.1 CASP's diagnostic study checklist

The proposed learning technique models the Cartesian trajectories achieved by the patients during their upper limb rehabilitation exercise and provides several numerical and visual instruments to evaluate the evolution of the recovery process. According to our results, the numerical metrics presented good to moderate clinimetrics properties, suggesting the feasibility of the instruments as evaluation measures.

The accuracy of the measurement allows detection of significant changes after at most three sessions into the training. The validation of the metrics has been approached using the operational data from a Rehabilitation Center at the University Hospital of Tlemcen in Algeria and based on data from the Armeo Spring exoskeleton. This data illustrates a real-world use case scenario of the developed tool and suggested the high likelihood of successfully measuring recovery progress. To compare the results a clinical golden standard was our baseline, it represented the FMA-UE scores at care discharge for a cohort of subjects with Acquired or Traumatic Brain (ABI/TBI) injuries.

A strong finding of the study is that the measures seemed to not be affected by the underlying pathology and we have achieved significance despite the differences in population. The proposed measures are not affected by our apriori knowledge of the FMA-UE significant trends since they are evaluated objectively through an optimization procedure.

The measures were calculated using the same thresholds and initial values to provide a common ground for comparison. The same baseline dynamic model was used for the EKS iterations to justify the use of errors measures. The standard methodology in clinical research was adopted using the latest recommendation in the statistical field in the form of BFB to ascertain the higher reproducibility of the statistical findings that were reported. BFB measure ascertains that our results speak in favor of the alternative when the odds are higher for instance than 16:1 as suggested by Bayesian statisticians.

The pointing task is a commonly used task of evaluation of the upper limb rehabilitation and with different motion capture systems, we can argue that the cartesian trajectories modeled using the proposed framework can indeed yield comparable results. Different exoskeleton platforms might as well find positive results using the same technique.

The overall coherence of the findings and the comparative trends noted on standard kinematics does follow suggestions in the literature and helps assist the practitioners and researchers by the addition of an evaluation instrument capable of detecting changes over shorter periods and which exhibit strong clinimetrics.

The results are primary and the study is observational so real causal deduction might not be robust enough considering a high number of lurking variables not accounted for. Nevertheless, this can be viewed as a positive claim since it coherently fits with the FMA-UE score and is evaluated with high randomness considering the framework implementation.

The remaining fact is that we evaluated these metrics on real patients data, data from a highly heterogeneous panel, following different therapy schedules and belonging to different age groups. Nevertheless, we could observe similar tendencies that could be viewed as strong evidence in favor of the proposed method.

6.4 Clinical implications

The presented metrics after statistically studying their evolution during patients' recovery are capable of evolving in both directions. Thus, the measures are characterizing the direction of recovery (i.e. improving or worsening). This property helps indicate the instantaneous level of recovery to the patients for feedback and to the practitioners for proper adjustments.

The regression tests are positive and describe the ability of the metrics to evolve monotonically with the recovery. The responsive measure can be used to follow up patients during the training. The monotonic nature also helps us make some small future interval predictions about the potential of the evolution for the given patient. Therefore it helps provide care to those that can benefit the most from it.

The monotonic nature of change permits us to state that the motor task relearning can successfully be tracked using the provided instruments.

Importantly, the consistently present monotonic relationship at different FMA-UE levels of recovery showed no ceiling effect relative to the skill acquisition plateau seen for instance on all the motor task performance curves. This would constitute a strong quality to distinguish between the performance increase and actual recovery attained through training [148]. We have seen for instance that some measures had a second-order fit while the rest remained conforming to a first fit. From the assessment perspective this can be related to the sensitivity of the measures to the learning effects in the case of the second-order fits.

The first order fits can reveal a potential learning free measures, a research question that merits further interest and cannot be immediately investigated given the available data. The evolutive significance and strong monotonic relationship reported for Q, var, cdf and L_T show an extremely valuable insight, in that, the metrics are direct quantities of the model proposed and shows us that the model evolves according to the performance changes during therapy.

The proportionality relationship between FMA-UE and the Q where we reported a Spearman ρ of $.61$ in addition to the second-order results of Fig. 6.2 showed that the metric evolution reached the plateau phase and thus fit the second-order dynamic for the patient with the highest FMA-UE score and not for the rest. Nevertheless, the $r - squared$ are low overall which suggests that a second order is not the best fit for data.

Additionally, the Spearman coefficient of the evolution of the metric longitudinally for all patients was found significant with high effect sizes. If we note that the metric Q is the contribution of the dynamic model and is learned through the Expectation-Maximization loop, we can be highly confident that the learning framework is providing a dynamic score of the physical rehabilitation of the upper limb.

In aggregate, the metric presented a similar dynamic to the evolution of the recovery and it can be concluded that it is reflective of the dynamic of the underlying motor task learning mechanism. Thus, computational modeling of the recovery is a strong value that can be derived from this proposed framework [81].

Specifically, when we consider that our model is the first to track these differences at the driving physical dynamic level. The remaining literature on the subject attempted to quantify performance dynamics through a proxy measure. Considering the challenges and limitations of these measures the pure dynamic formulation of our framework puts it in a strong position to investigate those cofounding effects for example. This for instance can be achieved through the type of data used at the entry of the model and through the introduction of further extraneous effects into the model itself. This ability can permit researchers for instance to investigate effects which can also be done through proxy measures. Meanwhile, the quantification of these effects from data is an added benefit to our formulation.

To our knowledge, only a few studies according to [147] and lately in [159, 21] have studied the evolutive quality of the assessment metric and showed significant trends with longitudinal data. We have attempted the same analysis on the classic kinematics and the results showed weak responsiveness (the reader is referred to the prestudy in chapter 3). This quality is of utmost importance for any individualized therapy planning. Not only that it allows individualization, the repeated measures within-subject nature of the tests helps us increase the robustness of the finding as each patient is acting as his control.

Consequently, we do not have to match for lurking variables as we ought to do in between-group statistics. Additionally, the metrics reported are mostly Bayesian measures of likelihood and probabilities and it would suffice according to the Bayesian statisticians, to report their monotonic increase or decrease to make conclusions about the hypothesis put forth [36]. Nevertheless, we sought to complement our results by providing traditional frequentist tests to help communicate to the clinical community and help make comparisons easily.

As a discriminative metric, the independent samples tests were also significant and thus ascertaining the ability of the measures to distinguish between the different stages of recovery and different pathologies. This should be put to further testing with detailed studies on homogeneous panels to be able to make decisive conclusions.

Analyzing the statistical results reported we can note that we are reporting fairly high BFB values and thus fairly high odds in favor of the alternative hypothesis. This should help reassure our conclusions that the actual alternatives for the tests are highly likely.

Compared to the standard FMA-UE measure, we hypothesize that we can make a fine-grained evaluation of the recovery using objective measurements. The measures are established based on the Armeo training sessions and thus are easy to administer and more precise to parameterize the recovery process. Based on the requirements that we reported in the literature review in section 2.3, we can be confident that the measures are longitudinally sensitive, easy to administer, and fast to obtain. In addition, the pathologies disparity seems to have no effects on the measures.

This can be due to the benefit of considering low-level physical implementations of the movement modeling as a proxy which is common regardless of the underlying causes. These causes manifest on the trajectories in different ways and this might be interesting to investigate for instance, as a classification problem.

The sensitivity of the measures may also help us provide a boost for patients' motivation. As is known during physical therapy, patient involvement and the psychological state play a central role in the progress of his recovery. Motivation governs the amounts of efforts put into training, so by providing finer measures of the progress and highlighting gains achieved, we hope this effect would constitute a further advantage to the framework.

Although we laid the results in the rehabilitation assessment context, other closely related applications might benefit from this technique. The framework is modular. For instance, all the assumptions on the trajectories, robotic or otherwise, can be encoded for at the dynamic modeling phase. Building a specific model for a predefined task by knowing its trajectory might allow us to assess the fidelity between the execution and the supposed model.

Research involving Human-Robot interactions can find the formulation extremely useful since subjects' movement can be extremely unpredictable and thus by providing more compliant robot controls we can allow patients freedom of movement while imposing softer constraints at the dynamic model or as a perturbation to the signal during EKS iterations. A more human trajectory planning module can be built on top of characterizing users' invariability into a dynamic formulation and accounting for subject-level differences at runtime by executing through a properly adapted EKS iteration.

Importantly, in any robotic application attempting to solve the proper model analytically can be unfeasible due to many constraints. One of the challenges is dimensionality. Moreover, the intricate details of the skills can easily escape the model. The formulation of the imitation learning problem helps bypass these challenges by extracting expertise into models.

The Bayesian formulation helps us account for natural variability and depending on the predefined expert model can help eliminate errors that fall outside of our task model. To provide definitive conclusions, a longitudinal case study with multiple period follow-ups using FMA-UE scores would help provide sufficient validation of the tool for the intended use case.

6.4.1 Shortcomings

Among the main limitations of the current study is the sampling irregularity. Each patient followed a different training schedule that was not intended for research studies. It relies on probabilistic estimations and modeling, a fact that makes it inherently stochastic and thus results in its high sensitivity to initialization and coefficients changes during the estimation loops.

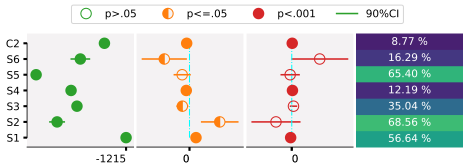
The results were calculated based on identical threshold and initial conditions to allow for comparison. The resampling of the data at the beginning of the procedure might be a source of undesirable artificial data, still, the transformations did not affect the assumptions of the modeling technique nor of the statistical tests used to derive conclusions.

The sample size is low for the statistical tests and a larger cohort can better address this concern. Nevertheless, this limitation concerns mainly the conclusion about the construct validity of the metrics since it requires measuring between-subjects correlations. The reported results also had significant effect sizes comforting the conclusions.

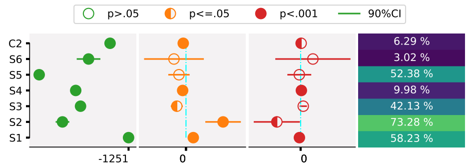
6.5 Conclusion

In this study, we investigated the clinimetric qualities of the developed metrics. This was consistently reported both as a need and as a recommendation for researchers to validate the proposed instruments and qualify their clinimetrics.

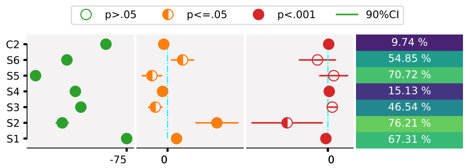
The results suggest good quality overall. The use of the scale to follow up on recovery is enabled and can be used in the context of motor learning modeling to parameterize influences and cofounding effects that interact with the recovery performance.



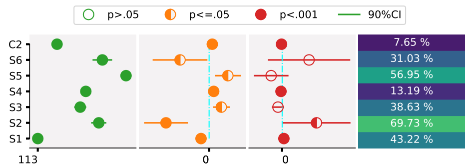
(a) Results for metric AIC_{z_1}



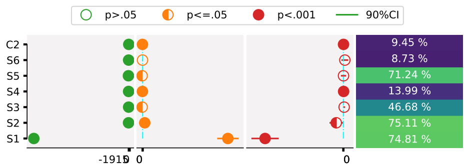
(b) Results for metric AIC_{z_2}



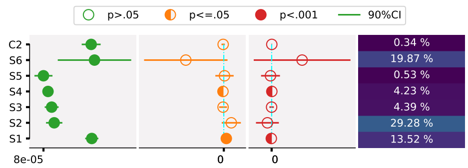
(c) Results for metric L_T



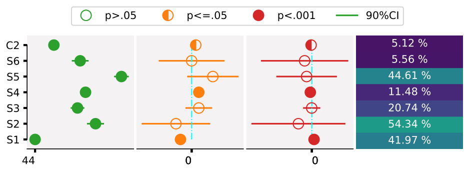
(d) Results for metric M



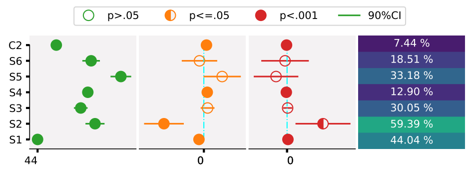
(e) Results for metric E



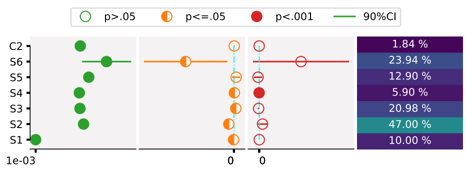
(f) Results for metric Q



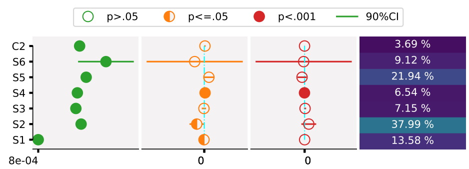
(g) Results for metric cdf_{z_1}



(h) Results for metric cdf_{z_2}

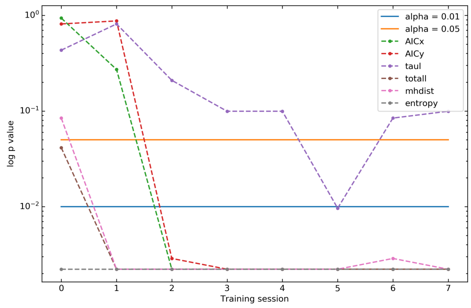


(i) Results for metric var_{z_1}

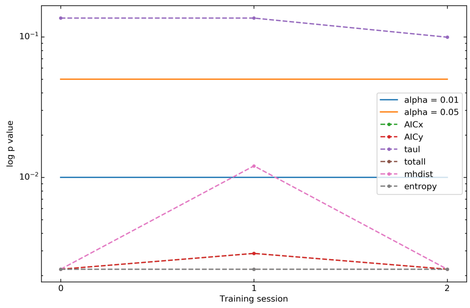


(j) Results for metric var_{z_2}

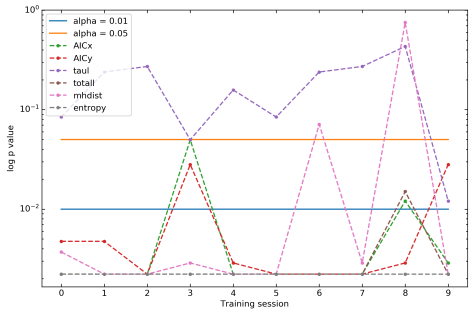
Fig. 6.2.: Regression charts is a graphic to visualize the 2^{nd} order regression results. The intercept, first order and second order coefficients are colored green, orange and red in order. Ticks on the x axis is placed at zero or at the nearest value to zero in the data, when at zero a blue line in the middle distinguishes between positive and negative coefficient. This is pertaining to the convexity or concavity of the parabolic fit. The heatmap represents the R^2 results refering to variability of the metric that is represented by the parabolic fit.



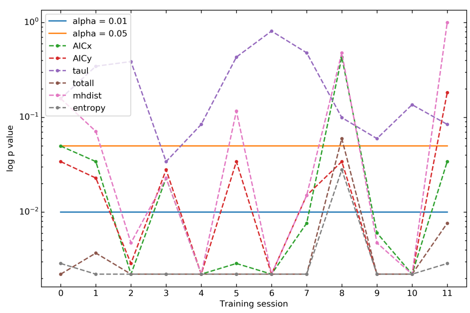
(a) Results for Patient $S1$



(b) Results for Patient $S3$

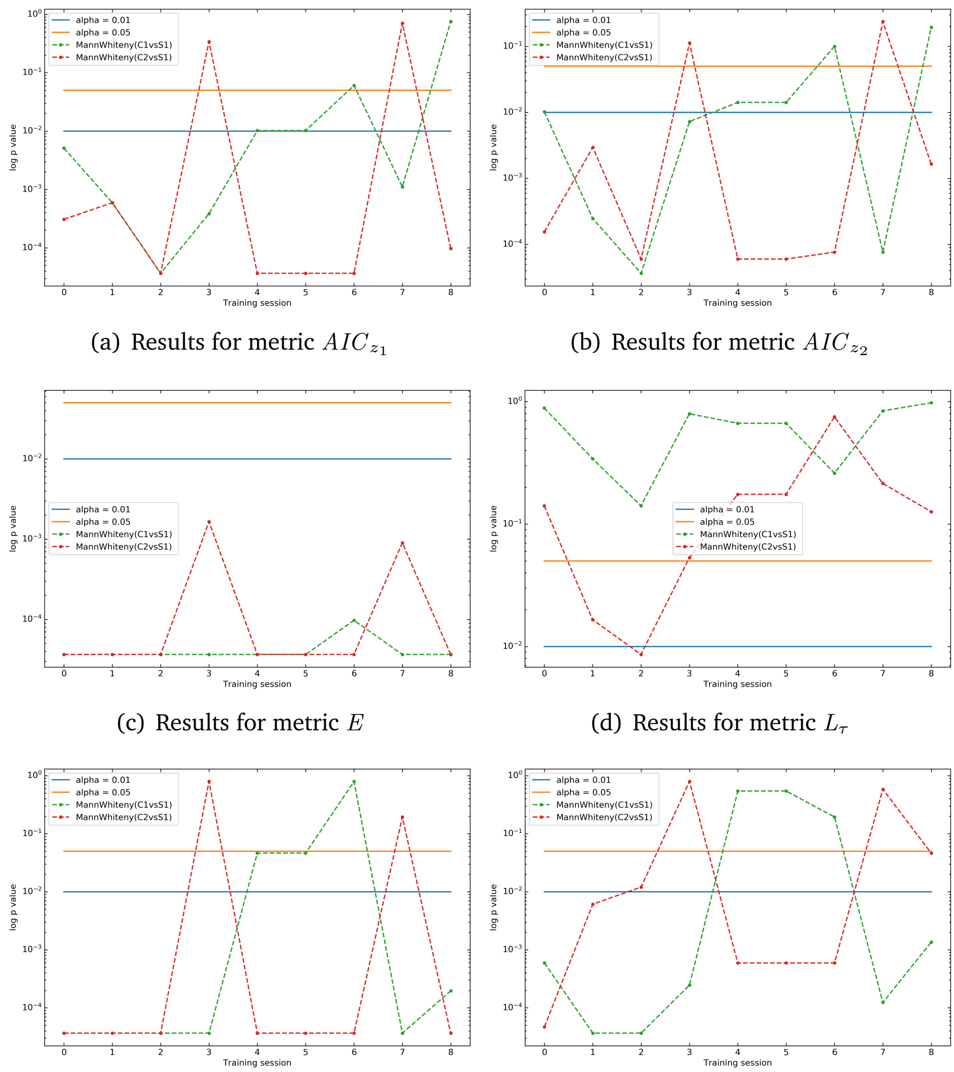


(c) Results for Patient $S4$



(d) Results for Patient $C2$

Fig. 6.3.: The evolution of the $p-values$ for the statistical Wilcoxon test for baseline metric measurements and those after each additional session of training compared to significance thresholds $.05$ and $.01$. The results show that except the L_τ all the metrics are capable of detecting significant differences between the performance for S1, S3 after some additional sessions. Same behavior is present for results of subjects S4 and C2 where tests indicated some insignificant results following a fluctuating performance. The results are moderately to strongly in favor of the alternative considering the BFB associated with the $p-values$ below 10^{-3}.



(a) Results for metric AIC_{z_1}

(b) Results for metric AIC_{z_2}

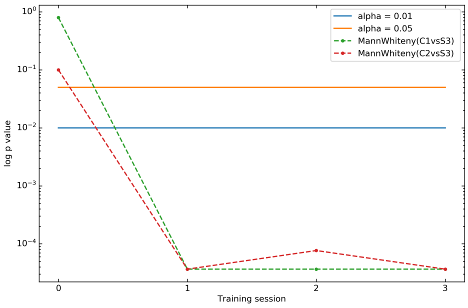
(c) Results for metric E

(d) Results for metric L_τ

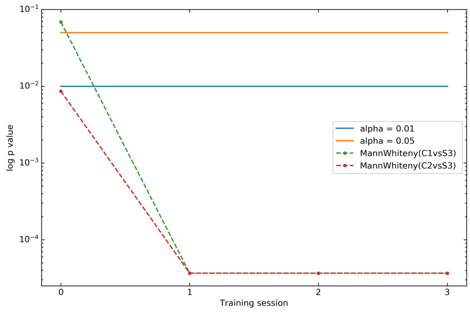
(e) Results for metric L_T

(f) Results for metric M

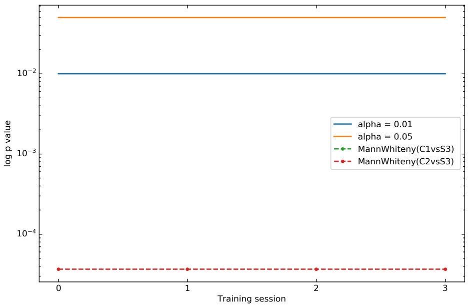
Fig. 6.4.: Plots of log $p-value$ of the Mann Whitney test for metrics evaluated on trajectories of subject S1 and a healthy control C1 (in green) and a non fully recovered patient C2 (in red) with significance thresholds in vertical lines.



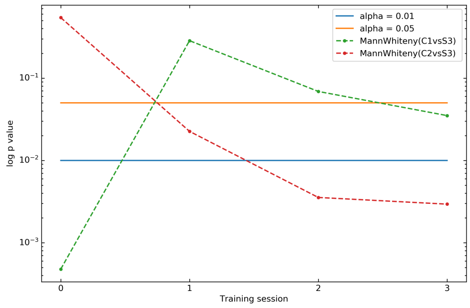
(a) Results for metric AIC_{z_1}



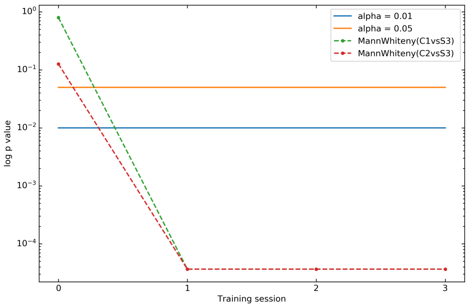
(b) Results for metric AIC_{z_2}



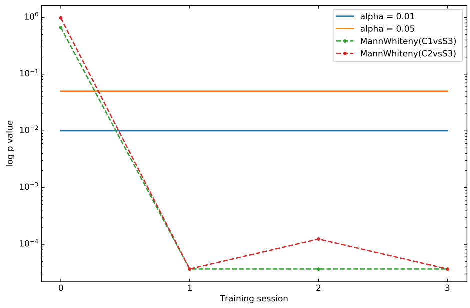
(c) Results for metric E



(d) Results for metric L_τ

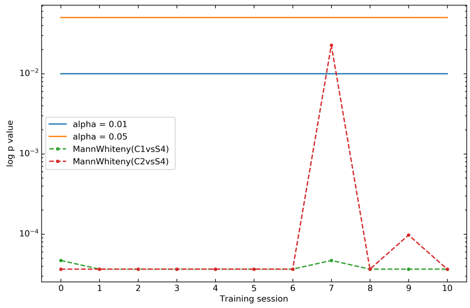


(e) Results for metric L_T

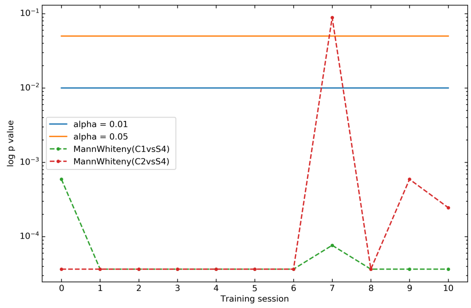


(f) Results for metric M

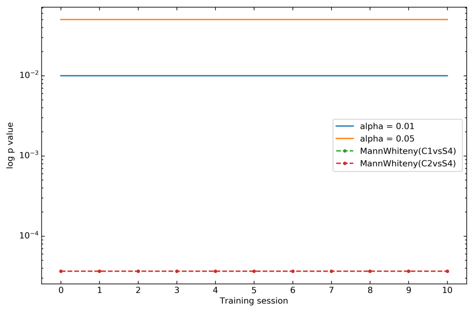
Fig. 6.5.: Plots of log $p-value$ of the Mann Whitney test for metrics evaluated on trajectories of subject S3 and a healthy control C1 (in green) and a non fully recovered patient C2 (in red) with significance thresholds in vertical lines.



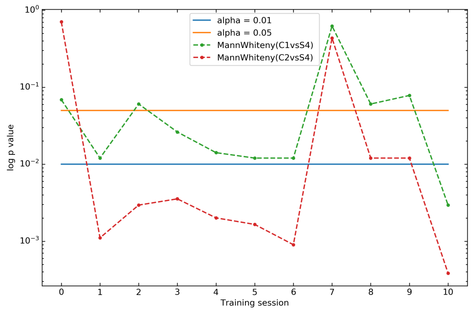
(a) Results for metric AIC_{z_1}



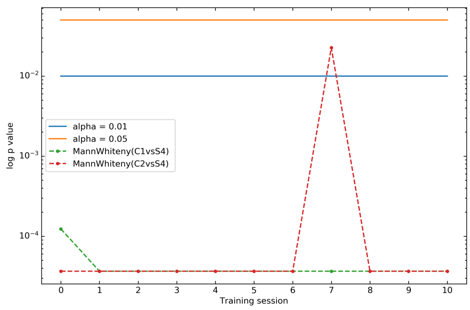
(b) Results for metric AIC_{z_2}



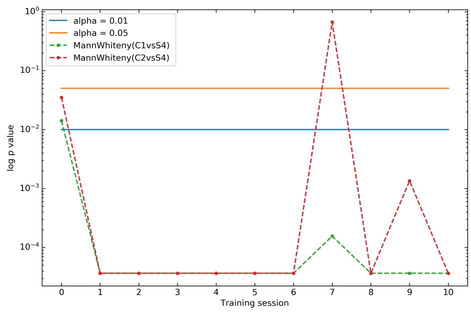
(c) Results for metric E



(d) Results for metric L_τ

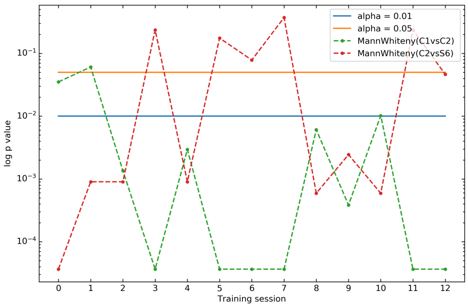


(e) Results for metric L_T

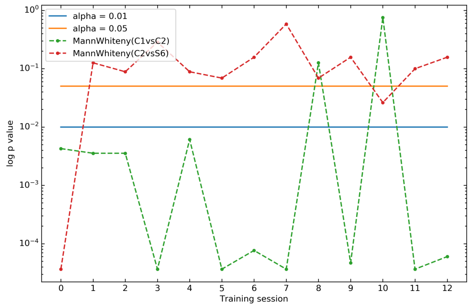


(f) Results for metric M

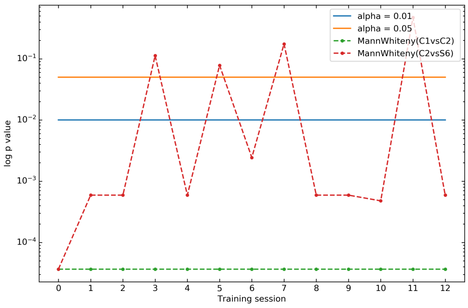
Fig. 6.6.: Plots of log $p{-}value$ of the Mann Whitney test for metrics evaluated on trajectories of subject S4 and a healthy control C1 (in green) and a non fully recovered patient C2 (in red) with significance thresholds in vertical lines.



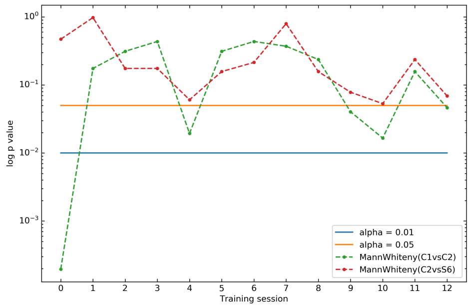
(a) Results for metric AIC_{z_1}



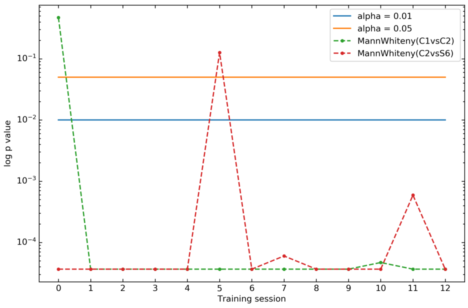
(b) Results for metric AIC_{z_2}



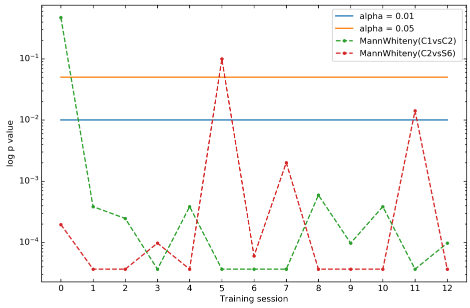
(c) Results for metric E



(d) Results for metric L_τ



(e) Results for metric L_T



(f) Results for metric M

Fig. 6.7.: Plots of log $p-value$ of the Mann Whitney test for metrics evaluated on trajectories of subject C2 and a healthy control C1 (in green) and a non fully recovered patient S6 (in red) with significance thresholds in vertical lines.

7 Procedural Generation of Rehabilitation Data Visualizations

> *The purpose of computing is insight not numbers.*
>
> — **R.W. Hamming**

This chapter takes a different route towards assessing rehabilitation training. The visualization techniques at the heart of all data-centered applications are rather scarce in the rehabilitation space. We formulate basic motor task learning assumptions into a visual model.

The model consists of presenting the data aggregated as a probability map. The probability map is validated using a derived measure. Different modalities of the visualization are presented and are advocated for use as patient-centered feedback forms.This chapter takes a different route towards assessing rehabilitation training.

The visualization techniques at the heart of all data-centered applications are rather scarce in the rehabilitation space. We formulate basic motor task learning assumptions into a visual model. The model consists of presenting the data aggregated as a probability map. The probability map is validated using a derived measure. Different modalities of the visualization are presented and are advocated for use as patient-centered feedback forms.

7.1 Scope of Data Visualizations in Rehabilitation

Rehabilitation research involves the collaboration between mixed teams often involving interdisciplinary scientists. The standardization of the technological means of communication would benefit all involved parties, a case that was made in [97] where a 3D visualization of patients' rehabilitation outcomes was shown to be favorable by the practitioners and clinical researchers. As the research in this field involves more technologies such as robotics, we see a clear rise in the data-driven approaches to the study of rehabilitation with more interest given to the interaction between devices and therapists. Thus, guidelines are deduced that serve to design software to address this gap especially in the visualization of patient data [53].

The visualization tools proved useful, for instance, in the case of brachial plexus rehabilitation to help practitioners analyse the movements and collate large dataset [29]. While data visualization has proved a very important tool mostly in the fields of data mining and data science, it starts to become interestingly useful in the context of rehabilitation. Examples have been shown where practitioners follow up on rehabilitation over long-term real life conditions was augmented by visualization methods addressed for the practitioners [131].

Heatmaps, line plots, and dots charts have also been proposed as tools to visualize data of body sensors network tracking systems [111]. In rheumatology, patient-facing visualization feedback in the form of linear charts was found to significantly affect the control of disease activity and to promote adherence to treatment [113].

In gait training after stroke, a 3D visualization has been proposed as well to provide visual feedback on performance to patients [157]. In following up activities in real-life conditions visuals are used as an intuitive tool to define patterns, detect changes and monitor patients. In [28] authors used a spiral plot to visualize in different colors the patterns of arms use through time in daily life conditions.

Visualization tools are gaining more interest in applications where overall health outcome or status of the population is a concern [154] and is proposed as feedback to support the healthy aging [9]. Applications of symptom visualizations have also been reported in the literature [101] with inconclusive optimal designs stressing again the importance of a user-centered approach to developing such tools.

Another fairly advanced application context is in sports analytics where follow-up on performances of players is of importance [158]. The visualization tools might help provide an easy to interpret, comprehensive report of the research findings as has been shown on gait analysis with visualization tools proposed in [130]. The tools have been shown to provide the clinicians using it, useful knowledge from complex datasets.

Visual outcome results are accessible to all the stakeholders in the rehabilitation context, from providers to decision-makers and practitioners and patients [64]. A universal communication tool can help overcome the technical disparity between the involved parties and provide a common ground to communicate results [163]. Another motivation is the current developments of telerehabilitation tools.

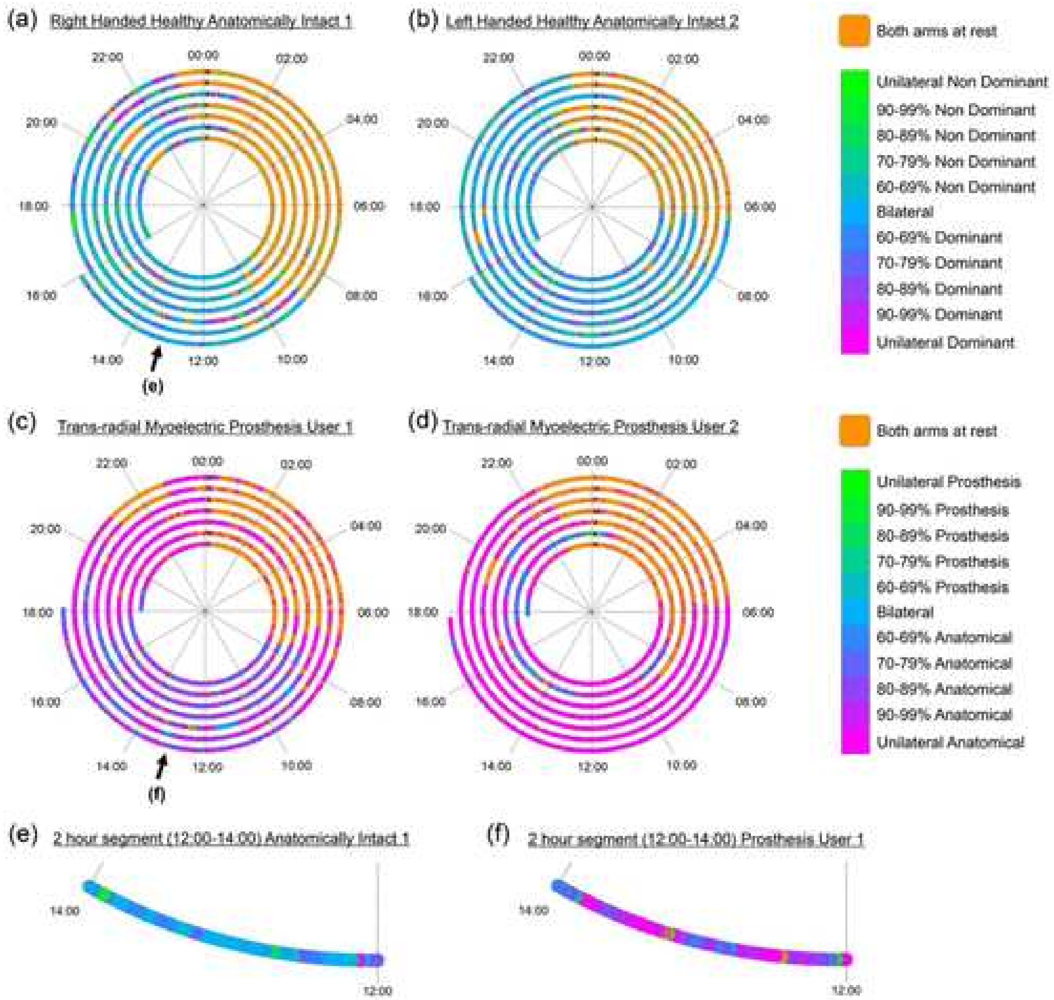


Fig. 7.1.: The authors [28] used visualization to condense a week's data into spiral graph representing the frequency of activity. Relating limb use to pathology to pattern of use in comparison to prosthesis users satisfaction.

As these tools are intended for personal unassisted use by patients, usually at home, users can highly benefit from an interpretable tool to identify the level of recovery and ensure the gains achieved are maintained [45, 127, 156]. Another gained benefit to this approach is the inclusiveness of the tools' design and their overall accessibility, since visual communication is better suited to convey the information without the need of the technical background involved with producing it and considering the older adults audience [95, 88].

A surface Electromyography topography was used to assess the lower back pain through visualizing the lumbar muscle activity. The resulting map representing the root mean square of the signal and plotted as a spatial distribution shown to better distinguish healthy and patient's results [80].

The possibility to involve all the stakeholders in the rehabilitation equation into the related discussions has been shown within a focus group made up with different backgrounds, the primary communication tool was a 3D visualization with a heatmap encoding biomedical data of the older adults during activities of daily living (ADLs) [104]. A review conducted on telerehabilitation's perceived benefit and challenges from the users' perspective has highlighted that the display of the recovery progress would constitute a major element of any proposed system [35].

Researchers proposed a visualization of the cognitive performance which is established in timelines and heatmaps and aggregated visuals co-designed with the carer [79]. The visuals are presented in a storytelling approach to rehabilitation progress feedback, primarily for patient's home use.

An adversarial autoencoder was trained to visually represent, in a latent space, high dimensional time series data collected using a pressure signal during different physical exercises. The network represents the performance and classifies movements as a promising result for monitoring at-home rehabilitation training [75]. Kinematic modeling of the upper limb serves as a tool to generate a reachable workspace graphic that can communicate the differences between healthy and pathological limbs [91].

In [10] the authors were able to use the density plots of the bilateral arm movements to differentiate between the healthy use pattern and those from different paretic dominant or non-dominant limbs during the daily life of stroke survivors. In terms of communicating assessment results included, research reported interesting techniques for instance: Radar charts of the kinematic measures in different directions comparing pre and post-training differences was proposed in [125].

Another approach for workspace visualization in joint space and with interesting projections on elbow wrist and shoulder elbow planes was presented in [139]. Trajectory drawing presented a specific context that could benefit live scoring. Scoring as feedback, using color-coded visualization of the trajectory execution was proposed in [58]. The authors used violet-red gradient markings to delimit the highly likely-unlikely parts of the trajectories assessed using a distribution over trajectories model learned with LfD.

As many optimization and learning techniques are adopted in the rehabilitation literature, the use of a surface map $F = f(m, R)$ where m was the upper limb mass and R was the radius of the movements was proposed. The surface is color-coded

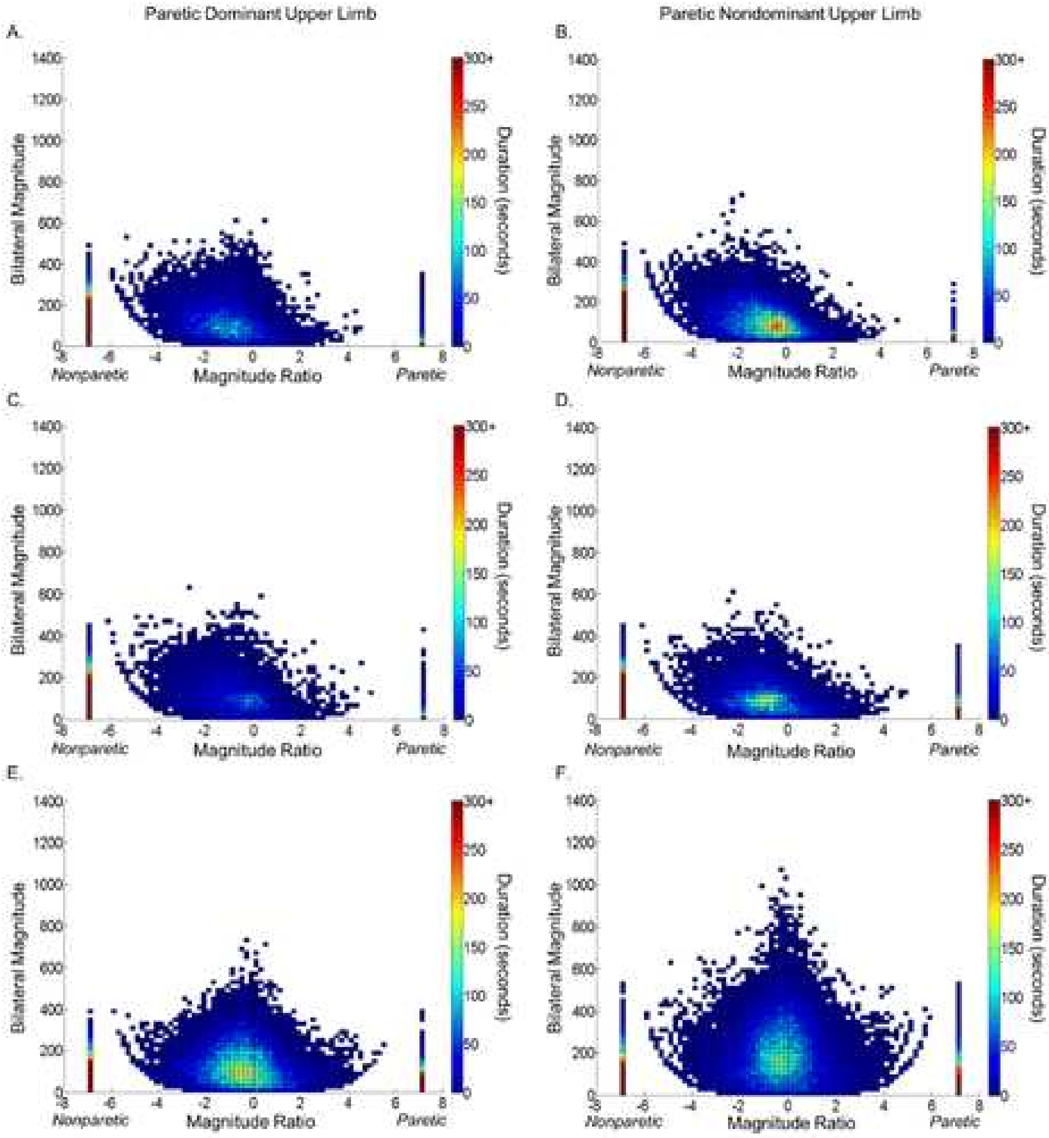


Fig. 7.2.: Visualizing the range of motion per side and by frequency of movement in time spent in a given area the authors [10] showed differences between patients sides and the six different function levels as measured with ARAT.

and visualizes optimization solution and patients' performance comparatively [5]. They also presented plots providing a simple clear representation of the subject performance by showing ellipsoids around the centers of the tracing task the subject performed showing the intra variability and comparing against healthy controls [48].

A graphic representation of the kinematic indices on a circle graph was proposed. This helps contrast the differences compared to normative data visually. The vertices from the circle center are drawn to identify directions of movement and provide

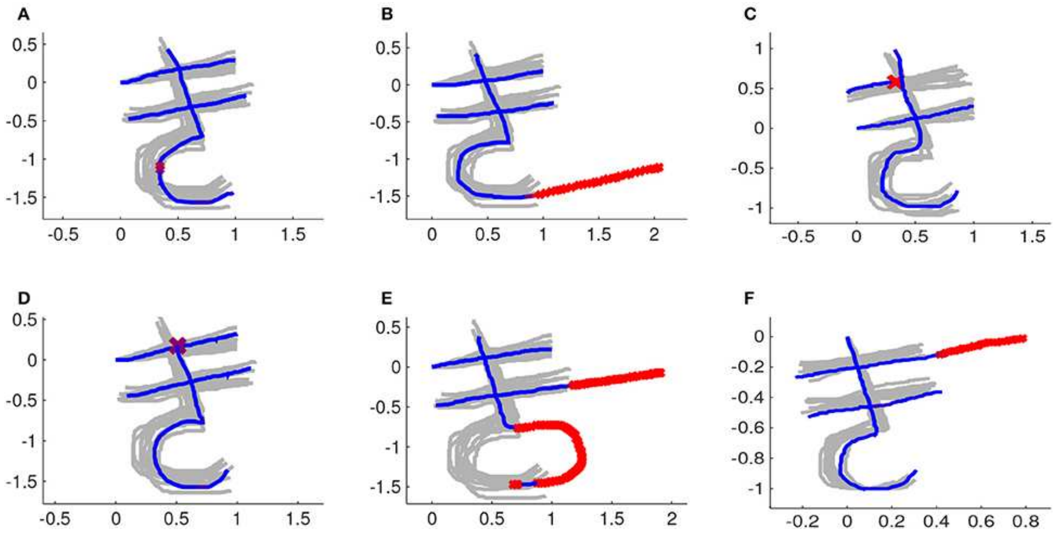


Fig. 7.3.: Real time scoring vizualization for rehabilitation tracking task. The blue parts of the trajectories are rated correct while moving away towards the red marks increasingly wrong movements [58].

additional insight into deficit in particular directions [100]. In [84], authors suggested ROM image visuals as a map of movement reflecting range and intensity of training.

7.2 Study Motivation

From the presented literature, we find that none of the methods presented visualizations of exoskeleton data during reaching moves. Very few visuals are intended to provide a form of visual feedback. Therefore, we proceed in the presented work to implement visual feedback in the form of heatmaps. These graphical tools are intended to communicate patient recovery status objectively using training data from the exoskeleton device. The graphics have the form of maps that represent a set of kinematic metrics reported in the literature as metrics correlated to the recovery progress [146, 100, 16].

Data visualization in rehabilitation studies was found to be very limited. In stark contrast to what was one of the most recurrent needs from patients and therapists. As we have seen many potential uses are presented in the literature while the most favorable by patients being the probability maps. We assume a similar visualization technique and this chapter presents the formulation and the derivation of the maps. We first define a preprocessing phase that enables us to transform the data into a condensed form we use afterward to plot the movement as probability maps.

Two distinct forms are presented as cellular histograms and a continuous contours map. The preprocessing phase and the algorithms used to generate the visualizations are presented with associated implementation details. We also list the range of motion visualizations that are generated without the preprocessing phase and are useful tools to investigate joint level variations.

7.3 Visualizations of Active Range of Motion

We have seen in the literature chapter that the range of motion during training, for post-stroke for instance, increases dramatically. Investigating ways to visualize this information was also attempted and the results are presented in simple form maps without any preprocessing necessary.

First, the 3D representation of the trajectories dataset for a given session presents the entire $[x - y]$ grid and plots the actual trajectories using a hue representing the z coordinate. Figure 7.4 represents an example map where we see the extended planar trajectories with depth information for each position on the plane.

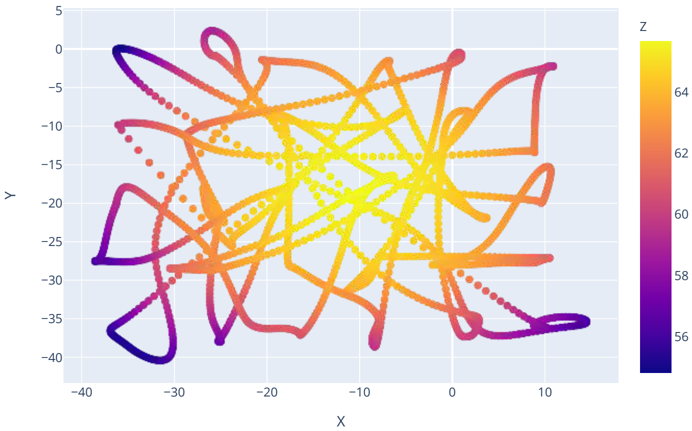


Fig. 7.4.: The 3D range of motion for subject C1, the planar trajectories are represented in 3D space by adding the depth information in form of Heatmap.

Second, the JointRom map is a circular graph representing the upper limb joints ROM using colored arcs on a circular graph. The joints are represented in their anatomical order from the inside out. Figure 7.5 represents the joint rom map. Each number from 1 to 7 correspond to a joint from proximal to distal. These DoF are related to the exoskeleton structure and does not map directly to the human upper limb reference frame. This does not obsecure the fact that the movements are completely subjected to the human joints tightly strapped to the exoskeleton device.

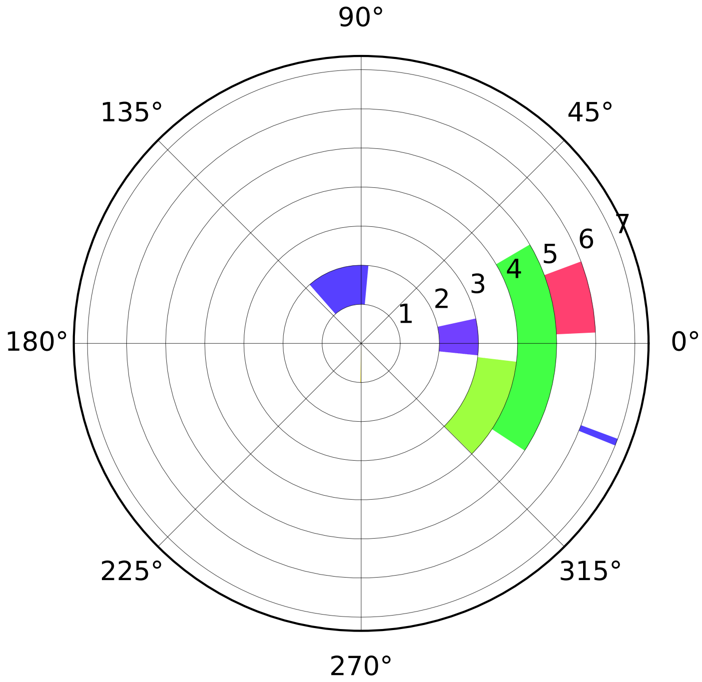


Fig. 7.5.: The joint range of motion for subject C1, each arc represent a corresponding exoskeleton DOF rom through a given exercise session.

These visualizations while undeniably informative, lacks standardization and requires -for the case of Joint ROM- expertise in the physical therapy to gain useful insight. As can be seen in the results in the next chapter the maps change drastically between subjects and between rehabilitation stages. To mitigate this remark we interest our work in designing standardized ways of visualization.

7.4 Visualization Technique

Generating the visuals consists of following a set of predefined steps. The first of which consists of triage of the ETs in a given training session, adopting the successful catch as a decision criterion. The successful ETs are then resampled to have the same number of timestamps for each. Dynamic Time Warping algorithm [143] (DTW) is employed afterward to match the pairs of coordinates according to the minimization of the in-between Euclidean distance.

The coordinates vectors representing each ET are then stacked according to their common indices as per DTW. The S-Map plot then reproduces the probability density of the coordinates given the stacked sample vector where color opacity represents the probability density value. The H-Map on the other hand shows a pixelized distribution map and provides additional marginal histograms for the x and y coordinates. The exact procedure is presented in Algorithm 3. A visual representation of the generation procedure is depicted in Fig. 7.6.

These tools are intended to encode standard kinematic features such as:

- *Velocity*: the movement's velocity can be seen in the form of condensed probability density on the extremities and low probabilities in the middle of the trajectory when velocity reaches its highest in a pointing task. The lower probabilities for higher velocities can be expected given the fixed sampling rate of the exoskeleton.
- *Smoothness*: which represents the continuity of the higher derivatives of the path trajectory of the movement during the task. The probability of moving near the x axis should be higher in the case of smooth direct movements.
- *Overshoot*: representing the distance traveled past the position of the target. The probability density area around the origin should become less as much fine control and coordination are attained. Thus less overshoot is present in pointing.

7.4.1 Algorithm

To perform the preprocessing phase on the exercise data file we can note the steps on Algorithm 3. The common length is first determined by considering the median of the ETs within a given session. Each ET is processed separately afterward, such that it is translated to the origin to account for randomly appearing targets on the vertical plane of the screen.

A resampling is performed to meet the predefined number of timestamps in sequence length. The result is rotated to emulate a catch on the horizontal axis. For ETs moving in the negative direction a mirroring step is applied to keep the trajectories in a unified direction. The ETs are also unit normalized to have comparable dimensions and standardize the procedure. Once all the ETs in a single data file are processed the first eighteen are picked. The DTW algorithm uses the

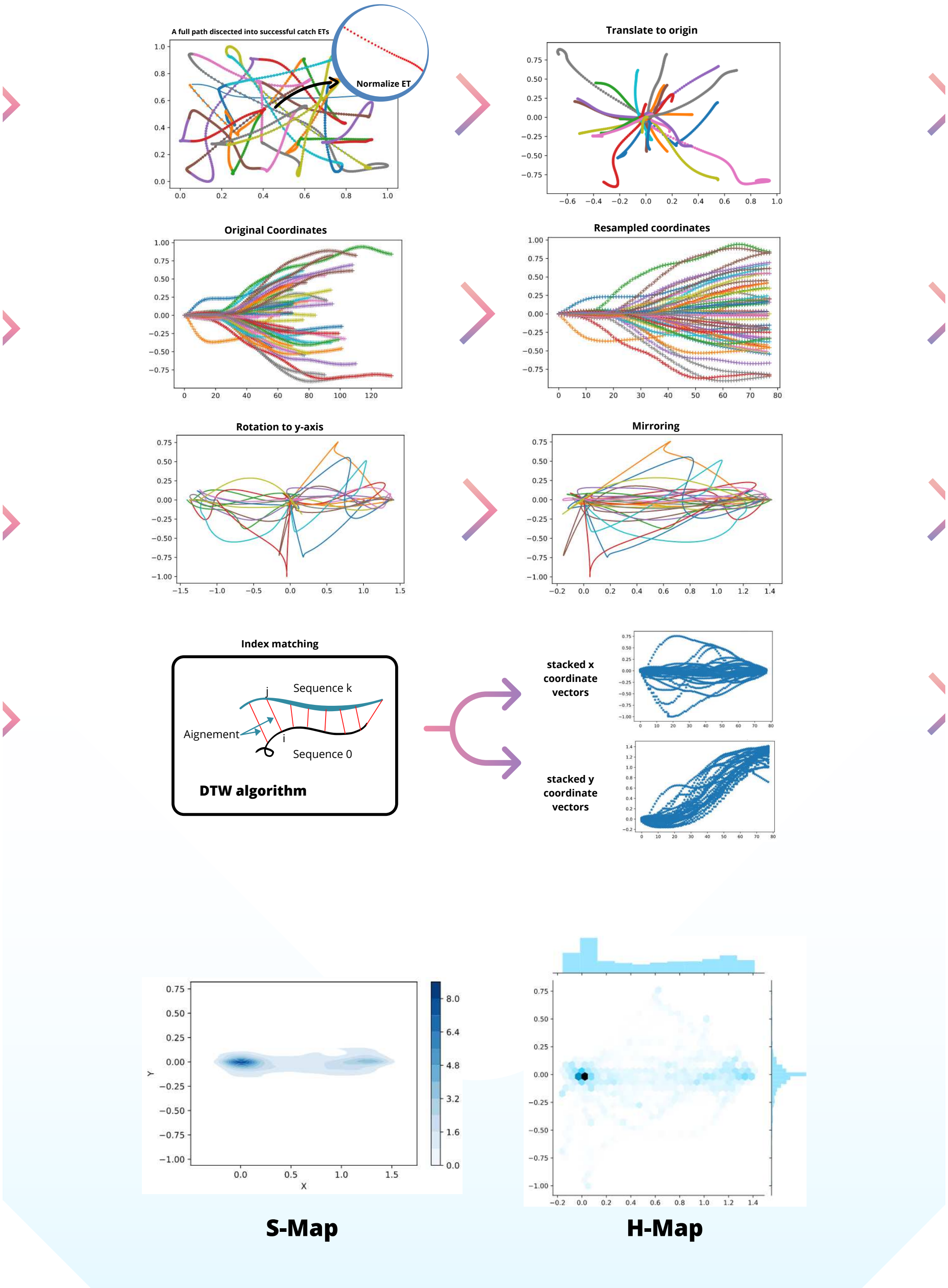


Fig. 7.6.: A schematic for the map generation procedure illustrating the steps using an example datafile and depicting examples of the resulting maps.

first trajectory as a template and matches successively the remaining trajectories resulting in alignment vectors A_i .

The DTW procedure is applied following a Sakoe-Chiba constraint of radius 2 which entails that the recursions in the search of the least costly path in the distance matrix are allowed in a radius of 2 in both directions from the current index [66]. The resulting alignment vectors A_i are then treated using the Algorithm 4.

The resulting alignment is then used to stack the coordinates into bins of the same indexes according to the alignment vectors. The results are then put into two separate condensed vectors: one holding the coordinates $[x, y]$ and the other containing the respective aligned indexes. These vectors are then passed to a plotting function to generate the map. Firstly, the S-Map uses a Kernel Density Estimation (KDE) technique to estimate the parameters of a bivariate normal probability density function from the sample vector that is passed on to it [69].

It then plots the resulting probability density function for the entire $[x - y]$ grid. Secondly, the H-Map is a histogram representation of frequencies of the apparition of a given coordinates bin in the presented sample. These are augmented by side histograms representing the marginal distributions of the coordinates X, and Y and give us additional insight on strategy. The algorithms were implemented using Python scripts and visuals are based on Seaborn package, namely on kdeplot and jointplot functions.

7.5 Conclusion

This chapter presented the modeling algorithms and procedures used to generate visualization maps. The definitions, as well as the implementation details of the maps, were detailed. The next chapter presents the results of these maps on our CHU-PRC clinical dataset.

Algorithm 3 Visual Maps Generator

Initialize: m, the median of ETs lengths in a single session
for ET in session : **do**
  Identifying the common frequency $F = \lfloor \frac{m}{len(ET)} \rfloor$
  Translate (ET) : translate the ET to the origin.
  Normalize (ET) : unit normalize the ET.
  Re-sample (ET) : either up-samples or down-samples to the predefined frequency.
  Rotate (ET): rotate the ET such that it ends on the y axis.
  Mirror (ET): switch the sign of x for ETs that have negative x values.
  DTW(ET_0, ET_i): find the proper alignment respecting the *Sakoe Chiba* radius of 2.
end for
Stack(ET_0,A_0,...,ET_n,A_n): stack the coordinate vectors according to their matched indices resulting from the DTW.
S-MAP(Z) : plot the probability density of the stacked vector $Z = [X, Y]$ as a hue representation.
H-MAP(Z) : plot the representation of the stacked vector as hex cells with marginal histograms of each coordinate.

Algorithm 4 Stacking algorithm

Stacker(l) : instantiate the Stacker class passing l the median length of an ET.
Stackit(τ, Sequences): call the stackit function passing the concatenated alignment vectors and respective sequences of (x, y) coordinates.
STACK$[k, depth]$: populates the k^{th} STACK column, with a sliding depth index indicating count of column visits, and returns STACK.
Packit : pack the original template STACK vector keeping only populated cells, and returns *stack*.
X : concatenate the first column of the *stack*.
Y : vertically stack the coordinate vectors in the second column of *stack*.

8 Visualizations as Objective Assessment & Feedback Tool

> *The most favorable form of feedback for patients, research suggests, is probability maps.*
>
> — **The Chapter Highlight**

From the literature review in the previous chapter, we found that studies that reported the use of probability maps were useful and favorable feedback form as patients reported. This chapter collates the results of the proposed visualization technique on clinical data. It also presents an attempt to validate the tool in a clinical sense by providing objective statistical evidence for the visual observations.

In similar, we list and detail the tools that we proposed in chapter 7 through the clinical dataset. We present the longitudinal sensitivity of the feedback through pertinent case series where longitudinal effects can be notable through the resulting maps. We then propose the report format for end-of-session feedback to patients and therapists. The report presents the previous or beginning of the training map in contrast to the current performance map and a healthy example for easy visual differentiation.

The workspace results are also eminent as research suggests the range of motion results to be a valid measure of recovery. The resulting maps provided in report form can be notably useful in differentiating strategies and constraints to movement at joint and workspaces through visual inspection of the results.

8.1 Patients Sample

This study involves a case series study where patients were chosen as a convenience sample retrieved from the CHU-PRC dataset. Patients demographics are listed in Table 8.1.

Tab. 8.1.: The panel of subjects included in the study.

Subject	Age	Gender	Affected arm	Pathology
C1	27	Male	Right	Healthy
C2	9	Female	Right	Infantile Hemiplegia
S1	28	Female	Right	Guillain Barre Syndrome
S2	24	Male	Right	Post Traumatic Hemiparesis
S3	38	Male	Right	Post Traumatic Tetra paresis
S4	30	Female	Right	Polyradiculonevrite
S5	13	Female	Right	Dystonia writer's cramp
S6	54	Male	Right	Post stroke
S7	45	Female	Right	Post stroke
S8	65	Male	Right	Post stroke
S9	60	Male	Right	Post stroke
S10	15	Male	Right	Post stroke
S11	51	Male	Right	Dystonia
S12	15	Male	Right	Plexus Brachial
S13	64	Male	Right	Post stroke
S14	11	Male	Right	
S15	45	Male	Right	Post stroke
S16	17	Male	Right	Elbow and wrist post traumatic injury
S17	53	Male	Right	Polytraumatic
S18	31	Male	Right	Tetra paresis
S19	54	Female	Right	Post stroke

8.2 Visualization as Objective Recovery Proxy

8.2.1 Is the maximum density value a useful measure of performance?

An interesting result from the maps suggested that the maximum value for the kernel density estimated from the data increases in value during training. This property was investigated and further tested. Using a locally weighted regression, a typical model trajectory was estimated that was then used to calculate the probability density function values. An example trajectory is represented in Fig. 8.2. The maximum value was used to compare pre and post-session performance for the dataset. The results are reported on Fig. 8.1. To further examine the measure a correlation with the FMA score for the sample of six patients showed a correlation factor of $(\rho(6) = .52, p = 0.28)$.

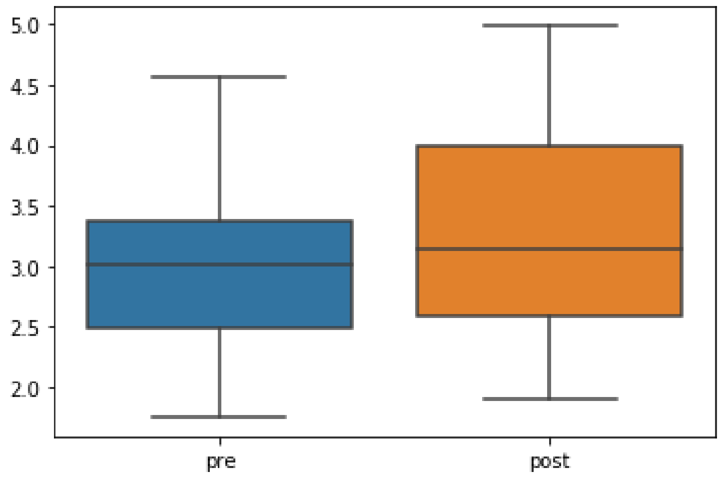
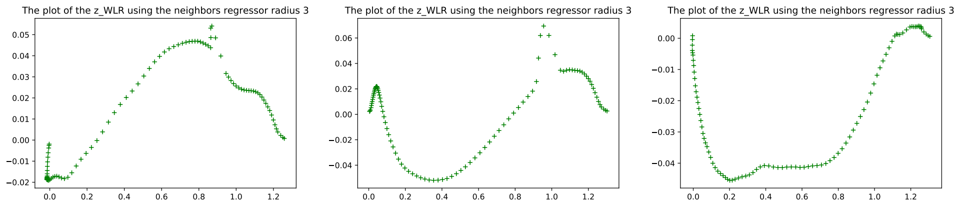


Fig. 8.1.: The boxplot representing the evolution of the maximum density as estimated using a locally weighted regression model from all the trajectories and the KDE Gaussian kernel estimated from the data. The maximum values of the probability density function is estimated given the model trajectory. The results were statistically significant with a Wilcoxon one-tailed test of pre-post data from all available patients $(W(30) = 133, 0.04, BFB = 3)$.

8.2.2 Patient's Longitudinal Evolution Through The Maps

The case of S1 The evolution of the trajectories during the training session for S1 is presented in Fig. 8.3. We can note that trajectories spread do decrease in mid-sessions with increased maximum value for the density while deteriorating by



(a) (b) (c)

Fig. 8.2.: An example of the locally weighted regression model trajectory for patient S7 (pre and post) compared to C1.

the sixth session. Yet the last session differs substantially from the beginning where we see that although movement was slow as seen on the middle region between the left and right extremities of the trajectories the direction errors were less apparent. This is because the movements were likely to be correctly directed from the start, yet being slow. This statement is reassured once we consider the marginals shown in Fig. 8.4. The y distribution is normally centered at the origin for the latest session. The case that midsession presented lower spreading and the low probabilities of middle regions showcase more well-directed movements at higher speeds. A result that is also reassured using the x histogram on Fig. 8.4 which shows the typical dip in the middle where the movement was more rapid. The tails of the y histograms get less extended in the mid-sessions with higher values at the center suggesting better-directed movements.

The case of S2 Longitudinal evolution of the maps on successive follow-ups with subject S2 are presented in Fig. 8.5 for the S-map graphic. On the results' figures, we can note the diminution of the spread in the probability density map as trajectories start to become precisely directed towards targets and moves tend to be smoother while mostly curved. Results on Fig. 8.6 represent the evolution of the H-map over the course of training for subject S2 where we can note the variability in the trajectories at the beginning of the training with a much wider spread on the y histograms. The x histograms show a tendency to dip in the middle of trajectories in sessions at the beginning of training and then start to increase until a target is caught. The x distribution becomes flatter with slightly higher values nearing the catch area where decelerations would make this noticeable.

The case of S3 The S-map evolution for patient S3 on Fig. 8.7 showed the transformation from mostly well-directed slow movements to increasingly faster moves

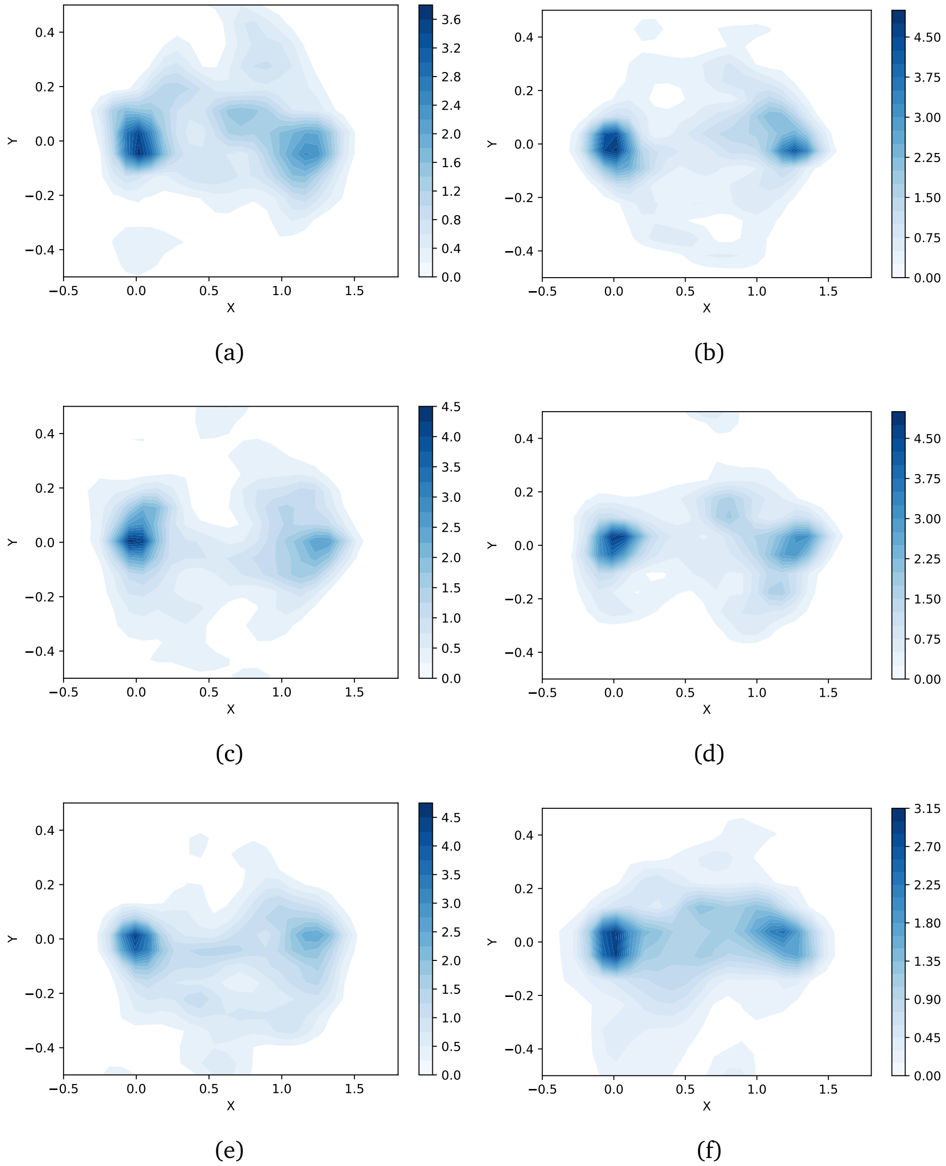


Fig. 8.3.: The longitudinal evolution as represented by the S-map for subject S1.

with higher maximum density values. The occasional dips in the speed of performances showed on the maximum values as well with sustained direction precision comparatively. Similarly, when remarking the H-map, the results on marginal histograms suggest that higher velocities coincide with higher maximal densities values. The directedness of the movements are suggested by the narrower y histograms in Fig. 8.8.

The case of S4 The persistence of the upper arced reaches for subject S4 is noticeable in Fig. 8.9. Later sessions seem to happen at lower speeds compared to the first three. This result can be noted in Fig. 8.10 where we can see the uniformly distributed x histograms during the later sessions, the same can be said about the upper skewed y histograms.

The case of S5 The spread of the maps for patient S5 in Fig. 8.11 becomes increasingly lower. The maximum density value continues to augment during follow-ups and so does the speed of the movement overall. The histograms comfort these remarks as we see typical distributions with more spread on third and fifth sessions where lower maximum densities are presented.

The case of S6 The findings on the S-map in 8.13 show fluctuating performances for patient S6 where we can note a reasonable start then deteriorated performances with a wider spread and eventually performance catching up by the end in form of arched reach with higher velocities and density maximums. Histograms on the H-map suggest that by the final session the velocity of the movement reached a high value with a slight upper skew on the y histograms.

The case of S7 The case of S7 show on Fig. 8.15 rather good performance overall starting and keeping relatively higher values for the density values and for the speed and overall spread of the movement. Typical histograms are presented on the H-map as well 8.16.

The case of S8 The case of S8 showed on Fig. 8.17 presents the evolution of the movement with a slight performance dip in sessions seven and eight. Overall the spread was shorter and higher values were attained for the distribution. Directions were well decided yet the speed of the movement varied suggesting a difference in performance. This is noted on the histograms as well as typical distributions for the marginals.

The case of C2 The case of the spread maps for patient C2 Fig. 8.19 presents a rather unique case where the overall performance tends to deteriorate as observed through the maps. The maximum values attained are getting lower and the spread increases during training sessions consistently. The H-map 8.20 presents a high increase in overshoot as seen near the origin of the maps.

8.2.3 Visual Maps Report Format Compared to a Healthy Control

Results of the visual indicators of the sample are reported on Figures 8.21 to 8.28 below. We can notice that the maps represented the change in variability and smoothness of the accomplished ET. This can be noted when comparing the S-map results of the pathological subjects with those of a healthy user for whom all the variation is restricted between the $(-.2, .2)$ interval along the y axis. The comparison shows that pathological users have more tendency at the beginning of the training to reach wider extents and stray from the smooth movements towards longer winding trajectories.

The recovery progress is shown on the post-training S-Maps where the probability densities became more centered around the x axis suggesting smoother ETs overall. The directness of the movement is also represented by the restricted trajectories closer to the x coordinate axis. Overshoot values do not seem to have a significant presence over the course of training for the included users. The speed of movement can be noted with the weak probability densities in the middle of the trajectories on the post-training S-Maps for most patients.

The results show that the most noticeable change can be noted in the smoothness and directness of the overall moves and this can be communicated on the maps. The H-map shows similar trends where we can note on the x axis histograms, zero centered normally distributed values suggesting the movement is mostly conducted near the origin and less spread is noticed on the post-training values. A remark that is commonly seen in the reported results. Results from subject C2 did not present any noticeable differences in smoothness contrary to the other cases. Nevertheless, the results showed a change in the concentration of moves in the upper arc and the direct path to the target.

8.2.4 Visual Maps of Active Range of Motion

The visuals representing the active range of motion are listed in the following Figures 8.29 to 8.33 . The results are for pre-post sessions and healthy control. The six cases presented in the respective figures present the evolution of each case between the beginning of training and at discharge with a comparison to a control. For 3Dmap the evolution of all patients showed an increase in the movement along the z-direction, where movement near the center of the frame was happening at higher z values where the extremities of the plane happen at lower depths.

This pattern is present more apparently in the post-session whereas for the pre-session we see the depth varying across the entire plane of movement. The joint movement resulting in these patterns presents a variety between subjects where we can see that compared to the healthy control, where most movements happen at the center of the reference frame for most joints, the pathological cases tends to have slight deviations towards one of the sides.

The amplitudes also tend to be for some more important than healthy control. This also highlights the fact that the healthy control had a relatively close per joint range for 3-6 without the need to move the wrist as the task does not depend on it. For the pathological cases, the wrist did present slight variations and the working span for each joint tends to differ from that of the healthy case.

8.3 Discussion

Visual feedback has been a neglected approach to rehabilitation assessment and we presented a framework to be able to generate two different types of heatmaps that would densely represent the kinematic metrics [140, 120]. These kinematics were shown in the literature as evaluation instruments of the recovery progress [161]. The diminution of the probability density in the middle of the trajectories, representing the higher movements velocities, are indicative of the finer motor control the patients exhibited by the end of their training. Movement smoothness is shown on both maps as the spread of the colored area. When patients gain more control over their limbs and can make more precise moves the resulting maps reflect this tendency.

Longitudinal sensitivity can be noted on the S-map Figures 8.3, 8.5, 8.7, 8.9, 8.11, 8.13, 8.15, 8.17 and 8.19 and on H-map Figures 8.4, 8.6, 8.8, 8.10, 8.12, 8.14, 8.16, 8.18 and 8.20. The evolutive quality of assessment measure is critical for close monitoring for a dynamic session of virtual reality robotic rehabilitation session. This behavior can be considered as a sensitivity property of this measure and merits further investigation to associate it to the functional level scales.

An interesting dimension of interpretation was the ability to visualize the different stategies adopted by the patient at different stages of recovery. We can note for instance ballistic launch movements with major overshoots, fragmentated moves, persistent curvatures in the movement path among other properties in the form of movements he executed. These strategical movements are presented visually and can thus be seen easily without the need to interpret the numerical values of speed or smoothness to identify their presence. This characteristic of the visualization is an important gained benefit in the study of the movement quality and might provide a missing link to the observational scales used in practice today. An interesting finding is an increase in the maximum values of the probability density function as the rehabilitation progressed as observed on the S-map.

On the example of subject C2, while smoothness was not a characteristic of the post- treatment results we could still note increased densities in some regions suggesting that for some targets an upper arched movement was a form of compensation to correct for directions and attain the targets. Another condensed area suggests that directed fast moves were still feasible.

A likely scenario would be the case where targets appeared in the same direction and thus permit the subject to continue moving without changing his direction to catch the next target and benefit from the momentum of the arm and exoskeleton to reach higher velocities. A result that is comforted by the H-Maps reporting typical x histograms with dips in the middle of trajectories and higher peaks near the targets area.

The fact that the movement became more compact and started concentrating on a restricted patch between the start and the target position permits the density function to estimate higher values compared to the beginning of training. This can be considered as an additional numerical measure that indicates the rehabilitation status. This property was present for most patients on the exception of subject S2 and C2 where the maximum probability density value was higher at pretreatment maps, which upon following on the course of their recovery seen through the visual maps revealed fluctuating values of the maximum density.

The benefit of visualization is the ability to make a simple distinction by simply comparing to an ideal outcome or those from previous stages, a comparison that does not necessitate neither elaborated training, nor advanced expertise in the domain of rehabilitation. Thus enabling it to serve the need through robotic or telerehabilitation systems as a patient first feedback tool. In a similar fashion authors [90] reported the patient's experience favoring the use of the proposed arm movement's density plots as a report on rehabilitation progress. We thus attempted to provide an examplary form of report to provide for patient's after training.

In Figures 8.21 to 8.28, we introduced a form of a report that is intended to provide previous training visuals (leftmost) and the current ones (in the middle) and those expected from a healthy subject (rightmost). A typical report that can be furnished to the patient or practitioner at the end of the training session. The report is aimed to give them a clear understanding of the performance progress or deterioration and the ideal they would be looking to achieve.

The investigation of the measure of maximum density value suggests, although not decisively, that the metric which represents a core assumption of the S-map visualization was significantly higher at the end of the training. This comforts the assumption that we made by design where we expected the movement to be more dense and smooth by end of training according to the kinematic studies in literature.

The attempt to justify the clinical validity of the tool was by correlating the measure to the FMA scores. While the result was not significant, the monotonic ρ coefficient was $.52$. This value is considered moderate correlations and is close to the reported values for other kinematics. The longitudinal evolutive quality and sensitivity can be noted by examining the results from the case series presented earlier. The effects are notable and even where densities are within similar ranges differences in strategies and movement preferences are notable on the maps.

Often when designing a study while using kinematics, researchers are supposed to pick the set of metrics that are known to relate to the particular aspect of the study they are interested in. The maps on the other hand condense many layers of information, representing kinematic measures such as velocities, smoothness, overshoot and spatio-temporal ranges in addition to the strategical choices and persistent constraints. When added to the visualization ability to illustrate by contrast help get both practitioners and patients to understand the maps with minimal training. From a mathematical stand point, these multidimensional variations are reduced to four dimensional model. Hypothetically, bringing variables of interest into the same plane adds the additional information on covariance of the variables which is an added dimension of interpretation.

These observations are not attainable using the traditional kinematic measurements and are obsecured by a single numerical value. The evolution of many patients showed fluctuations of the density function and thus of the patients' recovery between sessions and this is validated against the FMA-UE. The constructive validity comparison with FMA can be a validation of its clinical usefulness in detecting the longitudinal changes between sessions and can prove useful for adapting training as well.

The computational motor learning dimension of the results presented in the chapter is also of importance. The smoothness of the movement besides the speed increase are characteristics of the learning phenomena as represented by effective motor control. The spread of the movement as estimated from the data generates a spread model in the form of the bivariate model estimated from the exercise data. The model density is then gauged for its highest value using the most likely trajectory here represented by the locally weighted regression model to define a measure of recovery.

The measure attests that the model increasingly associated higher densities with the most likely observed path. This effect is notable since the smoothness of the movement is assumed as the quality of ML-driven recovery. At the same time accounting for the data-driven approach relying solely upon the observed data to define the likely trajectory. The likely trajectory is then associated with the highest density value monotonically.

One limitation to the study is the lack of validation of the resulting tools by patients and practitioners. The visuals have to meet the needs of its end-users. This cannot be achieved without actively integrating them into the design loop and properly conducting evaluations of the resulting graphics [167].

Another limitation would be the data quality since the approach is dependent on the entry data, results are as representative as the data that is fed into the generation algorithm. Similar challenges have been reported as a major concern for the practitioners to adopt such tools [50]. In our study data were collected from an operational exoskeleton system without any supervision to the training sessions which might limit the depth of justifications we can make on outlier or unexpected results.

8.4 Conclusion

In this chapter, we presented a set of visual instruments intended as patient-centered graphics for performance feedback during upper limb rehabilitation training using an orthotic exoskeleton device. The graphics represent many kinematic metrics that are shown to be reflective of the state of the rehabilitation in the upper limb assessment literature.

The results showed that the maps are representative of longitudinal changes in visual forms, the quantitative measure of the maximum probability density was found to significantly increase by end of training. This also might be used as a standalone objective measure of recovery. The results which is the main property of the presented visual maps, would help ascertain its clinical validity although lacking significant correlations to the FMA-UE.

The visual format permits an easily communicated report for the use of patients and practitioners. We argue that the presented form of performance feedback report figuring the prior, current, and objective outcome as maps can be easily understood by the patient and practitioners without the need for training. This form of feedback relies on the visual differentiation between the maps to inform about the progress made and the objective to be approached. We hope this can bridge the gap between the patient, the robot, and the practitioners by leveraging data to build an easy to interpret and accessible feedback form.

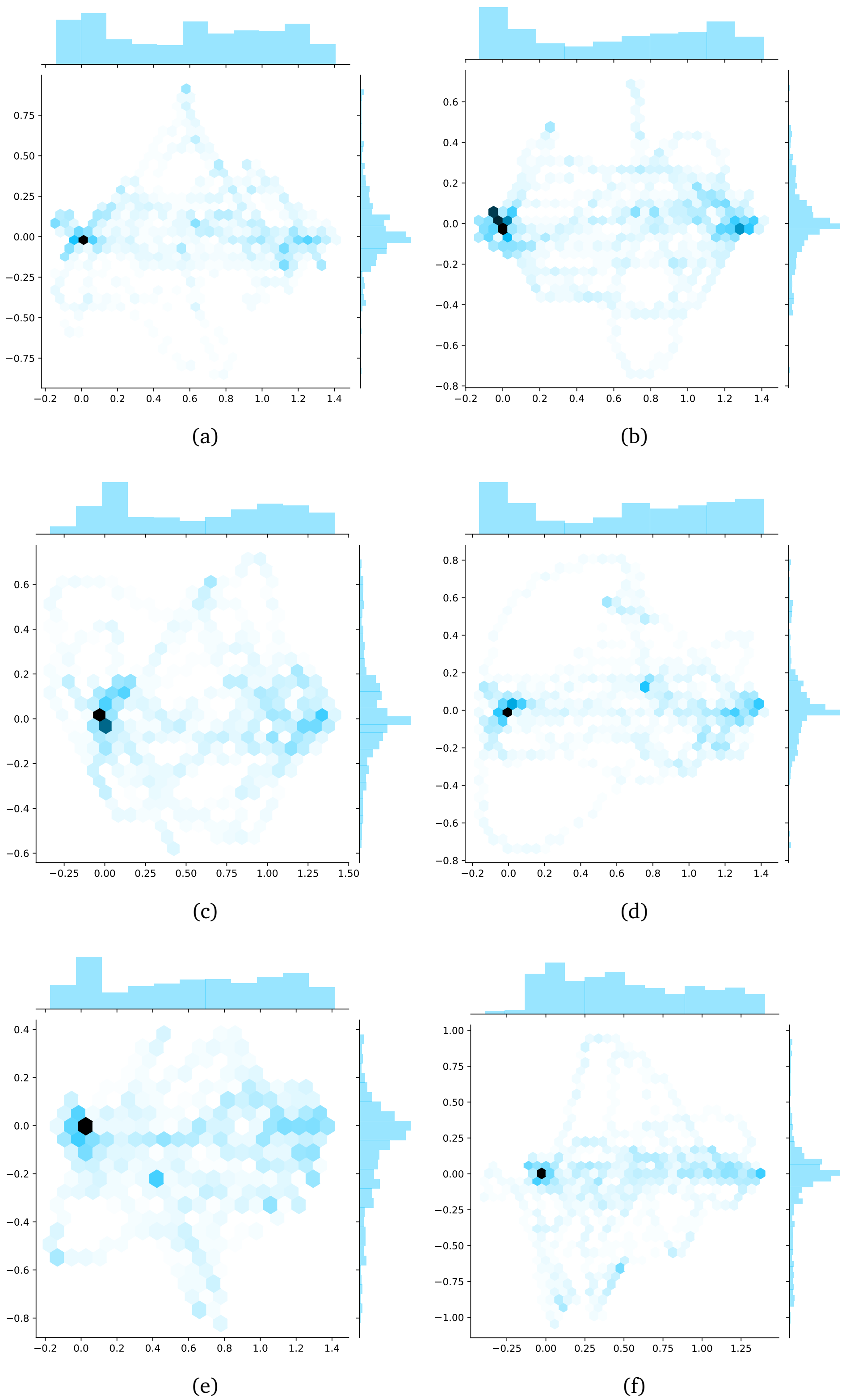


Fig. 8.4.: The longitudinal evolution as represented by the H-map for subject S1.

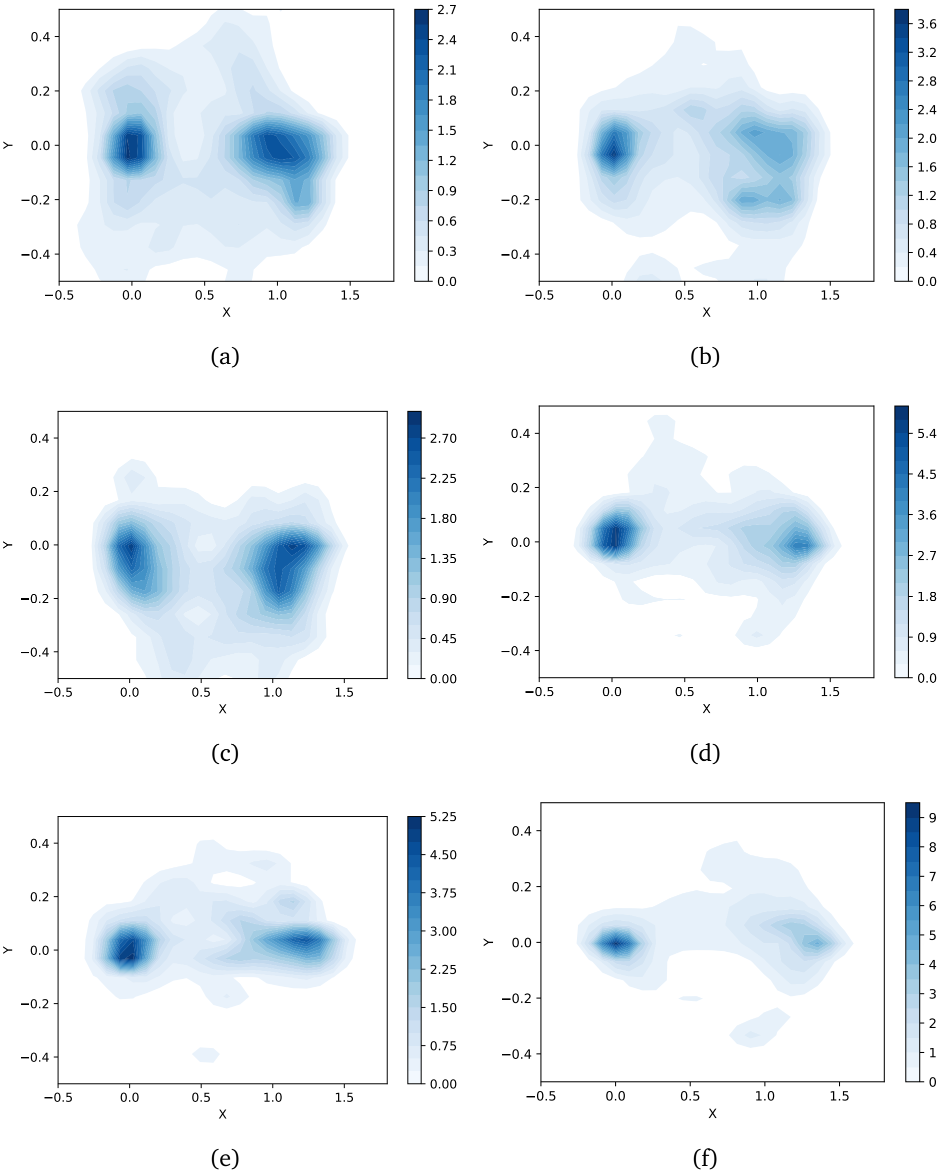


Fig. 8.5.: The longitudinal evolution as represented by the S-map for subject S2.

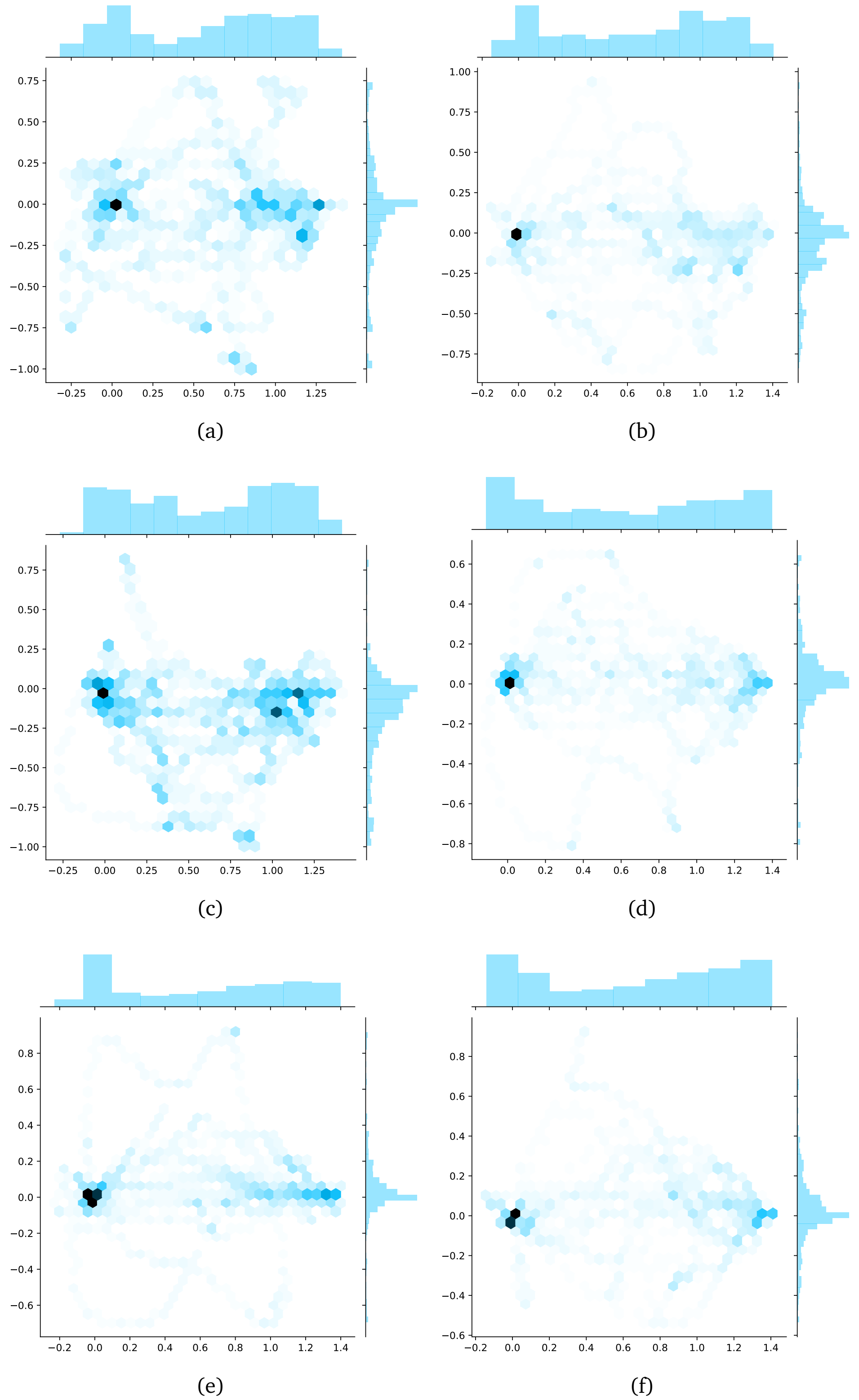
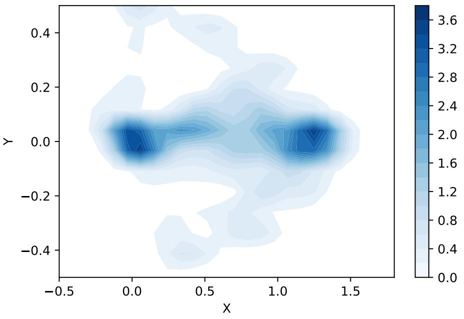
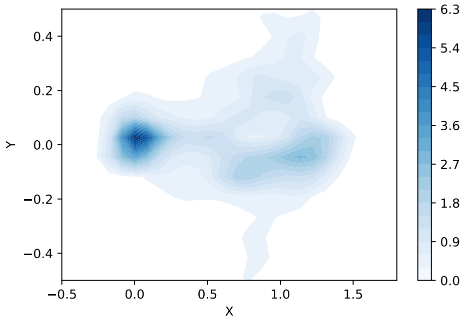


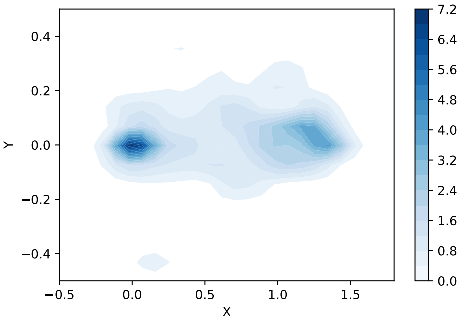
Fig. 8.6.: The longitudinal evolution as represented by the H-map for subject S2.



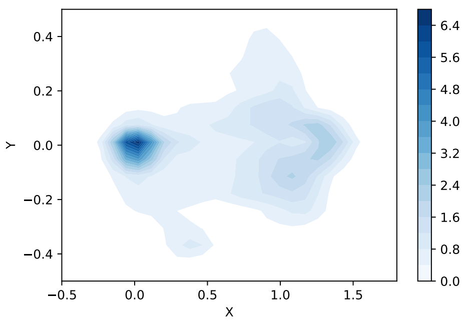
(a)



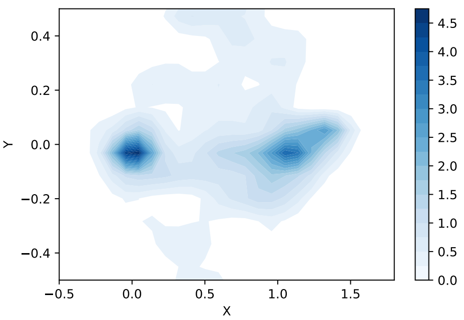
(b)



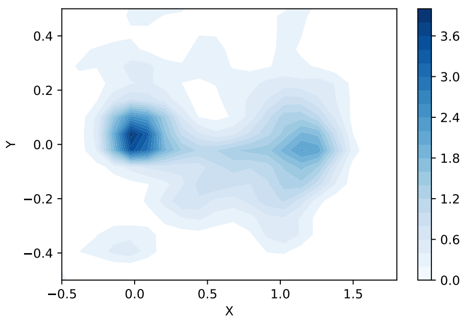
(c)



(d)



(e)



(f)

Fig. 8.7.: The longitudinal evolution as represented by the S-map for subject S3.

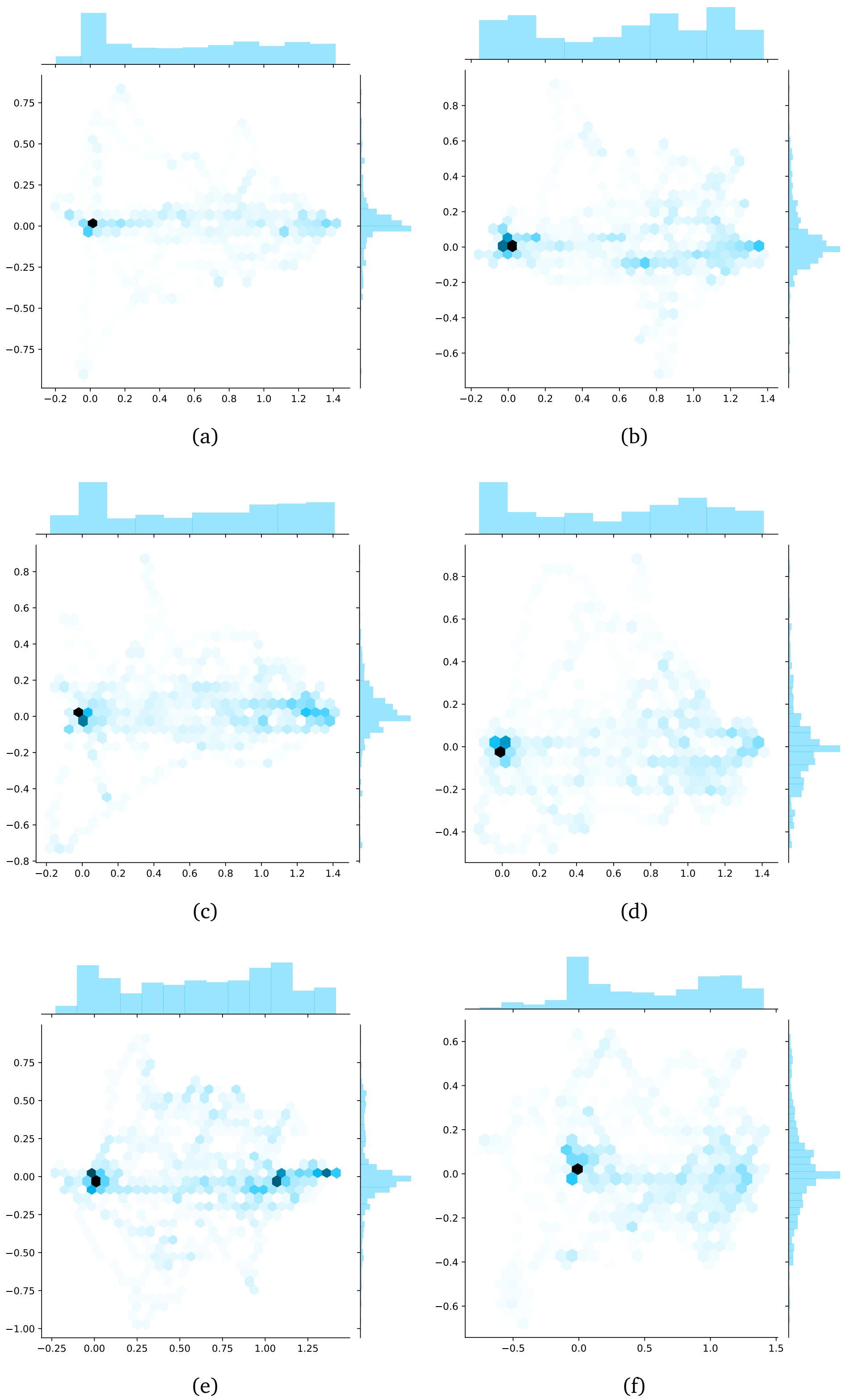


Fig. 8.8.: The longitudinal evolution as represented by the H-map for subject S3.

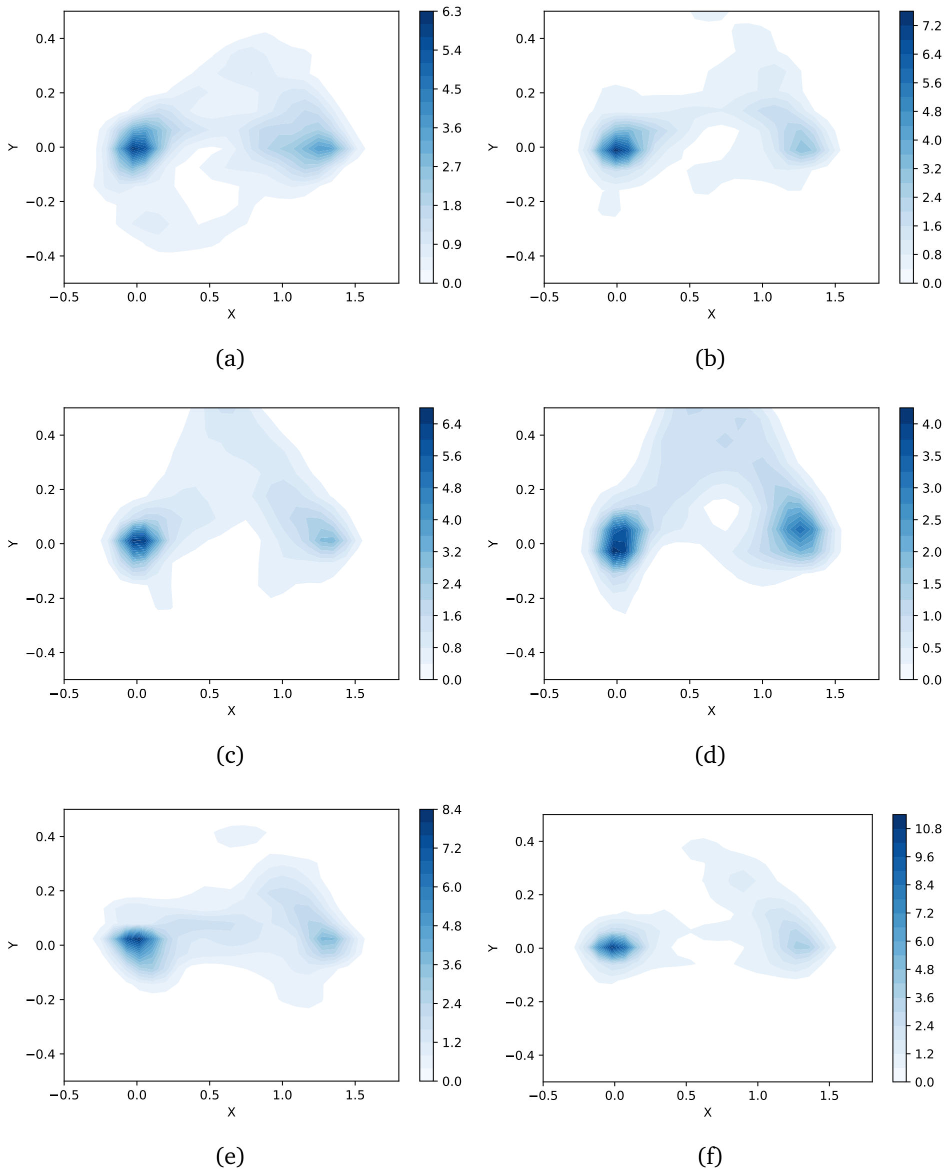


Fig. 8.9.: The longitudinal evolution as represented by the S-map for subject S4.

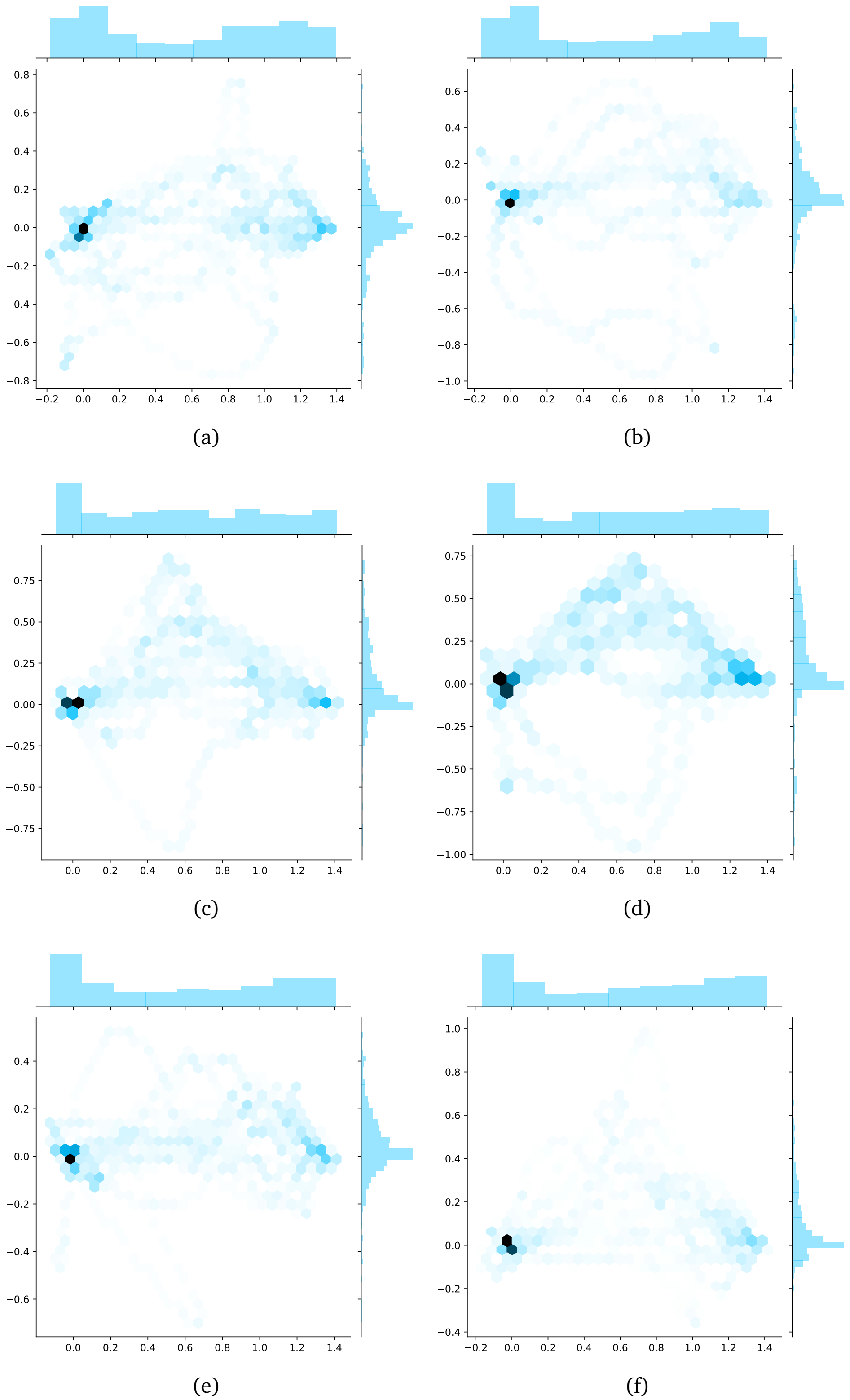


Fig. 8.10.: The longitudinal evolution as represented by the H-map for subject S4.

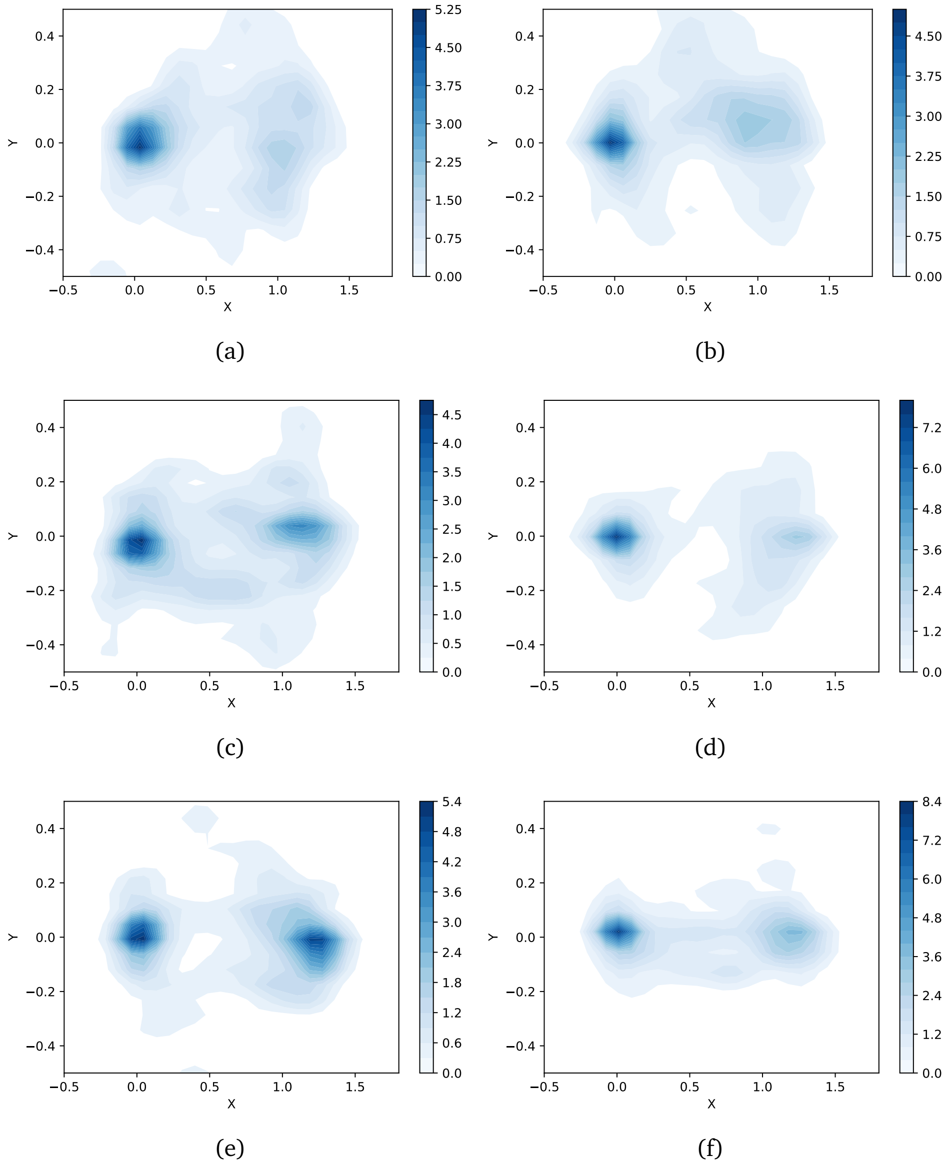


Fig. 8.11.: The longitudinal evolution as represented by the S-map for subject S5.

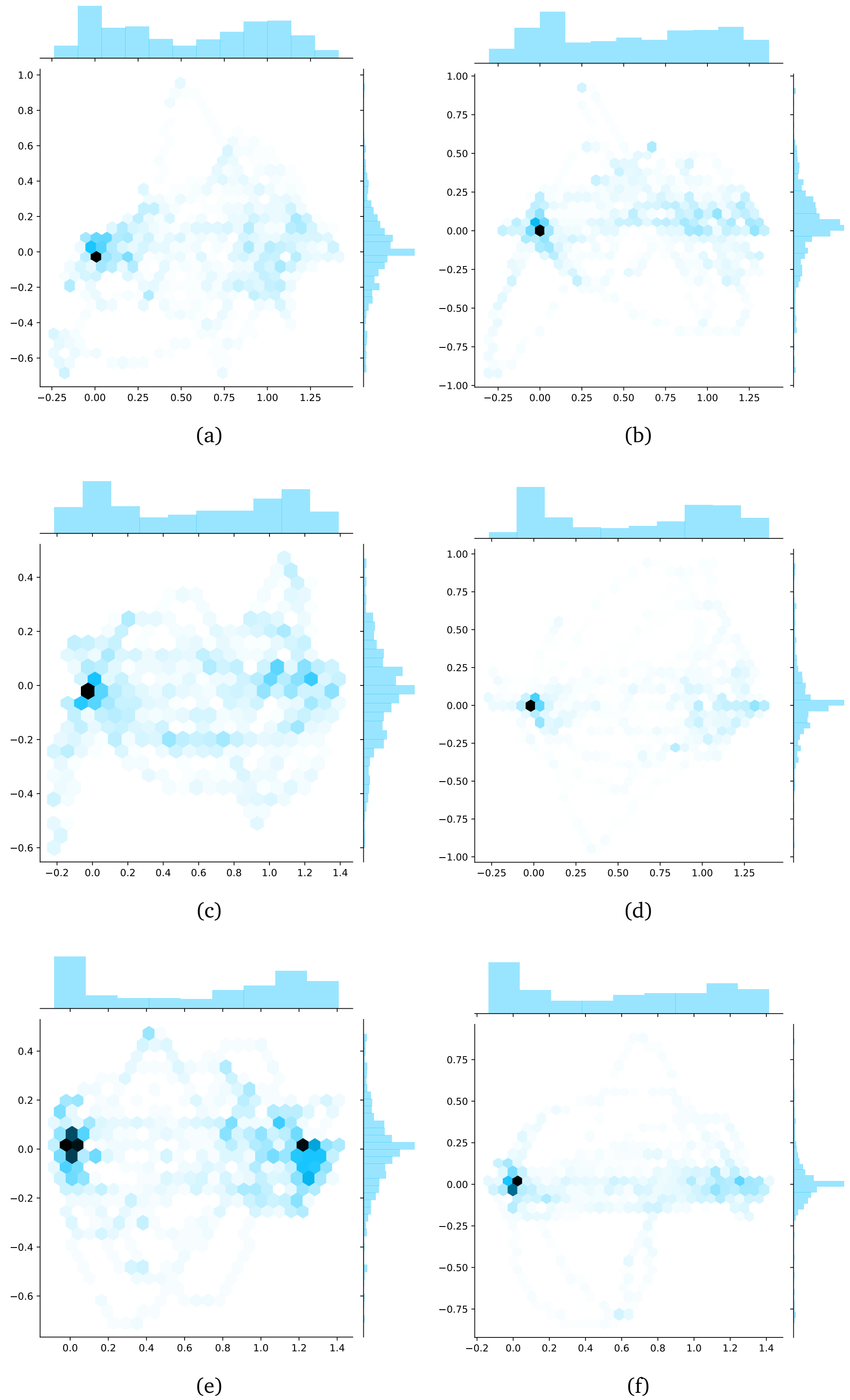


Fig. 8.12.: The longitudinal evolution as represented by the H-map for subject S5.

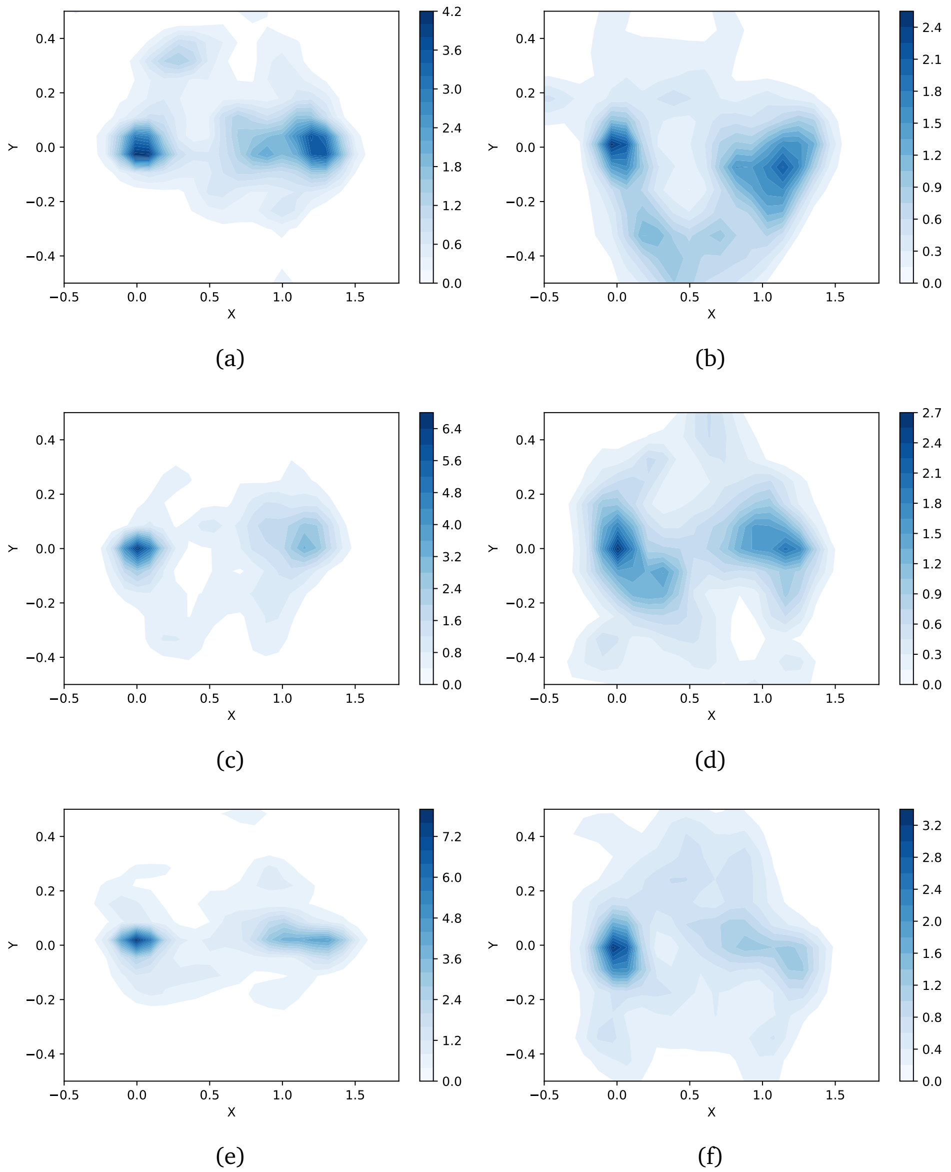


Fig. 8.13.: The longitudinal evolution as represented by the S-map for subject S6.

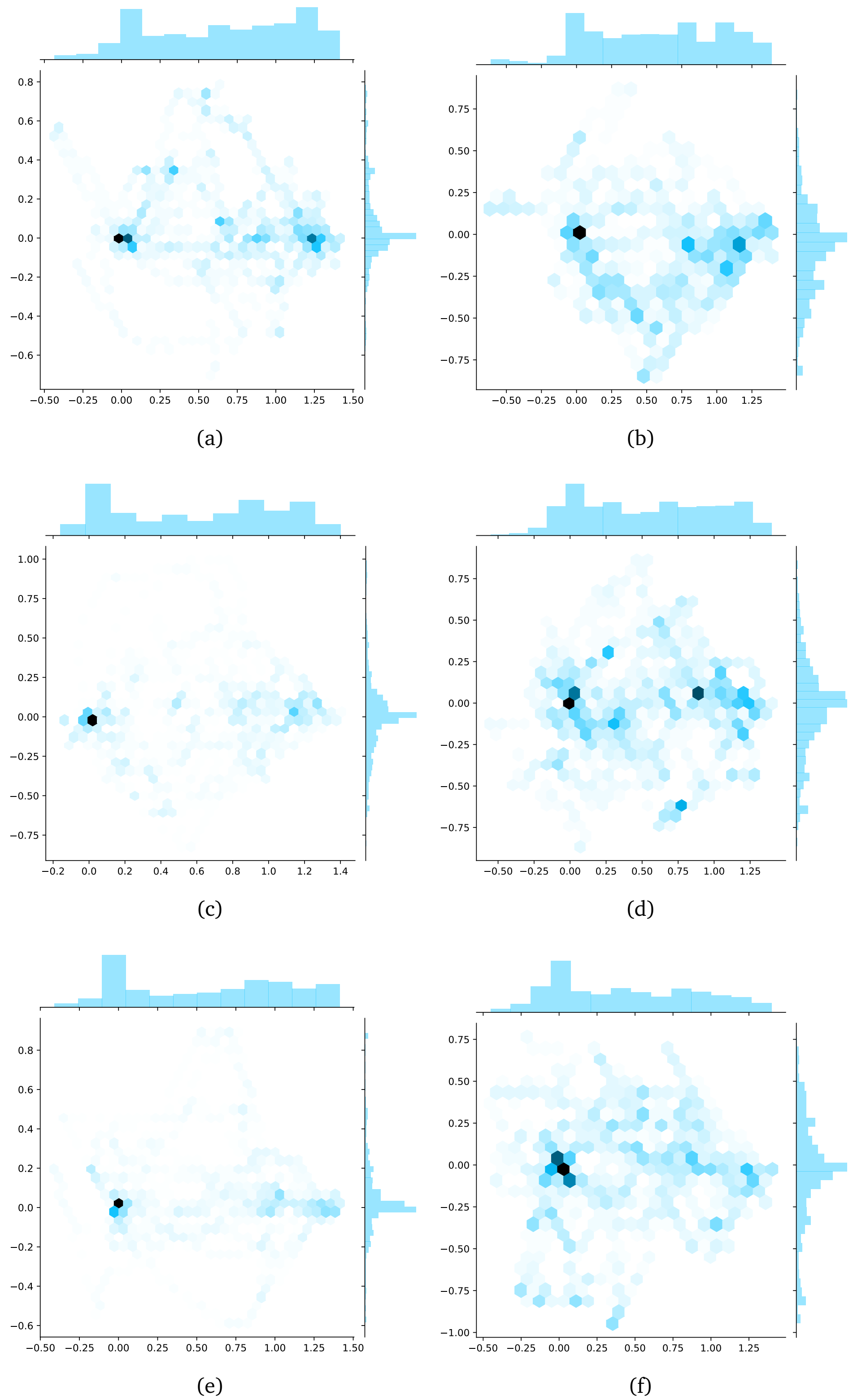


Fig. 8.14.: The longitudinal evolution as represented by the H-map for subject S6.

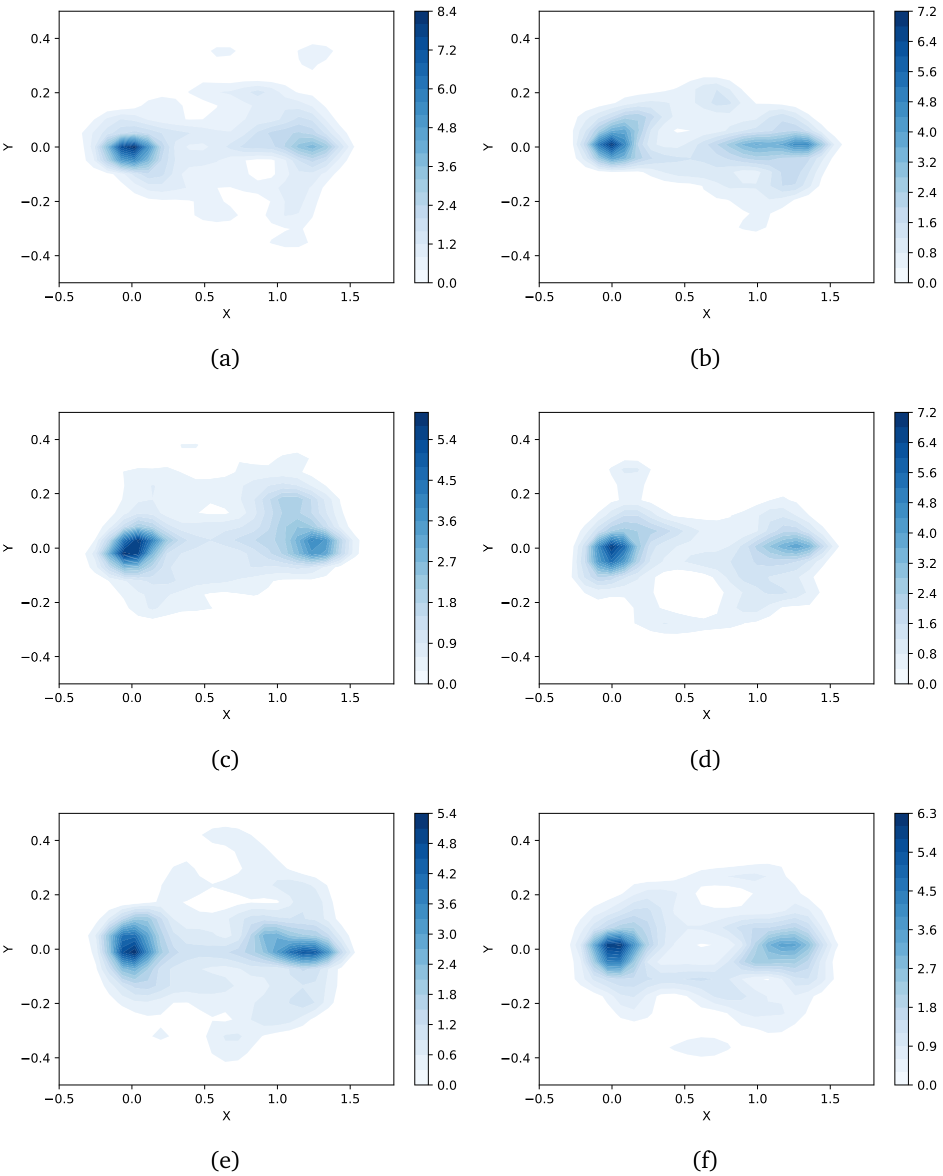


Fig. 8.15.: The longitudinal evolution as represented by the S-map for subject S7.

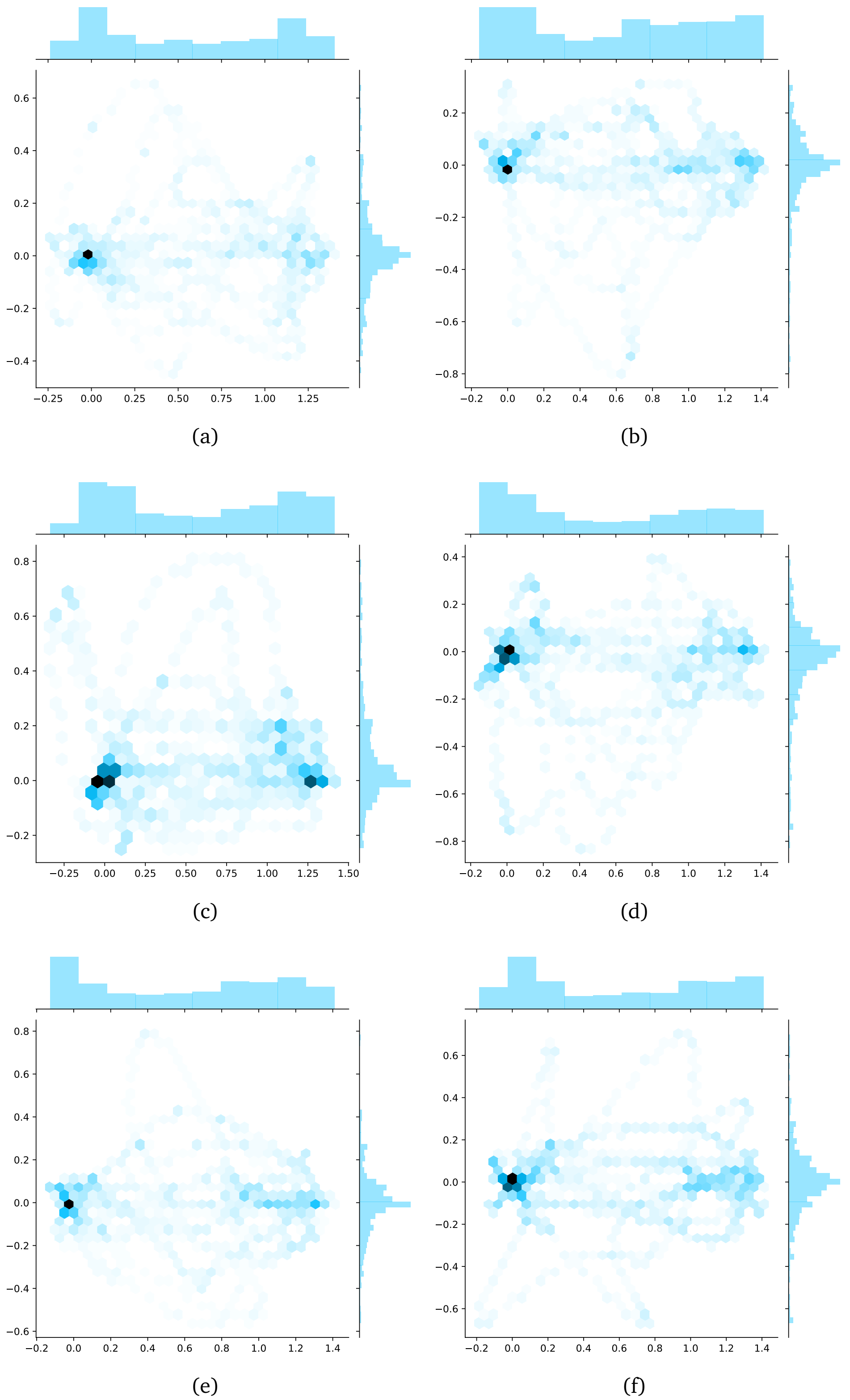


Fig. 8.16.: The longitudinal evolution as represented by the H-map for subject S7.

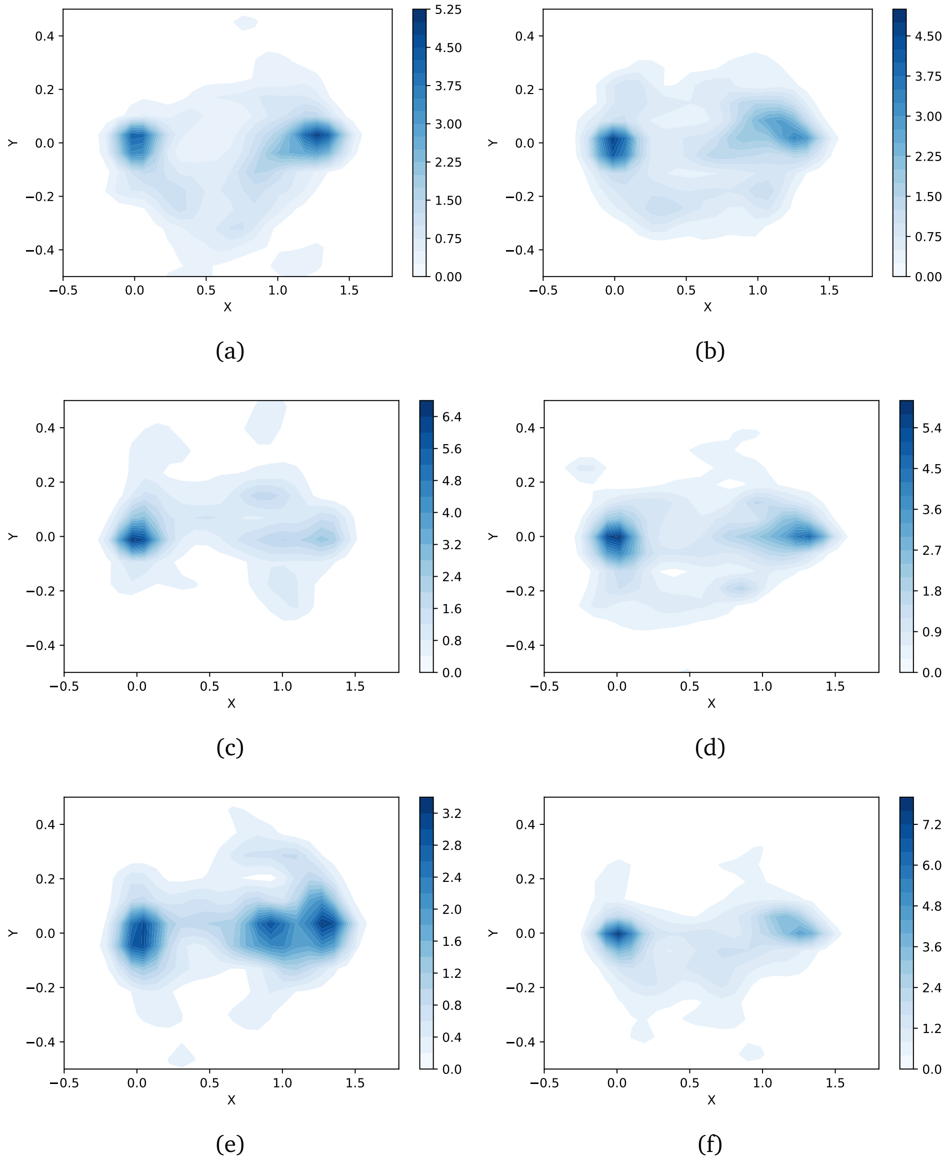


Fig. 8.17.: The longitudinal evolution as represented by the S-map for subject S8.

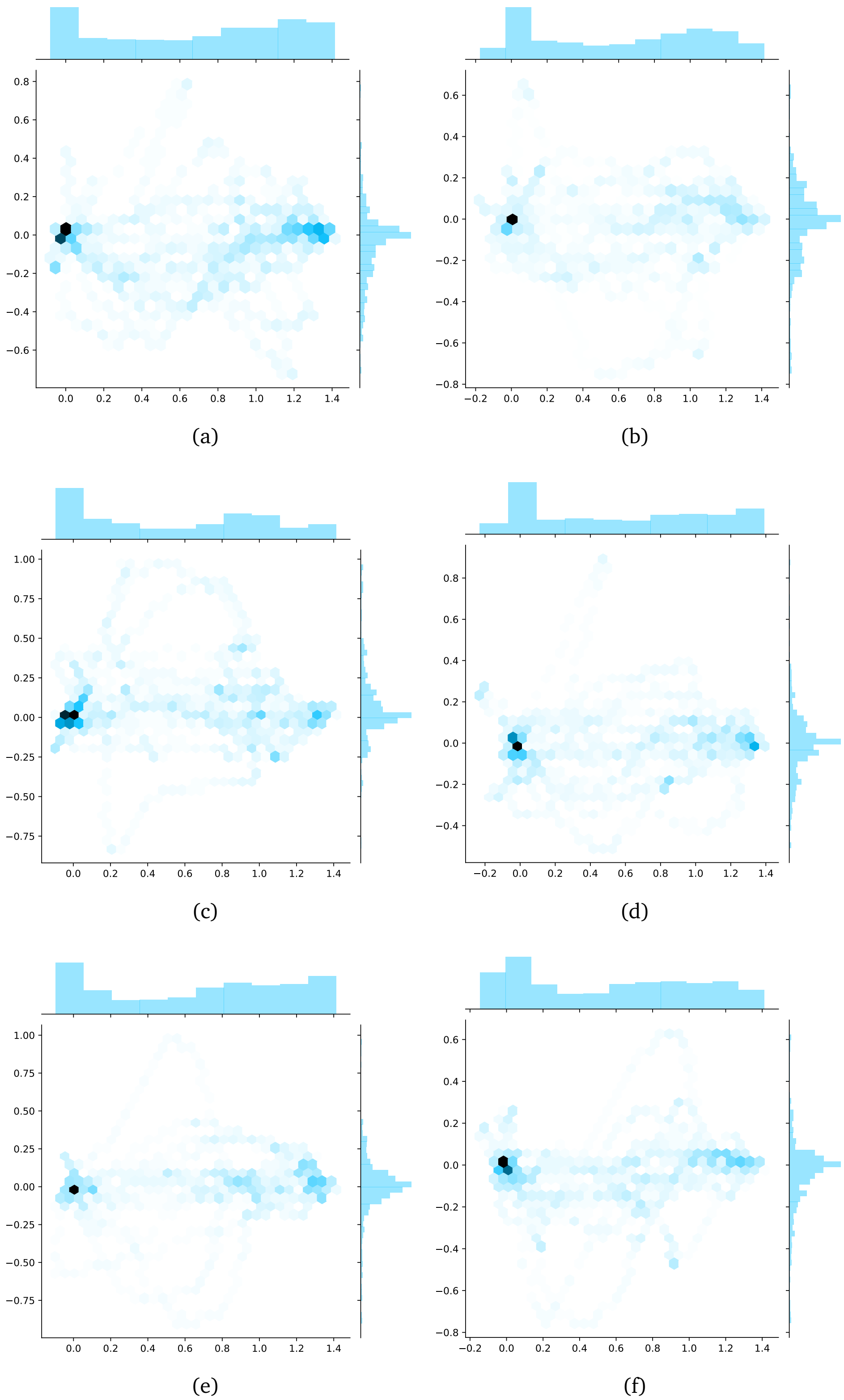


Fig. 8.18.: The longitudinal evolution as represented by the H-map for subject S8.

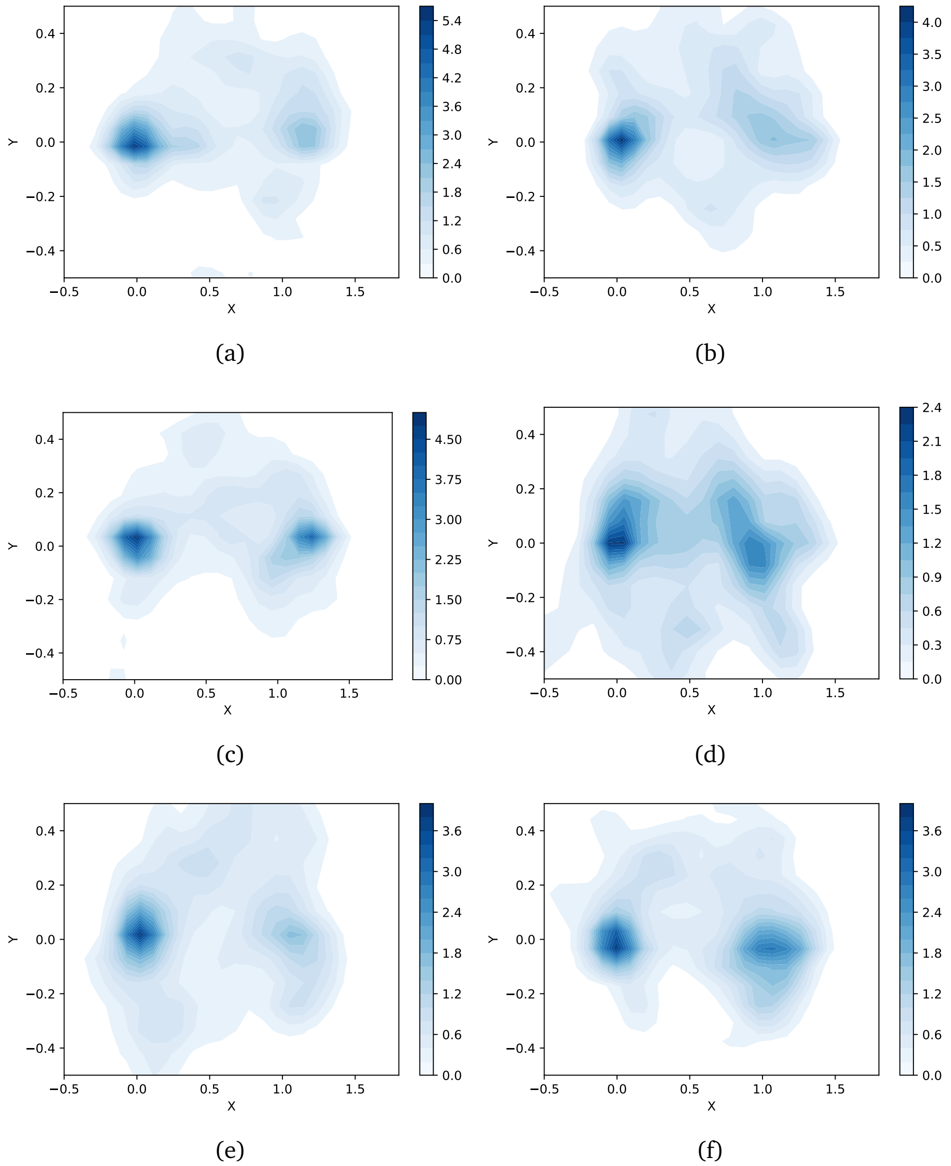
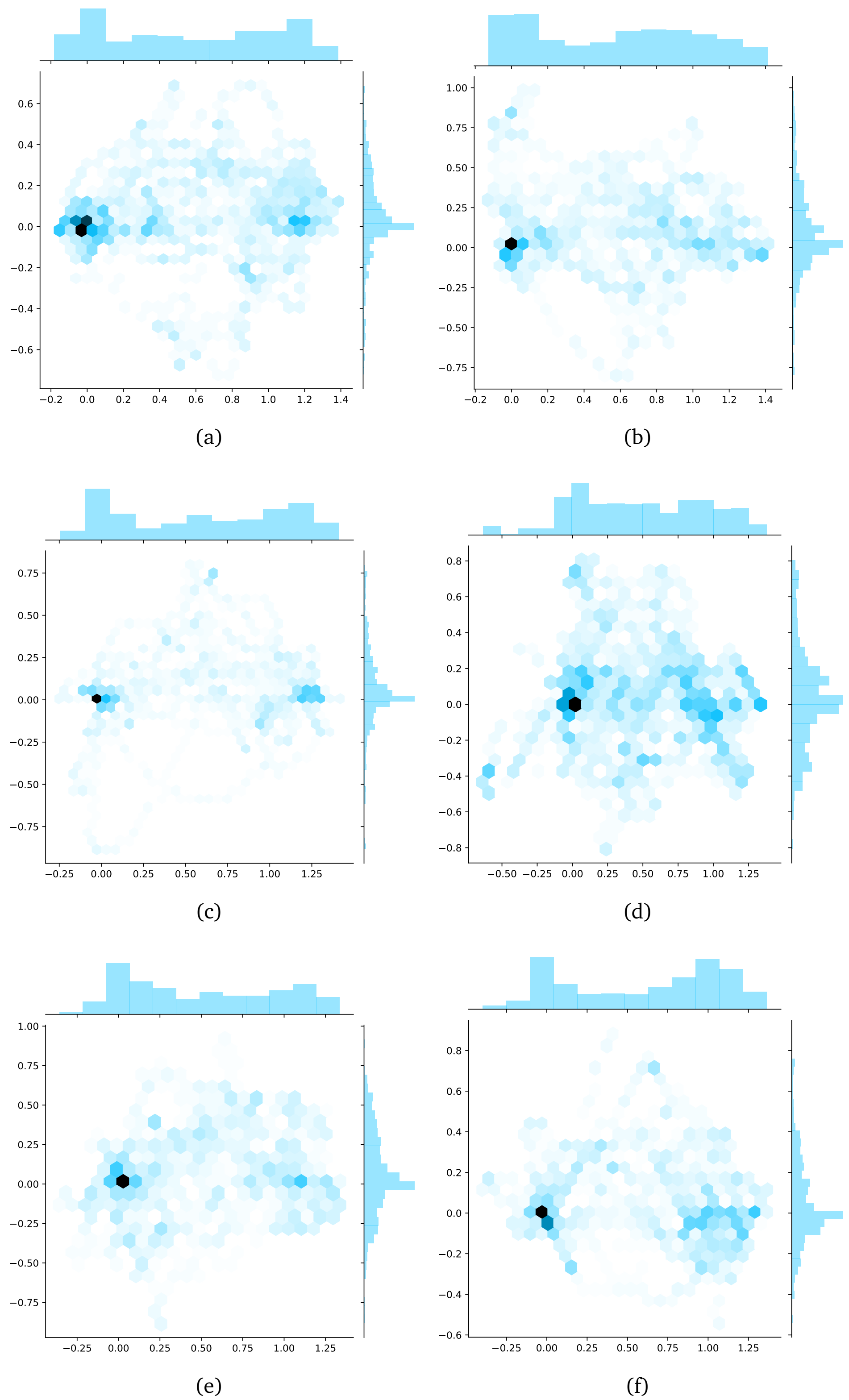


Fig. 8.19.: The longitudinal evolution as represented by the S-map for subject C2.



(a) (b) (c) (d) (e) (f)

Fig. 8.20.: The longitudinal evolution as represented by the H-map for subject C2.

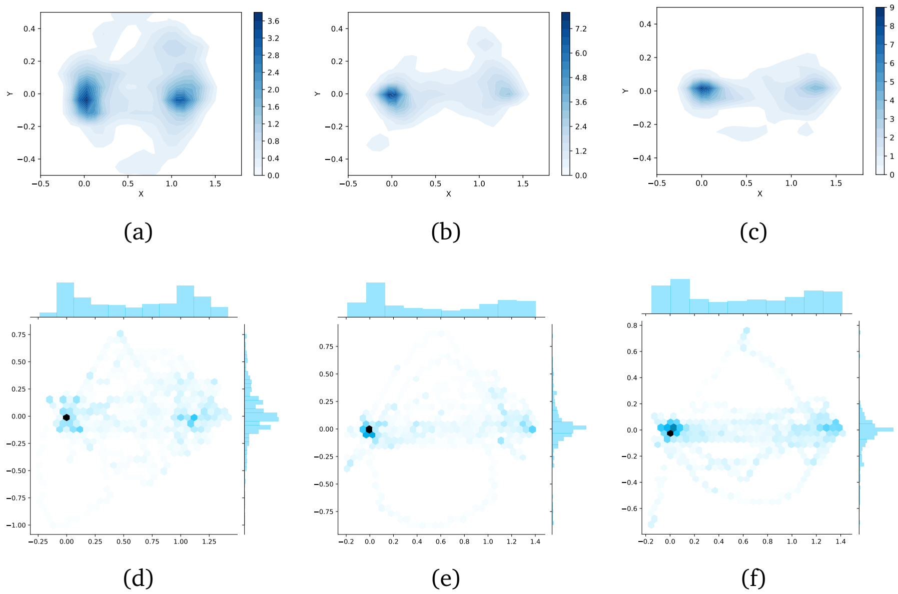


Fig. 8.21.: The maps results at pre-treatment (a,d), post-treatment (b,e) for patient S3 and healthy user C1 (c,f) (left to right).

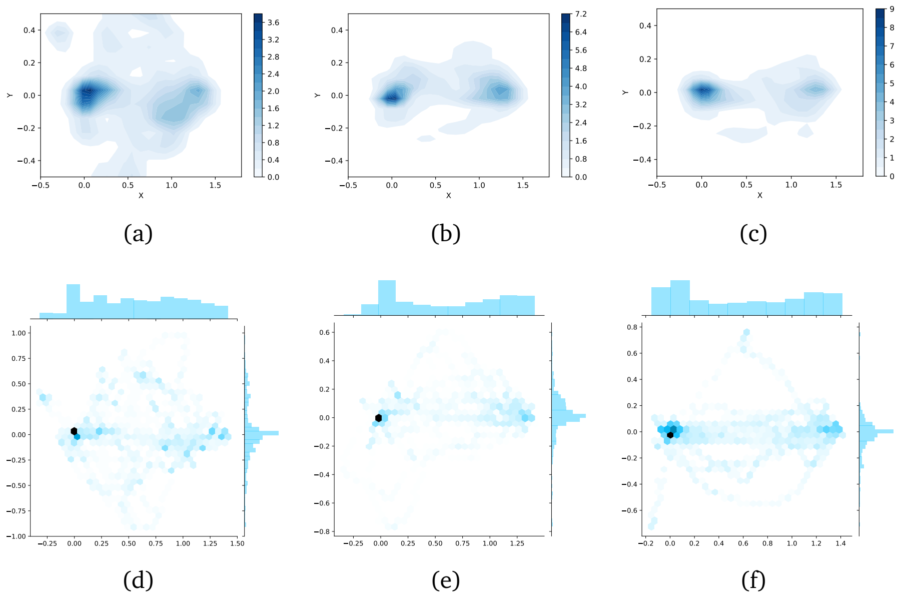


Fig. 8.22.: The maps results at pre-treatment (a,d), post-treatment (b,e) for patient S4 and healthy user C1 (c,f) (left to right).

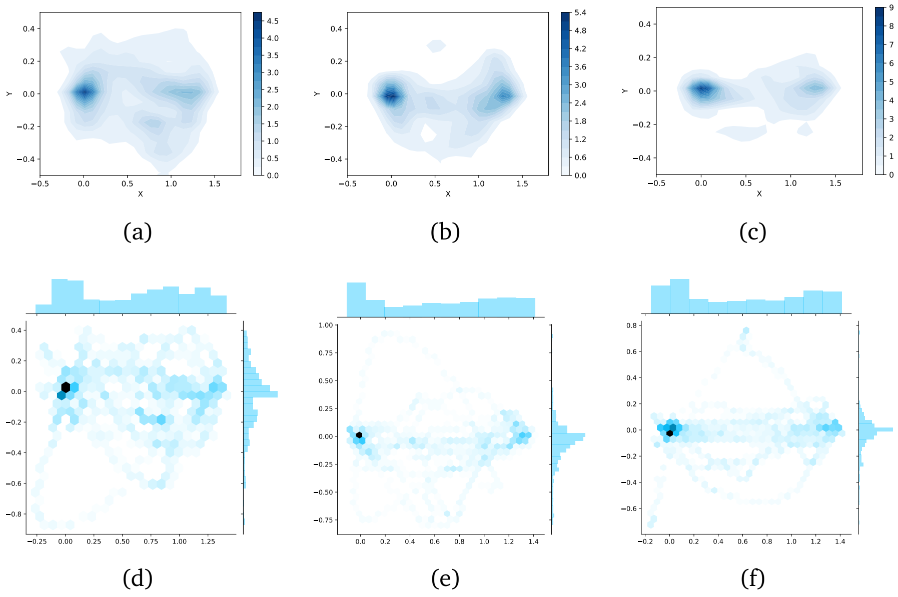


Fig. 8.23.: The maps results at pre-treatment (a,d), post-treatment (b,e) for patient S5 and healthy user C1 (c,f) (left to right).

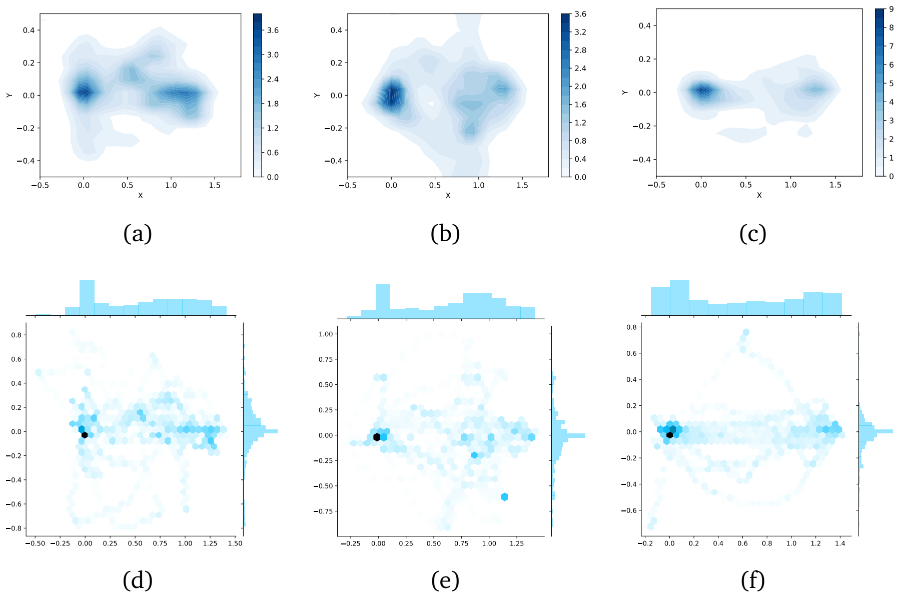


Fig. 8.24.: The maps results at pre-treatment (a,d), post-treatment (b,e) for patient S12 and healthy user C1 (c,f) (left to right).

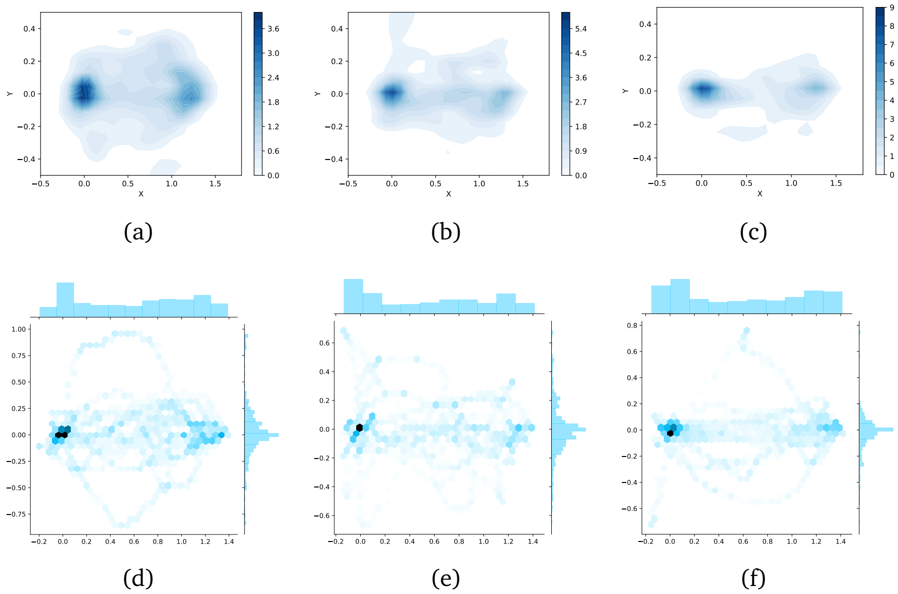


Fig. 8.25.: The maps results at pre-treatment (a,d), post-treatment (b,e) for patient S14 and healthy user C1 (c,f) (left to right).

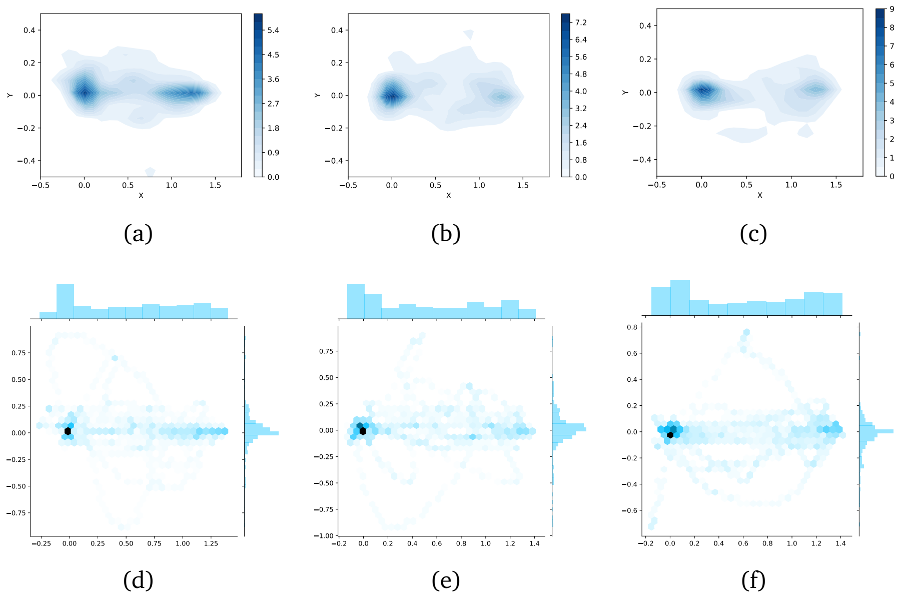


Fig. 8.26.: The maps results at pre-treatment (a,d), post-treatment (b,e) for patient S15 and healthy user C1 (c,f) (left to right).

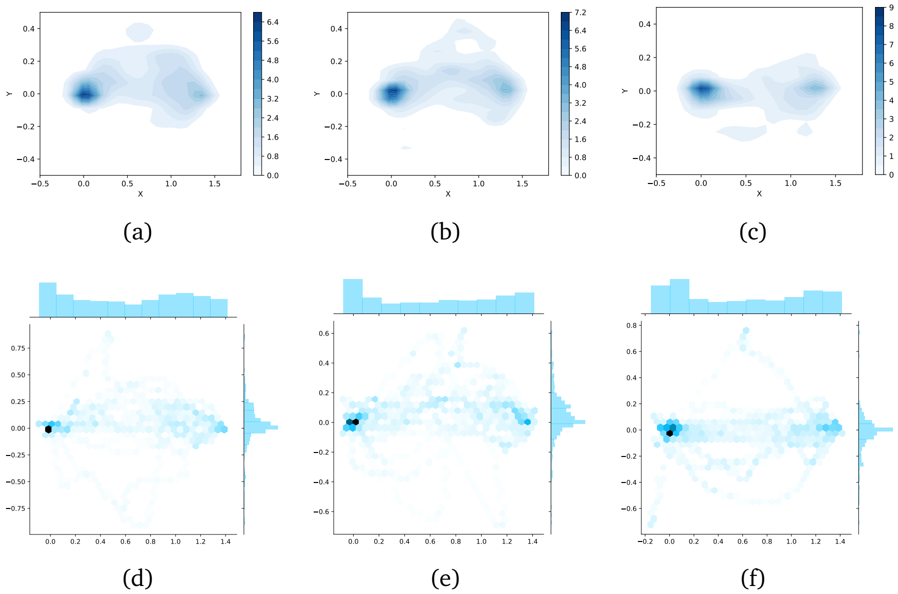


Fig. 8.27.: The maps results at pre-treatment (a,d), post-treatment (b,e) for patient S16 and healthy user C1 (c,f) (left to right).

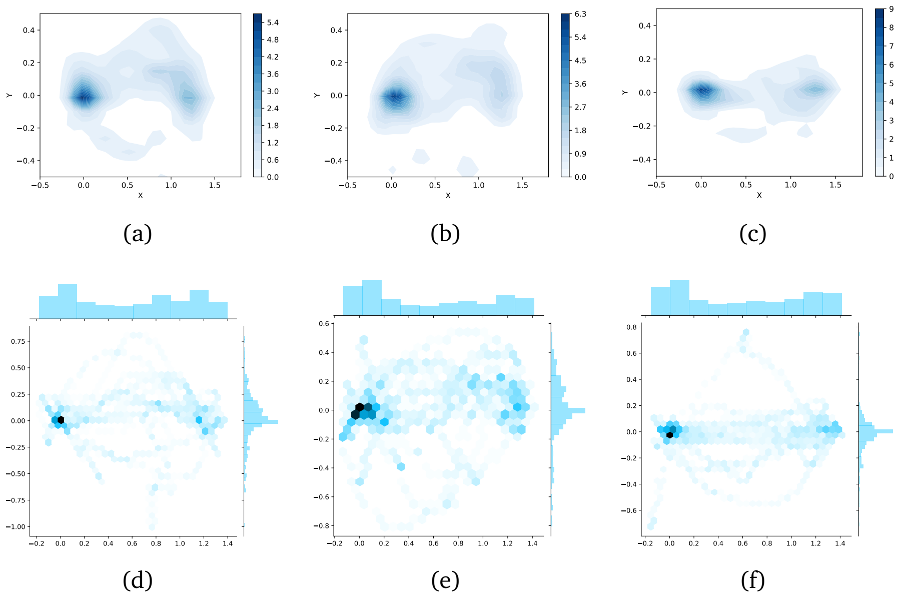


Fig. 8.28.: The maps results at pre-treatment (a,d), post-treatment (b,e) for patient S17 and healthy user C1 (c,f) (left to right).

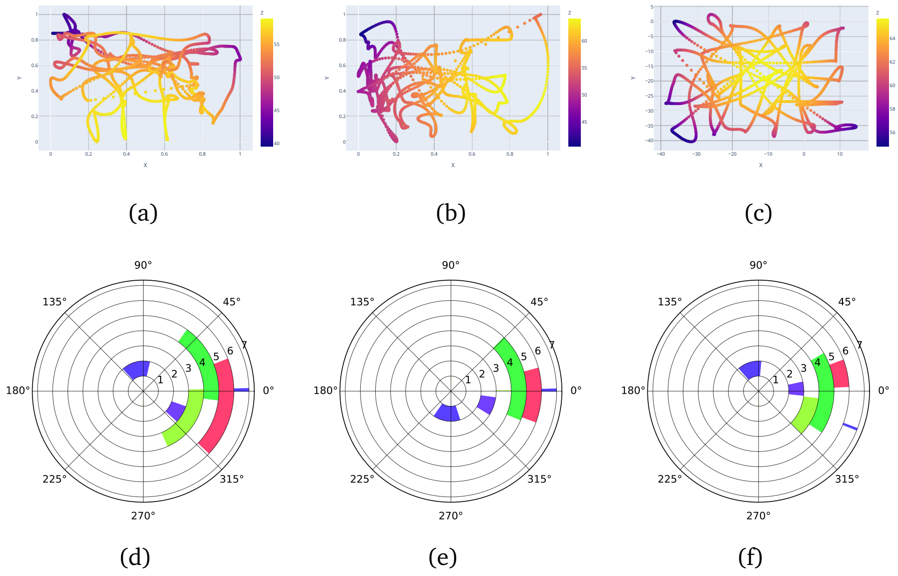


Fig. 8.29.: The 3D representation of training trajectories with color coded depth (1^{st} row). JointRom maps representing joints from distal(1) to proximal(7)(2^{nd} row). Reported results are at pre-treatment (a,d), post-treatment (b,e) for patient C2 and healthy user C1 (c,f) (left to right).

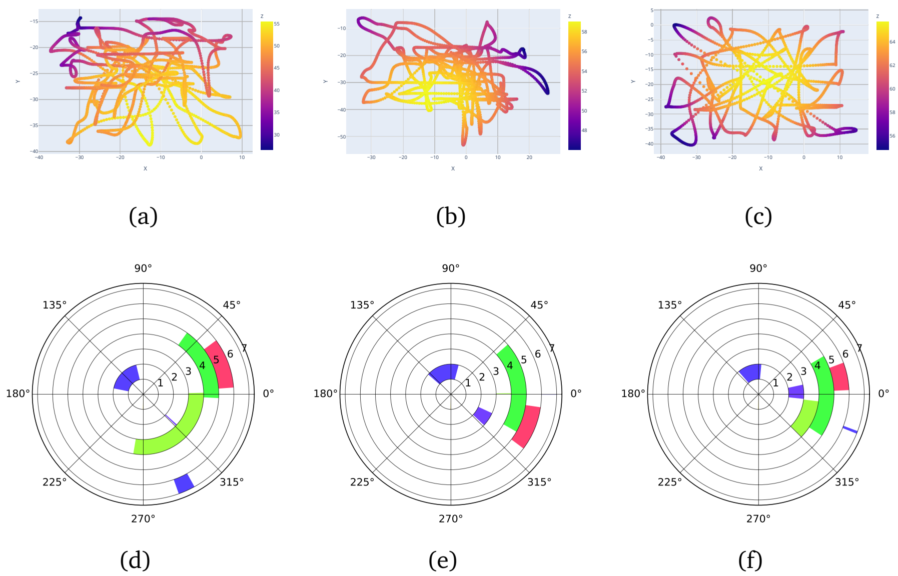


Fig. 8.30.: The 3D representation of training trajectories with color coded depth (1^{st} row). JointRom maps representing joints from distal(1) to proximal(7)(2^{nd} row). Reported results are at pre-treatment (a,d), post-treatment (b,e) for patient S1 and healthy user C1 (c,f) (left to right).

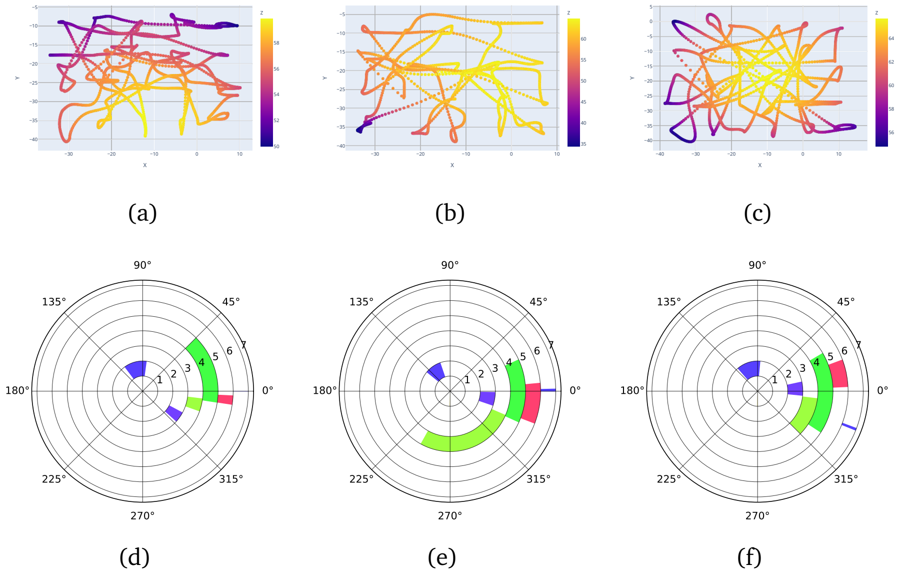


Fig. 8.31.: The 3D representation of training trajectories with color coded depth (1^{st} row). JointRom maps representing joints from distal(1) to proximal(7)(2^{nd} row). Reported results are at pre-treatment (a,d), post-treatment (b,e) for patient S4 and healthy user C1 (c,f) (left to right).

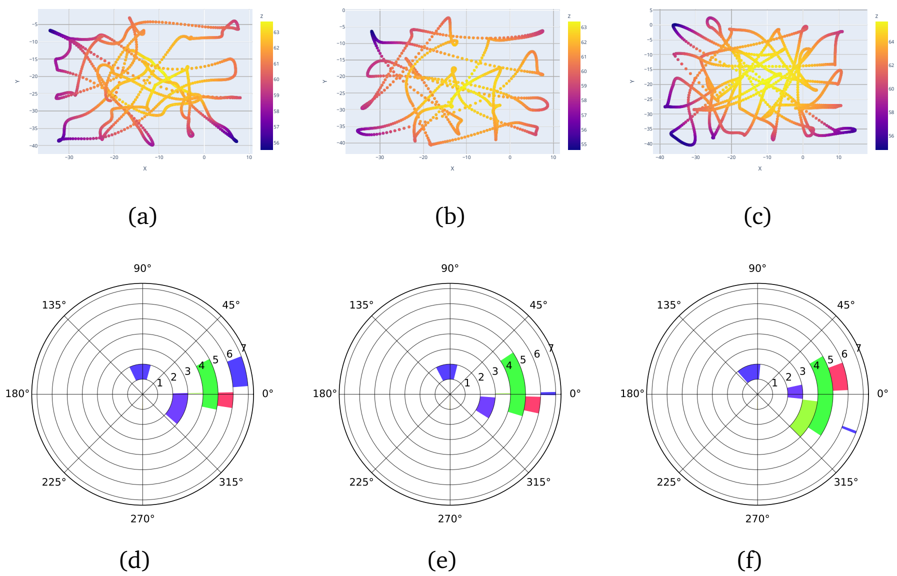


Fig. 8.32.: The 3D representation of training trajectories with color coded depth (1^{st} row). JointRom maps representing joints from distal(1) to proximal(7)(2^{nd} row). Reported results are at pre-treatment (a,d), post-treatment (b,e) for patient S7 and healthy user C1 (c,f) (left to right).

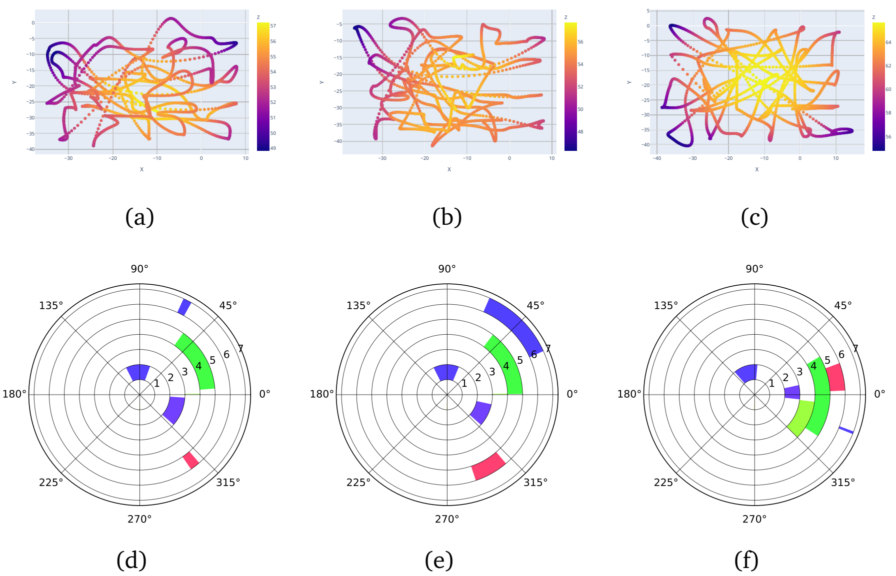


Fig. 8.33.: The 3D representation of training trajectories with color coded depth (1^{st} row). JointRom maps representing joints from distal(1) to proximal(7)(2^{nd} row). Reported results are at pre-treatment (a,d), post-treatment (b,e) for patient S18 and healthy user C1 (c,f) (left to right).

General conclusion

9

> *The field of rehabilitation has much in common with the field of motor learning ... therapists are involved in the treatment of patients with neurologic injury are concerned with issues related to motor relearning*
>
> — **Ann Shumway**

9.1 Overview of the main contributions

The research into robot mediated rehabilitation calls upon many different fields and knowledge backgrounds. As such, contributions of the current work can be viewed through different lenses. We first used operational dataset from the university Hospital Center in Tlemcen Algeria to investigate responsiveness of the major consensus kinematic scales reported in the clinical trial literature.

The results helped contrast the literature gaps, characterize the measures properties and highlight their limitations. This is specifically true in comparison to the concurrent research directions in the field namely adaptive rehabilitation, computational modeling and individualized patient therapy. Longitudinal sensitivity and evolutive assessment instruments seemed to be a prerequisite to enabling the adaptive and individualized care protocols and to reply to the therapist needs for more patient centered scale. We investigated data derived model identification to estimate a dynamic state space model representing the course evolution of the coordinates during exercise and validated it using unseen data and achieving reasonable prediction RMSE.

The continuous computational estimation of model parameters, was then investigated using an Expectation Maximization iteration to estimate a Hidden Markov Model representation of the trajectory. Spatial data modeling was approached using an extension of the dynamic model learned based on data and extended further to allow Kalman Smoothing of trajectory data. Once the Extended Kalman smoother run the dynamic time warping allowed the alignment of the observation to the hidden model maximizing the total likelihood.

Considering the framework as a computational approach to modeling a motor task learning and the significant evolution of the model parameter, namely the covariance which enabled proper estimation of the trajectories during smoothing phase, provides validation evidence for a new contribution to this research topic. The parameter reinjection allowed us to restrain the bias observed on the prediction to the within 4% norm, while the covariance changed longitudinally in significance suggesting that the underlying driving mechanisms can be approached deterministically.

The learned model is then used to estimate a set of metrics that were shown to have good responsiveness qualities among the investigated clinical properties. The clinical validation of the results also showed overall a consistent nature in the results and a promising validity compared to the FMA-UE although non-significant.

The learned covariance presented high correlation to the FMA and had significant monotonicity which consists a main advantage to the current model. It relies on basic motor learning principles in the formulation of the dynamic model and the covariance represents the amount of variation that is not accounted for by the model and different from the prediction error as this is compensated for by the bias term.

In relating needs to the findings of our research, basing the model on Motor Learning Properties and leveraging sensitivity of the metrics enabled detecting changes regardless of the impairment causing injury which has been related by practitioners as a major challenge for the current scale. The clinical assessments has for each pathology population its major scale of reference. Pending further clinical validations, the pathology-agnostic nature of the scale might prove a useful property to relate to the therapist as caseloads varies significantly during practice. Adapting to this would necessitate no further changes than considering the underlying phenomenon driving the change observed through the scale. This is also of importance as we can note that a unified outcome measure is one of the recurrent needs in clinical practice to enable standardization and proper management, evaluation and planification within care facilities.

From another perspective, the patients often require -in as much as the practitioners do-, tools for feedback that are intuitive and can help them communicate internally as well as with patients while setting therapy goals for instance. This was approached using some principles of motor learning where we assume that smoothness, direction error and movement velocity increase with recovery as a result of the learning phenomenon and we model this through exercise data. The visualizations represent a number of maps of preprocessed or raw exercise data. The longitudinal sensitivity as well as the overall pre post changes were investigated through a derived measure of maximum probability. The results showed that the observation is significant between pre and post data and had a moderate correlation to the FMA scale adding to the evidence in favor of the visualization proposed.

A list of the contributions is as follows:

- Dynamic State Space model estimated through exercise data;
- An HMM model of the reaching trajectory;
- An EM learning framework to adaptively learn model parameters;
- A set of probabilistic Bayesian measurements as assessment instruments of rehabilitation;
- Validation and clinimetric investigation of the properties of the new instruments;
- Primary results were pathology-agnostic (significant effects regardless of the incident causing impairment);
- A computational Motor Learning Task Modeling with validation on clinical data;
- A strong evolutive property capable of enabling per session adaptation;
- A modular framework with possibility for additional signals estimation and dynamics incorporation;
- Data visualizations techniques capable of longitudinal evolution assessment;
- A quantitative objective measure based on the visual model validated statistically and against the golden standard.

9.2 Executive summary

In this work we approached rehabilitation assessment research. While the use of robotic enabled too much data there is still next to none research taking full advantage of this data to enable further understanding and capabilities for research and clinics. The study of exoskeleton device data from an operational dataset constituted our research object.

Firstly, the needs presented by the researchers in the field, the patients, therapists as well as stakeholders aims at enabling much of the recent guidelines pertaining to client centered therapy, individualized care and adaptation for efficiency motivations. The current objective scales while providing some solutions to the major clinical scales limitations are themselves limited compared to these pending needs.

We investigated modeling trajectories as a mean of assessment and evolution parameterization. The data driven approach to research enables us to have little intervention and rely more closely on what the patient is doing in estimating our scales. A Hidden Markov Model was proposed as a latent trajectory model, and was then estimated using a Kalman Extended Smoother. Dynamic Time Warping was also used to ensure proper alignment due to the higher model resolution. The model enabled us to define a set of measurements that showed consistent evolutive property with longitudinal sensitivity to detect between sessions differences essential for the adaptive training philosophy.

Computational approaches to Motor Learning Modeling are very few in the literature and few still are based on the dynamic assumptions derived from ML literature. This was a strong contribution of the current model as it uses basic properties in its model formulation and relies heavily on data to estimate the rest of the parameters that were shown to have a clear evolution tendency. This is important result as this is the effect outside of the observed bias noted in the prediction model. This tells us that the measure is representative of all the variations not accounted for by our assumptions. This is for instance notable once we visually inspect the trajectories and look how different they are yet fail to detect these changes using usual end point kinematics.

Visual data representation as a feedback form or as a data mining tool has not been fully investigated in the rehabilitation context. We approached the visual feedback using the same basic ML assumptions and relied on the literature's promising response of the patients to probability maps visuals to adopt them as a visualization form. The model was shown to densely present the longitudinal changes and to significantly differ between pre and post assessments as related by a model derived measure representing the maximum density value. This can be expected as recovery results in better directed movements and more smooth trajectories that would necessarily increase the probability around the start position where speed drops and directions change are initiated.

A holistic approach to the needs reported as well to the current challenges was attempted and the results were investigated using highly heterogenous dataset from operational rehabilitation center which makes the results very promising and indeed potentially useful to both research and clinical practitioners.

9.3 Perspective research

The adaptation of rehabilitation task especially during the exergame sessions can be of crucial importance as it not only holds the keys to enabling more efficient training but also relates to the patients safety. Excessive training and involvement, unnoticeable challenges to movement or psychological state of the users are all phenomenon affecting the performance and can benefit from using finer scales to detect and account for them either semi automatically or automatically to increase the usefulness and the autonomy of the user.

Many approaches to motion capture are available nowadays as solutions approaching telerehabilitation are increasingly investigated and developed into dedicated spin offs and startups. The tools presented here are a wide spanning to fulfill many of the recurrent challenges and are built with modularity to ease adaptation and hopefully adoption into other adjacent use case scenarios.

Collaborative robotics is another domain that can find the modeling scheme useful to define and preadjust motion model and users profiles into a single motion planning framework. Although this might be a bit laborious, we did note that even the ideals used to generate human like trajectories are not accounting for the natural variability in human trajectories let alone those with impairments.

In terms of potential future research many possibilities can be investigated namely:

- The use of the measurements to drive game engine training session adaptation by dynamically changing properties during training which was seen to further enhance the efficiency of training. This is a tangible possibility that would benefit instantaneously the device users without incurring any additional expenses;
- The validation of the tools on homogeneous datasets can be useful to establish its clinical usefulness and validity. As we have listed in the literature review chapter, the evidence from the advanced kinematic assessment techniques are intriguing in terms of the dimensions of extension to the currently used tools. The findings applies to the results in our own modeling framework since immediate validation can benefit from existing assessment trial dataset

to provide extra clinical validity at minimum overhead of both expenses and time;

- Investigating the time alignment properties and different chains model length. The model parameters were not heavily assessed as the training method's sensitivity to parameters change seemed to be high. A proper systematic approach to assessing the parameters choices and potential simplifications can prove beneficial to the deployement version of the software as a research tool and as assessment tool;
- Looping in practitioners and patients to further the design and the validation of the visualizations is crucial after establishing the tools potential and qualitative information its provides. Working to extend the tools accessibility is an indispensable phase before moving the tool to use at the rehabilitation center;
- Investigating further effects using the dynamic model proposition as noted by the computational modeling of motor learning. We denoted that the covariance parameter estimated consistent trends while all other parameters were held constant. The investigation of potential mechanisms apparent in the form of deterministic variation of the parameter can help link between the different factors influencing the trajectory of the patient into a single modeling framework. Quantifying the different effects on trajectories can open doors to proper dissociation of the different cofounding factors that affect a proper objective assessment of the motor capacity.

Bibliography

[1] P. Abbeel, Varun Ganapathi and A. Ng. “Learning vehicular dynamics, with application to modeling helicopters”. In: *NIPS*. 2005 (cit. on pp. 88, 91, 102).

[2] Hyo-Sung Ahn, Yang Quan Chen and Kevin L. Moore. “Iterative Learning Control: Brief Survey and Categorization”. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37 (2007), pp. 1099–1121 (cit. on p. 59).

[3] euRobotics Aisbl. *Strategic Research Agenda for Robotics in Europe.* 2013 (cit. on p. 22).

[4] Atif Alamri, Heung-Nam Kim and Abdulmotaleb El Saddik. “A decision model of stroke patient rehabilitation with augmented reality-based games”. In: *2010 International Conference on Autonomous and Intelligent Systems, AIS 2010* (2010), pp. 1–6 (cit. on p. 60).

[5] Thiago Alves and Rogério Sales Gonçalves. “Predictive equation for a circular trajectory period in a cable-driven robot for rehabilitation”. In: *Journal of the Brazilian Society of Mechanical Sciences and Engineering* 42 (2020), pp. 1–11 (cit. on pp. 28, 30, 31, 62–66, 68, 145).

[6] Anna Aminov, Jeffrey M. Rogers, Sandy Middleton, Karen Caeyenberghs and Peter H. Wilson. “What do randomized controlled trials say about virtual rehabilitation in stroke? A systematic literature review and meta-analysis of upper-limb and cognitive outcomes”. In: *Journal of NeuroEngineering and Rehabilitation* 15 (2018) (cit. on p. 26).

[7] Christopher G. Atkeson. “Using locally weighted regression for robot learning”. In: *Proceedings. 1991 IEEE International Conference on Robotics and Automation* (1991), 958–963 vol.2 (cit. on p. 58).

[8] Giuseppe Averta, Cosimo Della Santina, Gaetano Valenza, Antonio Bicchi and Matteo Bianchi. “Exploiting upper-limb functional principal components for human-like motion generation of anthropomorphic robots”. In: *Journal of NeuroEngineering and Rehabilitation* 17 (2020) (cit. on p. 53).

[9] Uba Backonja, Nai-Ching Chi, Yong Kyung Choi et al. “Visualization approaches to support healthy aging: A systematic review”. In: *Journal of innovation in health informatics* 23 (2016), pp. 860 –860 (cit. on p. 142).

[10] Ryan R. Bailey, Joseph W. Klaesner and Catherine E. Lang. “Quantifying Real-World Upper-Limb Activity in Nondisabled Adults and Adults With Chronic Stroke”. In: *Neurorehabilitation and Neural Repair* 29 (2015), pp. 969 –978 (cit. on pp. 144, 145).

[11] Sivakumar Balasubramanian, Roberto Colombo, Irma Sterpi, Vittorio Sanguineti and Etienne Burdet. “Robotic Assessment of Upper Limb Motor Function After Stroke”. In: *American Journal of Physical Medicine & Rehabilitation* 91 (2012), S255–S269 (cit. on pp. 49, 52, 55).

[12] Francesco Bartolucci, Alessio Farcomeni and Fulvia Pennoni. “Latent Markov Models for Longitudinal Data”. In: *METRON* (2012) (cit. on p. 93).

[13] Angelo Basteris, Sharon M. Nijenhuis, Arno H. A. Stienen et al. “Training modalities in robot-mediated upper limb rehabilitation in stroke: a framework for classification based on a systematic review”. In: *Journal of NeuroEngineering and Rehabilitation* 11 (2013), pp. 111 –111 (cit. on pp. 13, 16, 23, 26, 27, 47).

[14] Kilian Baur, Nina Rohrbach, Joachim Hermsdörfer, Robert Riener and Verena Klamroth-Marganska. “The “Beam-Me-In Strategy” - remote haptic therapist-patient interaction with two exoskeletons for stroke therapy”. In: *Journal of NeuroEngineering and Rehabilitation* 16 (2019) (cit. on pp. 13, 21, 31).

[15] Arturo Bertomeu-Motos, Andrea Blanco, Francisco J. Badesa et al. “Human arm joints reconstruction algorithm in rehabilitation therapies assisted by end-effector robotic devices”. In: *Journal of NeuroEngineering and Rehabilitation* 15 (2018) (cit. on p. 55).

[16] Emilia Biffi, Cristina Maghini, Beatrice Cairo et al. “Movement Velocity and Fluidity Improve after Armeo Spring Rehabilitation in Children Affected by Acquired and Congenital Brain Diseases: An Observational Study”. In: *BioMed Research International* 2018 (2018) (cit. on pp. 11, 34, 39, 43, 146).

[17] Matteo Bigoni, Silvia Baudo, Veronica Cimolin et al. “Does kinematics add meaningful information to clinical assessment in post-stroke upper limb rehabilitation? A case report”. In: *Journal of Physical Therapy Science* 28 (2016), pp. 2408 –2413 (cit. on p. 25).

[18] Amy A. Blank, James A. French, Ali Utku Pehlivan and Marcia Kilchenman O’Malley. “Current Trends in Robot-Assisted Upper-Limb Stroke Rehabilitation: Promoting Patient Engagement in Therapy”. In: *Current Physical Medicine and Rehabilitation Reports* 2 (2014), pp. 184–195 (cit. on pp. 18, 23, 48, 59).

[19] Nunzio Alberto Borghese, Manuel Pezzera, Renato Mainetti et al. “A cloud-based platform for effective supervision of autonomous home rehabilitation through exergames”. In: *2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH)* (2018), pp. 1–6 (cit. on pp. 60, 61).

[20] Peter Brezany, Markéta Janatová, Olga Stepánková, Miroslav Uller and Marek Lenart. “Towards precision brain disorder rehabilitation”. In: *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (2018), pp. 0228–0233 (cit. on p. 60).

[21] Nabila Brihmat, Isabelle Loubinoux, Evelyne Castel-Lacanal, Philippe Marque and David Gasq. “Kinematic parameters obtained with the ArmeoSpring for upper-limb assessment after stroke: a reliability and learning effect study for guiding parameter use”. In: *Journal of NeuroEngineering and Rehabilitation* 17 (2020) (cit. on p. 131).

[22] Aamani Budhota, Karen S. G. Chua, Asif Hussain et al. “Robotic Assisted Upper Limb Training Post Stroke: A Randomized Control Trial Using Combinatory Approach Toward Reducing Workforce Demands”. In: *Frontiers in Neurology* 12 (2021) (cit. on pp. 24, 61).

[23] Marianna Capecci, Maria Gabriella Ceravolo, Francesco Ferracuti et al. “A Hidden Semi-Markov Model based approach for rehabilitation exercise assessment”. In: *Journal of biomedical informatics* 78 (2018), pp. 1–11 (cit. on p. 61).

[24] Janet H. Carr and Roberta B. Shepherd. *Stroke rehabilitation : guidelines for exercise and training to optimize motor skill*. 2003 (cit. on p. 22).

[25] Maura Casadio and Vittorio Sanguineti. “Learning, Retention, and Slacking: A Model of the Dynamics of Recovery in Robot Therapy”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 20 (2012), pp. 286–296 (cit. on pp. 18, 52, 56, 62–67).

[26] *CASP checklists - CASP - critical appraisal skills programme*. 2021 (cit. on pp. 123, 223).

[27] Andrew Centen, Catherine R. Lowrey, Stephen H. Scott, Ting Ting Yeh and George Mochizuki. “KAPS (kinematic assessment of passive stretch): a tool to assess elbow flexor and extensor spasticity after stroke using a robotic exoskeleton”. In: *Journal of NeuroEngineering and Rehabilitation* 14 (2017) (cit. on p. 39).

[28] Alix Chadwell, Laurence P. J. Kenney, Malcolm H. Granat et al. “Visualisation of upper limb activity using spirals: A new approach to the assessment of daily prosthesis usage”. In: *Prosthetics and Orthotics International* 42 (2018), pp. 37 –44 (cit. on pp. 142, 143).

[29] Gromit Yeuk-Yin Chan, Luis Gustavo Nonato, Alice Chu et al. “Motion Browser: Visualizing and Understanding Complex Upper Limb Movement Under Obstetrical Brachial Plexus Injuries”. In: *IEEE Transactions on Visualization and Computer Graphics* 26 (2020), pp. 981–990 (cit. on p. 142).

[30] Jyh-Jong Chang, Wen lin Tung, Wen-Lan Wu, Mao-Hsiung Huang and F. C. Su. “Effects of robot-aided bilateral force-induced isokinetic arm training combined with conventional rehabilitation on arm motor function in patients with chronic stroke.” In: *Archives of physical medicine and rehabilitation* 88 10 (2007), pp. 1332–8 (cit. on pp. 15, 17, 22, 23, 38, 40–42).

[31] Samprit Chatterjee and Jeffrey S. Simonoff. *Handbook of Regression Analysis*. 2012 (cit. on p. 83).

[32] Yevgen Chebotar, Karol Hausman, Marvin Zhang et al. “Combining Model-Based and Model-Free Updates for Trajectory-Centric Reinforcement Learning”. In: *ICML*. 2017 (cit. on p. 59).

[33] Radhika Chemuturi, Farshid Amirabdollahian and Kerstin Dautenhahn. “Adaptive training algorithm for robot-assisted upper-arm rehabilitation, applicable to individualised and therapeutic human-robot interaction”. In: *Journal of NeuroEngineering and Rehabilitation* 10 (2012), pp. 102 –102 (cit. on pp. 38, 60).

[34] Wenbin Chen, Caihua Xiong, Xiaolin Huang, Ronglei Sun and Youlun Xiong. “Manipulator Workspace Boundary Extraction and Its Application in Workspace Analysis of the Human’s Upper Extremity”. In: *Advanced Robotics* 23 (2009), pp. 1393 –1410 (cit. on pp. 55, 93).

[35] Yu Chen, Yunan Chen, Kai Zheng et al. “A qualitative study on user acceptance of a home-based stroke telerehabilitation system”. In: *Topics in Stroke Rehabilitation* 27 (2019), pp. 81 –92 (cit. on p. 144).

[36] Dr. Ton J. Cleophas and Prof. Aeilko H. Zwinderman. “Modern Bayesian Statistics in Clinical Research”. In: *Springer International Publishing*. 2018 (cit. on p. 131).

[37] Adam Coates, P. Abbeel and A. Ng. “Learning for control from multiple demonstrations”. In: *ICML ’08*. 2008 (cit. on pp. 88, 89, 93, 96).

[38] Roberto Colombo, Fabrizio Pisano, Carmen Delconte et al. “Comparison of exercise training effect with different robotic devices for upper limb rehabilitation: a retrospective study.” In: *European journal of physical and rehabilitation medicine* 53 2 (2017), pp. 240–248 (cit. on pp. 43, 47).

[39] Roberto Colombo, Fabrizio Pisano, Silvestro Micera et al. “Assessing Mechanisms of Recovery During Robot-Aided Neurorehabilitation of the Upper Limb”. In: *Neurorehabilitation and Neural Repair* 22 (2008), pp. 50 –63 (cit. on pp. 17–19, 26, 28, 46, 55).

[40] Roberto Colombo, Irma Sterpi, Alessandra Mazzone, Carmen Delconte and Fabrizio Pisano. “Improving proprioceptive deficits after stroke through robot-assisted training of the upper limb: a pilot case report study”. In: *Neurocase* 22 (2016), pp. 191 –200 (cit. on pp. 14, 17, 26).

[41] Roberto Colombo, Irma Sterpi, Alessandra Mazzone, Carmen Delconte and Fabrizio Pisano. “Robot-aided neurorehabilitation in sub-acute and chronic stroke: does spontaneous recovery have a limited impact on outcome?” In: *NeuroRehabilitation* 33 4 (2013), pp. 621–9 (cit. on p. 49).

[42] Roberto Colombo, Irma Sterpi, Alessandra Mazzone, Carmen Delconte and Fabrizio Pisano. “Taking a Lesson From Patients’ Recovery Strategies to Optimize Training During Robot-Aided Rehabilitation”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 20 (2012), pp. 276–285 (cit. on pp. 20, 26, 60, 62–67).

[43] Sara Contu, Asif Hussain, Simone Kager et al. “Proprioceptive assessment in clinical settings: Evaluation of joint position sense in upper limb post-stroke using a robotic manipulator”. In: *PLoS ONE* 12 (2017) (cit. on p. 14).

[44] Camilo Cortés, Aitor Ardanza, Francisco Molina-Rueda et al. “Upper Limb Posture Estimation in Robotic and Virtual Reality-Based Rehabilitation”. In: *BioMed Research International* 2014 (2014) (cit. on p. 47).

[45] Michelle A. Cottrell, Olivia A Galea, Shaun P O’Leary, Annie J. Hill and Trevor Glen Russell. “Real-time telerehabilitation for the treatment of musculoskeletal conditions is effective and comparable to standard practice: a systematic review and meta-analysis”. In: *Clinical Rehabilitation* 31 (2017), pp. 625 –638 (cit. on p. 143).

[46] Vincent Crocher, Anis Sahbani, J. Robertson, Agnès Roby-Brami and Guillaume Morel. “Constraining Upper Limb Synergies of Hemiparetic Patients Using a Robotic Exoskeleton in the Perspective of Neuro-Rehabilitation”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 20 (2012), pp. 247–257 (cit. on p. 53).

[47] Peter R Culmer, Andrew E. Jackson, S. G. Makower et al. “A novel robotic system for quantifying arm kinematics and kinetics: Description and evaluation in therapist-assisted passive arm movements post-stroke”. In: *Journal of Neuroscience Methods* 197 (2011), pp. 259–269 (cit. on p. 29).

[48] Ivana Cusmano, Irma Sterpi, Alessandra Mazzone et al. “Evaluation of upper limb sense of position in healthy individuals and patients after stroke.” In: *Journal of healthcare engineering* 5 2 (2014), pp. 145–62 (cit. on pp. 14, 39, 55, 145).

[49] Zachary C Danziger, Alon Fishbach and Ferdinando A. Mussa-Ivaldi. “Learning Algorithms for Human-Machine Interfaces”. In: *IEEE Transactions on Biomedical Engineering* 56 (2009), pp. 1502–1511 (cit. on p. 61).

[50] Alan Davies, Marisa Cunha, Kamilla Kopec-Harding et al. “Biomarker data visualisation for decision making in clinical trials”. In: *International journal of medical informatics* 132 (2019), p. 104008 (cit. on p. 163).

[51] Seyed Ali Mohamad Dehghan, Hamid Reza Koofigar and Mohsen Ekramian. “An adaptive arm’s mechanical impedance estimator for rehabilitation robots without force and acceleration sensors”. In: *International Journal of Systems Science* 49 (2018), pp. 2784 –2796 (cit. on pp. 26, 28, 31).

[52] Arthur P Dempster, Nan M Laird and Donald B Rubin. “Maximum likelihood from incomplete data via the EM algorithm”. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 39.1 (1977), pp. 1–22 (cit. on p. 227).

[53] Tyler Desplenter, Shrikant Chinchalkar and Ana Luisa Trejos. “Enhancing the Therapist-Device Relationship: Software Requirements for Digital Collection and Analysis of Patient Data”. In: *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)* (2019), pp. 1091–1096 (cit. on p. 142).

[54] Jan Kleine Deters. *Therapeutic exercise assessment automation, a hidden Markov model approach.* 2018 (cit. on p. 61).

[55] Pamela W Duncan, H. S. Jorgensen and Derick T. Wade. “Outcome measures in acute stroke trials: a systematic review and some recommendations to improve practice.” In: *Stroke* 31 6 (2000), pp. 1429–38 (cit. on p. 51).

[56] Christophe Duret, Ophélie Courtial, Anne-Gaëlle Grosmaire and Emilie Hutin. “Use of a Robotic Device for the Rehabilitation of Severe Upper Limb Paresis in Subacute Stroke: Exploration of Patient-Robot Interactions and the Motor Recovery Process”. In: *BioMed Research International* 2015 (2015) (cit. on pp. 17, 24, 42, 48).

[57] Michael D. Ellis, Yiyun Lan, Jun Yao and Julius P. A. Dewald. “Robotic quantification of upper extremity loss of independent joint control or flexion synergy in individuals with hemiparetic stroke: a review of paradigms addressing the effects of shoulder abduction loading”. In: *Journal of NeuroEngineering and Rehabilitation* 13 (2016) (cit. on pp. 13, 15, 20, 40).

[58] Marco Ewerton, David Rother, Jakob Weimar et al. "Assisting Movement Training and Execution With Visual and Haptic Feedback". In: *Frontiers in Neurorobotics* 12 (2018) (cit. on pp. 57, 62–67, 144, 146).

[59] Kobeissy Fh. "Models of Posttraumatic Brain Injury Neurorehabilitation Brain Neurotrauma: Molecular, Neuropsychological, and Rehabilitation Aspects". In: *Frontiers in Neuroengineering* (2015) (cit. on pp. 16, 23, 24, 33).

[60] Edelle Edee Field-Fote. "Mediators and Moderators, Confounders and Covariates: Exploring the Variables That Illuminate or Obscure the "Active Ingredients" in Neurorehabilitation." In: *Journal of Neurologic Physical Therapy* (2019) (cit. on p. 51).

[61] Emile Fiesler and Russell Beale. *Handbook of Neural Computation*. 1996 (cit. on p. 97).

[62] Tamar Flash and Neville Hogan. "The coordination of arm movements: an experimentally confirmed mathematical model". In: *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 1985 (cit. on p. 46).

[63] Jason Fong, Renz Ocampo, Douglas P. Gross and Mahdi Tavakoli. "Intelligent Robotics Incorporating Machine Learning Algorithms for Improving Functional Capacity Evaluation and Occupational Rehabilitation". In: *Journal of Occupational Rehabilitation* (2020), pp. 1–9 (cit. on pp. 12, 19, 25, 28, 29, 33, 35, 36).

[64] Catherine Ford, Donna Malley, Andrew Bateman et al. "Selection and visualisation of outcome measures for complex post-acute acquired brain injury rehabilitation interventions." In: *NeuroRehabilitation* 39 1 (2016), pp. 65–79 (cit. on p. 142).

[65] Axel R. Fugl-Meyer, L Jääskö, I. A. Leyman, Soren Olsson and S Steglind. "The post-stroke hemiplegic patient. a method for evaluation of physical performance." In: *Scandinavian journal of rehabilitation medicine* 7 1 (1975), pp. 13–31 (cit. on pp. 122, 123).

[66] Zoltan Geler, Vladimir Kurbalija, Mirjana Ivanović, Milovs Radovanović and Weihui Dai. "Dynamic Time Warping: Itakura vs Sakoe-Chiba". In: *2019 IEEE International Symposium on INnovations in Intelligent SysTems and Applications (INISTA)* (2019), pp. 1–6 (cit. on p. 151).

[67] Maxime Gilliaux, Thierry Lejeune, Christine Detrembleur et al. "Using the robotic device REAplan as a valid, reliable, and sensitive tool to quantify upper limb impairments in stroke patients." In: *Journal of rehabilitation medicine* 46 2 (2014), pp. 117–25 (cit. on pp. 20, 21, 42).

[68] R. A. R. C. Gopura, D. S. V. Bandara, Kazuo Kiguchi and George K. I. Mann. "Developments in hardware systems of active upper-limb exoskeleton robots: A review". In: *Robotics Auton. Syst.* 75 (2016), pp. 203–220 (cit. on p. 47).

[69] P. J. Green, Allan H. Seheult and Bernard W. Silverman. "Density Estimation for Statistics and Data Analysis." In: *Applied statistics* 37 (1988), p. 120 (cit. on p. 151).

[70] Marco Guidali, Alexander Duschau-Wicke, Simon Broggi et al. "A robotic system to train activities of daily living in a virtual environment". In: *Medical & Biological Engineering & Computing* 49 (2011), pp. 1213–1223 (cit. on p. 30).

[71] Marco Guidali, Urs Keller, Verena Klamroth-Marganska, Tobias Nef and Robert Riener. "Estimating the patient's contribution during robot-assisted therapy." In: *Journal of rehabilitation research and development* 50 3 (2013), pp. 379–94 (cit. on p. 48).

[72] Susan R. Harris and Carolee J. Winstein. "The Past, Present, and Future of Neurorehabilitation: From NUSTEP Through IV STEP and Beyond." In: *Journal of Neurologic Physical Therapy* (2017) (cit. on p. 51).

[73] Fatmah Hasani, Joy Macdermid, Ada Tang et al. "Knowledge, Attitude and Implementation of Evidence-Based Practice among Physiotherapists Working in the Kingdom of Saudi Arabia: A Cross-Sectional Survey". In: *Healthcare* 8 (2020) (cit. on p. 51).

[74] Vincent T. van Hees, Lukas Gorzelniak, Emmanuel Carlos Dean León et al. "Separating Movement and Gravity Components in an Acceleration Signal and Implications for the Assessment of Human Daily Physical Activity". In: *PLoS ONE* 8 (2013) (cit. on p. 45).

[75] Vincent Hernandez, Dana Kulić and Gentiane Venture. "Adversarial autoencoder for visualization and classification of human activity: Application to a low-cost commercial force plate." In: *Journal of biomechanics* (2020), p. 109684 (cit. on p. 144).

[76] Renée J. F. Heuvel, Monique A. S. Lexis and Luc P. de Witte. "ZORA Robot Based Interventions to Achieve Therapeutic and Educational Goals in Children with Severe Physical Disabilities". In: *International Journal of Social Robotics* 12 (2020), pp. 493–504 (cit. on p. 31).

[77] Neville Hogan and Dagmar Sternad. "Dynamic primitives of motor behavior". In: *Biological Cybernetics* 106 (2012), pp. 727–739 (cit. on p. 46).

[78] Maurits H. J. Hoonhorst, Rinske Nijland, Jan S.P. van den Berg et al. "How Do Fugl-Meyer Arm Motor Scores Relate to Dexterity According to the Action Research Arm Test at 6 Months Poststroke?" In: *Archives of physical medicine and rehabilitation* 96 10 (2015), pp. 1845–9 (cit. on p. 50).

[79] Bastian Ilso Hougaard and Hendrik Knoche. "Telling the Story Right: How Therapists Aid Stroke Patients Interpret Personal Visualized Game Performance Data". In: *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare* (2019) (cit. on p. 144).

[80] Yong Hu, Jnf Mak and K. D. K. Luk. "Application of Surface EMG Topography in Low Back Pain Rehabilitation Assessment". In: *2007 3rd International IEEE/EMBS Conference on Neural Engineering* (2007), pp. 557–560 (cit. on p. 144).

[81] Vincent S. Huang and John W. Krakauer. "Robotic neurorehabilitation: a computational motor learning perspective". In: *Journal of NeuroEngineering and Rehabilitation* 6 (2009), pp. 5 –5 (cit. on pp. 18, 19, 26, 27, 34, 35, 46, 48, 130).

[82] Luigi Iuppariello, Paolo Bifulco, Maria Romano, Mario Cesarelli and Gianni D'Addio. "Submovements composition and quality assessment of reaching movements in subjects with Parkinson's Disease". In: *2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings* (2015), pp. 329–334 (cit. on p. 46).

[83] Ashesh Jain, Brian Wojcik, Thorsten Joachims and Ashutosh Saxena. “Learning Trajectory Preferences for Manipulators via Iterative Improvement”. In: *NIPS*. 2013 (cit. on p. 59).

[84] Michelle Jillian Johnson, Xin Feng, L.M. Johnson and Jack M. Winters. “Potential of a suite of robot/computer-assisted motivating systems for personalized, home-based, stroke rehabilitation”. In: *Journal of NeuroEngineering and Rehabilitation* 4 (2006), pp. 6 –6 (cit. on pp. 15, 20, 21, 47, 48, 50, 59, 146).

[85] Fabian Just, Özhan Özen, Stefano Tortora et al. “Human arm weight compensation in rehabilitation robotics: efficacy of three distinct methods”. In: *Journal of NeuroEngineering and Rehabilitation* 17 (2020) (cit. on pp. 26, 28, 45).

[86] Patricia Kan, Rajibul Huq, Jesse Hoey, Robby Goetschalckx and Alex Mihailidis. “The development of an adaptive upper-limb stroke rehabilitation robotic system”. In: *Journal of NeuroEngineering and Rehabilitation* 8 (2010), pp. 33 –33 (cit. on p. 60).

[87] Christoph M. Kanzler, Mike D. Rinderknecht, Anne Schwarz et al. “A data-driven framework for selecting and validating digital health metrics: use-case in neurological sensorimotor impairments.” In: *NPJ digital medicine* 3 1 (2020), p. 80 (cit. on p. 37).

[88] Nam Wook Kim, Eston Schweickart, Zhicheng Liu et al. “Data-Driven Guides: Supporting Expressive Design for Information Graphics”. In: *IEEE Transactions on Visualization and Computer Graphics* 23 (2017), pp. 491–500 (cit. on p. 143).

[89] Tomoko Kitago, Jeff Goldsmith, Michelle D. Harran et al. “Robotic therapy for chronic stroke: general recovery of impairment or improved task-specific skill?” In: *Journal of neurophysiology* 114 3 (2015), pp. 1885–94 (cit. on pp. 15, 35, 53, 62–68).

[90] Bart Klaassen, Bert-Jan F. van Beijnum, Jeremia Philipp Oskar Held et al. “Usability Evaluations of a Wearable Inertial Sensing System and Quality of Movement Metrics for Stroke Survivors by Care Professionals”. In: *Frontiers in Bioengineering and Biotechnology* 5 (2017) (cit. on pp. 48, 161).

[91] Nives Klopcar and Jadran Lenarcic. “Kinematic Model for Determination of Human Arm Reachable Workspace”. In: *Meccanica* 40 (2005), pp. 203–219 (cit. on p. 144).

[92] Hermano Igo Krebs, Michael Krams, Dimitris K. Agrafiotis et al. “Robotic Measurement of Arm Movements After Stroke Establishes Biomarkers of Motor Recovery”. In: *Stroke* 45 (2014), pp. 200–204 (cit. on pp. 56, 62–68).

[93] Gert Kwakkel, Boudewijn J Kollen and Hermano Igo Krebs. “Effects of Robot-Assisted Therapy on Upper Limb Recovery After Stroke: A Systematic Review”. In: *Neurorehabilitation and Neural Repair* 22 (2008), pp. 111 –121 (cit. on pp. 26, 58).

[94] Clemente Lauretti, Francesca Cordella, Eugenio Guglielmelli and Loredana Zollo. “Learning by Demonstration for Planning Activities of Daily Living in Rehabilitation and Assistive Robotics”. In: *IEEE Robotics and Automation Letters* 2 (2017), pp. 1375–1382 (cit. on pp. 57, 63, 65–68).

[95] Thai Le, Nai-Ching Chi, Shomir Chaudhuri, Hilaire J. Thompson and George Demiris. “Understanding Older Adult Use of Data Visualizations as a Resource for Maintaining Health and Wellness”. In: *Journal of Applied Gerontology* 37 (2018), pp. 922 –939 (cit. on p. 143).

[96] Alberto Leardini, Claudio Belvedere, Fabrizio Nardini et al. “Kinematic models of lower limb joints for musculo-skeletal modelling and optimization in gait analysis.” In: *Journal of biomechanics* 62 (2017), pp. 77–86 (cit. on p. 58).

[97] Shijuan Li, S. Kay and Stuart Porter. “A 3D assessment and feedback tool for Ankylosing Spondylitis from the perspective of healthcare professionals”. In: *Informatics for Health and Social Care* 42 (2017), pp. 274 –289 (cit. on p. 141).

[98] I-Hsien Lin, Han-Ting Tsai, Chien-Yung Wang et al. “Effectiveness and Superiority of Rehabilitative Treatments in Enhancing Motor Recovery Within 6 Months Poststroke: A Systemic Review.” In: *Archives of physical medicine and rehabilitation* 100 2 (2019), pp. 366–378 (cit. on p. 51).

[99] Ho Shing Lo and Sheng Quan Xie. “Exoskeleton robots for upper-limb rehabilitation: state of the art and future prospects.” In: *Medical engineering & physics* 34 3 (2012), pp. 261–8 (cit. on pp. 28, 47).

[100] Maria Longhi, Andrea Merlo, Paolo Prati, Meris Giacobbi and Davide Mazzoli. “Instrumental indices for upper limb function assessment in stroke patients: a validation study”. In: *Journal of NeuroEngineering and Rehabilitation* 13 (2016) (cit. on pp. 34, 35, 40–42, 49, 146).

[101] Maichou Lor, Theresa A. Koleck and Suzanne Bakken. “Information visualizations of symptom information for patients and providers: a systematic review”. In: *Journal of the American Medical Informatics Association* 26 (2019), pp. 162–171 (cit. on p. 142).

[102] Rui C. V. Loureiro, William S. Harwin, Kiyoshi Nagai and Michelle Jillian Johnson. “Advances in upper limb stroke rehabilitation: a technology push”. In: *Medical & Biological Engineering & Computing* 49 (2011), pp. 1103–1118 (cit. on pp. 25, 48–50).

[103] Catherine R. Lowrey, Teige C Bourke, Stephen D. Bagg, Sean P. Dukelow and Stephen H. Scott. “A postural unloading task to assess fast corrective responses in the upper limb following stroke”. In: *Journal of NeuroEngineering and Rehabilitation* 16 (2019) (cit. on pp. 19, 35–38, 51).

[104] Alastair S. Macdonald, David Loudon and Philip John Rowe. “Visualisation of biomechanical data to assist therapeutic rehabilitation.” In: *Gerontechnology : international journal on the fundamental aspects of technology to serve the ageing society* 9 2 (2010), pp. 98–99 (cit. on p. 144).

[105] Pawel Maciejasz, Jörg Eschweiler, Kurt Gerlach-Hahn, Arne Jansen-Troy and Steffen Leonhardt. “A survey on robotic devices for upper limb rehabilitation”. In: *Journal of NeuroEngineering and Rehabilitation* 11 (2012), pp. 3 –3 (cit. on pp. 2, 11, 24, 26, 27).

[106] Sulaiman bin Mazlan, Hisyam Abdul Rahman, Yeong Che Fai, Babul Salam Ksm Kader Ibrahim and M. Saiful Huq. “Kinematic variables for upper limb rehabilitation robot and correlations with clinical scales: A review”. In: *Bulletin of Electrical Engineering and Informatics* 9 (2020), pp. 75–82 (cit. on p. 37).

[107] Stefano Mazzoleni, Patrizio Sale, Micol Tiboni et al. “Upper Limb Robot-Assisted Therapy in Chronic and Subacute Stroke Patients: A Kinematic Analysis”. In: *American Journal of Physical Medicine & Rehabilitation* 92 (2013), e26–e37 (cit. on p. 42).

[108] Alison McKenzie, Lucy Dodakian, Jill See et al. “Validity of Robot-Based Assessments of Upper Extremity Function.” In: *Archives of physical medicine and rehabilitation* 98 10 (2017), 1969–1976.e2 (cit. on pp. 24, 39, 44).

[109] Jan Mehrholz, Alex Pollock, Marcus Pohl, Joachim Kugler and Bernhard Elsner. “Systematic review with network meta-analysis of randomized controlled trials of robotic-assisted arm training for improving activities of daily living and upper limb function after stroke”. In: *Journal of NeuroEngineering and Rehabilitation* 17 (2020) (cit. on pp. 24, 49).

[110] Rafael Mendoza-Crespo, Diego Torricelli, Joel C. Huegel et al. “An Adaptable Human-Like Gait Pattern Generator Derived From a Lower Limb Exoskeleton”. In: *Frontiers in Robotics and AI* 6 (2019) (cit. on p. 58).

[111] Fokke B van Meulen, Bart Klaassen, Jeremia Philipp Oskar Held et al. “Objective Evaluation of the Quality of Movement in Daily Life after Stroke”. In: *Frontiers in Bioengineering and Biotechnology* 3 (2016) (cit. on p. 142).

[112] Silvestro Micera, Jacopo Carpaneto, Federico Posteraro et al. “Characterization of upper arm synergies during reaching tasks in able-bodied and hemiparetic subjects.” In: *Clinical biomechanics* 20 9 (2005), pp. 939–46 (cit. on pp. 19, 54).

[113] Yasser El Miedany, Maha El Gaafary and Deborah Palmer. “Assessment of the utility of visual feedback in the treatment of early rheumatoid arthritis patients: a pilot study”. In: *Rheumatology International* 32 (2011), pp. 3061–3068 (cit. on p. 142).

[114] Matjavz Mihelj. “Human arm kinematics for robot based rehabilitation”. In: *Robotica* 24 (2005), pp. 377 –383 (cit. on pp. 30, 55).

[115] Jennifer L. Moore, Kirsten Potter, Kathleen Blankshain et al. “A Core Set of Outcome Measures for Adults With Neurologic Conditions Undergoing Rehabilitation”. In: *Journal of Neurologic Physical Therapy* 42 (2018), pp. 174 –220 (cit. on pp. 20, 21, 32–34, 36).

[116] Margit Alt Murphy, Carin Willén and Katharina Stibrant Sunnerhagen. “Responsiveness of Upper Extremity Kinematic Measures and Clinical Improvement During the First Three Months After Stroke”. In: *Neurorehabilitation and Neural Repair* 27 (2013), pp. 844 –853 (cit. on p. 49).

[117] Seyedeh Somayeh Naghibi, Ali Fallah, Ali Maleki and Farnaz Ghassemi. “Elbow angle generation during activities of daily living using a submovement prediction model”. In: *Biological Cybernetics* 114 (2020), pp. 389–402 (cit. on pp. 46, 57, 62–67).

[118] Tobias Nef, Gabriela Quinter, Roland Müller and Robert Riener. “Effects of Arm Training with the Robotic Device ARMin I in Chronic Stroke: Three Single Cases”. In: *Neurodegenerative Diseases* 6 (2009), pp. 240 –251 (cit. on p. 30).

[119] Karl M. Newell. “What are Fundamental Motor Skills and What is Fundamental About Them?” In: *Journal of Motor Learning and Development* 8 (2020), pp. 280–314 (cit. on pp. 17, 18, 27, 36, 46, 47).

[120] Nurdiana Nordin, Sheng Quan Xie and Burkhard Claus Wünsche. “Assessment of movement quality in robot- assisted upper limb rehabilitation after stroke: a review”. In: *Journal of NeuroEngineering and Rehabilitation* 11 (2014) (cit. on pp. 13, 15, 18, 36–38, 40, 45–49, 160).

[121] Takayuki Osa, Joni Pajarinen, Gerhard Neumann et al. “An Algorithmic Perspective on Imitation Learning”. In: *Found. Trends Robotics* 7 (2018), pp. 1–179 (cit. on p. 87).

[122] Eri Otaka, Yohei Otaka, Shoko Kasuga et al. “Clinical usefulness and validity of robotic measures of reaching movement in hemiparetic stroke patients”. In: *Journal of NeuroEngineering and Rehabilitation* 12 (2015) (cit. on pp. 14, 34, 36, 38, 40, 42).

[123] Miguel A Padilla-Castañeda, Edoardo Sotgiu, Michele Barsotti et al. “An Orthopaedic Robotic-Assisted Rehabilitation Method of the Forearm in Virtual Reality Physiotherapy”. In: *Journal of Healthcare Engineering* 2018 (2018) (cit. on pp. 3, 12, 26, 60).

[124] Lizheng Pan, Aiguo Song, Suolin Duan and Baoguo Xu. “Robot-assisted humanized passive rehabilitation training based on online assessment and regulation.” In: *Biomedical materials and engineering* 26 Suppl 1 (2015), S655–64 (cit. on pp. 2, 51).

[125] Alessandro Panarese, Roberto Colombo, Irma Sterpi, Fabrizio Pisano and Silvestro Micera. “Tracking Motor Improvement at the Subtask Level During Robot-Aided Neurorehabilitation of Stroke Patients”. In: *Neurorehabilitation and Neural Repair* 26 (2012), pp. 822 –833 (cit. on pp. 60, 62–67, 144).

[126] Eugenia Papaleo, Loredana Zollo, Nicolás García Aracil et al. “Upper-limb kinematic reconstruction during stroke robot-aided therapy”. In: *Medical & Biological Engineering & Computing* 53 (2015), pp. 815–828 (cit. on pp. 2, 52, 57).

[127] Alessandro Peretti, Francesco Amenta, Seyed Khosrow Tayebati, Giulio Nittari and Syed Sarosh Mahdi. “Telerehabilitation: Review of the State-of-the-Art and Areas of Application”. In: *JMIR Rehabilitation and Assistive Technologies* 4 (2017) (cit. on p. 143).

[128] Elisabetta Peri, Emilia Biffi, Cristina Maghini et al. “Quantitative Evaluation of Performance during Robot-assisted Treatment.” In: *Methods of information in medicine* 55 1 (2016), pp. 84–8 (cit. on pp. 34, 36, 38, 50, 60).

[129] Michele Pirovano, Renato Mainetti, Gabriel Baud-Bovy, Pier Luca Lanzi and Nunzio Alberto Borghese. “Self-adaptive games for rehabilitation at home”. In: *2012 IEEE Conference on Computational Intelligence and Games (CIG)* (2012), pp. 179–186 (cit. on p. 61).

[130] Bernd Ploderer, Justin Fong, Marlena Klaic et al. “How Therapists Use Visualizations of Upper Limb Movement Information From Stroke Patients: A Qualitative Study With Simulated Information”. In: *JMIR Rehabilitation and Assistive Technologies* 3 (2016) (cit. on p. 142).

[131] Loreen Pogrzeba, Thomas Neumann, Markus Wacker and Bernhard Jung. “Analysis and Quantification of Repetitive Motion in Long-Term Rehabilitation”. In: *IEEE Journal of Biomedical and Health Informatics* 23 (2019), pp. 1075–1085 (cit. on p. 142).

[132] Patrizia Poli, Giovanni Morone, Giulio Rosati and Stefano Masiero. “Robotic Technologies and Rehabilitation: New Tools for Stroke Patients’ Therapy”. In: *BioMed Research International* 2013 (2013) (cit. on pp. 2, 16, 17, 21, 24, 27–29, 47).

[133] Federico Posteraro, Simona Crea, Stefano Mazzoleni et al. “Technologically-advanced assessment of upper-limb spasticity: a pilot study.” In: *European journal of physical and rehabilitation medicine* 54 4 (2018), pp. 536–544 (cit. on pp. 14, 24, 39).

[134] Federico Posteraro, Stefano Mazzoleni, Sara Aliboni et al. “Upper limb spasticity reduction following active training: a robot-mediated study in patients with chronic hemiparesis.” In: *Journal of rehabilitation medicine* 42 3 (2010), pp. 279–81 (cit. on pp. 26, 48).

[135] M. J. D. Powell. “An efficient method for finding the minimum of a function of several variables without calculating derivatives”. In: *Comput. J.* 7 (1964), pp. 155–162 (cit. on p. 92).

[136] Gerdienke B. Prange, Michiel J. A. Jannink, Arno H. A. Stienen et al. “Influence of Gravity Compensation on Muscle Activation Patterns During Different Temporal Phases of Arm Movements of Stroke Patients”. In: *Neurorehabilitation and Neural Repair* 23 (2009), pp. 478 –485 (cit. on p. 45).

[137] Robinson Ramírez-Vélez, M. Caridad Bagur-Calafat, Jorge Enrique Correa-Bautista and Montserrat Girabent-Farrés. “Barriers against incorporating evidence-based practice in physical therapy in Colombia: current state and factors associated”. In: *BMC Medical Education* 15 (2015) (cit. on p. 51).

[138] Jeffrey A. Reinbolt, Jaco F. Schutte, Benjamin J. Fregly et al. “Determination of patient-specific multi-joint kinematic models through two-level optimization.” In: *Journal of biomechanics* 38 3 (2005), pp. 621–6 (cit. on p. 58).

[139] Yupeng Ren, Sang Hoon Kang, Hyung-Soon Park, Yi-Ning Wu and Li-Qun Zhang. “Developing a Multi-Joint Upper Limb Exoskeleton Robot for Diagnosis, Therapy, and Outcome Evaluation in Neurorehabilitation”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 21 (2013), pp. 490–499 (cit. on pp. 29, 62–68, 144).

[140] Ana de los Reyes-Guzmán, Iris Dimbwadyo-Terrer, Fernando Trincado-Alonso et al. “Quantitative assessment based on kinematic measures of functional impairments during upper extremity movements: A review.” In: *Clinical biomechanics* 29 7 (2014), pp. 719–27 (cit. on pp. 37, 160).

[141] Brandon Rohrer, Susan E. Fasoli, Hermano Igo Krebs et al. “Submovements grow larger, fewer, and more blended during stroke recovery.” In: *Motor control* 8 4 (2004), pp. 472–83 (cit. on p. 46).

[142] Orlando Romero, Sarthak Chatterjee and Sérgio Daniel Pequito. "Convergence of the Expectation-Maximization Algorithm Through Discrete-Time Lyapunov Stability Theory". In: *2019 American Control Conference (ACC)* (2019), pp. 163–168 (cit. on p. 97).

[143] Hiroaki Sakoe and Seibi Chiba. "Dynamic programming algorithm optimization for spoken word recognition". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 26 (1978), pp. 159–165 (cit. on pp. 97, 148).

[144] Eduardo Salas, Katherine A. Wilson, C. Shawn Burke and Heather A. Priest. "Using simulation-based training to improve patient safety: what does it take?" In: *Joint Commission journal on quality and patient safety* 31 7 (2005), pp. 363–71 (cit. on p. 115).

[145] L Santisteban, Maxime Térémetz, J P Bleton et al. "Upper Limb Outcome Measures Used in Stroke Rehabilitation Studies: A Systematic Literature Review". In: *PLoS ONE* 11 (2016) (cit. on p. 49).

[146] Maria Laura Schiefelbein, Ana Paula Salazar, Ritchele Redivo Marchese et al. "Upper-limb movement smoothness after stroke and its relationship with measures of body function/structure and activity - A cross-sectional study". In: *Journal of the Neurological Sciences* 401 (2019), pp. 75–78 (cit. on p. 146).

[147] Anne Schwarz, Christoph M. Kanzler, Olivier Lambercy, Andreas Rüdiger Luft and Janne Marieke Veerbeek. "Systematic Review on Kinematic Assessments of Upper Limb Movements After Stroke". In: *Stroke* 50 (2019), pp. 718–727 (cit. on pp. 37, 75, 131).

[148] Nicolas Schweighofer, Chunji Wang, Denis Mottet et al. "Dissociating motor learning from recovery in exoskeleton training post-stroke". In: *Journal of NeuroEngineering and Rehabilitation* 15 (2018) (cit. on p. 130).

[149] Camila Shirota, Jelka Janvsa, Javier Diaz et al. "On the assessment of coordination between upper extremities: towards a common language between rehabilitation engineers, clinicians and neuroscientists". In: *Journal of NeuroEngineering and Rehabilitation* 13 (2016) (cit. on pp. 14, 21, 23, 24, 33–35, 48, 50).

[150] Tadhg Stapleton and Ciara McBrearty. "Use of Standardised Assessments and Outcome Measures among a Sample of Irish Occupational Therapists Working with Adults with Physical Disabilities". In: *The British Journal of Occupational Therapy* 72 (2009), pp. 55 –64 (cit. on pp. 20, 33, 34, 36).

[151] Cathy M. Stinear. "Prediction of recovery of motor function after stroke". In: *The Lancet Neurology* 9 (2010), pp. 1228–1232 (cit. on p. 49).

[152] Thomas Sugar, Jiping He, E J Koeneman et al. "Design and Control of RUPERT: A Device for Robotic Upper Extremity Repetitive Therapy". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 15 (2007), pp. 336–346 (cit. on pp. 17, 27, 28).

[153] Jane E Sullivan, Beth E Crowner and Patricia M Kluding. "Case Report Outcome Measures for Individuals With Stroke : Process and Case Description". In: *Physical Therapy* (2013) (cit. on p. 51).

[154] Shelagh M. Szabo, Karissa M. Johnston and Andrew J. Lloyd. “Advanced Data Visualisation in Health Economics and Outcomes Research: Opportunities and Challenges”. In: *Applied Health Economics and Health Policy* 17 (2019), pp. 433–441 (cit. on p. 142).

[155] Edward Taub, Peter S. Lum, Phillip Hardin, Victor W. Mark and Gitendra Uswatte. “AutoCITE: Automated Delivery of CI Therapy With Reduced Effort by Therapists”. In: *Stroke* 36 (2005), pp. 1301–1304 (cit. on p. 60).

[156] Huidi Tchero, Maturin Tabue Teguo, Annie Lannuzel and Emmanuel Rusch. “Telerehabilitation for Stroke Survivors: Systematic Review and Meta-Analysis”. In: *Journal of Medical Internet Research* 20 (2018) (cit. on pp. 18, 143).

[157] Heather Thikey, Madeleine Grealy, Frederike van Wijck, Mark Barber and Philip John Rowe. “Augmented visual feedback of movement performance to enhance walking recovery after stroke: study protocol for a pilot randomised controlled trial”. In: *Trials* 13 (2012), pp. 163 –163 (cit. on p. 142).

[158] Heidi Rose Thornton, Jace A. Delaney, Grant M. Duthie and Ben J. Dascombe. “Developing Athlete Monitoring Systems in Team Sports: Data Analysis and Visualization.” In: *International journal of sports physiology and performance* (2019), pp. 1–8 (cit. on p. 142).

[159] Gyrd Thrane, Katharina Stibrant Sunnerhagen and Margit Alt Murphy. “Upper limb kinematics during the first year after stroke: the stroke arm longitudinal study at the University of Gothenburg (SALGOT)”. In: *Journal of NeuroEngineering and Rehabilitation* 17 (2020) (cit. on p. 131).

[160] Marco Toigo, Martin Flück, Robert Riener and Verena Klamroth-Marganska. “Robot-assisted assessment of muscle strength”. In: *Journal of NeuroEngineering and Rehabilitation* 14 (2017) (cit. on p. 45).

[161] Vi Do Tran, Paolo Dario and Stefano Mazzoleni. “Kinematic measures for upper limb robot-assisted therapy following stroke and correlations with clinical outcome measures: A review.” In: *Medical engineering & physics* 53 (2018), pp. 13–31 (cit. on pp. 2, 23, 35, 37–39, 45, 48–50, 160, 221).

[162] Konstantinos Tsiakas, Manfred Huber and Fillia Makedon. “A multimodal adaptive session manager for physical rehabilitation exercising”. In: *Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments* (2015) (cit. on p. 60).

[163] Meghan Reading Turchioe, Annie C. Myers, Samuel Isaac et al. “A Systematic Review of Patient-Facing Visualizations of Personal Health Data”. In: *Applied clinical informatics* 10 4 (2019), pp. 751–770 (cit. on p. 142).

[164] Kathrin Tyryshkin, Angela M Coderre, Janice I. Glasgow et al. “A robotic object hitting task to quantify sensorimotor impairments in participants with stroke”. In: *Journal of NeuroEngineering and Rehabilitation* 11 (2013), pp. 47 –47 (cit. on pp. 36–38).

[165] Wang Wendong, Li Hanhao, Xiao Menghan et al. “Design and verification of a human-robot interaction system for upper limb exoskeleton rehabilitation.” In: *Medical engineering & physics* (2020) (cit. on pp. 59, 62–67).

[166] Kimberly J. Wisneski and Michelle Jillian Johnson. "Quantifying kinematics of purposeful movements to real, imagined, or absent functional objects: Implications for modelling trajectories for robot-assisted ADL tasks**". In: *Journal of NeuroEngineering and Rehabilitation* 4 (2006), pp. 7 –7 (cit. on p. 58).

[167] Danny T. Y. Wu, Annie T. Chen, John D. Manning et al. "Evaluating visual analytics for health informatics applications: a systematic review from the American Medical Informatics Association Visual Analytics Working Group Task Force on Evaluation". In: *Journal of the American Medical Informatics Association* 26 (2019), pp. 314–323 (cit. on p. 163).

[168] Evagoras G. Xydas and Loucas S. Louca. "Planar conformity of movements in 3D reaching tasks for persons with Multiple Sclerosis." In: *Human movement science* 62 (2018), pp. 221–234 (cit. on pp. 54, 62–68).

[169] H. Zhang, Chunyan Miao and Han Yu. "Fuzzy logic based assessment on the adaptive level of rehabilitation exergames for the elderly". In: *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)* (2017), pp. 423–427 (cit. on p. 115).

[170] Wenbing Zhao, Roanna Lun, Deborah D. Espy and M. Ann Reinthal. "Realtime Motion Assessment For Rehabilitation Exercises: Integration Of Kinematic Modeling With Fuzzy Inference". In: *Journal of Artificial Intelligence and Soft Computing Research* 4 (2014), pp. 267 –285 (cit. on p. 61).

[171] Shou-Han Zhou, Justin Fong, Vincent Crocher et al. "Learning control in robot-assisted rehabilitation of motor skills - a review". In: *Journal of Control and Decision* 3 (2016), pp. 19–43 (cit. on p. 56).

List of Figures

List of Tables

List of Algorithms

A

CHU-PRC Dataset Description

> *It is a capital mistake to theorize before one has data.*
>
> — **Sherlock Holmes**
> A study in scarlet Arthur C. Doyle

A.1 Exoskeleton Dataset Description

A.1.1 Raw Device Data

The device presents raw exercise recording sampled at 45Hz. Dump files contain the following columns: Timestamp; x; y; z; Inner Shoulder Angle; Outer Shoulder Angle; Upper Arm Angle; Elbow Angle; Forearm Angle; Pro-/Supination Angle; Flex-/Extension Angle; Grip. This file represents the raw sensory input of the device retrieved via the acquisition card to the accompanying PC.

Data files contain the following columns: Timestamp; x; y; Object number; Start; Object caught; Object disappeared; Ideal path; Sum of current path; Path ratio. The file contains the data relative to the exercise, logging events and task specific information. Exercise data is sampled at a rate of 64Hz.

The data was processed using Python scripts into a single database (CSV file) containing the following columns: Patient, Date, ID, dx, dy, d, time, vx, vy, v, pathSum, idealPath, pathRatio, result. The rows were filled by columns variables calculated relative to a single ET Fig. 1.4.

A.1.2 Population Data

The following table describes the population represented in the CHU-PRC exoskeleton dataset. The patient pathologies represented in the dataset are the following:

- UL Dystonia Writer's Cramp;
- Hemiplegia;
- Tetra paresis;
- Plexus brachial;
- Traumatic elbow;
- Post traumatic tetra paresis;
- Guillain Barré Syndrome (Polyradiculonevrite) tetra paresis;
- Rhizarthrose of the thumb;
- Hemiparesis post traumatic;
- Tetra paresis sequelae;
- Polytraumatized;
- Hemiparesis post traumatic;
- Ridder post traumatic.

Only 27 subjects had data for the assessment task and were available for use in the following studies.

Tab. A.1.: Description of the population at CHU-PRC

Variable	Details
Age	37(19) years
Target Arm	31(Right)
Post CVA	20
Pathology	17
Vision	2
Gender	27 Males
ADL Capable	12
Inpatient	5

B ICF Classification Models

> *Functioning status is not dependent on the cause of the health condition.*
>
> — **WHO ICF**

B.1 ICF categories

The model provides a comprehensive overview of functioning and disability in relation to the biological, individual, and social perspectives of health. The model maintains a neutral position towards the cause of health condition. Rather it focuses on the dynamics between the functionning and disability and the health condition including personal and environmental influences.

The ICF categories defined by the World Health Organization [161] are:

- Body Functions and Structure;
- Activities the execution of task or action;
- Body function refer to physiological body function;
- Body structure refer to anatomical organs or body parts;
- Impairments are problems in body function such as loss or degradation;
- Activity limitations are the difficulties; individuals have executing tasks;
- Participation is the involvment in different life situations;
- Restrictions or barriers are the problems hindering the involvement of the person in life activities;
- Environmental factors is made up of the outer environment in which individual carry out tasks and live daily;

- Personal factors features related to the individual. These are features not pertaining to the health condition such as lifestyle, education, social background, etc.

C

CASP Diagnostic Study Checklist

> *Could the results of the test have been influenced by the results of the reference standard?*
>
> — **CASP Checklist**

C.1 CASP Checklist Items

The CASP checklist items according to [26] are the following:

- Was there a clear question for the study to address?
- Was there a comparison with an appropriate reference standard?
- Did all patients get the diagnostic test and reference standard?
- Could the results of the test have been influenced by the results of the reference standard?
- Is the disease status of the tested population clearly described?
- Were the methods for performing the test described in sufficient detail?
- What are the results?
- How sure are we about the results? Consequences and cost of alternatives performed?
- Can the results be applied to your patients/the population of interest?
- Can the test be applied to your patient or population of interest?
- Were all outcomes important to the individual or population considered?
- What would be the impact of using this test on your patients/population?

D Concepts, Definitions & Prerequisites

> *Users do not care about what is inside the box, as long as the box does what they need done.*
>
> — **Jef Raskin**
> about Human Computer Interfaces

D.1 Bayes Theorem

Bayes rule is the assumption that a probability of a given event can be updated once we consider the available evidence in addition to our prior knowledge of assumptions. Let A, B be two events with probabilities $P(A)$ and $P(B)$ known as prior probabilities of each happening independently. The posterior probability introduces $P(A|B) = \frac{P(A)P(B|A)}{P(B)}$, where $P(B|A)$ is the conditional probability that B happens $P(B)$ given A happened.

D.2 The Joint Probability

The probability of two events A and B happening together is given by the formulation: $P(A, B) = P(A|B)P(B) = P(B|A)P(A)$.

Similarly $P(A, B, C) = P(A|B, C)P(B|C)P(C)$.

D.3 The law of total probability

The probability of an event conditional of mutually exclusive events B_i is: $P(A) = P(A|B_1)P(B_1) + \cdots + P(A|B_i)P(B_i)$

D.4 Hidden Markov Model

The hidden Markov model is composed of Markov states $Z_{1:T}$. The model is represented by:$Z_0 \in \mu_0$ the initial distribution.

$P(Z_k = j|Z_{k-1} = i) = T_k(i, j)$,

$Y_k|(Z_k = i) \sim g_k(.|i)$ emission distribution.

In general state space models we assume for state Z_0 has initial density P_0.

$$Z_k|(Z_{k-1} = z_{k-1}) \sim f_k(Z_k|Z_{k-1}) \tag{D.1}$$
$$Y_k|(Z_k = z_k) \sim h_k(y_k|z_k) \tag{D.2}$$

Where f_k, h_k are conditional densities for each k. The case is further extended to account for the Gaussian state space with non linearity assuming $Z_0 \sim N(\mu_0, \sigma_0)$ with $\mu \in \mathbb{R}^d$, $\sigma_0 \in \mathbb{R}^{d\times d}$.

$$Z_k = f_k(Z_{k-1}) + W_k; W \sim N(0, Q_k) \tag{D.3}$$
$$Y_k = h_k(Z_k) + U_k; U_k \sim N(0, R_k) \tag{D.4}$$

D.5 Extended Kalman Smoother

Kalman filter is considered an optimal estimation method when used in conjunction to the Gaussian distributions. It attains minimum-variance optimality. In a Bayesian sense the inference of model parameters using usual maximum likelihood estimates may present high levels of uncertainty which affects directly the estimates made using the model.

The case calls for the use of a bayesian approach where a posterior expectation is set for the inference. The joint state space model is represented by the joint probability: $p(\theta, x_{1:T}, y_{1:t}) = p_{prior}(\theta)p(x_{1:T}, y_{1:T}|\theta)$ Where θ is a parameters vector. The missing information in an continuous space discrete time HMM models can be estimated using the Extended Kalman Smoother. The Kalman filter is:

$$\mu_{k|k-1} = F_k \mu_{k|k_1} \tag{D.5}$$
$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^\intercal + Q_k \tag{D.6}$$

$$v_k = y_k - H\mu_{k|k-1} \tag{D.7}$$
$$S_k = H_k P_{k|k-1} H_k^\intercal + R_k \tag{D.8}$$
$$\mu_{k|k} = \mu_{k|k-1} + P_{k|k-1} H_k^\intercal S_k^{-1} v_k \tag{D.9}$$
$$P_{k|k} = (I - P_{k|k-1} H_k^\intercal S_k^{-1} H_k) P_{k|k-1} \tag{D.10}$$

The smoothed means and covariances can be calculated using the backward pass as follows:

$$\mu_{k|T} = \mu_{k|k} + C_k(\mu_{k+1|T} - \mu_{k+1|k}) \tag{D.11}$$
$$P_{k|T} = P_{k|k} + C_k(P_{k+1|T} - P_{k+1|k} C_k^\intercal) \tag{D.12}$$

The case of an extended kalman filter for continuous non linear dynamics models the previous implementation is updated such that: $\mu_{k|k-1} = f(\mu_{k-1|k-1})$ and $v_k = y_k - h_k(\mu_{k|k1})$. The Jacobian matrices $F_k = f_k'(\mu_{k-1|k-1})$ and $H_k = h_k'(\mu_{k|k-1})$. The case of general models assume that f and h are non linear dynamics and assuming noise is Gaussian. The system might be locally linearized to approximate the distribution which remain thus Gaussian. The algorithm is called extended Kalman filter.

D.6 Expectation maximization

The expectation maximization, EM, algorithm was first introduced by Baum and Petrie and further developed by Dempster et al [52]. Parameter estimation for the HMM model can be done using the Expectation Maximization algorithm. The idea behind it, is maximization of the pseudo log-likelihood and estimate new parameters based on the recursive state estimations. It defines the likelihood function for computing θ a parameter vector based on information in dataset Y as follows:

$$L(\theta) = E_0\left[\frac{dP_\theta}{dP_0}|Y\right] \tag{D.13}$$
$$L(\theta) = logP(Y|U,\theta) = log\int_Z P(Z,Y|U,\theta)dZ \tag{D.14}$$

The maximum likelihood estimate of the parameters vector is $\hat{\theta} = argmax_{\theta \in \Theta} L(\theta)$. θ is assumed to be the parameter vector that maximizes the expectation of the density. Assuming the hidden state is governed by $P(Z)$ using the Jensen inequality we can state:

$$
\begin{aligned}
log \int_Z P(Z,Y|U,\theta) &= log \int_Z P(Z) \frac{P(Z,Y|U,\theta)}{P(Z)} dZ && \text{(D.15)} \\
&\leq \int_Z P(Z) log \frac{P(Z,Y|U,\theta)}{P(Z)} dZ && \text{(D.16)} \\
&= \int_Z P(Z) log P(Z,Y|U,\theta) dZ - \int_Z P(Z) log P(Z) && \text{(D.17)} \\
&= Q(P,\theta) && \text{(D.18)}
\end{aligned}
$$

This allows us to say $Q(P,\theta) \leq L(\theta)$ A lower bound on the likelihood function. The EM step alternatively maximizes the quantity Q with respect to the marginals P and the parameters vector θ.

$$
\begin{aligned}
E_{step} : P_{k+1} &\leftarrow argmax_P Q(P,\theta) && \text{(D.19)} \\
M_{Step} : \theta_{k+1} &\leftarrow argmax_\theta Q(P_{K+1},\theta) && \text{(D.20)}
\end{aligned}
$$

Maximizing E step coincides with the conditional of the hidden state Z such that $P_{k+1}(Z) = P(Z|Y,U,\theta_k)$. This also is shown to verify the bound equality such as $Q(P_{k+1}(Z),\theta_k) = L(\theta_k)$.

Maximizing the M step coincides with maximizing the second term namely: $\theta_{k+1} \leftarrow \int_Z P(Z) log P(Z,Y|U,\theta) dZ$ Where $P_{k+1}(Z) = P(Z|Y,U,\theta_k)$. To solve this optimization an iterative approach is proposed as:

Algorithm 5 Basic EM procedure

Step 1: set $i = 0$, initilize $\theta = \theta_0$
Step 2: E-step let $\theta^* = \theta_i$ and find $Q(.,\theta^*)$
$Q(\theta,\theta^*) = E_{\theta^*}\left[log \frac{dP_\theta}{dP_{\theta^*}}|Y\right]$
M-step: $\theta_{i+1} = argmax Q(\theta,\theta^*)$
replace i by $i+1$ and repeat from step 2.

This procedure is suggested to iteratively increase our pseudo loglikelihood $Q(\theta,\theta^*)$.

Assuming the observations are independent and the noise signals are independent and assuming the Markov property of the state Z, the loglikelihood of the joint distribution of the form:

$$logL(Y, Z; \theta) = -\sum_{1}^{T} \frac{1}{2}(y_t - h(z_t))^{\intercal} R^{-1}(y_t - h(z_t)) \tag{D.21}$$

$$-\frac{T}{2} log\,|R| - \sum_{1}^{T} \frac{1}{2}(z_t - f(z_t))^{\intercal} Q^{-1}(z_t - h(z_t))$$

$$-\frac{T}{2} log\,|Q| \tag{D.22}$$

D.7 Dynamic Time Warping

A dynamic programming procedure for optimal time axis alignment. Optimality is decided using a cost function usually refering to Euclidean distances. Considering a time series sampled at the same rate, we consider asymmetric cases where observation series are transformed into the latent state time scale.

Let's assume X_i and Y_j to be our times series for $i \leq I, j \leq J \in N$; for a matrix $D \in \mathbb{R}^{I \times J}$ the alignment consists of finding a mapping $A(j) = i$, A is called a time warping function.

Publication

- Meziani Y., Hadj-Abdelkader A., Morere Y., Bourhis G., Karaouzène M., Benmansour M., Preanalysis of Exoskeleton Data for Physical Rehabilitation Parameterization, *Ecole de Printemps IVAR School 2019 Alger Algérie*.

- Meziani Y., Hadj-Abdelkader A., Morere Y., Bourhis G., Karaouzène M., Benmansour M., In Search of a Quantifable Shoulder Rehabilitation Measure : Evaluating Movement Kinematics Using an Orthotic Exoskeleton - Case Study, *JCJC2019 June 23, Paris, France*.

- Meziani Y., Hadj-Abdelkader A., Morere Y., Bourhis G., Karaouzène M., Benmansour M., Clinimetric Evaluation of Kinematic Measures of Upper Extremity Functional Rehabilitation, *Actes Conférence Handicap 2020, Paris, France*.

- Meziani Y., Hadj-Abdelkader A., Morere Y., Bourhis G., Karaouzène M., Benmansour M., [SMap H-Map]:Visual indicators of rehabilitation as patient's accessible feedback, case study of upper limb rehabilitation using an orthotic exoskeleton, *Actes Conférence Handicap 2020, Paris, France*.

- Meziani Y., Morere Y., Hadj-Abdelkader A., Benmansour M., Bourhis G., Towards adaptive and finer rehabilitation assessment: A learning framework for kinematic evaluation of upper limb rehabilitation on an Armeo Spring exoskeleton, *Control Engineering Practice, June, 2021*

- Meziani Y., Hadj-Abdelkader A., Morere Y., Bourhis G., Can visualization be Used as Objective Upper Limb Rehabilitation Assessment Instruments ? Patient-centered Graphics S-Map and H-Map Using Kinematic Data Captured by an Orthotic Exoskeleton - *Under Review*.

- Meziani Y., Hadj-Abdelkader A., Morere Y., Bourhis G., Karaouzène M., Benmansour M., Personalized Robotic Physical Therapy and Assessment Through Patient-Specific Kinematic Modeling : Evidence From a Clinimetrics Study - *Under Review*.

- Meziani Y., Hadj-Abdelkader A., Morere Y., Bourhis G., Advanced Objective Assessment Techniques for Robot Mediated Therapy : A Systematic Review - *Under Review*