BEN GURION UNIVERSITY OF THE NEGEV
FACULTY OF ENGINEERING SCIENCES
DEPARTMENT OF INDUSTIRLA ENGINEERING AND MANAGMENT

STATISTICAL MOVEMENT MODELING FOR SKILL ENCODING AND ANALYSIS

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIRMENETS FOR THE M.Sc. DEGREE

By: Isgav Davidowitz

SEPTEMBER 2017

BEN GURION UNIVERSITY OF THE NEGEV
FACULTY OF ENGINEERING SCIENCES
DEPARTMENT OF INDUSTIRLA ENGINEERING AND MANAGMENT

STATISTICAL MOVEMENT MODELING FOR SKILL ENCODING AND ANALYSIS

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIRMENETS FOR THE M.Sc. DEGREE

By: Isgav Davidowitz

Supervised By: Prof. Sigal Berman

Author:…… Date:……..

Supervisor:……..……. Date:……..

Chairman of Graduate Studies Committee:…… Date:……..

SEPTEMBER 2017

ABSTRACT

Data-driven stochastic movement modeling offers a powerful method for both movement analysis and skill generation. Such models facilitate integration of time, space, and variability into a single coherent representation. The model utility can be enhanced by adapting the space within which it is generated, e.g., the model can be generated in the original space or in a latent space found using dimensionality reduction methods. Within the current thesis we fit two dimensional, spatio-temporal, Gaussian mixture models to sets of repetitive human motions. We apply the modeling method to motion analysis in patients with stroke and to enhancing robotic motion learning efficiency. For analysis of motion in patients with stroke, we fit the model in the original motion space, to reach-to-grasp motion of patients with stroke and control subjects. A suitable goodness of fit measure, the symmetric Kullback-Liebler Divergence, was used for measuring model similarity between patients with stroke and the control group. The similarity is used to assess the effect of spasticity on motion kinematics. The similarity of elbow extension motion between patients with stroke and control subjects was found to be related to the Biceps modified Ashworth scale (which is a commonly used measure of spasticity). Moreover, the similarity was higher for a target positioned closer to the subject, as expected for patients with spasticity, which have a limited muscle control zone. This is the first study to relate kinematic aspects of movement disorders to spasticity. For robotic motor skill generation, the model was developed in a latent space and was combined with Gaussian mixture regression and Reinforcement Learning to facilitate efficient motion adaptation to task variations, based on the method of Zhang, Zhang and Parker (2015). For enhancing the efficiency of the method, we suggest dynamically changing the search space during run-time and using an adaptive probability for a random action, along with expending the action space based on the feature contribution to the latent space. The method was applied to teach a RV-2F-1D Mitsubishi industrial robot to perform a reach-to-grasp task and facilitate adaptation to different goal locations. The suggested method required less iterations to adapt to new goals when compared to the original method.

Key words: Gaussian Mixture Models, Kullback-Liebler Divergence, Spasticity, Modified Ashworth Scale, Reinforcement Learning, Robotics, Stroke.

ACKNOLEDGMENTS

Foremost, I wish to express my thanks to my advisor Prof. Sigal Berman, for imparting her knowledge and expertise in this research, with great patience and willingness to help in any circumstance, always heartening with a smile. It is difficult to find words adequate to express my gratitude, thanks for everything you taught me.

Special thanks to Prof. Yisrael Parmet, for teaching and making me passionate about in-depth statistics, sharing his wisdom and providing advice.

I would like to thank Prof. Dario G. Liebermann and Dr. Mindy F. Levin for giving professional advice whenever asked, encouraging to continue.

Another thanks to Rotem Duani, providing friendly encouragement, assisting in brainstorming and ideas.

Thanks to Silvi Frenkel-Toledo, Melanie C. Baniña, Dr. Nachum Soroker, Dr. John M. Solomon, Rotem Duani and Akash Shah for their assistance with the data collection. Thanks to my lab colleagues, Noam Peles, Gil Baron, Moshe Bardea, and Yossi Zehavi, who accepted any request, solved technical issues of great difficulty and contributed to this research success.

Finally, thanks to family and friends, for tolerating me after sleepless nights and providing encouragement at all times.

This research was supported by the Canada-Israel Binational cooperation (Mindy F. Levin, Dario G. Liebermann) Joint Canada-Israel Health Research Program, funded by the Israel Science Foundation (ISF) grant number 2392\15, the Azrieli foundation, the Canadian Institutes of Health Research (CIHR), the International Development Research Center (IDRC) grant number 108186-001, and the Helmsley Charitable Trust through the Agricultural, Biological and Cognitive (ABC) Robotics Center of Ben-Gurion University of the Negev.

Table of Contents

List of Figures

List of Tables

Abbreviations

ABAD	Abduction-Adduction
AI	Artificial Intelligence
AS	Action Space
BIC	Bayesian Information Criteria
CNS	Central Nervous System
CoD	Curse of Dimensionality
DoF	Degree of Freedom
DR	Dimensionality Reduction
EE	Elbow Extension
EM	Expectation Maximization
ET	Eligibility Traces
FA	Final Angle
FC	Far Center
FCL	Far Contralateral
FE	Flexion-Extension
FIL	Far Ipsilateral
FMA	Fugl-Mayer Assessment
ASD	Action Space Decomposition
GMM	Gaussian Mixture Model
GLMM	General Linear Mixed Model
GRNN	General Regression Neural Network
KLD	Kullback-Liebler Divergence
MAS	Modified Ashworth Scale
MC	Monte Carlo
MCP	Metacarpophalangeal Joints
MoC	Magnitude of Contribution
NC	Near Center
PbD	Programming by Demonstration
PC	Principal Components
PCA	Principal Component Analysis
PS	Pronation-Supination
RL	Reinforcement Learning
ROT	Rotation
SA	Shoulder Abduction
SARSA	State-Action-Reward-State-Action
SE	Shoulder Extension
ST-GMM	Spatio-Temporal GMM
ST-KLD	Spatio-Temporal KLD
TD	Temporal Difference
TSRT	Tonic Stretch Reflex Threshold
UCM	Uncontrolled Manifold
UL	Upper Limb

1. Introduction

1.1 Background

Motor control is the process by which humans activate and coordinate muscles and limbs as part of the performance of a motor skill. This process requires interaction between the central nervous system (CNS) and the musculoskeletal system, where the components involved are required to act in unison to produce movement. This coordination problem encapsulates issues of hard weird reflexes, synergies, redundancies, and variability, making the understanding of the process involved complex.

The ability to model human motion can benefit a variety of applications: for purposes of motion analysis, quantifying motion quality, or for motion generation of a robotic system based on human data. As variability, in both time and space, is an inherent characteristic of human motion, encapsulating variability by a generated model can lead to a better representation.

Stochastic models represent probability distributions of potential outcomes by allowing random variations in the input variables. Such models have been used in many fields, e.g., pattern recognition, economical modeling, and quality engineering (Cohn, Ghahramani, and Jordan 1996; Chen and Gopalakrishnan 1998). Spatio-temporal stochastic models can similarly be used to model spatial signals representing motion. Previous work in motor control applied such stochastic models in robotics, for developing robotic motion for humanoid robots based on generalization of human demonstration (Calinon, Guenter, and Billard 2007; Zhang, Zhang, and Parker 2015). Multi-dimensional spatial stochastic models were used for learning grasp success distributions (grasp affordance) from human demonstrations (Granville, Fagg, and Southerland 2006).

1.2 Objectives and Contribution

The presented research focuses on modeling human motion using spatio-temporal Gaussian mixture models (ST-GMM). The main advantage of mixture models is that hidden parameters are modeled without explicit assumptions. Hence, the model can easily be applied to different applications without requiring additional assumptions. The use of a stochastic mixture model like ST-GMM allows flexibility in modeling variability due to the unconstrained covariance structure in comparison to deterministic models. When compared to other stochastic models such as stochastic neural network models, ST-GMM require far less data and computation power, and can model probability distributions to any required level of accuracy with enough

components (Hinton et al. 2012). In the current work the model is adapted for two separate applications, spasticity analysis and robotic motion generation.

We developed a method for analysis of the similarity of kinematic scaled motion of patients with stroke to healthy controls based on the stochastic motion models. We show that this similarity is related to spasticity. Spasticity is characterized by the hyper excitability of the velocity-dependent stretch reflex (Calota and Levin 2009). The commonly used clinical measures for spasticity are based on assessing the resistance to a passive stretch of the muscle. Such a measurement does not capture all aspects of spasticity, which is velocity dependent, and its effects on motion quality. Moreover, the existing measures are inherently subjective. In the current work ST-GMM is used to form an objective, data-driven measure of spasticity using advanced goodness of fit analysis based on the Kullback-Liebler Divergence (Priyadharshini, Devi S., and Askerunisa 2014). The relationship of the developed similarity measure, spatio-temporal Kullback-Liebler Divergence (ST-KLD) to spasticity, is validated for Biceps muscle spasticity, with respect to modified Ashworth scale (MAS), which is a clinical spasticity measure (Bohannon and Smith 2014). ST-KLD is developed as part of the ENHANCE clinical project (Levin et al., submitted). The goal of ENHANCE is to identify effective upper limb interventions for recovery of voluntary movement control after stroke. The developed measure, ST-KLD, will be one of the secondary outcome measures used in the ENHANCE project for measuring treatment efficiency through decrease in upper limb spasticity. As part of the required pre-processing, arm joint centers were computed using linear regression analysis (Biryukova et al. 2000; O'Brien et al. 1999) (Appendix A3). The spasticity analysis was summarized in a poster presented at the "Progress in Motor Control XI" conference, Davidowitz et al., Stochastic Spatio-Temporal Movement Modeling, Miami, 2017. A journal publication is currently under development.

We improved a method for efficient learning based on combining ST-GMM with reinforcement learning. Reinforcement learning (RL) requires notoriously long convergence lags, especially in multi-dimensional spaces which are typically encountered in robotic environments. Therefore, improving learning efficiency is of importance, especially in scenarios where the robot must adapt to changes in the environment during run-time. Zhang, Zhang, and Parker (2015) proposed a method for reducing the search space. Their method is based on decomposing task dimensions into principal and non-principal components, and then using the non-principal dimensions as an initial search space for the RL algorithm, following the logic that an adaptation to task variants will more probably happen in within the non-principal component space. To ensure algorithm convergence, the search is re-initiated

using a search space with additional dimensions in case a solution is not found in the reduced space. In the current study, we propose a modification of this algorithm based on using a dynamically growing space rather than the previous abort-and-re-learn methodology. Moreover, we suggest an improvement to the search rate by changing the probability of a greedy action dynamically. The proposed algorithm was tested using an industrial RV-2F-1D Mitsubishi robot. As part of the development a software environment for controlling the robot was developed using C# (detailed in Appendix B1). This work was presented in the "12th Karniel Computational Motor Control Workshop": Davidowitz and Berman, Robot Motion Learning and Adaptation, Beer Sheva, 2016.

1.3 Work Scope

In the current work we developed two different methods based on ST-GMM. The developed similarity measure, ST-KLD, was tested using data recorded from subjects with stroke before they underwent training. Subject motion analyzed included 13 healthy control subjects that were recorded in August 2016, and 16 subjects with stroke that were recorded from August 2016 to March 2017 in Canada, India and Israel. According to the study protocol, 20 additional patients with stroke from the three countries will be recorded by October 2018. These additional patients will not be analyzed in the current work. Additionally, recordings of the post treatment will not be analyzed in the current work as these are currently sealed.

The developed robotic learning methods were tested for a reach-to-grasp task using a six degrees of freedom robot. Demonstration data used was obtained from an experiment conducted by Sagi et al., (2015) in which subjects remotely controlled a UP6 Motoman robot. These demonstration retain only Cartesian endpoint coordinates. Accordingly, the modeled profiles include only the arm trajectory, for which joint motion could be extracted from the demonstrated profiles through inverse kinematics, and not the wrist or hand motion. Adaption to include additional degrees of freedom is straightforward, but demonstrations which retain robot joint coordinates are required, e.g. kinesthetic demonstrations.

1.4 Outline

The rest of this thesis is organized as follows: Chapter 2 presents a literature review. Reviewed topics include motion modeling using Gaussian mixture models, spasticity following stroke and spasticity analysis, and robotic motion learning including frameworks for efficient and adaptive trajectory learning. Chapter 3 describes the spasticity modeling and analysis and chapter 4 describes the adaptive robot trajectory learning method and experiment validating it.

2. Literature Review

2.1 Overview

The literature review scans concepts and methods related to stochastic motion modeling. Chapter 2.2 describes Gaussian mixture models for spatio-temporal modeling of motion. The concepts reviewed include model definitions and parameter estimation. A regression based method to utilize the stochastic model to produce generalized movement, and a goodness of fit measure to utilize the stochastic model for analysis and comparison. The model and methods were used within this thesis in two application fields, spasticity analysis in patients with stroke, and robotic motion learning. Accordingly, literature regarding both domains is reviewed. Section 2.3 provides background on spasticity in patients with stroke. The section additionally reviews existing measures for spasticity. Chapter 2.4 describes robotic motion learning through programming by demonstration (PbD) and reinforcement learning (RL). The section reviews previous work regarding the use of stochastic models for improving efficiency of RL.

2.2 Gaussian Mixture Models

2.2.1 Model and definitions

Gaussian Mixture Model (GMM) is a common method in which multidimensional data can be represented as a set of multivariate Gaussians. A GMM is defined as the weighted sum of K multivariate Gaussians probability densities denoted $g(x|\mu_i, \sum_i)$, where μ_i is a matrix of the first moments for all dimensions, and $\sum_i$ is the covariance matrix. The mixture model is:

$$p(x) = \sum_{i=1}^{K} w_i \cdot g(x|\mu_i, \sum_i) \tag{2.1}$$

Where w_i is the weight of Gaussian *i*, and each multivariate Gaussian density *g* is defined by:

$$g(x|\mu_i, \sum_i) = \frac{1}{(2\pi)^{N+1/2}\sqrt{|\sum_i|}} \cdot \exp\left\{-\frac{1}{2}(x-\mu_i)^T {\sum_i}^{-1}(x-\mu_i)\right\} \tag{2.2}$$

Where N+1 is the total number of dimensions, and μ_i and $\sum_i$ are defined by:

$$\mu_i = \left\{\mu_{t,i}, \mu_{S_j,i}\right\}, \quad \sum_i = \begin{pmatrix} \sum_{t,i} & \sum_{ts_1,i} & \cdots & \sum_{ts_N,i} \\ \sum_{s_1 t,i} & \sum_{s_1,i} & \cdots & \sum_{s_1 s_N,i} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{s_N t,i} & \sum_{s_N s_1,i} & \cdots & \sum_{s_N,i} \end{pmatrix}, \text{j} = 1 \ldots \text{N} \tag{2.3}$$

Where $\mu_{S_j,i}$ is the spatial expectation of dimension *j* and $\mu_{t,i}$ is the temporal expectation of the *i* component. $\sum_i$ is the covariance matrix of the *i* component.

When using GMMs to model spatial and temporal data, each data point, $\xi_i = \{\xi_{t,i}, \xi_{S_1,i}, \dots, \xi_{S_N,i}\}$ includes one temporal dimension, $\xi_{t,i}$, and N spatial dimensions, $\xi_{S_j,i}$; $j = 1..N$. For motion along a single spatial dimension, each data point will be defined by $\xi_i = \{\xi_{t,i}, \xi_{s,i}\}$. The expectation vector μ_i will be $\mu_i = \{\mu_{t,i}, \mu_{s,i}\}$ and the covariance matrix will be $\sum_i = \begin{pmatrix} \sum_{t,i} & \sum_{ts,i} \\ \sum_{st,i} & \sum_{s,i} \end{pmatrix}$.

2.2.2 Parameter estimation

Due to the structure of the GMM distribution, parameters cannot be estimated analytically based on maximum likelihood. Expectation-Maximization (EM) algorithm can be used to find a local optimal fit of the GMM to the data (Bishop 2006). The EM algorithm requires an initial estimate of the model parameters μ_i and $\sum_i$, along with the number of Gaussians *K*. K-Means clustering technique can be used for calculating such an estimate given *K* (Guenter et al. 2007). The best model among the calculated models with different numbers of Gaussians *K*, can be determined by a model selection criterion, such as the Bayesian Information Criterion (BIC):

$$BIC = -2\mathrm{L} + T \cdot \ln(N) \tag{2.4}$$

Where *L* is the log-likelihood of the model, *T* is the number of independent parameters in the GMM, and *N* is the number of observations used in fitting the model. Moreover, in order to achieve models with good fit, Guenter et al., (2007) and Billard et al., (2008) suggest temporal normalization. In certain cases, the normalization can be of the form of Dynamic Time Warping in order to handle with time variations between multiple demonstrations made by human demonstrator. However, in most cases a simple scaling of the spatial function can be sufficient.

2.2.3 Gaussian Mixture Regression

Gaussian Mixture Regression (GMR) (Cohn, Ghahramani, and Jordan 1996) can be applied in order to build a general representation of data, following the computation of a GMM. Spatial values are estimated through regression, and temporal values are used as query points. In the specific case of one spatial dimension, the estimated value of $\xi_{s,i}$ denoted $\hat{\xi}_{s,i}$ and the estimated covariance $\hat{\sum}_{s,i}$ are given by:

$$\hat{\xi}_{s,i} = \mu_{s,i} + \sum_{st,i}(\sum_{st,i})^{-1}(\xi_t - \mu_{t,i}) \tag{2.5}$$

$$\hat{\sum}_{s,i} = \sum_{s,i} + \sum_{st,i}(\sum_{t,i})^{-1}\sum_{ts,i}$$

These estimators are mixed by the probability that a time step ξ_t is related to a Gaussian component i for $i = \{1, \dots, k\}$:

$$r_i = \frac{p(\xi_t|i)}{\sum_{j=1}^{K} p(\xi_t|j)} \tag{2.6}$$

Calculating $\{\hat{\xi}_s, \hat{\sum}_s\}$ given ξ_t, for all time steps *t*, form a generalized trajectory of the data, and is achieved by combining equations (2.5) and (2.6):

$$\begin{aligned} \hat{\xi}_s &= \sum_{i=1}^{K} r_i \hat{\xi}_{s,i} \\ \hat{\sum}_s &= \sum_{i=1}^{K} {r_i}^2 \hat{\sum}_{s,i} \end{aligned} \tag{2.7}$$

2.2.4 Measuring goodness of fit for mixture models

A GMM distribution cannot be easily compared using classical fit measures like Chi-squared due to the multivariate nature of the distribution. Even log likelihood based fit measures cannot be used directly since no closed form exists for the log likelihood of a GMM (when two or more Gaussians are mixed). An alternative log likelihood-based measure which can be used for GMMs is the Kullback-Liebler Divergence (KLD), also called relative entropy. KLD is a measure of how one probability distribution diverges from a second expected probability distribution (Hershey and Olsen 2007), useful especially when the difference between the distributions cannot be captured by geometric distances (Priyadharshini, Devi S., and Askerunisa 2014; Goldberger and Aronowitz 2005). Given a probability distribution, Q, and an expected distribution, P, KLD is denoted by $D_{KL}(P||Q)$ and is defined in its general form as:

$$D_{KL}(P||Q) = E_P[log\, p(x) - log\, q(x)] \tag{2.8}$$

When P and Q are continues variables, KLD is defined by:

$$D_{KL}(P||Q) = \int_{-\infty}^{\infty} p(x)\, log\left(\frac{p(x)}{q(x)}\right) \mathrm{d}x \tag{2.9}$$

KLD goodness of fit measure can be used as similarity measure between distributions. The higher $D_{KL}(P||Q)$ is, the less similar P and Q are. KLD is always non negative, hence $D_{KL}(P||Q) \geq 0$ and $D_{KL}(P||Q)$=0 if and only if P=Q. KLD is not a distance measure as it is not symmetric, $D_{KL}(P||Q) \neq D_{KL}(Q||P)$, not does it satisfy the triangle inequality.

Since KLD is based on the log likelihood, a Monte Carlo simulation or a lower bound approximations can be used, as described by Hershey and Olsen, (2007). In order to overcome the asymmetry of the measure, a symmetric variant can be applied:

$$D_{Symmetric-KL}(P||Q) = \frac{D_{KL}(P||Q) + D_{KL}(Q||P)}{2} \qquad (2.10)$$

Various alternatives exists to KLD, among them are the distance measures- normalized L2 norm and Earth Movers Distance measure. However, KLD performs superior to those measures especially when the number of components in the mixtures is higher than two (Jensen et al. 2007).

2.3 Spasticity following stroke

Stroke is characterized by a sudden onset of clinical signs related to the site in the brain where the morbid process occurs (Sommerfeld et al. 2004). Upper motor neuron syndrome are changes that can occur in skeletal muscle after stroke, where affected muscles may potentially have features of altered performance such as weakness, decrease in accuracy and dexterity, paralysis, and exaggerated tendon reflexes. The combination of tonic contraction of the muscles in response to a stretching force, referred to as tendon reflex, along with paralysis and the damage to the CNS, is often referred to as spasticity. Spasticity is a "motor disorder characterized by a velocity-dependent increase in tonic stretch reflexes (muscle tone) with exaggerated tendon jerks, resulting from hyper excitability of the stretch reflex"(Sommerfeld et al. 2004). Spasticity is one of the most common sequelae of CNS lesions and stroke. Levin et.al (2000) suggested that spasticity may be characterized by the limitation of the CNS to regulate the range of stretch-reflex thresholds in flexor and extensor muscles, and showed that for patients with stroke the ability to regulate muscle force in all part of the physiological range (control zone) may be lost due narrowing of the limits of stretch-reflex regulation thresholds. Measuring spasticity can benefit physical therapy treatments and patients with stroke condition evaluation. The following sections will present existing measures for spasticity, and the use of the ST-GMM combined with ST-KLD method as an alternative.

2.3.1 Modified Ashworth Scale

The Modified Ashworth Scale (MAS) is a discrete subjective measure for spasticity (Bohannon and Smith 2014). MAS is a 7 grades ordinal scale, which assigns grades to a manually determined resistance of muscle to passive stretching (Figure 1). Experiment conducted by

Bohannon and Smith (2014) showed that two raters which measured MAS on thirty patients had a significantly correlated ratings and 86.7% agreement on ratings. However, the subjective nature of MAS may require a few experienced testers to achieve accuracy in spasticity measure.

Grade	Description
0	No increase in muscle tone
1	Slight increase in muscle tone, manifested by a catch or by minimal resistance at the end of the range of motion (ROM) when the affected part(s) is moved in flexion or extension
1+	Slight increase in muscle tone, manifested by a catch, followed by minimal resistance throughout the remainder (less than half) of the ROM
2	More marked increase in muscle tone through most of the ROM, but affected part(s) easily moved
3	Considerable increase in muscle tone, passive movement difficult
4	Affected part(s) rigid in flexion or extension
9	Unable to test

Figure 1 -Modified Ashworth Scale grades [1]

2.3.2 Tonic Stretch Reflex Threshold

Stretch Reflex Threshold (SRT) is a way of characterizing spasticity based on equilibrium point hypothesis of motor control (Lambda Model) (Calota and Levin 2009). SRT Threshold (TSRT) is a threshold joint position. When TSRT is measured accurately enough, it can be used as a measure for spasticity. TSRT evaluate spasticity based on the excitability of motoneurons resulting from both descending and segmental influences (Calota and Levin 2009). The measure of those influences is the stretch reflex threshold. The study done by Jobin and Levin (2000) showed that when spasticity is present in the elbow, TSRT lies within the biomechanical range of motion elbow. Moreover, TSRT was found inversely correlated with the degree of clinical spasticity measured. TSRT can perform well for measuring spasticity especially for subjects with moderate to high levels of spasticity (Calota and Levin 2009). However, Low intra-evaluator reliability was showed in the experiments conducted by Calota and Levin, (2009) when measuring TSRT at test and re-test for three evaluators.

[1] https://i.pinimg.com/originals/58/ab/8a/58ab8ad74acf44d8305e009bb7a432cd.jpg

2.3.3 Alternative measures

The are several additional measures of spasticity in use, where each measure has some limitation (Burridge et al. 2005) compared to MAS. The Disability Rating Scale, is measured by a self-report scale filled by the patient. The report reflects how difficult it is to handle the arm. However, this measure relies on the patients report abilities that may be hindered due to the stroke. Some measures are muscle specific, such as the Wartenburg pendulum test which is used to test quadriceps muscle spasticity only, or the Hand-held dynamometer which can be used to test the calf-muscle spasticity. The H-Reflex measure assesses the response to electrical or mechanical stimulation. This technique is simple to perform and easy to use in neurology setting, yet it have low correlations with other clinical scales (Burridge et al. 2005).

2.4 Robotic motion learning

2.4.1 Programming by Demonstration

In robotics, the combination of demonstrations is a framework called Programming by Demonstrations (PbD). PbD is used for teaching a robot how to perform a desired task using samples of a human performing the task, without the need to explicitly program every detail of the task (Billard et al. 2008). PbD methods create state-action mapping which is referred to as Policy (Argall et al. 2009). In movement planning tasks, the goal is to achieve a desired trajectory. GMR can be used for representing a trajectory, in case a GMM is learned based on the demonstrations. Downside of relying on a set of demonstrations for GMR is the problem of dealing with undemonstrated states (Argall et al. 2009) and adapt to new variations of the demonstrated task. A solution is to integrate Reinforcement Learning (RL) methods (Kober and Peters 2011) that will allow to learn optimal trajectory based on the GMR results. The use of prior knowledge helps RL algorithms to achieve better results (Strosslin and Gerstner 2003).

2.4.2 Reinforcement Learning

2.4.2.1 *Reinforcement Learning in Continuous spaces*

Reinforcement learning defines a learning problem. In this framework, a mapping from environment situations to actions is created using a numerical reward signal. The RL problem separates between the decision maker (agent) and between the environment (everything else). The environment is represented at each time step t by a state, $S_t \in S$, where S is the set of all possible states. At each time step, the agent can take an action $A_t \in A(S_t)$, where $A(S_t)$ is the set of possible actions in state S_t. As a result from the action taken, the agent receives a numerical reward $R_{t+1} \in \mathbb{R}$ one time step later. A policy is a mapping from state to probabilities of choosing an action, denoted $\pi_t(a|s)$ is the probability of $A_t = a$ given $S_t = s$.

The agent goal is to maximize the cumulative rewards (Sutton and Barto 2012). RL algorithm is a method to solve RL problems.

2.4.2.2 *Gradient Descent SARSA (λ)*

Temporal Difference (TD) methods are RL algorithms which use temporal errors, i.e., the difference between an old estimate and the new estimate of the value function, after each time step *t* (Sutton and Barto 2012; Kober, Bagnell, and Peters 2013). A TD method that can be used for on-policy control is the *SARSA* (State-action-reward-state-action) algorithm, which performing TD updates on action value functions $Q(s,a)$ (Rummery and Niranjan 1994; Sutton and Barto 2012):

$$Q(S_t,A_t) \leftarrow Q(S_t,A_t) + \alpha[R_{t+1} + \lambda Q(S_{t+1},A_{t+1}) - Q(S_t,A_t)] \tag{2.11}$$

Where α is a learning rate (higher α gives more weight to recent information over old information), R_{t+1} is the next reward, and λ is a discount factor. An issue with TD methods is delayed rewards, i.e., if the rewards are not accepted right after a state is visited. In order to overcome delayed rewards, Eligibility Traces (ET) are suggested by Singh and Sutton (1996). Every time a state is visited, a temporal memory trace is created. The trace decays over time gradually. The conventional accumulating trace as shown in (Singh and Sutton 1996; Sutton and Barto 2012), when adjusted for action-value predictions using SARSA:

$$e_{t+1}(s,a) = \begin{cases} \gamma\lambda e_t(s,a) + 1 & , if\ S = s_t\ and\ a = A_t \\ \gamma\lambda e_t(s,a) & , else \end{cases} \tag{2.12}$$

Where λ is the discount factor, γ is trace decay parameter, and e_t is the current trace. As shown by Sutton and Barto (2012), gradient descent as a function approximation method can be combined with the SARSA algorithm and eligibility traces, formalized as gradient descent SARSA(λ):

$$\theta_{t+1} = \theta_t + \alpha\delta_t e_t \tag{2.13}$$

The parameter θ_t defines parametrized function for $Q(s,a)$, where δ_t is the SARSA update and e_t is the eligibility traces, as follows:

$$\begin{aligned} \delta_t &= R_{t+1} + \gamma\, Q(S_{t+1},A_{t+1}) - Q(S_t,A_t) \\ e_t &= \gamma\lambda e_{t-1} + \nabla_{\theta_t} Q(S_t,A_t) \end{aligned} \tag{2.14}$$

Reinforcement Learning algorithms in continuous settings usually suffer from “Curse of

Dimensionality" (CoD) where the number of iterations required for convergence is growing exponentially in high dimensional spaces (Schaal 2002). RL convergence time depends on the state and action space size (Sutton and Barto 2012). Hence, methods to increase learning efficiency are required.

2.4.3 Learning efficiency

2.4.3.1 *Dimensionality reduction*

"Curse of Dimensionality" refers to problems that arise when trying to analyze data set in high dimensional spaces. In movement planning, the number of dimensions is set by the number of active Degrees of Freedom (DoF). Data volumes grow exponentially with the number of dimensions, which lead to an exponential need in memory and computation time. Dimensionality Reduction (DR) methods deal with CoD by reducing the dimensions of the problem for computation purposes.

Principal Component Analysis (PCA) is a statistical method for data analysis (Pearson 1901). The goal of PCA is to identify main components in the data that might represent some theoretical constructs. To achieve this goal, PCA computes a new set of latent variables called Principal Components (PC). Those variables are a linear combination of the original variables, hence making PCA a linear DR method. The new variables are orthogonal to each other so they can span a projected latent space (Abdi and Williams 2010). PCA is the simplest and most popular method to perform DR effectively (Bishop 2006), especially in PbD and RL frameworks. DR with PCA achieve the best mean-square error compared to other linear DR methods (Jackson 1991).

The first principle component is required to have the largest amount of variance, which implies it contains most of the information compared to other components. The second principal component is under the constraint of orthogonality to the first component and to have the largest amount of variance, and so on. In the studies of Calinon and Billard (2005) and Calinon, Guenter, and Billard (2007), the chosen threshold for variance extraction was 98% and in the study of Zhang, Zhang, and Parker (2015) the chosen threshold was 95%. So far, this method has been proven to be an effective way to reduce the dimensionality of human motion data, thus why it is popular in PbD applications.

2.4.3.2 *Action Space Decomposition*

Action Space Decomposition (ASD), as the method presented by Zhang, Zhang, and Parker (2015), propose an algorithm that manipulates the RL problem action space to reduce learning

time by integrating PCA in the learning process. An initial policy is accepted through GMM computation over Principal Components identified by PCA (later re-projected and combined with GMR), and an initial action space is built from the least important components, referenced as non-principal components. The non-principal components are obtained empirically using a lower bound threshold ratio for contributing the least to the latent space, using a Magnitude of Contribution (MoC) formula:

$$MoC(J_i) = \sqrt{\sum_{j=1}^{p} \left(w_{j_i} \lambda_j\right)^2} \tag{2.15}$$

Where J_i are the robot degrees of freedom in original space, *p* the total number of principal components, w_{j_i} and λ_j are eigenvectors and eigenvalues respectively. To guarantee optimality, first search is applied using a decreased action space with decent capabilities to find the optimal solution following the logic of non-principal components should have the highest relevance when trying to adapt to a new situation. In case no optimal solution is found, the action space is scaled up randomly to include more components.

3. Spasticity analysis in patients with stroke

3.1 Overview

We developed a measure of kinematic motion similarity based on stochastic motion modeling with GMM (ST-GMM) and model comparison with KLD. The suggested similarity measure is based on reference demonstrations and kinematic data only, without external raters. A relationship between the motion similarity and spasticity was established using arm motion data recorded as part of the ENHANCE project. The pre-processing and analysis of the data required several stages, part of which will also be used in future analysis of the kinematic data within the ENHANCE project. These stages are depicted in Figure 2. Section 3.7 presents the results of the experiment and section 3.8 discusses them.

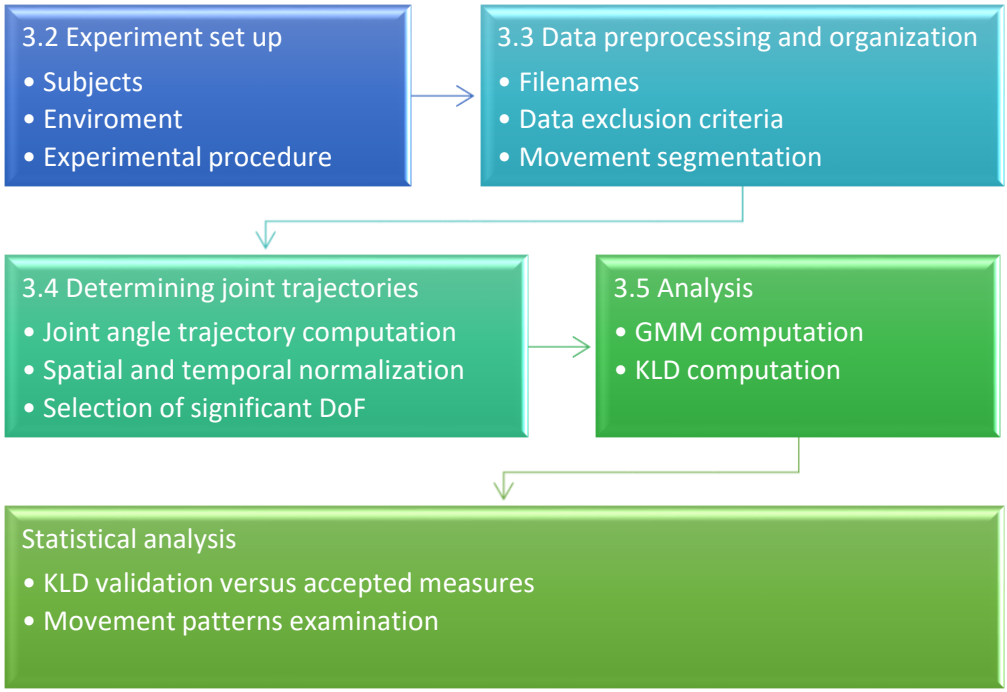

Figure 2 – Pre-processing and analysis stages

3.2 Experiment set up

3.2.1 Subjects

16 subjects with stroke (stroke group, 9 males, age 57.4±11 years), and 13 healthy age-matched controls (control group, 9 males, age 60.46±8.68) with no other neurological, sensorimotor, or orthopedic impairments participated in the experiment. Subjects with stroke were included in the experiment if they have a first ever stroke in the middle cerebral artery area territory, aged 25-75 years, in the sub-acute stage of the stroke (three weeks to six

months post stroke), have arm paresis, able to perform voluntary elbow flexion extension movement of at least 30 degrees, have elbow spasticity and are able to provide informed consent. Subjects were excluded due to clinical issues such as orthopedic problem or pain, major cognitive deficits, history of psychiatric disorders or under medicine treatment. For the stroke group, Upper-Limb (UL) impairment was measured using the Fugl-Meyer Arm Assessment (FMA), which is an integer index to assess the sensorimotor impairment. For UL impairment the score ranges between 9 to 66 points, with lower scores meaning higher impairment. In this project, the maximal score was 56 points as the test did not include reflex testing (10 points). Elbow spasticity was measured using the Modified Ashworth Scale (MAS) for the Biceps and for the Triceps. For the stroke group FMA values were 36.3±9.3, and Biceps MAS was in range of [0, 2]. Demographic data of the subjects are presented in Appendix A1.

3.2.2 Environment and apparatus

Subjects sat on a chair with feet supported and their hand resting alongside body (elbow extended to 180 degrees). The chair had a back support that did not restrict trunk movements. Four targets (standard hollow cones about 10 mm radius × 30 mm height) were placed on the table in front of the subject. Two target locations was in the mid-sagittal plane, one at 2/3rd of arm's length (NC - near center) and one at arm's length (FC – far center). Contralateral and ipsilateral targets was placed at arm's length, about 20-30 cm to the left (FCL – far contralateral) and right (FIL – far ipsilateral) of the mid-sagittal plane, depending on and within reaching distance (Figure 3).

Movements were recorded with a wireless electromagnetic tracking system, G4™ Polhemus (system details are given in Appendix A2). Each sensor has 6 degrees-of-freedom and is tracked at 120Hz. Five sensors (denoted M1-M5) were used to track the position of the upper limb, shoulder girdle, and trunk in real-time. Sensors were placed on the metacarpi-phalangeal (MCP) joint of the index finger (M1), on the dorsal surface of the forearm (1/3 of the length of the forearm proximal to the head of the ulna) (M2), on the lateral surface of the upper arm at about the middle of the upper-arm (M3), on the mid-point of the superior-lateral border of the acromion (M4), and on the mid-sternum (M5). All experiment recordings were saved to Microsoft Excel™ files using Matlab™.

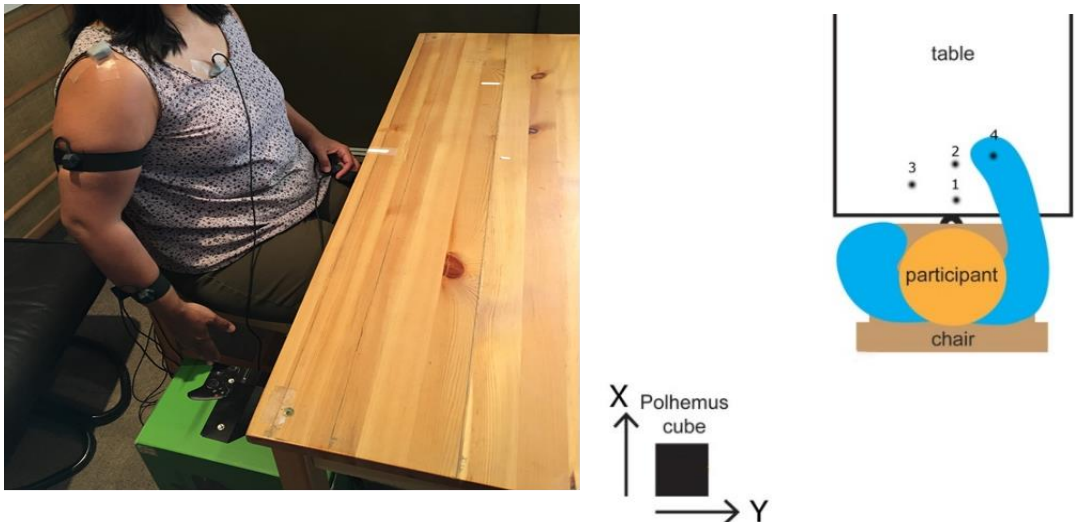

Figure 3 - Sensors M5-M2 positions on arm (without MCP sensor) and experimental setup

3.2.3 Experimental procedure

Each recording session started with calibration of the sensors for the markers locations on the subject's body, which included seven movement types: elbow flexion-extension (FEe), elbow supination-pronation (SPe), wrist flexion-extension (FEw), wrist abduction-adduction (ABADw), shoulder abduction-adduction (ABADs), shoulder flexion-extension (FEs), and shoulder internal-external rotation (ROTs). Moreover, a static calibration was performed for the 4 targets positions and for the chair. Subjects were instructed to perform a reach-to-grasp movement to the targets, based on a visual signaling. They were requested to rest between the sets and allowed to rest when needed between trials. Two sets of 40 trials (10 trials per target, fixed random order) were recorded for a total of 80 movements per subject. Sets order was counter-balanced between subjects.

3.3 Data preprocessing and organization

3.3.1 Filenames

The recording took place in three different centers in Israel, Canada, and India, and involved a large and diverse group of researchers and physicians. To regulate database creation, a naming convention was established and documented. Imposing naming regulations across the group was a difficult task and although most files were recorded correctly, in about 15% of the recorded subjects a typing error caused an invalid filename format. Files were checked for validity per subject and per recording session.

Filenames base name format was "g4_CCNNnS1T1XXX" where "g4_" is fixed, "CC" is the country code ("CA" for Canada, "IL" for Israel, "IN" for India), "NNn" is the subject's name

initials (two capital letters of the forename, one letter of surname), “S1T1” is the set number (1/2) and trial number (1…40), followed by “XXX” for recording session type, “PRE” for pre-treatment trials, “POST” for post treatment trials or “FU” for follow-up trials. Any deviation from this format required a filename change during post processing. Errors were encountered in the subject name and in recording session type. An automated script in Python was developed to automatically change the invalid file names.

3.3.2 Data exclusion criteria

Consecutive sensor measurements with the value “0” indicates a faulty measurement and files with more than 10% faulty measurements were considered damaged and removed from further analysis. Such errors typically occur when the sensor hubs fail to communicate the data to the computer. Recorded movements were determined as erroneous in several cases: the experimenter noted during task execution that the subject did not wait after grasping the cone, the target was misplaced, or the experimenter determined that the subject did not perform the task well (hand collided with the table, task not completed).

3.3.3 Movement Segmentation

Motion segmentation was conducted in order to identify the reach-to-grasp segment of the recorded movements. Segmentation was performed semi-autonomously, and was implemented and executed by Mrs. Rotem Duani[2]. An automatic procedure was developed for initial segmentation and the segmentation results were all manually screened. Movement trajectories were filtered using a Butterworth filter with 6 Hz cutoff frequency (to highlight main trajectory changes). Tangential velocity was computed by differentiating position samples and averaging linear velocity components, and angular velocity was similarly computed for angular components. Motion onset and offset were defined as the times at which the wrist (forward arm sensor, M2) tangential velocity exceeded and remained above, or decreased and remained below 10% peak wrist tangential velocity. Subjects typically performed two sub-movements, they raised their arm and then reached forward towards the target. The sub-movement interchange point was determined between movement onset and offset, when the elbow (upper arm sensor, M3) tangential velocity reached a local minimum.

[2]Industrial Engineering and Management Department, Ben Gurion University of the Negev.

3.4 Determining joint trajectories

3.4.1 Angles definition

Arm kinematics in the form of angles of rotations in the joints were reconstructed from the sensor data (Figure 4). In order to properly define joint rotations, a homogenous transformation matrix T_0^{Mi} was built for each sensor M_i to transform from the global task coordinate system to the sensor coordinate system:

$$T_0^{Mi} = \begin{pmatrix} & R_{zyx}(Ox_{Mi}, Oy_{Mi}, Oz_{Mi}) & & \begin{matrix} X_{Mi} \\ Y_{Mi} \\ Z_{Mi} \end{matrix} \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.1}$$

Where $R_{zyx}(Ox_{Mi}, Oy_{Mi}, Oz_{Mi})$ is the rotation component of the matrix, based on Tait-Bryan angles convention (Diebel 2006) and with accordance to the sensor specifications:

$$R_{zyx}(a,e,r) = \begin{pmatrix} \cos(a)\cos(e) & \cos(a)\sin(e)\sin(r) - \sin(a)\cos(r) & \cos(a)\sin(e)\cos(r) + \sin(a)\sin(r) \\ \sin(a)\cos(e) & \cos(a)\cos(r) + \sin(a)\sin(e)\sin(r) & \sin(a)\sin(e)\cos(r) - \cos(a)\sin(r) \\ -\sin(e) & \cos(e)\sin(r) & \cos(e)\cos(r) \end{pmatrix} \tag{3.2}$$

Where *a* is azimuth, *e* is elevation, and *r* is roll. Joint angles were extracted based on transformed coordinate systems (Detailed computations can be found in Appendix A4).

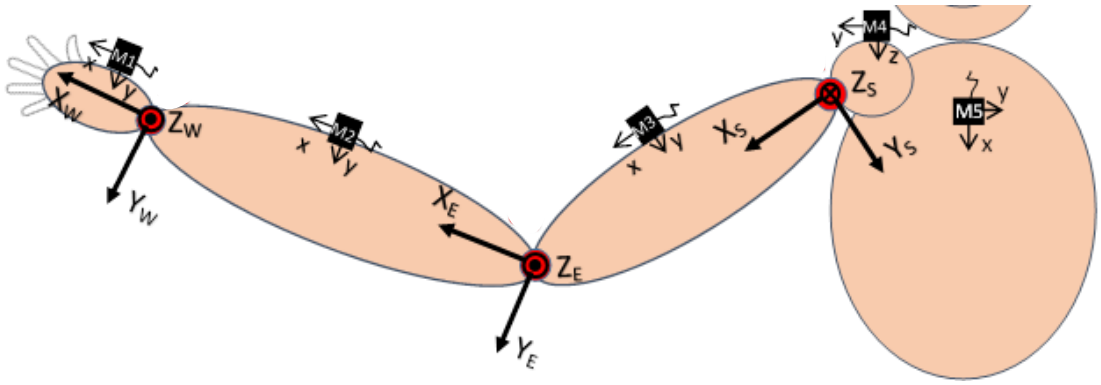

Figure 4 – Joint centers theoretical coordinate systems, and sensor locations

3.4.2 Spatial and temporal scaling

The temporal and spatial dimensions of the trajectory are of very different dimensions. Different scales in the modeling process is equivalent to giving larger weight to the dimension

in the higher scale, which in turn leads to biased modeling (Bishop 2006). We therefore scaled the joint angles to a common range [-1, 1] which is similar to the temporal range:

$$x_{i,new} = 2 \cdot \frac{x_i - \min(x)}{\max(x) - \min(x)} - 1 \quad (3.3)$$

Where x is the trajectory vector and $x_{i,new}$ is the transformed data point.

For creating a model for each subject for each joint, per target, all the joint trajectories (for each subject and target) are required to be of the same length. To achieve this, average task duration was calculated, and a spatial function representing the movement was approximated using General Regression Neural Networks (GRNN) (Specht, 1990). The functions was then used for resampling the data in the new scaled time frame.

3.4.3 Selection of significant degrees of freedom

The main joints expected to perform the reach-to-grasp task are the elbow and shoulder joints. PCA was performed on the trajectories of the control group to identify the main degrees of freedom which contribute to the task and validate the expected motion. The examination showed that the main degrees of freedom were indeed elbow extension (EE), shoulder extension (SE), and shoulder abduction (SA). Moreover, Matlab™ animations of two trials per subject per target were created to validate this result visually. All further analysis will include only these degrees-of-freedom.

3.5 Analysis

The GMMs were computed per subject for each joint (EE, SE, and SA) for each target, i.e., three models per subject for each target, a total of 12 models per subject. Each model included one spatial dimension (angle) and one temporal dimension (time). For each model, parameters were estimated with EM, initialized with K-Means. K different Gaussians were tested per model, for K between 2 to 30. Best matching model was chosen based on minimum estimated BIC.

KLD was computed based on variational approximation (Hershey and Olsen 2007). KLD was computed between each two subject's models per degree of freedom and per target. For the control group, KLD was computed within the group (for each control versus all other controls). KLD score was chosen as the minimum KLD per subject based on the nearest neighbor method. For the stroke group, KLD was computed between groups (for each subject with stroke versus all controls). KLD score was chosen as the minimum KLD per subject based on the nearest neighbor method.

The KLD between the groups was used for assessing movement similarity between subjects with stroke and control subjects. Additionally, the average final angle (FA) of the elbow extension (EE), shoulder abduction (SA), and shoulder extension (SE), average forward trunk movement (displacement of the trunk compared to rest position), and average total movement time (from start to end of reach), were calculated per target for both groups.

3.6 Statistical analysis

Statistical analysis was performed with R Studio IDE for R (version 3.4.2). P-values of <.05 were used for inclusion. Analysis was performed using Generalized Linear Mixed Models (GLMM) with REML criterion for convergence, and significance values were computed with analysis of variance table of type III with Satterthwaite approximation for degrees of freedom. All GLMMs included subjects as random effect intercept. Whenever needed, multiple comparisons were computed with confidence level of 95% adjusted using the Holm-Bonferroni correction.

Three GLMMs for the KLD of both within and between the groups were used to test the differences in the two KLD types. The models included group type (stroke or control), target (near center, far center, far contralateral, far ipsilateral) and their interaction as factors. Five GLMMs for final angle per degree of freedom (EE, SA, SE), trunk forward displacement and movement time, were used to test differences in those variables between the groups and between targets. The models included group type, target, and their interaction as factors.

Four GLMMs were used to test the effect of MAS and FMA on the KLD in the stroke group. One model included Biceps and Triceps MAS, target, and their interaction as factors for explaining the elbow extension KLD. Another model included FMA, target and their interaction as factors for explaining the elbow extension KLD. Both models included movement time and elbow extension final angle as explanatory variables. Since FMA is not related specifically to the elbow joint (unlike Biceps and Triceps MAS), two more models were used to test the effect of FMA on the shoulder abduction and shoulder extension KLD. Both included the corresponding final angle, time and their interaction as covariates. One ordinal logistic regression model was used to test the effects of EE final angle, time, and EE KLD on Biceps MAS.

Finally, three GLMMs to test movement of the stroke group to different targets, included the KLD of the stroke group only per degree of freedom as response, with target as factor. Those models include only the stroke group KLD since it already encapsulates the differences from control subjects.

3.7 Results

For the healthy controls, 9% (0.7% task failure), and for subjects with stroke 28% (15% task failure) of all trials were discarded. Reaches made by subjects with stroke were qualitatively slower, more variable, and less smooth compared with the healthy controls (Figure 6). GLMM for the trunk forward displacement had a significant main effect of group type ($P < 0.001$), significant main effect of target ($P < 0.001$) and significant interaction ($P < 0.001$). Subjects with stroke used more of their trunk in their movement, with an estimate of 5.08 cm more than the control group for near center target, and about 6.35 centimeters more than the control group to each other target (target, $P < 0.001$). There was no significant difference within the control group between targets.

GMMs for subjects with stroke included on average more Gaussian components compared to controls. Moreover, the components had more overlaps in both space and time (Figure 7-8).

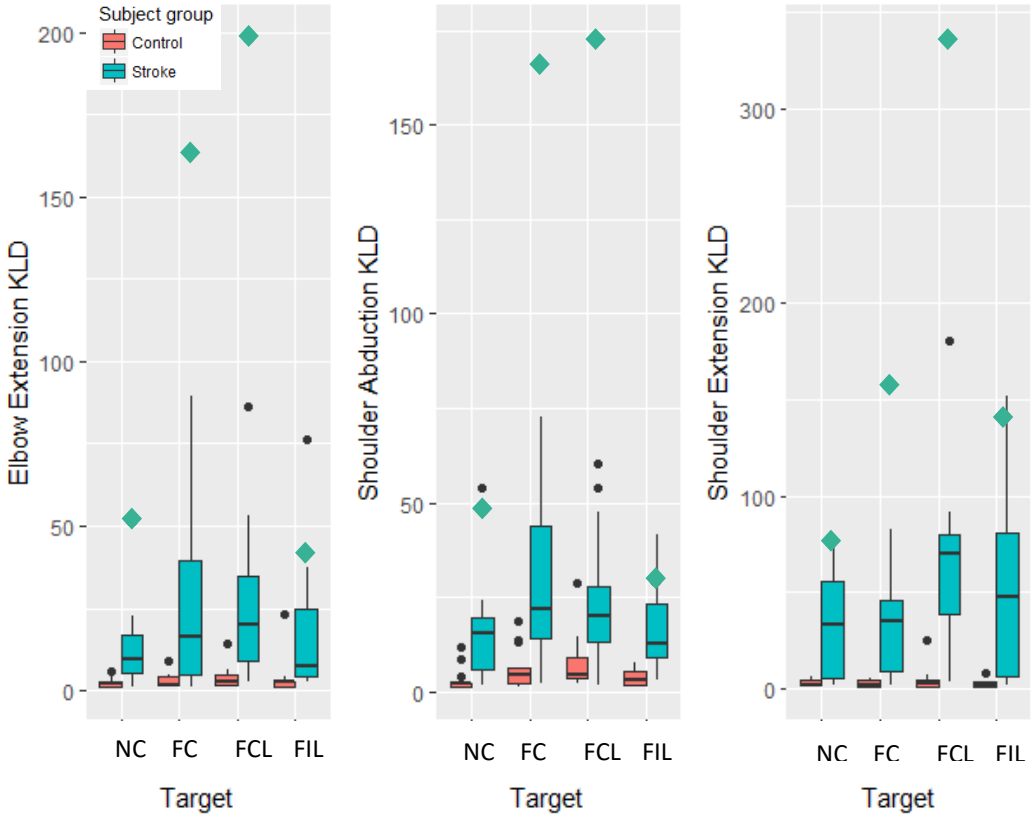

Figure 5 - KLD for EE, SA, SE per target for control subjects and subjects with stroke.

Exceptionally high KLD scores of Subject 5 are highlighted with greenish rhombus. NC- Near center; FC- Far center; FCL- Far contralateral; FIL- Far ipsilateral.

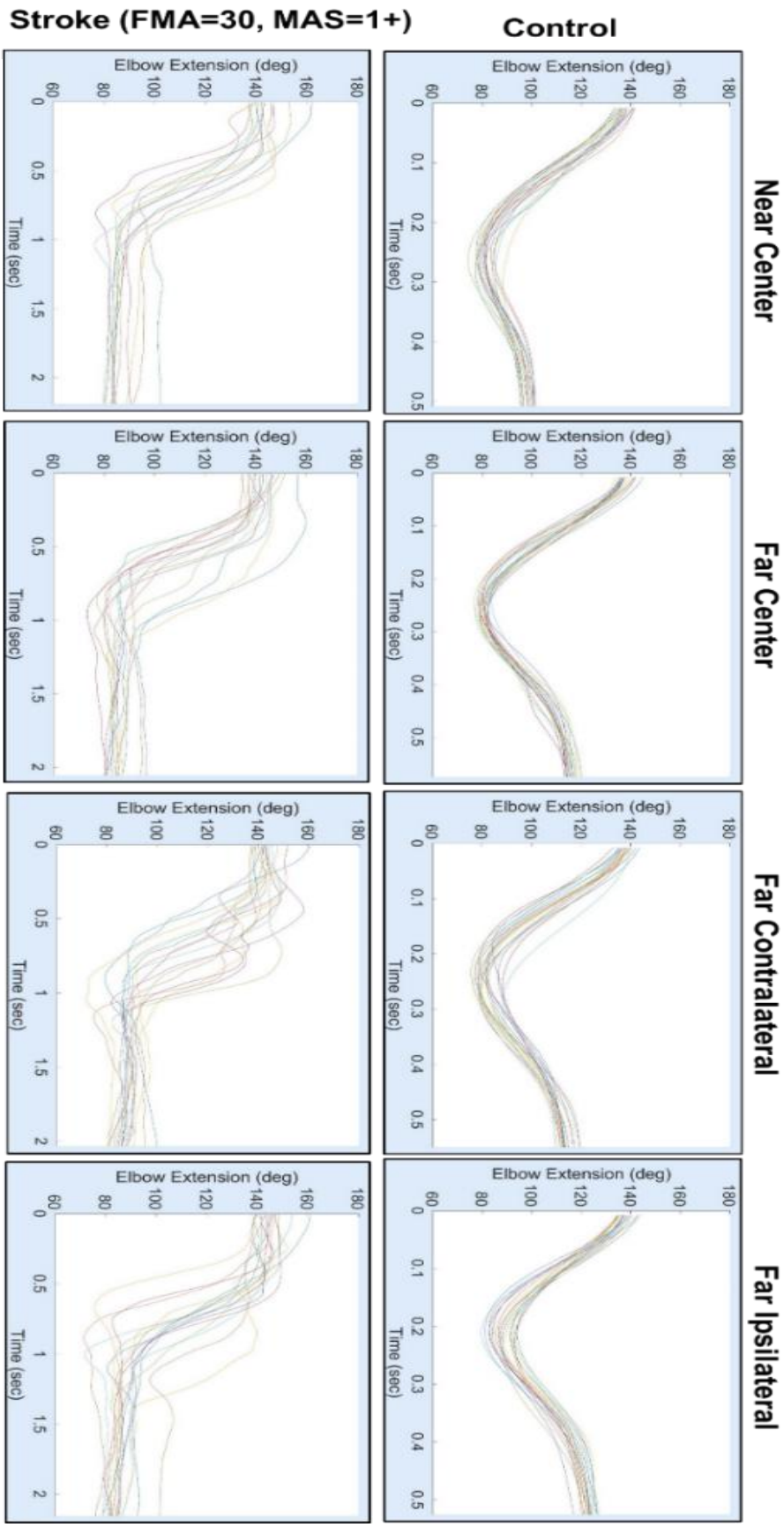

Figure 6 - Elbow extension by time, per target for a control subject (top) and a subject with stroke (bottom)

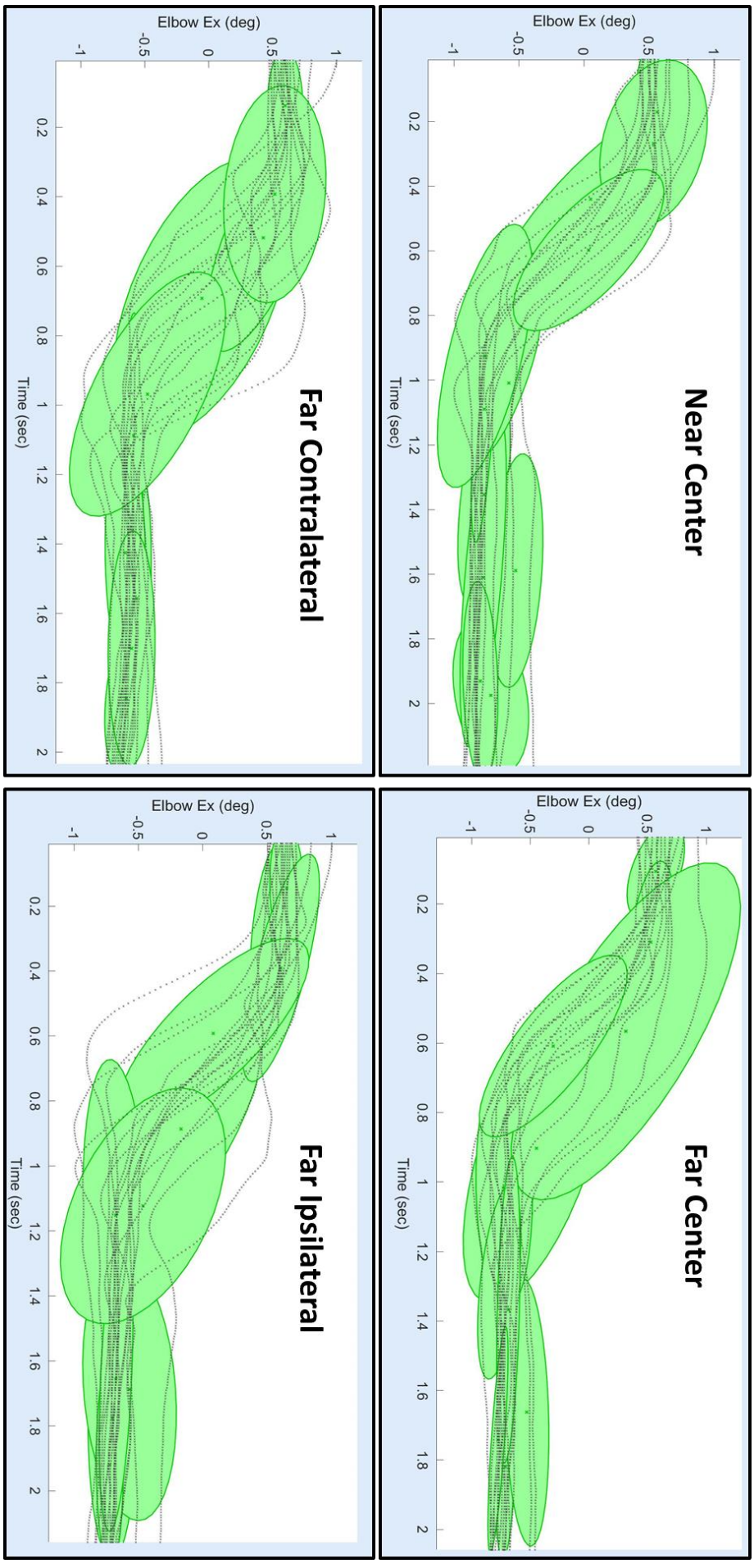

Figure 7 - Elbow extension GMM example for a subject with stroke (scaled angle, adapted duration)

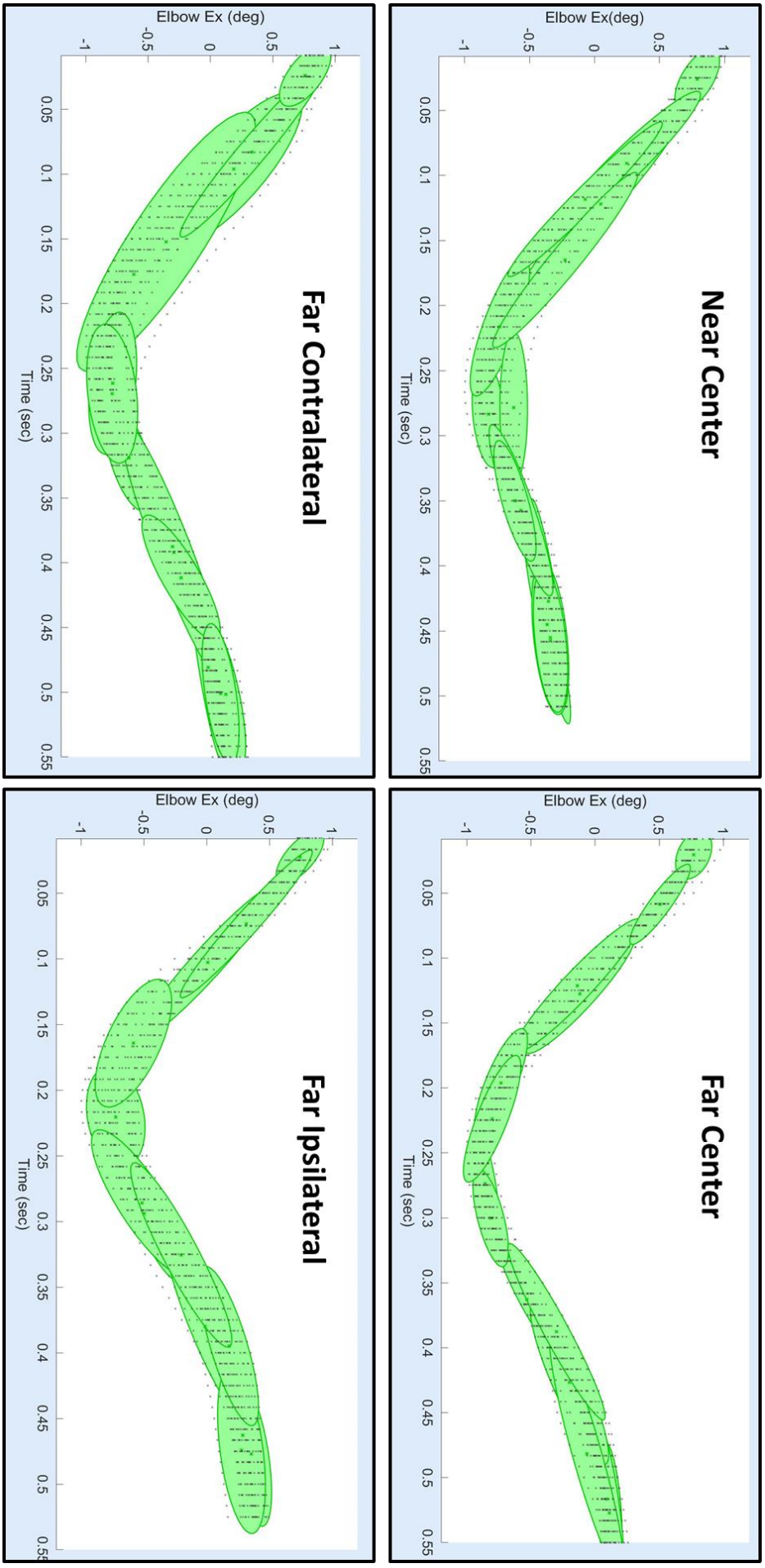

Figure 8 - Elbow extension GMM example for a control subject (scaled angle, adapted time)

All three KLD GLMMs (EE, SA, SE) which included both control group KLD and stroke group KLD as response, and target, group type and their interaction as factors, had a significant effect of group type ($P < 0.001$), where the stroke group had higher KLD values than the control group (Figure 5). For the elbow extension KLD there was an estimated difference of 21.61 ($P < 0.05$). For the shoulder abduction KLD there was an estimated difference of 20.25 ($P < 0.01$). For the shoulder extension KLD there was an estimated difference of 47.22 ($P < 0.001$). There was no main effect of target and there was no interaction.

One subject from the stroke group (Subject 5), had a relatively low FMA score (FMA=30), and exceptionally high spasticity measure (MAS=2). The subject's elbow and shoulder KLD values were also exceptionally high, especially for targets FC and FCL. These results effected the differences in the average KLD scores between the stroke and control groups. However, even when removing Subject 5, there is a significant difference in the KLD values between the stroke and control groups ($P < 0.001$). Stroke group KLD, final angle and movement time for both groups per target can be seen in Table 1.

Table 1 - Mean estimates per target and group

Type	*Measure Mean(SD)*	*Near center*	*Far center*	*Far contralateral*	*Far ipsilateral*
Stroke group	EE KLD***	14.8 (15.8)	**32.5** (42.8)	**33.3** (48.6)	**18.8** (19.2)
	SA KLD***	16.8 (15.1)	**36.0°** (39.9)	30.3 (41.3)	18.1 (11.6)
	SE KLD***	33.1 (27.0)	39.9 (38.9)	**76.7** (81.3)	52.9 (43.6)
	EE FA**	96.6 (15.0)	**106.2** (16.7)	**106.1** (15.8)	**107.8** (16.5)
	SA FA	92.4 (46.2)	89.4 (42.7)	83.5 (41.3)	93.2 (52.2)
	SE FA*	50.2 (43.8)	51.4 (38.5)	53.6 (37.5)	51.5 (39.6)
	Time***	1.96 (0.64)	1.96 (0.63)	1.98 (0.72)	1.84 (0.61)
Control group	EE FA	117.3 (10.3)	**138.0** (12.0)	**110.1** (12.0)	**110.1** (9.8)
	EE KLD	2.53 (1.46)	2.91 (2.26)	3.94 (3.63)	3.72 (6.01)
	SA KLD	3.18 (3.27)	6.23 (5.48)	7.53 (7.31)	3.47 (2.12)
	SE KLD	3.18 (1.56)	2.71 (1.81)	4.71 (6.43)	2.87 (2.40)
	SA FA	57.7 (33.8)	**84.5** (25.7)	**98.7** (30.1)	42.8 (24.8)
	SE FA	10.6 (10.9)	**21.1** (11.7)	**20.0** (12.4)	**16.7** (9.8)
	Time	0.65 (0.14)	**0.68** (0.15)	**0.70** (0.15)	**0.67** (0.16)

*FA- Final Angle; EE-Elbow Extension; SA-Shoulder Abduction; SE- Shoulder Extension; Time – Movement time; In bold are the measures who differed significantly between the targets (at least P<0.05). Bold with underline signifies a difference from the targets already in bold. Mark of ° on the table values means marginal difference (P<=0.1). Significance levels on the measure names are for difference of the stroke group from the control group (main effect), by the codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 ' ° ' 0.1 ' ' 1.*

3.7.1 Stroke KLD validation with MAS and FMA

The GLMM for the elbow extension KLD with Biceps and Triceps MAS, had a significant effect of the Biceps MAS ($P < 0.05$), but no effect of the Triceps MAS. Higher Biceps MAS values correlated to higher elbow extension KLD. There was significant main effect of target ($P < 0.001$) and significant interaction between MAS level and target ($P < 0.001$) (Figure 9 A). There was no effect of FMA on any of the KLD (EE, SA, SE) (Figure 9 B), yet there were significant effects of time ($P < 0.001$) and final angle ($P < 0.05$). The ordinal logistic regression model for the Biceps MAS had only the EE KLD significant ($P < 0.001$), where EE final angle and movement time were not significant.

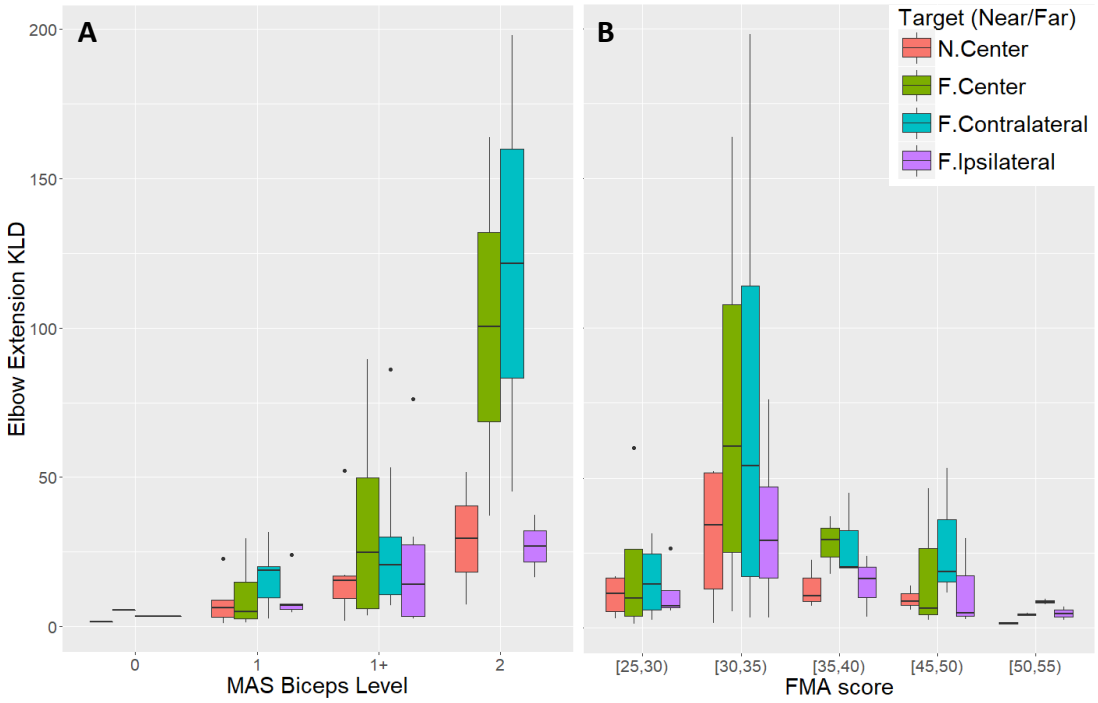

Figure 9 – Biceps MAS (A) and FMA score (B) versus elbow extension KLD

There were only two subjects with a Biceps MAS value of 2 (which are plotted at the boxes tails), and only 1 subject with a Biceps MAS value of 0.

3.7.2 Movement per target

GLMM for the total movement time with target, group type and their interaction as factors, had only significant effect of the group type ($P < 0.001$). Subjects with stroke had higher movement time to all targets, with no differences between the targets within each group (Table 1, Figure 10 A). GLMM for the elbow extension final angle with target, group type, and their interaction as factors, had a significant main effect of group type ($P < 0.01$), target ($P < 0.001$), as well as a significant interaction ($P < 0.001$). Control subjects had lower final angle than subjects with stroke to all targets (Figure 10 B). Within the stroke group, there was lower

final angle for the NC target than all other targets, with no differences between the far targets. GLMM for the shoulder abduction final angle had no effect of group type. However, there were significant interaction and main effect of target ($P < 0.001$ for both), which forms a crossover interaction (Figure 10 C), due to differences between the far targets (FC, FCL, and FIL). GLMM for the shoulder extension final angle with target, group type and their interaction as factors, had a significant main effect of group type ($P < 0.001$), target ($P < 0.05$), as well as a significant interaction ($P < 0.05$). Subjects with stroke reached higher average final angle to all targets, with no difference between the targets. The effect of target and the interaction can be attributed to the differences within the control group (Figure 10 D). Differences are summarized in Table 1. For detailed results and multiple comparisons on the center targets (NC, FC) see Appendix A5, and for the far targets (FC, FCL, FIL) see Appendix A6.

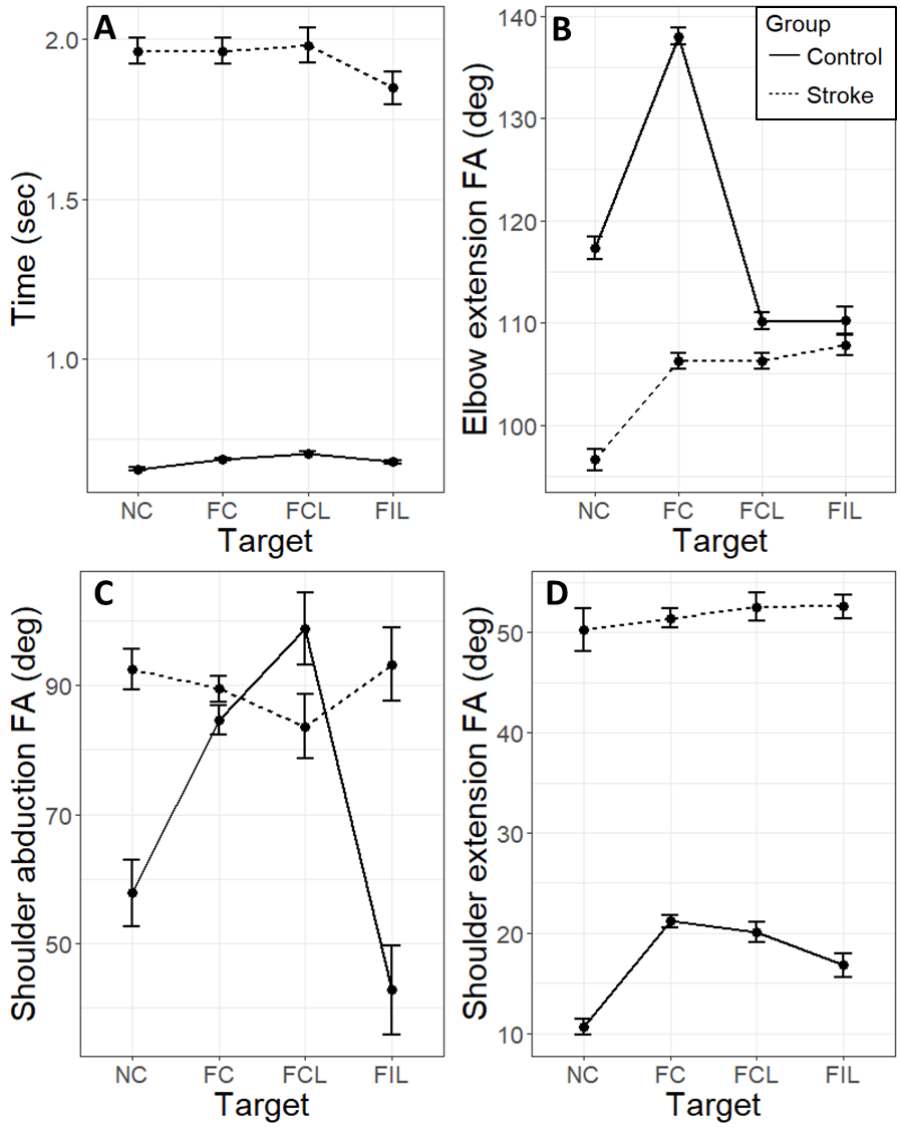

Figure 10 - Mean estimates for the final angle and time per target for each group

NC- Near center; FC- Far center; FCL- Far contralateral; FIL- Far ipsilateral

Plot A - Significant main effect of the group type ($P < 0.001$).

Plot B - Significant main effect of group type ($P < 0.01$), target ($P < 0.001$), and significant interaction ($P < 0.001$).

Plot C - Significant main effect of target ($P < 0.001$) and significant interaction ($P < 0.001$).

Plot D - Significant main effect of group type ($P < 0.001$), target ($P < 0.05$), and significant interaction ($P < 0.05$).

3.8 Discussion

The KLD values were much larger between the subjects with stroke and the control group than within in the control group. These large differences suggest that subjects in the control group use similar motions while that the subjects with stroke use motions that are very different from those used by the control group members. This result validates the use of the spatio-temporal KLD for examining the motion of patients with stroke. KLD values of patients with stroke for all joints were related to movement time and to final angle with significant interaction. This also validates the suitability of using a spatio-temporal model.

Movement time differed between groups for all targets, and similarly the final angles (EE, SA, SE) also differed between groups. The lack of effect of target on movement time is in line with the well-known concept of isochrony, i.e., the subjects moved within a similar time frame to all targets (Viviani and Flash, 1995). While movement time was not affected by target, the final angles differed between the targets and there was a significant interaction between target and group. When analyzing the control and stroke groups separately there were much more differences in final angles between targets in the control than the stroke group, e.g., while for the stroke group the elbow extension final angle differed only between the center targets (NC and FC), in the control group there were a differences between the NC, FC and FCL targets.

Elbow extension KLD was related to Biceps MAS. This validates the expected effect of spasticity on spatio-temporal scaled motion. The lack of relationship between Elbow extension KLD and Triceps MAS is also expected since the triceps is less relevant for performing the reach-to-grasp task. The KLD of all tested joints (EE, SA, SE) was not related to FMA. Unlike the MAS, which measures the properties of a specific muscle group, Fugl-Mayer assessment (FMA) measures a performance impairment at a higher level, and therefore is also effected by motion compensation and inter-joint coordination (Levin et al. 2016). This can explain the lack of relationship between the joint level KLD and the FMA. Moreover, FMA most

of the subjects participating in the ENHANCE project had relatively low FMA scores (Severe impairment) limiting the explanatory capability of the FMA measure. The usefulness of KLD as spasticity measure over end point measures highlighted by the chosen model for explaining the Biceps MAS which included only the EE KLD over final angle and time.

The elbow extension final angle for the near target was lower than the final angle used for all three far targets, which were similar to each other. This can explain the finding that the elbow extension KLD differed between the near and all three far targets, but was similar for all far targets. When using a lower final angle the subjects may be still operating within their control zone, and therefore have a motion that is more similar to that of the control group (lower KLD). When moving towards the higher final angle, i.e., within their spastic zone the effects of spasticity increase the difference from the scaled motion of the control group.

Fitting GMMs offers a methodological way for incorporating spatio-temporal motion variability with trajectory representation. The symmetric KLD form can quantify distance from a control model and measure motor deficiency, e.g., elbow extension KLD offers an objective measure of Biceps spasticity as corroborated by its effect with the Biceps Modified Ashworth Scale (MAS). This is the first study to relate kinematic aspects of movement disorders to spasticity, and we can conclude that the KLD successfully reflect the effects of spasticity on patients with stroke kinematic abilities.

4. Robot adaptive learning

4.1 Overview

This chapter details two improvements to enhance the efficiency of RL learning, based on the ASD method (Zhang, Zhang, and Parker 2015). Our method includes two components: we suggest a dynamically growing action space as appose to an abort and re-learn policy, and additionally we suggest an adaptive learning exploration coefficient (epsilon) rather than a fixed coefficient. The suggested dynamic learning space manipulation methodology is detailed in section 4.2. Section 4.3 presents an experiment conducted to test the methodology, section 4.4 presents the experimental results and section 4.5 discuss the improved performance of the method.

4.2 Method

4.2.1 Dynamically growing action space

As described in section 2.4.3.2, the ASD method is based on reducing the size of the learning space based on identifying non principal components using PCA. According to (Sutton and Barto 2012; Zhang, Zhang, and Parker 2015) for RL algorithm to converge to the optimal solution the algorithm must allow all actions a chance to be selected. To ensure optimal solution is achieved, in case a solution is not found in the reduced search space, additional degrees of freedom (DoF) are added to the search space, and the learning is re-initialized. However, although a final solution may not have been reached, maintaining the meaningful learning that was attained during the run can be used to reduce the search time in the following attempts to reach a solution. Accordingly we suggest a dynamically growing action space method, denoted Dynamic_AS, rather than the abort and re-learn approach (ASD).

At first, the action space is constructed from the non-principal components (same as in the ASD method). The accumulated reward from each run is retained for deciding if an increase in the action space is required. If the accumulated reward remains static, i.e., does not change by more than $\pm$C (where C is constant) for several trials, the action space is then expanded by a selector. Each DoF outside of the initial action space J_i is fitted with a selector, $\mu_i = 1 - MoC(J_i)$, where $MoC(J_i)$ represents the magnitude of contribution to the latent space of the demonstrated task of DoF J_i, as described in section 2.4.3.2. The selectors were used to decide the entrance order of DoF to the RL algorithm action space, based on highest μ_i first. This differs from the ASD method, where the action space is scaled up randomly.

4.2.2 Adaptive Epsilon Greedy with MoC

Another improvement, enhancing the dynamically expending action space, is to perform adaptive exploration, by using an adaptive epsilon greedy combined with MoC. This approach allows the probability for a random action to change dynamically, based on the degrees of freedom currently included in the action space. The learning will be adapted such that, as the action space expends, the longer the learning process will take to converge, the probability for a random action will be higher. This method will assist in escaping from local minima. The adaptive epsilon greedy, ε_t, for feature ranking and selection is of the form:

$$\varepsilon_{t+1} = (1-\delta)\varepsilon_t + \delta \cdot f_{t+1}(s,a) \tag{4.1}$$

Where δ is a weight given to the latest information ($0 < \delta < 1$), and $f_{t+1}(s,a)$ represents the latest information to update epsilon:

$$f_{t+1}(s,a) = \frac{1 - e^{-|Q_{t+1}(s,a) - Q_t(s,a)|}}{1 + e^{-|Q_{t+1}(s,a) - Q_t(s,a)|}} * \sum_{J_i \in A} MoC(J_i) \tag{4.2}$$

Where the sum of $MoC(J_i)$ represents the total magnitude of contribution to the latent space of the demonstrated task of degrees of freedom J_i that are currently part of the action space *A*, *s* and *a* are state and action pair, and $Q_t(s,a)$ is the state-action value function as described in section 2.4.2.2. The function $f_{t+1}(s,a)$ reduce the exploration as the process approach convergence (due to the fraction term), yet increase the exploration by a fixed multiplier ($\sum_{J_i \in A} MoC(J_i)$ ranges from 0% to 100%) when more degrees of freedom are added to the action space. Since algorithm behavior is changed, a proof of convergence is supplied below.

Lemma: An adaptive epsilon greedy of the form presented above allows Q-Learning algorithms and specifically gradient descent SARSA algorithm to converge.

Proof: In order to achieve convergence in a Q-Learning algorithm defined by the update rule: $Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha[R_{t+1} + \gamma Q(S_{t+1}, A_{t+1}) - Q(S_t, A_t)]$, the learning algorithm should allow infinitely many visits to all state-action pairs. This holds for any $0 < \varepsilon \leq 1$ (Melo and Ribeiro 2007). For any *x* general function $f(x) = \frac{1-e^{-|x|}}{1+e^{-|x|}}$ exists in the bounds $0 < f(x) < 1$**.** Thus $0 < f_{t+1}(s,a) < 1$ and the same is also true for ε_{t+1}. The two following conditions regarding the learning rate α must hold (Melo and Ribeiro 2007): $\sum_t^\infty \alpha_t = \infty$ and $\sum_t^\infty {\alpha_t}^2 < \infty$. When the learning rate holds $0 < \alpha < 1$ and is constant for all state-action pairs and for all time steps, both conditions hold.

4.3 Experiment

4.3.1 Hypotheses

The experiment goal was to compare the suggested Dynamic_AS method to the ASD method suggested by Zhang, Zhang, and Parker (2015), denoted for experimental purposes as Static_AS. We define two research hypotheses:

H1**:** A dynamic action space, composed initially of non-principal components for the gradient descent SARSA(λ) algorithm and expending between algorithm episodes (i.e. trials) based on identification of a non-converging trend (Dynamic_AS) will have a different learning process compared to a static action space where the algorithm is stopped when not converging and restarts (Static_AS).

H2**:** Learning goal which involves learning within the principal components will require more learning resources (i.e. algorithm iterations) over a learning goal which does not. The principal components define the basics of the task, and when an adaptation is required within the principal components space the task change in a more dramatic manner than an adaptation in an orthogonal space.

4.3.2 Environment

The experiment was conducted on a Melfa RV-2F-1D Mitsubishi industrial robot with 6 degrees of freedom. The robot is connected to a compatible Mitsubishi CR750 controller. The controller is connected to a Personal Computer (PC) with Windows operating system. Robot control is done by issuing position commands and movement speed or acceleration. Position commands are issued in joint angles. The PC issues commands to the controller through a middleware – MelfaRXM (Appendix B1).

All algorithms used in the experiment were coded in Matlab™ version 2015b and ran on a laptop PC with Intel core i7-3610QM 2.3GHz CPU, 6 GB of ram and 64 bit Windows operating system. Algorithm outputs were transferred to the PC connected to the robot controller for robot task execution.

4.3.3 Experimental procedure

The task chosen to test the research hypotheses was a simple reach-to-grasp task. Since the robot had no demonstration system for recording human data, there was a requirement for secondary data. The data was acquired from demonstrations taken in an experiment done as part of another research conducted on a Motoman robot by Sagi et al. (2015). Participants were recorded while performing a reach-to-grasp task using two phantom operators

controlling the robot gripper. In order to perform our research, a set of six random demonstrations were chosen from the whole demonstration set (Appendix B2).

Five learning goals were chosen to test the research hypotheses, defined as the experiment goals. The learning goals were to adapt to perform the reach-to-grasp task learned, with different grasp positions. The different grasp points chosen were divided to two condition groups- height and side, where both conditions are relative to the original grasp point from the demonstrations. The height condition includes categories *low* and *high*. The side condition includes categories *right*, *middle* and *left*. The five goals (labeled uniquely as A2, B1, B2, C1, C2), as seen in Figure 11, were divided to the conditions, as described in Appendix B3. Each grasp point defined a trajectory the robot needs to adapt to. The original grasp point generated from the demonstrations was labeled as *A1*. The grasp points were chosen with consideration of the task demonstrated. For example, a point classified as high and middle should involve only learning in orthogonal space (since original demonstration point was in category low and middle) and hence by definition involve only learning outside the principal component space (within the non-principal component space).

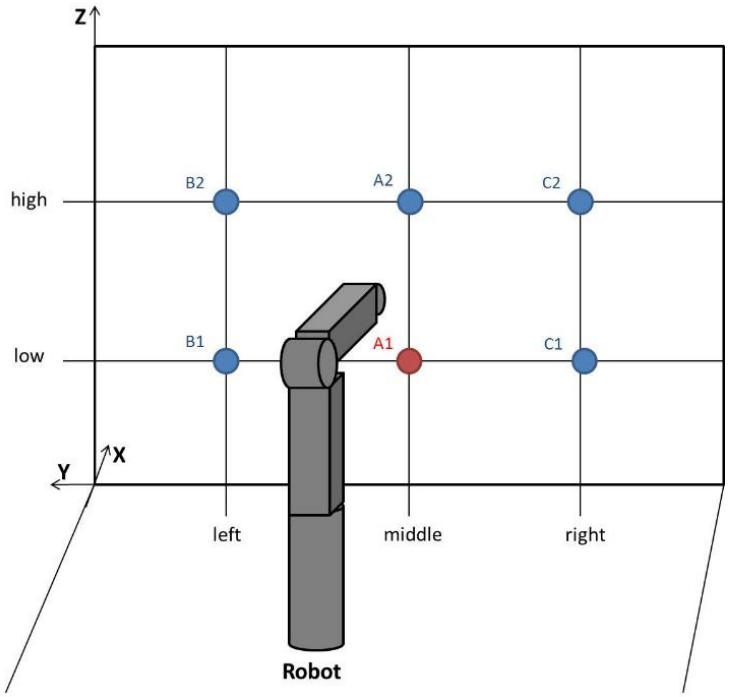

Figure 11 – Learning Goals Sketch

4.3.4 Analysis

The six demonstrations chosen were given as input to the algorithm. The PCA extraction criterion chosen was to extract up to 98% of variance explained. Rotation method chosen was Varimax. Non-principal components were identified by a total MoC value lower than 25%.

In order to fit a GMM model to the principal components in the latent space, EM algorithm was performed on the data, initialized with K-Means algorithm (K tested between 2 and 30). After obtaining a GMM model for the latent space, the model was re-projected to original space and GMR was performed both on all degrees of freedom to obtain a smooth signal for the robot. This process output is a generalized trajectory representing the task defined by the demonstrations. The learning goal and the generalized trajectory are given as input to both of the RL algorithms, Dynamic_AS and Static_AS, as initial policy. Shared parameters chosen for both of the RL algorithms, based on the literature are as follows: step size (learning rate) $\alpha = 0.5$, trace decay parameter $\lambda = 0.9$, and discount rate $\gamma = 0.9$. Reward function chosen was as in the literature $R_t = -\|s^* - s_t\|$.

For the Static_AS algorithm, maximum number of trials before aborting was chosen to be 2000, with epsilon greedy of 0.1, as done by Zhang, Zhang, and Parker 2015. For the Dynamic_AS algorithm suggested above, adaptive epsilon greedy was used with $\delta = 0.3$, and C=150 (number of trials before adding more DoF to the action space), based on highest $\mu_i = 1 - MoC(J_i)$.

The learning part of the procedure (i.e. RL algorithms) was executed for each learning goal (Appendix B4) 10 times per algorithm (Dynamic_AS and Static_AS), a total of 100 learning procedures. The output data of the RL algorithms differs from run to run due to exploration part of the algorithms. The success of each algorithm is validated by letting the robot perform the learned task. The measure variable chosen to compare between the effectiveness of the RL algorithms is the number of trials until convergence (i.e. algorithm iterations until a solution is found). Number of trials until convergence is the most common measure for comparisons for algorithm efficiency due to the homogeneity of this measure over different algorithms, and it is independent of the computer the algorithms are running on. Analysis overview can be seen in Figure 12.

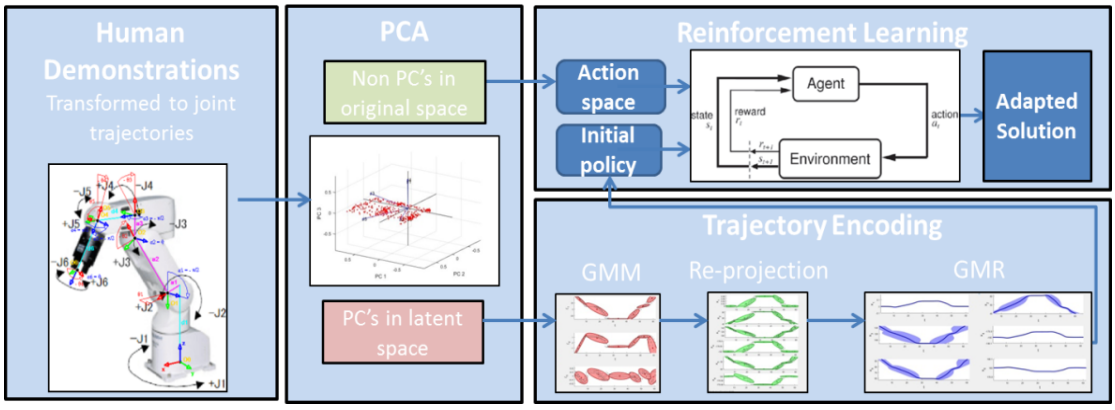

Figure 12 – Robot adaptive learning method overview

4.3.5 Statistical Analysis

In order to test the hypotheses, the dependent variable is measured for each of the five grasp point learning goals for each of the two algorithms. A two-way ANOVA was performed with number of trials until convergence as response variable, with two fixed factors of grasp point (A2, B1, B2, C1, and C2) and algorithm (Dynamic_AS and Static_AS). A full factorial model was tested. By using this test we can determine whether there is any effect of algorithm and grasp point on algorithm trials and if so what is the size of the effect.

In a case of significant interaction, another two-way ANOVA will be performed to determine main effects and perform post hoc tests to determine differences between the targets (using Bonferroni correction). The learning goals causing the interaction will be included in a one-way ANOVA with number of trials until convergence as response and algorithm as factor, to determine significance and effect size between the algorithms. All analysis was made in a 95% confidence level, using IBM SPSS version 22.

4.4 Results

Executing the above method resulted in the identification of 3 Principal components, containing 99.57% of the variance extracted (Figure 13) mainly from joints *J2*, *J3*, *J5* (Figure 14). The non-principal components evaluated were joints *J1*, *J4*, *J6* with total MoC values of 5.3% (Table 2). Full PCA results in Appendix B9.

Table 2 - Joints magnitude of contribution

MoC Values					
J1	J2	J3	J4	J5	J6
5.16%	19.80%	44.81%	0.09%	30.09%	0.04%

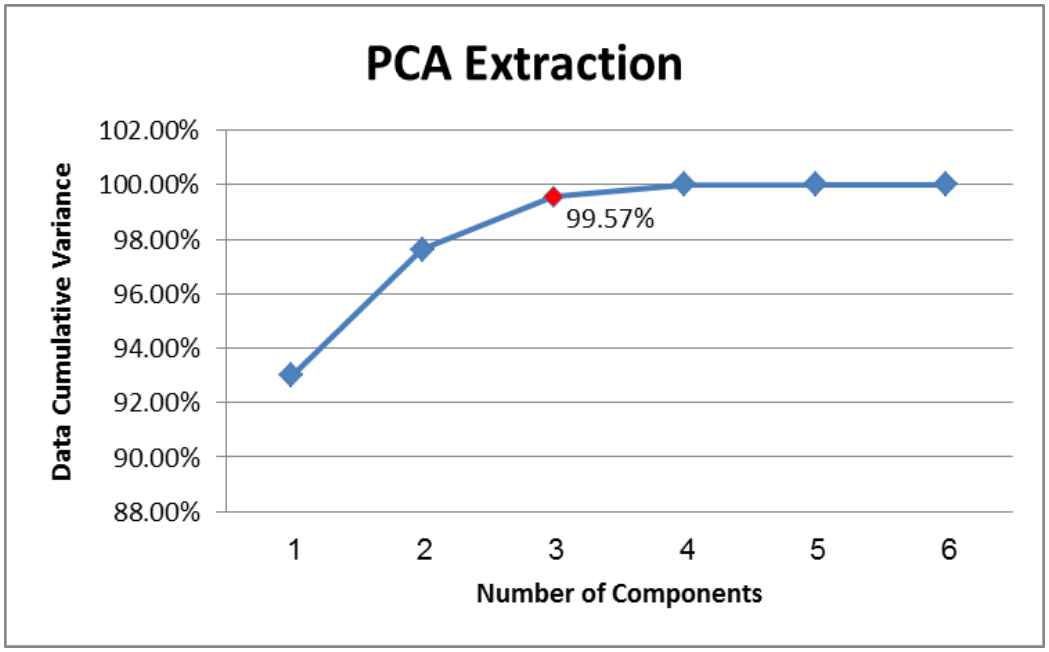

Figure 13 – Variance Extracted by PCA

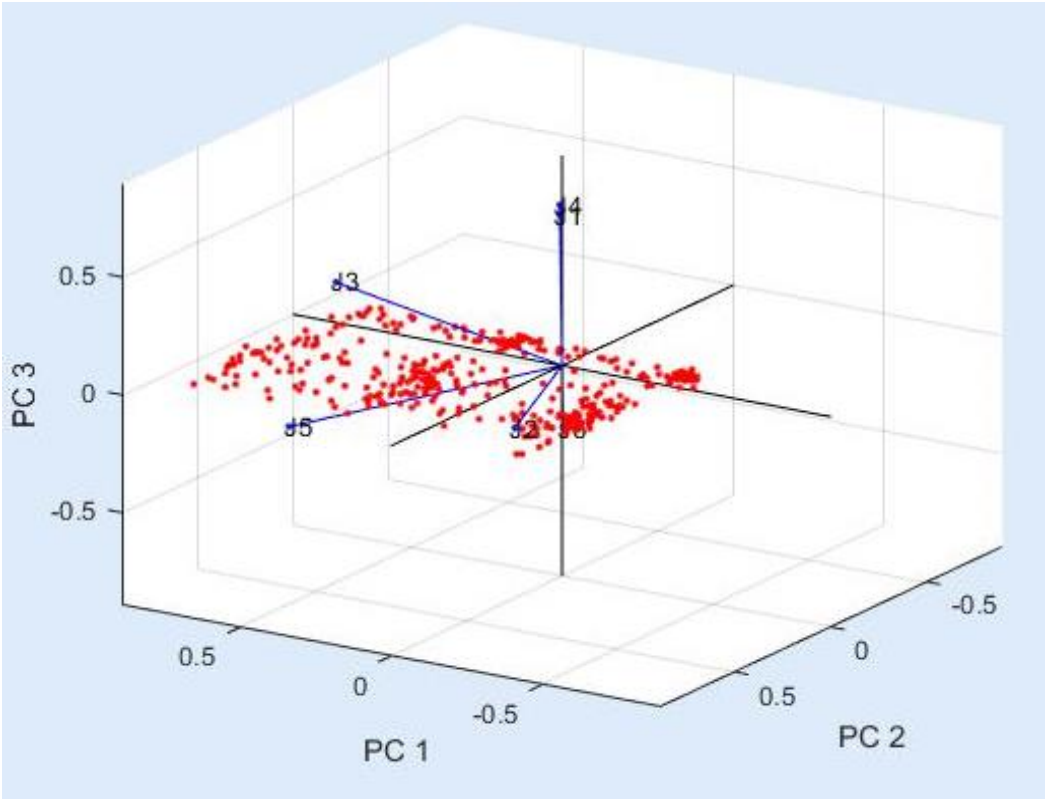

Figure 14 – Data distribution in normalized latent space

The EM algorithm best fit was by fitting 6 Gaussians per feature (Figure 15-16). The GMR trajectories provided smooth signal to the robot control (Figure 17).

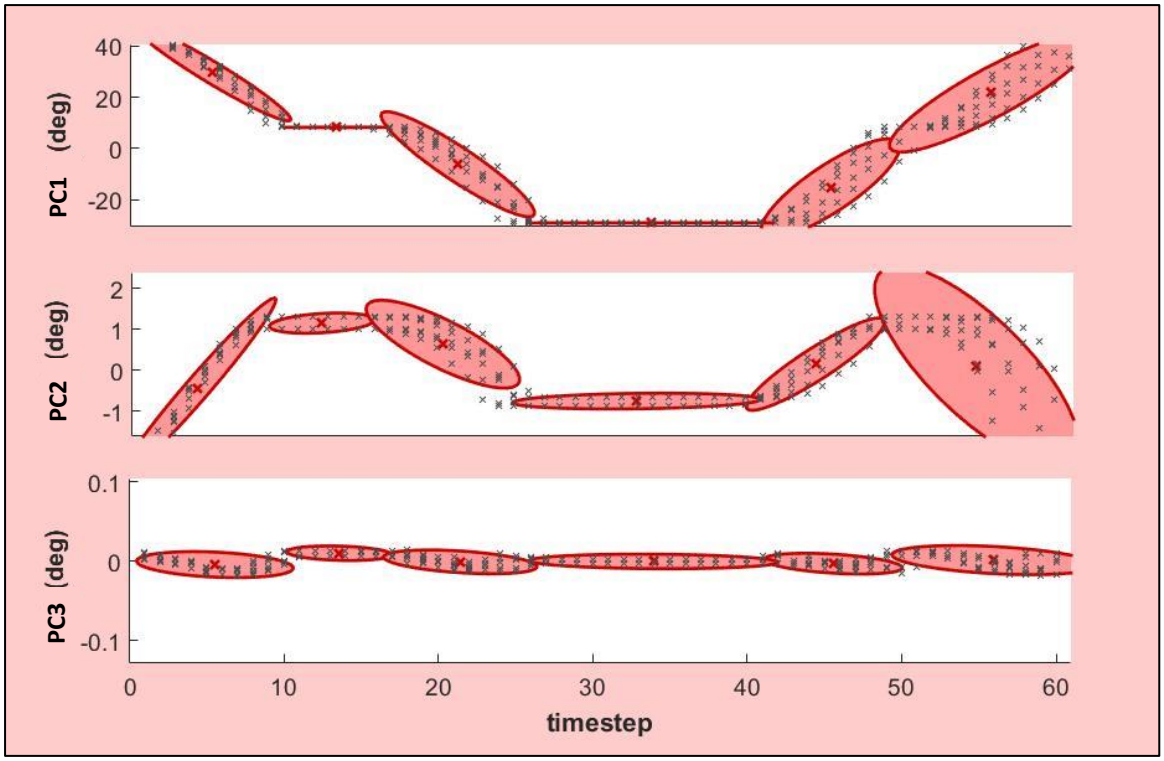

Figure 15 – GMM model on the centered latent space per feature (PC1-PC3)

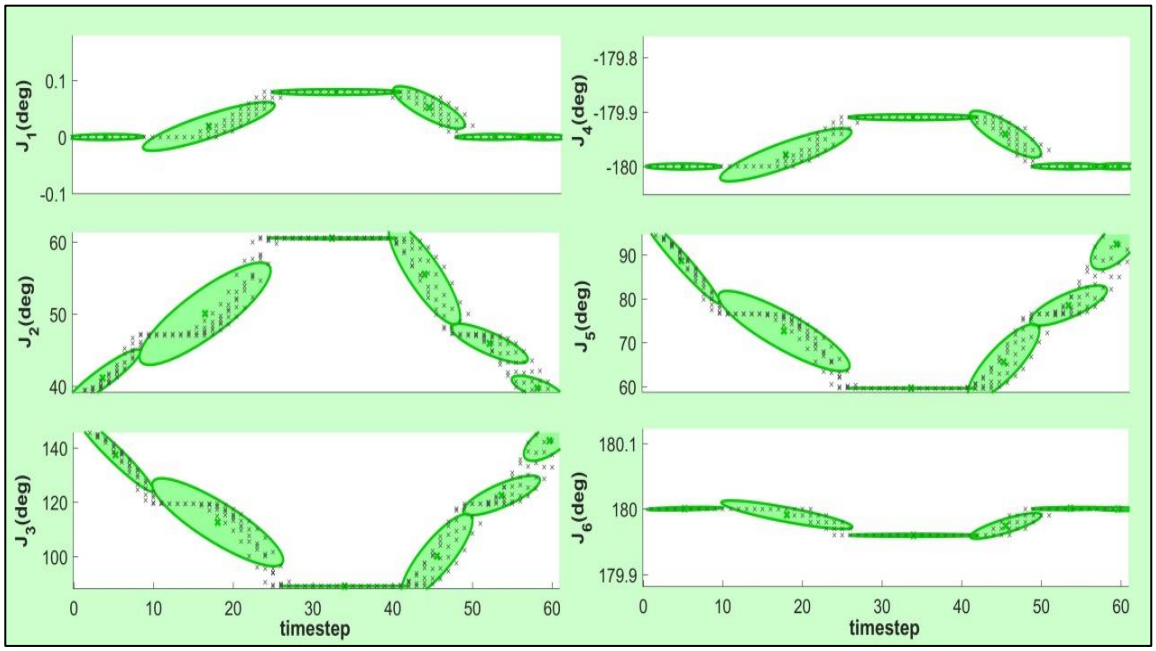

Figure 16 - GMM model reprojected to original space per DoF (J1-J6)

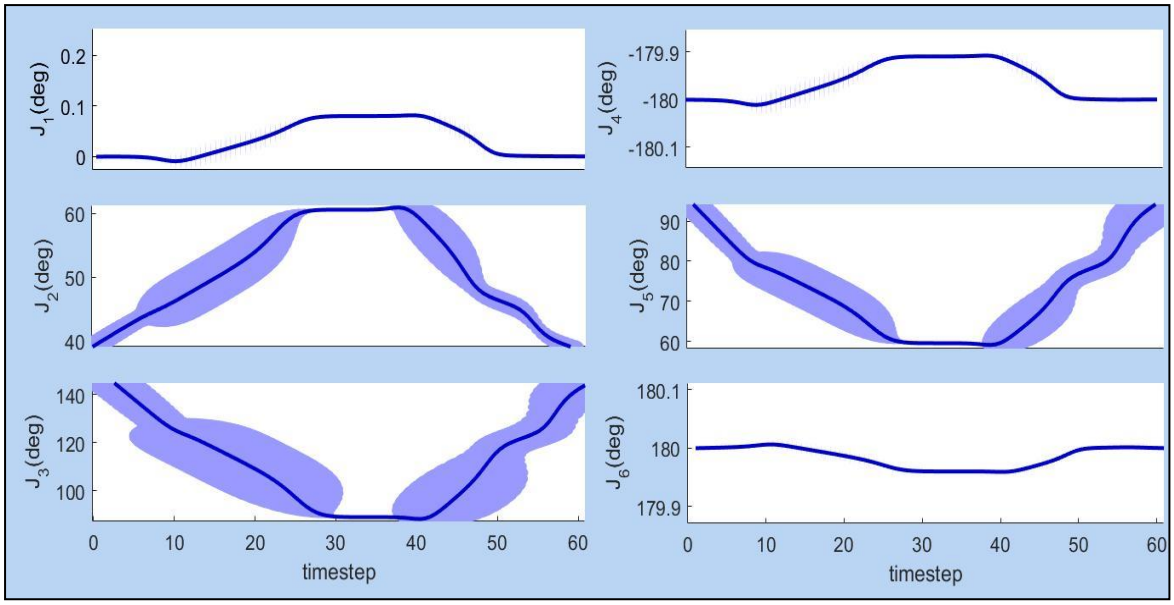

Figure 17 - Generalized trajectory with GMR per DoF (J1-J6)

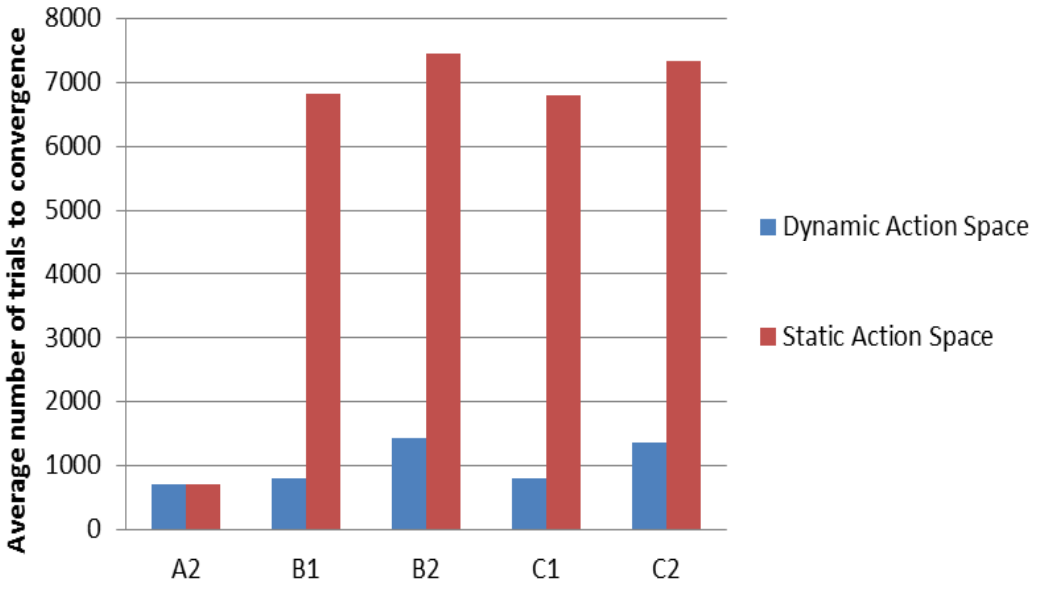

Figure 18 – Dynamic versus static action space (abort and re-learn)

The mean number of trials for Dynamic_AS algorithm was 1021 with a standard deviation of 346 and the mean number of trials for Static_AS algorithm was 5828 with a standard deviation of 2872 (Figure 18). The coefficient of variation for the Static_AS algorithm is considered high with a value of 0.49 due to low runtime only for goal *A2*, a matter which is discussed below. A two-way ANOVA test was performed and test results showed a significant interaction

between grasp point and algorithm (Figure 19) with a statistical estimate $F_{4,90} = 3717.251$ ($p < 0.0005$). Full test results are described in (Appendix B5).

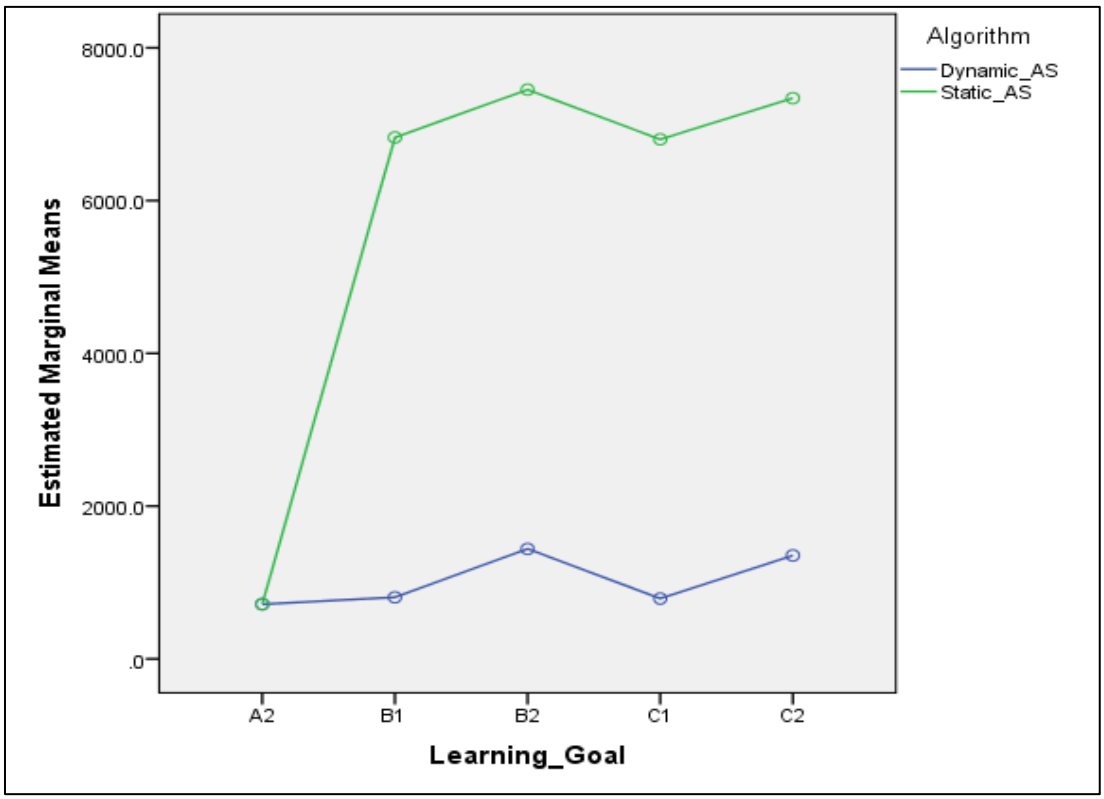

Figure 19 – Profile plot of mean number of trials for all learning goals

In order to determine the main effects of learning goal and algorithm, based on the graphic results, the two-way ANOVA test was re-performed without learning goal A2. All the learning goals besides A2 require learning within the principal component space (Full test results in Appendix B6). Test results for a full factorial model showed no significant interaction (Figure 20) with $F_{3,72} = 0.064$ ($p = 0.979$). The main effect of algorithm was significant with $F_{1,72} = 45094.313$ ($p < 0.0005$). Effect size of algorithm showed a mean difference in the number of trials of the Static_AS algorithm over the Dynamic_AS algorithm with a confidence interval of *[5951.798, 6064.602]*. The main effect of learning goal was significant with $F_{3,72} = 147.55$ ($p < 0.0005$). Post hoc tests were performed using the Bonferroni correction to compare between the effects of the different learning goals on the number of trials. Post hoc showed no significant difference between goals B1 and C1 (confidence interval of *[-87.85, 129.25]*) and no significant difference between goals B2 and C2 (confidence interval of *[-9.45, 207.66]*). There were significant difference between B1 and C2, B1 and B2, C1 and C2, C1 and B2 (Full post hoc results are in Appendix B7). Example of learning curves for one run can be seen in Figure 21.

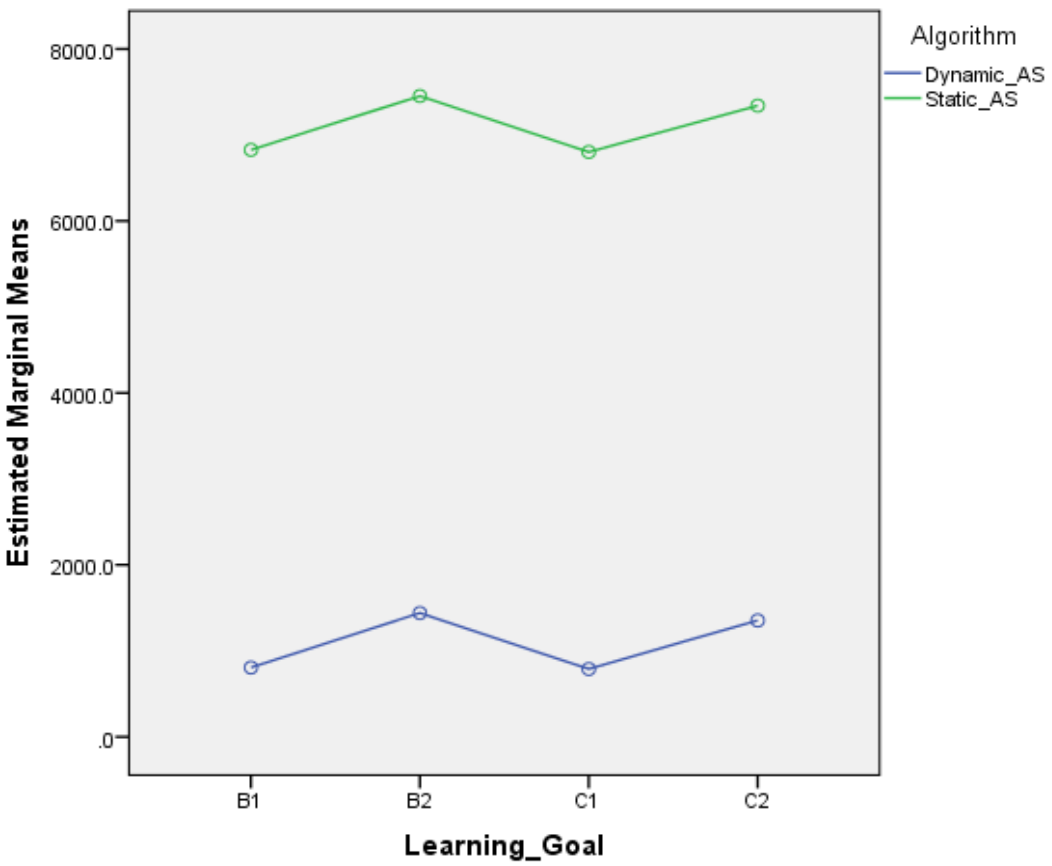

Figure 20 – Profile plot of mean number of trials without goal A2

To complete the analysis, a separate one-way ANOVA test where the target is trials and the factor tested is algorithm was performed for the 20 measurements of learning goal A2 only (10 measurements for each algorithm). The Learning goal A2 requires learning the demonstrated generalized trajectory in a new height. This learning process involves only learning in orthogonal space, hence outside the principal component space. Test results show no significant difference between the algorithms with $F_{1,18} = 0.129$ $(p = 0.724)$ (Appendix B8). The robot was able to perform the task successfully with both algorithms in all trials for all learning goals.

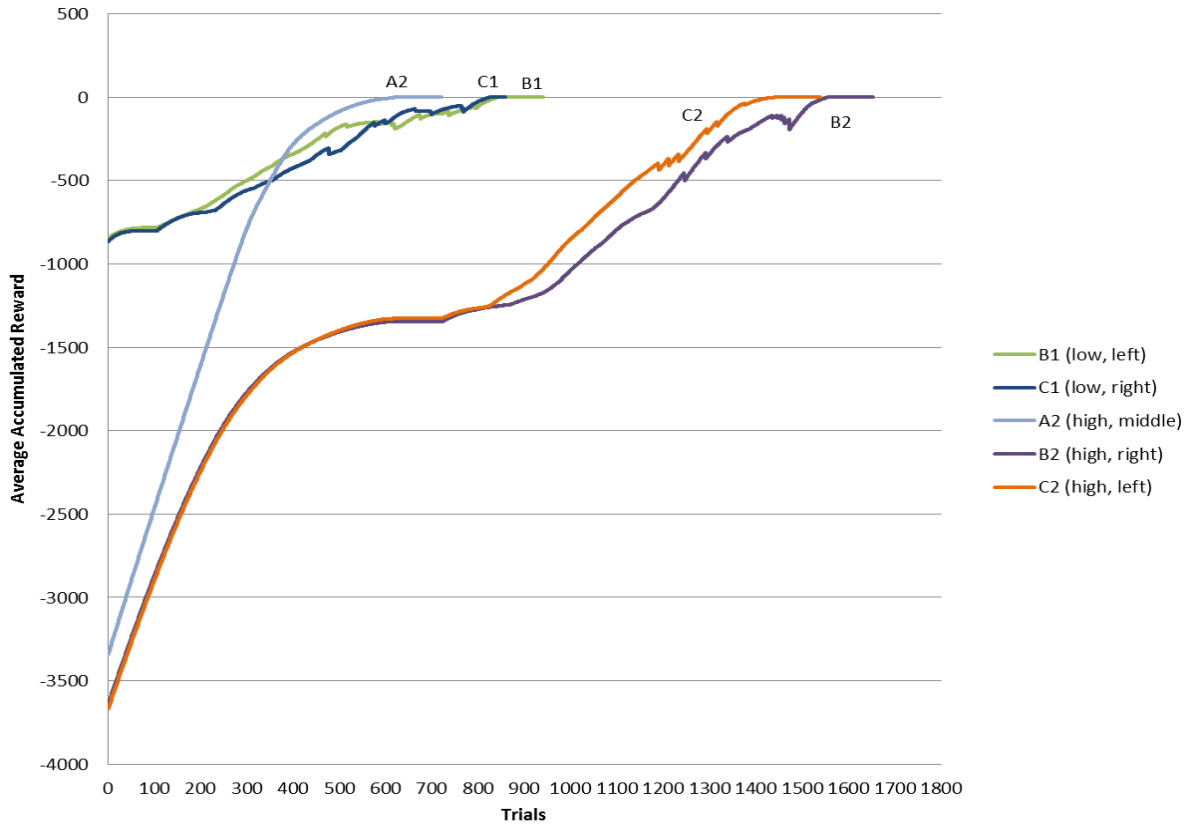

Figure 21 – Example of learning curves after algorithm run per learning goal

4.5 Discussion

The hypothesis that the learning process of the Dynamic_AS algorithm will be different from the Static_AS algorithm (H1) is not rejected. The learning process was different between the algorithms, and took less trials to reach optimal solution to all goals aside from goal A2, which was similar in both algorithms. Overall, a dynamic expending action space during algorithm run is preferred over abort and re-learn static action space. The one-way ANOVA test validated there is no performance difference between the two algorithms for the goal which required learning only within the non-principal component space, since the action space does not require expansion.

The hypothesis that learning towards a goal which required learning inside the principal component space will cost more learning resources (H2) cannot be rejected. The two-way ANOVA test showed differences between learning towards goals in different positions. When the demonstrated generalized trajectory is in position labeled as *low* and *middle*, learning towards a goal in the *middle* and *high* (goal *A2*) required the lowest number of iterations,

where the learning process happens only within the non-principal component space. All other learning processes required more trials than goal A2.

The experiments conducted showed that the physical robot is able to adapt and perform the learned trajectories with success, for both algorithms. We can conclude that the general approach of ST-GMM for combined with PbD and RL for robot learning works well with both RL algorithms tested for reach-to-grasp task on industrial robot, yet the suggested dynamically expending action space is preferred to increase learning efficiency.

References

Abdi, Hervé, and Lynne J Williams. 2010. “Principal Component Analysis.” *Wiley Interdisciplinary Reviews: Computational Statistics* 2(4): 433–59.

Argall, Brenna D., Sonia Chernova, Manuela Veloso, and Brett Browning. 2009. “A Survey of Robot Learning from Demonstration.” *Robotics and Autonomous Systems* 57(5): 469–83.

Billard, Aude, Sylvain Calinon, R. Dillmann, and Stefan Schaal. 2008. “Robot Programming by Demonstration.” *Handbook of robotics*: 1371–94.

Billard, Aude G., Sylvain Calinon, and Florent Guenter. 2006. “Discriminative and Adaptive Imitation in Uni-Manual and Bi-Manual Tasks.” *Robotics and Autonomous Systems* 54(5): 370–84.

Biryukova, E. V., A. Roby-Brami, A. A. Frolov, and M. Mokhtari. 2000. “Kinematics of Human Arm Reconstructed from Spatial Tracking System Recordings.” *Journal of Biomechanics* 33(8): 985–95.

Bishop, Christopher M. 2006. 4 Pattern Recognition *Pattern Recognition and Machine Learning*.

Bohannon, Richard, and Melissa Smith. 2014. “Interrater Reliability of a Modified Ashworth Scale.pdf.” : 415–17.

Burridge, J.H. et al. 2005. “Theoretical and Methodological Considerations in the Measurement of Spasticity.pdf.” : 27: 69-80.

Calinon, Sylvain, and Aude Billard. 2005. “Recognition and Reproduction of Gestures Using a Probabilistic Framework Combining PCA, ICA and HMM.” *Proceedings of the 22nd international conference on Machine learning ICML 05*: 105–12.

Calinon, Sylvain, Florent Guenter, and Aude Billard. 2007. “On Learning, Representing, and Generalizing a Task in a Humanoid Robot.” *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 37(2): 286–98.

Calota, Andra, and Mindy F. Levin. 2009. “Tonic Stretch Reflex Threshold as a Measure of Spasticity- Implications for Clinical Practice.” : 177–88.

Chen, S, and P Gopalakrishnan. 1998. “Speaker, Environment and Channel Change Detection and Clustering via the Bayesian Information Criterion.” *Proc. DARPA Broadcast News Transcription and Understanding Workshop* 6: 127–32. http://www.aquaphoenix.com/presentation/candidacy/chen98.pdf.

Cohn, David a, Z Ghahramani, and Michael I Jordan. 1996a. “Active Learning with Statistical Models.” *Journal of Artificial Intelligence Research* 4: 129–45.

Cohn, David, Zoubin Ghahramani, and Michael Jordan. 1996b. “Active Learning with Statistical Models.” *Journal of Artiicial Intelligence Research* 4: 129–45.

Diebel, James. 2006. “Representing Attitude: Euler Angles, Unit Quaternions, and Rotation Vectors.” *Matrix* 58: 1–35.

F. Levin, Mindy et al. “Enhancing Brain Plasticity for Sensorimotor Recovery in Spastic Hemiparesis.”

Goldberger, Jacob, and Hagai Aronowitz. 2005. “A Distance Measure between GMMs Based on the Unscented Transform and Its Application to Speaker Recognition.” *Engineering* (x): 1985–88.

Granville, Charles De, Andrew H Fagg, and Joshua Southerland. 2006. “Learning Grasp Affordances Through Human Demonstration.” : 1–6.

Guenter, Florent, Micha Hersch, Sylvain Calinon, and Aude Billard. 2007. “Reinforcement Learning for Imitating Constrained Reaching Movements.” *Advanced Robotics* 21(13): 1521–44.

Hershey, John R, and Peder A Olsen. 2007. “Approximating the Kullback Leibler Divergence Between Gaussian Mixture Models.” : 317–20.

Hinton, Geoffrey et al. 2012. “Deep Neural Networks For Acoustic Modeling in Speech Recognition.” : 29(6): 82-97.

Jackson, J. Edward. 1991. A user’s guide to principal components *A User’s Guide to Principal Components*. Wiley Interscience Publication.

Jensen, Jesper, Daniel Ellis, Mads Christensen, and Soren Jensen. 2007. “Evaluation of Distance Measures Between Gaussian Mixture Models of MFCCS.” : 107–8.

Jobin, Annik, and Mindy F. Levin. 2000. “Regulation of Stretch Reflex Threshold in Elbow Flexors in Children Wit Hcerebral Palsy - a New Measure of Spasticity.pdf.” : 531–40.

Kober, Jens, J. Andrew Bagnell, and Jan Peters. 2013. “Reinforcement Learning in Robotics : A Survey.” *The International Journal of Robotics Research* 32: 1238–74.

Kober, Jens, and Jan Peters. 2011. “Policy Search for Motor Primitives in Robotics.” *Machine Learning* 84(1–2): 171–203.

Latash, M L, and J Greg Anson. 2006. “Synergies in Health and Disease: Relations to Adaptive Changes in Motor Coordination.” *Physical therapy* 86: 1151–60.

Melo, Francisco S, and M Isabel Ribeiro. 2007. “Convergence of Q -Learning with Linear Function Approximation.” : 2671–78.

O’Brien, James, Bobby Boddenheimer, Gabriel Brostow, and Jessica Hodgins. 1999. “Automatic Joint Parameter Estimation from Magnetic Motion Capture Data.”

Pearson, Karl. 1901. “On Lines and Planes of Closest Fit to Systems of Points in Space.” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2(1): 559–72.

Press, William, Saul Teukolsky, William Vetterling, and Brian Flannery. 1987. 29 *Numerical Recipes: The Art of Scientific Computing*. Press Syndicate of the University of Cambridge.

Priyadharshini, J., Akila Devi S., and A. Askerunisa. 2014. “Kullback-Leibler Divergence Measurement for Clustering Based On Probability Distribution Similarity.” : 1944–50.

Rummery, G a, and M Niranjan. 1994. “On-Line Q-Learning Using Connectionist Systems.” (September).

Schaal, Stefan. 2002. “Learning Robot Control.” *Learning* 2: 983–87.

Scholz, John P., and Gregor Schöner. 1999. “The Uncontrolled Manifold Concept: Identifying Control Variables for a Functional Task.” *Experimental Brain Research* 126(3): 289–306.

Singh, S P, and R S Sutton. 1996. “Reinforcement Learning with Replacing Elibibility Traces.” *Machine Learning* 22: 123–58.

Sommerfeld, Disa et al. 2004. “Spasticity After Stroke.pdf.” : 134–40.

Levin, M. F., Selles, R. W., Verheul, M. H., & Meijer, O. G. (2000). Deficits in the coordination of agonist and antagonist muscles in stroke patients: implications for normal motor control. *Brain research*, *853*(2), 352-369.

Specht, Donald. 1990. “General Regression Neural Network 191.” *Network*: 1991–1991.

Strosslin, Thomas, and Wulfram Gerstner. 2003. “Reinforcement Learning in Continuous State and Action Space.” *Artifical Neural Networks - ICANN* 2003(January): 2004–7.

Sutton, Richard S., and Andrew G. Barto. 2012. “Reinforcement Learning.” *Learning* 3(9): 322.

Zhang, Chi, Hao Zhang, and Lynne E Parker. 2015. “Feature Space Decomposition for Effective Robot Adaptation.” *International Conference on Inteligent Robots and Systems*: 441–48.

Sagi N., Afgin O., Pal T., Ganel T., Nisky I., and Berman S. (2015) “Weber's law during remote grasping”, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, October 2, Hamburg, Germany

Viviani, P., & Flash, T. (1995). Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 32.

Levin, M. F., Liebermann, D. G., Parmet, Y., & Berman, S. (2016). Compensatory versus noncompensatory shoulder movements used for reaching in stroke. *Neurorehabilitation and neural repair*, *30*(7), 635-646.

Appendix A1 - ENHANCE Subject details

Subject details are detailed in Table A1.1. Effected side is the effected hand the subjects performed the experiment with.

Table A1.1 - ENHANCE subject details

Demographic Data of all subjects and Clinical Measures of Subjects with Stroke

	Age (Years)	Country	Gender	Effected Side	FMA (Arm 56)	MAS (Biceps)
1	59	IL	F	R	26	1
2	46	IL	M	R	46	1+
3	62	IL	M	R	52	1
4	66	CA	M	L	30	0
5	74	IN	F	R	30	2
6	38	CA	M	L	37	2
7	50	CA	F	L	38	1
8	77	IL	M	R	26	1
9	54	IN	M	R	34	1+
10	71	IN	M	R	46	1+
11	50	IN	M	L	50	1+
12	62	IL	F	R	39	1+
13	59	IN	F	R	30	1+
14	48	IN	F	L	26	1+
15	57	IL	F	R	46	1
16	46	IL	M	R	25	1+
Mean	57.4				36.3	
SD	10.9				9.3	
1	57	IL	M			
2	57	IL	F			
3	50	IL	M			
4	52	IL	M			
5	65	IL	F			
6	72	IL	F			
7	54	IL	M			
8	55	IL	M			
9	50	IL	M			
10	66	IL	F			
11	76	IL	M			
12	62	IL	M			
13	70	CA	M			
Mean	60.4					
SD	8.5					

Appendix A2 - G4 Polhemus Tracking System

G4 Polhemus is a wireless electromagnetic tracking system. The reported root mean square static accuracy of this system is 0.08 inches for position and 0.50 degrees for orientation when used within 1 meter of the source (G4, User Manual). The equipment consists of a RF/USB module, transmitter (the hub), wireless sensors (6 DoF each) and a magnetic field generator as the source (Figure A2.1). Sensors are tracked at 120Hz.

Figure A2.1- G4 Polhemus wireless hub and 3 sensors

Appendix A3 – Joint centers algorithm

A3.1 Method

Computing joint centers of the wrist, elbow, and shoulder is required in order to determine direction vectors for the coordinate systems of the joints, which can help determine joint motion. The centers are specifically crucial for the rotation angles of the elbow (SPe) and the shoulder (ROTs). The algorithm implemented for finding the joint centers is based on the work by O'Brien et al. (1999) and by Press et al. (1987). In this method, the human arm is modeled as an articulated hierarchy of bodies connected by joints.

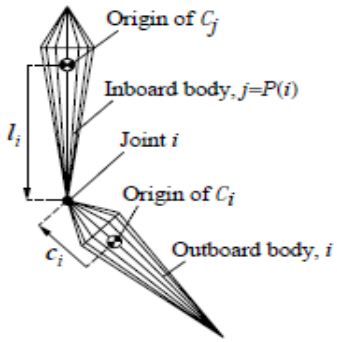

Figure A3.1 - Arm modeled by rigid bodies connected by joints (O'Brien et al. 1999)

A body i can be referenced by a parent body $j = P(i)$. A tracking system on a body defines a coordinate system and origin O_i. A joint center positioned between bodies i and $P(i)$ can be defined by vector displacements originated from the origins c_i and l_i respectively. A point x^i in the i-th coordinate system can be expressed in the j-th coordinate system using the following equation:

$$x^j = R^{i\to j}x^i + d^{i\to j} \qquad \text{(A3.1)}$$

In this expression $R^{i\to j}$ is the rotational component of the transformation, and $d^{i\to j}$ is the translational component. Using the articulated hierarchy model of the arm, allow the following equation to describe the same transformations using displacement vectors at time frame k (out of n discrete time frames of motion):

$$x^{P(i)} = {R_k}^{i\to P(i)}(x^i - c_i) + l_i \qquad \text{(A3.2)}$$

By comparing equations *A3.1* and *A3.2* and changing to a matrix form the following equation can be obtained:

$$Q_k{}^{i\to P(i)}\begin{pmatrix} c_i \\ l_i \end{pmatrix} = d_k{}^{i\to P(i)} \quad \text{(A3.3)}$$

The output of the algorithm is two vectors per joint center per subject: c_i is the vector between joint i center and the sensor below it (origin of c_i) and l_i is the vector between joint i center and the sensor above it (origin of c_j) (Figure A3.1). Using these vectors, each point determined in the sensor coordinate system can be represented in the corresponding joint center coordinate system. In order to isolate c_i and l_i, motion data from all time frames is used in combination with a numerical solution obtained from solving for least squares with singular value decomposition, as described in (O'Brien et al. 1999; Press et al. 1987). After obtaining the vector displacement c_i and l_i for the wrist, elbow, and shoulder joint centers, joint centers at each time step can be computed by adding the arm sensors (M1- MCP, M2- forearm, M3- upper arm) Cartesian positions at each time step with the relative vector displacements. The algorithm was implemented using Matlab ™.

In order to validate the results of the transformation, we recorded calibration movements of a control subject, and then recorded him statically holding his arm in 36 pre-defined positions. Using the calibration files, joint centers displacement vectors were computed. For each of the pre-defined positions, the wrist, elbow and shoulder joint centers were computed twice, once using the vector l_i and once using the vector c_i. Single coordinate errors, defined as X error, Y error, and Z errors were computed based on one dimensional Euclidean distance between the two vectors. Center total error was defined as the total distance between the two center position vectors:

$$error_i = \sqrt{\left[\left(Origin_{c_i} + c_i\right) - \left(Origin_{c_j} + l_i\right)\right] \cdot \left[\left(Origin_{c_i} + c_i\right) - \left(Origin_{c_j} + l_i\right)\right]} \quad \text{(A3.4)}$$

Where c_i is the vector between joint i center and the sensor below it (origin of c_i) and l_i is the vector between joint i center and the sensor above it (origin of c_j), as described above (Figure A3.1). This error was computed for each joint center, for each coordinate, and was averaged over the 36 recorded positions.

A3.2 Results

An example of joint centers estimated Y coordinate position (averaged over 36 recordings for 5 seconds) using the two computation methods (denoted *C* for c_i and *L* for l_i) is presented in Figure A3.2. It can be seen that for the Y coordinate, the average error in the shoulder and wrist positions were lower than the elbow (Table A3.1). When computing the total error for the joint by weighing the X and Z coordinates, the shoulder and wrist average errors were all

below 1 cm, but the elbow center had higher errors with an average total error 2.6 cm (Table A3.1).

Table A3.3 - Joint centers average errors

Measure averages	Wrist center	Elbow center	Shoulder center
X error [cm]	0.17	3.88	0.28
Y error [cm]	0.23	2.94	0.59
Z error [cm]	0.62	0.91	0.27
Total error [cm]	0.32	2.57	0.38

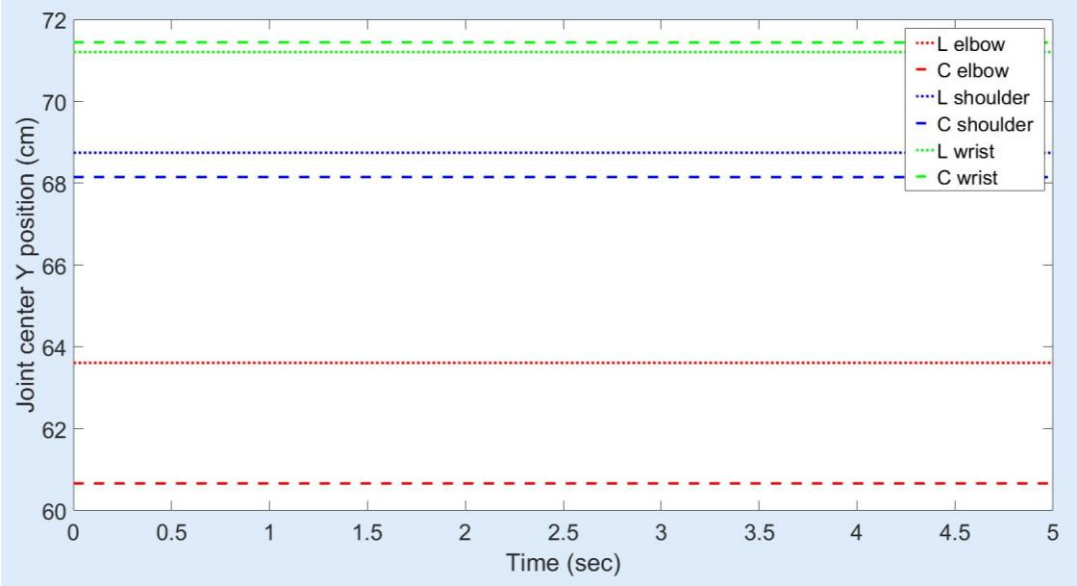

Figure A3.2 - Example of the joint center Y coordinate position recording

A3.3 Discussion

The average total error for the elbow center was higher than the shoulder and the wrist. When performing the elbow calibration movements (SPe and FEe), the distal arm movement around the elbow (SPe) moves the forearm sensor (M2) only slightly compared to the elbow center. Hence, most of the data for the joint center algorithm for the elbow comes from the second movement, the flexion-extension (FEe). Since most of the movement is around one axis, the algorithm output is an axis on which the center lies and not an exact point. This result cause the high error in the center estimation. Due to this error, the joint centers computation algorithm was used only for the shoulder and wrist centers.

The extraction of joint angles using joint centers requires the elbow-wrist and elbow-shoulder vectors (which cannot be computed without the elbow center). Five of the seven angles can be extracted based on the direction of the sensors without computing the joint centers at all, and since the SPe and ROTs angles are not expected to be relevant to the task, no alternatives

were tested for the computation of the elbow center. Wrist and shoulder joint centers were computed once per experiment subject, and results were recorded for the ENHANCE protocol, yet none of the joint centers were used further in this study.

Appendix A4 - Joint angles definition

Joint angles were defined according to Table A4.1. Notice that five out of the seven angles detailed below were computed for the ENHANCE project, yet only the main degrees of freedom (Elbow extension, shoulder abduction, shoulder extension) were used in the work.

Table A4.1 - Joint angles definitions

Name	Explanation	Units
Shoulder Flexion-Extension	Angle between upper arm marker X_{M3} and sternum marker X_{M5} in M5 XZ plane. Fix to 0 when arm pointing down alongside body X_{M3} and X_{M5} orthogonal.	Deg
Shoulder Abduction-Adduction	Angle between upper arm marker X_{M3} and sternum marker $-Z_{M5}$ in M5 ZY plane. Fix to 0 when arm pointing down alongside body X_{M3} and $-Z_{M5}$ orthogonal	Deg
Shoulder External-Internal Rotation	Angle around the shoulder elbow vector X_S. 0 in according to rest position, positive direction external.	Deg
Elbow Flexion – Extension	Angle between upper arm marker X_{M3} and lower arm marker X_{M2}. 180 deg- Full arm extension	Deg
Elbow Supination-Pronation	Angle around wrist elbow vector X_E. 0 is according to rest position. Positive direction external.	Deg
Wrist abduction adduction	Rotation around Z_{M2} (right hand rule), Angle between X_{M1} and X_{M2}. 0 when X_{M1} and X_{M2} are parallel.	Deg
Wrist Flexion-Extension	Angle between Z_{M2} and Z_{M1}	Deg
Plane angle	Angle between normal to arm plane and vertical body direction set by torso marker. 90 When arm plane normal pointing upwards in body direction	Deg

Notice that due to error in the elbow center (for details see Appendix A3), the vectors wrist-elbow and shoulder-elbow could not be computed, and therefore the directions of X_E and X_S as well. Hence, the shoulder external-internal rotation and elbow supination-pronation angles were not computed.

Appendix A5 – Movement to the center targets

Movement to the center targets (near, far) in the stroke group shows differences from the control group in the elbow and the shoulder abduction (significant difference in KLD between the targets). Significance values are described in Table A5.1.

Table A5.1 – KLD and final angle estimates, differences between center targets

Target / **Measure**	**Near center (NC) target**	**Far center (FC) target**
EE FA [deg]	96.616	106.247***
SA FA [deg]	92.491	89.426
SE FA [deg]	50.272	51.455
EE KLD	14.874	32.51**
SA KLD	16.896	36.032°
SE KLD	33.119	39.997

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 ' ° ' 0.1 ' ' 1

FA- Final Angle; EE-Elbow Extension; SA-Shoulder Abduction; SE- Shoulder Extension; Time – Movement time;

Significance values are for the difference NC target from FC target.

Due to the differences in the EE KLD, Table A5.2 presents the results of the difference the EE FA and time between the groups and for each target.

Table A5.2 – Elbow extension FA and time results, subjects with stroke compared to controls

Group / **Measure**	**Controls difference from subjects with stroke**			
	NC Control – NC, Stroke	NC, Control – FC, Stroke	FC, Control – NC, Stroke	FC, Control – FC, Stroke
EE FA [deg]	+20.69***	+11.06*	41.38***	+31.75***
Time [sec]	-1.3***	-1.3***	-1.27***	-1.27***

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 ' ° ' 0.1 ' ' 1

NC – Near center; FC- Far center; EE- Elbow extension; FA- Final angle.

Appendix A6 – Movement to the far targets

Movement to the far targets (FC –far center, FCL – far contralateral, FIL – far ipsilateral) in the stroke group shows significant difference from the control group in the shoulder abduction and the shoulder flexion. No significant difference in the elbow (Table A6.1).

Table A6.1 - KLD estimates, differences between far targets (FC, FCL, FIL)

Target / Measure	FC	FCL	FIL	FCL- FIL
EE KLD	32.51	34.92	17.28	+17.63
SA KLD	36.03	31.70	16.80*	+14.90
SE KLD	40.00	78.31**	50.96	+27.30°

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 ' ° ' 0.1 ' ' 1

FC- Far center; FCL- Far Contralateral; FIL – Far Ipsilateral;

Significance values for targets FCL and FIL are for their difference from target FC. Difference between targets FCL and FIL is in the last column.

The differences in the final angle and movement time with the corresponding significance values can be seen in Table A6.2.

Table A6.2 - Final angle and time estimates, differences between far targets (FC, FCL, FIL)

Group / Measure	Control subjects estimates			Subjects with stroke estimates		
	FC	FCL	FIL	FC	FCL	FIL
EE FA [deg]	138.00	110.13***	110.18***	106.24	106.27	107.84
SA FA [deg]	84.58	98.73	42.81***	89.42	83.58	93.22
SE FA [deg]	21.17	20.07	16.79°	51.45	52.57	52.59
Time [sec]	0.68	0.70	0.67	1.96	1.98	1.84

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 ' ° ' 0.1 ' ' 1

FC- Far center; FCL- Far Contralateral; FIL – Far Ipsilateral;

Significance values for differences between the targets per group are within the group columns, and are for the difference of targets FCL and FIL from target FC.

Appendix B1 - Robot Framework Description

The MelfaRXM middleware includes an API for C#. A robot control framework was built in .NET environment to allow maximum programming and calculations flexibility. The framework developed as an Object Oriented (OO) environment (Table B1.1).

Table B1.1 - Robot control object oriented framework

Object	Overview
Position	An abstract position object
Point	Inherits position, defines position of Cartesian coordinates
Joint	Inherits position, defines position in joint variables
Trajectory	A trajectory built from a set of positions
Cmd	Represent a command sent to the robot - e.g., create position, move to position, etc.
RobotProgram	A program that can execute commands to the robot controller
RoboController	Environment to create programs in and execute from. Can monitor robot state

All the robot programming through the experiment were programmed into *RobotProgram* objects and executed from the *RobotController* object.

Appendix B2 - Data Acquisition

The data that was acquired from the experiment done by Sagi et al. (2015), where demonstrations performed on a Motoman robot using two phantom grippers. In this experiment each subject performed reach-to-grasp 40 times. Valid demonstration data (without trials which failed) of one of the subjects were taken, and a total of 6 demonstration were chosen randomly from that set as a learning data. During a demonstration the Cartesian coordinates of the phantoms were recorded 240 times per second.

The Motoman Cartesian coordinate system was different from the RV-2F Cartesian coordinate system, and hence a transformation was required. A linear transformation was applied, since only linear movements were used in the demonstrations. The following formula was used to obtain the transformation:

$$RV - 2F(x, y) = Motoman(x, y) \cdot 20 \qquad \text{(B2.1)}$$

In order to execute control through joint variables, Inverse Kinematics (IK) model for the RV-2F was needed. The IK was built using the standard Denavit–Hartenberg (DH) parameters convention. The following parameters were calculated:

Table B2.1 - RV-2F DH parameters

index	Joint (deg)	α (deg)	d (mm)	a (mm)
1	J1	-90	295	0
2	J2	0	0	230
3	J3	90	0	50
4	J4	90	270	0
5	J5	-90	0	0
6	J6	0	70	0

Using Matlab™ robotic toolbox the DH parameters were validated (robot model in Figure B2.1-B2.2) and the IK were calculated.

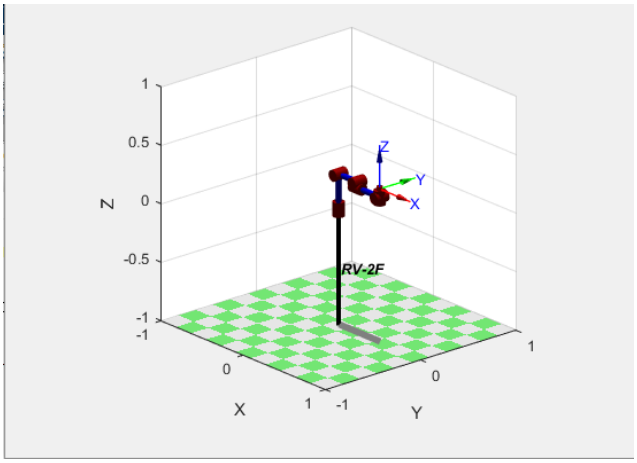

Figure B2.1 – Matlab™ RV-2F model

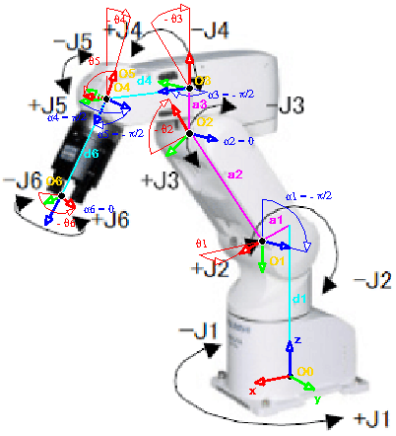

Figure B2.2 – RV-2F Analytical Model Sketch

The joint coordinates obtained from this process were used as a demonstration data for the experiment. Overall of 20 demonstrations were successfully obtained through this process, out of 50 different demonstrations tested. Some demonstrations failed to transform due to different physical constraints which are different in the two robots. Out of the 20 demonstrations, 6 demonstrations that represented the reach-to-grasp task with enough variability were chosen.

Appendix B3 - Learning Goals

The five grasp points chosen as learning goals labels, experiment conditions (relative to the anchor original point from demonstrations) and matching Cartesian coordinates are described in the following table:

Table B3.1 - Robotic experiment Learning Goals definitions

Learning Goal	X	Y	Z	Side	Height
A1 (Anchor)	450	0	200	middle	low
B1	450	65	200	left	low
C1	450	-65	200	right	low
A2	450	0	400	middle	high
B2	450	65	400	left	high
C2	450	-65	400	right	high

Appendix B4 - RL Runtime Results

The runtime results of the RL algorithm per learning goal and per algorithm are described in the following table. Each learning goal with a different experimental category (height and side) was measured 10 times (10 different algorithm runs). The measure itself is the number of algorithm iterations until convergence (Table B4.1).

Table B4.1 - Reinforcement Learning results per goal

Learning Goal	Measure Number	Algorithm 1 (Dynamic action space)	Algorithm 2 (Static action space)
B1 (low, left)	1	783	6875
	2	735	6736
	3	880	6882
	4	793	6794
	5	940	6945
	6	917	6931
	7	674	6674
	8	932	6798
	9	515	6891
	10	891	6740
C1 (low, right)	1	763	6762
	2	846	6685
	3	858	6893
	4	597	6704
	5	697	6927
	6	663	6592
	7	923	6923
	8	781	6663
	9	766	6876
	10	1002	6991
A2 (high, middle)	1	712	713
	2	714	712
	3	715	717
	4	720	711
	5	713	722
	6	718	724
	7	714	710
	8	718	718
	9	716	716
	10	715	718
B2 (high, left)	1	1445	7447
	2	1653	7701
	3	1338	7337
	4	1471	7420
	5	1454	7483
	6	1244	7231
	7	1435	7438
	8	1291	7276
	9	1609	7649
	10	1461	7546
C2 (high, right)	1	1463	7456
	2	1188	7179
	3	1469	7356
	4	1315	7322
	5	1539	7543
	6	1293	7222
	7	1211	7287
	8	1374	7371
	9	1452	7440
	10	1232	7235

Appendix B5 - Two-Way ANOVA results- all learning goals

The two-way ANOVA test was conducted to test the effects of algorithm and learning goal over number of trials. The descriptive statistics and test results in Tables B5.1-B5.2.

Table B5.1 – descriptive statistics per algorithm and learning goal

Descriptive Statistics

Dependent Variable: Trials

Algorithm	Learning_Goal	Mean	Std. Deviation	N
Dynamic_AS	A2	715.500	2.5055	10
	B1	806.000	136.0956	10
	B2	1440.100	127.6275	10
	C1	789.600	121.9309	10
	C2	1353.600	123.5873	10
	Total	1020.960	331.3026	50
Static_AS	A2	716.100	4.6536	10
	B1	6826.600	91.5013	10
	B2	7452.800	150.8817	10
	C1	6801.600	136.6489	10
	C2	7341.100	115.4089	10
	Total	5827.640	2597.6010	50
Total	A2	715.800	3.6505	20
	B1	3816.300	3090.5644	20
	B2	4446.450	3087.4474	20
	C1	3795.600	3086.6656	20
	C2	4347.350	3073.7267	20
	Total	3424.300	3037.8283	100

Table B5.2 – Two-Way ANOVA all learning goals and algorithm

Tests of Between-Subjects Effects

Dependent Variable: Trials

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	912458669[a]	9	101384296.5	7913.872	.000	.999
Intercept	1172583049	1	1172583049	91529.678	.000	.999
Algorithm	577604315.6	1	577604315.6	45086.732	.000	.998
Learning_Goal	190486237.3	4	47621559.32	3717.251	.000	.994
Algorithm * Learning_Goal	144368115.7	4	36092028.93	2817.277	.000	.992
Error	1152986.400	90	12810.960			
Total	2086194704	100				
Corrected Total	913611655.0	99				

a. R Squared = .999 (Adjusted R Squared = .999)

There was a significant interaction between algorithm and learning goal, and the main effects of algorithm and learning goal were both significant as well. Due to the significant interaction, main effects of algorithm and learning goal were analyzed separately.

Appendix B6 - Two-Way ANOVA results without A2

The two-way ANOVA was conducted to test the effect of algorithm and learning goal on number of trials when learning goal A2 is not included, due to the significant interaction prior ANOVA test showed (Table B6.1-B6.3).

Table B6.1 – descriptive statistics per algorithm and learning goal

Descriptive Statistics

Dependent Variable: Trials

Algorithm	Learning_Goal	Mean	Std. Deviation	N
Dynamic_AS	B1	806.000	136.0956	10
	B2	1440.100	127.6275	10
	C1	789.600	121.9309	10
	C2	1353.600	123.5873	10
	Total	1097.325	328.6308	40
Static_AS	B1	6826.600	91.5013	10
	B2	7452.800	150.8817	10
	C1	6801.600	136.6489	10
	C2	7341.100	115.4089	10
	Total	7105.525	321.4877	40
Total	B1	3816.300	3090.5644	20
	B2	4446.450	3087.4474	20
	C1	3795.600	3086.6656	20
	C2	4347.350	3073.7267	20
	Total	4101.425	3040.2617	80

Table B6.2 – Two-Way ANOVA all goals without A2, and algorithm

Tests of Between-Subjects Effects

Dependent Variable: Trials

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	729059361[a]	7	104151337.2	6505.308	.000	.998
Intercept	1345734962	1	1345734962	84054.806	.000	.999
Algorithm	721969344.8	1	721969344.8	45094.313	.000	.998
Learning_Goal	7086931.050	3	2362310.350	147.550	.000	.860
Algorithm * Learning_Goal	3084.700	3	1028.233	.064	.979	.003
Error	1152735.000	72	16010.208			
Total	2075947058	80				
Corrected Total	730212095.6	79				

a. R Squared = .998 (Adjusted R Squared = .998)

The main effects of algorithm and learning goal are both significant, with no interaction.

Table B6.3– Pairwise comparison of the algorithms

Pairwise Comparisons

Dependent Variable: Trials

(I) Algorithm	(J) Algorithm	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
Dynamic_AS	Static_AS	-6008.200*	28.293	.000	-6064.602	-5951.798
Static_AS	Dynamic_AS	6008.200*	28.293	.000	5951.798	6064.602

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Appendix B7 - Post Hoc tests- learning goals comparison

Post hoc with the Bonferroni correction were used to determine difference between the learning goals (without A2) (Table B7.1).

Table B7.1– Pairwise comparison of the learning goals

Multiple Comparisons

Dependent Variable: Trials

Bonferroni

(I) Learning_Goal	(J) Learning_Goal	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
B1	B2	-630.150*	40.0128	.000	-738.709	-521.591
	C1	20.700	40.0128	1.000	-87.859	129.259
	C2	-531.050*	40.0128	.000	-639.609	-422.491
B2	B1	630.150*	40.0128	.000	521.591	738.709
	C1	650.850*	40.0128	.000	542.291	759.409
	C2	99.100	40.0128	.094	-9.459	207.659
C1	B1	-20.700	40.0128	1.000	-129.259	87.859
	B2	-650.850*	40.0128	.000	-759.409	-542.291
	C2	-551.750*	40.0128	.000	-660.309	-443.191
C2	B1	531.050*	40.0128	.000	422.491	639.609
	B2	-99.100	40.0128	.094	-207.659	9.459
	C1	551.750*	40.0128	.000	443.191	660.309

Based on observed means.
The error term is Mean Square(Error) = 16010.208.

*. The mean difference is significant at the .05 level.

It can be seen that the pairs (B1, C1) and (B2, C2) have no significant difference in the number of trials, yet all other pairs have significant differences.

Appendix B8 - One-Way ANOVA results

The one-way ANOVA test was conducted to check whether there is a difference between the algorithms (1-Dynamic_AS, 2-Statis_AS) when the learning process occurs only outside the principal component space (learning goal A2) (Table B8.1-B8.2).

Table B8.1– Descriptive statistics for algorithm performance goal A2 only

Descriptives

Trials

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	10	715.500	2.5055	.7923	713.708	717.292	712.0	720.0
2	10	716.100	4.6536	1.4716	712.771	719.429	710.0	724.0
Total	20	715.800	3.6505	.8163	714.092	717.508	710.0	724.0

Table B8.2– One-Way ANOVA for difference in algorithms for goal A2

ANOVA

Trials

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.800	1	1.800	.129	.724
Within Groups	251.400	18	13.967		
Total	253.200	19			

It can be seen there is no significant difference between the algorithms.

Appendix B9 - PCA Results

The principal component analysis was conducted to identify principal components. The extraction criterion used was extract up to 98% of the variance explained. Rotation method chosen was Varimax (Table B9.1).

Table B9.1– PCA extraction results

Total Variance Explained

Component	Initial Eigenvalues[a]			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	546.69	93.00%	93.00%	546.690	93.00%	93.00%	177.411	0.302	0.302
2	27.18	4.62%	97.62%	27.175	4.62%	97.62%	401.328	0.683	0.984
3	11.45	1.95%	99.57%	1.145E+01	1.95%	99.57%	9.127	.016	1.000
4	2.53	0.43%	100.00%						
5	1.69E-02	0.00%	100.00%						
6	6.79E-06	0.00%	100.00%						

It can be seen that the actual variance extracted was 99.57%. The component matrix shows the effect of each joint. The 3 principal components were joints 2, 3 and 5 (Table B9.2).

Table B9.2– PCA component matrix

Component Matrix[a]

	Raw			Rescaled		
	Component			Component		
	1	2	3	1	2	3
X1	-2.032	.013	.005	-.915	.375	.142
X2	-7.795	.814	.001	-.995	.104	.000
X3	17.638	-.104	.001	1.000	-.005	.000
X4	-.037	.015	.005	-.915	.380	.134
X5	11.844	.708	-.001	.998	.060	.000
X6	.016	-.007	-.002	.903	-.401	-.115

Extraction Method: Principal Component Analysis.

a. 3 components extracted.

The non-principal components were calculated based on their contribution to the principal components, according to MoC formula (Section 2.4.3.2). The limit value chosen was below 25% total contribution. The MoC values:

MoC Values					
J1	J2	J3	J4	J5	J6
5.16%	19.80%	44.81%	0.09%	30.09%	0.04%

It can be seen that with components J1, J4 and J6 we get a total MoC value of 5.3%. Adding component J2, total MoC is 25.1%, and hence the non-principal components are J1, J4 and J6.

תקציר

מידול תנועה סטוכסטי המבוסס נתונים הינו שיטה רבת עצמה עבור ניתוח ויצירת תנועה. מודלים מסוג זה מאפשרים אינטגרציה של זמן, מרחב ושונות לייצוג קוהרנטי. תועלת המודל יכולה להתעצם על ידי אדפטציה של המרחב בו המודל נוצר, למשל ניתן לייצר את המודל במרחב הנתונים המקורי או במרחב חבוי המאותר באמצעות שיטות הפחתת מימד. אנו מיישמים את שיטת המידול עבור ניתוח תנועה של חולי שבץ ועבור שיפור יעילות הלמידה של מערכת רובוטית. לניתוח התנועה בחולי השבץ, אנו מתאימים את המודל במרחב התנועה המקורי לתנועות הושטה לאחיזה של חולי שבץ ונבדקי בקרה בריאים. מדד טיב התאמה המתאים לניתוח זה, גרסה סימטרית של שונות קאלבק-לייבלר, משמש למדידת הדמיון בין חולי שבץ נבדקים בקבוצת הבקרה. הדמיון משמש להערכת ההשפעה של ספסטיות על קינמטיקה של תנועה. הדמיון בתנועת יישור המרפק בין חולי השבץ לנבדקים הבריאים נמצא בעל קשר מובהק למדד סולם אשוורת' מותאם בשריר הזרוע (סולם אשר הינו מדד מקובל לספסטיות). יתרה מכך, הדמיון היה גבוה יותר עבור מטרות שהוצבו קרוב יותר לנבדק, כמצופה אצל חולי שבץ עם אפקט של ספסטיות בשריר הזרוע, כיוון שתחום השליטה בשריר מוגבל. זהו המחקר הראשון לקשר בין אפקטים קינמטיים של תנועה לקויה לספסטיות.

ליצירת יכולת מוטורית ברובוט, פותח מודל במרחב חבוי המאותר באמצעות הפחתת מימד עם ניתוח גורמים עיקריים, ושולב עם רגרסיית תערובת גאוסיאנים ולמידה בשיטה החיזוקים בכדי לאפשר אדפטציה יעילה של תנועה לווריאציות של המשימה, לפי השיטה של Zhang, Zhang and Parker (2015). בכדי להעצים את יעילות השיטה, אנו מציעים להרחיב מרחב החיפוש באופן דינאמי בזמן ריצת האלגוריתם ולהשתמש בהסתברות דינאמית לבחירת פעולה רנדומלית, בנוסף לבחירת הגורמים שיכנסו למרחב החיפוש לפי תרומתם למרחב החבוי. השיטה יושמה בכדי ללמד רובוט מיצובישי מדגם RV-2F-1D משימה של הושטה לאחיזה ולאפשר אדפטציה למטרות אחיזה שונות. הניסויים שבוצעו הראו כי השיפורים המוצעים מאפשרים להפחית את זמן הלמידה לעומת השיטה המקורית.

מילות מפתח: מודל תערובת גאוסיאנים, שונות קאלבק-לייבלר, ספסטיות, סולם אשוורת' מותאם, למידה בשיטת החיזוקים, רובוטיקה, שבץ

אוניברסיטת בן גוריון בנגב

הפקולטה למדעי ההנדסה

המחלקה להנדסת תעשיה וניהול

מידול תנועה סטוכסטי לצורך קידוד יכולת וניתוח

חיבור זה מהווה חלק מהדרישות לקבלת תואר מגיסטר בהנדסה

מאת: ישגב דוידוביץ

מנחה: פרופ' סיגל ברמן

חתימת המחבר:..... תאריך:.....

אישור המנחה:..... תאריך:.....

אישור יו"ר ועדת תואר שני מחלקתית:..... תאריך:.....

אלול תשע"ז ספטמבר 2017

אוניברסיטת בן גוריון בנגב

הפקולטה למדעי ההנדסה

המחלקה להנדסת תעשיה וניהול

מידול תנועה סטוכסטי לצורך קידוד יכולת וניתוח

חיבור זה מהווה חלק מהדרישות לקבלת תואר מגיסטר בהנדסה

מאת: ישגב דוידוביץ

אלול תשע"ז

ספטמבר 2017