

BEN-GURION UNIVERSITY OF THE NEGEV
FACULTY OF ENGINEERING SCIENCES
DEPARTMENT OF INDUSTRIAL ENGINEERING AND MANAGEMENT

Using Hellinger's distance for quantifying the effects of spasticity following stroke on
voluntary motor control

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

By: Hadar Lackritz

SEPTEMBER 2019

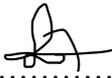
BEN-GURION UNIVERSITY OF THE NEGEV
FACULTY OF ENGINEERING SCIENCES
DEPARTMENT OF INDUSTRIAL ENGINEERING AND MANAGEMENT

Using Hellinger's distance for quantifying the effects of spasticity following stroke on
voluntary motor control

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

By: Hadar Lackritz

Supervised by: Prof. Sigal Berman, Prof. Yisrael Parmet

Author:.....Hadar Lackritz.......... Date: 25.09.2019.....

Supervisor: Sigal Berman ........ Date: 25.09.2019.....

Supervisor: Yisrael Parmet ...... Date: 25.09.2019.....

Chairman of Graduate Studies Committee:..... Date: 25.09.2019.....

SEPTEMBER 2019

Abstract

Spasticity is a common motor deficiency caused by stroke. It is characterized by a velocity-dependent increase in the stretch-reflex. There are existing clinical measures for ascertaining the presence of spasticity and for assessing motor deficits, however, the relationship between voluntary movement disorders and spasticity is not fully understood. This is due, in part, to the complexity and multidimensionality (space and time) of both phenomena, and the variability inherent in motion. In a previous work of our group, an innovative measure, based on spatio-temporal stochastic modeling, was suggested for investigating this relationship. Motion deficits were quantified by the distance between spatio-temporal Gaussian mixture models, constructed from motion trajectories of subjects with stroke and those of healthy controls. This distance was assessed using the bidirectional Kullback-Leibler divergence (BKLD). In the current work, we reinforce and add to the results and conclusions drawn in the previous work, using nearly three times as many subjects. We suggest a different distance measure: the Hellinger's distance (HD) measure and compare it to the BKLD. We show that a larger Hellinger's distance between the models is associated with a higher level of spasticity of the patient with stroke. HD has advantages over BKLD. It is a metric that satisfies the triangle inequality, with values bounded between 0 and 1. HD is symmetric, simple to interpret, is less susceptible to ceiling effects, less computationally intensive, and is shown to be invariant, consistent, asymptotically normal, and robust, compared to the BKLD and to other distribution distance measures.

The analysis in the current work included 13 controls and 42 subjects with stroke which performed reach-to-grasp movements toward 4 targets. Arm motion during reaching was recorded using electromagnetic sensors. Elbow spasticity was quantified using the tonic stretch-reflex threshold (TSRT), the velocity dependency of the spasticity (slope), and the Modified Ashworth Scale (MAS) (a common spasticity clinical measure). Upper limb motor function was quantified using the Fugl-Meyer assessment (FMA). Results suggest that HD is strongly related to the TSRT, the slope, their interaction, and the FMA, and has the best generalizability to the larger stroke population comparing with the BKLD measure and the kinematic measures tested.

BKLD was related to the FMA, marginally related to the slope, and not related to the TSRT measure. Thus, HD can be used as a robust, objective measure of the relationship between spasticity and reaching kinematics. Results also show that the TSRT is statistically superior to the MAS as a clinical spasticity measure.

Keywords: Hellinger's distance, Kullback-Liebler divergence, Stroke rehabilitation, Spasticity, Gaussian mixture models.

Acknowledgements

Foremost, I would like to express my thanks to my advisors Prof. Sigal Berman and Prof. Yisrael Parmet. Sigal, one page will not suffice for all my appreciation and gratitude to you, so I will use this opportunity to summarize briefly and say that I could not ask for more. Thank you for your patience, willingness to give, always pushing forward and teach, always with a smile. Thank you for your availability and for investing your time in my success, our research success, our success. I don't take it for granted. Yisrael, thank you for the calm, for succeeding making me passionate about statistics, always making me laugh. Thanks for sharing your wisdom and experience with me, and providing advice. Thank you for being reachable even overseas, always making me feel comfortable contacting you.

I would like to thank Dr. Silvi Frenkel-Toledo, Melanie C. Baniña, Dr. Nachum Dr. Soroker, Dr. John M. Solomon, Prof. Dario G. Liebermann, and Prof. Mindy F. Levin for giving professional advice and meaningful comments whenever asked.

Finally, I would like to thank my family and friends, for providing encouragements at all times and for always knowing what the right word is.

This research was supported by the Canada-Israel Health Research Program (MFL and DGL), a program that is jointly funded by the Canadian Institutes of Health Research, the Azrieli Foundation, the International Development Research Center (IDRC, grant number 108186-001), and the Israel Science Foundation (ISF, grant number 2392\15), and by the Helmsley Charitable Trust through the BGU Agricultural, Biological and Cognitive Robotics Center (HL and SB).

Table of Contents

1. Introduction	10
1.1 Stroke	10
1.2 ENHANCE project	11
1.3 Motivation and objectives.....	11
1.4 Innovations.....	12
1.5 Work scope.....	13
1.6 Thesis structure.....	14
2. Literature review	15
2.1 Overview	15
2.2 Spasticity	15
2.2.1 Spasticity following stroke	15
2.2.2 Clinical spasticity measures	16
2.3 Using stochastic mixture models for motion modeling.....	18
2.4 Distribution distance measures	20
2.4.1 Distance measures for mixture models.....	20
2.4.2 Hellinger's distance	22
3 .Method.....	25
3.1 Hypotheses	25
3.2 Subjects	25
3.3 Experimental Procedure	26
3.4 Pre-processing	28
3.4.1 Data validation	28
3.4.2 Movement Segmentation.....	29
3.4.3 Spatial and temporal scaling.....	30
3.4.4 Finding the coordinate frames	30
3.4.5 Finding joint angles	32
3.5 Analysis	33
3.6 Statistical analysis	35
4. Results	37
4.1 Hellinger's distance and Bidirectional Kullback-Liebler divergence.....	37
4.2 Clinical measures.....	41

4.2.1 Relationship with TSRT and FMA.....	41
4.2.2 Relationship with MAS and FMA	42
4.2.3 FMA distribution	44
4.2.4 TSRT and MAS	45
4.3 Movement time, final angle, average velocity, and velocity smoothness.....	46
5. Discussion and conclusions.....	48
5.1 HD and BKLD	48
5.2 Functional ability, joint spasticity, and muscle resistance	49
5.3 Kinematic measures	51
5.4 Using stochastic mixture models for modeling human motion	51
5.5 Future work	52
References.....	53
Appendix A - Lebesgue measure	59
Appendix B - Principal component analysis (PCA)	60
Appendix C - Un-controlled manifold (UCM).....	62
Appendix D - Joint Angles definitions	64
תקציר.....	65

List of Figures

Figure 1 : Grades and descriptions of the Modified Ashworth scale	17
Figure 2 : The tonic stretch reflex threshold regression line.....	18
Figure 3 : Experimental setup	27
Figure 4 : Targets locations	28
Figure 5 : Sensors positions.....	28
Figure 6 : Joint diagram.....	31
Figure 7 : Joint angles example.....	38
Figure 8 : GMM example.....	39
Figure 9 : HD and BKLD box-plots.....	40
Figure 10 : Relationship of HD and BKLD with MAS.....	43
Figure 11 : Relationship of HD and BKLD with FMA	44
Figure 12 : Clinical spasticity measures graphs	46
Figure 13 : Kinematic measures - box plots with line over means for (A) final elbow angle; (B) movement time; (C) elbow velocity; (D) the number of acceleration zero-crossings, per target for healthy control (red) and stroke (blue) groups.....	47

List of Tables

Table 1 : Mean (SD) estimates of demographic and clinical data.....	26
Table 2 : Mean (SD) estimates of kinematic characteristics	40
Table 3 : Wald chi-square values (significance levels) for LMM: Measure~FMA+(Slope+TSRT+Target)^2+ (1 ID).....	41
Table 4 : Wald chi-square values (significance levels) for LMM: Measure~FMA+(MAS+Target)^2+ (1 ID)	42
Table 5 : Wald chi-square values (significance levels) for LMM: Measure~(MAS+Target)^2+ (1 ID).....	43
Table 6 : Log likelihood values for different distributions fitted to the FMA values	45
Table 7 : Beta parameters estimates (SD) for the FMA values.....	45
Table 8 : Joint angles definition	64

List of abbreviations

ANOVA	Analysis Of Variance
AS	Ashworth Scale
BIC	Bayesian Information Criteria
BKLD	Bidirectional Kullback-Leibler divergence
CL	Contralateral
CNS	Central Nervous System
EM	Expectation-Maximization
EMD	Earth Mover distance
FA	Final elbow Angle
FC	Far Center
FMA	Fugl-Meyer assessment
GMMs	Gaussian Mixture Models
HD	Hellinger's Distance
IL	Ipsilateral
KLD	Kullback-Leibler Divergence
LL	Log likelihood
LMM	Linear Mixed effect Model
MAS	Modified Ashworth Scale
ML	Maximum-Likelihood
MT	Movement Time
MV	Mean elbow Velocity
NC	Near Center
REML	Restricted Maximum Likelihood
ST	spatial threshold
ST-GMM	Spatio-Temporal Gaussian Mixture Models
TCT	Threshold Control Theory
tDCS	transcranial Direct Current Stimulation
TSRT	Tonic Stretch Reflex Threshold
UL	Upper Limb
VR	Virtual Reality
VS	elbow Velocity Smoothness

1. Introduction

1.1 Stroke

Stroke is currently the leading cause of long-term sensorimotor disability (Zhang et al., 2002). A stroke occurs when a blood clot blocks an artery (a blood vessel that carries blood from the heart to the body) or when a blood vessel (a tube through which the blood moves through the body) breaks, interrupting blood flow to an area of the brain. When either of these happen, brain cells begin to die and brain damage occurs. The two major categories of stroke are ischemic stroke (lack of blood and hence oxygen to an area of the brain) and hemorrhagic stroke (bleeding from a burst or leaking blood vessel in the brain) (Gund et al., 2013).

When brain cells die during a stroke, abilities controlled by that area of the brain are lost. Disrupted functionalities and motor deficits following stroke may include speech, basic movements, confusion, loss of memory, muscle weakness, or paralysis of the face, arm, or leg (usually just on one side - the opposite side of the brain injury side). All these factors contribute to a low overall quality of life.

Motor disorders after stroke are treated by surgery, drugs, or rehabilitation therapy. However, stroke is currently the leading cause of long-term sensorimotor disability (Zhang et al., 2002). Motor deficits induced by stroke persist into the chronic stage in a large proportion of survivors (Langhorne et al., 2009). One of the most common motor disorders resulting from stroke is spasticity (Sommerfeld et al., 2004) (details regarding spasticity in 2.2 Spasticity). During the first year following the stroke, 20-50% of the patients suffer from spasticity. Spasticity is a motor disorder characterized by a velocity-dependent increase in the tonic stretch reflexes (muscle tone) with exaggerated tendon jerks, resulting from hyper excitability of the stretch reflex (Lance, 1980). This work focuses on quantifying the effects of spasticity on motion kinematics during voluntary movement.

1.2 ENHANCE project

This study is a part of the ENHANCE project (Levin et al., 2018). The ENHANCE project is an international collaboration project between Israel, India, and Canada. The project deals with enhancing brain plasticity for sensorimotor upper limb (UL) recovery in spastic hemiparesis of patients after stroke.

Enhance proposes a training program that combines current knowledge about brain plasticity and motor control, and includes virtual reality (VR) combined with non-invasive brain stimulation to enhance motor learning. The training incorporates personalized transcranial direct current stimulation (tDCS), which is a form of neurostimulation that uses constant, low current delivered to the brain area of interest via electrodes on the scalp, to balance cortical hypo/hyperexcitability. In addition, it involves personalized reaching training, based on the identification of the individual's disorders in spatial threshold (ST). The training approach is guided by identification of the elbow angular zone in which spasticity occurs ('spasticity zone') and limiting reaching training to the zone in which active control is preserved ('active control zone'), in each participant.

The first goal of the ENHANCE project is testing the effectiveness of personalized training programs to enhance UL motor, by increasing the range of regulation of STs in the elbow during reaching. The second goal is determining the effects of tDCS aimed to decrease spasticity and improve motor function of the arm. The third goal of the project is determining the feasibility of implementing personalized training programs in high and low-to-middle income countries. In order to test the treatment outcomes, the reach-to-grasp task has been chosen for kinematic assessment. This task has been chosen since it represents a functional reaching task which relies on the coordination of UL and trunk segments.

1.3 Motivation and objectives

Current clinical spasticity measures and clinical motor deficits measures do not capture the relationship between voluntary movement disorders and spasticity (Malhotra et al., 2009; Gregson et al., 1999; Calota et al., 2008). One of the reasons is that both phenomena are complex and consist of two (spatio and temporal)

dimensions and variability. The use of stochastic models offers an important tool for investigating this relationship.

In a previous work of our group, Davidowitz et al. (2019) have suggested using Gaussian mixture models (GMM) and the bi-directional Kullback-Leibler Divergence (BKLD) for quantifying the effects of spasticity on voluntary motor control. They have shown that for elbow reaching motion, in a cohort of 16 subjects with stroke, motion models of patients with higher MAS (more spasticity) were more distant (larger BKLDs) from motion models of healthy individuals. This work aims at further developing a more robust measure for quantifying the influence of spasticity on motor performance during voluntary movement, based on motion kinematics. In the current work we suggest a new stochastic distance measure for quantifying the distance between the models: the Hellinger's distance (HD) measure. HD has advantageous over the BKLD (further information in 2.4.2 Hellinger's distance), which we demonstrate in this work. In addition, we show the relations of this distance with spasticity quantified by the MAS, the tonic stretch reflex threshold (TSRT), and the regression line slope which gives an indication regarding the velocity dependency of the spasticity (Calota & Levin, 2009) (details regarding the clinical measures in 2.2.2 Clinical spasticity measures). We compare the HD and the BKLD as measures of the influence of spasticity on motor disorders.

1.4 Innovations

In order to quantify the influence of spasticity on motion kinematics with a robust measure, we used the stochastic HD measure (details regarding HD in 2.4.2 Hellinger's distance), based on the method developed in a previous work of our group (Davidowitz et al., 2019). The measure is based on spatio-temporal Gaussian mixture models (ST-GMMs) (details regarding ST-GMMs in 2.3 Using stochastic mixture models for motion modeling) constructed from motion trajectories. Using stochastic models offers a comprehensive representation of data which facilitates integrating multiple process dimensions along with variability, within a single generalized model. These can be of importance when representing motion data for examining motion quality, e.g., for monitoring rehabilitation progress.

We show that HD can be a good biomarker of motion disruptions and can serve as an objective, robust measure for the influence of spasticity on reaching movements. When compared to other similarity measures and kinematic characteristics, e.g. BKLD, HD has advantages which are shown in this work. We demonstrate that HD is strongly related to the patient functional ability and joint spasticity, is more robust, and has the best generalizability to the larger stroke population.

The TSRT is an objective, innovative, relatively new measure for quantifying spasticity. In this research we demonstrate that the TSRT is statistically superior to the most commonly used clinical spasticity measure - the MAS. We show that TSRT has high resolution and entropy, and is independent of the Fugl-Meyer assessment (FMA) for UL motor dysfunctions. This is while MAS is subjective, has low resolution and entropy, and has multicollinearity effects with FMA when modeling kinematic measures.

The study was summarized in a poster presented at the “Progress in Motor Control XII” conference, Lackritz et al., Quantifying the effects of spasticity on reaching movement patterns using stochastic spatiotemporal modeling, Holland, Amsterdam, 2019. This work was additionally presented in the “15th Karniel Computational Motor Control Workshop”: Lackritz et al., Stochastic Spatiotemporal Modeling and Spasticity, Beer Sheva, 2019, and at a Microsoft data science club talk, Herzelia, September 2019. A journal publication is currently under development. In addition, our developed measure will be one of the secondary outcome measures used in the ENHANCE project for measuring treatment efficiency through decrease in upper limb spasticity.

1.5 Work scope

Subject motion analyzed included 13 healthy control subjects that were recorded in August 2016 and 42 subjects with stroke that were recorded from August 2016 to January 2019, in Canada, India, and Israel. As part of this study, a complete kinematic analysis of the ENHANCE project database was performed for all pre, post (2 weeks post-intervention), and follow-up (1 month post-intervention) data. The analysis included a construction of the method for calculating the joint centers of the wrist,

elbow, and shoulder, based on the works by O'Brien et al. (1999) and Davidowitz et al. (2019). Post and follow-up recordings were not statistically analyzed in the current work. Elbow spasticity is analyzed in this work since it has a lot of motion contribution in the reach-to-grasp motion task presented to the participants and the joint for which there are reference spasticity clinical measures.

1.6 Thesis structure

The rest of this report is organized as follows: Chapter 2 presents a literature review. Reviewed topics include spasticity following stroke, clinical spasticity measures, motion modeling using stochastic mixture models, and GMMs in particular. In addition, the chapter presents HD measure, other similarity measures for GMMs, and a comparison between them. Chapter 3 describes the modeling method and the research hypotheses. The chapter includes the required pre-processing, the analysis of the raw data, and the calculation method of the different measures and models. Chapter 4 presents the results and the statistical analysis. Chapter 5 presents a discussion consider these results, the conclusions from this research, and suggested future work.

2. Literature review

2.1 Overview

The literature review includes concepts used in this study and scans measures, methods, algorithms, and previous researches related to stochastic motion modeling, similarity measures, and spasticity. Section 2.2 provides background on spasticity in patients with stroke. The section additionally reviews existing clinical measures for spasticity and their limitations. Chapter 2.3 describes the use of stochastic models for representing spatio-temporal motion data, especially spatio-temporal Gaussian mixture models, and their use for examining motion quality. The concepts reviewed include previous work of our group, model definitions, and parameter estimation. Chapter 2.4 presents distance measures between distributions, suitable for multivariate mixture models. The chapter describes the main distance measure examined in the current work - the Hellinger's distance measure.

2.2 Spasticity

2.2.1 Spasticity following stroke

Spasticity is a motor disorder characterized by a velocity-dependent increase in the tonic stretch reflex (tonic contraction of the muscles in response to a stretching force) with exaggerated tendon reflexes, resulting from the hyper excitability of the stretch reflex, as one component of the upper motor neuron syndrome (Lance, 1980). It is one of the most common disorders caused by stroke. During the first year following the stroke 20-50% of the patients suffer from spasticity (Sommerfeld et al., 2004). It leads to difficulty in daily activities and to reduced quality of life (Nichols-Larsen et al., 2005). Spasticity is often medically treated with an injection of botulinum toxin A (botox) or the drug baclofen.

Clinically, spasticity is assessed during passive rather than voluntary motion. The functional state of the motor system during voluntary motion is more complex than under passive conditions. Therefore, phenomena such as hypertonia displayed by a passive muscle following imposed stretch will not necessarily appear when the

muscle is stretched as part of a voluntary movement. Thus, the role of spasticity in the disruption of voluntary movement remains controversial (Fellows et al., 1994).

According to Jobin and Levin (2000), spasticity may be characterized by the limitation of the central nervous system (CNS) to regulate the range of stretch-reflex thresholds in flexor and extensor muscles. They showed that for patients with stroke the ability to regulate muscle force throughout the physiological range (control zone) may be lost due to narrowing of the stretch-reflex regulation thresholds. Measuring spasticity can benefit physical therapy treatments and patients with stroke condition and rehabilitation evaluation. As the spasticity measure will be more accurate, the quality of the treatments, and the contribution to the rehabilitation of the patients, can be measured more accurately. Furthermore, treatments could be better adapted to the patients. The following sections will present existing measures for spasticity, and the use of two methods combining ST-GMM with ST-HD and ST-KLD for quantifying the effects of spasticity on motor control.

2.2.2 Clinical spasticity measures

The primary clinical measure used to measure spasticity is the Modified Ashworth Scale (MAS) (Bohannon & Smith, 1987). This is a discrete, subjective measure which grades the resistance felt during stretching of passive muscles on a 6-point ordinal scale (Charalambous, 2014). The MAS uses a 1+ scoring category, which was added to the original Ashworth scale (AS) to indicate resistance through less than half of the movement and therefore increasing its sensitivity with 6 instead of 5 levels, as shown in Figure 1 (Bohannon & Smith, 1987).

One of the main problems of the MAS is that the resistance to passive movement and its range are complex variables that normally vary with the level of activity (voluntary and reflex). These variables may be influenced by many factors, e.g. temperature, only one of which could be spasticity. Another factor that influences the MAS score is the therapist's experience (Lee et al., 1989). The MAS assessment is subjective and therefore may be inconsistent and could affect the efficacy of the rehabilitation process (Puzi et al., 2017).

Grade	Description
0	No increase in muscle tone
1	Slight increase in muscle tone, manifested by a catch and release or by minimal resistance at the end range of motion when the affected parts 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 range of motion
2	More marked increase in muscle tone through most of the range of motion, but the affected part is easily moved
3	Considerable increase in muscle tone, passive movement is difficult
4	Affected part is rigid in flexion or extension

Figure 1 : Grades and descriptions of the Modified Ashworth scale.

Another problem regarding using the MAS as a spasticity measure is that evidence suggests that the resistance to passive movement is only an effect of spasticity and is not an exclusive measure of it (Pandyan et al., 1999). The assessment of resistance to passive muscle stretch does not capture all aspects of spasticity, such as velocity dependent and its effects on motion quality. Furthermore, evidence suggests that the resistance to passive movement is not significantly influenced by reflex neural activity unless the velocity of the passive stretching is high, although the MAS does not check the movement at high velocity only (Pandyan et al., 1999). Another issue is that the MAS takes no account of the relation of abnormal tone with posture and associated reaction, both of which may be important for the measurement of tone and its impact on function (Gregson et al., 1999).

The tonic stretch reflex threshold (TSRT) (Calota & Levin, 2009) is an objective, continuous, innovative, relatively new measure for quantifying spasticity. TSRT measurement and its relationship with spasticity are based on threshold control theory (TCT) of motor control (Feldman, 2015). According to the TCT, voluntary movement is generated by regulating the STs at which muscle activation begins. EMG emerges based on the interaction of the biomechanics of the system with the environment. The TSRT, i.e., the ST at zero velocity, is extrapolated based on regression from measurement of STs at different velocities. In addition to the TSRT, which is the regression line intercept, the regression line slope gives an indication regarding the velocity dependency of the spasticity (Figure 2).

There are some more common existing spasticity measures, with different limitations. For example, the disability rating scale, which is a self-reported scale filled by the patient that notes how difficult it is for him to handle his affected limb. This measure relies on the patient's report abilities that may be hindered due to the stroke (Francisco et al., 2005). Another example is the Hand-held dynamometer, which can be used to test only the calf-muscle spasticity (Boiteau et al., 1995), or the Wartenberg pendulum test which is used to test quadriceps muscle spasticity only (Nordmark & Andersson, 2002). Another method is the H-Reflex measure that assesses the response to electrical or mechanical stimulation. This technique is simple to perform and easy to use in neurology setting, yet it has low correlations with other clinical scales (BurrIDGE et al., 2005).

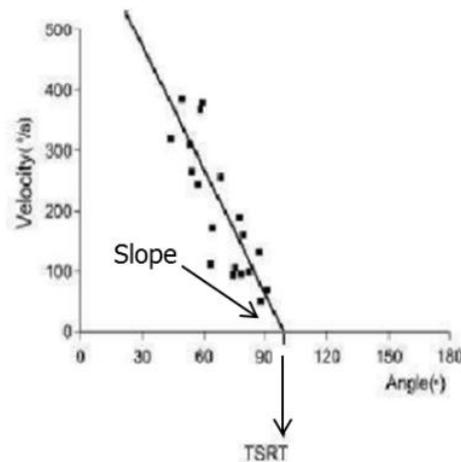


Figure 2 : The tonic stretch reflex threshold regression line. $velocity = Slope * angle + TSRT$.

2.3 Using stochastic mixture models for motion modeling

Using stochastic models rather than single moments, e.g., the sample mean, offers a comprehensive representation of data. The facilitate representation of multiple data moments, along multiple dimensions giving indications of underlining creation processes. These can be helpful when representing motion data, for examining motion quality. The additional representation detail comes with a considerable cost, both in the initial derivation from the data and in the subsequent interpretation of the results.

Stochastic Gaussian mixture models (GMM) can capture the behavior of many complex, multi-dimensional processes. They are a particularly attractive modeling option since there are readily available methods for estimation of GMM parameters. For example, expectation maximization algorithm (EM) (Dempster et al., 1977) for estimating the models weights, means, and covariances, or non-parametric Bayesian estimation (Rasmussen, 2000) for determining the number of Gaussians and making a selection between models.

A GMM is a stochastic model assumes that all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. A GMM is parameterized by two types of values - the mixture component weights and the component means and variances or covariances. For a GMM with K components, the K -th component has a mean of $\vec{\mu}_k$ and covariance matrix of Σ_k for the multivariate case. The mixture component weights are defined as ω_k for component c_k , with the constraint that $\sum_{i=1}^K \omega_i = 1$ so that the total probability distribution normalizes to 1 (Dinov, 2008).

GMM is computed by the weighted sum of K Gaussians probability densities, as given by the equation:

$$p(x) = \sum_{i=1}^K \omega_i \cdot g_i(x|\vec{\mu}_i, \Sigma_i) \quad (2.1)$$

where the multivariate Gaussian density $g_i(x|\vec{\mu}_i, \Sigma_i)$ is defined by:

$$g_i(x|\vec{\mu}_i, \Sigma_i) = \frac{1}{(2\pi)^{2/2} \sqrt{|\Sigma_i|}} \cdot \exp\left\{-\frac{1}{2}(x - \vec{\mu}_i)^T \Sigma_i^{-1} (x - \vec{\mu}_i)\right\} \quad (2.2)$$

$$\mu_i = \{\mu_{t,i}, \mu_{s,i}\}, \quad \Sigma_i = \begin{pmatrix} \Sigma_{tt,i} & \Sigma_{ts,i} \\ \Sigma_{st,i} & \Sigma_{ss,i} \end{pmatrix} \quad (2.3)$$

where $\mu_{s,i}$ is the spatial expectation and $\mu_{t,i}$ is the temporal expectation of the i -th component, and $\Sigma_{xx,i}$ is the variance or covariance of the i -th component.

One of the main advantages of using spatio-temporal Gaussian mixture models (ST-GMM) is that hidden parameters are modeled without explicit assumptions and therefore the model can easily be applied to varied applications without requiring

additional assumptions. This, along with the unconstrained covariance structure of stochastic mixture models comparing to deterministic models, allows flexibility in the modeling of the covariance structure. In addition, ST-GMM can model probability distributions to any required level of accuracy with enough components (Hinton et al., 2012).

For reach-to-grasp motion, ST-GMM models were constructed in a previous study of our group for the elbow joint (Davidowitz et al., 2019), and the distance between models of subjects with stroke and healthy controls was quantified using the Bidirectional Kullback-Leibler divergence (BKLD). This distance was related to the Modified Ashworth scale (MAS) spasticity measure, where patients with higher MAS had a higher BKLD value, indicating a GMM model and movement pattern that were more distant from those of controls. In the current study we examine BKLD distance measure for mixture models, and focus on the Hellinger's distance measure (HD) which has advantageous over the BKLD and other distribution distance measures.

2.4 Distribution distance measures

2.4.1 Distance measures for mixture models

Common classical maximum-likelihood based goodness-of-fit measures, e.g. Chi-squared test, cannot be used for comparing mixture models due to their multivariate nature. Furthermore, log likelihood based fit measures cannot be used directly since no closed form exists for the asymptotic distribution of the log-likelihood ratio statistic of mixture of two or more Gaussians. This can be used to produce confidence intervals for maximum-likelihood estimates or as a test statistic for performing the Likelihood-ratio test. Different methods have been developed for measuring distance between the GMM models, thus facilitating the use of this modeling technique for measuring various motion related phenomena (Jensen et al., 2007; Kristan et al., 2011). Selecting the suitable method is important for attaining improved performance and since computation cost is typically non-negligible.

One of the most commonly used methods for measuring similarity between GMMs is the Kullback-Leibler Divergence (KLD) (Kullback & Leibler, 1951; Jensen et al., 2007; Goldberger & Aronowitz, 2005). KLD is an information-based measure of disparity

among two probability distributions. It measures the dissimilarity between two probabilistic variables defined over the same set of outcomes (Polani, 2013). KLD can be expressed as the difference in the amount of additional information needed to reconstruct the probability distribution p with probability distribution q (Hyun et al., 2019). Very often in probability and statistics the observed data or complex distributions are replaced with a simpler, approximating distribution. KLD measures how much information is lost when the approximation is chosen. In other words, it is the expectation of the log likelihood ratio between the probability distributions p and q , where the expectation is taken using the probability p (Altman, 1992).

Specifically, the KLD of q from p , denoted $D_{KL}(p||q)$, is a measure of the information lost when q is used to approximate p . Typically p represents the "true" distribution of data, observations, or a precisely calculated theoretical distribution. The measure q typically represents a theory, model, description, or approximation of p . When p and q are continuous variables, KLD is defined by:

$$D_{KL}(p||q) = \int_{-\infty}^{\infty} p(x) \log\left(\frac{p(x)}{q(x)}\right) dx \quad (2.4)$$

KLD goodness of fit measure can be used as a similarity measure between distributions, so that the higher $D_{KL}(p||q)$ is, the less similar p and q are. KLD is always non negative with no upper bound, hence $D_{KL}(p||q) \geq 0$, while $D_{KL}(p||p) = 0$ indicating identical behavior of the two distributions. KLD is not a distance measure as it is not symmetric, not does it satisfy the triangle inequality. In order to overcome the asymmetry of the KLD, a symmetric variant can be applied: the Bidirectional-KLD, and thus, can be used as a distance measure:

$$BKLD(p||q) = \frac{D_{KL}(p||q) + D_{KL}(q||p)}{2} \quad (2.5)$$

Since the measure is based on log likelihood, KLD between GMMs cannot be computed analytically, it can be estimated using methods such as the variational approximation, Monte Carlo simulation, Gaussian approximation, or a lower-bound approximation. The variational approximation is relatively quick to compute, comparing with the Monte Carlo and Gaussian approximations for example. Therefore, when computation time is an issue, the variational approximation may be useful. When compared to other simple, closed-form approximations of the

similarity between GMMs, e.g. the lower-bound approximation and the Gaussian approximation, the variational approximation is the most accurate (Hershey & Peder, 2007; Hershey, Olsen, & Rennie, 2007).

2.4.2 Hellinger's distance

The HD is used to quantify the distance and separability between two probability distributions on the same set of outcomes. To define the HD, let p and q denote two probability measures with respect to a third probability measure λ . The square of the HD between p and q is defined as the quantity:

$$H^2(p, q) = \frac{1}{2} \int \left(\sqrt{\frac{dp}{d\lambda}} - \sqrt{\frac{dq}{d\lambda}} \right)^2 d\lambda \quad (2.6)$$

where multiplying by $1/2$ ensures $0 \leq H^2(p, q) \leq 1$, thus its values are simple to interpret (Lindsay, 1994). The HD between p and q does not change if λ is replaced with a different probability measure - it does not depend on the choice of the measure λ . Thus, the maximum distance 1 is achieved when p assigns probability zero to every set to which q assigns a positive probability, and vice versa (Cutler & Cordero-Brana, 1996). To define the HD in terms of elementary probability theory, take λ to be Lebesgue measure (Lebesgue, 1902). The Lebesgue measure is the standard way of assigning a measure to subsets of n -dimensional Euclidean space. It is used to define Lebesgue integration (Bartle & Bartle, 1995) (more details regarding Lebesgue measure in Appendix A - Lebesgue measure). When λ is Lebesgue measure, $dp/d\lambda$ and $dq/d\lambda$ are simply probability density functions. Denoting the densities as p and q , the squared HD can be expressed as a standard calculus integral:

$$H^2(p, q) = \frac{1}{2} \int (\sqrt{p(x)} - \sqrt{q(x)})^2 dx = 1 - \int \sqrt{p(x)q(x)} dx \quad (2.7)$$

where the second form can be obtained by expanding the square and using the fact that the integral of a probability density over its domain equals 1.

Due to the multivariate nature of the GMM distribution, HD between GMMs cannot be computed analytically, thus, in order to compute a highly accurate estimate, in this research we examined the Unscented Hellinger's distance between two GMM

distributions, as defined by Kristan et al. (2011). The unscented transform is a special case of a Gaussian quadrature, which, similarly to Monte Carlo integration, relies on evaluating integrals using carefully placed points, called sigma points, over the support of the integral.

$$H^2(p, q) = \frac{1}{2} \int g(x) p_0(x) dx \quad (2.8)$$

where

$$g(x) = \frac{(\sqrt{p(x)} - \sqrt{q(x)})^2}{p_0(x)} \text{ and } p_0(x) = \sum_{i=1}^n w_i g(x|\mu_i, \Sigma_i) \quad (2.9)$$

$p_0(x)$ is a combination of the two density functions, w_i is the weight of each of the Gaussian distributions that make up p and q , Thus $n=k_1+k_2$. By multiplying and dividing by $p_0(x)$ the function does not change. This creates an expectation function of a nonlinear transduction within the integral. Approximation of the expectation is much easier and faster due to existing generic solutions by the unscented transformation (Kristan et al., 2011).

Various alternatives to HD exist, for example, the Earth Mover distance (EMD), the L2-norm distance measure, and the KLD measure. For EMD distance, large computation overhead hinders its popularity. Furthermore, it is hard to set the required parameter of the basic distance in EMD. In addition, HD behaves more accurate and scalable than the EMD distance (Bishop, 2006; Ni et al., 2013). As for the L2-norm, it provides a bigger weight to farther points in the distributions and is thus susceptible to outliers (Ni et al., 2013). KLD outperforms EMD and L2-norm measures in terms of accuracy when measuring similarity between mixture models, especially when the number of components in the mixtures is higher than two (Jensen et al., 2007).

The HD measure has advantageous over other similarity measures, e.g. KLD, maximum likelihood, L2-norm, or EMD (Cutler & Cordero-Brana, 1996). One of its advantages is that HD is a metric, therefore is symmetric and satisfies the triangle inequality, allowing faster data localization as well as speeded up data clustering and nearest neighbor search (Weller-Fahy et al., 2014; Jensen et al., 2007). HD is not computationally intensive comparing to the BKLD (Sengar et al., 2008) and is shown

to be invariant, consistent, asymptotically normal, and robust (Tamura & Boos, 1986; Lindsay, 1994; Simpson, 1987). In addition, HD gives little weight to counts that are improbable relative to the model. It does not give a large weight for the distributions tails and to outliers that can have a substantial impact on wrong experimental conclusions (Lindsay, 1994), in contrast to KLD and L2-norm for example, which are susceptible to outliers (Simpson, 1987).

3 .Method

3.1 Hypotheses

We define 2 research hypotheses:

H1: Both the TSRT and the slope will be related to the movement disorders measures: the HD and the BKLD. We additionally hypothesize that HD will be a more robust measure than the BKLD for the effects of spasticity on motion deficits.

H2: The TSRT will be statistically superior to the MAS as a clinical spasticity measure.

3.2 Subjects

42 participants with stroke (28 males, age 53.3 [10.5 SD] years, 21 left-hemiparesis), medically stable in the sub-acute phase (3 weeks to 6 month post-stroke) and 13 healthy controls of similar age (9 males, 60.5 [8.7 SD] years) participated in the experiment (Table 1). Participants with stroke sustained a first ever stroke in the midcerebral artery territory, confirmed by medical resonance imaging/computed tomography, had arm paresis (Chedoke-McMaster Stroke Assessment 2-6/7) (Gowland et al., 1993), were able to perform voluntary elbow extension/flexion movement of at least 30° per direction, had elbow flexor/extensor spasticity, and were able to provide informed consent. Individuals were excluded if they had additional neurological, neuromuscular or orthopedic problems, pain, difficulty comprehending instructions, or if they were under antispasticity medication. Participants signed informed consent forms approved by institutional review boards of Loewenstein Rehabilitation Hospital, Raanana, Israel; Center for Interdisciplinary Research in Rehabilitation, Montreal, Canada; and Kasturba Hospital, Manipal, India.

UL impairment was assessed with the Fugl-Meyer assessment (FMA) (Fugl-Meyer et al., 1975). The FMA is a 66 point scale for performance-based sensorimotor assessment of UL motor function in patients with stroke. FMA scores between 0 and 30, 31 and 50, and 51 and 66, represent severe, moderate, and mild motor impairment, respectively (Duncan et al., 1994). Clinical spasticity was assessed with the MAS and the TSRT. For the patients with stroke the FMA values ranges from 14

to 57, elbow flexor and extensor MAS were in range of [0, 2], and TSRT values were in range of [54,162] degrees.

Table 1 : Mean (SD) estimates of demographic and clinical data

Group	Age	Participants (Male)	TSRT	Slope	FMA (/66)	Days since stroke
Stroke	53.3 (10.5)	42 (28)	101.4 (22.8)	-31.2 (58.4)	33.2 (12.2)	62.7 (37.0)
Control	60.5 (8.7)	13 (9)	-	-	-	-

Abbreviations: TSRT- Tonic stretch reflex threshold; FMA- Fugl-Meyer Assessment.

3.3 Experimental Procedure

The recordings took place in three different centers in Israel, Canada, and India. Participants performed reach-to-grasp motion toward a hollow cone (6-cm diameter base) placed on a table at 4 target locations which require coordination of UL segments (Levin et al., 2018; Davidowitz et al., 2019). The target locations were at two-third arm's length (near-center) and at one arm's length (far-center) in the midsagittal plane and ~30 cm to the right/left (depending on hemiparetic side: contralateral/ipsilateral) (Figure 4). Arm length was measured with the elbow extended from the medial axillary border to the distal wrist crease. Arm motion was recorded by a wireless electromagnetic tracking system G4 (Polhemus, Colchester, VT) with 5 sensors, each measuring 6 degrees of freedom with respect to a base calibration frame and is tracked at 120Hz. Sensors were placed on the midsternum, midpoint of the acromial superior-lateral border, midpoint of the ventrolateral arm, dorsal forearm (1/3 forearm length proximal to ulnar head), and the index metacarpophalangeal joint (Figure 5). All experiment recordings were saved to Microsoft Excel™ files using Matlab™.

Participants sat on an armless, wooden chair, in front of a table, with feet supported but unrestricted trunk movement (Figure 3). Initial arm posture was set with 30° elbow flexion by placement of the third fingertip on an ipsilateral seat height-support (Figure 3). The experiment began with recording session of calibration movements which included seven movement types: elbow extension, elbow supination, wrist extension, wrist adduction, shoulder adduction, shoulder extension, and shoulder rotation. Moreover, a static calibration was performed for the chair. The target

locations suitable for the participants were determined and marked by colored tape. The sensors were attached to the participants. Before motion was recorded, participants practiced reaching each target twice (a total of 8 reaching movements). During the recording participants performed 2 sets of 40 semi-randomized reach-to-grasp movements (10 trials toward each of the 4 targets, for 2 sets, total 80 trials) balanced in blocks across targets (so even when not all trials were completed, e.g. due to fatigue, a balanced number of trails per target was maintained). Each trial consisted of a series of movements starting from the initial position: reaching to grasp the cone "as fast and as precisely as possible"; holding the cone for 2 s; lifting the cone toward the chin; returning it to the target position on the table, and finally returning the hand to the initial position (Figure 3-right). Only the first segment of the sequence, reaching to grasp the cone, was analyzed. The full sequence was conducted to maintain a functional reaching task. Participants rested between trials and blocks as needed. Events during trials were logged, e.g. when the participant didn't succeed in grasping the cone, collided with the table, or dropped the cone during the movement.



Figure 3 : Experimental setup. Left: participant grasping a target cone (near-center target). Right: participant with the arm in the initial position and the elbow slightly bent.



Figure 4 : Targets locations. 4 targets on the table in front of the participant. The far targets (contralateral, far-center, and ipsilateral) were at arm's length, and the near target (near-center) was at two-third the arm's length. The center targets were in the midsagittal plane. Contralateral and ipsilateral targets for right-hand hemiparetic participants were 30 cm along the horizontal axis to the left (contralateral) and right (ipsilateral) of the center targets, respectively (Davidowitz et al., 2019).



Figure 5 : Sensors positions. Movement was recorded using 5 sensors, each measuring 6 degrees of freedom with respect to a base calibration frame and is tracked at 120Hz. Sensors were placed on the midsternum, midpoint of the acromial superior-lateral border, midpoint of the ventrolateral arm, dorsal forearm (1/3 forearm length proximal to ulnar head), and the index metacarpophalangeal joint.

3.4 Pre-processing

3.4.1 Data validation

Due to the fact that the data are collected by three international centers, creating uniformity in the process was difficult. This difficulty led to errors and inconsistencies, which added a lot of work to the analyzing process of the data. The main causes of the data errors were sensor hubs that fail to communicate the data

to the computer, reverse connection of some of the sensors which caused disruption of the axes systems' directions, and switching the locations of the contralateral and ipsilateral targets. These problems created a regression at the analysis of the data. In order to reveal these errors and find the failed recordings, we had to analyze the data and the graphs of the movements. Data reparation was performed manually, tests were performed using Matlab™ code and by analysis of graphs of sensor outputs.

Sensor measurements with the value "0" indicated a faulty measurement and files with more than 10% faulty measurements were considered damaged and removed from further analysis. In addition, 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, motion started prior to the cue).

3.4.2 Movement Segmentation

Segmentation was performed semi-autonomously as part of this work. The motion segmentation was conducted in order to identify the reach-to-grasp segment for each of the recorded movements of each subject. The automatic procedure was developed for the initial segmentation, and the segmentation results were all manually screened. Movement trajectories were filtered using a Butterworth filter with 6 Hz cutoff frequency. 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) tangential velocity exceeded and remained above, or decreased and remained below threshold of 10% peak wrist tangential velocity. The threshold for the movement was iteratively increase by 1% in case a hand closure was not identified. Subjects typically performed two sub-movements; they raised their arm and then reached forward toward the target. The sub-movement interchange point was determined between movement onset and offset, when the elbow (upper arm sensor) tangential velocity reached a local minimum.

3.4.3 Spatial and temporal scaling

Spatial and temporal dimensions of movement trajectories differ in magnitude. Using different value ranges in stochastic modeling would result in a construction of a biased model, where the dimension with larger values would have a larger weight (Bishop, 2006). To avoid this, joint angles were linearly scaled to the range of $[-1, 1]$, similar to the average task duration:

$$x_{i,new} = 2 * \frac{x_i - \min(x)}{\max(x) - \min(x)} - 1 \quad (3.1)$$

where x is the original joint angle trajectory vector, x_i is a point along the trajectory, and $x_{i,new}$ is the transformed point. To create a model per participant and target, trajectory lengths were scaled to a uniform length for all the trials of each target per participant. A function representing each movement was approximated using general regression neural networks (Specht, 1991). To equalize the number of samples for each trial per target, the function was resampled at a constant rate determined for each participant and target, based on the average trial length originally sampled at 120 Hz.

3.4.4 Finding the coordinate frames

The purpose of computing the joint centers and the directions of the wrist, elbow, and shoulder is computing the true joint angles of the motion trajectory. The algorithm implemented for finding the joint centers is based on the work by O'Brien et al. (1999). In this method, the human arm is modeled as an articulated hierarchy of bodies connected by joints.

A joint center positioned between body i and its parent body, can be defined by vectors originated from the origins c_i and l_i respectively (Figure 6). A point x_i in the i -th coordinate system can be expressed in the j -th coordinate system. The transformation from the i -th body's coordinate system to the coordinate system of the j -th body consists of a rotational component and a translational component. The transformation can be done using the following equation:

$$x^j = R^{i \rightarrow j} x^i + d^{i \rightarrow j} \quad (3.2)$$

where R is a multiplicative, invertible, 3X3 matrix component and d is a length 3 vector component. $R^{i \rightarrow j}$ in this expression is referred as the rotational component of the transformation and $d^{i \rightarrow j}$ in this expression is referred as the translational component of the transformation.

The usage of the articulated hierarchy model of the arm, allows describing the same transformations by the following equation, using vectors at time frame k (out of n discrete time frames of motion):

$$x^{P(i)} = R_k^{i \rightarrow P(i)}(x^i - c_i) + l_i \quad (3.3)$$

where $P(i)$ is the parent body of body i . By comparing these two equations and changing to a matrix form, the following equation can be obtained:

$$Q_k^{i \rightarrow P(i)} \begin{pmatrix} c_i \\ l_i \end{pmatrix} = d_k^{i \rightarrow P(i)} \quad (3.4)$$

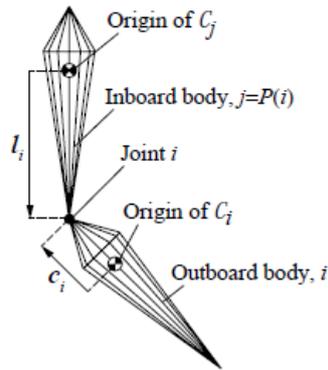


Figure 6 : Joint diagram. The location of the joint i is between bodies i and j . The location of the joint is defined by a vector c_i , relative to the coordinate system of body i , and a second vector l_i , in the coordinate system of body j .

In the current work, we used the described algorithm in order to compute the elbow, wrist, and shoulder centers for each participant. The sensors on each participant's arm, below and above each joint center, were used as the vectors origins (c_i and l_i respectively). The output of the algorithm is two vectors per joint center per participant: 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 6). Using these vectors, each point determined in the sensor coordinate system can be represented in the corresponding joint center coordinate system.

Motion data from all n time frames was used in order to calculate c_i and l_i . In addition, the calculation required a numerical solution obtained from a least squares solution using singular value decomposition, as described in O'Brien et al. (1999) and Press et al. (2007). After obtaining the vectors c_i and l_i for each joint, the joint centers can be computed by using the sensors Cartesian positions at each time step with the relative vectors. Having the location of each sensor and a vector from the sensor to the joint center, the center location can be obtained in relation to a global coordinate system. The algorithm was implemented using Matlab™.

In order to calculate the joints centers vectors according to the specified method, we used calibration files recorded for each participant at the beginning of the experiment – one for each degree of freedom. The calibration files contain the position (x, y, z) and the angle (x, y, z) of each of the 5 sensors at each time step during the movement. The sensors' locations were sampled using repeated local movements of each degree of freedom. For each joint, the center was computed twice, once using the vector l_i and once using the vector c_i . Center total error was defined as the total Euclidean distance between the two center position vectors.

Performing the distal arm movement around the elbow (supination), the forearm sensor moves only slightly compared to the elbow center. Hence, most of the data for the elbow joint center algorithm comes from the extension movement. Since most of the movement is around one axis, the output of the algorithm is an axis on which the center lies and not an exact point. This fact caused a large error in this center's estimation. Due to this error, the elbow center was calculated by obtaining the location that remains at the same distance from the shoulder center and the wrist center at all time steps. This location was calculated using optimization algorithm.

3.4.5 Finding joint angles

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

$$T_{0Mi} = \begin{pmatrix} R_{zyx}(Ox_{Mi}, Oy_{Mi}, Oz_{Mi}) & X_{Mi} \\ 0 & Y_{Mi} \\ & Z_{Mi} \\ & & 1 \end{pmatrix} \quad (3.5)$$

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

$$R_{zyx}(a, e, r) = \begin{pmatrix} \cos(a)\cos(e) & \cos(a)\sin(e)\sin(r) - \sin(a)\cos(r) & \cos(a)\sin(e)\cos(r) + \sin(a)\sin(r) \\ \sin(a)\cos(e) & \cos(a)\cos(r) + \sin(a)\sin(e)\sin(r) & \sin(a)\sin(e)\cos(r) - \cos(a)\sin(r) \\ -\sin(e) & \cos(e)\sin(r) & \cos(e)\cos(r) \end{pmatrix} \quad (3.6)$$

where a is azimuth, e is elevation, and r is roll.

In order to validate the joint trajectories, we recorded and filmed movements of 1 control participant. The participant conducted all calibration movements and 10 reach-to-grasp movements toward the 4 targets. Using the calibration files, joint centers were computed. The joint angles were calculated per time step. The angles were then represented by a graph of joint angle against time in seconds from the beginning of the movement. The graphs were compared to the films manually.

7 angles (shoulder extension, shoulder adduction, shoulder rotation, elbow extension, elbow supination, wrist extension, and wrist adduction) were computed for the ENHANCE project, yet only the elbow extension angle was used and analyzed in this project. Joint angles definition can be found in Appendix D - Joint Angles definitions.

3.5 Analysis

Trials were discarded in case of a recording error or task failure. Sensor data were filtered using a standard 2-way (zero lag), low-pass, third-order Butterworth filter with a 6-Hz cutoff. The first movement segment's onset and offset were defined as the times at which the forearm tangential velocity exceeded and remained above, or decreased and remained below 10% of peak forearm tangential velocity for 0.1 s. Tangential velocity was computed by differentiating position samples. The joints' centers were calculated for each subject and joints' angles were calculated per trial.

The GMMs were computed per subject for each target. Each of the 4 models per subject included spatial dimension (angle) and temporal dimension (time). The parameters for each model were initialized using K-means algorithm and estimated using EM algorithm. The best fitted model was chosen according to the "knee method" (Zhao et al., 2008) – GMMs were computed for different values of K (number of GMM components) varying from 2 to 25 Gaussians, for each K the model's BIC score was calculated. The curve of the BIC scores according to K was plotted, and the location of the bend (knee), where the BIC score ceases to improve significantly, was chosen. The BIC estimations were calculated according to 7 repetitions for each k .

HDs (calculated using the unscented transform) and BKLDs (calculated using the variational approximation) were computed between GMMs of different subjects per target. Within-control group values were computed for all controls (for each control subject versus all other control subjects), and between-group values were computed between patients with stroke and control subjects. BKLD is a unitless measure, therefore within-control group BKLDs were used as normative comparators. The final HD and BKLD scores for each participant (for both within and between computations) were determined as the minimal HD and minimal BKLD score using the nearest-neighbor methodology (Komaty et al., 2013). The between-group values were used in order to assess movement similarity between subjects with stroke and control subjects, and the within-group values were used in order to examine the similarity of the movement patterns of healthy subjects.

Additionally to the HD and BKLD values, kinematic measures were calculated for the reach-to-grasp movements. The measures included final elbow angle, movement time, average elbow velocity, and elbow velocity smoothness. Each measure was calculated per each trial and the mean was calculated for each subject per target. Average elbow final angle was calculated using circular mean calculation of the movements. Velocity was computed by differentiating joint angles at each time step. Subjects typically performed two sub-movements; they raised their arm and then reached forward toward the target. Therefore, the velocity was computed by averaging only samples that have passed the threshold of 0.1 of the maximum

sample value - ignoring the pauses. Mean movement time was calculated as the difference between movement offset and onset (from start to end of reach). For assessing discontinuities in the motion trajectory, elbow velocity smoothness was calculated by the number of acceleration zero-crossings. All of these measures were calculated for both stroke and control groups.

3.6 Statistical analysis

Statistical analysis was performed using R Studio IDE for R (version 3.4.2). Analysis was performed using Linear Mixed Models (LMMs) with restricted maximum likelihood (REML) criterion for convergence (Satterthwaite, 1946). REML is a particular form of maximum likelihood (ML) estimation that does not base estimates on a ML fit of all the information, but instead uses a likelihood function calculated from a transformed set of data (Cheung, 2013). All LMMs included subjects as random effect intercept. The LMM equation is as follows:

$$y_{i,j} = \beta X_{i,j} + \alpha_j + \varepsilon_{i,j} \quad (3.7)$$

where $y_{i,j}$ is the outcome value of measurement i , for participant j . β is the fixed effects vector, $X_{i,j}$ is the explanatory variable vector, $\alpha_j \sim N(0, \sigma_\alpha^2)$ is the random effect of participant j , and $\varepsilon_{i,j} \sim N(0, \sigma_\varepsilon^2)$ is the random error.

LMMs are an extension of simple linear models in order to allow both fixed and random effects. The LMM model is suitable for data sampled from normal distributions (Cnaan et al., 1997). LMM can be thought of as a trade-off between aggregating all the data coming from each subject and analyzing data from one unit at a time. The individual regressions have many estimates and lots of data, but are noisy. The aggregate is less noisy, but may lose important differences by averaging all samples within each subject. In our model the subject's values were treated as the sum of fixed and random effects.

LMM yields asymptotically efficient estimators - that is, tend toward being optimal (minimal variance) as the sample size increases for both balanced and unbalanced research designs. In contrast, analysis of variance (ANOVA) produces an optimal estimator only for balanced designs, and is not exact as the REML of the LMMs. In LMM, maximum likelihood and REML methods are used for estimating model

parameters. REML produces variance component estimators with a smaller bias than maximum likelihood and is, thus, more similar to the traditional ANOVA analysis.

In order to determine differences between within-control group and between-groups HD and BKLD values, we used two LMMs. Each LMM included HD or BKLD type (within-control group, between-group), target (near center, far center, contralateral, ipsilateral), and their interaction as factors. LMMs were computed for final elbow angle, mean movement time, average movement velocity, and velocity smoothness as independent variables as well. The models included group (control, stroke), target location, and their interaction as factors. Separate LMMs were used for each group in case of a significant interaction between groups. In addition, in order to examine the influence of elbow spasticity measured by TSRT, the regression slope, and MAS, with the upper limb impairment level measured by FMA, on kinematic characteristics (HD, BKLD, final elbow angle, mean movement time, average movement velocity, and velocity smoothness), LMMs were computed. In all analyses, target location and its interactions with the clinical measure (TSRT, slope, MAS, FMA) were defined as factors. Furthermore, in order to examine the influence of elbow spasticity measured by MAS separately from FMA, LMM including MAS, target location, and its interaction was used for each measure.

Conditional R^2 (R_c^2) and marginal R^2 (R_m^2) values were evaluated for all models (Nakagawa & Schielzeth, 2013). The R_c^2 represents the variance explained by both fixed and random factors, and thus indicates how the model fits the participant group. The R_m^2 represents the variance explained by fixed factors only, and thus indicates how the model fits the general population of people affected by stroke.

Target location was used as an explanatory variable in the models due to the fact that different target locations made the participants move using different elbow angle zones. The various angles affected the movements and were therefore expected to affect the measures as well. The FMA was used as an explanatory variable in the models because upper limb motor function is likely to affect the movement patterns and therefore the measures. However, the measure is general and does not measure spasticity, thus clinical spasticity measures were also included in the models in order to explain the measures values.

4. Results

Healthy controls completed 98% of trials, whereas participants with stroke completed 79% (22% SD) of trials. For controls, 6% of all completed trials were discarded (0.7% task failure), and for participants with stroke 22% (16% task failure). Control group made smoother and faster movements than the stroke group, with smaller variance between movements (Figure 7). GMMs were constructed based on the movement patterns, thus the amount of components of the GMMs, in addition to the width of each Gaussian component, reflected the variability of the movements of each subject (Figure 8). Therefore, control group had less Gaussian components comparing to subjects with stroke (5.77 ± 0.90 for control, 11.04 ± 1.31 for stroke; $\chi^2 = 601.89$, $P < .001$), with difference between targets ($\chi^2 = 19.62$, $P < .001$). The amount of Gaussians in the models of patients with stroke while reaching to the near center target was more similar to that of the control subjects (estimated difference of 4.4 components), comparing to the amount while reaching to the far targets (estimated component differences: far center: 5.82, $P < .01$; contralateral: 5.29, $P < .05$, ipsilateral: 5.94, $P < .01$). Within and between-group HDs and BKLDs, movement times, final angles, average elbow velocities, and velocity smoothness are listed in Table 2.

4.1 Hellinger's distance and Bidirectional Kullback-Liebler divergence

Between-group HD and BKLD values were an order of magnitude higher than within-control group values (HD: $\chi^2 = 17.96$, $P < .001$, $R^2_m = 0.13$, $R^2_c = 0.88$; BKLD: $\chi^2 = 5.40$, $P < .05$, $R^2_m = 0.07$, $R^2_c = 0.86$) (Figure 9). There were no significant interaction effects of HD type and target location on the HD values. There was no main effect of target location. Within-group BKLD values did not change significantly between targets while for the between-group BKLDs there were significant differences between targets ($\chi^2 = 50.97$, $P < .001$). Participants with stroke had lower BKLD values (e.g., were more similar to controls) for reaches to the near center target than for all other targets ($P < .001$ for each of the 3 far targets) (Figure 9).

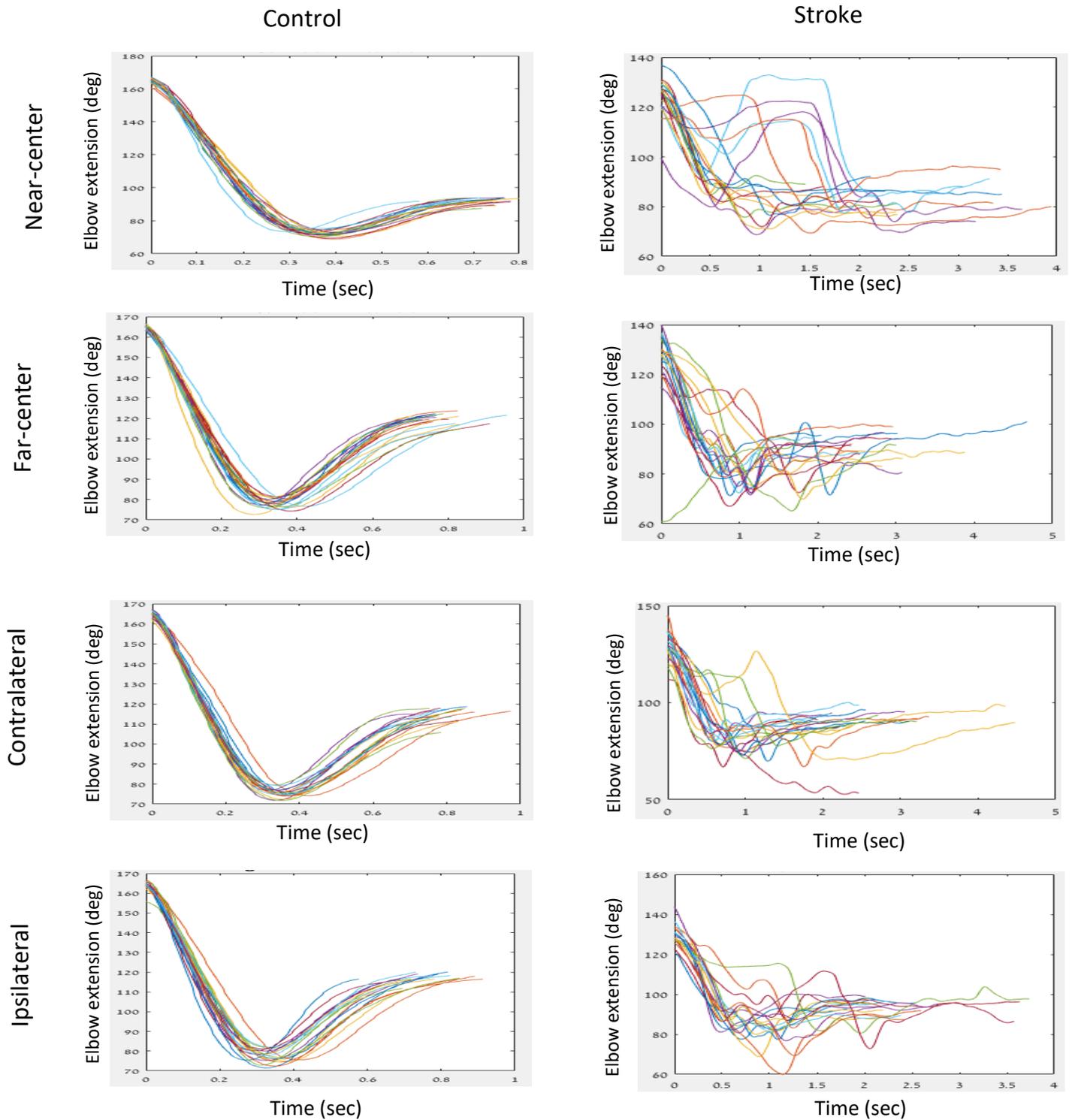


Figure 7 : Joint angles example. Elbow extension angle examples for 1 control participant (left) and 1 participant with stroke (right) for reaches toward each target.

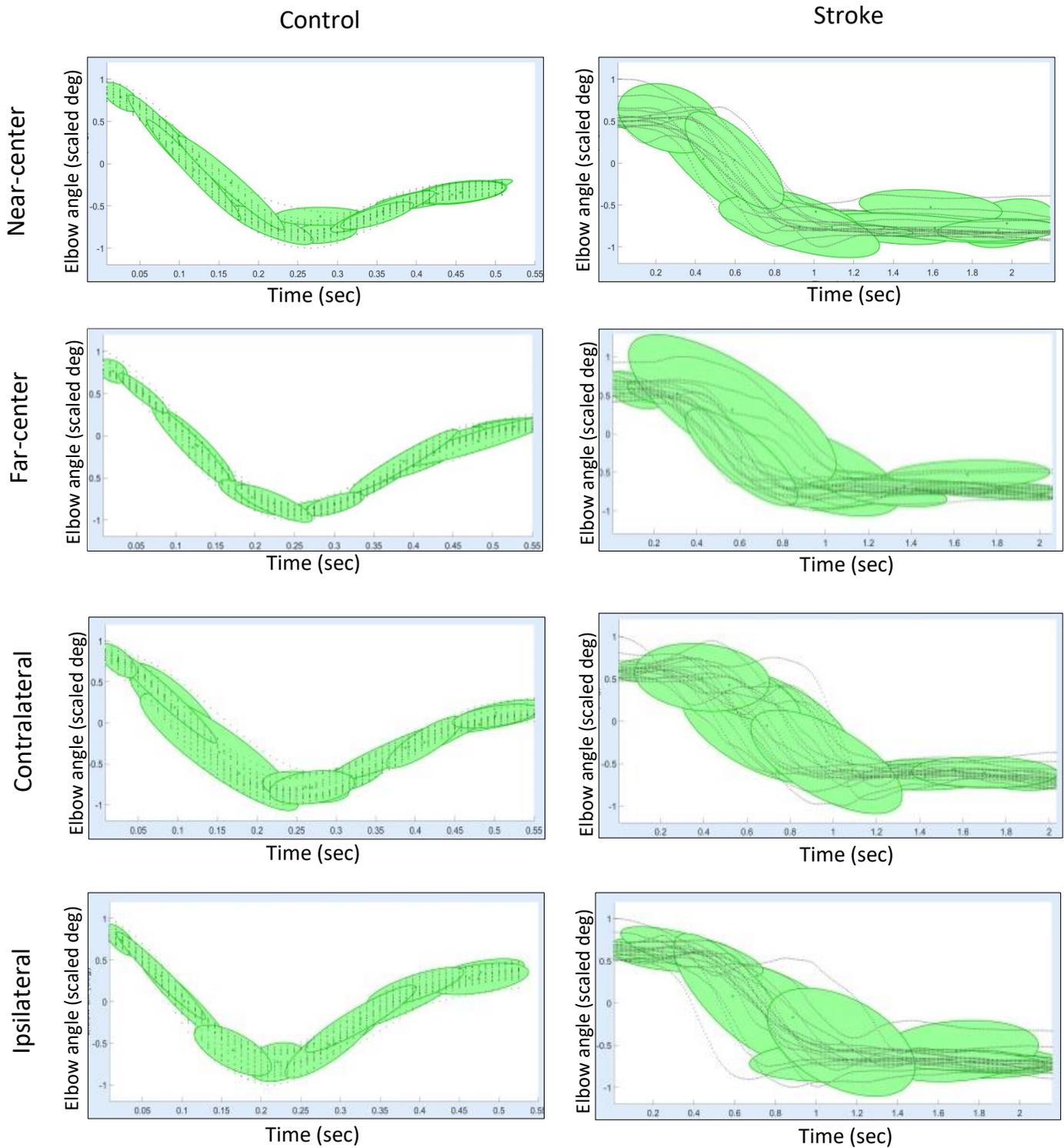


Figure 8 : GMM example. Elbow extension angles (scaled to [-1, 1]) Gaussian mixture models examples for 1 control participant (left) and 1 participant with stroke (right) for reaches toward each target (time scaled to the average movement time).

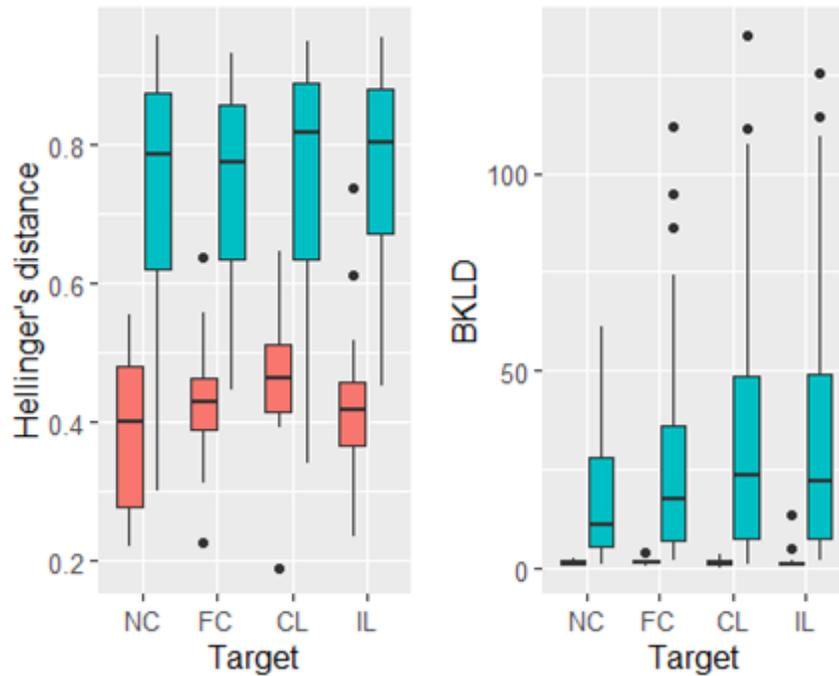


Figure 9 : HD and BKLD box-plots. Box plots of Hellinger's distance values (left) and BKLD values (right) for elbow extension for reaches to each of the 4 targets in control participants (red) and participants with stroke (blue). The large range of the BKLD values, along with the low variance within the control group, caused the control boxplots in the BKLD graph to be barely visible. Abbreviations: BKLD, bidirectional Kullback-Liebler divergence; NC, near-center; FC, far-center; CL, contralateral; IL, ipsilateral.

Table 2 : Mean (SD) estimates of kinematic characteristics

Measure\ Target	Stroke				Control			
	NC	FC	CL	IL	NC	FC	CL	IL
HD	0.74 (0.17)	0.75 (0.13)	0.76 (0.14)	0.76 (0.14)	0.39 (0.11)	0.42 (0.11)	0.44 (0.13)	0.42 (0.14)
BKLD	18.40 (16.55)	27.68 (27.60)	33.19 (33.07)	32.74 (32.28)	1.42 (0.78)	1.78 (1.10)	1.75 (1.11)	2.34 (3.60)
FA	89.34 (11.86)	98.07 (14.59)	96.48 (15.16)	99.98 (13.72)	80.95 (8.63)	101.58 (12.07)	98.41 (12.48)	100.11 (10.64)
MT	1.64 (0.51)	1.71 (0.56)	1.80 (0.59)	1.67 (0.57)	0.65 (0.15)	0.68 (0.15)	0.70 (0.16)	0.67 (0.17)
MV	89.37 (40.46)	85.99 (36.48)	86.81 (36.26)	80.66 (34.35)	244.29 (50.44)	224.85 (36.12)	22613 (41.11)	218.84 (50.82)
VS	7.36 (3.36)	7.88 (3.83)	8.34 (3.94)	7.41 (3.62)	1.83 (0.97)	1.79 (1.12)	1.82 (0.97)	1.83 (1.15)

Abbreviations: NC- Near center; FC- Far center; CL- Contralateral; IL- Ipsilateral; HD- Hellinger's distance; BKLD- Bidirectional Kullback-Leibler divergence; FA- Elbow extension final angle; MT- Movement time; MV- Mean velocity; VS- Velocity smoothness.

4.2 Clinical measures

4.2.1 Relationship with TSRT and FMA

HD values were strongly related to the clinical measures and had the highest value of R^2_m comparing to all other measures ($R^2_m = 0.55$, $R^2_c = 0.89$) (Table 3). The HD values were related to FMA ($\chi^2 = 46.07$, $P < .001$) (Figure 11), to the TSRT ($\chi^2 = 8.51$, $P < .01$) and to the interaction of TSRT and slope ($\chi^2 = 8.13$, $P < .01$). BKLD values were marginally related to the slope values ($\chi^2 = 3.35$, $P = 0.07$) and were associated with FMA ($\chi^2 = 17.23$, $P < .001$) (Figure 11). They were also associated with target location ($\chi^2 = 51.04$, $P < .001$) but not with its interactions. In addition, all kinematic measures (final elbow angle, mean movement time, average velocity, velocity smoothness) had significant main effect of FMA as explanatory variable as well ($P < .01$), and of target location ($P < .001$). Each of mean movement times, average velocity, and velocity smoothness, had significant effect of one of the measures TSRT or slope (mean time: slope: $\chi^2 = 3.22$, $P < .1$; average velocity: TSRT: $\chi^2 = 10.79$, $P < .01$; velocity smoothness: slope: $\chi^2 = 4.24$, $P < .05$). Final elbow angles were not related to the TSRT or to the slope. The interactions with target location were not significant for any of the measures. All relations with FMA, TSRT, slope, target location, and the interactions are listed in Table 3.

Table 3 : Wald chi-square values (significance levels) for LMM: Measure~FMA+(Slope+TSRT+Target)^2+ (1|ID)

Measure	FMA	Target	Slope	TSRT	Slope*TSRT	R^2_m R^2_c
HD	46.07 (***)	4.55 (-)	0.23 (-)	8.51 (**)	8.13 (**)	0.55 0.89
BKLD	17.24 (***)	51.04 (***)	3.35 (.)	0.50 (-)	1.97 (-)	0.37 0.87
FA	9.53 (**)	101.24 (***)	0.30 (-)	0.07 (-)	1.29 (-)	0.26 0.88
MT	13.05 (***)	31.10 (***)	3.22 (.)	1.73 (-)	1.66 (-)	0.34 0.96
MV	11.44 (***)	27.60 (***)	0.03 (-)	10.79 (**)	1.82 (-)	0.35 0.96
VS	9.32 (**)	36.67 (***)	4.24 (*)	0.75 (-)	1.28 (-)	0.30 0.95

*The age, days since stroke, country, gender, and the interactions with target were not significant for any of the measures. Abbreviations: HD- Hellinger's distance measure; BKLD- Bidirectional Kullback-Leibler divergence; FA- elbow extension final angle; MT- movement time; MV- mean velocity; VS- velocity smoothness; FMA- Fugl-Meyer Assessment; TSRT- Tonic stretch reflex threshold; Significance levels: . $P < .1$; * $P < .05$ ** $P < .01$ *** $P < .001$*

4.2.2 Relationship with MAS and FMA

When examining a model containing both FMA and MAS values as explanatory variables, all measures (HD, BKLD, mean movement time, final elbow angle, average movement velocity, velocity smoothness) had strong relationships with FMA ($P < .01$ each) and no relationship with MAS or with the interaction with target location. HD had the highest R^2_m of all models ($R^2_m = 0.42$, $R^2_c = 0.88$). All relations with FMA, MAS, target location, and its interactions are listed in Table 4.

When examining the models without FMA effect, MAS had marginally significant main effect ($P < .1$) for HD, BKLD, final elbow angle, and movement time. Mean movement time and velocity smoothness only had target location main effects. None of the models had a significant effect of the interaction of MAS with target location. All relations with MAS, target location, and its interactions are listed in Table 5.

Relationship of HD and BKLD with MAS are presented in Figure 10. Relationship of HD and BKLD with FMA are presented in Figure 11. We can see in Figure 10 that the relation of HD and BKLD with MAS is not consistent, as patients with MAS=2 had smaller HD or BKLD values compared to those with MAS=1+.

Table 4 : Wald chi-square values (significance levels) for LMM: Measure~FMA+(MAS+Target)^2+ (1|ID)

Measure	FMA	Target	MAS	MAS*Target	R^2_m R^2_c
HD	23.53 (***)	4.34 (-)	2.76 (-)	5.54 (-)	0.42 0.88
BKLD	12.53 (***)	52.53 (***)	3.67 (-)	12.75 (-)	0.34 0.87
FA	7.47 (**)	101.67 (***)	0.86 (-)	6.71 (-)	0.25 0.88
MT	8.79 (**)	31.80 (***)	3.84 (-)	10.85 (-)	0.29 0.95
MV	8.49 (**)	27.79 (***)	1.54 (-)	8.09 (-)	0.19 0.96
VS	7.41 (**)	37.33 (***)	3.11 (-)	9.49 (-)	0.24 0.95

*The age, days since stroke, country, gender, and the interactions with target location were not significant for any of the measures. The interaction between FMA and target was not significant as well. Abbreviations: HD- Hellinger's distance measure; BKLD- Bidirectional Kullback-Leibler divergence; FA- elbow extension final angle; MT- movement time; MV- mean velocity; VS- velocity smoothness; FMA- Fugl-Meyer Assessment; TSRT- Tonic stretch reflex threshold; Significance levels: . $P < .1$; * $P < .05$ ** $P < .01$ *** $P < .001$*

Table 5 : Wald chi-square values (significance levels) for LMM: Measure~(MAS+Target)^2+ (1|ID)

Measure	Target	MAS	MAS*Target	R ² _m R ² _c
HD	4.34 (-)	7.56 (.)	5.54 (-)	0.15 0.90
BKLD	52.53 (***)	7.15 (.)	12.75 (-)	0.17 0.87
FA	52.53 (***)	7.15 (.)	12.75 (-)	0.17 0.87
MT	31.80 (***)	7.05 (.)	10.85 (-)	0.15 0.94
MV	27.79 (***)	0.88 (-)	8.08 (-)	0.03 0.95
VS	37.33 (***)	4.99 (-)	9.49 (-)	0.11 0.95

Abbreviations: HD- Hellinger's distance measure; BKLD- Bidirectional Kullback-Leibler divergence; FA- elbow extension final angle; MT- movement time; MV- mean velocity; VS- velocity smoothness; MAS- Modified Ashworth Scale; TSRT- Tonic stretch reflex threshold; Significance levels: . P<.1; * P<.05 ** P<.01 *** P<.001

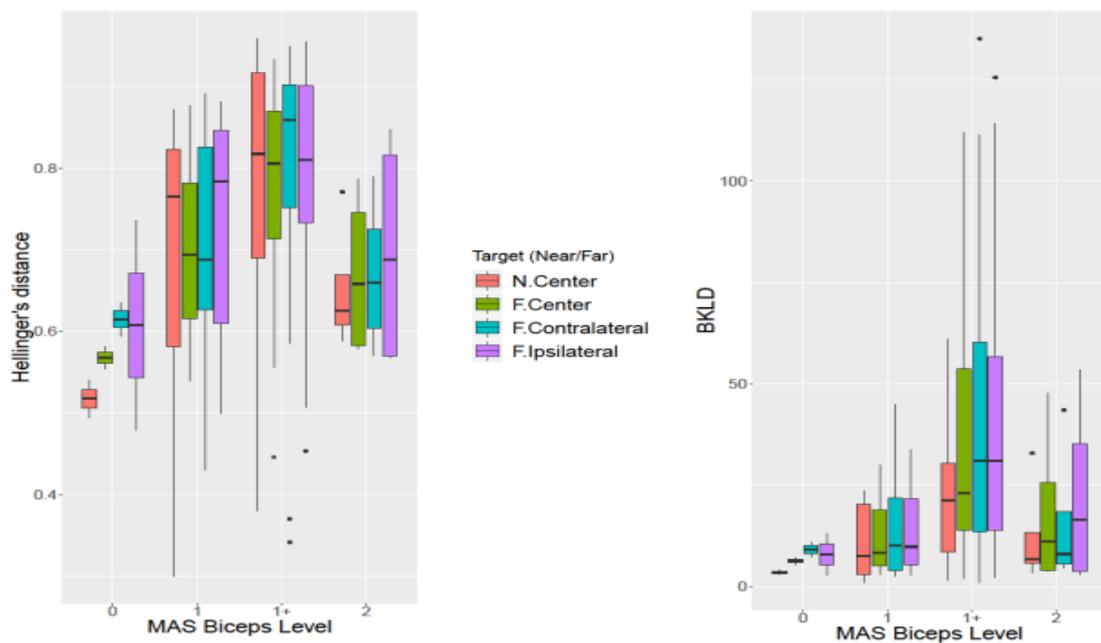


Figure 10 : Relationship of HD and BKLD with MAS. Box plots of Hellinger's distance values (left) and Bidirectional Kullback-Leibler divergence values (right) for elbow extension related to Modified Ashworth scale (MAS) scores for reaches to each target in participants with stroke.

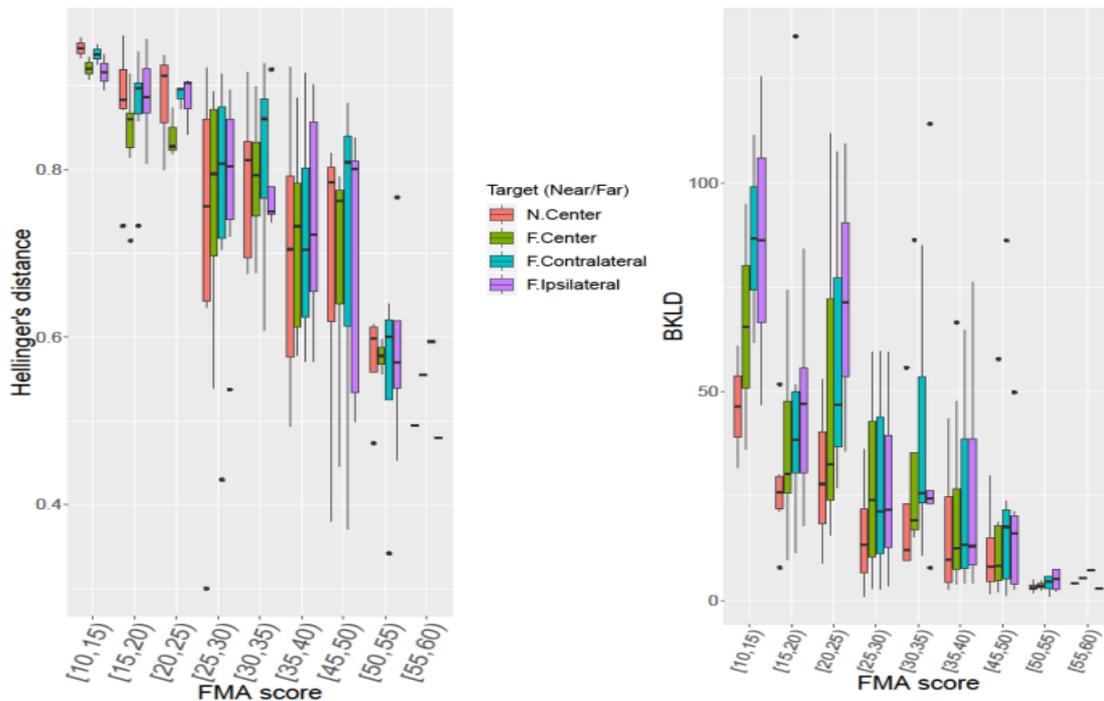


Figure 11 : Relationship of HD and BKLD with FMA. Box plots of Hellinger's distance values (left) and Bidirectional Kullback-Leibler divergence values (right) for elbow extension related to Fugl-Meyer assessment (FMA) scores for reaches to each target in participants with stroke. FMA scores between 0 and 30, 31 and 50, and 51 and 66, represent severe, moderate, and mild motor impairment, respectively.

4.2.3 FMA distribution

In a previous research of our group (Davidowitz et al., 2019) the relationship of BKLD and the kinematic measures with MAS and FMA was tested. FMA had no significant effect when explaining the BKLD values. In order to understand the reason for the lack of importance of FMA values in the models of the previous study considering our results, we conducted goodness-of-fit tests. The aim of the tests was to examine which distribution best fits the FMA values of the current study (42 values) and of the previous one (16 values). Normal, gamma, and beta distributions were examined. Results showed that beta distribution best suites the FMA values in ours and in the previous researches based on the log likelihood values (Table 6). We conducted a Kolmogorov-Smirnov test (Massey Jr, 1951) which confirmed the hypothesis - the values came from the beta distribution with the estimated parameters (KS statistic=0.10), and accordingly the current study's standard deviation (SD) was larger than the SD of the previous one. Results presented in Table 7. We conducted a Levene test (Levene, 1961) for testing equality of the variances of

the two studies and found that the variances are indeed not equal ($P < 0.05$). We then conducted a Welch test (Welch, 1947) for testing the equality of the means (under the assumption of un-equal variances) and found that the means are marginally not equal ($P = 0.05$).

Table 6 : Log likelihood values for different distributions fitted to the FMA values

Study	Beta	Normal	Gamma
Previous	38.42	36.67	37.96
Current	52.17	44.92	45.92

Table 7 : Beta parameters estimates (SD) for the FMA values

Study	shape1	shape2	Mean	SD
Previous	6.99 (1.22)	5.68 (0.98)	36.41	8.85
Current	3.37 (0.36)	3.31 (0.35)	33.30	11.43

Mean and standard deviation (SD) of the distributions were calculated by: $\mu = \frac{1}{1 + \frac{\text{shape2}}{\text{shape1}}} =$

$$\frac{\text{shape1}}{\text{shape1} + \text{shape2}} ; \text{Var}(X) = \frac{\text{shape1shape2}}{(\text{shape1} + \text{shape2})^2 (\text{shape1} + \text{shape2} + 1)} \quad (\text{Johnson et al., 1995})$$

4.2.4 TSRT and MAS

As can be seen in Figure 12, the majority of participants (28 out of 42) were classified as MAS=1+. It is also shown that the TSRT values are scattered. Shannon entropy (Shannon, 1948) was calculated for each of the measures. Shannon entropy represents the average rate at which information is produced by a stochastic source of data, and is calculated by: $S = -\sum_i P_i \log_4 P_i$. The log was calculated with 4 as the log base in order to get $0 \leq S \leq 1$ while there are 4 values of MAS. $S=1$ means no information at all (uniform distribution, all probabilities are the same), so that every sample gets its own value. $S=0$ means full information, so that the values are expected and all the samples can get the same value. For the continuous TSRT measure, the values were divided into 4 equal-length groups from the minimum to the maximum values. The results were $S_{MAS} = 0.59$; $S_{TSRT} = 0.97$.

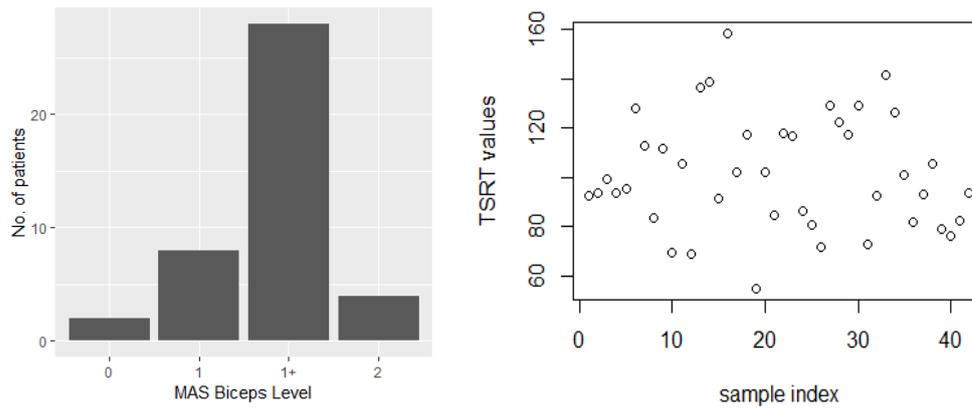


Figure 12 : Clinical spasticity measures graphs. Left panel: Histogram of the Modified Ashworth Scale (MAS) values (total of 42 participants). Right panel: Scatter plot of the Tonic Stretch reflex (TSRT) values of the participants against the sampling index.

4.3 Movement time, final angle, average velocity, and velocity smoothness

Final elbow extension angle (Figure 13A) was greater in controls than patients with stroke for all targets but the near center target ($\chi^2 = 4.51$, $P < .05$). Elbow extension in both groups was affected by target location (stroke: $\chi^2 = 103.59$, $P < .001$; controls: $\chi^2 = 209.06$, $P < .001$). Controls used less elbow extension for the near center compared with all far targets ($P < .001$ for each target) with no difference between the far targets. Similarly, participants with stroke used less elbow extension for the near center compared with all other targets ($P < .001$ for each target), whereas the other elbow ranges were not modified by target location.

Participants with stroke had longer movement times (Figure 13B) than controls for all targets ($\chi^2 = 38.48$, $P < .001$), with no differences between targets and with no interactions between targets and groups.

Average elbow velocity (Figure 13C) was higher in controls than stroke for all targets ($\chi^2 = 415.44$, $P < .001$). Controls reached the near center target with the highest average velocity (far center: $P < .01$; contralateral: $P < .01$; ipsilateral: $P < .001$) with no difference between the 3 far targets. Participants with stroke reached at different speeds to different targets ($\chi^2 = 28.00$, $P < .001$). The average velocity of patients with stroke for movement to the near center target was higher than to the ipsilateral target ($P < .001$) and marginally higher than that to the far center target ($P = .05$). The

average velocity of patients with stroke toward the ipsilateral target was the lowest (near center: $P<.001$; far center: $P<.01$; contralateral: $P<.001$).

Velocity smoothness (Figure 13D) was higher (lower number of acceleration zero-crossings) in controls than participants with stroke for all targets ($\chi^2=63.48$, $P<.001$). Reaches in controls had similar smoothness to all targets, whereas it differed according to target location in participants with stroke ($\chi^2=37.18$, $P<.001$). Average velocity smoothness for the near center target was similar to that toward the ipsilateral but lower than that for the far center target ($P<.01$) and the contralateral target ($P<.001$).

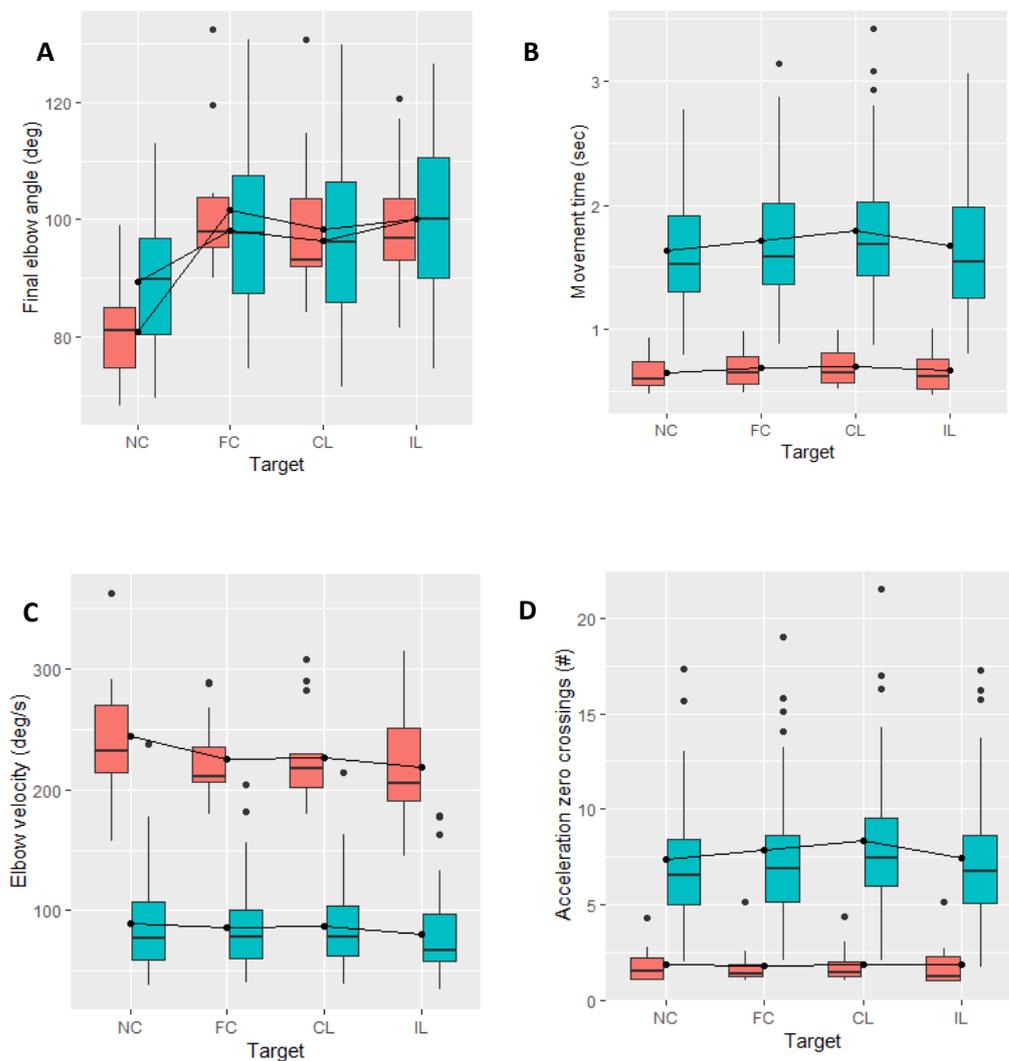


Figure 13 : Kinematic measures - box plots with line over means for (A) final elbow angle; (B) movement time; (C) elbow velocity; (D) the number of acceleration zero-crossings, per target for healthy control (red) and stroke (blue) groups. Abbreviations: NC, near-center; FC, far-center; CL, contralateral; IL, ipsilateral.

5. Discussion and conclusions

5.1 HD and BKLD

It has been shown that BKLD can be a useful tool for characterizing movement disorders (Davidowitz et al., 2019). In the current work, when examining BKLD, the results were in line with the previous finding. The BKLD values between the stroke and control groups were an order of magnitude higher than the values within the control group. Moreover, this work's results indicate that HD between participants with stroke and controls was an order of magnitude higher than the HD between healthy controls, as well. This suggests that HD may also be a good biomarker of the disruption in movement patterns in people with stroke. Furthermore, compared with the BKLD, HD has higher generalization to the population of people affected by stroke (R^2_m) and higher fit to the examined group (R^2_c).

In addition, our results show that the BKLD values for the reach-to-grasp movements of the patients with stroke to the near center target were lower than for the far targets. The results show that those movements were more similar to those of the control subjects comparing to the movements to the far targets. It is likely that the far targets required motion beyond the patient's TSRT angle into the spasticity zone, resulting in greater differences from healthy motion patterns expressed as greater BKLD values. HD values were not influenced by target location. This is likely because of the way each measure is calculated - HD using a first norm with value range of 0 to 1, and BKLD using a second norm with no upper bound. This allows BKLD values to differentiate the movement characteristic within and outside the patient's control zone. On the other hand, this makes BKLD values influenced by every sample including noises and outliers, and gives a larger weight to the tails of the distribution, especially in short movements. The first norm and the range of 0 to 1 also makes HD values easier to interpret, and less susceptible to outliers. This work shows that HD is a better biomarker of the general disruption in movement patterns in people with stroke than BKLD. However, BKLD is a better biomarker when examining the differences between the control and the spasticity zones.

Furthermore, HD was strongly related to the TSRT measure of spasticity, and to the interaction of TSRT and the slope, while BKLD was only marginally related to the slope, with no relation to the TSRT. The generalization of the models which explained the BKLD and the kinematic measures (mean movement time, average velocity, elbow final angle, and velocity smoothness) using the spasticity measures, was not very high (R^2_m values were below 0.37). In contrast, patient functional ability (FMA) and joint spasticity (TSRT and slope) strongly explained HD values for the examined group ($R^2_c=0.90$) and had the best generalizability to the larger stroke population ($R^2_m=0.55$). This shows that HD can serve as a robust, objective measure of the influence of spasticity on motor kinematics during voluntary movement. We demonstrated that HD has advantages over BKLD, and showed that HD is a better measure for quantifying the relationship between spasticity and movement disorders than BKLD.

5.2 Functional ability, joint spasticity, and muscle resistance

Generally, current clinical spasticity indexes measure biomechanical variables (e.g., resistance to passive muscle stretch) that are effects rather than causes of spasticity (Malhotra et al., 2009). This, together with the multidimensionality of both spasticity and motion, may explain why determination of the relationship between spasticity and movement disorders has, thus far, been elusive.

Our results show that MAS, the most commonly used clinical spasticity measure, did not explain the kinematic measures and our motion deficits measures when tested with the patient functional ability index (FMA). When tested without FMA, the elbow flexor passive muscle resistance (MAS) was marginally related to the HD, BKLD, final elbow angle, and mean movement time. Furthermore, the models generalization was very low ($R^2_m<0.17$). In contrast, spasticity quantified by TSRT and the slope was related to the movement disorders characterized by HD, BKLD, and the kinematics measures (mean movement time, average velocity, and velocity smoothness). This suggests that spasticity quantified by the TSRT and the slope has much better capability to explain the movement patterns of patients with stroke.

In order to examine the differences between the results of the two spasticity clinical measures, Shannon entropy was calculated. Results show that the MAS entropy was much lower than the entropy of the TSRT (MAS: 0.57; TSRT: 0.97). These results indicate that the MAS values are centered, and thus most of the patients get the same MAS score, and the measure cannot explain different levels of spasticity, comparing to the TSRT. As an explanatory variable, it is preferred that the entropy would be high, thus the distribution is not centered to one value and each patient can get a suitable value for his spasticity level.

Our results also show that FMA was necessary in order to explain all measures. This may be due to the fact that the total motor function of the upper limb is likely to impact the elbow reaching movements. However, the FMA index measures individual and combined joint movements and does not focus specifically on the elbow or a specific reaching task. Thus, the TSRT measure was significantly necessary in the models in order to explain the relationship between spasticity and reaching kinematics, and is independent of FMA. In contrast, all the models that included MAS in addition to the patient functional ability index as explanatory variables had no effect of MAS on the measures. This suggests that MAS, as a spasticity measure, does not have additional information to the patient upper limb functional ability in order to explain the voluntary movement disorders. FMA has more information than the MAS and it reflects more of the differences between subjects.

This work shows that the subjective MAS measure has low resolution and entropy, low capability of explaining different levels of spasticity, and has multicollinearity effects with FMA when modeling kinematic characteristics, whereas the TSRT is independent of the FMA. In addition, MAS only measures the resistance to a passive muscle stretch and not spasticity exclusively. We showed that TSRT is statistically superior to MAS as a clinical spasticity measure.

It is also shown that in this study FMA was necessary in order to explain all of the measures, whereas in the previous research of our group (Davidowitz et al., 2019) it was not related to the BKLD measure. This is probably due to the lower variance of the FMA distribution in the previous research compared to the variance in this research. As shown in this work - the FMA values in both researches came from the

beta distribution which has both upper and lower bounds, with SD of 11.43 in the current work (42 participants) and SD of 8.85 in the previous one (16 participants), thus making the FMA values less meaningful. In addition, most of the previous study's participants had low FMA scores (severe impairment) limiting the explanatory capability of the FMA measure. Our results suggest that the general upper limb motor function in patients with stroke affects the motor impairments of the elbow during voluntary reaching movements.

5.3 Kinematic measures

All kinematic measures (final elbow angle, mean movement time, average velocity, and velocity smoothness) differed significantly between groups (control subjects and patient with stroke). Generally, results were in line with the results of the previous study of our group which tested a smaller amount of participants (Davidowitz et al., 2019).

Final angles differed between targets and there was a significant interaction between target and group. When analyzing each group separately, the elbow extension final angle for the near center target was lower than the final angles used for all 3 far targets, which were similar to each other. This finding is in line with the significantly lower BKLD values and lower amount of GMM components, for the near center target compared to the far targets. This shows that the motion to the near center target is within the participants control zone, and therefore is more similar to that of the control group.

While final angles were affected by target location, movement times were not influenced by the target. 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 & Flash, 1995).

5.4 Using stochastic mixture models for modeling human motion

In this work, the amount of GMM components differed significantly between models of controls and of patients with stroke. In addition, patients with stroke had lower amount of Gaussian components when modeling reaching toward the near center target, where the movement is more similar to that of controls. Spasticity effects

both temporal and spatial axes, and the model reflects the combination of these two dimensions. However, the differences between the models may also be effected by the influence of motion time on smoothness. Subjects with stroke had significantly longer movements, and longer movements tend to be less smooth. It is difficult to isolate the influence of the movement time from the influence of the motor disorders.

In the present work we reinforce the conclusions drawn in the previous work of our group (Davidowitz et al., 2019), and show that stochastic mixture models, e.g. GMMs, which can incorporate time, space, and variability within a single representation, facilitating integration of specific joint variability as part of the task-specific motion model. Thus, representing an advantage over methods using global variability measures, such as Principal Component Analysis (PCA) (details regarding PCA in Appendix B - Principal component analysis (PCA)) (Pearson, 1901) or Uncontrolled Manifold (UCM) (details regarding UCM in Appendix C - Un-controlled manifold (UCM)) (Scholz & Schöner, 1999). They are therefore suitable for modeling human motion. This modeling method permits the identification of key joint motion deficits, such as limitations in elbow extension caused by elbow flexor spasticity in patients with post-stroke hemiparesis and can be used to model any movement without imposing excessive constraints on initial and final arm or joint configurations.

5.5 Future work

Elbow spasticity is the only joint analyzed, whereas patients with stroke can also have spasticity in the shoulder girdle, wrist or finger flexors. It is possible that spasticity in muscles spanning adjacent joints may have affected the elbow reaching movement. Shoulder and wrist joint centers, angles, and GMMs were calculated, and it is possible to answer this question in future studies. In addition, post and follow-up recordings were not statistically analyzed in the current work. In order to answer ENHANCE research questions it is necessary to analyze this data and examine the differences between the results of pre, post, and follow-up of the ENHANCE training.

References

- Abdi, H., & Williams, L. (2010). Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4), 433-459.
- Altman, N. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 46(3), 175-185.
- Bartle, R., & Bartle, R. (1995). *The elements of integration and Lebesgue measure*. New York: Wiley (Vol. 27).
- Bishop, C. (2006). *Pattern recognition and machine learning*. Springer.
- Bohannon, R., & Smith, M. (1987). Interrater reliability of a modified Ashworth scale of muscle spasticity. *Physical therapy*, 67(2), 206-207.
- Boiteau, M., Malouin, F., & Richards, C. (1995). Use of a hand-held dynamometer and a Kin-Com dynamometer for evaluating spastic hypertonia in children: a reliability study. *Physical therapy*, 75(9), 796-802.
- Bro, R., & Smilde, A. (2014). Principal component analysis. *Analytical Methods*, 6(9), 2812-2831.
- Burridge, J., Wood, D., Hermens, H., Voerman, G., Johnson, G., Wijck, F., & Pandyan, A. (2005). Theoretical and methodological considerations in the measurement of spasticity. *Disability and rehabilitation*, 27(1-2), 69-80.
- Calinon, S., Guenter, F., & Billard, A. (2007). On learning, representing, and generalizing a task in a humanoid robot. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(2), 286-298.
- Calota, A., & Levin, M. F. (2009). Tonic stretch reflex threshold as a measure of spasticity: implications for clinical practice. *Topics in stroke rehabilitation*, 16(3), 177-188.
- Calota, A., Feldman, A., & Levin, M. (2008). Spasticity measurement based on tonic stretch reflex threshold in stroke using a portable device. *Clinical Neurophysiology*, 119(10), 2329-2337.
- Charalambous, C. (2014). *Interrater reliability of a modified Ashworth scale of muscle spasticity*. London: In *Classic Papers in Orthopaedics*. Springer.
- Cheung, M. (2013). Implementing restricted maximum likelihood estimation in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 20(1), 157-167.
- Cnaan, A., Laird, N., & Slasor, P. (1997). Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Statistics in medicine*, 16(20), 2349-2380.

- Cutler, A., & Cordero-Brana, O. (1996). Minimum Hellinger's distance estimation for finite mixture models. *Journal of the American Statistical Association*, 91(436), 1716-1723.
- Davidowitz, I., Parmet, Y., Frenkel-Toledo, S., Baniña, M., Soroker, N., Solomon, J., & Berman, S. (2019). Relationship Between Spasticity and Upper-Limb Movement Disorders in Individuals With Subacute Stroke Using Stochastic Spatiotemporal Modeling. *Neurorehabilitation and neural repair*, 33(2), 141-152.
- de Freitas, S., & Schol, J. (2010). A comparison of methods for identifying the Jacobian for uncontrolled manifold variance analysis. *Journal of biomechanics*, 43(4), 775-777.
- Dempster, A., Laird, N., & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1), 1-22.
- Diebel, J. (2006). Representing attitude: Euler angles, unit quaternions, and rotation vectors. *Matrix*, 58(15-16), 1-35.
- Dinov, I. (2008). Expectation maximization and mixture modeling tutorial.
- Dobkin, B. H. (2005). Rehabilitation after stroke. *New England Journal of Medicine*, 352(16), 1677-1684.
- Duncan, P., Goldstein, L., Horner, R., Landsman, P., Samsa, G., & Matchar, D. (1994). Similar motor recovery of upper and lower extremities after stroke. *Stroke*, 25(6), 1181-1188.
- Feldman, A. (2015). Referent control of action and perception. *Challenging Conventional Theories in Behavioral Neuroscience*.
- Fellows, S., Kaus, C., & Thilmann, A. (1994). Voluntary movement at the elbow in spastic hemiparesis. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 36(3), 397-407.
- Francisco, G., Hu, M., Boake, C., & Ivanhoe, C. (2005). Efficacy of early use of intrathecal baclofen therapy for treating spastic hypertonia due to acquired brain injury. *Brain injury*, 19(5), 359-364.
- Fugl-Meyer, A., Jaasko, L., Leyman, I., Olsson, S., & Steglind, S. (1975). The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. *Scandinavian journal of rehabilitation medicine*, 7(1), 13-31.
- Goldberger, J., & Aronowitz, H. (2005). A distance measure between GMMs based on the unscented transform and its application to speaker recognition. *In Ninth European Conference on Speech Communication and Technology*.
- Gowland, C., Stratford, P., Ward, M., Moreland, J., Torresin, W., Van Hullenaar, S., & Plews, N. (1993). Measuring physical impairment and disability with the ChedokeMcMaster Stroke Assessment. *Stroke*, 24(1), 58-63.

- Gregson, J., Leathley, M., Moore, A., Sharma, A., Smith, T., & Watkins, C. (1999). Reliability of the Tone Assessment Scale and the modified Ashworth scale as clinical tools for assessing post stroke spasticity. *Archives of physical medicine and rehabilitation*, 80(9), 1013-1016.
- Gund, B., Jagtap, P., Ingale, V., & Patil, R. (2013). Stroke: A brain attack. *IOSR Journal of Pharmacy*, 3(8), 01-23.
- Hershey, J., & Peder, A. (2007). Approximating the Kullback Leibler Divergence Between Gaussian Mixture Models. 317-20.
- Hershey, J., Olsen, P., & Rennie, S. (2007, December). Variational Kullback-Leibler divergence for hidden Markov models. *In 2007 IEEE Workshop on Automatic Speech Recognition & Understanding (ASRU)*, (pp. 323- 328).
- Heuschling, A., Gazagnes, M., & Hatem, S. (2013). Stroke: from primary care to rehabilitation. *Revue medicale de Bruxelles*, 34(4), 205-210.
- Higgins, J., Mayo, N., Desrosiers, J., Salbach, N., & Ahmed, S. (2005). Upper-limb function and recovery in the acute phase poststroke. *Journal of Rehabilitation Research & Development*, 42(1).
- Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A., Jaitly, N., & Sainath, T. (2012). Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal processing magazine*, 29.
- Hyun, M., Choi, J., & Kwak, N. (2019). Disentangling Options with Hellinger Distance Regularizer. *arXiv preprint arXiv*, 1904.06887.
- Jensen, J., Ellis, D., Christensen, M., & Jensen, S. (2007). Evaluation distance measures between Gaussian mixture models of mfccs.
- Jobin, A., & Levin, M. (2000). Regulation of stretch reflex threshold in elbow flexors in children with cerebral palsy: a new measure of spasticity. *Developmental medicine and child neurology*, 42(8), 531-540.
- Johnson, N., Kotz, S., & Balakrishnan, N. (1995). beta distributions. *Continuous Univariate Distributions*, 2.
- Kelso, J. (2008). Synergies. *Scholarpedia*, 3(10), 1611.
- Komaty, A., Boudraa, A., Augier, B., & Dare-Emzivat, D. (2013). EMD-based filtering using similarity measure between probability density functions of IMFs. *IEEE Transactions on Instrumentation and Measurement*, 63(1), 27-34.
- Kristan, M., Leonardis, A., & Skocaj, D. (2011). Multivariate online kernel density estimation with Gaussian kernels. *Pattern Recognition*, 44(10-11), 2630-2642.
- Kullback, S., & Leibler, R. (1951). On information and sufficiency. *The annals of mathematical statistics*, 22(1), 79-86 .

- Lance, J. W. (1980). Pathophysiology of spasticity and clinical experience with baclofen. *Spasticity: disordered motor control*, 185-204.
- Langhorne, P., Coupar, F., & Pollock, A. (2009). Motor recovery after stroke: a systematic review. *The Lancet Neurology*, 8(8), 741-754.
- Latash, M., & Anson, J. (2006). Synergies in health and disease: relations to adaptive changes in motor coordination. *Physical therapy*, 86(8), 1151-1160.
- Lebesgue, H. (1902). Intégrale, longueur, aire. *Annali di Matematica Pura ed Applicata*, (1898-1922), 7(1), 231-359.
- Lee, K., Carson, L., Kinnin, E., & Patterson, V. (1989). The Ashworth scale: a reliable and reproducible method of measuring spasticity. *Journal of Neurologic Rehabilitation*, 3(4), 205-209.
- Levene, H. (1961). Robust tests for equality of variances. *Contributions to probability and statistics. Essays in honor of Harold Hotelling*, 279-292.
- Levin, M., Banina, M., & Frenkel-Toledo, S. (2018). Personalized upper limb training combined with anodal-tDCS for sensorimotor recovery in spastic hemiparesis: study protocol for a randomized controlled trial. *Trials*, 19(1):7.
- Lindsay, B. (1994). Efficiency versus robustness: the case for minimum Hellinger distance and related methods. *The annals of statistics*, 22(2), 1081-1114.
- Lindsay, B. (1994). Efficiency versus robustness: the case for minimum Hellinger distance and related methods. *The annals of statistics*, 22(2), 1081-1114.
- Malhotra, S., Pandyan, A., Day, C., Jones, P., & Hermens, H. (2009). Spasticity, an impairment that is poorly defined and poorly measured. *Clinical rehabilitation*, 23(7), 651-658.
- Massey Jr, F. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253), 68-78.
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in ecology and evolution*, 4(2), 133-142.
- Ni, W., Sun, F., Weng, G., & Xu, L. (2013, November). A Hellinger Distance Based Anonymization Method for Weighted Social Networks. . In *2013 10th Web Information System and Application Conference* (pp. 259-264). IEEE.
- Nichols-Larsen, D., Clark, P., Zeringue, A., Greenspan, A., & Blanton, S. (2005). Factors influencing stroke survivors' quality of life during subacute recovery. *Stroke*, 36(7), 1480-1484.
- Nordmark, E., & Andersson, G. (2002). Wartenberg pendulum test: objective quantification of muscle tone in children with spastic diplegia undergoing selective dorsal rhizotomy. *Developmental Medicine and Child Neurology*, 44(1), 26-33.

- O'Brien, J., Bodenheimer Jr, R., Brostow, G., & Hodgins, J. (1999). Automatic joint parameter estimation from magnetic motion capture data. *Georgia Institute of Technology*.
- Pandyan, A., Johnson, G., Price, C., Curless, R., Barnes, M., & Rodgers, H. (1999). A review of the properties and limitations of the Ashworth and modified Ashworth Scales as measures of spasticity. *Clinical rehabilitation*, 13(5), 373-383.
- Party, I. (2012). *National clinical guideline for stroke (Vol. 20083)*. London: Royal College of Physicians.
- Pearson, K. (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(11), 559-572.
- Polani, D. (2013). *Kullback-leibler divergence*. New York: Springer .
- Powers, R., Marder-Meyer, J., & Rymer, W. (1988). Quantitative relations between hypertonia and stretch reflex threshold in spastic hemiparesis. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 23(2), 115-124.
- Press, W., Teukolsky, S., Vetterling, W., & Flannery, B. (2007). *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press.
- Puzi, A., Sidek, S., Rosly, H., Daud, N., & Yusof, H. (2017, November). Modified Ashworth Scale (MAS) Model based on Clinical Data Measurement towards Quantitative Evaluation of Upper Limb Spasticity. *In IOP Conference Series: Materials Science and Engineering* (pp. Vol. 260, No. 1, p. 012024). IOP Publishing.
- Rasmussen, C. (2000). The infinite Gaussian mixture model. *In Advances in neural information processing systems* , 554-560.
- Satterthwaite, F. (1946). An approximate distribution of estimates of variance components. *Biometrics bulletin*, 2(6), 110-114.
- Scholz, J., & Schöner, G. (1999). The uncontrolled manifold concept: identifying control variables for a functional task. *Experimental brain research*, 126(3), 289-306.
- Sengar, H., Wang, H., Wijesekera, D., & Jajodia, S. (2008). Detecting VoIP floods using the Hellinger distance. *IEEE transactions on parallel and distributed systems*, 19(6), 794-805.
- Shannon, C. (1948). A mathematical theory of communication. *Bell system technical journal*, 27(3), 379-423.
- Simpson, D. (1987). Minimum Hellinger distance estimation for the analysis of count data. *Journal of the American statistical Association*, 82(399), 802-807.
- Sommerfeld, D., Eek, E., & Svensson. (2004). Spasticity after stroke. *Stroke*, 35(1), 134-139.

- Specht, D. (1991). A general regression neural network. *IEEE transactions on neural networks*, 2(6), 568-576.
- Tamura, R., & Boos, D. (1986). Minimum Hellinger distance estimation for multivariate location and covariance. *Journal of the American Statistical Association*, 81(393), 223-229.
- Upper limb management after stroke fact sheet*. (n.d.). Retrieved from Stroke Foundation: <https://strokefoundation.org.au/About-Stroke/Help-after-stroke/Stroke-resources-and-fact-sheets/Upper-limb-management-after-stroke-fact-sheet>
- 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.
- Welch, B. (1947). The generalization of student's problem when several different population variances are involved. *Biometrika*, 34(1/2), 28-35.
- Weller-Fahy, D., Borghetti, B., & Sodemann, A. (2014). A survey of distance and similarity measures used within network intrusion anomaly detection. *IEEE Communications Surveys & Tutorials*, 17(1), 70-91.
- Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3), 37-52.
- Zhang, L., Schallert, T., Zhang, Z., Jiang, Q., Arniago, P., Li, Q., & Chopp, M. (2002). A test for detecting long-term sensorimotor dysfunction in the mouse after focal cerebral ischemia. *Journal of neuroscience methods*, 117(2), 207-214.
- Zhao, Q., Hautamaki, V., & Fränti, P. (2008, October). Knee point detection in BIC for detecting the number of clusters. . In *International conference on advanced concepts for intelligent vision systems* (pp. 664-673). Berlin, Heidelberg: Springer.

Appendix A - Lebesgue measure

The Lebesgue measure (Lebesgue, 1902) is the standard way of assigning a measure to subsets of n -dimensional Euclidean space. For $n=1, 2, \text{ or } 3$, it coincides with the standard measure of length, area, or volume. It is used throughout real analysis, in particular to define Lebesgue integration.

Given a subset $E \subset \mathbb{R}$, with the length of interval $I=[a,b]$ given by $l(I) = b - a$, the Lebesgue $\lambda(E)$ is defined as

$$\lambda(E) = \inf \left\{ \sum_{k=1}^{\infty} l(I_k) : (I_k)_{k \in \mathbb{N}} \text{ is a sequence of intervals with open boundaries with } E \subset \bigcup_{k=1}^{\infty} I_k \right\} \quad (\text{A.1})$$

The Lebesgue measure is defined on the Lebesgue σ -algebra, which is the collection of all sets E which satisfy that for every $A \subset \mathbb{R}$, $\lambda(A) = \lambda(A \cap E) + \lambda(A \cap E^c)$.

Appendix B - Principal component analysis (PCA)

Principal Component Analysis (PCA) is a statistical method for data analysis (Pearson, 1901). PCA aims to compress the size of a data set and evaluate it in terms of the main components that capture the essential data patterns of the data set. To achieve this goal, PCA computes a new set of latent variables called Principal Components (PC) (Wold, Esbensen, & Geladi, 1987). PCA is the simplest and most popular method to perform dimensionality reduction (DR) effectively- the process of reducing the number of random variables. Using PCA, the original space is reduced (with data loss, but hopefully retaining the most important variance) to the new space.

The PCs 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. DR with PCA achieves the best mean-square error compared to other linear DR methods (Bro & Smilde, 2014).

The transformation that is done to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables, the PC, is defined in such a way that the first principal component is required to have the largest amount of variance and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. Larger amount of variance implies that the component contains more information compared to other components. The components are computed in this manner, until the PC contains most of the information of the data, as defined by an empirical demand - a set proportion of the data variability (a threshold). There are a few more methods to determine when to stop adding components – according to the predicted residual sum of squares or to a plot of the size of the eigenvalues. Another standard tradition is to keep only the components whose eigenvalue is larger than the average of the eigenvalues (Abdi & Williams, 2010).

Let the matrix X be defined as a column matrix $X = \{x_1, x_2, \dots, x_d\}$, where d is the original space dimensionality and x_i denotes a set of observations of the variable i in the original data space. First, the columns of X are centered by subtracting the

means- \bar{x}_i , from each column i for each $i = \{1, \dots, d\}$. Now let \bar{X} be the centered matrix.

After that, covariance matrix of the data is calculated by:

$$C = E(\bar{X}\bar{X}^T) \quad (\text{B.1})$$

The matrix of eigenvalues λ_j , and corresponding eigenvectors w_j are calculated as follows: $Cw_j = \lambda_j w_j$. The value λ_j denotes the importance of component j , and w_j denotes a weight vector. The eigenvalues λ_j are sorted in descending order- from the most important component to the least. Matrix W composed of the columns w_j , where the columns are ordered corresponding to the order of the eigenvalues. This matrix can be used to project the original features into the latent space encoded by the latent variables as follows:

$$Y = W \cdot X \quad (\text{B.2})$$

Where Y contains the new latent variables ε_j , $j = \{1, \dots, d\}$. The principal components are the set of ε_j , $j = \{1, \dots, p\}$ where $\sum_{j=1}^p \lambda_j > T1$.

$T1$ is set to be the threshold for data variability explanation. If $T1=100\%$ than all components will be selected and dimensionality will not be reduced. The number of principal components is less than or equal to the smaller of the number of original variables or the number of observations, the new space contains equal or less dimensions of the original dataset space.

PCA is simple to apply; takes polynomial computation time and the linear transformation applied allow re-projecting the data to the original space. PCA applied on motion data shows inconsistencies in the latent spaces created when different coordinate systems are used (Calinon, Guenter, & Billard, 2007).

Appendix C - Un-controlled manifold (UCM)

The Uncontrolled Manifold (UCM) is a motor coordination theory. The theory suggests a hypothesis about the human central nervous system and how it achieves motor coordination. UCM theory does not eliminate the redundant degrees of freedom which exist in motion, but instead it uses all of them to ensure flexible and stable performance of motor tasks. The central nervous system makes use of this abundance from the redundant systems instead of restricting them like other hypothesized. The concept can be described using synergies (Scholz & Schöner, 1999).

Any ordinary human activity requires cooperation among many structurally diverse elements. One hypothesis claims that in such complex living systems the elements are organized into synergies (also known as coordinative structures) defined as functional groupings of structural elements (e.g. neurons, muscles, joints) that are temporarily constrained to act as a single coherent unit (Kelso, 2008).

Synergies relates to two types of variables – elemental variables and performance variables. Elemental variables are the smallest sensible variables that can be used to describe the system mechanics. On the other hand, performance variables refer to important variables produced by the system as a whole, some might be goal or target related. According to the UCM hypothesis, the controller (the human brain) acts in the space of elemental variables (for example-the 7 major rotations shared by the shoulder, elbow, and wrist joints), and selects in that space sets of values corresponding to a required value of a performance variable. This is the Uncontrolled Manifold- Little to no control over these elements is required since they do not affect the performance variables.

The UCM subspace is orthogonal to the subspace consisting of all the variables which affect the performance variables. Most of the variability among the elemental variables is limited to the UCM, which allows flexibility in the performance. The UCM hypothesis suggests that variability can be “bad”- affecting an important performance variable and causing larger errors, or “good”- variability within the UCM which is maintaining a successful outcome with more motion flexibility (Latash

& Anson, 2006). In order to accept the UCM hypothesis, the “Good” variability in the UCM (referred to as V_{GOOD}) needs to be greater than the variability orthogonal to it (referred to as V_{BAD}). If $V_{GOOD} \geq V_{BAD}$ it can be concluded that a synergy exists, stabilizing a performance variable for which the UCM was computed. In contrast, in a case where $V_{BAD} \gg V_{GOOD}$, the variability indicates that abundance in DOF's was used for achieving performance variable goal. The UCM is not a linear subspace and therefore the UCM analysis for joint configurations requires derivation of the Jacobian which will be used for linear estimation of the UCM. Using regression for estimating the Jacobian is preferable over analytical approach (de Freitas & Schol, 2010). UCM analysis can also be applied in order to achieve effects of DR methods.

UCM requires the analysis of the Jacobian matrix, which is not always possible to define and depends on the data. If possible, the UCM method result can be stronger than the PCA since it allows more complex tasks and flexibility in the movement instead of simply disregarding the redundant degrees of freedom as in the PCA algorithm.

Appendix D - Joint Angles definitions

Joint angles were defined according to Table 8 : Joint angles definition. All the seven angles detailed below were computed for the ENHANCE project, yet only the elbow extension angle was analyzed in this project.

Table 8 : Joint angles definition

Name	Explanation	Units
Shoulder 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 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 Rotation	Angle around the shoulder elbow vector X_S . 0 according to rest position, positive direction external.	Deg
Elbow Extension	Angle between upper arm marker X_{M3} and lower arm marker X_{M2} . 180 deg- Full arm extension.	Deg
Elbow Supination	Angle around wrist elbow vector X_E . 0 is according to rest position. Positive direction external.	Deg
Wrist 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 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

תקציר

ספסטיות הינה הפרעה מוטורית שכיחה הנובעת משבץ מוחי. התופעה מאופיינת בעלייה תלויה מהירות ברפלקס המתיחה. קיימים מדדים קליניים לבדיקת קיום של ספסטיות ולהערכת פגיעות מוטוריות, עם זאת, הקשר בין ההפרעות בתנועה רצונית לבין ספסטיות אינו ברור לחלוטין. חלקית בגלל המורכבות, הרב-ממדיות (מרחב וזמן) של הפגישה ובגלל השונות המובנית הקיימת בתנועה. בעבודה קודמת של קבוצתנו הוצע מדד חדשני, המבוסס על מידול סטוכסטי של תנועה בזמן ובמרחב, לבחינת הקשר הזה. ההפרעות בתנועה כומתו באמצעות המרחק בין מודלי התערובת הגאוסיאנים, שנבנו על פי מסלולי תנועה של נבדקים לאחר שבץ לבין המודלים של נבדקי בקרה בריאים. מרחק זה הוערך באמצעות מדד שונות קאלבק-לייבלר בגרסתו הסימטרית. בעבודה הנוכחית אנו מחזקים ומוסיפים לתוצאות והמסקנות שנבעו מהעבודה הקודמת, תוך שימוש בכמעט פי שלושה נבדקים. אנו מציעים מדד מרחק שונה: מדד המרחק של הלינגר ומשווים אותו לשונות קאלבק-לייבלר. אנו מראים כי במדד הלינגר מרחק גדול יותר בין המודלים של התנועה קשור ברמה גבוהה יותר של ספסטיות של החולה. למרחק הלינגר יש יתרונות על פני שונות קאלבק-לייבלר. המדד מקיים את אי שוויון המשולש ובעל ערכים תחומים בין 0 ל-1. המרחק של הלינגר סימטרי, פשוט לפרשנות, פחות חשוף להשפעות דגימות חריגות, פחות אינטנסיבי חישובית וידוע כאינווריאנטי, יציב, נורמלי אסימפטוטית וחסין, בהשוואה למדד קאלבק-לייבלר ולמדדי מרחק בין התפלגויות אחרים.

הניתוח שבוצע בעבודה הנוכחית כלל 13 נבדקים בריאים ו-42 נבדקים לאחר שבץ אשר ביצעו תנועות הושטה לאחיזה לעבר 4 מטרות. תנועות הזרוע בזמן ההושטה הוקלטו באמצעות חיישנים אלקטרומגנטיים. ספסטיות המרפק כומתה באמצעות מדד סף רפלקס המתיחה במנוחה, שיפוע קו הרגרסיה המייצג את התלות במהירות של הספסטיות ומדד סולם אשוורת' המותאם (סולם אשר הינו מדד מקובל לספסטיות). התפקוד המוטורי של הגפיים העליונות כומת באמצעות הערכת פוגל-מאייר. התוצאות מראות כי המרחק של הלינגר מתקשר באופן חזק למדד סף רפלקס המתיחה במנוחה, לשיפוע, לאינטראקציה שלהם ולמדד התפקוד המוטורי של הגפיים העליונות, וכי הוא בעל יכולת ההכללה הטובה ביותר לאוכלוסיית השבץ הכללית, בהשוואה למדד שונות קאלבק-לייבלר ולמדדים הקינמטיים שנבדקו. מדד שונות קאלבק-לייבלר התקשר למדד התפקוד המוטורי, התקשר גבולית לשיפוע קו הרגרסיה וללא קשר למדד סף רפלקס המתיחה במנוחה. לפיכך, המרחק של הלינגר יכול לשמש מדד חסין ואובייקטיבי לקשר של ספסטיות לתנועות ההושטה לאחיזה. התוצאות הראו גם כי מדד סף רפלקס המתיחה במנוחה עדיף סטטיסטית על סולם אשוורת' המותאם כמדד ספסטיות קליני.

מילות מפתח: מדד המרחק של הלינגר, שונות קאלבק-לייבלר, שיקום לאחר שבץ, ספסטיות, מודל תערובת גאוסיאנים.

אוניברסיטת בן-גוריון בנגב
הפקולטה למדעי ההנדסה
המחלקה להנדסת תעשייה וניהול

שימוש במדד המרחק של הלינגר לכימות השפעת ספסטיות לאחר שבץ על בקרת תנועה רצונית

חיבור זה מהווה חלק מהדרישות לקבלת תואר מגיסטר בהנדסה

מאת : הדר לקריץ

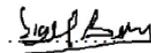
מנחים : פרופ' סיגל ברמן, פרופ' ישראל פרמט

תאריך ..25.9.2019



חתימת המחבר... הדר לקריץ...

תאריך ..25.9.2019



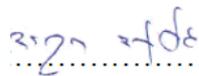
אישור המנחים... סיגל ברמן...

תאריך ..25.9.2019



...ישראל פרמט.....

תאריך ..25.09.2019



אישור יו"ר ועדת תואר שני מחלקתית.....

אוניברסיטת בן-גוריון בנגב
הפקולטה למדעי ההנדסה
המחלקה להנדסת תעשייה וניהול

שימוש במדד המרחק של הלינגר לכימות השפעת ספסטיות לאחר שבץ על בקרת תנועה רצונית

חיבור זה מהווה חלק מהדרישות לקבלת תואר מגיסטר בהנדסה

מאת: הדר לקריץ

ספטמבר 2019

אלול ה'תשע"ט