Automated Pose Estimation and Assessment of Dynamic Knee Valgus and Risk of Knee Injury during the Single Leg Squat

by

Rezvan Kianifar

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2017

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

The text for the pose estimation part of the related work (Section 2.1.2) was partially adapted from a course report completed for ME780. Also, part of the material in Chapter 3 regarding the Extended Kalman Filter tuning (Section 3.2.3) was originally developed for the ECE780 course project.

The text for the pilot study sections in Chapter 5 is adapted from the paper published in IEEE Engineering in Medicine and Biology Conference 2016 entitled "Classification of Squat Quality with Inertial Measurement Units in the Single Leg Squat Mobility Test" [48]. This paper originated as a course project in STAT 841, and was enriched later for publication. Dr. Alex Lee and Mr. Sachin Raina were coauthors and clinical collaborators of this research who contributed with data collection, discussion, and manuscript review. Professor Dana Kulić was the thesis advisor and contributed with ideas, discussion, and manuscript review.

The text for the single leg squat assessment part of the related work and the clinical study part in Chapter 5 are adapted from a manuscript entitled "Automated Assessment of Dynamic Knee Valgus and Risk of Knee Injury During the Single Leg Squat", which has been submitted to the Journal of Translational Engineering in Health and Medicine [2]. This paper started as a course project for the fulfilment of ECE657 requirements. It was enriched with further analysis and discussion and submitted for publication. Dr. Alex Lee and Mr. Sachin Raina were coauthors. Dr. Lee contributed with clinical data assessment and labelling, ideas, discussion, and reviewing the manuscript. Mr. Raina performed the clinical data collection and contributed with ideas and reviewing the work. Professor Dana Kulić was the thesis advisor and contributed with the development of ideas, discussion, and manuscript review.

Abstract

Many clinical assessment protocols rely on the evaluation of functional movement tests such as the Single Leg Squat (SLS), which are often assessed visually. Visual assessment is subjective and depends on the experience of the clinician. Developing a reliable automatic human motion tracking and assessment system can improve the accuracy of SLS clinical assessments and provide objective results that can be tracked and monitored over time to guide rehabilitation and determine an individual's response to an intervention. In this study, an Inertial Measurement Unit (IMU) based method for automated assessment of squat quality is proposed to provide clinicians with a quantitative measure of SLS performance.

First, an automated pose estimation method is applied to SLS motion data. A set of three IMUs is used to estimate the joint angles, velocities and accelerations of the squatting leg. To tackle noisy sensor measurements and gyro drift, a 7 degree of freedom (DOF) kinematic model of the lower leg was applied together with a constant acceleration assumption to approximate the angular velocity and linear acceleration at each sensor location. The kinematic model predictions of the angular velocity and linear acceleration and sensor measurements were fused via an Extended Kalman Filter (EKF). The position, velocity, and acceleration of each DOF were defined as the states to be estimated by the EKF. The pose estimation results showed successful extraction of joint angles with an average RMS error of 3.2°, 5.5°, 7° compared to joint angles estimated from motion capture for the ankle, knee, and hip joints, respectively. For this estimation, the required parameters for the kinematic model, including information about the sensor placement and orientation as well as the kinematic link lengths, were extracted from the marker data.

However, in clinical applications of the proposed method, when marker data is not available, these parameters need to be measured. Measuring these parameters is time consuming in the clinical setting, which limits application of IMUs for clinical purposes. With the motivation to make this procedure easier and faster, a method for approximating the parameters using placement assumptions and body measures was described. A sensitivity analysis was performed to detect those parameters which most affect pose estimation accuracy. The sensitivity analysis results revealed that sensor orientation is the most critical factor for accurate pose estimation. In this thesis, a simple and easy to use method is proposed for sensor orientation calibration, based on a systematic placement of sensors and using gyroscope information for orientation estimation. This protocol was evaluated experimentally and pose estimation error with approximated parameters before and after applying the calibration protocol were compared. The comparison results showed that the estimate of the sensor orientation increases the pose estimation accuracy by 6.5° for the

knee joint angle and with an average of 1.8° for other joints without the need for time consuming calibration.

In the second part of the thesis, an algorithm for automated assessment of the SLS in terms of dynamic knee valgus and risk of knee injury is developed. After applying the pose estimation algorithm to IMU data of SLS motions, the estimated time series data of joint angles, velocities and accelerations for consecutive squats were segmented into individual squat repetitions. Statistical time domain features were generated from each repetition. The most informative features were selected using a combination of 18 feature selection techniques. Six common classifiers in including SVM, Linear Multinomial Logistic Regression, Decision Tree, Naïve Bayes, K Nearest Neighbourhood, and Random Forests were applied to the full dimensional data, the subset of selected features, and extracted features by supervised principal component analysis. The proposed approach was evaluated in two trials. First, a pilot study was conducted on a small dataset, followed by analysis on a larger clinical data set, collected by our clinical collaborator. For the clinical study, a dataset of SLS performed by healthy participants was collected and labelled by three expert clinical raters using two different labelling criteria: "observed amount of knee valgus" and "overall risk of injury". Labels included "good", "moderate", and "poor" squat quality or "high risk", "mild to moderate risk", and "no risk" of injury. Feature selection results showed that both flexion at the hip and knee, as well as hip and ankle internal rotation are discriminative features, and that participants with "poor" squats bend the hip and knee less than those with better squat performance. Furthermore, improved classification performance was achieved by training separate classifiers stratified by gender. Classification results showed excellent accuracy, 93.1% for classifying squat quality as "poor" or "good" and 95.3% for differentiating between high and no risk of injury.

Acknowledgements

I would like to thank my supervisor, Professor Dana Kulić for her patient guidance, encouragement, immense knowledge, insightful advice, and support. Her positive outlook and motivation inspired me and gave me confidence. Her careful editing contributed enormously to the production of this thesis. Thank you for prompted response to all my emails, dedicated weekly meetings, reviewing my writings, being so supportive when things did not go well, and smiling all the time. I could not have imagined having a better supervisor for my MAS study.

My sincere thanks also goes to my clinical collaborator MSK Metrics, Dr. Alex Lee, Mr. Sachin Raina, and Ms. Wendy Porter for partial funding of the project, spending hours of their time with me in Motion Capture Laboratory for data collection, providing me with high quality clinical data, and asking insightful questions. I benefited from their advice, experience, and support at many stages of this project.

Thanks to my thesis reviewers, Prof.Karray and Prof.Tung for the time they spent to read my thesis and provide me with their insightful comments. I have also benefited from their knowledge and advice through participating in their classes and completing some parts of my research through their course project. I would also like to thank other UW professors, Prof.Ghodsi, Prof.Heunis, and Prof.Crowley for their amazing courses which helped me in enriching the required background for this thesis. Also, thanks to all ECE staff who helped me along this journey.

Thanks to all my fellow graduate students in Adaptive Systems Lab for volunteering in my study and for sharing their knowledge with me during the group meetings. My special thanks goes to Jonathan Lin for his generous help in data collection at Motion Capture Lab, being available for discussion and making me feel confident about what I was doing.

Thanks to Canada's Natural Science and Engineering Research Council (NSERC) for their financial support which allowed me to undertake this research.

Finally, I would like to thank my family, particularly my husband for their unconditional love and support.

Dedication

To my husband Ali.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACL	Anterior Cruciate Ligament
Add/Abd	Adduction/Abduction
ASIS	Anterior Superior Iliac Spine
CV	Cross Validation
DH	Denavit Hartenberg
DKV	Dynamic Knee Valgus
DOF	Degree of Freedom
DT	Decision Tree
EKF	Extended Kalman Filter
Flex/Ext	Flexion/Extension
IMU	Inertial Measurement Unit
IR/ER	Internal Rotation/External Rotation
IRR	Inter Rater Reliability
KNN	K Nearest Neighborhood
LL	Leg Length
LMLR	Linear Multinomial Logistic Regression
LOSO	Leave One Subject Out
MAD	Mean Absolute Deviation
MOCAP	Motion Capture
NB	Naïve Bayes
PD	Pelvic Depth
PSIS	Posterior Superior Iliac Spine
PW	Pelvic Width
RMS	Root Mean Square
ROM	Range of Motion
SLS	Single Leg Squat
SPCA	Supervised Principle Component Analysis
STD	Standard Deviation
SVM	Support Vector Machine
VAR	Variance
ZVC	Zero-Velocity Crossing

Chapter 1

Introduction

1.1 Motivation

Many clinical assessment protocols rely on functional movement tests, where the patient is asked to perform a target movement while the clinician observes and assesses the movement in terms of balance, stability, dynamic alignment and motion coordination [47]. The single leg squat (SLS) is an example of a functional movement test, and is commonly used in rehabilitation, sports medicine and orthopedic settings [39]. Many daily or athletic activities share similarities with the SLS, making it a powerful test to identify compromised muscle function [23]. Correct performance of the SLS can provide an indication of knee function and assessment of recovery. An important component of the rating of the quality of a performed SLS is the degree of inward movement of the knee, known as medial knee displacement or dynamic knee valgus (DKV), as shown in Figure 1.1.

DKV correlates with non-contact Anterior Cruciate Ligament (ACL) injury and patello femoral pain and is believed to be related to increased lower extremity risk of injury [67]. ACL injury mostly occurs during direction change, lateral pivoting, rapid deceleration, or landing tasks [43], [92] and is frequent among athletes involved in high risk sports such as soccer, football, basketball, and lacrosse [33]. ACL injury will result in knee instability and destruction of the menisci and articulated surfaces [44]. More than 120,000 ACL injuries occur annually, most during the high school years [33]. Treatment in 90% of patients includes reconstruction surgery, followed by a rehabilitation period [69]. The estimated average annual treatment cost of ACL rupture in the United States is more than 2 billion dollars [79]. Return to play for professional athletes following ACL surgery can take almost one year [88]. More than 50% will not return to their pre injury level of performance [69] and

Figure 1.1: Left: "good" SLS performance. Right: inward movement of the knee during "poor" SLS called Dynamic Knee Valgus (DKV).

between 50% to 100% develop osteoarthritis within 5 to 10 years after surgery. Moreover, the risk of re-injury increases up to 5 times in those who have undergone initial surgery [69]. All these statistics highlight the importance of early screening of individuals at higher risk, through functional movement tests such as the SLS.

1.2 Problem Statement

Current clinical assessment of the SLS involves clinicians visually observing patients conducting the movement and qualitatively rating performance utilizing clinical rating tools [91], [9], [71], [7], [23]. The rating tool can include several criteria such as knee position during the motion (pattella pointing past inside of foot is regarded as significant valgus), trunk alignment (degree of leaning), pelvic plane orientation (amount of rotation away from the horizontal plane), thigh motion (degree of hip adduction), and steady balance (stance leg wobbling) [42]. While qualitative rating is common in current practice, subjective evaluations are not always accurate, and assessor error can impact test validity and reliability [90], [25]. Reliability of visual assessment of the SLS by novice clinicians has

also been the subject of several studies [71], [90],[38]. These studies tend to confirm that, as expected, novice clinicians have lower reliability scores compared to experienced ones. However, Weeks et al. [90] and Harris-Hayes et al. [38] concluded that novice clinicians can still assess the SLS performance reliably while Poulsen et al. [71] found the novice clinicians reliability score lower than the level necessary for clinical applications. Also, Weeks et al. [90] found experienced clinicians were more sensitive to hip joint motion.

Moreover, qualitative assessment tends to rely primarily on the visual estimation of joint range of motion (ROM) and limb position, and often neglects the assessment of higher order kinematics (such as velocity and acceleration), since it is difficult for clinicians to assess these parameters without instrumentation.

In addition, the clinician selects the treatment protocol based on a short term consultation with the patient; however, better assessment and treatment may be possible when the patients' motor function can be monitored during long periods of activity and their entire rehabilitation regime, which currently is not possible. In-game monitoring of the athletes can also lead to better understanding of the injury mechanism.

Finally, since clinicians see a large number of patients each day, it may be difficult for them to remember the previous condition of each patient without a quantitative history.

The development of an automated SLS knee assessment system can improve the accuracy of SLS clinical assessments, and can provide objective results that can be tracked and monitored over time to guide rehabilitation and determine an individual's response to an intervention. Such a system can also be used to perform large population screenings to identify individuals with DKV and those at risk of knee injury.

In this thesis, our objective is to design a system that can (1) accurately measure human pose during the performance of the SLS, (2) provide a quantitative assessment of the SLS performance, and (3) is portable and easy to set up and use so that it can be deployed in clinical and field settings.

1.3 Thesis Contributions

1.3.1 Applying IMU-based Pose Estimation to Extract Lower Body Pose during the SLS

An IMU-based pose estimation algorithm based on kinematic chain modelling and EKF fusion, previously developed by Lin and Kulić [55], was adapted to estimate the human

lower body pose during the SLS. The following modification to the original algorithm were made:

- In the SLS, the hip is moving while the ankle has a fixed position on the ground so the ankle was set as the base frame.
- To improve sensor stability during the motion, unlike [55], the position of the tibia sensor was changed to the flat part on the side of tibia with considerable shift from frontal orientation. This shift introduced an additional calibration challenge, as it had to be estimated and compensated for in the algorithm.
- Because off-sagittal plane motion is important for poor SLS detection, the ankle and hip were modelled as 3 degree of freedom (DOF) joints. To avoid drift and get acceptable accuracy for all joints, similar to [46], a virtual yaw sensor was added to the hip.
- Moreover, for each IMU, a separate EKF was applied and a systematic way of EKF tuning was proposed to estimate process and measurement noise covariance parameters through optimization. Despite the importance of appropriate noise parameter selection, similar EKF-based studies [24],[81] usually do not describe how these parameters are selected. In this thesis, a structure for measurement and process noise covariance matrices based on movement characteristics and modelling assumptions is proposed, which offers a heuristic way of finding good initial values and speeds up further optimization.

1.3.2 Experimental Validation of the Pose Estimation Algorithm

The pose estimation algorithm was validated by comparing the estimated pose and joint positions to those obtained using marker-based motion capture. Ten participants were recruited to perform SLS and joint angles were obtained from both the proposed IMU-based algorithm and from the marker positions. Using marker-based pose as the ground truth, the average estimation error for the ankle, knee, and hip joints were 3.2°, 5.5°, and 7°, respectively. Considering the 3 dimensional estimation of hip and ankle motion and the fact that the proposed approach does not use magnetometers, this error is acceptable and comparable to similar studies [81], [24], [27].

1.3.3 Proposing a Clinical Calibration Method for Improved Pose Estimation from IMU Data with Minimum Required Measurements

The proposed pose estimation algorithm requires several measurements including kinematic link lengths as well as IMU position and orientation estimation with respect to the joint coordinate system, which is a concern in clinical applications. A sensitivity analysis was performed over all the required measurements to identify critical parameters. It was found that the most critical parameter for estimation accuracy was IMU orientation. A simple protocol is proposed to estimate IMU orientation in clinical settings. Using the experimental data collected in our validation study, we demonstrate that by extracting the sensor orientations using the proposed protocol and estimating the remaining parameter values using anthropometric tables, an accurate pose estimation can be achieved while minimizing the need for manual parameter measurements or performing calibration movements.

1.3.4 Automated DKV and Risk of Injury Assessment

In the second part of the thesis, the estimated joint parameters were utilized to distinguish between three different levels of squat quality and three different levels for risk of injury.

Squat quality is classified based on features extracted from joint angles, velocities or acceleration rather than raw sensor data, which has the benefit of providing interpretable information to the clinicians.

The estimated time series data of joint angles, velocities and accelerations for consecutive squats were segmented into individual squat repetitions. Statistical time domain features were generated from each repetition. The most informative features were selected using a combination of 18 feature selection techniques. Different classifiers were applied to the full dimensional data, the subset of selected features, and extracted features by supervised principal component analysis. The proposed approach was evaluated in two trials. First, a pilot study was conducted on a small dataset, followed by analysis on a larger clinical data set.

1.3.5 Validation on a Clinical Dataset

The proposed classification approach was verified on a clinical data collected by our clinical collaborator.

First, the feature selection results for both full agreement data and those with split decision of the raters were analyzed to identify the best predictors of DKV and overall risk of injury. The Feature selection results highlighted the discriminative role of the hip/ankle IR features as well as hip Flex features for DKV and risk of injury assessment which was in agreement with clinical findings.

Second, the performance of the proposed algorithm was characterized. The leave one subject out cross validation results suggest that discriminating "poor" squats from "good" ones is achievable with an accuracy of 90%, while the three class problem (adding "moderate" squats) achieves an accuracy of 68%. Screening participants at high risk of injury from those at no risk can be achieved with 87% accuracy and adding mild risk subjects drops the accuracy by 12%.

In addition, gender-specific differences were investigated by making male-only and female-only datasets. Applying feature selection and classification methods on these data sets revealed meaningful differences between DKV and risk of injury correlates in males and females, which was in agreement with findings in clinical and sports medicine studies. The results of gender specific classifiers suggest that developing separate classifiers for men and women improves classification results.

1.4 Thesis Outline

This thesis consists of two parts: pose estimation from IMU data during SLS and automated assessment of SLS based on labelled repetitions of SLS. Chapter 2 reviews the related studies to each of these parts. Regarding the pose estimation, available techniques for static and dynamic pose measurement are described and those using IMU measurements are examined and compared. In terms of automated assessment, recent findings about predictors of DKV during the SLS as well as studies which have attempted to detect DKV automatically are overviewed.

Chapter 3 describes the pose estimation algorithm including details of kinematic modeling and EKF tuning. The experimental data collection and the estimation of required parameters from marker data are described. The joint angles estimated from IMU measurements using the proposed algorithm are compared to angles extracted from marker information and the proposed algorithm performance is analyzed.

In Chapter 4 a sensitivity analysis is performed over all model parameters required by the pose estimation algorithm. This analysis identifies the IMUs' orientation as the most sensitive parameter set. Therefore, a simple protocol suitable for clinical settings

is proposed to extract full sensor orientation for more accurate pose estimation. Other, less-sensitive parameters were replaced by average values from anthropometric tables. The proposed calibration protocol is validated by comparing the pose estimation results with results when same parameters were extracted from marker positions for each person.

Chapter 5 proposes an approach for automated assessment of the SLS and its validation. The proposed method was first applied to a pilot study and subsequently expanded to larger clinical dataset. The data collection and labelling procedures for both the pilot and clinical studies are described. Second, we reported and compared the results with clinical findings.

Chapter 6 concludes the thesis and suggests directions for future work.

Chapter 2

Related Work

2.1 Pose Estimation

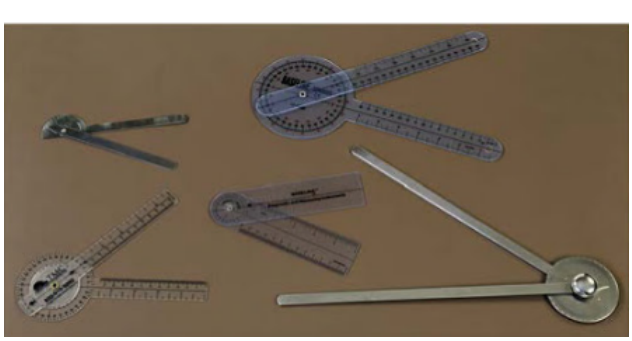
Human pose estimation is the process of estimating the configuration of the body and has many applications in rehabilitation, human-computer interaction, computer vision, and sports medicine.

2.1.1 Range of Motion Measurement in Clinical Settings

Pose estimation in clinical applications is based on visual assessment [39] and measurement tools are limited to range of motion (ROM) measurement, which is an essential part of physical therapy [32].

ROM refers to body joint movement limits. Full ROM is essential to move efficiently and with minimal effort [74]. Restricted ROM will result in impaired muscle extensibility, decreasing the body's ability to adapt to imposed stress on the body [74]. ROM may be impaired as a result of injury, degenerative disease or surgery. A crucial part of the rehabilitation process is to recover the full extent of the joint's ROM to avoid complications such as muscular dystrophy or contracture [80].

In current clinical practice, the most commonly available clinical tool is a goniometer, which is used for ROM measurement. A goniometer is a protractor-like device limited to single joint angle measurement in static positions (Figure 2.1(a))[32]. There are special types of goniometer designed for specific joints but in general, it is hard to apply a goniometer to all joints. Measuring several joints' ROM with a goniometer is time consuming.



(a)

(b)

Figure 2.1: Different styles and sizes of universal goniometer (a), cervical ROM device with three inclinometers (b)[74]

More importantly, it is not possible to measure ROM during dynamic movements with a goniometer. An inclinometer is a type of goniometer which measures the angle of the limb with respect to gravity and can measure motion at several joints [74](Figure 2.1(b)), but it only can be applied to motion directions which are not parallel to gravity. According to Norkin and White [65] several studies have shown lower reliability of visual assessment for lower extremity motion compared to upper extremity motion. Moreover, measurements are shown to be less reliable when motion is affected by adjacent joints or there are multi-joint muscles.

Rachkidi et al. [73] examined the reliability of visual assessment of the lower extremity passive ROM. According to their results, visual assessment of hip/knee Flex angle is reliable even for less experienced examiners. However, accurate assessment of hip Abd/IR needs experience and for the hip Add and ankle Flex/Ext, even with experienced clinicians reliability was poor. Therefore for those ROMs, using the goniometer was recommended. Youdas et al. [97] have also reported poor reliability for active ROM measurement of ankle Flex/Ext.

2.1.2 Dynamic Pose Estimation Methods

Development of an automated and objective pose estimation tool which can be used during movement has been an active research area within the past decades. Dynamic joint angle measurement techniques include camera-based, magnetic -based and IMU-based methods.

Camera-based Methods

Camera based systems can be categorized into two groups: marker-based and marker less techniques [24]. In the marker-based approach, reflective markers are mounted on bony land marks of the body, and movements are recorded by cameras installed in a Motion Capture Studio. Joint angles are then recovered by further analysis of the camera data. The camera-based system gives an accurate estimation of joint angles; however, it is not suitable for clinical applications due to cost and setup limitations. Also, lighting and capture space limitations limits its application for gait analysis, outdoor settings and sport analysis. Certain exercises and view angles may be difficult to capture if line-of-sight between markers and the cameras cannot be ensured. The process of mounting markers itself is time consuming and needs experience.

Marker-less methods are based on image processing techniques. Recently many researchers have taken advantage of Microsoft Kinect, which is inexpensive and provides depth and brightness/color information [21]. However, marker-less methods also work based on line-of-sight sensing. As a result, they suffer from the same limitations as marker-based systems.

Electromagnetic Field-based Methods

Electromagnetic-field based methods use a local magnetic field to track objects in the field [4]. These methods are accurate but subject to interference by ferromagnetic objects. The magnetic field generator also covers a limited space which limits its application for outdoor settings.

IMU-based Methods

An IMU is a compact package composed of an accelerometer which measures linear acceleration, a gyroscope which measures angular velocity and a magnetometer which measures earth magnetic field (Figure 2.2).

IMUs are very well suited to motion measurement in clinical settings, as they are lightweight, small, capable of long term data recording, wearable and have low cost and low power consumption [30]. However, they are not as accurate as camera-based systems. Accurate estimation of joint angles from IMU output is a challenging problem due to several issues including gyro drift, sensor to segment misalignment and motion artifacts.

Figure 2.2: An IMU is a compact package composed of an accelerometer, a gyroscope, and a magnetometer [6]

Joint angle data is of interest to clinicians and has to be recovered from IMU data, for which several methods have been proposed in the literature [89]. The majority use strap-down integration of angular velocity to estimate the orientation of the limb to which the IMU is attached with respect to a world frame and subsequently extract the joint angle from relative orientation of the two adjacent limbs [29], [24], [78].

However, IMU sensor readings are noisy and may have bias. When position estimation is done by integrating angular velocity, even a small amount of bias will grow over time and cause considerable errors in estimation [81], a problem known as *drift*.

To tackle drift, a common approach is to fuse the accelerometer and gyroscope data. Examples include applying the complimentary filter [18] or Kalman filter [58], [55], [24], [81] for data fusion. A comparison between different adaptive filtering techniques for sensor fusion and drift removal is provided in [66]. Drift can lead to physically unrealizable joint angle estimates. To ensure that the estimated joint angles are physically feasible, a number of studies introduced kinematic constraints to their estimation model [17], [55], [81].

Boonstra et al. [18] applied a complimentary filter including a low pass filter to accelerometer data to eliminate noise, and a high pass filter to gyroscope data to eliminate the DC offset. They estimated the sagittal inclination of the thorax, upper leg and the lower leg by combining an estimate from the arctangent of the accelerometer with integration of the gyroscope. Their method worked well for sagittal plane motions. However, since the accelerometer is used as an inclinometer, the motion needs to be constrained in

the gravity- acting plane, so a generalization of the approach to 3D motion is not trivial. Also, some information about the movement may be lost as a result of filtering.

Bonnet et al. [17] estimated joint angles of the hip, knee, and ankle during a double leg squat exercise using a single IMU placed on the lower trunk. They used a planar chain model of the body consisting of the trunk, upper and lower leg and calculated the angular velocity and linear acceleration in the sagittal plane. Taking advantage of the periodic pattern of movement, they represented joint angles by an N-harmonic Fourier series. The parameters for this representation were calculated by minimization of least square error between the measured and estimated linear accelerations and the angular velocity. An advantage of their method is that no integration is needed to recover joint angles; therefore, the drift problem is negligible. Moreover, joint angle physiological limits are introduced in the cost function. However, their estimation restricts the motion to the sagittal plane, and cannot be generalized to 3D angle measurement. Also joint angle representation by Fourier series is not applicable for non-rhythmic movements.

Lin and Kulić [55] developed an automated human motion measurement method capable of knee and hip pose estimation during typical physiotherapy and training exercises. The accelerometer and gyro data were input to an Extended Kalman Filter using a kinematic model to recover joint angles. Gyroscope drift was limited by introducing kinematic constraints to the model. In addition, employing filter noise adaptation helped in dealing with model uncertainty resulting from different motion speeds. However, the EKF assumed a constant acceleration motion profile and assumed a known sensor alignment.

Joukov et al. [46] improved Lin and Kulić's pose estimation method [55] for gait analysis. They attached IMUs to the waist, knees and ankles to extract 3D leg motion during gait. They combined EKF with a canonical dynamical system which was capable of learning rhythmic motion from estimated joint velocities and estimating the frequency and phase of the movement. The estimated frequency and phase was then used to improve EKF state estimation by providing additional information about motion acceleration. Moreover, virtual yaw sensors were applied to the waist to reduce drift. According to their results, rhythmic EKF needs at least 15 seconds for accurate frequency measurement. Therefore, it performs no better than the EKF-based method for movements of short duration. Moreover, it is not applicable to non-rythmic motions.

Some commercial position estimation systems based on IMUs are also available (e.g. Xsens MVN BIOMECH system [5]). In these systems, a large number of IMUs are usually mounted in a suit to estimate the full body pose. Although these systems may be useful for animation applications, they are not well suited for clinical applications because the suit is expensive and cannot be fitted to all patients. It also requires a long calibration

process and cleaning is also an issue.

Another major issue with IMUs is their sensitivity to misalignment [78]. Joint angles should be measured in the anatomical joint coordinate system and since the sensor local frame may not be matched to the anatomical joint frame, measurements are subject to error. An exact positioning of the sensor or a calibration procedure is needed for acceptable results. Calibration techniques rely on either pose or functional calibration or a combination of both [61]. In pose calibration, the subject is asked to stand in a known posture while in functional calibration the sensor alignment is found using limb movements [28], [57], [61], [77],[54], [66]. Although these methods can improve accuracy, they are time consuming and the precision depends on the accuracy of the movements/pose executed by the subject. Moreover, some of these methods need a motion with a wide range which might not be possible for patients with limited ROM. Recently, Seel et al. [78] have proposed a functional calibration method which works based on arbitrary movements; however, their method requires at least two IMUs around the target joint and arbitrary motions have to excite all DOFs.

Favre et al. [29], used two IMUs located on the thigh and shank to estimate knee joint angles. They proposed a fusion algorithm to combine accelerometer and gyroscope data and calculate the orientation of upper and lower leg segments. To convert orientations to joint angles, they aligned IMUs relative to the same world frame so that they could calculate differential orientation which gives joint angles. This alignment was based on predefined calibration movements done by the subject and its accuracy depends on the accuracy of performing the calibration movements. This method is not suitable for patients as they may have limited ROM which prevents them to perform the prescribed calibration movements. Moreover, no strategy for drift elimination is proposed.

Daponte et al. [24] used magnetometer and accelerometer information to calculate sensor orientation quaternions and fused these quaternions as well as gyroscope output into Extended Kalman Filter to compensate for gyro drift. Calibration methods were applied to compensate for both magnetic interference and sensor to segment misalignment. Relative quaternions were used to recover elbow and shoulder joint angles. Deponte's algorithm treats each body segment as an independent rigid body, without considering the mobility constraints between body segments, so adjacent body segments may drift and move with respect to each other in ways that are not physically realizable.

Soft tissue artifacts are another source of error in IMU-based pose estimation, which is less often addressed in the literature. McGinnis et al. [76] proposed a method for accurate femur angle estimation during Femoroacetabular Impingement diagnosis test using a thigh mounted IMU. They also investigated the effect of soft tissue artifacts using two IMUs

mounted around the thigh, one fixed to the bone directly and the other attached to the skin. According to their results, soft tissue artifacts were responsible for 3% and 2% overestimate in hip flexion/extension (Flex/Ext) and internal/external rotation (IR/ER) angles respectively and 17% underestimate in hip adduction/abduction (Add/Abd) angle. However, they concluded that mounting on skin would still be enough to show differences between normal and limited hip motions due to Femoroacetabular Impingement. Their study is limited to static range of motion and does not take into account the effect of muscle contractions in dynamic tests or during exercise.

Seel et al. [78] estimated the Flex /Ext angle of knee and ankle in gait trials of a transfemoral amputee using three IMUs. Despite calibrating the IMU local frame to the body segment anatomical frame, they reported almost 4° difference in knee and 1° degree difference in ankle joint angle estimates between the prosthesis and contralateral sides due to skin and muscle motion artifacts.

A summary of the advantages and limitations of the IMU-based pose estimation methods are listed in table 2.1.

2.2 Sensitivity Analysis

As mentioned in the previous section, IMUs are sensitive to correct positioning, and misalignment is a source of error in pose estimation. While several studies have mentioned the importance of sensor positioning and suggested methods for more accurate estimation of the sensor orientation [28], [78], [54], very few studies have quantified the sensitivity of pose estimation to sensor misplacement.

Trojaniello et al. [85] investigated the sensitivity of four different single IMU-based gait initial contact estimation methods to IMU misplacement. The lower back IMU was virtually rotated around the medio-lateral axis between $\pm 10°$. Two methods had acceptable performance only in a limited range of IMU orientation change. One method was quite insensitive but also had poor accuracy. Only one method showed acceptable performance in terms of a compromise between good accuracy and least sensitivity. This method was applied to a waist mounted IMU and wavelet transform was used to extract initial and final contact gait time events [59].

Leardini et al. [52] compared a magnetic IMU based rehabilitation assistive system to the optical motion capture analysis gold standard. The proposed system applies 3 IMUs to the thorax, thigh and shank to estimate hip, knee and thorax inclination angles. The sensitivity of the hip Abd/Add to frontal plane misorientation of the thigh IMU within

$\pm 15^\circ$ of the optimal orientation and the sensitivity of the hip Flex/Ext to mispositioning of the thigh IMU within ± 7 cm in mediolateral direction of the correct position was analysed. Their results showed more error due to mediolateral displacement than due to frontal misorientation. They concluded that overall error due to introduced misplacement is less than 5° and is still acceptable. However, misconfigurations were limited to two scenarios tested on one of the sensors, and the effect was investigated only on a single hip DOF estimation.

Table 2.1: Comparison between IMU-based joint angle measurement methods

IMU-based joint angle measurement methods		
Method	*Advantages*	*Limitations*
Boonstra et al. 2006 [18]	- Drift removal using high pass filter	- Motion information may be lost during filtering - Not applicable to off-sagittal plane - Estimate inclination with respect to world frame not joint angles - Joint physiological motion constraints are not considered - No sensor misalignment calibration
Favre et al. 2008 [29]	- Estimates 3D joint angles based on inclination - Sensor misalignment calibration	- No drift removal strategy - Needs exact calibration movements - Joint physiological motion constraints are not considered
Bonnet et al. 2012 [17]	- Drift removal and estimating joint angles using Fourier series - Motion constraints are considered	- Not applicable to off-sagittal plane motions - Not applicable for non-rhythmic movements - No sensor misalignment compensation
Lin et al. 2012 [55]	- Estimate 3D joint angles using kinematic model - Motion constraints are considered - Drift compensation using EKF, kinematic model and kinematic constraints	- No sensor misalignment calibration - Assumes constant acceleration in motion dynamics - Using EKF introduces additional challenge of appropriate noise parameter selection which is discussed
Daponte et al. 2014 [24]	- Sensor misalignment calibration - Drift removal using magnetometer and EKF - Estimating 3D joint angles using relative quaternions	- Body segments are independent of each other without applying motion constraints - Using magnetometer has added the challenge of magnetic calibration - Using EKF introduces additional challenge of appropriate noise parameter selection which is not discussed - Needs exact calibration movements

IMU-based joint angle measurement methods		
Method	*Advantages*	*Limitations*
Joukov et al. 2015 [46]	- Drift compensation - Estimate 3D joint angles using kinematic model - Motion constraints are considered	- Not applicable to non-rythmic motions - Needs time to converge and perform well
Seel et al. 2014 [78]	- Sensor misalignment calibration with arbitrary motions - Gyro drift removal - Accurate pose estimation	- Two sensors around each joint is needed - Joint physiological motion constraints are not considered - The calibration method estimates only main joint axis (is the one with greatest ROM) - Precise estimation in one direction (1DOF)
Xsens Suite [5]	- Sensor misalignment calibration - Gyro drift removal - Whole body pose estimation	- Expensive - Suit does not feet to all patients - Cleaning problem - Long calibration process because of large number of sensors

2.3 Single Leg Squat Assessment

SLS and other functional movement tests such as the double leg squat and double leg jump have been widely investigated in clinical and sport medicine studies. DKV occurrence during these tests is not due to torsion of the knee, but due to movement of the kinematic chain, particularly the hip/ankle inward movement. The main purpose of the majority of these studies is to find relationships between the occurrence of knee valgus during the mobility test and factors such as age, gender, body mass index, history of injury, and kinematic or neuromuscular characteristics of the subjects (usually athletes) [41], [98], [15], [67], [84], [86]. Studying these predictors aids in the development of appropriate injury prevention strategies. For example, if it is found that hip abductor weakness correlates with poor performance (DKV occurrence) in SLS, then specific exercises can be prescribed to improve the strength and function of this muscle group.

2.3.1 Finding DKV Correlates

Zeller. et al. [98] investigated the kinematics and muscular activity of nine men and nine women athletes during the SLS. According to their results, women exhibited more knee valgus, which was associated with greater ankle dorsiflexion (7°) and pronation (2.7°), less trunk lateral flexion (16.6°), and greater hip adduction (3.2°), flexion (9.1°), and rotation (6.6°). Rectus femoris muscle activation was also greater in women.

Hip and foot contributions to high DKV were investigated by Bittencourt et al. [15]. They examined 173 athletes during the SLS and at the landing moment of a double leg jump. Data was collected in a Motion Capture Studio and the frontal plane knee projection angle was measured at 60° of knee flexion and during a static single-leg stance. Four other measures, including the passive ROM of the hip internal rotation, the isometric strength of the dominant-limb hip abductors, the shank-forefoot alignment and participants gender were defined as features to be input into a classification and regression tree. Their results suggest that high DKV can be predicted by decreased hip abductor torque (below $1.03 Nm/Kg$) and increased passive ROM of the hip internal rotation (greater than 43°) for both the SLS and double leg jump landing.

Padua et al. [67] compared the neuromuscular characteristics of a group of 18 individuals with excessive knee valgus with a control group of 19 healthy individuals during double leg squat performance. Electromyography (EMG) was used for muscle activation measurement. Individuals were assigned to either the control or DKV group based on an evaluation by an expert rater. A correlation between DKV and increased hip-adductor activation as well as increased coactivation of the gastrocnemius (by 42%) and tibialis anterior muscles (by 25%) was reported.

In a similar study, Stiffler et al.[84] compared kinematic characteristics including ROM and postural alignment of 97 healthy individuals during the double leg jump, in order to find differences between those with and without excessive DKV. Motion labeling was based on the total Landing Error Scoring System (LESS)[68]. Their results showed associations between DKV and less ankle dorsiflexion (3.6°), as well as higher quadriceps angle (3.6°).

The relationship between the occurrence of DKV with age, gender, and body mass index was studied by Ugalde et al. [86]. They investigated 142 middle and high school athletes while performing the SLS and drop jump tests. They defined knee-hip ratio as the distance between the knees at maximum flexion divided by the distance between hips at a quiet stance during the drop-jump test. Their results showed significantly lower knee-hip ratio for individuals with DKV during SLS. However, they found no relationship between DKV and age, gender, or body mass index during the SLS test.

2.3.2 Automated DKV Detection

The investigations described above have focused on identifying correlates of DKV. Generally, these studies detect positive DKV occurrence based on expert clinician observations or manual measurements extracted from video frames. Very few studies have tried to develop an algorithm for automated DKV detection. In one such study, Whelan et al. [94] classified SLS repetitions of 19 healthy participants into correct and incorrect categories using a single lumbar-mounted IMU. They extracted time domain features from accelerometer and gyroscope measurements, the IMU orientation (represented as roll, pitch, yaw), and accelerometer magnitude. Using the generated feature vector and labels provided by an expert, they trained a Random Forest classifier, which achieved 92.1% accuracy with repeated random-sample validation. Despite these promising results, they may be difficult to interpret clinically, as features were defined based on direct acceleration and gyroscope output signals, where clinical assessment of the SLS includes the visual estimation and interpretation of kinematic joint parameters, especially the joint angles. Developing a classifier which works based on these parameters; therefore, has the advantage of interpretability. Furthermore, Whelan et al., did not perform a Leave One Subject Out Cross Validation (LOSO-CV), therefore it is not clear how well the classifier would generalize to subjects unseen during training, which is critical for clinical applications.

2.4 Summary

IMU-based pose estimation is a challenging problem because of gyro drift, sensitivity to mispositioning, and motion artifacts. Several studies have attempted to solve these issues and some of them have reported very low estimation errors [29], [78]. However, many of the proposed methods have limitations like one dimensional joint angle estimation or exhaustive calibration requirements which make clinical application of the method time consuming or technically complicated and demanding for the end user (see Table 2.1).

While sport medicine, rehabilitation, and other clinical applications are among the most important purposes of developing IMU-based motion tracking systems, in most of the studies, clinical applicability or performance are not evaluated. There is a gap in the work to date regarding the performance of IMU-based pose estimation methods in realistic clinical settings, where there are time constraints, data collection is to be performed by clinicians with limited technical background, and motion data is being generated by clinical populations.

In this thesis, we propose an accurate IMU-based pose estimation method for clinical

applications. The proposed approach addresses most of the existing issues with IMUs, and at the same time simplifies the calibration and setup procedures to make it usable in clinical settings without compromising accuracy or generalizability.

For clinical applications, in addition to pose estimation, there is a need for automated motion assessment during functional movement tests. There are very few efforts for automated assessment of clinical movement tests using IMUs in the literature. In this thesis, we propose and validate a system for assessment of the single leg squat mobility test.

Our proposed system provides an integrated solution for both pose estimation and assessment of the motion for the SLS, which makes it ideal for clinical applications.

Chapter 3

Automated Pose Estimation

In this chapter, the algorithm for continuous pose estimation from IMU measurements is developed. Many existing methods on human pose estimation provide body limb orientations [29],[24], [78], which need to be transformed into joint angles using kinematic information in a subsequent step. On the other hand, the algorithm proposed by Lin and Kulić [55] incorporates the kinematic model in the estimation process, to directly produce joint angles. Moreover, their proposed algorithm estimates joint angles during movement in any direction. There is no limitation to movement only in the sagittal plane as in many other studies [18], [17]. For these reasons we have adapted the pose estimation approach proposed in [55] to measure joint parameters including angle, velocity and acceleration during the SLS.

Since the IMU data is noisy and can suffer from drift, similar to [55], a kinematic model of the lower leg is applied to calculate angular velocity and linear acceleration at each time step to be used for correction of sensor estimates of these values. The kinematic model predictions and sensor measurements are fused via an Extended Kalman filter (EKF). The algorithm proposed in [55] is modified as follows:

- Degrees of freedom of the kinematic model was increased to 7 to simultaneously model the 3 dimensional hip, the 1 dimensional knee and the 3 dimensional ankle
- Drift was minimized by introducing a virtual yaw sensor at the hip
- Tibia mounted sensor location was changed to the flat part of the tibia to improve stability

- For each of the ankle, knee, and hip joints a separate EKF was applied to reduce joint parameters estimation error and simplify filter tuning

3.1 Kinematic Modeling

The human body can be modelled as a branched articulated chain of rigid bodies, and the mobility of the body joints can be modelled as a series of single degree of freedom joints, where each DOF represents joint motion in one direction. A common way to specify such a model is to assign local frames to each DOF. Body segment rotations can then be quantified by transformation matrices from one frame to the other. Each transformation matrix has 9 parameters representing rotation and 3 parameters representing transformation. By forcing some constraints on frame assignment, number of parameters in transformation matrix can be reduced [83].

3.1.1 DH Convention

The Denavit Hartenberg (DH) convention [87] is a common way of assigning frames to rigid body links. The convention allows the transformation from one link to the other to be fully defined using 4 parameters. These parameters include:

1. Joint angle (θ_i): angle that forms between the X_{i-1} and the X_i axes while rotating about the Z_{i-1} axis (i refers to frame order)
2. Link offset (d_i): distance from frame origin $i-1$ to the X_i axis along the Z_{i-1} axis
3. Link length (a_i): distance between the Z_{i-1} and the Z_i axes, along the X_i axis
4. Link twist (α_i): angle from Z_{i-1} to the Z_i axis, about the X_i axis

In addition, the DH fame assignment includes two constraints on axis assignment: axis X_i should be perpendicular to axis Z_{i-1}, and axis X_i intersects axis Z_{i-1}. Using these rules, the transformation matrix from frame $i-1$ to frame i can be defined as following:

$$T_{i-1,i} = \begin{bmatrix} cos(\theta_i) & -sin(\theta_i)cos(\alpha_i) & sin(\theta_i)sin(\alpha_i) & \alpha_i cos(\theta_i) \\ sin(\theta_i) & cos(\theta_i)cos(\alpha_i) & -cos(\theta_i)sin(\alpha_i) & \alpha_i sin(\theta_i) \\ 0 & sin(\alpha_i) & cos(\alpha_i) & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

3.1.2 Forward Kinematics

The developed kinematic model for the human leg is composed of a 3 DOF ankle joint, 1 DOF knee joint, and 3 DOF hip joint. Frame assignment is carried out according to the DH convention as depicted in Figure 3.1(left). Frames 0, 1, and 2 correspond to ankle internal/external rotation (IR/ER), adduction/abduction (Add/Abd) and flexion/extension (Flex/Ext) respectively, followed by the shank link. Frame 3 corresponds to knee Flex/Ext, followed by the thigh link. Frames 4, 5, 6 correspond to hip Flex/Ext, Add/Abd, and IR/ER respectively, followed by the pelvis link. Frame 7 is the final frame located at the back sensor. Frame 0 is the base frame as it is stationary during the SLS motion. To generate the measurement equations of the EKF, the angular velocity and linear acceleration measurements at each sensor location are related to the joint positions and velocities using the kinematic model, which can be done by iteratively computing the angular velocities and linear accelerations as contributions from each joint in the kinematic chain:

$$\omega_i = {R^T}_{i-1,i}\ \omega_{i-1} + {R^T}_{i-1,i}\ \dot{q}_i \tag{3.1}$$

$$\alpha_i = {R^T}_{i-1,i}\ \alpha_{i-1} + {R^T}_{i-1,i}\ \ddot{q}_i + \omega_i \times ({R^T}_{i-1,i}\ \dot{q}_i) \tag{3.2}$$

$$\ddot{x}_i = {R^T}_{i-1,i}\ \ddot{x}_{i-1} + \alpha_i \times r_i + \omega_i \times (\omega_i \times r_i)\ for\ i \neq si(sensorframe) \tag{3.3}$$

$$\ddot{x_{si}} = {R^T}_{i-1,i}\ \ddot{x}_{i-1} + \alpha_i \times r_{si} + \omega_i \times (\omega_i \times r_{si}) + R_{0,i}g\ for\ i = si \tag{3.4}$$

Where ω_i is the angular velocity of frame i, $R_{i-1,i}$ is the rotation matrix from frame $i-1$ to i , α_i is the angular acceleration of frame i, $R_{0,i}$ is the rotation matrix from the base frame to frame i, and r is the displacement vector from the origin of the current frame to the origin of the previous frame. x_i is the linear acceleration of the origin of frame i, and x_{si} is the linear acceleration at the sensor i location.

The linear acceleration at the tibia sensor location will be x_{S3}, where r_{S3} is the vector from the sensor location to frame 2 (origin at the ankle). r_2, r_1, and r_0 are zero as frames 2,1,0 are located at the same position. For the thigh sensor, x_{S4} is the linear acceleration at the thigh sensor location, and r_{S4} is the vector from the thigh sensor to frame 3 (origin at the knee). r_3 for this sensor is the vector from frame 3 (knee) to frame 2 (ankle). Similarly, for the back sensor, x_{S7} is the linear acceleration at the back sensor location. r_{S7} is the vector from the back sensor to the hip center. r_6 and r_5 are zero (all at hip center), r_4 is the vector from hip to knee. The displacement vectors are shown in Figure 3.1(middle and right).

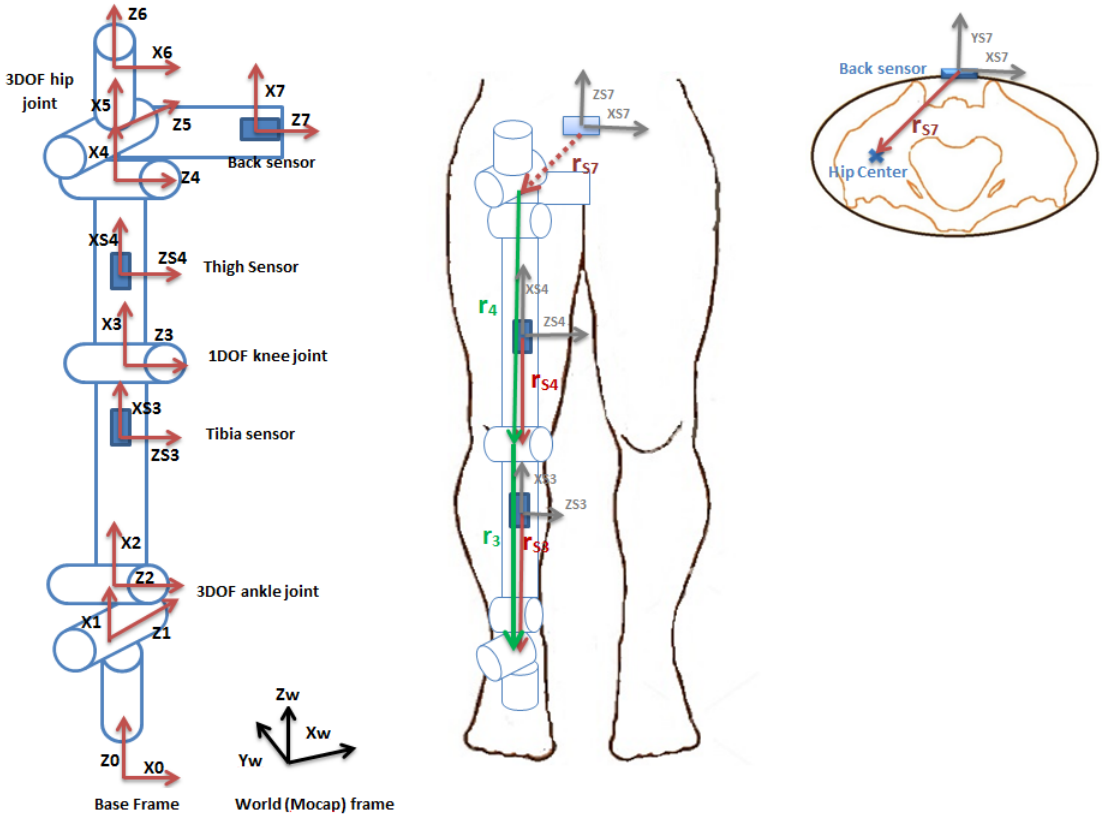


Figure 3.1: 7 DOF kinematic model of the right leg showing sensor positions and frame assignments (left) and displacement vectors frontal view (middle) and transversal view (right).

The resultant estimates of the angular velocity and linear acceleration from the kinematic model and the IMU measurements of these parameters are then fused into an EKF to recover the joint angles [55]. The position, velocity, and acceleration of each DOF are defined as the states to be estimated by the EKF. A constant acceleration model is used for the state propagation.

3.2 Extended Kalman Filter (EKF)

The Kalman Filter is an estimation technique which combines noisy observations and a model of the system to produce a more precise estimation of unknown variables [93].

The Kalman filter assumes that both the measurement and system model are linear, and that the measurement and model uncertainty are zero-mean Gaussian distributed. The Extended Kalman Filter is a nonlinear version of the Kalman Filter which linearizes the state and observation models about the current estimated mean and covariance [93].

In our model, the state propagation model is linear while the observation is a nonlinear function of the states. The discrete time model of our system can be formulated with state and observation equations as follows:

$$q_k = Aq_{k-1} + w_k \tag{3.5}$$
$$Z_k = h(q_k) + v_k \tag{3.6}$$

Where A is the linear state transition equation which relates previous state (q_{k-1}) to the next state (q_k). w_k and v_k are zero mean Gaussian process and measurement noises with covariance Q and R, respectively. Q relates to the states and has the same size as A while R relates to the observations. Z_k is the current observation which is a nonlinear function (h) of the state.

Our desired states to be estimated are position, velocity and acceleration of each DOF. Assuming a constant joint acceleration, the relationships between position, velocity, and acceleration can be written as follows:

$$q_k = q_{k-1} + \dot{q}_{k-1}\Delta k + \ddot{q}_{k-1}\frac{\Delta k^2}{2} \tag{3.7}$$
$$\dot{q}_k = \dot{q}_{k-1} + \ddot{q}_{k-1}\Delta k \tag{3.8}$$
$$\ddot{q}_k = \ddot{q}_{k-1} \tag{3.9}$$

Where Δk represents the sampling time interval. Using this model, for the pose estimation problem, the A matrix in the equations will have the following format:

$$A = \begin{bmatrix} 1 & \Delta k & \frac{\Delta k^2}{2} \\ 0 & 1 & \Delta k \\ 0 & 0 & 1 \end{bmatrix}$$

The EKF formulation consists of predict and update equations [93]. Prediction includes projecting the state ($\hat{q_k}^-$) and the error covariance ($P^-{}_k$) ahead based on the state model:

$$\hat{q_k}^- = A\hat{q}_{k-1} \tag{3.10}$$

$$P_k{}^- = AP_{k-1}A^T + Q \tag{3.11}$$

The update state includes calculating the innovation (residual) and the covariance (Equation 3.13) by linearizing the measurement model to obtain the jacobian C numerically (Equation 3.12), computation of the Kalman gain K (Equation 3.14) and updating state estimation (Equation 3.15) and state error covariance (Equation 3.16) using the Kalman gain.

$$C = \frac{\partial h(q_k)}{\partial q_k} \tag{3.12}$$

$$S = CP_k{}^-C^T + R \tag{3.13}$$

$$K_k = P_k{}^-C^T(S)^{-1} \tag{3.14}$$

$$\hat{q_k} = \hat{q_k}^- + K_k(Z_k - h(q_k)) \tag{3.15}$$

$$P_k = (I - K_kC)P_k{}^- \tag{3.16}$$

Before the filter starts to run, some of its parameters including the initial state value (q_0), the initial state error covariance (P_0), the measurement noise covariance (R), and the process noise covariance (Q) need to be initialized.

The measurement noise initialization can be carried out empirically, because the sensor noise profile is usually measurable by taking offline samples [93]. However, the correct value for the process noise, which compensates for the model inaccuracies, is difficult to estimate. Poor values for the process noise can affect the accuracy of estimation significantly.

3.2.1 Adding Virtual Yaw Sensor

For rotation angles parallel to gravity, drift due to gyro bias cannot be compensated by the accelerometer and can result in large IR/ER angle errors. The drift problem is most prevalent in the hip IR/ER due to accumulation of error from previous states. To alleviate this issue, similar to [46] a virtual yaw sensor was assumed at the hip location to measure hip internal rotation only. The output of this sensor was set to zero to avoid excessive internal rotation due to drift. To fuse this additional measurement to EKF, the measurement model

needed to be updated, through a rotation matrix from the base frame to the hip IR frame ($R_{0,7}$), calculated as shown below:

$$Hip_{IR} = tan^{-1}\left[\frac{R_{0,7}(2,1)}{R_{0,7}(1,1)}\right] \tag{3.17}$$

This additional information was added to $h(q_k)$

3.2.2 1-EKF versus 3-EKF Structure

Two different approaches can be implemented for state estimation.

1-EKF Structure

The first approach is to estimate all joint angles using the information of all sensors via a single EKF. In this approach, the state vector has a dimension of 21×1 (including 3 states for each of the 7 DOFs) and the observation vector has a dimension of 19×1 (including 6 readings for each IMU and 1 reading for virtual yaw sensor). As a result, the A matrix is a 21×21 matrix with the following structure:

$$A = \begin{bmatrix} A_1 & O & \dots & O \\ O & A_2 & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & A_7 \end{bmatrix}, \; where \; A_1 = A_2 = \dots = A_7 = \begin{bmatrix} 1 & \Delta k & \frac{\Delta k^2}{2} \\ 0 & 1 & \Delta k \\ 0 & 0 & 1 \end{bmatrix} \tag{3.18}$$

Where O is a 3×3 zero matrix.

With this approach Q has the same size of 21×21 and R is a 19×19 matrix.

3-EKF Structure

In this approach a separate EKF is assigned to each IMU. Therefore, the observation vector for the ankle and thigh EKF has 6 elements while for the hip EKF, it has 7 elements

including the virtual sensor. This will result in a 6×6 R matrix for the first two EKFs and a 7×7 R matrix for the third one.

Ankle joint states are estimated using the tibia sensor EKF, knee joint states are estimated using the thigh sensor EKF, and hip joint states are estimated using the back sensor EKF. This means a 9×9 A matrix for ankle states and hip states and a 3×3 A matrix for knee states in the prediction equation 3.10. Q follows the same size as A for each joint. The three filters are run sequentially, so that the estimate of the ankle joints' state is used as input to the knee estimate, and the ankle and knee estimates are used as inputs to the hip EKF estimator.

The one EKF approach simultaneously estimates the ankle states based on all sensors' measurements, and the knee state based on both thigh and back sensor measurements, while in the three EKF approach, these joint angles are estimated using only one of the sensors (tibia sensor for the ankle joints, thigh sensor for the knee joints). Therefore, we expect to have a more accurate estimation for these values compared to the 3-EKF model. However, after implementing the two approaches and comparing the results, we found that the three EKF model provides better joint angle estimates. These results and their interpretation are detailed in Sections 3.6 and 3.7. In the rest of the thesis, we use the better performing 3-EKF formulation throughout, unless otherwise mentioned.

3.2.3 EKF Tuning

EKF has been a very successful estimation method in a wide variety of research fields [53]. However, good performance of the EKF depends on the proper tuning of its noise statistics. Despite the importance of a well-tuned filter, the estimation of optimal filter parameters is less often discussed in the literature. Tuning can be done based on ad hoc or heuristic methods or using supervised or unsupervised optimization [14].

Ad hoc methods are based on trial and error, while heuristic methods rely on empirical analysis of the system and using a priori knowledge of the system.

Unsupervised optimization methods try to adapt noise parameters over time using the covariance of the filter innovation term. There are four main categories of adaptive tuning techniques including Bayesian, maximum likelihood, covariance matching and correlation techniques [64]. These adaptive methods are advantageous because they can respond to change in the system properties over time.

Supervised methods typically perform offline optimization of the filter parameters using a reference or ground truth. This method gives optimal parameters; however, those parameters may not be useful any more when system changes.

Adaptive tuning techniques are most suitable for long runs to allow adaptation. Since SLS trials are rather short and we had access to ground truth via motion capture, the best choice for filter tuning was supervised optimization. For this purpose, first, Q and R matrix structures are simplified to reduce the number of parameters to be tuned. Then, they are initialized to values selected based on a priori knowledge of the system. Finally an objective function is defined based on RMS error between joint angle estimates obtained on motion capture and from the algorithm and final values of the parameters are found by optimization. Details of the Q and R matrix initialization are described in the next section.

EKF Tuning by Optimization

Our objective is to identify the optimal elements of measurement noise R and process noise Q covariance. For each IMU a separate EKF was applied. The tibia IMU estimates the first 9 states including the 3 DOF ankle joint angles, positions, and accelerations. Therefore, the number of states to be estimated by the first EKF was 9, while the measurements were the angular velocity (x, y, z) and linear acceleration (x, y, z) from the tibia IMU, which resulted in a 9×9 Q matrix and 6×6 R matrix.

The thigh IMU estimates the 3 states of the 1 DOF knee joint, the angle, velocity and acceleration. For the second EKF therefore, we have a 3×3 Q matrix and a 6×6 R matrix.

Finally, the back IMU estimates 9 states of the 3 DOF hip: the joint angles, positions, and accelerations. Because of the additional virtual yaw sensor added at the hip, the R matrix for hip EKF has a size of 7×7.

R and Q in general are full matrices. To simplify the matrix structures and reduce the number of parameters to be optimized, we assume that measurement noise v_k and process noise w_k elements are uncorrelated with each other, so we can define R and Q as diagonal matrices. We assign positive values to diagonal elements to keep them positive definite.

For simplicity, we assume all accelerometer dimensions have the same noise profile and that the tibia, thigh and back accelerometers are also similar. The same assumption holds for the gyro. In this way, we can reduce the 19 parameters of R (6 for tibia, 6 for thigh and 7 for hip) into three parameters, one related to gyro noise covariance (β_1), one related to accelerometer noise covariance (β_2) and one related to virtual yaw sensor noise covariance (β_3). From investigation of raw sensor data, it was observed that the accelerometer is noisier than the gyro; therefore, the accelerometer coefficient in R is initialized to a higher value. The structure of R for the tibia and thigh sensors is the following:

$$R_{ankle,knee} = \begin{bmatrix} \beta_1^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \beta_1^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta_1^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta_2^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \beta_2^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \beta_2^2 \end{bmatrix}, \ where \ \beta_2 > \beta_1.$$

β_1 corresponds to gyroscope noise and β_2 corresponds to accelerometer noise. For the hip EKF there is one more element corresponding to the virtual yaw sensor (β_3):

$$R_{hip} = \begin{bmatrix} \beta_1^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \beta_1^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta_1^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta_2^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \beta_2^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \beta_2^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \beta_3^2 \end{bmatrix}$$

Each element in Q corresponds to one of the states. Whenever there is higher uncertainty in the state, it means that the corresponding diagonal element of Q for that state should have a higher value. Due to the constant acceleration assumption, we hypothesize that acceleration will be more inaccurate than velocity, and velocity will be more inaccurate than position for all of the states. Therefore, we define a 3×3 diagonal matrix D such that its first element describes position noise, the second element defines velocity noise, and the third element defines acceleration noise:

$$D = \begin{bmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3 \end{bmatrix}, \ where \ \alpha_3 > \alpha_2 > \alpha_1 > 0.$$

To come up with a good initial guess for the diagonal elements of Q, we can benefit from our knowledge of the system. In the case of pose estimation for SLS, the motion is known which means we know the direction and relative range of motion in each DOF.

The motion (SLS) is primarily in the sagittal plane so flexion/extension joints have the biggest ROM. Therefore, the constant acceleration assumption will most likely be violated

in the joints moving in the sagittal plane so coefficients in Q for joints moving in that direction are initialized with higher magnitudes.

Using this approach, the following structure is proposed for the process noise covariance matrix for the ankle:

$$Q_{ankle} = \begin{bmatrix} D \times {\gamma_1}^2 & O & O \\ O & D \times {\gamma_2}^2 & O \\ O & O & D \times {\gamma_3}^2 \end{bmatrix}$$

Where γ_1 corresponds to the uncertainty in the first DOF parameters (ankle IR/ER) estimation, γ_2 corresponds to uncertainty in the second DOF (ankle Abd/Add) and γ_3 corresponds to uncertainty in the third DOF (ankle Flex/Ext).

The same structure is used for the hip, such that γ_5, γ_6, γ_7 correspond to Flex/Ext, Abd/Add, and IR/ER, respectively:

$$Q_{hip} = \begin{bmatrix} D \times {\gamma_5}^2 & O & O \\ O & D \times {\gamma_6}^2 & O \\ O & O & D \times {\gamma_7}^2 \end{bmatrix}$$

For the knee we only have one DOF so Q will be:

$$Q_{knee} = \begin{bmatrix} D \times {\gamma_4}^2 \end{bmatrix}$$

The knee, ankle and hip Flex/Ext have the biggest movements in the SLS, so we can assume that γ_3, γ_4, and γ_5 are greater than other parameters. Also, ankle IR/ER and hip IR/ER are parallel to gravity so their estimates rely only on gyro integration. To prevent drift and inaccurate estimation, we trust the model more for these values so we assign low coefficients for them in Q. Therefore, we can set $\gamma_4 > (\gamma_3, \gamma_5) > (\gamma_2, \gamma_6) > (\gamma_1, \gamma_7)$.

Using this structure for R and Q matrices, the number of parameters to be tuned will be 13 parameters in total.

To find the optimum values of the parameters, the objective function was defined as the root mean square error between the estimated joint angles from the EKF-based algorithm and the same angles obtained from the marker-based ground truth, averaged over all joint angles as shown in Equation 3.19

$$J = argmin_{(\beta_1,\beta_2,\beta_3,\alpha_1,\alpha_2,\alpha_3,\gamma_1,\gamma_2,\gamma_3,\gamma_4,\gamma_5,\gamma_6,\gamma_7)} \left[\frac{\sqrt{\frac{\sum_{t=0}^{n}(\hat{q}_{1,t}-q_{1,t})^2}{n}} + \ldots + \sqrt{\frac{\sum_{t=0}^{n}(\hat{q}_{7,t}-q_{7,t})^2}{n}}}{7} \right] \tag{3.19}$$

where $\hat{q}_{1,t}$ refers to estimated joint angle for the first DOF at time t using the EKF-based pose estimation and $q_{1,t}$ refers to the same angle estimated from marker information. Marker-based estimations are extracted by applying inverse kinematics to the developed lower body model using the approach proposed by Joukov [45].

This minimization problem was solved by global search optimization implemented using the MatlabR2016a optimization toolbox. Global search applies the fmincon solver, which is a gradient based solver started from a number of random initializations. If the new point returns smaller cost, then it keeps that point and removes the previous guess. At the end, it returns the best global solution. Starting from a good initial guess helps with faster convergence in optimization. Such a guess was obtained by defining very small or zero values for α_1 , α_2 and considering the described relative relationship between other parameters in assigning values.

It should be noted that the optimal parameters obtained for one subject may not necessarily be optimal for a different subject. To report the best achievable accuracy of the method, tuning was repeated for each subject separately and subject-specific optimal parameters were used in reporting the errors. In general, however, when there is no reference information available, an average value of the all subjects' parameters can be utilized.

3.3 IMUs Characteristics

The IMUs used for data collection were obtained from Yostlabs [6] and are depicted in Figure 3.2.

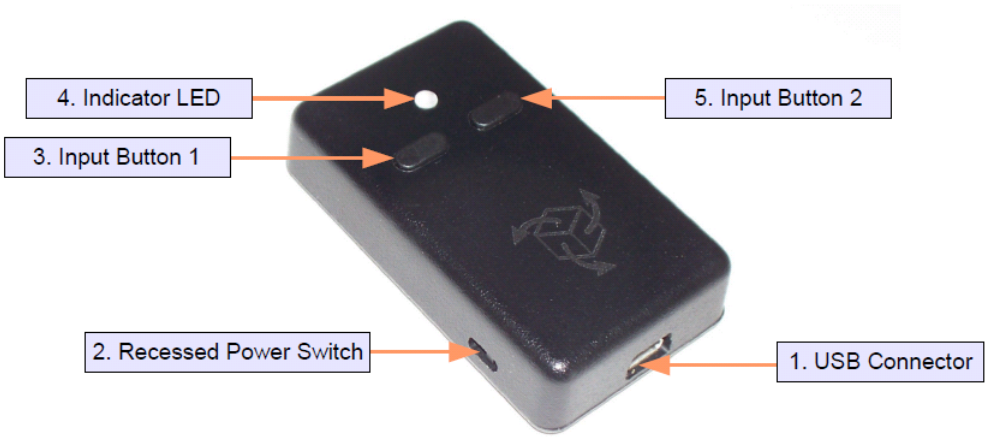


Figure 3.2: Yost IMU used for data collection [6]

Each IMU consists of a triaxial accelerometer, gyroscope and compass sensors with a 2.4GHz DSSS communication interface and a rechargeable lithium-polymer battery. The IMUs include on-board filtering algorithms. Calibration was possible through the 3 Space Suite firmware which allowed for calibrating accelerometer, gyroscope, or compass using a gradient-descent-based algorithm and an easy to use wizard (Figure 3.3)[1].

Moreover, it was possible to have access to various types of outputs including raw/corrected or normalized data and filtered absolute and relative orientation outputs in quaternion, Euler angles (pitch/roll/yaw), or rotation matrix. It was also possible to define the desired frame for each IMU and save output data in CSV format via an interface developed by MSK-Metrics. To increase the sampling rate, we used one dongle for each sensor.

The specifications of the gyroscope and accelerometer are listed in Figure 3.4. The magnetometer was not used in this study.

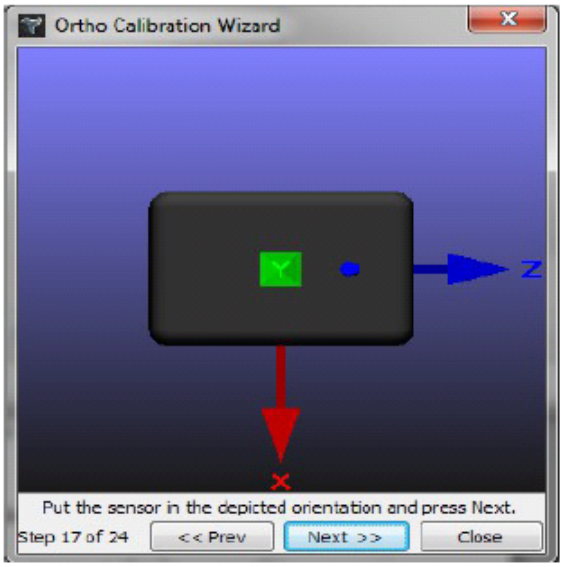


Figure 3.3: Yost IMU wizard for accelerometer calibration[1]

Sensor	
Orientation range	360° about all axes
Orientation accuracy[2]	±1° for dynamic conditions & all orientations
Orientation resolution	<0.08°
Orientation repeatability	0.085° for all orientations
Accelerometer scale	±2g / ±4g / ±8g selectable for standard models ±6g / ±12g / ±24g selectable for HH models ±100g / ±200g / ±400g selectable for H3 models
Accelerometer resolution	14 bit, 12 bit(HH), 12 bit(H3)
Accelerometer noise density	99µg/√Hz, 650µg/√Hz(HH), 15mg/√Hz(H3)
Accelerometer sensitivity	0.00024g/digit-0.00096g/digit 0.003g/digit-0.012/digit(HH) 0.049g/digit-0.195g/digit(H3)
Accelerometer temperature sensitivity	±0.008%/°C, ±0.01%/°C(HH, H3)
Gyro scale	±250/±500/±1000/±2000 °/sec selectable
Gyro resolution	16 bit
Gyro noise density	0.009°/sec/√Hz
Gyro bias stability @ 25°C	2.5°/hr average for all axes
Gyro sensitivity	0.00833°/sec/digit for ±250°/sec 0.06667°/sec/digit for ±2000°/sec
Gyro non-linearity	0.2% full-scale
Gyro temperature sensitivity	±0.03%/°C

Figure 3.4: Yost IMU specifications [1]

3.4 Experiments

To evaluate the accuracy of pose estimation, SLS data was collected with marker-based motion capture (MOCAP) and IMUs simultaneously. 10 participants (5 men, 5 women, mean age: 28.5 ± 6.37) were recruited. Inclusion criteria were adults without any lower back or leg injuries within the past six months. The experiment was approved by the University of Waterloo Research Ethics Board, and all participants signed a consent form prior to the start of data collection. Data from 3 participants were excluded from the analysis due to unexpected local frame change in the sensors which was identified after collection.

3.4.1 Mocap Data Collection

Three Yost [6] IMUs were attached to the participants' low back at the level of the first sacral vertebra, the anterior thigh 10 cm above the patella aligned with the sagittal plane, and the flat surface of the shank at the level of the tibial tuberosity using hypoallergenic tape. Sensor placement locations are illustrated in Figure 3.6. Data was communicated to a nearby computer via Bluetooth communication with an average sampling rate of 90 ± 10 Hz. Data were interpolated and resampled to the same rate of 200 Hz (equal to the MOCAP camera frame rate) before subsequent analysis.

At the same time, 8 reflective markers were attached to bony landmarks including: right and left anterior superior iliac spine (ASIS), right and left posterior superior iliac spine (PSIS), medial and lateral knee, and medial and lateral ankle of the squatting leg. Moreover, three markers were attached to the thigh and tibia sensors to enable sensor orientation recovery from the marker data. Due to the vicinity of the back sensor to the right and left PSIS markers, attaching three markers on the back sensor resulted in marker swapping; hence only one marker was attached to the back sensor. MOCAP data collection was performed with eight Eagle cameras and Motion Analysis Cortex software for data collecting and pre-processing.

Figure 3.5: Sensor and marker placement for the single leg squat experiment in the Motion Capture Lab

Participants were instructed to remove their shoes and perform five continuous cycles of SLS with their toes pointing forward and arms crossed in front of the body. They were asked to perform SLS with their dominant leg (the one they kick the ball with) without moving the foot or lifting the heel. In instances where subjects lost their balance, their legs contacted each other, or the non-weight bearing leg touched the ground, the trial was deemed unsuccessful and all cycles were repeated.

Before starting the exercise, subjects were asked to lift their squatting leg up and back down and then stay for a few seconds in a rest position. This additional motion was used for synchronization of the IMUs and MOCAP and was not included in the data analysis.

3.5 Forward Kinematic Parameter Estimation

In order for the kinematic model to be accurate, the displacement vectors have to be specified, which means that the exact position and orientation of the sensors are required.

3.5.1 IMU Orientations

Misorientation of the thigh and tibia sensors results from the fact that the leg anatomical orientation is not vertical while in the kinematic model (see Figure 3.1) a perfectly vertical position for these sensors is assumed. The back sensor is also assumed to be mounted perfectly horizontal with no angular offset with respect to the hip frame. The transformation matrices from the sensors' orientation on the body to the desired vertical orientation can be calculated by defining each sensor's frame using three markers on the sensor plane.

$$O_d = {}^{d}R_b O_b \tag{3.20}$$

Where O_d and O_b are the IMU desired and body mounted orientations, respectively. ${}^{d}R_b$ is the transformation from the body frame to the desired frame calculated by the average of marker data for the first five seconds before exercise begins, when participants were asked to stand still on both feet.

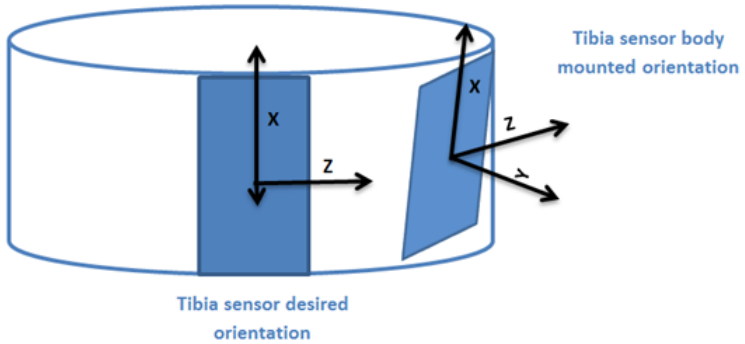


Figure 3.6: tibia sensor desired and body mounted orientations

${}^{d}R_b$ was calculated for each sensor using the three markers on each sensor plane. Each sensor measurement was rotated by ${}^{d}R_b$ before being input into the EKF. For the back

sensor, only one marker was mounted so it was not possible to extract back sensor misorientation from marker data.

3.5.2 Displacement Vectors and DH Link Lengths

Sensor location vectors and limb lengths were calculated using marker data of the first five seconds before exercise begins, when participants were asked to stand still on both feet. The displacement vectors were computed as follows:

$$\boldsymbol{r_{S3}} = \boldsymbol{P_{ankle}} - \boldsymbol{P_{tibia}} \tag{3.21}$$
$$\boldsymbol{r_{3}} = \boldsymbol{P_{ankle}} - \boldsymbol{P_{knee}} \tag{3.22}$$
$$\boldsymbol{r_{S4}} = \boldsymbol{P_{knee}} - \boldsymbol{P_{thigh}} \tag{3.23}$$
$$\boldsymbol{r_{4}} = \boldsymbol{P_{knee}} - \boldsymbol{P_{hip}} \tag{3.24}$$
$$\boldsymbol{r_{S7}} = \boldsymbol{P_{back}} - \boldsymbol{P_{hip}} \tag{3.25}$$

where $\boldsymbol{P_{ankle}}, \boldsymbol{P_{knee}}$, and $\boldsymbol{P_{hip}}$ refer to joint center position vectors and $\boldsymbol{P_{tibia}}, \boldsymbol{P_{thigh}}$, and $\boldsymbol{P_{back}}$ refer to IMU position vectors obtained from markers.

$\boldsymbol{P_{ankle}}$ and $\boldsymbol{P_{knee}}$ were calculated by computing the average of their medial and lateral markers.

$\boldsymbol{P_{tibia}}$ and $\boldsymbol{P_{thigh}}$ were obtained by computing the average of their corresponding three markers. The $\boldsymbol{P_{back}}$ was obtained from the marker attached to the back sensor.

For $\boldsymbol{P_{hip}}$ estimation, Harrington et al.'s method [37] was applied, which estimates the hip center location based on leg length (LL), pelvic depth (PD) and pelvic width (PW), shown in Figure 3.7. According to Harrington, LL is the distance from the right ASIS marker to the right medial ankle marker (averaged over first 5 seconds). In this study, all subjects were right legged so all figures and calculations are described for the right leg.

$$LL = |\boldsymbol{P_{RASIS}} - \boldsymbol{P_{RMA}}| \tag{3.26}$$
$$PD = |(\boldsymbol{P_{RASIS}} + \boldsymbol{P_{LASIS}})/2 - (\boldsymbol{P_{RPSIS}} + \boldsymbol{P_{LPSIS}})/2| \tag{3.27}$$
$$PW = |\boldsymbol{P_{RASIS}} - \boldsymbol{P_{LASIS}}| \tag{3.28}$$

Where, $\boldsymbol{P_{RMA}}$ refers to the right medial ankle marker position vector. $\boldsymbol{P_{RASIS}}$ and $\boldsymbol{P_{LASIS}}$ refer to right and left ASIS marker position vectors, and $\boldsymbol{P_{RPSIS}}$ and $\boldsymbol{P_{LPSIS}}$

refer to right and left PSIS marker position vectors, respectively. Thigh and shank link lengths were defined as the norm of $\boldsymbol{r_4}$ and $\boldsymbol{r_3}$ vectors, respectively (Figure 3.1). Hip link length was defined as the length of the X component of the $\boldsymbol{r_{S7}}$ vector.

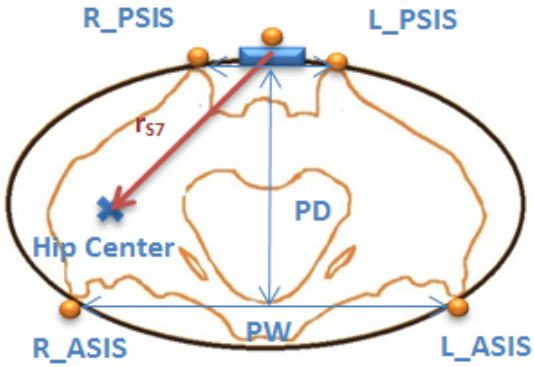


Figure 3.7: PD and PW calculation from markers

The calculated r vectors, orientation compensated sensor readings, and DH link lengths were then used in the pose estimation algorithm to recover the hip, knee, and ankle joint angles.

3.6 Pose Estimation Results

Table 3.1 reports RMS errors between MOCAP and joint angles estimated by the algorithm. To produce these results, all required parameters including displacement vectors and link lengths as well as thigh and tibia sensor orientations were extracted from markers. The only concern was that the back sensor orientation was unknown. We developed a calibration protocol to retrieve sensor orientations from gyro data which is explained in Section 4.4, this approach was used for the back sensor orientation. Figure 3.8 summarizes the workflow of the pose estimation algorithm used to generate the results shown in Table 3.1.

The joint angle estimates for the participant with the smallest error (subject 1) and largest error (subject 7) are depicted in Figures 3.9 to 3.14. Other participants' joint angle estimates can be found in Appendix A.

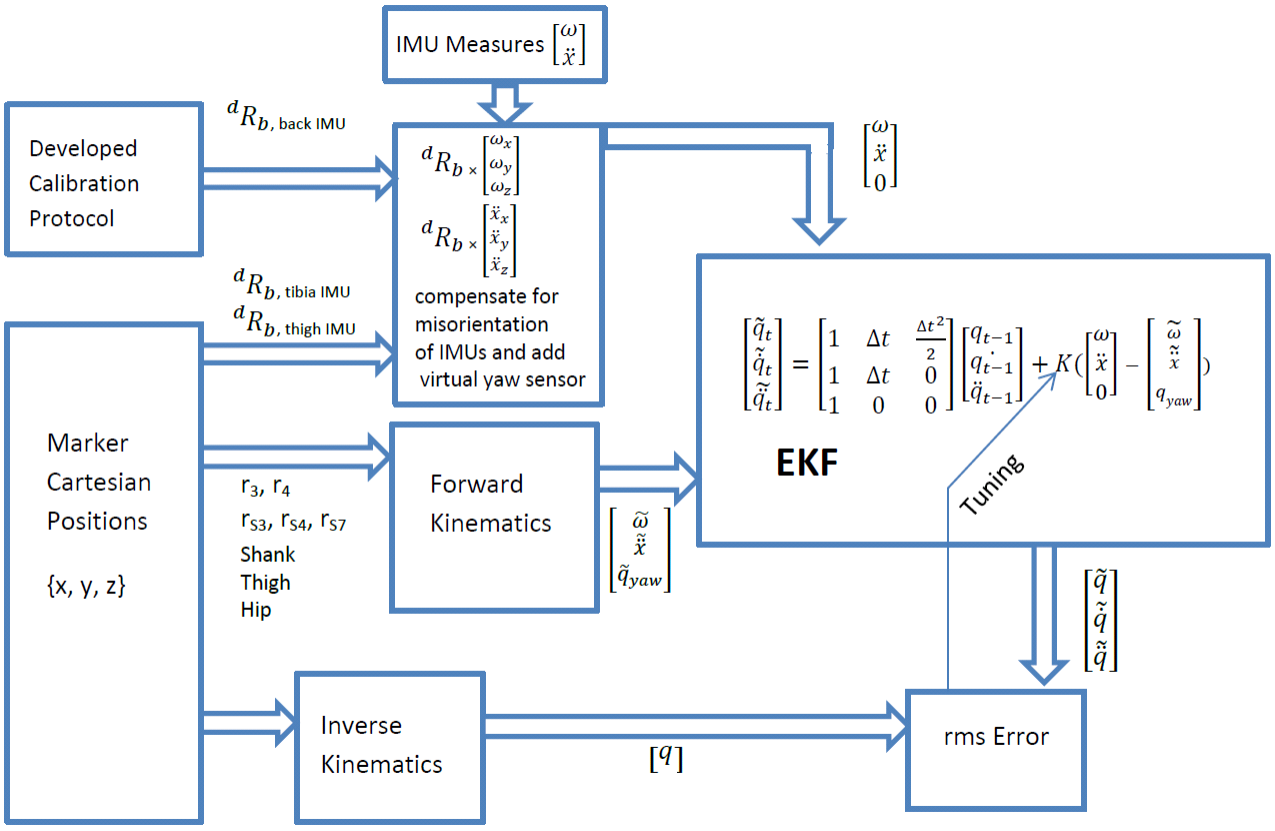


Figure 3.8: Pose estimation algorithm overview

Table 3.1: RMS error between IMU and Mocap when kinematic parameters and tibia and thigh sensor orientations are extracted from markers and back sensor orientation is obtained from the calibration protocol developed in Section 4.4

RMS error between Mocap and IMU-based estimated joint angles (degree)								
Subject	Ankle IR/Er	Ankle Abd/Add	Ankle Flex/Ext	Knee Flex/Ext	Hip Flex/Ext	Hip Abd/Add	Hip IR/ER	Average Error
1	2.6	1.3	3.2	4	5.5	2.1	3.3	3.1
2	2.4	2.4	3.9	4.8	14.3	3.1	2.5	4.8
3	1.6	2.4	4.9	5.7	11.9	5.1	4.6	5.2
4	3.8	1.4	5.2	7.5	15.3	3.7	6.3	6.2
5	5.2	4.4	2.3	4.8	9.4	5	5.5	5.2
6	5.1	1.7	3.5	3.6	6.5	8.4	6.4	5
7	2.6	2.5	4.2	8.3	13.5	3.9	10.2	6.4
Average	**3.3**	**2.3**	**3.9**	**5.5**	**10.9**	**4.5**	**5.5**	**5.1**

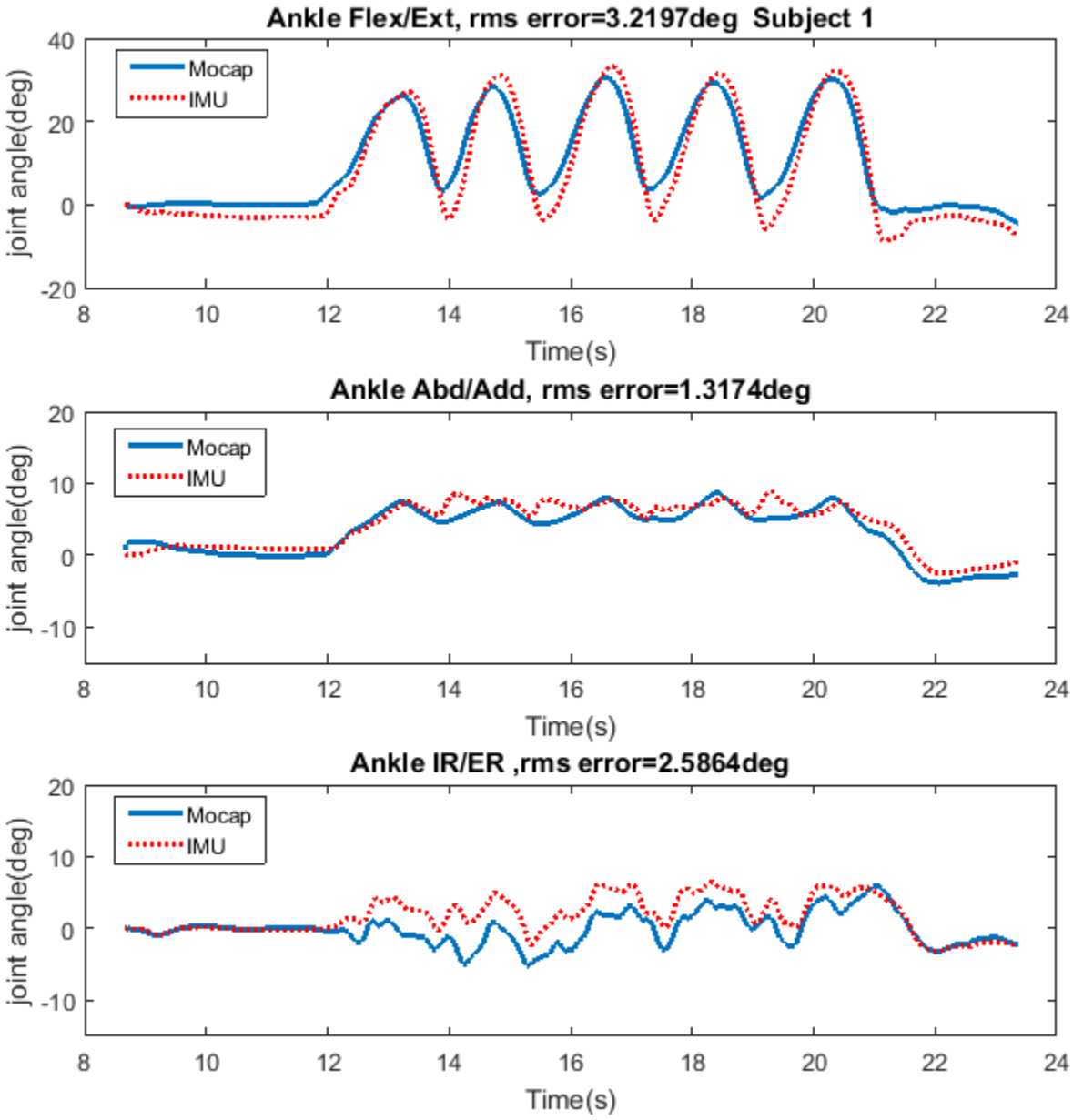


Figure 3.9: Subject 1 ankle joint angles RMS error using marker-derived values for tibia and thigh sensors orientation as well as displacement vectors

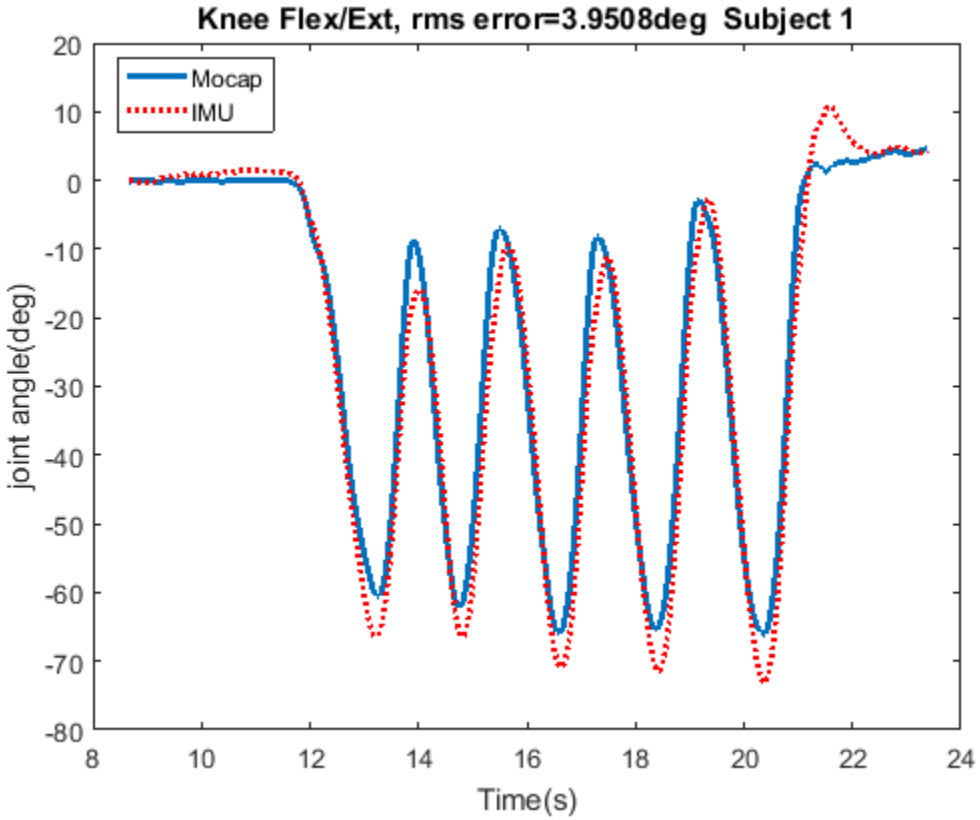


Figure 3.10: Subject 1 knee joint angle RMS error using marker-derived values for tibia and thigh sensors orientation as well as displacement vectors

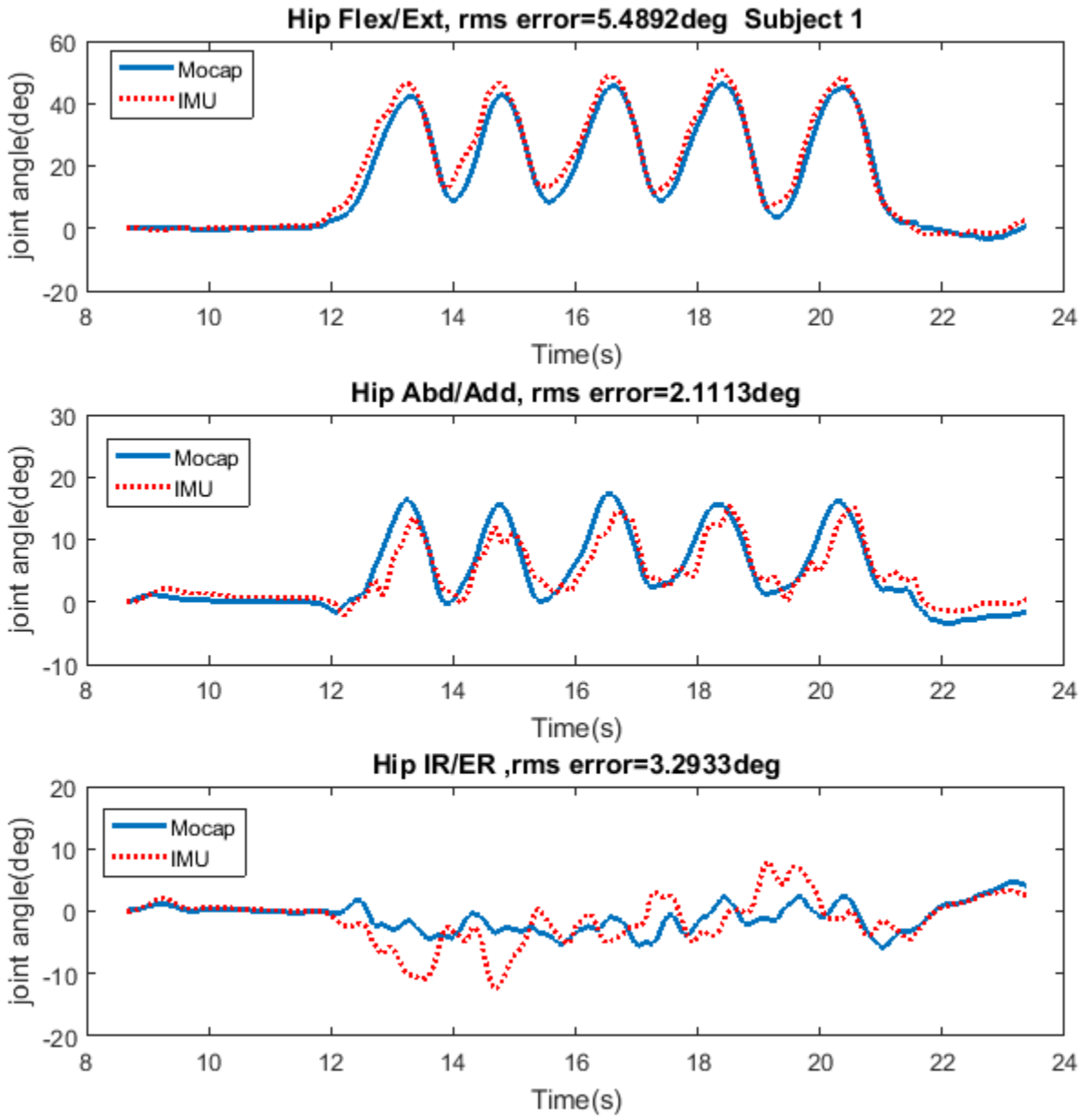


Figure 3.11: Subject 1 hip joint angles RMS error using marker-derived values for tibia and thigh sensors orientation as well as displacement vectors. Back sensor orientation obtained from calibration protocol developed in Section 4.4

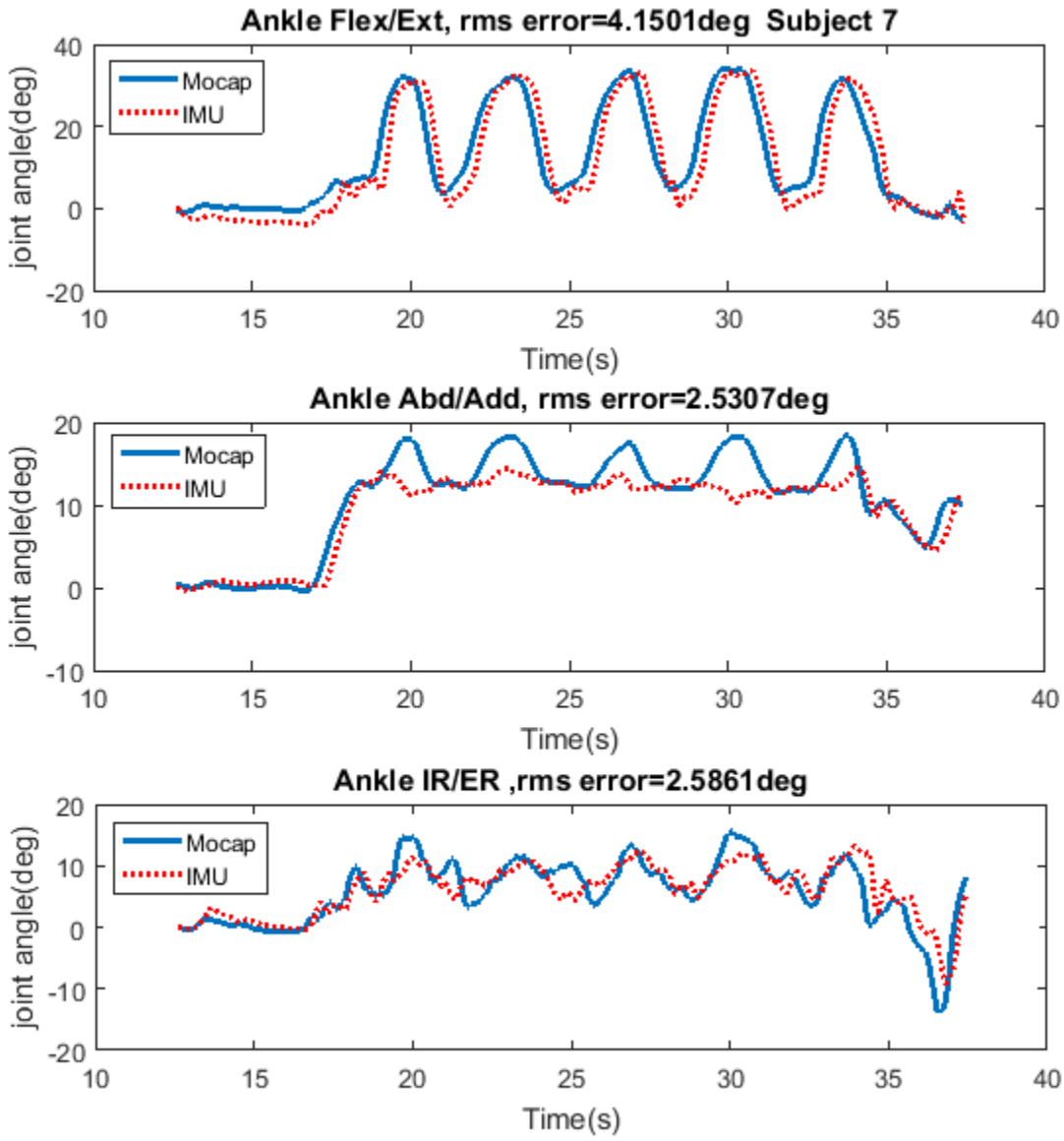


Figure 3.12: Subject 7 ankle joint angles RMS error using marker-derived values for tibia and thigh sensors orientation as well as displacement vectors

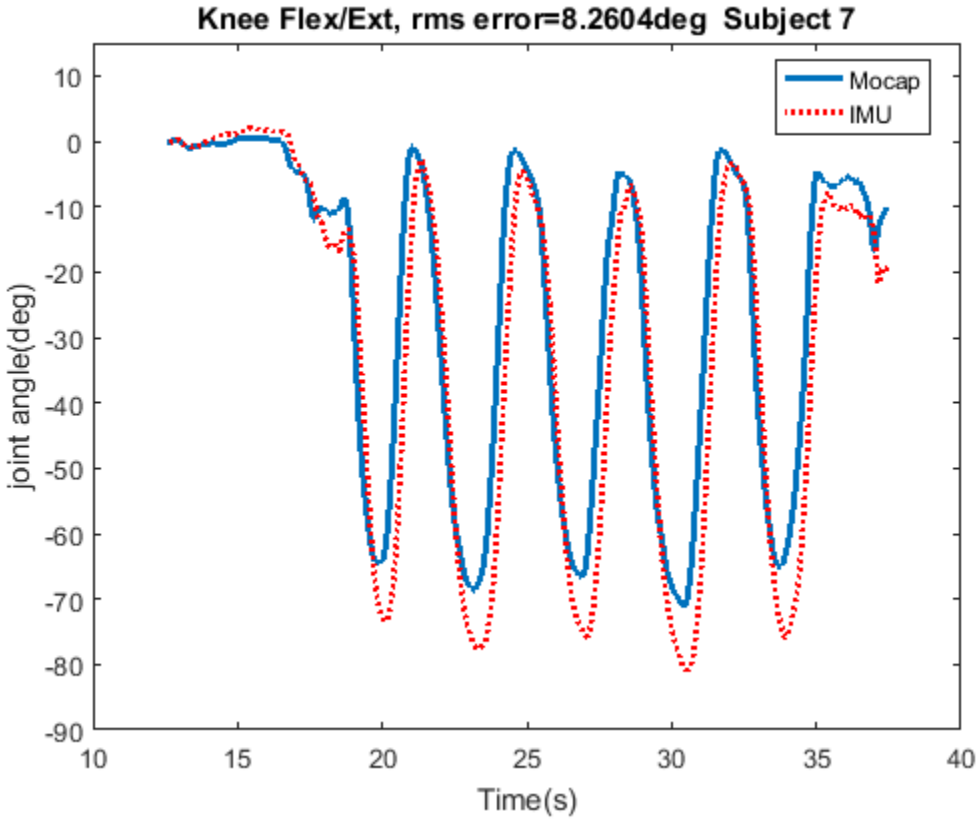


Figure 3.13: Subject 7 knee joint angle RMS error using marker-derived values for tibia and thigh sensors orientation as well as displacement vectors

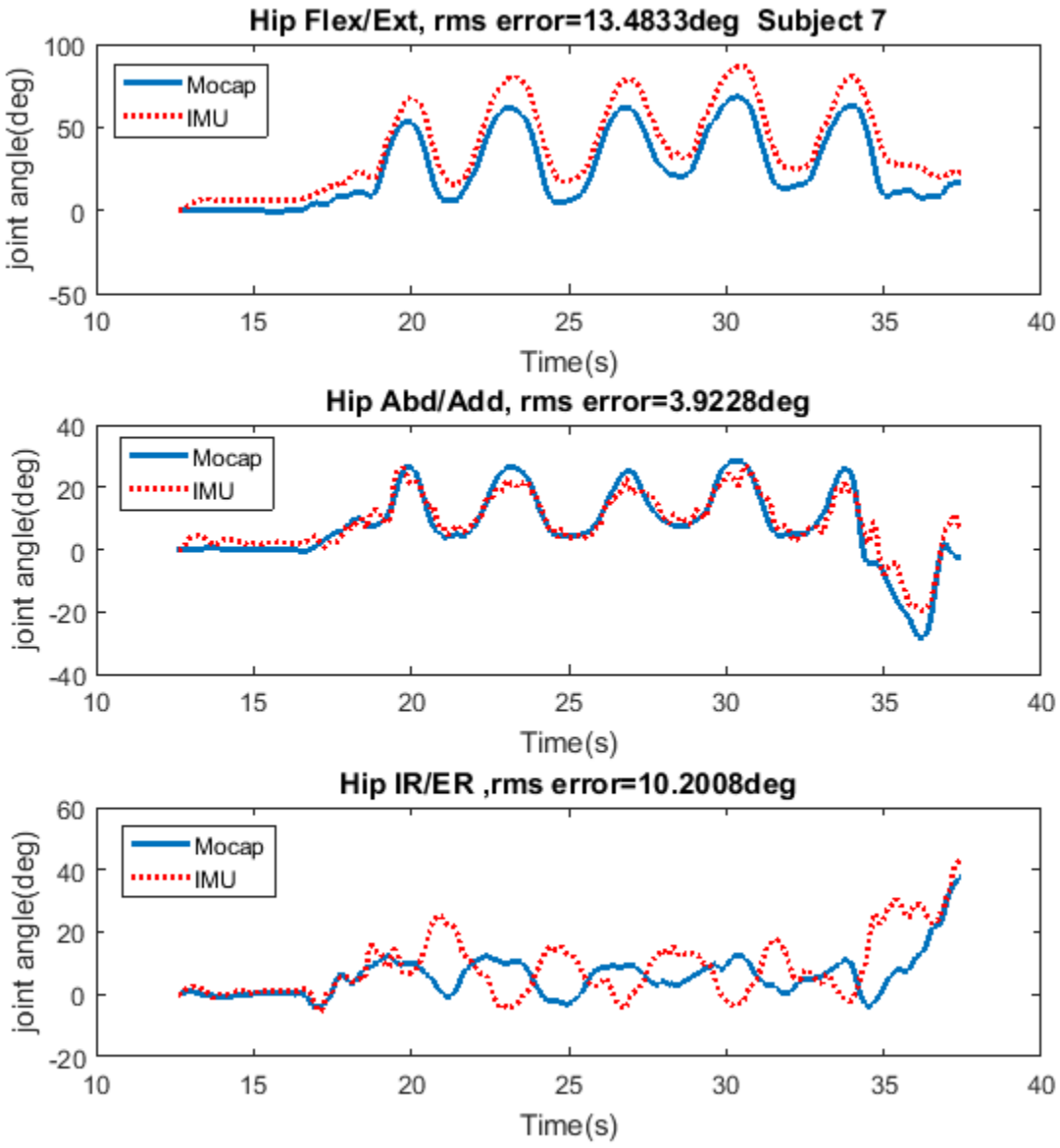


Figure 3.14: Subject 7 hip joint angles RMS error using marker-derived values for tibia and thigh sensors orientation as well as displacement vectors. Back sensor orientation obtained from calibration protocol developed in 4.4

Table 3.1 shows the results obtained when subject-specific optimal parameters were used for EKF noise statistics. In field application of the method, however, these optimal values are not available for each subject. Therefore, an approximate value should be used for these parameters. This will impact accuracy of the pose estimation. To examine the effect, we have averaged over the optimal parameters found for all subjects and used the same average value for all of them. Table 3.2 shows the estimated joint angle errors for this condition.

Table 3.2: RMS error between IMU and Mocap when the EKF noise parameter statistics are set to a fixed same value equal to average of optimal values of all subjects

RMS error between Mocap and IMU-based estimated joint angles (degree)								
Subject	Ankle IR/Er	Ankle Abd/Add	Ankle Flex/Ext	Knee Flex/Ext	Hip Flex/Ext	Hip Abd/Add	Hip IR/ER	Average Error
1	2.5	1.3	2.2	5.5	6.8	1.8	7.4	3.9
2	2.3	2.5	3.8	4.9	14.3	4.4	6.1	5.5
3	1.5	2.4	4.9	5.9	11.6	5.4	4.7	5.2
4	3.9	1.7	5.2	8.2	13.9	7.5	16.7	8.1
5	5.4	4	2.7	4.7	8.4	8	9.6	6.1
6	4.5	2.1	3.6	3.3	7.2	9.8	10.1	5.8
7	2.6	2.6	4.2	8	13.4	7.3	13.9	7.4
Average	**3.2**	**2.4**	**3.8**	**5.8**	**10.8**	**6.3**	**9.8**	**6**

In Section 3.2.2 we argued that 3-EKF implementation results were better than 1-EKF implementation results. Table 3.3 compares the results of pose estimation for subject 1 using both implementations. In order for the comparison to be fair, both methods were tuned separately and method-specific optimal noise parameter values were utilized. The results show almost the same error in all joint angles except IR angles. It can be observed from Figure 3.15 that despite using a hip virtual yaw sensor, the large observed error in IR angles in 1-EKF implementation is due to drift. The drift in the hip and ankle is enabled due to the redundancy of the 7 DOF model and the fact that these two angles are indeterminate when the leg is straight (i.e., when the leg is in a singular configuration). Since the hip IR/ER and ankle IR/ER are drifting in opposite directions, it may be possible that if we add another virtual yaw sensor to the ankle, the error will reduce. However, the virtual yaw sensor may suppress internal rotations which are real and not due to drift as well. Also it adds one more parameter to be tuned in the EKF. The 3-EKF configuration

achieves driftless results for the same subject (Figures 3.9, 3.11), so it is preferable.

Table 3.3: RMS error between IMU and Mocap for Subject 1 compared between 1-EKF and 3-EKF implementations

RMS error between Mocap and IMU-based estimated joint angles (degree)								
Method	Ankle IR/Er	Ankle Abd/Add	Ankle Flex/Ext	Knee Flex/Ext	Hip Flex/Ext	Hip Abd/Add	Hip IR/ER	Average Error
3-EKF	2.5	1.3	2.2	5.5	6.8	1.8	7.4	3.9
1-EKF	17.8	1.3	1.4	5.9	6.8	2.7	20.2	8

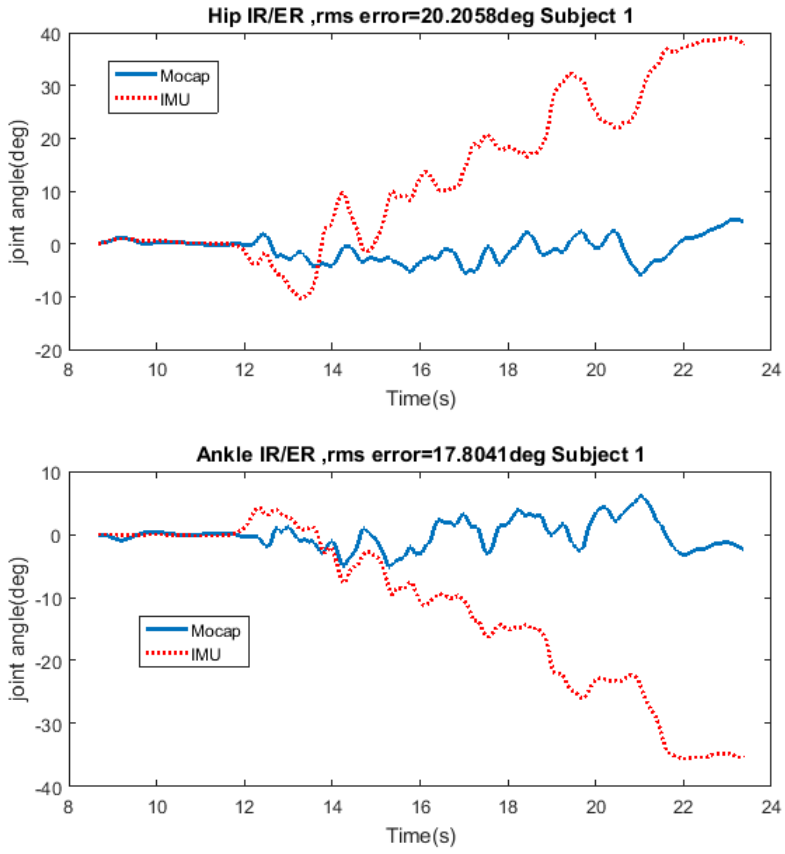


Figure 3.15: Subject 1 hip and ankle IR joint angle estimation with 1-EKF implementation of the algorithm

3.7 Discussion

Table 3.1 shows successful extraction of joint angles. The average estimated errors for the ankle, knee and hip joints of seven participants compared to optical measurement and using marker-extracted kinematic model parameters are 3.2°, 5.5°, and 7°, respectively. The total estimated error averaged over all subjects and all angles is 5.1°, which is comparable to similar IMU-based pose estimation studies [81], [24], [27].

These results were based on using subject-specific optimal EKF noise parameters. How-

ever, Table 3.2 shows that if the same fixed value obtained by taking the average of all subjects' optimal values for each parameter was used instead, the error increase is 1° which is still promising. Comparing the results of Table 3.1 to Table 3.2 shows that the hip IR joint angle is most affected by the change in the filter noise parameters, which was expected as it is the last parameter in the kinematic chain and its error cannot be compensated by the accelerometer.

The reason for achieving better results for the the 3-EKF structure is that the 7 DOF model is a kinematically redundant structure, and that the SLS motion occurs near the leg singularity. In the singular configuration (when the leg is fully extended), motion in the hip and ankle IR joints is aligned and drift can occur in opposing directions. By separating the hip and ankle models, this source of error is removed. A second cause of error can be due to the fact the 7 DOF model has a much larger parameter space, requiring longer for optimization and introducing the possibility of more local minima. IR angle estimation is more sensitive to poor process noise selection because its estimation is based on model states and gyro readings. For other states, there is corrective information available from the accelerometer which improves the results.

Comparing initial values of the filter parameters with the average of optimum values over all subjects (reported in Table 3.4) showed that the optimal values of $\beta_2 > \beta_1$ and $\alpha_3 > \alpha_2 > \alpha_1$ but the order of magnitude for the γ parameters was not the same as the initial setting. The obtained order is: γ_7 $(hip\ IR) > \gamma_2$ $(ankle\ Add) > \gamma_4$ $(knee\ Flex) > \gamma_5$ $(hip\ Flex) > \gamma_6$ $(hip\ Add) > \gamma_1$ $(ankle\ IR) > \gamma_3$ $(ankle\ Flex)$. We expected to see the largest coefficients for the flexion angles $(\gamma_3, \gamma_4, \gamma_5)$ and the smallest ones for the IR angles (γ_1, γ_7) and set the initial parameters to reflect that. However, the optimal ankle flexion coefficient was lower, while the hip IR coefficient was higher than our a priori initialization. For the hip IR, we hypothesize that model uncertainty was increased due to the distality of the hip IR from the base of the kinematic chain as well as due to the hip center approximation, which resulted in a large value of the process noise model for γ_7. Ankle Flex/IR were more reliable than all other angle estimates because of the proximity to the base frame.

Table 3.4: Optimal values obtained for EKF noise statistics averaged over all subjects

Optimal values obtained for EKF noise statistics averaged over all subjects						
Gyro	Accel	VirYaw	q	$\dot{q}$	$\ddot{q}$	
β_1	β_2	β_3	α_1	α_2	α_3	
18.6 ± 5.1	69.5 ± 3.8	46.9 ± 4.5	0	0.001	78.7 ± 0.6	
ankle IR	ankle Add	ankle Flex	knee Flex	hip Flex	hip Add	hip IR
γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7
6.4 ± 6	13.1 ± 5.4	0.5 ± 0.2	10.8 ± 5	9.1 ± 8.8	8.8 ± 5.1	28.7 ± 0.8

α_1 and α_2 always hit the lowest boundary so they were set to low fixed values

The highest errors in our algorithm correspond to hip joint angles and particularly hip flexion. Looking into hip Flex/Add estimates of all subjects (Appendix A) reveals that a big portion of this error is due to a fixed offset. The main reason for this offset is the inaccurate extraction of sensor orientation on the body. The exact estimation of back sensor orientation from markers was not possible. Although the estimated orientation from the calibration protocol helps in improving the hip joint angle estimations, it is not yet as accurate as marker-derived orientation. Moreover, since the hip center location is not directly measurable, approximation of the hip center based on other estimated parameters such as the pelvic depth, width and leg length also contributes to decreased accuracy.

In general, estimated flexion angle errors are higher because the motion is mainly performed in the sagittal plane and flexion angles have the biggest ROM in SLS. Also knee flexion has higher error than ankle flexion because of bigger ROM.

Currently available tools limit clinicians to passive ROM measurements. Reliability of goniometric ROM depends on the joint and type of motion [65]. Due to the subjective nature of visual assessment or goniometry, reliability of the measurements or visual assessments is an issue [65].

The proposed method in this study offers significant benefits to clinicians as it provides objective and therefore reliable pose measurements during dynamic movements. This enables not only estimation of ROM but also maximal flexion and extension associated with that ROM as well as information about velocity and acceleration.

Reported mean errors of visuial assessment in Rachkidi et al. are up to 6° for hip Add/rotation angles and up to 3° in flexion angles. Edwards et al. [26] have reported

visual knee flexion error of 5°. Allington et al. [8] also reported $5° \pm 5$ error for ankle Flex/Ext visual assessment. Our results show $3.9° \pm 1$ for ankle Flex/Ext, $5° \pm 2.3$ for hip Add/IR and $5.5° \pm 1.8$ and $10.9° \pm 3.9$ for knee and hip flexions, respectively. This ankle error is better than visual assessment, knee flex and hip Add/IR error is almost the same but hip Flex is higher than the reported values for human visual assessment capabilities. Considering the fact that the comparison is made between passive clinical ROM reports and active pose value measurements, the achieved results are promising.

Moreover, examining the flexion errors particularly hip error, revealed that most of the error is related to a fixed offset (see Appendix A). This means that if an absolute joint angle value is not required but range of motion is desired (which is the case in many clinical applications), a more accurate estimation for ROM may be attainable with our method.

3.8 Summary

In this chapter, an IMU-based pose estimation algorithm was developed. The proposed approach addresses many common issues related to IMUs including drift removal, estimation of joint angles in all directions, compensation for sensor misorientation and direct extraction of joint angles.

To achieve the above mentioned objectives, we combined kinematic chain modelling with an Extended Kalman Filter. The Kinematic model helps to estimate 3 dimensional and physically realizable joint angles as well as reducing the drift. EKF also improves accuracy by filtering sensor measurements. By defining joint angles, velocities and accelerations as EKF states, we can estimate these parameters directly. For better drift removal from hip internal rotation, a virtual yaw sensor with zero measurement was added to the observation vector. To address sensor to segment alignment, three markers were placed on tibia and thigh sensors and exact sensor orientations were extracted from these markers.

One of the limitations of the proposed method is that the EKF includes parameters such as the process and measurement noise covariances, which need to be carefully selected. Process noise reflects model inaccuracies such as the constant acceleration assumption, while the measurement noise reflects measurement inaccuracies. Poor selection of these values will result in inaccurate estimations. These parameters were initialized using apriori knowledge of the motion and further tuned by optimization.

Another limitation is that the location of each IMU on the link (displacement vector) needs to be determined for the pose estimation algorithm. The same is true for sensor orientations on the body. When marker data is available, the displacement vectors and sensor

orientations can be obtained from the marker data, as explained in this chapter. However, when marker data is not available, as would be the case in clinical use, translations and orientations have to be identified using manual measurements and calibration procedures. This limitation is addressed in the next chapter.

Chapter 4

Sensitivity Analysis

Clinicians usually have a busy appointment schedule, which does not allow time for extensive sensor calibration or accurate measurement of the required parameters for pose estimation (e.g sensor orientations and segment lengths) for each individual. One solution is to only measure those parameters whose variation affects the pose estimation accuracy the most and to use anthropometric data (e.g., [60], [31]) for less sensitive parameters. Therefore, a sensitivity analysis is required to identify the most sensitive parameters influencing the accuracy of pose estimation.

While several studies have mentioned the importance of sensor positioning and suggested methods for more accurate estimation of the sensor orientation [28], [78], [54], very few studies have quantified the sensitivity of pose estimation to sensor misplacement [85], [52]. Even in these studies, the main focus was pose estimation; sensitivity analysis was done for limited scenarios as a subsidiary part.

In this chapter, the sensitivity of pose estimation to the variations resulting from inaccurate sensor placement is quantified, and a practical protocol for estimating the sensitive parameters in clinical settings is proposed.

Before performing sensitivity analysis, a list of parameters for which sensitivity needs to be measured has to be identified. These parameters include those needed for the forward kinematic model, including the displacement vectors ($r_3, r_{S3}, r_4, r_{S4}, r_{S7}$), link lengths ($L_{shank}, L_{thigh}, L_{hip}$), and sensor orientations. If marker data is available, displacement vector elements can be obtained directly. However, in clinical applications the marker data is not available. Therefore, an alternative method is needed to approximate these quantities. For this reason, auxiliary parameters such as the pelvic width, depth, leg length, or leg

circumference are needed. Therefore, the list of parameters required is first identified in this chapter, and subsequently sensitivity analysis is applied.

4.1 Kinematic Model Parameter Approximation

Since the IMU-based pose estimation algorithm is to be applied in the absence of marker information, an estimation method for the sensor orientation, displacement vectors, and kinematic link lengths has to be defined.

4.1.1 Approximating the Displacement Vectors

To obtain displacement vectors and kinematic link lengths, one approach could be manual measurement of vector elements, which is time consuming and may be dependent on clinician expertise. An alternative approach is to make assumptions about sensor positions and replace these measures by an average value from available anthropometric tables.

If we assume that sensors are placed exactly in the middle of the shank and thigh and that the knee and ankle centers are aligned, then the X, Y, Z components of the displacement vectors with respect to the sensor frame can be approximated according to Table 4.1. In addition, if the leg is assumed to have a conical or cylindrical shape, the Y component of r_{S3} and r_{S4} would be equal to the radius of the leg at the sensor location.

$Vector$	X	Y	Z
$\boldsymbol{r_{S3}}$	$-(L_{shank} - L_{Tib2knee})$	$-r_{shank} = -(\frac{C_{Tib}}{2\pi})$	0
$\boldsymbol{r_3}$	$-L_{shank}$	0	0
$\boldsymbol{r_{S4}}$	$-L_{Thi2knee}$	$-r_{thigh} = -(\frac{C_{Thi}}{2\pi})$	0
$\boldsymbol{r_4}$	$-L_{thigh}$	0	0

$L_{shank} : Shank\ length$
$L_{thigh} : Thigh\ length$
$r_{shank} : Shank\ radius\ at\ tibia\ sensor\ location$
$C_{Tib} : Shank\ circumference\ at\ tibia\ sensor\ location$
$C_{Thi} : Thigh\ circumference\ at\ tibia\ sensor\ location$
$r_{thigh} : Thigh\ radius\ at\ thigh\ sensor\ location$
$L_{Tib2knee} : Vertical\ distance\ from\ middle\ of\ the\ tibia\ sensor\ to\ the\ knee\ center$
$L_{Thi2knee} : Vertical\ distance\ from\ middle\ of\ the\ thigh\ sensor\ to\ the\ knee\ center$

Table 4.1: Sensor displacement vectors approximated based on thigh and shank lengths, sensor vertical distance to previous joint center and leg radius at sensor location.

$L_{Tib2knee}$ and $L_{Thi2knee}$ distances have to be specified during sensor placement but have a fixed value for all subjects.

The $\boldsymbol{r_{S7}}$ vector can be estimated as the summation of $\boldsymbol{V1}$ and $\boldsymbol{V2}$ vectors as depicted in Figure 4.1. The Harrington method [37] gives the vector $\boldsymbol{V2}$, for which LL, PD, and PW need to be measured.

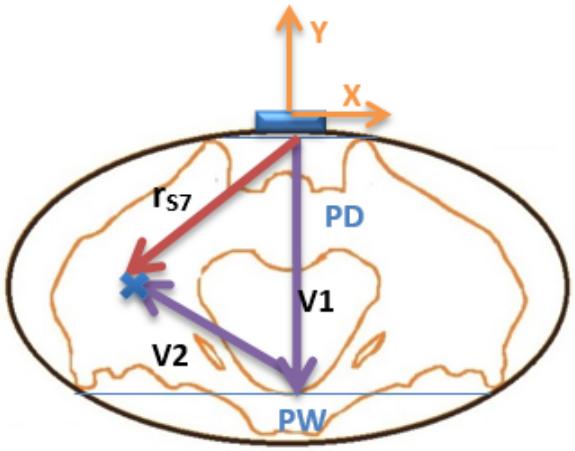


Figure 4.1: The $\boldsymbol{r_{S7}}$ vector can be estimated as the summation of vectors $\boldsymbol{V1}$ and $\boldsymbol{V2}$ where $\boldsymbol{V2}$ is estimable using PD, PW and LL. $\boldsymbol{V1}$ is assumed to have only Y component equal to PD.

Given that the back sensor is placed at the midpoint between the right and left ASIS, the Z and X components of $\boldsymbol{V1}$ are assumed to be zero and the Y component is equal to PD with negative sign (according to the back sensor frame shown in the figure), fully defining $\boldsymbol{V1}$.

Using the above mentioned assumptions, the required measurements are: LL, PD, PW, shank length, thigh length, tibia and thigh circumference at sensor locations (C_{Tib}, C_{Thi}), and the tibia and thigh sensor distances to the knee ($L_{Tib2Knee}, L_{Thi2Knee}$). These parameters are shown in Figure 4.2.

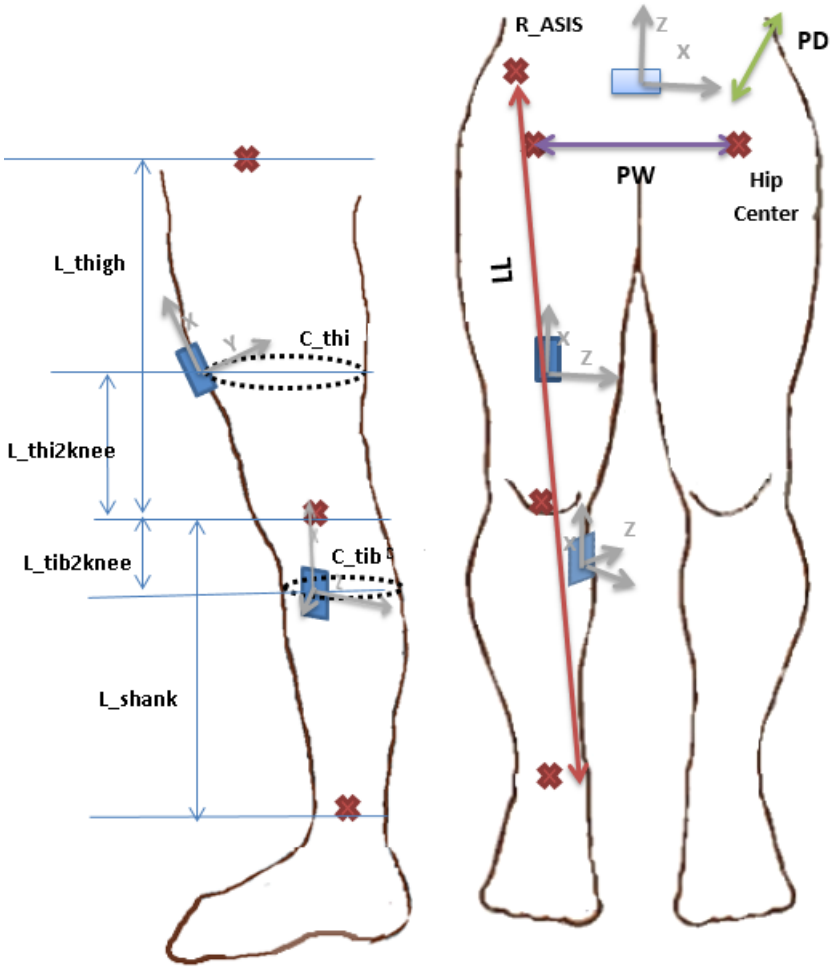


Figure 4.2: Required parameters for pose estimation including PW, PD, LL, shank length, thigh length, tibia and thigh circumference at sensor locations (C_{Tib}, C_{Thi}), and the tibia and thigh sensor distances to the knee ($L_{Tib2Knee}, L_{Thi2Knee}$).

To avoid manual measurement of these values, the pelvic width, shank and thigh lengths were estimated as a fraction of participant height following Winter et al.[95] (Figure 4.3). The circumference of the thigh and tibia at sensor locations were estimated as the "MidThigh Circumference" and "Maximal Calf Circumference" values from McDowell et al. [60]. For the four remaining values, the average value of all participants was used for analysis.

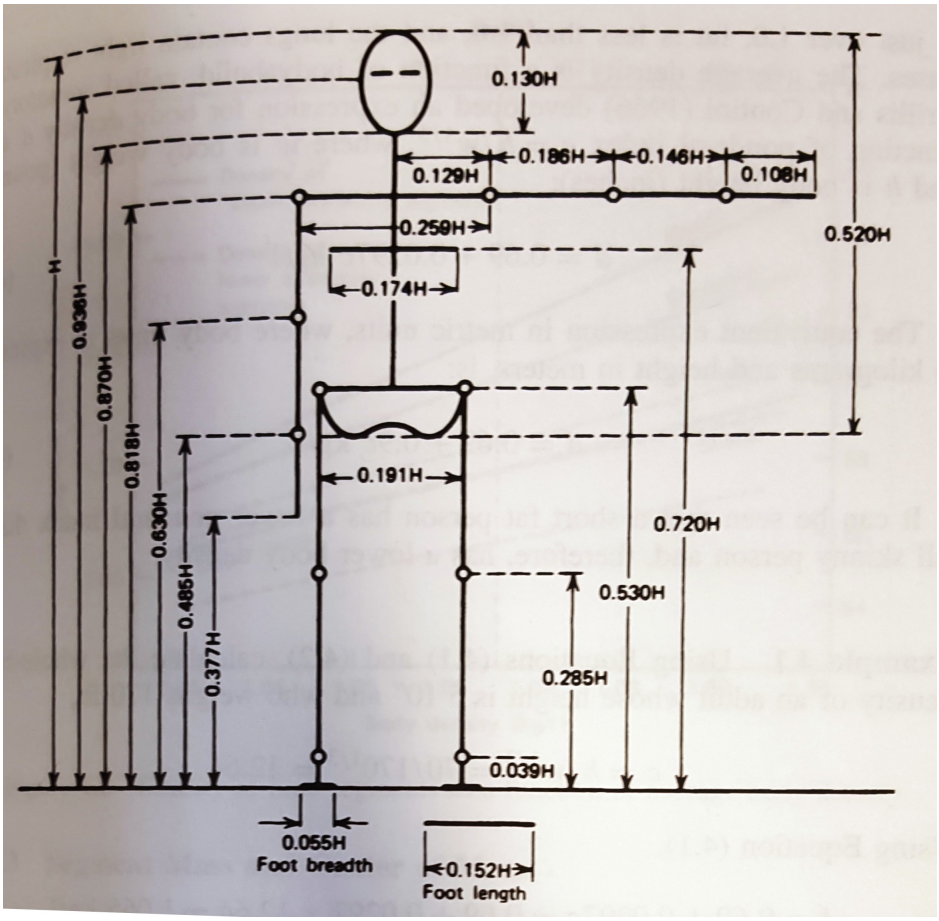


Figure 4.3: Body segment lengths expressed as fraction of body height (H)[95]

4.1.2 IMU Orientations

We assume that the back and thigh sensors are perfectly aligned with the sagittal plane. However; since the tibia sensor is placed on the flat part of the tibia, it has a significant offset angle from the frontal plane about the X axis which cannot be neglected. We assume that it is a fixed value for all participants and set it to $-45°$ which is the roll angle value we obtained from marker data averaged over all subjects. Therefore, in the absence of marker data, we only compensate for the tibia sensor roll rotation and assume that the other possible misorientations can be neglected.

4.2 Pose Estimation using Approximated Parameters

Table 4.2 shows the RMS error for the pose estimation when approximated values described in Section 4.1 are used in the pose estimation algorithm. Referring to Figure 3.8, we have removed the marker information block as well as the calibration block and approximated the kinematic model lengths, and the sensor displacement vectors, and orientations instead.

Table 4.2: RMS error between IMU and Mocap when kinematic parameters are approximated from anthropometric tables and only tibia sensor orientation is compensated for using a fixed shift angle

RMS error between Mocap and IMU-based estimated joint angles (degree)								
Subject	Ankle IR/Er	Ankle Abd/Add	Ankle Flex/Ext	Knee Flex/Ext	Hip Flex/Ext	Hip Abd/Add	Hip IR/ER	Average Error
1	2.9	1	10.1	11.9	29.1	7	7.8	10
2	3.2	2.3	6.9	17.1	11.6	5.5	7.3	7.7
3	1.8	1.5	3.8	7.9	16.4	4.8	5.1	5.9
4	5.9	8.7	2.4	12.5	7	10.9	17.6	9.3
5	6.6	7.8	16	9.3	12.5	12.1	6.2	10.1
6	10	12.8	4.3	18.8	6	8.2	9.1	9.9
7	4.4	10.8	3.8	12.3	4.7	9.5	10.9	8.1
Average	**5**	**6.4**	**6.8**	**12.8**	**12.5**	**8.3**	**9.2**	**8.7**

Joint angle estimates for participants 1 and 7 are depicted in Figures 4.4 to 4.9. These can be compared to the performance when using marker-measured parameters for the same participants depicted in Figures 3.9 to 3.14.

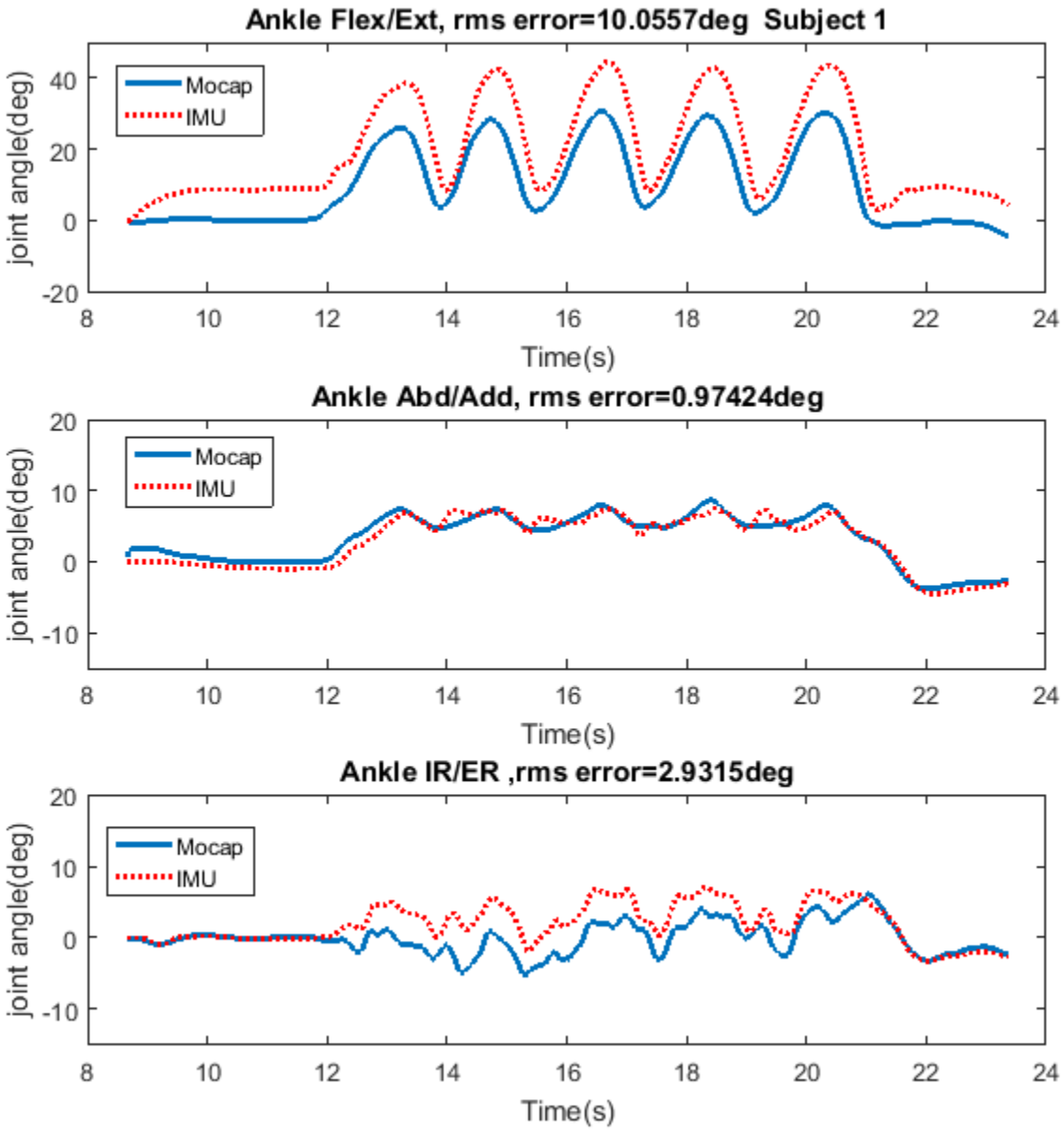


Figure 4.4: Subject 1 ankle joint angles RMS error using approximated values for tibia sensor orientation, kinematic link lengths as well as displacement vectors

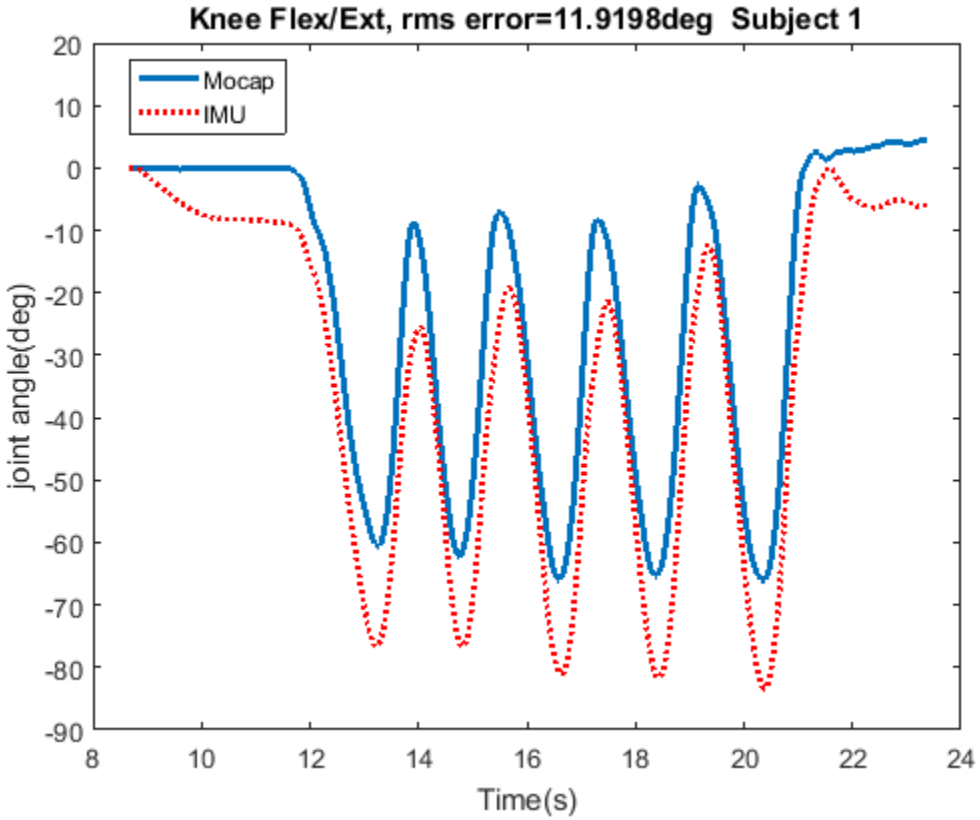


Figure 4.5: Subject 1 knee joint angle RMS error using approximated values for tibia sensor orientation, kinematic link lengths as well as displacement vectors

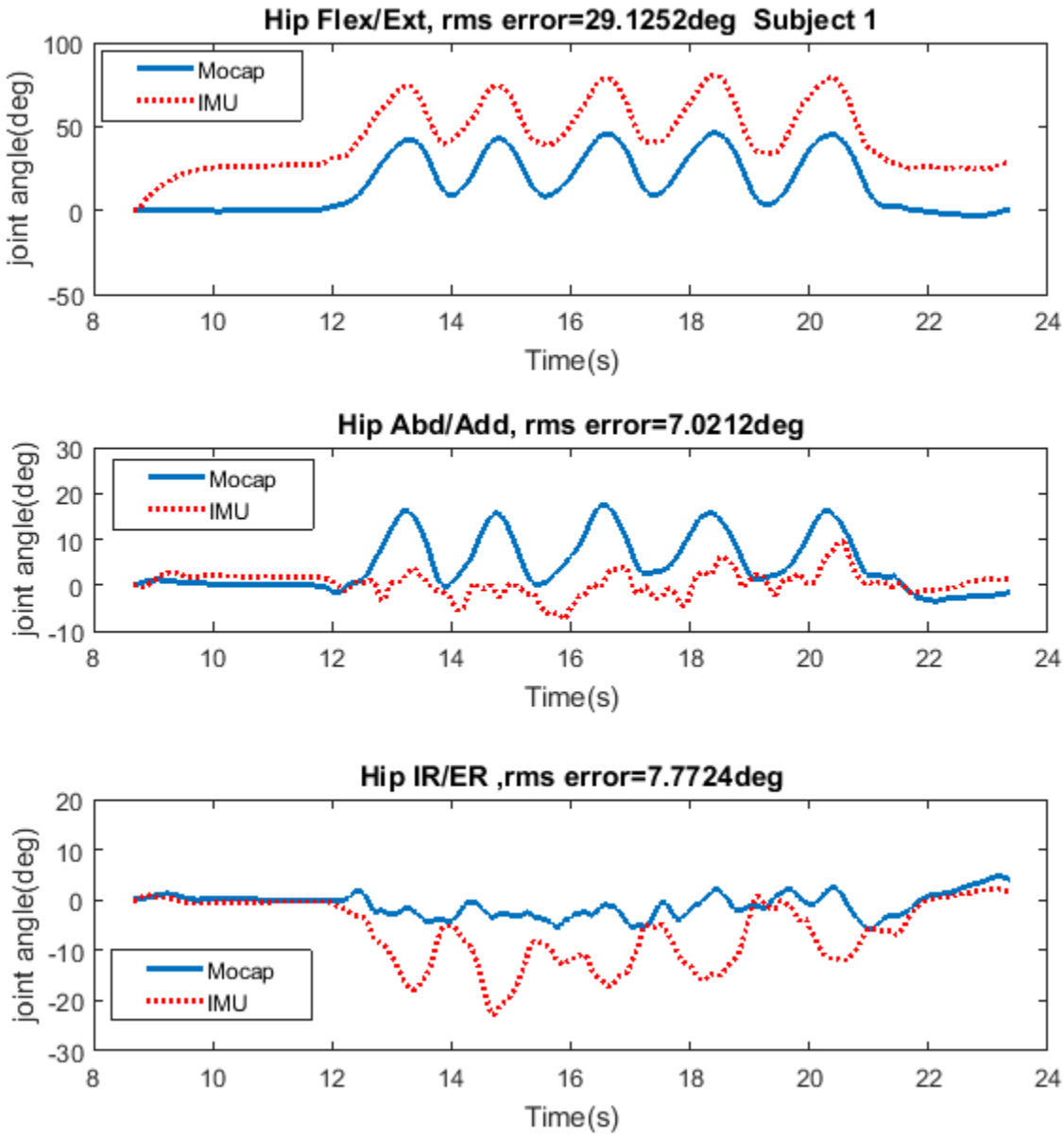


Figure 4.6: Subject 1 hip joint angles RMS error using approximated values for tibia sensor orientation, kinematic link lengths as well as displacement vectors

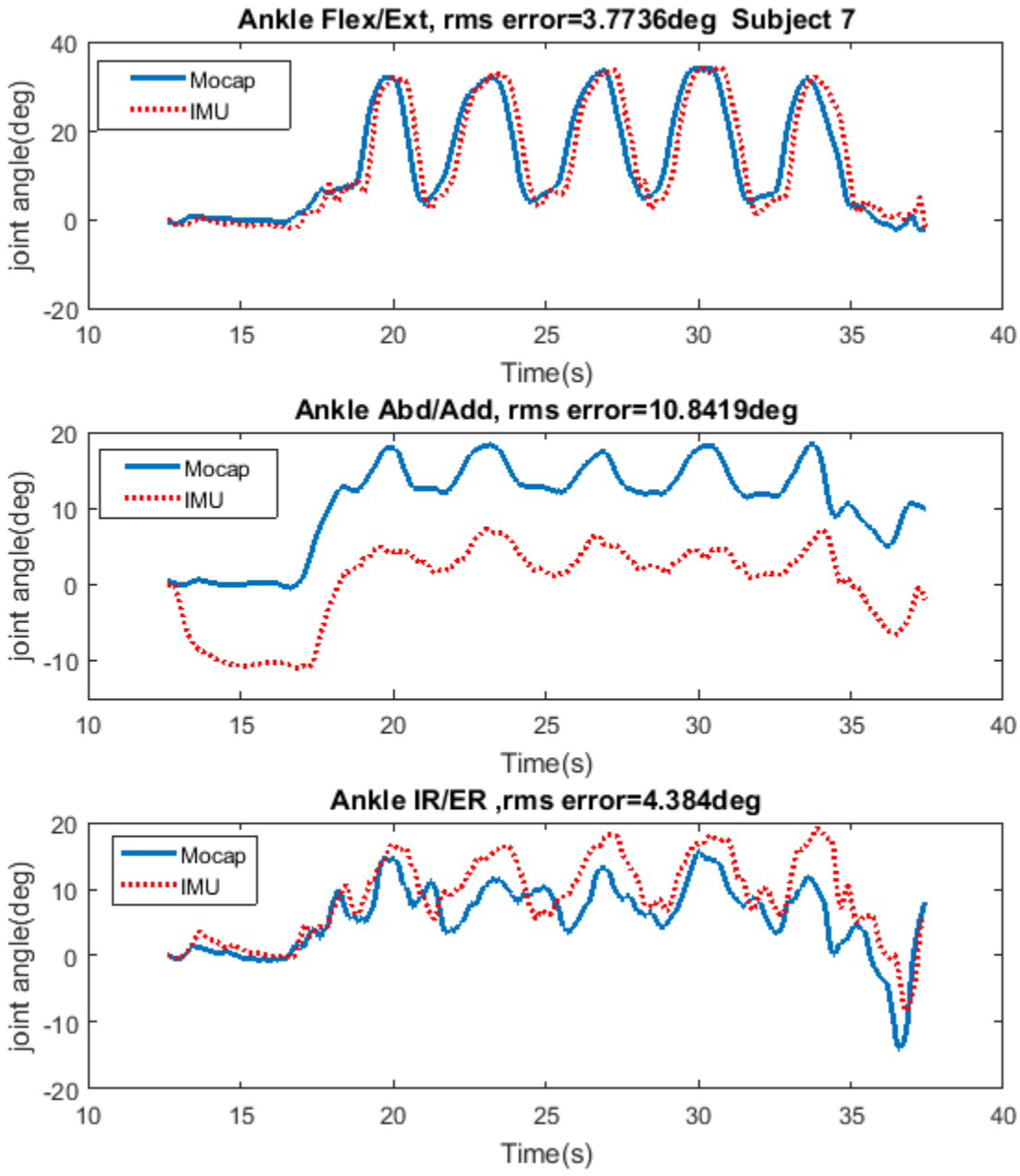


Figure 4.7: Subject 7 ankle joint angles RMS error using approximated values for tibia sensor orientation, kinematic link lengths as well as displacement vectors

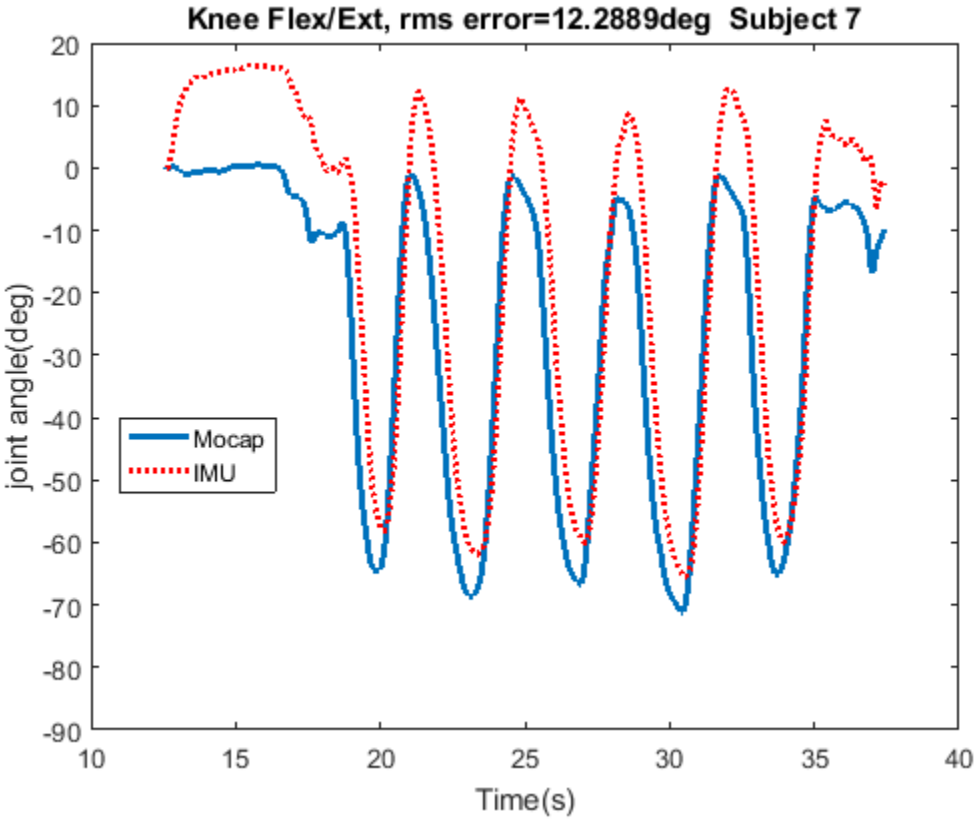


Figure 4.8: Subject 7 knee joint angle RMS error using approximated values for tibia sensor orientation, kinematic link lengths as well as displacement vectors

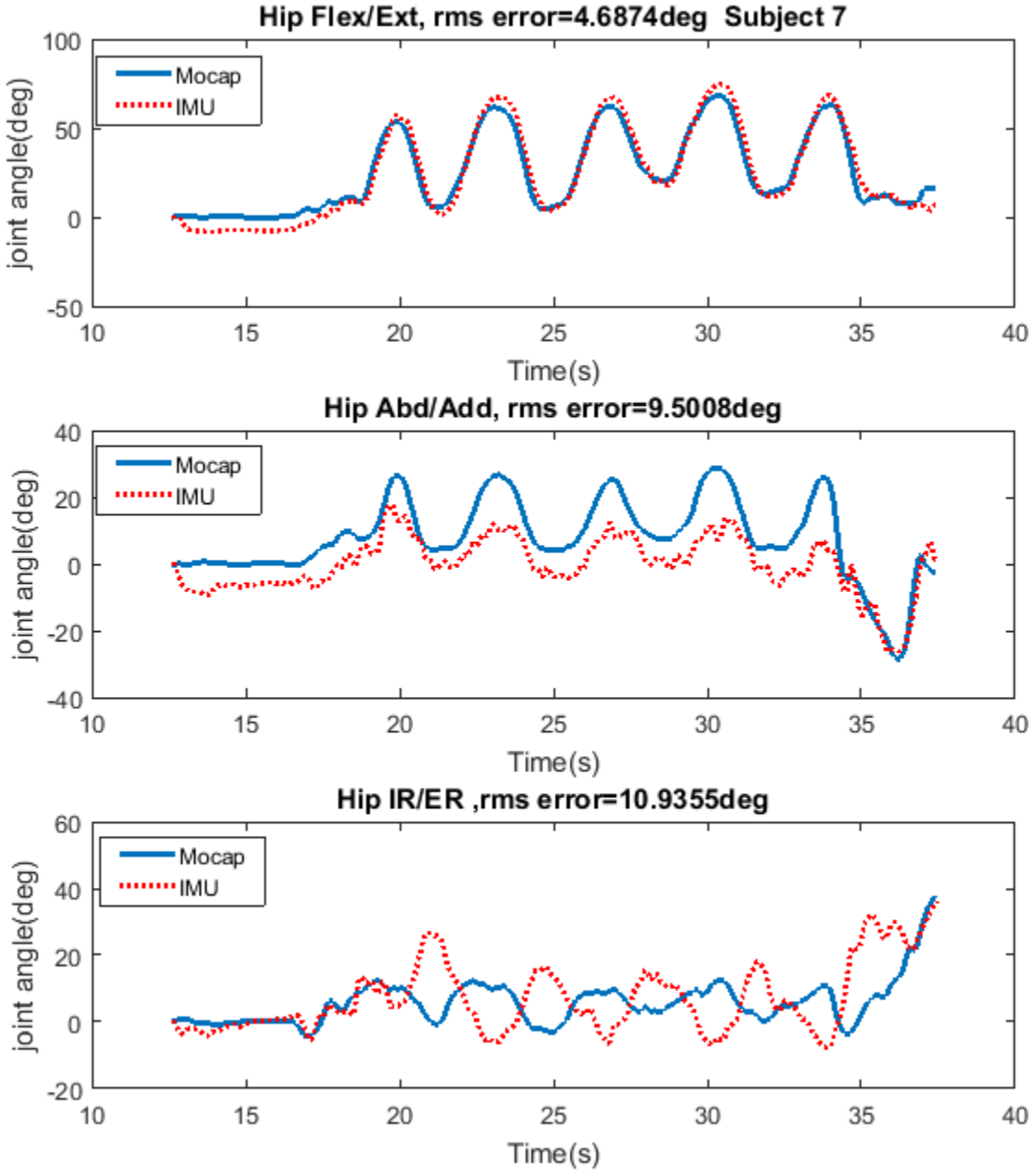


Figure 4.9: Subject 7 hip joint angles RMS error using approximated values for tibia sensor orientation, kinematic link lengths as well as displacement vectors

Comparing Table 4.2 to Table 3.1, the results show increased error in all joint angle estimates when using approximated values, as expected. The most affected angle is knee Flex/Ext, with a 7.3° increase in error, while the ankle IR/ER and hip Flex/Ext are less affected, with 1.7° and 1.6° increase, respectively. The overall increase in error, averaged over all joints and all subjects is 3.6°.

4.3 Sensitivity Analysis

To find out how pose estimation is impacted by variations in the kinematic parameters and sensor alignment, a sensitivity analysis is performed. The needed parameters for the forward kinematics (described in Table 4.1 and Section 4.1.2 and depicted in Figure 4.2) include :

$L_{Tib2Knee}$, C_{Tib}, $Z_{r_{S3}}$: *describing* $\boldsymbol{r_{S3}}$
L_{shank}, Y_{r_3}, Z_{r_3} : *describing* $\boldsymbol{r_3}$
$L_{Thi2Knee}$, C_{Thi}, $Z_{r_{S4}}$: *describing* $\boldsymbol{r_{S4}}$
L_{thigh}, Y_{r_4}, Z_{r_4} : *describing* $\boldsymbol{r_4}$
PW, PD, LL : *describing* $\boldsymbol{r_{S7}}$
Roll-tibia, *Pitch-tibia*, *Yaw-tibia* : *describing tibia sensor orientation*
Roll-thigh, *Pitch-thigh*, *Yaw-thigh* : *describing thigh sensor orientation*
Roll-back, *Pitch-back*, *Yaw-back* : *describing back sensor orientation*

There are a total of 24 parameters, which can be obtained from marker data if available, or must be approximated. The set of parameters obtained from markers are called Pm_1. These parameters were changed one-by-one by ±5% of their nominal value and used to calculate the one-at-the time sensitivity analysis [36].

Sensitivity analysis is performed based on the resulting joint angle errors of the IMU-based method described in Chapter 3 using Pm_1 and Pm_2 and according to the following formula:

$$Sensitivity = 100 \times \left| \frac{\frac{Err_2 - Err_1}{Err_1}}{\frac{Pm_2 - Pm_1}{Pm_1}} \right| \tag{4.1}$$

Pm_1 are the accurate parameter values obtained from markers, Pm_2 are the modified parameter values equal to $Pm_1 \pm 5\%$, Err_1 is the error between Motion Capture and IMU-based joint angle estimates when Pm_1 values are used in the forward kinematic model, and Err_2 is the error between the Mocap and IMU joint angle estimates when Pm_2 values are used in the forward kinematic model.

4.3.1 Results

Sensitivity analysis was performed for both increasing the parameters by 5% and decreasing them by 5%. Tables 4.3 and 4.4 show the sensitivity of the pose estimation accuracy to each of the defined parameters averaged for both decrease and increase and over all subjects for the ankle, knee, and hip joints.

Table 4.3: Sensitivity analysis results for the ankle and knee joints

Sensitivity %					
Parameter	Ankle IR/ER	Ankle Abd/Add	Ankle Flex/Ext	Knee Flex/Ext	Ankle Average
$L_{Tib2Knee}$	3.4	15.1	5.8	3.1	8.3
C_{Tib}	2.9	2.5	2.8	2.3	2.7
$Z_{r_{3S}}$	6.1	10	1.1	1	5.7
L_{shank}	10.7	**43.6**	16.3	9.5	23.5
Y_{r_3}	0	0	0	1.4	0
Z_{r_3}	0	0	0	0.1	0
$L_{Thi2Knee}$	0	0	0	8.5	0
C_{Thi}	0	0	0	4.4	0
$Z_{r_{4S}}$	0	0	0	0.1	0
L_{thigh}	0	0	0	0	0
Y_{r_4}	0	0	0	0	0
Z_{r_4}	0	0	0	0	0
PW	0	0	0	0	0
PD	0	0	0	0	0
LL	0	0	0	0	0
Roll-tib	**47.7**	**309.5**	28.5	**38.9**	**128.6**
Pitch-tib	24.6	**107.8**	2.2	3.7	**44.9**
Yaw-tib	4	1.3	**113.5**	**51.4**	**39.6**
Roll-thi	0	0	0	8.3	0
Pitch-thi	0	0	0	1.6	0
Yaw-thi	0	0	0	**67.5**	0
Roll-bac	0	0	0	0	0
Pitch-bac	0	0	0	0	0
Yaw-bac	0	0	0	0	0
Average	**4.2**	**20.4**	**7.1**	**8.4**	

Table 4.4: Sensitivity analysis results for the hip joint

Sensitivity %				
Parameter	Hip Flex/Ext	Hip Abd/Add	Hip IR/ER	Hip Average
$L_{Tib2Knee}$	2.7	12.9	30.8	15.5
C_{Tib}	0.8	13.4	21.2	11.8
$Z_{r_{3S}}$	1.8	7.8	10.6	6.7
L_{shank}	2.3	22.8	37.9	21
Y_{r_3}	0.6	3.8	9	4.5
Z_{r_3}	0.7	2.7	3.7	2.4
$L_{Thi2Knee}$	4.2	18.1	**50.6**	24.3
C_{Thi}	4.1	4.2	8.3	5.5
$Z_{r_{4S}}$	0.2	2.9	2.5	1.9
L_{thigh}	12.3	14.1	28.7	18.4
Y_{r_4}	0.3	0.9	2	1.1
Z_{r_4}	0.2	1.1	2.2	1.2
PW	1.5	5.6	20.6	9.2
PD	3.5	18.1	30.4	17.3
LL	1.1	5.4	4.9	3.8
Roll-tib	7.7	**167.5**	**55.8**	**77**
Pitch-tib	1.2	**44.8**	12.3	19.4
Yaw-tib	1.2	3.1	3.3	2.5
Roll-thi	7.6	7.9	13.5	9.6
Pitch-thi	0.8	0.2	1	0.7
Yaw-thi	**89.4**	2.5	7.6	**33.2**
Roll-bac	**92.5**	33.4	26.6	**50.8**
Pitch-bac	6.3	36.5	4.7	15.9
Yaw-bac	1.8	23.4	5.5	10.2
Average	**10.2**	**18.9**	**16.4**	

4.3.2 Discussion

According to Tables 4.3 and 4.4, the most sensitive parameters for ankle joint estimation are the tibia sensor orientation parameters, specially the tibia roll angle. The most sensitive

joints include the hip and ankle Abd/Add as well as the hip IR/ER.

Knee joint angle estimation is sensitive to the tibia sensor roll and yaw angles as well as the thigh sensor yaw angle.

The most sensitive parameters for the hip joint are the tibia and back sensor roll angles as well as the thigh sensor yaw angle. Hip joints, especially hip IR/ER, are sensitive to more parameters because they are the last joints in the kinematic chain. This analysis also reveals that approximations using the population-average leg circumference and assuming alignment between joint centers will not impact the estimation results considerably.

By using the 3-EKF structure, as can be seen in Table 4.3, the ankle joints are not sensitive to any parameter beyond the tibia sensor in the kinematic chain and the knee joint is not sensitive to any parameter beyond the thigh sensor. This may partly explain the better joint angle estimation using 3-EKF structure compared to 1-EKF structure.

In addition, an explanation for the obtained order of magnitude for the optimal values of the EKF noise parameters reported in Section 3.7 can be provided based on the sensitivity analysis results. By looking to the optimal parameters obtained for each subject, we found that for majority of them $\gamma_7, \gamma_6, \gamma_2$ have the highest values and $\gamma_1, \gamma_5, \gamma_3$ have the lowest values. This pattern is similar to the sensitivity of the corresponding joint angles. According to Table 4.4, ankle Add, hip Add, and hip IR are the most sensitive ones while ankle IR and ankle Flex are the least sensitive ones.

In general, the sensitivity analysis results suggest that having a good estimation of the sensors' orientation can improve joint angle estimation results considerably. If hip IR/ER angle is of high interest then other measures such as the distance between the knee center and the tibia and thigh sensors, shank and thigh lengths and PD should also be carefully measured.

4.4 IMU Orientation Estimation Calibration Protocol

According to the sensitivity analysis results, knowing the exact sensor orientation on the body is a key factor for accurate pose estimation. We would like to extract full sensor orientations without requiring patients/subjects to perform any calibration movement or posture.

For short duration movements, orientation can be estimated by gyro measurement integration. Sensor orientations can therefore be retrieved from gyro data under specific

considerations for sensor placement. For this purpose, a protocol for sensor placement was developed as follows:

1. All sensors were placed on the table in the same known orientation (Figure 4.10-right).
2. Sensor locations were marked on the thigh and tibia using a double sided tape attached to the desired sensor location.
3. The outer side of the tape was removed and the participant was asked to stand still in a defined frontal orientation with respect to the table as depicted in Figure 4.10.
4. Data collection was started and sensors were moved to the defined locations one by one (Figure 4.10-left).
5. After a few seconds at the final position, data collection was stopped. This process took less than one minute which is reasonable to avoid gyro drift.

Figure 4.10: Different steps of performing the calibration protocol

Please note that the calibration protocol does not require any marker information. However, since the participant was going to perform the SLS task in the Mocap Lab, there are markers on the body and sensors in Figure 4.10. These are used only for ground truth data collection and are not required during clinical use.

In the next step, the Rodriguez method [75] was applied to gyroscope data to calculate rotation matrices from the start to the final position for each sensor according to Equations 4.3 to 4.5. ${}^{B}R_0(t)$ is the rotation matrix between the initial position on the table and the sensor at time t. To get the final orientation, we averaged the last 200 samples which is equal to the last second of data collection when all sensors were in their assigned position on the body.

$$\omega(t) = \sqrt{\omega_x(t)^2 + \omega_y(t)^2 + \omega_z(t)^2} \tag{4.2}$$

$$S(t) = \begin{bmatrix} 0 & -\omega_z(t)/\omega(t) & \omega_y(t)/\omega(t) \\ \omega_z(t)/\omega(t) & 0 & -\omega_x(t)/\omega(t) \\ -\omega_y(t)/\omega(t) & \omega_x(t)/\omega(t) & 0 \end{bmatrix} \tag{4.3}$$

$${}^{B}R_0(t+1) = S(t)sin(\omega(t)\delta t) + (S(t))^2(1 - cos(\omega(t)\delta t)) + I \tag{4.4}$$

$${}^{B}R_0(t) = {}^{B}R_0(t-1){}^{B}R_0(t+1) \tag{4.5}$$

Where $\omega_x(t)$, $\omega_y(t)$, $\omega_z(t)$ refer to X, Y, Z components of the angular velocity at time t, respectively. ω is the magnitude of angular velocity. S and I are the skew-symmetric and identity matrices, and δt is the sampling interval.

When the tibia and thigh sensors were to be placed in the desired orientation as depicted in Figure 3.1-left, it was only necessary to rotate them $+90^\circ$ about the Z axis (+90 yaw rotation) which was equal to ${}^{D}R_0(tibia, thigh)$ in Equation 4.6 (rotation matrix from initial position to the desired position). For the back sensor; however, two rotations were needed from the table to the desired back position: a -90° rotation about Y followed by -90° rotation about X which is equal to ${}^{D}R_0(back)$ in Equation 4.7.

$${}^{D}R_0(tibia, thigh) = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.6}$$

$${}^{D}R_0(back) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{4.7}$$

Any difference between the ideal rotation matrices and those obtained from the gyro integration procedure described above is as a result of sensor offset. Therefore, by obtaining

this offset as a rotation matrix and applying it to sensor readings, we are able to compensate for the offsets from the ideal position. This offset matrix is shown in Equation 4.8 where ${}^{B}R_{0}{}^{T}$ is the transpose of ${}^{B}R_{0}$.

$$ {}^{D}R_{B}(offset) = {}^{D}R_{0}{}^{B}R_{0}{}^{T} \tag{4.8} $$

4.5 Joint Angle Estimation Results After Applying Calibration

Table 4.5 shows the RMS error of pose estimation when approximated values are used for displacement vectors and sensor orientations are extracted from the calibration protocol. Comparing the results with Section 4.2 reveals significant improvement in accuracy for ankle and hip Add/Abd angles (2.8° and 3°) and knee flexion angle (6.5°). The results have improved for all subjects and the overall improvement is 2.5°. Joint angles for subjects 1 and 7 are depicted in Figures 4.11 to 4.16.

Table 4.5: RMS error between IMU and Mocap when kinematic parameters are extracted from markers and sensor orientations are obtained from the developed calibration protocol

RMS error between Mocap and IMU-based estimated joint angles (degree)								
Subject	Ankle IR/Er	Ankle Abd/Add	Ankle Flex/Ext	Knee Flex/Ext	Hip Flex/Ext	Hip Abd/Add	Hip IR/ER	Average Error
1	2	4.3	3.4	3.6	4.6	3.5	3.2	3.5
2	2.4	3.1	4.6	5.4	16	4.6	6.7	6.1
3	1.8	1.8	3.7	12.6	16	3.2	5.7	6.4
4	5.1	5.9	4.3	7.8	13.5	7.8	12.5	8.1
5	5.4	3	4.6	3	11.5	5.3	8	5.8
6	6.4	3.1	9.7	5.3	5.1	7.3	9	6.6
7	3.7	3.8	5.1	6.6	14.2	5.3	10.6	7
Average	**3.8**	**3.6**	**5.1**	**6.3**	**11.6**	**5.3**	**8**	**6.2**

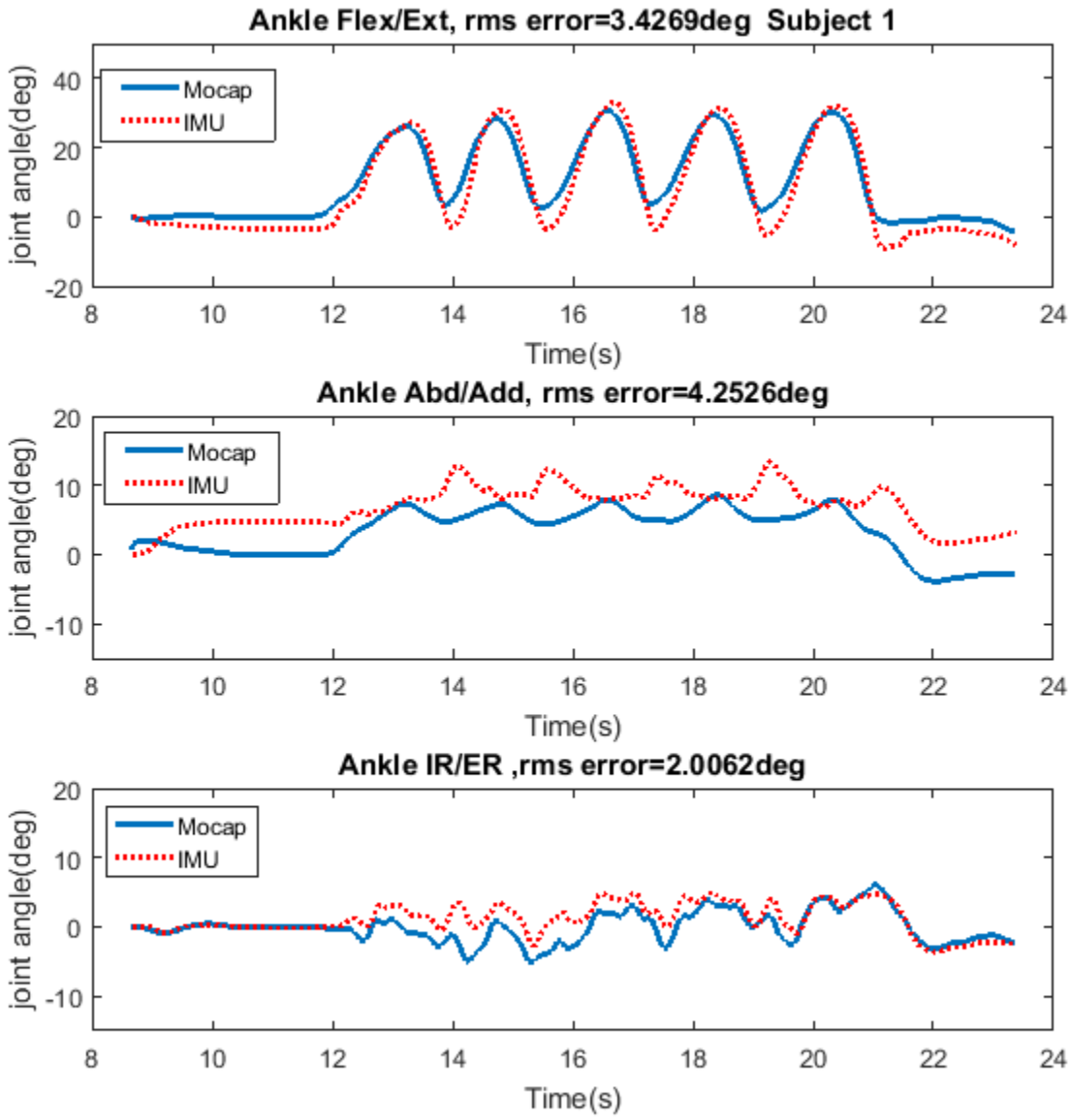


Figure 4.11: Subject 1 ankle joint angles RMS error using approximated values for displacement vectors and estimated sensors' orientations from the developed calibration protocol

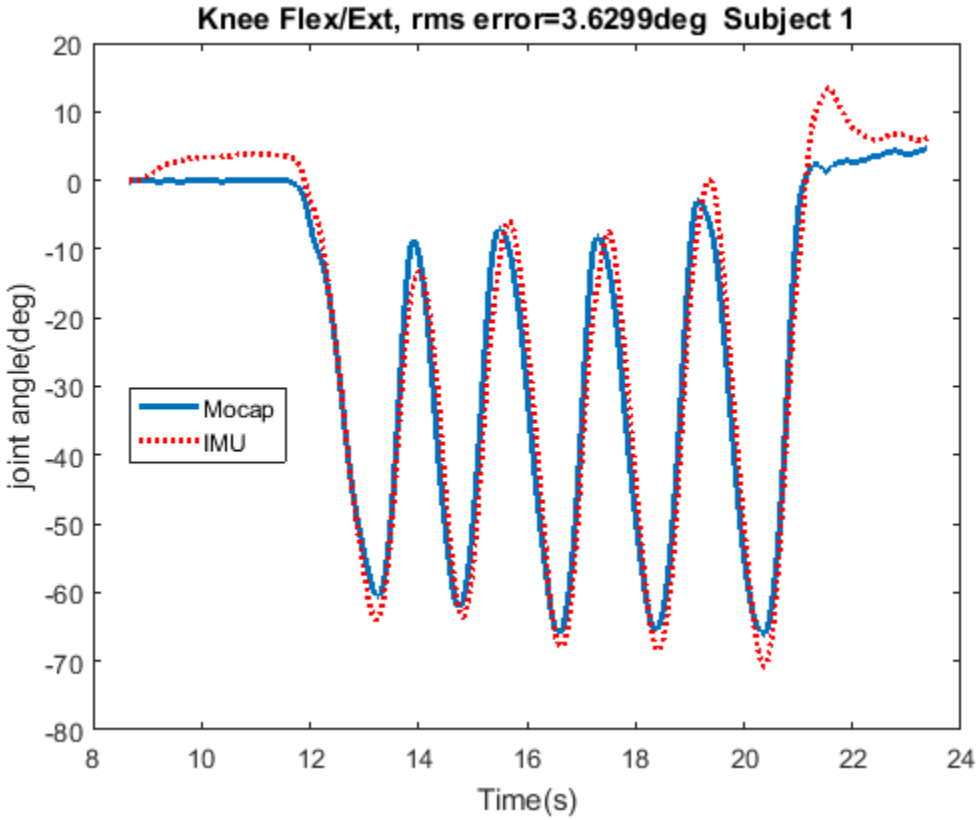


Figure 4.12: Subject 1 knee joint angle RMS error using approximated values for displacement vectors and estimated sensors' orientations from the developed calibration protocol

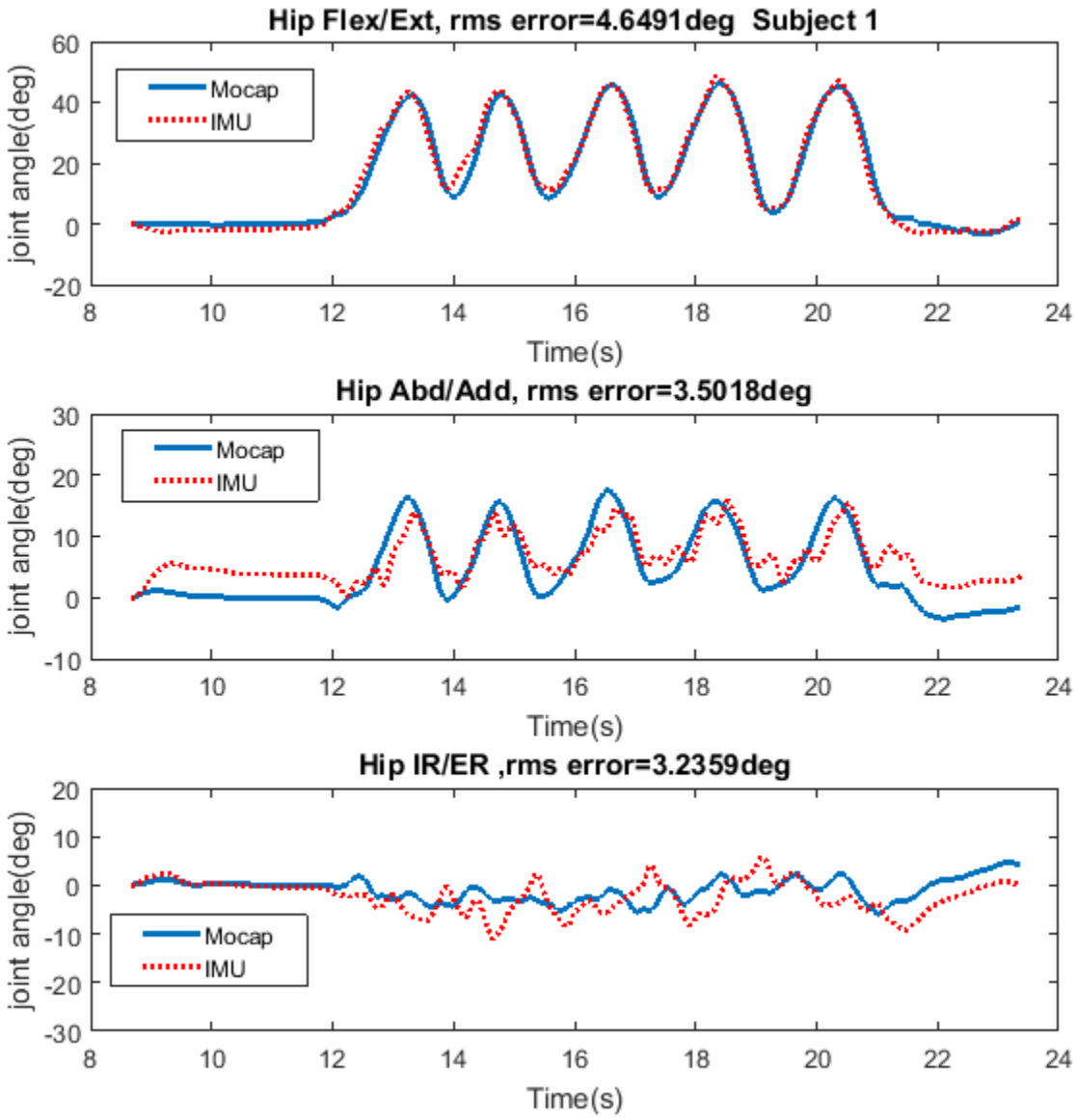


Figure 4.13: Subject 1 hip joint angles RMS error using approximated values for displacement vectors and estimated sensors' orientations from the developed calibration protocol

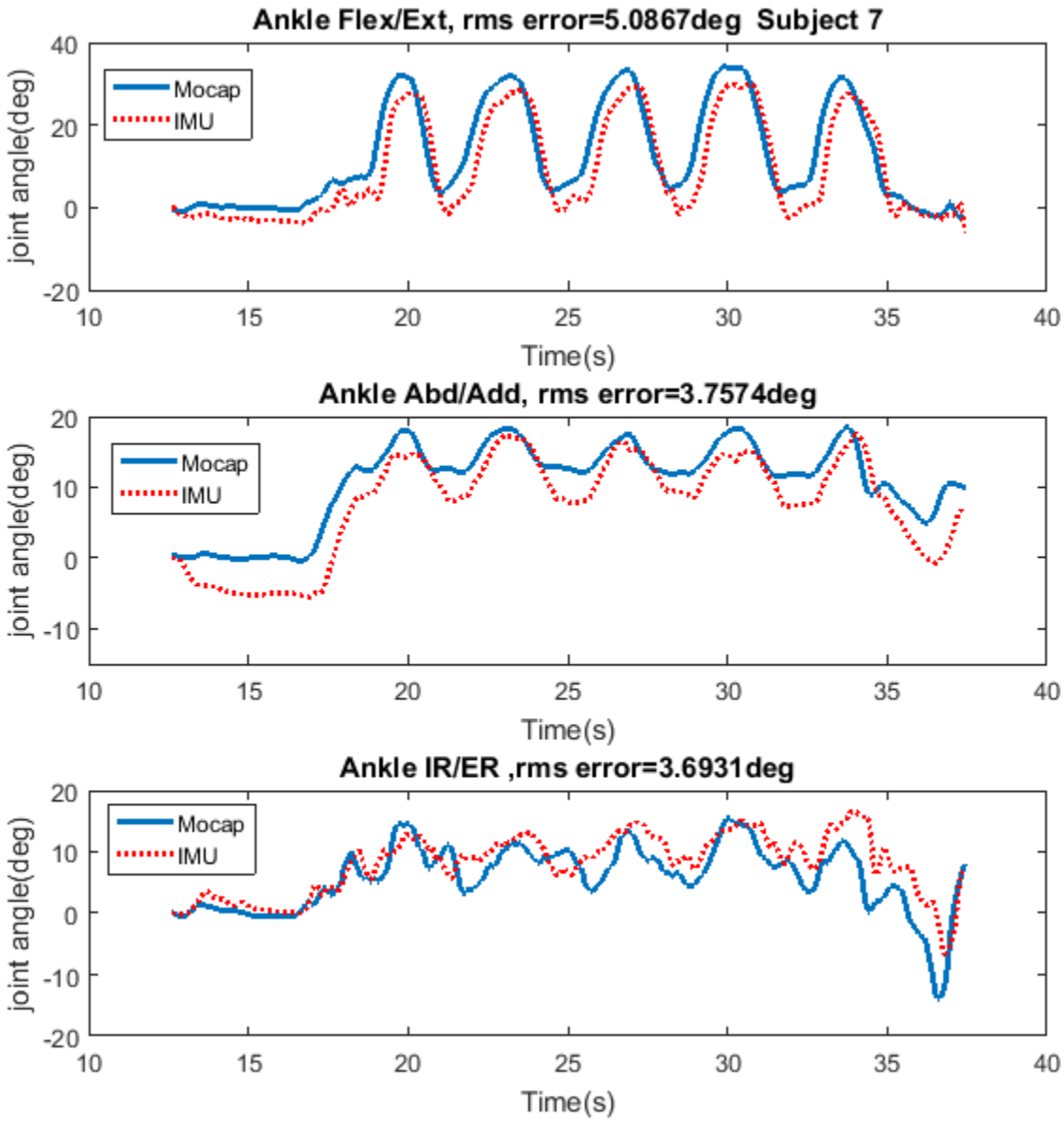


Figure 4.14: Subject 7 ankle joint angles RMS error using approximated values for displacement vectors and estimated sensors' orientations from the developed calibration protocol

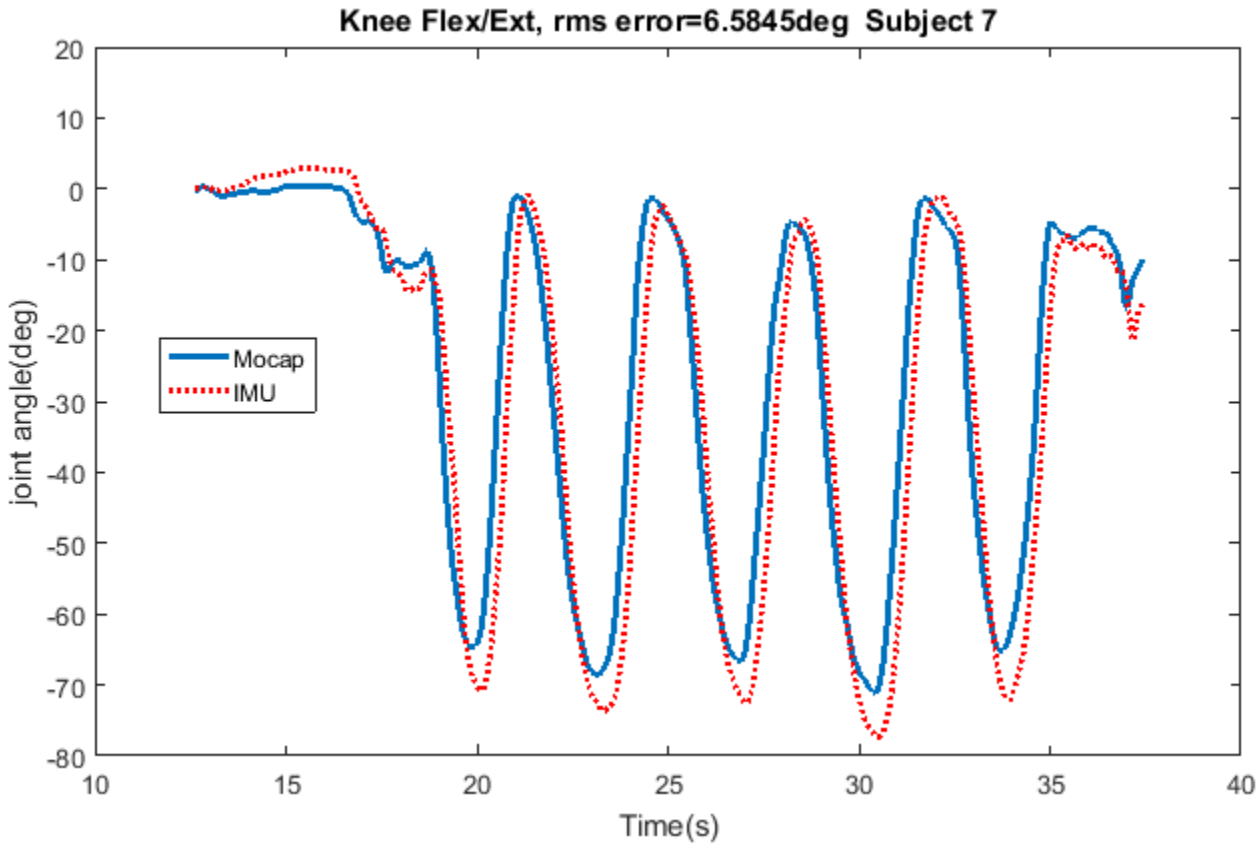


Figure 4.15: Subject 7 knee joint angle RMS error using approximated values for displacement vectors and estimated sensors' orientations from the developed calibration protocol

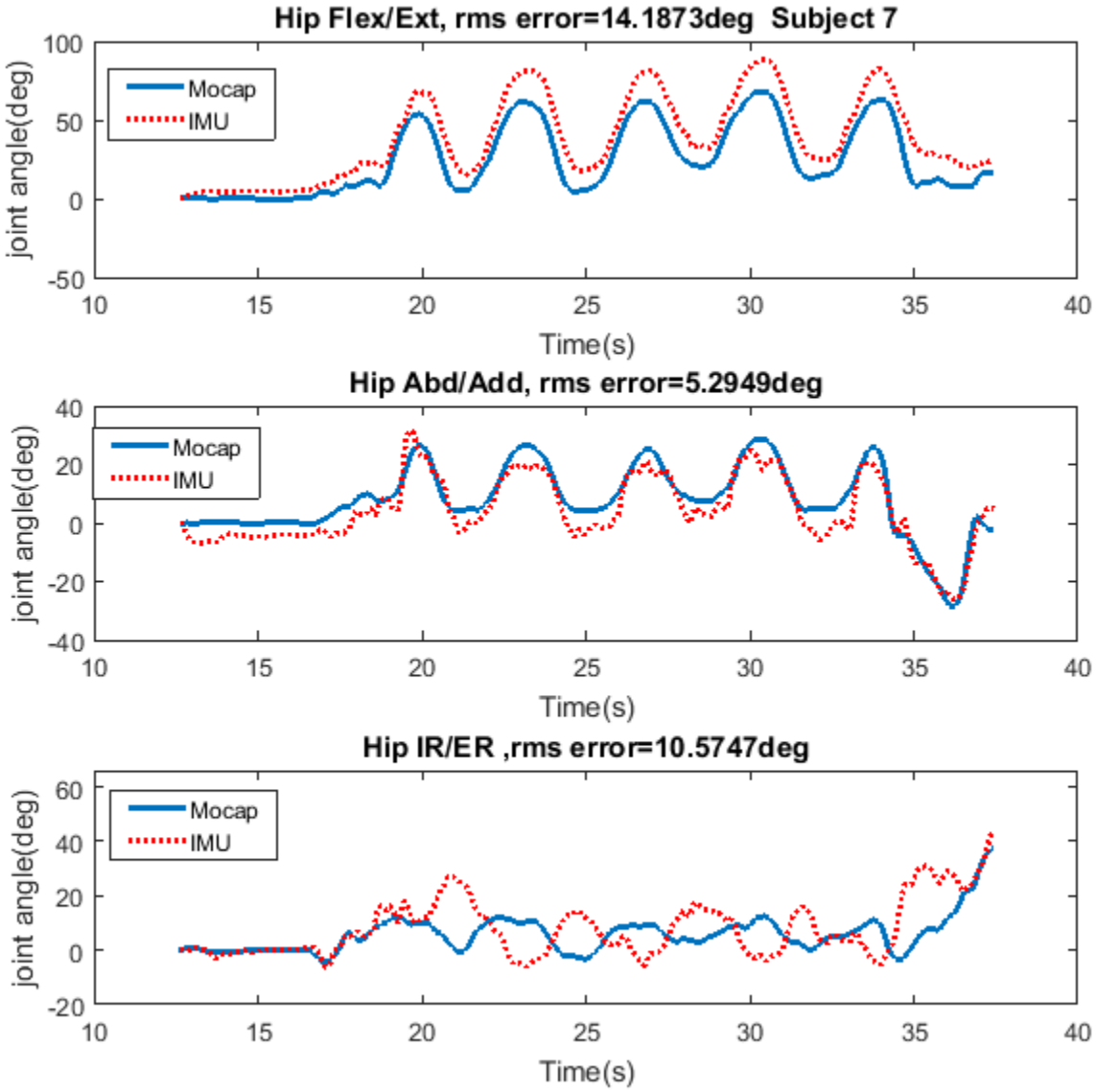


Figure 4.16: Subject 7 hip joint angles RMS error using approximated values for displacement vectors and estimated sensors' orientations from the developed calibration protocol

4.6 Soft Tissue Artifacts

To investigate the influence of soft tissue artifacts on the pose estimation, joint angles were calculated by applying inverse kinematics (IK) to the marker positions using the algorithm developed by Joukov [45]. In addition to the knee, ankle and hip joints, two 6 DOF joints including three prismatic and three revolute joints were defined at the tibia and thigh sensor locations and a 3 DOF prismatic joint was assigned to the back sensor in the IK model (shown in Figure 4.17), which enabled us to measure sensor movements with respect to their initial location during the SLS performance. Any movement in these additional joints corresponds to sensor motion relative to the skeleton due to soft tissue artifacts.

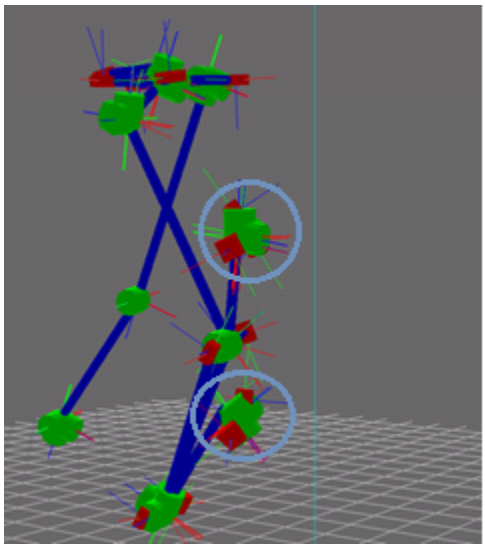


Figure 4.17: Kinematic model of the skeleton developed by [45]. 6 DOF joints are defined at the tibia and thigh sensor locations.

The movement of the tibia/thigh sensor included the displacement in X, Y, Z directions with respect to the skeletal frame for prismatic joints, and roll, pitch, and yaw for the revolute joints. Since ankle joint angles' estimates are based on tibia sensor measurements only, if there was any error introduced by tibia sensor motion, we expected to see a correlation between ankle joint angle errors and tibia sensor motion. For this reason, we analyzed the correlation between the obtained movements of the tibia sensor and the RMS error of each ankle joint during the whole period of the motion. Similarly, the knee joint angle estimate is based on thigh sensor measures and previously estimated ankle joint angles. Therefore, thigh sensor motion should be observed primarily in the knee joint angle error.

We also analyzed the correlation between the obtained movements of the thigh sensor and the RMS error of the knee joint angle. Similarly, the correlation between the back sensor displacement and each of the hip joint angles was analyzed. The results, however, revealed no significant correlation between the joint angle estimation errors and sensor motion for the collected data.

We hypothesise that sensor motion artifact may have been partially compensated for through the measurement noise covariance of the EKF, illustrating another potential benefit of EKF-based methods.

4.7 Summary

In this chapter, approximation methods for kinematic chain parameters required for pose estimation were proposed. Approximations relied on body geometry, anthropometric information and use of average fixed values for measures which were not available through anthropometric tables.

Next, those parameters whose variation contributed to the joint angle error the most was identified through sensitivity analysis.

Finally, a fast and simple calibration protocol based on gyro data was proposed to extract sensor orientations on the body, and it was shown that the protocol results in improved pose estimation while it does not demand any specific calibration posture or movement from the participant.

A correlation analysis was also performed to determine if there is any relationship between sensor motion artifacts and joint angle errors. The results showed no significant correlation for the collected dataset.

The main objective of the pose estimation in this study was to use the estimated pose for the SLS motion analysis in order to detect faulty movement patterns. In the next chapter, we will use the developed pose estimation algorithm to measure pose during the performance of SLS and apply feature selection and classification techniques to the estimated pose data for DKV detection as well as risk of injury evaluation.

Chapter 5

Automated DKV and Risk of Injury Assessment

To develop an automated assessment method for SLS, the developed IMU-based pose estimation algorithm was applied to new participants while participants, both in the Motion Capture Studio and in the field, were performing the SLS test. The estimated hip, knee, and ankle joint position, velocity and acceleration was then used for assessment of squat quality using the feature selection and classification techniques which will be elaborated in this chapter.

The automated assessment method was first developed for a pilot data set. After successful initial results with the pilot study, a larger dataset was collected and labelled by our clinical collaborator and used for analysis.

In this chapter, we will first explain the methodology, then describe experiments and results of both pilot study as well as clinical study.

5.1 Methodology

The IMU-based pose estimation was applied to a labelled SLS dataset. Time domain features were extracted from segmented pose data; the most informative features were selected via feature selection, and classifiers were trained to classify squat quality based on the amount of observed DKV as well as overall risk of injury. Figure 5.1 shows the steps of the automated assessment system. Each step is explained in the following sections.

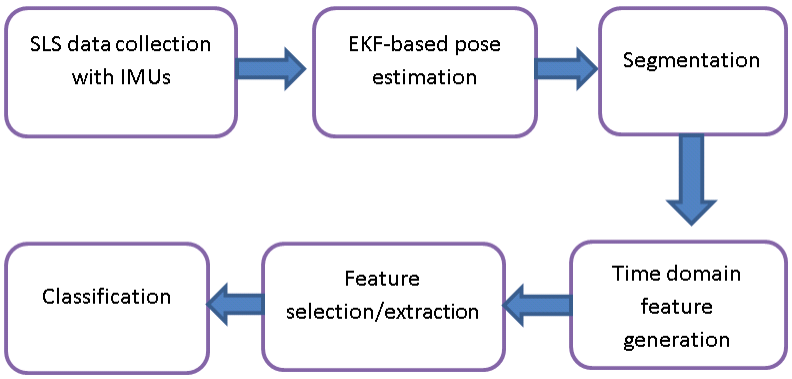


Figure 5.1: Overview of the automated assessment system

5.1.1 Pose Estimation

The pose estimation algorithm explained in Chapter 3 was applied to the collected data. The required set of parameters including tibia and thigh sensors to knee distances ($L_{Tib2Knee}$ and $L_{Thi2Knee}$), thigh and shank lengths (L_{shank}, L_{thigh}), tibia and thigh circumferences at sensors location (C_{Tib}, C_{Thi}), pelvic width and depth and leg length (PW, PD, LL) were measured manually. It should be noted that, as discussed in Chapter 4, these manual measurements can be omitted and replaced with anthropomorphic values in future data collection efforts. Sensor orientations were extracted from the calibration protocol described in Section 4.4.

5.1.2 Segmentation

To extract a single SLS repetition from continuous time series data, the joint angle, velocity, and acceleration trajectories needed to be segmented before feature extraction. Two different approaches were applied for the pilot study and clinical data analysis.

Zero Velocity Crossing

The zero velocity crossing criterion [56], shown in Figure 5.4, was used for segmenting the estimated pose data of the pilot study.

The knee flexion velocity was chosen for segmentation. A first order Butterworth filter with a cutoff frequency of 0.3 Hz was applied to the knee joint trajectory prior to segmentation. This filter was applied only for segmentation and not for the subsequent feature extraction. Every time velocity changed sign (crossed zero), a segment point was reported. Additional criteria including the minimum interval between subsequent detected segments and minimum amplitude were applied to remove false segment points.

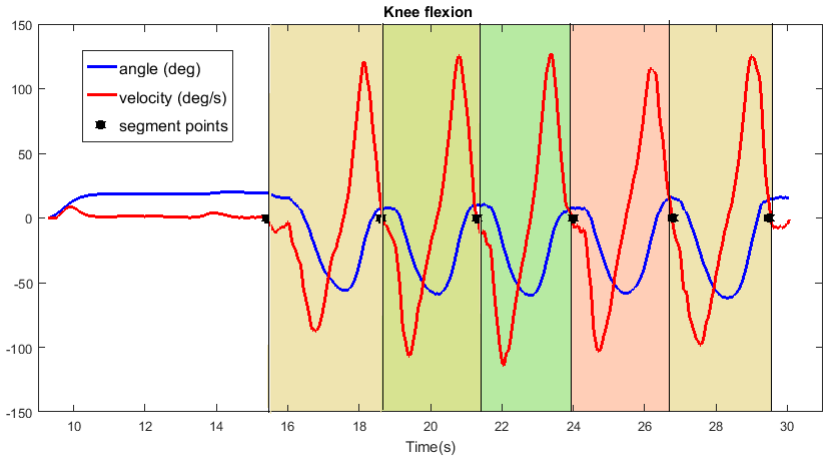


Figure 5.2: Zero velocity crossing criterion applied to knee joint velocity for detecting the segment points

Segmenting based on Joint Angle Peaks

Zero velocity crossing is sensitive to near zero fluctuation of the velocity trajectory and additional criteria were necessary to remove false segment points.

For segmentation of the estimated joint pose for the clinical dataset, a peak detection method developed by [40] was applied to the knee flexion angle. The knee flexion was chosen for segmentation because the knee has a large ROM, and its peaks are easily detectable. A first order Butterworth filter with cutoff frequency of 0.3 Hz was applied to the knee joint trajectory prior to segmentation. This filter is applied only for segmentation and not for the subsequent feature extraction. The midpoints between peaks were then calculated and used as segmenting points as depicted in Figure 5.3. Figure 5.4 shows an example of segmented joint angles used for feature extraction (without low pass filtering).

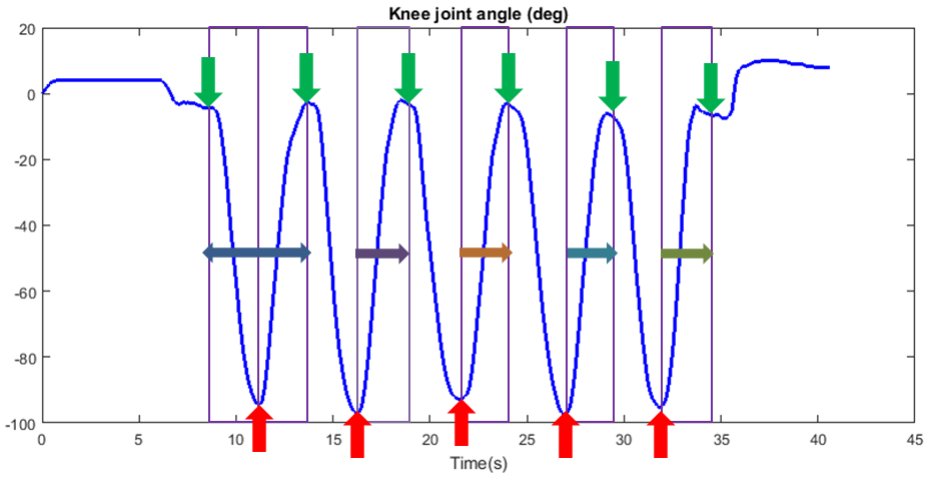


Figure 5.3: Segment points (green arrows) were found by detecting peaks (red arrows) of low pass filtered knee joint angle and computed the midpoint of the peak to peak distances (horizontal arrows).

5.1.3 Feature Generation

Various feature extraction methods have been used for human activity recognition [51]. These methods are categorized into time domain or frequency domain methods.

The mean, standard deviation (STD), variance (VAR), interquartile range (IQR), mean absolute deviation (MAD), correlation between axes, entropy, and kurtosis are among the time domain features commonly used for activity recognition from the acceleration signal [51]. Similarly, Preece et al. [72] have identified the mean, median, variance, skewness, kurtosis and interquartile range as commonly used time domain features.

Common frequency domain methods include Fourier transform (FT) and discrete cosine transform (DCT)[51].

Due to easier clinical interpretability and better temporal localization, we applied only time domain feature extraction methods including the root mean square (RMS), STD, VAR, mean, MAD, skewness, kurtosis, range, minimum, and maximum of the joint angle, velocity and acceleration of each DOF for each segment of the data. Therefore, for each repetition of the squat, a feature vector of 210 different features was generated.

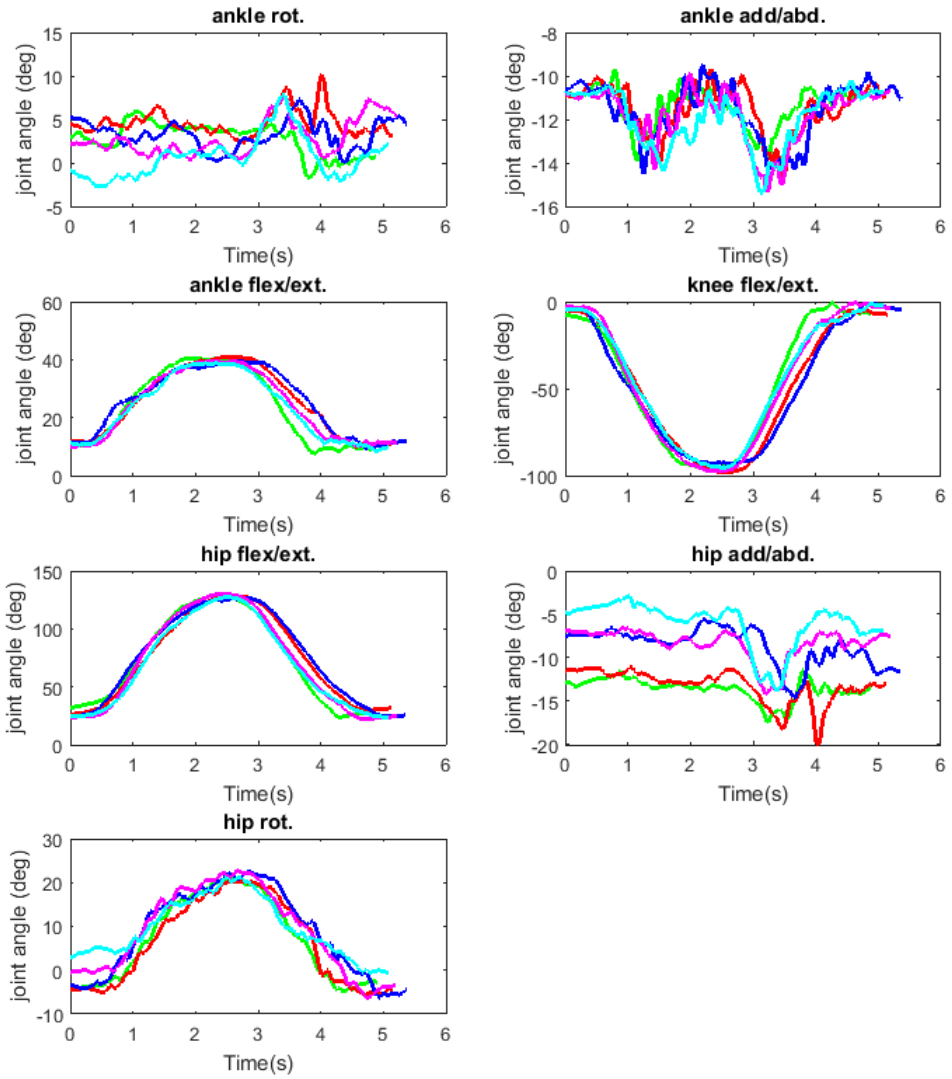


Figure 5.4: An example of segmented joint angles without low pass filtering used for feature extraction.

5.1.4 Feature Selection/Extraction

We do not know which of the defined features better predicts DKV. Moreover, some features might be redundant or irrelevant, which may degrade the classification results. Selecting the most appropriate features not only helps with dimensionality reduction but also

suggests the best predictors of DKV to clinicians.

A large number of feature selection techniques are available in the literature, usually categorized as filter, wrapper or embedded techniques [99]. Filter techniques select relevant features based on statistical tests. Wrapper techniques use the performance of a predefined learning algorithm as the selection criterion. In embedded techniques, feature selection occurs in parallel to model learning, so that feature selection is embedded within a classification model [99].

For this study, we applied 18 different feature selection techniques from all three categories. Matlab packages available from the Arizona State University [99] repository and from Pohjalainen et al. [70] were used for implementation.

Wrapper methods included Random Subset Feature Selection, Sequential Forward Selection, and Sequential Floating Forward Selection.

Filter methods were Mutual Information, Statistical Dependency, Correlation based Feature Selection, Chi-Sqaure, Fast Correlation-Based Filter, Fisher Score, Gini Index, Information Gain, Kruskal- Wallis, Minimum-Redundancy-Maximum-Relevance selection, Relief-Feature selection strategy, and T-test.

From embedded techniques, Sparse Multinomial Logistic Regression via Bayesian L1 Regularization, Bayesian logistic regression, and Least Absolute Shrinkage and Selection Operator (LASSO) were utilized.

Features ranked among the top ten by a majority of the methods are reported as selected features. In addition to subset feature selection, feature extraction using supervised principal component analysis (SPCA) was also applied. Matlab code developed by Barshan et al. [13] was used for the SPCA implementation.

5.1.5 Classification

For classification purposes, six different methods were applied: Support Vector Machine (SVM), Linear Multinomial Logistic Regression (LMLR), Decision Tree (DT), Naïve Bayes (NB), K Nearest Neighborhood (KNN), and Random Forests.

For the pilot study, a subset of methods including SVM, LMLR, and DT were tried. All classifiers were implemented using MATLAB R2014b. SVM was selected as it is robust to small training data size. For 3-class classification, one-versus-all and one-versus-one SVM with linear kernel were implemented. SVM multi-label results were computed by majority vote between one-vs-one classification results. The Decision Tree was selected as

it provides threshold values (cutoff points) in the selected features which can be informative for clinical interpretation of the classifier.

For the clinical dataset, all six classification techniques were implemented using Matlab 2016a. The results showed that SVM, KNN, and NB always outperform other classifiers for this dataset. Therefore, in the clinical data analysis, classification results are reported for these three classifiers only.

5.2 Pilot Study

To develop and initially validate our proposed approach, we first conducted a pilot study. In the pilot study, the automated assessment method described in Section 5.1 was applied to a labelled SLS dataset collected from seven participants. Classifiers were trained to distinguish between "poor", "moderate", and "good" squat qualities.

5.2.1 Experiment

Seven participants (6 male, 1 female, mean age 32.3±11.6 years) took part in the pilot study. Inclusion criteria were adults not having any lower back or leg injuries in the past six months. The experiment was approved by the University of Waterloo Research Ethics Board, and all participants signed a consent form prior to the start of data collection.

Data Collection

Three Yost [6] IMU sensors were affixed to the participant using hypoallergenic tape. Sensor placement sites included the low back at the level of the first sacral vertebra, the anterior thigh 10 cm above the patella aligned with the sagittal plane, and the flat surface of the shank at the level of the tibial tuberosity, as illustrated in Figure 5.5. Due to wireless communication, sampling rates were not consistant or identical for all sensors. The average sampling rate was 90±10 Hz. All sensors were interpolated and resampled to the same rate of 100 Hz.

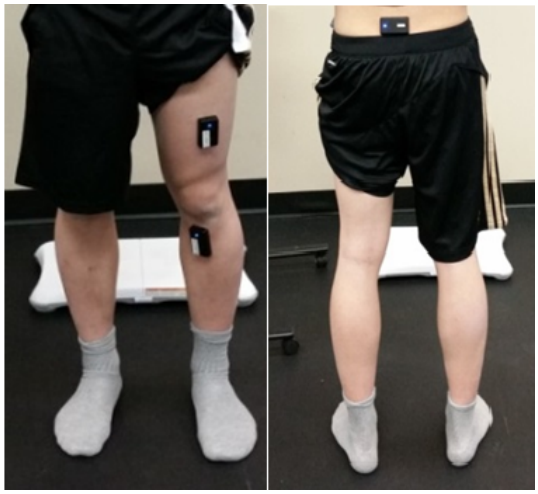


Figure 5.5: Sensor placement during SLS pilot data collection

Participants were instructed to remove their shoes and stand on their dominant leg (the leg they would kick a ball with) with toes pointing straight ahead, while keeping their arms crossed in front of their body. In each trial, participants performed five consecutive cycles of the SLS movement. For the SLS collection to be deemed successful, the subject had to perform the squat without allowing the legs to contact each other, and without losing balance (ie. without having the non-weight bearing leg touch the ground).

Data Labelling

Three of the participants replicated good, poor, and moderate squats under the instruction and supervision of an expert clinician; the other participants performed the squats naturally. The naturally performed squats were labelled by an experienced movement scientist using a clinical knee valgus rating scale that was modified from a qualitative SLS clinical rating tool [9]. A SLS was rated "good" if DKV did not occur during the squat or DKV occurred, but the patella did not have a trajectory that pointed towards the second toe; "moderate" if the patella pointed toward or past the second toe, but did not point past the inside aspect of the foot; and "poor" if the patella pointed past the inside aspect of the foot.

The number of trials used for analysis was not the same for all participants. There were 7 labelled trials available from participant 2 (3 good, 1 moderate and 3 poor), 6 from participant 1 and participant 3 (1 good, 1 poor, and 1 moderate for each), 1 from

participant 4 (poor), 2 from participant 7 (moderate), 3 from participant 5 (2 poor and 1 good) and 1 from participant 6 (moderate). Each trial consisted of 5 consecutive squats, which resulted in 100 examples of SLS including 30 examples of good, 30 examples of moderate, and 40 examples of poor squats.

5.2.2 Pilot Study Results

Given the 7 DOF kinematic model, where each DOF includes an estimate of its position, velocity, and acceleration, the total number of features for each segment or observation was 210. Therefore, our final data set had 100×210 dimensions. Another dataset was also produced with the same features, but including only good and poor data (i.e., excluding the moderate SLS data), which had 70 observations. All data was normalized to bring the values in [0 1] range.

The feature selection results are summarized in Table 5.1. The feature selection results highlight the importance of the ankle IR angle features for differentiating good, moderate and poor squats. Although according to clinical studies [15], [98], the hip plays an important role in DKV, the feature selection results in our pilot study suggested that good classification can be performed based on only the ankle kinematics. Possible explanations for the finding that the hip data was not as informative as the ankle for classification include the large variability in hip joint movements between different subjects (independent of squat quality) or a larger error in the pose estimation for the hip parameters.

The finding that all of the selected features are related to the ankle suggests that it may be possible to achieve good classification of the SLS by using only a simple 3 DOF model to estimate ankle joint kinematics. This is advantageous, as it simplifies the pose estimation and reduces the number of sensors from 3 to 1, reducing the complexity of the measurement apparatus and the setup and computation procedure. Similar clinical studies [15] used time consuming manual measurements and focused on only the feature selection part, while the proposed method in this study is completely automated and simple to apply, and therefore more easy to apply in the clinical setting.

For the pilot study a subset of methods including SVM, LMLR, and DT were tried for both the 2-class ("good" vs "poor") and 3-class ("good" vs "moderate" vs "poor") classification problems. The classification results are reported for both 10 fold cross validation (10 F-CV) and leave one subject out (LOSO) cross validation methods in Table 5.2 and Table 5.3 for 2-class and 3-class problems, respectively. For reporting the accuracy percentage, the number of selected features or Principal Components (PCs) in SPCA was set to one first and accuracy was calculated. Then, the number of features or PCs was increased one

Table 5.1: features ranked as top ten by more than 8 feature selection techniques.

Selected features For 2-class problem	N_r	Selected features For 3-class problem	N_r
ROM of ankle IR	14	STD of ankle IR angle	13
STD of ankle IR angle	11	VAR of ankle IR angle	13
MAD of ankle IR angle	11	MAD of ankle IR angle	13
VAR of ankle IR angle	10	ROM of ankle IR	12
RMS of ankle IR velocity	9		
MAD of ankle IR velocity	9		
RMS of ankle Add acceleration	9		

N_r : Number of times ranked as top ten features

by one up to the point that further increases did not improve performance. The reported accuracies are the best performance each classifier achieved. Matrix inversion with the full dimensional dataset was not possible with LMLR; therefore no results are reported for this condition.

Analysis of the decision tree results using majority selected features shows that for both LOSO and 10 fold validation, the best performance was achieved using only the ROM of ankle IR feature for the 2-class problem, while for the three class problem, ROM and MAD of the ankle IR angle resulted in best accuracy for LOSO CV and STD and MAD of the ankle IR angle for 10 fold CV. The decision tree structure shown in Fig. 5.6 for the 2-class problem suggests an association between poor squats and ROM of the ankle IR greater than 20.63°. For the 3-class problem, MAD of ankle IR angle less than 14.9° differentiates good squats. MAD of ankle IR angle greater than 14.9° indicates either moderate or poor squats, which are again differentiated based on ROM of ankle IR.

Table 5.2: Accuracies (%) for the 2-class classification problem using three classifiers and two different cross-validation methods

2-class classification accuracy (%)						
Validation method	10 Fold CV			LOSO CV		
Dimensionality reduction method	Majority Selected features	SPCA	No reduction	Majority Selected features	SPCA	No reduction
SVM	95.71	98.57	**99.71**	88.57	**98.57**	75.7143
Logistic Regression	93.57	98.57	–	91.43	**98.57**	—
Decision Tree	92.43	98.57	95.86	87.14	**98.57**	81.43

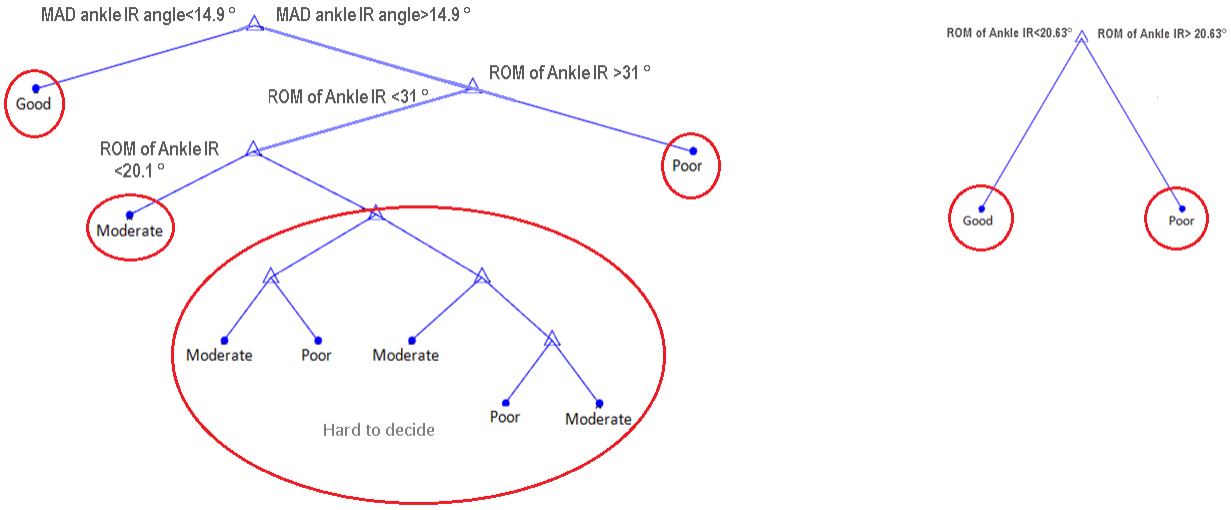


Figure 5.6: Decision Tree structure for LOSO-CV for the 3-class (left) and 2-class (right) problems. For 2-class problem, poor squats are detected when ROM of ankle IR $> 20.63°$. For 3-class problem, MAD of ankle IR angle$> 14.9°$ indicates either poor or moderate squats, which are again differentiated based on ankle IR ROM. This also shows that deciding between poor and moderate squats is harder than good and moderate for this dataset

Table 5.3: Accuracies (%) for the 3-class classification problem using three classifiers and two different cross-validation methods

3 class classification accuracy (%)						
Validation method	10 Fold CV			LOSO CV		
Dimensionality reduction method	Majority Selected features	SPCA	No reduction	Majority Selected features	SPCA	No reduction
SVM good vs all	91.8	84.7	98.1	91	83	79
SVM poor vs all	77.8	87.4	96.9	75	82	72
SVM moderate vs all	64.8	77.5	87.2	62	74	42
SVM good vs moderate	91.67	100	99	91.66	73.33	62.33
SVM poor vs moderate	70.14	90	96.71	67.14	80	50
SVM good vs poor	93.71	97.57	99.43	91.43	97.14	75.71
Logistic Regression	73.6	93.1	—	68	68	—
Decision Tree	70.2	83.5	77.6	62	**73**	68
SVM majority vote	74.2	93.2	**96.6**	72	70	46.4

5.3 Clinical Study

The promising results of the pilot study showed that automated assessment of the SLS based on pose data and expert labels is possible. However, there were some limitations in the pilot study:

- Variability of motion generation and labelling in the dataset. The collected data were a mixture of naturally performed squats labelled by an expert and replicated "poor", "moderate", and "good" squats. We found no meaningful difference between LOSO error of the two groups. However, feature selection/extraction and subsequent

classification results could have been different between the two groups if we developed different classifiers for each. Due to the small number of available naturally performed squats and unbalanced distribution of "good", "moderate", and "poor" exemplars, this was not possible with the pilot dataset.

- The small size of the training set. Although we had 100 examples of squats, many of them were generated by the same subjects, limiting our ability to test the generalization capabilities of the method.

To investigate the generalizability and reliability of the assessment method, a second dataset of SLS was collected and provided by the clinical collaborator of the project.

The following are the differences between the pilot and field datasets:

- Larger number of participants with the same number of males and females were recruited.
- The exclusion criteria were different between the two studies. For the pilot study, participants were excluded if they had lower back or leg injuries in the past six months whereas, in the clinical study, they must not have had a medical history restricted participation in a standard musculoskeletal clinical examination of the lower extremity.
- There were two labelling criteria: amount of observed DKV as well as overall risk of injury.
- All squats were performed naturally with both legs and labelled by three expert clinical raters.
- Reliabality analysis of raters was performed.
- Additional training datasets (differing based on features, gender, or agreement between raters) were generated.

5.3.1 Clinical Experiments

The dataset was collected by our clinical partner, MSK Metrics. 14 participants including 7 males and 7 females with mean age of 30.8 ± 5.5, mean height of 173.8 ± 12 cm, and mean weight of 70.4 ± 10.4 kg participated in the study. For two participants, the dominant

leg (the leg they would kick a soccer ball with) was the left; the other participants were right legged. To be included in the study, subjects had to be between the ages of 18-65 years, and must not have had a medical history that restricted participation in a standard musculoskeletal clinical examination of the lower extremity. This would include the clinical suspicion of an emergent health issue, severe neurological compromise, or the presence of an acute fracture, dislocation or severe knee ligament instability. Subjects were excluded from the study if they did not meet any of the above inclusion criteria.

Ethics approval from Institutional Review Board Services was obtained prior to the start of the study. All participants signed a consent form prior to the start of data collection.

Given our inclusion criteria, it is possible that some individuals with active knee pain or discomfort could have been recruited. Considering we were primarily interested in analyzing data from individuals without active knee injury, we identified subjects with active knee pain or discomfort by having participants fill out the International Knee Documentation Committee (IKDC) subjective knee evaluation form for each knee [10]. Data from all 14 subjects were used in the Inter and Intra-rater Reliability (IRR) analyses, while only data from knees that scored over 95% on the IKDC were included in training and cross validation.

Data Collection

Three Yost [6] IMUs were attached to the participants' low back at the level of the first sacral vertebra, the anterior thigh 10 cm above the patella aligned with the sagittal plane, and the flat surface of the shank at the level of the tibial tuberosity using hypoallergenic tape. Sensor placement locations are depicted in Figure 5.7. Data was communicated to a nearby computer via Bluethooth communication with an average sampling rate of 90 ± 10 Hz. Data were interpolated and resampled to the same rate of 200 Hz before subsequent analysis.

Participants were instructed to take off their shoes and perform five continuous cycles of SLS with their toes pointing forward and arms crossed in front of the body. They were asked to perform SLS with both the right and left legs without moving the foot or lifting the heel. In instances where subjects lost their balance, their legs contacted each other, or the non-weight bearing leg touched the ground, the trial was deemed unsuccessful and all cycles were repeated.

Figure 5.7: Sensor placement during SLS clinical data collection

Data Labelling

The participants' performance was videotaped during the tests. Videos were then reviewed by three expert clinicians with advanced training in a sports sciences fellowship, with an average of 9 years clinical experience.

Raters were asked to label each squat repetition. The clinical rating criteria were adapted and modified from [42] and included 2 items: "Knee Valgus" and "Rater's Subjective Overall Knee Injury Risk". We aimed to discriminate between "good" and "poor" squats or screen "high risk" subjects from "no risk" ones (2-class classification), and to assess if a finer grained assessment is possible by adding a "moderate" grading level (3-class classification). For this purpose, each criterion was comprised of a three-level rating scale of "0", "1" or "2". For the knee valgus criterion [42], a score of "0" was defined as no valgus, "1" as moderate knee valgus, and "2" as severe knee valgus. For the overall knee risk of injury criterion, a score of "0" was defined when the individual was at no risk and no intervention was required, a score of "1" was defined when there was mild/low risk and moderate intervention was required, and a score of "2" was defined when the individual was at moderate to high risk and significant intervention was required. To determine the

rating for the "overall knee injury risk assessment" item, the clinical raters were asked to base their assessment of the overall whole body motion that occurred when the subject was performing the SLS, not just the subject's knee position.

The 14 participants performed 5 SLS repetitions with both left and right legs resulting in 140 squat repetitions to be labelled. Three categories were created from the labelled samples: samples which were unanimous (U) among raters, samples with a split (S) decision among raters, where two raters gave the same score and one gave a different score, and samples for which there was no consensus among raters, where each rater gave a different score. Labeled data statistics for each of the two criteria are summarized in Table 5.4. Samples that came from participants who scored less than 95% on the IKDC are referred to as unhealthy in the table and are excluded from the analysis.

For split decision ratings, a final label based on majority vote was given to the samples. For feature selection, 4 different datasets were generated: two with combinations of both healthy unanimous and healthy split decision samples (for the two different criteria), and the others with only healthy unanimous samples (again for the two criteria).

Table 5.4: Labeled Data Information

U: unanimous S: split decision H: healthy	Labeled with knee valgus criterion		Labeled with overall risk of knee injury criterion	
	Male#	Female#	Male#	Female#
Good(U,H)	7	5	1	5
Good(S,H)	11	16	7	8
Moderate(U,H)	10	5	9	1
Moderate(S,H)	18	16	18	15
Poor(U,H)	6	4	5	14
Poor(S,H)	11	10	22	12
No Consensus(H)	2	4	3	5
Unhealthy	5	10	5	10
Total	70	70	70	70

For classification, only the datasets which included both split decision and unanimous samples were utilized. We made two different datasets for each labelling index: one included combined healthy split decision and unanimous samples of the three categories to be used

for 3-class classification. The other was generated by removing "moderate" exemplars from the previous set to implement 2-class classification. Details of the training datasets are summarized in Table 5.5.

Table 5.5: Training dataset details

		Labelled with knee valgus criterion	Labelled with overall risk of injury criterion
Training and validation sets	Healthy Unanimous or Split decision	119 exemplars (39 "good", 49 "moderate", 31 "poor")	117 exemplars (21 "good", 43 "moderate", 53 "poor")
	Healthy Unanimous	37 exemplars (12 "good", 15 "moderate", 10 "poor")	35 exemplars (6 "good", 10 "moderate", 19 "poor")
Removed samples	Unhealthy and no - consensus	21 exemplars	23 exemplars

5.3.2 Inter and Intra Rater Reliability (IRR) Analysis

Since we have more than one rater in this study, we analyzed the degree of agreement between them (inter rater reliability), and the consistency of the ratings by each of the raters (intra rater reliability).

Three raters ranked SLS data on a three point scale so the ratings were ordinal. IRR assessment was performed using the two-way mixed, consistency, average-measures ICC test. Details of the test selection criteria can be found in Appendix B. Calculations were made using MedCalc [3]. The resulting ICC value was 0.80 for the knee valgus criterion (Figure 5.8(a)) and 0.84 for the risk of injury criterion (Figure 5.8(b)). This indicates excellent agreement between raters according to Cicchetti guidelines [22].

To assess intra-rater reliability, 15 of the 140 squat samples were randomly selected and duplicated in the dataset provided to the raters for labeling. The two-way mixed,

Intraclass correlation coefficient

Number of subjects (n)	140
Number of raters (k)	3
Model	The same raters for all subjects. Two-way model.
Type	Consistency
Measurements	AL AR PK

Intraclass Correlation Coefficient

	Intraclass correlation [a]	95% Confidence Interval
Single measures [b]	0.5636	0.4716 to 0.6493
Average measures [c]	0.7949	0.7281 to 0.8474

[a] The degree of consistency among measurements.
[b] Estimates the reliability of single ratings.
[c] Estimates the reliability of averages of *k* ratings.

(a) IRR results for knee valgus labeling criterion

Intraclass correlation coefficient

Number of subjects (n)	140
Number of raters (k)	3
Model	The same raters for all subjects. Two-way model.
Type	Consistency
Measurements	AL AR PK

Intraclass Correlation Coefficient

	Intraclass correlation [a]	95% Confidence Interval
Single measures [b]	0.6302	0.5462 to 0.7064
Average measures [c]	0.8364	0.7831 to 0.8783

[a] The degree of consistency among measurements.
[b] Estimates the reliability of single ratings.
[c] Estimates the reliability of averages of *k* ratings.

(b) IRR results for risk of injury labeling criterion

Figure 5.8: Inter-rater reliability evaluation by MedCalc

consistency, average-measures ICC test was applied to two ratings provided for the original and duplicated samples by each rater. Intra-rater reliability results for the three raters were 1, 0.96, and 0.88 suggesting excellent reliability for all raters.

Intraclass correlation coefficient

Number of subjects (n)	15
Number of raters (k)	2
Model	The same raters for all subjects. Two-way model.
Type	Consistency
Measurements	score1 score2

Intraclass Correlation Coefficient

	Intraclass correlation [a]	95% Confidence Interval
Single measures [b]	1.0000	
Average measures [c]	1.0000	

[a] The degree of consistency among measurements.
[b] Estimates the reliability of single ratings.
[c] Estimates the reliability of averages of *k* ratings.

(a) rater 1

Intraclass correlation coefficient

Number of subjects (n)	15
Number of raters (k)	2
Model	The same raters for all subjects. Two-way model.
Type	Consistency
Measurements	score1 score2

Intraclass Correlation Coefficient

	Intraclass correlation [a]	95% Confidence Interval
Single measures [b]	0.7761	0.4541 to 0.9188
Average measures [c]	0.8739	0.6245 to 0.9577

[a] The degree of consistency among measurements.
[b] Estimates the reliability of single ratings.
[c] Estimates the reliability of averages of *k* ratings.

(b) rater 2

Intraclass correlation coefficient

Number of subjects (n)	15
Number of raters (k)	2
Model	The same raters for all subjects. Two-way model.
Type	Consistency
Measurements	score1 score2

Intraclass Correlation Coefficient

	Intraclass correlation [a]	95% Confidence Interval
Single measures [b]	0.9187	0.7760 to 0.9720
Average measures [c]	0.9577	0.8739 to 0.9858

[a] The degree of consistency among measurements.
[b] Estimates the reliability of single ratings.
[c] Estimates the reliability of averages of *k* ratings.

(c) rater 3

Figure 5.9: Intra-rater reliability evaluations by MedCalc

The IRR assessment results suggest that the measurement error introduced by individual raters is minimal and that SLS ratings are suitable for the purpose of classification.

5.3.3 Clinical Study Results

All Joints' Features

Tables 5.6 to 5.9 show the feature selection results for the four datasets and two different classification problems (2-class versus 3-class).

Table 5.6: Feature selection results for 2-class problem and knee valgus criterion

Knee Valgus criterion- 2class ("good" vs "poor")			
Healthy (Unanimous)	N_r	Healthy (Unanimous+Split decision)	N_r
RMS of ankle IR angle	12	Mean of hip Flex. angle RMS of hip Flex. angle	10
Mean of ankle IR angle	9	Max of hip Flex. angle	9

N_r: Number of times ranked as top ten features

Table 5.7: Feature selection results for 2-class problem and risk of injury criterion

Risk of injury criterion- 2class ("good" vs "poor")			
Healthy (Unanimous)	N_r	Healthy (Unanimous+Split decision)	N_r
Mean of hip Flex. angle	7	Mean of hip Flex. angle	14
RMS of ankle IR angle	6	Mean of knee Flex. angle	11
RMS of ankle IR acceleration		RMS of ankle IR angle	10
RMS of hip IR angle		Max of hip Flex. angle	9

N_r: Number of times ranked as top ten features

The feature selection results for the unanimous data in both 2-class and 3-class problems reveal that ankle IR/Add features are the most important predictors of DKV, while in terms of risk of injury, hip IR and Flex and ankle IR features are more discriminative. Another observation from the unanimous data is that in the 2-class problem, joint angle features appear as predictors, while for the 3-class problem, joint velocity plays a significant role. These results are in agreement with our pilot study.

Table 5.8: Feature selection results for 3-class problem and knee valgus criterion

Knee valgus criterion- 3 class ("good" vs "moderate" vs "poor")			
Healthy (Unanimous)	N_r	Healthy (Unanimous+Split decision)	N_r
Kurtosis of ankle Add. angle	12	Max of hip IR. angle	15
Mean of ankle IR angle RMS of ankle Add. velocity Min of ankle IR. velocity Max of ankle IR. velocity STD of ankle IR. acceleration	7	Mean of hip Flex. angle	13
		Min of knee Flex. angle	11
		Range of hip Flex. angle RMS of hip Flex. angle	9

N_r: Number of times ranked as top ten features

Table 5.9: Feature selection results for 3-class problem and risk of injury criterion

Risk of injury criterion- 3 class ("good" vs "moderate" vs "poor")			
Healthy (Unanimous)	N_r	Healthy (Unanimous+Split decision)	N_r
STD of hip IR velocity	11	Max of hip Flex. angle	10
Range of hip Flex. angle MAD of hip IR velocity	10	Mean of hip Flex. angle VAR of ankle Add. velocity	9
VAR of hip IR velocity RMS of hip IR angle	9		

N_r: Number of times ranked as top ten features

On the other hand, when the data includes both unanimous and split decision samples, flexion angles, particularly hip Flex, frequently appear as predictors of the knee valgus or risk of injury.

The dimensionality of the training data was reduced by keeping only the identified important features. Classification techniques are applied to the reduced-dimensionality dataset including both unanimous and split decision data, using the labels from the two criteria. Results for both 10 fold and LOSO cross-validations are reported in Tables 5.10 to 5.13. SPCA dimensionality reduction method results are also provided.

Table 5.10: Classification results for 2-class problem and knee valgus criterion

2-class classification problem accuracy(%) **Knee Valgus criterion**				
Validation method	10F-CV		LOSO-CV	
Dim. Red. method	Subset of selected features	SPCA	Subset of selected features	SPCA
SVM	92.71	**93.14**	87.5	**89.77**
NB	92.71	90.42	87.5	84.1
KNN	92.85	92.57	86.36	86.36

Table 5.11: Classification results for 3-class problem and knee valgus criterion

3-class classification problem accuracy(%) **Knee Valgus criterion**				
Validation method	10F-CV		LOSO-CV	
Dim. Red. method	Subset of selected features	SPCA	Subset of selected features	SPCA
SVM	66.87	**73.5**	65.4	67.4
NB	60.99	67.7	57	59.6
KNN	70.12	72.15	**67.6**	67.2

Table 5.12: Classification results for 2-class problem and risk of injury criterion

2-class classification problem accuracy(%) **Risk of Injury criterion**				
Validation method	10F-CV		LOSO-CV	
Dim. Red. method	Subset of selected features	SPCA	Subset of selected features	SPCA
SVM	92.11	86.9	77.46	84.1
NB	**92.25**	85.39	**87.39**	82.25
KNN	93.7	87.32	85.87	79.35

Table 5.13: Classification results for 3-class problem and risk of injury criterion

3-class classification problem accuracy(%) **Risk of Injury criterion**				
Validation method	10F-CV		LOSO-CV	
Dim. Red. method	Subset of selected features	SPCA	Subset of selected features	SPCA
SVM	67.17	74.4	61.67	**74.87**
NB	66.34	68.67	66.67	61.27
KNN	67.86	**76.3**	66.26	73.27

Classification results for 10F-CV showed that distinguishing between "good" and "poor" squats is achievable with a promising accuracy (93%). For the 3- class problem; however, the best achieved accuracy was 74%. LOSO-CV results were slightly lower, with best accuracy of 90% for 2-class and 68% for the 3-class problems. With respect to predicting the risk of injury, the best achieved accuracy using 10F-CV was 95% for the 2-class and 76% for the 3-class problem. Using the LOSO-CV, the best accuracy for 2-class was 87% and for 3-class problem was 75%.

Ankle Only Features

In the pilot data analysis, we found the ankle IR features to be the best predictors of the DKV, which led us to suggest that it is possible to use only one sensor on the tibia (saving time and simplifying the test protocol) and still have good classification accuracy. To confirm this hypothesis with the larger datasets, we used feature selection on only ankle extracted features (90 out of 210 features) and found that ankle IR velocity, angle and acceleration features are the best predictors in the absence of hip or knee information. We also repeated the classification using ankle only features. The best achieved results using ankle only features and the percentage of change in accuracy in comparison to the best reported results using all joints features are shown in Tables 5.14 and 5.15.

Table 5.14: Best achieved classification results for 10F-CV using ankle features

10F-CV accuracy(%)				
	Knee Valgus criterion		**Risk of Injury criterion**	
Best results	ankle only features	change in accuracy	ankle only features	change in accuracy
2-Class	84.14	-9%	91.67	-3.6%
3-Class	67.5	-6%	77.13	+1.23%

The results from Tables 5.14 and 5.15 indicate that there is less than 4% drop in accuracy for risk of injury detection using only ankle information (one tibia sensor), suggesting that one sensor can be used to simplify the data collection procedure, particularly if overall risk of injury is of interest.

Table 5.15: Best achieved classification results for LOSO-CV using ankle features

LOSO-CV accuracy(%)				
	Knee Valgus criterion		**Risk of Injury criterion**	
Best results	ankle only features	change in accuracy	ankle only features	change in accuracy
2-Class	85.23	-4.54%	83.7	-3.69%
3-Class	60.4	-7.2%	73.5	-1.37%

Gender-specifc

We hypothesized that men and women might have different biomechanical characteristics and movement strategies which result in different predictors. To test this hypothesis, we separated the healthy subject data based on gender, resulting in female only data (60 samples) and male only data (65 samples). Feature selection methods were applied to combined unanimous and split decision samples of both datasets separately. The results reported in Tables 5.16 to 5.19 showed that different features are selected when the data is segregated by gender. For the male dataset, the features selected were the hip and knee flexion features. For females, hip and ankle IR features were selected. Based on this finding, we also tested whether male-specific and female-specific classifiers might work better than a general classifier for both genders. The SVM classifier was used for the two data set and results are compared to general classifier results (developed in previous section) in tables 5.20 and 5.21.

Classification results show that for women, in all cases, the female-specific classifier works better than the general classifier. For men, the same holds for risk of injury index. The only noticeable exception is the 2-class classification with knee valgus criterion, for which the general classifier is better.

Table 5.16: Gender-specific feature selection results for 2-class problem and knee valgus criterion

Knee Valgus criterion- 2 class ("good" vs "poor")			
Males	N_r	Females	N_r
Mean of knee Flex. angle	13	RMS of ankle IR velocity	10
Max of hip Flex. velocity	11	STD of ankle IR velocity	7
RMS of hip Flex. angle STD of hip Flex. angle MAD of hip Flex. angle	7	RMS of ankle IR acceleration MAD of ankle IR acceleration VAR of ankle IR velocity MAD of ankle IR velocity Mean of ankle Add. velocity	6

N_r: Number of times ranked as top ten features

Table 5.17: Gender-specific feature selection results for 2-class problem and Risk of Injury criterion

Risk of Injury criterion- 2 class ("good" vs "poor")			
Males	N_r	Females	N_r
MAD of hip Flex. velocity	10	STD of hip IR velocity	11
Kurtosis of hip Flex. velocity	8	Mean of knee Flex. angle	9
RMS of hip Flex. velocity	7	Mean of hip Flex. angle VAR of hip IR velocity	8

N_r: Number of times ranked as top ten features

Table 5.18: Gender-specific feature selection results for 3-class problem and Knee Valgus criterion

Knee Valgus criterion- 3 class ("good" vs "moderate" vs"poor")			
Males	N_r	Females	N_r
RMS of hip Flex. angle	13	RMS of hip Add. velocity	8
Max of hip Flex. velocity	12	MAD of hip IR acceleration	7
Mean of hip Flex. velocity	9	VAR of ankle IR velocity	6
STD of hip Flex. angle VAR of hip Flex. angle MAD of hip Flex. angle Mean of hip Flex. angle	8	MAD of ankle IR velocity STD of hip Add. velocity	

N_r: Number of times ranked as top ten features

Table 5.19: Gender-specific feature selection results for 3-class problem and Risk of Injury criterion

Risk of Injury criterion- 3 class ("good" vs "moderate" vs"poor")			
Males	N_r	Females	N_r
STD of hip Flex. angle	9	STD of hip IR velocity MAD of ankle IR acceleration	9
Mean of hip Flex. velocity Kurtosis of knee Felx. acceleration	8	STD of ankle IR velocity RMS of ankle IR acceleration	
Kurtosis of knee Flex. velocity	7		8

N_r: Number of times ranked as top ten features

Table 5.20: Gender-specific classification results for 10F-CV

10F-CV accuracy(%)				
Classifier type	2-Class		3-Class	
	Valgus	Risk	Valgus	Risk
Male only- SVM	90.5	97.3	84	73.6
Female only- SVM	94.3	99.8	74.7	84.9
General- best results	93.1	95.3	73.5	74.4

Table 5.21: Gender-specific classification results for LOSO-CV

LOSO-CV accuracy(%)				
Classifier type	2-Class		3-Class	
	Valgus	Risk	Valgus	Risk
Male only- SVM	72.7	91.7	82.7	80
Female only- SVM	93.2	100	74.6	86.8
General- best results	89.8	87.4	67.6	74.9

5.3.4 Discussion

The identification of ankle IR/Add and hip IR/Flex features for the determination of DKV and risk of injury are consistent with previous reports evaluating the kinematic characterization of DKV, since ankle IR/Add and hip IR/Flex represent component joint motions that contribute to the multi-joint DKV movement [98], [96], [43], [92]. Moreover, knee flexion angle features were also identified as predictors of dynamic valgus and knee injury risk in the present study. This finding is also consistent with previous investigations that have identified a shallow knee flexion angle during single leg loading to be correlated with dynamic knee valgus loading, ACL injury, and patellofemoral pain syndrome [92]. Knee flexion angles less than 30° have been shown to cause a large strain force on the ACL caused by quadriceps contraction, and shallow knee flexion angles coupled with hip internal rotation may also increase patellofemoral contact forces [63], [62], [82], [44].

Analyzing the flexion joint angles of the SLS repetitions revealed that those labelled as “good” tended to have increased knee and torso flexion during the motion. Previous

research investigating the effect of forward trunk lean on predicted anterior cruciate ligament (ACL) strains that occur during the SLS movement indicate that a more moderate forward trunk lean of approximately 40 degrees can lower ACL strains and increase muscle activation of the hamstring muscles that assist in preventing anterior tibial translation and lower activation of the quadriceps muscles that can increase anterior tibial translation [49]. It is possible that the torso flexion observed in the "good" subjects in the present study may be a result of a movement strategy utilized to minimize internal knee loads and optimize the co-contraction of the quadriceps and hamstring musculature.

The achieved performance in the 2-class problem is comparable to Whelan et al. [94]. We further showed that the classification generalizes to unseen participants and investigate 3-class classification. Unlike Whelan et al., joint angle, velocity and acceleration features are used, which are clinically interpretable parameters.

The results of gender specific classifiers suggest that developing separate classifiers for men and women improves classification results and strengthens our hypothesis about different biomechanical characteristics or movement strategies in men and women. Previous literature investigating gender differences in the SLS movement test has identified that females perform the SLS with more "valgus collapse" which involves more pelvic rotation, hip internal rotation, femoral adduction, knee external rotation and abduction, and ankle pronation compared to males [34], [91], [92], [98]. Females also perform the SLS with less trunk, hip and knee flexion compared to their gender counterparts [98], [92], [34]. As a result of these gender differences, it is not surprising that our gender-specific classifiers for females primarily involved hip and ankle IR/Add features, and the male classifier primarily involved hip and knee flexion features.

5.4 Summary

In this chapter, an automated assessment method to evaluate SLS quality was developed for a small pilot dataset and expanded to a larger clinical dataset. Two criteria were used for labelling by expert clinician raters: amount of inward knee movement that occurred during the task (knee valgus) and perceived overall knee injury risk. SLS data from 14 volunteers were collected and two data sets were generated: one included the data with unanimous agreement among raters and the other dataset was a combination of full and partial agreement of labelled data. 18 feature selection methods were applied to the datasets to find the best predictors of knee valgus and risk of knee injury. The feature selection results for only unanimous data suggested ankle IR/Add and hip IR/Flex features to be correlated with DKV and risk of injury, respectively. However, for combined unanimous and split

decision data, hip/knee flexion angle features were highlighted as the predictors of both DKV and risk of injury.

Three common classification techniques were applied to the datasets. The LOSO-CV results suggest that discriminating of "poor" squats from "good" ones is achievable with promising accuracy of 90%. Changing the problem to multiclass (adding "moderate" squats) drops the accuracy by 22%. Screening participants at high risk of injury from those at no risk can be achieved with 87% accuracy, while adding mild risk subjects drops the accuracy by 12%. Gender-specific classifiers were also developed and showed improved classification accuracy.

The unanimous cases represent instances where 100% agreement between all three clinician raters occurred, and likely represent cases where the motion characteristics can be clearly identified. By combining both the unanimous and split decision cases, it is possible that some labelling inaccuracies may have been introduced. However, split decision cases may also represent borderline cases where a labelling judgement may be difficult. Considering borderline cases are likely to occur in the population, it is important to understand the impact these cases may have on feature selection and classification results.

In the present study, the participants were not instructed to keep their torso upright during the data collection. The fact that hip flexion angle features appeared as best predictors of DKV and risk of injury in the full dataset indicates that other motion behaviors are also associated with knee valgus and knee injury risk, and that different test protocols and instructions can lead to different results. SLS test protocols do vary between studies, with some authors constraining the squat depth, time duration of the SLS, and upper and lower extremity position [34], [47], [98]. Since the purpose of the SLS test is to assess how an individual functions during single leg loading, which is a foundational movement that is encountered in everyday life and athletic instances, we chose not to constrain the rate and depth of the SLS, to study the subjects own inherent movement preferences while performing the SLS. Our decision to not implement some of these constraints may have affected feature selection results, which has to be considered in the clinical application of the developed tool.

Given our results demonstrating improved classification with gender specific classifiers, and the consistency of previous investigations identifying gender differences in SLS movement performance, we propose that automated SLS assessments should include gender-specific classifiers.

Chapter 6

Conclusions and Future Work

6.1 Summary of Findings

In this thesis, an IMU-based automated pose estimation and assessment method for assessing the single leg squat in terms of amount of observed dynamic knee valugus and risk of knee injury was proposed. The pose estimation accuracy was validated for 7 participants in the Motion Capture Lab and the developed SLS assessment method was tested for a group of 14 healthy participants under clinical test conditions.

Automated Pose Estimation

While IMU-based pose estimation has received extensive attention in the literature, the majority of IMU-based pose estimation techniques are not yet validated for clinical applications due to various issues such as lack of generalization to 3 dimensional motion measurement, limited number of joints measured, indirect joint angle information provision, or the need for complicated and time consuming calibration procedures for accurate results. The proposed approach in this study was adapted from [55] and modified to provide 3 dimensional joint angle, velocity and acceleration estimates of the lower extremity during the SLS. The pose estimation results were validated against marker-based pose estimation and showed acceptable average accuracy of 5.1°. This accuracy is acceptable for assessment purposes, as according to the decision tree analysis (Figure 5.6), the classification thresholds are significantly greater than this error. The proposed method has advantages over visual assessment, which is currently the main method of dynamic movement evaluation in the clinic, as it provides quantitative measures of absolute joint angle

values as well as velocity and acceleration (which are hard to assess visually) during the lower extremity mobility tests.

Sensitivity Analysis

The pose estimation method was simplified for easier clinical application, by replacing the required manual measurements by approximated values from anthropometric tables and making assumptions about sensor placement. The amount of error introduced by these simplifications was evaluated. A sensitivity analysis was performed to identify the most critical parameters. Finally, a simple and easy to apply calibration method was proposed for better estimation of sensitive parameters, which resulted in considerable improvement in accuracy. The proposed calibration method was successfully applied in the clinical setting by clinicians, which shows the effectiveness and feasibility of the method.

Automate Dynamic Knee Valgus and Risk of Injury Assessment

The proposed pose estimation method was applied to a dataset collected from 14 healthy participants and labelled by three expert raters in a clinical setting. The estimated hip, knee, and ankle joint position, velocity and acceleration were then used for assessment of squat quality using a number of feature selection and classification techniques. The results showed that classification of squat quality into two classes of good and poor or low risk and high risk is possible with a high accuracy of $93 - 95\%$. Increasing the number of classes by adding moderate squats drops the accuracy by $12 - 22\%$. Further analysis was performed to identify DKV correlates from feature selection and to develop more accurate gender-specific classifiers.

To our knowledge, the present study is the first to investigate an automated SLS 3-class classification, which would be beneficial for clinicians, as this would allow a determination of not just the presence or absence of DKV and overall knee injury risk, but it would also provide an assessment of the severity of these parameters. This stratification could assist clinicians in developing interventions that could be tailored to an individual's severity level of DKV and knee injury risk. Our study is also unique in that we developed a classifier for the identification of overall knee injury risk based on the SLS movement test. The successful performance of a SLS requires the precise coordination and control of movement about multiple joints (the trunk, hip, knees and ankles) while simultaneously maintaining balance over a small base of support. Considering variables other than DKV, such as lateral trunk position [92] and control of the non-squatting leg [47], has been shown to influence

knee loading during the SLS, we had expert clinical raters judge the entire composite SLS movement to rate individuals on knee injury risk.

6.2 Future Work

6.2.1 Sensor Orientation Estimation

The developed calibration protocol in this study extracts the sensors' orientation on the body based on gyro data, which improves the pose estimation accuracy. However, it depends on the correct performance of the protocol by the clinician. Other possible orientation extraction methods which can be used with minimum effort in the clinics need to be explored. With the current method, visual guidance for ensuring the correct implementation can be developed through computer interfaces and wizards. One source of error in this method can be the initial non zero orientation of the participant with respect to the sagittal plane, which will be interpreted as a roll angle of the tibia and thigh sensors and yaw angle in the back sensor. An additional step can be added to the method to detect this offset.

6.2.2 Analysis of Motion Artifacts

The effect of soft tissue artifacts on the pose estimation accuracy can be better investigated by designing the following experiments:

- Trying different sensor placements and comparing the amount of sensor motion resulting from each placement.
- Using a fixed sensor placement and examining the effectiveness of different pose estimation methods (e.g. EKF-based vs non-EKF based) in combatting motion artifacts.

6.2.3 Automated Segmentation

A challenge in the developed assessment method was correct segmentation of the squat samples. Two different methods were tried. However, no method can guarantee segmentation with 100% accuracy, so a final visual check and tuning of the segmentation algorithm were necessary, which remains an impediment to full automation of the technique. A more

advanced motion segmentation technique [56] suitable for the SLS with high accuracy of segmentation can be utilized to fully automate the process. Alternatively, an assessment method can be developed for the whole period of the motion (5 continuous squats in this study) and not for individual squats.

6.2.4 Larger Sample Size including Rehabilitation Patients

A limitation of the present study includes the small sample size. In addition, our sample was comprised of healthy individuals with no current knee injury. There likely exist many different movement strategies to perform a SLS within the population; therefore, future work should consider a larger sample, and classification of subject samples with knee pathology.

6.2.5 Providing Continuous Score for SLS Assessment

The current assessment method classifies the squats into three categories. Another useful approach for clinical purposes would be an evaluation based on a continuous score. This can be done through regression analysis. However, for better results of the regression, a continuous way of labelling or a larger number of grading categories for the movement is preferred.

6.2.6 Extending the Assessment Method to Other Mobility Tests

There are several other mobility tests such as the Double Leg Jump, Lunge, and Single Leg Hop which can be analysed using the same proposed approach. Some of these tests, like the Single Leg Hop or Double Leg Jump introduce additional challenges due to sensor vibrations resulting from impact. Also the motion consists of flying periods when the foot is no longer in contact with the ground, during which the current approach for pose estimation would be unreliable because it assumes a fixed base. Extending the proposed method to those tests can be an interesting direction for future work.

References

[1] 3 Space Sensor User's Manual-Yost Labs. yostlabs.com/wp/wp-content/.../3.../YEI_TSS_Users_Manual_3.0_r1_4Nov2014.pdf.

[2] Journal of Translation Engineering in Health and Medicine. http://health.embs.org/.

[3] MedCalc Statistical Software. https://www.medcalc.org/.

[4] Polhemus. http://polhemus.com/motion-tracking/overview/.

[5] Xsens. http://www.xsens.com/.

[6] Yost-IMU. http://yostlabs.com/.

[7] Eva Ageberg, Kim L Bennell, Michael a Hunt, Milena Simic, Ewa M Roos, and Mark W Creaby. Validity and inter-rater reliability of medio-lateral knee motion observed during a single-limb mini squat. *BMC musculoskeletal disorders*, 11(1):265, 2010.

[8] Nanni J Allington, Nathalie Leroy, and Carole Doneux. Ankle joint range of motion measurements in spastic cerebral palsy children: intraobserver and interobserver reliability and reproducibility of goniometry and visual estimation. *Journal of Pediatric Orthopaedics B*, 11(3):236–239, 2002.

[9] Adel Almangoush, Lee Herrington, and Richard Jones. A preliminary reliability study of a qualitative scoring system of limb alignment during single leg squat. *Physical Therapy and Rehabilitation*, 1(1):2, 2014.

[10] Allen F Anderson, James J Irrgang, MS Kocher, Barton J Mann, and John J Harrast. The International Knee Documentation Committee Subjective Knee Evaluation Form: normative data. *The American journal of sports medicine*, 34(1):128–135, 2006.

[11] Andreas Aristidou and Joan Lasenby. Real-time marker prediction and CoR estimation in optical motion capture. *The Visual Computer*, 29(1):7–26, 2013.

[12] Robert Bailey, James Selfe, and Jim Richards. The Single Leg Squat Test in the Assessment of Musculoskeletal Function: a Review. *Physiotherapy Practice and Research*, 31(1):18–23, 2011.

[13] Elnaz Barshan, Ali Ghodsi, Zohreh Azimifar, and Mansoor Zolghadri Jahromi. Supervised principal component analysis: Visualization, classification and regression on subspaces and submanifolds. *Pattern Recognition*, 44(7):1357–1371, 2011.

[14] Vinay A Bavdekar, Anjali P Deshpande, and Sachin C Patwardhan. Identification of process and measurement noise covariance for state and parameter estimation using extended kalman filter. *Journal of Process control*, 21(4):585–601, 2011.

[15] Natalia F N Bittencourt, Juliana M Ocarino, Luciana D M Mendonça, Timothy E Hewett, and Sergio T Fonseca. Foot and hip contributions to high frontal plane knee projection angle in athletes: a classification and regression tree approach. *The Journal of orthopaedic and sports physical therapy*, 42(12):996–1004, 2012.

[16] Lori Bolgla, Naomi Cook, Kyle Hogarth, Jennifer Scott, and Cary West. Trunk and Hip Electromyographic Activity During Single Leg Squat Exercises Do Sex Differences Exist? *International Journal of Sports Physical Therapy*, 9(6):756–764, 2014.

[17] Vincent Bonnet, Claudia Mazzà, Philippe Fraisse, and Aurelio Cappozzo. A least-squares identification algorithm for estimating squat exercise mechanics using a single inertial measurement unit. *Journal of Biomechanics*, 45(8):1472–1477, 2012.

[18] Miranda C. Boonstra, Rienk M A Van Der Slikke, Noël L W Keijsers, Rob C. Van Lummel, Maarten C. De Waal Malefijt, and Nico Verdonschot. The accuracy of measuring the kinematics of rising from a chair with accelerometers and gyroscopes. *Journal of Biomechanics*, 39(2):354–358, 2006.

[19] A Brennan, K Deluzio, and Q Li. Assessment of anatomical frame variation effect on joint angles: a linear perturbation approach. *Journal of biomechanics*, 44(16):2838–2842, 2011.

[20] Julien Chardonnens, Julien Favre, and Kamiar Aminian. An effortless procedure to align the local frame of an inertial measurement unit to the local frame of another motion capture system. *Journal of Biomechanics*, 45(13):2297–2300, 2012.

[21] Chien-Yen Chang, B. Lange, Mi Zhang, S. Koenig, P. Requejo, Noom Somboon, A.a. Sawchuk, and A.a. Rizzo. Towards Pervasive Physical Rehabilitation Using Microsoft Kinect. In *6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, pages 159–162, 2012.

[22] Domenic V Cicchetti. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment*, 6(4):284, 1994.

[23] Kay M Crossley, Wan-Jing Zhang, Anthony G Schache, Adam Bryant, and Sallie M Cowan. Performance on the single-leg squat task indicates hip abductor muscle function. *The American journal of sports medicine*, 39(4):866–73, 2011.

[24] Pasquale Daponte, Luca De Vito, Maria Riccio, and Carmine Sementa. Design and validation of a motion-tracking system for ROM measurements in home rehabilitation. *Measurement*, 55:82–96, 2014.

[25] Mario A DiMattia, Ann L Livengood, Tim L Uhl, Carl G Mattacola, and Terry R Malone. What are the validity of the single-leg-squat test and its relationship to hip-abduction strength? *Journal of Sport Rehabilitation*, 14(2):108–123, 2005.

[26] John Z Edwards, Kenneth A Greene, Robert S Davis, Mark W Kovacik, Donald A Noe, and Michael J Askew. Measuring flexion in knee arthroplasty patients. *The Journal of arthroplasty*, 19(3):369–372, 2004.

[27] Mahmoud El-Gohary and James McNames. Human joint angle estimation with inertial sensors and validation with a robot arm. *IEEE Transactions on Biomedical Engineering*, 62(7):1759–1767, 2015.

[28] J. Favre, R. Aissaoui, B. M. Jolles, J. A. de Guise, and K. Aminian. Functional calibration procedure for 3D knee joint angle description using inertial sensors. *Journal of Biomechanics*, 42(14):2330–2335, 2009.

[29] J. Favre, B. M. Jolles, R. Aissaoui, and K. Aminian. Ambulatory measurement of 3D knee joint angle. *Journal of Biomechanics*, 41(5):1029–1035, 2008.

[30] Daniel Tik Pui Fong and Yue Yan Chan. The use of wearable inertial motion sensors in human lower limb biomechanics studies: A systematic review. *Sensors*, 10(12):11556–11565, 2010.

[31] C D Fryar, Q Gu, and C L Ogden. Anthropometric reference data for children and adults: United States, 2007-2010. *Vital Health Statistics*, (11):1–40, 2012.

[32] Richard L Gajdosik and Richard W Bohannon. Clinical measurement of range of motion. *Phys ther*, 67(12):1867–1872, 1987.

[33] Alex L Gornitzky, Ariana Lott, Joseph L Yellin, Peter D Fabricant, J Todd Lawrence, and Theodore J Ganley. Sport-Specific Yearly Risk and Incidence of Anterior Cruciate Ligament Tears in High School Athletes: A Systematic Review and Meta-analysis. *The American journal of sports medicine*, pages 0363546515617742–, 2015.

[34] G Graci, V; Van Dillen, L; Salsich. Gender differnces in trunk pelvis and lower limb kinematics during a single leg squat. *Gait Posture*, 36(3):461–466, 2013.

[35] Kevin A Hallgren. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutor Quant Methods Psychol*, 8(1):23–34, 2012.

[36] DM Hamby. A review of techniques for parameter sensitivity analysis of environmental models. *Environmental monitoring and assessment*, 32(2):135–154, 1994.

[37] M. E. Harrington, A. B. Zavatsky, S. E M Lawson, Z. Yuan, and T. N. Theologis. Prediction of the hip joint centre in adults, children, and patients with cerebral palsy based on magnetic resonance imaging. *Journal of Biomechanics*, 40(3):595–602, 2007.

[38] Marcie Harris-Hayes, Karen Steger-May, Christine Koh, Nat K Royer, Valentina Graci, and Gretchen B Salsich. Classification of lower extremity movement patterns based on visual assessment: reliability and correlation with 2-dimensional video analysis. *Journal of athletic training*, 49(3):304–310, 2014.

[39] P Hattam and A Smeatham. *Special tests in musculoskeletal examination: an evidence-based guide for clinicians*. Elsevier Health Sciences.

[40] Tom O Haver. A Pragmatic Introduction to Signal Processing.

[41] Lee Herrington. Knee valgus angle during single leg squat and landing in patellofemoral pain patients and controls. *Knee*, 21(2):514–517, 2014.

[42] Lee Herrington and Allan Munro. A preliminary investigation to establish the criterion validity of a qualitative scoring system of limb alignment during single-leg squat and landing. *J Exerc Sports Orthop*, 1(2):1–6, 2014.

[43] Timothy E Hewett, Gregory D Myer, Kevin R Ford, Robert S Heidt, Angelo J Colosimo, and Scott G McLean. Biomechanical Measures of Neuromuscular Control and Valgus Loading of the Knee Predict Anterior Cruciate Ligament Injury Risk in Female Athletes A Prospective Study Biomechanical Measures of Neuromuscular Control and Valgus Loading of the Knee Predict Ant. *The American journal of sports medicine*, 33(4):492–501, 2005.

[44] Gerwyn Hughes. A review of recent perspectives on biomechanical risk factors associated with anterior cruciate ligament injury. *Research in sports medicine*, 22(2):193–212, 2014.

[45] Vladimir Joukov. Pose estimation and segmentation for rehabilitation. Master's thesis, University of Waterloo, 2015.

[46] Vladimir Joukov, Vincent Bonnet, Michelle Karg, Gentiane Venture, and Dana Kulić. Rhythmic ekf for pose estimation during gait. In *Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*, pages 1167–1172. IEEE, 2015.

[47] Anne Khuu, Eric Foch, and Cara L Lewis. Not All Single Leg Squats Are Equal: a Biomechanical Comparison of Three Variations. *International journal of sports physical therapy*, 11(2):201–11, 2016.

[48] Rezvan Kianifar, Alex Lee, Sachin Raina, and Dana Kulić. Classification of Quality of Squats with Inertial Measurement Units in the Single Leg Squat Mobility Test. In *IEEE Engineering in Medicine and Biology Society*, 2016.

[49] Anthony S. Kulas, Tibor Hortobágyi, and Paul Devita. Trunk position modulates anterior cruciate ligament forces and strains during a single-leg squat. *Clinical Biomechanics*, 27(1):16–21, 2012.

[50] Agnes W K Lam, Danniel Varona-Marin, Yeti Li, Mitchell Fergenbaum, and Dana Kulić. Automated Rehabilitation System: Movement Measurement and Feedback for Patients and Physiotherapists in the Rehabilitation Clinic. *HumanComputer Interaction*, 31(3-4):294–334, 2016.

[51] Oscar D Lara and Miguel A Labrador. A Survey on Human Activity Recognition using Wearable Sensors. *Communications Surveys & Tutorials*, 15(3):1192–1209, 2013.

[52] Alberto Leardini, Giada Lullini, Sandro Giannini, Lisa Berti, Maurizio Ortolani, and Paolo Caravaggi. Validation of the angular measurements of a new inertial-measurement-unit based rehabilitation system: comparison with state-of-the-art gait analysis. *Journal of NeuroEngineering and Rehabilitation*, 11(1), 2014.

[53] Qiang Li, Ranyang Li, Kaifan Ji, and Wei Dai. Kalman filter and its application. In *Intelligent Networks and Intelligent Systems (ICINIS), 2015 8th International Conference on*, pages 74–77. IEEE, 2015.

[54] Qingguo Li and Jun-tian Zhang. Post-trial anatomical frame alignment procedure for comparison of 3D joint angle measurement from magnetic/inertial measurement units and camera-based systems. *Physiological measurement*, 35(11):2255–2268, 2014.

[55] Jonathan F S Lin and Dana Kulić. Human pose recovery using wireless inertial measurement units. *Physiological measurement*, 33(12):2099–115, 2012.

[56] Jonathan Feng-Shun Lin, Michelle Karg, and Dana Kulić. Movement primitive segmentation for human motion modeling: A framework for analysis. *IEEE Transactions on Human-Machine Systems*, 46(3):325–339, 2016.

[57] H. J. Luinge, P. H. Veltink, and C. T M Baten. Ambulatory measurement of arm orientation. *Journal of Biomechanics*, 40(1):78–85, 2007.

[58] H. J. Luinge and Peter H. Veltink. Measuring orientation of human body segments using miniature gyroscopes and accelerometers. *Medical and Biological Engineering and Computing*, 43(2):273–282, 2005.

[59] John McCamley, Marco Donati, Eleni Grimpampi, and Claudia Mazzà. An enhanced estimate of initial contact and final contact instants of time using lower trunk inertial sensor data. *Gait & posture*, 36(2):316–318, 2012.

[60] Margaret A McDowell, Cheryl D Fryar, Cynthia L Ogden, and Katherine M Flegal. Anthropometric reference data for children and adults : united states , 2003-2006. *National Health Statistics Reports*, (10):1–48, 2008.

[61] Lee Morton, Lynne Baillie, and Roberto Ramirez-Iniguez. Pose calibrations for inertial sensors in rehabilitation applications. In *1st International Workshop on e-Health Pervasive Wireless Applications and Services (eHPWAS'13)*, pages 204–211, 2013.

[62] Gregory D Myer, Kevin R Ford, Kim D Barber Foss, Arlene Goodman, Adrick Ceasar, Mitchell J Rauh, M P H Facsm, G Jon, and Timothy E Hewett. The Incidence and Potential Pathomechanics of Patellofemoral Pain in Female Athletes. *Clinical biomechanics*, 25(7):700–707, 2010.

[63] Yasuharu Nagano, Hirofumi Ida, Masami Akai, and Toru Fukubayashi. Gender differences in knee kinematics and muscle activity during single limb drop landing. *Knee*, 14(3):218–223, 2007.

[64] Naren Naik, RMO Gemson, MR Ananthasayanam, et al. Introduction to the kalman filter and tuning its statistics for near optimal estimates and cramer rao bound. *arXiv preprint arXiv:1503.04313*, 2015.

[65] Cynthia C Norkin and D Joyce White. *Measurement of joint motion: a guide to goniometry*. FA Davis, 2016.

[66] Fredri Ohberg, Ronnie Lundström, and Helena Grip. Comparative analysis of different adaptive filters for tracking lower segments of a human body using inertial motion sensors. *Measurement science and technology*, 24(8):85703, 2013.

[67] Darin A. Padua, David R. Bell, and Micheal A. Clark. Neuromuscular characteristics of individuals displaying excessive medial knee displacement. *Journal of Athletic Training*, 47(5):525–536, 2012.

[68] Darin A Padua, Stephen W Marshall, Michelle C Boling, Charles A Thigpen, William E Garrett, and I Anthony. The Landing Error Scoring System (LESS) Is a Valid and Reliable Clinical Assessment Tool of Jump-Landing Biomechanics: The JUMP-ACL Study. *The American Journal of Sports Medicine*, 37:1996–2002, 2009.

[69] Mark V. Paterno. Incidence and predictors of second anterior cruciate ligament injury after primary reconstruction and return to sport. *Journal of Athletic Training*, 50(10):1097–1099, 2015.

[70] Jouni Pohjalainen, Okko Rasanen, and Serdar Kadioglu. Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits. *Computer Speech and Language*, 29(1):145–171, 2015.

[71] Daniel R Poulsen and C Roger James. Concurrent validity and reliability of clinical evaluation of the single leg squat. *Physiotherapy theory and practice*, 27(8):586–94, 2011.

[72] Stephen J Preece, John Y Goulermas, Laurence PJ Kenney, Dave Howard, Kenneth Meijer, and Robin Crompton. Activity identification using body-mounted sensorsa review of classification techniques. *Physiological measurement*, 30(4):R1, 2009.

[73] Rami Rachkidi, Ismat Ghanem, Ibrahim Kalouche, Samer El Hage, Fernand Dagher, and Khalil Kharrat. Is visual estimation of passive range of motion in the pediatric lower limb valid and reliable. *BMC musculoskeletal disorders*, 10(1):126, 2009.

[74] Nancy Berryman Reese and William D Bandy. *Joint range of motion and muscle length testing*. Elsevier Health Sciences, 2016.

[75] Olinde Rodrigues. *Des lois géométriques qui régissent les déplacements d'un système solide dans l'espace: et de la variation des cordonnées provenant de ces déplacements considérés indépendamment des causes qui peuvent les produire*. 1840.

[76] McGinnis R.S., Cain S.M., Tao S., Whiteside D., Goulet G.C., Gardner E.C., Bedi A., and Perkins N.C. Accuracy of femur angles estimated by IMUs during clinical procedures used to diagnose femoroacetabular impingement. *IEEE Transactions on Biomedical Engineering*, 62(6):1503–1513, 2015.

[77] Sarvenaz Salehi, Gabriele Bleser, Attila Reiss, and Didier Stricker. Body-IMU autocalibration for inertial hip and knee joint tracking. *Proceedings of the 10th EAI International Conference on Body Area Networks (Bodynets-2015)*, pages 51–57, 2015.

[78] Thomas Seel, Jörg Raisch, and Thomas Schauer. IMU-based joint angle measurement for gait analysis. *Sensors*, 14(4):6891–6909, 2014.

[79] Holly Jacinda Silvers and Bert R Mandelbaum. Prevention of anterior cruciate ligament injury in the female athlete. *British journal of sports medicine*, 41(Suppl 1):i52–i59, 2007.

[80] Andrew J Skalsky and Craig M McDonald. Prevention and management of limb contractures in neuromuscular diseases. *Physical medicine and rehabilitation clinics of North America*, 23(3):675–687, 2012.

[81] S. Šlajpah, R. Kamnik, and M. Munih. Kinematics based sensory fusion for wearable motion assessment in human walking. *Computer Methods and Programs in Biomedicine*, 116(2):131–144, 2014.

[82] Richard B Souza and Christopher M Powers. Differences in hip kinematics, muscle strength, and muscle activation between subjects with and without patellofemoral pain. *The Journal of orthopaedic and sports physical therapy*, 39(1):12–19, 2009.

[83] Mark W Spong, Seth Hutchinson, and Mathukumalli Vidyasagar. *Robot modeling and control*, volume 3. Wiley New York, 2006.

[84] Mikel R Stiffler, Anthony P Pennuto, Mason D Smith, Matt E Olson, and David R Bell. Range of Motion, Postural Alignment, and LESS Score Differences of Those With and Without Excessive Medial Knee Displacement. *Clinical journal of sport medicine*, 25(1):61–66, 2014.

[85] D. Trojaniello, S. Cao, A. Cereatti, and U. Della Croce. Single IMU gait event detection methods: Sensitivity to imu positioning. *Gait & Posture*, 37:S24, 2013.

[86] Viviane Ugalde, Chuck Brockman, Zach Bailowitz, and Christine D. Pollard. Single Leg Squat Test and Its Relationship to Dynamic KneeValgus and Injury Risk Screening. *Physical Medicine and Rehabilitation*, 7(3):229–235, 2015.

[87] JJ Uicker, J Denavit, and RS Hartenberg. An iterative method for the displacement analysis of spatial mechanisms. *Journal of Applied Mechanics*, 31(2):309–314, 1964.

[88] Markus Waldén, Martin Hägglund, Henrik Magnusson, and Jan Ekstrand. ACL injuries in men's professional football: a 15-year prospective study on time trends and return-to-play rates reveals only 65% of players still play at the top level 3 years after ACL rupture. *British Journal of Sports Medicine*, 50:744–750, 2016.

[89] Qi Wang, Wei Chen, and Panos Markopoulos. Literature review on wearable systems in upper extremity rehabilitation. In *2014 IEEE-EMBS International Conference on Biomedical and Health Informatics, BHI 2014*, pages 551–555, 2014.

[90] Benjamin K Weeks, Christopher P Carty, and Sean A Horan. Kinematic Predictors of Single-leg Squat Performance. A Comparison of Experienced Physiotherapists and Student Physiotherapists. *BMC musculoskeletal disorders*, 13:207, 2012.

[91] Benjamin K. Weeks, Christopher P. Carty, and Sean a. Horan. Effect of sex and fatigue on single leg squat kinematics in healthy young adults. *BMC Musculoskeletal Disorders*, 16(1):271, 2015.

[92] Kaitlyn Weiss and Chris Whatman. Biomechanics Associated with Patellofemoral Pain and ACL Injuries in Sports. *Sports Medicine*, 45(9):1325–1337, 2015.

[93] Greg Welch and Gary Bishop. An introduction to the kalman filter. 1995.

[94] Darragh Whelan, Martin O Reilly, Tomás Ward, and Brian Caulfield. Evaluating Performance of the Single Leg Squat Exercise with a Single Inertial Measurement Unit. In *Workshop on ICTs for improving Patients Rehabilitation Research Techniques*, pages 144–147, 2015.

[95] David A Winter. *Biomechanics and motor control of human movement*. John Wiley & Sons, 2009.

[96] Narelle Wyndow, Amy De Jong, Krystal Rial, Kylie Tucker, Natalie Collins, Bill Vicenzino, Trevor Russell, and Kay Crossley. The relationship of foot and ankle mobility to the frontal plane projection angle in asymptomatic adults. *Journal of foot and ankle research*, 9:3, 2016.

[97] James W Youdas, Connie L Bogard, and Vera Jean Suman. Reliability of goniometric measurements and visual estimates of ankle joint active range of motion obtained in a clinical setting. *Archives of physical medicine and rehabilitation*, 74(10):1113–1118, 1993.

[98] Brian L Zeller, Jean L McCrory, W Ben Kibler, and Timothy L Uhl. Differences in kinematics and electromyographic activity between men and women during the single-legged squat. *The American journal of sports medicine*, 31(3):449–56, 2003.

[99] Zheng Zhao, Fred Morstatter, Shashvata Sharma, Salem Alelyani, Aneeth Anand, and Huan Liu. Advancing Feature Selection Research. *ASU Feature Selection Repository Arizona State University*, pages 1 – 28, 2010.

APPENDICES

Appendix A

Pose Estimation Results of all Subjects

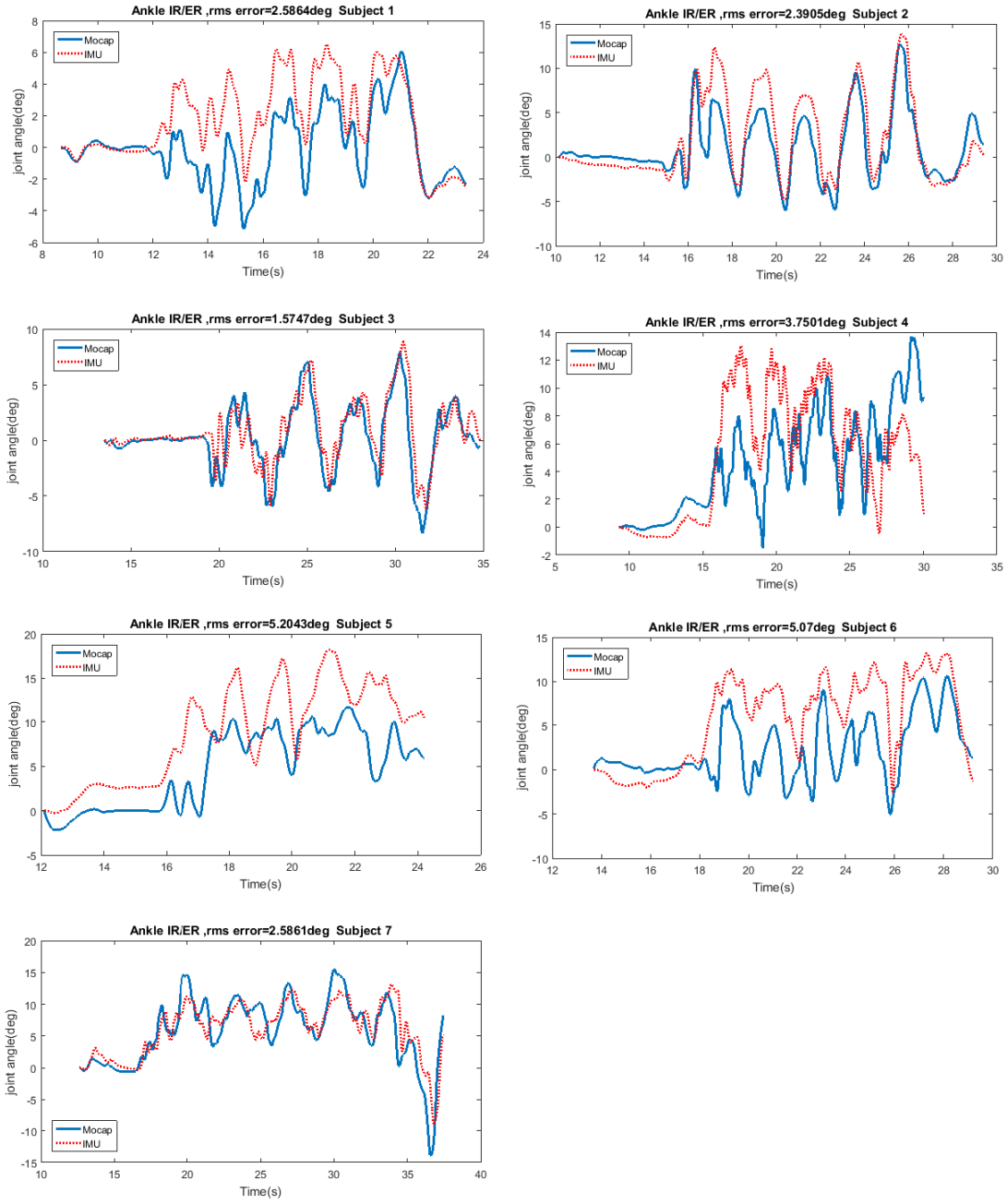


Figure A.1: Estimated ankle IR/ER joint angles related to Table 3.1

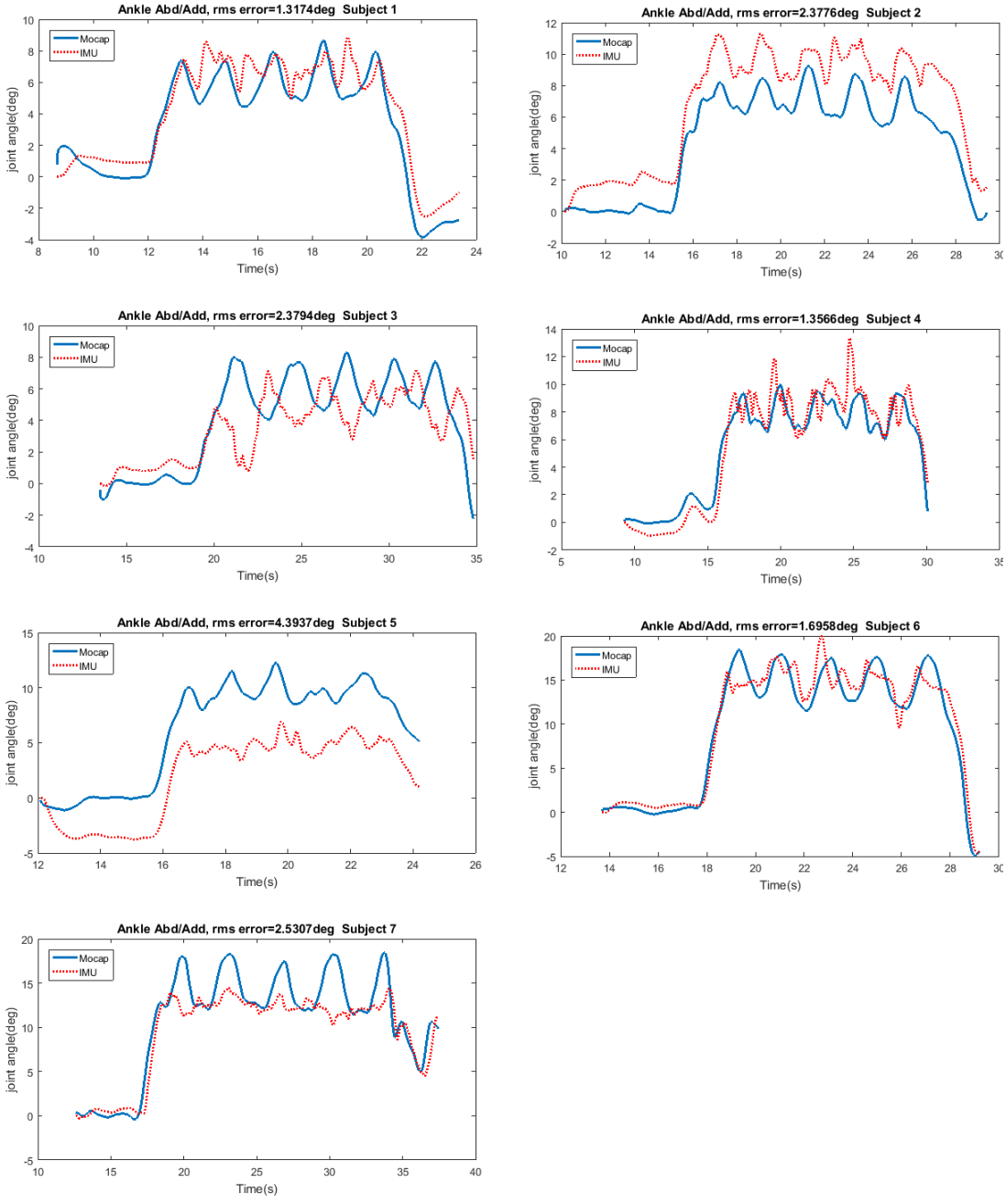


Figure A.2: Estimated ankle Add/Abd joint angles related to Table 3.1

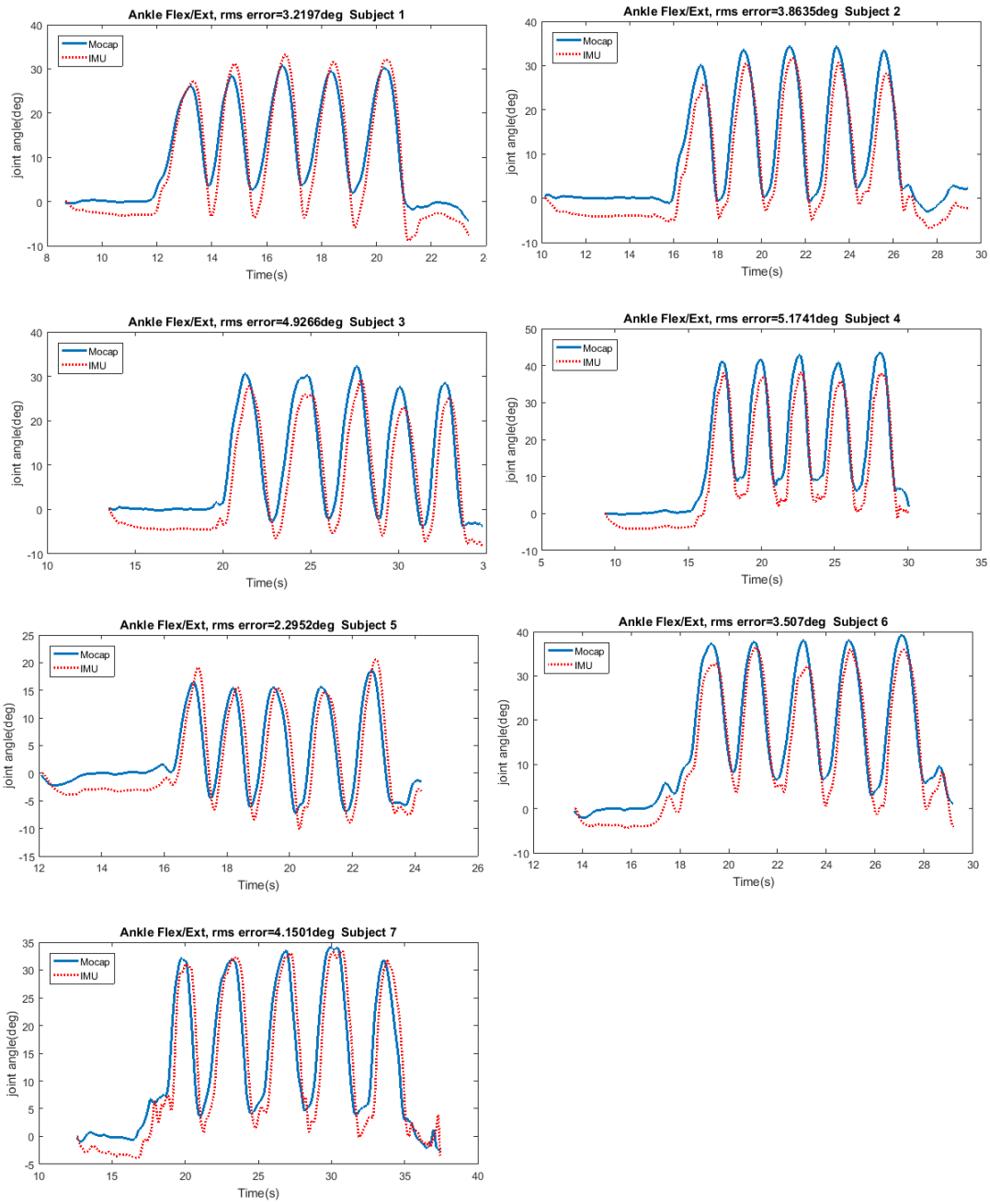


Figure A.3: Estimated ankle Flex/Ext joint angles related to Table 3.1

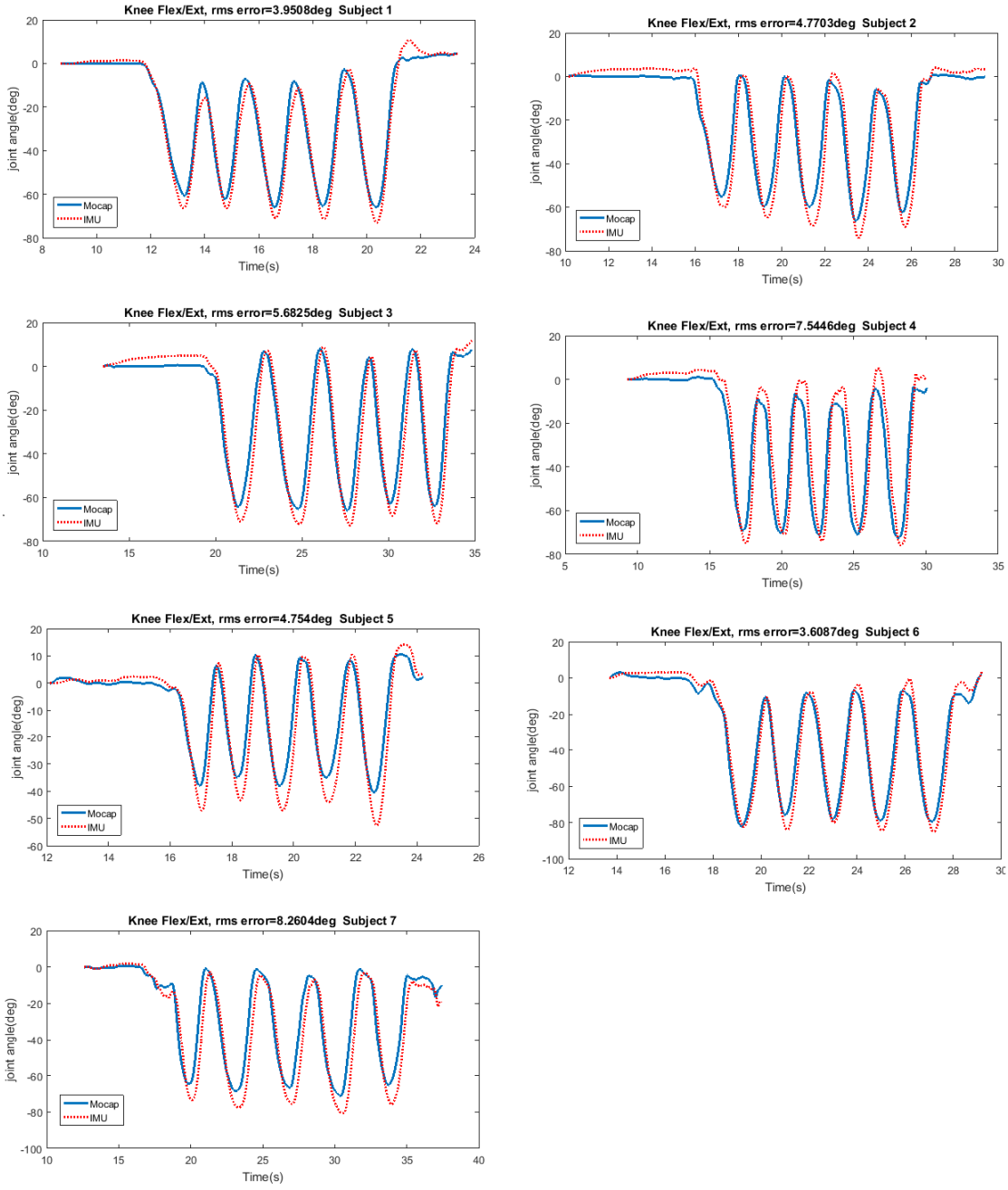


Figure A.4: Estimated knee Flex/Ext joint angles related to Table 3.1

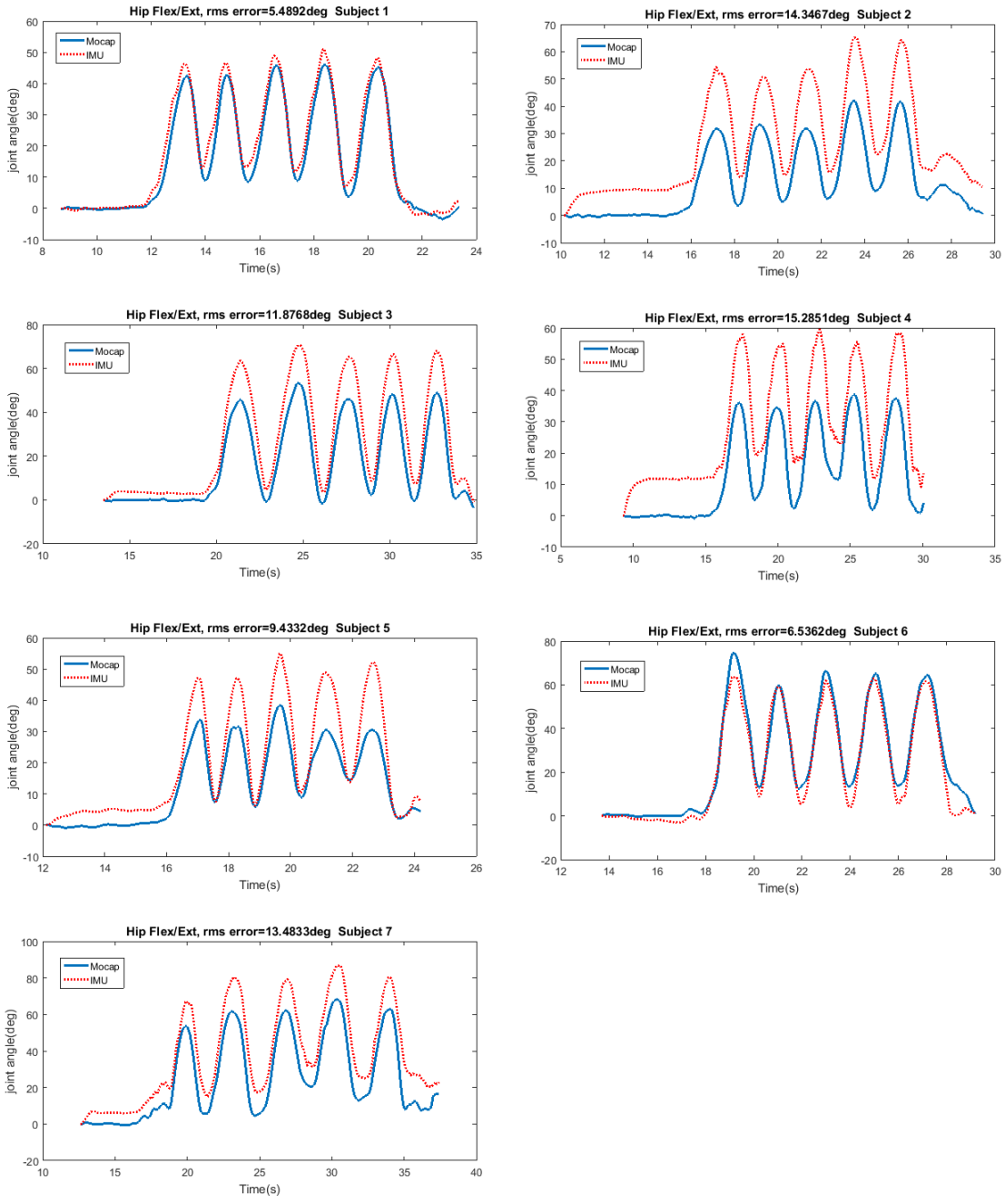


Figure A.5: Estimated hip Flex/Ext joint angles related to Table 3.1

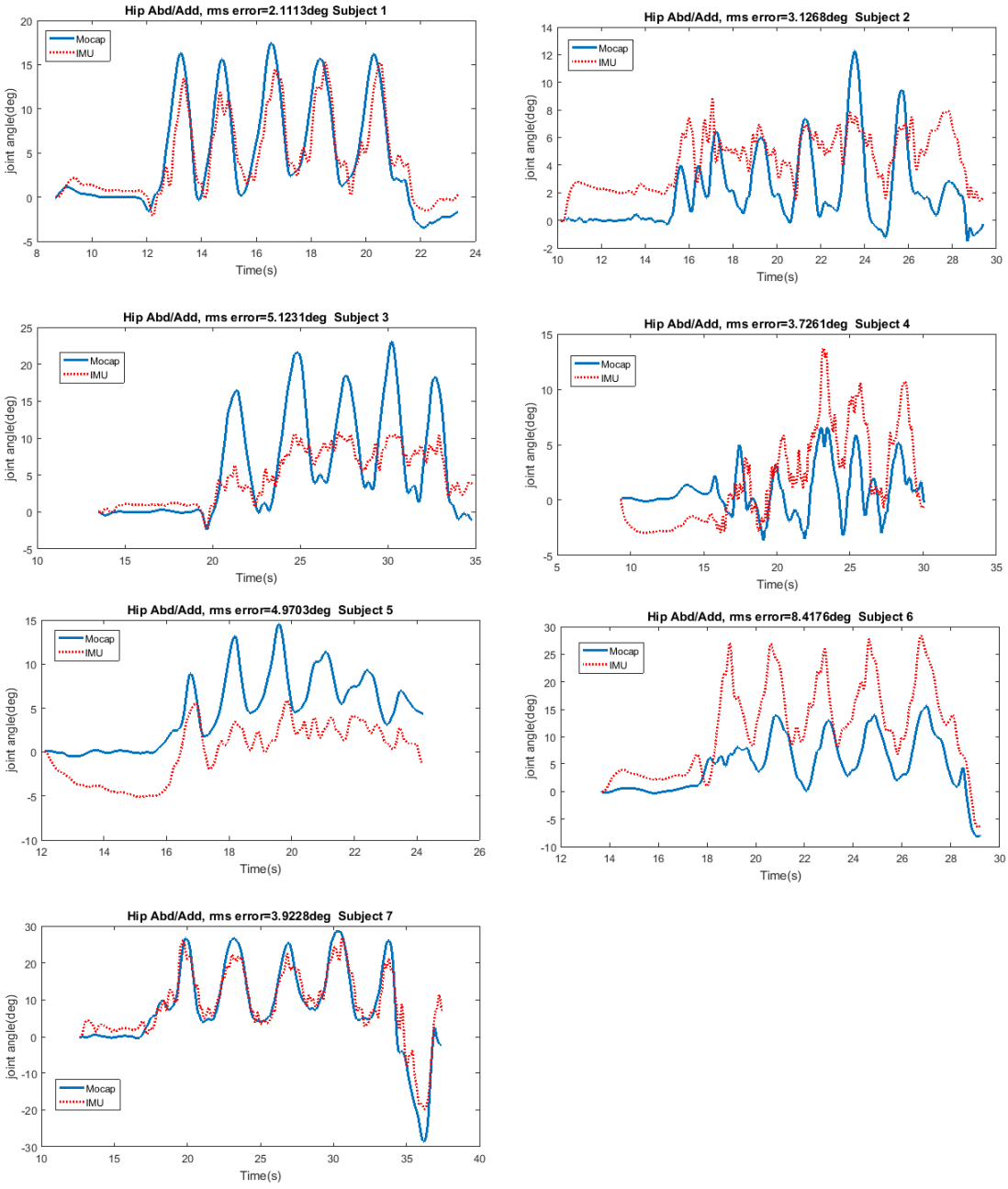


Figure A.6: Estimated hip Add/Abd joint angles related to Table 3.1

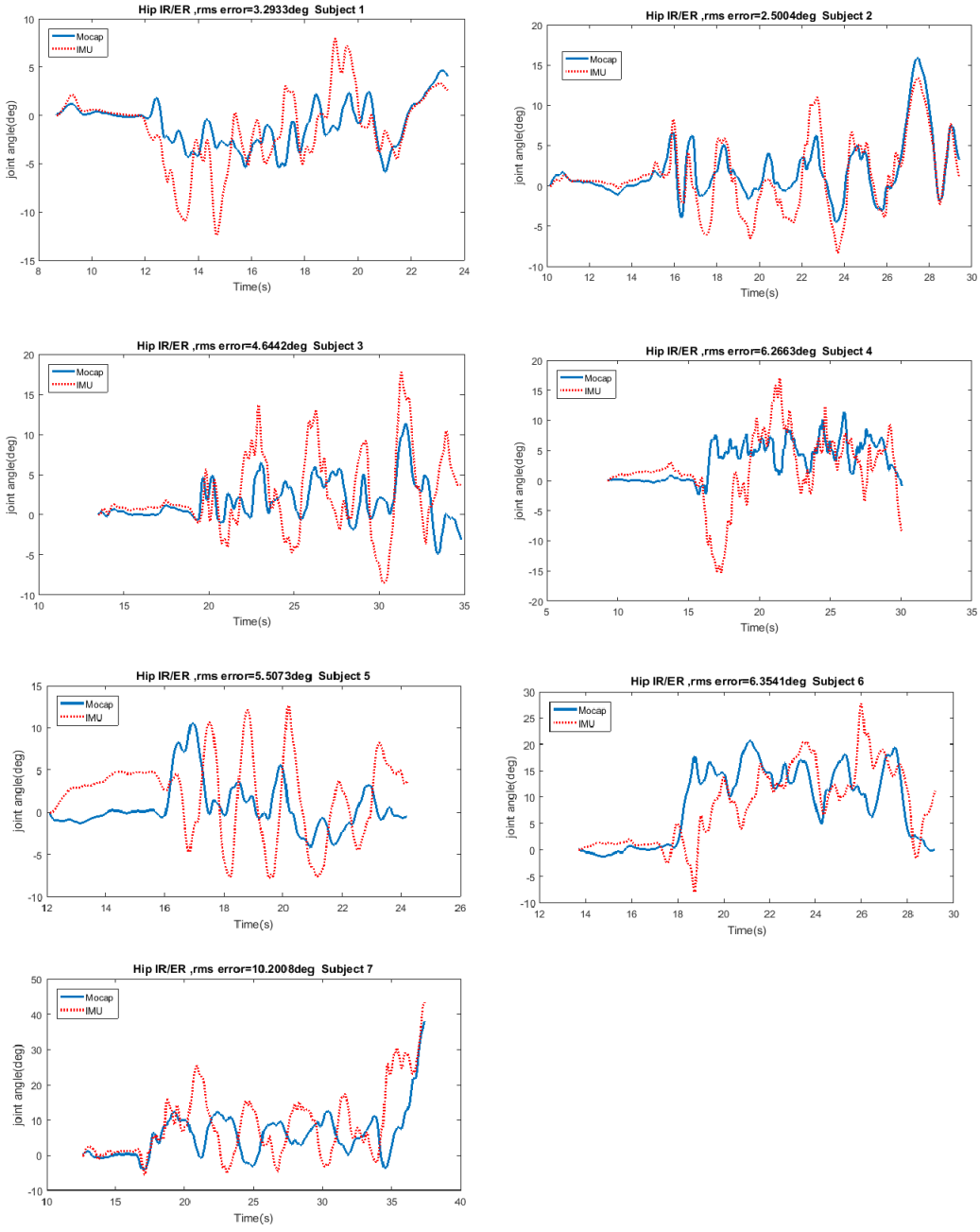


Figure A.7: Estimated hip IR/ER joint angles related to Table 3.1

Appendix B

IRR Test Selection Criteria

Inter rater reliability (IRR) provides an assessment of whether the scale or rating criteria were appropriate and the experience of the raters [35]. Percentage of agreement in the ratings (number of consensus ratings divided by all ratings) is not a proper index of raters agreement because it does not take into account agreements made by chance [35].

There are a number of statistical tests for IRR estimation. Two common tests include Cohonens Kappa test and Intra Class Correlation (ICC) test. Cohonens Kappa test is used for IRR assessment when the data is nominal (categorical) and is suitable for two raters. For more than two raters, the Kappa test can be used for each pair of raters, and the arithmetic mean of all Kappas can be reported as for the overall IRR estimation [35].

ICC is the most common test used for IRR assessment when the ratings are ordinal, interval, or ratio. It can be used for more than two raters and non-fully crossed designs as well [35].

In contrast to the Kappa test which defines IRR based on all or no agreement, ICC takes into account the degree of disagreement such that lower ICC corresponds to the larger disagreement. ICC is calculated by analysis of variance (ANOVA) by modeling each rating as a sum of the true score and the measurement error as follows:

$$X_{ij} = \mu + r_i + e_{ij} \tag{B.1}$$

Where X_{ij} is the rating of subject i by rater j, μ and r_i are the mean and deviation of the true score for X and e_{ij} is the measurement error. The ICC score is calculated by

combination of variance of Equation B.1 components. The combination formula depends on the study design [35].

An ICC value of 1 corresponds to perfect agreement and -1 corresponds to perfect disagreement, and in case of more than two raters ICC can have values less than -1. Cicchetti's guideline [22] for interpreting ICC values is summarized in Table B.1.

ICC value	*IRR*
$0.75 - 1.00$	*Excellent*
$0.60 - 0.74$	*good*
$0.40 - 0.59$	*fair*
< 0.40	*poor*

Table B.1: Cicchetti(1994) guidelines for interpreting ICC score [22]

Different variants of Cohonen's Kappa and ICC tests are available for different study designs. In order to select the most appropriate variant of the existing statistical tests, the study design should be considered. For example, it should be taken into account if the ratings were nominal, ordinal, interval, or ratios, whether the study is fully crossed or non-fully crossed, and whether it is one-way or two-way [35].

If a subset of raters are selected randomly from a large number of raters to rate each subject then ICC model is one-way, but if fixed number of raters have rated all subjects (fully-crossed), two-way model should be used [35].

Moreover, it has to be specified whether absolute agreement between raters is of interest or similarity in the rank ordering (consistency of rating) is enough. If one rater generates lower ratings than another in general, absolute agreement between them would be low but consistency of the ratings could still be high if rank order is the same [35].

Two different units of analysis are also available to apply in ICC. If multiple raters have rated all subjects then reliability should be quantified based on average of the ratings. However, if multiple raters have rated a subset of subjects and IRR has to be generalized to other subjects, a single-measure ICC is more appropriate [35].

Finally, if a subset of coders are randomly selected to rate each subject and their IRR will be generalized to other coders then the ICC model is random- effect whereas if raters are not randomly sampled then ICC is mixed-effect [35].