

Adaptive Extended Kalman Filtering for IMU-based Human Relative Orientation Estimation

Valentijn van der Hoeden

Master of Science Thesis

Adaptive Extended Kalman Filtering for IMU-based Human Relative Orientation Estimation

MASTER OF SCIENCE THESIS

For the degree of Master of Science in Systems and Control at Delft
University of Technology

Valentijn van der Hoeden

October 15, 2026

Faculty of Mechanical Engineering (ME) · Delft University of Technology



Abstract

Inertial measurement units (IMUs) offer a portable alternative to optical motion capture, but the relative orientation of connected body segments is difficult to estimate when the quality of the measurements changes over time. This thesis investigates how online adaptation of the covariance of the joint-centre acceleration update in a drift-free relative-orientation Extended Kalman Filter affects the estimation accuracy, and compares it with rejecting inconsistent updates based on the Normalized Innovation Squared (NIS). Adaptive scaling and NIS rejection are each applied to a magnetometer-free estimator and to an extension that conditionally uses magnetic information (MAJIC).

Monte Carlo simulations, in which one source of uncertainty is varied at a time, show that both mechanisms respond to every source, but lower the error only when the acceleration update remains informative and is systematically inconsistent with its model. A fixed covariance with a larger value reproduces most of this benefit.

On lower-body human motion recordings, tuning the acceleration covariance of the baseline lowered its mean overall angular RMSE from 15.5° to 9.9° , and the magnetic update lowered it to 5.9° . NIS rejection gave a small but consistent further improvement of 0.4° and 0.8° , which made MAJIC with NIS rejection the best variant (5.1°). With the literature settings, adaptive scaling lowered the post-convergence error by about 2.4° for both estimators, and for MAJIC it increased the convergence success from 73% to 89%, without any tuning of the covariance. Relative to a baseline that is tuned on the same recordings, however, adaptive scaling increased the overall error (to 27.4° and 9.7°).

Adaptive scaling is therefore a self-adjusting alternative when the covariance cannot be tuned, but with a tuned baseline it does not improve relative-orientation estimation on recorded motion, whereas a statistically gated acceleration update and the conditional magnetic update do.

Table of Contents

Preface	vii
1 Introduction	1
1-1 Motivation	1
1-2 Problem Statement and Research Gap	3
1-3 Research Objective	4
1-4 Contributions	5
1-5 Thesis Outline	6
2 Background and Related Work	7
2-1 Coordinate Frames and Orientation Representation	7
2-2 Inertial Measurement Units	8
2-2-1 Gyroscope	8
2-2-2 Accelerometer	8
2-2-3 Magnetometer	9
2-3 Orientation Estimation from Inertial Measurements	9
2-4 Connected-Segment Kinematics	9
2-4-1 Rigid-Body Acceleration	10
2-4-2 Relative-Orientation Information from a Common Joint Centre	10
2-5 Related Work on Relative-Orientation Estimation	10
2-5-1 Drift-Free Relative Orientation Using Joint-Centre Acceleration	10
2-5-2 Observability of Relative Orientation	11
2-5-3 Conditional Magnetic Information	12
2-6 Extended Kalman Filtering and Measurement Consistency	12
2-6-1 Normalized Innovation Squared and Measurement Rejection	13
2-6-2 Adaptive Covariance Estimation	13

3	Methods for IMU-Based Relative-Orientation Estimation	15
3-1	Relative-Orientation Estimation Problem	15
3-2	Magnetometer-Free Relative-Orientation Estimator	17
3-2-1	Orientation Deviation State	17
3-2-2	Gyroscope-Driven Propagation	17
3-2-3	Joint-Centre Acceleration Measurement Model	18
3-2-4	Deviation-State EKF Formulation	19
3-3	Conditional Magnetic-Direction Update	20
3-4	Innovation-Based Adaptive Acceleration Covariance	21
3-4-1	Empirical Innovation Statistics	21
3-4-2	Covariance Matching	22
3-5	NIS-Based Acceleration Measurement Rejection	22
3-6	Estimator Variants	23
3-7	Unified Estimation Procedure	24
3-8	Estimator Configuration and Design Choices	25
3-8-1	Estimator Initialization	25
3-8-2	Process Covariance	25
3-8-3	Nominal Acceleration Covariance	25
3-8-4	Adaptive-Covariance Parameters	26
3-8-5	NIS Threshold	26
3-8-6	MAJIC Configuration	26
3-8-7	Tuned Configuration	27
3-8-8	Numerical Differentiation	27
4	Simulation Study	29
4-1	Simulation Framework	30
4-1-1	General Setup	30
4-1-2	Synthetic Motion and Sensor Measurements	31
4-1-3	Simulation Parameters	32
4-1-4	Estimator Settings in the Simulations	32
4-2	Performance Evaluation	33
4-2-1	Convergence	33
4-2-2	Orientation Error After Convergence	33
4-2-3	Diagnostics of the Reliability-Handling Mechanisms	34
4-3	Influence of Motion Excitation	34
4-4	Sensor Errors: Accelerometer Noise and Gyroscope Bias	36
4-5	Joint-Centre Geometry Mismatch	38
4-6	Sustained Acceleration-Model Disturbance	39
4-7	Short Accelerometer Outliers	40
4-8	Inconsistent Magnetic Field	42
4-9	Control Experiment with a Tuned Fixed Covariance	44
4-9-1	Results	45
4-9-2	Error over Time	47
4-9-3	Dependence on the Error Measure	49
4-9-4	Interpretation	49
4-10	Conclusions from the Simulation Study	50

5	Evaluation on Recorded Human Motion	53
5-1	Dataset and Processing	53
5-1-1	OpenSense Dataset	53
5-1-2	Data Selection	54
5-1-3	Synchronization and Reference Orientation	54
5-1-4	Constant Sensor–Marker Frame Alignment	55
5-2	Joint-Centre Geometry	56
5-2-1	Inertially Estimated Geometry	56
5-2-2	Marker-Derived Sensor-to-Joint-Centre Geometry	56
5-3	Estimator Tuning and Evaluation	57
5-3-1	Tuning of the Estimators	57
5-3-2	Performance Measures	58
5-4	Results	59
5-4-1	Primary Estimator Comparison	59
5-4-2	Paired Comparison per Recording	60
5-4-3	Joints and Motion Types	62
5-4-4	Behaviour of the Adaptive Factor	62
5-4-5	Error over Time	63
5-4-6	Effect of the Tuning	64
5-4-7	Joint-Centre Geometry	65
5-4-8	Behaviour with the Literature Settings	66
5-5	Discussion	67
5-6	Summary of the Recorded-Data Evaluation	68
6	Discussion and Conclusions	69
6-1	Discussion	69
6-1-1	What the Mechanisms React to and When This Helps	69
6-1-2	Tuning and Magnetic Information	70
6-1-3	Limitations	70
6-2	Conclusions	71
6-3	Recommendations and Future Work	72
A	Appendix	75
A-1	Stationary Smartphone IMU Measurements	75
A-2	Motion-Excitation Parameters	75
A-3	Tuning of the Estimator Parameters	75
A-3-1	Procedure	76
A-3-2	Acceleration Covariance of the Baseline	77
A-3-3	Magnetic Covariance of MAJIC	77
A-3-4	NIS Thresholds	78
A-3-5	Adaptive Covariance Scaling	80
A-4	Sensitivity of the Convergence Criterion	81

Preface

This thesis marks the completion of my MSc in Systems and Control at Delft University of Technology. Over the course of this project, I have had the opportunity to work on a topic that brought together several subjects that I particularly enjoyed during my studies, including estimation, sensor fusion and human-motion analysis.

I would first like to thank my main supervisor, Dr.ir. Manon Kok, for her guidance, expertise and feedback throughout this thesis. Her suggestions repeatedly helped me to look more carefully at both the technical details and the broader interpretation of the work.

I would also like to thank my daily supervisor, Ruiyuan Li, for her continued support throughout the project. Our regular discussions were extremely helpful, both when working through technical problems and when deciding how to structure the thesis as a whole.

Finally, I would like to thank my family, friends and everyone around me for their support and encouragement throughout my studies and during the completion of this thesis.

Chapter 1

Introduction

1-1 Motivation

Human motion analysis plays an important role in clinical, scientific and engineering applications. Quantitative measurements of movement are used in, among others, gait analysis, rehabilitation, sports performance assessment and wearable monitoring [20, 11, 15]. Across these applications, accurate estimation of the motion of individual body segments and, in particular, their relative orientation is required to characterize movement between adjacent segments.

Marker-based optical motion capture (OMC) is widely used as a reference method for biomechanical motion analysis. Markers attached to the body are tracked by multiple calibrated cameras, allowing their three-dimensional positions and orientations to be reconstructed with high accuracy [6, 23]. However, optical systems require dedicated measurement environments, extensive setup and calibration, and an unobstructed line of sight between cameras and markers. These requirements limit their applicability outside the laboratory and make long-duration or unconstrained measurements more difficult.

Wearable inertial measurement units (IMUs) provide an attractive alternative. Modern IMUs are compact and can be attached directly to body segments. They typically contain a three-axis gyroscope and accelerometer, and often also a three-axis magnetometer, which measure angular velocity, specific force and the local magnetic field, respectively [9]. Their portability and relatively low cost have therefore motivated widespread use in wearable human-motion estimation [2, 11].

The use of IMUs introduces a more challenging estimation problem because orientation is not measured directly. Integrating gyroscope measurements gives accurate short-term orientation changes, but gyroscope noise and bias accumulate through integration and lead to orientation drift [9]. Accelerometer measurements can provide complementary orientation information, for example through the direction of gravity under sufficiently low non-gravitational acceleration, but this assumption is often violated during dynamic human motion [22, 13].

Magnetometers can additionally provide heading information, but nearby ferromagnetic materials and electronic equipment may distort the local magnetic field and introduce substantial orientation errors [4, 18, 10].

For many biomechanical applications, the quantity of interest is not the absolute orientation of an individual sensor, but the relative orientation between sensors attached to adjacent body segments. Because these segments are physically connected through a joint, their inertial measurements are related by the underlying rigid-body kinematics. This relationship provides additional information that can be exploited without requiring an absolute orientation reference [5, 12, 19].

A particularly relevant method was proposed in [25]. Their Extended Kalman Filter (EKF) calculates the acceleration of the common joint centre independently from two IMUs attached to connected segments. Since both calculations describe the acceleration of the same physical point, their difference provides information about the relative sensor orientation. This joint-centre acceleration measurement model allows relative drift to be corrected without magnetometer measurements and therefore provides the principal estimation framework used in this thesis.

The magnetometer-free framework in [25] was extended in [21] by conditionally incorporating magnetic-field direction information when acceleration-based observability is low, for example, during prolonged standing or other nearly stationary human motion [8]. This extension retains the same joint-centre acceleration measurement model and provides an additional source of orientation information during periods of poor acceleration-based observability. Comparing the proposed and alternative method in both estimators therefore makes it possible to determine whether these methods are only beneficial when the estimator depends strongly on the acceleration update, or whether they remain useful when another correction source is available.

The information available from this measurement model is nevertheless motion-dependent. In [8] it was shown that magnetometer-free relative orientation becomes observable when the joint-centre specific force and its temporal variation provide sufficiently independent directional information. During poorly exciting motion, relative orientation may therefore be temporarily unobservable, whereas richer motion provides more information for correcting relative orientation drift.

A related but distinct issue is that the statistical reliability of the joint-centre acceleration measurement model may also vary. Sensor noise, gyroscope bias, inaccuracies in the assumed IMU-to-joint-centre geometry and violations of the rigid-body model can all change the disagreement between the predicted and observed measurements. Conventional EKF implementations represent such uncertainty using covariance matrices selected before estimation [24, 9]. In the estimator from [25], these covariance parameters remain fixed throughout a recording. Consequently, the filter assigns the same nominal statistical confidence to the acceleration update even when the agreement between the measurement model and the observed data changes.

Adaptive covariance estimation provides one possible way of addressing this limitation. Innovation-based adaptive Kalman filtering uses discrepancies between predicted and observed measurements to adjust estimator uncertainty online [26, 27]. Such methods allow the weighting between prediction and measurement information to change when the observed innovation statistics no longer agree with the assumed model.

In this thesis, *adaptive filtering* refers specifically to online adaptation of the covariance associated with the joint-centre acceleration measurement update. Recent acceleration innovations are used to calculate a scalar covariance factor that continuously changes the influence of this update. The method is compared with the original fixed-covariance estimator and with Normalized Innovation Squared (NIS)-based acceleration rejection, which provides a binary alternative in which sufficiently inconsistent acceleration updates are omitted. Both reliability-handling strategies are evaluated in the magnetometer-free estimator in [25] and in the magnetometer-aided estimator in [21].

The resulting six estimator variants are evaluated using controlled Monte Carlo simulations and recorded human-motion data. The simulations isolate individual sources of estimator degradation, while the experimental evaluation assesses the behaviour of the methods under the combined uncertainties present in recorded lower-extremity motion.

1-2 Problem Statement and Research Gap

The joint-centre acceleration model provides a mechanism for correcting relative-orientation drift without relying (continuously) on magnetometer measurements. Its effectiveness, however, depends not only on the information contained in the performed motion but also on the validity of the measurement model and its assumed uncertainty.

These two effects should be distinguished. Reduced motion excitation limits the information available to the estimator and can make particular relative orientation directions unobservable [8]. In contrast, sensor errors or model mismatch can cause the joint-centre acceleration update to become statistically inconsistent even when the performed motion remains informative. Examples include increased accelerometer noise, uncompensated gyroscope bias, inaccurate IMU-to-joint-centre geometry and unmodelled acceleration disturbances.

The fixed acceleration measurement covariance used in the baseline EKF cannot distinguish between such changing conditions. Adaptive covariance scaling can modify the influence of the acceleration update when its recent innovation statistics change, but a large innovation does not uniquely identify its cause. It may result from an unreliable measurement model, but also from an inaccurate state estimate that requires correction. Downweighting the measurement in these two situations does not necessarily have the same effect.

NIS-based rejection introduces a related trade-off. Rejecting an isolated strongly inconsistent update can protect the estimator from an outlier, whereas rejecting measurements persistently may remove the information required to correct orientation drift. The usefulness of either reliability-handling strategy can therefore be expected to depend on the source and duration of the observed inconsistency.

Improving relative-orientation estimation accuracy is also relevant from an application perspective because relative sensor orientations are commonly used to derive biomechanical joint angles. In [14], it was reported that joint-angle errors of several degrees can already affect clinical interpretation, with errors above approximately 5° potentially becoming problematic. These thresholds apply to errors in individual joint-angle components and should not be interpreted directly as thresholds for the three-dimensional relative-orientation error considered in this thesis. A given relative-orientation error can affect the resulting joint-angle components differently depending on its direction, the joint configuration and the sensor-to-segment

calibration. In [25], the filtering implementation resulted in average knee joint-angle RMSE values between 2.42° and 4.47° , with maximum values up to 5.38° . This shows that errors in the underlying relative-orientation estimate can still have a meaningful effect on the final joint-angle estimates, motivating further improvement of relative-orientation estimation.

In [25], the proposed method was validated on a robotic manipulator, demonstrating that arbitrary three-dimensional relative orientation can be estimated from inertial measurements. Its validation on human motion, however, was limited to prolonged walking trials evaluating knee-joint kinematics. Consequently, the performance of the method for other lower-limb joints and during non-walking movements remains less well established.

The resulting research gap is that, although the reliability of the joint-centre acceleration measurement model can vary over time, existing relative-orientation estimators based on this model do not explicitly account for such changes in the measurement update. At the same time, it has not been established whether online reliability handling, such as adaptive covariance scaling or innovation-based measurement rejection, improves relative-orientation estimation when applied to this measurement model.

The central problem addressed in this thesis is therefore to determine how online reliability handling of the joint-centre acceleration update affects relative-orientation estimation for various connected body-segments, and under which conditions continuous covariance adaptation or binary measurement rejection is beneficial.

1-3 Research Objective

The objective of this thesis is to investigate whether online adaptation of the joint-centre acceleration measurement covariance can improve the accuracy and reliability of IMU-based relative-orientation estimation for connected human segments. In particular, the aim is to determine whether reliability-aware handling of the acceleration update can reduce the remaining relative-orientation estimation errors of the baseline filtering approach, when the baseline and the adaptive method are tuned in the same way.

In this context, *accuracy* refers to the angular error of the estimated relative orientation over a complete recording, whereas *estimator reliability* refers to the occurrence of large errors when measurement disturbances or model errors are present. Both are evaluated with the overall angular RMSE over a complete recording and its distribution over the recordings, because an error that is evaluated only after convergence excludes the recordings in which an estimator fails.

The proposed adaptive method calculates a scalar covariance factor from recent acceleration innovation statistics and applies this factor to the nominal joint-centre acceleration measurement covariance. Its behaviour is compared with fixed-covariance estimation and with NIS-based acceleration rejection. The same acceleration reliability-handling mechanisms are evaluated within both the baseline estimator frameworks, while the magnetic update itself remains unchanged.

The main research question is formulated as:

How does online adaptation of the joint-centre acceleration measurement covariance affect the accuracy and reliability of IMU-based relative-orientation estimation?

tion, under which conditions is it beneficial, and how does its performance compare with NIS-based measurement rejection?

This question is addressed through the following subquestions:

1. How do adaptive acceleration covariance scaling and NIS-based acceleration rejection affect relative orientation estimation accuracy, convergence success and convergence time under controlled changes in motion excitation and sensor uncertainty?
2. Can adaptive acceleration covariance scaling and NIS-based acceleration rejection reduce estimation degradation caused by errors in the assumed IMU-to-joint-centre geometry?
3. Do adaptive acceleration covariance scaling and NIS-based acceleration rejection improve the accuracy of relative orientation in recorded human motion, including walking and non-walking activities, when all estimators are tuned on the same recordings?

The first two subquestions are addressed using controlled Monte Carlo simulations, in which motion excitation, sensor uncertainty and joint-centre geometry mismatch are varied independently. The third subquestion is addressed using recorded lower-extremity human-motion data from [1], where the estimator variants are compared in terms of the overall relative-orientation angular RMSE.

1-4 Contributions

The main contributions of this thesis are:

- The implementation of the proposed innovation-based adaptive covariance method for the joint-centre acceleration update of the drift-free relative-orientation EKF in [25] and the magnetometer-aided extension in [21].
- A systematic comparison of fixed covariance, continuous adaptive covariance scaling and binary NIS-based rejection for handling inconsistency in the joint-centre acceleration update.
- A controlled Monte Carlo simulation study that isolates the effects of motion excitation, accelerometer measurement-noise mismatch, gyroscope bias, joint-centre geometry mismatch, sustained acceleration-model disturbance, magnetic-direction inconsistency and short accelerometer outliers.
- A tuning of all estimator variants on the recorded data, so that the reliability-handling methods are compared with baselines that are tuned in the same way (Appendix A-3).
- An extended evaluation of relative-orientation estimation on recorded lower-extremity human motion using the *OpenSense* [1] dataset. In contrast to the human validation in [25], which considered knee-joint kinematics during prolonged walking, the evaluation includes hip, knee and ankle joints and both walking and non-walking lower-body movements.

Main findings. In the controlled simulations, both mechanisms respond to every source of uncertainty, but they lower the error only when the acceleration update remains informative and is systematically inconsistent with its model, and a fixed covariance with a larger value reproduces most of this benefit. On recorded human lower-body motion, adaptive covariance scaling lowers the post-convergence error with the literature settings, making it a self-adjusting option when the covariance cannot be tuned. With all estimators tuned on the same data, however, it does not improve the overall accuracy, NIS-based rejection gives a small but consistent improvement, and the conditional magnetic update is the largest single difference between the estimators (Chapters 4 to 6).

1-5 Thesis Outline

The remainder of this thesis is organized as follows.

Chapter 2 defines the estimation problem, coordinate frames and sensor models, and introduces the joint-centre acceleration measurement. It reviews the magnetometer-free estimator of [25], the motion-dependent observability of the measurement, and the estimator of [21] that conditionally incorporates magnetic information. Statistical measurement rejection and adaptive covariance estimation are reviewed at the end.

Chapter 3 derives the proposed innovation-based adaptive covariance scaling and the NIS-based rejection method as a binary alternative. Both are applied to the two existing estimators, and the resulting estimator variants and common design choices are defined.

Chapter 4 presents controlled simulation studies that isolate individual factors, such as motion excitation, sensor noise and bias, joint-centre geometry mismatch, acceleration disturbances and magnetic-direction mismatch. They show what the two mechanisms react to and when this helps or hurts. A control experiment with a tuned fixed covariance then tests whether the benefit of adaptive scaling survives a stronger baseline.

Chapter 5 evaluates the variants on recorded lower-extremity motion from [1]. It describes the recordings, the preprocessing and reference orientation, the tuning, and it analyses the orientation error by motion type, joint and joint-centre geometry in relation to the simulation findings.

Chapter 6 discusses the limitations, answers the research question and subquestions, and gives recommendations for future work.

Background and Related Work

This chapter introduces the theoretical background and previous work relevant to IMU-based relative-orientation estimation. First, the coordinate frames and orientation representations are defined. The measurements provided by inertial measurement units are then introduced, followed by the principles of orientation estimation from inertial measurements and the associated drift problem. Subsequently, rigid-body kinematics and the use of the physical connection between adjacent segments are discussed as a source of relative-orientation information. Existing approaches for magnetometer-free relative-orientation estimation, their motion-dependent observability and the conditional use of magnetic information are then reviewed. Finally, statistical measurement consistency, innovation-based measurement rejection and adaptive covariance estimation are introduced.

2-1 Coordinate Frames and Orientation Representation

Two types of coordinate frames are required to describe the motion of inertial sensors. The global frame G is an inertial reference frame fixed to the environment, while each sensor frame S_i is rigidly attached to the corresponding IMU, with its axes aligned with the physical sensor axes [9]. For a pair of sensors attached to two adjacent body segments, the sensor frames are denoted by S_1 and S_2 .

The orientation of sensor frame S_i with respect to the global frame G is represented by the rotation matrix

$$R^{GS_i} \in SO(3). \quad (2-1)$$

A vector v expressed in the sensor frame is transformed to the global frame according to

$$v^G = R^{GS_i} v^{S_i}, \quad (2-2)$$

while the inverse transformation is given by

$$R^{S_i G} = \left(R^{GS_i}\right)^\top. \quad (2-3)$$

Orientations can equivalently be represented using unit quaternions. A quaternion is written as

$$q = \begin{bmatrix} q_0 & q_1 & q_2 & q_3 \end{bmatrix}^\top, \quad \|q\|_2 = 1, \quad (2-4)$$

where q_0 denotes the scalar component. The quaternion corresponding to R^{GS_i} is denoted by q^{GS_i} , and quaternion multiplication is denoted by $\odot$.

The relative orientation of sensor frame S_2 with respect to sensor frame S_1 is independent of their common orientation with respect to the global frame and can be written as

$$R^{S_1 S_2} = \left(R^{GS_1}\right)^\top R^{GS_2}. \quad (2-5)$$

Equivalently, in quaternion form,

$$q^{S_1 S_2} = \left(q^{GS_1}\right)^{-1} \odot q^{GS_2}. \quad (2-6)$$

The relative orientation therefore describes the rotational relationship between the two sensors without requiring knowledge of their absolute orientation with respect to G . When orientations vary over time, a subscript t is added, as in $R_t^{GS_i}$ and $q_t^{S_1 S_2}$.

2-2 Inertial Measurement Units

An IMU typically contains a three-axis gyroscope and accelerometer and may additionally contain a three-axis magnetometer. These sensors measure different physical quantities in the local sensor frame and are affected by measurement noise, bias and other disturbances [9].

2-2-1 Gyroscope

A gyroscope measures angular velocity expressed in the sensor frame. A commonly used measurement model is

$$y_{\omega,t}^{S_i} = \omega_t^{S_i} + \delta_{\omega,t}^{S_i} + e_{\omega,t}^{S_i}, \quad (2-7)$$

where $\omega_t^{S_i}$ is the true angular velocity, $\delta_{\omega,t}^{S_i}$ denotes the gyroscope bias and $e_{\omega,t}^{S_i}$ denotes zero-mean measurement noise [9]. The measurement noise is commonly modelled as $e_{\omega,t}^{S_i} \sim \mathcal{N}(0, \Sigma_\omega)$, where Σ_ω denotes the gyroscope measurement-noise covariance.

2-2-2 Accelerometer

An accelerometer measures specific force. Its measurement can be modelled as

$$y_{a,t}^{S_i} = R_t^{S_i G} \left(a_t^G - g^G\right) + \delta_{a,t}^{S_i} + e_{a,t}^{S_i}, \quad (2-8)$$

where a_t^G is the translational acceleration of the sensor origin, g^G is the gravity vector, $\delta_{a,t}^{S_i}$ denotes the accelerometer bias and $e_{a,t}^{S_i}$ denotes measurement noise. The latter is commonly modelled as $e_{a,t}^{S_i} \sim \mathcal{N}(0, \Sigma_a)$, with accelerometer measurement-noise covariance Σ_a [9].

2-2-3 Magnetometer

A magnetometer measures the local magnetic-field vector expressed in the sensor frame. Its measurement model is

$$y_{m,t}^{S_i} = R_t^{S_i G} m_t^G + e_{m,t}^{S_i}, \quad (2-9)$$

where m_t^G is the local magnetic-field vector expressed in G and $e_{m,t}^{S_i}$ represents measurement uncertainty [9].

2-3 Orientation Estimation from Inertial Measurements

The angular velocity measured by a gyroscope provides information about the change in sensor orientation. Given an orientation estimate at a previous time instant, successive angular-velocity measurements can be integrated to propagate the orientation forward in time [9]. Gyroscope integration can provide accurate short-term orientation changes, but it has two important limitations.

First, integration requires an initial orientation from which to propagate. Errors in this initial orientation are not removed by gyroscope integration alone. Second, the measurement in (2-7) contains both noise and bias. These errors are integrated together with the angular velocity and therefore result in an orientation error that grows over time.

Accelerometers and magnetometers can provide complementary orientation information. Accelerometers can stabilize orientation with respect to gravity when non-gravitational acceleration is sufficiently small, while magnetometers can provide an absolute heading reference when the local magnetic field is reliable. Combining gyroscope propagation with such complementary information is commonly performed using filtering or optimization-based sensor-fusion methods [9].

For magnetometer-free orientation estimation, the heading component remains particularly challenging. Gravity provides no information about rotations around the vertical direction, and residual gyroscope bias therefore causes heading drift. When multiple sensors are attached to physically connected segments, however, their motion is not independent. Information about their physical connection can consequently be exploited to estimate their *relative* orientation even when an absolute heading reference is unavailable.

2-4 Connected-Segment Kinematics

Human body segments can be approximated as rigid bodies connected by joints. This physical connection introduces relationships between the motion measured by IMUs attached to adjacent segments. Such relationships have been used in several inertial-motion-estimation methods to reduce drift or to estimate joint kinematics without relying continuously on magnetometer measurements [5, 19, 12].

2-4-1 Rigid-Body Acceleration

For a rigid body, the acceleration at one point can be related to the acceleration measured at another point using the body's angular velocity, angular acceleration and the fixed displacement between the two points [5, 19]. This principle allows acceleration measured by an IMU to be translated from the sensor location to another point on the same body segment. Because an accelerometer measures specific force, the quantity obtained at another point of the body is also a specific force. Following common usage, it is referred to as the acceleration at that point and, for the joint centre, denoted by a_{jc} in the remainder of this thesis.

For two adjacent body segments, a particularly useful point is their common joint centre. Although IMUs attached to the two segments generally undergo different motion, both segments share the same physical joint centre. Consequently, the acceleration of this point can be calculated independently from the measurements of either IMU. After expressing both quantities in a common coordinate frame, their relationship can provide information about the relative motion of the connected segments [5, 19, 25].

This approach requires knowledge of the position of the joint centre relative to each sensor. The vector from the origin of sensor S_i to the joint centre, expressed in the corresponding sensor frame, is denoted by $r_i^{S_i}$. Methods for estimating such sensor-to-joint-centre vectors from inertial measurements have previously been proposed, for example in [19]. Errors in these vectors change the joint-centre acceleration that is calculated from the sensor measurements and therefore add to its uncertainty.

2-4-2 Relative-Orientation Information from a Common Joint Centre

The use of a common joint centre provides a mechanism for coupling measurements from two adjacent IMUs. In [5], accelerometer and gyroscope measurements were used to construct virtual accelerometers at the centre of rotation of a joint. Because both virtual sensors represent the acceleration of the same physical point, their relationship provides information about the relative motion of the connected segments.

Related approaches have subsequently exploited kinematic relationships between adjacent segments for joint-axis estimation, joint-position estimation and relative-orientation estimation [19, 12]. An important advantage of these approaches is that the information is derived from the connection between the segments rather than from an external heading reference. This makes magnetometer-free relative-orientation estimation possible under suitable motion conditions.

The amount of orientation information obtained from such relationships is, however, dependent on the performed motion. Motion excitation and observability are therefore important considerations when exploiting connected-segment kinematics.

2-5 Related Work on Relative-Orientation Estimation

2-5-1 Drift-Free Relative Orientation Using Joint-Centre Acceleration

A tightly coupled approach for estimating the three-dimensional relative orientation between two IMUs attached to connected segments was proposed in [25]. Instead of estimating each

sensor orientation independently and combining the results afterwards, the method incorporates the shared joint-centre motion directly into the estimation procedure.

The acceleration of the common joint centre is calculated from the accelerometer and gyroscope measurements of the two sensors, independently, using rigid-body kinematics. Since both calculations represent the acceleration of the same physical point, their agreement provides information about the relative orientation of the sensors. This information is used to correct the relative drift caused by gyroscope integration without requiring continuous magnetometer measurements.

Both an optimization-based smoothing implementation and an online filtering implementation were presented in [25]. The method was validated using an industrial robotic manipulator under different movement conditions and using recorded human gait data. The robotic experiments showed that arbitrary three-dimensional relative sensor orientation can be estimated, while the human evaluation demonstrated long-term estimation of knee-joint kinematics during walking.

2-5-2 Observability of Relative Orientation

The joint-centre acceleration relationship does not provide the same amount of relative-orientation information for every performed motion. The observability of magnetometer-free relative orientation from connected-segment inertial measurements was analysed in [8].

The analysis shows that relative orientation becomes observable when the joint-centre specific force and its temporal variation provide sufficiently independent directional information. Intuitively, repeatedly observing the joint-centre acceleration in different directions restricts the set of relative orientations that are compatible with the measurements. In contrast, static motion or motion for which the relevant directions change insufficiently may leave one or more relative-orientation directions unobservable.

An observability-related quantity used in [8] is

$$o_{i,t} = \left\| a_{jc,t}^{S_i} \times \left(\dot{a}_{jc,t}^{S_i} + \omega_t^{S_i} \times a_{jc,t}^{S_i} \right) \right\|_2. \quad (2-10)$$

The quantity is evaluated at time t . In the estimator of Chapter 3 it is calculated using the measured angular velocity $y_{\omega,t}^{S_i}$. For two adjacent sensors, the minimum of the two quantities, $o_{\min,t} = \min(o_{1,t}, o_{2,t})$, can be used as a conservative indication of the available acceleration-based information [21].

The magnitude of $o_{i,t}$ indicates the level to which the measured motion contains the directional variation required for relative-orientation estimation. An example of low observable relative motion is prolonged standing, where the knee joint-centre acceleration provides little changing directional information for resolving the relative orientation between the thigh and the shank. Richer motion, such as walking or movements involving larger changes in segment acceleration, provides more information for correcting relative-orientation drift. Temporary periods of lower observability do not necessarily prevent accurate estimation over an entire recording if sufficiently informative motion occurs at other times.

2-5-3 Conditional Magnetic Information

The magnetometer-free framework of [25] was extended in [21] by conditionally incorporating relative magnetic-field direction information. The method, referred to as MAJIC (Magnetic Field at Inertial Joint Centre), is intended to provide additional directional information during periods in which the joint-centre-based measurement provides insufficient information for accurately constraining the relative orientation.

Rather than relying on a single globally known magnetic-field direction, the approach compares the directions of the magnetic-field measurements from the two neighbouring sensors. This reduces the dependence on knowledge of the absolute magnetic field and instead relies on the assumption that the two sensors experience sufficiently similar local magnetic-field directions.

The magnetic information is not applied continuously. It is incorporated when the acceleration-based observability falls below a selected threshold, so that the additional measurement is primarily used during periods in which the joint-centre acceleration provides limited relative-orientation information.

This introduces a different source of uncertainty from the joint-centre acceleration relationship. When the magnetic-field directions at the two sensor locations are sufficiently consistent, the additional information can assist relative-orientation estimation during poorly observable motion. When the local magnetic fields differ, however, the magnetic information itself becomes inconsistent and may introduce orientation error.

2-6 Extended Kalman Filtering and Measurement Consistency

Kalman filtering combines a model-based state prediction with measurement information while explicitly representing uncertainty. For nonlinear systems, the EKF applies the Kalman-filtering framework after linearizing the relevant process and measurement models around the current state estimate [24, 9].

Consider a measurement model

$$y_t = h(x_t) + v_t, \quad (2-11)$$

where the measurement noise is modelled as

$$v_t \sim \mathcal{N}(0, R_t). \quad (2-12)$$

For a predicted state $\hat{x}_{t|t-1}$, the measurement innovation is

$$\nu_t = y_t - h(\hat{x}_{t|t-1}), \quad (2-13)$$

with predicted innovation covariance

$$S_t = H_t P_{t|t-1} H_t^\top + R_t, \quad (2-14)$$

where $P_{t|t-1}$ is the predicted state covariance and H_t is the measurement-model Jacobian.

The corresponding Kalman gain is

$$K_t = P_{t|t-1} H_t^\top S_t^{-1}. \quad (2-15)$$

Consequently, the measurement covariance R_t directly influences the relative weighting of prediction and measurement information. A larger measurement covariance reduces the influence of the corresponding measurement update, whereas a smaller covariance achieves the opposite.

The Kalman-filtering equations rely on the assumed process and measurement uncertainties adequately representing the observed system. If the actual measurement uncertainty changes while R_t remains fixed, the statistical weight assigned to the measurement can become inappropriate. Innovation statistics provide one way of assessing whether the measurements remain consistent with these assumptions.

2-6-1 Normalized Innovation Squared and Measurement Rejection

The Normalized Innovation Squared (NIS) is a statistical measure of innovation consistency. It is defined as

$$\epsilon_t = \nu_t^\top S_t^{-1} \nu_t. \quad (2-16)$$

For a statistically consistent filter with an n_y -dimensional Gaussian innovation, ϵ_t follows a chi-squared distribution with n_y degrees of freedom [17]. This allows unexpectedly large innovations to be identified using a confidence threshold.

One possible response to such an innovation is statistical gating, in which a measurement update is omitted when its NIS exceeds a selected threshold. Measurement rejection is attractive because it is straightforward and computationally inexpensive and can prevent an isolated strongly inconsistent measurement from causing a large state correction.

The approach is inherently binary. A measurement immediately below the threshold is accepted, whereas a measurement immediately above it is discarded completely. Moreover, a large innovation does not uniquely imply an unreliable measurement. It may also result from an inaccurate state estimate or from model errors. Rejecting the measurement in such circumstances may remove information that would otherwise help the estimator recover.

Related threshold-based strategies have been used in inertial orientation estimation to reduce accelerometer influence during dynamic motion [13] and to suppress magnetic updates when disturbances are detected [10].

2-6-2 Adaptive Covariance Estimation

Adaptive covariance estimation provides a continuous alternative to binary measurement rejection. Instead of accepting or discarding a measurement entirely, an adaptive Kalman filter modifies the covariance assumptions online. Depending on the considered source of inconsistency, adaptive methods may modify the process covariance Q , the measurement covariance R , or both [26, 27].

Innovation-based covariance adaptation exploits the fact that the observed innovation statistics contain information about the statistical consistency of the filter. According to (2-14), the predicted innovation covariance depends on both the propagated state uncertainty and the assumed measurement covariance. An empirical estimate of the innovation statistics

can therefore be compared with the predicted covariance and used to modify the assumed uncertainty.

Several adaptive filtering strategies have been proposed. In [26], an adaptive factor is calculated from discrepancies between predicted and observed quantities and used to modify the relative influence of model and measurement information. In [27], innovation statistics are estimated over a sliding window for adaptive magnetic-inertial human-motion tracking.

Sliding-window methods introduce a trade-off between responsiveness and statistical averaging. Short windows can react quickly to changes but produce noisier covariance estimates, whereas longer windows provide smoother estimates but respond more slowly and may retain the effect of a disturbance after it has ended [27, 22].

These methods establish that innovation statistics can be used to modify measurement weighting online. Their application to the specific joint-centre acceleration update considered in the following chapter requires defining how the innovation is formed, which covariance is adapted and how the adaptation is constrained.

Methods for IMU-Based Relative-Orientation Estimation

This chapter presents the relative-orientation estimation methods investigated in this thesis. The estimation problem is first formulated for two IMUs attached to adjacent rigid body segments. The magnetometer-free estimator proposed in [25] is then defined, referred to as “Weygers”, including the gyroscope-driven orientation propagation and the joint-centre acceleration measurement update. Subsequently, the conditional magnetic-direction update from [21] is introduced as a second estimator framework.

The proposed innovation-based adaptive covariance method and the NIS-based (normalized innovation squared) measurement-rejection method are then incorporated into the joint-centre acceleration update. Both mechanisms are applied without modifying the underlying propagation model or, for the magnetometer-aided estimator, the magnetic update. This results in six estimator variants that differ only in their treatment of the joint-centre acceleration measurement and in whether conditional magnetic information is available.

3-1 Relative-Orientation Estimation Problem

Consider two rigid body segments connected by a joint, with one IMU rigidly attached to each segment. The corresponding sensor frames are denoted by S_1 and S_2 , while G denotes a common global reference frame. The orientations of the two sensor frames with respect to G are represented by the unit quaternions $q_t^{GS_1}$ and $q_t^{GS_2}$, with corresponding rotation matrices $R_t^{GS_1}$ and $R_t^{GS_2}$, where $t \in \{1, \dots, N\}$ denotes the discrete time-step index and N is the total number of time steps.

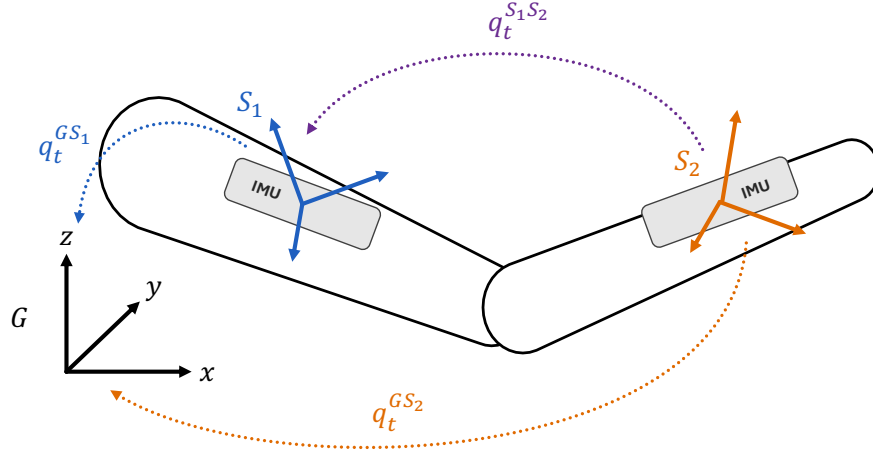


Figure 3-1: Schematic of the relative-orientation estimation problem. Two rigid segments are connected by a joint, with an IMU rigidly attached to each segment. The sensor frames S_1 and S_2 have orientations $q_t^{GS_1}$ and $q_t^{GS_2}$ with respect to the common global frame G . The relative orientation $q_t^{S_1S_2}$ between the two sensor frames is the quantity to be estimated.

The quantity of interest is the three-dimensional relative orientation of sensor frame S_2 with respect to sensor frame S_1 , given by (2-6). Estimation of $q_t^{S_1S_2}$ therefore does not require the individual sensor orientations to be known with respect to an externally meaningful absolute orientation, provided that their rotational relationship can be estimated accurately.

Each IMU provides gyroscope and accelerometer measurements, $y_{\omega,t}^{S_i}$ and $y_{a,t}^{S_i}$, respectively. Magnetometer measurements $y_{m,t}^{S_i}$ are additionally available for the magnetometer-aided estimator. Gyroscope measurements provide the short-term orientation change of each sensor but, when integrated over time, measurement errors and bias cause the individual orientation estimates to drift. If this drift differs between the two sensors, it also appears in the relative orientation.

The physical connection between the two body segments provides additional information that can be used to correct this relative drift. Both segments share the same joint centre. Consequently, the acceleration of this common point can be calculated independently from the inertial measurements of either sensor. After transformation to a common coordinate frame, both calculations should represent the same physical acceleration. The resulting disagreement therefore provides information about the relative orientation between the two sensors. This principle forms the joint-centre acceleration measurement used in [25] and constitutes the primary measurement update considered throughout this thesis.

Two distinct properties of this measurement are relevant. First, the amount of relative-orientation information provided by the measurement depends on the performed motion. Insufficient excitation can make one or more relative orientation directions weakly observable or temporarily unobservable [8]. Second, even when the motion is informative, the statistical agreement between the measurement model and the observed inertial data can deteriorate due to sensor uncertainty, errors in the assumed sensor-to-joint-centre geometry, numerical differentiation errors or violations of the rigid-body model.

The methods investigated in this thesis address the second of these properties. They modify

the treatment of the joint-centre acceleration update according to its observed statistical consistency. They cannot, however, create orientation information that is absent because of insufficient motion excitation.

3-2 Magnetometer-Free Relative-Orientation Estimator

The magnetometer-free estimator used as the primary baseline is the online filtering formulation proposed in [25]. It is formulated as a deviation-state EKF [9]. The nominal orientations of both sensors are represented by unit quaternions, while small orientation deviations around these quaternion estimates are represented in a six-dimensional deviation state that is expressed in the global frame.

3-2-1 Orientation Deviation State

For sensor S_i , the orientation is locally represented as

$$q_t^{GS_i} = \exp_q\left(\frac{1}{2}\eta_{i,t}\right) \odot \tilde{q}_t^{GS_i}, \quad (3-1)$$

where $\tilde{q}_t^{GS_i}$ is the current quaternion linearization point with corresponding rotation matrix $\tilde{R}_t^{GS_i}$, $\eta_{i,t} \in \mathbb{R}^3$ is a local orientation deviation expressed in the global frame G , and $\exp_q(\cdot)$ denotes the quaternion exponential map. The subscripts $t|t-1$ and $t|t$ denote the linearization point (and the covariance P below) before and after the measurement update at time t . The measurement update at time t is evaluated at the prediction, so that $\tilde{q}_t^{GS_i} = \tilde{q}_{t|t-1}^{GS_i}$ and $\tilde{R}_t^{GS_i} = \tilde{R}_{t|t-1}^{GS_i}$ in all measurement models of this chapter.

For the two-sensor system, the combined deviation state is

$$\eta_t = \begin{bmatrix} \eta_{1,t}^\top & \eta_{2,t}^\top \end{bmatrix}^\top \in \mathbb{R}^6. \quad (3-2)$$

The quaternion orientations therefore remain the nominal representation of the sensor attitudes, whereas the EKF estimates a local six-dimensional correction around the current quaternion linearization points.

3-2-2 Gyroscope-Driven Propagation

Between measurement updates, the two sensor orientations are propagated independently using their gyroscope measurements. For sensor S_i , the discrete-time propagation is

$$\tilde{q}_{t|t-1}^{GS_i} = \tilde{q}_{t-1|t-1}^{GS_i} \odot \exp_q\left(\frac{T_s}{2}y_{\omega,t-1}^{S_i}\right), \quad (3-3)$$

where T_s denotes the sampling interval. The result is normalized to unit norm.

Because the deviation is expressed in the global frame, its linearized dynamics are the identity, $F_t = I_6$ [9], and the gyroscope noise enters only through G_t . The covariance of the deviation state is therefore propagated as

$$P_{t|t-1} = P_{t-1|t-1} + G_t Q_\omega G_t^\top, \quad G_t = T_s \begin{bmatrix} \tilde{R}_{t|t-1}^{GS_1} & 0 \\ 0 & \tilde{R}_{t|t-1}^{GS_2} \end{bmatrix}, \quad (3-4)$$

where $Q_\omega \in \mathbb{R}^{6 \times 6}$ is the process covariance associated with the gyroscope-driven propagation.

3-2-3 Joint-Centre Acceleration Measurement Model

In [25], the shared acceleration of the joint centre is used to couple the orientation estimates of two sensors attached to adjacent segments. Consider two rigid segments carrying sensor frames S_1 and S_2 and connected at a common joint centre. Let $r_i^{S_i} \in \mathbb{R}^3$ denote the vector from the origin of sensor i to the joint centre, expressed in frame S_i . Using rigid-body kinematics, the accelerometer measurement can be translated from the sensor origin to the joint centre. Following [25],

$$a_{jc,t}^{S_i} = y_{a,t}^{S_i} + C_t^{S_i} r_i^{S_i}, \quad C_t^{S_i} = \left[\dot{y}_{\omega,t}^{S_i} \times \right]^2 + \left[\dot{y}_{\omega,t}^{S_i} \times \right], \quad (3-5)$$

where $\dot{y}_{\omega,t}^{S_i}$ is the angular acceleration, obtained by numerical differentiation of the gyroscope measurements. Here, $[\cdot \times]$ denotes the skew-symmetric cross-product matrix. The two terms of $C_t^{S_i}$ describe the centripetal and tangential acceleration associated with the offset between the sensor and the joint centre. Both $a_{jc,t}^{S_1}$ and $a_{jc,t}^{S_2}$ contain the same gravity contribution when expressed in G , which cancels in (3-6).

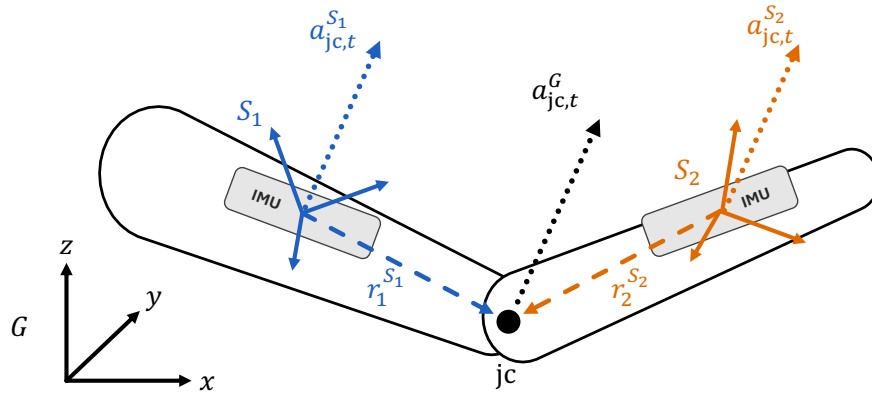


Figure 3-2: Joint-centre acceleration model. Two rigid segments with sensors S_1 and S_2 are connected at the joint centre (jc). The vector $r_i^{S_i}$ points from the origin of sensor i to the joint centre, expressed in S_i . The joint-centre acceleration $a_{jc,t}^{S_i}$ in (3-5) is calculated separately from each sensor and expressed in its own sensor frame (dotted, blue and orange). Both describe the acceleration of the same physical point, which is $a_{jc,t}^G$ when expressed in the global frame G (dotted, black), so that the two calculated vectors can be compared after rotation to G , as in (3-6).

Although $a_{jc,t}^{S_1}$ and $a_{jc,t}^{S_2}$ are expressed in different sensor frames, they describe the acceleration of the same physical point (Figure 3-2). After transformation to a common frame, they are

therefore related through the measurement model

$$R_t^{GS_1} a_{jc,t}^{S_1} - R_t^{GS_2} a_{jc,t}^{S_2} = e_{\text{link},t}, \quad (3-6)$$

where $e_{\text{link},t}$ represents the uncertainty associated with the joint-centre acceleration measurement model. Its covariance is denoted by $R_{a,t}$.

3-2-4 Deviation-State EKF Formulation

In [25], the gyroscope propagation and the joint-centre acceleration measurement model are combined in a deviation-state EKF. The measurement model (3-6) is nonlinear in the sensor orientations. It is linearized with respect to the deviation state around $\eta_{i,t} = 0$, corresponding to the predicted orientations $\tilde{R}_t^{GS_i}$. Using

$$R_t^{GS_i} \approx (I_3 + [\eta_{i,t} \times]) \tilde{R}_t^{GS_i} \quad (3-7)$$

and $[\eta \times]v = -[v \times]\eta$, this gives

$$R_t^{GS_i} a_{jc,t}^{S_i} \approx \tilde{R}_t^{GS_i} a_{jc,t}^{S_i} - [\tilde{R}_t^{GS_i} a_{jc,t}^{S_i} \times] \eta_{i,t}. \quad (3-8)$$

Substituting this in (3-6) results in the linearized model $\nu_{a,t} \approx H_{a,t} \eta_t + e_{\text{link},t}$, with the acceleration innovation

$$\nu_{a,t} = \tilde{R}_t^{GS_1} a_{jc,t}^{S_1} - \tilde{R}_t^{GS_2} a_{jc,t}^{S_2} \quad (3-9)$$

and the measurement Jacobian

$$H_{a,t} = \left[[\tilde{R}_t^{GS_1} a_{jc,t}^{S_1} \times] \quad - [\tilde{R}_t^{GS_2} a_{jc,t}^{S_2} \times] \right]. \quad (3-10)$$

The innovation represents the disagreement between the two estimates of the acceleration of the same physical joint centre when expressed in a common frame, and $H_{a,t}$ is the corresponding sensitivity to the deviation state.

The covariance $R_{a,t}$ of the measurement error represents the assumed uncertainty associated with the joint-centre acceleration update and therefore directly affects its influence on the orientation correction. This uncertainty is not limited to accelerometer measurement noise, as the joint-centre acceleration also depends on gyroscope measurements, angular-acceleration estimation and the assumed sensor-to-joint-centre geometry. In [25], it is a fixed matrix, $R_{a,t} = R_a$, referred to as the nominal acceleration covariance. The adaptive method of Section 3-4 lets $R_{a,t}$ vary over time.

The innovation covariance, Kalman gain and deviation estimate of the acceleration update are

$$S_{a,t} = H_{a,t} P_{t|t-1} H_{a,t}^\top + R_{a,t}, \quad (3-11)$$

$$K_{a,t} = P_{t|t-1} H_{a,t}^\top S_{a,t}^{-1}, \quad (3-12)$$

$$\hat{\eta}_t = K_{a,t} \nu_{a,t}. \quad (3-13)$$

The deviation estimate is applied to the linearization points,

$$\tilde{q}_{t|t}^{GS_i} = \exp_q \left(\frac{1}{2} \hat{\eta}_{i,t} \right) \odot \tilde{q}_{t|t-1}^{GS_i}, \quad (3-14)$$

where $\hat{\eta}_{i,t}$ is the part of $\hat{\eta}_t$ that belongs to sensor i . The result is normalized to unit norm, serves as the new linearization point, and the deviation state is reset to zero. The covariance is updated in Joseph form,

$$P_{t|t} = (I_6 - K_{a,t}H_{a,t})P_{t|t-1}(I_6 - K_{a,t}H_{a,t})^\top + K_{a,t}R_{a,t}K_{a,t}^\top. \quad (3-15)$$

The covariance is not transformed further by the reset of the deviation state: for a deviation expressed in the global frame, the Jacobian of this reset is the identity [9, Eq. (4.30)]. The algorithm in [25] contains an additional matrix J_t in this step, which is not defined there.

3-3 Conditional Magnetic-Direction Update

A second estimator framework is obtained by extending the magnetometer-free estimator with the conditional magnetic-direction update proposed in [21]. The joint-centre acceleration update is retained, while magnetic information is additionally used during periods of low acceleration-based observability.

For each sensor, the magnetometer measurement is normalized according to

$$\hat{m}_t^{S_i} = \frac{y_{m,t}^{S_i}}{\|y_{m,t}^{S_i}\|_2}. \quad (3-16)$$

The resulting measurement therefore depends on magnetic-field direction rather than magnitude. Under the assumption that both neighbouring sensors experience approximately the same magnetic-field direction, the transformed normalized measurements should agree in a common coordinate frame. The corresponding magnetic innovation is

$$\nu_{m,t} = \tilde{R}_t^{GS_1}\hat{m}_t^{S_1} - \tilde{R}_t^{GS_2}\hat{m}_t^{S_2}. \quad (3-17)$$

Linearizing in the same way as for the acceleration measurement gives the measurement Jacobian

$$H_{m,t} = \left[\left[\tilde{R}_t^{GS_1}\hat{m}_t^{S_1} \times \right] \quad - \left[\tilde{R}_t^{GS_2}\hat{m}_t^{S_2} \times \right] \right]. \quad (3-18)$$

The magnetic update is not applied continuously. Following [21], its availability is determined using the acceleration-based observability of both sensors. For sensor S_i , the quantity

$$o_{i,t} = \left\| a_{jc,t}^{S_i} \times \left(\dot{a}_{jc,t}^{S_i} + y_{\omega,t}^{S_i} \times a_{jc,t}^{S_i} \right) \right\|_2 \quad (3-19)$$

is calculated,

and the conservative observability indicator is

$$o_{\min,t} = \min(o_{1,t}, o_{2,t}). \quad (3-20)$$

The magnetic measurement update is requested when $o_{\min,t} < o_{\text{threshold}}$. It is then used together with the acceleration measurement in a single update, with fixed magnetic measurement covariance R_m . The measurements that are used at time t are stacked,

$$\nu_t = \begin{bmatrix} \nu_{a,t} \\ \nu_{m,t} \end{bmatrix}, \quad H_t = \begin{bmatrix} H_{a,t} \\ H_{m,t} \end{bmatrix}, \quad R_t = \text{blkdiag}(R_{a,t}, R_m), \quad (3-21)$$

where the blocks of measurements that are not used are omitted. An acceleration measurement is not used when it is rejected (Section 3-5), and the magnetic measurement is not used when it is not requested. The update then follows Section 3-2-4 with $\nu_{a,t}$, $H_{a,t}$ and $R_{a,t}$ replaced by ν_t , H_t and R_t ,

$$S_t = H_t P_{t|t-1} H_t^\top + R_t, \quad K_t = P_{t|t-1} H_t^\top S_t^{-1}, \quad \hat{\eta}_t = K_t \nu_t, \quad (3-22)$$

followed by (3-14) and (3-15) with $K_{a,t}$, $H_{a,t}$ and $R_{a,t}$ replaced by K_t , H_t and R_t . If no measurement is used, $\tilde{q}_{t|t}^{GS_i} = \tilde{q}_{t|t-1}^{GS_i}$ and $P_{t|t} = P_{t|t-1}$.

3-4 Innovation-Based Adaptive Acceleration Covariance

The proposed adaptive method modifies the influence of the joint-centre acceleration measurement according to its recent innovation statistics. The nominal acceleration measurement covariance R_a is multiplied by a time-varying scalar factor,

$$R_{a,t} = \alpha_t R_a, \quad (3-23)$$

where α_t denotes the covariance scaling factor applied at time t .

This scalar formulation preserves the directional structure of the nominal covariance while allowing its overall magnitude, and therefore the statistical weight assigned to the acceleration update, to vary over time.

From (3-11), the predicted innovation covariance consists of a contribution from the propagated state uncertainty and a contribution from the assumed acceleration measurement uncertainty. The adaptive method estimates whether the nominal value R_a adequately represents the latter contribution by comparing the predicted innovation statistics with those recently observed by the estimator.

3-4-1 Empirical Innovation Statistics

An empirical second moment of the acceleration innovation is calculated over a sliding window containing the most recent N_w innovations,

$$\hat{C}_{\nu,t} = \frac{1}{N_w} \sum_{j=t-N_w+1}^t \nu_{a,j} \nu_{a,j}^\top. \quad (3-24)$$

The innovations are not explicitly mean-centred. Consequently, (3-24) represents an empirical second moment rather than a sample covariance in the strict statistical sense. When the innovation is approximately zero mean, it provides an empirical estimate of the innovation covariance.

3-4-2 Covariance Matching

Using (3-23) as an approximation for (3-11) gives

$$\hat{C}_{\nu,t} \approx H_{a,t} P_{t|t-1} H_{a,t}^\top + \alpha_t R_a. \quad (3-25)$$

A single scalar α_t cannot generally make both sides of (3-25) equal element by element. Instead, the overall covariance magnitude is matched using the trace,

$$\text{tr}(\hat{C}_{\nu,t}) \approx \text{tr}(H_{a,t} P_{t|t-1} H_{a,t}^\top) + \alpha_t \text{tr}(R_a). \quad (3-26)$$

Solving this scalar relationship gives the unconstrained estimate

$$\alpha_t^{\text{raw}} := \frac{\text{tr}(\hat{C}_{\nu,t} - H_{a,t} P_{t|t-1} H_{a,t}^\top)}{\text{tr}(R_a)}. \quad (3-27)$$

The trace operation reduces the three-dimensional covariance discrepancy to a single scalar. The method therefore adapts the overall assumed uncertainty of the joint-centre acceleration measurement without independently estimating all elements of a complete 3×3 covariance matrix.

The unconstrained estimate can become very small or negative when the empirically observed innovation magnitude is lower than the contribution predicted from the state uncertainty. A lower bound is therefore imposed, $\alpha_t = \max(\alpha_{\min}, \alpha_t^{\text{raw}})$.

The effective acceleration measurement covariance is then obtained from (3-23). For $\alpha_t > 1$, the acceleration measurement is assigned greater uncertainty and receives less weight during the EKF correction. For $\alpha_t < 1$, the filter assigns greater statistical confidence to the acceleration measurement.

The resulting $R_{a,t}$ should be interpreted as an effective covariance of the complete joint-centre acceleration measurement rather than as an estimate of accelerometer sensor noise. An increased innovation can result from several different mechanisms, including sensor errors, incorrect joint-centre geometry, state-estimation error or violations of the rigid-body model. The scalar factor does not identify which of these mechanisms caused the observed covariance mismatch.

Before a complete innovation window is available, the nominal covariance is retained. The window only contains innovations obtained at or after a start-up delay $T_d \geq 0$ (Section 3-8), so that

$$R_{a,t} = R_a, \quad t < N_w + \frac{T_d}{T_s}. \quad (3-28)$$

3-5 NIS-Based Acceleration Measurement Rejection

As a binary alternative to continuous covariance adaptation, the joint-centre acceleration measurement can be rejected based on its Normalized Innovation Squared (NIS). For the acceleration update,

$$\epsilon_{a,t} = \nu_{a,t}^\top S_{a,t}^{-1} \nu_{a,t}, \quad (3-29)$$

where $S_{a,t}$ is the predicted innovation covariance in (3-11), evaluated with $R_{a,t} = R_a$.

The nominal covariance R_a is used for this consistency test. Under the assumptions of a statistically consistent filter and Gaussian innovations, the three-dimensional acceleration innovation results in an NIS that follows a chi-squared distribution with three degrees of freedom.

The acceleration update is accepted when $\epsilon_{a,t} \leq \epsilon_{\max}$, and rejected when $\epsilon_{a,t} > \epsilon_{\max}$. When rejected, the acceleration block is omitted from the update of that time step. A requested magnetic measurement is still used, and if no measurement is used, the propagated estimate is retained. Rejection is only applied from time T_{nis} onward. Before that, the NIS is calculated but the measurement is not rejected.

The NIS-based method therefore represents a fundamentally different response to inconsistency from adaptive covariance scaling. The adaptive method changes the measurement weight continuously, whereas NIS rejection either applies the nominal update or removes it completely.

3-6 Estimator Variants

The two reliability-handling mechanisms are incorporated into two underlying relative orientation estimator frameworks: the magnetometer-free estimator in [25], and the magnetometer-aided estimator in [21].

For the magnetometer-free framework, the three variants differ only in the treatment of the joint-centre acceleration measurement: fixed covariance, adaptive covariance scaling or NIS-based rejection.

For the magnetometer-aided framework, the same three acceleration-update variants are used while the conditional magnetic-direction update remains unchanged. Restricting the reliability-handling modifications to the acceleration update ensures that differences between the three magnetometer-aided variants cannot be attributed to changes in the magnetic update.

The resulting six estimator variants are summarized in Table 3-1.

Table 3-1: Estimator variants considered in this thesis. Adaptation and NIS rejection affect only the joint-centre acceleration update.

Method	Acceleration update	Adaptive R_a	Magnetic update
Weygers	Fixed covariance	No	No
Weygers + Adaptive	Adaptive covariance	Yes	No
Weygers + NIS	NIS gated	No	No
MAJIC	Fixed covariance	No	Observability dependent
MAJIC + Adaptive	Adaptive covariance	Yes	Observability dependent
MAJIC + NIS	NIS gated	No	Observability dependent

3-7 Unified Estimation Procedure

The common structure of the six estimator variants is summarized in Algorithm 1. The selected variant determines how the joint-centre acceleration measurement is treated and whether the conditional magnetic-direction update is available.

Algorithm 1 Relative-orientation estimation with optional adaptive covariance scaling, NIS rejection and conditional magnetic updates.	
Require:	IMU measurements $\{y_{\omega,t}^{S_i}, y_{a,t}^{S_i}\}_{t=1}^N$, $i = 1, 2$; for MAJIC, $\{y_{m,t}^{S_i}\}_{t=1}^N$; $T_s, r_i^{S_i}, \hat{q}_0^{GS_i}, P_0, Q_\omega, R_a$; adaptive: $N_w, T_d, \alpha_{\min}$; NIS: $\epsilon_{\max}, T_{\text{nis}}$; MAJIC: $R_m, o_{\text{threshold}}$
Ensure:	Relative-orientation estimate $\hat{q}_t^{S_1 S_2}$, $t = 1, \dots, N$
1:	Set $\hat{q}_{1 1}^{GS_i} = \hat{q}_0^{GS_i}$ and $P_{1 1} = P_0$
2:	for $t = 2, \dots, N$ do
3:	Time update: $\tilde{q}_{t t-1}^{GS_i} = \tilde{q}_{t-1 t-1}^{GS_i} \odot \exp_q\left(\frac{T_s}{2} y_{\omega,t-1}^{S_i}\right)$, $i = 1, 2$ $\triangleright$ (3-3)
4:	$P_{t t-1} = P_{t-1 t-1} + G_t Q_\omega G_t^\top$ $\triangleright$ (3-4)
5:	Acceleration measurement: calculate $a_{jc,t}^{S_i}, \nu_{a,t}$ and $H_{a,t}$ $\triangleright$ (3-5), (3-9), (3-10)
6:	$R_{a,t} = R_a$
7:	if adaptive covariance is enabled and N_w innovations since T_d are available then
8:	Calculate $\hat{C}_{\nu,t}$ and α_t^{raw} $\triangleright$ (3-24), (3-27)
9:	$\alpha_t = \max(\alpha_{\min}, \alpha_t^{\text{raw}})$, $R_{a,t} = \alpha_t R_a$ $\triangleright$ (3-23)
10:	end if
11:	if NIS rejection is enabled and $(t-1)T_s \geq T_{\text{nis}}$ then
12:	Calculate $\epsilon_{a,t}$ $\triangleright$ (3-29), with R_a
13:	if $\epsilon_{a,t} > \epsilon_{\max}$ then discard the acceleration measurement
14:	end if
15:	if MAJIC is enabled and $o_{\min,t} < o_{\text{threshold}}$ then $\triangleright$ (3-19)
16:	Magnetic measurement: calculate $\hat{m}_t^{S_i}, \nu_{m,t}$ and $H_{m,t}$ $\triangleright$ (3-16), (3-17), (3-18)
17:	end if
18:	if at least one measurement is used then
19:	Measurement update: stack the used measurements, $\nu_t = [\nu_{a,t}^\top, \nu_{m,t}^\top]^\top$, $H_t = [H_{a,t}^\top, H_{m,t}^\top]^\top$, $R_t = \text{blkdiag}(R_{a,t}, R_m)$ $\triangleright$ unused blocks omitted
20:	$S_t = H_t P_{t t-1} H_t^\top + R_t$, $K_t = P_{t t-1} H_t^\top S_t^{-1}$, $\hat{\eta}_t = K_t \nu_t$
21:	$\tilde{q}_{t t}^{GS_i} = \exp_q\left(\frac{1}{2} \hat{\eta}_{i,t}\right) \odot \tilde{q}_{t t-1}^{GS_i}$, $i = 1, 2$ $\triangleright$ followed by normalization
22:	$P_{t t} = (I_6 - K_t H_t) P_{t t-1} (I_6 - K_t H_t)^\top + K_t R_t K_t^\top$
23:	else
24:	$\tilde{q}_{t t}^{GS_i} = \tilde{q}_{t t-1}^{GS_i}$ and $P_{t t} = P_{t t-1}$
25:	end if
26:	$\hat{q}_t^{S_1 S_2} = \left(\tilde{q}_{t t}^{GS_1}\right)^{-1} \odot \tilde{q}_{t t}^{GS_2}$ $\triangleright$ for $t = 1$ use $\tilde{q}_{1 1}^{GS_i}$
27:	end for

The adaptive and rejection mechanisms should therefore not be interpreted as independent relative-orientation estimators. Both modify only the treatment of the joint-centre acceleration measurement within the same underlying deviation-state EKF. This common structure allows their influence to be evaluated without simultaneously changing the orientation propagation or, for MAJIC, the conditional magnetic update.

3-8 Estimator Configuration and Design Choices

The preceding sections define the estimator structure independently of the numerical parameters used in the evaluation. Two configurations are used. The *literature configuration* described in this section follows [25] and [21] and is used in the simulation study of Chapter 4, where the response of the estimators to controlled changes is investigated without tuning. The *tuned configuration* is used for the recorded data of Chapter 5. It keeps the structure and the parameters of this section that are not mentioned in Section 3-8-7, and replaces the covariances, the NIS thresholds, the adaptive window and the start-up delay by values tuned on the recorded data.

3-8-1 Estimator Initialization

Both sensor orientations are initialized using the identity quaternion,

$$\tilde{q}_0^{GS_1} = \tilde{q}_0^{GS_2} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^\top. \quad (3-30)$$

No prior knowledge of the initial sensor orientations is therefore assumed.

The initial deviation-state covariance is $P_0 = I_6$, following the initialization used in [25]. The same initialization is used for all six estimator variants.

3-8-2 Process Covariance

The gyroscope-driven propagation uses $Q_\omega = 10^{-4}I_6$, following the configuration in [21]. This covariance is kept fixed for every estimator variant. It is not derived from the sensor noise levels used in the evaluations. The proposed adaptive method therefore modifies only the uncertainty associated with the joint-centre acceleration measurement and does not adapt the propagation model.

3-8-3 Nominal Acceleration Covariance

The nominal joint-centre acceleration measurement covariance is $R_a = 2.5 \times 10^{-3}I_3$, based on the configuration in [21].

For the baseline estimators, this covariance remains fixed. For the adaptive variants, it forms the reference covariance scaled by α_t . For the NIS-rejection variants, the same nominal covariance is used to calculate the innovation covariance used by the consistency test. Using the same R_a across the estimator variants ensures that their comparison concerns the reliability-handling mechanism rather than different nominal covariance settings. Like Q_ω , R_a is adopted from the literature and is not matched to the noise levels of the evaluations, so that these capture the mismatch between the assumed covariance and the data.

3-8-4 Adaptive-Covariance Parameters

The empirical innovation second moment is estimated over a temporal window of $T_w = 0.5$ s. For sampling frequency f_s , the corresponding number of samples is $N_w = T_w f_s$. At $f_s = 100$ Hz, this gives $N_w = 50$.

The window length introduces a trade-off between responsiveness and statistical averaging. A shorter window responds more rapidly to changes but produces a more variable estimate of the innovation statistics. A longer window provides greater averaging but responds more slowly and retains the influence of a disturbance for a longer period. A duration of 0.5 s is selected based on the covariance-matching approach in [3], where the same duration was used for online measurement-covariance estimation.

The adaptive factor is lower bounded by $\alpha_{\min} = 0.1$. This prevents the effective covariance from approaching zero when the empirical innovation second moment is smaller than predicted. The lower bound is therefore introduced as a numerical and estimator-confidence safeguard rather than as a theoretically optimal value.

No upper bound is imposed. Consequently, sufficiently large innovation statistics can strongly reduce the influence of the joint-centre acceleration measurement.

The adaptation can additionally be delayed. Innovations obtained during a start-up delay T_d are not used in (3-24), and the nominal covariance is retained until a complete window of N_w innovations has been collected after the delay. The motivation is that the innovations at the start of a recording are dominated by the error of the initial orientation, which is not caused by an unreliable measurement. The delay is $T_d = 0$ in the literature configuration and is tuned for the recorded data.

3-8-5 NIS Threshold

The acceleration innovation is three-dimensional. The NIS-rejection method therefore uses the 99% upper chi-squared threshold with three degrees of freedom, $\epsilon_{\max} = \chi_{3,0.99}^2 = 11.3449$.

This provides a statistically defined rejection criterion rather than an arbitrary threshold on the Euclidean innovation magnitude.

Rejection is not active during the first $T_{\text{nis}} = 10$ s of a recording. The large innovations caused by the initial orientation error would otherwise be rejected together with the measurements that are needed to remove that error. The NIS is still calculated during this interval. Any deviations from this nominal configuration required for individual evaluations are described in the corresponding chapter.

3-8-6 MAJIC Configuration

For all MAJIC variants, the acceleration-based observability threshold is $o_{\text{threshold}} = 150$, following [21].

The magnetic measurement covariance is fixed at $R_m = (0.03)^2 I_3$, also following [21].

Neither adaptive covariance scaling nor NIS rejection is applied to the magnetic update. The purpose of the present methodology is to investigate reliability handling of the joint-centre acceleration measurement while retaining the original mechanism in [21] for introducing additional directional information during periods of low acceleration-based observability.

3-8-7 Tuned Configuration

The literature parameters were chosen either for other recordings than the ones used in this thesis, or were taken from the open source code from [21]. For the recorded-data evaluation, the following parameters were therefore tuned on the OpenSense recordings: the acceleration covariance R_a of the fixed-covariance estimators, the magnetic covariance R_m of MAJIC, the NIS thresholds, and for the adaptive estimators the covariance R_a , the window length N_w and the start-up delay T_d . The gyroscope process covariance Q_ω , the initial covariance, the lower bound $\alpha_{\min}$ and the observability threshold are unchanged. The objective is the overall angular RMSE over the complete recordings. The procedure, the tuned values and the results of the tuning are given in Section 5-3 and Appendix A-3.

3-8-8 Numerical Differentiation

Calculation of the joint-centre acceleration requires an estimate of angular acceleration. In the implementation, $\dot{y}_{\omega,t}^{S_i}$ is obtained numerically from the gyroscope measurements using a five-point derivative, following [25].

For the MAJIC variants, calculation of the observability metric additionally requires the derivative of the joint-centre acceleration. This derivative is obtained using the numerical differentiation procedure in [21].

The same differentiation procedures are used for all estimator variants so that differences in estimator behaviour cannot be attributed to different preprocessing of the inertial measurements.

Chapter 4

Simulation Study

This chapter addresses the first two subquestions of this thesis (Section 1-3). The first asks how adaptive acceleration covariance scaling and NIS-based rejection affect relative-orientation estimation under controlled changes in motion excitation and sensor uncertainty. The second asks whether they reduce the degradation caused by errors in the assumed IMU-to-joint-centre geometry. Recorded human motion cannot answer these questions on its own, because several sources of uncertainty are always present at the same time and the true relative orientation is only available through an optical reference. In simulation, the ground-truth orientations and sensor signals are known exactly, so that one source of uncertainty can be introduced at a time and its effect on the six estimator variants of Chapter 3 can be isolated.

Section 1-2 distinguished two reasons why the joint-centre acceleration update can fail to correct the relative orientation. Either the motion does not contain sufficient information, or the motion is informative but the update is statistically inconsistent with its assumed model. The simulations are organized along this distinction and along the origin of the inconsistency, as summarized in Table 4-1. For every source of uncertainty the same two questions are asked. First, do the reliability-handling mechanisms respond, that is, does the adaptive factor α_t increase and does the NIS test reject updates? Second, in which direction does the relative-orientation error of the adaptive and NIS variants change compared with that of the fixed-covariance estimator of the same family? The questions are separated deliberately, because a large innovation does not identify its cause, and a response of a mechanism therefore does not necessarily reduce the error.

How to read the results. The simulations are not an accuracy benchmark. The absolute error values depend on settings that were chosen for the purpose of the experiments (noise levels, motion, geometry and the size of the imposed errors) and are not values that can be expected in practice. It is also not the aim to determine which estimator variant is best. The simulations are used to observe what the mechanisms react to, as visible in α_t and in the rejection percentage, and whether this reaction becomes visible in the estimation error. The error is therefore discussed qualitatively and relative to the fixed-covariance estimator of the same family, and numerical values are quoted only to illustrate the size of an effect. All

Table 4-1: Overview of the simulation studies. All studies use the six estimator variants of Chapter 3 and 100 paired Monte Carlo realizations per condition.

Section	Source of uncertainty	Quantity varied
4-3	Insufficient motion excitation	Excitation level (low, medium, rich)
4-4	Accelerometer noise	Noise covariance multiplier λ_a
4-4	Gyroscope bias	Bias multiplier λ_δ
4-5	Joint-centre geometry	Length error κ_r , direction error θ_r
4-6	Sustained acceleration disturbance	Disturbance amplitude
4-7	Short accelerometer outliers	Number of outliers N_{out}
4-8	Inconsistent magnetic field (MAJIC)	Direction mismatch θ_m

simulations concern idealized conditions: two perfectly rigid segments connected by an ideal joint, one source of uncertainty varied at a time, and all other quantities at their nominal values. Whether the observed behaviour is also found on recorded motion is the subject of Chapter 5.

Section 4-3 first establishes what the mechanisms cannot do, namely create information that the motion does not contain. Sections 4-4 to 4-6 consider errors that affect an update which still carries information, from random (accelerometer noise) to persistent (gyroscope bias, joint-centre geometry, sustained disturbance). Section 4-7 considers the opposite temporal extreme, isolated outliers, for which hard rejection is designed. Section 4-8 considers the MAJIC estimators when the magnetic update, which neither mechanism modifies, is itself inconsistent. Section 4-9 repeats the conditions in which adaptive scaling lowered the error with a tuned fixed covariance and compares two error measures. Section 4-10 summarizes what the simulations show for the subquestions.

Unless stated otherwise, all simulation conditions are evaluated using paired Monte Carlo trials. Within each realization, the same generated motion, sensor errors, and initial conditions are supplied to all applicable estimator variants. Differences between variants can therefore be attributed to the estimator formulation rather than to different random realizations.

4-1 Simulation Framework

4-1-1 General Setup

A synthetic two-segment rigid-body system was developed in MATLAB. The two rigid segments are connected by a three-degree-of-freedom rotational joint, with one simulated IMU rigidly attached to each segment. All signals are generated at $f_s = 100$ Hz and each trial has a duration of $T = 600$ s unless specified otherwise, which equal the sampling frequency and the maximum duration of the recorded data used in Chapter 5. For each simulation condition, $N_{\text{MC}} = 100$ independent Monte Carlo realizations are generated (a design choice that gives stable medians and interquartile ranges without a formal sample-size analysis). A deterministic base random seed is used such that all simulations are reproducible. Conditions that are compared directly are paired by retaining the same underlying random realization and changing only the investigated variable. The simulated system provides the ground-truth orientations $q_t^{GS_1}$ and $q_t^{GS_2}$ of both sensor frames. The corresponding ground-truth relative

orientation serves as the reference for the estimates $\hat{q}_t^{S_1 S_2}$ of the six estimator variants of Chapter 3.

4-1-2 Synthetic Motion and Sensor Measurements

The notation of Chapters 2 and 3 is used: superscripts denote the coordinate frame (G for the global frame and S_i for the frame of sensor $i \in \{1, 2\}$), and subscripts denote the type of quantity and the sample index t , so that tT_s is the time in seconds. Quantities that are only needed to generate the data are introduced below in the same style.

Each simulated trial is defined by three prescribed quantities in the global frame: the specific force of the joint centre, the angular velocity common to both segments, and their relative angular velocity. A sum of sinusoids is used because it gives smooth signals of which the derivatives are known analytically, and because the number, amplitude, frequency and directional richness of the components directly control the excitation. It is a design choice and is not meant to resemble a particular human movement. The dynamic component of the specific force is generated independently along the three global axes as

$$f_{\text{dyn},j,t}^G = \sum_{\ell=1}^{N_f} A_{j,\ell}^f \sin \left(2\pi \gamma_{j,\ell}^f tT_s + \phi_{j,\ell}^f \right), \quad j \in \{x, y, z\}, \quad (4-1)$$

where $T_s = 1/f_s$ is the sampling period, N_f is the number of sinusoidal components of the considered motion condition, and $A_{j,\ell}^f$, $\gamma_{j,\ell}^f$ and $\phi_{j,\ell}^f$ are the amplitude, frequency and phase of component ℓ along axis j . These parameters are sampled independently from ranges defined by the motion condition. The complete specific force of the joint centre is $f_{\text{jc},t}^G = -g^G + f_{\text{dyn},t}^G$ with $g^G = [0 \ 0 \ -9.81]^\top \text{ m/s}^2$, so that the constant part is the specific force at rest (see the accelerometer model in Chapter 2).

The angular velocities are generated with the same procedure. The common component $\omega_{\text{com},t}^G$ describes the angular motion shared by both segments and the relative component $\omega_{\text{rel},t}^G$ their angular motion relative to one another, both expressed in G , with angular accelerations $\dot{\omega}_{\text{com},t}^G$ and $\dot{\omega}_{\text{rel},t}^G$ obtained analytically. The angular velocity and angular acceleration of sensor i in its own frame are

$$\omega_t^{S_i} = R_t^{S_i G} \left(\omega_{\text{com},t}^G + \frac{(-1)^i}{2} \omega_{\text{rel},t}^G \right), \quad \dot{\omega}_t^{S_i} = R_t^{S_i G} \left(\dot{\omega}_{\text{com},t}^G + \frac{(-1)^i}{2} \dot{\omega}_{\text{rel},t}^G \right). \quad (4-2)$$

The orientations $R_t^{GS_i} = (R_t^{S_i G})^\top$, and thereby $q_t^{GS_i}$, follow from integrating $\omega_t^{S_i}$ from independently sampled random initial orientations.

The specific force of the joint centre in the frame of sensor i is $f_{\text{jc},t}^{S_i} = R_t^{S_i G} f_{\text{jc},t}^G$. Let $\bar{C}_t^{S_i} = [\omega_t^{S_i} \times]^2 + [\dot{\omega}_t^{S_i} \times]$ be the noise-free counterpart of the matrix $C_t^{S_i}$ of the measurement model in Chapter 3, which the estimators compute from the measured angular velocity. The measurements are then generated according to the sensor models of Chapter 2 as

$$y_{\omega,t}^{S_i} = \omega_t^{S_i} + \delta_{\omega,t}^{S_i} + e_{\omega,t}^{S_i}, \quad (4-3)$$

$$y_{a,t}^{S_i} = f_{\text{jc},t}^{S_i} - \bar{C}_t^{S_i} r_i^{S_i} + \delta_{a,t}^{S_i} + e_{a,t}^{S_i}, \quad (4-4)$$

$$y_{m,t}^{S_i} = R_t^{S_i G} m^G + e_{m,t}^{S_i}, \quad (4-5)$$

with the bias terms $\delta_{\omega,t}^{S_i}$ and $\delta_{a,t}^{S_i}$ constant in time and the noise terms zero-mean Gaussian, as specified in the next section. Without bias and noise, the accelerometer and gyroscope signals are exactly consistent with the joint-centre measurement model used by the estimator.

For the MAJIC estimators, the constant magnetic-field direction m^G is derived from Delft's local magnetic inclination $\beta_{\text{inc}} \approx 67.2^\circ$ and declination $\beta_{\text{dec}} \approx +2.56^\circ$ [16]. Resolving these angles into the $+x$ (East), $+y$ (North) and $+z$ (Up) axes yields

$$m^G = \begin{bmatrix} \cos(\beta_{\text{inc}}) \sin(\beta_{\text{dec}}) \\ \cos(\beta_{\text{inc}}) \cos(\beta_{\text{dec}}) \\ -\sin(\beta_{\text{inc}}) \end{bmatrix} \approx \begin{bmatrix} 0.0173 \\ 0.3872 \\ -0.9219 \end{bmatrix}.$$

Since only the direction is used, m^G has unit norm.

4-1-3 Simulation Parameters

The parameters of the estimators are those of Chapter 3. The parameters below concern only the generation of the data. The noise terms are $e_{\omega,t}^{S_i} \sim \mathcal{N}(0, \Sigma_\omega)$, $e_{a,t}^{S_i} \sim \mathcal{N}(0, \Sigma_a)$ and $e_{m,t}^{S_i} \sim \mathcal{N}(0, \Sigma_m)$ with $\Sigma_\omega = \sigma_\omega^2 I_3$, $\Sigma_a = \sigma_a^2 I_3$ and $\Sigma_m = \sigma_m^2 I_3$, and the biases are $\delta_{\omega,t}^{S_i} \sim \mathcal{N}(0, \sigma_{\delta_\omega}^2 I_3)$ and $\delta_{a,t}^{S_i} \sim \mathcal{N}(0, \sigma_{\delta_a}^2 I_3)$, drawn once per sensor and trial.

The standard deviations are $\sigma_\omega = 1 \times 10^{-3}$ rad/s and $\sigma_a = 1 \times 10^{-2}$ m/s². They are of the same order as the values measured in a stationary recording of a smartphone IMU (Appendix A-1: approximately 1.2×10^{-3} rad/s and 8×10^{-3} m/s²). The simulations are not meant to reproduce this specific sensor. The bias standard deviations $\sigma_{\delta_\omega} = 2 \times 10^{-3}$ rad/s and $\sigma_{\delta_a} = 1 \times 10^{-2}$ m/s² are design choices of the order of the noise level. Since the bias is multiplied by a factor in the studies that follow, the nominal value only sets the scale. The magnetometer noise is $\sigma_m = 0.1$ per component of the unit-norm field. This is a design choice, deliberately larger than the relative noise of the stationary recording (approximately 0.007), so that the magnetic information is noticeably noisy without dominating the field direction.

The nominal sensor-to-joint-centre vectors are

$$r_1^{S_1} = \begin{bmatrix} 0.06 & -0.04 & 0.28 \end{bmatrix}^\top \text{ m}, \quad r_2^{S_2} = \begin{bmatrix} -0.05 & 0.03 & -0.23 \end{bmatrix}^\top \text{ m},$$

with lengths of 0.29 m and 0.24 m, which are of the order of segment dimensions near a knee. The non-zero components along all three axes avoid an axis-aligned special case. These vectors are the true geometry in the data generation. When the estimator is supplied with different vectors, these are denoted $\hat{r}_i^{S_i}$, as for the estimated vectors in Chapter 5.

4-1-4 Estimator Settings in the Simulations

All estimators use the literature configuration of Section 3-8. No tuning is applied in the simulations, so that the response of the estimators to the controlled changes is not mixed with the effect of tuning. The tuned configuration is used for the recorded data in Chapter 5. Two settings are specific to the simulations. First, the estimators are initialized with identity orientations, whereas the true initial orientations are sampled randomly, so that every trial

starts with a large relative-orientation error. Second, NIS-based rejection is switched off during the first 10 s of every trial. In preliminary tests, rejection from the first sample frequently prevented convergence, because the large innovations caused by the initial orientation error were rejected together with the measurements needed to remove that error. The NIS is still computed during this interval. The adaptive factor α_t is active from the start of the trial, with the nominal covariance retained until the first innovation window is complete. The rejection threshold is the theoretical value $\epsilon_{\max} = \chi_{3,0.99}^2 = 11.3449$.

4-2 Performance Evaluation

The estimated relative orientation $\hat{q}_t^{S_1 S_2}$ is compared with the ground truth through the geodesic angular distance between the two unit quaternions,

$$e_{\theta,t} = 2 \arccos \left(\left| \left(\hat{q}_t^{S_1 S_2} \right)^\top q_t^{S_1 S_2} \right| \right), \quad (4-6)$$

where the absolute value accounts for the equivalence of q and $-q$. The angle is reported in degrees.

4-2-1 Convergence

A trial is considered converged when $e_{\theta,t} \leq 10^\circ$ continuously for at least 5 s. The convergence time t_{conv} is the beginning of the first interval satisfying this criterion. This definition was chosen after a sensitivity analysis of the angular-error threshold (5° to 15°) and the hold duration (2 to 20 s), reported in Appendix A-4 (computed on the recorded data with the literature parameters). It is a compromise between avoiding classification based on brief transient reductions in error and retaining sufficient converged trials for the evaluation after convergence. Acquisition is summarized by the percentage of trials that converge and by the distribution of t_{conv} among the converged trials.

4-2-2 Orientation Error After Convergence

For converged trials, the error after acquisition is quantified by the root-mean-square error from t_{conv} to the end of the trial,

$$e_{\text{RMSE}}^{\text{post}} = \sqrt{\frac{1}{N_{\text{post}}} \sum_{t=t_{\text{conv}}}^N e_{\theta,t}^2}, \quad (4-7)$$

where N_{post} is the number of samples from t_{conv} to the final sample N . Since t_{conv} differs between estimators and realizations, the evaluation interval differs as well, so $e_{\text{RMSE}}^{\text{post}}$ is read together with the convergence rate and time. The same expression evaluated over a fixed interval gives $e_{\text{RMSE}}^{\text{dist}}$ (the disturbed middle third of the trial, from 200 s to 400 s, with smooth onset and offset ramps) and $e_{\text{RMSE}}^{\text{after}}$ (the final third) for the studies with a time-limited disturbance. In most realizations the estimators have converged long before the disturbance starts (median convergence times of at most approximately 10 s), so these quantities are largely

independent of the acquisition phase. For the magnetic disturbance, a trial has *recovered* when $e_{\theta,t}$ remains below 10° for 5 s after the disturbance has ended, and the recovery time is measured from the end of the disturbance.

Performance metrics are summarized over the realizations by the median and the interquartile range, reported as median [P25, P75]. These statistics are used because some conditions produce skewed distributions or non-converged trials. As explained in the introduction of this chapter, the values of e_{RMSE} are used to compare the behaviour of the variants under the same condition, and not as accuracy figures.

4-2-3 Diagnostics of the Reliability-Handling Mechanisms

The response of the mechanisms is quantified by the median of α_t over the evaluation interval of the study (the disturbed interval for the disturbance studies) for the adaptive variants, and by the percentage of samples at which the acceleration update is omitted for the NIS variants, calculated from 10 s onwards (the moment rejection is active) or over the disturbed interval. Both are summarized by their median over the realizations. Since $R_{a,t} = \alpha_t R_a$, $\alpha_t = 1$ corresponds to the nominal covariance. Under nominal conditions with rich excitation, the median α_t is approximately 0.47 and the NIS test rejects approximately 0.02% of the updates, far below the 1% implied by the 99% threshold for a consistent filter. This indicates that the nominal covariance R_a is conservative with respect to the innovations of the simulated data, and these nominal values are the reference for the responses in the following sections.

4-3 Influence of Motion Excitation

The joint-centre acceleration update can only correct the relative orientation when the measured acceleration provides sufficiently independent directional information. This is a limit on the available information and not a matter of how reliable the measurement is, so it is fundamentally different from the sources of uncertainty in the remainder of this chapter. In [25] it was found that highly dynamic movements result in faster convergence than slow movements for the Weygers estimator. Here this is verified for the baseline Weygers and MAJIC estimators, and it is evaluated whether the adaptive and NIS-based variants change this dependence.

Three excitation conditions are considered, denoted *low*, *medium* and *rich*. They are generated with the procedure of Section 4-1-2 by progressively increasing the number, amplitude, frequency range and directional richness of the sinusoidal components (parameters in Appendix A-2). For the low and medium conditions, unequal axis-scaling factors reduce the directional richness, and their assignment to the global axes is randomly permuted per realization. The primary outcomes are convergence success and convergence time, together with $e_{\text{RMSE}}^{\text{post}}$ of the converged trials.

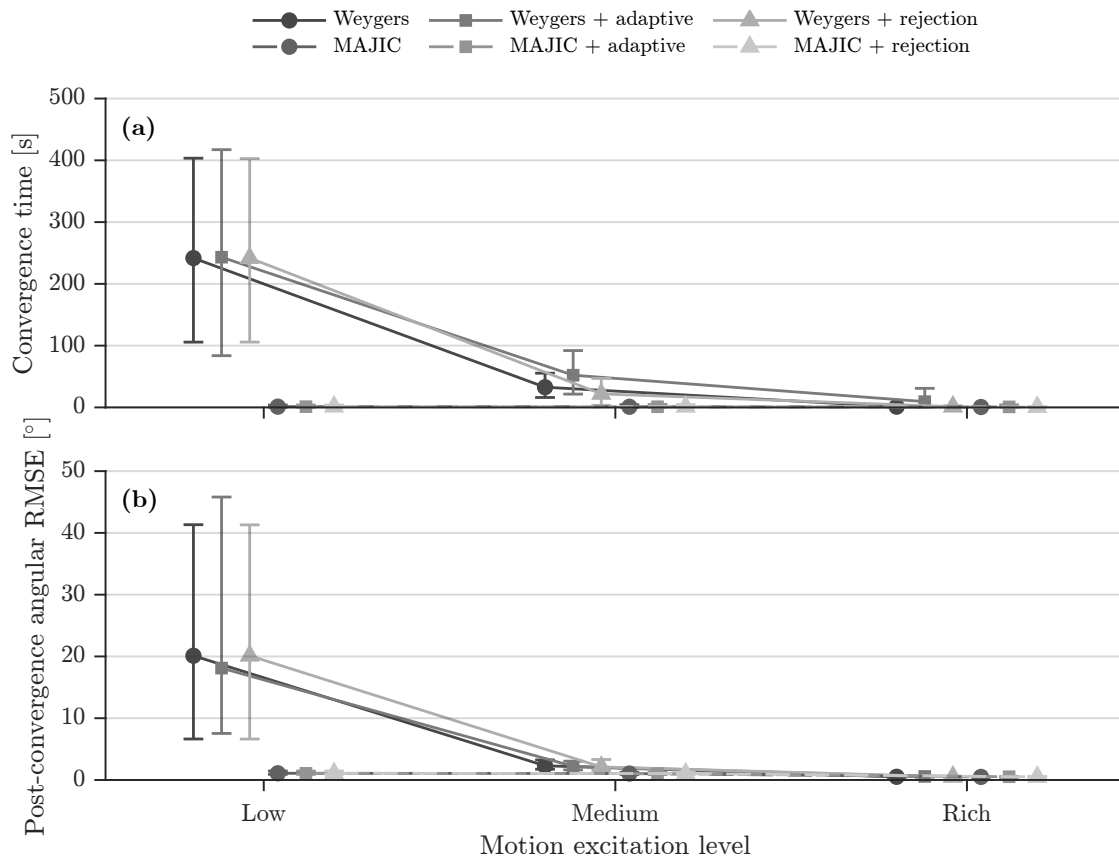


Figure 4-1: Convergence time and post-convergence angular RMSE as a function of motion excitation. Markers show the median over the Monte Carlo realizations and error bars indicate the interquartile range. Panel (a) shows convergence time and panel (b) shows post-convergence angular RMSE. Statistics are calculated over trials for which the corresponding metric is finite.

Results. Figure 4-1 shows that the three Weygers variants behave almost identically under low excitation: only about 30 of 100 realizations converge (30, 31 and 30), after a median of about 240s, whereas all realizations of the MAJIC variants converge within about a second. Under medium and rich excitation the Weygers variants converge in nearly all realizations, but the adaptive variant needs clearly more time (median 52s against 33s under medium and 9.4s against 1.2s under rich excitation), and Weygers + NIS converges in only 64 of 100 realizations under medium excitation. The MAJIC variants converge with a median of about one second in all conditions. After convergence, the error decreases with increasing excitation for all variants, and under rich excitation all six variants are at about 0.5° , so that their differences are small compared with the effect of excitation.

Interpretation. The dependence on excitation reported in [25] is reproduced for the baseline Weygers estimator, and neither mechanism removes it: under low excitation the convergence rate and error of the three Weygers variants are nearly the same. This is consistent with the mechanisms only modifying the weight of information that is present in the update. The MAJIC variants show that the magnetic update provides directional information when the acceleration-based observability is low.

The mechanisms do affect acquisition. During acquisition a large innovation is caused mainly by the initial orientation error and not by an unreliable measurement, and neither mechanism can distinguish these causes. Increasing $R_{a,t}$ or rejecting the update then reduces the influence of the update that is needed to remove the initial error. For rejection, the lower convergence rate under medium excitation is consistent with updates being rejected after the first 10 s in realizations that have not converged by then, although the data do not isolate this cause. The remaining studies use rich excitation, so that the information content of the motion is not a confounding factor, except for the magnetic study in which the observability is varied deliberately.

4-4 Sensor Errors: Accelerometer Noise and Gyroscope Bias

The first sources of uncertainty that affect an update which still carries information originate in the sensors. Accelerometer noise enters the joint-centre acceleration directly through $y_{a,t}^{S_i}$. It is zero mean and uncorrelated in time, and therefore increases the innovation randomly. Gyroscope bias acts differently. It drives a persistent drift in the propagated orientation, and it also enters $C_t^{S_i}$, so that it produces an innovation that is systematic and persists. Both change the innovation statistics in a way that α_t and the NIS test can detect. The question is whether acting on this detection changes the error, and whether this differs between an error in the measurement (noise) and an error that mainly affects the prediction (bias).

Setup. Rich excitation is used and the estimator tuning is unchanged. For the noise study the accelerometer noise covariance used to generate the data is $\lambda_a \Sigma_a$ with $\lambda_a \in \{0.1, 1, 10, 100\}$, while R_a is kept nominal. For the bias study the bias of sensor i is $\delta_{\omega,t}^{S_i} = \lambda_\delta \sigma_{\delta_\omega} z_\omega^{S_i}$ with $z_\omega^{S_i} \sim \mathcal{N}(0, I_3)$ shared across levels and $\lambda_\delta \in \{0, 1, 5\}$, which gives per-axis standard deviations of 0, 2×10^{-3} and 1×10^{-2} rad/s. The multiplier ranges are design choices that span from no error to levels at which the response of the mechanisms is clearly visible.

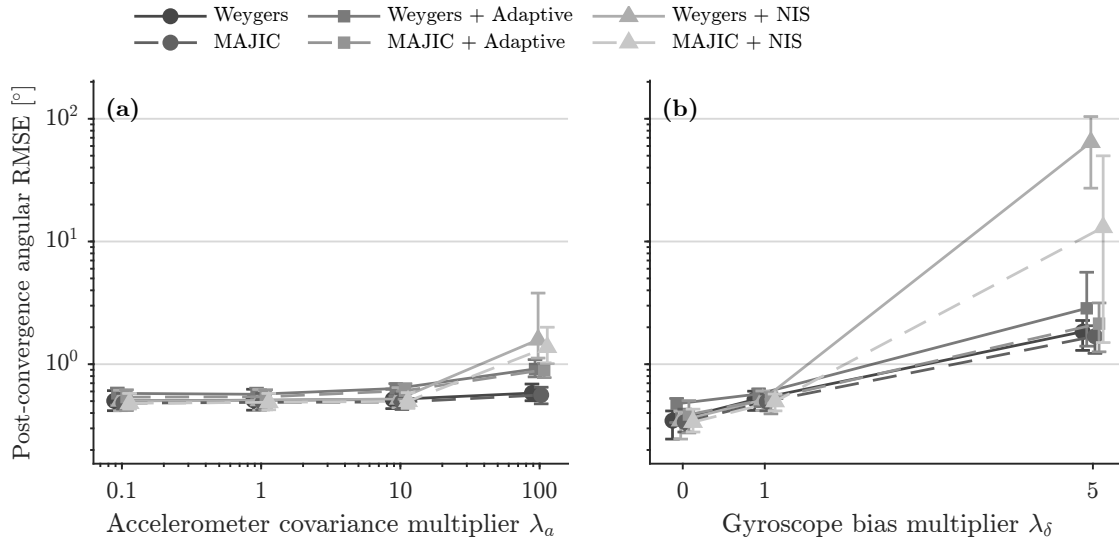


Figure 4-2: Post-convergence angular RMSE under increased sensor uncertainty. Panel (a) shows the effect of the accelerometer covariance multiplier λ_a , while panel (b) shows the effect of the gyroscope bias multiplier λ_δ . Markers indicate the median over the Monte Carlo realizations and error bars show the interquartile range. The vertical axis is shown on a logarithmic scale.

Results. Figure 4-2 shows $e_{\text{RMSE}}^{\text{post}}$. All variants converge in at least 97 of 100 realizations in every condition. For accelerometer noise, all variants remain at about 0.5° up to $\lambda_a = 10$, with the adaptive variants slightly above their baselines. At $\lambda_a = 100$ the baselines hardly change (about 0.6°), whereas the adaptive variants rise to about 0.9° and the NIS variants to about 1.4° to 1.6° . The mechanisms do respond. The median α_t increases from 0.47 at the nominal level to 1.2 at $\lambda_a = 10$ and 8.6 at $\lambda_a = 100$, and the rejection percentage from 0.02% to 2.6% and approximately 75%.

For gyroscope bias the baselines increase moderately, from about 0.5° at the nominal bias to about 1.8° (Weygers) and 1.7° (MAJIC) at $\lambda_\delta = 5$, and the adaptive variants are slightly above them (2.9° and 2.1°). The NIS variants diverge from their baselines: 64° [27° , 104°] for Weygers + NIS and 13° [1.5° , 50°] for MAJIC + NIS. The adaptive factor increases only moderately (0.46 to 1.45 and 1.25), whereas the NIS test rejects 95% [58%, 99%] (Weygers + NIS) and 54% [3%, 90%] (MAJIC + NIS) of the updates; without bias and at the nominal bias the rejection stays below 0.02%.

Interpretation. Both mechanisms react to both sensor errors, but this reaction does not lower the error relative to the fixed-covariance estimator. For white noise the update remains informative, and a plausible explanation is that its zero-mean error averages out over successive samples whereas the gyroscope-driven propagation remains accurate between updates, so that down-weighting or removing updates only discards information. For gyroscope bias, the acceleration update is what continuously counteracts the drift. When the NIS test removes most updates, as for Weygers + NIS, the drift is no longer counteracted, which is consistent with the divergence of that variant. For MAJIC + NIS a possible reason for the smaller effect is that the magnetic update remains available when the acceleration-based observability is low. The adaptive factor rises only moderately and the update stays active, consistent with

the error of the adaptive variants staying close to the baselines. An increase in the innovation therefore does not by itself indicate that the corresponding measurement should be trusted less, since here the measurement is the means by which the persistent error is corrected.

4-5 Joint-Centre Geometry Mismatch

The joint-centre acceleration depends on the vectors $r_1^{S_1}$ and $r_2^{S_2}$ from each sensor origin to the joint centre. In practice these are not measured directly, but estimated or assumed, for example using the method of [19]. An error in these vectors changes $C_t^{S_i} r_i^{S_i}$ and therefore creates an error in the joint-centre acceleration that depends on the angular motion and does not vanish with time. In contrast to sensor noise, the resulting inconsistency is systematic, while the acceleration remains informative about the relative orientation. This makes it the main case for the second subquestion.

Setup. The data are always generated with the true vectors and only the vectors supplied to the estimator are perturbed. Rich excitation is used. For the length error, $\hat{r}_i^{S_i} = (1 + \kappa_r) r_i^{S_i}$ with $\kappa_r \in \{-0.20, -0.10, 0, +0.10, +0.20\}$. For the direction error, $\hat{r}_i^{S_i} = \mathcal{R}(\theta_r, n_i) r_i^{S_i}$, the rotation of $r_i^{S_i}$ by $\theta_r \in \{0^\circ, 5^\circ, 15^\circ, 30^\circ\}$ about a unit axis n_i perpendicular to it, which preserves its length; the axes are drawn once per realization and reused for all levels. These ranges are design choices. They displace the assumed joint centre by roughly 2 to 6 cm for the length errors and 2 to 15 cm for the direction errors.

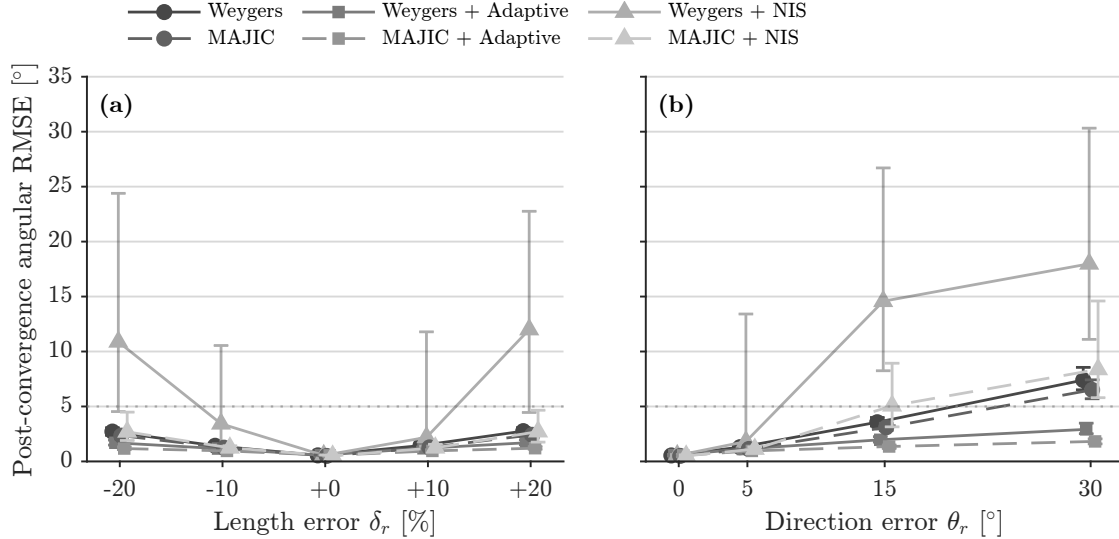


Figure 4-3: Post-convergence angular RMSE under errors in the assumed sensor-to-joint-centre vectors. Panel (a) shows the effect of the relative length error κ_r , while panel (b) shows the effect of the direction error θ_r . Markers indicate the median over the Monte Carlo realizations and error bars show the interquartile range. Values exceeding the displayed vertical-axis range are annotated with their corresponding median RMSE.

Results. Figure 4-3 shows that the error of all variants increases with the mismatch. For the baselines it reaches about 2.4° to 2.8° at $\kappa_r = \pm 0.20$ and 6.5° to 7.4° at $\theta_r = 30^\circ$. The

adaptive variants have a lower error than the baselines in every mismatch condition, by a factor of 1.15 to 3.6 (for example from 7.4° to 2.9° for Weygers and from 6.5° to 1.8° for MAJIC at $\theta_r = 30^\circ$), but their error still increases with the mismatch. The NIS variants do not lower the error: Weygers + NIS is higher than Weygers in every mismatch condition (for example 18° at $\theta_r = 30^\circ$, where only 83 of 100 realizations converge), and MAJIC + NIS is within 3% of MAJIC at $\kappa_r = \pm 0.10$ and $\theta_r = 5^\circ$ and higher otherwise.

The mechanisms respond strongly and symmetrically to positive and negative length errors. The median α_t is between 4 and 5 at $\kappa_r = \pm 0.10$ and about 20 (Weygers + Adaptive) and 18 (MAJIC + Adaptive) at ± 0.20 , and for direction errors 4 at 5° , 32 to 35 at 15° and 123 to 142 at 30° . The NIS test rejects 55% to 65% of the updates at $\kappa_r = \pm 0.10$ and 89% to 97% at ± 0.20 , 47% to 55% at $\theta_r = 5^\circ$ and more than 95% from 15° onwards. The adaptive variants also converge more slowly as the mismatch grows (the median convergence time of Weygers + Adaptive increases from 12 s without mismatch to 72 s at $\theta_r = 30^\circ$, against 1 s to 8 s for Weygers).

Interpretation. A geometry error is the first case in which a mechanism consistently lowers the error relative to the fixed-covariance estimator. The update stays informative but is systematically inconsistent with the assumed model, and the adaptive method reduces its weight without removing it. It does not remove the model error, and the lower weight is accompanied by slower acquisition. The NIS test instead removes almost all updates already at moderate mismatches. The estimator then relies on gyroscope propagation, and for MAJIC on the magnetic update, and the error is equal to or larger than that of the fixed-covariance estimator. The difference between the mechanisms is consistent with continuous down-weighting versus binary removal of an update that is informative but biased.

4-6 Sustained Acceleration-Model Disturbance

The rigid-body model of the joint-centre acceleration is an idealization. In recorded motion, effects such as relative movement between sensor and segment can add acceleration components that this model does not describe, and that vary smoothly over time instead of being isolated. The question is how the mechanisms behave during a sustained but temporary degradation of the acceleration model and whether the estimators recover afterwards. The study is not meant to reproduce a specific physical effect.

Setup. Rich excitation is used. A disturbance $d_t^{S_i} = \sum_{\ell=1}^2 A_{i,\ell}^d \sin(2\pi\gamma_{i,\ell}^d tT_s + \phi_{i,\ell}^d) u_{i,\ell}$ is added to the accelerometer measurement $y_{a,t}^{S_i}$ of each sensor during the middle third of the trial, where $u_{i,\ell}$ is a random unit vector, $\gamma_{i,\ell}^d \sim \mathcal{U}(0.2, 2)$ Hz (periods of 0.5 to 5 s) and $A_{i,\ell}^d$ is drawn from $\mathcal{U}(0.1, 0.4)$ m/s² (moderate) or $\mathcal{U}(0.5, 1.2)$ m/s² (strong). These are design choices, with amplitudes comparable to the dynamic specific-force amplitudes of the rich excitation (0.4 to 1.5 m/s², Appendix A-2). The frequencies overlap with, and extend above, those of the simulated motion (0.1 to 0.8 Hz), so that the disturbance cannot be separated from the motion by frequency alone. The disturbance is added to the accelerometer only. Effects on the orientation of the sensor, which relative movement of the sensor would also cause, are not simulated.

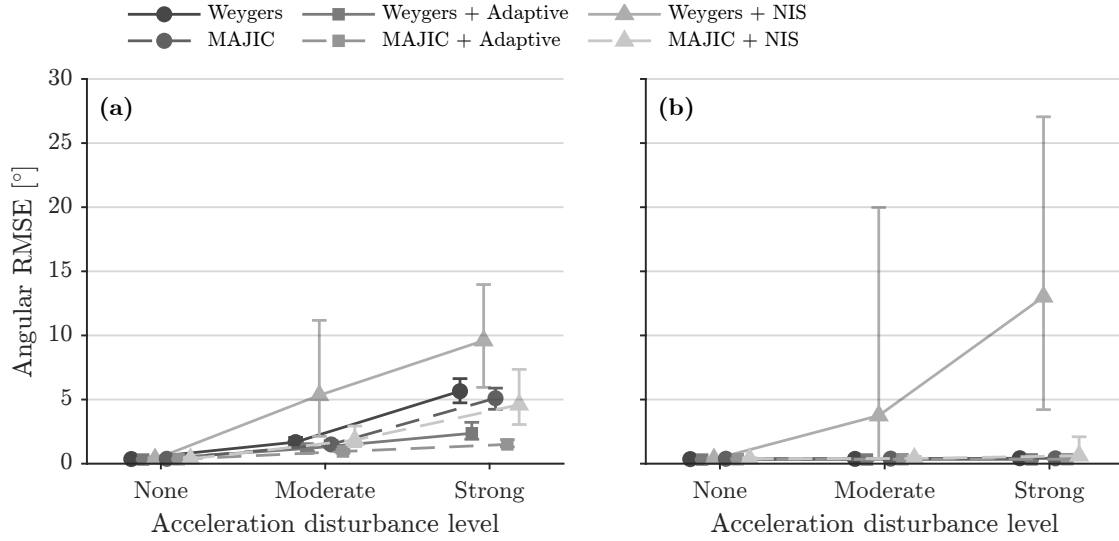


Figure 4-4: Angular RMSE during and after a sustained acceleration disturbance of increasing amplitude. Panel (a) shows the RMSE during the disturbed interval, while panel (b) shows the RMSE after the disturbance has ended. Markers indicate the median over the Monte Carlo realizations and error bars show the interquartile range.

Results. Figure 4-4 shows $e_{\text{RMSE}}^{\text{dist}}$ and $e_{\text{RMSE}}^{\text{after}}$. Without disturbance all variants are between 0.3° and 0.4° . During the strong disturbance the baselines rise to about 5° to 5.7° , the adaptive variants to 2.4° (Weygers) and 1.5° (MAJIC), Weygers + NIS to 9.6° , and MAJIC + NIS to 4.6° , which is similar to its baseline given the overlapping interquartile ranges. The moderate disturbance shows the same order. The mechanisms respond strongly: the median α_t is 17 (moderate) and 183 (strong), and the NIS test rejects 88% to 94% (moderate) and 99% (strong) of the updates. After the disturbance has ended, the error of all variants returns to at most about 0.4° , except for Weygers + NIS, which retains 3.8° [0.3° , 20°] after the moderate and 13° [4.2° , 27°] after the strong disturbance, and MAJIC + NIS after the strong disturbance (0.6°).

Interpretation. The result is similar to that for the geometry error. The update remains partly informative but is temporarily inconsistent with the model, and adaptive scaling lowers the error relative to the fixed-covariance estimator without removing the update. The NIS test removes almost all updates during the disturbance. For Weygers + NIS no other correcting information is available and the accumulated error remains after the disturbance has ended. A likely explanation is that this error itself produces large innovations that keep the gate closed, but the summary statistics do not verify this. MAJIC + NIS is less affected, consistent with the magnetic update continuing to provide a reference while the field is consistent. Section 4-8 considers what happens when it is not.

4-7 Short Accelerometer Outliers

Section 3-5 introduced statistical gating as a way to protect the estimator against isolated inconsistent measurements. Short accelerometer outliers, for example caused by impacts, are

the case for which this strategy is designed. This study considers the opposite extreme of the sustained disturbance and shows how both mechanisms behave when the inconsistency is isolated in time.

Setup. Rich excitation is used. For $N_{\text{out}} \in \{0, 10, 50\}$ events, event k adds the vector $\xi_k = A_k^{\text{out}} u_k$ to the accelerometer measurement of a randomly chosen sensor for one to three samples (10 to 30 ms), with $A_k^{\text{out}} \sim \mathcal{U}(2, 6) \text{ m/s}^2$ and a random unit vector u_k . These design choices make the outliers much larger than the nominal noise ($\sigma_a = 0.01 \text{ m/s}^2$) but short. Events are separated by at least 1.5 s and do not occur during the first 15 s or the last 5 s, and the conditions use nested subsets of the same events. In addition to $e_{\text{RMSE}}^{\text{post}}$, the error is evaluated from 0.1 s before to 1 s after each outlier onset, and the percentages of injected outlier samples and of nominal samples rejected by the NIS test are determined.

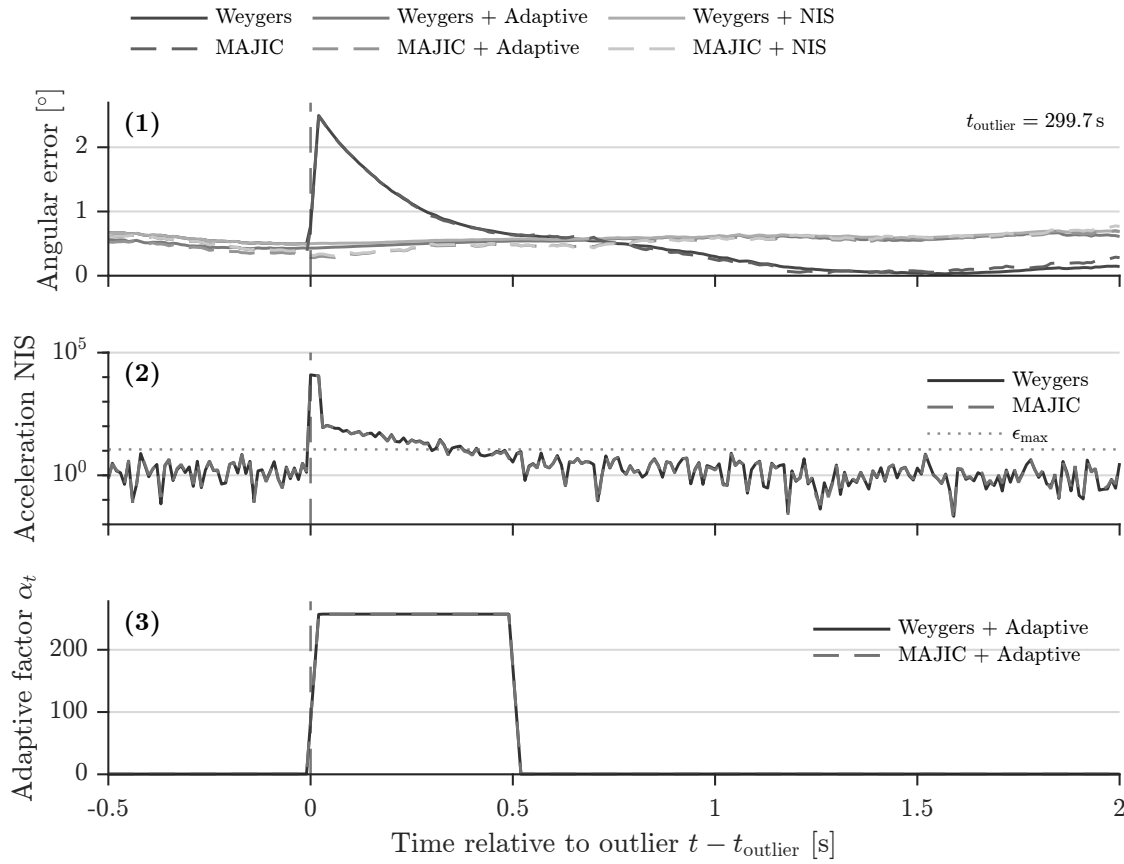


Figure 4-5: Estimator response around a representative short accelerometer outlier event. Panel (1) shows the angular error of the six estimator variants, panel (2) the acceleration NIS with the rejection threshold ϵ_{max} , and panel (3) the adaptive factor α_t for the adaptive variants. Time is shown relative to the outlier occurrence, with the dashed vertical line indicating $t - t_{\text{outlier}} = 0$.

Results. Figure 4-5 shows one representative outlier. It produces an NIS of 1.2×10^4 in the baseline estimators, about a thousand times the threshold of 11.3, and the error of Weygers and MAJIC rises from about 0.5° to a peak of 2.5° within 20 ms. For the adaptive and NIS variants the error stays below 0.8° . The adaptive factor increases from 0.48 to a peak of 257

within 50 ms and remains above half of its peak for 0.50 s, the length of the innovation window of $N_w = 50$ samples, whereas the NIS variants omit the update only for the affected samples.

Over all events, the median peak error in the window is 2.4° to 3.3° for the baselines and about 0.7° to 1.2° for the NIS variants and for MAJIC + Adaptive. Weygers + Adaptive has a similar median (1.1° to 1.2°) but an upper quartile of 42.6° , most likely because events start at 15 s whereas its upper-quartile convergence time is approximately 38 s. The NIS test rejects the injected outlier samples in all realizations and rejects a median of 0.017% of the nominal samples (with a worst case of 49% in the worst realization of Weygers + NIS, which is also present without outliers). $e_{\text{RMSE}}^{\text{post}}$ hardly depends on the number of outliers: from zero to 50 outliers it increases from 0.50° to 0.57° for Weygers and from 0.51° to 0.55° for MAJIC, remains at about 0.58° and 0.54° for the adaptive variants, and does not change for the NIS variants.

Interpretation. For isolated outliers, both mechanisms prevent the error peak that the baselines show. Hard rejection achieves this by omitting only the affected samples, so that no information is lost otherwise. The adaptive method suppresses the peak as well, but lowers the weight of the update for the whole innovation window of 0.5 s, of which only the first one to three samples are affected (about 25 s of the trial with 50 outliers). The differences in $e_{\text{RMSE}}^{\text{post}}$ are at most approximately 0.1° , so this study demonstrates how the mechanisms act on an isolated inconsistency and not a large difference in error. The relevant result is the contrast with the persistent sources in the preceding sections: the same rejection that leaves the estimate unaffected by isolated outliers led to errors above those of the fixed-covariance estimator under gyroscope bias, accelerometer noise and geometry errors, and, for Weygers, under a sustained disturbance.

4-8 Inconsistent Magnetic Field

Neither mechanism modifies the magnetic update of MAJIC. This study nevertheless investigates how the reliability handling of the acceleration update interacts with it. The magnetic update is only requested when the acceleration-based observability is low, so it takes over the role of the acceleration update when that update is least informative. It assumes that both sensors observe the same magnetic-field direction, which is violated near ferromagnetic material. Section 1-2 stated that applying the mechanisms in both estimators makes it possible to determine whether they remain useful when another correction source is available. This study considers the case in which that other source is inconsistent. The Weygers variants do not use the magnetometer and serve as a reference.

Setup. Sensor 1 observes the nominal field direction m^G throughout. During the middle third of the trial the direction observed by sensor 2 is rotated by $\theta_m \in \{0^\circ, 15^\circ, 30^\circ, 60^\circ, 90^\circ, 180^\circ\}$ about a random axis perpendicular to the field, with 1 s onset and offset ramps. These angles are not meant to represent realistic distortions. They are chosen as a stress test, from a small mismatch up to the extreme case in which the two sensors observe opposite field directions, so that the response of the mechanisms becomes clearly visible. The study is not meant to

quantify the error that a realistic distortion would cause, and the large errors at large angles indicate what the estimators react to and not the errors to be expected.

The motion is rich in the first and final third. In the middle third, two conditions are used. In the low-observability condition the specific force of the joint centre keeps a constant direction in G , $f_{jc,t}^G = [0, 0, 9.81 + \sin(2\pi 0.25 t T_s)]^\top \text{ m/s}^2$, while the angular motion remains rich. With a fixed global direction of the specific force the acceleration-based observability remains low even though the sensor frames rotate, which is the situation in which MAJIC relies on the magnetic update. In the high-observability condition the direction of the specific force varies continuously in three dimensions. Because the measured observability fluctuates due to sensor noise, the magnetic update is requested in a median of 85% of the disturbed interval in the low-observability condition and in 42% [35%, 51%] in the high-observability condition. $e_{\text{RMSE}}^{\text{dist}}$, $e_{\text{RMSE}}^{\text{after}}$ and the recovery metrics defined above are used.

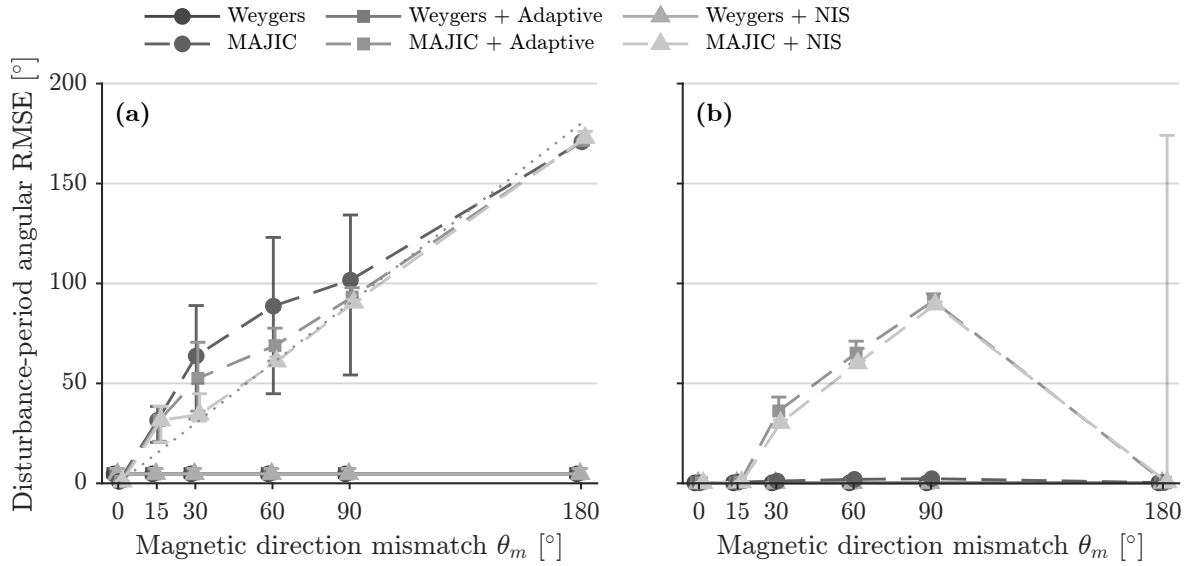


Figure 4-6: Disturbance-period angular RMSE under an inconsistent magnetic field. Panel (a) shows the low-observability condition and panel (b) the high-observability condition. Markers indicate the median over the Monte Carlo realizations and error bars show the interquartile range.

Results. Figure 4-6 shows $e_{\text{RMSE}}^{\text{dist}}$. In the low-observability condition the Weygers variants are unaffected by the mismatch (about 4.7°), since they do not use the magnetometer. The MAJIC variants reach 0.94° without mismatch, but already about 31.5° at 15° , and their error then keeps growing with the mismatch to between 34° and 64° at 30° and 171° to 173° at 180° , so it is of the order of, or larger than, the imposed mismatch. The differences between the three MAJIC variants are small compared with this effect, and the NIS test rejects more than 90% of the acceleration updates from 30° onwards. After the disturbance, the baseline MAJIC recovers in all realizations (median recovery time at most 4.6 s), whereas the recovery rate of MAJIC + Adaptive drops to 68% at 180° and that of MAJIC + NIS is 47% to 66% for mismatches up to 90° and 17% at 180° .

In the high-observability condition the baseline MAJIC stays below 2.4° for mismatches up to 90° (against 0.23° without mismatch). The adaptive and NIS variants are close to it at 15° ,

but show a sharp transition between 15° and 30° , to about 30° to 36° at 30° and about 90° at 90° , with a lower recovery rate for MAJIC + NIS (90%, 74% and 60% at 30° , 60° and 90°). The NIS test rejects 99% of the acceleration updates over this range. At 180° the median error of all MAJIC variants returns to below 0.6° , although the upper quartile of MAJIC + NIS is 174° .

Interpretation. When the acceleration-based observability is low, the acceleration update contains little information about the relative orientation, and the magnetic update, which is requested most of the time, is the dominant correction. An inconsistent field then produces large errors in all three MAJIC variants, and no modification of the acceleration update removes this effect, which is consistent with the conclusion of Section 4-3 that the mechanisms cannot create missing information. The magnetic update is also only beneficial while the field is consistent: without mismatch the MAJIC variants are below the Weygers variants, and with a mismatch they are above.

The high-observability condition shows an interaction that is not visible in the preceding studies. With an informative acceleration update, the fixed covariance keeps that update influential enough to counteract the errors introduced by the magnetic update, whereas the adaptive and NIS variants do not. A plausible explanation is that a magnetic update that perturbs the orientation increases the acceleration innovation, and that both mechanisms respond by reducing the weight of the acceleration update (for NIS supported by its rejection percentage), which removes the path that would otherwise correct the error. The innovation then increases because the state estimate is wrong and not because the acceleration measurement is unreliable, so reducing its weight is the wrong response, as anticipated in Section 1-1. The mechanisms thus cannot distinguish an unreliable acceleration measurement from an error that originates elsewhere. The return to a low error at 180° is non-monotonic in the mismatch. For two unit vectors separated by θ_m , the innovation norm $\|\nu_{m,t}\| = 2\sin(\theta_m/2)$ is maximal at 180° , whereas the first-order correction $\|H_{m,t}^\top \nu_{m,t}\| = \sqrt{2}|\sin \theta_m|$ vanishes, so that the magnetic update then has little first-order effect. This is a possible explanation only for this condition, since in the low-observability condition the error at 180° is the largest.

4-9 Control Experiment with a Tuned Fixed Covariance

In the preceding sections all estimators use the literature configuration, in which the fixed acceleration covariance is $R_a = 2.5 \times 10^{-3} I_3$. If this value is too small for a given condition, a method that increases the covariance can improve on the baseline only because it compensates for this mismatch, and not because it adapts to the changing reliability of the measurement. The recorded data of Chapter 5 showed that a much larger covariance (36.5 times the literature value) is better for the fixed-covariance estimator. This section therefore repeats the conditions in which adaptive scaling lowered the error, a direction error of the joint-centre vectors of 30° , a length error of +20% and a strong sustained disturbance, with a stronger baseline, and it compares the conclusions for two error measures.

Setup. The same data generator, motion and seeds are used as in the original studies, with 50 instead of 100 Monte Carlo realizations per condition, for both estimator families. Besides

the fixed covariance with the literature R_a , three fixed-covariance baselines with R_a multiplied by 10, 36.5 (the value tuned on the recordings) and 100 are evaluated, and the best of these four for each condition is used as the reference. Because this choice is made on the same realizations, it favours the baseline. Adaptive scaling is evaluated with the literature settings and with the tuned settings of the recordings (R_a multiplied by 1.54, $N_w = 0.3$ s, start-up delay 10 s). NIS rejection is evaluated with the theoretical threshold and with the tuned setting of the recordings (R_a multiplied by 36.5 and threshold 4000 for the magnetometer-free estimator, 751 for MAJIC). For MAJIC, the tuned variants also use the tuned magnetic covariance. The error is the overall angular RMSE over the complete trial, and, for comparison with the preceding sections, the median RMSE after convergence.

4-9-1 Results

Table 4-2: Magnetometer-free estimator: mean [median] overall angular RMSE in degrees over 50 Monte Carlo realizations. The scale factor of the best fixed covariance is given in brackets.

Estimator	Nominal	Direction 30°	Length +20%	Disturbance (strong)
Fixed, literature R_a	3.2 [2.5]	10.8 [9.7]	6.1 [5.3]	5.0 [4.5]
Fixed, $R_a \times 10$	3.1 [2.6]	9.3 [8.3]	6.3 [5.4]	4.3 [3.3]
Fixed, $R_a \times 36.5$ (tuned)	3.3 [2.7]	9.9 [7.9]	5.8 [4.5]	4.3 [3.4]
Fixed, $R_a \times 100$	3.3 [2.7]	11.2 [8.3]	6.2 [5.1]	3.6 [3.1]
Best fixed covariance	3.1 [2.6] ($\times 10$)	9.3 [8.3] ($\times 10$)	5.8 [4.5] ($\times 36.5$)	3.6 [3.1] ($\times 100$)
Adaptive, literature settings	12.6 [4.4]	29.2 [21.5]	25.7 [19.8]	14.2 [5.3]
Adaptive, tuned settings	3.2 [2.4]	8.0 [6.8]	5.8 [4.8]	4.2 [3.0]
NIS, theoretical threshold	3.2 [2.5]	25.3 [19.3]	18.2 [15.4]	14.3 [12.6]
NIS, tuned thresholds	3.3 [2.7]	9.9 [7.9]	5.8 [4.5]	4.3 [3.4]

Table 4-3: MAJIC: mean [median] overall angular RMSE in degrees over 50 Monte Carlo realizations. The scale factor of the best fixed covariance is given in brackets.

Estimator	Nominal	Direction 30°	Length +20%	Disturbance (strong)
Fixed, literature R_a	1.9 [1.2]	7.9 [7.3]	3.7 [2.9]	3.7 [3.2]
Fixed, $R_a \times 10$	1.4 [1.3]	3.2 [3.0]	1.9 [1.5]	2.0 [1.7]
Fixed, $R_a \times 36.5$ (tuned)	1.7 [1.5]	2.3 [1.8]	1.8 [1.6]	1.7 [1.5]
Fixed, $R_a \times 100$	1.9 [1.6]	2.0 [1.8]	1.8 [1.6]	1.7 [1.6]
Best fixed covariance	1.4 [1.3] ($\times 10$)	2.0 [1.8] ($\times 100$)	1.8 [1.6] ($\times 36.5$)	1.7 [1.5] ($\times 36.5$)
Adaptive, literature settings	4.0 [1.2]	11.0 [3.1]	11.7 [2.2]	7.0 [1.6]
Adaptive, tuned settings	1.7 [1.0]	2.9 [2.4]	2.2 [1.7]	2.0 [1.7]
NIS, theoretical threshold	1.9 [1.2]	11.6 [8.1]	7.2 [3.8]	6.4 [3.7]
NIS, tuned thresholds	1.7 [1.5]	2.3 [1.8]	1.8 [1.6]	1.7 [1.5]

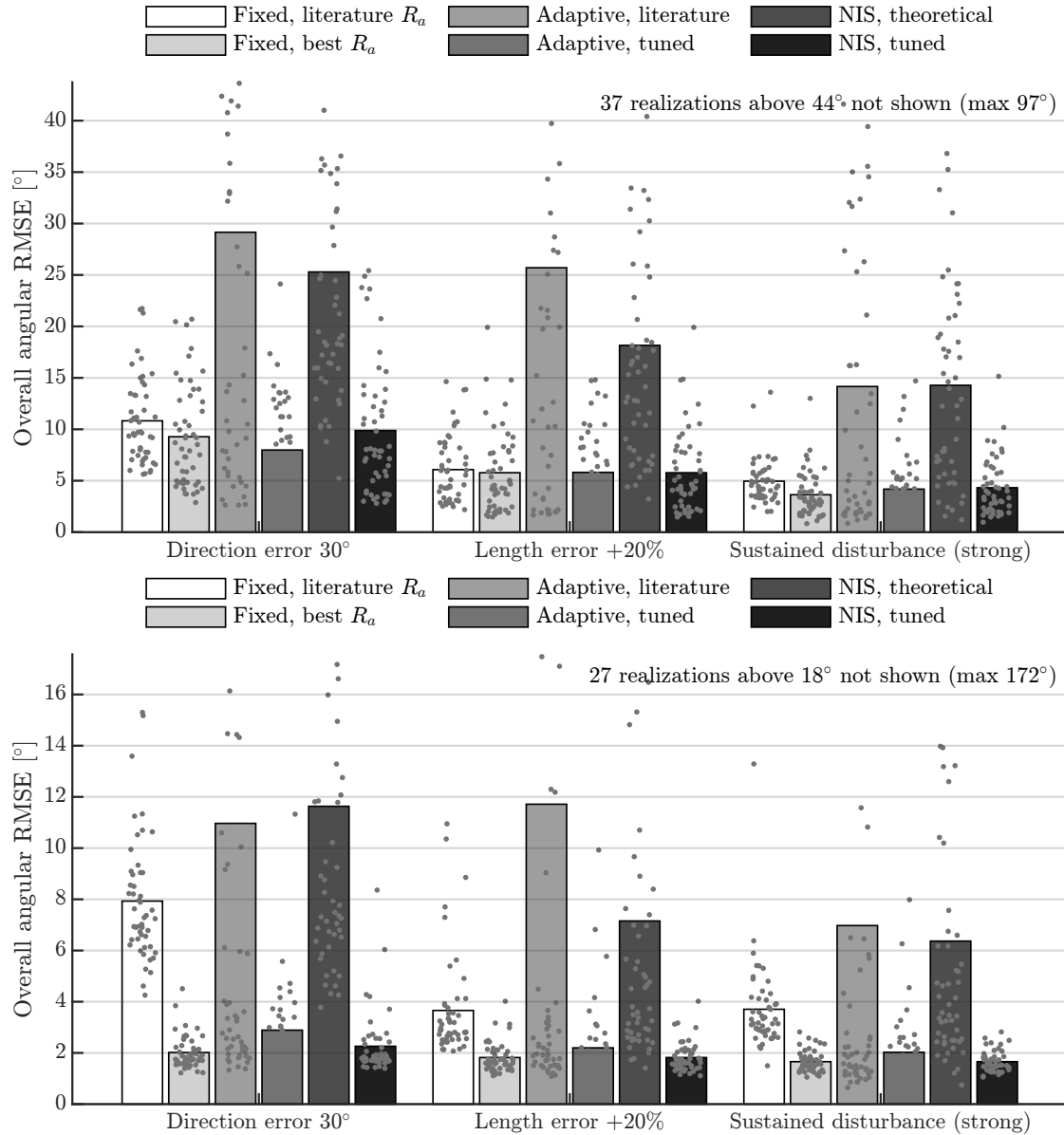


Figure 4-7: Overall angular RMSE in the three stressed conditions for the magnetometer-free estimator (top) and MAJIC (bottom). Bars show the mean over the 50 Monte Carlo realizations and dots the individual realizations.

Tables 4-2 and 4-3 and Figure 4-7 give the overall error of all variants. For the magnetometer-free estimator, the best fixed covariance (scale factors 10, 36.5 and 100) lowers the error of the fixed covariance with the literature R_a in all three conditions. Adaptive scaling with the literature settings has a higher error than the best fixed covariance in all conditions, including without any error (12.6° against 3.2° for the literature baseline), and with the tuned settings it is close to it. The paired difference between the best fixed covariance and adaptive scaling with the tuned settings (positive when adaptive scaling has the lower error) is +1.29° (95% interval [+0.52°, +2.14°]), -0.04° ([-0.75°, +0.63°]) and -0.55° ([-1.34°, +0.16°]) for the direction error, the length error and the disturbance. With the literature settings the differences are

-19.88° , -19.93° and -10.52° .

For MAJIC, adaptive scaling with the literature and with the tuned settings has a higher error than the best fixed covariance in all three conditions. The paired differences of adaptive scaling with the tuned settings relative to the best fixed covariance are -0.87° ($[-1.21^\circ, -0.59^\circ]$), -0.38° ($[-0.75^\circ, -0.09^\circ]$) and -0.37° ($[-0.72^\circ, -0.09^\circ]$), and with the literature settings -8.95° , -9.89° and -5.32° .

NIS rejection with the tuned thresholds is identical or very close to the fixed covariance with the tuned R_a in all conditions, because the tuned thresholds (4000 and 751) are rarely exceeded by the innovations of the simulated data. NIS rejection with the theoretical threshold gives 25.3° , 18.2° and 14.3° for the magnetometer-free estimator, which is larger than the fixed covariance in all three conditions.

4-9-2 Error over Time

Figure 4-8 shows the median RMSE in consecutive bins of 25 s for the fixed covariance with the literature and with the best R_a and for adaptive scaling with the literature and the tuned settings. It separates the acquisition at the start of the trial from the later behaviour, without a convergence criterion. In all conditions every estimator starts with a large error in the first bin. For the magnetometer-free estimator with the literature settings, adaptive scaling needs about 100 to 150 s to reach the level of the fixed covariances under a direction error or a length error, whereas the fixed covariances reach it within the first 50 s. This slower acquisition is the reason for its higher overall error. With the tuned settings, adaptive scaling follows the fixed covariances from the start. During the sustained disturbance, both adaptive variants and the best fixed covariance reach a lower plateau than the fixed covariance with the literature R_a , and these plateaus are similar. MAJIC shows the same pattern at a lower error level.

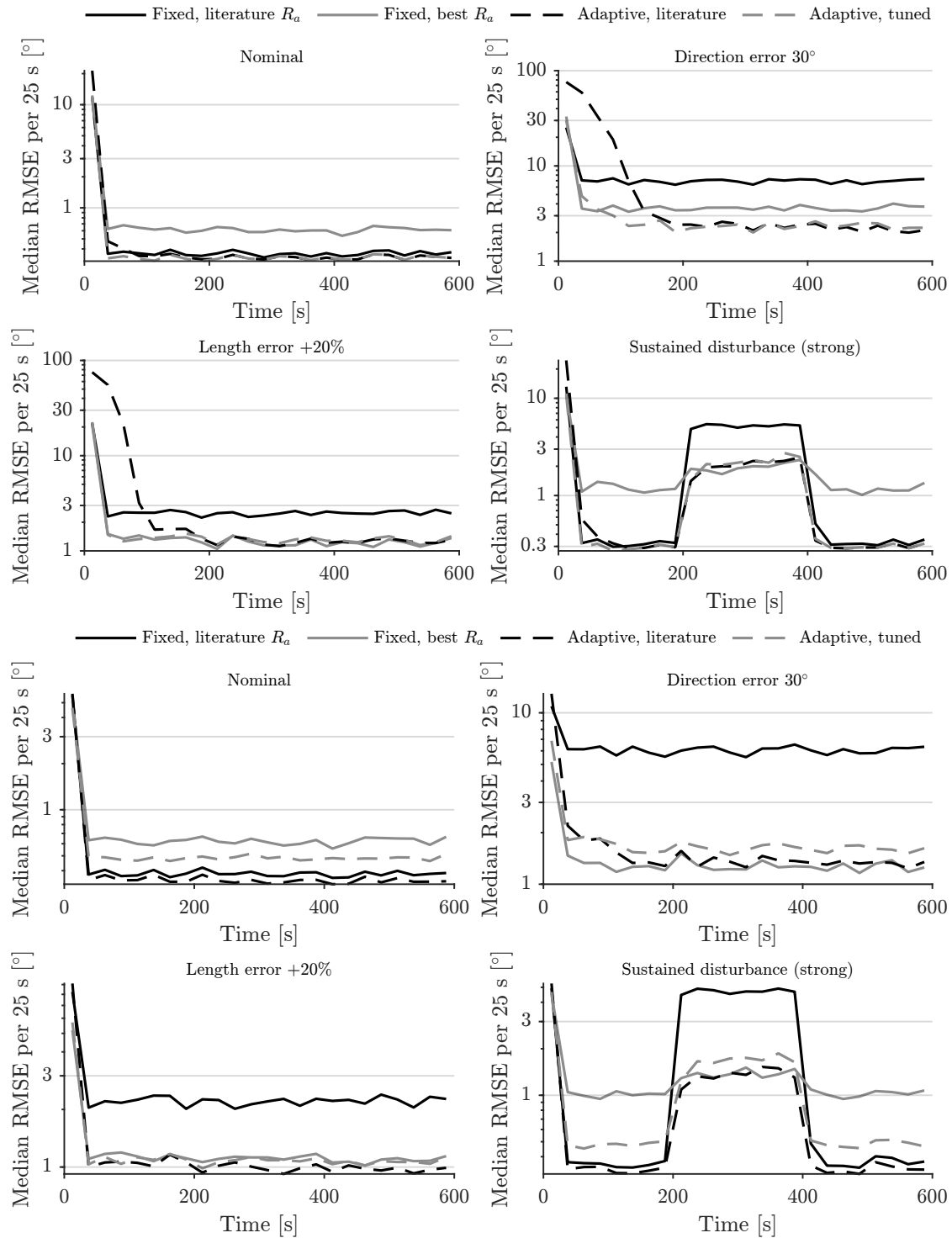


Figure 4-8: Median RMSE in bins of 25 s over the Monte Carlo realizations for the magnetometer-free estimator (top) and MAJIC (bottom), for the nominal condition and the three stressed conditions.

4-9-3 Dependence on the Error Measure

The preceding sections report the error after convergence. Table 4-4 compares this measure with the overall RMSE for the original simulations with the literature configuration. For the magnetometer-free estimator, adaptive scaling lowers the error after convergence under geometry errors and sustained disturbance, but not the overall error. The reason is its slower acquisition: without any error, the median overall RMSE is 3.71° against 2.19° for the baseline. For MAJIC, adaptive scaling lowers both measures under geometry errors and sustained disturbance, and it increases both for accelerometer noise and gyroscope bias. NIS rejection does not lower either measure in any condition, except marginally after convergence for MAJIC under the sustained disturbance. The conclusion that adaptive scaling reduces the error therefore depends on the error measure, on the estimator, and, as shown above, on the tuning of the fixed-covariance baseline.

Table 4-4: Ratio of the median RMSE of the method to that of the corresponding baseline in the original simulations (literature configuration, 100 realizations), for the overall RMSE (all) and the RMSE after convergence (conv.). Values below one indicate that the method lowers the error.

Condition	Weygers				MAJIC			
	Adaptive		NIS		Adaptive		NIS	
	all	conv.	all	conv.	all	conv.	all	conv.
Accelerometer noise ($100\times$)	2.08	1.57	1.62	2.72	1.33	1.61	1.63	2.45
Gyroscope bias ($5\times$)	2.23	1.55	20.81	34.97	1.29	1.27	5.99	7.72
Length error $+20\%$	2.89	0.62	3.24	4.35	0.69	0.50	1.27	1.12
Direction error 30°	1.72	0.39	2.45	2.43	0.46	0.28	1.28	1.29
Sustained disturbance (strong)	1.07	0.46	2.49	3.14	0.44	0.33	1.07	0.93

4-9-4 Interpretation

Magnetometer-free estimator. With the literature settings, adaptive scaling has a larger overall error than a fixed covariance with a larger R_a in all conditions, including the nominal one, because of its slower acquisition. With the tuned settings, in particular the start-up delay, this penalty disappears, and the overall error is then similar to that of the best fixed covariance for the length error and the sustained disturbance, and slightly lower for the direction error. After convergence, the best fixed covariance reaches the error of adaptive scaling with the literature settings for the direction and length errors, and is higher only for the sustained disturbance. The benefit of adaptive scaling that was found in the preceding sections is therefore largely a response relative to a fixed covariance that is too small. Compared with a fixed covariance that is tuned for the condition, it has an advantage only for the direction error and, after convergence, for the sustained disturbance.

MAJIC. For MAJIC, a fixed covariance with a larger R_a has a slightly lower overall error than adaptive scaling with the tuned settings in all three conditions. The larger improvement of adaptive scaling with the literature settings over the fixed covariance with the literature R_a is therefore explained by the small covariance of the baseline. After convergence, adaptive

Table 4-5: Response of the mechanisms and direction of the change in error relative to the fixed-covariance estimator of the same family, at the highest severity of each source of uncertainty. Values are given as Weygers variant / MAJIC variant. $\uparrow$ ($\downarrow$) indicates a higher (lower) median error than the baseline, with the ratio of the two medians in parentheses; $\approx$ indicates overlapping interquartile ranges. The post-convergence error is used, except for the sustained disturbance, for which the error during the disturbance is used. The overall error is compared in Table 4-4.

Source of uncertainty	Median α_t	Rejected [%]	Adaptive	NIS
Accelerometer noise, $\lambda_a = 100$	8.6 / 8.6	75 / 75	$\uparrow$ (1.6 / 1.6)	$\uparrow$ (2.7 / 2.5)
Gyroscope bias, $\lambda_\delta = 5$	1.45 / 1.25	95 / 54	$\uparrow$ (1.6 / 1.3)	$\uparrow$ (35 / 7.7)
Length error, $\kappa_r = \pm 0.20$	20 / 18	97 / 89–90	$\downarrow$ (0.62 / 0.50)	$\uparrow$ (4.0–4.4 / 1.1)
Direction error, $\theta_r = 30^\circ$	142 / 123	99.8 / 99.3	$\downarrow$ (0.39 / 0.28)	$\uparrow$ (2.4 / 1.3)
Sustained disturbance, strong	183 / 183	99 / 99	$\downarrow$ (0.42 / 0.29)	$\uparrow$ (1.7) / $\approx$ (0.90)

scaling with the literature settings is lower than the best fixed covariance only for the sustained disturbance.

Consequence. In both estimators, the lower error of adaptive scaling in the preceding sections is largely reproduced by a fixed covariance with a larger R_a , and adaptive scaling with settings that avoid the start-up penalty does not outperform a fixed covariance that is tuned for the condition, except for small advantages in single conditions. The simulations therefore show how the mechanisms respond to the conditions, but not that adaptation is preferable to tuning. NIS rejection with the tuned thresholds is identical to the fixed covariance with the tuned R_a , because the thresholds are never exceeded by the simulated innovations, so that this setting cannot be judged in the simulations.

Limitations. The check uses 50 realizations, a selection of conditions at the largest simulated severity, and a scale of the fixed covariance that is selected on the same realizations, which favours the baseline. The tuned settings come from the recorded data and are not tuned per simulated condition. It is a control for the effect of the tuning, and it does not replace the studies above.

4-10 Conclusions from the Simulation Study

Table 4-5 compares the response of the mechanisms with the direction of the change in error at the highest severity of each persistent source of uncertainty.

The first subquestion asks how the two mechanisms affect estimation under controlled changes in motion excitation and sensor uncertainty. For the idealized conditions considered here, the simulations show the following.

First, the mechanisms respond to every source of uncertainty that was considered (Table 4-5), but this response does not by itself indicate whether the error decreases. When the update stays informative but is systematically inconsistent with the model (geometry errors, sustained disturbance), the adaptive variants had a lower error after convergence than the fixed-covariance estimator with the literature covariance. For Weygers this does not hold for the overall error, because of the slower acquisition, and a fixed covariance with a larger R_a

reproduces most of the benefit (Sections 4-9-3 and 4-9). For accelerometer noise and gyroscope bias, the adaptive variants had a slightly higher error and the NIS variants a higher error. A consistent reading is that reducing the weight of the acceleration update helps when the error lies in the measurement model itself and is systematic, and does not help when the error is zero mean (noise) or lies in the prediction (gyroscope bias), where the update provides the information that counteracts it.

Second, the temporal structure of the inconsistency determines how NIS rejection behaves. For isolated outliers it removes only the affected samples, and the error does not depend on the number of outliers; the adaptive method provides a similar suppression of the error peak while lowering the weight of the update for 0.5 s per event. For persistent inconsistency, rejection removed most of the updates at the highest severity of each source, the error of Weygers + NIS was higher than that of Weygers in all these conditions, and that of MAJIC + NIS in four of the five (with overlapping interquartile ranges in the fifth). After a sustained disturbance, Weygers + NIS did not return to the error of the baseline.

Third, neither mechanism changes the information available in the motion. Under low excitation the Weygers variants rarely converged, against all realizations of the MAJIC variants, and the adaptive variants converged more slowly under rich excitation and geometry errors. When the inconsistency originates in another measurement source, as for the magnetic update, both mechanisms led to a larger error than the fixed-covariance MAJIC when the acceleration update was informative, because they reduced the weight of the update that would correct the error.

The second subquestion asks whether the mechanisms reduce the degradation caused by errors in the assumed IMU-to-joint-centre geometry. For the tested length errors (up to 20%) and direction errors (up to 30°), adaptive scaling reduced the error after convergence relative to the fixed-covariance estimator with the literature covariance (Table 4-5 and Section 4-5). It did not remove the effect of the geometry error, since the error still increased with the mismatch, and the convergence times of the adaptive variants increased as well. NIS-based rejection did not reduce the error, and increased it for Weygers in every mismatch condition. The reduction by adaptive scaling does not hold for the overall error of Weygers, and a fixed covariance with a larger R_a reached a similar error after convergence (Section 4-9).

These conclusions concern the behaviour of the estimators under a single source of uncertainty at a time, with two ideal rigid segments and one joint geometry. They describe what the mechanisms react to and in which direction this shows in the error, and not which accuracy can be expected in practice. Chapter 5 evaluates the same variants on recorded lower-extremity motion, in which these sources are present simultaneously.

Evaluation on Recorded Human Motion

The simulation study in Chapter 4 isolated individual sources of uncertainty and showed what adaptive covariance scaling and NIS-based rejection react to. Recorded human motion contains all of these sources at the same time, together with soft-tissue artifacts and imperfect joint-centre geometry. This chapter therefore evaluates the six estimator variants of Chapter 3 on recorded lower-extremity motion from the OpenSense dataset [1].

The estimators are tuned on the recorded data before they are compared. The default parameters in the literature were chosen for other recordings, and a comparison between a tuned and an untuned estimator would mainly show the effect of the tuning. Every estimator variant therefore received its own tuned parameters, using the same recordings and the same objective (Section 5-3). Accuracy is measured as the overall angular RMSE over the complete recording. This measure needs no convergence criterion, it includes every recording, and failures of an estimator therefore remain visible in the result.

The sensor-to-joint-centre vectors required by the acceleration measurement model are estimated from the inertial data (sensor-only setting) for the primary comparison. Marker-derived vectors are used as a secondary condition.

5-1 Dataset and Processing

5-1-1 OpenSense Dataset

The publicly available OpenSense dataset of [1] contains synchronized inertial and optical motion-capture measurements from 11 healthy participants. Eight Xsens MTw Awinda IMUs were mounted on rigid marker plates attached to the upper and lower back and to the bilateral thighs, shanks and feet. Optical marker trajectories were recorded at 100 Hz.

Two approximately 10-minute motion conditions were recorded. The first consists of repeated walking and turning referred to as “Walking”, while the second contains a sequence of more

varied movements including sitting, standing, stair negotiation, side-stepping, walking and running referred to as “Complex Tasks” [1].

This study considers the left and right hip, knee and ankle joints. The corresponding sensor pairs are the pelvis and thigh for the hip, thigh and shank for the knee, and shank and foot for the ankle.

5-1-2 Data Selection

The exclusions follow the supporting information of [1]. Because ferromagnetic disturbances in the laboratory distorted the IMU orientation estimates, the authors pre-screened the IMU data and excluded IMUs of which the orientation differences in their inverse kinematics exceeded 45° in the first 10 s of a trial, or of which the average range over 60 ms bins exceeded 30° . They list the excluded IMUs per subject.

Subjects 06 and 10 were excluded because their IMU measurements were recorded at 40 Hz, whereas the processing pipeline used in this study assumes the 100 Hz optical time base. Subject 05 was excluded because the IMUs of the right side did not transmit data, the left foot IMU was excluded, and all data stopped transmitting during the sequence of lower-extremity movements. Subject 09 was excluded: both foot IMUs were excluded by the OpenSense authors and the provided dataset contains no complex-task recording of this subject.

In addition to these subject-level exclusions, several ankle recordings were unavailable because the corresponding foot IMUs were excluded by the OpenSense authors: both feet for Subjects 01 and 11, and the left foot for Subject 08. Since ankle estimation requires a shank–foot sensor pair, these missing foot recordings reduce the number of available ankle trials relative to the hip and knee joints. Complex-task recordings are not contained in the provided dataset for Subjects 07 and 08, and the right-knee recording of the complex tasks is missing for Subject 11.

The final dataset therefore contains seven subjects, {01, 02, 03, 04, 07, 08, 11}, and 62 unique subject–motion–joint recordings. Of these, 37 correspond to walking and 25 to the complex-task condition, and 24, 23 and 15 correspond to the hip, knee and ankle joints, respectively.

5-1-3 Synchronization and Reference Orientation

The inertial measurements and optical marker trajectories are synchronized before constructing the datasets for each sensor pair. From the marker trajectories, the orientation of the marker plate is reconstructed and used to obtain its angular-speed magnitude. This signal and the magnitude of the pelvis gyroscope measurement are low-pass filtered at 6 Hz, and their temporal offset is determined as the lag that maximizes their cross-correlation. The resulting offset is applied to all sensor pairs within the recording. The pelvis is therefore used only as a common synchronization anchor, synchronization is not performed independently for each joint. The 6 Hz filtering is applied to remove measurement spikes, and is used only for determining the temporal offset. After synchronization, only the interval available in both recordings is retained, with a maximum duration of 600 s.

The orientations of the rigid marker plates obtained from the optical motion-capture system are subsequently used to construct the reference relative orientation. Let $R_t^{BR_1}$ and $R_t^{BR_2}$

denote the orientations of the marker frames R_1 and R_2 with respect to the optical reference frame B . The reference relative orientation of the second marker frame with respect to the first is then given by

$$R_t^{R_1 R_2} = \left(R_t^{B R_1} \right)^\top R_t^{B R_2}. \quad (5-1)$$

5-1-4 Constant Sensor–Marker Frame Alignment

The optical marker frames R_1 and R_2 are rigidly connected to the corresponding IMUs, with sensor frames S_1 and S_2 , respectively. However, the marker and sensor coordinate axes are not necessarily aligned. Consequently, the IMU-based relative orientation cannot be compared directly with the optical reference relative orientation defined in Section 5-1-3. The constant rotations between each marker frame and its corresponding sensor frame must first be determined.

In [25], this coordinate-frame misalignment is estimated from the optical and inertial relative-orientation trajectories. The approach uses the relative-orientation calibration proposed in [7], in which two constant rotations are jointly obtained using an SVD-based solution. Following this approach, let $q^{R_1 S_1}$ and $q^{R_2 S_2}$ denote the constant orientations of sensor frames S_1 and S_2 with respect to their corresponding marker frames R_1 and R_2 . For each time step t , these transformations relate the optical reference relative orientation $q_t^{R_1 R_2}$ to the estimated sensor relative orientation $\hat{q}_t^{S_1 S_2}$ according to

$$q_t^{R_1 R_2} \odot q^{R_2 S_2} \approx q^{R_1 S_1} \odot \hat{q}_t^{S_1 S_2}. \quad (5-2)$$

The constant rotations $q^{R_1 S_1}$ and $q^{R_2 S_2}$ are estimated from the orientation trajectories using the SVD-based solution described in [7]. The estimated relative orientation can subsequently be expressed in the optical marker frames as

$$\hat{R}_t^{R_1 R_2} = \hat{R}^{R_1 S_1} \hat{R}_t^{S_1 S_2} \hat{R}^{S_2 R_2}, \quad (5-3)$$

where $\hat{R}^{S_2 R_2} = \left(\hat{R}^{R_2 S_2} \right)^\top$. This aligned relative orientation can then be compared directly with the optical reference $R_t^{R_1 R_2}$.

The constant alignment rotations were estimated separately for each estimator. This is required because the calibration uses the relative-orientation trajectory produced by the estimator itself. Applying calibration parameters obtained from one estimator to the remaining methods would therefore couple their evaluation to the errors of that particular estimator.

Only the final two-thirds of each recording were used to estimate the alignment rotations. The first third was excluded to reduce the influence of the initial convergence period of the orientation estimators on the estimated constant transformations.

The alignment accounts only for constant coordinate-frame misalignment between the IMUs and their marker frames. It cannot compensate for time-varying estimation errors. Relative-orientation drift, temporary divergence, and other time-dependent errors therefore remain present after alignment.

5-2 Joint-Centre Geometry

The joint-centre acceleration measurement model requires the constant vectors $r_i^{S_i}$ from each IMU origin to the corresponding joint centre. In recorded human motion these vectors are not known a priori. Two estimates are therefore considered: an inertial estimate based on [19] and a marker-derived estimate obtained from the optical motion-capture data.

5-2-1 Inertially Estimated Geometry

Inertially estimated geometry is used for the primary evaluation. For each subject–motion–joint recording, one pair of constant joint-centre vectors is estimated directly from the synchronized accelerometer and gyroscope measurements using an implementation based on [19]. The optimization settings were adopted from [25], where this method was used to obtain the sensor-to-joint vectors.

The method exploits the fact that the two adjacent sensors observe the acceleration of the same physical joint centre. Although the corresponding joint-centre accelerations are expressed in different sensor frames, their magnitudes are invariant to rotation. The vectors $r_1^{S_1}$ and $r_2^{S_2}$ are therefore estimated by minimizing the disagreement

$$\rho_t = \left\| a_{jc,t}^{S_1} \right\|_2 - \left\| a_{jc,t}^{S_2} \right\|_2, \quad (5-4)$$

where $a_{jc,t}^{S_i}$ is calculated using the joint-centre acceleration model defined in Chapter 3. The resulting nonlinear least-squares problem is solved with damped Gauss–Newton iterations [25], using the settings from [25]: The damping factor of 0.7 limits the step size, and number of iterations is fixed at 25. Tests with slightly different settings did not give better results, and additional iterations only increased the computation time. Both vectors are initialized as $[0.1, 0.1, 0.1]^\top$ m, which corresponds to a length of 17 cm, the order of magnitude of the distance between the sensor and the joint centre. One pair of vectors is estimated independently for every recording using the complete synchronized interval, up to the maximum retained duration of 600 s. No anatomical bounds or vector-length constraints are imposed.

Because this procedure requires only inertial measurements, these vectors represent the geometry available in an inertial-only implementation and are therefore used for the primary estimator comparison.

5-2-2 Marker-Derived Sensor-to-Joint-Centre Geometry

Although the sensor-to-joint-centre vectors can be estimated from inertial data, marker-based estimates are additionally computed to provide an independent geometric reference for the evaluation. In [21], the sensor-to-joint-centre vectors are estimated from the optical marker trajectories by exploiting the common joint-centre geometry of the two connected segments. For sensor i , let $p_{i,t}^B$ denote the position of the corresponding marker plate in the optical reference frame B , obtained as the mean position of its four markers. Its orientation is denoted by $R_t^{BR_i}$.

Because the marker frame R_i and sensor frame S_i are rigidly connected but not necessarily aligned, their constant relative orientation is first determined. An angular-velocity trajectory $\omega_t^{R_i}$ is reconstructed from the optical orientation trajectory $R_t^{BR_i}$ and compared with the measured gyroscope signal $y_{\omega,t}^{S_i}$. The constant rotation $R^{R_i S_i}$ is obtained from

$$\omega_t^{R_i} \approx R^{R_i S_i} y_{\omega,t}^{S_i}. \quad (5-5)$$

The orientation of the sensor frame with respect to the optical reference frame is then

$$R_t^{BS_i} = R_t^{BR_i} R^{R_i S_i}. \quad (5-6)$$

Let $r_i^{S_i}$ denote the constant vector from sensor i to the joint centre, expressed in sensor frame S_i . The joint-centre position computed from either sensor must be identical, such that

$$p_{1,t}^B + R_t^{BS_1} r_1^{S_1} = p_{2,t}^B + R_t^{BS_2} r_2^{S_2}. \quad (5-7)$$

Rearranging gives

$$-R_t^{BS_1} r_1^{S_1} + R_t^{BS_2} r_2^{S_2} = p_{1,t}^B - p_{2,t}^B. \quad (5-8)$$

Since $r_1^{S_1}$ and $r_2^{S_2}$ are constant over the recording, stacking (5-8) over all time samples results in an overdetermined linear system. The sensor-to-joint-centre vectors $\hat{r}_1^{S_1}$ and $\hat{r}_2^{S_2}$ are obtained from its least-squares solution.

It is important to note that the marker-derived geometry assumes the sensor origin to coincide with the origin of the corresponding marker plate. Since this assumption is only approximate, these vectors should not be regarded as exact ground truth. It can therefore not be concluded that the marker-derived vectors are necessarily more accurate than the inertially estimated vectors obtained using the method of [19].

5-3 Estimator Tuning and Evaluation

5-3-1 Tuning of the Estimators

The default parameters of the estimators (Section 3-8) are taken from the literature, and in particular from the configuration in [21]. They were selected for other recordings, and using them unchanged would leave the comparison open to the objection that one estimator happens to suit these values better than another. Each estimator variant was therefore tuned on the OpenSense recordings. The tuning is limited to the parameters that determine how the acceleration and magnetic updates are weighted. The gyroscope process covariance Q_ω , the initial covariance, the observability threshold and the numerical differentiation are not changed.

The objective of every tuning step is the mean, over the 62 recordings, of the overall angular RMSE defined in Section 5-3-2, obtained with the inertially estimated geometry. The mean is used because it keeps recordings with large errors visible, which the median would hide. The parameters are tuned one group at a time and in the following order, so that each estimator is compared with a baseline that has been tuned in the same way:

1. the acceleration covariance $R_a = r I_3$ of the fixed-covariance magnetometer-free estimator (Weygers). The same value is used for all other variants except the adaptive ones;
2. the magnetic covariance $R_{\text{mag}} = r_m I_3$ of the MAJIC baseline, with R_a fixed;
3. the NIS threshold of Weygers + NIS and of MAJIC + NIS, each with the tuned R_a (and R_{mag}) fixed;
4. for the adaptive estimators, the acceleration covariance R_a , the window length N_w and a start-up delay during which the adaptation is inactive. MAJIC + Adaptive uses the settings found for Weygers + Adaptive.

Each sweep starts at the literature value and continues in both directions until the objective no longer improves, followed by a refinement around the best value. The sweeps, the leave-one-subject-out check and the figures are given in Appendix A-3.

Table 5-1: Parameters used for the recorded-data evaluation. The literature values are those of Chapter 3. R_a and R_{mag} are given as the scalar multiplying the identity matrix.

Parameter	Estimators	Literature	Tuned
R_a	Weygers, + NIS, MAJIC, + NIS	2.5×10^{-3}	9.13×10^{-2}
R_a	+ Adaptive (both)	2.5×10^{-3}	3.85×10^{-3}
R_{mag}	MAJIC variants	9.0×10^{-4}	2.85×10^{-4}
NIS threshold ϵ_{max}	Weygers + NIS	11.3449	4000
NIS threshold ϵ_{max}	MAJIC + NIS	11.3449	751
Window N_w	+ Adaptive (both)	50 (0.5 s)	30 (0.3 s)
Start-up delay of the adaptation	+ Adaptive (both)	0 s	10 s

The tuning is performed on the same recordings that are used for the evaluation. It is therefore in-sample, and the tuned estimators have an advantage over the literature configuration that would not be present on new data. This is partly checked with a leave-one-subject-out procedure (Appendix A-3), in which the parameter is selected without the held-out subject.

5-3-2 Performance Measures

The relative-orientation angular error $e_{\theta,t}$ in (4-6) is evaluated between the aligned estimate $\hat{q}_t^{R_1 R_2}$ and the optical reference $q_t^{R_1 R_2}$. The primary accuracy measure is the overall angular RMSE over the complete recording,

$$e_{\text{RMSE}} = \sqrt{\frac{1}{N} \sum_{t=1}^N e_{\theta,t}^2}, \quad (5-9)$$

where N is the number of samples of the recording. No convergence criterion is applied: a metric that is evaluated only after convergence excludes the recordings in which an estimator fails, so that estimators with different failure rates cannot be compared fairly.

The distribution over the 62 recordings is summarized by the mean (which is also the tuning objective), the median, the 90th percentile and the maximum. To compare a modified estimator with its fixed-covariance baseline on the same recording, the paired difference

$$\Delta e_{\text{RMSE}} = e_{\text{RMSE}}^{\text{baseline}} - e_{\text{RMSE}}^{\text{modified}} \quad (5-10)$$

is used, which is positive when the modified estimator has the lower error. For each comparison the number of recordings with $\Delta e_{\text{RMSE}} > 0$, the median and the mean of Δe_{RMSE} and the number of subjects whose mean difference is positive are reported. The mean difference is also given with a 95% percentile interval from a subject-level bootstrap, in which the seven subjects are resampled with replacement together with all of their recordings (10 000 resamples). This takes the dependence between the recordings of a subject into account (left and right joints are included and adjacent joints share an IMU). With seven subjects these intervals remain approximate. A two-sided sign test is added as a rough indication only, and its p -values are not used as formal evidence.

5-4 Results

5-4-1 Primary Estimator Comparison

The primary comparison uses the inertially estimated geometry of Section 5-2-1. Figure 5-1 shows the overall angular RMSE of every recording for the six variants, and Table 5-2 summarizes the distributions together with the paired comparison with the corresponding baseline.

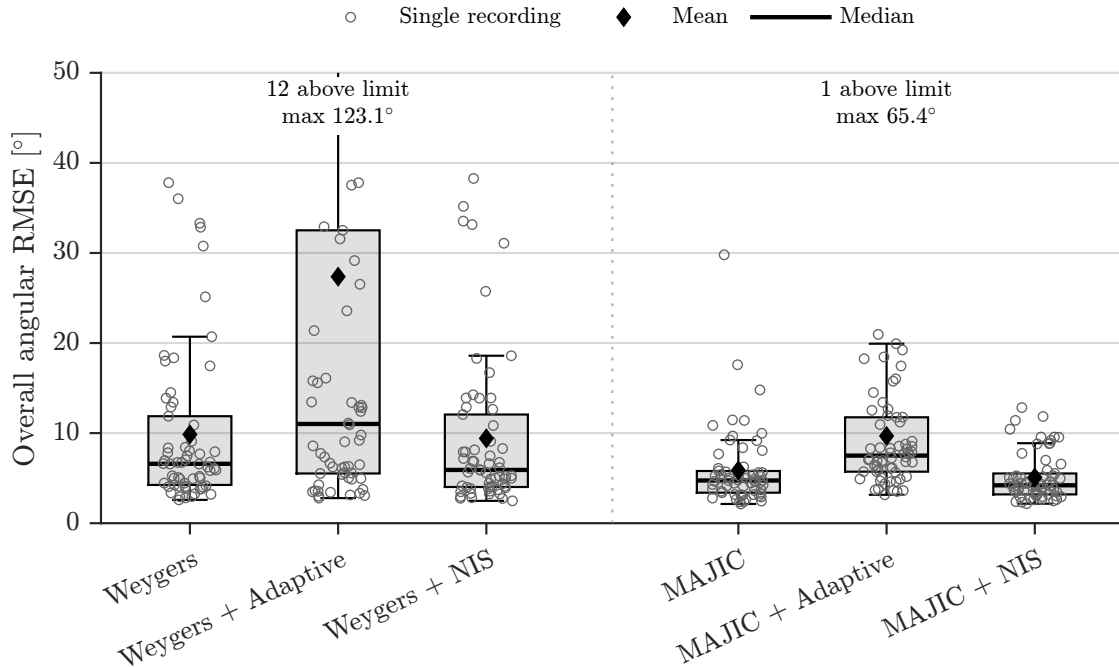


Figure 5-1: Overall angular RMSE of the six estimator variants for the 62 recordings, using the inertially estimated geometry. Boxes show the quartiles and the median, whiskers the data within 1.5 times the interquartile range, diamonds the mean and circles the individual recordings. Values above the displayed range are indicated by their number and maximum.

Table 5-2: Overall angular RMSE over the 62 recordings (inertially estimated geometry), and paired comparison with the corresponding baseline. “Better” is the number of recordings with $\Delta e_{\text{RMSE}} > 0$, and “Subj.” the number of the seven subjects with a positive mean Δe_{RMSE} . Gains are in degrees; positive is better than the baseline.

Method	Overall RMSE [°]				Compared with baseline			
	Mean	Median	P90	Max	Better	Median	Mean	Subj.
Weygers	9.85	6.59	22.03	37.8	–	–	–	–
+ Adaptive	27.37	11.02	90.90	123.1	21/62	–1.48	–17.51	1/7
+ NIS	9.42	5.91	20.73	38.3	42/62	+0.07	+0.44	5/7
MAJIC	5.87	4.74	10.24	29.8	–	–	–	–
+ Adaptive	9.68	7.50	17.70	65.4	1/62	–2.51	–3.81	0/7
+ NIS	5.08	4.21	9.55	12.8	45/62	+0.09	+0.79	6/7

The magnetometer-free baseline has a mean overall RMSE of 9.85° (median 6.59°). Adding the conditional magnetic update (MAJIC) lowers this to 5.87° (median 4.74°), which is the better estimator on 50 of the 62 recordings and for all seven subjects (mean difference 3.99°). NIS-based rejection lowers the mean error of both estimators, by 0.44° for Weygers and by 0.79° for MAJIC. MAJIC + NIS has the lowest mean, 90th percentile and maximum of all variants (5.08° , 9.55° and 12.8°). The median gains of NIS-based rejection (0.07° and 0.09°) are much smaller than the mean gains, so that most recordings change little and a few improve more.

Adaptive covariance scaling increases the error in both estimators. For Weygers + Adaptive the mean overall RMSE is 27.37° , with a 90th percentile of 90.9° and a maximum of 123.1° . Twelve recordings, ten of them hip recordings, have an overall RMSE above 45° , whereas no recording of the baseline does. For MAJIC + Adaptive the mean is 9.68° and only one of the 62 recordings improves.

5-4-2 Paired Comparison per Recording

Figure 5-2 shows the paired differences (5-10) for every recording. For NIS-based rejection the differences are small and mostly positive: 42 of 62 recordings improve for Weygers and 45 of 62 for MAJIC. The improvement is found in five and six of the seven subjects, respectively. The overall RMSE changes by more than 1° in only a minority of the recordings: it decreases by more than 1° in 8 (Weygers) and 14 (MAJIC) recordings and increases by more than 1° in one recording in each case. For MAJIC, a single recording contributes 20.2° of the summed gain, which is 0.33° of the mean gain of 0.79° . For adaptive scaling the picture is different: a minority of the recordings improves slightly (21 of 62 for Weygers), while the remaining recordings lose between a few degrees and more than 100° .

Table 5-3 gives the mean paired differences with subject-level bootstrap intervals for both geometry sources. All intervals exclude zero in the direction of the main findings: the magnetic update and NIS-based rejection improve the estimate, and adaptive scaling worsens it. The interval of the gain of NIS-based rejection for the magnetometer-free estimator has the narrowest margin, with its lower end close to zero, and it holds for the threshold of 4000 that was tuned on all subjects: when the threshold is selected without the held-out subject, the

gain is not found (Appendix A-3). The intervals for the marker-derived geometry are further above zero.

Table 5-3: Mean paired difference in overall angular RMSE (baseline minus modified; positive is better) with the 95% subject-level bootstrap interval in brackets.

Comparison	Inertial geometry [°]	Marker-derived geometry [°]
Weygers + NIS vs Weygers	+0.44 [+0.13, +0.70]	+1.24 [+0.62, +1.91]
Weygers + Adaptive vs Weygers	-17.51 [-30.27, -3.80]	-13.86 [-22.53, -3.58]
MAJIC + NIS vs MAJIC	+0.79 [+0.21, +1.68]	+1.17 [+0.67, +1.74]
MAJIC + Adaptive vs MAJIC	-3.81 [-4.79, -2.74]	-2.65 [-3.52, -1.64]
MAJIC vs Weygers	+3.99 [+2.05, +6.37]	+2.09 [+1.01, +3.44]
MAJIC + NIS vs Weygers + NIS	+4.34 [+2.76, +6.59]	+2.01 [+0.78, +3.53]

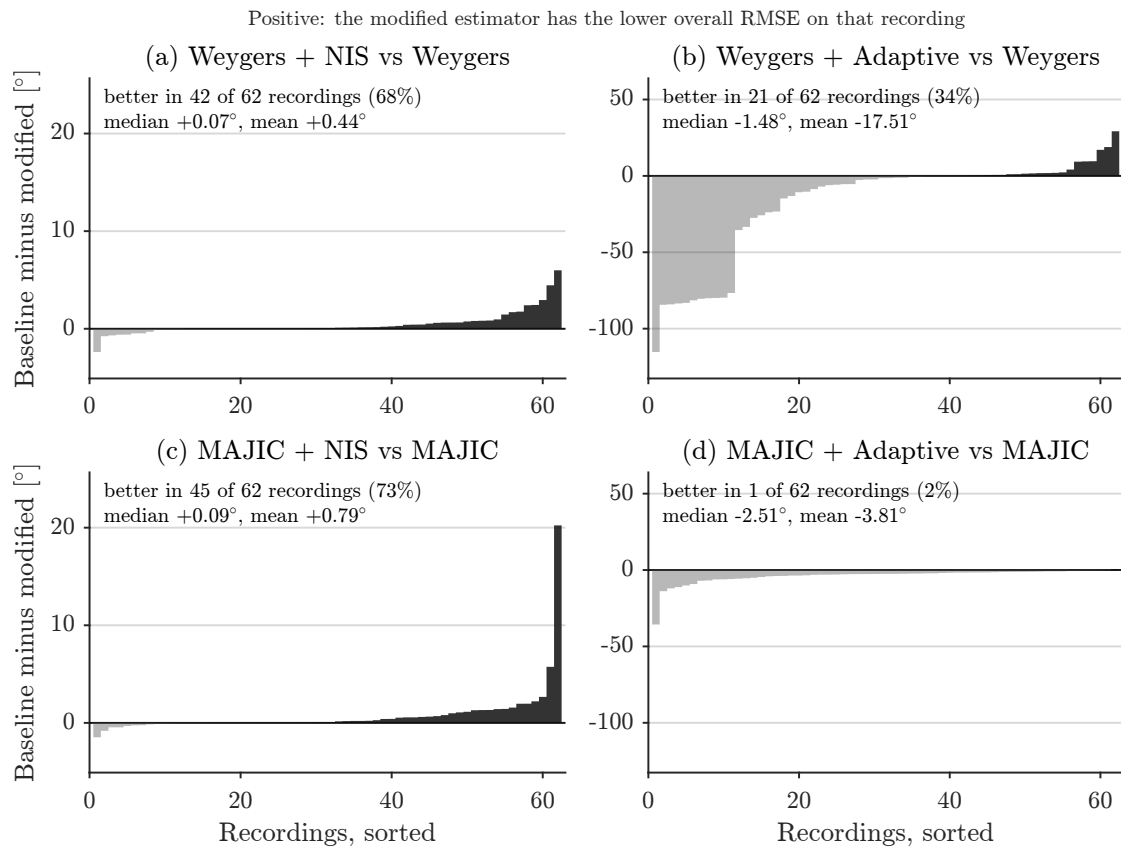


Figure 5-2: Paired difference in overall angular RMSE between each estimator and its baseline for every recording, sorted per panel. Positive values (dark bars) indicate that the modified estimator has the lower error on that recording. Top row: Weygers, bottom row: MAJIC; left: NIS-based rejection, right: adaptive covariance scaling.

5-4-3 Joints and Motion Types

Table 5-4: Mean overall angular RMSE [$^{\circ}$] per joint and per motion type (inertially estimated geometry). The number of recordings is given in brackets.

	Weygers	+ Adaptive	+ NIS	MAJIC	+ Adaptive	+ NIS
Hip (24)	11.32	48.30	11.05	4.51	7.52	4.26
Knee (23)	7.01	10.93	6.52	5.66	8.88	4.41
Ankle (15)	11.87	19.10	11.26	8.35	14.37	7.39
Walking (37)	7.20	20.85	7.03	4.81	9.07	4.03
Complex tasks (25)	13.79	37.01	12.96	7.43	10.59	6.62

Table 5-4 gives the results per joint and per motion type. The advantage of the magnetic update is largest at the hip, where the mean error is 4.51° for MAJIC and 11.32° for Weygers, and smaller at the knee (5.66° against 7.01°) and the ankle (8.35° against 11.87°). MAJIC + NIS is the best variant at all three joints. The failures of Weygers + Adaptive are concentrated at the hip (48.30° mean, ten recordings above 45°), but the adaptive estimators are also worse than their baselines at the knee and the ankle. All variants have a higher error during the complex tasks than during walking, and within each estimator the ordering of the variants (NIS rejection, baseline, adaptive scaling) is the same in both motion types. The ankle group contains only 15 recordings, so that differences between joints should be interpreted with caution.

5-4-4 Behaviour of the Adaptive Factor

The adaptive factor α_t is far above one on the recorded data (Figure 5-3). The median of α_t over the recordings is 743 at the hip, 1158 at the knee and 2568 at the ankle for Weygers + Adaptive, and of the same order for MAJIC + Adaptive (677, 1154 and 2675). With the tuned $R_a = 3.85 \times 10^{-3}$ of the adaptive estimators, this corresponds to an effective acceleration covariance $\alpha_t R_a$ of about 2.9, 4.5 and 9.9, which is 31, 49 and 108 times the tuned fixed covariance of the baseline (9.13×10^{-2}). The acceleration update therefore receives a very small weight for most of the recording.

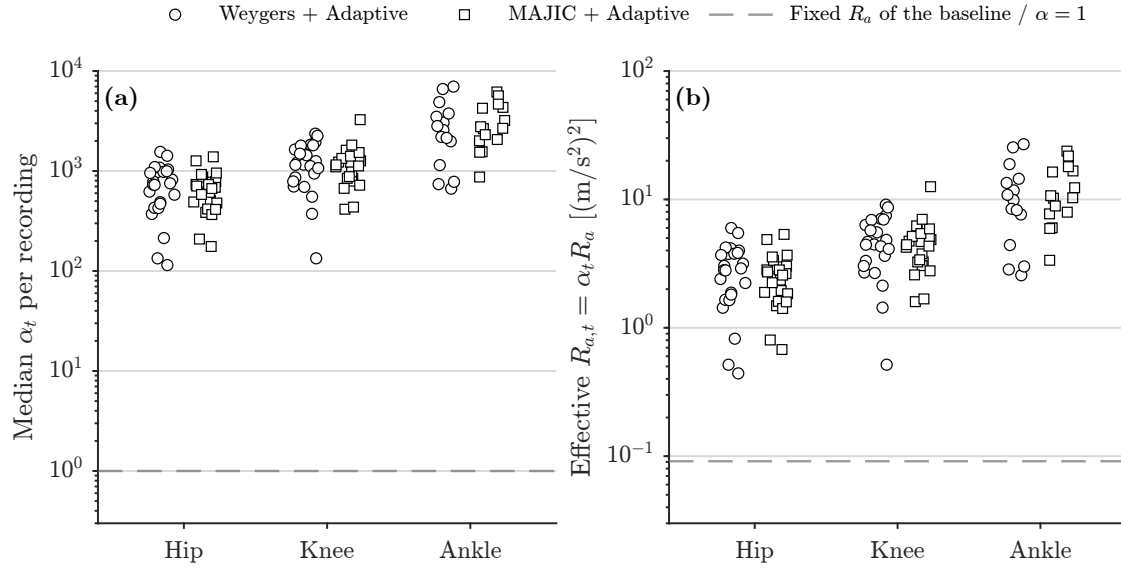


Figure 5-3: Median adaptive factor α_t per recording (a) and the resulting effective acceleration covariance (b), per joint, for the adaptive estimators. The dashed lines show $\alpha = 1$ and the tuned fixed covariance of the baseline.

An upper bound on α_t was investigated as a possible remedy (Appendix A-3). The lowest overall RMSE of the tuned bounded estimator is obtained for the bound $\alpha_{\max} = 1$ (9.81°), which is practically equal to the baseline (9.85°). The adaptive factor can then only reduce the covariance, and the result shows that this gives no measurable improvement. The bounded estimator is not part of the main comparison.

5-4-5 Error over Time

Figure 5-4 shows the RMSE in consecutive bins of 25 s, as median and mean over the 62 recordings, and Table 5-5 summarizes three segments of the recording: the first 25 s, the following 75 s and the remainder. This shows when the error is made, and separates the acquisition at the start of a recording from the later behaviour without a convergence criterion.

For NIS-based rejection, the benefit builds up over time. In the first 25 s, in which the rejection is inactive during the first 10 s, the magnetometer-free estimator is close to its baseline (better in 23 of 62 recordings, median difference 0.00°). It improves 36 of 62 recordings in the following 75 s and 41 of 62 in the remainder (mean difference 0.52° and 0.53°). For MAJIC, rejection improves 41, 43 and 43 of 62 recordings in the three segments, with a mean difference of 1.14° already in the first 25 s.

Adaptive scaling does not catch up. For the magnetometer-free estimator, it has a lower error than the baseline in 18, 19 and 19 of 62 recordings in the three segments. The median difference decreases in magnitude from -6.41° via -2.81° to -1.23° , so that the typical recording recovers part of its loss, but the mean difference does not (-13.98° , -17.07° and -18.09°): the error of a minority of the recordings persists or grows, which is the failure described in Section 5-4-4. The mean error of Weygers + Adaptive is not monotonic in time (Figure 5-4c): it decreases around 400 s and increases again. This is consistent with a heading

error that drifts through zero when the estimate is driven by the gyroscope alone, but this was not checked per recording. For MAJIC, the difference decreases in both statistics (median -4.59° , -2.99° and -1.80° , mean -7.68° , -5.34° and -2.58°), so that the loss is mainly due to slower acquisition, but adaptive scaling has the lower error in only 9, 7 and 5 of 62 recordings and does not reach the baseline in the remainder of the recording.

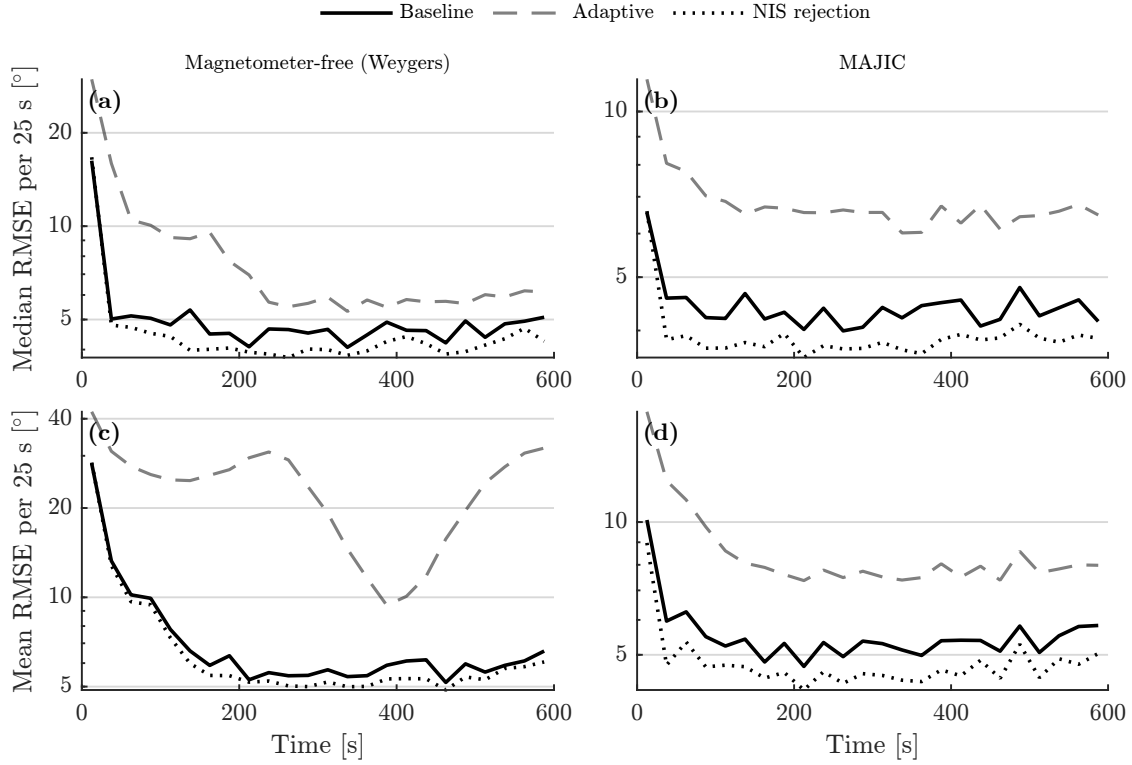


Figure 5-4: Median (top) and mean (bottom) RMSE per bin of 25 s over the 62 recordings for the magnetometer-free estimator (left) and MAJIC (right), with the inertially estimated geometry.

Table 5-5: Mean [median] RMSE in degrees over the recordings in three segments of the recording (inertially estimated geometry).

Method	0–25 s	25–100 s	100–600 s
Weygers	28.2 [16.3]	11.7 [5.1]	6.3 [5.2]
+ Adaptive	42.2 [29.8]	28.8 [12.5]	24.4 [6.5]
+ NIS	28.5 [16.7]	11.2 [4.9]	5.8 [4.7]
MAJIC	10.1 [6.5]	6.1 [4.6]	5.4 [4.5]
+ Adaptive	17.8 [11.4]	11.4 [8.1]	8.0 [6.7]
+ NIS	9.0 [6.6]	5.0 [3.9]	4.7 [3.9]

5-4-6 Effect of the Tuning

Figure 5-5 shows the effect of each tuning step on the mean overall RMSE. Tuning the acceleration covariance of the magnetometer-free baseline lowers its mean error from 15.46°

to 9.85° (36%), which is a larger change than that of any reliability-handling method. For MAJIC, with the tuned R_a , tuning the magnetic covariance lowers the error from 6.41° to 5.87° . The subsequent NIS thresholds lower the error further to 9.42° and 5.08° . Adaptive scaling remains above the baselines at its best tuning.

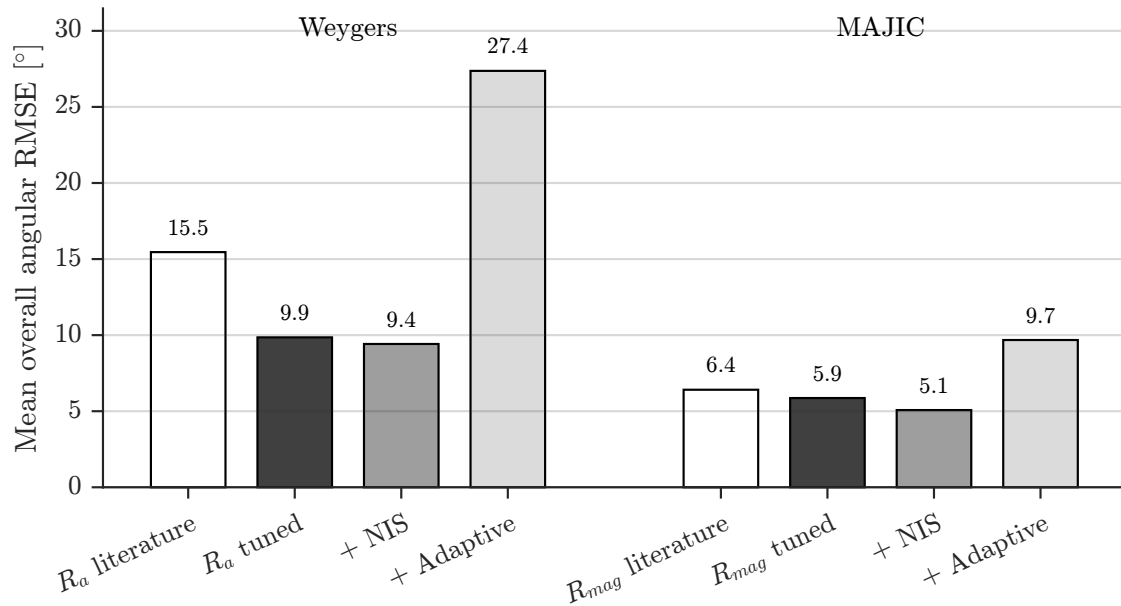


Figure 5-5: Mean overall angular RMSE after each tuning step. For MAJIC the acceleration covariance is already tuned in the first bar, so that only the effect of the magnetic covariance is shown.

5-4-7 Joint-Centre Geometry

The inertially estimated joint-centre vectors are compared with the marker-derived estimates from Section 5-2-2. This comparison characterizes the geometry disagreement present in the sensor-only configuration, and it is used to see whether the results depend on the geometry source.

Table 5-6: Difference between inertially estimated and marker-derived joint-centre vectors. Both segment vectors are included within each joint class. Values are reported as median [IQR].

Joint	Vector difference [cm]	Direction difference [°]
Hip	11.66 [6.11]	21.14 [15.70]
Knee	12.56 [18.79]	15.24 [113.97]
Ankle	10.64 [18.15]	11.77 [11.35]

For the vector from the thigh sensor at the knee, the direction differs by more than 90° from the marker-derived vector in 14 of 23 recordings (median 123°), and its median length is 5.9 cm against 24.3 cm. The corresponding vector from the shank sensor differs by a median of 10° . For the vector of the proximal sensor, the direction differs by more than 90° in 4 of 24 hip and 2 of 15 ankle recordings, and none of the vectors of the distal sensor does.

Table 5-7: Overall angular RMSE for the two geometry sources, mean [median] over the 62 recordings. The last column is the difference of the means (marker-derived minus inertial); negative values indicate a lower error with marker-derived vectors.

Method	Inertial [°]	Marker-derived [°]	Difference [°]
Weygers	9.85 [6.59]	7.66 [5.79]	−2.20
+ Adaptive	27.37 [11.02]	21.52 [6.84]	−5.85
+ NIS	9.42 [5.91]	6.42 [4.48]	−3.00
MAJIC	5.87 [4.74]	5.57 [4.77]	−0.29
+ Adaptive	9.68 [7.50]	8.23 [6.51]	−1.45
+ NIS	5.08 [4.21]	4.40 [3.48]	−0.67

Table 5-7 shows that every variant has a lower mean error with the marker-derived vectors. The difference is largest for the magnetometer-free estimators (2.20° for the baseline and 3.00° with NIS rejection) and small for MAJIC (0.29° and 0.67°), which is consistent with the magnetic update reducing the dependence on the acceleration update. The ranking of the variants does not change: MAJIC + NIS is the best and the adaptive variants are the worst, also with marker-derived vectors. The benefit of NIS-based rejection is larger and more consistent with the marker-derived vectors: 50 of 62 recordings (Weygers) and 54 of 62 (MAJIC) improve, in all seven subjects, with mean gains of 1.24° and 1.17°. Adaptive scaling again increases the error (mean −13.86° and −2.65° for Weygers and MAJIC, respectively).

5-4-8 Behaviour with the Literature Settings

The comparisons above use estimators that are tuned on the same recordings. In practice, such tuning is often not possible, so that the question remains how the variants behave with the settings taken from the literature. Table 5-8 gives the post-convergence error, the convergence success and the retention of the six variants with these settings, for the same 62 recordings and the inertially estimated geometry. The retention is the number that do not subsequently lose convergence, where convergence is lost when the error remains above 10° continuously for at least 5 s.

Table 5-8: Recorded data with the literature settings (no tuning). Post-convergence RMSE is given as median [IQR]; Δe is the paired reduction relative to the fixed-covariance estimator of the same family (positive is better). Convergence is the number of recordings that converged and retention the number of those that did not subsequently lose convergence.

Method	$e_{\text{RMSE}}^{\text{post}}$ [°]	Δe [°]	Convergence	Retention
Weygers	8.07 [5.21]	—	44/62 (71.0%)	17/44 (38.6%)
Weygers + Adaptive	5.06 [3.24]	2.44 [3.09]	40/62 (64.5%)	31/40 (77.5%)
Weygers + NIS	7.10 [3.94]	1.17 [2.10]	50/62 (80.6%)	28/50 (56.0%)
MAJIC	8.10 [5.12]	—	45/62 (72.6%)	21/45 (46.7%)
MAJIC + Adaptive	5.55 [4.10]	2.38 [3.76]	55/62 (88.7%)	41/55 (74.5%)
MAJIC + NIS	6.88 [3.15]	0.97 [2.17]	50/62 (80.6%)	31/50 (62.0%)

With the literature settings, adaptive scaling gives the largest reduction of the post-convergence error for both estimators, and it raises the retention from 38.6% to 77.5% and from 46.7% to 74.5%. For MAJIC it also increases the convergence success from 72.6% to 88.7%, so that the improvement does not come from excluding recordings. For the magnetometer-free estimator, fewer recordings converge (64.5% against 71.0%), so that its benefit is limited to the recordings that do. Adaptive scaling therefore partly compensates for an acceleration covariance that is too small without any tuning of the covariance, which is the situation in which a self-adjusting mechanism is most useful.

This benefit does not carry over to the tuned comparison of the preceding sections. A fixed covariance that is tuned on the recordings (36.5 times the literature value) lowers the error of the baseline by more than the adaptive scaling does, and the adaptive factor reaches values of the order of 10^3 on the recordings, so that its effective covariance becomes larger than that of the best fixed covariance. The post-convergence error also excludes the recordings that do not converge, and the slower acquisition of the adaptive method is therefore not reflected in Table 5-8; it is included in the overall RMSE of the preceding sections. The two comparisons answer different questions: Table 5-8 shows what adaptation offers when the covariance cannot be tuned, and the tuned comparison shows what it offers when it can.

5-5 Discussion

Consistency with the simulations. NIS-based rejection behaved as in the simulations: it removed only a small share of the updates and gave small but consistent improvements, which fits the removal of occasional inconsistent updates. Its threshold had to be far above the theoretical value, which shows that the innovations on the recorded data are not described by the assumed covariance. Adaptive scaling did not reproduce the simulated benefit. Its factor was far larger than in the simulations (Figure 5-3), so that the acceleration update received very little weight. This agrees with the control experiment of Section 4-9, in which a fixed covariance with a larger R_a reproduced most of the simulated benefit.

Why adaptive scaling fails on this data. Covariance matching attributes all excess innovation to the measurement. On recorded motion the innovation also contains the error of the state estimate and model errors that are not represented in $P_{t|t-1}$, so that a possible explanation is that the acceleration update is down-weighted precisely when it is needed to correct the state. Without an upper bound on α_t the update can then be effectively disabled, and the estimate drifts with the gyroscope. This explanation was not tested directly, but it is consistent with the time-resolved error (Section 5-4-5), with the failures at the hip, where the share of the time with low acceleration-based observability is highest (median 14.6% against 9.1% at the knee and 7.1% at the ankle), and with the finding that bounding the factor only recovers the baseline (Section 5-4-4).

Magnetic information, tuning and geometry. The largest and most consistent difference between the estimators is the conditional magnetic update, which is used only in a small share of the time. The tuning of the baseline has an effect of similar size, so that reliability-handling methods can only be assessed against a baseline that is tuned in the same way. The

marker-derived geometry lowers the error of every variant without changing the ranking, and the dependence on the geometry is smallest for MAJIC, which suggests that the magnetic update makes the estimate less sensitive to errors in the joint-centre vectors.

Limitations. The tuning is in-sample, and the evaluation contains 62 recordings of seven subjects from one dataset, so that the small gain of NIS-based rejection in particular should be seen as an upper estimate. The limitations are discussed in Section 6-1-3.

5-6 Summary of the Recorded-Data Evaluation

On the recorded data, the conditional magnetic update and the tuning of the baseline have the largest effect on the accuracy. NIS-based rejection gives a small but consistent further improvement for both estimators, and MAJIC with NIS-based rejection is the best variant. With a baseline that is tuned in the same way, adaptive covariance scaling does not improve the overall error and causes large failures in a minority of the recordings, mainly at the hip. Without tuning of the covariance, it lowers the post-convergence error (Table 5-8), so that its value lies in reducing the dependence on tuning and not in replacing it. These findings are combined with the simulation results, and used to answer the third subquestion and the main research question, in Chapter 6.

Discussion and Conclusions

This thesis investigated how online reliability handling of the joint-centre acceleration update affects IMU-based relative-orientation estimation. Adaptive covariance scaling and NIS-based rejection were applied to a magnetometer-free estimator and to an estimator that conditionally uses magnetic information, and were evaluated in controlled simulations (Chapter 4) and on recorded lower-extremity motion with all estimators tuned on the same data (Chapter 5). This chapter combines the findings of both evaluations, discusses their limitations, answers the research question and subquestions, and gives recommendations for future work.

6-1 Discussion

6-1-1 What the Mechanisms React to and When This Helps

In the simulations, both mechanisms responded to every source of uncertainty that was considered: the adaptive factor α_t and the rejection percentage increased with its severity. A response did not by itself lower the error. When the acceleration update remained informative but was systematically inconsistent with its model, as for errors in the assumed geometry and for a sustained disturbance, adaptive scaling lowered the error after convergence, relative to a fixed covariance with the literature value, without removing the update. A fixed covariance with a larger R_a reached a similar error after convergence (Section 4-9), and with tuned settings the overall error of adaptive scaling was equal to that of a tuned fixed covariance, except for a small advantage (1.3°) for the magnetometer-free estimator under the direction error and a small disadvantage (0.4° to 0.9°) for MAJIC. When the error was zero mean (accelerometer noise) or acted mainly on the prediction (gyroscope bias), the update provided the information that was needed, and reducing its weight did not lower the error. Neither mechanism could replace information that the motion did not contain. These cases illustrate the central problem of this thesis: a large innovation does not identify its cause, and the appropriate response depends on that cause.

The recorded data did not reproduce the simulated benefit of adaptive scaling. On the recordings the adaptive factor reached median values of 10^3 , one to two orders of magnitude

above the simulated values, so that the effective covariance of the acceleration update was 30 to 110 times that of the best fixed covariance. A possible explanation is that on recorded motion the innovation also contains state errors and model errors that the predicted innovation covariance does not represent, so that the covariance-matching rule attributes them to the measurement and down-weights the update that is needed to correct the state. It is consistent with the failures at the hip, with the finding that bounding the factor only recovers the baseline, and with the simulations in which the same method worked when the cause of the inconsistency was a systematic measurement error.

This does not make the self-adjusting property without value. A fixed covariance has to be chosen before the data are seen, and the tuning that gave the best baseline (a covariance 36.5 times the literature value) was found by a sweep over the recordings themselves. Adaptive scaling needs no such sweep: with the literature settings it lowered the post-convergence error and improved the retention without tuning the covariance, and in the simulations it lowered the error under geometry errors and sustained disturbances by a factor between 1.15 and 3.6 relative to the fixed covariance with the literature value. The comparison with the tuned baseline shows the limit of this property: on recorded motion, the adaptation does not reach the accuracy of a tuned covariance, and it fails for a minority of recordings.

NIS-based rejection behaved as in the simulations: it removed a small share of the updates (median 0.5% and 6.9% for the two estimators) and gave small but consistent improvements, which is consistent with the removal of occasional inconsistent updates. Its threshold had to be orders of magnitude above the theoretical value, which shows that the innovations on the recorded data are not described by the assumed covariance model.

6-1-2 Tuning and Magnetic Information

The largest effects in the recorded-data evaluation were not those of the reliability-handling methods. Tuning the acceleration covariance of the magnetometer-free baseline lowered its mean overall RMSE by 36%, and the conditional magnetic update lowered it further by about 4°, for every subject. The conclusions on adaptive scaling and NIS-based rejection are therefore valid relative to a baseline that has been tuned in the same way. A comparison with an untuned baseline would mainly measure the quality of the baseline, and an error that is computed only after convergence excludes the recordings in which an estimator fails, so that both can overstate the benefit of a method. The overall RMSE, which includes every recording, avoids the second effect.

6-1-3 Limitations

- *Simulations.* The simulations used two ideal rigid segments, one joint geometry, Gaussian noise, rich excitation for most studies, one source of uncertainty at a time and the literature parameters. The magnitudes of the imposed errors are design choices, and the absolute error values do not predict the accuracy in practice.
- *Recorded data.* Seven subjects and 62 recordings of one dataset were used, and the recordings are not independent. The subject-level bootstrap intervals take this into account, but with seven subjects they are approximate, and they do not account for

the in-sample tuning. The tuning was done on the same recordings, and the gain of NIS-based rejection for the magnetometer-free estimator did not hold in the leave-one-subject-out check (Appendix A-3), so that it should be seen as an upper estimate. The mean gain of MAJIC + NIS is partly due to a single recording (Section 5-4-2), whereas the consistency over recordings and subjects does not depend on it.

- *Adaptive tuning.* The adaptive method was tuned for the magnetometer-free estimator, and the MAJIC variant uses the same settings. The adaptive factor is a single scalar with a fixed lower bound of 0.1, and the covariance of the update is isotropic.
- *Metric.* The overall RMSE does not distinguish a slow start from a later loss of orientation, and the ranking of the estimators was not tested with other metrics.
- *Geometry.* The inertially estimated vectors were obtained from the complete recording, so that the sensor-only setting of the evaluation is offline, and the marker-derived vectors are not an exact reference.

6-2 Conclusions

Subquestion 1. The first subquestion asked how adaptive acceleration covariance scaling and NIS-based acceleration rejection affect estimation accuracy, convergence success and convergence time under controlled changes in motion excitation and sensor uncertainty. Neither mechanism changed the dependence on motion excitation: under low excitation only about 30 of 100 realizations of the three magnetometer-free variants converged, against all realizations of the MAJIC variants. Both mechanisms reacted to increased accelerometer noise and gyroscope bias, but neither lowered the error relative to the fixed-covariance estimator. NIS-based rejection led to much larger errors under gyroscope bias (median errors of approximately 64° and 13° at five times the nominal bias, with 95% and 54% of the updates rejected), whereas adaptive scaling stayed close to the baselines. Adaptive scaling slowed acquisition, with a median convergence time of 9.4 s against 1.2 s under rich excitation, and NIS-based rejection reduced the convergence success under medium excitation to 64 of 100 realizations.

Subquestion 2. The second subquestion asked whether the mechanisms can reduce the degradation caused by errors in the assumed IMU-to-joint-centre geometry. For the tested length errors of up to 20% and direction errors of up to 30° , adaptive scaling reduced the error relative to the fixed-covariance estimator by a factor between 1.15 and 3.6, for example from 7.4° to 2.9° for the magnetometer-free estimator and from 6.5° to 1.8° for MAJIC at a direction error of 30° , but did not remove the effect of the error and slowed acquisition. These reductions are relative to a fixed covariance with the literature R_a and are measured after convergence. A fixed covariance with a larger R_a reached a similar error after convergence (2.7° and 2.8° for the magnetometer-free estimator at a direction error of 30°), and with tuned settings the overall error of adaptive scaling was equal to that of the best fixed covariance for the magnetometer-free estimator, except for a small advantage under the direction error, and slightly higher for MAJIC. NIS-based rejection did not reduce the degradation.

Subquestion 3. The third subquestion asked whether the mechanisms improve accuracy in recorded human motion. With all estimators tuned on the same 62 OpenSense recordings and the inertially estimated geometry, adaptive scaling increased the mean overall RMSE from 9.85° to 27.37° for the magnetometer-free estimator and from 5.87° to 9.68° for MAJIC, with the largest failures at the hip. NIS-based rejection lowered the mean overall RMSE by 0.44° and 0.79° (9.85° to 9.42° and 5.87° to 5.08°), improving 68% and 73% of the recordings. With marker-derived vectors the improvement was larger and more consistent (81% and 87% of the recordings, all subjects). MAJIC with NIS-based rejection was the best variant, and the conditional magnetic update was the largest single difference between the estimators.

Main research question. The main research question asked how online adaptation of the joint-centre acceleration measurement covariance affects the accuracy and reliability of relative-orientation estimation, under which conditions it is beneficial, and how it compares with NIS-based rejection. In controlled conditions, adaptation lowers the error after convergence, relative to a fixed covariance with the literature value, when the acceleration update is informative but systematically inconsistent with its model, and not when the error is random, acts on the prediction, or when information is missing. Most of this benefit is reproduced by a fixed covariance with a larger value, so that adaptation mainly compensates for a covariance that is too small, and its slower acquisition increases the overall error unless a start-up delay is used. On recorded human motion, adaptation lowered the post-convergence error with the literature settings, without any tuning of the covariance. With a baseline tuned in the same way, it did not improve the overall error and caused failures in a minority of recordings, whereas NIS-based rejection gave a small but consistent improvement. Adaptation is therefore best seen as a way to reduce the dependence on tuning, and not as a replacement for it.

6-3 Recommendations and Future Work

- *Tune the baseline before adding reliability handling.* The acceleration and magnetic covariances should be tuned on data representative of the application, and reliability-handling methods should be compared with the tuned baseline using a metric that includes every recording.
- *Use the conditional magnetic update and a gated acceleration update.* On the recordings, the magnetic update and NIS-based rejection with a threshold calibrated to the innovation statistics gave the lowest errors. A separate validation set should be used to select the threshold.
- *Revisit the adaptive formulation.* Covariance matching that attributes all excess innovation to the measurement is not suitable on this data. Options are to include the state uncertainty in the attribution, to adapt per direction or per joint, to limit the adaptation to the periods in which the state is known to have converged, or to combine the adaptation with the observability measure. On the recorded data, adaptive scaling depended little on the nominal covariance and improved on the untuned baseline after convergence, so that it can be used when R_a cannot be tuned, but it requires a start-up delay and its factor should be monitored, because it does not replace the tuning.

- *Improve the joint-centre geometry.* The inertially estimated vector from the thigh sensor at the knee differed strongly from the marker-derived vector. More robust estimation, or estimating the vectors online, should be investigated together with its effect on reliability-handling methods.
- *Coordinate with the magnetic update.* The simulations showed that reducing the weight of the acceleration update can worsen the error caused by an inconsistent magnetic field. Handling of both updates together should be investigated.
- *Strengthen the evaluation.* More subjects, datasets and joints, a subject-level split for tuning and testing, and a statistical analysis that accounts for the dependence between recordings are needed to confirm the small effects reported here.

Appendix A

Appendix

A-1 Stationary Smartphone IMU Measurements

Table A-1: Stationary smartphone IMU measurement statistics. The reported standard deviations were calculated after subtracting the temporal mean of each sensor axis. The stationary recording duration after removal of the start and end segments was 246 s.

Sensor	Quantity	x	y	z	Mean std.
Accelerometer	Mean [m/s ²]	-0.0079	0.1640	9.7502	–
	Std. [m/s ²]	0.0066	0.0069	0.0109	0.00815
Gyroscope	Mean [rad/s]	-0.0015	-0.0004	-0.0033	–
	Std. [rad/s]	0.0013	0.0013	0.0011	0.00121
Magnetometer	Mean [μ T]	-25.1734	-22.6747	-6.2950	–
	Std. [μ T]	0.1883	0.1946	0.2968	0.2266

A-2 Motion-Excitation Parameters

Table A-2 specifies the parameters used to generate the three motion-excitation conditions according to the sum-of-sinusoids procedure described in Section 4-1-2.

A-3 Tuning of the Estimator Parameters

This appendix documents how the parameters of the estimators were tuned for the recorded-data evaluation of Chapter 5, and why the resulting values are used. The values are summarized in Table 5-1.

Table A-2: Parameters used for the low, medium, and rich motion-excitation conditions. Amplitude units are m/s^2 for f_{dyn}^G and rad/s for the angular-velocity signals; frequency ranges are given in Hz.

Level	Signal	N_f	Amplitude	Frequency	Axis scaling
Low	f_{dyn}^G	2	[0, 0.08]	[0.02, 0.08]	[1, 0.15, 0.10]
	ω_{com}^G	2	[0.01, 0.08]	[0.02, 0.10]	[1, 0.20, 0.10]
	ω_{rel}^G	2	[0, 0.03]	[0.02, 0.10]	[1, 0.15, 0.10]
Medium	f_{dyn}^G	3	[0.15, 0.60]	[0.05, 0.35]	[1, 0.70, 0.35]
	ω_{com}^G	3	[0.10, 0.45]	[0.05, 0.40]	[1, 0.75, 0.40]
	ω_{rel}^G	3	[0.05, 0.25]	[0.05, 0.40]	[1, 0.70, 0.35]
Rich	f_{dyn}^G	4	[0.40, 1.50]	[0.10, 0.80]	[1, 1, 1]
	ω_{com}^G	4	[0.25, 1.00]	[0.10, 0.90]	[1, 1, 1]
	ω_{rel}^G	4	[0.15, 0.70]	[0.10, 0.90]	[1, 1, 1]

A-3-1 Procedure

The objective of every step is the mean, over the 62 recordings, of the overall angular RMSE (5-9), obtained with the inertially estimated geometry. The mean is used because it keeps recordings with large errors visible. Each step evaluates all 62 recordings for every candidate value. The runs of a candidate are independent, and a run that failed or diverged would be assigned an error of 180° , so that it cannot lower the objective unnoticed.

The sweeps are one-dimensional, except for the adaptive estimator, and they do not cover regions that were clearly worse than their neighbours. A sweep starts at the literature value (or at the optimum of the preceding step) and continues in both directions in steps of half a decade (or, for the thresholds, a factor 1.25 to 1.5) until the objective has not improved for two consecutive values. A refinement around the best value follows with the midpoints between the best value and its neighbours.

The tuning is in-sample. As a check, the best value was also selected with each subject held out in turn, and the held-out subject was evaluated with the value chosen on the other six (leave-one-subject-out). The pooled held-out objective in the last column of Table A-3 shows how optimistic the in-sample optimum is.

Table A-3: Summary of the tuning steps: literature value (or estimator without the mechanism) $\rightarrow$ best value, and the mean overall angular RMSE over the 62 recordings in degrees for both. LOSO: pooled leave-one-subject-out objective of the best value.

Parameter	Start $\rightarrow$ best	Objective $[\circ]$	LOSO $[\circ]$
R_a (Weygers)	$2.5 \times 10^{-3} \rightarrow 9.13 \times 10^{-2}$	$15.46 \rightarrow 9.85$	9.90
R_{mag} (MAJIC)	$9.0 \times 10^{-4} \rightarrow 2.85 \times 10^{-4}$	$6.41 \rightarrow 5.87$	6.18
NIS threshold (Weygers)	$11.34 \rightarrow 4000$	$65.49 \rightarrow 9.42$	10.11
NIS threshold (MAJIC)	$\text{none} \rightarrow 751$	$5.87 \rightarrow 5.08$	5.08
N_w , delay (adaptive)	$50, 0 \text{ s} \rightarrow 30, 10 \text{ s}$	$- \rightarrow 27.37$	27.37

A-3-2 Acceleration Covariance of the Baseline

The acceleration covariance $R_a = r I_3$ of the magnetometer-free baseline was varied from 0.01 to 10^4 times the literature value 2.5×10^{-3} . Figure A-1 shows the result. The mean overall RMSE decreases from 15.46° at the literature value to 9.85° at $r = 9.13 \times 10^{-2}$, which is 36.5 times larger. The error is clearly higher at both ends of the range (50° at 0.01 and 10^4 times the literature value), so that the optimum is inside the tested range. The optimum is narrow in the refinement (r between 9.13×10^{-2} and 9.81×10^{-2} is within 1% of the best), but values between 10 and 100 times the literature value all give a mean error between 10.4° and 11.4° , so that the choice is not critical within this plateau. The leave-one-subject-out error is 9.90° .

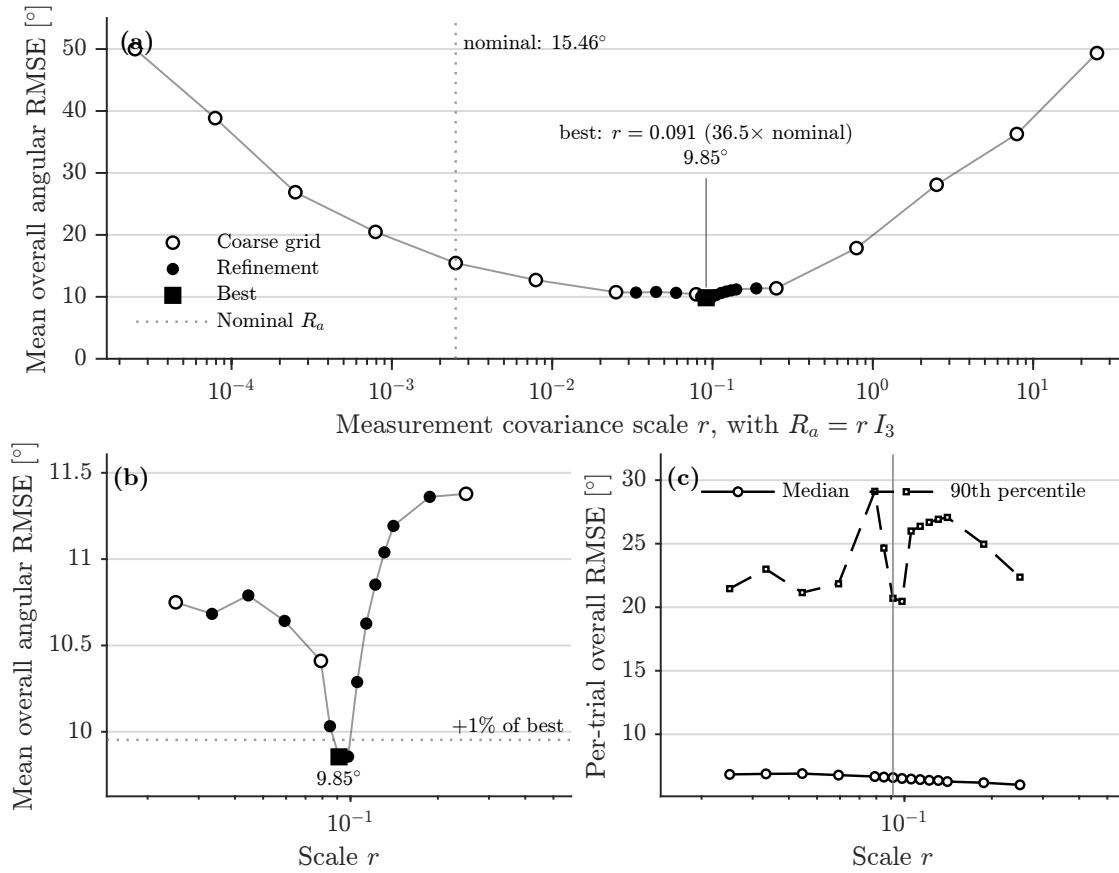


Figure A-1: Tuning of the acceleration covariance $R_a = r I_3$ of the baseline. (a) Mean overall RMSE for all tested values, (b) zoom around the optimum, (c) median and 90th percentile of the per-recording error around the optimum.

A-3-3 Magnetic Covariance of MAJIC

With R_a fixed at the tuned value, the magnetic covariance $R_{\text{mag}} = r_m I_3$ was varied from 0.03 to 10 times the literature value 9.0×10^{-4} (Figure A-2). The mean overall RMSE decreases from 6.41° to 5.87° at $r_m = 2.85 \times 10^{-4}$ (0.32 times the literature value). The objective is flat around the optimum (r_m between 2.1×10^{-4} and 3.8×10^{-4} is within 1% of the best). The median paired gain relative to the literature value is close to zero, and the improvement

is found in 30 of the 62 recordings, so that the gain is due to a minority of the recordings. The leave-one-subject-out error is 6.18° .

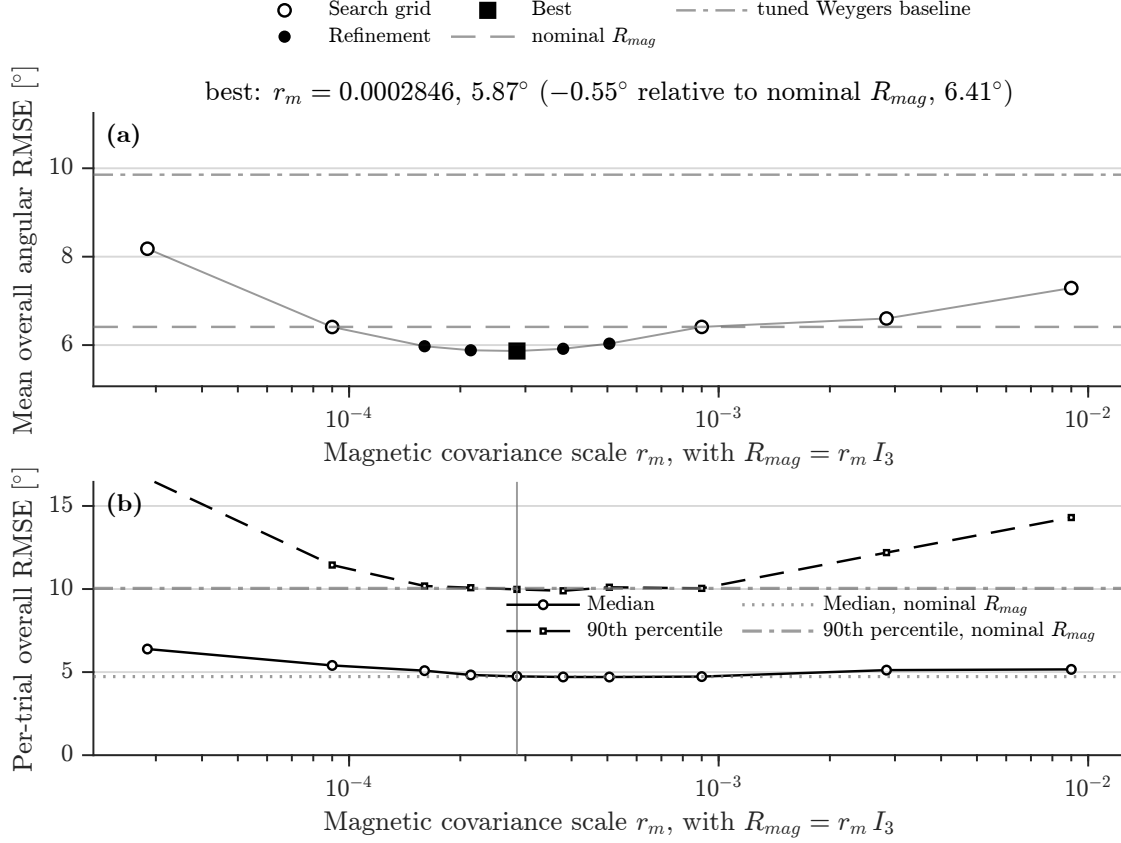


Figure A-2: Tuning of the magnetic covariance $R_{mag} = r_m I_3$ of the MAJIC baseline. (a) Mean overall RMSE, (b) median and 90th percentile.

A-3-4 NIS Thresholds

The theoretical 99% threshold (11.34) rejects the acceleration update almost continuously on the recorded data and gives a mean overall RMSE of 65.5° for the magnetometer-free estimator. The thresholds were therefore tuned as a free parameter, with R_a (and R_{mag}) fixed at their tuned values. Figure A-3 shows the magnetometer-free estimator. The best threshold is 4000 (9.42° against 9.85° without rejection), and all thresholds between 3000 and 13000 are within 1.5% of the best. Low thresholds reduce the median error but increase the 90th percentile and the maximum, which means that too much rejection removes updates that are needed. The improvement is found in 42 of the 62 recordings, but the leave-one-subject-out error (10.11°) is above the error without rejection, so that the gain for this estimator is not robust to the choice of the threshold.

For MAJIC (Figure A-4) the best threshold is 751 (5.08° against 5.87°). The objective is very flat between 538 and 1050 (within 0.6%), the improvement is found in 45 of the 62 recordings, and the leave-one-subject-out error equals the in-sample value. With the tuned covariances, the share of rejected updates is 0.5% (median) for Weygers and 6.9% for MAJIC.

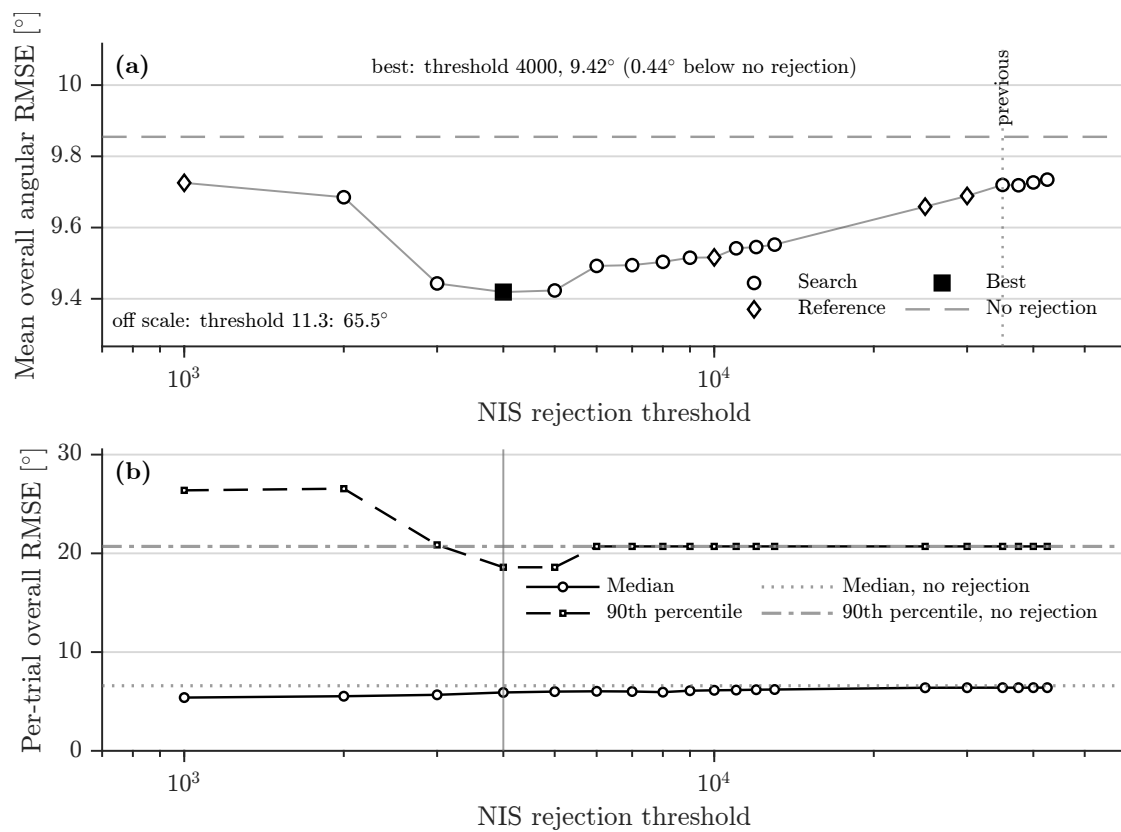


Figure A-3: Tuning of the NIS threshold of Weygers + NIS. (a) Mean overall RMSE for all tested thresholds. The theoretical threshold (off scale) gives 65.5°. (b) Median and 90th percentile.

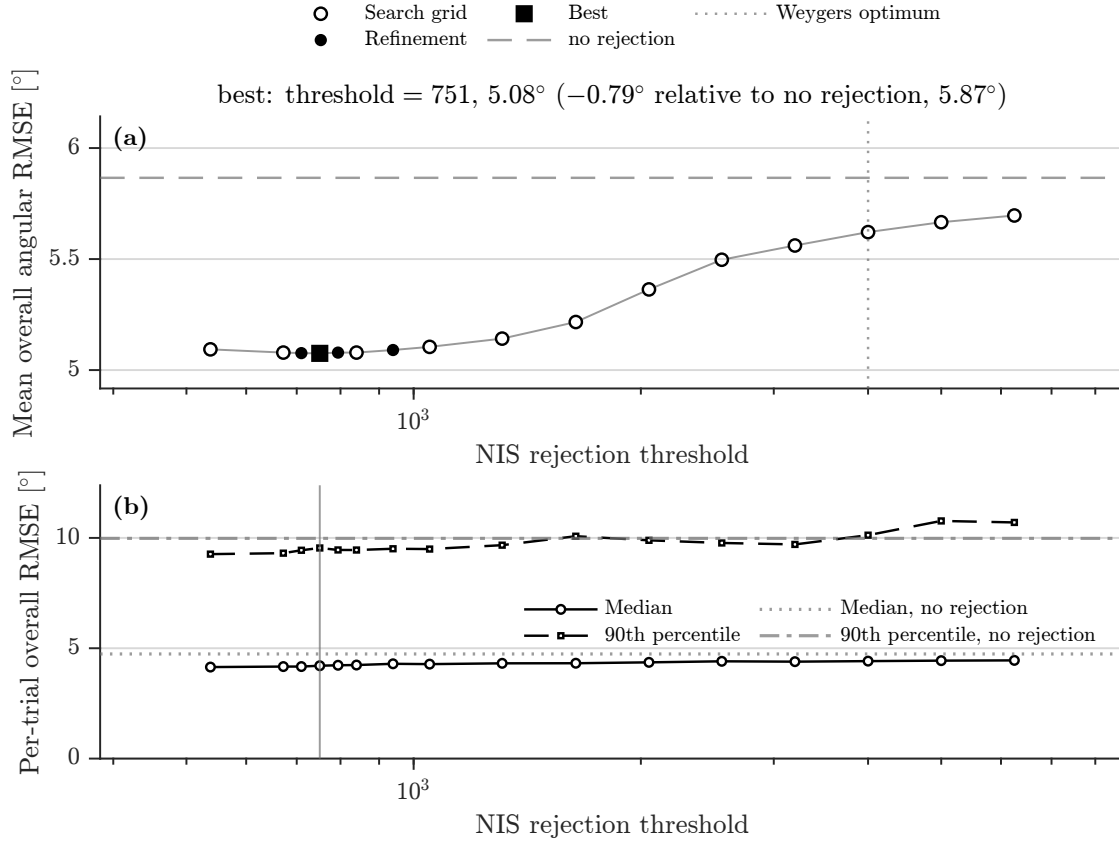


Figure A-4: Tuning of the NIS threshold of MAJIC + NIS.

A-3-5 Adaptive Covariance Scaling

The adaptive estimator has more parameters than the baseline, and these were tuned jointly: the acceleration covariance R_a (from 0.01 to 10 times the baseline optimum), the window length N_w (10 to 75 samples, with refinement) and the start-up delay of the adaptation (0, 5 and 10 s). Figure A-5 shows the map over R_a and N_w . The window and the delay have a larger effect than R_a (all 119 candidates give a mean overall RMSE between 27.4° and 45.6°), and the lowest mean overall RMSE (27.37° for $R_a = 3.85 \times 10^{-3}$, $N_w = 30$ and a delay of 10 s) is still 17.5° above the baseline. The estimator improves on the baseline in 21 of the 62 recordings, with a median paired difference of −1.48°.

An upper bound on the adaptive factor was tested as a remedy. The bound $\alpha_{\max}$ was varied from 0.5 to 20 together with the window and the delay. The best bound is $\alpha_{\max} = 1$ ($N_w = 75$, delay 10 s), which gives a mean overall RMSE of 9.81° against 9.85° for the baseline. The factor can then only decrease the covariance, and the result is practically equal to the baseline, so that the bounded estimator is not used in the main comparison.

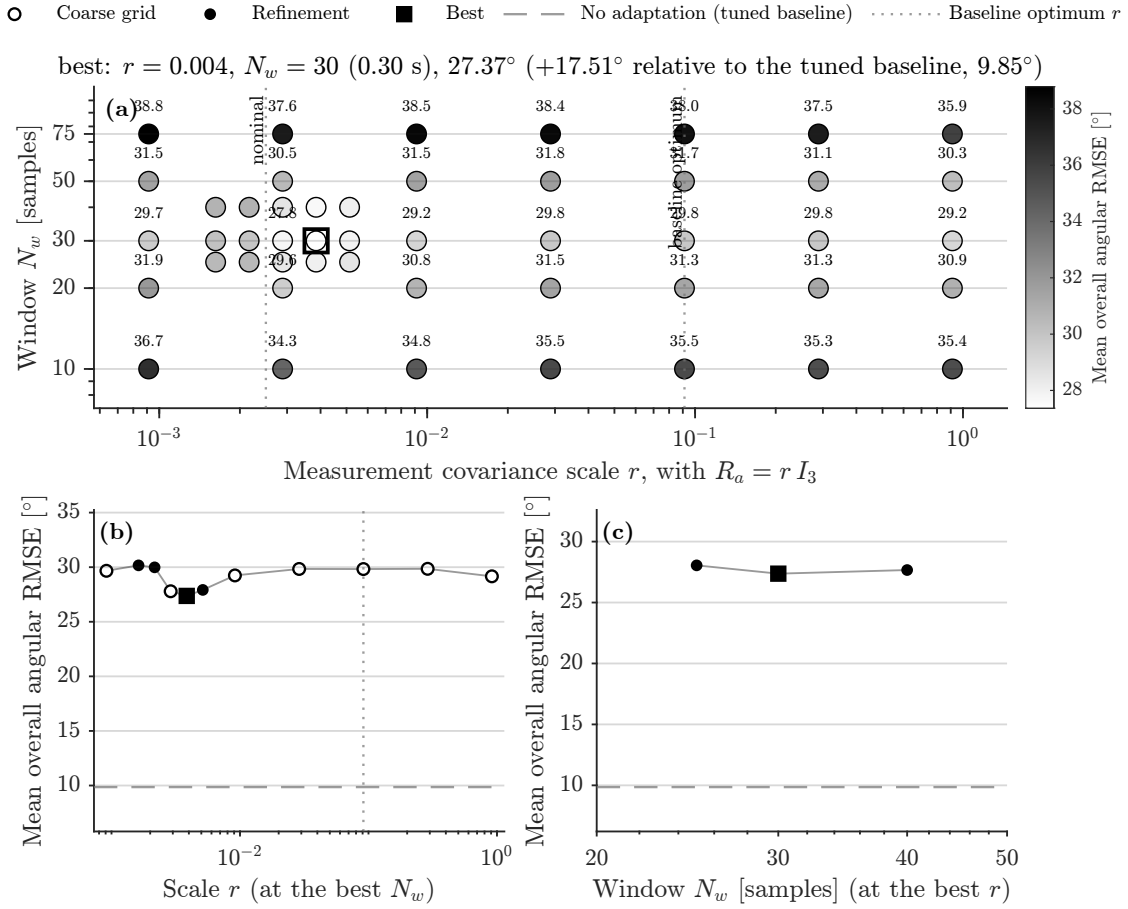


Figure A-5: Tuning of the adaptive estimator. (a) Mean overall RMSE for all tested combinations of R_a and N_w (darker is larger error), (b) and (c) slices through the best combination.

A-4 Sensitivity of the Convergence Criterion

This analysis was performed on the recorded data at an early stage of the work, with the literature parameters of the estimators and before the tuning described in Appendix A-3. It supports the convergence criterion that is used for the convergence rate and the convergence time in the simulation study (Chapter 4). Convergence is not used as a performance measure in Chapter 5: the strong dependence of the convergence rate on its definition, which is shown here, is one reason why the overall angular RMSE is used there. The analysis includes all 71 available recordings, also those of the two subjects recorded at 40 Hz, which are excluded from Chapter 5.

In the simulation study and in this analysis, convergence was defined as an angular error below 10° continuously for at least 5 s. To assess the sensitivity of the results to this definition, angular-error thresholds of 5° , 7.5° , 10° , 12.5° , and 15° were combined with required hold durations of 2, 5, 10, and 20 s. The analysis was performed using the Weygers and MAJIC baseline estimators for both marker-derived and inertially estimated sensor-to-joint geometry.

Figure A-6 shows the fraction of trials classified as converged for each angular-error threshold

and hold duration. As expected, convergence becomes more likely for a larger error threshold and less likely when the estimate is required to remain below the threshold for longer. The same behaviour is summarized in Figure A-7, where the selected $10^\circ/5$ s criterion is highlighted.

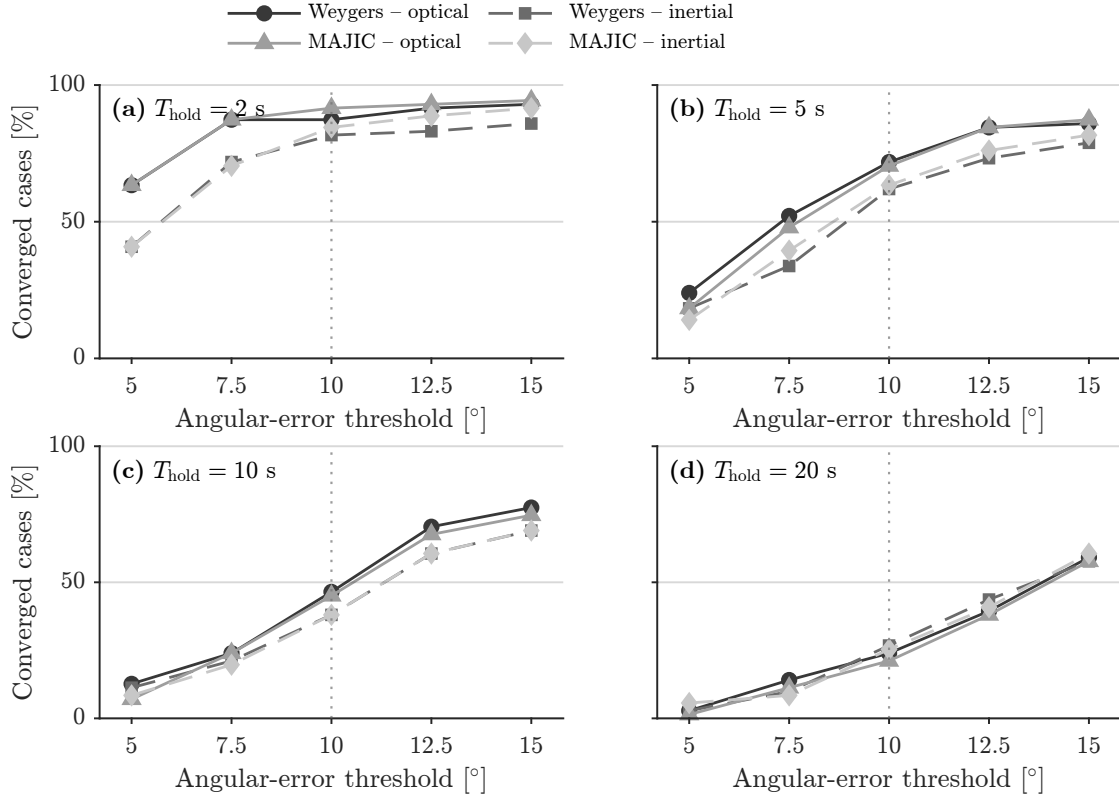


Figure A-6: Sensitivity of the baseline convergence rate to the angular-error threshold and required hold duration. Results are shown for the Weygers and MAJIC baseline estimators using marker-derived and inertially estimated sensor-to-joint geometry.

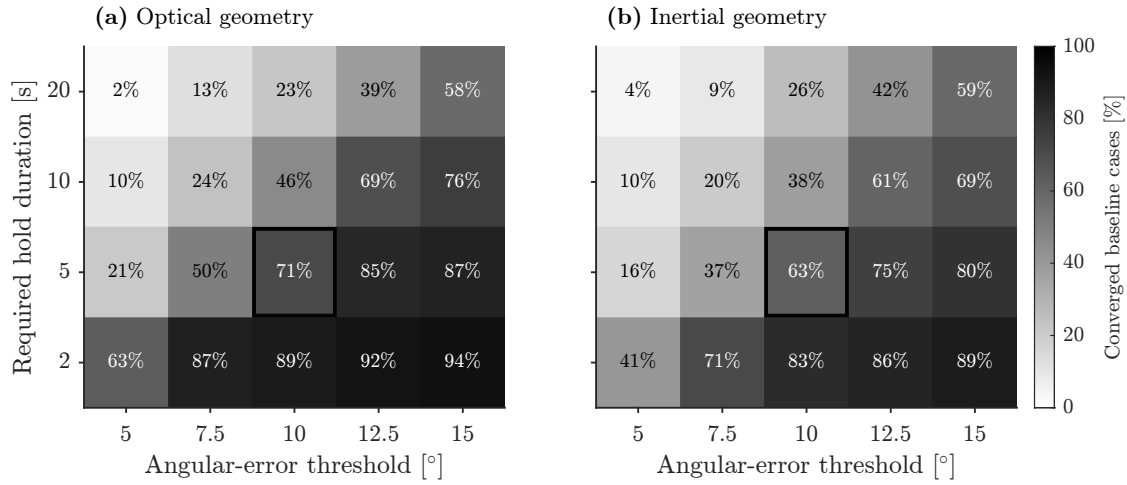


Figure A-7: Baseline convergence rate for the investigated combinations of angular-error threshold and required hold duration. The selected $10^\circ/5$ s convergence definition is outlined.

Among the trials classified as converged, stricter definitions also result in later convergence times, as shown in Figure A-8. This is a direct consequence of requiring the angular error to remain within a smaller bound and/or to satisfy that bound for a longer uninterrupted period. Consequently, very strict definitions both reduce the number of trials available for post-convergence analysis and delay the start of the corresponding evaluation interval.

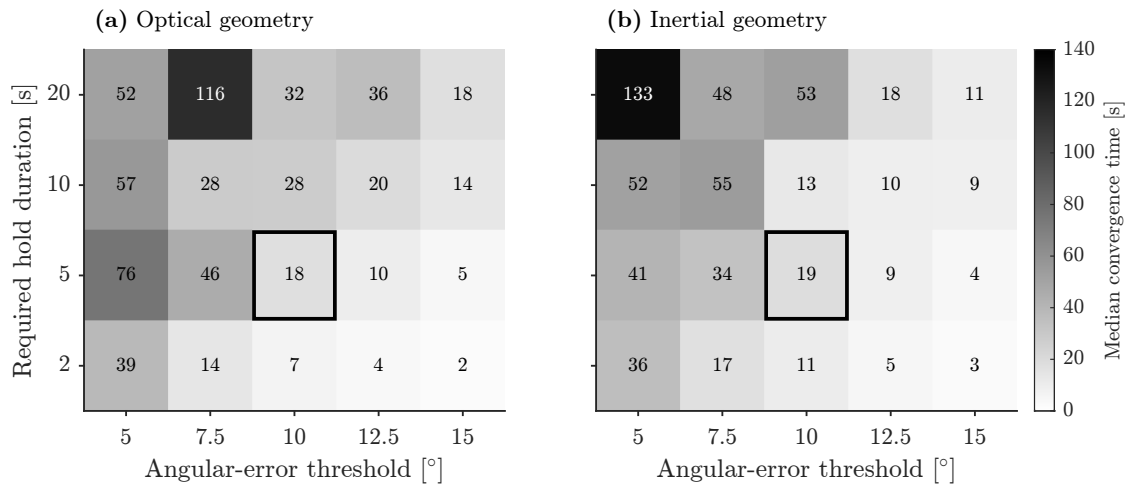


Figure A-8: Median convergence time among converged baseline trials for each candidate convergence definition. The selected $10^\circ/5$ s definition is outlined.

Overall, the selected $10^\circ/5$ s criterion provides a compromise between the more permissive definitions, which can classify brief reductions in error as convergence, and the stricter definitions, which substantially reduce the number of converged trials. This criterion was therefore used for the convergence rate and time in the simulation study.

Acknowledgement of AI Use

I acknowledge the use of Claude and ChatGPT to support the development of this thesis by assisting with writing and refining language, and by assisting with programming and debugging.

Bibliography

- [1] M. Al Borno, J. O'Day, V. Ibarra, J. Dunne, A. Seth, A. Habib, C. Ong, J. Hicks, S. Uhlrich, and S. Delp. OpenSense: An open-source toolbox for inertial-measurement-unit-based measurement of lower extremity kinematics over long durations. *Journal of NeuroEngineering and Rehabilitation*, 19(1):22, 2022.
- [2] C. M. N. Brigante, N. Abbate, A. Basile, A. C. Faulisi, and S. Sessa. Towards miniaturization of a MEMS-based wearable motion capture system. *IEEE Transactions on Industrial Electronics*, 58(8):3234–3241, 2011.
- [3] A. C. B. Chiella, B. O. S. Teixeira, and G. A. S. Pereira. Quaternion-based robust attitude estimation using an adaptive unscented Kalman filter. *Sensors*, 19(10):2372, 2019.
- [4] W. H. K. De Vries, H. E. J. Veeger, C. T. M. Baten, and F. C. T. van der Helm. Magnetic distortion in motion labs, implications for validating inertial magnetic sensors. *Gait & Posture*, 29(4):535–541, 2009.
- [5] H. Dejnabadi, B. M. Jolles, and K. Aminian. A new approach to accurate measurement of uniaxial joint angles based on a combination of accelerometers and gyroscopes. *IEEE Transactions on Biomedical Engineering*, 52(8):1478–1484, 2005.
- [6] E. Dorschky, M. Nitschke, A.-K. Seifer, A. J. van den Bogert, and B. M. Eskofier. Estimation of gait kinematics and kinetics from inertial sensor data using optimal control of musculoskeletal models. *Journal of Biomechanics*, 95:109278, 2019.
- [7] J. D. Hol. *Sensor fusion and calibration of inertial sensors, vision, ultra-wideband and GPS*. PhD thesis, Linköping University, 2011.
- [8] M. Kok, K. Eckhoff, I. Weygers, and T. Seel. Observability of the relative motion from inertial data in kinematic chains. *Control Engineering Practice*, 125:105206, 2022.
- [9] M. Kok, J. D. Hol, and T. B. Schön. Using inertial sensors for position and orientation estimation. *Foundations and Trends in Signal Processing*, 11(1–2):1–153, 2017.

- [10] D. Laidig and T. Seel. VQF: Highly accurate IMU orientation estimation with bias estimation and magnetic disturbance rejection. *Information Fusion*, 91:187–204, 2023.
- [11] C. J. Lee and J. K. Lee. Inertial motion capture-based wearable systems for estimation of joint kinetics: A systematic review. *Sensors*, 22(7):2507, 2022.
- [12] J. K. Lee and T. H. Jeon. IMU-based but magnetometer-free joint angle estimation of constrained links. In *Proceedings of the IEEE SENSORS Conference*, pages 1–4, 2018.
- [13] J. K. Lee, E. J. Park, and S. N. Robinovitch. Estimation of attitude and external acceleration using inertial sensor measurement during various dynamic conditions. *IEEE Transactions on Instrumentation and Measurement*, 61(8):2262–2273, 2012.
- [14] J. L. McGinley, R. Baker, R. Wolfe, and M. E. Morris. The reliability of three-dimensional kinematic gait measurements: A systematic review. *Gait & Posture*, 29(3):360–369, 2009.
- [15] T. McGrath and L. Stirling. Body-worn IMU-based human hip and knee kinematics estimation during treadmill walking. *Sensors*, 22(7):2544, 2022.
- [16] National Oceanic and Atmospheric Administration. Magnetic field calculators (WMM). <https://www.ngdc.noaa.gov/geomag/calculators/magcalc.shtml>, 2026. National Centers for Environmental Information (NCEI). Accessed: 2026-09-28.
- [17] R. Piché. Online tests of Kalman filter consistency. *International Journal of Adaptive Control and Signal Processing*, 30(1):115–124, 2016.
- [18] D. Roetenberg, H. J. Luinge, C. T. M. Baten, and P. H. Veltink. Compensation of magnetic disturbances improves inertial and magnetic sensing of human body segment orientation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 13(3):395–405, 2005.
- [19] T. Seel, T. Schauer, and J. Raisch. Joint axis and position estimation from inertial measurement data by exploiting kinematic constraints. In *Proceedings of the IEEE International Conference on Control Applications*, pages 45–49, 2012.
- [20] P. B. Shull, W. Jirattigalachote, M. A. Hunt, M. R. Cutkosky, and S. L. Delp. Quantified self and human movement: A review on the clinical impact of wearable sensing and feedback for gait analysis and intervention. *Gait & Posture*, 40(1):11–19, 2014.
- [21] S. Skov, K. Werling, J. O’Day, J. Hicks, C. K. Liu, and S. Delp. Estimating magnetic field at joint centers reduces kinematic errors in inertial motion capture. *Research Square*, 2026. Preprint.
- [22] Y. S. Suh. Orientation estimation using a quaternion-based indirect Kalman filter with adaptive estimation of external acceleration. *IEEE Transactions on Instrumentation and Measurement*, 59(12):3296–3305, 2010.
- [23] L. Sy, M. Raitor, M. Del Rosario, H. Khamis, L. Kark, N. H. Lovell, and S. J. Redmond. Estimating lower limb kinematics using a reduced wearable sensor count. *IEEE Transactions on Biomedical Engineering*, 68(4):1293–1304, 2020.

-
- [24] M. Verhaegen and V. Verdult. *Filtering and System Identification: A Least Squares Approach*. Cambridge University Press, 2007.
 - [25] I. Weygers, M. Kok, H. De Vroey, T. Verbeerst, M. Versteyhe, H. Hallez, and K. Claeys. Drift-free inertial sensor-based joint kinematics for long-term arbitrary movements. *IEEE Sensors Journal*, 20(14):7969–7979, 2020.
 - [26] Y. Yang and W. Gao. An optimal adaptive Kalman filter. *Journal of Geodesy*, 80(4):177–183, 2006.
 - [27] J.-H. Zhang, P. Li, C.-C. Jin, W.-A. Zhang, and S. Liu. A novel adaptive Kalman filtering approach to human motion tracking with magnetic-inertial sensors. *IEEE Transactions on Industrial Electronics*, 67(10):8659–8669, 2019.

