Check for updates

METHOD ARTICLE

REVISED JointTracker: Real-time inertial kinematic chain tracking with joint position estimation

[version 2; peer review: 3 approved]

Bertram Taetz[1,2], Michael Lorenz[1], Markus Miezal[1], Didier Stricker[1], Gabriele Bleser-Taetz[1,2]

[1]Augmented Vision, German Research Center for Artificial Intelligence, Kaiserslautern, Rhineland-Palatinate, 67663, Germany
[2]IT & Engineering, International University of Applied Sciences, Erfurt, Thuringia, 99084, Germany

v2 **First published:** 19 Feb 2024, **4**:33
https://doi.org/10.12688/openreseurope.16939.1
Latest published: 22 Jan 2025, **4**:33
https://doi.org/10.12688/openreseurope.16939.2

Abstract

In-field motion capture is drawing increasing attention due to the multitude of application areas, in particular for human motion capture (HMC). Plenty of research is currently invested in camera-based markerless HMC, however, with the inherent drawbacks of limited field of view and occlusions. In contrast, inertial motion capture does not suffer from occlusions, thus being a promising approach for capturing motion outside the laboratory. However, one major challenge of such methods is the necessity of spatial registration. Typically, during a predefined calibration sequence, the orientation and location of each inertial sensor are registered with respect to an underlying skeleton model. This work contributes to calibration-free inertial motion capture, as it proposes a recursive estimator for the simultaneous online estimation of all sensor poses and joint positions of a kinematic chain model like the human skeleton. The full derivation from an optimization objective is provided. The approach can directly be applied to a synchronized data stream from a body-mounted inertial sensor network. Successful evaluations are demonstrated on noisy simulated data from a three-link chain, real lower-body walking data from 25 young, healthy persons, and walking data captured from a humanoid robot.

Keywords

inertial motion capture, 3D kinematics, 3D human pose estimation, online joint position estimation, calibration-free, recursive state estimation, parameter estimation

Open Peer Review

Approval Status ✓ ✓ ✓

	1	2	3
version 2 (revision) 22 Jan 2025	✓ view	✓ view	
version 1 19 Feb 2024	? view	? view	✓ view

1. **Manon Kok**, Delft University, Delft, The Netherlands
2. **Simon Bachhuber**, FAU Erlangen-Nuernberg, Nürnberg, Germany
3. **Lauro Armando Contreras Rodríguez**, Universidad Autonoma de Sinaloa, Culiacán, Mexico

Any reports and responses or comments on the article can be found at the end of the article.



This article is included in the Horizon Europe gateway.



This article is included in the Horizon 2020 gateway.

Corresponding author: Bertram Taetz (Bertram.Taetz@iu.org)

Author roles: Taetz B: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Lorenz M**: Data Curation, Investigation, Validation, Writing – Review & Editing; **Miezal M**: Data Curation, Software, Visualization, Writing – Review & Editing; **Stricker D**: Funding Acquisition, Project Administration, Resources, Writing – Review & Editing; **Bleser-Taetz G**: Funding Acquisition, Methodology, Project Administration, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No [826304](Personalized Body Sensor Networks with Built-In Intelligence for Real-Time Risk Assessment and Coaching of Ageing workers, in all types of working and living environments [BIONIC]) and from the European Union's Horizon Europe research and innovation programme under grant agreement No [101092889] (Embodied Social Experiences in Hybrid Shared Spaces [SHARESPACE]) as well as the BMBF projects OrthoSuPer (Sichere Datenplattform und intelligente Sensorik für die Versorgung der Zukunft in der Orthopädie, project ID FKZ 13 GW0564 A-F) and DECODE (Continual learning zur visuellen und multimodalen Erkennung menschlicher Bewegungen und des semantischen Umfeldes in alltäglichen Umgebungen, project ID 01IW21001).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2025 Taetz B *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Taetz B, Lorenz M, Miezal M *et al.* **JointTracker: Real-time inertial kinematic chain tracking with joint position estimation [version 2; peer review: 3 approved]** Open Research Europe 2025, **4**:33
https://doi.org/10.12688/openreseurope.16939.2

First published: 19 Feb 2024, **4**:33 https://doi.org/10.12688/openreseurope.16939.1

REVISED **Amendments from Version 1**

We sincerely thank the reviewers for providing detailed and valuable comments, which have greatly helped us to revise and improve the article! Based on these comments, the following main revisions were made: 1) Reduced emphasis on camera-based tracking in the abstract and introduction. 2) Reformulated the contributions for greater clarity. 3) Clarified the differences of the proposed approach compared to previous joint estimation approaches. 4) Justified the mathematical derivation of the recursive estimator as a theoretical contribution of the article. 5) Clarified that no IMU-to-segment calibrations are required by the proposed algorithm. 6) Explained the handling of IMU biases. 7) Added a more detailed description of test persons in the human lower body scenario. 8) Removed Figure 5 and Figure 7 and presented the summarized results in the main text. 9) Justified the usage of joint connection measurement models on both positional and velocity levels and added an experimental comparison with a variant of the proposed algorithm using an acceleration-level model. 10) Better explained "real-time capable" and added exemplified processing times to the evaluation.

Any further responses from the reviewers can be found at the end of the article

1 Introduction

Motion capture, in particular human motion capture (HMC), is a well-studied field with multiple application areas like healthcare, sports[1,2], human-machine interaction[3], workplace analysis[4,5], robotics, teleworking, additive manufacturing and human safety[6]. Numerous HMC methods utilize camera-based approaches, which, however, face limitations such as occlusions caused by external factors or self-occlusions. Inertial HMC (IHMC) does not suffer from occlusions and is a promising approach for in-field measurements. It uses multiple inertial sensors (inertial measurement units, IMUs) mounted on body segments, using e.g., straps or clothing integration[7–9]. The motion is typically deduced via a suitable state estimation approach in combination with a personalized biomechanical model[10]. IHMC is an approach for in-field HMC that does not suffer from occlusions. However, it comes with other challenges, such as integration drift[11], sensitivity to magnetic disturbances[12] or calibration issues. The latter are typically related to segment lengths and sensor poses with respect to the anatomical reference frames of the associated body segments, the so-called IMU-to-segment (I2S) calibrations[13–16]. Some of the mentioned challenges are connected, e.g. omitting magnetometer measurements can lead to orientation drift, while magnetometer usage can violate the often-used constant magnetic field assumption and lead to incorrect state estimates. The I2S calibration is a major concern for assessing and monitoring anatomical joint angles since I2S orientation errors linearly propagate into joint-angle errors[13]. I2S calibration is currently considered the main challenge for IHMC in biomechanical motion analysis[17], and an automated (self-calibrating) approach can be regarded as crucial for reliable in-field usage of the inertial tracking technology. Self-calibrating IHMC approaches include the direct I2S alignment estimation. Some promising approaches for I2S orientation and position calibration for different body parts are proposed in 14,15,18,19. Direct estimation of biomechanical joint angles without I2S alignment is proposed in 20,21, but via exploiting a specific calibration motion and assumptions / heuristics concerning the human skeleton and walking kinematics.

This work derives and proposes an algorithm for self-calibrating inertial motion capture based on a general kinematic chain model representation consisting of global IMU poses with local (IMU-centered) joint positions, as shown in Figure 1 (left). The model is applicable to any kinematic chain and not restricted to or specialized to a human skeleton. Note, joint position refers here to the actual positional offset between joint and adjacent IMU, not to joint angles. Hence, the focus is on self-calibration of the joint positions in the IMU frames or, in other words, on a calibration-free approach with respect to the I2S orientations and positions. The proposed algorithm can be used to track joint positions, i.e. joint trajectories, for motion analysis, and it delivers consistent information with camera-based markerless 3D joint tracking, where the majority of approaches relies on joint position detections on the image plane (2D) with a subsequent lifting of the joint positions into 3D space[22]. The body pose can then be illustrated by connecting the segments at the joints, which is common practice in the vision community[23], and which can be done with the tracking results of our proposed algorithm as well. Moreover, the proposed algorithm allows estimating IMU trajectories in one coordinate system with minimal magnetometer usage (usually one measurement sample only for initialization). For the skeleton state estimation and joint parameter identification, we propose a consistent recursive estimation of IMU poses in a global coordinate frame and joint positions in the IMU coordinate frames.

Previous approaches to joint position estimation typically work offline, independently per joint and exploit the measured accelerations[18,25,26]. A recent integration-based (using velocities and positions) example operating on a complete kinematic chain is 20. It estimates IMU poses, multiple joint positions, and hinge axes, however, supported by segment length information and in an offline manner. In contrast to these methods, our

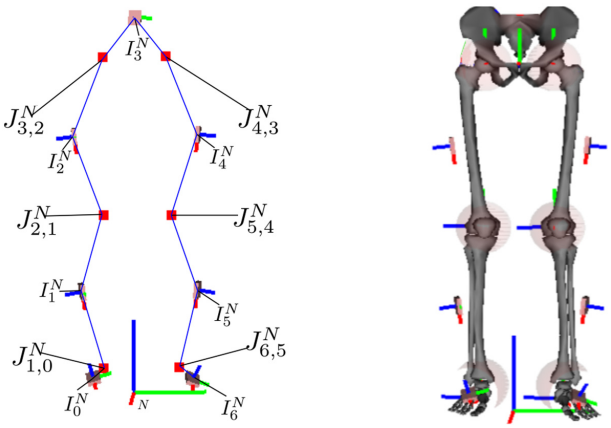

Figure 1. The left image is an illustration of the kinematic model with joints (red squares), IMU poses, and connections between IMU and joint centers (blue lines), which is tracked using the proposed approach. The right image shows the corresponding biomechanical skeleton model with I2S calibrations, based on 24, where the joints are emphasized with semi-transparent spheres.

integration-based approach works online for a complete kinematic chain without requiring prior information. Moreover, we also investigate an indicator for convergence of the joint position parameters, which none of the previous approaches did.

In summary, the main contributions of this work are:

- mathematical derivation of a consistent optimization-based recursive (real-time capable) state estimator, which uses only data from the current timestep, for
- calibration-free simultaneous online estimation of IMU kinematics and joint (or fixed) positions for three degrees of freedom (DoF) joints for kinematic chain models (so that the chain pose can be illustrated by connecting the segments at the joints). This results in online segment length estimation for all segments connecting two joints.
- convergence indicator for the motion-dependent joint position parameters
- minimal magnetometer usage (only for initialization).

In the following, Section 2 introduces the mathematical notation and derives the recursive state estimator, including the corresponding models, while also addressing initialization and observability. Section 3 evaluates the method on three different experimental setups: 1) simulated IMU data from a three-link chain fixed in space, 2) real lower-body walking data from a group of young, healthy persons, and 3) real lower-body walking data from a humanoid robot. Section 4 concludes this work and discusses open problems and further research.

2 Methods

In this section, we present the notation (Section 2.1), the problem statement, where a batch estimator is defined (Section 2.2), the derivation of the optimization-based recursive state estimator from the batch estimator (Section 2.3), and the motion (Section 2.4) and measurement models (Section 2.5) used to estimate the state of the kinematic chain model, illustrated in Figure 1 and Figure 2. Note, the mathematical derivation of the recursive state estimator extends the work presented in 27 for a nonlinear motion model and is a theoretical contribution of this article.

2.1 Notation

The following description uses Figure 2 as a reference illustration. The navigation frame N has its z-axis aligned with gravity, and its x-axis pointing towards magnetic north. It is the reference frame to which all IMUs are registered. Each IMU $i \in \mathcal{I}$, $|\mathcal{I}| = N_I$, has its own coordinate system, I_i, which is related to the navigation frame via a pose transformation, e.g., using a homogeneous matrix

$$T^{NI_i} = \begin{pmatrix} R^{NI_i} & I_i^N \\ 0_3^T & 1 \end{pmatrix}. \tag{1}$$

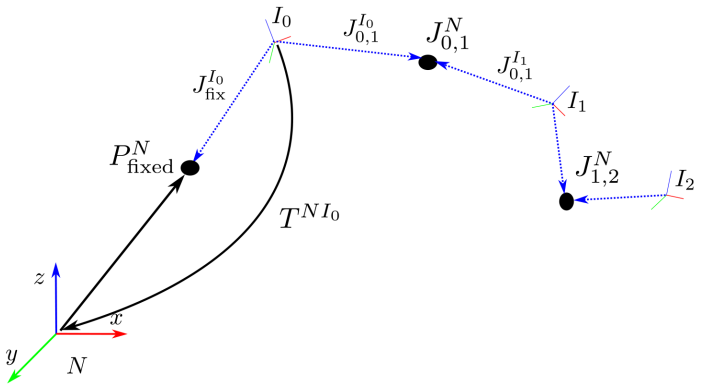

Figure 2. Coordinate systems and notation.

Here, $I_i^N \in \mathbb{R}^3$ denotes the position of the IMU in the navigation frame. The orientation matrix $R^{NI_i} \in \mathscr{SO}(3)$ denotes the IMU rotation relative to the navigation frame. Hence, a vector that is right multiplied to R^{NI_i} is transformed from I_i to N. The inverse of the homogeneous matrix is

$$\left(T^{NI_i}\right)^{-1} = T^{I_i N} = \begin{pmatrix} R^{I_i N} & -R^{I_i N} I^N \\ \mathbf{0}_3^T & 1 \end{pmatrix}. \tag{2}$$

The inverse rotation matrix is the transpose and denoted by $(R^{NI_i})^T = R^{I_i N}$. As rotation parametrizations we use Modified Rodriguez Parameters (MRP) χ, unit quaternions $q = (q_0, q_1, q_2, q_3)$, where q_0 represents the scalar part, and rotation matrices R. Conversions are indicated with e.g. $\chi(q)$, refer to 28 for the equations. The position of a joint connecting two neighbouring IMUs I_i and I_j is denoted in the navigation frame by $J^N_{i,j}$, $(i, j) \in \mathscr{J}$, the set of all joint indices. This joint position is, however, estimated separately in the frames of the two adjacent IMUs as $J^{I_i}_{i,j}$ and $J^{I_j}_{i,j}$. For our first test case of a simulated three link chain, one segment is connected to the environment with a fixed point, which is also estimated. The indices for additional (optional) fixed point estimates, $J^{I_i}_{\text{fix}}$ are collected in $i \in \mathscr{F}$.

The symbol X_t denotes the collection of all hidden states to be estimated at time step t and $X_{1:N_T} = \{X_t\}_{t=1}^{N_T}$ denotes a complete time sequence of states consisting of N_T time instances. Hence, our kinematic model is defined via IMUs and connecting joints only.

This should not be confused with a biomechanical model, where segments S_i are defined via anatomical landmarks and are associated with the respective IMUs, typically via a rigid I2S transformation $T^{S_i I_i}$. The latter would need an I2S rotation calibration ($R^{S_i I_i}$) in addition to the estimates given via the proposed algorithm. Note that the I2S position could be computed from the joint position estimates in case the I2S rotation would be known, via $I^{S_i} = -R^{S_i I_i} J^{I_i}_{i,j}$. The associated biomechanical skeleton could then be directly computed via $T^{NS_i} = T^{NI_i} T^{I_i S_i}$. A lower body skeleton with respective joints and IMUs is illustrated in Figure 1 (left). Figure 1 (right) shows a corresponding biomechanical skeleton with a given I2S alignment (which is, however, not needed or estimated by the proposed algorithm).

2.2 Problem statement

Synchronized measurements from a network with N_I IMUs at time t are denoted

$$Y_t = \left(\begin{Bmatrix} y^{I_i}_{a,t} \\ y^{I_i}_{\omega,t} \\ y^{I_i}_{m,t} \end{Bmatrix}_{i \in \mathscr{I}} \right). \tag{3}$$

Here, $y^{I_i}_{a,t} \in \mathbb{R}^3$ is the accelerometer measurement, $y^{I_i}_{\omega,t} \in \mathbb{R}^3$ is the gyroscope measurement, and $y^{I_i}_{m,t} \in \mathbb{R}^3$ is the magnetometer measurement, all in the IMU frame I_i. We consider a sensor network with $|\mathscr{I}| > 1$ IMUs connecting $|\mathscr{J}| \geq 1$ joints. Moreover, we assume one IMU being mounted on each segment of the considered kinematic chain.

The aim is to recursively estimate the time dependent state of a kinematic chain of N_I IMUs together with the joint and optionally fixed position parameters. For each IMU i, the state includes its position $I_{i,t}^N$ velocity $\dot{I}_{i,t}^N$, linear acceleration $\ddot{I}_{i,t}^N$, orientation $\chi_{i,t}^{IN}$, and angular velocity $\omega_{I_i,t}^{NI_i}$, all with respect to the navigation frame N and element of $\mathbb{R}^3$. Together with the previously mentioned IMU centered joint and fixed position parameters, the state can be written as

$$X_t = \begin{pmatrix} \left\{ \begin{matrix} \omega_{I_i}^{NI_i} \\ \ddot{I}_i^N \end{matrix} \right\}_{i \in \mathscr{I}} \\ \left\{ \begin{matrix} \chi^{NI_i} \\ I_i^N \\ \dot{I}_i^N \end{matrix} \right\}_{i \in \mathscr{I}} \\ \left\{ \begin{matrix} J_{(i,j)}^{I_i} \\ J_{(i,j)}^{I_j} \end{matrix} \right\}_{(i,j) \in \mathscr{J}} \\ \left\{ J_{\text{fix}}^{I_i} \right\}_{i \in \mathscr{F}} \end{pmatrix}_t . \tag{4}$$

where the specific grouping will be explained later.

Mathematically speaking, we want to obtain a maximum a-posteriori (MAP) state estimate, given the sensor measurements from (3) and statistical models (priors) regarding the IMU motion and the kinematic structure of the skeleton (joint connections). To obtain a formal estimation problem, we utilize Bayes' theorem and the fact that a monotonic function does not change the estimates for extreme points. Thus, the estimation objective can be written as

$$\begin{aligned} \hat{X}_{1:N_T} &= \underset{X_{1:N_T}}{\operatorname{argmax}}\, p(X_{1:N_T} | Y_{1:N_T}) \\ &= \underset{X_{1:N_T}}{\operatorname{argmin}} - \log(p(Y_{1:N_T} | X_{1:N_T})) - \log(p(X_{1:N_T})) + \log(p(Y_{1:N_T})) . \end{aligned} \tag{5}$$

Note, additional parameter estimates, like sensor biases, often denoted as θ[29], could be included, but are left out for brevity of notation. Using the Markov property, the prior can be further subdivided into a prior distribution at the first time step and a temporal prior between subsequent time steps, i.e.

$$\log(p(X_{1:N_T})) = \log(p(X_1)) + \sum_{t=2}^{N_T} \log(p(X_t | X_{t-1})) . \tag{6}$$

Assuming independent measurements and applying the Markov assumption also to the likelihood, the objective becomes

$$\hat{X}_{1:N_T} = \underset{X_{1:N_T}}{\arg\min} \underbrace{-\sum_{t=1}^{N_T} \log(p(Y_t | X_t))}_{\text{likelihood}} \underbrace{-\log(p(X_1))}_{\text{initial prior}} \underbrace{-\sum_{t=2}^{N_T} \log(p(X_t | X_{t-1}))}_{\text{motion prior}} \underbrace{+\log(p(Y_{1:N_T}))}_{\text{constant}} . \tag{7}$$

The following state splitting is used in this work:

$$X_t = \begin{Bmatrix} X_t^1 \\ X_t^2 \end{Bmatrix}, \quad \text{with} \quad X_t^1 = \left(\left\{ \begin{matrix} \omega_{I_i}^{NI_i} \\ \ddot{I}_i^N \end{matrix} \right\}_{i \in \mathscr{I}} \right)_t , \; X_t^2 = \begin{pmatrix} \left\{ \begin{matrix} \chi^{NI_i} \\ I_i^N \\ \dot{I}_i^N \end{matrix} \right\}_{i \in \mathscr{I}} \\ \left\{ \begin{matrix} J_{i,j}^{I_i} \\ J_{i,j}^{I_j} \end{matrix} \right\}_{(i,j) \in \mathscr{J}} \\ \left\{ J_{\text{fix}}^{I_i} \right\}_{i \in \mathscr{F}} \\ \\ \end{pmatrix}_t . \tag{8}$$

The splitting is motivated by the following discrete state-space model

$$X_t^1 = F_t^1(X_{t-1}) + w_t, \tag{9}$$

$$X_t^2 = F_t^2(X_{t-1}), \tag{10}$$

$$Y_t = H_t(X_t) + v_t, \tag{11}$$

where $w_t \sim N(0, Q_t)$, $v_t \sim N(0, \Sigma_t)$ denote normally distributed additive process and measurement noises for the motion model F_t^1 and the measurement model H_t. Note that there is no process noise modeled for the states in X_t^2. The time-dependent states $\omega_{I_i}^{NI_i}$ and $\ddot{I}_i^N$ have directly associated measurement models (see Section 2.5), and thus a process noise is crucial in the model. The joint and fixed positions, $J_{(i,j)}^{I_i}, J_{(i,j)}^{I_j}, J_{\text{fix}}^{I_i}$, are modeled as parameters. Since they are linear with respect to the other states in the joint (49) and velocity (50) connection models, their estimation can be viewed as a linear Bayesian regression jointly estimated with the other states. The remaining states in X_t^2 are time-dependent but do not have an associated measurement model. Thus, they need to satisfy the proposed motion model strictly to avoid increasing estimation noise. The proposed subdivision is crucial to obtain a consistent and long-term stable recursive estimator (see Section 2.3). Leaving out the constants, the estimation problem with the model Equation (9) to Equation (11) can generally be written as the following constrained non-linear weighted least squares problem

$$\begin{aligned} \hat{X}_{1:N_T} = \min_{X_{1:N_T}} & \frac{1}{2}\|X_1 - \hat{X}_1\|^2_{P_1^{-1}} + \frac{1}{2}\sum_{t=2}^{N_T}\sum_{i\in\mathscr{I}}\|w_t\|^2_{Q_t^{-1}} + \frac{1}{2}\sum_{t=1}^{N_T}\sum_{i\in\mathscr{I}}\|v_t\|^2_{\Sigma_t^{-1}} \\ s.t. \quad & X_t^2 = F^2(X_{t-1}), \end{aligned} \tag{12}$$

where $X_t \sim N(X_t;\ \hat{X}_t,\ P_t)$ are state distributions for each time step and $\hat{X}_{1:N_T}$ is the collection of all these state distributions.

2.3 Derivation of optimization-based recursive Bayesian estimator

Since we are interested in online estimation with a low computational cost, a consistent recursive estimator of (12) is derived in the following. It estimates the states for the current time step and reduces to the Extended Kalman Filter (EKF) as one special case. For the derivation, we rewrite the general model Equation (9) to Equation (11) as

$$X_t^1 = F_t^1(X_{t-1}) + w_t, \tag{13}$$

$$X_t^2 = \underbrace{\lim_{\varepsilon_t \to 0} \tilde{F}_t^2(X_{t-1}, \varepsilon_t)}_{=F_t^2(X_{t-1})}, \tag{14}$$

$$Y_t = H_t(X_t) + v_t, \tag{15}$$

with $w_t \sim N(0, Q_t^1)$ and $\varepsilon_t \sim N(0, Q_t^2)$. $\tilde{F}_t^2$ is a slightly modified version of F_t^2, where small artificial noise (ε_t) is added to the variables X_{t-1}^2 that emerge in F_t^2. Let the complete state $X_t = ((X_t^1)^T, (X_t^2)^T)^T \in \mathbb{R}^{N_1+N_2}$ consist of $X^1 \in \mathbb{R}^{N_1}$ and $X^2 \in \mathbb{R}^{N_2}$. We can now linearize Equation (14) with respect to the noise at $\varepsilon_t = 0$ to obtain

$$X_t^1 = F_t^1(X_{t-1}) + w_t, \tag{16}$$

$$X_t^2 = \tilde{F}_t^2(X_{t-1}, 0) + \lim_{\varepsilon_t \to 0} \underbrace{W_t Q_t^2 W_t^T}_{=\tilde{Q}_t^2}, \tag{17}$$

$$Y_t = H_t(X_t) + v_t, \tag{18}$$

for $W_t = \dfrac{\partial \tilde{F}^2(X_{t-1}, 0)}{\partial \varepsilon_t}$ with the property $\lim_{\varepsilon_t \to 0} Q_t^2 = 0_{N_2 \times N_2}$ i.e. Q_t^2 converges to the zero matrix for decreasing ε_t. Without loss of generality, we assume $\tilde{Q}_t^2$ to be invertible for all $\varepsilon_t > 0$ (otherwise replace $\tilde{Q}_t^2$ with $\tilde{Q}_t^2 + \varepsilon_t I_{N_2 \times N_2}$). Analogously to the derivation of the (E)KF in 30, we can write Equation (16) to Equation (18) as linear equation system for time steps $t = 1 \ldots k$, with measurements up to time step $m \leq k$, and with the collective state vector $X_{1:k} = (X_1^T, \ldots, X_k^T)^T$ (corresponds to z_k in 30). The least squares solution to this equation system is the minimizing solution to the following objective function

$$J_{k|m}(X_{1:k}) = \frac{1}{2}\|X_1 - \hat{X}_1\|^2_{P_1^{-1}} + \underbrace{\frac{1}{2}\sum_{t=2}^{k}\|X_t^1 - F_t^1(X_{t-1})\|^2_{(Q_t^1)^{-1}} + \lim_{\varepsilon_t\to 0}\frac{1}{2}\sum_{t=2}^{k}\|X_t^2 - F_t^2(X_{t-1})\|^2_{(\tilde{Q}_t^2)^{-1}}}_{=\lim_{\varepsilon_t\to 0}\frac{1}{2}\sum_{t=2}^{k}\|X_t - F_t(X_{t-1})\|^2_{\tilde{Q}_t^{-1}}} + \frac{1}{2}\sum_{t=1}^{m}\|Y_t - H_t(X_t)\|^2_{\Sigma_t^{-1}}, \quad (19)$$

with

$$F_t(X_{t-1}) = \begin{pmatrix} F_t^1(X_{t-1}) \\ F_t^2(X_{t-1}) \end{pmatrix}, \qquad Q_t = \begin{pmatrix} Q_t^1 & 0 \\ 0 & Q_t^2 \end{pmatrix}. \quad (20)$$

In (19), $J_{k|m}(X_{1:k})$ describes the objective with motion model residuals up to time step k and measurement residuals up to time step m and thus state estimates up to time step k.

Note that the limes of ε must not be taken in (19), since this would lead to a singular matrix for the residual weighting of the third term. Instead, a recursive estimation form is derived in the following. This allows taking the limes in the final form and naturally leads to a consistent prediction. To obtain a recursive formulation, similarly to 30, we proceed by considering the case $k = m$, where there are measurements available for each time step.

The prediction step consists of determining the estimate $\hat{X}_{k|k-1}$ based on $\hat{X}_{k-1|k-1} \sim N(X_{k-1|k-1};\ \hat{X}_{k-1|k-1},\ P_{k-1|k-1})$, the latter including the measurements up to time step $k - 1$. To derive the prediction step, we write the objective (19) in recursive form as

$$J_{k|k-1}(X_{1:k}) = J_{k-1|k-1}(X_{1:k-1}) + \lim_{\varepsilon_k\to 0}\frac{1}{2}\|X_k - F_k(X_{k-1})\|^2_{(Q_k)^{-1}}. \quad (21)$$

Here, $J_{k-1|k-1}(X_{1:k-1})$ relates to the prior distribution $p(X_{k-1}|Y_{1:k-1}) \approx N(X_{k-1};\hat{X}_{k-1|k-1},\ P_{k-1|k-1})$ and, exploiting the Markov assumption, could be represented as

$$J_{k-1|k-1}(X_{1:k-1}) \underbrace{=}_{\text{Markov}} J_{k-1|k-1}(X_{k-1}) = \|X_{k-1|k-1} - \hat{X}_{k-1|k-1}\|^2_{P^{-1}_{k-1|k-1}}. \quad (22)$$

The gradient of (21) is

$$\nabla J_{k|k-1}(X_{1:k}) \underbrace{=}_{\text{Markov}} \nabla J_{k|k-1}(X_{k-1}, X_k) = \frac{\partial J_{k|k-1}(X_{1:k})}{\partial(X_{k-1}, X_k)} \quad (23)$$

$$= \lim_{\varepsilon_k\to 0}\begin{pmatrix} \nabla J_{k-1|k-1}(X_{k-1}) + \nabla F_k(X_{k-1})^T Q_k^{-1}(X_k - F_k(X_{k-1})) \\ Q_k^{-1}(X_k - F_k(X_{k-1})) \end{pmatrix}, \quad (24)$$

with $\nabla J_{k-1|k-1}(X_{k-1}) = \dfrac{\partial J_{k-1|k-1}(X_{k-1})}{\partial X_{k-1}}$ and $\nabla F_k(X_{k-1}) = \dfrac{\partial F_k(X_{k-1})}{\partial X_{k-1}}$.

The Hessian of (21) is

$$\nabla^2 J_{k|k-1}(X_{k-1}, X_k) = \lim_{\varepsilon_k\to 0}\begin{pmatrix} \nabla^2 J_{k-1|k-1}(X_{k-1}) + G(X_{k-1}, X_k, \varepsilon_k) & \nabla F_k(X_{k-1})^T Q_k^{-1} \\ Q_k^{-1}\nabla F_k(X_{k-1})^T & Q_k^{-1} \end{pmatrix}, \quad (25)$$

with

$$G(X_{k-1}, X_k, \varepsilon_k) = \nabla^2 F_k(X_{k-1}) Q_k^{-1}(X_k - F_k(X_{k-1})) + \nabla F_k(X_{k-1})^T Q_k^{-1}\nabla F_k(X_{k-1}).$$

Since (21) is a quadratic objective function and $\nabla^2 J_{k|k-1}(X_{k-1},X_k)$ is positive definite, a single iteration of Newton's method yields the minimum. Thus, the prediction can be written as

$$\begin{pmatrix} \hat{X}_{k-1} \\ \hat{X}_k \end{pmatrix} = \begin{pmatrix} X_{k-1} \\ X_k \end{pmatrix} - \nabla^2 J_{k|k-1}(X_{k-1}, X_k)^{-1}\nabla J_{k|k-1}(X_{k-1}, X_k), \quad (26)$$

for any initial guess for $\begin{pmatrix} X_{k-1} \\ X_k \end{pmatrix}$.

Now, observing that if the estimate of the previous time step was a minimum $\nabla J_{k-1|k-1}(\hat{X}_{k-1}) = 0$ and, since choosing $\hat{X}_k = F_k(\hat{X}_{k-1})$ sets all other terms in (24) to zero, the initial guess $\begin{pmatrix} \hat{X}_{k-1} \\ F_k(\hat{X}_{k-1}) \end{pmatrix}$ yields a minimum of (21). Considering this in (26), i.e. that the second term on the right-hand side equals zero, the estimate is this initial guess and the prediction mean reads

$$\hat{X}_{k-1|k} = \hat{X}_k = F_k(\hat{X}_{k-1}). \tag{27}$$

The prediction covariance can now be computed by evaluating the Hessian (25) at the prediction mean $\begin{pmatrix} \hat{X}_{k-1} \\ F_k(\hat{X}_{k-1}) \end{pmatrix}$, which gives

$$\begin{aligned} &\nabla^2 J_{k|k-1}(\hat{X}_{k-1}, \hat{X}_k) = \\ &\lim_{\varepsilon_k \to 0} \begin{pmatrix} \nabla^2 J_{k-1|k-1}(\hat{X}_{k-1}) + \nabla F_k(\hat{X}_{k-1})^T Q_k^{-1} \nabla F_k(\hat{X}_{k-1}) & \nabla F_k(\hat{X}_{k-1})^T Q_k^{-1} \\ Q_k^{-1} \nabla F_k(\hat{X}_{k-1})^T & Q_k^{-1} \end{pmatrix}. \end{aligned} \tag{28}$$

With the inverse of Schur's complement the inverse of (28) can be written as

$$\begin{aligned} &\nabla^2 J^{-1}_{k|k-1}(\hat{X}_{k-1}, \hat{X}_k) = \\ &\lim_{\varepsilon_k \to 0} \begin{pmatrix} \nabla^2 J^{-1}_{k-1|k-1}(\hat{X}_{k-1}) & \nabla^2 J^{-1}_{k-1|k-1}(\hat{X}_{k-1}) \nabla F_k(\hat{X}_{k-1})^T \\ \nabla F_k(\hat{X}_{k-1}) \nabla^2 J^{-1}_{k-1|k-1}(\hat{X}_{k-1}) & Q_k + \nabla F_k(\hat{X}_{k-1}) \underbrace{\nabla^2 J^{-1}_{k-1|k-1}(\hat{X}_{k-1})}_{=P_{k-1|k-1}} \nabla F_k(\hat{X}_{k-1})^T \end{pmatrix}. \end{aligned} \tag{29}$$

Since the inverse Hessian (29) is computed for both time steps, we can extract the covariance for the predicted mean (27) from the lower right corner of (29). Therefore,

$$P_{k|k-1} = \lim_{\varepsilon_k \to 0} Q_k + \nabla F_k(\hat{X}_{k-1}) P_{k-1|k-1} \nabla F_k(\hat{X}_{k-1})^T. \tag{30}$$

In this form Q_k does not need to be inverted and can be singular, thus we can now take the limes $\varepsilon_k \to 0$ to obtain

$$P_{k|k-1} = Q_k + \nabla F_k(\hat{X}_{k-1}) P_{k-1|k-1} \nabla F_k(\hat{X}_{k-1})^T \tag{31}$$

with

$$Q_k = \begin{pmatrix} Q_k^1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Thus, a well defined, consistent and recursive prediction for the model (9) and (10) is given via the mean (27) and the covariance (31). With this we have derived the prediction

$$p(X_k | Y_{1:k-1}) \approx N(X_k; \hat{X}_{k|k-1}, P_{k|k-1}). \tag{32}$$

To obtain the posterior distribution

$$p(X_k | Y_{1:k}) \approx N(X_k; \hat{X}_{k|k}, P_{k|k}),$$

we have to include the measurement update for time step k. From the distributional point of view the posterior is proportional to the product of the prediction with the likelihood

$$p(X_k | Y_{1:k}) \propto p(Y_k | X_k) p(X_k | Y_{1:k-1}).$$

From the objective function point of view we have to approximate

$$J_{k|k}(X_k) = \underbrace{\|Y_k - H_k(X_k)\|^2_{\Sigma_k^{-1}}}_{\text{likelihood}} + \underbrace{J_{k|k-1}(X_{k-1:k})}_{\text{prediction}}. \tag{33}$$

This can be rewritten as

$$J_{k|k}(X_k) = \|Y_k - H_k(X_k)\|^2_{\Sigma_k^{-1}} + \|X_k - \hat{X}_{k|k-1}\|^2_{P^{-1}_{k|k-1}}, \tag{34}$$

where the prediction yields a prior on the states for the measurement update. With the non-linear least squares problem (34), and the exact prediction (27) and (31), we can derive multiple EKF variants, like the Iterated EKF (via the Gauss-Newton method) or a Levenberg-Marquardt Iterated EKF, depending on the numerical optimization scheme used to minimize this objective. This is, in detail, described in 27 and can be analogously applied here. In this work, we use a Gauss-Newton method with line-search for the measurement models to obtain the estimate

$$\hat{X}_{k|k} = \min_{X_k} J_{k|k}(X_k). \tag{35}$$

The covariance is computed as

$$P_{k|k} = (I - KH)P_{k|k-1}, \tag{36}$$

$$H = \left.\frac{\partial H_k(s)}{\partial s}\right|_{s=\hat{X}_{k|k-1}}, \tag{37}$$

$$K = P_{k|k-1}H^T(HP_{k|k-1}H^T + \Sigma_k)^{-1}, \tag{38}$$

with $P_{k|k-1}$ from (31).

2.4 Motion models

This section specifies the equations for constructing F_t in (9) and (10). The IMU motion models assume constant acceleration and constant angular velocity from time $t-1$ to t, with sampling time Δt, hence

$$I^N_{i,t} = I^N_{i,t-1} + \Delta t \dot{I}^N_{i,t-1} + \frac{(\Delta t)^2}{2}\ddot{I}^N_{i,t-1} \tag{39}$$

$$\dot{I}^N_{i,t} = \dot{I}^N_{i,t-1} + \Delta t \ddot{I}^N_{i,t-1} \tag{40}$$

$$\ddot{I}^N_{i,t} = \ddot{I}^N_{i,t-1} + \Delta t\, w^a_t \tag{41}$$

$$\chi^{NI_i}_t = \chi\left(q^{NI_i}_{t-1} \odot \exp\left(\frac{\Delta t}{2}\omega^{NI_i}_{I_i,t-1}\right)\right) \tag{42}$$

$$\omega^{NI_i}_{I_{i,t}} = \omega^{NI_i}_{I_i,t-1} + \Delta t w^\omega_t, \tag{43}$$

where $w^a_t \sim N\left(0, Q^a_t\right)$, $w^\omega_t \sim N\left(0, Q^\omega_t\right)$ denote independent normally distributed process noises, $\odot$ denotes the quaternion multiplication and exp denotes the quaternion exponential. The transition models for the joint $(i, j) \in \mathcal{J}$ and fixed point $l \in \mathcal{S}$ parameters are

$$\{J^{Ip}_{i,j,t} = J^{Ip}_{i,j,t-1}\}_{p\in\{i,j\}} \tag{44}$$

$$J^{I_l}_{\text{fix},t} = J^{I_l}_{\text{fix},t-1}. \tag{45}$$

2.5 Measurement models

This section presents the different measurement models used for constructing H_t in (11). These comprise models for exploiting the actual sensor measurements, i.e. the IMU data, as well as, models to incorporate assumptions concerning the kinematic chain (connected joints) and its interaction with the environment (fixed position,

zero velocity detections). The latter are required for positional drift compensation. In this work, we evaluate two optional measurement models for positional drift compensation: 1) fixed position (assuming that one IMU moves around a fixed point in the navigation frame) is evaluated on the simulated three link chain (see Section 3.2), 2) zero velocity detections are evaluated on the lower body walking data from humans (see Section 3.3) and a humanoid robot (see Section 3.4). Note, for the purpose of positional drift compensation, position or velocity based information needs to be induced into the estimation at least for one IMU, since the joint connection models ((49, 50) below) ensure the propagation of such induced information through the kinematic chain. Moreover, the magnetometer measurement model is also optional, and evaluation results are presented with and without using this model.

2.5.1 IMU data. Let g^N be the gravity vector in the navigation frame, the accelerometer and gyroscope measurement models are[13]

$$y_{a,t}^{I_i} = \left(R_{i,t}^{NI}\right)^T \left(\ddot{I}_i^N - g^N\right) + v_t^a \tag{46}$$

$$y_{\omega,t}^{I_i} = \omega_{I_i,t}^{NI_i} + v_t^\omega, \tag{47}$$

where $v_t^a \sim N\left(0, \Sigma_t^a\right)$, $v_t^\omega \sim N\left(0, \Sigma_t^\omega\right)$ denote independent normally distributed measurement noises. To omit heading drift, the magnetometer measurements can be used. The following (implicit) model is constructed to only provide information in the heading direction and omit information about the local dip angle, which is a common way to reduce the influence of magnetic disturbances[31]:

$$0 = \operatorname{atan2}\left[\frac{\left(R_t^{NI_i} y_{m,t}^{I_i}\right)_y}{\left(R_t^{NI_i} y_{m,t}^{I_i}\right)_x}\right] + v_t^{\text{mag}} \tag{48}$$

where $(\cdot)_x$, $(\cdot)_y$ denote the x and y component of the respective vector and $v_t^{mag} \sim N\left(0, \Sigma_t^{\text{mag}}\right)$ denotes normally distributed measurement noises.

2.5.2 Joint connections. The skeleton structure, i.e. the fact that the segments of the tracked skeleton (kinematic chain) are connected by joints, is also incorporated via two measurement models. The information is operationalized by equating the two IMU-centered position states per joint in the navigation frame, as well as, the joint velocities:

$$\left\{0 = I_{i,t}^N + R_t^{NI_i} J_{i,j,t}^{I_i} - \left(I_{j,t}^N + R_t^{NI_j} J_{i,j,t}^{I_j}\right) + v_t^{\text{cp}}\right\}_{(i,j)\in\mathcal{J}} \tag{49}$$

$$\left\{0 = \dot{I}_{i,t}^N + R_t^{NI_i}\left(\left[\omega_{I_i,t}^{NI_i}\times\right] J_t^{I_i}\right) - \left(\dot{I}_{j,t}^N + R_t^{NI_i}\left(\left[\omega_{I_j,t}^{NI_j}\times\right] J_t^{I_j}\right)\right) + v_t^{\text{cv}}\right\}_{(i,j)\in\mathcal{J}}, \tag{50}$$

where $[a\times]$ denotes the cross-product matrix of vector a and $v_t^{\text{cp}} \sim N\left(0, \Sigma_t^{\text{cp}}\right)$, $v_t^{\text{cv}} \sim N\left(0, \Sigma_t^{\text{cv}}\right)$ denote independent normally distributed measurement noises. Note, from a theoretical point of view, the joint velocity constraint could be left out, since IMU position and velocity are completely linked through zero process noise in the time update. However, the experimental evaluation on the three-link chain (Section 3.2) supports using both models. Moreover, note that the above measurement models represent time integrations of the acceleration-level joint connection measurement model proposed in 26. Comparative results are also provided in the experimental evaluation.

2.5.3 Fixed positions. This model assumes an IMU i moving around a fixed point P_{fix}^N in the navigation frame (e.g. the upper arm IMU moving around the shoulder joint in the simulated three-link chain). The fixed point is then estimated in the respective IMU coordinate system I_i as additional IMU-centered joint position state $J_{\text{fix},t}^{I_i}$. The associated measurement model is:

$$0 = P_{\text{fix}}^N - \left(I_{i,t}^N + R_t^{NI_i} J_{\text{fix},t}^{I_i}\right) + v_t^{\text{fix}}, \tag{51}$$

where $v_t^{\text{fix}} \sim N\left(0, \Sigma_t^{\text{fix}}\right)$ denotes normally distributed measurement noise.

2.5.4 Zero velocity detections. For all IMUs $c \in \mathscr{C} \subset \mathscr{I}$, zero velocity $m_{c,t} = 1$, $m_{c,t} \in \{0, 1\}$ is detected, based on a sliding window of inertial measurements $y^{I_c}_{a,t-b:t}$ and $y^{I_c}_{\omega,t-b:t}$, using the stance hypothesis optimal (SHOE) detector[32]. Note, in case of multiple detections, only the IMU with the lowest SHOE value is kept. The measurement model for a zero velocity detection ($m_{c,t} = 1$) is

$$0 = \dot{I}^N_{c,t} + v^{\text{zvd}}_t, \tag{52}$$

where $v^{\text{cvd}}_t \sim N\left(0, \Sigma^{\text{cvd}}_t\right)$ denotes normally distributed measurement noise. Note, in a locomotion scenario, zero velocities are typically regularly detected at the foot IMUs during the stance phases. In locomotion activities with flight phases, such as running or jumping, these events become less frequent, and the estimation can show drift, if there is no other drift compensating mechanism.

2.6 State initialization

There exist different approaches for state initialization in the context of IHMC, e.g. initially assumed poses[33,34] or orientation initializations $R^{NI_i}, i \in \mathscr{I}$, while the other quantities are assumed to be zero[35].

In this work, we also assume the kinematic chain to be initially stationary, i.e. IMU states $\dot{I}^N_{i,0}, \ddot{I}^N_{i,0}, \omega^{NI_i}_{I_i,0}, \forall i$ are initialized with $N(0, \Sigma^v_{init}), N(0, \Sigma^a_{init}), N(0, \Sigma^\omega_{init})$. Moreover, the IMU positions $I^N_{i,0}$ are also initialized with $N(0, \Sigma^p_{init}), \forall i$.

The IMU orientations $\chi_0^{NI_i}, \forall i$ are initialized using the accelerometer and the magnetometer measurements via the triad method[36] as e.g. in 13. Note, this is the only point in time, where magnetometer measurements are required.

In order to support Monte Carlo simulations in the evaluation, the joint positions are initialized randomly with $J^{I_i}_{i,j,0}, J^{I_j}_{i,j,0} \sim N(0, \Sigma^J_{init})$, with $\Sigma^J_{init} = l^2 I_{3\times3}$ and l is the maximal segment length in order to obtain a reasonable distribution ($l = 0.4$ in the experiments). The fixed positions are treated analogously.

2.7 On observability of the estimated quantities

Following the notation about structural and practical observability, based on 37, we can identify the following structural non-observabilities. The Gauge freedom[20,38] in initial position $I^N_{i,0}, \forall i$ and initial velocity $\dot{I}^N_{i,0}, \forall i$ (before zero velocity updates) are correct only up to additive offsets, since no absolute references exist for these estimates. To mitigate this problem, we define prior distributions for the initial position and velocity, with the assumptions stated in Section 2.6.

Regarding practical observability, the joint position estimates need "sufficient" excitation to be observable. Ideally, all degrees of freedom of the relative rotations between the neighboring IMUs have to be excited. This can be analyzed using an approach similar to 26. Alternatively, we propose a scalar joint position uncertainty measure based on the covariance of our state estimator. This can be exploited to indicate sufficient motion excitation for a given joint (e.g. by applying a threshold), or, in other words, as convergence indicator for joint position estimation. Let $\Sigma(J^I_{i,j,t})$ be the average of the marginalised (3 × 3) covariance matrices of joint (i, j) as estimated in both IMU coordinate systems I_i, I_j. We then calculate the scaled largest eigenvalue of this matrix with

$$s^{max}_{J^I_{i,j,t}} = s_c \sqrt{\text{EV}^{max}_{\Sigma(J^I_{i,j,t})}}, \tag{53}$$

where $s_c \approx 3.37$ corresponds to the square root of the 99%-quantile of the Chi-squared distribution with three degrees of freedom. This approximates the calculation of the 99% credibility region using a scaled spectral norm. Respective evaluations for different movement scenarios can be found in Section 3.

While, in this work, magnetometers are always used to initialize the IMU orientations via the triad method (see Section 2.6), the magnetometer measurement model (48) can be used to obtain observable orientation estimates with respect to the heading direction in the navigation frame. However, if only relative rotations are relevant, and sufficient motion is present to allow for local observability at each joint, as formalized in 39, the magnetometer measurements can be neglected during tracking. According to 39, local non-observability is given, if acceleration and jerk at a joint are co-linear. This only happens in case of no motion or motion along the gravitational vector. In Section 3, we also evaluate the effects of using or omitting magnetometer measurements during tracking.

3 Evaluation and results

This section presents the evaluation of the proposed approach with respect to different estimated quantities (IMU orientations, joint positions) and aspects (convergence speed, accuracy, drift behaviour). The evaluation metrics are summarized in Section 3.1. Evaluation is performed on three different motion scenarios: simulated data of a three-link kinematic chain fixed in space (sinusoidal motion in all degrees of freedom) (Section 3.2), real lower body human gait data from 25 healthy persons measured at two points in time (Section 3.3), and gait data captured from a humanoid robot (Section 3.4). The covariance settings used in the proposed recursive Bayesian estimator are listed in Table 1 (in the order they are introduced in Section 2.4 to Section 2.6). Sensor biases were subtracted from the real IMU data as pre-processing step.

3.1 Evaluation metrics

The joint position estimates are evaluated via the segment length errors. Let sl_i be the length of segment S_i with corresponding IMU I_i and connecting joints $J_{h,i}$ and $J_{i,j}$. The estimated segment length at time t is then

$$sl_{i,t} = \| J^{I_i}_{i,j,t} - J^{I_i}_{h,i,t} \|_2, \tag{54}$$

and the error can be calculated with respect to the reference $sl_{i,ref}$ as

$$E^{sl}_{i,t} = \| sl_{i,t} - sl_{i,ref} \|_2. \tag{55}$$

The reference is either chosen as the ground truth length of the segment (simulation study, Section 3.2), the estimated segment length at the end of a reference motion (gait study, Section 3.3), or a measured reference (robot, Section 3.4). In the second case, the reliability of the joint position estimates is additionally evaluated exploiting gait data from the same person measured at two different days, with different IMU mountings. In the first case (simulation study), we also evaluate the end-point errors for each estimated joint position in each IMU coordinate system, e.g.

$$E^{J}_{i,j,t} = \| J^{I_i}_{i,j,t} - J^{I_i}_{i,j,ref} \|_2, \tag{56}$$

where reference joint position $J^{I_i}_{i,j,ref}$ is taken from the constructed model.

Table 1. Estimator covariances used in the experimental evaluation.

Covariance	Value
Q^a	$3 \cdot 10^5 \cdot I_{3\times3}$
Q^ω	$10^4 \cdot I_{3\times3}$
Σ^a	$10^{-2} \cdot I_{3\times3}$
Σ^ω	$10^{-3} \cdot I_{3\times3}$
Σ^{mag}	$10^{-2} \cdot I_{3\times3}$
Σ^{cp}	$10^{-4} \cdot I_{3\times3}$
Σ^{cv}	$10^{-3} \cdot I_{3\times3}$
Σ^{fix}	$10^{-4} \cdot I_{3\times3}$
Σ^{cvd}	$10^{-4} \cdot I_{3\times3}$
Σ^v_{init}	$1 \cdot I_{3\times3}$
Σ^a_{init}	$1 \cdot I_{3\times3}$
Σ^ω_{init}	$1 \cdot I_{3\times3}$
Σ^χ_{init}	$10^{-6} \cdot I_{3\times3}$
Σ^p_{init}	$1 \cdot I_{3\times3}$
Σ^J_{init}	$1.6 \cdot 10^{-1} \cdot I_{3\times3}$

The estimated relative IMU orientations $R_t^{I_i I_j} = R_t^{I_i N} R_t^{N I_j}, \forall (i,j)$ are evaluated on the angle errors

$$E_{i,j,t}^{rel} = |\phi(R_t^{I_i I_j} R_{ref,t}^{I_j I_i})|, \tag{57}$$

where $\phi(R) = \arccos\left(\frac{\text{trace}(R)-1}{2}\right)$[13]. Analogously, the global IMU orientations are evaluated in the simulation study. To summarize the errors (deviations from the references) over a complete sequence in one number, the root mean squared error (RMSE) is exploited as

$$E_X^{RMSE} = \frac{1}{N_T} \sqrt{\sum_{t=1}^{N_T} \sum_{l \in X} (E_{l,t})^2}, \tag{58}$$

where X is the index set of the respective measure, e.g. $l = (i,j) \in \mathcal{J}$ for joint (i,j) and E denotes an error, e.g. E^J or E^{rel} as defined above.

To evaluate the convergence behaviour of the joint position estimates, we also consider the sample standard deviation (RMSE Std.) of the RMSEs for different initial values obtained via Monte Carlo sampling.

The validity of the relative IMU orientations is furthermore evaluated via the coefficient of determination R^2 [40]. Moreover, systematic biases and drift in IMU orientation estimates are evaluated via the ordinary least product regression (OLP) coefficients[41].

3.2 Simulation study on a three-link chain

This section presents the evaluation results of the proposed approach using the fixed position measurement model for positional drift compensation (see Section 2.5.3) on simulated (noisy) IMU data. The results regard the observability of the joint positions as well as the validity of the IMU orientation estimates.

3.2.1 Experimental setup. Tests are performed on simulated IMU data from a three-link chain fixed in space, as illustrated in Figure 2. The skeleton consists of three segments S_i, $i = 0, 1, 2$ with IMUs I_i, $i = 0, 1, 2$. The I2S calibrations assumed for simulation are:

$$T^{S_0 I_0} = \begin{pmatrix} 0 & 0 & 1 & 0.1 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0.15 \\ 0 & 0 & 0 & 1 \end{pmatrix}, T^{S_1 I_1} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0.1 \\ -1 & 0 & 0 & 0.2 \\ 0 & 0 & 0 & 1 \end{pmatrix}, T^{S_2 I_2} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0.1 \\ -1 & 0 & 0 & 0.05 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{59}$$

The first segment is fixed in space at the starting point and the other two segments are attached via two three-dimensional rotary joints.

The motion and IMU data simulation are taken from 13 (Section 2.6.3, *sim-fast-artificial* sequence). In short, the motion is obtained by assuming sinusoidal angle profiles for all joint degrees of freedom. This satisfies the need for sufficient motion for the observability of the joint positions. Based on the above, the IMU data, ground truth IMU poses and joint positions are calculated.

The average measured execution time per timestep (for three IMUs) was 1.38 ms when running our single-threaded unoptimized C++ implementation on an octa-core AMD Ryzen 7 7800X3D (Base Clock: 4.2 GHz).

3.2.2 Joint position estimates. The convergence speed of the joint position estimation is evaluated with $N = 10$ Monte Carlo samples for the initial joint positions as well as for the noises applied during IMU data simulation. Comparable results were obtained with and without magnetometer model.

Figure 3 shows the convergences of the joint position estimates. In the figure, the red lines correspond to the end-point errors (56) for each estimated joint position in each IMU coordinate system. The dark grey shaded areas (labelled init uncertainty) correspond to the sample standard deviations over the errors due to different joint initializations from the Monte Carlo samples, and the light green shaded areas (labelled max model uncertainty) correspond to the maximum estimated joint position uncertainties (53) over the Monte Carlo samples. The latter is our currently proposed scalar joint position uncertainty measure to indicate the motion excitation for the respective joint. Note, during online estimation, we have no Monte Carlo samples, and the uncertainty is based on the estimation covariance.

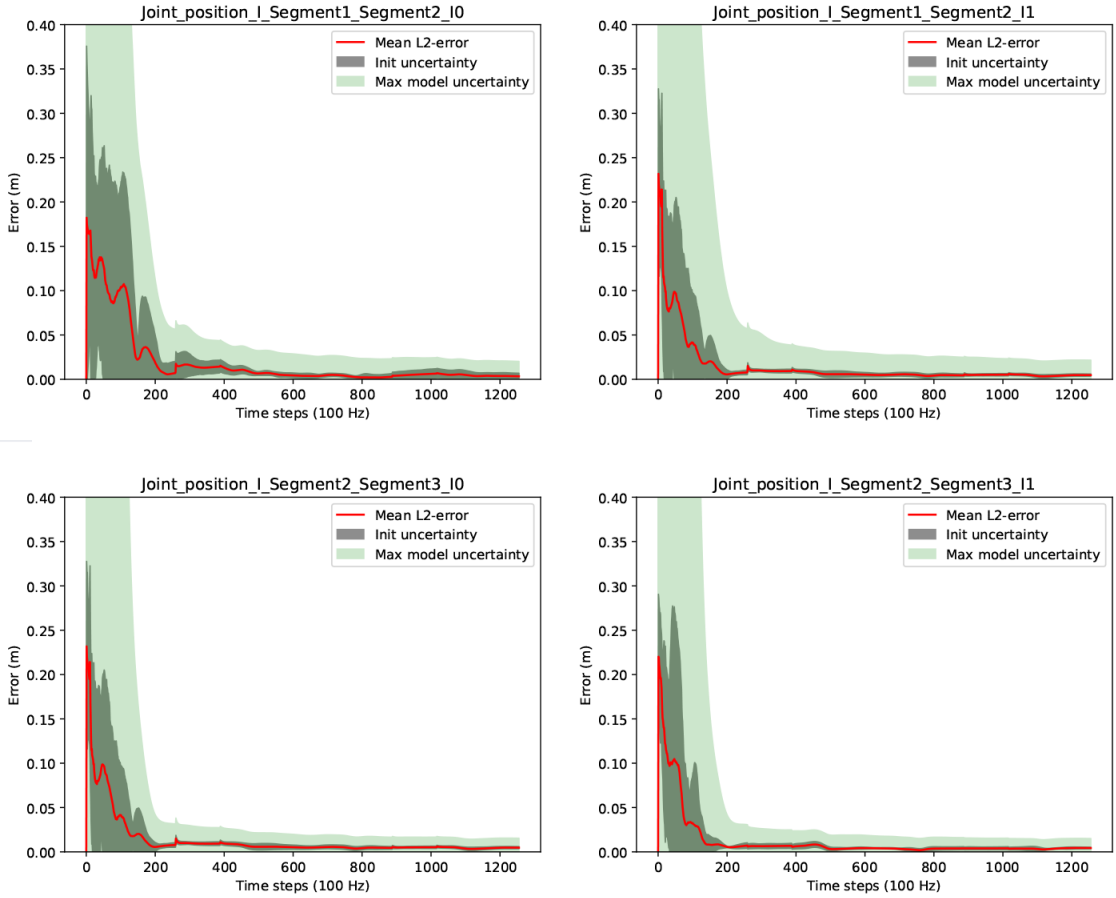

Figure 3. Three-link chain: Joint position estimation errors and uncertainties over two motion cycles, i.e. 1256 time steps at 100 Hz.

It can be observed that the joint position estimates are converged with very low init uncertainty after about 200 samples, which corresponds to about 2 seconds of data. Given the Monte Carlo sampling, this can be considered a robust result for random initializations for the current experimental setup. It is also visible, that the proposed joint position uncertainty measure (max model uncertainty) is a valid convergence indicator in this experiment (the light green shaded areas include the red lines). Comparable results were obtained for the estimation of the fixed position.

Figure 4 shows results concerning the segment length errors as calculated via (55). Note, since the segment length computation requires the joint positions, only two lengths can be calculated. For the third (distal) segment of the three link chain, where no fixation in space is assumed, the segment length calculation is not possible. In the figure, the upper row shows the results from our proposed online method operating on a complete kinematic chain and using position- and velocity-level joint connection measurement models. With this, the final absolute segment length errors (after time step 1256) are very low with 1.1 millimeters for the first and 1.5 millimeters for the second segment. As mentioned in Section 2.5.2, a previous offline and per-joint method for joint position estimation employed an acceleration-level joint connection measurement model[26]. For comparison, we tested a variant of our proposed method, where an acceleration-level joint connection measurement model was used instead of the position- and velocity-level models. The latter is based on Equation (9) in 26, whereas the angular accelerations were calculated using a difference quotient on the angular velocities in the state, and the measurement model uncertainty was set to $I_{3\times3}$. The results in the lower row of Figure 4 show no convergence and significantly higher errors and uncertainties compared to our proposed method. This is due to the missing drift correction on IMU positions and velocities over the complete kinematic chain, when purely using the acceleration-level measurement model. Note, segment lengths are calculated based on multiple IMU poses in our method. We also evaluated a variant using only the position-level connection measurement

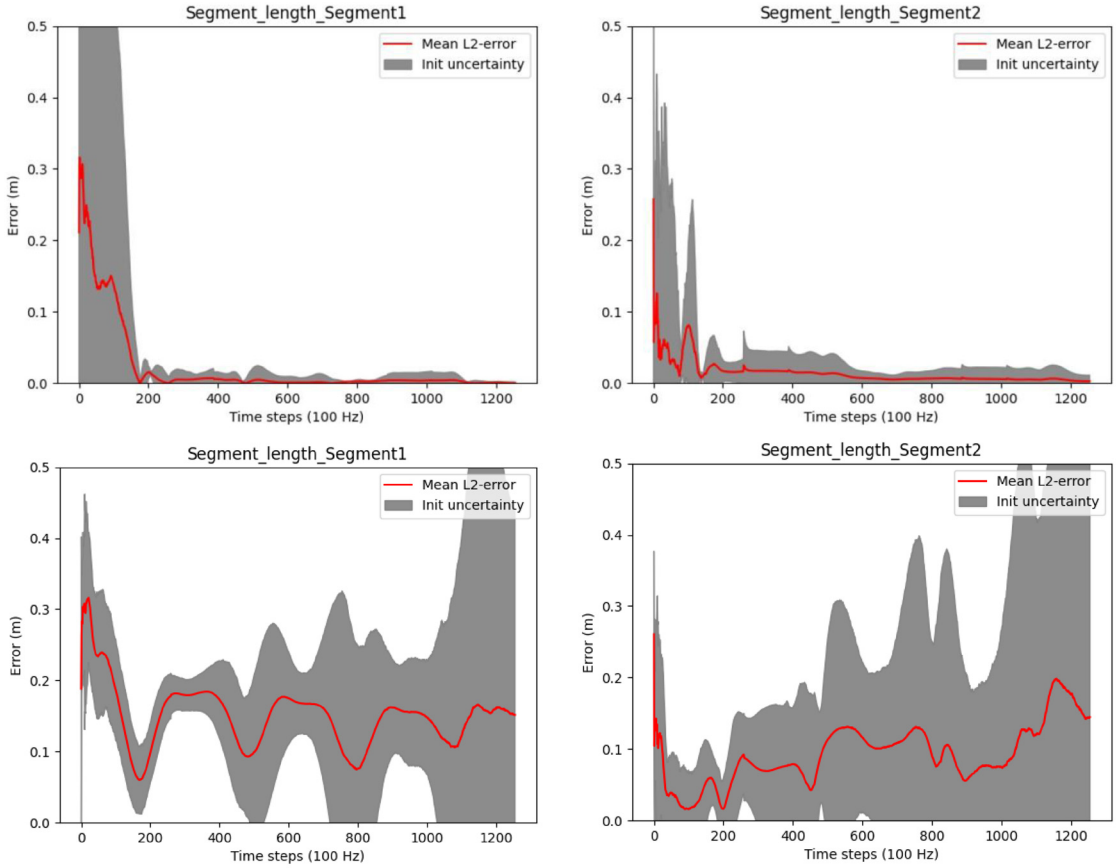

Figure 4. Three-link chain: Segment length estimation errors and uncertainties over two motion cycles, i.e. 1256 time steps at 100 Hz. The upper row shows the results when using our proposed method (joint connection measurement models on position and velocity level). The lower row shows the results when using only an acceleration-level joint connection measurement model.

model (omitting the velocity-level model), which resulted in slightly lower accuracy and slower convergence. However, using the acceleration-level model in addition to the other two, did not improve the results but increase the computational cost. The following experiments only consider our proposed method (using position- and velocity-level joint connection measurement model).

3.2.3 IMU orientation estimates. To assess the IMU orientation estimates and the influence of the joint position estimation on these with and without magnetometers, we also investigated the angle errors of the estimated relative and global IMU orientations (57) for the randomly initialized joint positions. The results for the global IMU orientations are shown in Figure 5. The results for the relative IMU orientations look similar and are therefore omitted here.

The figure shows that the IMU orientation estimates with or without magnetometers, are not significantly affected by the joint position estimation. The errors stay mostly below 1°, independently of joint position initialization and magnetometer usage. This holds also for the relative IMU orientations. Slightly higher init uncertainties can be observed without the magnetometer model, which is expected due to the lower amount of information fed to the estimator when omitting magnetometer information. Note that, even if not visible in this experimental setting, without magnetometers, the global heading direction is not observable, which might result in drift in the heading direction when processing longer data sequences with other disturbances (besides noise), see e.g. 33. This is, however, not the focus of the present work.

3.2.4 Long term estimation. To assess the long term stability of the proposed approach after convergence, we investigated the relative and global IMU orientation estimates over 300 motion cycles (i.e. $N = 300 \cdot 629$ samples

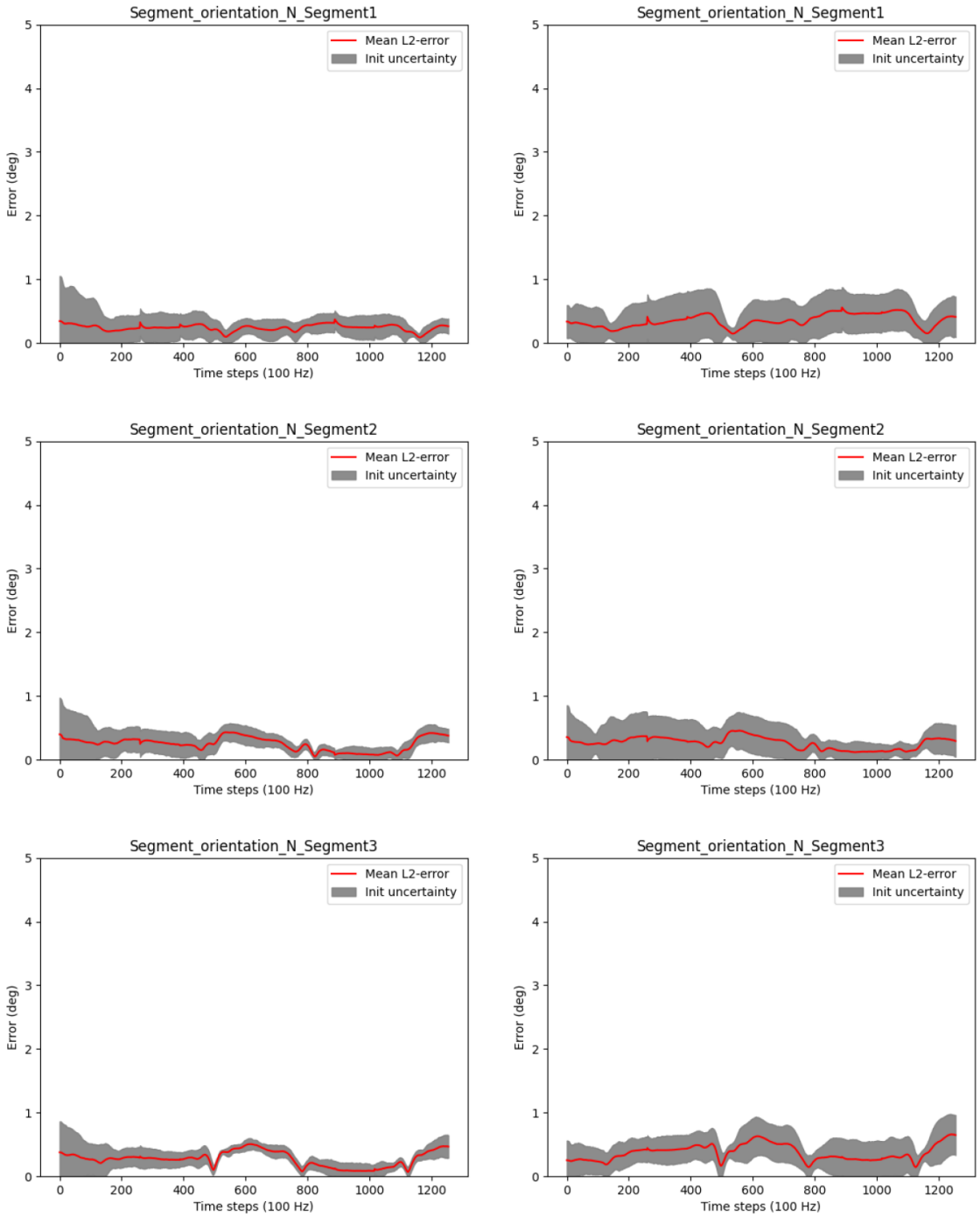

Figure 5. Three-link chain: Angle errors and init uncertainties of estimated global IMU orientations with magnetometers (left) and without magnetometers (right) for randomly initialized joint positions.

at 100 Hz, which corresponds to 31.45 minutes) with respect to systematic errors. For this, the joint positions and initial IMU poses were initialized correctly. Moreover, the relative IMU orientation estimates are without the magnetometer model, while the global IMU orientation estimates are with the magnetometer model (since the global heading direction is otherwise not observable, as pointed out in the previous section). As introduced in Section 3.1, we computed the OLP coefficients, the R^2 scores and the RMSEs between the estimated orientations and the ground truth orientations, both reduced to single angles (57), considering all samples. The

results show that the proposed approach performs very well in all measures: The absolute additive OLP coefficients are all far below 1° (0.11° on average over all joints and segments) indicating only very minor biases, whereas the OLP scaling factors are all 1.0 indicating no systematic error like drift. These results are confirmed by the R^2 scores (all equal to 1.0) and the very low RMSEs all far below 1° (0.20° on average over all joints and segments). These results demonstrate the consistency and long term stability of the proposed approach for the noisy simulated IMU data in a scenario with sufficient motion excitation.

3.3 Human lower body scenario

This section presents the evaluation results of the proposed approach using the zero velocity detections measurement model for positional drift compensation (see Section 2.5.4) on real lower body human gait data, namely the TUK-6-minute-walking dataset[42], first introduced in 33.

3.3.1 Experimental setup. The experimental setup is shown in Figure 6. The dataset contains synchronized lower body IMU and optical reference data from 25 young, healthy test persons (15 females, age 24.60 ± 2.70 years; 69.44 ± 12.17 kg and 1.75 ± 0.08 m in height) and two sessions (6.75 ± 2.26 days between test and retest). The hardware setup consisted of twelve OptiTrack Prime 13 cameras (NaturalPoint, Inc., Corvallis, OR, USA) and seven XSens MTw Awinda (Xsens Technologies BV, Enschede, The Netherlands) IMUs. The study was approved by the ethical committee of the Rheinland-Pfälzische Technische Universität (RPTU) (date of approval 5.12.2017), formerly known as Technische Universität Kaiserslautern (TUK) and meets the criteria of the declaration of Helsinki. After receiving all relevant study information, the participants signed an informed consent to the study including a permission to publish the data.

In contrast to the previously described simulated motion, there is no exact ground truth data available for the lower body dataset. In particular, the joint position reconstruction from skin markers typically involves approximations, such as regression models[43]. Therefore, we focus our evaluation of the joint position estimates on test-retest reliability of the derived segment lengths (rather than joint position validity). For the IMU orientations, we focus on validity (comparison with the optical data), since the optical marker bodies rigidly attached to the IMUs can be very accurately tracked to deduce the IMU orientations. The subsequent results are all without the magnetometer model. Comparable results were obtained with the magnetometer model.

The average measured execution time per timestep (for seven IMUs) was 15.3 ms when running our single-threaded unoptimized C++ implementation on an octa-core AMD Ryzen 7 7800X3D (Base Clock: 4.2 GHz).



Figure 6. Experimental setup of the TUK-6-minute-walking dataset with IMUs and optical markers.

3.3.2 Joint position estimates. The convergence speed of the joint position estimation is exemplified in Figure 7 for one test person via the segment lengths derived from the joint position estimates (red lines). The meaning of the shaded areas corresponds to the description in Section 3.2.2. However, note that the red lines indicate the segment lengths rather than the estimation errors and that the max model uncertainty reflects the maximum uncertainty of the two joint positions.

It can be observed that the segment lengths (and hence the joint position estimates) reach steady states after about 300 samples, which corresponds to about 5 seconds of the 60 Hz data. Moreover, the estimated lengths seem feasible for a human body. Also the drops in the proposed joint position uncertainty measure (max model

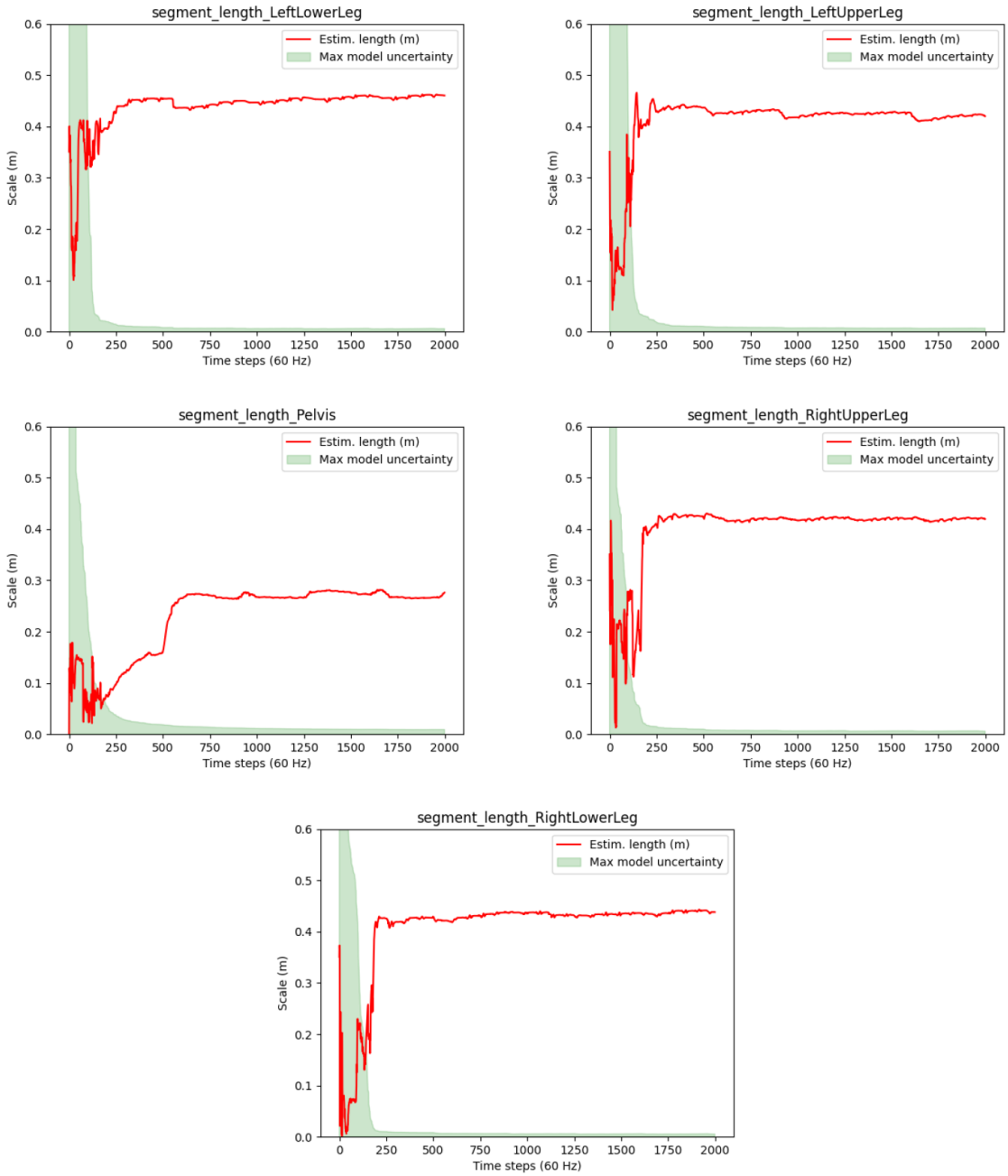

Figure 7. Human lower body: Estimated segment lengths and uncertainties for a selected test person (P1) of the TUK-6-minute-walking dataset.

uncertainty) roughly correspond with reaching the steady states (given that the max model uncertainty is associated to only one of the joint positions). However, the red lines keep showing visible swings between frames 200 and 300. A deeper analysis showed that sudden changes in the estimates correlate with the zero velocity detections, which might be inaccurate, but at the same time significantly reduce the model uncertainties.

For quantifying the reliability, we computed statistical differences in segment length estimates after 2000 samples between test and retest data. The assumption is, that the segment lengths (per person) do not change between test and retest, where the same person is measured with the same hardware performing the same walking activity on two different days. However, note that the IMU mountings, which heavily influence the estimated quantities in the local IMU coordinate systems, are different between test and retest, though they were not explicitly varied. The I2S differences between test and retest (approximated from the optical data) are quantified in Table 2. Table 3 shows the rather small RMSEs (below 2 centimeters) over all persons of the final segment length estimates between test and retest and the sample standard deviations due to Monte Carlo sampling (init uncertainties).

Altogether, it can be concluded that straight walking provides sufficient motion excitation for obtaining reasonable joint position estimates via the proposed approach, though motion excitation is significantly lower in the frontal and transversal plane compared to the sagittal plane.

3.3.3 IMU orientation estimates. To evaluate the validity of the relative IMU orientation estimates, we computed the OLP coefficients, the R^2 scores and the RMSEs between the estimated orientations and the orientations obtained from the optical data during the test session, as before, both reduced to single angles and considering all samples. The results are shown in Figure 8. The red OLP regression lines correlate well with the blue dots (indicating the individual angle samples), though the latter show higher variances compared to the results from the three-link chain. In detail, the absolute additive OLP coefficients are all far below 1° indicating only minor biases, whereas the OLP scaling factors are all close to 1.0 indicating no systematic errors such as drift. These results are confirmed by the R^2 scores (all close to 1.0) and the low RMSEs (all below 1°). The detailed numbers are given in the figure.

3.4 Humanoid robot lower body scenario

This section presents the evaluation results of the proposed approach using the zero velocity detections measurement model for positional drift compensation (see Section 2.5.4) on real lower body data captured from a humanoid robot[44].

Table 2. Human lower body: Mean and maximum I2S position and orientation (angle) differences over all segments and all test persons between test and retest in the TUK-6-minute-walking dataset.

I2S difference	Mean	Maximum
I^S [m]	0.0192	0.0716
$\phi(R^{SI})$ [°]	7.4206	88.2683

Table 3. Human lower body: Reliability of the estimated segment lengths on the TUK-6-minute-walking dataset.

Segment name	RMSE [m]	Init uncertainty
LeftLowerLeg	0.00160116	3.29466e-08
LeftUpperLeg	0.01426190	6.05016e-07
Pelvis	0.01485470	3.99128e-07
RightLowerLeg	0.00907722	1.66203e-08
RightUpperLeg	0.0187878	3.46729e-07

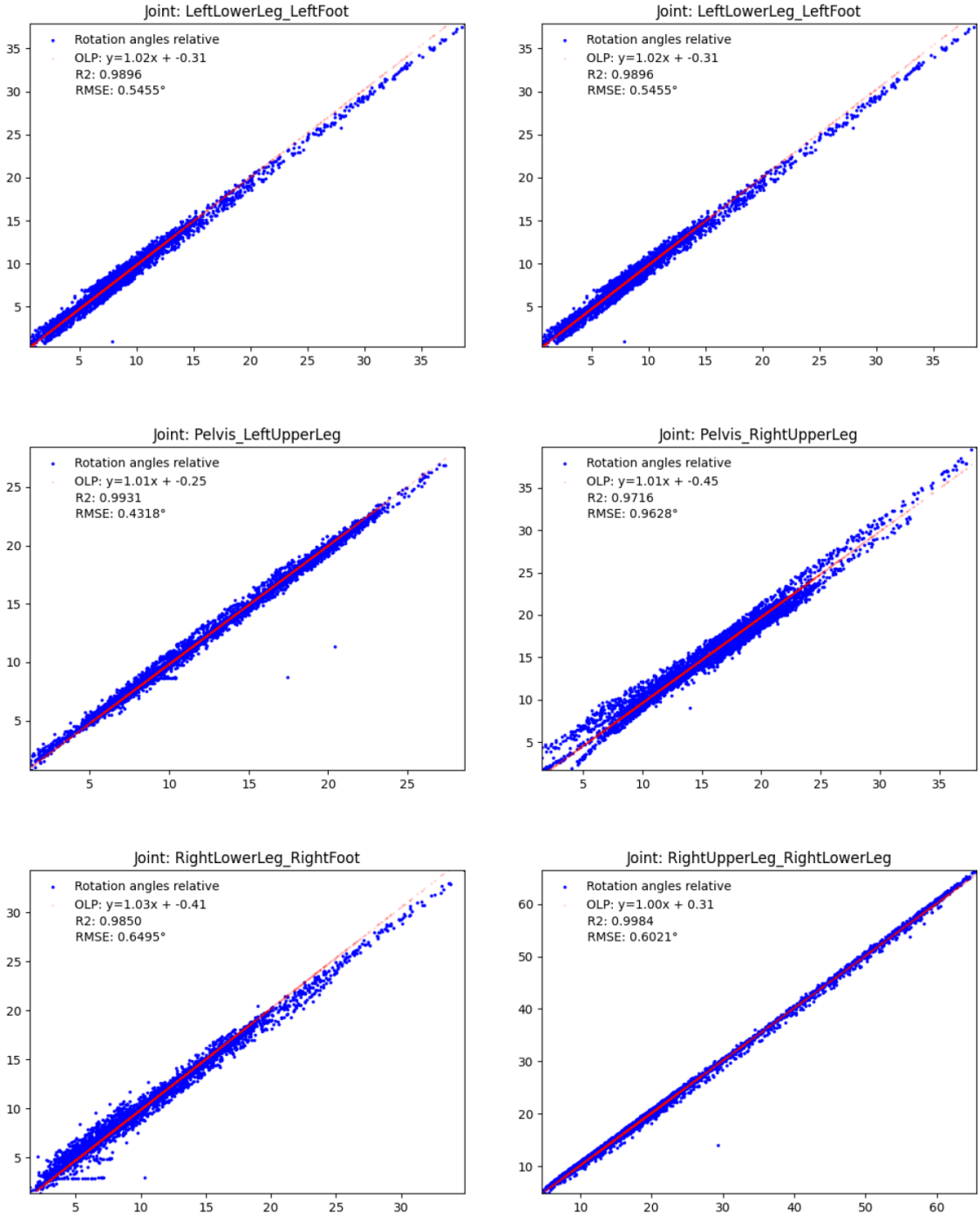

Figure 8. Human lower body: Validity of the estimated IMU orientations on the TUK-6-minute-walking dataset (without magnetometer information): estimated joint angles (y-axis) vs. joint angles obtained from the optical data (x-axis), with OLP coefficients, RMSEs and R^2 scores. In the plots, the orange dots represent the OLP coefficients and the blue dots represent the individual angle pairs (estimated over optical reference). A perfect estimation results in blue dots on the line through the orange dots.

3.4.1 Experimental setup. The same IMU hardware and a comparable IMU setup were used as in the previous experiment. In particular, seven IMUs were attached on the foot, lower leg, upper leg and pelvis segments of the humanoid robot Reem-C from Pal Robotics (Barcelona, Spain). IMU data was collected at 100 Hz. Moreover, the robot motion was synchronously captured with a marker-based optical system (Qualisys AB, Göteborg, Sweden) at 150 Hz. The setup is shown in Figure 9. The captured activity consists of the robot

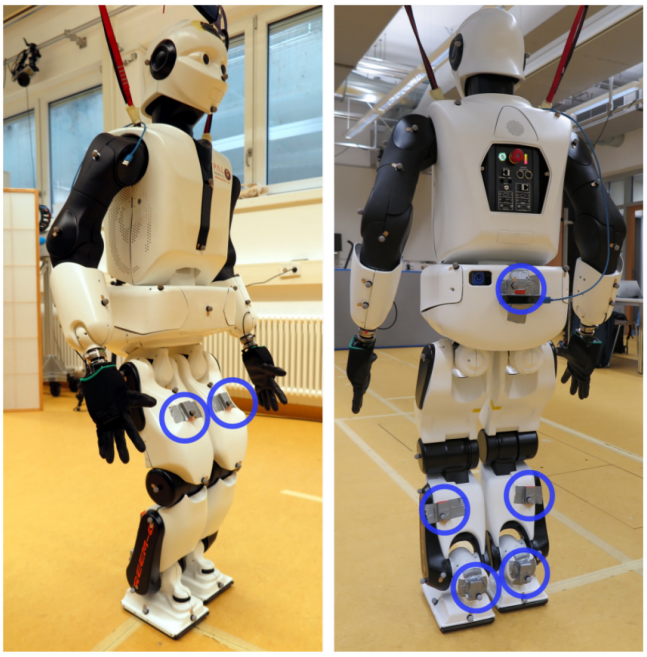

Figure 9. Humanoid robot Reem-C from Pal Robotics (Barcelona, Spain) equipped with optical markers and IMUs. The location of the IMUs are highlighted with blue circles.

walking in a circle along the edge of the recording volume of the optical motion capture system. The considered IMU data sequence has 40.000 samples corresponding to about 6.67 minutes.

The Reem-C robot is equipped with electrically driven joints. The ankles have two degrees of freedom, the knees are hinge joints (one degree of freedom) and the hips are ball-and-socket joints (three degrees of freedom). Compared to the previous experiment with humans, this setup has some advantages but also challenges with respect to the evaluation of the proposed approach. Advantages are that the robot has rigid segments which do not induce soft tissue artefacts, that it has precisely known segment lengths and that the mechanical joints allow rotations around fixed and predefined axes.

According to the manual, lower and upper legs have each a length of 0.3 meters, and the pelvis has a width (distance from hip joint center to hip joint center) of 0.15 meters.

Potential challenges are the very low acceleration and velocity of the robot motion, given that the joint position estimation requires sufficient motion excitation for convergence (see Section 2.7), and the hinge joints (knees). In case of a perfect hinge joint, the relative rotation between the two associated segments only takes place around a fixed axis, which results in the fact that the joint position estimates are not identifiable, since each position on the hinge axis represents a valid solution to the models in 2.5.2[18,20,45]. The following results are only without magnetometer model.

3.4.2 Joint position estimates. The convergences of the segment length estimates are illustrated in Figure 10. The plots show slower convergences as compared to the human walking experiment (cf. Figure 7). Table 4 shows the accuracy of the estimated segment lengths with respect to the above mentioned reference measures, with the RMSEs and standard deviations over the last 100 samples of the motion sequence. The segment length estimate at the pelvis is most accurate with a RMSE of about 1.5 centimeters. This might be due to the higher motion variability at the hip joints. The RMSEs of the other segments range between 1.5 and 3.7 centimeters. Altogether, reasonably stable estimates close to the ground truth measurements are reached for all segments despite the overall slow motion and the hinge joints of the robot.

However, note the occasionally appearing abrupt changes in the segment length estimates in Figure 10, e.g. around frame 1000 at the left upper leg. A deeper analysis showed that these result from false positive zero velocity detections (due to the slow robot motion), where erroneous information is fed into the estimator via

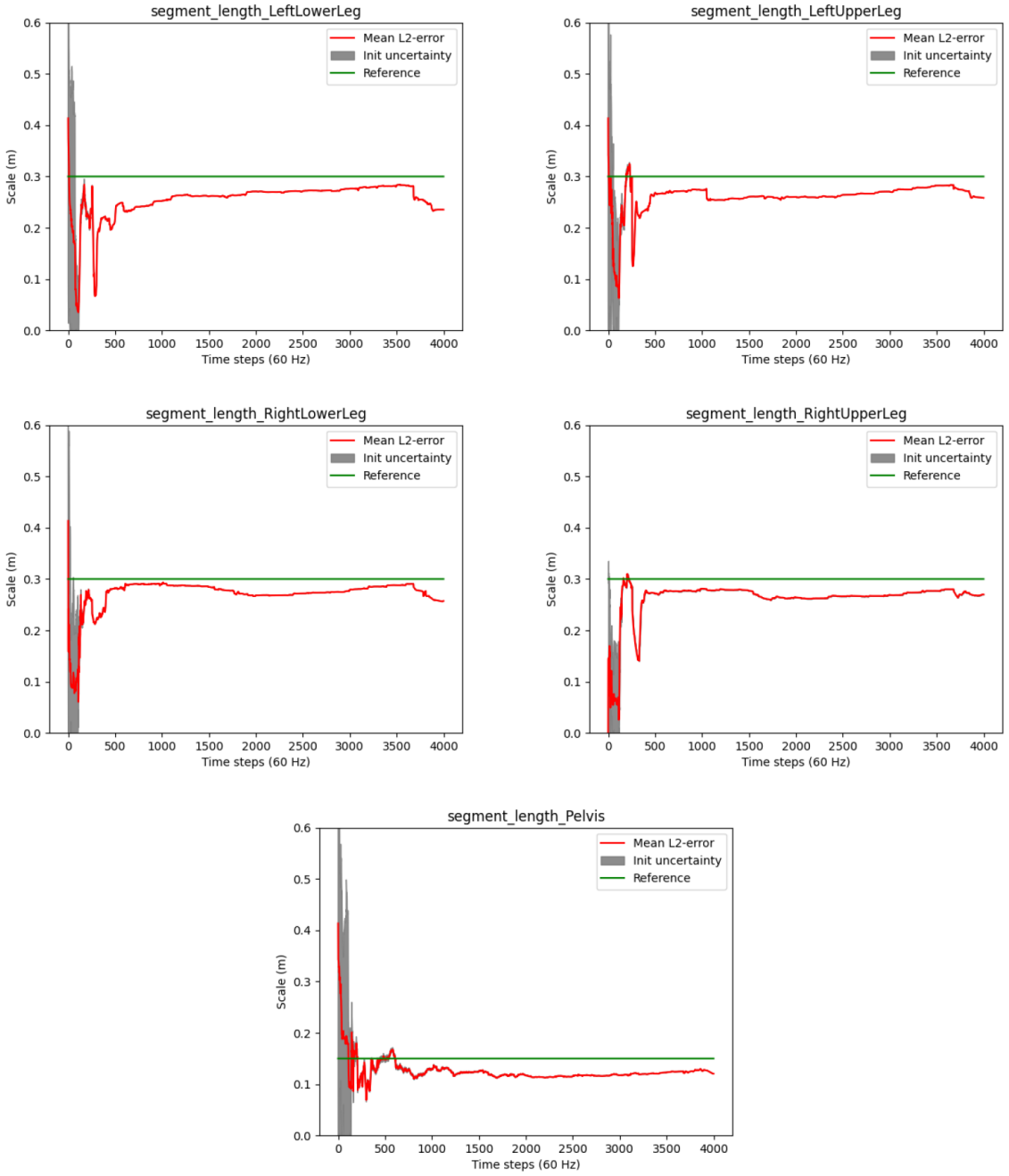

Figure 10. Humanoid robot: Estimated segment lengths (red lines) and init uncertainties in comparison to the ground truth lengths (green lines).

the model described in Section 2.5.4. Similar artefacts were observed in the previous experiment, so that is seems desirable to decouple the joint position estimation from the global positional drift reduction mechanism.

3.4.3 IMU orientation estimates. The accuracy of the relative IMU orientations with respect to the optical reference, reduced to single angles, is shown in Table 5. Altogether, the RMS angle errors are very low (below 2°).

4 Conclusions and open challenges

This section summarizes and draws conclusions from the evaluation in Section 3 and then outlines limitations, open challenges and future work.

Table 4. Humanoid robot: Accuracy of the estimated segment lengths (without magnetometer model).

Segment name	RMSE [m]	Std.
LeftLowerLeg	0.0333680	1.00993e-05
LeftUpperLeg	0.0157264	5.00482e-06
Pelvis	0.0147733	2.14394e-05
RightLowerLeg	0.0355856	1.30683e-05
RightUpperLeg	0.0368549	6.38446e-06

Table 5. Humanoid robot: Accuracy of the relative IMU orientation estimates (without magnetometer model).

Joint name	RMSE [°]	Std.
LeftFoot_LeftLowerLeg	0.855757	0.382451
LeftLowerLeg_LeftUpperLeg	1.981640	0.654190
LeftUpperLeg_Pelvis	1.362980	0.600270
Pelvis_RightUpperLeg	0.586756	0.344031
RightFoot_RightLowerLeg	1.186340	0.646329
RightLowerLeg_RightUpperLeg	0.847667	0.514045

4.1 Summary and conclusions

In this work we derived and evaluated an optimization-based recursive Bayesian estimation approach for simultaneous estimation of IMU poses (orientation and position) and joint positions for kinematic chains of different configurations. To the best of our knowledge, this is the first real-time capable approach for simultaneous IMU pose and joint (and fixed point) position estimation. The approach is recursive and uses only the data of the current timestep. Moreover, it was shown to run in real-time (below IMU framerate) in the experimental evaluation (15.3 ms average measured processing time per timestep on 60 Hz data from seven IMUs). Note that the used C++ implementation is single-threaded and not optimized; e.g. it uses numerical derivatives and no sparse matrices. We also propose an indicator for joint or fixed point position estimation convergence (53). We evaluated the proposed approach in different experimental scenarios: simulated IMU data from an arm-like three-link kinematic chain fixed in space as well as human and robotic lower body walking data.

The proposed estimation approach performs excellently on noisy simulated IMU data from a three link chain with a fixed position and with sufficient motion in all degrees of freedom. This concerns the convergence speed (below 200 samples at 100 Hz) and accuracy of the joint positions (final segment length estimation error below 1.5 millimeters) as well as the accuracy of the IMU orientations (all errors below 1°), even without magnetometer usage and during a longer data sequence (over 30 minutes). This demonstrates the overall consistency of the derived estimator.

Also on real IMU data from different locomotion scenarios with zero velocity detections, very good results were obtained for the accuracy of the magnetometer-free relative IMU orientation estimates (RMSE below 2°). Moreover, reasonable results were obtained for the joint position estimates in terms of convergence speed (about 300 samples at 60 Hz for the humans and between 500 and 1000 samples at 100 Hz for the robot), reliability (RMSE of segment length estimates below 2 centimeters between test and retest for the humans) and accuracy (RMSE of segment length estimates below 4 centimeters for the robot). This demonstrates the successful application of the proposed approach on real data and in scenarios with sub-optimal conditions, such as the comparably planar nature of walking movements in general, in particular for the robot with

perfect hinges at the knees, in combination with a very slow walking speed. Altogether, the proposed approach is a promising method for real-time IMU pose and joint position estimation without the need for I2S calibration. Hence, in this sense, the approach is calibration-free. Note that a pose based I2S calibration procedure (e.g. as in 46) could of course be used, with the additional advantage that the I2S position can be directly computed from the joint position estimates, as mentioned in Section 2.1.

Casting this in the direction of visual human pose tracking, our approach allows for IMU mounting independent estimation of joint positions in a navigation frame. Hence, it paves the way for the effortless estimation of segment lengths (in physical units), as well as the integration with joint position estimates from computer vision systems.

4.2 Limitations, open challenges and future work

As mentioned in Section 2.6, in the present work, a single magnetometer sample per IMU is always used for initializing the IMU orientations in one common reference frame. Hence, magnetometer-free refers to not using the magnetometer model in the tracking (after initialization). A magnetometer-free initialization is therefore an open challenge and part of our future work. First work in this direction can be found in 47.

The above summarized results refer to magnetometer-free relative IMU orientation estimates, since drift-free global heading estimation without magnetometers is not in the focus of this work. Moreover, we did not focus on critical cases such as completely stationary phases or phases with motion only along the gravitational vector. In these cases, relative IMU orientations are also not observable in the proposed approach without using magnetometers[39]. However, an intelligent usage of magnetometer information to address these cases is currently investigated.

Moreover, consistent joint position convergence depends on sufficient motion excitation and thus on the movement. In the case of insufficient motion excitation for joint position convergence, prior knowledge, e.g. a prior on the segment-length like in 20, could be included, or the joint position uncertainty could be used to guide the person to perform motions that target the uncertain degrees of freedom.

Despite the non-observability of the joint position of a perfect hinge via the proposed approach, as described in Section 3.4.1, reasonable joint position estimates were obtained for both the human and the robot dataset. At the same time, in particular for the robot with the mechanical hinges at the knees, improved results might be obtained when injecting this prior knowledge via specific hinge measurement models.

The proposed scalar joint position uncertainty measure proved to be a good indicator for joint position convergence in the experiments, in particular in the simulation study. On real data, a scaling calibration might be required as an additional step. However, the approach is not robust against outliers. In particular, imprecise or even false positive zero velocity detections (as observed in the human and robot experiments) prior to convergence decrease the uncertainty and at the same time significantly harm the joint position estimates. For locomotion scenarios, it is therefore desirable to decouple the joint position estimation from the global position drift reduction mechanism.

Also, since the covariances of the joint position estimates typically decrease with each incoming measurement, stationary phases (e.g. if the person stands still) can lead to overly certain estimates in a specific direction. This can slow down the convergence (even if it would not decrease the proposed convergence indicator).

Finally, the integration of the proposed approach with inside-out or outside-in visual pose reconstruction is part of our future work.

Ethics and consent

The TUK-6-minute-walking dataset[42] study was approved by the ethical committee of the Rheinland-Pfälzische Technische Universität (RPTU), formerly known as Technische Universität Kaiserslautern (TUK) (date of approval 5.12.2017), and meets the criteria of the declaration of Helsinki. After receiving all relevant study information, the participants signed a written informed consent to the study including a permission to publish the data.

Data availability

Underlying data

Zenodo: Underlying data for 'JointTracker: Real-time inertial kinematic chain tracking with joint position estimation', https://doi.org/10.5281/zenodo.10253111[42] (TUK-6-minute-walking dataset)

This project contains the following underlying data:

- Description.pdf
- BVH 6 min walking test.zip
- HDF5 6 min walking test.zip

Zenodo: Underlying data for 'JointTracker: Real-time inertial kinematic chain tracking with joint position estimation', https://doi.org/10.5281/zenodo.10253284[44] (IMU and marker-based optical motion capture from a humanoid robot)

This project contains the following underlying data:

- Description.pdf
- Robot circle run 0003.h5
- Robot circle run 0004.h5
- Robot turning 0003.h5

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0)

Software availability

Source code for data visualization and inspection available from: https://github.com/BTaetz/Visualizer-SkeletalMotionHDF5

License: MIT

References

1. Colyer SL, Evans M, Cosker DP, *et al.*: **A Review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system.** *Sports Med Open.* 2018; **4**(1): 24.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**
2. Yamamoto M, Shimatani K, Ishige Y, *et al.*: **Verification of gait analysis method fusing camera-based pose estimation and an IMU sensor in various gait conditions.** *Sci Rep.* 2022; **12**(1): 17719.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**
3. Tagliamonte NL, Peruzzi A, Accoto D, *et al.*: **Assessment of lower limbs kinematics during human-robot interaction using inertial measurement unit.** *Gait Posture.* 2014; **40**(Supplement 1): S24–S25.
Publisher Full Text
4. Bleser G, Damen D, Behera A, *et al.*: **Cognitive learning, monitoring and assistance of industrial workflows using egocentric sensor networks.** *PLoS One.* 2015; **10**(6): e0127769.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**
5. Wong C, Zhang ZQ, Lo B, *et al.*: **Wearable sensing for solid biomechanics: a review.** *IEEE Sens J.* 2015; **15**(5): 2747–2760.
Publisher Full Text
6. Menolotto M, Komaris DS, Tedesco S, *et al.*: **Motion capture technology in industrial applications: a systematic review.** *Sensors (Basel).* 2020; **20**(19): 5687.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**
7. **Xenoma.** (accessed 30. Jan. 2023).
Reference Source
8. Xiao X, Zarar S: **Machine learning for placement-insensitive inertial motion capture.** In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. Brisbane, QLD, IEEE, 2018; 6716–6721.
Publisher Full Text
9. Lorenz M, Bleser G, Akiyama T, *et al.*: **Towards artefact aware human motion capture using inertial sensors integrated into loose clothing.** In: *2022 International Conference on Robotics and Automation (ICRA)*. 2022; 1682–1688.
Publisher Full Text
10. Roetenberg D, Luinge H, Slycke P: **Xsens MVN: full 6DOF human motion tracking using miniature inertial sensors.** 2009.
Reference Source
11. Harle R: **A survey of indoor inertial positioning systems for pedestrians.** *IEEE Communications Surveys and Tutorials.* 2013; **15**(3): 1281–1293.
Publisher Full Text
12. Ligorio G, Sabatini AM: **Dealing with magnetic disturbances in human motion capture: a survey of techniques.** *Micromachines (Basel).* 2016; **7**(3): 43.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**
13. Miezal M, Taetz B, Bleser G: **On inertial body tracking in the presence of model calibration errors.** *Sensors (Basel).* 2016; **16**(7): 1132.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**
14. Taetz B, Bleser G, Miezal M: **Towards self-calibrating inertial body motion capture.** In: *19th International Conference on Information Fusion*. FUSION, Heidelberg, Germany, July 5–8, IEEE, 2016; 1751–1759.
Reference Source

15. Zimmermann T, Taetz B, Bleser G: **Imu-to-segment assignment and orientation alignment for the lower body using deep learning.** *Sensors (Basel).* 2018; **18**(1): 302.
PubMed Abstract | Publisher Full Text | Free Full Text

16. Pacher L, Chatellier C, Vauzelle R, *et al.*: **Sensor-to-segment calibration methodologies for lower-body kinematic analysis with inertial sensors: a systematic review.** *Sensors (Basel).* 2020; **20**(11): 3322.
PubMed Abstract | Publisher Full Text | Free Full Text

17. Vitali RV, Perkins NC: **Determining anatomical frames via inertial motion capture: a survey of methods.** *J Biomech.* 2020; **106**: 109832.
PubMed Abstract | Publisher Full Text

18. Seel T, Raisch J, Schauer T: **IMU-based joint angle measurement for gait analysis.** *Sensors (Basel).* 2014; **14**(4): 6891–909.
PubMed Abstract | Publisher Full Text | Free Full Text

19. Laidig D, Weygers I, Seel T: **Self-Calibrating magnetometer-free inertial motion tracking of 2-dof joints.** *Sensors (Basel).* 2022; **22**(24): 9850.
PubMed Abstract | Publisher Full Text | Free Full Text

20. McGrath T, Stirling L: **Body-Worn IMU human skeletal pose estimation using a factor graph-based optimization framework.** *Sensors (Basel).* 2020; **20**(23): 6887.
PubMed Abstract | Publisher Full Text | Free Full Text

21. McGrath T, Stirling L: **Body-Worn IMU-based human hip and knee kinematics estimation during treadmill walking.** *Sensors (Basel).* 2022; **22**(7): 2544.
PubMed Abstract | Publisher Full Text | Free Full Text

22. Martinez J, Hossain R, Romero J, *et al.*: **A simple yet effective baseline for 3d human pose estimation.** In: *Proceedings IEEE International Conference on Computer Vision (ICCV)*. Piscataway, NJ, USA, IEEE, October, 2017.
Publisher Full Text

23. Cao Z, Hidalgo G, Simon T, *et al.*: **Openpose: realtime multi-person 2d pose estimation using part affinity fields.** *IEEE Trans Pattern Anal Mach Intell.* 2021; **43**(1): 172–186.
PubMed Abstract | Publisher Full Text

24. Rajagopal A, Dembia CL, DeMers MS, *et al.*: **Full-Body musculoskeletal model for muscle-driven simulation of human gait.** *IEEE Trans Biomed Eng.* 2016; **63**(10): 2068–2079.
PubMed Abstract | Publisher Full Text | Free Full Text

25. Seel T, Schauer T, Raisch R: **Joint axis and position estimation from inertial measurement data by exploiting kinematic constraints.** In: *2012 IEEE International Conference on Control Applications.* 2012; 45–49.
Publisher Full Text

26. Olsson F, Halvorsen K: **Experimental evaluation of joint position estimation using inertial sensors.** In: *2017 20th International Conference on Information Fusion (Fusion).* Xi'an, China, IEEE, 2017; 1–8.
Publisher Full Text

27. Skoglund MA, Hendeby G, Axehill D: **Extended Kalman Filter modifications based on an optimization view point**. In: *18th International Conference of Information Fusion.* 2015.
Reference Source

28. Markley FL, Crassidis JL: **Fundamentals of spacecraft attitude determination and control.** Springer, 2014.
Reference Source

29. Kok M: **Probabilistic modeling for positioning applications using inertial sensors.** licentiate thesis, 2014.
Reference Source

30. Humpherys J, Redd P, West JM: **A fresh look at the kalman filter.** *SIAM Rev.* 2012; **54**(4): 801–823.
Publisher Full Text

31. de Vries WHK, Veeger HEJ, Baten CTM, *et al.*: **Magnetic distortion in motion labs, implications for validating inertial magnetic sensors.** *Gait Posture.* 2009; **29**(4): 535–541.
PubMed Abstract | Publisher Full Text

32. Skog I, Händel P, Nilsson JO, *et al.*: **Zero-velocity detection—an algorithm evaluation.** *IEEE Trans Biomed Eng.* 2010; **57**(11): 2657–2666.
PubMed Abstract | Publisher Full Text

33. Teufl W, Miezal M, Taetz B, *et al.*: **Validity, test-retest reliability and long-term stability of magnetometer free inertial sensor based 3D joint kinematics.** *Sensors (Basel).* 2018; **18**(7): 1980.
PubMed Abstract | Publisher Full Text | Free Full Text

34. Teufl W, Lorenz M, Miezal M, *et al.*: **Towards inertial sensor based mobile gait analysis: event-detection and spatio-temporal parameters.** *Sensors (Basel).* 2019; **19**(1): 38.
PubMed Abstract | Publisher Full Text | Free Full Text

35. Kok M, Hol JD, Schön TB: **An optimization-based approach to human body motion capture using inertial sensors.** *IFAC Proceedings.* 2014; **47**(3): 79–85.
Publisher Full Text

36. Black HD: **A passive system for determining the attitude of a satellite.** *AIAA J.* 1964; **2**(7): 1350–1351.
Publisher Full Text

37. Raue A, Becker V, Klingmüller U, *et al.*: **Identifiability and observability analysis for experimental design in nonlinear dynamical models.** *Chaos.* 2010; **20**(4): 045105.
PubMed Abstract | Publisher Full Text

38. Triggs B, McLauchlan PF, Hartley RI, *et al.*: **Bundle adjustment - a modern synthesis.** In: *Proceedings of the International Workshop on Vision Algorithms: Theory and Practice, ICCV '99.* London, UK, UK, Springer-Verlag, 2000; 298–372.
Reference Source

39. Kok M, Eckhoff K, Weygers I, *et al.*: **Observability of the relative motion from inertial data in kinematic chains**. *arXiv: 2102.02675 [cs, eess]*. arXiv: 2102.02675, February, 2021.
Publisher Full Text

40. Casella G, Berger RL: **Statistical inference**. Duxbury advanced series. CENGAGE Learning, Andover Melbourne Mexico City Stamford, CT Toronto Hong Kong New Delhi Seoul Singapore Tokyo, second edition, 2002.
Reference Source

41. Ludbrook J: **Linear regression analysis for comparing two measurers or methods of measurement: but which regression?** *Clin Exp Pharmacol Physiol.* 2010; **37**(7): 692–699.
PubMed Abstract | Publisher Full Text

42. Teufl W, Miezal M, Taetz B, *et al.*: **TUK-6-minute-walking dataset.** *Zenodo.* [Dataset], 2023.
http://www.doi.org/10.5281/zenodo.1025311

43. Kainz H, Carty CP, Modenese L, *et al.*: **Estimation of the hip joint centre in human motion analysis: a systematic review.** *Clin Biomech (Bristol, Avon).* 2015; **30**(4): 319–329.
PubMed Abstract | Publisher Full Text

44. Lorenz M, Aller F, Bleser-Taetz G, *et al.*: **IMU and marker-based optical motion capture from a humanoid robot.** *Zenodo.* [Dataset], 2023.
http://www.doi.org/10.5281/zenodo.10253283

45. Olsson F, Kok M, Seel T, *et al.*: **Robust plug-and-play joint axis estimation using inertial sensors.** *Sensors (Basel).* 2020; **20**(12): 3534.
PubMed Abstract | Publisher Full Text | Free Full Text

46. Bouvier B, Duprey S, Claudon L, *et al.*: **Upper limb kinematics using inertial and magnetic sensors: comparison of sensor-to-segment calibrations.** *Sensors (Basel).* 2015; **15**(8): 18813–18833.
PubMed Abstract | Publisher Full Text | Free Full Text

47. Lorenz M, Taetz B, Bleser G: **An approach to magnetometer-free on-body inertial sensors network alignment.** *IFAC-PapersOnLine.* 2020; **53**(2): 15982–15989.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔ ✔

Version 2

Reviewer Report 27 March 2025

https://doi.org/10.21956/openreseurope.20759.r50175

© **2025 Kok M.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Manon Kok**
Delft University, Delft, The Netherlands

Thank you for the detailed rebuttal and congrats on a very nice paper!

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Sensor fusion

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 04 February 2025

https://doi.org/10.21956/openreseurope.20759.r50173

© **2025 Bachhuber S.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Simon Bachhuber**
FAU Erlangen-Nuernberg, Nürnberg, Germany

Thank you for the point-by-point response and the revision of your article.

While I believe the document would benefit from improvements in figure quality and equation typesetting, the contributions and results remain comprehensible and merit indexing. However, not open-sourcing the code seems like a missed opportunity, and the suggestion to self-implement the method is not satisfactory.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Applied Machine Learning in the field of inertial motion capture

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 28 June 2024

https://doi.org/10.21956/openreseurope.18306.r41351

© 2024 Contreras Rodríguez L. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Lauro Armando Contreras Rodríguez
Universidad Autonoma de Sinaloa, Culiacán, Sinaloa, Mexico

Overall, the article adequately presents the development of the evaluation of an algorithm for real-time monitoring of the position of joints using inertial sensors.

Taking into consideration the problem posed in the introduction section, the methodological process describes step-by-step the algorithm used to capture the information through the optimization of the derivation of a recursive Bayesian estimator for each IMU sensor to subsequently adapt it to a kinematical model.

The results shown provide a clear explanation of what was obtained for the assessment scenarios, as well as the conclusions show the virtues and drawbacks of the proposal.

However, there are some points that must be addressed:

- Section 3.1 (Evaluation metrics) provides information that describes how the proposal will be assessed, this information belongs to the methodology section.
- Figure 12, Table 4, and Table 5 correspond to the section 3.4, not the Results section.
- Section 3.3.1 (Experimental setup) requires further information related to the test persons such as: age, height, and weight. The word “young” used to describe the population of this study could lead to interpretation of the use of children, adolescents, and/or young adults. Height and weight help to estimate other movement conditions that are affected by these parameters (the angular velocity is different according to the body segment’s size and soft tissues could lead to misalignments of the IMU sensors that modify the interpretation of the body movement).
- A direct comparison between the measurements of the humanoid robot and the test persons could provide information related to the sensor's agreement for its use in both systems. Bland-Altman analysis could be used to quantify this agreement.

Is the rationale for developing the new method (or application) clearly explained?
Yes

Is the description of the method technically sound?
Yes

Are sufficient details provided to allow replication of the method development and its use by others?
Yes

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Biomechanical movement; IMU sensor; Sport science.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 19 Dec 2024

Bertram Taetz

We thank the reviewer for reading the article in depth and for providing detailed and helpful comments! Below, we address the review point by point, with our answers being underlined and bold. Note, the figure and reference numbers in the answers refer to version 1 (first submission). Section 3.1 (Evaluation metrics) provides information that describes how the proposal will be assessed, this information belongs to the methodology section. **We agree that this is one option. However, in our article, the developed algorithm and its mathematical basis are considered the method and are therefore described in the methods section. The evaluation metrics are not a part of the method (algorithm) and are therefore described in the evaluation section, together with the experimental setups and results. To make this clearer, we changed the section title from "Results" to "Evaluation and results". We believe that this is a comprehensible distribution of information over the sections and would like to keep it that way.** Figure 12, Table 4, and Table 5 correspond to the section 3.4, not the Results section. **We agree that the mentioned figure and tables belong to Section 3.4, which is a subsection of Section 3 ("Results", now "Evaluation and Results"). Note that the figure and tables are also referenced correctly in Section 3.4.**
Note, based on a comment from another reviewer, we removed Figure 5 and Figure 7, so that the mentioned parts should appear closer to the references in the text in the

revised article. Section 3.3.1 (Experimental setup) requires further information related to the test persons such as: age, height, and weight. The word “young” used to describe the population of this study could lead to interpretation of the use of children, adolescents, and/or young adults. Height and weight help to estimate other movement conditions that are affected by these parameters (the angular velocity is different according to the body segment’s size and soft tissues could lead to misalignments of the IMU sensors that modify the interpretation of the body movement). **We added the missing information to the article: The dataset contains synchronized lower body IMU and optical reference data from 25 young, healthy test persons (15 females, age 24.60 ± 2.70 years; 69.44 ± 12.17 kg and 1.75 ± 0.08 m in height) and two sessions (6.75 ± 2.26 days between test and retest).** A direct comparison between the measurements of the humanoid robot and the test persons could provide information related to the sensor's agreement for its use in both systems. Bland-Altman analysis could be used to quantify this agreement. **A direct comparison is not possible, due to different IMU mountings and movements between humans and robot (note, we are estimating IMU and joint trajectories). Instead, we evaluated the tracking results with the respective reference data using the same metrics, and we reported these evaluation results, e.g. in Tables 3,4 for the humans and in Tables 5,6 for the robot. The results are then summarized and discussed in Section 4.1.**

Competing Interests: No competing interests were disclosed.

Reviewer Report 18 June 2024

https://doi.org/10.21956/openreseurope.18306.r41352

© 2024 Bachhuber S. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Simon Bachhuber
FAU Erlangen-Nuernberg, Nürnberg, Germany

This paper proposes a method for calibration-free motion capture of kinematic chains using body-attached IMUs.
Overall, the paper is well written and the motivation clear. However, particularly concerns regarding novelty and contributions should be clarified and alternative, baseline methods should be added to the experimental validation. Specifically, addressing the following issues will help to increase the overall quality of the manuscript:

Content
A) The novelty of the work is not distinctly outlined. The authors acknowledge the existence of various I2S calibration approaches but do not explicitly compare their method to these predecessors to highlight the specific advancements made. Furthermore, while there are calibration-free IHMC methods available, the authors have not explained why these were not

considered suitable for the task at hand. Consequently, it remains uncertain whether this work addresses an unresolved issue or if it offers a solution that outperforms existing methods, which would necessitate a comparative analysis with alternative approaches.
As an example, the authors should add sentences that compare the proposed approach, e.g., "Previous approaches to joint position estimation typically work offline, often also independently per joint, and require high-frequency inertial data; e.g., 32,37,38, though drift-free, typically require a higher sampling rate for obtaining accurate results as compared to approaches working on integrated states." -> [...] In contrast to these methods, our approach does not ...
Also directly referring to the cited sentence, it seems that the authors are motivating the contribution by: 1) online-capable, 2) not independently per joint, 3) not requiring high-frequency inertial data. But then they do not properly address these points, regarding
- 1) The ability of the proposed method to be applicable online (and in real-time?) is not further discussed in the manuscript.
- 2) The advantage of the proposed method compared to previous work that uses a independently per joint approach is not shown because the authors then later do not compare the proposed method to alternative, baseline methods
- 3) This work seems to be using 60Hz+ IMUs which, whilst not being very high-frequency, certainly are also not low-frequency IMUs (like 25Hz).

B) From the problem statement, it is unclear whether or not the IMU-to-segment orientation must be known or if sensors can be attached arbitrarily.

C) The quality and potential of figures is not yet fully utilized. Specifically, the large amount of figures make the results difficult to comprehend. The figures seem to be presented in an exhaustive manner rather than as an highlighting tool. For example, Figure 7 seems highly space inefficient. The compressed information value of the figure is too low to justify its inclusion. A more focused presentation with well chosen figures to highlight exemplary trials and rely on tables for point estimates may be advantageous, or instead of exemplary trials, the authors could plot the average error and average uncertainty over time.

D) It is unclear why the joint position estimates for human trials can not be validated. The authors could test the joint position estimates by placing an additional IMU to a segment with a fixed and known offset. This way the joint position estimate must be different by exactly this known offset value for the two IMUs.

E) There is a lack of alternative, baseline methods. The authors should compare themselves to prior work. For example, for comparing the joint estimation performance, the authors could use [Seel, et al., 2012 (Ref 1)], for comparing orientation estimation performance a simple magnetometer-reliant, joint-independent baseline can be straightforwardly computed for all scenarios. Moreover, by comparing yourself to the independent baselines (in a sense that they are applied for each joint independently) you can then then discuss the difference and advantage to the proposed kinematic-model-based approach.

F) The code for JointTracker should be made openly available.

Formatting, Typesetting
A) The formatting of figures is suboptimal.
- The figure titles use a rather unattractive underscore notation even though the authors have

defined mathematical symbols for these quantities.
- The x-axis should be given in units of second and labeled time for a more comprehensible presentation. The sampling rate can be included in the caption.
- The figures should be embedded into the documents as vector graphics.
- The font sizes, tick sizes, tick labels, titles, labels are all too small and inconsistent with the manuscript font size and font.

Minor
A) The authors motivate using *human* motion capture but this choice seems confusing. The proposed method is applicable to any kinematic chain, you even shows this applicability to nonhuman systems during experimental validation. Thus, why the restriction to human motion capture in the motivation?

B) The acronym IMU seems incorrectly introduced.

C) Some sentences are overly lengthy and hard to understand, e.g., "Previous approaches to joint position estimation typically work offline, often also independently per joint, and require high-frequency inertial data; e.g., 32,37,38, though drift-free, typically require a higher sampling rate for obtaining accurate results as compared to approaches working on integrated states." Tools like Grammarly can detect such sentences and rewrite them accordingly.

D) Some hyperrefs are not working, e.g., "In Section 3, we also evaluate ...".

E) The choice of reporting length scales in meters in this context seems suboptimal. A more natural choice might be centimeters, consider for example Figure 9.

References
1. Seel,T Schauer,T: Joint axis and position estimation from inertial measurement data by exploiting kinematic constraints. 2012. 45-49 Publisher Full Text

Is the rationale for developing the new method (or application) clearly explained?
Yes

Is the description of the method technically sound?
Yes

Are sufficient details provided to allow replication of the method development and its use by others?
Partly

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Applied Machine Learning in the field of inertial motion capture

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 19 Dec 2024

Bertram Taetz

We thank the reviewer for reading the article in depth and for providing detailed and helpful comments! Below, we address the review point by point, with our answers being underlined and bold. Note, the figure and reference numbers in the answers refer to version 1 (first submission). We hope that with our revisions and explanations, the reviewer can approve the article.

Content

A) The novelty of the work is not distinctly outlined. The authors acknowledge the existence of various I2S calibration approaches but do not explicitly compare their method to these predecessors to highlight the specific advancements made. Furthermore, while there are calibration-free IHMC methods available, the authors have not explained why these were not considered suitable for the task at hand. Consequently, it remains uncertain whether this work addresses an unresolved issue or if it offers a solution that outperforms existing methods, which would necessitate a comparative analysis with alternative approaches.
As an example, the authors should add sentences that compare the proposed approach, e.g., "Previous approaches to joint position estimation typically work offline, often also independently per joint, and require high-frequency inertial data; e.g., 32,37,38, though drift-free, typically require a higher sampling rate for obtaining accurate results as compared to approaches working on integrated states." -> [...] In contrast to these methods, our approach does not ...
Also directly referring to the cited sentence, it seems that the authors are motivating the contribution by: 1) online-capable, 2) not independently per joint, 3) not requiring high-frequency inertial data. But then they do not properly address these points, regarding
- 1) The ability of the proposed method to be applicable online (and in real-time?) is not further discussed in the manuscript.
- 2) The advantage of the proposed method compared to previous work that uses a independently per joint approach is not shown because the authors then later do not compare the proposed method to alternative, baseline methods
- 3) This work seems to be using 60Hz+ IMUs which, whilst not being very high-frequency, certainly are also not low-frequency IMUs (like 25Hz).

We agree that our contribution was not formulated clearly enough, and that the description of previous work on joint position estimation was too condensed. Our work addresses an unresolved issue (rather than offering a solution that outperforms existing methods), since it is, to the best of our knowledge, the first real-time capable method for simultaneous joint (or fixed) position and IMU kinematics estimation for

general kinematic chains (as mentioned in Section 4.1). We modified the formulation of the contributions at the end of Section 1 accordingly. Moreover, in the following, we answer to the numbered list above:

1. **Real-time capable here means that the estimation is recursive and uses only the data of the current time step, rather than being a batch approach (offline) or using a moving horizon with a window size > 1. The mathematical derivation in Section 2.3 addresses this from a theoretical perspective. This is now also clearly stated in Section 4.1. Moreover, we added exemplary processing times for the three-IMU and seven-IMU setup on the hardware running the experiments.**
2. **As mentioned above, it is not our aim to outperform independently per joint methods (concerning joint position estimation), or independently per IMU orientation estimation methods, but to enable online simultaneous estimation of the states of a moving skeleton (or kinematic chain), plug and play, and with only minimal magnetometer usage, which was a previously unresolved issue. By using our approach, we can obtain consistent relative positional (and rotational) information between all the joints, so that the overall pose of the kinematic chain can be visualized in 3D (by connecting the joint centers, similar to how it is done based on markerless camera-based tracking). The latter is now also explained at the end of Section 1.**
3. **We agree that "require a higher sampling frequency" was not sufficiently analyzed in the article, and we therefore removed this formulation. The more important aspect is that integrated states (velocities and positions), as used in our proposed algorithm (Section 2.5.2), provide drift correction on IMU position and velocity level, which is required to obtain an overall consistent kinematic tracking, and that they provide smoother signals compared to accelerations (as used for joint position estimation in the mentioned previous work).**

We modified the paragraph about previous work on joint position estimation in Section 1 accordingly. We also added a comparison to the acceleration-level joint connection model (in our algorithmic framework) to the experimental evaluation.

B) From the problem statement, it is unclear whether or not the IMU-to-segment orientation must be known or if sensors can be attached arbitrarily.
The IMU-to-segment orientation is not required. This is now more clearly stated at the end of Section 2.1. The sensors can be attached arbitrarily.

C) The quality and potential of figures is not yet fully utilized. Specifically, the large amount of figures make the results difficult to comprehend. The figures seem to be presented in an exhaustive manner rather than as an highlighting tool. For example, Figure 7 seems highly space inefficient. The compressed information value of the figure is too low to justify its inclusion. A more focused presentation with well chosen figures to highlight exemplary trials and rely on tables for point estimates may be advantageous, or instead of exemplary trials, the authors could plot the average error and average uncertainty over time.
We agree that there is a rather large amount of figures in the article, even though we had reduced them already before the first submission. However, we believe that most of the selected figures are needed to demonstrate the convergence behaviors, uncertainties and temporal progressions of estimates. For Figure 7, we removed this and included the averaged results in the main text. We also removed Figure 5 (Three-

link chain: Angle errors and init uncertainties of estimated relative IMU orientations), since it is very similar to Figure 6 (same with global IMU orientations).

D) It is unclear why the joint position estimates for human trials can not be validated. The authors could test the joint position estimates by placing an additional IMU to a segment with a fixed and known offset. This way the joint position estimate must be different by exactly this known offset value for the two IMUs.
We agree that the method described by the reviewer could be used for validation. However, in experimental setup 2, we worked with an existing dataset containing test and retest measurements from 25 test persons, and we believe that it is an interesting evaluation to analyze reliability based on this, in addition to the other two validation experiments based on simulated data and the robot. With this, we cover both, validity and reliability. Also note that the article doesn't state that joint position estimates for human trials cannot be validated in general, just that we do not have exact ground truth data for the available dataset and focus our analysis on reliability. Please excuse if we miss something.

E) There is a lack of alternative, baseline methods. The authors should compare themselves to prior work. For example, for comparing the joint estimation performance, the authors could use [Seel, et al., 2012 (Ref 1)], for comparing orientation estimation performance a simple magnetometer-reliant, joint-independent baseline can be straightforwardly computed for all scenarios. Moreover, by comparing yourself to the independent baselines (in a sense that they are applied for each joint independently) you can then then discuss the difference and advantage to the proposed kinematic-model-based approach.
As explained above (A), it is not our aim to outperform independently per joint or offline methods, but to enable a new functionality (online, calibration-free, kinematic chain tracking (simultaneous IMU kinematics and joint positions)), and to demonstrate that the proposed algorithm works in principle (using simulated data in experiment 1) and under different realistic and partly suboptimal conditions (using data from humans and a robot in experiments 2 and 3). However, to address this comment, as also mentioned above, we added a comparison of the proposed approach (using joint connection measurement models on positional and velocity level) with a variant using an acceleration level joint connection measurement model as proposed in [38], for the simulated data (three-link chain). This demonstrates that our algorithmic framework operating on a complete kinematic chain requires working on integrated states (position and velocity), since significantly better results are obtained.

F) The code for JointTracker should be made openly available.
The implementation of JointTracker is integrated (in a distributed way) into a bigger C++ software library. Unfortunately, there are no resources available to isolate the code and make it openly available. However, the article includes a detailed mathematical description of the algorithm, and the datasets have been made available. To nevertheless address this comment, we could offer an executable of JointTracker on individual request of other researchers.

Formatting, Typesetting

A) The formatting of figures is suboptimal.
- The figure titles use a rather unattractive underscore notation even though the authors have defined mathematical symbols for these quantities.
- The x-axis should be given in units of second and labeled time for a more comprehensible presentation. The sampling rate can be included in the caption.
- The figures should be embedded into the documents as vector graphics.
- The font sizes, tick sizes, tick labels, titles, labels are all too small and inconsistent with the manuscript font size and font.
We agree that the formatting of the figures could be optimized. However, we believe that the figures and figure texts (titles, ticks, labels) are in their current form reasonably well readable, comprehensible and consistent with respect to each other (none of the other reviewers commented on this). We therefore hope for the reviewer's understanding of our decision to keep the figures as they are and invest our efforts into addressing the other comments.

Minor
A) The authors motivate using *human* motion capture but this choice seems confusing. The proposed method is applicable to any kinematic chain, you even shows this applicability to nonhuman systems during experimental validation. Thus, why the restriction to human motion capture in the motivation?
The emphasis on human motion capture as motivation comes mainly from the fact that most (previous) motion capture research is in the field of human motion capture. However, it is a valid comment, and we reduced the emphasis on "human" in the abstract as well as in the introduction (beginning of Section 1). We also added the following sentence to Section 1 "The model is applicable to any kinematic chain and not restricted to or specialized to a human skeleton." (which is also an advantage of the proposed method over [34,35])

B) The acronym IMU seems incorrectly introduced.
We introduced it now as "inertial measurement unit" in the second paragraph of Section 1.

C) Some sentences are overly lengthy and hard to understand, e.g., "Previous approaches to joint position estimation typically work offline, often also independently per joint, and require high-frequency inertial data; e.g., 32,37,38, though drift-free, typically require a higher sampling rate for obtaining accurate results as compared to approaches working on integrated states." Tools like Grammarly can detect such sentences and rewrite them accordingly.
We modified this sentence.

D) Some hyperrefs are not working, e.g., "In Section 3, we also evaluate ...".
This specific hyperlink seems to work in the PDF and DOC we downloaded. However, we found other incorrect hyperlinks, which we corrected (Section 2.2., references to equations (49), (50)).

E) The choice of reporting length scales in meters in this context seems suboptimal. A more natural choice might be centimeters, consider for example Figure 9.

We agree that centimeters could have been a better choice and will consider this for future submissions. However, since all positional results are consistently reported in meters (figures and tables), we would like to keep it in this article.

Competing Interests: No competing interests were disclosed.

Reviewer Report 09 May 2024

https://doi.org/10.21956/openreseurope.18306.r39589

© 2024 Kok M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Manon Kok
Delft University, Delft, The Netherlands

The paper presents a method for recursive inertial human motion estimation in which the joint position is estimated simultaneously with the pose, thereby resulting in a calibration-free method. The paper presents an interesting approach and recursive, calibration-free motion estimation is an important topic. According to me the contribution is definitely worth publishing but I think the presentation needs improvement before doing so. I specifically think the contribution needs to be presented more clearly throughout the whole paper. More specifically:
1) The contribution is not clear from the abstract and introduction. More specifically:
a) Both the abstract and the introduction first seem to suggest the paper is about a camera-based approach, while it is actually about an inertial sensor-based approach. To aid the clarity to a reader, I would make this clear right from the start of the paper.
b) The last sentence of the abstract suggests that the presented approach can be used in visual-inertial human motion capture. There are no results included related to this though, so I would suggest to only mention this as a direction of future work.
c) In section 1, the sentences “This work contributes to … e.g. 20” are not very clear. For instance, it is not clear what “contributes to” exactly means (since it doesn’t specify that you present an algorithm), it is not clear what type of kinematic model you mean and what you refer to with visual 3D joint tracking (optitrack since that’s what you evaluate against?).
d) You estimate joint positions. It is not uncommon that this wording is used for joint angles but you actually mean the position (in meters), right? This may be worth specifying clearly.

2) Also in Section 2 the contribution is not completely clear. More specifically:
a) In the problem statement you introduce your estimation problem but while your contribution is a recursive estimator your problem statement presents a batch approach.
b) Section 2.3 until Equation (32) presents the time update of your filter. This is a very long derivation which as far as I can see ends up with a standard Kalman filter update? If this is the case, I would suggest to just present your approach much more concisely just as an iterated EKF. I think this will help the reader and hence increase the impact of the work and the theoretical presentation in terms of an optimization problem and the limits can be included in an appendix.

c) You mention below Equation (11) that the states X_t^2 do not have an associated measurement model. This is not clear at that time in the text since it's not yet known what measurement models you use.
d) In Section 2.5 it's not clear that you test your algorithm with and without some measurement models (e.g. the magnetometer) but that some of them you include in all tests.
e) The fixed positions that you describe in Section 2.5.3 are not clear, especially also not since I didn't see them mentioned anywhere in Section 3.
e) Only at the end of section 2.1 you mention "Joint Tracker" but this has not been introduced anywhere. In fact, throughout the paper you refer to your method at Joint Tracker, an iterated EKF and an optimization-based method. I would suggest to stick to one naming.
f) You mention I2S calibration multiple times, e.g. at the end of section 2.1 and in Equation (59). It is not clear how this relates to the estimates from your approach as you don't estimate the relative orientation of the sensor with respect to the segments?
g) The title of section 2.6 is observability and identifiability but throughout the text you only mention observability. In Section 3.2 you use the wording identifiability though.

3) The relation between your main contribution and the results is also not completely clear:
a) It is not clear what you want to show exactly with the experiments. For instance, you test a lot how global / relative orientation estimates depend on including the magnetometer or not. But in the rest of the paper it is not clear that this is an important part of your study.
b) To me the legend entries "init uncertainty" and "max model uncertainty" are not clear. With init I expect that this is related to the initialization which I don't think is the case. And why don't you refer to the max model uncertainty as your uncertainty measure?
c) Table 2 shows very large rotation errors but I don't think this is discussed in the text?
d) 10 Monte Carlo simulations seems very little.
e) How does your method handle biases in the inertial sensor signals?
f) Why did you correctly initialize the positions and poses in Section 3.2.4?

4) Some comments about the conclusions:
a) You conclude that your method can be run real-time, but this is not evaluated in Section 3 (no computational times are included).
b) You also draw a conclusion about convergence speed, but this is highly dependent on the movement, right?
c) You mention that you obtain reasonable estimates of unobservable variables. Why is this the case and why does constraining unobservable variables lead to improved estimates?

5) Also a more general comment: You make an interesting design choice in Section 2 and I think it's worth commenting on the reason for that. You choose to not include process noise on the position and velocity states but then you use a "loose" joint constraint for both the position and the velocity in Section 2.5.2. Why do you not take the limit of those noises to zero as well? And why do you need a constraint on both if both are completely linked through the zero process noise in the time update?

6) Some additional comments:
a) The formatting of the paper is not intuitive to me, though of course it could be part of the template: Whenever there are numbers in the paper, these are citations. I would expect some brackets around these. Furthermore, e.g. on page 4 (2.1) refers to a section while I would expect this to be a reference to an equation.

b) Some parts of equations / matrices, e.g. in Equations (2) and (3) are highlighted. It is not clear why.
c) Shouldn't it be argmax instead of max in Equations (5) and (7)?
d) Shouldn't the epsilon three lines below (18) be a different symbol than the epsilon of which you just took the limit to zero?
e) There's a dot missing on the I in Equation (8) in the definition of X^2.
f) The notation of the exp and the quaternion multiplication in (42) have not been introduced.
g) Was E_{l,t} in (58) defined?
h) In Figure 7 your legend suggests that there are yellow dots in the figure that are barely visible. The red dots, however, are not mentioned in the legend.

Is the rationale for developing the new method (or application) clearly explained?

Yes

Is the description of the method technically sound?

Yes

Are sufficient details provided to allow replication of the method development and its use by others?

Yes

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Sensor fusion

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 19 Dec 2024

Bertram Taetz

We thank the reviewer for reading the article in depth and for providing detailed and helpful comments! Below, we address the review point by point, with our answers being underlined. Note, the figure and reference numbers in the answers refer to version 1 (first submission). We hope that with our revisions and explanations, the reviewer can approve the article.

1) The contribution is not clear from the abstract and introduction. More specifically:
a) Both the abstract and the introduction first seem to suggest the paper is about a camera-based approach, while it is actually about an inertial sensor-based approach. To aid the clarity to a reader, I would make this clear right from the start of the paper.
Response: True, this was driven by our current research to fuse the JointTracker results with markerless visual approaches, which also typically detect joint positions in the images and lift them to 3D. However, this is work in progress and shouldn't be over-emphasized in this article. Hence, we removed the state-of-the-art of visual body tracking in the introduction and reduced it in the abstract.

b) The last sentence of the abstract suggests that the presented approach can be used in visual-inertial human motion capture. There are no results included related to this though, so I would suggest to only mention this as a direction of future work.
Response: We removed this part from the abstract.

c) In section 1, the sentences "This work contributes to ... e.g. 20" are not very clear. For instance, it is not clear what "contributes to" exactly means (since it doesn't specify that you present an algorithm), it is not clear what type of kinematic model you mean and what you refer to with visual 3D joint tracking (optitrack since that's what you evaluate against?).
Response: We modified this paragraph to be clearer. We specified that we contribute an algorithm and its derivation. The kinematic model is shown in Figure 1 (left). It is represented by global IMU poses with local (IMU-centered) joint positions. We further explained the markerless visual 3D joint tracking approaches we are referring to.

d) You estimate joint positions. It is not uncommon that this wording is used for joint angles but you actually mean the position (in meters), right? This may be worth specifying clearly.
Response: Right, we specified this clearly in Section 1.

2) Also in Section 2 the contribution is not completely clear. More specifically:
a) In the problem statement you introduce your estimation problem but while your contribution is a recursive estimator your problem statement presents a batch approach.
Response: The batch approach is first presented, since the recursive estimator is derived from this. This is now more clearly explained at the beginning of Section 2. All mentioned previous work on joint position estimation uses a batch approach.

b) Section 2.3 until Equation (32) presents the time update of your filter. This is a very long derivation which as far as I can see ends up with a standard Kalman filter update? If this is the case, I would suggest to just present your approach much more concisely just as an iterated EKF. I think this will help the reader and hence increase the impact of the work and the theoretical presentation in terms of an optimization problem and the limits can be included in an appendix.
Response: We understand this suggestion. In fact, we had included the mathematical derivation in the appendix to start with (before submission), both for the batch and the recursive estimator. Internal feedback from other researchers in the field suggested that we should include the derivation into the main body, since it represents a theoretical contribution in terms of an extension of [42]. We would therefore like to keep the derivation in the main text, and we added an explanation to the beginning of Section 2.

c) You mention below Equation (11) that the states X_t^2 do not have an associated measurement model. This is not clear at that time in the text since it's not yet known what measurement models you use.
Response: We added a forward reference to the measurement model section.

d) In Section 2.5 it's not clear that you test your algorithm with and without some measurement models (e.g. the magnetometer) but that some of them you include in all tests.
Response: We added a sentence at the beginning of Section 2.5, explaining that the magnetometer measurement model is optional and its effects are evaluated in the article.

e) The fixed positions that you describe in Section 2.5.3 are not clear, especially also not since I didn't see them mentioned anywhere in Section 3.
Response: The fixed positions measurement model is used for the simulation study on the three-link chain. This is explained at the beginning of Section 3.2 "This section presents the evaluation results of the proposed approach using the fixed position measurement model for positional drift compensation (see Section 2.5.3) on simulated (noisy) IMU data". For the other two experiments, the zero velocity detections and measurement model are used instead, since these movements involve locomotion. This is also explained at the beginning of Section 2.5 "In this work, we evaluate two optional measurement models for positional drift compensation: 1) fixed position (assuming that one IMU moves around a fixed point in the navigation frame) is evaluated on the simulated three link chain (see Section 3.2), 2) zero velocity detections are evaluated on the lower body walking data from humans (see Section 3.3) and a humanoid robot (see Section 3.4). Note, for the purpose of positional drift compensation, position or velocity based information needs to be induced into the estimation at least for one IMU, since the joint connection models ((49, 50) below) ensure the propagation of such induced information through the kinematic chain". Note, for the fixed position in the world-frame (P^N_fix), an arbitrary 3D coordinate can be assumed (e.g. the origin).

e) Only at the end of section 2.1 you mention "Joint Tracker" but this has not been introduced anywhere. In fact, throughout the paper you refer to your method at Joint Tracker, an iterated EKF and an optimization-based method. I would suggest to stick to one naming.
Response: We replaced "JointTracker" with "proposed algorithm".

f) You mention I2S calibration multiple times, e.g. at the end of section 2.1 and in Equation (59). It is not clear how this relates to the estimates from your approach as you don't estimate the relative orientation of the sensor with respect to the segments?
Response: Right, for the proposed algorithm, the I2S calibrations are irrelevant (neither estimated nor assumed given). We extended the explanations at the end of Section 2.1 to clarify this. In (59), the I2S calibrations are those used for simulation to generate the test IMU data, but they are not used in the algorithm. This was also added to the text.

g) The title of section 2.6 is observability and identifiability but throughout the text you only mention observability. In Section 3.2 you use the wording identifiability though.

Response: We used the term identifiability in the context of the joint and fixed positions, since these can be considered as parameters of the system. However, since these quantities are estimated as part of the state together with the IMU kinematics in the final algorithm, observability is the correct term to use, and we replaced all mentions of identifiability with observability.

3) The relation between your main contribution and the results is also not completely clear:
a) It is not clear what you want to show exactly with the experiments. For instance, you test a lot how global / relative orientation estimates depend on including the magnetometer or not. But in the rest of the paper it is not clear that this is an important part of your study.
Response: As mentioned above, we revised the formulation of the contributions. In summary, these are: mathematical derivation of real-time capable state estimator, calibration-free simultaneous online IMU kinematics and joint (or fixed) position estimation for kinematic chains (leads to segment length estimation for all segments connecting two joints), convergence indicator for the motion-dependent joint position parameters, minimal magnetometer usage (only for initialization). The above-mentioned tests with and without magnetometer measurement model activation address the last bullet point.

b) To me the legend entries “init uncertainty” and “max model uncertainty” are not clear. With init I expect that this is related to the initialization which I don’t think is the case. And why don’t you refer to the max model uncertainty as your uncertainty measure?
Response: The term “init uncertainty” was chosen, since it represents the uncertainty due to different joint initializations from the Monte Carlo sampling. We extended the explanation in Section 3.2.2. Moreover, in Section 3.2.2, we explain, that the max model uncertainty is used as uncertainty measure: “The latter is our currently proposed scalar joint position uncertainty measure to indicate the motion excitation for the respective joint.” We added the sentence: “Note, during online estimation, we have no Monte Carlo samples, and the uncertainty is based on the estimation covariance.”

c) Table 2 shows very large rotation errors but I don’t think this is discussed in the text?
Response: The numbers in Table 2 are not our evaluation results, but they quantify the IMU mounting differences between the test and retest data in the TUK-6-minute-walking dataset, as explained in Section 3.3.2. We further clarified this in the text. This shows that the IMU data was different per person between test and retest (among other things, due to different IMU mountings), while the differences in estimated segment lengths were quite small, which speaks for the reliability of the proposed algorithm.

d) 10 Monte Carlo simulations seems very little.
Response: In a related offline approach [34], which is mentioned in the introduction, evaluation was performed with only one initialization (no Monte Carlo sampling) using assumptions: “Static vectors from IMUs to neighboring joint centers and knee axes were initialized from an assumed nominal alignment.”. Compared to this, we use a zero initialization of the joint centers (without assumptions), with additional 10 offset Monte Carlo samples, as explained in Section 2.6. Together with the variety of test sequences (1 simulated+25 test persons+1 robot), we believe that this amount of sampling can be considered a reasonable trade-off between computational efforts and sampling variance.

e) How does your method handle biases in the inertial sensor signals?
Response: This information was indeed missing. We added the following sentence to the beginning of Section 3: “Sensor biases were subtracted from the real IMU data as a pre-processing step.” As mentioned in Section 2.2, the proposed algorithm can be extended for biases, but we didn’t focus on this in this article: “Note, additional parameter estimates, like sensor biases, often denoted as θ40, could be included, but are left out for brevity of notation”.

f) Why did you correctly initialize the positions and poses in Section 3.2.4?
Response: We did this to specifically test for long-term stability after convergence. This explanation was added to the article.

4) Some comments about the conclusions:
a) You conclude that your method can be run real-time, but this is not evaluated in Section 3 (no computational times are included).
Response: The conclusion states “real-time capable” and “promising method for real-time”, since, by design, the proposed algorithm uses only the data of the current timestep. This explanation was added to Section 4.1. Moreover, we added exemplary processing times for the three-IMU and seven-IMU setup on the hardware running the experiments.

b) You also draw a conclusion about convergence speed, but this is highly dependent on the movement, right?
Response: This is correct. In Section 4.1, we do not draw general conclusions concerning the convergence speed, but in relation to the different experimental setups. To make the dependence on the movement even clearer, we extended the explanation in Section 4.2: “Moreover, consistent joint position convergence depends on sufficient motion excitation and thus on the movement. In the case of insufficient motion excitation for joint position convergence, prior knowledge, e.g. a prior on the segment-length like in 34, could be included, or the joint position uncertainty could be used to guide the person to perform motions that target the uncertain degrees of freedom.”

c) You mention that you obtain reasonable estimates of unobservable variables. Why is this the case and why does constraining unobservable variables lead to improved estimates?
Response: We changed “constraining” to “adding prior knowledge”, i.e.: “At the same time, in particular for the robot with the mechanical hinges at the knees, improved results might be obtained when injecting this prior knowledge via specific hinge measurement models.” We have an initial implementation of such a specific hinge measurement model, which we; however, didn’t include in this article, since 1) the article would become too long, 2) the measurement model didn’t bring significant improvements for the considered datasets (only for a perfectly simulated hinge joint).

5) Also a more general comment: You make an interesting design choice in Section 2 and I think it’s worth commenting on the reason for that. You choose to not include process noise on the position and velocity states but then you use a “loose” joint constraint for both the position and the velocity in Section 2.5.2. Why do you not take the limit of those noises to zero as well? And why do you need a constraint on both if both are completely linked

through the zero process noise in the time update?
Response: The measurement noises for the joint connection measurement models are required to be able to cast the estimation problem as least squares estimation (and estimate the joint positions as part of the state). Otherwise, the connection model must be formulated with additional Lagrange multipliers to obtain a constrained optimization problem. This, however, increases computational complexity and therefore processing time. It is true that the velocity measurement model could be left out from a theoretical point of view. However, during empirical tests on the three-link chain, we observed a slower convergence and slightly lower accuracy when leaving it out. We added this explanation to Sections 2.5.2 and 3.2.2.

6) Some additional comments:
a) The formatting of the paper is not intuitive to me, though of course it could be part of the template: Whenever there are numbers in the paper, these are citations. I would expect some brackets around these.
Response: Brackets were included in our initial submission, but they were removed during the production process of the article.
Furthermore, e.g. on page 4 (2.1) refers to a section while I would expect this to be a reference to an equation.
Response: We added “Section” in front of these numbers.

b) Some parts of equations / matrices, e.g. in Equations (2) and (3) are highlighted. It is not clear why.
Response: The highlighting came from the Latex template used for initial submission, and we were not able to remove it. We asked, whether this could be removed in the revision.

c) Shouldn’t it be argmax instead of max in Equations (5) and (7)?
Response: Yes, these are mistakes! We corrected them.

d) Shouldn’t the epsilon three lines below (18) be a different symbol than the epsilon of which you just took the limit to zero?
Response: No, this is correct, since ϵ does not reach zero.

e) There’s a dot missing on the I in Equation (8) in the definition of X^2.
Response: Thanks, we corrected this.

f) The notation of the exp and the quaternion multiplication in (42) have not been introduced.
Response: We added an introduction of this notation in Section 2.4.

g) Was E_{l,t} in (58) defined?
Response: The l is explained below (58) and E can denote an error, such as the ones defined in (56), (57). We added an explicit explanation for this in Section 3.1.

h) In Figure 7 your legend suggests that there are yellow dots in the figure that are barely visible. The red dots, however, are not mentioned in the legend.
Response: We adapted the caption in Figure 8 (Figure 7 was removed based on a comment

from another reviewer). The light orange dots appear red over the blue dots.

Competing Interests: No competing interests were disclosed.