

Estimation and observability analysis of human motion on Lie groups

Vladimir Joukov, Josip Ćesić, Kevin Westermann, Ivan Marković, Ivan Petrović and Dana Kulic

Abstract—This paper proposes a framework for human pose estimation from wearable sensors that relies on Lie group representation to model the geometry of human movement. Human body joints are modeled by matrix Lie groups, using special orthogonal groups SO(2) and SO(3) for joint pose and special Euclidean group SE(3) for base link pose representation. To estimate the human joint pose, velocity and acceleration, we develop the equations for employing the Extended Kalman Filter on Lie Groups (LG-EKF), to explicitly account for the non-Euclidean geometry of the state space. We present the observability analysis of an arbitrarily long kinematic chain of SO(3) elements based on a differential geometric approach, representing a generalization of kinematic chains of a human body. The observability is investigated for the system using marker position measurements. The proposed algorithm is compared to two competing approaches, the EKF and unscented KF (UKF) based on Euler angle parametrization, in both simulations and extensive real-world experiments. The results show that the proposed approach achieves significant improvements over the Euler angle based filters. It provides more accurate pose estimates, is not sensitive to gimbal lock, and more consistently estimates covariances.

Index Terms—human body kinematics, motion estimation on Lie groups, marker measurements, IMUs, observability analysis

I. INTRODUCTION

HUMAN bodies can perform a variety of manipulation and locomotion tasks in different contexts. Different aspects of human body capabilities have been a focus of research in biomechanics and medicine, human-machine interaction, and have served as inspiration for humanoid robot design [1]. Many applications, including rehabilitation, athlete performance monitoring, imitation learning or human-machine interaction are enabled by accurate measurement of human pose.

Several different measurement technologies have been used for human pose estimation, including (i) wearable sensors, e.g., inertial measurement units (IMUs), and (ii) motion capture using cameras placed in the environment. While the former are cheaper and more portable, they are usually less accurate. With optical motion capture systems, the position of body worn markers is measured by stationary cameras; these measurements are processed to compute the pose [2]. Motion capture is typically considered the gold standard for accurate

This work has been supported from the Unity Through Knowledge Fund under the project Cooperative Cloud based Simultaneous Localization and Mapping in Dynamic Environments (cloudSLAM) and by the FLAG-ERA JTC 2016 and the Ministry of Science and Education of the Republic of Croatia under the project Rethinking Robotics for the Robot Companion of the future (RoboCom++).

motion measurement. However, motion capture requires line of sight between the cameras and the markers, and therefore cannot be utilized in large or outdoor spaces [3].

When relying on marker position measurements, a kinematic model of the human body is defined, assuming rigid attachment between the markers and the skeleton. Since a closed form solution for the inverse kinematics (IK) problem for a full body skeletal model is not available, the joint angles can be estimated using the Jacobian pseudoinverse [4]. As the Jacobian can become non-invertible in singular configurations, damping can be introduced in the least squares formulation to ensure full rank [5]. Nevertheless, methods based on the Jacobian inverse do not take uncertainty of marker position measurements into account, are susceptible to outliers, and cannot predict future poses. To account for these drawbacks, stochastic estimation techniques can be utilized by defining the skeleton pose as the system state and the marker positions as measurements. These techniques can leverage information from joint positions, velocities, and acceleration in the filter system model, which aids in keeping the correct pose estimate during short marker occlusions. Various stochastic estimation approaches for IK were developed based on this approach, such as the Smart Sampling Kalman Filter [6] and the Unscented Kalman Filter [7].

When relying on IMU measurements, gyroscope data can be integrated to yield the orientation of each limb; however, the drift accumulates and induces error over time [8]. To reduce the drift, accelerometer signals can be used in combination with the gyroscope within a stochastic filtering framework. Without taking the human body kinematic model into account, the orientation of each limb can be estimated separately [9], [10], and kinematic constraints can be incorporated for estimating joint states via optimization in post processing [11]. However, to estimate the joint angles directly, human body model constraints need to be incorporated in the estimation process. In [12] and [13] the authors used the EKF and UKF to estimate arbitrary 3D leg and arm motions by using a set of rigid links connected with hinge joints as the human skeleton model. For tracking a 3-segment inverted pendulum motion, in [14] the authors used a model-based extended Kalman filter with quaternion rotation parametrization. Combining the skeleton model with kinematic constraints or knowledge of the motion profile can further enhance pose estimation accuracy [8], [15].

Kinematic models considered in human motion applications rely on rigid links articulated with rotational and/or translational joints with arbitrary number of degrees of freedom. Since rotations are generally non-Euclidean values,

the range of the associated variables is most appropriately represented using curved geometries called manifolds. By accounting for the manifold geometry, the pose estimation performance can be significantly improved [16]. Examples of applications of manifolds in human-oriented applications include visual object tracking [17], action recognition [18], and gait modeling [19].

Typically, rotation is described using special orthogonal group $\text{SO}(n)$, while pose (a combination of rotation and translation) is described using special Euclidean group $\text{SE}(n)$, both having their 2 and 3-dimensional counterparts ($n = 2, 3$)¹, where $\text{SO}(2)$ and $\text{SO}(3)$, $\text{SE}(2)$ and $\text{SE}(3)$, represent rotation and transformation matrices in 2D and 3D, respectively. Notably, both SO and SE, as well as many types of manifolds appearing in physical sciences and engineering, belong to a family of matrix Lie groups. Several studies have developed the uncertainty representation and association to Lie groups, such as $\text{SO}(3)$ [20], $\text{SE}(2)$ [21], $\text{SE}(3)$ [22]. Recently, a number of groups have developed theoretically rigorous approaches for filtering on Lie groups. In [23] an EKF respecting the geometry of matrix Lie groups was proposed, while [24] proposed the unscented transform-based filtering approach and the particle-based approach was presented in [25].

In this paper we present a novel algorithm for full body human pose estimation on Lie groups, capable of using either 3D marker position measurements and/or IMU measurements. The contributions of the paper are the following: First, we develop an approach for stochastic inference of human pose, where the state space is defined to reside on a Lie group. The geometry of the state space is explicitly accounted for by applying LG-EKF, where the prediction step is assumed to follow a constant acceleration model. In preliminary work, we presented the derivation of Jacobians on Lie groups needed for the evaluation of the update step for 3D marker position measurements [2], and for gyro and accelerometer measurements [3]. Second, we extend the preliminary work to provide an extensive experimental evaluation of the proposed estimation approach in both simulations and with real-world human motion data. Third, we address the problem of filter observability. For this purpose we use an approach relying on the evaluation of Lie derivatives, and analyze the observability of a chain of $\text{SO}(3)$ elements, a generalization of human body kinematics. The results of the observability analysis are validated by emulating a human arm.

The paper is organized as follows. Section II presents the preliminaries of Lie groups and describes marker and IMU based human motion estimation approach on Lie groups. Section III provides the observability analysis of a generalized chain model relying on markers. Validation results in simulations and in real-world experiments are given in Sections IV and V, while Section VI concludes the paper.

II. MOTION ESTIMATION ON LIE GROUPS

¹Note that the matrix dimension does not directly match the number of degrees of freedom, e.g., $\text{SO}(2)$ has 1 degree of freedom

A. Lie groups and Lie algebra

A Lie group G is a group which also has the structure of a smooth manifold, resulting in the nice property that a curved object, like a Lie group G , can be almost completely represented by a flat approximation. This flat representation, that can be associated to each point $X \in G$ is called the Lie algebra of G , which we denote by \mathfrak{g} [26]. More formally, the Lie algebra $\mathfrak{g} \subset \mathbb{R}^{n \times n}$ associated to a p -dimensional matrix Lie group $G \subset \mathbb{R}^{n \times n}$ is a p -dimensional vector space defined by a basis consisting of p real matrices E_r , $r = 1, \dots, p$, often referred to as generators [27]. This property enables us to carry out sophisticated analysis of the group elements by transferring them to the algebra and using familiar mathematical tools of the Euclidean vector space. To switch between the two representations, we use the matrix exponential \exp_G and logarithm \log_G

$$\exp_G : \mathfrak{g} \rightarrow G \quad \text{and} \quad \log_G : G \rightarrow \mathfrak{g}. \quad (1)$$

Furthermore, a linear isomorphism between the p -dimensional Lie algebra \mathfrak{g} and the Euclidean space \mathbb{R}^p is established

$$[\cdot]_G^\vee : \mathfrak{g} \rightarrow \mathbb{R}^p \quad \text{and} \quad [\cdot]_G^\wedge : \mathbb{R}^p \rightarrow \mathfrak{g}. \quad (2)$$

For brevity, we will use the following notation [28]

$$\exp_G^\wedge(x) = \exp_G([x]_G^\wedge) \quad \text{and} \quad \log_G^\vee(X) = [\log_G(X)]_G^\vee, \quad (3)$$

where $x \in \mathbb{R}^p$ and $X \in G$.

Lie groups are generally non-commutative, i.e., $XY \neq YX$, which necessitates the use of the adjoint representation of G on \mathfrak{g} , Ad_G , which can be interpreted as representing the elements of the group as linear transformations of the group's algebra. For a more thorough treatment of the subject the reader is directed to [26], [29], [30].

B. Construction of the state space

Our human body modeling approach and the corresponding state space follow the approach proposed for a human leg in [2] and a human arm in [3]. Each joint of interest in the state vector is represented by a corresponding Lie group. For one dof revolute joints, the special orthogonal group $\text{SO}(2)$ is used, while 3 dof spherical joints are modelled with the special orthogonal group $\text{SO}(3)$. To localize the human in an inertial frame, a special euclidean group member $\text{SE}(3)$ is used to connect the origin frame with the base of the body, modeling both translational and rotational motion. Finally, the Lie group members are concatenated via a Cartesian product, starting with $\text{SE}(3)$, and extending with either $\text{SO}(2)$ or $\text{SO}(3)$ groups, to form the system state vector.

As an example we consider a state space model of the full human body employing either Lie groups or Euler angles, as illustrated in Fig. 1. A potential representation for the

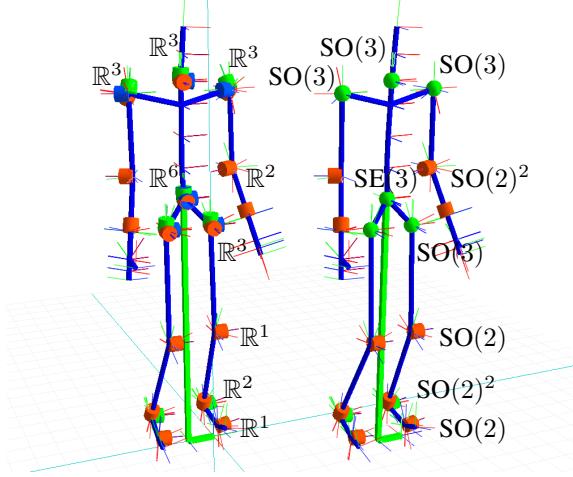


Fig. 1: Comparison of Euler angle (left) and Lie group (right) skeleton models. Revolute $\text{SO}(2)$ joints are red cylinders, spherical $\text{SO}(3)$ joints are green spheres, prismatic $\text{SE}(3)$ joints are green square rods.

positional variables of this model is

$$\begin{aligned} \text{G}_{\text{pos}} = & \underbrace{\text{SE}(3)}_{\text{base pose}} \times \underbrace{\text{SO}(3)}_{\text{L hip}} \times \underbrace{\text{SO}(3)}_{\text{R hip}} \times \underbrace{\text{SO}(2)}_{\text{L knee}} \times \underbrace{\text{SO}(2)}_{\text{R knee}} \times \\ & \times \underbrace{\text{SO}(2)^2}_{\text{L ankle}} \times \underbrace{\text{SO}(2)^2}_{\text{R ankle}} \times \underbrace{\text{SO}(2)}_{\text{L toes}} \times \underbrace{\text{SO}(2)}_{\text{R toes}} \times \underbrace{\text{SO}(3)}_{\text{neck}} \times \\ & \times \underbrace{\text{SO}(3)}_{\text{L shoulder}} \times \underbrace{\text{SO}(3)}_{\text{R shoulder}} \times \underbrace{\text{SO}(2)^2}_{\text{L elbow}} \times \underbrace{\text{SO}(2)^2}_{\text{R elbow}}, \end{aligned} \quad (4)$$

where $\text{SO}(2)^2 = \text{SO}(2) \times \text{SO}(2)$. Alongside positional variables, in this work we also want to estimate joint velocities and accelerations. For this purpose we associate velocity and acceleration to each dof of each joint of the body, where each component is represented as a real-numbered value. The full state of the system X_k at time instant k is then of the form

$$X_k = \text{blkdiag}\{\theta_k, \omega_k, \alpha_k\} \in \text{G} \quad (5)$$

$$\theta_k = \text{blkdiag}\{\theta_k^1, \dots, \theta_k^n\} \in \text{G}_{\text{pos}}$$

$$\omega_k = \text{blkdiag}\{\omega_k^1, \dots, \omega_k^n\} \in \mathbb{R}^{p_1} \times \dots \times \mathbb{R}^{p_n}$$

$$\alpha_k = \text{blkdiag}\{\alpha_k^1, \dots, \alpha_k^n\} \in \mathbb{R}^{p_1} \times \dots \times \mathbb{R}^{p_n},$$

where θ_k^i is the position of the i -th joint, ω_k^i is the velocity of the i -th joint, $\alpha_k^i \in \mathbb{R}^{p_i}$ is the acceleration of the i -th joint, n is the number of joints of a body, while p_i is the number of dofs of the i -th joint. Note that any Euclidean vector can be viewed as a pure translation and thus is an element of a subgroup of $\text{SE}(n)$ [16]. In order to include velocities and accelerations in the state $X_k \in \text{G}$ we use their matrix representation obtained by simple matrix embedding.

C. Motion prediction and measurement update steps

The prediction and update steps of the filter strongly resemble the standard EKF procedure. First, we assume that the motion model of the system can be described as an action on the group [23]

$$X_{k+1} = f(X_k, n_k) = X_k \exp_G^\wedge(\hat{\Omega}_k + n_k), \quad (6)$$

where $X_k \in \text{G}$ is the state of the system at time k (5), $n_k \sim \mathcal{N}_{\mathbb{R}^p}(\mathbf{0}^{p \times 1}, Q_k)$ is zero mean white Gaussian noise with covariance Q_k , and $\hat{\Omega}_k = \Omega(X_k) : \text{G} \rightarrow \mathbb{R}^p$ is a non-linear \mathcal{C}^2 function. Thus, a noiseless mean and covariance estimates can be propagated as

$$\mu_{k+1|k} = \mu_k \exp_G^\wedge(\hat{\Omega}_k), \quad (7)$$

$$P_{k+1|k} = \mathcal{F}_k P_k \mathcal{F}_k^T + \Phi_{\text{G}}(\hat{\Omega}_k) Q_k \Phi_{\text{G}}(\hat{\Omega}_k)^T. \quad (8)$$

Where \mathcal{F}_k is the matrix Lie group Jacobian equivalent of $f(X_k, n_k)$ and Φ_{G} appears due to the displacement of the tangential space during the prediction step.

The discrete measurement model on the matrix Lie group is given as

$$Z_{k+1} = h(X_{k+1}) \exp_{\text{G}'}^\wedge(m_{k+1}), \quad (9)$$

where $Z_{k+1} \in \text{G}'$, $h : \text{G} \rightarrow \text{G}'$ is a \mathcal{C}^1 function, G' is a p' -dimensional Lie group and $m_{k+1} \sim \mathcal{N}_{\mathbb{R}^{q \times 1}}(\mathbf{0}^{q \times 1}, R_{k+1})$ is zero-mean white Gaussian noise with covariance R_{k+1} . The Kalman gain K_{k+1} and innovation vector ν_{k+1} are calculated as

$$\begin{aligned} K_{k+1} &= P_{k+1|k} \mathcal{H}_{k+1}^T (\mathcal{H}_{k+1} P_{k+1|k} \mathcal{H}_{k+1}^T + R_{k+1})^{-1} \\ \nu_{k+1} &= K_{k+1} \log_{\text{G}'}^\vee(h(\mu_{k+1|k})^{-1} Z_{k+1}). \end{aligned} \quad (10)$$

The matrix \mathcal{H}_{k+1} can be seen as the matrix Lie group equivalent of the Jacobian of $h(X_{k+1})$, and is given as

$$\mathcal{H}_{k+1} = \frac{\partial}{\partial \epsilon} \log_{\text{G}'}^\vee(h(\mu_{k+1|k})^{-1} h(\mu_{k+1|k}^\epsilon))|_{\epsilon=0}, \quad (11)$$

where $h(\mu_{k+1|k}^\epsilon) = h(\mu_{k+1|k} \exp_G^\wedge(\epsilon))$, describes the variation of measurements for an infinitesimal motion ϵ . Finally, the measurement update step is calculated as

$$\mu_{k+1} = \mu_{k+1|k} \exp_G^\wedge(\nu_{k+1}) \quad (12)$$

$$P_{k+1} = \Phi_{\text{G}}(\nu_{k+1}) (I - K_{k+1} \mathcal{H}_{k+1}) P_{k+1|k} \Phi_{\text{G}}(\nu_{k+1})^T.$$

For a more formal derivation, the reader is referred to [23]. In our previous works we derived a constant acceleration motion model and the necessary Jacobian \mathcal{H}_{k+1} to perform motion estimation using optical markers, gyroscopes, and accelerometers as measurements [3], [31]. Note that we use the discrete LG-EKF, which has been extended to continuous discrete version in [23], wherein analysis was performed on the state and measurement residing on $\text{SO}(3)$; however, even though continuous prediction can potentially improve the accuracy of the motion prediction step, in our case for the full body skeleton the numerical integration would induce higher computation costs and, for a constant update rate, result in slightly improved estimation given typical motion capture sample rates.

III. OBSERVABILITY ANALYSIS

A system is observable if its state at a certain time instant can be uniquely determined given a finite sequence of its input and outputs [32]. In the case of an estimation problem dealing with the kinematic model of the human body, system observability translates to the property that the

joint states can be determined based on the measurements, e.g., markers or IMUs. Practically, observability provides information about the minimal sufficient measurement setup needed for determining all the joint states of interest. In the following, we focus on the joint orientation observability, since joint velocities and accelerations are observable via direct dependency, once it is shown that the orientation is observable.

Approaches for observability analysis of linear time-invariant systems include well-established tests including the rank of the Gramian matrix [33] or the Popov–Belevitch–Hautus (PBH) test [34]. The PBH test cannot standardly be applied to linearized systems because of the time-invariance requirement, although some recent generalizations of the PBH explore the use of nonlinear eigenvalues [35]. The test based on the Gramian matrix could theoretically be applied, but practically becomes intractable for 6dof applications; hence, evaluation via the Gramian matrix is usually performed numerically [36].

The observability analysis in this work relies on a differential geometric characterization of observability, which leads to the evaluation of the observability rank condition based on Lie derivatives [37]. This approach has been used in various applications, beginning with the observability of a map-based single robot localization [38] and cooperative localization of pairs of robots [39] in 2D, and subsequently with a full SLAM system [40] and camera-odometry extrinsic calibration in 2D [41]. Although early applications dealt with 2D state space, the approach was also successfully applied for observability properties of camera–IMU extrinsic parameters calibration [36], [42] dealing with 3D applications, i.e., 6dof transformations. Furthermore, in [43] the same approach is used for evaluating estimator inconsistency in a vision aided inertial navigation system.

To the best of the authors’ knowledge, this is the first paper to analyze the observability of an arbitrary kinematic chain by accounting for its structure and length based on marker position measurements. The analysis performed herein can then be directly applied to the full body human motion estimation task.

A. Nonlinear observability

In [37], the state space is assumed to be a smooth manifold and the considered non-linear system has the following form

$$\begin{cases} \dot{x} = f(x, u) \\ z = h(x) \end{cases}, \quad (13)$$

where $x \in \mathbb{R}^p$ is the state vector, $z \in \mathbb{R}^q$ is the measurement vector, u is the control input, $f(\cdot)$ is the nonlinear system state equation, and $h(\cdot)$ is the nonlinear measurement equation. Some previous works relied on incorporating a subset of measurements through the input signal [36], [42], [43], while in our approach we assume zero control inputs, i.e., $u = 0$. The zeroth-, first- and second-order Lie derivatives L of the

function h with respect to f at x are formed recursively as follows

$$\begin{aligned} L^0 h_s(x) &= h_s(x) \\ L_f^1 h_s(x) &= \nabla h_s(x) \cdot f(x) \\ L_f^2 h_s(x) &= L_f^1 \left(L_f^1 h_s(x) \right) = \nabla L_f^1 h_s(x) \cdot f(x) \\ &\vdots \\ L_f^{i+1} h_s(x) &= \nabla L_f^i h_s(x) \cdot f(x), \end{aligned} \quad (14)$$

where ‘ \cdot ’ denotes the vector inner product, s represents the measurement of the s -th sensor, and ‘ ∇ ’ is the gradient operator (partial derivative). Then, the observability matrix is defined as the matrix with rows

$$\mathcal{O} = \left\{ \nabla L_f^l h_s(x) \mid s = 1, \dots, m; l \in \mathbb{N} \right\}, \quad (15)$$

where l is the order of the Lie derivative. If \mathcal{O} is full rank, the system is locally weakly observable [36].

B. Nonlinear observability on Lie groups

Since Lie algebra is tangential to the pertaining Lie group at every point of the state space, analysis leading to locally weak observability can be considered in the Lie algebra space without loss of generality. Therefore, similarly to the derivations of the Jacobians in Sec. II-C, instead of evaluating gradients directly over the variables on the group, we rather redefine the problem such that the system is described in terms of local Lie algebra coordinates of the current state X . The continuous-time system used for the analysis is then given as

$$\begin{cases} \dot{x} = \Omega(X) \\ z = \log_{G'}^\vee(h(X)) \end{cases}, \quad (16)$$

where the motion model $\Omega(X)$ is given in the Lie algebra local coordinates. The zeroth-, first- and second-order Lie derivatives \mathcal{L} used for the Lie group based formulation of the system, with measurement function h with respect to Ω at X are then given as

$$\begin{aligned} \mathcal{L}^0 h_s(X) &= \log_{G'}^\vee(h_s(X)) \\ \mathcal{L}_\Omega^1 h_s(X) &= \nabla_G \log_{G'}^\vee(h_s(X)) \cdot \Omega(X) \\ &= \frac{\partial \log_{G'}^\vee(h_s(X \exp_G^\wedge(\xi)))}{\partial \xi} \Big|_{\xi=0} \cdot \Omega(X) \\ \mathcal{L}_\Omega^2 h_s(X) &= \mathcal{L}_\Omega^1 \left(\mathcal{L}_\Omega^1 h_s(X) \right) = \nabla_G \mathcal{L}_\Omega^1 h_s(X) \cdot \Omega(X) \\ &\vdots \\ \mathcal{L}_\Omega^{i+1} h_s(X) &= \nabla_G \mathcal{L}_\Omega^i h_s(X) \cdot \Omega(X), \end{aligned} \quad (17)$$

where ∇_G is a gradient operator such that some noisy perturbation ξ is added to the current state and then partial derivatives are evaluated as ξ approaches 0. Note that a similar idea was used for the evaluation of Lie group Jacobian (11). Based on these expressions, the observability matrix is defined with rows

$$\mathcal{O} = \left\{ \nabla_G \mathcal{L}_\Omega^l h_s(X) \mid s = 1, \dots, m; l \in \mathbb{N} \right\}. \quad (18)$$

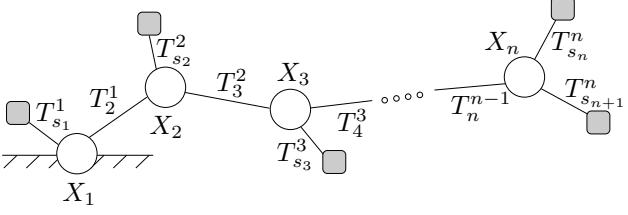


Fig. 2: Illustration of an $\text{SO}(3)$ joints chain. A single marker is attached to each joint, while two markers are attached to the last joint.

Again, if \mathcal{O} is full rank, the system is locally weakly observable. Both matrices (15) and (18) can have an infinite number of rows; however, to prove that they are full rank, it is sufficient to show that a subset of rows are linearly independent [36]. There exists no systematic method for selecting suitable Lie derivatives for designing the observability matrix. Hence, this matrix is formed by sequentially considering directions of the state space along which the gradient of each of the candidate Lie derivatives provides information [36].

C. Marker based observability

We now consider observability of joint states relying on markers attached to an arbitrarily long chain of $\text{SO}(3)$ joints. An illustration of a general $\text{SO}(3)$ joints chain is given in Fig. 2, where all but the last joint have a single marker attached, while the last joint has two markers. The marker position through forward kinematics is calculated as

$$h_{s_i}(X) = \mathcal{K}_{s_i}^0 \vec{o}, \quad (19)$$

where the term $\mathcal{K}_{s_i}^0 = \mathcal{K}_{s_i}^0(X) \in \text{SE}(3)$ represents the transformation from marker s_i frame to frame 0, and $\vec{o} \in \mathbb{R}^4$ is the origin represented in homogeneous coordinates. In the proof of the observability of this system, we will need only the zeroth-order Lie derivative $\mathcal{L}^0 h_{s_i}(X)$.

$$\nabla_G^\theta \mathcal{L}^0 h_{s_i}(X) = \frac{\partial \mathcal{L}^0 h_{s_i}(X \exp_G^\wedge(\xi))}{\partial \xi} \Big|_{\xi=0} = \mathcal{H}_{s_i}, \quad (20)$$

where \mathcal{H}_{s_i} is a matrix Jacobian as described in [31], and $X \in \text{SE}(3) \times \dots \times \text{SE}(3)$ with all translational components set to zero. If the considered marker s_i is affected by the joint j , the respective Jacobian is given as

$$\mathcal{H}_{s_i}^j = \mathcal{K}_j^0 X_j E(\mathcal{K}_{s_i}^j \vec{o}) \quad (21)$$

where \mathcal{K}_j^0 and $\mathcal{K}_{s_i}^j$ are transformations from the joint j to frame 0 and from marker s_i frame to the joint frame respectively, otherwise $\mathcal{H}_{s_i}^j = 0$. The operator $E(\cdot)$ is constructed so that $\mathfrak{so}(3)$ generators are embedded into $\mathfrak{se}(3)$ in the positions of rotational generators, while the translational generators of $\mathfrak{se}(3)$ are set to zero:

$$E(\vec{t}) = \begin{bmatrix} E_x t & E_y t & E_z t & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \in \mathfrak{se}(3), \quad (22)$$

where \vec{t} are homogeneous coordinates, t are its first three elements, and E_x, E_y, E_z are rotational generators of $\mathfrak{so}(3)$.

We re-write (21) by applying properties of the operator $E(\cdot)$ as follows

$$\begin{aligned} \mathcal{H}_{s_i}^j &= \mathcal{K}_j^0 X_j E(\mathcal{K}_{s_i}^j \vec{o}) = R_j^0 X_j E(\mathcal{K}_{s_i}^j \vec{o}) \\ &= E(R_j^0 X_j \mathcal{K}_{s_i}^j \vec{o}) R_j^0 X_j. \end{aligned} \quad (23)$$

We first applied the property that $\mathcal{K}E(\vec{t}) = RE(\vec{t})$, where $\mathcal{K} \in \text{SE}(3)$ is the full kinematics and $R \in \text{SE}(3)$ is the rotational kinematics with the translation elements set to zero. This property holds due to the form of $E(\cdot)$, i.e., the translation components of all the transformation matrices preceding E can be discarded and filled with zeros, since the last row/column of $E(\cdot)$ are zero, hence eliminating translational components via multiplication. Second, we apply the property $RE(\vec{t}) = E(R\vec{t})R$, since because of the construction of X the product $R_j^0 X_j$ represents rotation with zero translational components.

The observability matrix \mathcal{O} is constructed such that $n+1$ is the number of markers, corresponding to $n+1$ rows, and n is the number of joints, corresponding to n columns:

$$\mathcal{O} = \begin{bmatrix} \mathcal{H}_{s_1}^1 & 0 & \dots & 0 \\ \mathcal{H}_{s_2}^1 & \mathcal{H}_{s_2}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{H}_{s_n}^1 & \mathcal{H}_{s_n}^2 & \dots & \mathcal{H}_{s_n}^n \\ \mathcal{H}_{s_{n+1}}^1 & \mathcal{H}_{s_{n+1}}^2 & \dots & \mathcal{H}_{s_{n+1}}^n \end{bmatrix}. \quad (24)$$

The observability matrix contains non-zero elements whenever a joint motion affects sensor measurements. Otherwise, a zero element appears if a joint motion does not affect the considered marker measurement. For example, by considering Fig. 2, we can see that motion of joint X_1 affects the measurements of all the markers, while motion of the joint X_3 affects only measurements of the last two markers.

In order to prove that (15) is full rank, we now proceed with Gaussian elimination. By inserting the last expression of (23), the observability matrix becomes

$$\mathcal{O} = \begin{bmatrix} E(R_1^0 X_1 \mathcal{K}_{s_1}^1 \vec{o}) R_1^0 X_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ E(R_1^0 X_1 \mathcal{K}_{s_n}^1 \vec{o}) R_1^0 X_1 & \dots & E(R_n^0 X_n \mathcal{K}_{s_n}^n \vec{o}) R_n^0 X_n \\ E(R_1^0 X_1 \mathcal{K}_{s_{n+1}}^1 \vec{o}) R_1^0 X_1 & \dots & E(R_n^0 X_n \mathcal{K}_{s_{n+1}}^n \vec{o}) R_n^0 X_n \end{bmatrix}$$

Since right multiplication of the whole column by the same transformation only performs a linear combination of row elements, we multiply the i -th column with $(R_i^0 X_i)^{-1}$ elements outside the brackets obtaining

$$\mathcal{O} = \begin{bmatrix} E(R_1^0 X_1 \mathcal{K}_{s_1}^1 \vec{o}) & \dots & 0 \\ \vdots & \ddots & \vdots \\ E(R_1^0 X_1 \mathcal{K}_{s_n}^1 \vec{o}) & \dots & E(R_n^0 X_n \mathcal{K}_{s_n}^n \vec{o}) \\ E(R_1^0 X_1 \mathcal{K}_{s_{n+1}}^1 \vec{o}) & \dots & E(R_n^0 X_n \mathcal{K}_{s_{n+1}}^n \vec{o}) \end{bmatrix}. \quad (25)$$

In order to proceed further with simplifying the observation matrix for Gaussian elimination, we look next at how we can decompose the kinematic transformations. Kinematics between the i -th joint, preceding the j -th sensor, is given as

$$\mathcal{K}_{s_j}^i = T_{i+1}^i X_{i+1} \cdots T_j^{j-1} X_j T_{s_j}^j. \quad (26)$$

It follows that

$$\begin{aligned}
E(R_i^0 X_i \mathcal{K}_{s_j}^i \vec{o}) &= E(R_i^0 X_i T_{i+1}^i X_{i+1} \cdots X_{j-1} T_j^{j-1} X_j T_{s_j}^j \vec{o}) \\
&= E(R_i^0 X_i \vec{t}_{i+1}^i + \cdots + R_{j-1}^0 X_{j-1} \vec{t}_j^{j-1} + R_j^0 X_j \vec{t}_{s_j}^j) \\
&= E\left(\sum_{k=i}^{j-1} R_k^0 X_k \vec{t}_{k+1}^k + R_j^0 X_j \vec{t}_{s_j}^j\right) \\
&= E\left(\sum_{k=i}^{j-1} \tau_{k+1}^k + \tau_{s_j}^j\right), \tag{27}
\end{aligned}$$

where \vec{t}_{i+1}^i represents the translation component of transformation matrix T_{i+1}^i , and $\tau_{k+1}^k = R_k^0 X_k \vec{t}_{k+1}^k$. We continue the Gaussian elimination procedure by writing the observability matrix elements in terms of sums as in (27), and consider the upper left corner of the matrix

$$\begin{bmatrix} E(\tau_{s_1}^1) & 0 & 0 & \cdots \\ E(\tau_2^1 + \tau_{s_2}^2) & E(\tau_{s_2}^2) & 0 & \cdots \\ E(\tau_2^1 + \tau_3^2 + \tau_{s_3}^3) & E(\tau_3^2 + \tau_{s_3}^3) & E(\tau_{s_3}^3) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{28}$$

Since it holds that $E(a+b) = E(a) + E(b)$, by applying several simple column and row subtractions, we obtain

$$\begin{bmatrix} E(\tau_{s_1}^1) & 0 & 0 & 0 & \cdots \\ E(\tau_2^1) & E(\tau_{s_2}^2) & 0 & 0 & \cdots \\ 0 & E(\tau_3^2 - \tau_{s_2}^2) & E(\tau_{s_3}^3) & 0 & \cdots \\ 0 & 0 & E(\tau_4^3 - \tau_{s_3}^3) & E(\tau_{s_4}^4) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} = A.$$

We now proceed by evaluating the nullspace of matrix A , i.e., to determine if there exists an $x \neq 0$ such that

$$Ax = 0. \tag{29}$$

which would indicate that A is not full rank. For this purpose we extract the 2-nd and 3-rd rows of the A matrix, and separate vector x into components as follows

$$\begin{bmatrix} E(\tau_2^1) & E(\tau_{s_2}^2) & 0 & 0 & \cdots \\ 0 & E(\tau_3^2 - \tau_{s_2}^2) & E(\tau_{s_3}^3) & 0 & \cdots \end{bmatrix} \begin{bmatrix} x_{1:3} \\ x_{4:6} \\ x_{7:9} \\ \vdots \end{bmatrix} = 0. \tag{30}$$

The analysis can then be generalized to other pairs of rows given the symmetry of A . The part of the x vector that corresponds to the 4-th to 6-th elements is

$$\begin{bmatrix} E(\tau_{s_2}^2) \\ E(\tau_3^2 - \tau_{s_2}^2) \end{bmatrix} x_{4:6} = 0. \tag{31}$$

We can observe that there is no solution $x_{4:6} \neq 0$ unless $\tau_{s_2}^2 = \alpha \tau_3^2$, where $\alpha \neq 0$. A similar observation then applies to the entire A and x , meaning that A is full rank except for some special cases that are further discussed below.

D. Three joint example

Using the results given in (31), we consider observability of a three joint chain emulating a human arm or leg. The illustration of the three joint chain is given in Fig. 2, when $n = 3$. The complete matrix A for this system evaluates to

$$\begin{bmatrix} E(\tau_{s_1}^1) & 0 & 0 \\ E(\tau_2^1) & E(\tau_{s_2}^2) & 0 \\ 0 & E(\tau_3^2 - \tau_{s_2}^2) & E(\tau_{s_3}^3) \\ 0 & E(\tau_3^2 - \tau_{s_2}^2) & E(\tau_{s_4}^4) \end{bmatrix}. \tag{32}$$

By decomposing terms, as in the previous section, we identify the conditions that lead to observability violations

- the terms $\tau_{s_1}^1$ and τ_2^1 correspond to

$$\tau_{s_1}^1 = R_1^0 X_1 t_{s_1}^1 \text{ and } \tau_2^1 = R_1^0 X_1 t_2^1,$$

hence observability would be violated if $t_{s_1}^1 = \alpha t_2^1$, i.e., if $t_{s_1}^1$ and t_2^1 are colinear, which occurs if sensor s_1 is placed on the axis defined by connection of joints 1–2.

Analogous results can be obtained relating sensor s_2 and the connection between joints 2 and 3, and relating the axes of sensors s_3 and s_4 . Generally, we can conclude that the observability is violated if a marker is placed on the axis connecting its associated joint with the succeeding one.

IV. SIMULATION VALIDATION

The proposed approach is extensively validated in simulation. We first compare the proposed approach to estimation utilizing a standard kinematic model defined with revolute and prismatic joints [44], as shown in Fig. 1 (left). We discuss the gimbal lock limitation of modeling kinematics with Euler angles and show that because group based models do not have gimbal lock, LG-EKF not only provides better estimation but is also easier to tune for best performance. To ensure that the estimation improvement is due to the difference in representation and not selected filtering method we employ the Unscented Kalman filter (UKF) in addition to EKF when estimating gimbal lock Euler angles. UKF propagates carefully chosen samples through the non linear state update and measurement functions to estimate the mean, covariance, and cross correlation of the state and observation. It captures the mean and covariance up to second order and does not require explicit calculation of the Jacobians [45]. The state of the EKF and UKF is defined as the position q , velocity \dot{q} , and acceleration \ddot{q} of the joints. Just as in LG-EKF, the constant acceleration model is used.

To compare the estimated rotation of each SO(3) joint with the ground truth, we use the deviation from the identity matrix as the distance metric [46]

$$\mathcal{D}_F = \|I - R_e^T R_{gt}\|_F, \tag{33}$$

where R_e and R_{gt} are the estimated and ground truth rotation matrices of each joint and $\|\cdot\|_F$ denotes the Frobenius norm, which is functionally equivalent to the geodesic on SO(3) [46]. Subsequently, we verify the observability analysis on a model of a human arm, and demonstrate the sensor setup under which the arm orientation is observable.

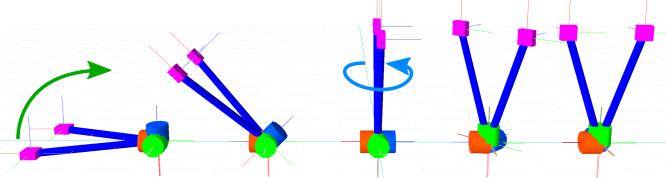


Fig. 3: Snapshots of the Euler model at different times throughout the gimbal lock validation motion. The green and light blue arrows represent the direction of model rotation about world y and z axis respectively. To ensure the system is observable, two markers (pink boxes) are attached at a fixed offset (illustrated by the blue bar) from a 3dof joint. In Euler angle representation as the second angle (green cylinder) approaches gimbal lock the other two axes (red and blue) align, removing a degree of freedom. To rotate about world z axis the model must first leave gimbal lock, which during estimation typically results in high velocities and increased error in the Euler joints.

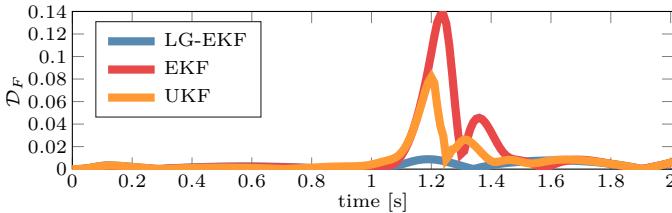


Fig. 4: LG-EKF, EKF, and UKF marker based estimation during gimbal lock. When the system enters the gimbal lock at 1 second UKF and EKF are unable to accurately track the state since the rotation is about the lost degree of freedom. Both filters incorrectly estimate high accelerations in joints to escape the lock and return to accurate tracking. LG-EKF is unaffected by gimbal lock.

A. Gimbal lock

When the Euler angle parametrization is used in the kinematic model to represent spherical joints, two of the axes can become aligned and a degree of freedom is lost; this is referred to as gimbal lock. Contrarily, the SO(3) representation maintains perpendicularity of axes allowing LG-EKF to inherently estimate all rotations accurately, as long as Lie algebra unique parametrization constraints are respected.

To demonstrate the advantages of SO(3) representation over Euler angles in gimbal lock, a single spherical joint with two markers is simulated. The markers are attached at an offset to ensure observability. Since we focus on the gimbal lock problem, no noise is added to marker measurements. The simulated Euler angle model and its motion is shown in Fig. 3, where a quintic polynomial was used to generate a smooth trajectory, sampling at 100 Hz. First, with zero initial velocity and acceleration, the model experiences a 1 s rotation about the world y axis starting from 0 rad and reaching $\frac{\pi}{2}$ rad with zero final velocity and acceleration. In the Euler angle model, this motion aligns the first and third revolute joint axes reaching singularity and removing a degree of freedom (gimbal lock). Second, the model experiences another 1 s rotation in the now locked world z axis.

Figure 4 shows the error in estimation for the filters. When Euler angles enter gimbal lock, its Jacobian is singular and thus the linearized system is no longer observable. While UKF does not rely on the Jacobian and outperforms

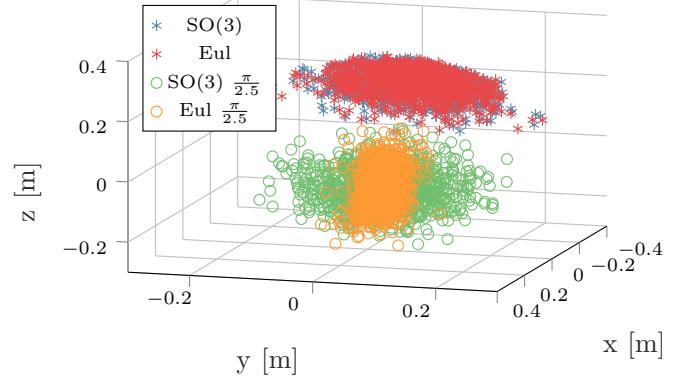


Fig. 5: Propagation of zero mean Gaussian noise with standard deviation of 0.2 through the Euler angle equations. When the revolute joints are perpendicular the Lie group (blue) and Euler angle (red) representations have the same distribution on the sphere. As soon as the second Euler angle experiences a rotation ($\frac{\pi}{2.5}$ rad) bringing the system closer to gimbal lock the resulting distribution changes (orange). The Lie group representation (green) retains the distribution properties through the rotation.

EKF, it requires high accelerations in the Euler angles to quickly come out of gimbal lock, and during this period the estimation error increases. Furthermore, propagation of the sample points through the non linear measurement function is computationally expensive and it is not feasible to run UKF in real time. For full body (30dof) estimation we would need to compute forward kinematics 181 times at each iteration since the state includes positions, velocities, and accelerations. Finally, while one expects a true spherical joint to have a consistent noise representation on the entire sphere, this is not possible with Euler angle representation. When UKF propagates the zero mean Gaussian samples through the Euler angle equations, the resulting distribution on the sphere significantly changes based on the middle joint which can bring it in and out of gimbal lock as shown in Figure 5. When estimating human motion this implies varying mobility in the hip or shoulder based on the selected Euler angle representation and current pose.

B. Observability

To verify our observability analysis we simulate the manipulator shown in Fig. 2 ($n = 3$) and investigate the convergence properties of LG-EKF in the observable and unobservable cases described in Sec. III. In the zero configuration, the manipulator is standing upright, each link length is 0.5 m. When the system is observable, the markers are attached at offsets $t_{s_1}^1, t_{s_2}^2, t_{s_3}^3, t_{s_4}^3$ of $[0.1 \ 0 \ 0.3], [0.1 \ 0 \ 0.25], [0.1 \ 0 \ 0.2], [-0.1 \ 0 \ 0.2]$ respectively. We allow LG-EKF to converge from the zero configuration to a random static pose, the observation noise, η , and initial covariance are set to 0.01, 1, and identity. Figure 6 shows the convergence of the observable case.

Next, the second marker is placed on the axis between joints two and three ($t_{s_2}^2 = [0 \ 0 \ 0.25]$). Following the similar discussion as given in subsection III-D, if $t_{s_2}^2$ and t_3^2 are colinear, which occurs if sensor s_2 is placed on the axis defined

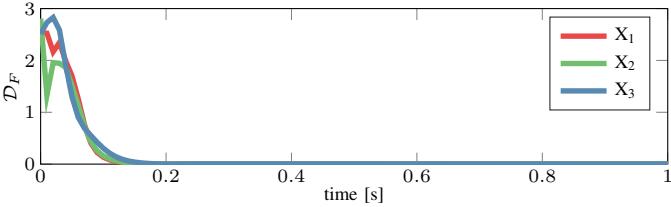


Fig. 6: LG-EKF correctly converges to the true pose of the manipulator when all of the observability criteria are satisfied

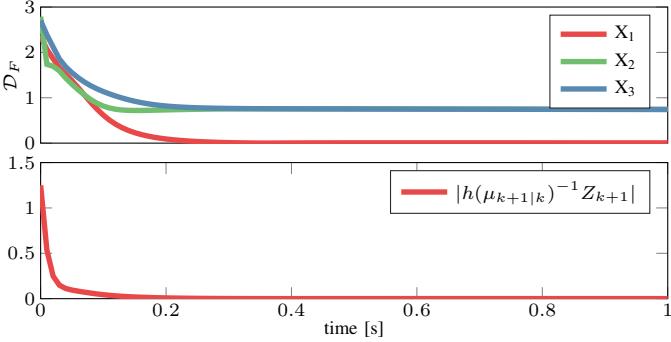


Fig. 7: When observability criteria is not satisfied, even though the measurement residuals are minimized, LG-EKF does not converge to the true pose. In this particular setup, state estimate of the second joint converges onto an arbitrary rotation about the axis between joints two and three, next to ensure the residuals of markers three and four are minimized this rotation is compensated for by the last joint.

by connection of joints 2–3, the observability criteria are not satisfied and we do not expect the filter to converge to the true pose. Figure 7 shows the convergence of each joint and the LG-EKF measurement residuals in this unobservable case.

V. REAL DATA VALIDATION

First we show that the proposed filtering method is applicable for marker based full body motion capture using real human motion data from the CMU Graphics Lab Motion Capture Database [47]. Next we verify that LG-EKF overcomes gimbal lock which can occur in complex arm movements allowing for accurate pose estimation using only wearable IMUs. Finally, we demonstrate that the consistent distribution of the SO(3) provides a better measure of movement variance.

A. CMU Dataset

Five different human actions were chosen from the CMU motion capture database [47] to cover a variety of human motions, consisting of boxing (B), dancing (D), running (R), stretching (St), and soccer kick movements (So). Two data sequences, performed by different participants, were chosen for each movement type for a total of 10 data sequences. Movement in the CMU database is captured at 120 Hz with a 12-camera Vicon motion capture system. The skeletal model of each participant is created with the Vicon BodyBuilder software and markers are attached at predetermined bony landmarks. We simplified the model by ignoring finger joints and extra joints in the spine the Vicon software generates in post processing.

TABLE I: Frequency at which each filtering strategy diverged. The divergence frequency of the LG-EKF is consistently lower than EKF.

Data seq.	B1	B2	D1	D2	R1	R2	So1	So2	St1	St2
EKF [Hz]	49	34	22	22	28	22	23	29	30	19
LG-EKF [Hz]	49	33	21	15	23	21	21	24	30	15

We used two models for pose estimation using the marker data (Fig. 1). The Euler angle model uses three orthogonal revolute joints at each 3dof joint of the simplified human skeleton (shoulders, hips, lower back and neck) and uses single revolute joints at hinge joints (elbows, forearms, knees, ankles). To position the model in space, three orthogonal prismatic joints describe the position of the model pelvis relative to the origin, followed by a 3dof Euler angle joint assembly to rotate the pelvis in space. The joint order of the Euler angle model matches the joint order of the skeleton generated by Vicon BodyBuilder. In the Lie groups model, each 3dof rotational joint assembly of the first model is replaced with an SO(3) joint group; hinge joints are represented with SO(2) groups.

Marker position sensors are rigidly attached to each model, then forward kinematics is used to generate the estimated marker positions. Three sets of marker position estimation strategies were used for each data sequence: (1) the Euler angle model and the Vicon IK joint angles, (2) the Euler angle model and a standard EKF algorithm [48], and (3) the Lie group model and the LG-EKF algorithm derived in Sec. III. For each algorithm and each data sequence, pose estimation and marker position estimation was run at multiple data sampling frequencies. The EKF and LG-EKF algorithms were run with an identity covariance matrix, marker observation noise of 0.01, and a process noise n value of 300, with parameters empirically determined to provide the best performance for each filter.

1) *Pose Estimation Divergence:* A model and filtering strategy can be considered more robust if the pose estimation algorithm can still converge at lower sampling frequencies. To evaluate the robustness of EKF and LG-EKF the 10 data sequences were resampled at decreasing frequencies. The mean absolute error (MAE) of each marker position was computed between the original motion capture data (ground truth) and the three estimation strategies. Pose estimation was considered to diverge when the overall MAE of the EKF or LG-EKF data increased above twice that of the Vicon MAE data, and never returned below this threshold. Table I shows the sampling frequencies at which each filtering strategy diverged for each data sequence.

The filters diverged at lower sampling rates for movements that were less dynamic and/or did not encounter gimbal lock as frequently. EKF pose estimation always diverged at the same or a higher sampling frequency than LG-EKF, and LG-EKF diverged at a significantly lower frequency for movements with slower, more consistent joint motion. The consistently lower divergence frequency of the LG-EKF suggests that the constant acceleration assumption on the group is better for human motion estimation than the constant acceleration assumption in the standard EKF algorithm.

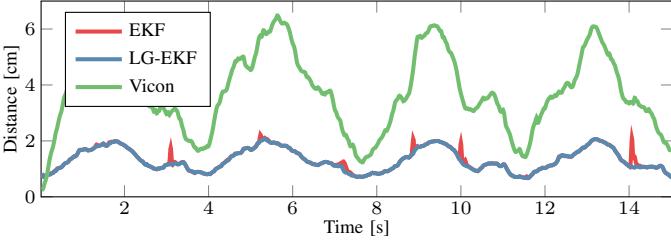


Fig. 8: Sample of average MAE of all markers for dancing data sequence, down-sampled to 30Hz. Notable spikes in position error are seen when the Euler angle model encounters gimbal lock.

TABLE II: Average MAE [mm] of all markers for all data sequences at 40Hz. EKF solutions have similar results, and both outperform the Vicon IK solution. Notation: right (r), left (l), arm (A), leg (L), end-effector (EE).

	Vicon	EKF	LG-EKF
r A	37.25	14.36	14.18
l A	34.68	15.00	14.93
Torso	23.69	18.60	18.58
Head	29.16	17.17	17.14
r L	23.93	14.57	14.44
l L	25.47	13.07	12.93
r A EE	38.20	9.19	9.00
l A EE	37.09	9.19	9.01
r L EE	20.87	10.04	9.77
l L EE	21.96	9.23	9.08
All	27.73	14.83	14.74

2) *Marker Position Mean Absolute Error:* Table II shows the average marker MAE of all data sequences down sampled to 40 Hz, averaged over all data frames. One data sequence diverged when the data was sampled lower than 50 Hz (Table I), so 9 of the 10 data sequences were used in the marker position MAE analyses. Figure 8 shows the average MAE of all markers in a sample data sequence. There is a significant difference in the MAE between the Vicon IK estimated marker positions and the two EKF implementations, with the EKF outperforming the Vicon IK solution.

Comparing the two EKF solutions, the primary difference in marker error occurs in regions where an Euler angle 3dof joint encounteres gimbal lock, and during certain highly dynamic motions where the constant acceleration assumption of each filter is severely violated. If a data sequence did not contain very dynamic movement or encounter gimbal lock in the Euler angle model, EKF and LG-EKF results showed negligible differences, especially at higher sampling frequencies.

3) *Gimbal Lock Marker Error:* We further investigate the estimation accuracy of the two filters when any of the model 3dof rotation joints were within 10% of gimbal lock. Marker error results were generated at sampling rates of 40 Hz, and then averaged over all data sequences and over all frames. The data was down sampled to 40 Hz to emphasize the difference between the two filters.

For all frames containing a 3dof joint within 10% of gimbal lock (second Euler angle within 9° of $\pm 90^\circ$), the marker MAE was tabulated for the EKF, LG-EKF, and Vicon pose estimation methods. The hip joints do not encounter gimbal lock in any of the data sequences. The shoulder joints

TABLE III: Marker position MAE [mm] during gimbal lock region of shoulders (hip is not shown due to smaller motion range and no gimbal lock). When a shoulder joint encounters gimbal lock the position error on the respective arm of the Euler angle model is significantly greater. This error propagates along the length of the arm and is largest at the end effector (EE). Notation: right (r), left (l), arm (A), leg (L), end-effector (EE).

	Gimbal lock region			Gimbal lock peak		
	Vicon	EKF	LG-EKF	Vicon	EKF	LG-EKF
r A	43.96	31.73	22.38	52.53	34.25	23.95
r A EE	39.7	29.96	15.62	53.75	33.05	14.68
l A	44.36	27.9	26.46	52.74	33.40	26.74
l A EE	45.68	19.42	18.22	53.26	33.36	17.51

TABLE IV: Marker position MAE [mm] during gimbal lock region of world-to-pelvis joint. The error difference between EKF and LG-EKF is notably higher in the EEs as compared to other regions. Notation: right (r), left (l), arm (A), leg (L), end-effector (EE).

	Gimbal lock region			Gimbal lock peak		
	Vicon	EKF	LG-EKF	Vicon	EKF	LG-EKF
r A	52.92	14.13	13.66	53.51	16.47	15.47
l A	43.15	10.5	9.98	42.90	13.25	11.65
Torso	41.3	24.25	24.21	40.59	25.02	24.93
Head	54.62	18.68	18.48	58.08	21.67	20.93
r L	26.99	13.36	12.79	29.00	15.39	14.10
l L	26.21	11.09	10.79	28.77	13.13	11.85
r A EE	53.58	5.39	4.62	54.38	8.40	6.71
l A EE	44.54	5.15	4.33	43.10	8.45	5.75
r L EE	19.8	8.15	7.54	23.22	10.67	8.88
l L EE	28.91	8.26	7.77	30.76	10.59	8.55
All	36.00	14.01	13.65	37.32	16.06	15.02

encounter gimbal lock (when arm is abducted to 90°) in the boxing, dancing and soccer kick data sequences (Table III).

The EKF error is notably higher than LG-EKF error for the right arm within the right shoulder gimbal lock region. Due to the variability of the motions in the dataset, different motions generate gimbal lock in the left and right shoulder. The right shoulder motions involve two axes of rotation, and the LG-EKF significantly outperforms EKF in this case. The left arm error difference in the left shoulder gimbal lock region is not as significant, because the left shoulder motions are primarily in the sagittal plane.

The world-to-pelvis joint encounters gimbal lock in data sequences where the participant rotates their entire body about the global vertical axis, as the second Euler angle was aligned with this axis. Since the torso revolute joint is at the beginning of the kinematic tree it affects the estimation accuracy of the entire body. Table IV (left) shows the MAE of all marker regions when the world-to-pelvis joint is near gimbal lock.

We also compare peak error in the gimbal lock region for all three methods. Table III (right) shows the arm peak marker error in shoulder gimbal lock regions and Table IV (right) shows all marker regions during the world-to-pelvis joint gimbal lock peaks.

As expected, peak marker errors are greater during gimbal lock as compared to the average error while in the gimbal lock region. The peak marker error with EKF is substantially higher than with LG-EKF at shoulder gimbal lock for both

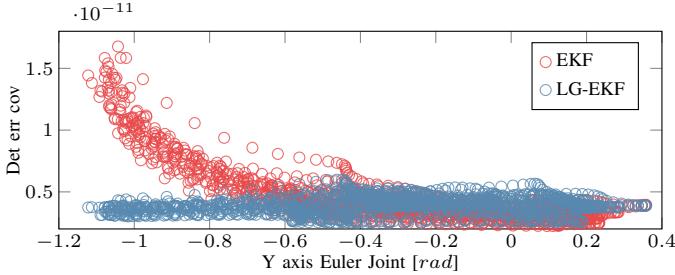


Fig. 9: Determinant of shoulder position error covariance plotted against the second Euler angle estimate. For the EKF, near the gimbal lock at $-\frac{\pi}{2}$, the error covariance increases not due to changes in the motion but due to kinematic modeling. LG-EKF covariance remains smooth through the entire state space.

arms, and still greater for each end-effector (EE). The world-to-pelvis gimbal lock peaks result in an error increase for the entire body, and notably in all end effectors.

B. Wearable IMU Dataset

The proposed approach for IMU-based pose estimation was validated through an experiment, where a dynamic figure eight human arm movement was simultaneously recorded with IMUs and motion capture. The IMUs were placed on the humerus and radius. Our IMUs utilize the MPU9250 sensors and sample at 100Hz. Prior to data collection they were calibrated with the method proposed in [49]. To compute the offset and rotation from the humerus and radius, three motion capture markers were placed on each IMU. For the ground truth data and to build a kinematic model of the participant, motion capture markers were placed on the shoulder and medial and lateral sides of the elbow and wrist.

In our earlier work we showed that since IMUs do not provide any reference measurement perpendicular to the world Z axis, EKF accumulates error about said axis with every motion pass near gimbal lock. Since LG-EKF is not affected by the gimbal lock the estimation in the locked region is improved and error accumulation is reduced [3]. We extend our analysis to demonstrate that group based representation of human motion leads to a better movement variance representation. Both LG-EKF and EKF provide a state error covariance measure which can be used for analyzing human motion. We extract the estimated shoulder position error covariance during the dynamic figure eight motion for both filters and plot its determinant against the second Euler angle estimate which causes gimbal lock (Fig. 9).

One expects that fast unpredictable motions will have a large error covariance while smooth continuous movements will maintain constant low error covariance. When utilizing the LG-EKF estimator the SO(3) shoulder position error covariance is always presented in the same three perpendicular axes aligned with the upper arm link. However, with EKF error covariance is represented about each of the Euler axes which change with the motion. Because of this, when working with EKF it is difficult to distinguish between gimbal lock and movement related variance sources. At each iteration the algorithms also project the error covariance into measurement

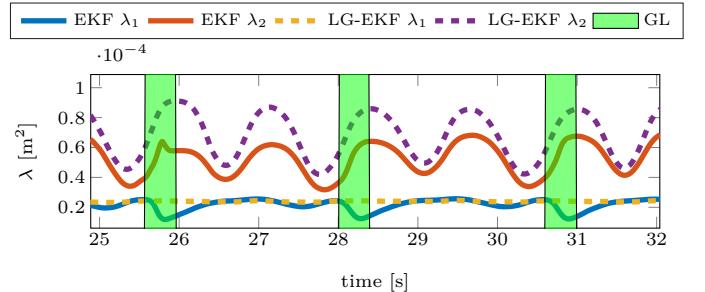


Fig. 10: Eigenvalues of the error covariance projected to the elbow marker, regions closest to gimbal lock are highlighted. Since the projection is into sensor frame, the first eigenvector is aligned with the direction of motion. Thus we expect variation in the first covariance eigenvalue throughout each motion repetition as the marker experiences directional changes as seen in the LG-EKF plot. However, it is clear that the EKF based estimator covariance is not only effected by the motion but by gimbal lock as well causing error in the covariance estimate in the direction perpendicular to motion.

space, which in motion capture applications can be used to detect missing and incorrectly labeled markers [44]. Since LG-EKF maintains a smooth covariance estimate over smooth motions, its projection will be consistent throughout and thus is better suited for such application. Figure 10 shows the main eigenvalues of the projected elbow marker covariance during marker based estimation.

VI. CONCLUSION

In this paper we proposed a novel algorithm for full body human motion estimation based on motion capture markers and wearable inertial measurement units, based on a Lie group representation of the full body skeleton. LG-EKF was used to estimate the body pose and joint velocities and accelerations, to explicitly consider the non-euclidean geometry of the state space. We show that in order for the motion of an arbitrary kinematic chain to be observable solely based on motion capture markers, a single marker on each link and two at the end effector are required. The proposed approach was extensively validated in both simulation and using real-world motion capture data. Our simulations show that the SO(3) representation of spherical joints is not affected by gimbal lock and provides a consistent error distribution, unlike Euler angle representation, which leads to higher estimation accuracy and easier tuning over regular EKF and UKF. The real world motion capture marker based experiments show that the proposed approach can successfully estimate full body pose at a lower sampling rate than EKF and significantly improves estimation near shoulder and pelvis gimbal lock regions. When estimating human motion using only wearable IMUs the proposed approach significantly outperforms EKF which accumulates more error over time. Furthermore, the estimated SO(3) joint error covariance is dependent only on the motion and not the state thus making it much more appealing for human motion analysis. The paper also presents the observability analysis of an arbitrarily long kinematic chain of SO(3) elements representing a generalization of a human body. The setup relies on marker position measurements, while the theoretical analysis rests on a differential geometric approach

based on Lie derivatives. The conclusions of the observability analysis are also further illustrated through simulations.

Future work will include extending the observability analysis to include IMU based measurements. We will also incorporate magnetometer as part of the measurement vector which can be useful for open outdoor environments. Finally, the proposed approach will be evaluated for full body motion estimation with simultaneous marker and IMU measurements.

REFERENCES

- [1] D. Kulić, G. Venture, K. Yamane, E. Demircan, I. Mizuuchi, and K. Mombaur, "Anthropomorphic movement analysis and synthesis: A survey of methods and applications," *IEEE Transactions on Robotics*, vol. 32, no. 4, pp. 776–795, 2016.
- [2] J. Ćesić, V. Joukov, I. Petrović, and D. Kulić, "Full body human motion estimation on Lie groups using 3D marker position measurements," in *International conference on Humanoid Robots (Humanoids)*. IEEE-RAS, 2016, pp. 826–833.
- [3] V. Joukov, J. Ćesić, K. Westermann, I. Marković, D. Kulić, and I. Petrović, "Human motion estimation on Lie groups using IMU measurements," in *International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 1012–1018.
- [4] K. Yamane and Y. Nakamura, "Natural motion animation through constraining and deconstraining at will," *IEEE Transactions on visualization and computer graphics*, vol. 9, no. 3, pp. 352–360, 2003.
- [5] T. Sugihara, "Solvability-unconcerned inverse kinematics by the Levenberg–Marquardt method," *IEEE Transactions on Robotics*, vol. 27, no. 5, pp. 984–991, 2011.
- [6] J. Steinbring, C. Mandery, N. Vahrenkamp, T. Asfour, and U. D. Hanebeck, "High-accuracy real-time whole-body human motion tracking based on constrained nonlinear kalman filtering," *CoRR*, vol. abs/1511.04278, 2015.
- [7] A. Aristidou and J. Lasenby, "Real-time marker prediction and CoR estimation in optical motion capture," *The Visual Computer*, vol. 29, no. 1, pp. 7–26, 2013.
- [8] V. Bonnet, C. Mazzà, J. McCamley, and A. Cappozzo, "Use of weighted Fourier linear combiner filters to estimate lower trunk 3D orientation from gyroscope sensors data," *Journal of neuroengineering and rehabilitation*, vol. 10, no. 1, p. 29, 2013.
- [9] X. Yun and E. R. Bachmann, "Design, implementation, and experimental results of a quaternion-based Kalman filter for human body motion tracking," *IEEE Transactions on Robotics*, vol. 22, no. 6, pp. 1216–1227, 2006.
- [10] G. Ligorio and A. M. Sabatini, "A novel Kalman filter for human motion tracking with an inertial-based dynamic inclinometer," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 8, pp. 2033–2043, 2015.
- [11] H. Zhou, T. Stone, H. Hu, and N. Harris, "Use of multiple wearable inertial sensors in upper limb motion tracking," *Medical engineering & physics*, vol. 30, no. 1, pp. 123–133, 2008.
- [12] J. F. Lin and D. Kulić, "Human pose recovery using wireless inertial measurement units," *Physiological Measurement*, vol. 33, 2012.
- [13] M. El-Gohary and J. McNames, "Shoulder and elbow joint angle tracking with inertial sensors," *Biomedical Engineering, IEEE Transactions on*, vol. 59, no. 9, pp. 2635–2641, 2012.
- [14] A. Szczesna and P. Pruszowski, "Model-based extended quaternion Kalman filter to inertial orientation tracking of arbitrary kinematic chains," *SpringerPlus*, vol. 5, no. 1, p. 1965, 2016.
- [15] V. Joukov, V. Bonnet, M. Karg, G. Venture, and D. Kulic, "Rhythmic extended kalman filter for gait rehabilitation motion estimation and segmentation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2017.
- [16] G. S. Chirikjian, *Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications*. Springer, 2012.
- [17] Z. H. Khan and I. Y.-H. Gu, "Nonlinear dynamic model for visual object tracking on Grassmann manifolds with partial occlusion handling," *IEEE Transactions on Cybernetics*, vol. 43, no. 6, pp. 2005–2019, 2013.
- [18] M. Devanne, H. Wannous, S. Berretti, P. Pala, M. Daoudi, and A. Del Bimbo, "3-D human action recognition by shape analysis of motion trajectories on Riemannian manifold," *IEEE Transactions on Cybernetics*, vol. 45, no. 7, pp. 1340–1352, 2015.
- [19] M. Ding and G. Fan, "Multilayer joint gait-pose manifolds for human gait motion modeling," *IEEE Transactions on Cybernetics*, vol. 45, no. 11, pp. 2413–2424, 2015.
- [20] T. Lee, M. Leok, and N. H. McClamroch, "Global symplectic uncertainty propagation on SO(3)," in *Conference on Decision and Control (CDC)*, no. 3. IEEE, 2008, pp. 61–66.
- [21] A. W. Long, K. C. Wolfe, M. J. Mashner, and G. S. Chirikjian, "The Banana Distribution is Gaussian: A Localization Study with Exponential Coordinates," in *Robotics: Science and Systems (RSS)*, Sydney, Australia, 2012.
- [22] T. D. Barfoot and P. T. Furgale, "Associating uncertainty with three-dimensional poses for use in estimation problems," *IEEE Transactions on Robotics*, vol. 30, no. 3, pp. 679–693, 2014.
- [23] G. Bourmaud, R. Mégret, M. Arnaudon, and A. Giremus, "Continuous-discrete extended Kalman filter on matrix Lie groups using concentrated Gaussian distributions," *Journal of Mathematical Imaging and Vision*, vol. 51, no. 1, pp. 209–228, 2015.
- [24] C. Hertzberg, R. Wagner, U. Frese, and L. Schröder, "Integrating generic sensor fusion algorithms with sound state representations through encapsulation of manifolds," *Information Fusion*, vol. 14, no. 1, pp. 57–77, 2013.
- [25] Q. Rentmeesters, P. A. Absil, P. Van Dooren, K. Gallivan, and A. Srivastava, "An efficient particle filtering technique on the Grassmann manifold," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2010, pp. 3838–3841.
- [26] J. Stillwell, *Naive Lie theory*. Springer, 2008.
- [27] W. Park, Y. Wang, and G. S. Chirikjian, "The path-of-probability algorithm for steering and feedback control of flexible needles," *The International Journal of Robotics Research*, vol. 29, no. 7, pp. 813–830, 2010.
- [28] G. Bourmaud, R. Mégret, and A. Giremus, "From intrinsic optimization to iterated extended Kalman filtering on Lie groups," *Journal of Mathematical Imaging and Vision*, vol. 55, no. 3, pp. 288–303, 2016.
- [29] J. M. Selig, "Lie groups and Lie algebras in robotics," in *Computational Noncommutative Algebra and Applications*, 2005, pp. 101–125.
- [30] K. Tapp, *Matrix Groups for Undergraduates*. American Mathematical Society, 2005.
- [31] J. Ćesić, V. Joukov, I. Petrović, and D. Kulić, "Full body human motion estimation on Lie groups using 3D marker position measurements," in *International Conference on Humanoid Robots (Humanoids)*. IEEE-RAS, 2016, pp. 826–833.
- [32] W. L. Brogan, *Modern Control Theory (3rd Ed.)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1991.
- [33] P. S. Maybeck, *Stochastic Models: Estimation and Control: Volume 1*, ser. Mathematics in Science and Engineering. Elsevier Science, 1979, no. vol. 1.
- [34] W. J. Rugh, *Linear System Theory (2Nd Ed.)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1996.
- [35] Y. Kawano and T. Ohsuka, "Nonlinear eigenvalue approach to differential Riccati equations for contraction analysis," *IEEE Transactions on Automatic Control (accepted for publication)*, vol. 62, no. 12, pp. 6497–6504, 2017.
- [36] F. M. Mirzaei and S. I. Roumeliotis, "A Kalman filter-based algorithm for IMU-camera calibration: Observability analysis and performance evaluation," *IEEE Transactions on Robotics*, vol. 24, no. 5, pp. 1143–1156, 2008.
- [37] R. Hermann and A. Krener, "Nonlinear controllability and observability," *IEEE Transactions on Automatic Control*, vol. 22, no. 5, pp. 728–740, 1977.
- [38] P. Bonnifait and G. Garcia, "Design and experimental validation of an odometric and goniometric localization system for outdoor robot vehicles," *IEEE Transactions on Robotics and Automation*, vol. 14, no. 4, pp. 541–548, 1998.
- [39] A. Martinelli and R. Siegwart, "Observability analysis for mobile robot localization," in *International Conference on Intelligent Robots and Systems (IROS)*. IEEE/RSJ, 2005, pp. 1471–1476.
- [40] K. W. Lee, W. S. Wijesoma, and J. Ibanez Guzman, "On the observability and observability analysis of SLAM," in *International Conference on Intelligent Robots and Systems (IROS)*. IEEE/RSJ, 2006, pp. 3569–3574.
- [41] A. Martinelli, D. Scaramuzza, and R. Siegwart, "Automatic self-calibration of a vision system during robot motion," in *International Conference on Intelligent Robots and Systems (IROS)*. IEEE/RSJ, 2006, pp. 43–48.
- [42] J. Kelly and G. S. Sukhatme, "Visual–inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration," *The*

- International Journal of Robotics Research*, vol. 30, no. 1, pp. 56–79, 2011.
- [43] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis, “Camera-IMU-based localization: observability analysis and consistency improvement,” *The International Journal of Robotics Research*, vol. 33, no. 1, pp. 182–201, 2013.
- [44] V. Joukov, R. D’Souza, and D. Kulić, “Human pose estimation from imperfect sensor data via the extended kalman filter,” in *International Symposium on Experimental Robotics*. Springer, 2016, pp. 789–798.
- [45] S. J. Julier and J. K. Uhlmann, “New extension of the kalman filter to nonlinear systems,” in *Signal processing, sensor fusion, and target recognition VI*, vol. 3068. International Society for Optics and Photonics, 1997, pp. 182–194.
- [46] D. Q. Huynh, “Metrics for 3D rotations: Comparison and analysis,” *Journal of Mathematical Imaging and Vision*, vol. 35, no. 2, pp. 155–164, 2009.
- [47] F. De la Torre, J. Hodgins, A. Bargteil, X. Martin, J. Macey, A. Collado, and P. Beltran, “Guide to the Carnegie Mellon University multimodal activity (CMU-MMAC) database,” p. 11, 2008.
- [48] V. Joukov, M. Karg, and D. Kulić, “Online tracking of the lower body joint angles using IMUs for gait rehabilitation,” in *IEEE Engineering in Medicine and Biology Conference*, 2014, pp. 2310–2313.
- [49] D. Tedaldi, A. Pretto, and E. Menegatti, “A robust and easy to implement method for IMU calibration without external equipments,” in *ICRA*. IEEE, 2014, pp. 3042–3049.

AUTHORS’ BIOGRAPHIES



Vladimir Joukov is a PhD candidate at the University of Waterloo Adaptive Systems Laboratory. His research focus is on human motion estimation and controller learning to reproduce natural looking movements on robotic manipulators.



Josip Ćesić is the Head of Autonomous Technologies at Gideon Brothers. Before joining the company he worked at the Laboratory for Autonomous Systems and Mobile Robotics at the University of Zagreb, Croatia. His research interests include autonomous systems, mobile robotics and estimation theory.



Kevin Westermann is a medical devices engineer at Abbott Point of Care. Before joining the company he worked at the University of Waterloo Adaptive Systems Laboratory. His research interests include natural human movement, sports movement analysis and mechanical design.



Ivan Marković is an Assistant Professor at the University of Zagreb Faculty of Electrical Engineering and Computing, Croatia. In 2018 he received the young scientists award "Vera Johanides" from the Croatian Academy of Engineering and in 2014, for outstanding doctoral dissertation, the Silver Plaque "Josip Lončar" faculty award. He was a visiting researcher at INRIA Rennes-Bretagne Atlantique, Rennes, France (Prof. François Chaumette). His research interests include estimation theory with applications to autonomous mobile robotics.



Ivan Petrović is a professor and the head of the Laboratory for Autonomous Systems and Mobile Robotics at the Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia. His research interests include advanced control and estimation techniques and their application in autonomous systems and robotics. He published about 60 journal and 200 conference papers. Among others, he is a full member of the Croatian Academy of Engineering, and the Chair of the IFAC Technical Committee on Robotics.



Dana Kulić leads the Adaptive System Laboratory at the University of Waterloo, Canada, and is a professor at Monash University, Australia. Her research interests include robot learning, human-robot interaction and human motion analysis.