Design and Evaluation of an Invariant Extended Kalman Filter for Trunk Motion Estimation with Sensor Misalignment

Zenan Zhu[1,*], Seyed Mostafa Rezayat Sorkhabadi[2,*], Yan Gu[3,†], Wenlong Zhang[2]

Abstract—Understanding human motion is of critical importance for health monitoring and control of assistive robots, yet many human kinematic variables cannot be directly or accurately measured by wearable sensors. In recent years, invariant extended Kalman filtering (InEKF) has shown a great potential in nonlinear state estimation, but its applications to human poses new challenges, including imperfect placement of wearable sensors and inaccurate measurement models. To address these challenges, this paper proposes an augmented InEKF design which considers the misalignment of the inertial sensor at the trunk as part of the states and preserves the group affine property for the process model. Personalized lower-extremity forward kinematic models are built and employed as the measurement model for the augmented InEKF. Observability analysis for the new InEKF design is presented. The filter is evaluated with three subjects in squatting, rolling-foot walking, and ladder-climbing motions. Experimental results validate the superior performance of the proposed InEKF over the state-of-the-art InEKF. Improved accuracy and faster convergence in estimating the velocity and orientation of human, in all three motions, are achieved despite the significant initial estimation errors and the uncertainties associated with the forward kinematic measurement model.

Index Terms—Extended Kalman filtering, human motion estimation, nonlinear state estimation, forward kinematics, observability analysis

I. INTRODUCTION

Wearable robots have gained growing interests over the past decades as they demonstrated great potentials in facilitating neurorehabiltiation, assisting in daily activities, and reducing work-related injuries [1]–[3]. Wearable robots have been designed with different actuation mechanisms (e.g., cable-driven and pneumatic-driven) and materials (e.g., carbon fibers and fabrics), and they have been applied to various human joints. Since wearable robots physically interact with humans, it is critical to develop control systems that can understand the human's intent and physical states to adaptively exert an appropriate amount of assistance. To this end, various human-centered controller designs have been applied in wearable robots, including adaptive impedance control [4], phase-based control [5], reinforcement learning [6], to name a few.

All the aforementioned robot controllers rely on wearable sensors and estimation algorithms to understand human's intent and physical states. Many existing controllers (e.g., finite-state machines) for wearable robots rely on classification of human motion into finite number of states, such as gait phases [7] and activity types [8]. However, this method will limit the humans to pre-defined activities and it can suffer from misclassification if the user behaves differently from the training data. On the other side of the spectrum, continuous human states, such as positions and orientations, present a promising direction for controller design since they contain more information about the user for fine adjustment of the robot assistance. A challenge for using continuous human states is that many of them cannot be directly measured, or the sensor measurement is too noisy to be used for robot control. This is particularly important for lower-extremity wearable robots, where the center of mass is critical for ensuring postural stability but is not directly measurable [9].

Various approaches have been proposed for the estimation of continuous human movement state. Earlier methods produce accurate, real-time estimation of stance-foot locations during human walking, by fusing the zero toe velocity during stance phase and the reading of inertial measurement units (IMUs) attached to a subject's shoes [10]. To monitor the movement of the body (e.g., pelvis) during walking and stair climbing, the stance leg's kinematics can be used to compute the movement under the assumptions that the stance feet is stationary and that the contact detection is sufficiently accurate [11]. This method is computationally efficient for real-time estimation, and its accuracy has been improved by explicitly handling the joint angle reading noise through Kalman filtering [12]. To further enhance the error convergence, extended Kalman filtering (EKF) has been introduced to fuse the forward kinematics of a subject's segments with the data returned by an IMU attached to the body during bicycle riding [13]. Still, the standard EKF methodology relies on system linearization whose accuracy depends on the estimation error, and thus its performance under large initial errors may not be satisfactory. To this end, the invariant EKF (InEKF) [14] has been created to ensure accurate, efficient estimation under large errors, by exploiting the accurate linearization of systems that meet the group affine condition and invariant observation form.

[1]Z. Zhu is with the College of Engineering, University of Massachusetts Lowell, Lowell, MA 01854, USA zenan_zhu@student.uml.edu.

[3]Y. Gu is with the School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA. This work was partly conducted while Y. Gu was with the University of Massachusetts Lowell. yan.gu.purdue@gmail.com.

[2]M. Rezayat and W. Zhang are with the School of Manufacturing Systems and Networks, Ira A. Fulton Schools of Engineering, Arizona State University Mesa, AZ 85212, USA {sm.rs, wenlong.zhang}@asu.edu.

∗ These two authors have equal contributions.

† Address all correspondence to this author.

Although the existing InEKF methods have achieved rapid convergence and accurate estimates of legged robots under large estimation errors [15]–[17], InEKF design for human movement state estimation poses complex challenges. One challenge is the inaccurate or unknown sensor placement relative to the subject's segments during subject movement. The relative inertial sensor placement is assumed to be accurately known in the existing fusion of inertial odometry and limb kinematics of legged robots [15]–[17]. Yet, this assumption may not be realistic for human movement estimation because it may require frequent calibration during relatively dynamic movement (e.g., squatting, walking, and ladder climbing) due to shift of sensors on the skin or garment. There is a rich body of work in fault detection, but most of them focus on detecting faults in the system rather than obtaining accurate state estimation with such faults [18]. Thus, an appropriate filter design should address the sensor placement imperfection. Its performance also needs to be examined during various types of common daily mobility tasks so that accurate, real-time state estimates could be used to inform the control of wearable devices for different human movement activities.

This paper introduces an InEKF design to address sensor misalignment, which is a very common problem in human and robot locomotion, so that sensor placement offset can be estimated and corrected to ensure accurate state estimation. The intended application is to use lower-extremity joint angles and forward kinematic models to augment the trunk velocity and orientation estimation by correcting misalignment of trunk sensors during the stance phase. A preliminary version of this work was presented in [19]. The present paper reports the following new, substantial contributions. (i) A forward kinematic model is developed in this work and used as the measurement model for the InEKF. The new measurement model is less accurate but much more practical since it relies on joint angles, compared to the 3-D vectors used in [19]. (ii) An observability analysis is presented for the new InEKF design and validated using experimental results with large initial errors. (iii) Two new activities, namely ladder climbing and rolling-foot walking, are evaluated in human testing along with squat motion in [19]. The new results extend the InEKF to more dynamic locomotion tasks with new challenges in the measurement model introduced by contact-foot switching. (iv) The proposed filter is evaluated with three participants in three tasks, and results are compared with the existing InEKF to draw insights on how to extend InEKF to human locomotion estimation.

The paper is structured as follows. Section II introduces the state representation and the process and measurement models of the proposed InEKF. Section III presents the propagation and update steps of the filter along with the observability analysis. Section IV reports and discusses the experiment validation results. Section V provides the conclusion remarks.

II. SYSTEM MODELING

A. Preliminaries

A matrix Lie group G is a subset of $n \times n$ invertible matrices possessing the following properties: $\forall \mathbf{X} \in G, \mathbf{X}^{-1} \in G$;
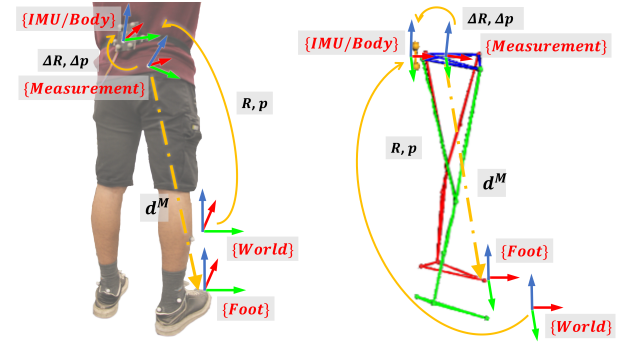

Fig. 1. Coordinate frames and estimation variables used in the proposed filter design. They are illustrated on the real human subject (left) and the subject's lower extremity skeleton in the motion capture software (right). $\mathbf{R}$ and $\mathbf{p}$ represent the orientation and position of the IMU (body) in the world frame, and $\mathbf{d}^M$ represents the vector from the measurement frame to the foot frame, expressed in the measurement frame. $\mathbf{\Delta p}$ and $\mathbf{\Delta R}$ represent the position and orientation offset of the IMU frame with respect to the measurement frame, expressed in the measurement frame.

$\forall \mathbf{X}_1, \mathbf{X}_2 \in G, \mathbf{X}_1 \mathbf{X}_2 \in G$; and $\mathbf{I} \in G$ with $\mathbf{I}$ the identity element. The associated Lie algebra $\mathfrak{g}$ with a dimension of $\dim \mathfrak{g}$ is the tangent space defined at $\mathbf{E}$, and is a set of $n \times n$ square matrices. Any vector $\boldsymbol{\zeta} \in \mathbb{R}^{\dim \mathfrak{g}}$ can be mapped onto the Lie algebra through the linear operator $(.)^\wedge : \mathbb{R}^{\dim \mathfrak{g}} \to \mathfrak{g}$. The exponential map, $\exp : \mathbb{R}^{\dim \mathfrak{g}} \to G$, is defined as: $\exp(\boldsymbol{\zeta}) \triangleq \mathrm{expm}(\boldsymbol{\zeta}^\wedge)$, where expm is the usual matrix exponential. The inverse operator of $(\cdot)^\wedge$ is denoted as $(\cdot)^\vee : \mathfrak{g} \to \mathbb{R}^{\dim \mathfrak{g}}$. The adjoint matrix $\mathbf{Ad}_\mathbf{X}$ at $\mathbf{X}$ for any vector $\boldsymbol{\zeta} \in \mathbb{R}^{\dim \mathfrak{g}}$ is defined as $\mathbf{Ad}_\mathbf{X} \boldsymbol{\zeta} = (\mathbf{X} \boldsymbol{\zeta}^\wedge \mathbf{X}^{-1})^\vee$, which is the linear mapping from the local tangent space (defined at $\mathbf{X}$) to the Lie algebra. A more detailed introduction to matrix Lie groups is given in [20].

B. Sensors considered

The sensors considered in the proposed filter design include: a) an inertial measurement unit (IMU) attached to the subject's body (e.g., lower back near the pelvis), which measures the linear acceleration $\mathbf{a} \in \mathbb{R}^3$ and angular velocity $\boldsymbol{\omega} \in SO(3)$ of the IMU/body with respect to the IMU frame, and b) a motion capture system (e.g., IMU sensors or markers for motion capture cameras) that measures the lower-extremity joint angles of the subject.

The particular issue this study focuses on addressing is the placement inaccuracy of the IMU that measures the subject's body movement. As reviewed in Sec. I, to achieve accurate state estimation during mobility tasks (e.g., walking and ladder climbing) despite sensor noises, an effective technique is to fuse the data returned by IMU and the subject's limb kinematics obtained based on motion capture data. For effective sensor fusion, the IMU should be static relative to a reference segment along the limb kinematics chain, and the relative pose needs to be accurately known [15]. Yet, the exact relative sensor placement is usually unknown without calibration.

To explicitly treat the sensor placement offset, the proposed filter design involves the body/IMU frame, which is fixed to the IMU, and the measurement frame, which is fixed to the reference segment of the kinematic chains sensed by the motion capture system (see Fig. 1). As the filter aims to

estimate subject movement in the environment, the world frame is considered. Also, the foot frame, which is rigidly attached to the stance foot, is used because the foot defines the far end of the kinematics chains of interest.

C. State representation

To estimate body movement in the world, one state variable of interest to this study is the subject's body position, denoted as $^W\mathbf{p}^{WB} \in \mathbb{R}^3$. Here, the superscript on the left-hand side represents the coordinate frame in which the variable is expressed, with W, B, M, and F denoting the world, body/IMU, measurement, and foot frames. The superscript WB on the right-hand side indicates the position vector points from the origin of the world frame to that of the body/IMU frame. The state variables also include the body velocity, $^W\mathbf{v}^{WB} \in \mathbb{R}^3$, which is the time derivative of $^W\mathbf{p}^{WB}$, as well as the body orientation $^W\mathbf{R}^B \in SO(3)$, which is the orientation of the body/IMU frame with respect to the world frame. In addition, to explicitly treat the IMU placement offset (i.e., the relative position and orientation of the IMU/body frame with respect to the measurement frame), they are also chosen as state variables and denoted as $^M\mathbf{p}^{MB} \in \mathbb{R}^3$ and $^M\mathbf{R}^B \in SO(3)$. For notational brevity, the rest of the paper uses $\mathbf{p}$, $\mathbf{v}$, $\mathbf{R}$, $\Delta\mathbf{p}$, and $\Delta\mathbf{R}$ to respectively denote $^W\mathbf{p}^{WB}$, $^W\mathbf{v}^{WB}$, $^W\mathbf{R}^B$, $^M\mathbf{p}^{MB}$, and $^M\mathbf{R}^B$.

We express the state variables on the matrix Lie group G as prescribed by the methodology of InEKF [14]:

$$\mathbf{X} = \begin{bmatrix} \mathbf{R} & \mathbf{v} & \mathbf{p} & \mathbf{0}_{3,3} & \mathbf{0}_{3,1} \\ \mathbf{0}_{1,3} & 1 & 0 & \mathbf{0}_{1,3} & \mathbf{0}_{3,1} \\ \mathbf{0}_{1,3} & 0 & 1 & \mathbf{0}_{1,3} & \mathbf{0}_{3,1} \\ \mathbf{0}_{3,1} & \mathbf{0}_{3,1} & \mathbf{0}_{3,1} & \Delta\mathbf{R} & \Delta\mathbf{p} \\ \mathbf{0}_{1,3} & 0 & 0 & \mathbf{0}_{1,3} & 1 \end{bmatrix} \in G. \tag{1}$$

Here, the matrix Lie group G is a combination of double direct spatial isometries $SE_2(3)$ [14] and a special Euclidean group $SE(3)$, and $\mathbf{0}_{n,m}$ represents an $n \times m$ zero matrix. It can be proved that G is a valid matrix Lie group.

D. Process model

This subsection introduces the proposed process model in Euclidean space as well as on the matrix Lie group G.

1) IMU motion dynamics: Given its accuracy and simplicity [15], the IMU motion dynamics is used to build a process model with the noisy IMU readings $\tilde{\mathbf{a}}$ and $\tilde{\boldsymbol{\omega}}$ serving as its input. Corrupted by white Gaussian zero-mean noise $\mathbf{w}_a \in \mathbb{R}^3$ and $\mathbf{w}_\omega \in SO(3)$, these readings are given by: $\tilde{\mathbf{a}} = \mathbf{a} + \mathbf{w}_a$ and $\tilde{\boldsymbol{\omega}} = \boldsymbol{\omega} + \mathbf{w}_\omega$. Note that for simplicity, the biases in the raw data returned by the accelerometer and gyroscope are not considered here. Then, the IMU motion dynamics can be expressed as: $\frac{d}{dt}\mathbf{R} = \mathbf{R}(\tilde{\boldsymbol{\omega}} - \mathbf{w}_\omega)_\times$, $\frac{d}{dt}\mathbf{p} = \mathbf{v}$, and $\frac{d}{dt}\mathbf{v} = \mathbf{R}(\tilde{\mathbf{a}} - \mathbf{w}_a) + \mathbf{g}$, where $(.)_\times$ is a skew-symmetric matrix and $\mathbf{g}$ is the gravitational acceleration.

2) IMU placement error dynamics: Given that the IMU, if appropriately attached, usually does not shift quickly on the subject, we choose to model the dynamics of the IMU placement offsets $\Delta\mathbf{p}$ and $\Delta\mathbf{R}$ as slowly time-varying, which is given by: $\frac{d}{dt}\Delta\mathbf{p} = \mathbf{w}_{\Delta p}$ and $\frac{d}{dt}\Delta\mathbf{R} = \Delta\mathbf{R}(\mathbf{w}_{\Delta R})_\times$, where $\mathbf{w}_{\Delta p} \in \mathbb{R}^3$ and $\mathbf{w}_{\Delta R} \in SO(3)$ are white Gaussian noise with zero mean.

3) Process model on G: At time t, these process models can be compactly expressed on the matrix Lie group G as:

$$\frac{d}{dt}\mathbf{X}_t = \begin{bmatrix} \mathbf{R}_t(\tilde{\boldsymbol{\omega}}_t)_\times & \mathbf{R}_t\tilde{\mathbf{a}}_t + \mathbf{g} & \mathbf{v}_t & \mathbf{0}_{3,3} & \mathbf{0}_{3,1} \\ \mathbf{0}_{1,3} & 0 & 0 & \mathbf{0}_{1,3} & \mathbf{0}_{3,1} \\ \mathbf{0}_{1,3} & 0 & 0 & \mathbf{0}_{1,3} & \mathbf{0}_{3,1} \\ \mathbf{0}_{3,3} & \mathbf{0}_{3,1} & \mathbf{0}_{3,1} & \mathbf{0}_{3,3} & \mathbf{0}_{3,1} \\ \mathbf{0}_{1,3} & 0 & 0 & \mathbf{0}_{1,3} & 0 \end{bmatrix}$$

$$- \mathbf{X}_t \begin{bmatrix} (\mathbf{w}_{\omega_t})_\times & \mathbf{w}_{a_t} & \mathbf{0}_{3,1} & \mathbf{0}_{3,3} & \mathbf{0}_{3,1} \\ \mathbf{0}_{1,3} & 0 & 0 & \mathbf{0}_{1,3} & \mathbf{0}_{3,1} \\ \mathbf{0}_{1,3} & 0 & 0 & \mathbf{0}_{1,3} & \mathbf{0}_{3,1} \\ \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & (\mathbf{w}_{\Delta R_t})_\times & \mathbf{w}_{\Delta p_t} \\ \mathbf{0}_{1,3} & 0 & 0 & \mathbf{0}_{1,3} & 0 \end{bmatrix} \tag{2}$$

$$\triangleq f_{u_t}(\mathbf{X}_t) - \mathbf{X}_t \mathbf{w}_t^\wedge,$$

where the vector $\mathbf{w}_t$ is defined in the IMU/body frame as $\mathbf{w}_t \triangleq \mathrm{vec}(\mathbf{w}_{\omega_t}, \mathbf{w}_{a_t}, \mathbf{0}_{3,1}, \mathbf{w}_{\Delta R_t}, \mathbf{w}_{\Delta p_t})$, and the vector $\mathbf{u}_t$ is the input to the process model defined as $\mathbf{u}_t \triangleq \mathrm{vec}(\tilde{\boldsymbol{\omega}}_t, \tilde{\mathbf{a}}_t)$.

Proposition 1 (Group affine system) The deterministic portion of the process model in (2) is group affine; that is, the deterministic dynamics $f_{u_t}(.)$ meets the following group affine condition for right-invariant cases [14]:

$$f_{u_t}(\mathbf{X}_1\mathbf{X}_2) = f_{u_t}(\mathbf{X}_1)\mathbf{X}_2 + \mathbf{X}_1 f_{u_t}(\mathbf{X}_2) - \mathbf{X}_1 f_{u_t}(\mathbf{I})\mathbf{X}_2. \tag{3}$$

Proof. By the definition of f_{u_t} in (2), we obtain

$$f_{u_t}(\mathbf{X}_1) = \begin{bmatrix} \mathbf{R}_1(\tilde{\boldsymbol{\omega}})_\times & \mathbf{R}_1\tilde{\mathbf{a}} + \mathbf{g} & \mathbf{v}_1 & \mathbf{0}_{3,4} \\ \mathbf{0}_{6,3} & \mathbf{0}_{6,1} & \mathbf{0}_{6,1} & \mathbf{0}_{6,4} \end{bmatrix},$$

$$f_{u_t}(\mathbf{X}_1\mathbf{X}_2) = \begin{bmatrix} \mathbf{R}_1\mathbf{R}_2(\tilde{\boldsymbol{\omega}})_\times & \mathbf{R}_1\mathbf{R}_2\tilde{\mathbf{a}} + \mathbf{g} & \mathbf{R}_1\mathbf{v}_2 + \mathbf{v}_1 & \mathbf{0}_{3,4} \\ \mathbf{0}_{6,3} & \mathbf{0}_{6,1} & \mathbf{0}_{6,1} & \mathbf{0}_{6,4} \end{bmatrix}, \text{ and}$$

$$f_{u_t}(\mathbf{I}) = \begin{bmatrix} (\tilde{\boldsymbol{\omega}})_\times & \tilde{\mathbf{a}} + \mathbf{g} & \mathbf{0}_{3,1} & \mathbf{0}_{3,4} \\ \mathbf{0}_{6,3} & \mathbf{0}_{6,1} & \mathbf{0}_{6,1} & \mathbf{0}_{6,4} \end{bmatrix}$$

for any $\mathbf{X}_1, \mathbf{X}_2 \in G$. Here $\mathbf{R}_i$ and $\mathbf{v}_i$ ($i = 1, 2$) denote elements of the state $\mathbf{X}_i$. It can be seen that these matrices satisfy the condition in (3).

E. Measurement model

1) Forward kinematics based measurement: During mobility tasks (e.g., squatting, walking, and stair climbing), when the stance foot has a static, secured contact with the ground, the pose of the measurement frame can be obtained through the leg kinematic chain that connects the foot and measurement frames. Note that the joint angles along the kinematic chain are measured by the motion capture system. The kinematic chain is built based on the Vicon lower-body Plug-in-Gait model [21], in which 3D joint angles (hip, knee, and ankle), along with subject lower-body segment length measurements, were used to obtain the desired 3D vector.

Let $\boldsymbol{\alpha}_t \in \mathbb{R}^k$ be the joint angles of the stance leg with k the number of joint angles. Then the measured joint angle $\tilde{\boldsymbol{\alpha}}_t$ can be expressed as $\tilde{\boldsymbol{\alpha}}_t = \boldsymbol{\alpha}_t + \mathbf{w}_{\alpha_t}$ with $\mathbf{w}_{\alpha_t} \in \mathbb{R}^k$ the zero-mean white Gaussian noise.

Let $^M\mathbf{d}_t^{MF}$ denote the 3-D position vector pointing from the measurement to the foot frame expressed in the measurement frame. For brevity, we denote it as $\mathbf{d}_t^M$. Let the function h_F be the forward kinematics representing $\mathbf{d}_t^M$; that is, $\mathbf{d}_t^M = h_F(\boldsymbol{\alpha}_t)$.

Given that $\mathbf{d}_t^M = (\Delta\mathbf{R}_t)\mathbf{R}_t^T(\mathbf{d}_t - \mathbf{p}_t) - \Delta\mathbf{p}_t$, we have $\mathbf{d}_t - \mathbf{p}_t = \mathbf{R}_t\Delta\mathbf{R}_t^T(\Delta\mathbf{p}_t + h_F(\boldsymbol{\alpha}_t))$. Taking the first derivative with respect to time t on both sides of this equation, we obtain:

$$\frac{d}{dt}(\mathbf{d}_t - \mathbf{p}_t) = \mathbf{R}_t\Delta\mathbf{R}_t^T(\mathbf{w}_{\Delta p_t} + J(\boldsymbol{\alpha}_t)(\dot{\boldsymbol{\alpha}}_t + \mathbf{w}_{\dot{\alpha}_t})) + \left(\mathbf{R}_t(\boldsymbol{\omega}_t + \mathbf{w}_{\omega_t})_{\times}\Delta\mathbf{R}_t^T + \mathbf{R}_t(\Delta\mathbf{R}_t(\mathbf{w}_{\Delta\mathbf{R}_t})_{\times})^T\right)(h_F(\boldsymbol{\alpha}_t) + \Delta\mathbf{p}_t), \tag{4}$$

where $J(\boldsymbol{\alpha}) \triangleq \frac{\partial h_F(\boldsymbol{\alpha})}{\partial\boldsymbol{\alpha}}$ is the forward kinematic Jacobian and $\mathbf{w}_{\dot{\alpha}_t}$ is the joint velocity measurement noise.

Note that $\dot{\mathbf{d}}_t = \mathbf{0}$ due to the stationary contact point. Also, $\dot{\mathbf{p}}_t = \mathbf{v}_t$ holds. Thus, the measurement model in (4) can be compactly rewritten as:

$$\mathbf{y} = h(\mathbf{X}_t) + \mathbf{n}_t, \tag{5}$$

where $\mathbf{y} = -J(\tilde{\boldsymbol{\alpha}}_t)\dot{\tilde{\boldsymbol{\alpha}}}_t$, $h(\mathbf{X}_t) = (\Delta\mathbf{R}_t)\mathbf{R}_t^T\mathbf{v}_t - (\Delta\mathbf{p}_t)_{\times}\Delta\mathbf{R}_t\tilde{\boldsymbol{\omega}}_t - (h_F(\tilde{\boldsymbol{\alpha}}_t))_{\times}\Delta\mathbf{R}_t\tilde{\boldsymbol{\omega}}_t$, and $\mathbf{n}_t = (\Delta\mathbf{p}_t + h_F(\tilde{\boldsymbol{\alpha}}_t))_{\times}\Delta\mathbf{R}_t(\mathbf{w}_{\Delta\mathbf{R}_t} + \mathbf{w}_{\omega_t}) + \mathbf{w}_{\Delta p_t} - J(\tilde{\boldsymbol{\alpha}}_t)\mathbf{w}_{\dot{\alpha}_t}$. For simplicity, $\mathbf{n}_t$ is treated as white, Gaussian, zero-mean noise in this study.

2) 3-D vector based measurement model: To compare the filtering performance under different kinematics measurements formed based on data returned by motion capture systems, we introduce a simplified measurement obtained based on the 3-D position vector between the measurement and foot frame. Both forward kinematics based and 3-D position vector based measurements return the relative position and velocity between the measurement and foot frames. Note that the 3-D position vector based measurement will have higher accuracy than the forward kinematics measurements.

Let $\mathbf{v}_t^M$ denotes the 3-D velocity vector pointing from the measurement to the foot frame expressed in the measurement frame, which is the time derivative of $\mathbf{d}_t^M$. Then the measured 3-D velocity vector $\tilde{\mathbf{v}}_t^M$ is expressed as $\tilde{\mathbf{v}}_t^M = \mathbf{v}^M + \mathbf{w}_{v_t^M}$, with $\mathbf{w}_{v_t^M}$ the noise associated with the 3-D velocity vector measurement.

Given that $\mathbf{d}_t^M = (\Delta\mathbf{R}_t)\mathbf{R}_t^T(\mathbf{d}_t - \mathbf{p}_t) - \Delta\mathbf{p}_t$, we have $\mathbf{d}_t - \mathbf{p}_t = \mathbf{R}_t\Delta\mathbf{R}_t^T(\Delta\mathbf{p}_t + \mathbf{d}_t^M)$. Taking the time derivative on both sides of this equation, we obtain:

$$\frac{d}{dt}(\mathbf{d}_t - \mathbf{p}_t) = \mathbf{R}_t\Delta\mathbf{R}_t^T(\mathbf{w}_{\Delta p_t} + (\mathbf{v}^M + \mathbf{w}_{v^M})) + \left(\mathbf{R}_t(\boldsymbol{\omega}_t)_{\times}\Delta\mathbf{R}_t^T + \mathbf{R}_t(\Delta\mathbf{R}_t(\mathbf{w}_{\Delta\mathbf{R}_t})_{\times})^T\right)(\mathbf{d}_t^M + \Delta\mathbf{p}_t). \tag{6}$$

Assume the contact foot velocity is zero due to the stationary contact point. Then the measurement model (6) becomes:

$$\mathbf{y} = h(\mathbf{X}_t) + \mathbf{n}_t \tag{7}$$

where $\mathbf{y} = -\tilde{\mathbf{v}}^M$, $h(\mathbf{X}_t) = \Delta\mathbf{R}_t\mathbf{R}_t^T\mathbf{v}_t - (\Delta\mathbf{p}_t)_{\times}\Delta\mathbf{R}_t\boldsymbol{\omega}_t - (\mathbf{d}_t^M)_{\times}\Delta\mathbf{R}_t\boldsymbol{\omega}_t + \mathbf{n}_t$, and the vector $\mathbf{n}_t$ is the lumped measurement noise term.

Remark 1: The two measurement models in (5) and (7) do not satisfy the right-invariant observation form (i.e., $\mathbf{y} = \mathbf{X}^{-1}\mathbf{b}$ for some known vector $\mathbf{b}$) as defined in the theory of InEKF [14]. This is because with the state defined in (1) and our measurements in (5) and (7), a vector $\mathbf{b}$ that is known and satisfies $\mathbf{y} = \mathbf{X}^{-1}\mathbf{b}$ does not exist. Then, by the theory of invariant filtering (Proposition 2 in [14]), the error dynamics during the measurement update is not independent of state trajectories. The effects of this property on filter performance are discussed in Sec. IV-D.

III. INEKF DESIGN

A. State propagation

The design of the proposed InEKF relies on the right-invariant error $\boldsymbol{\eta}_t$ between the true and estimated values:

$$\boldsymbol{\eta}_t = \bar{\mathbf{X}}_t\mathbf{X}_t^{-1} \in G, \tag{8}$$

where $\overline{(.)}$ denotes the estimated value of the variable $(.)$. Based on this definition and the process model (2), one can obtain:

$$\frac{d}{dt}\boldsymbol{\eta}_t = f_{u_t}(\boldsymbol{\eta}_t) - \boldsymbol{\eta}_t f_{u_t}(\mathbf{I}_d) + Ad_{\bar{\mathbf{X}}_t}\mathbf{w}_t^{\wedge} \triangleq g_{u_t}(\boldsymbol{\eta}_t) + \bar{\mathbf{w}}_t^{\wedge}. \tag{9}$$

Remark 2: The deterministic portion of the error dynamics in (9) is independent of the state trajectories because the process model in (2) is group affine, as predicted by the theory of InEKF [14]. This property is drastically different from the standard EKF whose error dynamics depends on the state trajectories. Furthermore, since the process model is group affine, by the theory of invariant filtering (Theorem 2 in [14]), the corresponding dynamics of the log of the invariant error $\boldsymbol{\eta}_t$ is exactly linear in the deterministic case, whose expression is derived next.

Let $\boldsymbol{\zeta}_t$ be the logarithmic error defined through $\boldsymbol{\eta}_t = \exp(\boldsymbol{\zeta}_t)$, and denote $\boldsymbol{\zeta}_t$ as $\boldsymbol{\zeta}_t = \text{vec}(\boldsymbol{\zeta}_{R_t}, \boldsymbol{\zeta}_{v_t}, \boldsymbol{\zeta}_{p_t}, \boldsymbol{\zeta}_{\Delta R_t}, \boldsymbol{\zeta}_{\Delta p_t}) \in \mathbb{R}^{\dim\mathfrak{g}}$. In order to linearize this error dynamics, we use the first-order approximation $\boldsymbol{\eta}_t = \exp(\boldsymbol{\zeta}_t) \approx \mathbf{I} + \boldsymbol{\zeta}_t^{\wedge}$, which allows us to obtain the Jacobian $\mathbf{A}_t$ of deterministic dynamics as [14]:

$$g_{u_t}(\exp(\boldsymbol{\zeta}_t)) = (\mathbf{A}_t\boldsymbol{\zeta}_t)^{\wedge} + \text{h.o.t.}(||\boldsymbol{\zeta}_t||) \approx (\mathbf{A}_t\boldsymbol{\zeta}_t)^{\wedge}, \tag{10}$$

and the Jacobian $\mathbf{A}_t$ will also define the linear dynamics of the log of the right invariant error as follows:

$$\frac{d}{dt}\boldsymbol{\zeta}_t = \mathbf{A}_t\boldsymbol{\zeta}_t + \bar{\mathbf{w}}_t. \tag{11}$$

To compute $\mathbf{A}_t$, we plug in the approximated right-invariant error into (10) as:

$$g_{u_t}(\exp(\boldsymbol{\zeta}_t)) \approx g_{u_t}(\mathbf{I}_d + \boldsymbol{\zeta}_t^{\wedge})$$
$$= \begin{bmatrix} \mathbf{0}_{3,3} & (\mathbf{g})_{\times}\boldsymbol{\zeta}_{R_t} & \boldsymbol{\zeta}_{v_t} & \mathbf{0}_{3,4} \\ \mathbf{0}_{1,3} & 0 & 0 & \mathbf{0}_{1,4} \\ \mathbf{0}_{1,3} & 0 & 0 & \mathbf{0}_{1,4} \\ \mathbf{0}_{3,3} & \mathbf{0}_{3,1} & \mathbf{0}_{3,1} & \mathbf{0}_{3,4} \\ \mathbf{0}_{1,3} & 0 & 0 & \mathbf{0}_{1,4} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{3,1} \\ (\mathbf{g})_{\times} \\ \boldsymbol{\zeta}_{v_t} \\ \mathbf{0}_{3,1} \\ \mathbf{0}_{3,1} \end{bmatrix}^{\wedge}, \tag{12}$$

which yields:

$$\mathbf{A}_t = \begin{bmatrix} \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} \\ (\mathbf{g})_{\times} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} \\ \mathbf{0}_{3,3} & \mathbf{I}_3 & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} \\ \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} \\ \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} \end{bmatrix}.$$

As a result, the state estimates can be propagated using (2) and covariance matrix can be updated using the Riccati equation associated with (11):

$$\frac{d}{dt}\bar{\mathbf{X}}_t = f_{u_t}(\bar{\mathbf{X}}_t), \quad \frac{d}{dt}\mathbf{P}_t = \mathbf{A}_t\mathbf{P}_t + \mathbf{P}_t\mathbf{A}_t^T + \bar{\mathbf{Q}}_t, \tag{13}$$

where $\bar{\mathbf{Q}}_t$ is the process noise covariance defined as $\bar{\mathbf{Q}}_t \triangleq \text{Cov}(\bar{\mathbf{w}}_t) = Ad_{\bar{\mathbf{X}}_t}\text{Cov}(\mathbf{w}_t)Ad_{\bar{\mathbf{X}}_t}$.

B. Measurement update

The nonlinear measurement model (5) does not follow the right-invariant form, so we use a first-order approximation to find the innovation as:

$$\mathbf{H}_t\boldsymbol{\zeta}_t + \text{h.o.t}(\boldsymbol{\zeta}_t) \triangleq h(\bar{\mathbf{X}}_t) - h(\mathbf{X}_t). \tag{14}$$

Since $\boldsymbol{\eta}_t \approx \bar{\mathbf{X}}_t\mathbf{X}_t^{-1} \approx \mathbf{I} + \boldsymbol{\zeta}_t^\wedge$, we can derive the following relationships between the true and estimated states with $\boldsymbol{\zeta}_t$:

$$\Rightarrow \mathbf{R}_t^T \approx \bar{\mathbf{R}}_t^T(\mathbf{I} + (\boldsymbol{\zeta}_{R_t})_\times), \ \Delta\mathbf{R}_t^T \approx \overline{\Delta\mathbf{R}}_t^T(\mathbf{I} + (\boldsymbol{\zeta}_{\Delta R_t})_\times)$$
$$\mathbf{v}_t \approx (\mathbf{I} - (\boldsymbol{\zeta}_{R_t})_\times)(\bar{\mathbf{v}}_t - \boldsymbol{\zeta}_{v_t}), \ \Delta\mathbf{p}_t \approx (\mathbf{I} - (\boldsymbol{\zeta}_{\Delta R_t})_\times)(\overline{\Delta\mathbf{p}}_t - \boldsymbol{\zeta}_{\Delta p_t}).$$

Now $\mathbf{H}_t$ can be computed by differentiating (14) after dropping the nonlinear terms:

$$\mathbf{H}_t = [\mathbf{0}_{3,3}, \ \overline{\Delta\mathbf{R}}_t\bar{\mathbf{R}}_t^T, \ \mathbf{0}_{3,3}, \ \mathbf{h}_4, \ (\overline{\Delta\mathbf{R}}_t\boldsymbol{\omega}_t)_\times],$$
$$\mathbf{h}_4 = -(\overline{\Delta\mathbf{R}}_t\bar{\mathbf{R}}_t^T\bar{\mathbf{v}}_t)_\times - (\overline{\Delta\mathbf{R}}_t\boldsymbol{\omega}_t)_\times(\overline{\Delta\mathbf{p}}_t)_\times \tag{15}$$
$$+ (\overline{\Delta\mathbf{p}}_t)_\times(\overline{\Delta\mathbf{R}}_t\boldsymbol{\omega}_t)_\times + (h_F(\boldsymbol{\alpha}_t))_\times(\overline{\Delta\mathbf{R}}_t\boldsymbol{\omega}_t)_\times.$$

For 3-D position based measurement model, the element $\mathbf{h}_4$ of the matrix $\mathbf{H}_t$ is slightly different from (15). The $h_F(\boldsymbol{\alpha}_t))$ needs to be replaced as $\mathbf{d}_t^M$:

$$\mathbf{H}_t = [\mathbf{0}_{3,3}, \ \overline{\Delta\mathbf{R}}_t\bar{\mathbf{R}}_t^T, \ \mathbf{0}_{3,3}, \ \mathbf{h}_4, \ (\overline{\Delta\mathbf{R}}_t\boldsymbol{\omega}_t)_\times],$$
$$\mathbf{h}_4 = -(\overline{\Delta\mathbf{R}}_t\bar{\mathbf{R}}_t^T\bar{\mathbf{v}}_t)_\times - (\overline{\Delta\mathbf{R}}_t\boldsymbol{\omega}_t)_\times(\overline{\Delta\mathbf{p}}_t)_\times \tag{16}$$
$$+ (\overline{\Delta\mathbf{p}}_t)_\times(\overline{\Delta\mathbf{R}}_t\boldsymbol{\omega}_t)_\times + (\mathbf{d}_t^M)_\times(\overline{\Delta\mathbf{R}}_t\boldsymbol{\omega}_t)_\times.$$

Now we can express the update equation for our InEKF based on the InEKF methodology [14]:

$$\bar{\mathbf{X}}_t^+ = \exp(\mathbf{K}_t(\mathbf{y}_t - h(\bar{\mathbf{X}}_t)))\bar{\mathbf{X}}_t,$$
$$\mathbf{P}_t^+ = (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_t^-(\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)^T + \mathbf{K}_t\mathbf{N}_t\mathbf{K}_t^T, \tag{17}$$

where $\bar{\mathbf{X}}_t^+$ and $\mathbf{P}_t^+$ are the updated values. Here, the Kalman gain $\mathbf{K}_t$ and measurement noise covariance $\mathbf{N}_t$ are defined as: $\mathbf{K}_t = \mathbf{P}_t\mathbf{H}_t^T\mathbf{S}_t^{-1}$, $\mathbf{S}_t = \mathbf{H}_t\mathbf{P}_t^-\mathbf{H}_t^T + \mathbf{N}_t$, and $\mathbf{N}_t = \bar{\mathbf{R}}_t\overline{\Delta\mathbf{R}}_t^T\text{Cov}(\mathbf{n}_t)\overline{\Delta\mathbf{R}}_t\bar{\mathbf{R}}_t^T$.

Algorithm 1 Proposed InEKF design

Initialize $\bar{\mathbf{X}} \in G$ and $\mathbf{P} = \mathbf{P}^T > 0$
for iteration$=1, 2, \ldots$ **do**
    if foot contact is detected **then**
        Propagation step:
        $\frac{d}{dt}\bar{\mathbf{X}}_t = f_{u_t}(\bar{\mathbf{X}}_t)$, $\frac{d}{dt}\mathbf{P}_t = \mathbf{A}_t\mathbf{P}_t + \mathbf{P}_t\mathbf{A}_t^T + \bar{\mathbf{Q}}_t$
        Measurement update step:
        $\mathbf{K}_t = \mathbf{P}_t\mathbf{H}_t^T\mathbf{S}_t^{-1}$,
        $\mathbf{S}_t = \mathbf{H}_t\mathbf{P}_t^-\mathbf{H}_t^T + \mathbf{N}_t$,
        $\mathbf{N}_t = \bar{\mathbf{R}}_t\overline{\Delta\mathbf{R}}_t^T\text{Cov}(\mathbf{n}_t)\overline{\Delta\mathbf{R}}_t\bar{\mathbf{R}}_t^T$
        if using forward kinematics based measurement **then**
            $\mathbf{y} = -J(\tilde{\boldsymbol{\alpha}}_t)\dot{\tilde{\boldsymbol{\alpha}}}_t$
        else if using 3-D vector based measurement **then**
            $\mathbf{y} = -\tilde{\mathbf{v}}_t^M$
        end if
        $\bar{\mathbf{X}}_t^+ = \exp(\mathbf{K}_t(\mathbf{y}_t - h(\bar{\mathbf{X}}_t^-)))\bar{\mathbf{X}}_t^-$
        $\mathbf{P}_t^+ = (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_t^-(\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)^T + \mathbf{K}_t\mathbf{N}_t\mathbf{K}_t^T$
    end if
end for

C. Observability analysis

With new states and measurement model introduced, we need to analyze the observability of the whole system. Here we only include the linear observability analysis, and leave the nonlinear observability analysis for future work. Similar to [22], we analyze the observability around the operating point (i.e., the latest estimate that the system is linearized about).

The discrete filter is expressed as: $\mathbf{X}_{k+1} = \boldsymbol{\Phi}_k\mathbf{X}_k$ and $\mathbf{y}_k = \mathbf{H}_k\mathbf{X}_k$, where $\Delta t = t_{k+1} - t_k$ is the duration of the propagation step with t_k the timing of the kth update. $\boldsymbol{\Phi}_k$ is the discrete-time state transition matrix, which can be computed as:

$$\boldsymbol{\Phi}_k = \exp_m(\mathbf{A}_k\Delta t) = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} \\ (\mathbf{g})_\times\Delta t & \mathbf{I}_3 & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} \\ \frac{1}{2}(\mathbf{g})_\times\Delta t^2 & \mathbf{I}_3\Delta t & \mathbf{I}_3 & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} \\ \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{I}_3 & \mathbf{0}_{3,3} \\ \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} & \mathbf{I}_3 \end{bmatrix}. \tag{18}$$

Then the observability matrix $\mathbf{O}$ can be computed as:

$$\mathbf{O} = \begin{bmatrix} \mathbf{H}_k^- \\ \mathbf{H}_{k+1}^-\boldsymbol{\Phi}_k^+ \\ \mathbf{H}_{k+2}^-\boldsymbol{\Phi}_{k+1}^+\boldsymbol{\Phi}_k^+ \\ \vdots \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{3,3} & \Delta\mathbf{R}_k^-\mathbf{R}_k^{-T} & \mathbf{0}_{3,3} & o_{0,4} & (\Delta\mathbf{R}_k^-\boldsymbol{\omega}_k)_\times \\ o_{1,1} & o_{1,2} & \mathbf{0}_{3,3} & o_{1,4} & o_{1,5} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ o_{4,1} & o_{4,2} & \mathbf{0}_{3,3} & o_{4,4} & o_{4,5} \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix},$$

where $(.)_k^+$ denotes the updated estimated state at time t_k, $(.)_k^-$ is the estimated state at time t_k after the propagation step. Here the terms $o_{i,1}$, $o_{i,2}$, $o_{i,4}$, and $o_{i,5}$ ($i \in \mathbb{N}^+$) are defined as:

$$o_{i,1} = i(\Delta\mathbf{R}_{k+i}^-\mathbf{R}_{k+1}^{-T}(\mathbf{g})_\times\Delta t); \quad o_{i,2} = \Delta\mathbf{R}_{k+i}^-\mathbf{R}_{k+1}^{-T};$$
$$o_{i,4} = -(\Delta\mathbf{R}_{k+i}^-\mathbf{R}_{k+1}^{-T}\mathbf{v}_{k+i}^-)_\times - (\Delta\mathbf{R}_{k+i}^-\boldsymbol{\omega}_{k+i})_\times(\Delta\mathbf{p}_{k+i}^-)_\times$$
$$\quad + (\Delta\mathbf{p}_{k+i}^-)_\times(\Delta\mathbf{R}_{k+i}^-\boldsymbol{\omega}_{k+i})_\times + (FK(\boldsymbol{\alpha}_{k+i}))_\times(\Delta\mathbf{R}_{k+i}^-\boldsymbol{\omega}_{k+i})_\times;$$
$$o_{i,5} = (\Delta\mathbf{R}_{k+i}^-\boldsymbol{\omega}_{k+1})_\times.$$

To analyze the observability for each variable of interest, we need to see if the corresponding column vectors in $\mathbf{O}$ are linearly independent. From the observability matrix $\mathbf{O}$, it can be seen that the position of the IMU in the world frame is completely non-observable. The yaw angle of the IMU is also non-observable as the third column of the matrix $\mathbf{g}_\times$ is always zero. These results are similar to what was reported in [15]. The observability of $\Delta\mathbf{p}$ and $\Delta\mathbf{R}$ depends on the rotational movement. For example, if the human is completely stationary ($\boldsymbol{\omega}, \mathbf{v} = \mathbf{0}$), both $\Delta\mathbf{p}$ and $\Delta\mathbf{R}$ are completely unobservable, as all corresponding columns in $\mathbf{O}$ are zero. If $\boldsymbol{\omega} = \mathbf{0}, \mathbf{v} \neq \mathbf{0}$, then $\Delta\mathbf{p}$ is completely non-observable, while it can be shown at least one of the components of $\Delta\mathbf{R}$ is observable. In general, if the direction of $\boldsymbol{\omega}$ is constant, then $\Delta\mathbf{p}$ will be non-observable in at least one direction.

IV. EXPERIMENTAL RESULTS

A. Experimental setup and protocol

Experiments were performed in a motion capture laboratory at Arizona State University (ASU) with 3 participants (two males and one female, 27 ± 3 years old, $174 \pm 17cm$, $77 \pm$
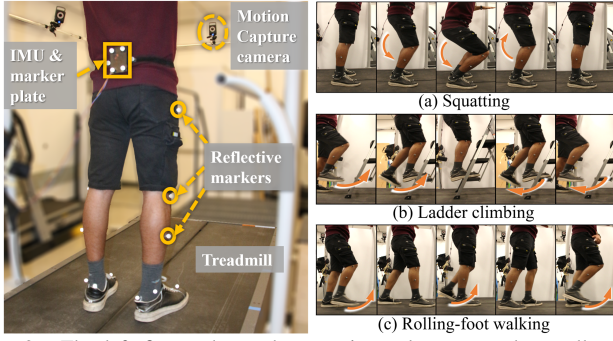
Fig. 2. The left figure shows the experimental setup used to collect the ground truth of the estimated state variables and the sensor data needed by the proposed filter. The right figure shows time-lapse figures of the three motion types. The arrows in the subplots indicate the direction of the pelvis movement (in subplot (a)) and swing foot movement (subplots (b) and (c)).

TABLE I
NOISE CHARACTERISTICS

Measurement type	Noise SD (proposed InEKF)	Noise SD (existing InEKF)
Linear acceleration	0.2 m/s^2	0.2 m/s^2
Angular velocity	0.05 rad/s	0.05 rad/s
Kinematics measurement	0.5 m	0.1 m
Placement offset ($\Delta\mathbf{p}, \Delta\mathbf{R}$)	(0.05 m, 0.05 rad)	NA
Contact velocity	NA	0.05 m/s

$17kg$). The study was approved by institutional review board at ASU (STUDY00011266) and University of Massachusetts Lowell (20-057-YAN-XPD).

Sensor setup. Twelve IR motion capture cameras (Vicon, Oxford, UK) and sixteen reflective markers were used to build a lower-limb model for each subject via Vicon Nexus 2.8. Using this model, 3-D joint angles (hip, knee, and ankle) and pelvic position and orientation were estimated. Four extra markers were attached to a plate that rigidly houses the IMU, which is used to build a rigid body model in Nexus to acquire the ground-truth pose of the IMU/body frame (as shown in the left figure of Fig. 2). The IMU (BNO085, New York, NY) was placed on the back of the subject close to the pelvis (the placement of the IMU with regard to the pelvic center is shown in Fig. 1). The accelerometer and gyroscope data were recorded using a data acquisition board (Arduino UNO, Boston, MA). The data were later synchronised with the motion captured data using a trigger signal from the Vicon system. The experiments were done on an instrumented dual-belt treadmill equipped with force plates (Bertec Corp., Columbus, OH) that record ground contact forces.

Movement types. Each participant was asked to perform three types of motion: squatting, ladder climbing, and rolling-foot walking (see the right figure of Fig. 2). Two trials were performed for each motion type, each for 1.5 minute.

B. Data Processing

Filters compared. The proposed filter is compared with a state-of-the-art InEKF [15]. The existing InEKF was originally designed for a Cassie series bipedal robot. In the existing filter, the IMU and measurement frames are well aligned. The state variables of the existing filter are the IMU orientation, velocity, and position and the contact foot position, all expressed in

the world frame. The kinematic measurements of the existing filter are the contact foot positions with respect to the measurement frame expressed in the measurement frame. Unlike the proposed filter, the measurement model of the existing filter has an invariant observation form, which, in combination with the exponential form of the measurement update, renders the deterministic error update equation to be independent of state trajectories [14]. Also, the kinematic measurements of the existing filter have smaller noises since the Cassie series bipedal robot uses highly accurate leg encoders to formulate the forward kinematics chain. With a human subject, however, the IMU and measurement frames are not aligned, and the kinematics measurements have relatively large noises.

Covariance settings. The noise characteristics for both filters are shown in Table I. The noise standard deviations of the linear acceleration and angular velocity are obtained from the IMU specifications provided by the manufacturer. To reach the better performance of both filters, these two noise standard deviations are slightly tuned around the nominal values. Note that covariance tuning is also reported in other InEKF designs [23]. It should also be noted that since the two filters use different measurement models ($\mathbf{d}^M$ is used as the measurement in [15]), different noise covariance values are used for the two filters. Moreover, the proposed filter considers the noise of the IMU angular velocity in the kinematics measurement noise term, while the same covariances are used for the common parameters of the two filters (linear acceleration and angular velocity). During the tuning process it was observed that the performance of the proposed filter does not vary significantly within a relatively wide range of parameters (e.g., the covariances matrices). Yet, the estimation performance degrades when the covariances are far from the optimal values. The placement offset noises are only used in the proposed filter. Given that the placement offsets are relatively constant, the noise standard deviations of the placement offsets are set as small values. The contact velocity noise is only accounted for in the existing filter, which is induced by contact foot slippage.

Initial estimation errors. To demonstrate the accuracy and convergence rate of both filters under large estimation errors, a relatively wide range of initial estimation errors of the IMU/body velocity and orientation are used. They are respectively chosen to be uniformly within $[-1, 1]$ m/s and $[-20, 20]$ degrees across 50 trials.

Filter performance indicators. To evaluate the filter performance, we choose to use three common indicators [14], [15], [17]: (a) computation time (for assessing the filter's capability in real-time implementation); (b) convergence rate (for evaluating how rapidly the estimation error reaches the steady state); and (c) estimation accuracy (for testing the accuracy during transient and steady-state periods).

C. Results

To illustrate the accuracy of the forward kinematics based measurement (5), Fig. 3 compares it with its reference obtained from the motion capture system. As noted in Sec. III-B, the measurement associated with a given leg is fed into the update step of the filter only when the leg is in contact with the ground (highlighted by yellow shaded areas in Fig. 3) .
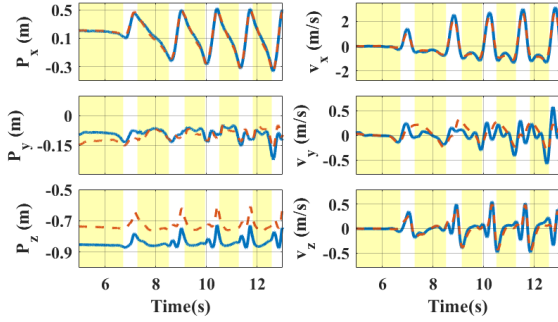
Fig. 3. Relative position (left) and velocity (right) of the right toe in the measurement/pelvis frame during one of the rolling-foot walking trials, obtained by using joint angle readings and forward kinematics (blue, solid line) and by using the marker positions of the toe and measurement plate returned by motion capture system (red, dashed line). The yellow shaded areas indicate the periods during which the right foot contacts the ground.

TABLE II
AVERAGE RMSE VALUES FOR THE THREE PARTICIPANTS AND MOTIONS. V AND O REFER TO THE BODY/IMU VELOCITY AND ORIENTATION.

Time period	Motion type	Variable	FK		3-D vector	
			Proposed InEKF	Existing InEKF	Proposed InEKF	Existing InEKF
Initial	Stand	V (m/s)	0.062	0.226	0.063	0.210
		O (deg)	4.484	6.616	4.737	6.643
Steady state	Squat	V (m/s)	0.067	0.147	0.041	0.152
		O (deg)	4.745	4.909	4.832	4.727
	Roll.-foot	V (m/s)	0.237	0.467	0.103	0.442
		O (deg)	2.909	2.984	2.685	2.838
	Ladder climb.	V (m/s)	0.072	0.135	0.047	0.14
		O (deg)	3.13	3.11	3.12	3.199

Table II displays the average root mean square error (RMSE) values of different motions and different kinematics measurements for all three subjects. During the initial period, subjects are standing still, and the estimation algorithms just start. During the steady-state motion period, subjects are doing continuous movements with different motion types. The average RMSE values of the variable "V" are obtained from the estimated IMU velocity of all three axes for all subjects with the same motion type, while the average RMSE values of the variable "O" are obtained from the estimated IMU orientation of roll and pitch angles for all subjects with the same motion type. The "FK" portion shows the RMSE values of the estimation results with the forward kinematics measurements in (5), and the "3-D vector" part indicates the RMSE values of the estimation results under the 3-D vector kinematics measurements in (7).

Figure 4 shows the comparison of the IMU velocity and orientation estimation results of one subject with forward kinematics measurement under different motions and filters but the same sensor data set. The grey shaded and white backgrounds indicate the initial period and steady-state motion period, respectively.

Computational time. MATLAB R2020b was used to process the experimental data sets with both filters. The average computational time for one filter loop of the proposed filter is about 0.6 ms, and the average computational time for one filter loop of the existing filter is about 3 ms. Both are sufficiently fast for typical human movement monitoring.

Convergence rate. By investigating the initial period of the estimation results figure (Fig. 4) and the RMSEs table (Table II), it is obvious that the proposed filter converges faster than the existing filter, driving the estimation error close to the ground truth within 0.6 s.

Estimation accuracy. The results during the steady-state periods in both Fig. 4 and Table II indicate that under both filters the estimated roll and pitch angles of the IMU converge to a small neighborhood around their ground truth. Yet, the yaw angle of the IMU is not observable under both filters. Also, the overall accuracy of the IMU velocity estimation under the proposed filter is better than the existing filter.

D. Discussion

This subsection discusses the performance of the proposed InEKF with respect to the existing InEKF.

Forward kinematics vs. 3-D vector measurement. Incorporating forward kinematics measurement (i.e., (5)) introduces considerable uncertainties to the filtering system for human motion estimation (as depicted in Fig. 3). The uncertainties in the forward kinematics measurement could be induced by various sources, such as imperfect marker placement, shift of markers on skin or garment, inaccurate parameters (e.g., body segment lengths) and structure of human kinematics chain. These factors lead to the less accurate estimation of both body velocity and orientation compared to directly using the 3-D vector measurement, for the proposed InEKF, as shown in Table II. Yet, the proposed filter with forward kinematics measurement has a better overall performance compared with the existing filter. This highlights the importance of modeling the sensor placement offset in ensuring effective filtering under relatively less accurate forward kinematics measurement.

Convergence rate. As discussed in Sec. IV-B, the process and measurement covariance matrices were well tuned for both filters. While the higher convergence rate can be a result of the explicit treatment of sensor placement offset in the proposed filter, additional human studies are needed to confirm that covariance matrices used in the two filters match the actual sensor characteristics and ensure a fair comparison.

Steady-state estimation. From the average RMSE results in Table II, we can see that the proposed filter reaches a higher steady-state accuracy in the velocity estimation compared with the existing filter while they achieve similar accuracy in orientation estimation, given properly tuned noise covariance (and kinematics parameters for the forward kinematics based measurement model). During the tuning process it was observed that both filters had better performance with relatively small angular velocity noise covariance. This indicates that orientation estimation was relied more on the process model for the steady-state period during which the large initial errors have already been corrected using the measurement model. Therefore, both the proposed and existing InEKFs show an almost similar performance in orientation estimation during the steady-state period. Nonetheless, velocity estimation is relatively more dependent on the measurements. Therefore, the proposed filter has a superior performance in steady-state velocity estimation as it benefits from more accurate measurement updates thanks to its offset treatment.
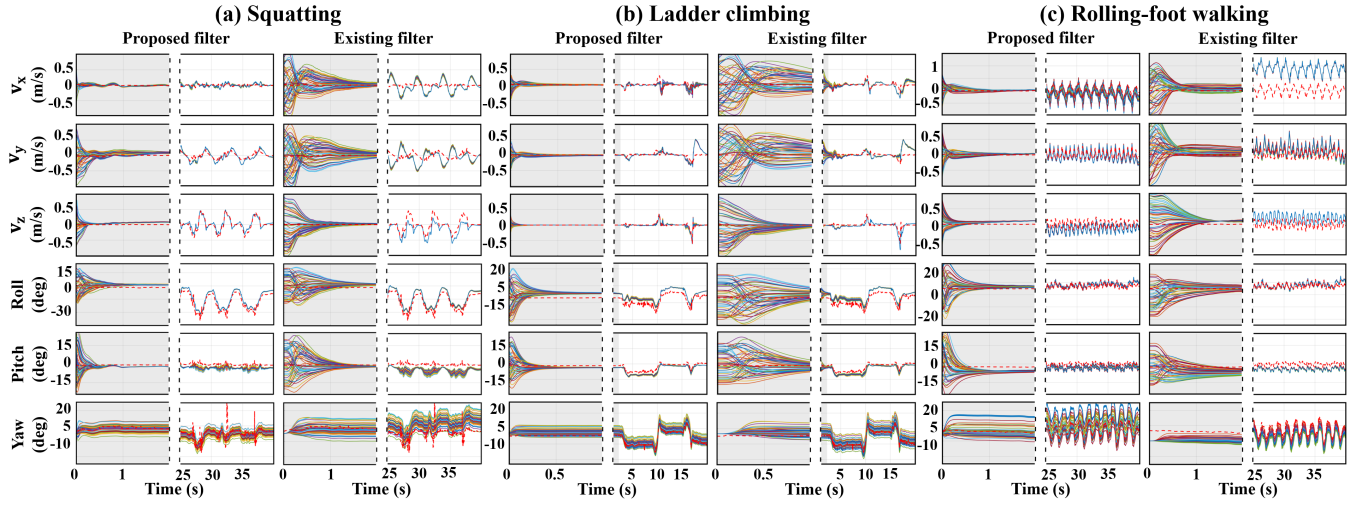
Fig. 4. Estimation results of the velocity and orientation of the subject 2's body/IMU frame during (a) squatting motion, (b) ladder climbing, and (c) rolling-foot walking under the proposed and existing filters. The same sensor data set is used, including the raw data returned by the IMU at the trunk and the leg joint angle data provided by the motion caption system. The solid lines are the state estimates corresponding to different initial errors. The red, dashed lines are the ground truth. The gray and white backgrounds indicate the periods of initial and steady-state movement, respectively. The x-, y-, and z-directions are the lateral, forward and vertical directions, respectively.

Different motion estimation. From Table. II, it is evident that RMSE of the velocity estimations in rolling-foot walking is higher compared to the other two motion types. Walking is a more dynamic task compared to the other two, making the estimation more challenging. The static-foot assumption is more likely to be violated [24] during foot-rolling. Specifically, in rolling-foot walking, we observed relatively high errors in the y-direction of forward kinematic measurement. When large errors are introduced in the measurement, it can also impact the offset estimation, weakening the advantage of the proposed InEKF with respect to the existing one. Since the existing filter does not consider the placement offsets, its velocity estimates have large final errors as shown in Fig. 4.

Limitations. One notable limitation of this work is the modeling of the IMU offset dynamics. It is assumed that the IMU offset has a slowly time-varying dynamics, which can be valid under some activities such as slow walking and stair climbing. However, under more aggressive movements such as running, the IMU might have large and sudden shifts relative to the body, which will violate this assumption. Other limitations include the practical difficulty in obtaining accurate joint angles for the forward kinematic model, and the validity of the assumption of the static foot-ground contact point (which was discussed through our results).

V. CONCLUSION

This paper introduced a right-invariant extended Kalman filter that explicitly considered the offsets between the IMU frame and the measurement frame. The proposed filter design is an "imperfect" invariant extended Kalman filter since the process model satisfied the group affine property but the measurement model does not have the right-invariant observation form. As demonstrated by experimental results among different subjects and motion types, the proposed filter has a low computational cost, and with properly tuned parameters

(e.g., noise covariance and leg kinematics), it improves the convergence rate and estimation accuracy of the IMU velocity estimation compared with the existing filter. This is largely because the IMU offset is treated as a noise source in the original filter while the proposed filter explicitly models and estimates it. The observability analysis shows that the IMU positions and the rotation about the gravity vector were not observable whereas the IMU velocities and the rotations about the other two axes were observable. The observability analysis matched with the experimental results.

In the future work, a more accurate forward kinematics model (e.g., obtained based on online estimation of limb lengths and other kinematics parameters) and explicitly treating IMU biases are needed to improve the filter performance. Also, it is time consuming to tune the process and measurement covariances for different subjects and motion types. A data-driven learning algorithm may be useful for solving this issue. Finally, the forward kinematics measurement in this study assumes a fixed contact point on the foot, which may not be valid for movements involving a nonstationary contact point (e.g., during rolling-foot walking) and thus needs to be relaxed for a more realistic forward kinematics model.

REFERENCES

[1] A. J. Young and D. P. Ferris, "State of the art and future directions for lower limb robotic exoskeletons," *IEEE Trans. Neur. Syst. Rehab. Eng.*, vol. 25, no. 2, pp. 171–182, 2016.

[2] N. Nordin, S. Q. Xie, and B. Wünsche, "Assessment of movement quality in robot-assisted upper limb rehabilitation after stroke: a review," *J. Neuroeng. Rehab.*, vol. 11, no. 1, pp. 1–23, 2014.

[3] E. Q. Yumbla, Z. Qiao, W. Tao, and W. Zhang, "Human assistance and augmentation with wearable soft robotics: a literature review and perspectives," *Cur. Rob. Rep.*, pp. 1–15, 2021.

[4] P. T. Chinimilli, Z. Qiao, S. M. R. Sorkhabadi, V. Jhawar, I. H. Fong, and W. Zhang, "Automatic virtual impedance adaptation of a knee exoskeleton for personalized walking assistance," *Rob. Auton. Syst.*, vol. 114, pp. 66–76, 2019.

[5] A. Martinez, B. Lawson, C. Durrough, and M. Goldfarb, "A velocity-field-based controller for assisting leg movement during walking with a bilateral hip and knee lower limb exoskeleton," *IEEE Trans. Rob.*, vol. 35, no. 2, pp. 307–316, 2018.

[6] S. Luo, G. Androwis, S. Adamovich, H. Su, E. Nunez, and X. Zhou, "Reinforcement learning and control of a lower extremity exoskeleton for squat assistance," *Front. Rob. AI*, vol. 8, 2021.

[7] A. J. Del-Ama, Á. Gil-Agudo, J. L. Pons, and J. C. Moreno, "Hybrid FES-robot cooperative control of ambulatory gait rehabilitation exoskeleton," *J. Neuroeng. Rehab.*, vol. 11, no. 1, pp. 1–15, 2014.

[8] S. Cheng, E. Bolívar-Nieto, and R. D. Gregg, "Real-time activity recognition with instantaneous characteristic features of thigh kinematics," *IEEE Trans. Neur. Syst. Rehab. Eng.*, vol. 29, pp. 1827–1837, 2021.

[9] R. Huang, Z. Peng, S. Guo, K. Shi, C. Zou, J. Qiu, and H. Cheng, "Estimating the center of mass of human-exoskeleton systems with physically coupled serial chain," in *Proc. 2021 IEEE/RSJ Int. Conf. Intel. Rob. Syst.*, pp. 9532–9539.

[10] L. Ojeda and J. Borenstein, "Personal dead-reckoning system for GPS-denied environments," in *2007 IEEE Int. Workshop Safe., Secur. Resc. Rob.*, 2007, pp. 1–6.

[11] Q. Yuan and I.-M. Chen, "3-D localization of human based on an inertial capture system," *IEEE Trans. Rob.*, vol. 29, no. 3, pp. 806–812, 2013.

[12] Q. Yuan and I. Chen, "Localization and velocity tracking of human via 3 IMU sensors," *Sens. Act. A: Phys.*, vol. 212, pp. 25–33, 2014.

[13] Y. Zhang, K. Chen, J. Yi, T. Liu, and Q. Pan, "Whole-body pose estimation in human bicycle riding using a small set of wearable sensors," *IEEE/ASME Trans. Mechatron.*, vol. 21, no. 1, pp. 163–174, 2015.

[14] A. Barrau and S. Bonnabel, "The invariant extended kalman filter as a stable observer," *IEEE Trans. Autom. Contr.*, vol. 62, no. 4, pp. 1797–1812, 2016.

[15] R. Hartley, M. Ghaffari, R. M. Eustice, and J. W. Grizzle, "Contact-aided invariant extended kalman filtering for robot state estimation," *Int. J. Rob. Res.*, vol. 39, no. 4, pp. 402–430, 2020.

[16] Y. Gao, C. Yuan, and Y. Gu, "Invariant extended Kalman filtering for hybrid models of bipedal robot walking," in *Proc. of IFAC Mod., Est, and Contr. Conf.*, vol. 54, no. 20, 2021, pp. 290–297.

[17] S. Teng, M. W. Mueller, and K. Sreenath, "Legged robot state estimation in slippery environments using invariant extended kalman filter with velocity update," in *Proc. IEEE Int. Conf. Rob. Autom.*, 2021, pp. 3104–3110.

[18] P. Cheng, H. Wang, V. Stojanovic, S. He, K. Shi, X. Luan, F. Liu, and C. Sun, "Asynchronous fault detection observer for 2-d markov jump systems," *IEEE Trans. Cyb.*, 2021, in press.

[19] Z. Zhu, M. Rezayat, Y. Gu, and W. Zhang, "Invariant extended kalman filtering for human motion estimation with imperfect sensor placement," in *Proc. 2022 Amer. Contr. Conf.*, accepted.

[20] J. Sola, J. Deray, and D. Atchuthan, "A micro lie theory for state estimation in robotics," *arXiv preprint arXiv:1812.01537*, 2018.

[21] H. Kainz, D. Graham, J. Edwards, H. P. Walsh, S. Maine, R. N. Boyd, D. G. Lloyd, L. Modenese, and C. P. Carty, "Reliability of four models for clinical gait analysis," *Gait & posture*, vol. 54, pp. 325–331, 2017.

[22] G. P. Huang, A. I. Mourikis, and S. I. Roumeliotis, "Observability-based rules for designing consistent EKF SLAM estimators," *Int. J. Rob. Res.*, vol. 29, no. 5, pp. 502–528, 2010.

[23] M. Brossard, A. Barrau, and S. abel, "Ai-imu dead-reckoning," *IEEE Trans. Intel. Veh.*, vol. 5, no. 4, pp. 585–595, 2020.

[24] R. Grönqvist, "Mechanisms of friction and assessment of slip resistance of new and used footwear soles on contaminated floors," *Erg.*, vol. 38, no. 2, pp. 224–241, 1995.