2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Legged Robot State Estimation within Non-inertial Environments

Zijian He[1], Sangli Teng[2], Tzu-Yuan Lin[2], Maani Ghaffari[2], Yan Gu[1]

***Abstract*—This paper investigates the robot state estimation problem within a non-inertial environment. The proposed state estimation approach overcomes the common assumption of static ground in the system modeling. The process and measurement models explicitly treat the movement of the non-inertial environments without requiring knowledge of its motion in the inertial frame or relying on GPS or sensing environmental landmarks. Further, the proposed state estimator is formulated as an invariant extended Kalman filter (InEKF) with the deterministic part of its process model obeying the group-affine property, leading to log-linear error dynamics. The observability analysis of the filter confirms that the robot's pose (i.e., position and orientation) and velocity relative to the non-inertial environment are observable. Hardware experiments on a humanoid robot moving on a rotating and translating treadmill demonstrate the high convergence rate and accuracy of the proposed InEKF even under significant treadmill pitch sway, as well as large estimation errors.**

***Index Terms*—state estimation, non-inertial environments, invariant filtering, legged robots.**

I. Introduction

Legged robots operating in non-inertial environments, such as moving vehicles on land, sea, and air, have significant applications in emergency response, inspection, and surveillance [1]. These environments present unique challenges for robot state estimation, as they deviate from the typical assumption of a static ground, showing continuous and time-varying ground movements [2]. These environments can also be GPS-denied and visually constrained, adding to the complexity of filter design. This paper aims to create an accurate real-time state estimator for legged locomotion in such challenging settings.

Existing filtering approaches for legged and general ground robots generally assume a static ground in the inertial frame, utilizing this static contact point as a pseudo measurement to ensure accurate state estimation through various odometry methods and extended Kalman filters [3]–[7]. However, these methods may fall short in non-inertial settings where ground movement is temporally persistent and multidirectional.

To relax the zero contact velocity condition, new approaches have been explored, such as visual-inertial odometry to account for foot or wheel slippage [8], and inertial-wheel odometry [9] to reject slippage as outliers. Also, to handle significant ground accelerations, previous research has relaxed the static ground assumption, integrating inertial and leg odometry without the need for GPS or fixed environmental landmarks [2]. Yet, this approach assumes accurately known ground pose and velocity in the inertial frame, which may not hold in practical scenarios where the ground movement in the inertial frame cannot be directly sensed or estimated.

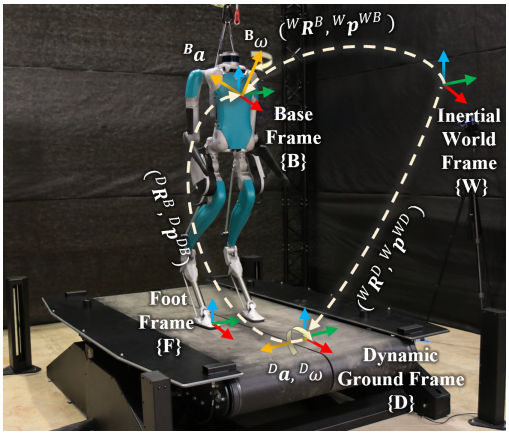

Fig. 1: Illustration of the reference frames used in the filter derivation.

State estimation in multi-agent systems, such as those involving unmanned aerial vehicles, reflects similar complexities encountered in non-inertial environments where reference frames move in the inertial frame. Although cameras and laser scanners can provide data for relative pose estimation [10], their typical slow data acquisition rates and high costs limit their applicability for real-world tasks.

Beyond state estimation in non-inertial environments, invariant extended Kalman filtering (InEKF) [11]–[13] has been introduced to enable fast error convergence under significant errors by exploiting the symmetry reduction for systems evolving on matrix Lie groups. By the InEKF theory [6], if the deterministic unbiased process model satisfies the group-affine property, then there exists an exactly log-linear error dynamics in the Lie algebra. Also, given an invariant observation, the filter is provably convergent under arbitrary initial error, and nonlinear error can be recovered exactly at any time. Although InEKF has been applied to solve the state estimation problem for legged [3] and wheeled [4], [14] robots, the applicability of the InEKF for state estimation in non-inertial environments remains under-explored due to reasons mentioned earlier.

This paper presents an InEKF approach that estimates the relative pose and velocity of a legged robot operating in non-inertial environments under significant estimation errors. The proposed filter overcomes the common assumption that the

This work was supported by the National Science Foundation under Grants 1934280 and 2046562 and by the Office of Naval Research under Grant N00014-21-1-2582. [1]Purdue University, West Lafayette, IN 47907, USA. E-mails: {he348,yangu}@purdue.edu. [2]University of Michigan, Ann Arbor, MI 48109, USA. E-mail: {sanglit,tzuyuan,maanigj}@umich.edu.

ground is static in the inertial frame, and its underlying models explicitly consider the ground movement without requiring knowledge of the ground motion in the inertial frame. The key contributions include: (a) We expand the standard leg odometry-based measurement model to accommodate non-inertial environments, utilizing the kinetic characteristics of robot movement on accelerating surfaces. (b) The deterministic part of the process model is structured to be group affine, ensuring that the logarithmic error equations are not dependent on the state trajectories and are exactly linear. (c) Our filter fuses leg odometry with data from an IMU attached to dynamic ground and another onboard the robot, rendering the robot's relative position, orientation, and velocity observable without relying on exteroceptive sensors such as cameras. (d) Hardware experiments on a humanoid robot moving on a rotating treadmill validate the theoretical results.

II. Mathematical Preliminaries and Notations

Consider a matrix Lie group $\mathcal{G} \subset \mathbb{R}^{n\times n}$. Its Lie algebra $\mathfrak{g}$ is the tangent space at the group identity element $\mathbf{I}_d$. The isomorphism, $(\cdot)^\wedge : \mathbb{R}^{\dim\mathfrak{g}} \to \mathfrak{g}$, maps any vector $\boldsymbol{\xi} \in \mathbb{R}^{\dim\mathfrak{g}}$ to the Lie algebra. The exponential map of the Lie group, $\exp : \mathbb{R}^{\dim\mathfrak{g}} \to \mathcal{G}$, is given by $\exp(\boldsymbol{\xi}) = \exp_m\left(\boldsymbol{\xi}^\wedge\right)$, where $\exp_m(\cdot)$ is the usual matrix exponential. For any $\boldsymbol{\xi} \in \mathbb{R}^{\dim\mathfrak{g}}$ and $\mathbf{X} \in \mathcal{G}$, the adjoint matrix $\mathrm{Ad}_{\mathbf{X}} : \mathfrak{g} \to \mathfrak{g}$ performs a change of basis for velocities to account for the change of observing frame, and is defined as: $(\mathrm{Ad}_{\mathbf{X}}\boldsymbol{\xi})^\wedge = \mathbf{X}\boldsymbol{\xi}^\wedge\mathbf{X}^{-1}$.

We use $\bar{(\cdot)}$ and $\tilde{(\cdot)}$ to denote the estimated value and measurement of $(\cdot)$, respectively. The right subscript t of $(\cdot)_t$ indicates the time t. We use $\{D\}$, $\{W\}$, $\{F\}$ and $\{B\}$ to denote the reference frames attached to the non-inertial dynamic ground, inertial world frame, robot's stance foot, and robot's base link (see Fig. 1). $\mathbf{R}$, $\mathbf{p}$, and $\mathbf{v}$ respectively denote the orientation, position, and velocity of a given object.

The left superscript of a position, orientation, or velocity variable denotes the coordinate system where the variable is expressed. If the right superscript contains two letters, then the first and the second letters respectively represent the reference frame and the object of interest. For instance, we use ${}^D\mathbf{p}^{DB}$ to represent the position of robot's base frame $\{B\}$ relative to the origin of the dynamic ground frame $\{D\}$, expressed in $\{D\}$. If the right superscript only has one letter, then it represents the object of interest. For example, ${}^D\mathbf{R}^B$ denotes the orientation of $\{B\}$ with respect to (w.r.t.) $\{D\}$.

III. Problem Formulation

Navigating robots within non-inertial environments requires that planners and controllers understand the robot's movement state relative to the dynamic ground, rather than the inertial frame. Standard proprioceptive sensors such as IMUs and encoders do not directly measure these states. Thus, our proposed filter estimates the robot's orientation, velocity, and position relative to the dynamic ground frame $\{D\}$.

A. Sensor Measurements

The sensors include (a) a robot's IMU mounted at the robot torso (i.e., the base link), which measures the angular velocity and linear acceleration of the base w.r.t. the base frame $\{B\}$, and (b) joint encoders, which measure the joint angles $\mathbf{q}_t$ of the robot. Additionally, we consider an external IMU attached to the non-inertial dynamic ground frame $\{D\}$ whose data is shared with the robot. This IMU can be placed at any location fixed to the dynamic ground. Such an external sensor setting is general and common as non-inertial platforms such as ships and airplanes are typically equipped with onboard IMUs.

Without loss of generality, we assume that the IMU frames of the robot and the dynamic ground are respectively aligned with the robot's base frame $\{B\}$ and the ground frame $\{D\}$. The joint angle data $\tilde{\mathbf{q}}_t$ returned by encoders is assumed to be corrupted by additive white Gaussian noise. The angular velocity and linear acceleration data from the two IMUs at time t are respectively denoted as ${}^i\tilde{\boldsymbol{\omega}}^{Wi}$ and ${}^i\tilde{\mathbf{a}}^{Wi}$ with $i \in \{B, D\}$. We assume the sensor data is corrupted by additive white Gaussian noise, ${}^i\mathbf{w}_t^g$ and ${}^i\mathbf{w}_t^a$. For brevity, let ${}^i\tilde{\boldsymbol{\omega}}_t := {}^i\tilde{\boldsymbol{\omega}}_t^{Wi}$ and ${}^i\tilde{\mathbf{a}}_t := {}^i\tilde{\mathbf{a}}_t^{Wi}$. Then, we can express the sensor data as: ${}^i\tilde{\boldsymbol{\omega}}_t = {}^i\boldsymbol{\omega}_t^{Wi} + {}^i\mathbf{w}_t^\omega$ and ${}^i\tilde{\mathbf{a}} = {}^i\mathbf{a}_t^{Wi} + {}^i\mathbf{w}_t^a$, where ${}^i\boldsymbol{\omega}^{Wi}$ and ${}^i\mathbf{a}^{Wi}$ are the true angular velocity and linear acceleration.

B. IMU Motion Dynamics

We use ${}^W\mathbf{R}_t^i$, ${}^W\mathbf{v}_t^{Wi}$, and ${}^W\mathbf{p}_t^{Wi}$ to respectively denote the absolute orientation, velocity, and position of the reference frame $\{i\}$ w.r.t. the world frame $\{W\}$, with $i \in \{B, D\}$ (see Fig. 1). The IMU dynamics for the frame $\{i\}$ are [3]: $\frac{d}{dt}\left({}^W\mathbf{R}_t^i\right) = {}^W\mathbf{R}_t^i[{}^i\tilde{\boldsymbol{\omega}}_t - {}^i\mathbf{w}_t^\omega]_\times$, $\frac{d}{dt}\left({}^W\mathbf{v}_t^{Wi}\right) = {}^W\mathbf{R}_t^i\left({}^i\tilde{\mathbf{a}}_t - {}^i\mathbf{w}_t^a\right) + \mathbf{g}$, and $\frac{d}{dt}\left({}^W\mathbf{p}_t^{Wi}\right) = {}^W\mathbf{v}_t^{Wi}$, where $[\,]_\times$ denotes the skew-symmetric matrix of a vector and $\mathbf{g}$ is the gravitational acceleration.

C. Leg Odometry

We denote the position of the robot's stance foot relative to the robot's base, expressed in the base frame as ${}^B\mathbf{p}_t^{BF}$. Since the robot's joint angles $\mathbf{q}_t$ are directly measurable, we introduce the forward kinematics function $s(\mathbf{q}_t)$ satisfying ${}^B\mathbf{p}_t^{BF} = s(\mathbf{q}_t)$, where $s(\mathbf{q}_t)$ is known for a given robot.

The usual measurement model built upon the leg odometry typically assumes a stationary ground in the inertial frame to ensure the observability of the robot's base orientation (roll and pitch) and linear velocity. However, this assumption breaks down for non-inertial environments. Thus, we will introduce a new measurement model based on the leg odometry.

IV. Process and Measurement Models

This section presents the proposed process and measurement models that serve as the basis of the proposed InEKF.

A. Process Model

The process model describes the propagation step, i.e., during the time period between successive instants of measurement updates. For brevity, let $\mathbf{R}_t := {}^D\mathbf{R}_t^B$ and $\mathbf{p}_t := {}^D\mathbf{p}_t^{DB}$.

Given the IMU motion dynamics in Sec. III and the relationship ${}^D\mathbf{R}_t^B = ({}^W\mathbf{R}_t^D)^\mathsf{T}({}^W\mathbf{R}_t^B)$, the dynamics of the robot's relative orientation ${}^D\mathbf{R}_t^B$ during the propagation step is:

$$\tfrac{d}{dt}\mathbf{R}_t = \mathbf{R}_t\left[{}^B\tilde{\boldsymbol{\omega}}_t - {}^B\mathbf{w}_t^\omega\right]_\times - \left[{}^D\tilde{\boldsymbol{\omega}}_t - {}^D\mathbf{w}_t^\omega\right]_\times\mathbf{R}_t. \quad (1)$$

Since the dynamic ground frame $\{D\}$ translates and rotates in the inertial frame, the dynamics of the robot's relative position during the propagation step are given by:

$$\frac{d}{dt}\mathbf{p}_t = -\left[{}^D\tilde{\boldsymbol{\omega}}_t - \mathbf{w}_D^{\omega}\right]_\times \mathbf{p}_t + \mathbf{v}_t, \tag{2}$$

where $\mathbf{v}_t := ({}^W\mathbf{R}_t^D)^\mathsf{T}\left({}^W\mathbf{v}_t^{WB} - {}^W\mathbf{v}_t^{WD}\right)$. Taking the first time derivative of both sides of this equation yields:

$$\begin{aligned}\frac{d}{dt}\mathbf{v}_t = &-\left[{}^D\tilde{\boldsymbol{\omega}}_t - {}^D\mathbf{w}_t^{\omega}\right]_\times \mathbf{v}_t + \mathbf{R}_t\left({}^B\tilde{\mathbf{a}}_t - {}^B\mathbf{w}_t^a\right)\\ &- \left({}^D\tilde{\mathbf{a}}_t - {}^D\mathbf{w}_t^a\right).\end{aligned} \tag{3}$$

The state variables $\mathbf{R}_t$, $\mathbf{v}_t$, and $\mathbf{p}_t$ can be expressed on the matrix Lie group $\mathcal{G} \subset \mathbb{R}^{9\times 9}$ as: $\mathbf{X}_t = \begin{bmatrix}\mathbf{R}_t & \mathbf{v}_t & \mathbf{p}_t\\ \mathbf{0}_{1,3} & 1 & 0\\ \mathbf{0}_{1,3} & 0 & 1\end{bmatrix}$, where $\mathbf{0}_{m,n}$ is an $m \times n$ zero matrix. Here the Lie group $\mathcal{G}$ is the direct isometries group $\mathrm{SE}_2(3)$ [15].

Defining the input $\mathbf{u}_t$ to the process model as: $\mathbf{u}_t = \left[({}^B\tilde{\boldsymbol{\omega}}_t)^\mathsf{T} \quad ({}^D\tilde{\boldsymbol{\omega}}_t)^\mathsf{T} \quad ({}^B\tilde{\mathbf{a}}_t)^\mathsf{T} \quad ({}^D\tilde{\mathbf{a}}_t)^\mathsf{T}\right]$, the process models in (1), (2), and (3) can be expressed as:

$$\begin{aligned}\frac{d}{dt}\mathbf{X}_t &= -{}^D\tilde{\mathbf{U}}_t\mathbf{X}_t + \mathbf{X}_t{}^B\tilde{\mathbf{U}}_t + ({}^D\mathbf{w}_t)^\wedge\mathbf{X}_t - \mathbf{X}_t({}^B\mathbf{w}_t)^\wedge\\ &=: f_{u_t}(\mathbf{X}_t) + ({}^D\mathbf{w}_t)^\wedge\mathbf{X}_t - \mathbf{X}_t({}^B\mathbf{w}_t)^\wedge,\end{aligned} \tag{4}$$

where ${}^i\mathbf{w}_t := \left[({}^i\mathbf{w}_t^g)^\mathsf{T} \quad ({}^i\mathbf{w}_t^a)^\mathsf{T} \quad \mathbf{0}_{1,3}\right]^\mathsf{T}$ and ${}^i\tilde{\mathbf{U}}_t := \begin{bmatrix}\left[{}^i\tilde{\boldsymbol{\omega}}_t\right]_\times & {}^i\tilde{\mathbf{a}}_t & \mathbf{0}_{3,1}\\ \mathbf{0}_{1,3} & 0 & 1\\ \mathbf{0}_{1,3} & 0 & 0\end{bmatrix}$, with $i \in \{D, B\}$.

Proposition 1: *The deterministic part of the system dynamics in* (4)*, i.e.,* $\frac{d}{dt}\mathbf{X}_t = f_{u_t}(\mathbf{X}_t)$*, is group affine.*

Proof: From the process model in (4), we know $f_{u_t}(\mathbf{X}_t) := -{}^D\tilde{\mathbf{U}}_t\mathbf{X}_t + \mathbf{X}_t{}^B\tilde{\mathbf{U}}_t$. Thus, for any $\mathbf{X}_1, \mathbf{X}_2 \in \mathcal{G}$, we have:

$$f_{u_t}(\mathbf{X}_1\mathbf{X}_2) = -{}^D\tilde{\mathbf{U}}_t\mathbf{X}_1\mathbf{X}_2 + \mathbf{X}_1\mathbf{X}_2{}^B\tilde{\mathbf{U}}_t. \tag{5}$$

Meanwhile, by the definition of f_{u_t}, the following expressions can be obtained: $f_{u_t}(\mathbf{X}_1)\mathbf{X}_2(-{}^D\tilde{\mathbf{U}}_t\mathbf{X}_1 + \mathbf{X}_1{}^B\tilde{\mathbf{U}}_t)\mathbf{X}_2$, $\mathbf{X}_1 f_{u_t}(\mathbf{X}_2) = \mathbf{X}_1(-{}^D\tilde{\mathbf{U}}_t\mathbf{X}_2 + \mathbf{X}_2{}^B\tilde{\mathbf{U}}_t)$, and $\mathbf{X}_1 f_{u_t}(\mathbf{I}_d)\mathbf{X}_2 = \mathbf{X}_1(-{}^D\tilde{\mathbf{U}}_t\mathbf{I}_d + \mathbf{I}_d{}^B\tilde{\mathbf{U}}_t)\mathbf{X}_2 = -\mathbf{X}_1{}^D\tilde{\mathbf{U}}_t\mathbf{X}_2 + \mathbf{X}_1{}^B\tilde{\mathbf{U}}_t\mathbf{X}_2$. Note that for the system in (4), the group element $\mathbf{I}_d$ becomes $\mathbf{I}_d = \mathbf{I}_9$ with $\mathbf{I}_m$ an $m \times m$ identity matrix. Combining these equations yields: $f_{u_t}(\mathbf{X}_1)\mathbf{X}_2 + \mathbf{X}_1 f_{u_t}(\mathbf{X}_2) - \mathbf{X}_1 f_{u_t}(\mathbf{I}_d)\mathbf{X}_2 = -{}^D\tilde{\mathbf{U}}_t\mathbf{X}_1\mathbf{X}_2 + \mathbf{X}_1\mathbf{X}_2{}^B\tilde{\mathbf{U}}_t = f_{u_t}(\mathbf{X}_1\mathbf{X}_2)$. Thus, the group affine condition defined in Theorem 1 of [6] is met, confirming the deterministic part of the proposed process model is group affine. □

B. Process Model Discretization

Since filters are implemented in a discrete-time fashion in real-world applications, the process model in (4) needs to be discretized in order to be used during the propagation step.

Let t_k denote the time instant of the k^{th} measurement update with $k \in \mathbb{N}_+$. With abuse of notation, we use $(\cdot)_k$ to represent the value of a variable $(\cdot)$ at t_k. Further, the real scalar Δt denotes the period between two successive measurement updates; i.e., $\Delta t := t_{k+1} - t_k$.

As the process model in (4) is a differential Sylvester equation [16], the closed-form solution of (4) is:

$$\mathbf{X}_{k+1} = {}^D\mathbf{Z}_k^{-1}\mathbf{X}_k{}^B\mathbf{Z}_k, \tag{6}$$

where the matrix ${}^i\mathbf{Z}_k$ with $i \in \{B, D\}$ is defined as [17], [18]:

$${}^i\mathbf{Z}_k = \begin{bmatrix}\boldsymbol{\Gamma}_0({}^i\boldsymbol{\omega}_k\Delta t) & \boldsymbol{\Gamma}_1({}^i\boldsymbol{\omega}_k\Delta t){}^i\mathbf{a}_k & \boldsymbol{\Gamma}_2({}^i\boldsymbol{\omega}_k\Delta t){}^i\mathbf{a}_k\Delta t^2\\ \mathbf{0}_{1,3} & 1 & \Delta t\\ \mathbf{0}_{1,3} & 0 & 1\end{bmatrix}. \tag{7}$$

Given the expression of ${}^i\mathbf{Z}_k$, we can use (6) to discretize the process model and propagate the estimated state $\bar{\mathbf{X}}_t$ during the propagation step of the filter, as explained later.

C. Measurement Model

When the robot's foot has static contact with the ground of the non-inertial environment (i.e., no foot slipping or rolling on the ground), the foot velocity satisfies: $\frac{d}{dt}({}^D\mathbf{p}_t^{DF}) = \mathbf{0}_{3,1}$. For brevity, we define $\mathbf{d}_t := {}^D\mathbf{p}_t^{DF}$.

Using the kinematics relationship associated with the leg odometry, we obtain: $\mathbf{d}_t - \mathbf{p}_t = \mathbf{R}_t s(\mathbf{q}_t)$. Taking the first time derivative of both sides of this equation gives: $\frac{d}{dt}(\mathbf{d}_t - \mathbf{p}_t) = \mathbf{R}_t\left(\left[{}^B\boldsymbol{\omega}_t\right]_\times + \left[{}^D\boldsymbol{\omega}_t\right]_\times\right)s(\mathbf{q}_t) + \mathbf{R}_t\mathbf{J}(\mathbf{q}_t)\dot{\mathbf{q}}_t$, where $\mathbf{J}(\mathbf{q}_t) = \frac{\partial s(\mathbf{q}_t)}{\partial \mathbf{q}_t}$ is the Jacobian of leg odometry $s(\mathbf{q}_t)$ and $\dot{\mathbf{q}}_t$ is the time derivative of the joint angle $\mathbf{q}_t$.

Combining the equations above gives the observation as:

$$\mathbf{y}_t = h(\mathbf{X}_t) + \mathbf{n}_f, \tag{8}$$

where $h(\mathbf{X}_t) = \mathbf{R}_t^\mathsf{T}\left(\left[{}^D\tilde{\boldsymbol{\omega}}_t\right]_\times \mathbf{R}_t s(\tilde{\mathbf{q}}_t) - \mathbf{v}_t + \left[{}^D\tilde{\boldsymbol{\omega}}_t\right]_\times \mathbf{p}_t\right)$, $\mathbf{y}_t = \left[{}^B\tilde{\boldsymbol{\omega}}_t\right]_\times s(\tilde{\mathbf{q}}_t) + \mathbf{J}\dot{\tilde{\mathbf{q}}}_t$, and $\mathbf{n}_f$ is the lumped white Gaussian noise of the uncertainty in the encoder reading $\tilde{\mathbf{q}}_t$ and foot slippage on the ground.

The deterministic portion of the measurement model in (8) does not satisfy the right-invariant observation form, which is defined as $\mathbf{y}_t = \mathbf{X}_t^{-1}\mathbf{b}$ with a known constant vector $\mathbf{b}$ [6]. Thus, the log-error equation associated with the proposed measurement model does not enjoy the attractive properties of an invariant observation and is thus not necessarily independent of state trajectories or exactly linear for the deterministic case.

Instead, we linearize the measurement model as follows:

$$\mathbf{z}_t = h(\bar{\mathbf{X}}_t) - h(\mathbf{X}_t) := \mathbf{H}_t\boldsymbol{\xi}_t + \text{h.o.t}(\boldsymbol{\xi}_t). \tag{9}$$

where $\mathbf{H}_t := \frac{\partial \mathbf{z}_t}{\partial \boldsymbol{\xi}_t}$. As $\boldsymbol{\eta}_t \approx \mathbf{I}_d + \boldsymbol{\xi}_t^\wedge$, the following first-order approximations hold: $\bar{\mathbf{R}}_t\mathbf{R}_t^\mathsf{T} \approx \mathbf{I}_3 + [\boldsymbol{\xi}_t^R]_\times$, $\bar{\mathbf{v}}_t - \bar{\mathbf{R}}_t\mathbf{R}_t^\mathsf{T}\mathbf{v}_t \approx \boldsymbol{\xi}_t^v$, and $\bar{\mathbf{p}}_t - \bar{\mathbf{R}}_t\mathbf{R}_t^\mathsf{T}\mathbf{p}_t \approx \boldsymbol{\xi}_t^p$, with $\boldsymbol{\xi}_t^R$, $\boldsymbol{\xi}_t^v$, and $\boldsymbol{\xi}_t^p$ defined as:

$$\boldsymbol{\xi}_t^\wedge =: \begin{bmatrix}[\boldsymbol{\xi}_t^R]_\times & \boldsymbol{\xi}_t^v & \boldsymbol{\xi}_t^p\\ \mathbf{0}_{1,3} & 0 & 0\\ \mathbf{0}_{1,3} & 0 & 0\end{bmatrix}. \tag{10}$$

We substitute the definition of $\mathbf{h}(\mathbf{X}_t)$ from (8) into (9) and simplify it using the first-order approximations mentioned above, while dropping the higher-order terms. By differentiating the resulting equation w.r.t. $\boldsymbol{\xi}_t$, we obtain the expression of the update matrix $\mathbf{H}_t$ as:

$$\mathbf{H}_t = \begin{bmatrix}\mathbf{C}_t & -\bar{\mathbf{R}}_t^\mathsf{T} & \bar{\mathbf{R}}_t^\mathsf{T}\left[{}^D\tilde{\boldsymbol{\omega}}_t\right]_\times\end{bmatrix} \tag{11}$$

with $\mathbf{C}_t := \bar{\mathbf{R}}_t^\mathsf{T}\left[\left[{}^D\tilde{\boldsymbol{\omega}}_t\right]_\times \bar{\mathbf{R}}_t s(\tilde{\mathbf{q}})\right]_\times - \bar{\mathbf{R}}_t^\mathsf{T}\left[{}^D\tilde{\boldsymbol{\omega}}_t\right]_\times \left[\bar{\mathbf{R}}_t s(\tilde{\mathbf{q}})\right]_\times + \bar{\mathbf{R}}_t^\mathsf{T}\left[\left[{}^D\tilde{\boldsymbol{\omega}}_t\right]_\times \bar{\mathbf{p}}_t\right]_\times - \bar{\mathbf{R}}_t^\mathsf{T}\left[{}^D\tilde{\boldsymbol{\omega}}_t\right]_\times [\bar{\mathbf{p}}_t]_\times$.

V. Filter Design

A. Propagation Step

1) Error Dynamics of Process Model: By the methodology of InEKF, the right-invariant estimation error $\boldsymbol{\eta}_t$ between the state $\mathbf{X}_t$ and its estimate $\bar{\mathbf{X}}_t$ is defined as: $\boldsymbol{\eta}_t = \bar{\mathbf{X}}_t \mathbf{X}_t^{-1}$.

Because of the group-affine property of the proposed process model [6], the right-invariant error dynamics in the absence of noise are independent of state trajectories and exactly log-linear in the deterministic case.

The dynamics of the right-invariant error $\boldsymbol{\eta}_t$ is given by [6]:

$$\tfrac{d}{dt}\boldsymbol{\eta}_t = g_{u_t}(\boldsymbol{\eta}_t) + \left(\bar{\mathbf{X}}_t ({}^B\mathbf{w}_t)^\wedge \bar{\mathbf{X}}_t^{-1}\right)\boldsymbol{\eta}_t + ({}^D\mathbf{w}_t)^\wedge \boldsymbol{\eta}_t, \quad (12)$$

where $g_{u_t}(\boldsymbol{\eta}_t) := f_{u_t}(\boldsymbol{\eta}_t) - \boldsymbol{\eta}_t f(\mathbf{I}_d)$. Note that by the InEKF theory, the deterministic part of the right-invariant error ($\tfrac{d}{dt}\boldsymbol{\eta}_t = g_{u_t}(\boldsymbol{\eta}_t)$) are state trajectory independent and accordingly independent of estimation errors.

By using the first-order approximation $\boldsymbol{\eta}_t = \exp(\boldsymbol{\xi}_t) \approx \mathbf{I}_d + \boldsymbol{\xi}_t^\wedge$, we linearize (12) to yield:

$$g_{u_t}(\exp(\boldsymbol{\xi}_t)) =: (\mathbf{A}_t\boldsymbol{\xi}_t)^\wedge + \mathrm{h.o.t}(\|\boldsymbol{\xi}_t\|) \approx (\mathbf{A}_t\boldsymbol{\xi}_t)^\wedge, \quad (13)$$

where $\mathrm{h.o.t}(\cdot)$ represents the higher-order terms of $(\cdot)$.

Then, the linearized log-error dynamics become:

$$\tfrac{d}{dt}\boldsymbol{\xi}_t = \mathbf{A}_t\boldsymbol{\xi}_t + \mathrm{Ad}_{\bar{\mathbf{X}}_t}{}^B\mathbf{w}_t + {}^D\mathbf{w}_t. \quad (14)$$

Since the deterministic part of the right-invariant error equation are state trajectory independent, the logarithmic error dynamics are naturally independent of state trajectories in the absence of noise, as indicated by (14). Further, the linear error equation (14) is exact in the absence of noise.

Proposition 2: *In the absence of the noise terms in the stochastic process model* (4)*, the deterministic portion of the logarithmic error dynamics* (14)*, i.e.,* $\tfrac{d}{dt}\boldsymbol{\xi}_t = \mathbf{A}_t\boldsymbol{\xi}_t$*, are exact and represent the true error dynamics during propagation.*

Proof: By Proposition 1, the deterministic part of the process model (4) is group affine. Then, by Theorem 2 in [6], the logarithmic error dynamics in the absence of noise ${}^B\tilde{\boldsymbol{\omega}}_t$ and ${}^D\tilde{\boldsymbol{\omega}}_t$ are exact, which completes the proof. □

By Proposition 2, the linear equation in (14) is the exact dynamics of the error $\boldsymbol{\xi}_t$ in the absence of noise terms. Such linearity is rare for nonlinear process models, and holds here because the deterministic portion of the process model is group affine for the deterministic case, as stated in the proof.

The log-error equation in (14) is used to form the propagation step of the proposed InEKF, and the advantage of its exactness is illustrated via experiment results.

To obtain the matrix $\mathbf{A}_t$, we substitute the right-invariant error dynamics (12) into (13), which yields:

$$\begin{aligned} g_{u_t}(\exp(\boldsymbol{\xi}_t)) &\approx f(\mathbf{I}_d + \boldsymbol{\xi}_t^\wedge) - \left(\mathbf{I}_d + \boldsymbol{\xi}_t^\wedge\right) f(\mathbf{I}_d) \\ &= \begin{bmatrix} -\left[{}^D\tilde{\boldsymbol{\omega}}_t\right]_\times \boldsymbol{\xi}_t^R \\ -\left[{}^D\tilde{\mathbf{a}}_t\right]_\times \boldsymbol{\xi}^R - \left[{}^D\tilde{\boldsymbol{\omega}}_t\right]_\times \boldsymbol{\xi}_t^v \\ \boldsymbol{\xi}_t^v - \left[{}^D\tilde{\boldsymbol{\omega}}_t\right]_\times \boldsymbol{\xi}_t^p \end{bmatrix}^\wedge. \end{aligned} \quad (15)$$

Then, based on (13), we obtain the matrix $\mathbf{A}_t$ as:

$$\mathbf{A}_t = \begin{bmatrix} -\left[{}^D\tilde{\boldsymbol{\omega}}_t\right]_\times & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} \\ -\left[{}^D\tilde{\mathbf{a}}_t\right]_\times & -\left[{}^D\tilde{\boldsymbol{\omega}}_t\right]_\times & \mathbf{0}_{3,3} \\ \mathbf{0}_{3,3} & \mathbf{I}_3 & -\left[{}^D\tilde{\boldsymbol{\omega}}_t\right]_\times \end{bmatrix}. \quad (16)$$

2) State and Covariance Propagation: Between two successive instants of measurement updates, i.e., $t \in [t_k, t_{k+1})$ ($k \in \mathbb{N}_+$), the estimated state $\bar{\mathbf{X}}_t$ can be propagated [18] using the discretized process model in (6): $\bar{\mathbf{X}}_{k+1} = {}^D\mathbf{Z}_k^{-1}\bar{\mathbf{X}}_k{}^B\mathbf{Z}_k$.

By the theory of the standard Kalman filtering for continuous-time systems, the covariance matrix $\mathbf{P}_t$ is propagated based on the following Riccati equation [19] associated with the linearized log-error equation in (14):

$$\tfrac{d}{dt}\mathbf{P}_t = \mathbf{A}_t\mathbf{P}_t + \mathbf{P}_t\mathbf{A}_t^\mathsf{T} + \bar{\mathbf{Q}}_t, \quad (17)$$

where $\bar{\mathbf{Q}}_t$ is the process noise covariance defined as $\bar{\mathbf{Q}}_t = \mathrm{Ad}_{\bar{\mathbf{X}}_t}\mathrm{Cov}({}^B\mathbf{w}_t)\mathrm{Ad}_{\bar{\mathbf{X}}_t}^\mathsf{T} + \mathrm{Cov}({}^D\mathbf{w}_t)$, with $\mathrm{Cov}({}^i\mathbf{w}_t)$ the covariance of ${}^i\tilde{\mathbf{w}}_t$ ($i \in \{B, D\}$).

In filter implementation, the discrete version of the Riccati equation (17) is used for covariance propagation.

B. Update Step

Based on the measurement model introduced in Sec. IV-C, the update equations of the proposed InEKF are: $\bar{\mathbf{X}}_t^+ = \exp\left(\mathbf{K}_t(\mathbf{y}_t - h(\bar{\mathbf{X}}_t))\right)\bar{\mathbf{X}}_t$ and $\mathbf{P}_t^+ = (\mathbf{I}_9 \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_t\left(\mathbf{I}_9 - \mathbf{K}_t\mathbf{H}_t\right)^\mathsf{T} + \mathbf{K}_t\mathbf{N}_t\mathbf{K}_t^\mathsf{T}$, where $\bar{\mathbf{X}}_t^+$ and $\mathbf{P}_t^+$ are the updated values of the state estimate $\bar{\mathbf{X}}_t$ and covariance matrix $\mathbf{P}_t$, respectively, $\mathbf{K}_t$ is the Kalman gain, and $\mathbf{N}_t$ is the measurement covariance matrix. The Kalman gain $\mathbf{K}_t$ is given by: $\mathbf{K}_t = \mathbf{P}_t\mathbf{H}_t^\mathsf{T}\mathbf{S}_t^{-1}$, $\mathbf{S}_t := \mathbf{H}_t\mathbf{P}_t\mathbf{H}_t^\mathsf{T} + \mathbf{N}_t$, and $\mathbf{N}_t := \bar{\mathbf{R}}_t\mathrm{Cov}(\mathbf{n}_f)\bar{\mathbf{R}}_t^\mathsf{T}$.

VI. Observability Analysis

Assuming that IMU measurements are constant over the propagation step on $[t_k, t_{k+1})$, the matrix $\mathbf{A}_k$ is constant. Thus, the discrete-time state-transition matrix, denoted as $\boldsymbol{\Phi}_k$, is given by [20]: $\boldsymbol{\Phi}_k = \exp_m(\mathbf{A}_k\Delta t)\begin{bmatrix} \boldsymbol{\phi}_{11}^k & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} \\ \boldsymbol{\phi}_{21}^k & \boldsymbol{\phi}_{22}^k & \mathbf{0}_{3,3} \\ \boldsymbol{\phi}_{31}^k & \boldsymbol{\phi}_{32}^k & \boldsymbol{\phi}_{33}^k \end{bmatrix}$, where $\boldsymbol{\phi}_{11}^k = \boldsymbol{\phi}_{22}^k = \boldsymbol{\phi}_{33}^k = \exp_m(-[{}^D\boldsymbol{\omega}_k]_\times\Delta t)$, $\boldsymbol{\phi}_{21}^k = -[{}^D\mathbf{a}_k]_\times\exp_m(-[{}^D\boldsymbol{\omega}_k]_\times\Delta t)\Delta t$, $\boldsymbol{\phi}_{31}^k = -\frac{1}{2}[{}^D\mathbf{a}_k]_\times \exp_m\left(-[{}^D\boldsymbol{\omega}_k]_\times\Delta t\right)\Delta t^2$, and $\boldsymbol{\phi}_{32}^k = \exp_m(-[{}^D\boldsymbol{\omega}_k]_\times\Delta t)\Delta t$.

Then, the local observability matrix $\mathcal{O}$ [21] at the state estimate $\bar{\mathbf{X}}_k$ is expressed as: $\mathcal{O} = [(\mathbf{H}_k^-)^\mathsf{T},\ (\mathbf{H}_{k+1}^-\boldsymbol{\Phi}_k^+)^\mathsf{T},\ (\mathbf{H}_{k+2}^-\boldsymbol{\Phi}_{k+1}^+\boldsymbol{\Phi}_k^+)^\mathsf{T},\ \ldots]^\mathsf{T}$. By definition, $\mathcal{O}$ can be computed as:

$$\mathcal{O} = \begin{bmatrix} \mathbf{C}_t & -\bar{\mathbf{R}}_k^\mathsf{T} & \mathbf{R}_k^\mathsf{T}\left[{}^D\boldsymbol{\omega}_k\right]_\times \\ \mathbf{o}_{21} & \mathbf{o}_{22} & \mathbf{R}_{k+1}^\mathsf{T}\left[{}^D\boldsymbol{\omega}_{k+1}\right]_\times\boldsymbol{\phi}_{33}^k \\ \mathbf{o}_{31} & \mathbf{o}_{32} & \bar{\mathbf{R}}_{k+2}^\mathsf{T}[{}^D\boldsymbol{\omega}_{k+2}]_\times\boldsymbol{\phi}_{33}^{k+1}\boldsymbol{\phi}_{33}^k, \\ \vdots & \vdots & \vdots \end{bmatrix}, \quad (18)$$

where $\mathbf{o}_{21} = \mathbf{C}_{k+1}\boldsymbol{\phi}_{11}^k - \bar{\mathbf{R}}_{k+1}^\mathsf{T}\boldsymbol{\phi}_{21}^k + \bar{\mathbf{R}}_{k+1}^\mathsf{T}[{}^D\boldsymbol{\omega}_k]_\times\boldsymbol{\phi}_{31}^k$, $\mathbf{o}_{22} = -\bar{\mathbf{R}}_{k+1}^\mathsf{T}\boldsymbol{\phi}_{22}^k + \bar{\mathbf{R}}_{k+1}^\mathsf{T}[{}^D\boldsymbol{\omega}_k]_\times\boldsymbol{\phi}_{33}^k$, $\mathbf{o}_{31} = \mathbf{C}_{k+2}\boldsymbol{\phi}_{11}^{k+1}\boldsymbol{\phi}_{11}^k - \bar{\mathbf{R}}_{k+2}^\mathsf{T}\boldsymbol{\phi}_{21}^{k+1}\boldsymbol{\phi}_{11}^k - \bar{\mathbf{R}}_{k+2}^\mathsf{T}\boldsymbol{\phi}_{22}^{k+1}\boldsymbol{\phi}_{21}^k + \bar{\mathbf{R}}_{k+2}^\mathsf{T}[{}^D\boldsymbol{\omega}_{k+2}]_\times\boldsymbol{\phi}_{31}^{k+1}\boldsymbol{\phi}_{11}^k + \bar{\mathbf{R}}_{k+2}^\mathsf{T}[{}^D\boldsymbol{\omega}_{k+2}]_\times\boldsymbol{\phi}_{32}^{k+1}\boldsymbol{\phi}_{21}^k + \bar{\mathbf{R}}_{k+2}^\mathsf{T}[{}^D\boldsymbol{\omega}_{k+2}]_\times\boldsymbol{\phi}_{33}^{k+1}\boldsymbol{\phi}_{31}^k$, and $\mathbf{o}_{32} = -\bar{\mathbf{R}}_{k+2}^\mathsf{T}\boldsymbol{\phi}_{22}^{k+1}\boldsymbol{\phi}_{22}^{k+2} + \bar{\mathbf{R}}_{k+1}^\mathsf{T}[{}^D\boldsymbol{\omega}_k]_\times\boldsymbol{\phi}_{32}^{k+1}\boldsymbol{\phi}_{22}^k + \bar{\mathbf{R}}_{k+1}^\mathsf{T}[{}^D\boldsymbol{\omega}_k]_\times\boldsymbol{\phi}_{33}^{k+1}\boldsymbol{\phi}_{32}^k$.

setup.png setup.png setup.bb setup.bb

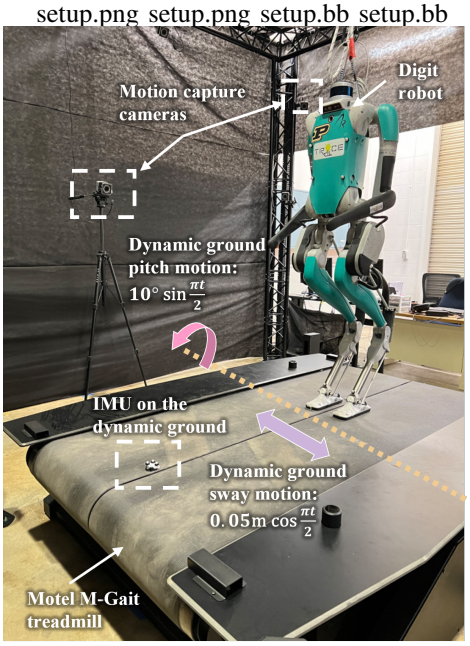

Fig. 2: Experimental setup that includes a Digit robot, motion capture cameras, a pitch sway treadmill, and an IMU mounted on the dynamic ground.

To evaluate the observability of each variable of interest, we examine whether the associated column vectors in the observability matrix $\mathcal{O}$ are linearly independent.

From the expression of $\mathcal{O}$, the observability of the state variables depends on the estimated relative orientation $\bar{\mathbf{R}}_t$ as well as the linear acceleration data ${}^D\mathbf{a}_t$ and angular velocity data ${}^D\boldsymbol{\omega}_t$ of the dynamic ground. The estimate $\bar{\mathbf{R}}_t$ is always a non-zero matrix. Thus, when the ground is rotating and translating (i.e., ${}^D\mathbf{a}_t \neq \mathbf{0}$ and ${}^D\boldsymbol{\omega}_t \neq \mathbf{0}$), all columns of $\mathcal{O}$ are linearly independent, indicating the relative orientation $\mathbf{R}_t$, velocity $\mathbf{v}_t$, and position $\mathbf{p}_t$ are observable.

When the ground is stationary, the angular velocity data ${}^D\tilde{\boldsymbol{\omega}}_t$ is zero in the absence of sensor noise, and thus the entire third column block becomes zeros, indicating the relative position $\mathbf{p}_t$ is no longer observable. Also, when the ground is not moving, ${}^D\tilde{\mathbf{a}}_t$ remains nonzero because ${}^D\tilde{\mathbf{a}}_t$ includes the vertical gravitational acceleration. Thus, the third column of $\left[{}^D\mathbf{a}_t\right]_\times$ is zero, indicating the yaw angle is non-observable when the ground is stationary.

VII. Experimental Validation

A. Experimental setup

Experiments are conducted on a Digit humanoid robot (Agility Robotics, Inc.) and a Motek M-Gait treadmill (Fig. 2). Digit is 1.6 m tall with 6 encoders on each leg. The robot stands on the treadmill commanded by its proprietary controller. The treadmill serves as a dynamic ground, simultaneously performing a pitch motion of $10^\circ \sin\frac{\pi t}{2}$ and a sway motion of $0.05\text{m}\cos\frac{\pi t}{2}$. An IMU (WT901BLECL from WitMotion Co.,Ltd) is attached to the treadmill and measures the angular velocity and linear acceleration of the dynamic ground frame at 200 Hz via Bluetooth. The robot IMU and encoders return data at 500 Hz. Additionally, a Vicon motion capture system gives the ground-truth value of the state $\mathbf{X}_t$.

TABLE I: NOISE STANDARD DEVIATION

Measurement types	SRS	Proposed
Robot linear acc. (m/s^2)	0.3	0.1
Ground linear acc. (rad/s)	NA	0.1
Robot angular vel. (m/s^2)	0.01	0.01
Ground angular vel. (rad/s)	NA	0.01
Encoder reading	1°	0.1 m/s
Contact vel. (m/s)	0.01	NA

The proposed filter is compared with an InEKF [3] designed for locomotion on a static, rigid surface (denoted as "SRS"), so as to highlight the advantage of explicitly treating the environment/ground motion in the filter formulation.

The key difference between the proposed and SRS filters is that the SRS filter assumes a static ground. Accordingly, the SRS filter aims to estimate the absolute base position, orientation, and velocity w.r.t. the world frame, which is different from the proposed filter. Although the process models of the two filters are different due to different choices of state variables, both models meet the group-affine property for the deterministic case. This indicates that both filters obey the attractive property of invariant filtering, such as the exact linearity and state independence of log-error dynamics for the deterministic part of the process model. Also, the measurement models of both filters exploit the leg odometry, with the SRS and proposed filters using position- and velocity-based ones, respectively. Yet, the baseline filter has a right-invariant measurement model, while the proposed one does not.

The setting of the standard deviation (SD) of both filters is shown in Table I. All the SD values are individually tuned based on the IMU specifications provided by the manufacturers for the two filters to achieve their respective best performance. Both filters are assessed using the same hardware sensor data. To highlight the proposed InEKF can handle large errors, 50 simulations of each filter were performed, using the same initial position, velocity, and orientation errors uniformly sampled from $[-3, 3]$ m, $[-1, 1]$ m/s, and $[-23, 23]$ deg, respectively.

B. Results

1) Convergence Rate: Figure 3 presents the results of the proposed filter for the relative velocity $\mathbf{v}_t$, orientation $\mathbf{R}_t$, and position $\mathbf{p}_t$ w.r.t. $\{D\}$ in subplot a), and the baseline filter results for the absolute velocity ${}^W\mathbf{v}^{WB}$, orientation ${}^W\mathbf{R}^B$, and position ${}^W\mathbf{p}^{WB}$ w.r.t. $\{W\}$ in subplot b). Both filters drive the errors of the base roll, pitch, and velocities close to zero, confirming the observability analysis results from Sec. VI and previous work [3]. Both filters show fast error convergence for their respective observable state even under large initial errors, thanks to InEKF's provable error convergence under the deterministic case. The proposed filter shows a much faster convergence rate than that of the SRS filter due to its explicit treatment of the ground motion.

2) Yaw and Position Observability: Notably, under the proposed filter, the robot's relative base yaw and position converge to the ground truth, confirming that they are indeed observable during ground motion. In contrast, the absolute

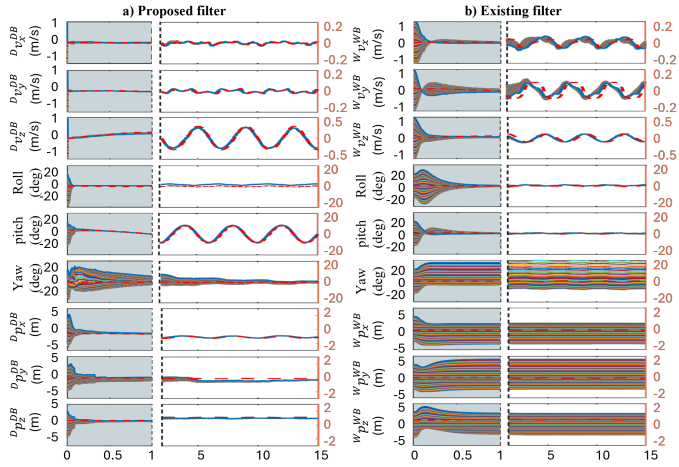

Fig. 3: Estimation results of a) relative velocity, orientation, and position under the proposed filter and b) absolute velocity, orientation, and position under the existing filter. The same set of robot on-board sensor data is used. The red, dashed lines are the ground truth. The solid lines are the state estimations corresponding to different initial errors. The light blue and white background denotes the transient and steady-state periods, respectively.

TABLE II: RMS ERROR COMPARISON

State variables	SRS	Proposed
(v_x, v_x, v_x) (m/s)	(0.048, 0.080, 0.041)	(0.017, 0.018, 0.040)
(roll, pitch, yaw) (°)	(1.889, 1.520, 10.91)	(1.886, 0.980, 2.871)
(p_x, p_y, p_z) (m)	(1.145, 1.688, 1.225)	(0.283, 0.336, 0.165)

yaw and position under the SRS filter are not observable as predicted by the previous study [3].

3) Estimation Accuracy: The estimation results with white background in Figure 3 show the steady-state periods on $t \in [2,\ 15]s$. Table II reports the comparison of the root-mean-square (RMS) errors between the state estimate and the ground truth for both filters. As the state variables estimated by the two filters have different physical meanings, directly comparing their specific accuracy may not be meaningful. Still, the smaller errors of the proposed method do highlight the need to explicitly consider the ground motion in the state estimation, especially under significant ground motions such as the tested treadmill movements. Without explicit treatment, the ground motion acts as temporally persistent, significant uncertainties that could notably degrade estimator performance.

VIII. Conclusion

This paper developed a real-time state estimation approach for legged locomotion inside a non-inertial environment with an unknown ground motion. The process and measurement models underlying the estimator were formulated to explicitly consider the movement of the non-inertial environment. A minimal suite of proprioceptive sensors and an inertial measurement unit attached to the environment were used to inform the proposed InEKF. The observability analysis revealed that all state variables (i.e., relative pose and linear velocity) are observable during environment translation and rotation. Hardware experiment results and comparison with a baseline InEKF demonstrated the fast convergence rate and high accuracy of the proposed filter under various ground motions and substantial estimation errors. The proposed system modeling can be readily used in filtering and optimization frameworks beyond InEKF, and can be combined with data returned by exteroceptive sensors such as cameras and LiDARs. Future work includes the study of fully onboard sensing and learning-aided methods to remove the need for an external IMU attached to the moving environment.

References

[1] A. Iqbal, Y. Gao, and Y. Gu, "Provably stabilizing controllers for quadrupedal robot locomotion on dynamic rigid platforms," *IEEE/ASME Trans. Mechatron.*, vol. 25, no. 4, pp. 2035–2044, 2020.

[2] Y. Gao, C. Yuan, and Y. Gu, "Invariant filtering for legged humanoid locomotion on a dynamic rigid surface," *IEEE/ASME Trans. Mechatron.*, vol. 27, no. 4, pp. 1900–1909, 2022.

[3] 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.

[4] T.-Y. Lin, T. Li, W. Tong, and M. Ghaffari, "Proprioceptive invariant robot state estimation," *arXiv preprint arXiv:2311.04320*, 2023.

[5] M. Bloesch, M. Hutter, M. A. Hoepflinger, S. Leutenegger, C. Gehring, C. D. Remy, and R. Siegwart, "State estimation for legged robots-consistent fusion of leg kinematics and IMU," *Rob.*, vol. 17, pp. 17–24, 2013.

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

[7] X. Yu, S. Teng, T. Chakhachiro, W. Tong, T. Li, T.-Y. Lin, S. Koehler, M. Ahumada, J. M. Walls, and M. Ghaffari, "Fully proprioceptive slip-velocity-aware state estimation for mobile robots via invariant Kalman filtering and disturbance observer," in *Proc. IEEE Int. Conf. Intel. Rob. Syst.*, 2023, pp. 8096–8103.

[8] 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.

[9] M. Bloesch, C. Gehring, P. Fankhauser, M. Hutter, M. A. Hoepflinger, and R. Siegwart, "State estimation for legged robots on unstable and slippery terrain," in *Proc. IEEE Int. Conf. Intel. Rob. Syst.*, 2013, pp. 6058–6064.

[10] F. Aghili and C.-Y. Su, "Robust relative navigation by integration of icp and adaptive Kalman filter using laser scanner and IMU," *IEEE/ASME Trans. Mechatron.*, vol. 21, no. 4, pp. 2015–2026, 2016.

[11] S. Bonnable, P. Martin, and E. Salaün, "Invariant extended Kalman filter: theory and application to a velocity-aided attitude estimation problem," in *Proc. IEEE Conf. Dec. Contr.*, 2009, pp. 1297–1304.

[12] A. Barrau and S. Bonnabel, "Intrinsic filtering on Lie groups with applications to attitude estimation," *IEEE Trans. Autom. Contr.*, vol. 60, no. 2, pp. 436–449, 2014.

[13] ——, "Extended kalman filtering with nonlinear equality constraints: A geometric approach," *IEEE Transactions on Automatic Control*, vol. 65, no. 6, pp. 2325–2338, 2019.

[14] M. Brossard, A. Barrau, and S. Bonnabel, "AI-IMU dead-reckoning," *IEEE Trans. Intel. Veh.*, vol. 5, no. 4, pp. 585–595, 2020.

[15] A. Barrau, "Non-linear state error based extended Kalman filters with applications to navigation," Ph.D. dissertation, Mines Paristech, 2015.

[16] M. Behr, P. Benner, and J. Heiland, "Solution formulas for differential Sylvester and Lyapunov equations," *Calcolo*, vol. 56, no. 4, p. 51, 2019.

[17] J. Sola, "Quaternion kinematics for the error-state kalman filter," *arXiv preprint arXiv:1711.02508*, 2017.

[18] M. A. Shalaby, C. C. Cossette, J. L. Ny, and J. R. Forbes, "Multi-robot relative pose estimation and IMU preintegration using passive UWB transceivers," *IEEE Trans. Rob.*, pp. 1–20, 2024.

[19] P. S. Maybeck, *Stochastic models, estimation, and control*. Academic press, 1982.

[20] Z. Huai and G. Huang, "Robocentric visualinertial odometry," *Int. J. Rob. Res.*, vol. 41, no. 7, pp. 667–689, 2022.

[21] Z. Chen, K. Jiang, and J. C. Hung, "Local observability matrix and its application to observability analyses," in *Proc. Ann. Conf. IEEE Ind. Electron. Soc.*, 1990, pp. 100–103.