Robustification of Bayesian-Inference-Based Gait Estimation for Lower-limb Wearable Robots

Ting-Wei Hsu[1], Robert D. Gregg[2], and Gray C. Thomas[3]

***Abstract*—Lower-limb wearable robots designed to assist people in everyday activities must reliably recover from any momentary confusion about what the user is doing. Such confusion might arise from momentary sensor failure, collision with an obstacle, losing track of gait due to an out-of-distribution stride, etc. Systems that infer a user's walking condition from angle measurements using Bayesian filters (e.g., extended Kalman filters) have been shown to accurately track gait across a range of activities. However, due to the fundamental problem structure and assumptions of Bayesian filter implementations, such estimators risk becoming 'lost' with little hope of a quick recovery. In this paper, we 1) introduce a Monte Carlo-based metric to quantify the robustness of pattern-tracking gait estimators, 2) propose strategies for improving tracking robustness, and 3) systematically evaluate them against this new metric using a publicly available gait biomechanics dataset. Our results, aggregating 2,700 trials of simulated walking of 10 able-bodied subjects under random perturbations, suggest that drastic improvements in robustness (from 8.9% to 99%) are possible using relatively simple modifications to the estimation process without noticeably degrading estimator accuracy.**

***Index Terms*—Prosthetics and Exoskeletons; Wearable Robotics; Rehabilitation Robotics.**

I. Introduction

LOWER-LIMB wearable robots, such as powered prostheses and exoskeletons, have great potential to enhance mobility for people with lower-limb disabilities. By inferring their user's intent (such as walking or climbing stairs) from onboard sensors and applying the corresponding torques to the user's biological or prosthetic joints, these devices aim to mechanically compensate for the disability and to allow users to perform the tasks comfortably. However, since typical onboard sensors of such robots, for example, their inertial measurement units (IMUs), joint encoders, and force/torque sensors, can only offer a limited picture of the user's true intent [1]–[12], the inference of human intent is challenging, and its failure can lead to unreliable device behavior. For safety-critical applications with a prosthetic leg, this unreliability is a major obstacle to system acceptance and adoption. As a result, reliably tracking the various activities of daily living has become a key challenge for lower-limb wearable robot controllers.

One way to estimate gait variation is to quantify the gait progression using a continuous gait phase variable [1]–[3], [13]. Gait phase is typically considered to increase at a constant rate from 0 (at heel strike) to 1 (right before the next ipsilateral heel strike) in a steady-state gait cycle. Continuous phase variables are continuous functions of measurable quantities that improve over this behavior so that control laws can be synchronized to motion, even within a stride. Some researchers utilize phase plane analysis and nonlinear mappings to calculate a phase variable from the user's lower-limb kinematics such as the global thigh angle [1], [14], hip angle [2], [15], or tibia motion [3]. This approach allows tracking of variable-period locomotion. However, the same phase variable will not necessarily work between different tasks, limiting the practical scalability of the approach.

Machine learning (ML) has also been applied to task variation in human locomotion. ML has been utilized to perform classification of locomotion mode identification and intent recognition for lower-limb wearable robots [4]–[9]. Deep convolutional neural networks have also been used to estimate continuous task-related gait parameters, such as stride length [16], [17]. ML-based gait phase estimators have demonstrated promising accuracy in gait phase estimation and adaptation to varying walking tasks [18]–[23]. However, ML-based methods usually require a significant amount of data for model training [5], [6], [8]. The black-box-like nature of neural networks also makes their behavior unpredictable and hard to debug.

Bayesian inference provides a more straightforward methodology for gait estimation [10]–[12]. Bayesian filters work by treating the gait state estimates as random variables that follow certain probability distributions, utilizing Bayes' theorem to update the probability as new information is observed. One common technique to implement Bayesian inference is to assume a Gaussian distribution. This leads to different variants of Kalman filters, such as the extended Kalman filter (EKF) and unscented Kalman filter (UKF). The main difference between the EKF and UKF is the method they utilize to propagate Gaussian random variables through the nonlinear dynamic and measurement models. The EKF linearizes the dynamic and measurement models about the current state estimate using Taylor approximation, thereby obtaining an

Manuscript received: August 21, 2023; Revised: November 29, 2023; Accepted: December 29, 2023.

This paper was recommended for publication by Editor Jee-Hwan Ryu upon evaluation of the Reviewers' comments. This work was supported by the National Institute of Child Health & Human Development of the NIH under Award Number R01HD094772. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

[1]Ting-Wei Hsu was with the Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109 USA twhsu@umich.edu. He is now with Bechamo LLC, Buffalo, NY 14203 USA

[2]Robert D. Gregg is with the Department of Robotics, University of Michigan, Ann Arbor, MI 48109 USA rdgregg@umich.edu.

[3]Gray C. Thomas was with the Department of Robotics, University of Michigan, Ann Arbor, MI 48109 USA. He is now with the J. Mike Walker '66 Department of Mechanical Engineering, Texas A&M University, College Station, TX 77843 USA gthomas@tamu.edu.

Digital Object Identifier (DOI): see top of this page.

analytical expression of the approximated filtering equation. The UKF utilizes unscented transform to approximate the mean and covariance of the Gaussian distribution that is mapped through the nonlinear dynamic and measurement models [24]. The equations and the derivation of the EKF and UKF can be found in [24], [25]. Recent work has shown promising results in tracking the stance phase for a powered prosthetic leg using an EKF [10] and state estimation for a powered prosthetic ankle using a suite of Kalman filters [26]. Moreover, the EKF and UKF have been used to estimate continuous task variables such as walking speed, stride length, and ground inclination in addition to phase [11], [12], building on the recent development of such continuous models and supporting datasets [27]–[29]. However, these Kalman filter variants all share the disadvantage of potentially losing track of the state estimates due to the assumptions they make. For gait estimation to reach safety-critical applications, the robustness of the estimators is a key requirement; however, this robustness issue of the Bayesian-filter-based gait estimators has yet to be properly investigated.

In this study, the objective is to evaluate, compare, and enhance the robustness of two Bayesian-filter-based gait estimators (EKF and UKF-based) *in silico*. Our simulation supplies the filters with kinematic measurements from a biomechanical dataset [27], and the filters are judged on their ability to predict a gait-state comprising phase, phase rate, normalized stride length, and ramp angle. Our contributions include 1) the introduction of a Monte Carlo-based metric for quantifying the gait estimators' robustness against loss-of-tracking scenarios, 2) three robustifying mechanisms for the gait estimators designed to achieve an almost-100% recovery rate from loss-of-tracking scenarios, and 3) systematic evaluations of the EKF and UKF implementations and comparisons of the proposed robustifying mechanisms using hypothesis testing. To the best of our knowledge, this paper is the first to systematically study and address the robustness issue of Bayesian-filter-based gait estimators for lower-limb wearable robots. By addressing this gap in robustness quantification and robustness-oriented filter design, we hope to enable the practical application of gait-state tracking controllers to safety-critical lower-limb wearable robots.

II. Gait Estimation Methods

A. Human Gait Model

We define a gait-state vector with four state variables to describe human gait, given by

$$x = [\phi \ \dot{\phi} \ l \ r]^T \in \mathbb{R}^{4\times 1}, \tag{1}$$

where ϕ denotes the gait phase, $\dot{\phi} = \mathrm{d}\phi/\mathrm{d}t$ denotes the phase rate, l denotes the normalized stride length (defined as stride length normalized by leg length), and r denotes the ramp angle. We model the evolution of the gait-state vector in time as a stochastic process: $\dot{\phi}$, l, and r integrate white noise, while ϕ integrates $\dot{\phi}$.

The measurement model simply predicts the gait kinematics according to phase and task variables—similar to existing models in the biomechanics literature [27]. In this work, we consider global thigh angles θ_{th}, global thigh angular velocities $\dot{\theta}_{th}$, and foot angles θ_f as measurements. The definition of θ_{th} and θ_f are shown in Fig. 1a. First, we define a set of continuous functions that map the gait-state x to θ_{th} and θ_f. This kinematic model is structurally linear in its parameters,

$$\theta_{th} = \Lambda(x)\Psi_{th}, \quad \theta_f = \Lambda(x)\Psi_f, \tag{2}$$

where $\Lambda : \mathbb{R}^{4\times 1} \to \mathbb{R}^{1\times M}$ is a single-row regressor vector and Ψ_{th} and $\Psi_f \in \mathbb{R}^{M\times 1}$ are single-column parameter vectors to be learned from model fitting. The definition of the regressor $\Lambda(x)$ uses the Kronecker product to compact simple relationships in ϕ, l, and r,

$$\Lambda(x) = \Lambda_\phi(\phi) \otimes \Lambda_l(l) \otimes \Lambda_r(r), \tag{3}$$

where $\Lambda_\phi(\phi)$, $\Lambda_l(l)$, and $\Lambda_r(r)$ are a Fourier series

$$\begin{aligned}\Lambda_\phi(\phi) =&[1 \ \cos(2\pi\phi) \ \sin(2\pi\phi) \ \cos(4\pi\phi) \\ &\ldots \cos(20\pi\phi) \ \sin(20\pi\phi)] \in \mathbb{R}^{1\times 21},\end{aligned} \tag{4}$$

and two Bernstein polynomial bases,

$$\Lambda_l(l) = [(1-l)^2 \quad 2(1-l)l \quad l^2] \in \mathbb{R}^{1\times 3}, \tag{5}$$

$$\Lambda_r(r) = [(1-r)^2 \quad 2(1-r)r \quad r^2] \in \mathbb{R}^{1\times 3}. \tag{6}$$

Thus, the width of the regressor $\Lambda(x)$ and size of the parameter vectors is $M = 189$. Since the Kronecker product is linear, we have convenient expressions for the phase-derivative of the kinematic signals, and we use this to produce velocity estimates, for example

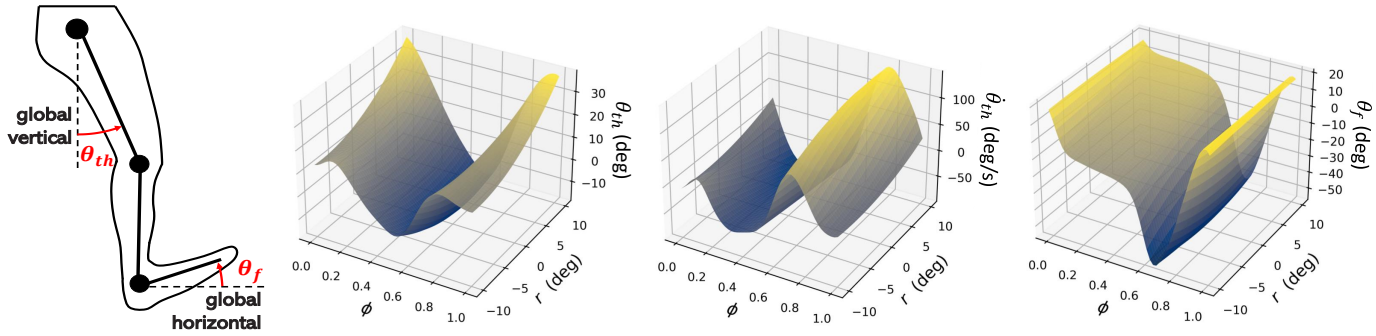
$$\dot{\theta}_{th} = \frac{\partial \Lambda}{\partial \phi}\dot{\phi} \ \Psi_{th} = \dot{\phi}\left(\frac{\partial \Lambda_\phi(\phi)}{\partial \phi} \otimes \Lambda_l(l) \otimes \Lambda_r(r)\right) \ \Psi_{th} \tag{7}$$

The parameter vectors, Ψ_{th} and Ψ_f, are fit using constrained least-squares to best predict the kinematics in a multi-activity dataset published by Embry et al. [27]. This dataset provides ground truth gait-state (phase, phase rate, normalized stride length, and ramp angle) along with kinematics data (including global thigh angles and foot angles) of 10 able-bodied subjects (5 male/5 female; mean age: 23 $\pm$ 2.8 (SD) years; mean height: 170 $\pm$ 8.2 cm) walking at different combinations of constant speeds (0.8, 1.0, and 1.2 m/s) and constant ramp angles (-10, -7.5, -5, -2.5, 0, 2.5, 5, 7.5, 10 deg). Using these labeled data, $(x, \theta_{th}, \theta_f)$, we can solve for the unknown parameter vectors Ψ_{th} and Ψ_f in Eq. (2) using least-squares. When implementing lest-squares, we impose constraints such that the global thigh angles are always zero when the normalized stride length is zero. The resulting gait kinematics model is visualized in Fig. 1b.

B. Real-time Gait State Estimation

Our simulation uses the EKF and UKF to estimate the gait-state from gait kinematic measurements. These filters use the following dynamics model alongside the previously mentioned measurement model. In discrete time, the gait state evolution is given by

$$x_k = A \ x_{k-1} + \xi_k, \tag{8}$$



(a) (b)

Fig. 1. (a) View of a human leg from the sagittal plane and the definitions of θ_{th} and θ_f. (b) The data-driven gait kinematic models of global thigh angles θ_{th}, global thigh angular velocities $\dot{\theta}_{th}$, and foot angles θ_f. Note that each model is multidimensional, depending on ϕ, l, r (and also $\dot{\phi}$ for the $\dot{\theta}_{th}$ model). For visualization, we plot the models with respect to ϕ and r by fixing $l = 1.1$ (and $\dot{\phi} = 0.8$ for the $\dot{\theta}_{th}$ model).

where $x_k = [\phi_k \ \dot{\phi}_k \ l_k \ r_k]^T \in \mathbb{R}^{4\times1}$ denotes the gait-state vector at time step k, $A \in \mathbb{R}^{4\times4}$ denotes the discrete-time state matrix, and $\xi_k \sim N(0, Q)$ denotes the additive Gaussian process noise. The phase ϕ is always wrapped to $[0, 1)$. The state matrix A describes how the gait-state propagates over time. In this work, we assume that in the absence of process noise, the phase rate, normalized stride length, and ramp angle are constant over time, resulting in a linear, discrete-time state matrix, given by

$$A = \begin{bmatrix} 1 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{9}$$

where Δt is the time step. For the process noise $\xi_k \sim N(0, Q)$, we assume that the process noise covariance is a diagonal matrix in the form of $Q = \text{diag}([0, \sigma_{\dot{\phi}}^2, \sigma_l^2, \sigma_r^2]) \times \Delta t$. In this work, we consider $\sigma_{\dot{\phi}}^2 = 10^{-3}$, $\sigma_l^2 = 5 \times 10^{-2}$, and $\sigma_r^2 = 5$, which are tuned such that the EKF and UKF both yield comparable level of root-mean-squared errors (RMSEs) with [12] and achieve similar RMSEs with each other.

We consider global thigh angles, global thigh angular velocities, and foot angles as measurements for the filters. The data-driven kinematics models discussed in (2) and (7) are used as the measurement model, given by

$$z_k = h(x_k) + \eta_k = \begin{bmatrix} \Lambda(x_k)\Psi_{th} \\ \frac{\partial \Lambda(x_k)}{\partial \phi}\dot{\phi}\Psi_{th} \\ \Lambda(x_k)\Psi_f \end{bmatrix} + \eta_k, \tag{10}$$

where $z_k = [\theta_{th_k} \ \dot{\theta}_{th_k} \ \theta_{f_k}]^T \in \mathbb{R}^{3\times1}$ denotes the measurements, and $\eta_k \sim N(0, \ R)$ denotes the measurement noise. The measurement noise covariance $R \in \mathbb{R}^{3\times3}$ is computed empirically using the residuals between the predicted values of the kinematics model and the actual data in the dataset.

In this work, both the EKF and UKF are initialized with prior mean $\mu_0 = [0.5, 0.8, 1.1, 0]^T$ (i.e., mid-stance) and prior state covariance $\Sigma_0 = \text{diag}([10^{-2}, 10^{-1}, 10^{-1}, 10^{-1}])$, and the parameters of the UKF are set to $(\alpha, \beta, \kappa) = (10^{-3}, 2, 0)$ as suggested by [24] and [12]. Singular value decomposition is employed to compute the matrix square root in the UKF algorithm.

III. Simulation Methods

In order to evaluate and compare the performance of the EKF and the UKF as gait estimators, we performed simulations using the gait data provided by the same dataset used for model training in Section II-A. This dataset contains a total of 270 trials of gait kinematics data and ground truth gait-state sampled at 100 Hz. In each trial of simulation, we used the global thigh angles, global thigh angular velocities, and foot angles as the measurements for the filters to estimate the gait states. Fig. 2 demonstrates the EKF and UKF estimation. As can be seen, both filters are able to track the ground truth.

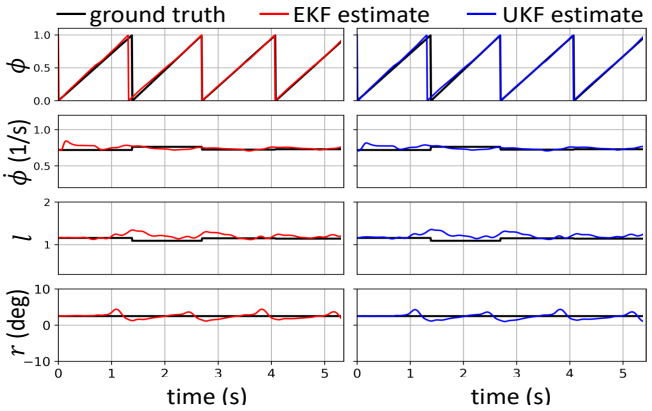


Fig. 2. An example of the EKF and UKF gait-state estimation for a randomly selected trail in which a user walked at 0.8 m/s on a 2.5-degree inclined ramp.

IV. Evaluation of Robustness

When Bayesian-filter-based gait-state estimators lose track, their local Gaussian simplification of the posterior distribution prevents them from noticing the superior tracking performance of more correct gait-state estimates. This results in the state estimate remaining 'lost' in a local minimum instead of converging to the best possible estimate of the gait state. The risk of this failure mode depends on both the filter's myopic linear assumptions and on the non-linearity of the estimation problem itself. In this work, we evaluate the robustness of the EKF and UKF empirically by two metrics: 1) the RMSE between

the estimation and the ground truth (a standard measure of tracking performance) and 2) a Monte Carlo-based kidnapped robot test that aims to quantify robustness against loss-of-tracking scenarios.

A. Root-Mean-Squared Error (RMSE)

We use the RMSE between the gait-state estimation and the ground truth to evaluate the performance of the filters. The RMSEs are shown in Table I. As can be seen, both the EKF and the UKF yield satisfactory tracking performance, and their RMSEs are similar. In addition, the RMSEs are comparable to their counterparts reported in [12], which calculated the average RMSEs across all individual strides for a representative trial in simulations (EKF: RMSE_ϕ = 0.01 $\pm$ 0.009 (SD), $\text{RMSE}_{\dot\phi}$ = 0.02 $\pm$ 0.02, RMSE_l = 0.08 $\pm$ 0.06, and RMSE_r = 0.72 $\pm$ 1.62; UKF: RMSE_ϕ = 0.01 $\pm$ 0.009, $\text{RMSE}_{\dot\phi}$ = 0.02 $\pm$ 0.02, RMSE_l = 0.08 $\pm$ 0.06, and RMSE_r = 0.73 $\pm$ 1.64).

TABLE I
RMSEs OF THE GAIT-STATE ESTIMATORS

	RMSE$_\phi$	**RMSE**$_{\dot\phi}$ (1/s)	**RMSE**$_l$	**RMSE**$_r$ (deg)
EKF	0.022	0.028	0.126	2.351
UKF	0.021	0.028	0.125	2.323

B. A Monte Carlo-based Kidnapping Test

The kidnapped robot problem refers to a scenario in which a mobile robot is suddenly moved to an unknown location while performing localization [30]. This is a challenging problem because the robot has a strong belief in where it is when being kidnapped, making it difficult to recover and locate itself again. Inspired by this idea, we propose a metric to empirically evaluate a gait-state estimator's robustness against staying lost by kidnapping the state estimates to random conditions and measuring the likelihood that the estimate returns to the correct value within a time limit.

The kidnapping simulates scenarios in which the state estimates deviate from tracking due to any possible disturbances (e.g., momentary sensor failure, striking an obstacle, etc.) and end up at some unexpected locations in the state space after the disturbance is over. For each gait-state variable, we kidnap the estimate during steady-state by setting its value to a uniformly distributed random variable:

$$\begin{aligned} &\hat{\phi}_k \sim U[0,1), \quad \hat{\dot{\phi}}_k \sim U[0,5], \\ &\hat{l}_k \sim U[0,2], \quad \hat{r}_k \sim U[-10,10] \text{ deg}, \end{aligned} \tag{11}$$

where $\hat{(.)}$ denotes the state estimate and k is the time step when kidnapping occurs. The bounds of the kidnapped states are prescribed values selected based on the loss-of-tracking scenarios from which we hope the estimators can always recover. For each trial in the dataset, we run 11 simulations. One of the simulations is the nominal case in which no kidnapping is performed. In the other 10 simulations, we randomly kidnap the state estimates during steady-state, and check if the state estimates converge back to the nominal case after the kidnapping event. Note that the time step k at which the kidnapping event occurs is also random. Fig. 3 shows an example of the kidnapping tests for the EKF. The condition of convergence is met if the differences of the state estimates between the kidnapped case and the nominal case is less than 1) 0.025 for phase, 2) 0.15 for normalized stride length, and 3) 2.5 degrees for ramp angle. These values can be any reasonably small numbers, depending on how strict we want the kidnapping tests to be. In this work, we define the thresholds using values slightly larger than the RMSEs in Table I. Because we have 270 trials of walking data for simulation as mentioned in Section III, we ultimately ran a total of 2,700 kidnapping tests.

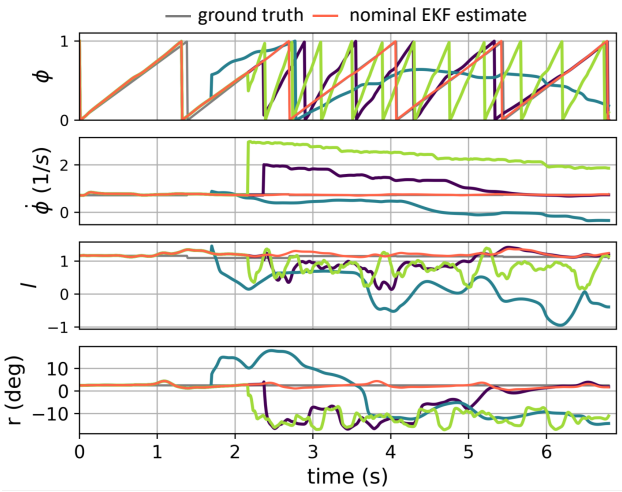


Fig. 3. An example of the kidnapping test (1 nominal simulation + 3 kidnapped simulations) for the original EKF using the same trail as in Fig. 2. The kidnapping occurs during the second stride. The lines with different colors denote the kidnapped simulations, showing that most trajectories are unable to recover back to the nominal case.

To complete the metric, we compute the percentage of the kidnapping tests in which the state estimates converge within i strides of the kidnapping events, and define this percentage as the robustness of the filter, denoted by R_i. Here, we show only $i = 1$ and 3. The percentage represents the likelihood of the estimator recovering from any disturbed state estimates (given by Eq. (11)) using available measurements after the disturbance is over. For example, an estimator with $R_1 = 99\%$ means that if its estimates deviate to some unexpected location in the state space due to some disturbances, it has a 99% chance of getting back on track after one regular stride using available measurements. In this work, R_1 emphasizes recovery speed, and R_3 emphasizes the recoverability (i.e., ability to recover at slower recovery speeds). For the gait estimator to be practically safe, it must recover within one stride after the kidnapping occurs. In other words, our goal is to ultimately enhance R_1, while R_3 is only used to better understand the structure of the problem.

Note that the values of R_1 and R_3 are the sample means of Bernoulli random variables (success or failure in each kidnapping test). By Central Limit Theorem, if we perform a sufficiently large number of the kidnapping tests, the values

of R_1 and R_3 will converge to Gaussian distributions,

$$R_1 \xrightarrow{d} \mathcal{N}(\mathcal{R}_1,\ \sigma_1^2/N); \quad R_3 \xrightarrow{d} \mathcal{N}(\mathcal{R}_3,\ \sigma_3^2/N), \tag{12}$$

where $\xrightarrow{d}$ denotes converging in distribution; $\mathcal{R}_1$ and $\mathcal{R}_3$ denote the true means of the Bernoulli random variables (i.e., true robustness); σ_1 and σ_3 denote the standard deviation of the Bernoulli random variables; and N denotes the total number of kidnapping tests. In this work, we will utilize hypothesis testing to compare the values of R_1 and R_3 in different cases.

V. Enhancement of Robustness

Applying the robustness test to the original EKF and UKF, neither the EKF ($R_1 = 8.9\%$) nor UKF ($R_1 = 9.8\%$) appears to be R_1-robust, and the results are still unsatisfactory after three steps (EKF: $R_3 = 20.8\%$, UKF: $R_3 = 22.5\%$). Looking closely at the recovery trajectories in Fig. 3, we observe that some of the estimates are completely lost, and some of the l and r estimates seem to reach values that are out of the range of the training dataset (i.e., $-10 \leq r \leq 10$ deg and $0.8 \leq l \leq 1.9$). In addition, some of the l estimates even become negative, which is physically impossible. These observed failure modes indicate that the measurement model does not provide unambiguous information regarding the gait states, and the state estimates might also reach some regions in the state space that are out of the training dataset. Therefore, we introduced the following three robustifying mechanisms, described in subsections V-A, V-B, and V-C, respectively, with the goal of achieving $R_1 \approx 100\%$.

A. Auxiliary Measurement: Phase Angle of Thigh

One factor that causes the filters to lose track is the ambiguity of the measurement model. For example, the relationship between the gait-state and the gait kinematics may not be unique in some regions in the state space. Therefore, to enhance the robustness, we introduce an auxiliary measurement that provides less-ambiguous information regarding the gait phase. To do this, we apply a phase-portrait-based estimate of gait progression [14], in which the phase angle of thigh motion acts as our auxiliary measurement. This auxiliary measurement, φ_a, is calculated the same way as in [14] (see Appendix).

We utilize the same regression approach to train the measurement model for this auxiliary measurement. Because we assume that φ_a depends only on phase, we construct this measurement model by a single sub-regressor of phase $\Lambda_\phi(\phi)$ of degree 10, that is

$$\varphi_a = \Lambda_\phi(\phi)\Psi_{\varphi_a}, \tag{13}$$

where $\Psi_{\varphi_a} \in \mathbb{R}^{21\times 1}$ is a column vector of unknown parameters. Then, we use the same dataset as in Section II-A to train this model, thereby solving for Ψ_{φ_a} using least-squares. Fig. 4 shows the comparison among the actual φ_a, the predicted φ_a by the least-square model, and the ground truth phase. This auxiliary measurement closely resembles the phase variable, which means it can provide practically unambiguous information regarding the user's position in the gait cycle. Finally, the new measurement model incorporating φ_a becomes

$$z_k = h(x_k) + \eta_k = \begin{bmatrix} \Lambda(x_k)\Psi_{th} \\ \frac{\partial \Lambda(x_k)}{\partial \phi}\dot{\phi}\ \Psi_{th} \\ \Lambda(x_k)\Psi_f \\ \Lambda_\phi(\phi_k)\Psi_{\varphi_a} \end{bmatrix} + \eta_k, \tag{14}$$

where $\eta_k \sim N(0, R)$ is the additive measurement noise. Here, the measurement noise covariance $R \in \mathbb{R}^{4\times 4}$ is also computed empirically using the errors between the predicted values of the model and the actual data in the dataset.

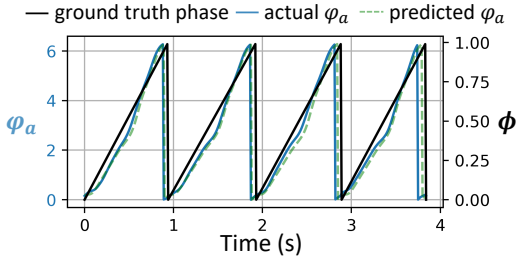


Fig. 4. The auxiliary measurement φ_a compared to the ground truth phase ϕ. The RMSE between φ_a and the ground truth ϕ is 0.06.

B. State Estimate Saturation

The second approach to enhance the robustness is to saturate the gait-state estimates within a certain range. Because the training dataset only covers gait-state variables within a certain range, our data-driven measurement model may become unpredictable when the estimates are outside this range, causing the filter to lose track. Therefore, saturating the state estimates may serve as a solution to this issue. In this study, we define the saturation bounds using the lower and upper bounds of the ground truth gait-state in the training dataset: [0.5, 1.2] for the $\dot{\phi}$ estimates, [0.8, 1.9] for the l estimates, and [-10, 10] degrees for the r estimates.

C. Resetting the Filter based on Failure Detection

Our third robustifying mechanism is resetting the filter whenever it detects itself losing track. To this end, we define a detector based on the Mahalanobis distance between the measurement and the predicted measurement, given by

$$MD_k = \sqrt{v_k^T S_k^{-1} v_k}, \tag{15}$$

where v_k and S_k are the innovation and innovation covariance of the filter at the k-th time step. We determine that the filter is lost if $MD_k > \overline{MD}$, where $\overline{MD}$ is a tunable threshold value. In this work, we consider $\overline{MD} = 4$. We reset the filter to the initial prior mean and covariance (i.e., $\mu_0 = [0.5, 0.8, 1.1, 0]^T$ and $\Sigma_0 = \text{diag}([10^{-2}, 10^{-1}, 10^{-1}, 10^{-1}])$) whenever it detects itself losing track (i.e., $MD_k > 4$).

D. Statistical Comparison Between Robustification Methods

To compare the robustness R_1 and R_3, which are Bernoulli random variables as mentioned in Section IV-B, associated with different robustifying mechanisms, we utilize one-sided

hypothesis testing. Let R_i and R_i' $(i = 1, 3)$ be the statistical robustness resulting from two different robustifying mechanisms. The null and alternative hypotheses are $H_0 : \mathcal{R}_i = \mathcal{R}_i'$ and $H_1 : \mathcal{R}_i > \mathcal{R}_i'$, where $\mathcal{R}_i$ and $\mathcal{R}_i'$ are the true robustness as described in Eq. (12). With a large number of kidnapping tests performed ($N = 2,700$), we can define the test statistics

$$z_{\text{TS}} = \frac{R_i - R_i'}{\sqrt{\overline{R}_i(1-\overline{R}_i)\frac{2}{N}}}, \tag{16}$$

where $\overline{R} = (R_i + R_i')/2$. Considering a significant level of 5%, we reject H_0 and accept H_1 if $z_{\text{TS}} > 1.65$.

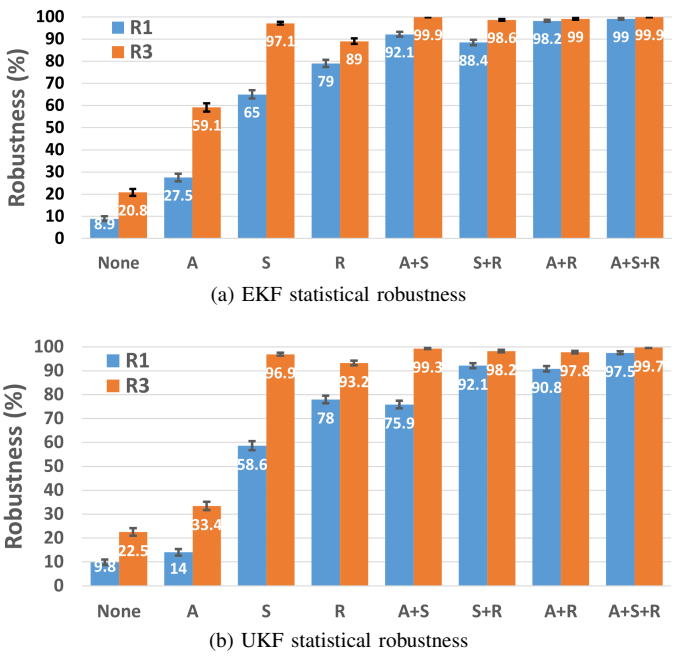
VI. Results

In general, all of the robustifying mechanisms statistically improved R_1 and R_3 for both the EKF and the UKF (see Fig. 5). In particular, if we apply only one of the three robustifying mechanisms, estimate saturation and filter resetting yield the most significant enhancement of the robustness. For example, the EKF with only 'S' improved the robustness from the original $R_1 = 8.9\%$ and $R_3 = 20.8\%$ to $R_1 = 65\%$ and $R_3 = 97.1\%$ (both $z_{\text{TS}} > 1.65$). With only 'R', the EKF's robustness also increased to $R_1 = 79\%$ and $R_3 = 89\%$ (both $z_{\text{TS}} > 1.65$).

The auxiliary measurement constructed by the phase angle of global thigh motion does not seem to be as effective as the other two mechanisms when it is applied alone. Compared to the original robustness, 'A' increased the robustness only to $R_1 = 27.5\%$ and $R_3 = 59.1\%$ for the EKF and $R_1 = 14\%$ and $R_3 = 33.4\%$ for the UKF (all $z_{\text{TS}} > 1.65$). However, the auxiliary measurement plays a crucial role in further enhancing the robustness when the other two mechanisms have been applied. For instance, with 'S+R', the EKF achieved $R_1 = 88.4\%$ and $R_3 = 98.6\%$. Although it had been significantly improved compared to the original robustness, it still failed to achieve an almost-100% recovery rate. With 'A+S+R', the robustness was further enhanced to $R_1 = 99\%$ and $R_3 = 99.9\%$ with statistical significance ($z_{\text{TS}} > 1.65$ comparing to 'S+R'). Similarly, the UKF with 'S+R' could only achieve $R_1 = 92.1\%$ and $R_3 = 98.2\%$, but with 'A+S+R' the robustness further increased to $R_1 = 97.5\%$ and $R_3 = 99.7\%$ ($z_{\text{TS}} > 1.65$). This enhancement of robustness with all three robustifying mechanisms applied can also be seen in the recovery trajectories (compare Fig. 6 to Fig. 3).

The RMSEs of both filters with all three robustifying mechanisms applied are shown in Table II. The RMSE of ϕ of the EKF (i.e., 0.027) roughly corresponds to 31 milliseconds at the typical speeds within our dataset. Compared to Table I, the RMSEs of the UKF do not degrade at the cost of applying the robustifying mechanisms; however, the RMSEs of the EKF degrade as a trade-off. In particular, the RMSE of the EKF's phase rate estimation degrades the most.

In our experiment, the EKF generally outperformed the UKF in the Monte Carlo-based kidnapping tests (compare Figs. 5a and 5b). Most of the R_1 and R_3 of the UKF are statistically lower than their counterpart of the EKF. With all three robustifying mechanisms applied, the UKF achieved $R_1 = 97.5\%$, which was still unsatisfactory compared to the



(a) EKF statistical robustness

(b) UKF statistical robustness

Fig. 5. Statistical robustness by the Monte Carlo-based kidnapping tests with 95% confidence intervals. 'A': auxiliary measurement; 'S': estimate saturation; 'R': filter resetting.

EKF's $R_1 = 99\%$ ($z_{\text{TS}} > 1.65$). Yet, with no robustifying mechanisms applied, the differences in robustness between the two filters were not significant. Compared to the EKF's $R_1 = 8.9\%$ and $R_3 = 20.8\%$, the UKF yielded slightly but not significantly higher $R_1 = 9.8\%$ ($z_{\text{TS}} = 1.13 < 1.65$) and $R_3 = 22.5\%$ ($z_{\text{TS}} = 1.51 < 1.65$).

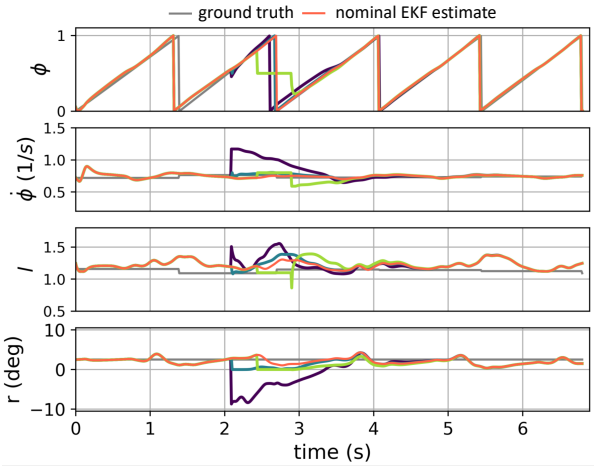


Fig. 6. An example of the kidnapping test (1 nominal simulation + 3 kidnapped simulations) for the robustified EKF using the same trail as in Figs. 2 and 3. All three robustifying mechanisms are applied. The lines with different colors denote the kidnapped simulations, showing enhanced recovery performance compared to Fig. 3.

VII. Discussion

The results of the kidnapping tests have shown that the proposed robustifying approaches can significantly enhance the estimator's robustness against loss-of-tracking scenarios. If only one robustifying mechanism is applied, saturation and

TABLE II
RMSEs OF THE GAIT-STATE ESTIMATORS WITH ALL ROBUSTIFYING MECHANISMS

	$\mathbf{RMSE}_{\phi}$	$\mathbf{RMSE}_{\dot{\phi}}$ (1/s)	$\mathbf{RMSE}_{l}$	$\mathbf{RMSE}_{r}$ (deg)
EKF	0.027	0.039	0.126	2.264
UKF	0.021	0.028	0.122	2.199

resetting are the two most effective mechanisms. When the filters lose track, saturation contains the estimates within the regions where the measurement model is trained, and filter resetting brings the estimate back to a state that is in the training dataset. In other words, both approaches prevent the state estimates from entering the subset of the state space that is not covered by the training dataset (i.e., where the measurement model is unpredictable), thereby enhancing the robustness. Comparing the efficacy of resetting to saturation for both the EKF and UKF (see Fig. 5), we observe that 'R' is statistically more effective than 'S' for R_1, whereas 'S' is more effective than 'R' for R_3. Resetting is more effective for R_1 because it resets the estimates to the prior immediately when failure is detected, enabling faster recovery. Saturation is more effective for R_3 because it structurally improves the recoverability, despite slower recovery speeds, by always forcing the state estimates to stay within the region where the measurement model is predictable.

The auxiliary measurement provides non-estimate-dependent information regarding the user's position in the gait cycles, i.e., the gait phase. Although it is not as effective as the other two mechanisms when it is applied alone, it does significantly further enhance the robustness when applied with saturation and/or resetting. For both the EKF and UKF, the statistical robustness values resulting from 'A+S' and 'A+R' are both significantly higher than those yielded by only 'S' and 'R' (Fig. 5), corroborating the effectiveness of the auxiliary measurement as a robustifying mechanism. Most importantly, without the auxiliary measurement, our EKF could not achieve its almost-100% recovery rate.

Among the three pairs of robustifying mechanisms ('A+S', 'S+R', and 'A+R'), 'A+S' demonstrates the highest R_3 for both the EKF and UKF. This is because both the auxiliary measurement and saturation structurally enhance the recoverability of the estimators by giving unambiguous gait information and limiting the state estimates in the predictable region. This is a property that resetting does not have.

Comparing the performance of the EKF and UKF, it has been shown in Section IV-A that they yield comparable levels of RMSEs in nominal tests. This is aligned with [12], [31]. In the kidnapping tests, however, the EKF demonstrates generally higher robustness than the UKF, indicating that our robustifying techniques are more useful for the EKF than the UKF under our parameter settings. Moreover, the EKF takes less average computation time than the UKF in our simulations (EKF's 1.4 ms/time step vs. UKF's 2.4 ms/time step), which is aligned with [25], [31]. Given that the EKF has comparable RMSEs to the UKF in nominal cases, better robustness than the UKF, and faster computation, we may conclude that the EKF is a better choice for gait-state estimation than the UKF.

This study utilizes the Monte Carlo-based kidnapping test as a metric to examine the gait-state estimators' robustness against loss of tracking in simulations. The kidnapping test enables random exploration of disturbed states in the estimator's state space (see (11)), and estimates the percentage of them that would converge back to the correct walking cycle. In this way, it provides a thorough evaluation of the estimator's ability to recover from loss-of-tracking scenarios.

This study focuses only on the robustness against loss-of-tracking scenarios caused by unexpected disturbances as mentioned in Section IV-B. Some other aspects of robustness are not captured by the kidnapping test. For example, the kidnapping test does not clearly quantify the robustness against model uncertainty, such as a new user or a slightly different task. Robustness against such uncertainty can be measured by leave-one-out cross-validation [11]. Kidnapping is also no stand-in for hardware reliability testing for sensor errors, actuation failures, or computational delays though it can help design systems that can recover quickly from these problems if they are intermittent.

This study provides a general framework to empirically quantify the robustness of a Bayesian-inference-based gait estimator in simulation. In this work, the gait estimator considers a four-dimensional gait-state vector to be estimated and the global thigh and foot motion as measurements. However, the proposed kidnapping tests and the robustifying mechanisms are not necessarily limited to this particular setting. For instance, additional measurements, such as global hip motion or ankle moments, could be appended to the measurement model. In addition, non-able-bodied gait data, such as from symptomatic gait or gait modified by use of a wearable robot, could also be explicitly evaluated for robustness using this framework, as future work. Gaits that are very similar to able-bodied gait could reasonably be expected to have similar levels of R_1 and R_3 performance once the robustness-enhancing features are applied.

In terms of application, the type of robustness this paper focuses on is key for the practical applications of gait-state estimation in lower-limb wearable robots. While some wearable robot applications, for example, lightweight backdrivable exoskeletons for augmentation of healthy people, can tolerate occasionally incorrect torques, applications like powered prosthetic legs or high torque exoskeletons for people with partial paralysis require high reliability in control to ensure safety. The goal of developing high-reliability gait-state estimation systems is to allow the benefits of task and phase inference to translate into these systems.

VIII. CONCLUSIONS

In this work, we studied the robustness of two Bayesian-inference-based gait-state estimators (EKF and UKF). In terms of the RMSE metric, both the EKF and the UKF were tracking the ground truth satisfactorily with low RMSEs. However, both filters were likely to lose track of the gait-state under the Monte Carlo-based kidnapping tests. Therefore, we proposed three robustifying mechanisms: 1) introducing the phase angle

of thigh motion as an auxiliary measurement, 2) state estimate saturation, and 3) resetting the filter based on failure detection. The results showed that all three mechanisms significantly improved the robustness of the EKF and the UKF without noticeably degrading the RMSE tracking performance. In particular, the robustness of the EKF with all three mechanisms recovered from unexpected tracking disturbances within one step with 99% reliability.

APPENDIX

The auxiliary measurement (the phase estimate from [14]) is generated from the global thigh angle and velocity as

$$\varphi_a(t) = \text{atan2}\bigg(-\Big(\dot{\theta}_{th}(t) + d(t)\Big),\ \Gamma(t)\Big(\theta_{th}(t) + c(t)\Big)\bigg), \tag{17}$$

where $\varphi_a \in [0,\ 2\pi)$, $c(t)$ and $d(t)$ are shifting parameters, and $\Gamma(t)$ is a scaling parameter. The shifting and scaling parameters are given by

$$\begin{aligned} c(t) = -\frac{\overline{\theta_{th}}(t) + \underline{\theta_{th}}(t)}{2},\ d(t) = -\frac{\overline{\dot{\theta}_{th}}(t) + \underline{\dot{\theta}_{th}}(t)}{2}, \\ \Gamma(t) = \frac{|\overline{\dot{\theta}_{th}}(t) - \underline{\dot{\theta}_{th}}(t)|}{|\overline{\theta_{th}}(t) - \underline{\theta_{th}}(t)|}, \end{aligned} \tag{18}$$

where $\overline{(.)}$ and $\underline{(.)}$ denote the maximum and minimum values in the previous stride.

REFERENCES

[1] D. Quintero, D. J. Villarreal, D. J. Lambert, S. Kapp, and R. D. Gregg, "Continuous-phase control of a powered knee–ankle prosthesis: Amputee experiments across speeds and inclines," *IEEE Transactions on Robotics*, vol. 34, no. 3, pp. 686–701, 2018.

[2] D. J. Villarreal, D. Quintero, and R. D. Gregg, "Piecewise and unified phase variables in the control of a powered prosthetic leg," in *Int. Conf. Rehab. Robot. (ICORR)*. IEEE, 2017, pp. 1425–1430.

[3] M. A. Holgate, T. G. Sugar, and A. W. Bohler, "A novel control algorithm for wearable robotics using phase plane invariants," in *Int. Conf. Robot. Automat. (ICRA)*. IEEE, 2009, pp. 3845–3850.

[4] H. Huang, F. Zhang, L. J. Hargrove, Z. Dou, D. R. Rogers, and K. B. Englehart, "Continuous locomotion-mode identification for prosthetic legs based on neuromuscular–mechanical fusion," *Trans. Biomed. Eng.*, vol. 58, no. 10, pp. 2867–2875, 2011.

[5] A. J. Young, A. M. Simon, N. P. Fey, and L. J. Hargrove, "Classifying the intent of novel users during human locomotion using powered lower limb prostheses," in *Int. IEEE/EMBS Conf. Neural Eng. (NER)*, 2013, pp. 311–314.

[6] A. J. Young and L. J. Hargrove, "A classification method for user-independent intent recognition for transfemoral amputees using powered lower limb prostheses," *Trans. Neural Sys. Rehab. Eng.*, vol. 24, no. 2, pp. 217–225, 2015.

[7] M. Liu, F. Zhang, and H. Huang, "An adaptive classification strategy for reliable locomotion mode recognition," *Sensors*, vol. 17, no. 9, p. 2020, 2017.

[8] B. Hu, A. M. Simon, and L. Hargrove, "Deep generative models with data augmentation to learn robust representations of movement intention for powered leg prostheses," *Trans. Med. Robot. Bionics*, vol. 1, no. 4, pp. 267–278, 2019.

[9] B.-Y. Su, J. Wang, S.-Q. Liu, M. Sheng, J. Jiang, and K. Xiang, "A cnn-based method for intent recognition using inertial measurement units and intelligent lower limb prosthesis," *Trans. Neural Sys. Rehab. Eng.*, vol. 27, no. 5, pp. 1032–1042, 2019.

[10] N. Thatte, T. Shah, and H. Geyer, "Robust and adaptive lower limb prosthesis stance control via extended kalman filter-based gait phase estimation," *Robot. Automat. Lett.*, vol. 4, no. 4, pp. 3129–3136, 2019.

[11] R. L. Medrano, G. C. Thomas, C. G. Keais, E. J. Rouse, and R. D. Gregg, "Real-time gait phase and task estimation for controlling a powered ankle exoskeleton on extremely uneven terrain," *Trans. Robot.*, vol. 39, no. 3, pp. 2170–2182, 2023.

[12] R. L. Medrano, G. C. Thomas, E. J. Rouse, and R. D. Gregg, "Analysis of the bayesian gait-state estimation problem for lower-limb wearable robot sensor configurations," *Robot. Automat. Lett.*, vol. 7, no. 3, pp. 7463–7470, 2022.

[13] C. L. Lewis and D. P. Ferris, "Invariant hip moment pattern while walking with a robotic hip exoskeleton," *J. Biomech.*, vol. 44, no. 5, pp. 789–793, 2011.

[14] D. Quintero, D. J. Lambert, D. J. Villarreal, and R. D. Gregg, "Real-time continuous gait phase and speed estimation from a single sensor," in *Conf. Control Tech. Appl. (CCTA)*. IEEE, 2017, pp. 847–852.

[15] D. J. Villarreal, H. A. Poonawala, and R. D. Gregg, "A robust parameterization of human gait patterns across phase-shifting perturbations," *Trans. Neural Sys. Rehab. Eng.*, vol. 25, no. 3, pp. 265–278, 2016.

[16] J. Hannink, T. Kautz, C. F. Pasluosta, K.-G. Gaßmann, J. Klucken, and B. M. Eskofier, "Sensor-based gait parameter extraction with deep convolutional neural networks," *J. Biomed. Health Informatics*, vol. 21, no. 1, pp. 85–93, 2016.

[17] H. Jin, I. Kang, G. Choi, D. D. Molinaro, and A. J. Young, "Wearable sensor-based step length estimation during overground locomotion using a deep convolutional neural network," in *Int. Conf. Eng. Med. Bio. Soc. (EMBC)*. IEEE, 2021, pp. 4897–4900.

[18] W. Choi, W. Yang, J. Na, J. Park, G. Lee, and W. Nam, "Unsupervised gait phase estimation with domain-adversarial neural network and adaptive window," *J. Biomed. Health Informat.*, vol. 26, no. 7, pp. 3373–3384, 2021.

[19] I. Kang, D. D. Molinaro, S. Duggal, Y. Chen, P. Kunapuli, and A. J. Young, "Real-time gait phase estimation for robotic hip exoskeleton control during multimodal locomotion," *Robot. Automat. Lett.*, vol. 6, no. 2, pp. 3491–3497, 2021.

[20] J. Lee, W. Hong, and P. Hur, "Continuous gait phase estimation using lstm for robotic transfemoral prosthesis across walking speeds," *Trans. Neural Sys. Rehab. Eng.*, vol. 29, pp. 1470–1477, 2021.

[21] M. K. Shepherd, D. D. Molinaro, G. S. Sawicki, and A. J. Young, "Deep learning enables exoboot control to augment variable-speed walking," *Robot. Automat. Lett.*, vol. 7, no. 2, pp. 3571–3577, 2022.

[22] K. Seo, Y. J. Park, J. Lee, S. Hyung, M. Lee, J. Kim, H. Choi, and Y. Shim, "Rnn-based on-line continuous gait phase estimation from shank-mounted imus to control ankle exoskeletons," in *Int. Conf. Rehab. Robot. (ICORR)*. IEEE, 2019, pp. 809–815.

[23] M. Shushtari, H. Dinovitzer, J. Weng, and A. Arami, "Ultra-robust real-time estimation of gait phase," *Trans. Neural Sys. Rehab. Eng.*, vol. 30, pp. 2793–2801, 2022.

[24] E. A. Wan and R. Van Der Merwe, "The unscented kalman filter for nonlinear estimation," in *Adapt. Sys. Sig. Proc. Comm. Con. Symp.* IEEE, 2000, pp. 153–158.

[25] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. MIT Press, 2005.

[26] G. Clark, J. Campbell, S. M. R. Sorkhabadi, W. Zhang, and H. B. Amor, "Predictive modeling of periodic behavior for human-robot symbiotic walking," in *Int. Conf. Robot. Automat. (ICRA)*. IEEE, 2020, pp. 7599–7605.

[27] K. R. Embry, D. J. Villarreal, R. L. Macaluso, and R. D. Gregg, "Modeling the kinematics of human locomotion over continuously varying speeds and inclines," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 12, pp. 2342–2350, 2018.

[28] E. Reznick, K. R. Embry, R. Neuman, E. Bolívar-Nieto, N. P. Fey, and R. D. Gregg, "Lower-limb kinematics and kinetics during continuously varying human locomotion," *Scientific Data*, vol. 8, no. 1, p. 282, 2021.

[29] J. Camargo, A. Ramanathan, W. Flanagan, and A. Young, "A comprehensive, open-source dataset of lower limb biomechanics in multiple conditions of stairs, ramps, and level-ground ambulation and transitions," *J. Biomech.*, vol. 119, p. 110320, 2021.

[30] S. P. Engelson and D. V. McDermott, "Error correction in mobile robot map learning," in *Int. Conf. Robot. Automat.* IEEE, 1992, pp. 2555–2556.

[31] J. LaViola, "A comparison of unscented and extended kalman filtering for estimating quaternion motion," in *Proceedings of the 2003 American Control Conference, 2003.*, vol. 3, 2003, pp. 2435–2440 vol.3.