Effects of Personalization on Gait-State Tracking Performance Using Extended Kalman Filters

José A. Montes-Pérez, Gray Cortright Thomas, and Robert D. Gregg

***Abstract*— Emerging partial-assistance exoskeletons can enhance able-bodied performance and aid people with pathological gait or age-related immobility. However, every person walks differently, which makes it difficult to directly compute assistance torques from joint kinematics. Gait-state estimation-based controllers use phase (normalized stride time) and task variables (e.g., stride length and ground inclination) to parameterize the joint torques. Using kinematic models that depend on the gait-state, prior work has used an Extended Kalman filter (EKF) to estimate the gait-state online. However, this EKF suffered from kinematic errors since it used a subject-independent measurement model, and it is still unknown how personalization of this measurement model would reduce gait-state tracking error. This paper quantifies how much gait-state tracking improvement a personalized measurement model can have over a subject-independent measurement model when using an EKF-based gait-state estimator. Since the EKF performance depends on the measurement model covariance matrix, we tested on multiple different tuning parameters. Across reasonable values of tuning parameters that resulted in good performance, personalization improved estimation error on average by 8.5 $\pm$ 13.8% for phase (mean $\pm$ standard deviation), 27.2 $\pm$ 8.1% for stride length, and 10.5 $\pm$ 13.5% for ground inclination. These findings support the hypothesis that personalization of the measurement model significantly improves gait-state estimation performance in EKF based gait-state tracking ($P \ll 0.05$), which could ultimately enable reliable responses to faster human gait changes.**

I. Introduction

Partial-assistance exoskeletons have the potential to aid varied populations of people to regain their mobility, augment their physical capabilities, and prevent workplace injuries. To accomplish this, exoskeletons need to synchronize their assistance to the user in whatever task they are doing. Many exoskeleton controllers use lower-limb kinematics to determine how much torque to apply to the joints. However, every person walks differently and the tasks (e.g., inclines or stairs) all have different patterns of motion. This difficulty has motivated a variety of exoskeleton control approaches.

Human-in-the-loop optimization can adapt assistance to each user over time, improving walking speed [1], [2], reducing muscle activation [3], and lowering the metabolic cost of walking [4]–[6]. Ref. [1] has shown human-in-the-loop optimization that fine-tunes torques online at varied walking speeds and fixed ground inclinations. However, torque profiles are only timing and velocity based, requiring online re-optimization for changing ground inclinations. This is problematic since optimization can take over 30 minutes to converge and tasks can change quickly. Timing-based torque profiles are less robust to speed and gait initiation/termination changes as they update walking speed estimates only at heel strike. Proper torque timing is crucial for metabolic cost reduction [5]–[7]. Therefore, quick task transitions without re-optimization necessitate a more elaborate gait-state estimator containing task information.

An emerging category of exoskeleton controllers are designed to be both task- and person-independent, such as energy shaping [8] methods. These controllers directly use joint kinematics to calculate assistive torques. They have been shown to reduce muscle activation across varied activities such as sit-to-stand, level ground walking, and stair climbing [9], but "one-controller-fits-all" strategies are limited in how accurately they can predict the appropriate torque across tasks and users. This can potentially force the user to adapt to the assistance, and not the other way around.

An alternative control philosophy parameterizes the assistive joint torques based on explicit estimates of gait phase and task variables. Deep and Convolutional Neural Networks (CNN) can estimate user phase (normalized stride time) [10]–[12] and ground inclination [13] directly from sensor measurements. Additionally, CNNs track phase robustly across various incline and speed conditions [12]. However, black box estimators lack interpretability and verifiability, which makes their behavior hard to predict—both for engineers and users—in new situations such as unseen users. This has led to some interest in personalizing neural-network-based estimators using subject-specific data, either learning an estimator that is completely subject-specific or an adjustment of an existing model trained on more subjects via fine-tuning or transfer learning [13], [14]. Personalized adjustments result in better phase tracking performance [14], at least within level ground walking. Personalized multi-task estimation is known to improve ground inclination tracking on ramps [13], but only when it is estimated in the stance phase (as detected by a foot contact sensor). In addition, both of these personalization results are complicated by the black-box nature of the estimator and the use of data collected from sensors on a particular exoskeleton, which makes it hard to tell which features are being learned, and how well those features would generalize to different sensors and exoskeletons.

Extended Kalman Filters offers a white-box alternative for quickly tracking gait behavior by leveraging apriori information about gait into a measurement model. While

This work was supported by the National Institute of Biomedical Imaging and Bioengineering of the NIH under Award Number R01EB031166. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

All authors are with the Department of Robotics, University of Michigan, Ann Arbor, MI 48109 USA. {jmontp,gcthomas,rdgregg}@umich.edu

the approach was originally introduced to estimate phase and phase rate [15], it has since been extended to include continuous task variables like ramp angle and stride length in its state definition (which we will refer to as the gait-state) [16], [17]. The addition of task variables improved phase prediction performance and enabled the adaptation of control torques as the task quickly varied [16]. These measurement models continue a research effort to measure [18] and model [19] the kinematics of human gait using continuous rather than discrete models. However, personalization is known to be an important source of error in continuous gait models [20], and prior work has yet to attempt personalization of the EKF's measurement model, relying instead on a subject-independent model to estimate the relationship between the joint angles and the gait-state. This was likely responsible for some of the persistent task estimation errors produced by the algorithms in leave-one-subject-out cross-validation [16].

This paper investigates how personalized, subject-specific measurement models influence the performance of EKF-based gait-state estimation. In particular, the contributions of this paper are 1) a statistical comparison of a personalized measurement model (PM) and an un-personalized leave-one-out cross-validation measurement model (UM), and 2) a sensitivity analysis of this comparison to the tuning parameters of the EKF estimator, namely the process model covariance matrix. The testing is based on an existing dataset of multi-speed, multi-ramp walking kinematics for n=10 able-bodied subjects [21]. Our results indicate that phase, ramp, and stride-length estimation are significantly improved by personalization ($P \ll 0.05$). This result held for a wide range of "reasonable" EKF tunings, indicating a lack of sensitivity to the tuning parameters. As a worst-case comparison, we looked at the best EKF tuning for a UM and compared the performance of the same tuning for the PM. The PM improved performance by 20.4% for stride length and 12.7% for ground inclination, with a small, 2.0% tradeoff in performance for phase tracking.

II. Method of State Estimation

A. Review of Extended Kalman Filter

The Extended Kalman Filter (EKF) is a type of Bayes Filter that is used to estimate a hidden state that is not directly observable from sensors. We define a hidden state (gait-state) corresponding to the measurement model which describes gait continuously across tasks as in [16], [17]. Specifically, the gait-state $x(t)$ is defined as

$$x(t) = \begin{bmatrix} p(t) & \dot{p}(t) & l(t) & r(t) \end{bmatrix}^T, \tag{1}$$

comprising phase p, phase rate $\dot{p}$, stride length l, and ground (or ramp) inclination r.

The EKF produces estimates according to a difference equation starting from an initial state, x_0. At each time-step thereafter, it uses the current state to predict the gait-state at the next time-step using a dynamic model (Sec. II-B). Feeding this gait-state into a measurement model (Sec. II-C, a continuous model of the kinematics of human locomotion) the filter attempts to predict the system's measurements (in our case, the global thigh, shank, and foot angles). The EKF algorithm then calculates a Kalman gain that it uses to adjust the predicted gait-state based on the error between the predicted angles and the measured angles.

B. Gait Dynamic Model

We define the state dynamics in discrete time, similar to the dynamics used in the gait-state estimation literature [15]–[17], as

$$x_{k+1} = Fx_k + \omega_Q, \quad \omega_Q \sim \mathcal{N}(0, \Sigma_Q), \tag{2}$$

where x_{k+1} is the predicted state at the next time-step, ω_Q is a multivariate normal random vector with mean 0 and Σ_Q covariance, Δt is the time-step, and the matrices F and Σ_Q are defined

$$F = \begin{bmatrix} 1 & \Delta t & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}, \; \Sigma_Q = \begin{bmatrix} \sigma^2_{Q,p} & & & \\ & \sigma^2_{Q,\dot{p}} & & \\ & & \sigma^2_{Q,l} & \\ & & & \sigma^2_{Q,r} \end{bmatrix}. \tag{3}$$

We term Σ_Q the process model noise covariance matrix and define $\sigma^2_{Q,p}$, $\sigma^2_{Q,\dot{p}}$, $\sigma^2_{Q,l}$, $\sigma^2_{Q,r}$ to correspond to independent process noise in phase, phase rate, stride length, and ramp, respectively.

The dynamic update matrix, F, updates the phase estimate according to a simple Euler integration of the phase rate. The evolution of the other states cannot be predicted since stride length depends on user choice, and ground inclination depends on the environment. Therefore, phase rate, stride length, and ground inclination only increase their uncertainty in the prediction step by the corresponding elements in the covariance matrix Σ_Q. This matrix is consequently used as a tuning matrix which determines how fast the states can change between consequent time steps. Similar to [15], we use a noiseless Euler-integration of phase (i.e. $\sigma^2_{Q,p} = 0$). Since we do not expect the states to be correlated with one another, the matrix is defined with a diagonal structure.

C. Measurement Model

The measurement model maps gait-state to expected measurements. For this paper, the measurement model will predict link angles and link angular velocities. In particular, we will use global shank, global thigh, and global foot angles and velocities. We use global link angles instead of joint angles since it is easier to obtain reliable information from IMUs than from joint encoders. The measurement model that predicts link angles and velocities based on the current state is therefore

$$h(x) = \begin{bmatrix} \theta(x) \\ \dot{\theta}(x) \end{bmatrix} + \omega_R, \quad \omega_R \sim \mathcal{N}(0, \Sigma_R), \tag{4}$$

which is a vector in $\mathbb{R}^6$, and ω_R is a multivariate normal random vector with 0 mean and Σ_R covariance. The link

angle model is

$$\theta(x) = \begin{bmatrix} \theta_{\text{thigh}}(x) \\ \theta_{\text{shank}}(x) \\ \theta_{\text{foot}}(x) \end{bmatrix} = \begin{bmatrix} \lambda(x) & & \\ & \lambda(x) & \\ & & \lambda(x) \end{bmatrix} \begin{bmatrix} \xi_{\text{thigh}} \\ \xi_{\text{shank}} \\ \xi_{\text{foot}} \end{bmatrix}, \tag{5}$$

where $\theta(x)$ is a vector in $\mathbb{R}^3$ of link angles, each θ_i represents a specific link, $\lambda(x)$ is a nonlinear model vector of the state (which is the same across all links), and ξ_i represents the model fit for a particular link. To calculate the corresponding link angle velocities, we have that

$$\dot{\theta}(x) = \begin{bmatrix} \dot{\theta}_{\text{thigh}}(x) \\ \dot{\theta}_{\text{shank}}(x) \\ \dot{\theta}_{\text{foot}}(x) \end{bmatrix} = \begin{bmatrix} \dot{\lambda}(x) & & \\ & \dot{\lambda}(x) & \\ & & \dot{\lambda}(x) \end{bmatrix} \begin{bmatrix} \xi_{\text{thigh}} \\ \xi_{\text{shank}} \\ \xi_{\text{foot}} \end{bmatrix}, \tag{6}$$

where the model is just the time derivative of θ. Finally, the measurement model covariance matrix Σ_R is defined as a 6×6 matrix with only diagonal entries. The $\sigma^2_{R,\text{thigh}}$, $\sigma^2_{R,\text{shank}}$, and $\sigma^2_{R,\text{foot}}$ are the first three diagonal components correspond to the link angles, and $\sigma^2_{R,\dot{\text{thigh}}}$, $\sigma^2_{R,\dot{\text{shank}}}$, and $\sigma^2_{R,\dot{\text{foot}}}$ are the last three components correspond to the link velocities.

D. Link Kinematic Model

To model the link kinematics as functions of the gait-state, we decided to use a linear model which consist of Fourier basis functions that are scaled by the tasks variables:

$$\begin{aligned} \lambda = \big[& f(p) \quad \dot{p}f(p) \quad rf(p) \quad lf(p) \\ & \dot{p}rf(p) \quad \dot{p}lf(p) \quad rlf(p) \big]. \end{aligned} \tag{7}$$

Here, $f(p)$ corresponds to the 20th order Fourier series basis

$$f(p) = \begin{bmatrix} 1 & \sin(\pi p) & \cos(\pi p) & \dots & \sin(20\pi x) & \cos(20\pi x) \end{bmatrix}. \tag{8}$$

We used a linear model of task and phase since these have been shown to be effective in [16], [21]. We decided against incorporating higher-order polynomial terms for task variables, anticipating that it might lead to overfitting and potentially create erratic model predictions in areas lacking prior training data.

E. Estimating the Measurement Model

1) Fitting Particular Individuals: The measurement model fits, ξ_i, are used to differentiate the link models from each other and between subjects. Least squares is used to calculate the personalized model (PM) for subject i:

$$\vec{\theta}_{k,i} = \Lambda(x_i)\xi_{k,i}, \tag{9}$$

Here, $\vec{\theta}_{k,i}$ is a $d_i \times 1$ vector of the k^{th} link angles, d_i represents the number of data points for the i^{th} subject, x_i represents their gait-state corresponding to the link angle in the same row, $\Lambda(x_i)$ is a $d_i \times m$ matrix, and finally $\xi_{\text{joint},i}$ is a $m \times 1$ vector representing the unknown model fit. Since the model fit vector $\xi_{\text{link},i}$ is the only unknown in this equation, LS can be used to solve for it. However, some task conditions in the dataset have less information than others (due to subjects taking fewer steps), so we want to make sure that these are weighted properly to not bias the model in the tasks that have more data points. Therefore, we weight each data point with the inverse of the number of data points for the given speed and ground inclination condition. All the weights are collected into a square matrix W. With this weighting in mind, the solution to the weighted least squares problem is

$$\xi_{\text{link},i} = \Lambda^+_W(x_i)\theta_{\text{link},i}, \tag{10}$$

where $\Lambda(x_i)^+$ is the weighted Moore–Penrose inverse.

2) Estimating the Average Person's Gait: To create a link angle model that can be used in the absence of subject-specific data, Medrano et al. use the inter-subject average model. This model is calculated by taking the mean of the model fits in a dataset,

$$\bar{\xi} = \frac{1}{n}\sum_{i=1}^{n} \xi_i. \tag{11}$$

However, since we want to understand how well $\bar{\xi}$ estimates a subject that it has not seen before, simulations will leave out the validation subject's data from the inter-subject average calculation. This leave-one-out inter-subject average model will be referred to as the un-personalized model (UM).

3) Setting the Measurement Model Covariance Matrix: The measurement model covariance matrix represents how error sensor measurements can have. A proper calibration would capture how much variance there are in kinematics across all humans for a given gait-state. We estimate this quantity as the average error for each link across subjects. That is, for link angle k,

$$\sigma^2_{R,k} = \frac{1}{10}\sum_{i=1}^{10} ||\theta_{k,i} - \Lambda^+(x_i)\xi_{k,i}||^2_2, \tag{12}$$

where $\sigma^2_{R,k}$ represents the covariance for the k^{th} link angle, and $||.||_2$ represents the L2 norm. The residual is averaged across all subjects. The link velocities share their corresponding link angle covariance $\sigma^2_{\dot{R},k} = \sigma^2_{R,k}$. To keep the meaning of the EKF tuning parameters consistent between subjects and models, we defined this matrix once and used it in all tests, rather than re-generating it according to leave-one-out cross-validation.

4) Constraining the Gait-State: Since the measurement model fitting does not contain data for very slow walking, very fast walking, or inclines at very steep angles, we decided to limit the state of the EKF. This is because when the link angle model has to extrapolate too far, it could provide nonsensical link angles and velocities, resulting in improper state predictions. Therefore, we constrained the phase rate to be above 0.6, stride length above 0.5 m, and ground inclination between ± 15 deg. The dataset had ground truth information for ground inclines between ± 10 deg, so this bound does permit some extrapolation in inclination.

F. Difference from Previous Work

Our labs previous work [16] has some key differences from our EKF implementation by having a different selection of basis functions for the measurement model, a different constrained least-squares model fitting, a saturating and non-linear transformation in the stride length state definition, and a heteroscedastic measurement model covariance matrix. We opted for a simpler EKF implementation because it has sufficient structure so that we can test our hypothesis about personalization, while not being overly complex.

III. Simulation Methods

A. Walking Dataset

An existing able-bodied walking dataset [21] was used to fit the models and validate their state-tracking performance. This dataset consists of kinematic and kinetic data for 10 able-bodied human subjects walking at combinations of 9 ramp inclines and 3 walking speeds (therefore 27 ramp/speed pairs). It also contains the ground truth phase, phase rate, ground inclination, and joint kinematics needed for the model fitting procedure and the EKF simulation. Stride length was calculated as the distance between foot placements at toe-off and heel strike; gait events were determined from the Bertec treadmill's force plate data. The dataset was divided into a training set and a validation set (80/20% split) so we could test the EKF's performance on untrained data. Because strides in the dataset were time-normalized, phase rate was always constant.

To better extrapolate to low-speed walking tasks that were not included in the dataset, we artificially created 9 steps worth of data to represent standing still (one 'step' per unique ground inclination in the dataset). The standing thigh and shank angles were taken from the average walking data across all users and conditions while at mid-stance. The foot angle corresponded to the ground inclination. The phase rate and stride length were both set to zero.

B. EKF Simulation

We hypothesized that the EKF would have better gait-state tracking performance with a PM than with a UM. To test this, we performed an EKF simulation with both the PM and UM. The EKF was simulated by playing back the joint angles and joint velocities in the validation dataset as measurement inputs to the EKF equations (see Appendix) to generate a new prediction of the gait state. The initial state was slightly different from the true initial state (to validate that the EKF can correct for that difference). The initial covariance was set to a constant value of 1e-5 for all the states. We empirically found that this allowed the state to account for initial errors without being large enough to reduce tracking performance.

We then compared the estimates of the gait-states to the ground truth gait-state for each time step. This allowed us to estimate the performance of the PM and UM based on unseen data for each subject. The metric used to quantify improvement was the root-mean-squared error (RMSE) between the ground truth gait-state and the estimated gait-state (the lower the error, the better). Although phase rate is important for the gait-state dynamics, it is not necessary for this analysis because it is not directly used for exoskeleton control. Therefore, we focused on the RMSE of phase, stride length, and ground inclination.

To simulate ground inclination and walking speed changes in real-time, we created the simulation dataset by splicing together all the speeds and ground inclinations from the validation dataset. Since we spliced data from different ramp and speed conditions together, the link angles and ground truth measurements are not necessarily continuous as they would be in real life. This discontinuity could cause the EKF to behave in a way that is not representative of real-world implementations. To counteract this, the measurement model covariance was boosted by a factor of 20 for link angles, and 100 for link velocities when the estimated phase is greater than 0.95 but lower than 0.05. This allowed the EKF to rely less on the measurements near the end and the start of steps to ignore the hard discontinuities. The metrics still relate to the EKF's ability to track quick changes between tasks, and in a dataset with continuous transitions between tasks, this feature would not be necessary.

To avoid potentially biasing the results by hand-picking the order of the ground incline and speed conditions, we randomly generated a set of ramp and speed conditions for the validation dataset. This is to avoid, for example, using all the ramp conditions in ascending order, which would unfairly remove the significance of quickly adapting to changing ramp conditions. The same ordering was used for all subjects in the training dataset. For each subject, 10 steps were included for each condition. If a subject did not have enough steps for a given condition, data was repeated.

C. Phase Error Calculation

Since phase is defined on the unit circle, we apply an alternative difference measure for phase to handle the discontinuity at $0 = 1$. For example, if the predicted phase is 0.99 and the actual phase is 0.01, the appropriate error is 0.02, not 0.98. We defined this difference measure as $\Delta p = p_{\text{actual}} - p_{\text{predicted}}$,

$$p_{\text{difference}} = \begin{cases} \Delta p, & \text{if} \quad \Delta p < 0.5 \\ 1 - \Delta p, & \text{else} \end{cases}. \tag{13}$$

This calculated $p_{\text{difference}}$ is what is used to calculate the RMSE difference.

D. Process Model Covariance Matrix Tuning

Empirically, the selection of the process model covariance matrix (i.e., the EKF tuning) notably influences the test results. To test our hypothesis to the greatest possible generality, we investigated it both for a wide range of viable tuning parameters, and in particular for one set of tuning parameters that was approximately optimal according to a coarse grid search. While properly optimizing the tuning would have been ideal, the EKF simulation takes long enough that this would be an impractical optimization. We started with an exploratory analysis to determine a proper process model covariance tuning range without worrying too much about

global optimality. Thus the grid search was broken into two stages: first determining a "reasonable" range of process model covariance tunings for the EKF (using a very course grid), and secondly, within the range of "reasonable" values, finding the "optimal" tuning parameters using a denser grid.

We defined a cost for each EKF tuning based on the performance of that tuning when using the UM. This cost was a weighted sum of RMSE values for the three primary gait-states:

$$cost = (\frac{p_{\text{RMSE}}}{p_{\text{RMSE,max}}})^2 + (\frac{l_{\text{RMSE}}}{l_{\text{RMSE,max}}})^2 + (\frac{r_{\text{RMSE}}}{r_{\text{RMSE,max}}})^2. \quad (14)$$

The cost function was designed to weight each metric by the maximum error for the UM across all the tuning parameters in the "reasonable values" to be able to compare the best overall performance improvement.

A 3D grid of process model covariance matrices was created for gait parameters (phase rate, stride length, ground inclination), sampled sparsely across multiple orders of magnitude. Considering no correlation between states, only diagonal matrices were evaluated. After 505 simulations, the optimal region was identified through visual inspection of error metrics. The lowest Phase RMSE was observed at phase rates of 1e-6 or 1e-7, irrespective of stride length or ramp tuning parameters. Stride length RMSE was robust to parameter changes, but increased if stride length noise exceeded 1e-6. Ground incline RMSE was sensitive to parameters, best at phase rate 1e-4, stride length 1e-7, and ramp 1e-5. The "reasonable" region, centered around these optimal values with one order of magnitude allowance in each direction, was further explored with denser linear interpolation of 216 samples to approximate the optimal solution. Based on these observations, the optimal region is selected as phase rate being 1e-6, stride length being in 1e-7, and ramp being 1e-4. The "reasonable" region was then determined to be centered at phase rate: [1e-7, 1e-5], stride length: [1e-8, 1e-5], ramp: [1e-5, 1e-3], with one order of magnitude added and subtracted for each. Inside the "reasonable" region, we performed a denser linear interpolation of 216 samples to find an approximately optimal solution (Fig. 2).

The "optimal" tuning was found by a grid search of the "reasonable" range. We minimized the cost function based only on the UM performance so that we could compare the best UM to its corresponding PM. But, this tuning corresponded to the best PM as well. The resulting "optimal" model tuning was phase rate covariance = 1e-7, stride length covariance = 1e-8, and ramp covariance = 2.1e-4.

IV. Results

To test our hypothesis relating to the effect of personalization on the EKF, we measured the difference in the performance of the PM when compared to the UM for each EKF tuning in the "reasonable" range grid. The distribution of results can be seen in the blue box plot in Fig. 3. Each sample in the box plot represents the RMSE of the difference in performance from UM to PM when averaged across subjects. There is one sample per set of tuning

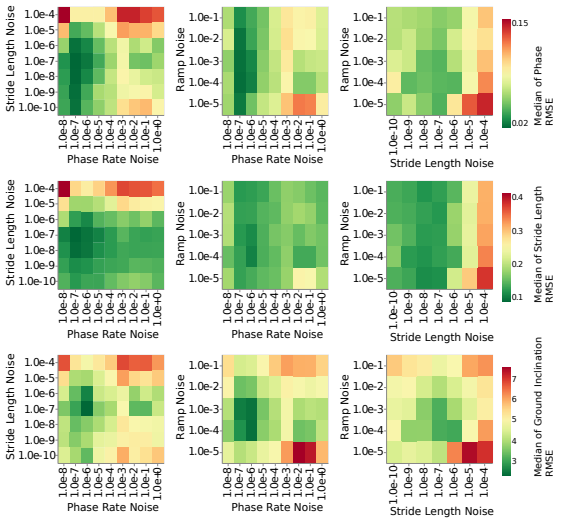

Fig. 1. Heat map of RMSE vs process model covariance tuning parameters. Each row is a different gait-state and every column is a different process noise model covariance matrix tuning parameter.

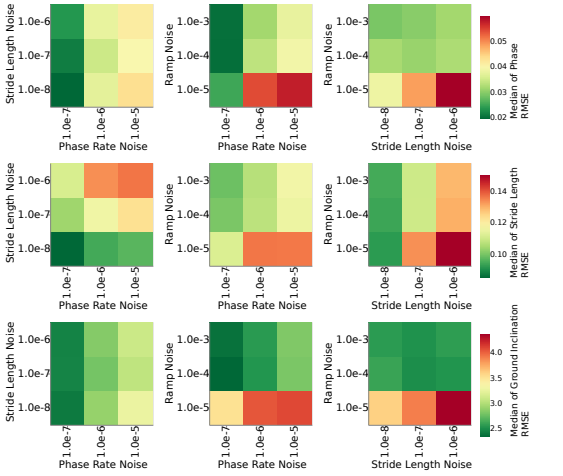

Fig. 2. Heat map of RMSE vs process model covariance tuning parameters for "reasonable" region. Each row is a different gait-state and every column is a different process noise model covariance matrix tuning parameter.

parameters. Since the values represent the UM RMSE minus the PM RMSE, higher values mean that personalization has improved the gait-state tracking performance. This plot has an alternative y-axis label in percentage improvement (right-hand axis), which normalizes this difference by the corresponding median RMSE of the UM method. All the states in the "reasonable difference" had statistically significant improvements according to our naive t-test (Figure 3, $P \ll 0.05$), though since samples with similar tuning parameters are likely correlated it might artificially inflate the confidence of this statistical test. Some process model tuning parameters did results in worse PM phase RMSE and ground incline RMSE, but every stride length RMSE was improved over the baseline.

The average improvement across all the "reasonable tunings" is 8.5 $\pm$ 13.8% for phase (mean $\pm$ standard deviation), 27.2 $\pm$ 8.1% for stride length, and 10.5 $\pm$ 13.5% for ground

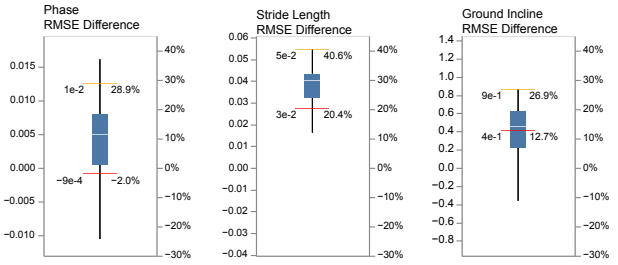

Fig. 3. Pairwise difference of the tests ran in the "reasonable process model covariance matrices" range (n=216, $P \ll 0.005$). Larger numbers represent higher performance improvements due to model personalization. Two special cases are also encoded as horizontal lines: the optimal process model covariance matrix (red), and the most-improved process model covariance matrix (orange).

TABLE I
EKF SIMULATION RESULTS

Measurement Model	Phase RMSE (%)	Stride Length RMSE (m)	Ground Inclination RMSE (°)
UM	2.19 ± 0.68	0.108 ± 0.0337	2.61 ± 0.652
PM	2.28 ± 0.960	0.081 ± 0.0223	2.21 ± 0.701

Gait-state RMSE averaged across all subjects when using the Optimal process model covariance matrix. The data is displayed as mean ± one standard deviation with n=10 datapoints.

inclination. For the optimal process covariance tuning, the RMSE for phase, stride length, and ground inclination is reported in Table I. Figure 3 encodes the relative improvement of the optimal process covariance matrix as a horizontal red line. The stride length estimation and ramp estimations significantly improved (20% and 12% respectively) while phase error was slightly worse (2%). We note that the process model tuning parameter that resulted in the highest improvement in gait-state tracking has 28.9% better phase RMSE, 40.6% better stride length RMSE, and 26.9% better ground inclination RMSE. This is also shown in Figure 3 as a horizontal orange line.

V. DISCUSSION

Our experimental results supported our hypothesis that the EKF would have better gait-state tracking performance with a PM than with a UM. For the "optimal tuning" parameters, where the UM model had the least overall error, the PM improved stride length RMSE by 20.4%, ground inclination RMSE by 12.7%, and slightly decreased phase RMSE by 2%. Even though the phase estimation is slightly worsened, the stride length and ground incline estimation were clearly improved. The improvement in tracking performance is most likely due to the reduced fitting error between the PM and a user's link angles while walking.

More broadly, the distribution in Figure 3 shows that in most cases phase and ground inclination were improved, and in all cases, stride length was improved. Median improvements of 11.5% for phase, 30.0% for stride length, and 14.3% for ground inclination were observed. Therefore, we can see that for sub-optimal tunings, the performance increases are even better. This is perhaps even stronger evidence supporting our hypothesis about the importance of personalization, since finding process model covariance matrix parameters that are optimal by sampling hundreds of tuning parameters is not feasible for online systems.

Other process model covariances yielded improvements as large as 28.9% for phase, 40.6% for stride length, and 26.9% for ground inclination (Fig. 3). These large improvements, which look like outliers, could be a result of the UM model EKF state estimate "getting lost," which is a potential failure mode for EKF systems in which the states become stuck in a local minimum behavior (due to the linearization of the EKF) that does not reflect accurate gait kinematic tracking. Having less fitting error with the PM could mean that it is more robust to "getting lost", which is a phenomenon when the EKF begins to track a different set of gait-states that can describe the observations from the sensors.

The gait-tracking performed comparably to other state of the art approaches using similar sensors and outcomes. Note that direct comparisons cannot be made since benchmark datasets for gait tracking are not yet standardized in the field, so each measure of RMSE refers to a different combination of subjects and tasks. Our PM phase tracking RMSE of 2.28% ± 0.960% compares favorably to the CNN solution in [14], which achieved 5.22% ± 0.81% phase RMSE when changing walking speeds. Our group's previous unpersonalized EKF implementations [16] achieved 1.6% ± 0.31% phase RMSE under different conditions. However, our stride length RMSE of 0.081m ± 0.0223m slightly improved over that implementation's 0.1m ± 0.02m [16]. Our performance on ground incline (2.21° ± 0.701° RMSE) was worse than the previous UM EKF (1.5° ± 0.23°). The difference in performance is likely due to the difference in our methods, which are discussed in Section II-F. Additionally, as is explained in Section III-B, we randomized the order of the treadmill speed and incline transitions. This would make replicating the simulation in [16] one-to-one very unlikely, which can also affect the simulation results. We believe that the personalization performance improvements would generalize to their methods as well.

Relative to a CNN-based ground incline estimator [13], our performance was worse than their personalized version, with the CNN having 0.61° ± 0.05° RMSE vs our PM EKF's 2.21° ± 0.701° RMSE. The same authors also reported results for the un-personalized CNN as 2.15° ± 0.29° RMSE, which is more in line with our PM EKF. While the different methodologies greatly weaken the comparative value of these examples, our gait-state estimator seems close enough to the state-of-the-art that our analysis of the personalization effects remains relevant.

One of the benefits of this algorithm is that it is very easy to add sensors to the measurement function. It is even possible to have a personalized model for some sensors and add an un-personalized model for other sensors, for example, if new sensing modalities (such as ground reaction forces) become more practical. Additionally, it might be possible to update the measurement model online just by swapping out the measurement model between subsequent predictions, presenting an opportunity to perform online personalization

as a subject walks. The EKF is particularly suited for online learning since it not only gives us state estimates but also state covariance. Having access to covariance information allows us to discard measurements that are not trustworthy, or give more weight to more certain measurements.

VI. Conclusion

Inferring what someone is doing from only IMU measurements is naturally sensitive to the model of how these measurements relate to the gait-state. This paper demonstrated that the personalization of these continuous, multi-task gait models significantly improves the performance of EKF-based gait-state estimation systems. However, despite our efforts to be thorough, these results are not guaranteed to generalize to all sensor configurations, sets of tasks, subject populations, and gait-model structures. Our subjects were all able-bodied, our tasks were limited to a finite set of speeds and treadmill inclinations, and the tasks were spliced together rather than continuously transitioned. Our results suggest the personalization of continuous gait-models could improve the performance of more general gait-state estimators. This paper motivates future research directions for EKF-based gait-state estimation systems such as learning the personalized model online, applying the EKF in faster-paced activities such as running, and quantifying how much the user benefits from the improved gait-state tracking.

Appendix

In the notation of Sec. II, the EKF updates its internal states to account for the passage of time using the "prediction" equations found in [22],

$$x_{k|k-1} = F x_{k-1|k-1}, \tag{15}$$

$$P_{k|k-1} = F P_{k-1|k-1} F^T + \Sigma_Q. \tag{16}$$

To account for measurements, its state changes according to the "update" equations

$$\begin{aligned} \tilde{y}_k &= z_k - h(\hat{x}_{k|k-1}), & H_k &= \tfrac{\partial h}{\partial x}|_{x_{k|k-1}} \\ S_k &= H_k P_{k|k-1} H_k^T + \Sigma_R, & K_k &= P_{k|k-1} H_k^T S_k^{-1} \\ x_{k|k} &= x_{k|k-1} + K_k \tilde{y}_k, & P_{k|k} &= (I - K_k H_k) P_{k|k-1}, \end{aligned} \tag{17}$$

where P is the estimate ofthe variance, $\tilde{y}_k$ is the prediction error, S_k is the covariance of the prediction, z_k is the sensor measurements, and K_k is the Kalman gain.

Acknowledgment

The authors thank Roberto Leonardo "Leo" Medrano for many fruitful discussions on EKF-based gait-state estimation.

References

[1] P. Slade, M. J. Kochenderfer, S. L. Delp, and S. H. Collins, "Personalizing exoskeleton assistance while walking in the real world," *Nature*, vol. 610, no. 7931, pp. 277–282, Oct. 2022.

[2] S. Song and S. H. Collins, "Optimizing Exoskeleton Assistance for Faster Self-Selected Walking," *IEEE Trans. Neural Syst. Rehabilitation Eng.*, vol. 29, pp. 786–795, 2021.

[3] X. Tu, M. Li, M. Liu, J. Si, and H. H. Huang, "A Data-Driven Reinforcement Learning Solution Framework for Optimal and Adaptive Personalization of a Hip Exoskeleton," in *2021 IEEE Int. Conf. on Robotics and Automation (ICRA)*, May 2021, pp. 10 610–10 616.

[4] J. Zhang, P. Fiers, K. A. Witte, R. W. Jackson, K. L. Poggensee, C. G. Atkeson, and S. H. Collins, "Human-in-the-loop optimization of exoskeleton assistance during walking," *Science*, vol. 356, no. 6344, pp. 1280–1284, Jun. 2017.

[5] S. Galle, P. Malcolm, S. H. Collins, and D. De Clercq, "Reducing the metabolic cost of walking with an ankle exoskeleton: Interaction between actuation timing and power," *J. Neuroeng. Rehabilitation*, vol. 14, no. 1, p. 35, Apr. 2017.

[6] I. Kang, H. Hsu, and A. Young, "The Effect of Hip Assistance Levels on Human Energetic Cost Using Robotic Hip Exoskeletons," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 430–437, Apr. 2019.

[7] A. J. Young, J. Foss, H. Gannon, and D. P. Ferris, "Influence of Power Delivery Timing on the Energetics and Biomechanics of Humans Wearing a Hip Exoskeleton," *Front. Bioeng. Biotechnol.*, vol. 5, 2017.

[8] J. Lin, N. V. Divekar, G. C. Thomas, and R. D. Gregg, "Optimally biomimetic passivity-based control of a lower-limb exoskeleton over the primary activities of daily life," *IEEE Open Journal of Control Systems*, vol. 1, p. 15–28, 2022.

[9] N. V. Divekar, J. Lin, C. Nesler, S. Borboa, and R. D. Gregg, "A Potential Energy Shaping Controller with Ground Reaction Force Feedback for a Multi-Activity Knee-Ankle Exoskeleton," in *2020 IEEE RAS/EMBS Int. Conf. for Biomedical Robotics and Biomechatronics (BioRob)*, Nov. 2020, pp. 997–1003.

[10] 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 *IEEE Int. Conf. on Rehabilitation Robotics*, Jun. 2019, pp. 809–815.

[11] J. Yang, T.-H. Huang, S. Yu, X. Yang, H. Su, A. M. Spungen, and C.-Y. Tsai, "Machine Learning Based Adaptive Gait Phase Estimation Using Inertial Measurement Sensors," in *ASME Design of Medical Devices Conf.*, Jul. 2019.

[12] 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," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, p. 3491–3497, Apr 2021.

[13] D. Lee, I. Kang, D. D. Molinaro, A. Yu, and A. J. Young, "Real-Time User-Independent Slope Prediction Using Deep Learning for Modulation of Robotic Knee Exoskeleton Assistance," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 3995–4000, Apr. 2021.

[14] I. Kang, P. Kunapuli, and A. J. Young, "Real-time neural network-based gait phase estimation using a robotic hip exoskeleton," *IEEE Trans. Med. Robot. Bionics*, vol. 2, no. 1, p. 28–37, Feb 2020.

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

[16] 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," *IEEE Transactions on Robotics*, vol. 39, no. 3, pp. 2170–2182, 2023.

[17] 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," *IEEE Robot. Autom. Lett.*, vol. 7, no. 3, pp. 7463–7470, Jul. 2022.

[18] 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, pp. 1–12, 2021.

[19] 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 Trans. Neural Syst. Rehabilitation Eng.*, vol. 26, no. 12, pp. 2342–2350, 2018.

[20] K. R. Embry and R. D. Gregg, "Analysis of continuously varying kinematics for prosthetic leg control applications," *IEEE Trans. Neural Syst. Rehabilitation Eng.*, vol. 29, pp. 262–272, 2020.

[21] K. Embry, "The effect of walking incline and speed on human leg kinematics, kinetics, and emg," Oct 2018. [Online]. Available: https://ieee-dataport.org/open-access/effect-walking-incline-and-speed-human-leg-kinematics-kinetics-and-emg

[22] M. I. Ribeiro, "Kalman and extended kalman filters: Concept, derivation and properties," *Institute for Systems and Robotics*, vol. 43, p. 46, 2004.