Compact Maximum Correntropy-Based Error State Kalman Filter for Exoskeleton Orientation Estimation

Shilei Li, Peihu Duan, Dawei Shi, *Senior Member, IEEE*, Wulin Zou, Pu Duan, and Ling Shi, *Senior Member, IEEE*

***Abstract*— This brief investigates the maximum correntropy-based Kalman filtering problem for exoskeleton orientation by fusing signals from accelerometers and gyroscopes. The conventional error state Kalman filter (ESKF) has been applied to many applications for orientation estimation. However, its performance degenerates remarkably with external acceleration. In this brief, the influence of the external acceleration is analyzed and the dilemma of the conventional ESKF is declared. To address this issue, a weighted correntropy and a novel correntropy-induced metric (CIM) are provided. Then, a compact maximum correntropy-based Kalman filter (CMC-KF) is derived based on the proposed metric, which performs well both with and without non-Gaussian noises. Finally, a compact maximum correntropy-based ESKF (CMC-ESKF) is designed for orientation estimation of exoskeletons. A series of experiments are conducted to verify the effectiveness of the proposed method. Results reveal that the proposed algorithm is significantly better than the conventional ESKF and the gradient descent (GD) method, especially with external accelerations.**

***Index Terms*— Inertial measurement units (IMUs), Kalman filter (KF), non-Gaussian noise, orientation estimation, weighted correntropy.**

I. Introduction

INERTIAL measurement units (IMUs) composed of accelerometers and gyroscopes have attracted continuing research efforts for the advantages of portability, low cost, and real-time communication [1], [2]. They have been applied to the fields of rehabilitation [3], gait assessment [4], human–robot interaction [5], motion animation [6], and virtual and augmented reality [7]. However, as one of the most critical issues of IMUs, orientation estimation is challenging and their performances are greatly affected by different algorithms [8]. One important reason is that the gyroscope contains a random bias term, and the accelerometer is sensitive to impact and external acceleration, leading to a nonlinear estimation problem with non-Gaussian noises. Note that a small gyroscope drift can cause an obvious orientation error by long-time integration, while the fast-varying external acceleration can result in an instantaneous estimation error. Therefore, the performances of algorithms largely depend on how they address the gyroscope drift and external acceleration.

Many algorithms have been developed to mitigate the gyroscope drift using the noisy but drift-free accelerometer signals, which can be summarized into three categories: the Kalman filter (KF) [8]–[12], the gradient descent (GD) [2], [13], [14], and the complementary filter [15]–[17]. The most popular method should be the KF, which can be divided into the extended KF (EKF) and the unscented KF (UKF). The EKF [8]–[11] is a common approach for IMUs, which has been widely utilized in many commercial inertial sensors, e.g., MicroStrain, Xsens, Crossbow, and InterSense. The UKF [12] is usually used together with other sensors, such as global positioning system (GPS) or compasses, and it costs more computation resources compared with the EKF. The GD proposed by Madgwick *et al.* [2] has a lower computation complexity compared with the KF, but it does not model the external acceleration and cannot guarantee the steepest descent. The complementary filter [17] designed on the special orthogonal group is more robust, while its performance is worse compared with the KF when without disturbance. Hence, the error state KF (ESKF) [8], [11], a type of EKF, has been a good remedy for IMUs without considering the fast-varying external acceleration.

Many studies report that orientation estimation is vulnerable to external accelerations [18]–[25]. Nevertheless, the explicit influence of the external acceleration is rarely mentioned and analyzed. Previous solutions to the external acceleration include: adding extra sensors, e.g., ultra-wideband (UWB), cameras, and GPS [22]–[24]; estimating the external acceleration using a first-order Markov model [1], [8], [10]; and rearranging the weights of prediction and correction based on the carriers' status [19]. However, these methods all possess some limitations. For example, the additional sensors increase the cost and complexity, and are not always available in some applications. The first-order acceleration assumption relies heavily on the assumed Markov model, which works well for the case of slow-varying kinematics but usually produces unsatisfactory performance for fast-varying

Manuscript received 27 December 2021; revised 31 May 2022; accepted 18 July 2022. The work of Shilei Li, Wulin Zou, and Ling Shi was supported by the Hong Kong Research Grants Council (RGC) General Research Fund under Grant 16206620. The work of Dawei Shi was supported by the National Natural Science Foundation of China under Grant 61973030. Recommended by Associate Editor A. Behal. *(Corresponding author: Peihu Duan.)*

Shilei Li and Wulin Zou are with the Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong, China, and also with the Control Department, Xeno Dynamics Company Ltd., Shenzhen 518055, China (e-mail: slidk@connect.ust.hk; wzouab@connect.ust.hk).

Peihu Duan is with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, China (e-mail: phduan@hku.hk).

Dawei Shi is with the School of Automation, Beijing Institute of Technology, Haidian 100811, China (e-mail: daweishi@bit.edu.cn).

Pu Duan is with the Control Department, Xeno Dynamics Company Ltd., Shenzhen 518055, China (e-mail: duanpu@xeno.com).

Ling Shi is with the Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong, China (e-mail: eesling@ust.hk).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCST.2022.3193760.

Digital Object Identifier 10.1109/TCST.2022.3193760

movements. As for the weight redistribution, it schedules the weights of prediction and correction based on the IMUs movement status, but an explicit acceleration estimation is ignored and the weight-scheduling is usually designed empirically. Aravkin *et al.* [26] reported that the ℓ_0 or the ℓ_1 norm was a better metric for heavy-tailed noise compared with the ℓ_2 norm, while the ℓ_2 norm-based cost function is optimal for Gaussian noise. In our application, non-Gaussian noises are occasionally involved in some channels of the system. Therefore, a time-varying norm-based objective function is preferable. Liu *et al.* [27] reported that the correntropy induced metric (CIM) was a time-varying norm and the scale of the norm can be controlled by the kennel bandwidths. Moreover, Chen *et al.* [28] designed a maximum correntropy KF (MCKF) which exhibited both robustness against heavy-tailed noises and good performance with Gaussian noises. However, the MCKF considers the situation of heavy-tailed noises existing in both the process and the measurement channels, which is not the case of our application where only parts of the channels contain non-Gaussian noises (i.e., the acceleration channel). Therefore, it is desirable to derive a novel correntropy-based algorithm for our application.

A motion tracking system, with an IMU mounted to the shank of an exoskeleton robot, usually contains both impulsive external accelerations and continuous external accelerations [29], [30] when walking. However, few existing approaches can handle this issue. In this brief, we first qualitatively analyze the influence of the external acceleration on the conventional ESKF and reveal that the existing algorithm is incapable of accurate orientation tracking with fast-varying external accelerations. Then, we introduce a weighted correntropy for *random vectors* and present a novel CIM that utilizes different kernel bandwidths for different channels. Further, we design a compact maximum correntropy-based KF (CMC-KF) that utilizes an infinite bandwidth for Gaussian channels and a unified bandwidth σ for non-Gaussian channels. Finally, we apply this method to orientation estimation and construct a compact maximum correntropy-based ESKF (CMC-ESKF) for orientation estimation. A series of experiments are conducted to verify the effectiveness of the proposed method, which reveals that its performance is significantly better than the GD [2] and the conventional ESKF [8], especially with impulsive accelerations and fast-varying accelerations. The contributions of this brief are summarized as follows.

1) The effect of the external acceleration on the conventional ESKF is investigated, and it is discovered that the conventional ESKF is incapable of accurate orientation tracking with fast-varying external accelerations.
2) A weighted maximum correntropy is introduced and the properties of the corresponding CIM are given. Then, a CMC-KF is designed which utilizes *different* bandwidths for different kinds of noises (i.e., Gaussian noises and heavy-tailed noises). Finally, a CMC-ESKF is designed by applying the CMC-KF to orientation estimation.
3) The performance of the proposed algorithm is validated by experiments. Results show that our algorithm is *significantly better* than the GD and ESKF.

The remainder of this brief is organized as follows. In Section II, some preliminaries and the problem statement are provided. In Section III, the CMC-KF is derived for state estimation of a system with different types of noises, and the CMC-ESKF is designed for orientation estimation of IMUs. In Section IV, experiments are conducted to verify the efficacy of the proposed method. In Section V, a conclusion is drawn.

Notations: For matrix A, A^T and A^{-1} denotes its transpose and inverse, respectively. For state x, x^- and x^+ denotes its the *a priori* and the *a posteriori* estimate. The operator $g\times$ converts a vector g to a skew symmetric matrix.

II. Preliminaries and Problem Statement

We first introduce the weighted correntropy for *random vectors* and provide its properties. Then, we present the sensor models of IMUs and provide the dilemma of the conventional ESKF.

A. Weighted Correntropy

Traditionally, the correntropy is a local similarity measure of two random variables $\mathrm{X}, \mathrm{Y} \in \mathbb{R}$ with joint distribution $\mathrm{F}_{\mathrm{XY}}(x, y)$ as

$$\mathrm{C}(\mathrm{X}, \mathrm{Y}) = \mathrm{E}[\kappa(\mathrm{X}, \mathrm{Y})] = \int \kappa(x, y) d\mathrm{F}_{\mathrm{XY}}(x, y) \tag{1}$$

where $\kappa(x, y)$ is a shift-invariant Mercer kernel. Here, we use the Gaussian kernel with

$$\kappa(x, y) = G_\sigma(x, y) = \exp\left(-\frac{e^2}{2\sigma^2}\right) \tag{2}$$

where $e = x - y$ and σ is the kernel bandwidth. In a practical application, the joint distribution $\mathrm{F}_{\mathrm{XY}}(x, y)$ is unknown and only samples $x(k)$, $y(k)$ are available. Then, the correntropy can be obtained by

$$\bar{\mathrm{C}}(\mathrm{X}, \mathrm{Y}) = \frac{1}{M}\sum_{k=1}^{M} \kappa(x(k), y(k)) = \frac{1}{M}\sum_{k=1}^{M} G_\sigma(x(k), y(k)) \tag{3}$$

where M is the sample number and k is the sample index.

In this brief, we define a weighted correntropy for *random vectors* $\mathcal{X}, \mathcal{Y} \in \mathbb{R}^l$ as follows:

$$\mathrm{C}(\mathcal{X}, \mathcal{Y}) = \sum_{i=1}^{l} \mathrm{E}[\tilde{\kappa}_i(\mathcal{X}_i, \mathcal{Y}_i)] = \sum_{i=1}^{l} \int \tilde{\kappa}_i(x_i, y_i) d\mathrm{F}_{\mathcal{X}_i\mathcal{Y}_i}(x_i, y_i) \tag{4}$$

with

$$\tilde{\kappa}_i(x_i, y_i) = 2\sigma_i^2 \kappa_{\sigma_i}(e_i) = 2\sigma_i^2 G_{\sigma_i}(e_i) = 2\sigma_i^2 \exp\left(-\frac{e_i^2}{2\sigma_i^2}\right) \tag{5}$$

where $\mathcal{X}_i$ and $\mathcal{Y}_i$ are the ith element of $\mathcal{X}$ and $\mathcal{Y}$, and $e_i = x_i - y_i$ is the corresponding error. In a practical application, the joint distribution $\mathrm{F}_{\mathcal{X}_i\mathcal{Y}_i}$ is unknown and the correntropy can be estimated by

$$\bar{\mathrm{C}}(\mathcal{X}, \mathcal{Y}) = \sum_{i=1}^{l} 2\sigma_i^2 \bar{\mathrm{C}}_{\sigma_i}(\mathcal{X}_i, \mathcal{Y}_i) \tag{6}$$

with

$$\bar{\mathrm{C}}_{\sigma_i}(\mathcal{X}_i, \mathcal{Y}_i) = \frac{1}{M}\sum_{k=1}^{M} G_{\sigma_i}(e_i(k)) \tag{7}$$

where $e_i(k) = x_i(k) - y_i(k)$. One can see that $\bar{\mathrm{C}}(\mathcal{X}, \mathcal{Y})$ actually is a weighted summation of $\bar{\mathrm{C}}_{\sigma_i}(\mathcal{X}_i, \mathcal{Y}_i)$ from $i = 1$ to l. Then, we design the corresponding CIM as

$$\begin{aligned}\overline{\mathrm{CIM}}(\mathcal{X}, \mathcal{Y}) &= \left(\sum_{i=1}^{l} 2\sigma_i^2\big(1 - \bar{\mathrm{C}}_{\sigma_i}(\mathcal{X}_i, \mathcal{Y}_i)\big)\right)^{1/2} \\ &= \left(\sum_{i=1}^{l} 2\sigma_i^2 \frac{1}{M}\sum_{k=1}^{M}\big(1 - G_{\sigma_i}(e_i(k))\big)\right)^{1/2}. \end{aligned} \tag{8}$$

Theorem 1: $\overline{\mathrm{CIM}}(\mathcal{X}, \mathcal{Y})$ defines a metric in M-dimensional sample vector space.

The proof of this theorem can be found in Appendix A.

Theorem 2: The metric $\overline{\mathrm{CIM}}(\mathcal{X}, \mathcal{Y})$ is identical to an ℓ_2 norm $\|\mathcal{X} - \mathcal{Y}\|_2$ when setting all kernels $\sigma_i \to \infty$ with $M = 1$.

Proof: Taking series expansion of the Gaussian Kernel, we have

$$G_{\sigma_i}(e_i) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n \sigma_i^{2n} n!} e_i^{2n}. \tag{9}$$

By setting $\sigma_i \to \infty$, we obtain

$$\lim_{\sigma_i \to \infty} 2\sigma_i^2(1 - G_{\sigma_i}(e_i)) = e_i^2. \tag{10}$$

In the case with $M = 1$, we have

$$\begin{aligned}\lim_{\sigma_i \to \infty} \overline{\mathrm{CIM}}(\mathcal{X}, \mathcal{Y}) &= \lim_{\sigma_i \to \infty}\left(\sum_{i=1}^{l} 2\sigma_i^2\big(1 - G_{\sigma_i}(e_i)\big)\right)^{1/2} \\ &= \left(\sum_{i=1}^{l} e_i^2\right)^{1/2} \\ &= \|\mathcal{X} - \mathcal{Y}\|_2. \end{aligned} \tag{11}$$

This completes the proof. ∎

B. Sensor Models and the Conventional ESKF

The gyroscope reading $y_{G,k} \in \mathbb{R}^{3\times 1}$ is modeled as the summation of the angular velocity ω_k, the offset b_k, and the white noise $v_{G,k}$. The offset term b_k is modeled as a first-order Markov process with the white noise $v_{b,k}$. Hence, the gyroscope model can be written as

$$\begin{cases} y_{G,k} = \omega_k + b_k + v_{G,k} \\ b_k = b_{k-1} + v_{b,k}. \end{cases} \tag{12}$$

The accelerometer reading $y_{A,k} \in \mathbb{R}^{3\times 1}$ is modeled as the subtraction of the gravity vector ${}^s g_{A,k}$ and the linear acceleration ${}^s a_k$ in the sensor frame, plus the white noise $v_{A,k}$. The linear acceleration in the global coordinate ${}^g a_k \in \mathbb{R}^{3\times 1}$ is modeled as a Markov process with a decay coefficient η and white noises $v_{a,k}$. Thus, the accelerometer model is written as

$$\begin{cases} y_{A,k} = {}^s g_{A,k} - {}^s a_k + v_{A,k} \\ {}^g a_k = \eta({}^g a_{k-1}) + v_{a,k}, \quad 0 < \eta < 1. \end{cases} \tag{13}$$

The linear acceleration in the global frame and the sensor frame can be transferred by a rotation matrix $R_k \in \mathbb{R}^{3\times 3}$ with

$${}^g a_k = R_k\big({}^s a_k\big). \tag{14}$$

The conventional ESKF models the error state rather than the true state, which includes the error orientation $\theta_{\varepsilon,k} \in \mathbb{R}^{3\times 1}$ (Euler angle), error offset $b_{\varepsilon,k} \in \mathbb{R}^{3\times 1}$, and error acceleration ${}^s a_{\varepsilon,k} \in \mathbb{R}^{3\times 1}$. The state transfer matrix $\Phi_k \in \mathbb{R}^{9\times 9}$ is a zero matrix since the *a priori* estimate $x_{\varepsilon,k}^-$ does not depend on the previous *a posteriori* estimate $x_{\varepsilon,k-1}^+$ [8], [11]. Thus, we have

$$x_{\varepsilon,k} = \Phi_k x_{\varepsilon,k-1} + w_{x,k} \tag{15}$$

with

$$x_{\varepsilon,k} = \begin{bmatrix} \theta_{\varepsilon,k} \\ b_{\varepsilon,k} \\ {}^s a_{\varepsilon,k} \end{bmatrix}, \quad w_{x,k} = \begin{bmatrix} w_{\theta,k} \\ w_{b,k} \\ w_{a,k} \end{bmatrix}$$

where $w_{\theta,k}$, $w_{b,k}$, $w_{a,k}$ is the orientation noise, offset noise, and acceleration noise, respectively. The process covariance is set to be Q.

The measurement signal $z_{\varepsilon,k} \in \mathbb{R}^{3\times 1}$ describes the difference of the *a priori* estimate of the gravity vector obtained by the accelerometer and by the *a priori* orientation. Thus, we have

$$z_{\varepsilon,k} = {}^s\hat{g}_{A,k}^- - {}^s\hat{g}_{G,k}^- = H_k x_{\varepsilon,k} + v_{z,k} \tag{16}$$

with

$$\begin{aligned} {}^s\hat{g}_{A,k}^- &= y_{A,k} + {}^s\hat{a}_k^- \\ {}^s\hat{g}_{G,k}^- &= \hat{R}_k^- g \\ H_k &= \big[-{}^s\hat{g}_{G,k}^-\times, \delta t\big({}^s\hat{g}_{G,k}^-\times\big), I_{3\times 3}\big] \end{aligned} \tag{17}$$

where ${}^s\hat{a}_k^-$ is the *a priori* estimate of the external acceleration, $\hat{R}_k^-$ is the *a priori* estimate of the orientation matrix, g is the universal gravity vector, δt is the sampling time, and H_k is the measurement matrix. In the conventional ESKF, the measurement covariance matrix is R_z, and the orientation $\theta_k \in \mathbb{R}^{3\times 1}$ can be expressed as $q_k \in \mathbb{R}^{4\times 1}$ in a quaternion form or $R_k \in \mathbb{R}^{3\times 3}$ in a rotation matrix form. The error covariance at step k is P_k. The detailed algorithm of the ESKF is summarized in Appendix B, and a more detailed description can be found in [8] and [11].

C. Effects of the External Acceleration

Based on Line 8 in Algorithm 3, the *a posteriori* estimate of the error state $\hat{x}_{\varepsilon,k}^+$ can be written as

$$\hat{x}_{\varepsilon,k}^+ = K_k z_{\varepsilon,k} \tag{18}$$

with

$$\hat{x}_{\varepsilon,k}^+ = \begin{bmatrix} \hat{\theta}_{\varepsilon,k}^+ \\ \hat{b}_{\varepsilon,k}^+ \\ {}^s\hat{a}_{\varepsilon,k}^+ \end{bmatrix}, \quad K_k = \begin{bmatrix} K_k(\theta_\varepsilon) \\ K_k(b_\varepsilon) \\ K_k({}^s a_\varepsilon) \end{bmatrix}.$$

According to (16)–(18), it follows that:

$${}^s\hat{a}_{\varepsilon,k}^+ = K_k\big({}^s a_\varepsilon\big) z_{\varepsilon,k} = K_k\big({}^s a_\varepsilon\big)\big({}^s\hat{g}_{A,k}^- - {}^s\hat{g}_{G,k}^-\big). \tag{19}$$

Based on (19), and Lines 5 and 15 in Algorithm 3, the closed form of the acceleration estimation can be derived as

$$\begin{aligned} {}^s\hat{a}_k^+ &= \eta\big({}^s\hat{a}_{k-1}^+\big) - {}^s\hat{a}_{\varepsilon,k}^+ \\ &= \eta\big({}^s\hat{a}_{k-1}^+\big) - K_k({}^s a_\varepsilon)\big({}^s\hat{g}_{A,k}^- - {}^s\hat{g}_{G,k}^-\big) \\ &= \eta\big({}^s\hat{a}_{k-1}^+\big) - K_k({}^s a_\varepsilon)\big(y_{A,k} + \eta\big({}^s\hat{a}_{k-1}^+\big) - {}^s\hat{g}_{G,k}^-\big) \\ &= (I - K_k({}^s a_\varepsilon))\eta\big({}^s\hat{a}_{k-1}^+\big) + K_k({}^s a_\varepsilon)\big({}^s\hat{g}_{G,k}^- - y_{A,k}\big). \end{aligned} \tag{20}$$

Remark 1: This estimator relies heavily on the acceleration model ${}^s a_k = \eta({}^s a_{k-1})$. Generally, this model does not match the real situation since the practical external acceleration model is time-varying. This model mismatch would inevitably affect the estimation accuracy of the ESKF.

D. Problem Statement

When attaching IMUs to the shank of exoskeletons, the impulsive acceleration occurs at the heel strike moment due to collision with the ground, whereas the continuous external acceleration occurs at the swing phase due to the movement of the leg. These accelerations do not match the first-order acceleration model, and the unmodeled dynamics (the difference between the true model and nominal model) can be seen as non-Gaussian noises. Then, it becomes a state estimation problem with part of the channels containing non-Gaussian noises. A conventional method for this issue is to increase the nominal noise covariance. However, it would significantly degrade the estimation accuracy [31]. Generally, this dilemma cannot be solved by the traditional ESKF. In this brief, we employ the property of the weighted correntropy to handle this issue.

III. Compact Maximum Correntropy-Based ESKF

To deal with the aforementioned issues, we derive a CMC-KF which uses infinite bandwidth for Gaussian channels and bandwidth σ for non-Gaussian channels. Then, we apply this method to the conventional ESKF and obtain the CMC-ESKF for orientation estimation.

A. Compact Maximum Correntropy-Based KF

We consider the following linear system with state transfer matrix $A \in \mathbb{R}^{n\times n}$ and observation matrix $C \in \mathbb{R}^{m\times n}$. Then, we have

$$\begin{aligned} x_k &= Ax_{k-1} + w_k \\ y_k &= Cx_k + v_k \end{aligned} \tag{21}$$

where w_k and v_k are white noises with covariances Q and R. We can rewrite (21) as

$$\begin{bmatrix} x_k^- \\ y_k \end{bmatrix} = \begin{bmatrix} I \\ C \end{bmatrix} x_k + q_k \tag{22}$$

where $I \in \mathbb{R}^{n\times n}$ is an identity matrix and q_k represents

$$q_k = \begin{bmatrix} x_k^- - x_k \\ v_k \end{bmatrix}$$

with

$$E\big(q_k q_k^T\big) = \begin{bmatrix} P_k^- & 0 \\ 0 & R \end{bmatrix} = \begin{bmatrix} B_{p,k}B_{p,k}^T & 0 \\ 0 & B_r B_r^T \end{bmatrix} = B_k B_k^T$$

where $B_{p,k}$ and B_r can be obtained by Cholesky decomposition. Left multiplying B_k^{-1} in the both sides of (22), we obtain

$$T_k = W_k x_k + \xi_k \tag{23}$$

with

$$T_k = B_k^{-1}\begin{bmatrix} x_k^- \\ y_k \end{bmatrix}, \quad W_k = B_k^{-1}\begin{bmatrix} I \\ C \end{bmatrix}, \ \xi_k = B_k^{-1} q_k.$$

The noise term ξ_k is white since $E[\xi_k \xi_k^T] = [(B_k^{-1}q_k)(B_k^{-1}q_k)^T] = I$.

In this brief, we consider the situation that parts of the system channels contain non-Gaussian noises and present the following weighted correntropy-based objective function:

$$J(x_k) = \sum_{i=1}^{n+m} 2\sigma_i^2 G_{\sigma_i}(t_{i,k} - w_{i,k}x_k) \tag{24}$$

where σ_i is the kernel bandwidth for channel i, $t_{i,k}$ is the ith element of T_k, and $w_{i,k}$ is the ith row of W_k. Then, the estimator becomes

$$x_k^+ = \arg\max_{x_k} \ J(x_k) = \arg\max_{x_k} \ \sum_{i=1}^{n+m} 2\sigma_i^2 G_{\sigma_i}(e_{i,k}) \tag{25}$$

where $e_{i,k} = t_{i,k} - w_{i,k}x_k$. It is equivalent with

$$x_k^+ = \arg\min_{x_k} \ \sum_{i=1}^{n+m} 2\sigma_i^2 (1 - G_{\sigma_i}(e_{i,k})) \tag{26}$$

which actually is identical to $(\overline{\mathrm{CIM}}(\mathcal{T},\mathcal{X}))^2$ with $M = 1$ where $\mathcal{T}$ and $\mathcal{X}$ are random vectors with realizations T_k and $W_k x_k$. The solution of (25) can be obtained by solving

$$\frac{\partial J(x_k)}{\partial x_k} = 0. \tag{27}$$

It follows that:

$$\sum_{i=1}^{n+m} G_{\sigma_i}(e_{i,k}) w_{i,k}^T \big(t_{i,k} - w_{i,k}x_k\big) = 0. \tag{28}$$

Then, we have

$$x_k = \left(\sum_{i=1}^{n+m}\big[G_{\sigma_i}(e_{i,k}) w_{i,k}^T w_{i,k}\big]\right)^{-1}\left(\sum_{i=1}^{n+m} G_{\sigma_i}(e_{i,k}) w_{i,k}^T t_{i,k}\right). \tag{29}$$

Since $e_{i,k} = t_{i,k} - w_{i,k}x_k$, (29) is a fix-point equation of x_k which can be solved by a fix-point iterative algorithm. It can be rewritten as

$$x_k = \big(W_k^T M_k W_k\big)^{-1} W_k^T M_k T_k \tag{30}$$

with

$$\begin{aligned} M_k &= \begin{bmatrix} \tilde{M}_p & 0 \\ 0 & \tilde{M}_r \end{bmatrix} \\ \tilde{M}_p &= \mathrm{diag}(G_{\sigma_p}(e_p)), \quad \tilde{M}_r = \mathrm{diag}(G_{\sigma_r}(e_r)) \\ e_p &= B_{p,k}^{-1}\big(x_k^- - x_k\big), \quad e_r = B_r^{-1}(y_k - Cx_k) \end{aligned} \tag{31}$$

Algorithm 1 CMC-KF

Initialization:
Choose a kernel bandwidth σ and positive number ε
State Prediction:
$\hat{x}_k^- = A\hat{x}_{k-1}^+$
$P_k^- = AP_{k-1}^+A^T + Q$
Obtain $B_{p,k}$ and B_r by $P_k^- = B_{p,k}B_{p,k}^T$ and $R = B_rB_r^T$
State Update:
$\hat{x}_{k,0}^+ = \hat{x}_k^-$
while $\frac{\|\hat{x}_{k,t}^+ - \hat{x}_{k,t-1}^+\|}{\|\hat{x}_{k,t-1}^+\|} > \varepsilon$ **do** ▷ t starts form 1
$\hat{x}_{k,t}^+ = \hat{x}_k^- + \tilde{K}_k(y_k - C\hat{x}_k^-)$
$\tilde{K}_k = \tilde{P}_k^-C^T(C\tilde{P}_k^-C^T + \tilde{R}_k)^{-1}$
$\tilde{P}_k^- = B_{p,k}\tilde{M}_p^{-1}B_{p,k}^T$
$\tilde{R}_k = B_r\tilde{M}_r^{-1}B_r^T$
$\tilde{M}_p = \Lambda_p I_{n\times n} + (I_{n\times n} - \Lambda_p)diag(G_\sigma(e_p))$
$\tilde{M}_r = \Lambda_r I_{m\times m} + (I_{m\times m} - \Lambda_r)diag(G_\sigma(e_r))$
$e_p = B_{p,k}^{-1}\hat{x}_k^- - B_{p,k}^{-1}\hat{x}_{k,t-1}^+$
$e_r = B_r^{-1}y_k - B_r^{-1}C\hat{x}_{k,t-1}^+$
$t = t + 1$
end while
$P_k^+ = (I - \tilde{K}_kC)P_k^-(I - \tilde{K}_kC)^T + \tilde{K}_kR\tilde{K}_k^T$

where $\sigma_p = [\sigma_1, \sigma_2, \ldots, \sigma_n]^T$ is the process bandwidth vector, $\sigma_r = [\sigma_{n+1}, \sigma_{m+2}, \ldots, \sigma_{n+m}]^T$ is the measurement bandwidth vector, e_p is the process error, and e_r is the measurement error. Equation (30) can rewritten as

$$\hat{x}_k^+ = \hat{x}_k^- + \tilde{K}_k\big(y_k - C\hat{x}_k^-\big) \tag{32}$$

with

$$\begin{cases} \tilde{K}_k = \tilde{P}_k^-C^T\big(C\tilde{P}_k^-C^T + \tilde{R}_k\big)^{-1} \\ \tilde{P}_k^- = B_{p,k}\tilde{M}_p^{-1}B_{p,k}^T \\ \tilde{R}_k = B_r\tilde{M}_r^{-1}B_r^T. \end{cases} \tag{33}$$

One can refer to [28] for a detailed derivation. To avoid involving too much parameters, we apply infinite bandwidth σ_{inf} for Gaussian channels, and bandwidth σ for non-Gaussian channels. In this case, $\tilde{M}_p$ and $\tilde{M}_r$ can be written as

$$\begin{aligned} \tilde{M}_p &= \Lambda_p I_{n\times n} + (I_{n\times n} - \Lambda_p)\text{diag}(G_\sigma(e_p)) \\ \tilde{M}_r &= \Lambda_r I_{m\times m} + (I_{m\times m} - \Lambda_r)\text{diag}(G_\sigma(e_r)) \end{aligned} \tag{34}$$

where Λ_p and Λ_r are diagonal indicator matrices where the diagonal entries are set to 1 for Gaussian channels, and are set to 0 for non-Gaussian channels. Then, we can obtain the CMC-KF in Algorithm 1.

Theorem 3: CMC-KF is identical to the traditional KF by setting all bandwidths as σ_{inf} and is equivalent to the conventional MCKF by setting all bandwidths as σ.

Proof: By designing all bandwidths as σ_{inf}, we have $\tilde{M}_p = I_{n\times n}$ and $\tilde{M}_r = I_{m\times m}$. Correspondingly, $\tilde{P}_k^- = P_k^-$ and $\tilde{R}_k = R$. In this case, the CMC-KF is identical to the KF. By setting all bandwidths as σ, we have $\tilde{M}_p = \text{diag}(G_\sigma(e_p))$ and $\tilde{M}_r = \text{diag}(G_\sigma(e_r))$. Then, it is the same with the MCKF. ■

TABLE I
ERROR PERFORMANCES OF DIFFERENT ALGORITHMS WITH IMPULSIVE ACCELERATIONS

Algorithm	GD	ESKF	**CMC-ESKF**
RMSE (°)	0.6520	0.6734	0.3425
ME (°)	1.4303	1.3371	0.8323

Remark 2: The KF is a well-known minimum mean square error estimator for linear systems with Gaussian noises. However, its performance would degenerate with non-Gaussian noises. On the contrary, conventional MCKF can suppress non-Gaussian noises. However, its performance on the Gaussian noises is bad. The CMC-KF combines the advantages of the KF and MCKF, which employs σ_{inf} to suppress Gaussian noises and σ to reject non-Gaussian noises.

B. Compact Maximum Correntropy-Based ESKF

As shown in (20), the first-order acceleration model cannot accurately represent the practical acceleration dynamics. The unmodeled dynamics can be seen as non-Gaussian noises. To mitigate its bad effects, we design the process bandwidth vector and measurement bandwidth vector as

$$\begin{aligned} \sigma_p &= [\sigma_{\text{inf}} I_{1\times 3}, \sigma_{\text{inf}} I_{1\times 3}, \sigma\, I_{1\times 3}]^T \\ \sigma_r &= [\sigma_{\text{inf}} I_{1\times 3}]^T. \end{aligned} \tag{35}$$

Then, we have

$$\begin{aligned} \Lambda_p &= \begin{bmatrix} I_{6\times 6} & 0_{3\times 3} \\ 0_{3\times 3} & 0_{3\times 3} \end{bmatrix}, \quad \Lambda_r = I_{3\times 3} \\ \tilde{M}_p &= \begin{bmatrix} I_{6\times 6} & 0_{3\times 3} \\ 0_{3\times 3} & \Upsilon \end{bmatrix}, \quad \tilde{M}_r = I_{3\times 3} \end{aligned} \tag{36}$$

with $\Upsilon = \text{diag}(G_\sigma(e_{p,a}))$ where $e_{p,a}$ is a component of e_p with $e_p = [e_{p,\theta}^T, e_{p,b}^T, e_{p,a}^T]^T$. In this case, we have $\tilde{R}_z = B_r\tilde{M}_rB_r^T = R_z$. Thus, the calculation of the B_r, e_r, and $\tilde{R}_z$ can be omitted.

We apply these kernel bandwidths to the conventional ESKF and construct the CMC-ESKF for orientation estimation. The details are summarized in Algorithm 2.

IV. EXPERIMENTS

A series of experiments are conducted to verify the efficacy of the CMC-ESKF, and its performance is compared with the GD [2] and ESKF [8].

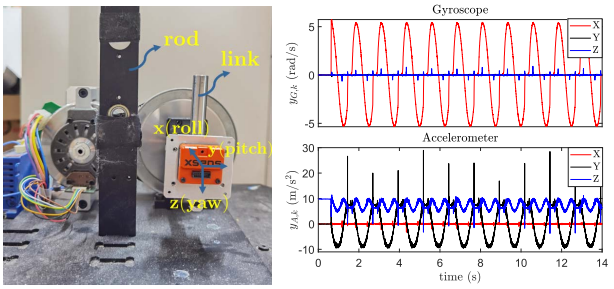
A. Dynamic Impact Test

The Xsens MTI-670 is utilized to test the effectiveness of the CMC-ESKF with impulsive accelerations. The IMU is mounted on a rotating shaft driven by a motor. The impulsive acceleration is generated by the link hitting the rod. The experimental setup is shown in Fig. 1(a) and the sensor signals are shown in Fig. 1(b).

The ground truth angle is obtained by the optical encoder and the corresponding error performances of the GD, ESKF, and CMC-ESKF are shown in Fig. 2. The root-mean-square error (RMSE) for the GD and ESKF is 0.6520° and 0.6734°,

Algorithm 2 CMC-ESKF

Initialization:
Choose bandwidth σ and small positive number ε
State Prediction:
$\hat{q}_k^- = \hat{q}_{k-1}^+ \Delta q(\hat{\omega}_k^- \delta t)$
$\hat{\omega}_k^- = y_{G,k} - \hat{b}_k^-$
$\hat{b}_k^- = \hat{b}_{k-1}^+$
${}^s\hat{a}_k^- = \eta({}^s\hat{a}_{k-1}^+)$
Error State Propagation:
$P_k^- = f(P_{k-1}^+, Q)$ ▷ details in (40)
Obtain $B_{p,k}$ by $B_{p,k}B_{p,k}^T = P_k^-$
$x_{\varepsilon,k,0}^+ = x_{\varepsilon,k}^- = 0$
while $\frac{x_{\varepsilon,k,t}^+ - x_{\varepsilon,k,t-1}^+}{x_{\varepsilon,k,t-1}^+} \leq \varepsilon$ **do** ▷ t starts from 1
$x_{\varepsilon,k,t}^+ = x_{\varepsilon,k}^- + \tilde{K}_k z_{\varepsilon,k}$ ▷ $z_{\varepsilon,k}$ is shown in (16) and (17)
$\tilde{K}_k = \tilde{P}_k^- H_k^T S_k^{-1}$ ▷ H_k is shown in (17)
$S_k = H_k \tilde{P}_k^- H_k^T + R_z$
$\tilde{P}_k^- = B_{p,k}\tilde{M}_p^{-1}B_{p,k}^T$
$\tilde{M}_p = \Lambda_p I_{n\times n} + (I_{n\times n} - \Lambda_p)diag(G_\sigma(e_p))$
$e_p = B_{p,k}^{-1}x_{\varepsilon,k}^- - B_{p,k}^{-1}x_{\varepsilon,k,t-1}^+$
$t = t + 1$
end while
State Update:
$\hat{q}_k^+ = \hat{q}_k^- \Delta q(-\hat{\theta}_{\varepsilon,k}^+)$
$\hat{b}_k^+ = \hat{b}_k^- - \hat{b}_{\varepsilon,k}^+$
${}^s\hat{a}_k^+ = {}^s\hat{a}_k^- - {}^s\hat{a}_{\varepsilon,k}^+$
$P_k^+ = (I - \tilde{K}_k H_k)P_k^-$



(a) (b)

Fig. 1. Experimental setup of dynamic impact test and the sensor signals. (a) Experimental setup. (b) Sensor signals.

while it is a mere 0.3425° for the CMC-ESKF. The RMSE and maximum error (ME) of different algorithms are summarized in Table I. One can see that the CMC-ESKF is significantly better than the GD and ESKF.

B. Exoskeleton Orientation Estimation

To further investigate the performance of the proposed algorithm on a microprocessor for orientation estimation, we apply our algorithm to a self-made low-cost IMU. To minimize the computation cost, the maximum iteration number in Algorithm 2 is set to be 2 (i.e., $t \leq 2$). The IMU is attached to the shank of an exoskeleton and the gait is designed to imitate human walking. The external acceleration grows with gait frequency f since the IMU has both rotations and translations when walking. A total of five frequencies ($f = 0.05$, 0.2, 0.4, 0.6, and 0.8 Hz) are involved in experiments to investigate the performance of the CMC-ESKF with different levels of external accelerations. The experimental setup is shown in Fig. 3(a) and the sensor signals at $f = 0.8$ Hz are shown in Fig. 3(b). The sampling frequency of the IMU is 500 Hz.

The ground truth angle is obtained by encoders of the exoskeleton. When $f = 0.8$ Hz, the ground truth angle and the corresponding angle errors of different algorithms are shown in Fig. 4. One can see that there is an obvious error for the GD (4.7885°) and ESKF (3.6346°), while it is a mere 1.1747° for the CMC-ESKF. The error performances

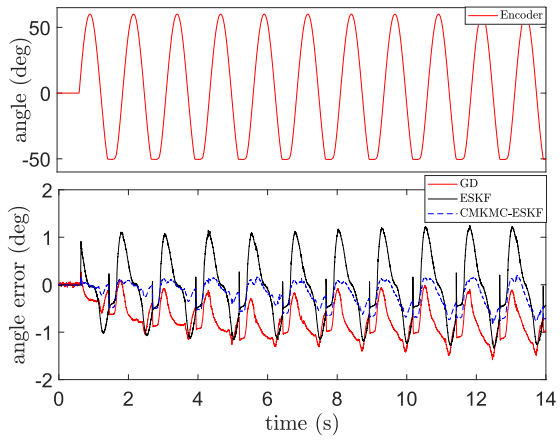
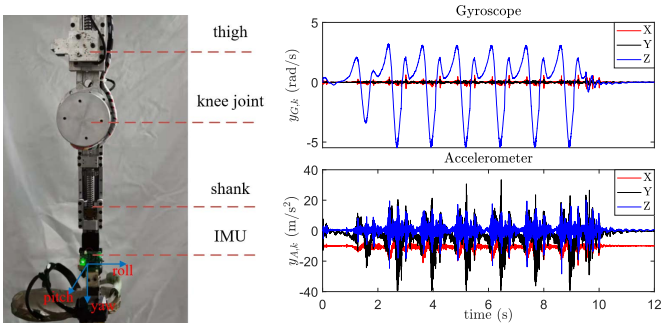


Fig. 2. Rotational angle obtained by an encoder and the angle errors of the GD, ESKF, and CMC-ESKF.



(a) (b)

Fig. 3. Experimental setup and sensor signals with $f = 0.8$ Hz. (a) Experimental setup. (b) Sensor signals.

TABLE II
ERROR PERFORMANCES OF DIFFERENT ALGORITHMS WITH DIFFERENT GAIT FREQUENCIES

f (Hz)	RMSE (°)			ME (°)		
	GD	ESKF	**CMC**	GD	ESKF	**CMC**
0.05	0.7452	0.5290	0.4786	1.1416	1.1089	0.9867
0.2	0.7321	0.3862	0.3268	1.5536	0.9738	0.8233
0.4	0.9707	0.7006	0.1977	2.5580	1.9172	0.5786
0.6	1.9972	1.3515	0.3609	5.2197	3.7455	1.0701
0.8	1.8236	1.2210	0.3641	4.7885	3.6346	1.1747

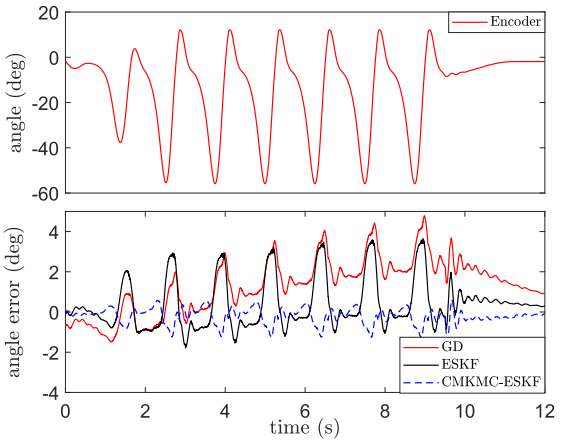


Fig. 4. Encoder angle and the corresponding errors of the GD, ESKF, and CMC-ESKF with $f = 0.8$ Hz.

with different frequencies are summarized in Table II (we use CMC to denote CMC-ESKF in this table). One can see that the performance of the CMC-ESKF is better than the GD and ESKF in all frequencies, especially with obvious external accelerations ($f \geq 0.4$ Hz).

V. Conclusion

In this brief, we first investigate the adverse effects of external accelerations on the traditional ESKF and reveal that the conventional ESKF is incapable of accurate orientation estimation with fast-varying external accelerations. Then, we construct a weighted correntropy and introduce a novel CIM which utilizes different bandwidths for different types of noises. Based on the proposed CIM, we design a novel CMC-KF that gives a time-varying norm penalty for non-Gaussian channels and an ℓ_2 norm penalty for Gaussian channels. Finally, we apply this method to orientation estimation and construct the CMC-ESKF. Compelling experiments verify the effectiveness of the proposed method.

The magnetometer signals and the magnetic disturbance are not considered in this brief, which may be our future work.

Appendix

A. Proof of Theorem 1

Proof: The $\overline{\text{CIM}}(\mathcal{X},\mathcal{Y})$ in (8) can be rewritten as

$$\overline{\text{CIM}}(\mathcal{X},\mathcal{Y}) = \left(\sum_{i=1}^{l} 2\sigma_i^2 \frac{1}{M}\sum_{k=1}^{M}\left(1-\kappa_{\sigma_i}(e_i(k))\right)\right)^{1/2} \tag{37}$$

with $\kappa_{\sigma_i}(e_i(k)) = \exp(-((e_i^2(k))/(2\sigma_i^2)))$. To be a metric, $\overline{\text{CIM}}$ has to fulfill the following properties:

1) *Negativity:* $\overline{\text{CIM}}(\mathcal{X},\mathcal{Y}) \geq 0$. This can be proved by $0 < \kappa_{\sigma_i}(e_i(k)) \leq 1$.
2) *Identities of Indiscernibles:* $\overline{\text{CIM}}(\mathcal{X},\mathcal{Y}) = 0$ iff $\mathcal{X} = \mathcal{Y}$. This can be proved by $\kappa_{\sigma_i}(e_i(k)) = 0$ iff $e_i(k) = x_i(k) - y_i(k) = 0$.
3) *Symmetry:* $\overline{\text{CIM}}(\mathcal{X},\mathcal{Y}) = \overline{\text{CIM}}(\mathcal{Y},\mathcal{X})$. This can be proved by $\kappa_{\sigma_i}(x_i,y_i) = \kappa_{\sigma_i}(y_i,x_i)$.

Algorithm 3 ESKF

State Prediction:
$\hat{q}_k^- = \hat{q}_{k-1}^+ \Delta q(\hat{\omega}_k^- \delta t)$
$\hat{\omega}_k^- = y_{G,k} - \hat{b}_k^-$
$\hat{b}_k^- = \hat{b}_{k-1}^+$
${}^s\hat{a}_k^- = \eta({}^s\hat{a}_{k-1}^+)$
Error State Propagation:
$P_k^- = f(P_{k-1}^+, Q)$, details in (40)
$\hat{x}_{\varepsilon,k}^+ = K_k z_{\varepsilon,k}$ ▷ $z_{\varepsilon,k}$ is shown in (16), (17)
$K_k = P_k^- H_k^T S_k^{-1}$
$H_k = [-{}^s\hat{g}_{G,k}^-\times, \delta t({}^s\hat{g}_{G,k}^-\times), I_{3\times 3}]$
$S_k = H_k P_k^- H_k^T + R_z$
Update:
$\hat{q}_k^+ = \hat{q}_k^- \Delta q(-\hat{\theta}_{\varepsilon,k}^+)$
$\hat{b}_k^+ = \hat{b}_k^- - \hat{b}_{\varepsilon,k}^+$
${}^s\hat{a}_k^+ = {}^s\hat{a}_k^- - {}^s\hat{a}_{\varepsilon,k}^+$
$P_k^+ = (I - K_k H_k) P_k^-$

4) *Triangle Inequity:* $\overline{\text{CIM}}(\mathcal{X},\mathcal{Z}) \leq \overline{\text{CIM}}(\mathcal{X},\mathcal{Y}) + \overline{\text{CIM}}(\mathcal{Y},\mathcal{Z})$. The Gaussian kernel induces a mapping function Φ from the input space to an infinite dimensional reproducing kernel Hilbert space (RKHS) $\mathcal{F}$ with $\kappa_\sigma(x_i,y_i) = \langle\Phi(x_i),\Phi(y_i)\rangle_{\mathcal{F}}$. For random pairs $\mathcal{X}_i$, $\mathcal{Y}_i$, we construct new vectors $\tilde{\mathcal{X}}_i = [\Phi(x_i(1)),\Phi(x_i(2)),\ldots,\Phi(x_i(M))]^T$, $\tilde{\mathcal{Y}}_i = [\Phi(y_i(1)),\Phi(y_i(2)),\ldots,\Phi(y_i(M))]^T$ in Hilbert space $\mathcal{F}^M$. Then, the Euclidean distance $D(\tilde{\mathcal{X}}_i,\tilde{\mathcal{Y}}_i)$ has

$$\begin{aligned}
&D(\tilde{\mathcal{X}}_i,\tilde{\mathcal{Y}}_i)\\
&= \left(\langle\tilde{\mathcal{X}}_i-\tilde{\mathcal{Y}}_i,\tilde{\mathcal{X}}_i-\tilde{\mathcal{Y}}_i\rangle\right)^{1/2}\\
&= \left(\langle\tilde{\mathcal{X}}_i,\tilde{\mathcal{X}}_i\rangle - 2\langle\tilde{\mathcal{X}}_i,\tilde{\mathcal{Y}}_i\rangle + \langle\tilde{\mathcal{Y}}_i,\tilde{\mathcal{Y}}_i\rangle\right)^{1/2}\\
&= \left(\sum_{k=1}^{M}\kappa_{\sigma_i}(0) - 2\sum_{k=1}^{M}\kappa_{\sigma_i}(x_i(k),y_i(k)) + \sum_{k=1}^{M}\kappa_{\sigma_i}(0)\right)^{1/2}\\
&= \sqrt{2M}\left(1-\bar{\text{C}}_{\sigma_i}(\mathcal{X}_i,\mathcal{Y}_i)\right)^{1/2}.
\end{aligned} \tag{38}$$

Then, we have $1-\bar{\text{C}}_{\sigma_i}(\mathcal{X}_i,\mathcal{Y}_i) = ((D^2(\tilde{\mathcal{X}}_i,\tilde{\mathcal{Y}}_i))/2M)$. Correspondingly, we obtain

$$\begin{aligned}
&\overline{\text{CIM}}(\mathcal{X},\mathcal{Z})\\
&= \left(\sum_{i=1}^{l} 2\sigma_i^2\left(1-\bar{\text{C}}_{\sigma_i}(\mathcal{X}_i,\mathcal{Z}_i)\right)\right)^{1/2}\\
&= \left(\sum_{i=1}^{l}\frac{\sigma_i^2}{M}D^2(\tilde{\mathcal{X}}_i,\tilde{\mathcal{Z}}_i)\right)^{1/2}\\
&\leq \left(\sum_{i=1}^{l}\frac{\sigma_i^2}{M}\left(D^2(\tilde{\mathcal{X}}_i,\tilde{\mathcal{Y}}_i)+D^2(\tilde{\mathcal{Y}}_i,\tilde{\mathcal{Z}}_i)\right)\right)^{1/2}\\
&\leq \left(\sum_{i=1}^{l}\frac{\sigma_i^2}{M}D^2(\tilde{\mathcal{X}}_i,\tilde{\mathcal{Y}}_i)\right)^{1/2} + \left(\sum_{i=1}^{l}\frac{\sigma_i^2}{M}D^2(\tilde{\mathcal{Y}}_i,\tilde{\mathcal{Z}}_i)\right)^{1/2}\\
&= \overline{\text{CIM}}(\mathcal{X},\mathcal{Y}) + \overline{\text{CIM}}(\mathcal{Y},\mathcal{Z}).
\end{aligned} \tag{39}$$

The first inequality utilizes the property of $D(\tilde{\mathcal{X}}_i, \tilde{\mathcal{Z}}_i) \leq D(\tilde{\mathcal{X}}_i, \tilde{\mathcal{Y}}_i) + D(\tilde{\mathcal{Y}}_i, \tilde{\mathcal{Z}}_i)$, while the second inequality utilizes the triangle inequality of the ℓ_2 norm $\|p+q\|_2 \leq \|p\|_2 + \|q\|_2$. This completes the proof. ■

B. Error State KF

The traditional ESKF is summarized in Algorithm 3. The calculation of the *a priori* estimate of error covariance P_k^- is as follows:

$$\begin{cases} P_k^-(\theta_\varepsilon, \theta_\varepsilon) \\ = P_{k-1}^+(\theta_\varepsilon, \theta_\varepsilon) + \delta t^2 \left[P_{k-1}^+(b_{\varepsilon,k}, b_{\varepsilon,k}) + Q_{vb} + Q_{vG} \right] \\ P_k^-(\theta_\varepsilon, b_\varepsilon) = -\delta t \left[P_{k-1}^+(\theta_\varepsilon, b_\varepsilon) + Q_{vb} \right] \\ P_k^-(\theta_\varepsilon, a_\varepsilon) = 0_{3\times 3} \\ P_k^-(b_\varepsilon, \theta_\varepsilon) = -\delta t \left[P_{k-1}^+(b_\varepsilon, \theta_\varepsilon) + Q'_{vb} \right] \\ P_k^-(b_\varepsilon, b_\varepsilon) = P_{k-1}^+(b_\varepsilon, b_\varepsilon) + Q_{vb} \\ P_k^-(b_\varepsilon, a_\varepsilon) = 0_{3\times 3} \\ P_k^-(a_\varepsilon, \theta_\varepsilon) = 0_{3\times 3} \\ P_k^-(a_\varepsilon, b_\varepsilon) = 0_{3\times 3} \\ P_k^-(a_\varepsilon, a_\varepsilon) = \eta^2 P_{k-1}^+(a_\varepsilon, a_\varepsilon) + Q_{va} \end{cases} \tag{40}$$

where Q_{vb}, Q_{vG}, Q_{va} is the covariance of the gyroscope offset, gyroscope signal, and linear acceleration, respectively. $P_k^-(x_1, x_2)$ and $P_k^+(x_1, x_2)$ represents the *a priori* and the *a posteriori* estimate of error covariance among states x_1 and x_2.

REFERENCES

[1] X. Yun and E. R. Bachmann, "Design, implementation, and experimental results of a quaternion-based Kalman filter for human body motion tracking," *IEEE Trans. Robot.*, vol. 22, no. 6, pp. 1216–1227, Dec. 2006.

[2] S. O. H. Madgwick, A. J. L. Harrison, and R. Vaidyanathan, "Estimation of IMU and MARG orientation using a gradient descent algorithm," in *Proc. IEEE Int. Conf. Rehabil. Robot.*, Jun. 2011, pp. 1–7.

[3] D. A. Bennett and M. Goldfarb, "IMU-based wrist rotation control of a transradial myoelectric prosthesis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 419–427, Feb. 2018.

[4] C. Caramia *et al.*, "IMU-based classification of Parkinson's disease from gait: A sensitivity analysis on sensor location and feature selection," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 6, pp. 1765–1774, Nov. 2018.

[5] W. Huo, S. Mohammed, Y. Amirat, and K. Kong, "Fast gait mode detection and assistive torque control of an exoskeletal robotic orthosis for walking assistance," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 1035–1052, Aug. 2018.

[6] L. Udawatta *et al.*, "Motion capturing of biomechanical systems in traditional dance using IMU sensors," in *Proc. Int. Conf. Electr. Comput. Technol. Appl. (ICECTA)*, 2017, pp. 1–5.

[7] F. Wittmann *et al.*, "Assessment-driven arm therapy at home using an IMU-based virtual reality system," in *Proc. IEEE Int. Conf. Rehabil. Robot. (ICORR)*, Aug. 2015, pp. 707–712.

[8] D. Roetenberg, H. J. Luinge, C. T. M. Baten, and P. H. Veltink, "Compensation of magnetic disturbances improves inertial and magnetic sensing of human body segment orientation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 13, no. 3, pp. 395–405, Sep. 2005.

[9] M. Kok, J. D. Hol, and T. B. Schön, "Using inertial sensors for position and orientation estimation," *Found. Trends Signal Process.*, vol. 11, nos. 1–2, pp. 1–153, 2017.

[10] A. M. Sabatini, "Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 7, pp. 1346–1356, Jul. 2006.

[11] M. S. M. Pedley and Z. Baranski. (2014). *Freescale Sensor Fusion Kalman Filter*. pp. 24–26. [Online]. Available: https://github.com/memsindustrygroup/Open-Source-Sensor-Fusion/tree/master/docs

[12] X. Yuan, S. Yu, S. Zhang, G. Wang, and S. Liu, "Quaternion-based unscented Kalman filter for accurate indoor heading estimation using wearable multi-sensor system," *Sensors*, vol. 15, no. 5, pp. 10872–10890, 2015.

[13] M. Admiraal, S. Wilson, and R. Vaidyanathan, "Improved formulation of the IMU and MARG orientation gradient descent algorithm for motion tracking in human-machine interfaces," in *Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst. (MFI)*, Nov. 2017, pp. 403–410.

[14] S. Madgwick, "An efficient orientation filter for inertial and inertial/magnetic sensor arrays," *Rep. X-Io Univ. Bristol*, vol. 25, pp. 113–118, Apr. 2010.

[15] P. Gui, L. Tang, and S. Mukhopadhyay, "MEMS based IMU for tilting measurement: Comparison of complementary and Kalman filter based data fusion," in *Proc. IEEE 10th Conf. Ind. Electron. Appl. (ICIEA)*, Jun. 2015, pp. 2004–2009.

[16] K. Kanjanapas, Y. Wang, W. Zhang, L. Whittingham, and M. Tomizuka, "A human motion capture system based on inertial sensing and a complementary filter," in *Proc. ASME Dyn. Syst. Control Conf.*, Oct. 2013, Art. no. V003T40A004.

[17] R. Mahony, T. Hamel, and J.-M. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," *IEEE Trans. Autom. Control*, vol. 53, no. 5, pp. 1203–1218, Jun. 2008.

[18] R. Zhang and L. M. Reindl, "Pedestrian motion based inertial sensor fusion by a modified complementary separate-bias Kalman filter," in *Proc. IEEE Sensors Appl. Symp.*, Feb. 2011, pp. 209–213.

[19] H. Rong, C. Peng, L. Zou, Y. Zhu, Y. Chen, and J. Lv, "Vector-optimized request used to MAG-μ IMU aiming at influence depression of linear acceleration," in *Proc. Int. Conf. Frontiers Sens. Technol. (ICFST)*, Apr. 2017, pp. 68–72.

[20] M. Liu, L. Xu, W. Zhang, J. Cao, G. Shi, and U. Wejinya, "Dynamic posture sensing module based on micro IMU and adaptive filter algorithm," in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Dec. 2017, pp. 2562–2566.

[21] H. Benzerrouk and A. V. Nebylov, "Robust IMU/UWB integration for indoor pedestrian navigation," in *Proc. 25th Saint Petersburg Int. Conf. Integr. Navigat. Syst. (ICINS)*, May 2018, pp. 1–5.

[22] J. A. Corrales, F. A. Candelas, and F. Torres, "Hybrid tracking of human operators using IMU/UWB data fusion by a Kalman filter," in *Proc. 3rd Int. Conf. Hum. Robot Interact. (HRI)*, 2008, pp. 193–200.

[23] A. Martinelli, "Vision and IMU data fusion: Closed-form solutions for attitude, speed, absolute scale, and bias determination," *IEEE Trans. Robot.*, vol. 28, no. 1, pp. 44–60, Feb. 2012.

[24] P. Zhang, J. Gu, E. E. Milios, and P. Huynh, "Navigation with IMU/GPS/digital compass with unscented Kalman filter," in *Proc. IEEE Int. Conf. Mechtron. Automat.*, vol. 3, Jul./Aug. 2005, pp. 1497–1502.

[25] A. Godwin, M. Agnew, and J. Stevenson, "Accuracy of inertial motion sensors in static, quasistatic, and complex dynamic motion," *J. Biomech. Eng.*, vol. 131, no. 11, pp. 1–5, Nov. 2009.

[26] A. Aravkin, J. V. Burke, L. Ljung, A. Lozano, and G. Pillonetto, "Generalized Kalman smoothing: Modeling and algorithms," *Automatica*, vol. 86, pp. 63–86, Dec. 2017.

[27] W. Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: Properties and applications in non-Gaussian signal processing," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5286–5298, Nov. 2007.

[28] B. Chen, X. Liu, H. Zhao, and J. C. Principe, "Maximum correntropy Kalman filter," *Automatica*, vol. 76, pp. 70–77, Feb. 2017.

[29] H. Ju, M. S. Lee, S. Y. Park, J. W. Song, and C. G. Park, "A pedestrian dead-reckoning system that considers the heel-strike and toe-off phases when using a foot-mounted IMU," *Meas. Sci. Technol.*, vol. 27, no. 1, Jan. 2016, Art. no. 015702.

[30] J. B. Bancroft and G. Lachapelle, "Estimating MEMS gyroscope G-sensitivity errors in foot mounted navigation," in *Proc. Ubiquitous Positioning, Indoor Navigat., Location Based Service (UPINLBS)*, 2012, pp. 1–6.

[31] L. Shi, K. H. Johansson, and R. M. Murray, "Kalman filtering with uncertain process and measurement noise covariances with application to state estimation in sensor networks," in *Proc. IEEE Int. Conf. Control Appl.*, Oct. 2007, pp. 1031–1036.