IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 22, 2019, accepted January 15, 2020, date of publication January 21, 2020, date of current version January 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2968363

Random Weighting-Based Nonlinear Gaussian Filtering

ZHAOHUI GAO[1,2], CHENGFAN GU[3], JIAHUI YANG[1], SHESHENG GAO[1], AND YONGMIN ZHONG[4]

[1]School of Automatics, Northwestern Polytechnical University, Xi'an 710072, China
[2]Research and Development Institute, Northwestern Polytechnical University, Shenzhen 710000, China
[3]*(Independent Researcher)* Rowville, VIC 3178, Australia
[4]School of Engineering, RMIT University, Melbourne, VIC 3000, Australia

Corresponding authors: Zhaohui Gao (alexandergao@mail.nwpu.edu.cn) and Yongmin Zhong (yongmin.zhong@rmit.edu.au)

This work was supported in part by the Science, Technology and Innovation Commission of Shenzhen Municipality, China, under Project JCYJ20180306171439979, and in part by the Shaanxi Key Development Program, China, under Project 2018ZDXM-GY-024.

ABSTRACT The Gaussian filtering is a commonly used method for nonlinear system state estimation. However, this method requires both system process noise and measurement noise to be white noise sequences with known statistical characteristics. However, it is difficult to satisfy this condition in engineering practice, making the Gaussian filtering solution deviated or diverged. This paper adopts the random weighting concept to address the limitation of the nonlinear Gaussian filtering. It establishes the random weighting estimations of system noise characteristics on the basis of the maximum a-posterior theory, and further develops a new Gaussian filtering method based on the random weighting estimations to restrain system noise influences on system state estimation by adaptively adjusting the random weights of system noise characteristics. Simulation, experimental and comparison analyses prove that the proposed method overcomes the limitation of the traditional Gaussian filtering in requirement of system noise characteristics, leading to improved estimation accuracy.

INDEX TERMS Nonlinear system state estimation, Gaussian filtering, system noise characteristics, random weighting.

I. INTRODUCTION

Nonlinear system state estimation is an important research topic in many science and engineering fields, such as vehicle navigation and guidance systems, robotic control, target recognition, radar tracking, information fusion, spacecraft orbit determination, and signal processing [5], [21]. The nonlinear Gaussian filtering is a representative technique for nonlinear system state estimation [1], [3], [20], [21], [28]–[30], [32]. It can be classified into two categories. One is the analytical approximation-based Gaussian filtering method. The typical example is the extended Kalman filter (EKF), which is the commonly used for nonlinear state estimation. EKF approximates the nonlinear system model by the first-order Taylor expansion [17], and then conducts nonlinear state estimation based on the linear structure of the traditional Kalman filter. It is simple and easy to implement. However, the filtering accuracy of EKF is degraded due to the truncation of the second-order and above terms in the system model [19]. It also requires computing the Jacobian matrix, which is a non-trivial calculation process [16].

The other category is the numerical approximation-based Gaussian filtering method by approximating the probability distribution of a nonlinear system rather than the nonlinear system itself. The typical examples include UKF and CKF, both in similar accuracy [1], [3], [21], [27]. UKF approximates the posterior probability density function of a nonlinear system state with a set of sigma points by unscented transformation [10], [22], [26], while CKF approximates the posterior mean and variance of a nonlinear system function by spherical-radial cubature rules [2], [27]. Comparing to the analytical approximation-based Gaussian filtering method such as EKF, the numerical approximation-based Gaussian filtering method has higher accuracy and eliminates the non-trivial calculation of the Jacobian matrix. However, due to the inheritance of the linear structure of the traditional Kalman

The associate editor coordinating the review of this manuscript and approving it for publication was Ehsan Asadi.

filter, the numerical approximation-based Gaussian filtering method requires the prior knowledge of system noise characteristics, which is generally unknown in engineering practice, making its filtering solution sensitive to system noises [1], [3], [21], [29], [30].

Research endeavors have been devoted to estimating system noise characteristics to overcome the limitation involved in the numerical approximation-based Gaussian filtering method. The maximum a-posteriori (MAP) is an approach to noise characteristics estimation based on the prior information of system noises [4], [7]. The Sage windowing estimates system noise characteristics at current time point by the arithmetic mean of residual vectors within a small time window [11], [25], [26], [31]. Since the residual vectors within the time window equally contribute to the estimation of noise characteristics at current time point, similar to the MAP approach, this method only works for the cases that system noises are constant or involve small variations, rather than the cases that system noises involve large variations. Cho et al. used the receding horizon, which is a well-established concept in predictive control, to handle noise characteristics estimation for nonlinear systems [6]. However, this technique involves an infinite covariance, and thus may lead to the singularity problem for filtering [8]. Li studied an adaptive method to estimate and further correct system noise characteristics according to measured output information [23]. Similar to the Sage windowing, this method can only handle constant or small-variation system noises due to the use of arithmetic mean for system noise estimation. Chen and Hu studied a method by combing matrix analysis with second-order Taylor series expansion to estimate the upper bound of nonlinear estimation error [36]. However, this method is only suitable for bounded noises.

Random weighting is a simple statistical method. It can provide unbiased estimations for large samples without requiring the probability distribution of variables [9], [12], [13]. This method has applied extensively in science and engineering for various problems [11], [14], [15], [24]. However, the research on using the random weighting method to improve the performance of the nonlinear Gaussian filtering is still very limited.

This paper proposes a novel Gaussian filtering method based on random weighting to overcome the shortcoming of the traditional Gaussian filtering for nonlinear system state estimation by adaptively estimating system noise characteristics. This method establishes random weighting theories to online estimate the means and covariance of both system process and measurement noises based on the MAP principle. Subsequently, it dynamically adjusts the random weights of system noise statistics to restrain the interferences of system noises on system state estimation, resulting in increased estimation accuracy. Simulations and experiments as well as comparison analysis with the traditional Gaussian filtering were conducted to comprehensively evaluate the performance of the proposed method.

II. PRINCIPLE OF RANDOM WEIGHTING

Suppose $(S_1, S_2, \cdots, S_n)$ are the independent random variables sharing the identical distribution function $\Psi(s)$, and their empirical distribution function is described by

$$\Psi_n(s) = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\chi}_{(S_i \le s)} \tag{1}$$

where $\boldsymbol{\chi}_{(S_i \le s)}$ denotes the characteristic function.

Then, the random weighting estimation of $\Psi_n(s)$ can be expressed as [11]

$$\Phi_n(s) = \sum_{i=1}^{n} \sigma_i \boldsymbol{\chi}_{(S_i \le s)} \tag{2}$$

where the random weighting vector $(\sigma_1, \sigma_2, \cdots, \sigma_n)$ is subject to Dirichlet distribution $D(1, 1, \cdots, 1)$, that is, $\sum_{i=1}^{n} \sigma_i = 1$ and the united density function of $(\sigma_1, \sigma_2, \cdots, \sigma_{n-1})$ is $f(\sigma_1, \sigma_2, \cdots, \sigma_{n-1}) = G(n)$, in which G denotes the Gamma function, $(\sigma_1, \sigma_2, \cdots, \sigma_{n-1}) \in D_{n-1}$ and

$$D_{n-1} = \left\{ (\sigma_1, \sigma_2, \cdots, \sigma_{n-1}) : \sigma_k \ge 0, \; k = 1, \cdots, n-1, \sum_{k=1}^{n-1} \sigma_i \le 1 \right\}.$$

From the above, it can be seen that the arithmetic mean estimation described by (1) applies one common factor to each sample. This method is not capable of precisely characterizing the real dynamic properties of system noises due to the equivalent contributions of each sample to the estimation. However, the random weighting estimation described by (2) applies random weights to each sample, enabling us to account for randomly variable characteristics of system noises.

III. NONLINEAR GAUSSIAN FILTERING

A. NONLINEAR GAUSSIAN SYSTEM MODELS

Consider the nonlinear Gaussian system

$$\begin{cases} \boldsymbol{x}_{k+1} = f_k(\boldsymbol{x}_k) + \boldsymbol{w}_k \\ \boldsymbol{y}_k = g_k(\boldsymbol{x}_k) + \boldsymbol{v}_k \end{cases} \tag{3}$$

where $\boldsymbol{x}_k \in \boldsymbol{R}^n$ is the n-dimensional system state vector, $\boldsymbol{y}_k \in \boldsymbol{R}^m$ is the m-dimensional measurement vector, $f(\cdot)$ and $g(\cdot)$ are the nonlinear system and measurement functions, and $\boldsymbol{w}_k \in \boldsymbol{R}^n$ and $\boldsymbol{v}_k \in \boldsymbol{R}^m$ are the process noise and measurement noise. It is assumed that $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are mutually uncorrelated Gaussian white noises with the following statistical characteristics

$$\begin{cases} E(\boldsymbol{w}_k) = \boldsymbol{a}_k, \quad cov(\boldsymbol{w}_k, \boldsymbol{w}_j) = \boldsymbol{A}_k \delta_{kj} \\ E(\boldsymbol{v}_k) = \boldsymbol{b}_k, \quad cov(\boldsymbol{v}_k, \boldsymbol{v}_j) = \boldsymbol{B}_k \delta_{kj} \\ cov(\boldsymbol{w}_k, \boldsymbol{v}_j) = 0 \end{cases} \tag{4}$$

where $cov(\cdot)$ denotes the covariance function, both $\boldsymbol{A}_k \ge 0$ and $\boldsymbol{B}_k > 0$ are a symmetric matrix, and δ_{kj} is the $Kronecker - \delta$ function.

Equation (4) can be rewritten as

$$\begin{cases} \boldsymbol{a}_k = E\left[\boldsymbol{w}_k\right] = E\left[\boldsymbol{x}_{k+1} - f(\boldsymbol{x}_k)\right] \\ \boldsymbol{b}_k = E\left[\boldsymbol{v}_k\right] = E\left[\boldsymbol{y}_k - g(\boldsymbol{x}_k)\right] \\ \boldsymbol{A}_k = cov\left[\boldsymbol{w}_k \boldsymbol{w}_k^T\right] \\ \quad = E\left[\left(\boldsymbol{x}_{k+1} - f(\boldsymbol{x}_k) - \boldsymbol{a}_k\right)\left(\boldsymbol{x}_{k+1} - f(\boldsymbol{x}_k) - \boldsymbol{a}_k\right)^T\right] \\ \boldsymbol{B}_k = cov\left[\boldsymbol{v}_k \boldsymbol{v}_k^T\right] \\ \quad = E\left[\left(\boldsymbol{y}_k - g(\boldsymbol{x}_k) - \boldsymbol{b}_k\right)\left(\boldsymbol{y}_k - g(\boldsymbol{x}_k) - \boldsymbol{b}_k\right)^T\right] \end{cases} \tag{5}$$

B. ANALYSIS OF TRADITIONAL GAUSSIAN FILTERING

For the nonlinear system described by (3), under the condition that the state probability density function is subject to the Gaussian distribution, the traditional Gaussian filtering estimates the system state based on the linear update structure of the standard Kalman filter to calculate the first-order moment (mean) and second-order moment (variance). It includes the following steps:

Step 1. Initialize estimated state $\hat{\boldsymbol{x}}_0$ and its associated error covariance $\boldsymbol{P}_0$

$$\begin{aligned} \hat{\boldsymbol{x}}_0 &= E[\boldsymbol{x}_0] \\ \boldsymbol{P}_0 &= cov\left(\boldsymbol{x}_0, \boldsymbol{x}_0^{\mathrm{T}}\right) = E[(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)^{\mathrm{T}}] \end{aligned} \tag{6}$$

Step 2. Calculate the predicted state and its error covariance matrix

$$\begin{aligned} \hat{\boldsymbol{x}}_{k+1|k} &= E[f(\boldsymbol{x}_k)] \\ &= \int f(\boldsymbol{x}_k) N(\boldsymbol{x}_k; \hat{\boldsymbol{x}}_k, \boldsymbol{P}_k) d\boldsymbol{x}_k \end{aligned} \tag{7}$$

$$\begin{aligned} \boldsymbol{P}_{k+1|k} &= E[(\boldsymbol{x}_{k+1} - \hat{\boldsymbol{x}}_{k+1|k})(\boldsymbol{x}_{k+1} - \hat{\boldsymbol{x}}_{k+1|k})^{\mathrm{T}}] \\ &= \int f(\boldsymbol{x}_k) f^{\mathrm{T}}(\boldsymbol{x}_k) N(\boldsymbol{x}_k; \hat{\boldsymbol{x}}, \boldsymbol{P}_k) d\boldsymbol{x}_k \\ &\quad - \hat{\boldsymbol{x}}_{k+1|k} \hat{\boldsymbol{x}}_{k+1|k}^{\mathrm{T}} + \boldsymbol{A}_{k+1} \end{aligned} \tag{8}$$

Step 3. Calculate the predicted measurement and its error covariance matrix

$$\begin{aligned} \hat{\boldsymbol{y}}_{k+1|k} &= E\left[g\left(\boldsymbol{x}_{k+1}\right)\right] \\ &= \int g(\boldsymbol{x}_{k+1}) N(\boldsymbol{x}_{k+1}; \hat{\boldsymbol{x}}_{k+1|k}, \boldsymbol{P}_{k+1|k}) d\boldsymbol{x}_{k+1} \end{aligned} \tag{9}$$

$$\begin{aligned} \boldsymbol{P}_{\tilde{\boldsymbol{y}}_{k+1}} &= E[(\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1})(\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1})^{\mathrm{T}}] \\ &= \int g\left(\boldsymbol{x}_{k+1}\right) g^{\mathrm{T}}\left(\boldsymbol{x}_{k+1}\right) N \\ &\quad \times \left(\boldsymbol{x}_{k+1}; \hat{\boldsymbol{x}}_{k+1|k}, \boldsymbol{P}_{k+1|k}\right) d\boldsymbol{x}_{k+1} \\ &\quad - \hat{\boldsymbol{y}}_{k+1} \hat{\boldsymbol{y}}_{k+1}^{\mathrm{T}} + \boldsymbol{B}_{k+1} \end{aligned} \tag{10}$$

where $\tilde{\boldsymbol{y}}_{k+1} = \boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1|k}$.

Step 4. Calculate the cross covariance matrix between the predicted state and measurement

$$N(\boldsymbol{x}; \hat{\boldsymbol{x}}, \boldsymbol{P}) = \frac{exp\left(-\frac{1}{2}(\boldsymbol{x} - \hat{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{P}^{-1} (\boldsymbol{x} - \hat{\boldsymbol{x}})\right)}{((2\pi)^n \det \boldsymbol{P})^{1/2}} \tag{11}$$

In (6)-(11), $N(\boldsymbol{x}; \hat{\boldsymbol{x}}, \boldsymbol{P})$ represents the Gaussian distribution with mean $\hat{\boldsymbol{x}}$ and covariance $\boldsymbol{P}$, i.e.

$$N(\boldsymbol{x}; \hat{\boldsymbol{x}}, \boldsymbol{P}) = \frac{exp\left(-\frac{1}{2}(\boldsymbol{x} - \hat{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{P}^{-1} (\boldsymbol{x} - \hat{\boldsymbol{x}})\right)}{((2\pi)^n \det \boldsymbol{P})^{1/2}} \tag{12}$$

Based on the linear update structure of the standard Kalman filter, the state estimation is calculated by

$$\begin{cases} \hat{\boldsymbol{x}}_{k+1} = \hat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{K}_{k+1}(\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1}) \\ \boldsymbol{K}_{k+1} = \boldsymbol{P}_{\tilde{\boldsymbol{x}}_{k+1}\tilde{\boldsymbol{y}}_{k+1}} (\boldsymbol{P}_{\tilde{\boldsymbol{y}}_{k+1}})^{-1} \\ \boldsymbol{P}_{k+1} = \boldsymbol{P}_{k+1|k} - \boldsymbol{K}_{k+1} \boldsymbol{P}_{\tilde{\boldsymbol{y}}_{k+1}} \boldsymbol{K}_{k+1}^T \end{cases} \tag{13}$$

where $\tilde{\boldsymbol{x}}_{k+1} = \boldsymbol{x}_{k+1} - \hat{\boldsymbol{x}}_{k+1|k}$.

It is known from (7)-(11) that the Gaussian filtering requires the calculation of the following multi-dimensional Gaussian integral

$$I(f) = \int_{\boldsymbol{R}^n} f(\boldsymbol{x}) \exp\left(-\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}\right) d\boldsymbol{x} \tag{14}$$

However, it is difficult to find the exact analytical solution for the Gaussian integral (13). Unscented transformation and spherical-radial cubature rules are the commonly used numerical solutions to approximate the Gaussian integral, resulting in UKF and CKF.

It is clear from the above that the Gaussian filtering accuracy relies on the prior knowledge of the process and measurement noises. If the process noise $\boldsymbol{w}_k$ is inaccurate or unknown, $\boldsymbol{A}_k$ given by (5) will be inaccurate. This will further make the state error covariance $\boldsymbol{P}_{k+1|k}$ described by (8) inaccurate, thus leading the predicted state $\hat{\boldsymbol{x}}_{k+1|k}$ to be biased. Similarly, if the measurement noise $\boldsymbol{v}_k$ is inaccurate or unknown, the predicted measurement $\hat{\boldsymbol{y}}_{k+1|k}$ will also be biased.

IV. NOISE STATISTICS ESTIMATION

A. MAP ESTIMATION

Theorem 1: Assume $h = p\left[\boldsymbol{x}_{0:k+1}, \boldsymbol{a}, \boldsymbol{A}, \boldsymbol{b}, \boldsymbol{B}, \boldsymbol{y}_{1:k+1}\right]$ is a joint probability density function, i.e.

$$\begin{aligned} h &= p\left[\boldsymbol{x}_{0:k+1}, \boldsymbol{a}, \boldsymbol{A}, \boldsymbol{b}, \boldsymbol{B}, \boldsymbol{y}_{1:k+1}\right] \\ &= p\left[\boldsymbol{y}_{1:k+1} | \boldsymbol{x}_{0:k+1}, \boldsymbol{a}, \boldsymbol{A}, \boldsymbol{b}, \boldsymbol{B}\right] \\ &\quad \times p\left[\boldsymbol{x}_{0:k+1} | \boldsymbol{a}, \boldsymbol{A}, \boldsymbol{b}, \boldsymbol{B}\right] p\left[\boldsymbol{a}, \boldsymbol{A}, \boldsymbol{b}, \boldsymbol{B}\right] \end{aligned} \tag{15}$$

where $p[\boldsymbol{a}, \boldsymbol{A}, \boldsymbol{b}, \boldsymbol{B}]$ can be calculated from pre-defined information.

Then, the MAP estimates $\hat{\boldsymbol{a}}$, $\hat{\boldsymbol{A}}$, $\hat{\boldsymbol{b}}$ and $\hat{\boldsymbol{B}}$ of $\boldsymbol{a}$, $\boldsymbol{A}$, $\boldsymbol{b}$ and $\boldsymbol{B}$ at time point $k+1$ for both process noise statistics and

measurement noise statistics are expressed as

$$\begin{cases} \hat{\boldsymbol{a}}_{k+1} = \dfrac{1}{k+1}\displaystyle\sum_{j=0}^{k}\left\{\hat{\boldsymbol{x}}_{j+1} - f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_j}\right\} \\ \hat{\boldsymbol{A}}_{k+1} = \dfrac{1}{k+1}\displaystyle\sum_{j=0}^{k}\left\{\left[\hat{\boldsymbol{x}}_{j+1} - f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_j} - \boldsymbol{a}_{k+1}\right]\right. \\ \qquad \left.\times\left[\hat{\boldsymbol{x}}_{j+1} - f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_j} - \boldsymbol{a}_{k+1}\right]^{\mathrm{T}}\right\} \\ \qquad = \dfrac{1}{k+1}\displaystyle\sum_{j=0}^{k}\left\{\left[\hat{\boldsymbol{x}}_{j+1} - \hat{\boldsymbol{x}}_{j+1|j}\right]\left[\hat{\boldsymbol{x}}_{j+1} - \hat{\boldsymbol{x}}_{j+1|j}\right]^{\mathrm{T}}\right\} \end{cases} \tag{16}$$

and

$$\begin{cases} \hat{\boldsymbol{b}}_{k+1} = \dfrac{1}{k+1}\displaystyle\sum_{j=0}^{k}\left\{\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}}\right\} \\ \hat{\boldsymbol{B}}_{k+1} = \dfrac{1}{k+1}\displaystyle\sum_{j=0}^{k}\left\{\left[\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}} - \boldsymbol{b}_{k+1}\right]\right. \\ \qquad \left.\times\left[\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}} - \boldsymbol{b}_{k+1}\right]^{\mathrm{T}}\right\} \\ \qquad = \dfrac{1}{k+1}\displaystyle\sum_{j=0}^{k}\left\{\left[\boldsymbol{y}_{j+1} - \hat{\boldsymbol{y}}_{j+1|j}\right]\left[\boldsymbol{y}_{j+1} - \hat{\boldsymbol{y}}_{j+1|j}\right]^{\mathrm{T}}\right\} \end{cases} \tag{17}$$

where $f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_j}$ denotes the mathematical expectation of estimated state $\hat{\boldsymbol{x}}_j$ via nonlinear system function $f_j(\cdot)$; $g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}}$ denotes the mathematical expectation of predicted state $\hat{\boldsymbol{x}}_{j+1|j}$ via nonlinear measurement function $g_{j+1}(\cdot)$; and

$$\hat{\boldsymbol{x}}_{j+1|j} = f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_j} + \boldsymbol{a}_{k+1} \tag{18}$$

and

$$\hat{\boldsymbol{y}}_{j+1|j} = g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}} + \boldsymbol{b}_{k+1} \tag{19}$$

Proof of Theorem 1: The objective is to solve the conditional probability density function

$$\begin{aligned} h^* &= p\left[\boldsymbol{x}_{0:k+1},\ \boldsymbol{a},\ \boldsymbol{A},\ \boldsymbol{b},\ \boldsymbol{B}\,|\boldsymbol{y}_{1:k+1}\right] \\ &= \frac{p\left[\boldsymbol{x}_{0:k+1},\ \boldsymbol{a},\ \boldsymbol{A},\ \boldsymbol{b},\ \boldsymbol{B},\ \boldsymbol{y}_{1:k+1}\right]}{p\left[\boldsymbol{y}_{1:k+1}\right]} \end{aligned} \tag{20}$$

for the MAP estimates $\hat{\boldsymbol{a}}$, $\hat{\boldsymbol{A}}$, $\hat{\boldsymbol{b}}$ and $\hat{\boldsymbol{B}}$. As $\hat{\boldsymbol{a}}$, $\hat{\boldsymbol{A}}$, $\hat{\boldsymbol{b}}$ and $\hat{\boldsymbol{B}}$ are independent of the denominator $p\left[\boldsymbol{y}_{1:k+1}\right]$, solving (20) for $\hat{\boldsymbol{a}}$, $\hat{\boldsymbol{A}}$, $\hat{\boldsymbol{b}}$ and $\hat{\boldsymbol{B}}$ is actually a problem to maximize the probability density function $p\left[\boldsymbol{x}_{0:k+1}, \boldsymbol{a}, \boldsymbol{A}, \boldsymbol{b}, \boldsymbol{B}, \boldsymbol{y}_{1:k+1}\right]$.

Using the rule of probability multiplication, we have

$$\begin{aligned} &p\left[\boldsymbol{x}_{0:k+1}\,|\boldsymbol{a}, \boldsymbol{A}, \boldsymbol{b}, \boldsymbol{B}\right] \\ &= p\left[\boldsymbol{x}_0\right]\prod_{j=0}^{k} p\left[\boldsymbol{x}_{j+1}\,|\boldsymbol{x}_j, \boldsymbol{a}, \boldsymbol{A}\right] \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}\left|\boldsymbol{P}_0\right|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}\left\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\right\|^2_{\boldsymbol{P}_0^{-1}}\right\} \\ &\quad\times\prod_{j=0}^{k}\frac{1}{(2\pi)^{\frac{n}{2}}\left|\boldsymbol{A}\right|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}\left\|\boldsymbol{x}_{j+1} - f_j\left(\boldsymbol{x}_j\right) - \boldsymbol{a}\right\|^2_{\boldsymbol{A}^{-1}}\right\} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}+\frac{n(k+1)}{2}}}\left|\boldsymbol{P}_0\right|^{-\frac{1}{2}}\left|\boldsymbol{A}\right|^{-\frac{k+1}{2}}\exp\left\{-\frac{1}{2}\left(\left\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\right\|^2_{\boldsymbol{P}_0^{-1}}\right.\right. \\ &\quad\left.\left.+\sum_{j=0}^{k}\left\|\boldsymbol{x}_{j+1} - f_j\left(\boldsymbol{x}_j\right) - \boldsymbol{a}\right\|^2_{\boldsymbol{A}^{-1}}\right)\right\} \\ &= C_1\left|\boldsymbol{P}_0\right|^{-\frac{1}{2}}\left|\boldsymbol{A}\right|^{-\frac{k+1}{2}}\exp\left\{-\frac{1}{2}\left(\left\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\right\|^2_{\boldsymbol{P}_0^{-1}}\right.\right. \\ &\quad\left.\left.+\sum_{j=0}^{k}\left\|\boldsymbol{x}_{j+1} - f_j\left(\boldsymbol{x}_j\right) - \boldsymbol{a}\right\|^2_{\boldsymbol{A}^{-1}}\right)\right\} \end{aligned} \tag{21}$$

where n denotes the size of the system state, the symbols "$|\cdot|$" and "$\|\cdot\|$" represent the determinant of a matrix and the magnitude of a vector, and $C_1 = \frac{1}{(2\pi)^{n(k+2)/2}}$.

If $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_{k+1}$ are the uncorrelated measurements, then

$$\begin{aligned} &p\left[\boldsymbol{y}_{1:k+1}\,|\boldsymbol{x}_{0:k+1}, \boldsymbol{a}, \boldsymbol{A}, \boldsymbol{b}, \boldsymbol{B}\right] \\ &= \prod_{j=0}^{k} p\left[\boldsymbol{y}_{j+1}\,|\boldsymbol{x}_{j+1}, \boldsymbol{b}, \boldsymbol{B}\right] \\ &= \prod_{j=0}^{k}\frac{1}{(2\pi)^{\frac{m}{2}}\left|\boldsymbol{B}\right|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}\left\|\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right) - \boldsymbol{b}\right\|^2_{\boldsymbol{B}^{-1}}\right\} \\ &= C_2\left|\boldsymbol{B}\right|^{\frac{k+1}{2}}\exp\left\{-\frac{1}{2}\sum_{j=0}^{k}\left\|\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right) - \boldsymbol{b}\right\|^2_{\boldsymbol{B}^{-1}}\right\} \end{aligned} \tag{22}$$

where m denotes the size of the system measurement and $C_2 = \frac{1}{(2\pi)^{m(k+1)/2}}$.

Replacing (21) and (22) into (15) yields

$$\begin{aligned} h &= C_1C_2\left|\boldsymbol{P}_0\right|^{-\frac{1}{2}}\cdot\left|\boldsymbol{A}\right|^{-\frac{k+1}{2}}\cdot\left|\boldsymbol{B}\right|^{-\frac{k+1}{2}}\cdot p\left[\boldsymbol{a}, \boldsymbol{A}, \boldsymbol{b}, \boldsymbol{B}\right] \\ &\quad\cdot\exp\left\{-\frac{1}{2}\left[\left\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\right\|^2_{\boldsymbol{P}_0^{-1}} + \sum_{j=0}^{k}\left\|\boldsymbol{x}_{j+1} - f_j\left(\boldsymbol{x}_j\right) - \boldsymbol{a}\right\|^2_{\boldsymbol{A}^{-1}}\right.\right. \\ &\quad\left.\left.+\sum_{j=0}^{k}\left\|\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right) - \boldsymbol{b}\right\|^2_{\boldsymbol{B}^{-1}}\right]\right\} \\ &= C\left|\boldsymbol{A}\right|^{-\frac{k+1}{2}}\cdot\left|\boldsymbol{B}\right|^{-\frac{k+1}{2}}\exp\left\{-\frac{1}{2}\left[\sum_{j=0}^{k}\left\|\boldsymbol{x}_{j+1} - f_j\left(\boldsymbol{x}_j\right) - \boldsymbol{a}\right\|^2_{\boldsymbol{A}^{-1}}\right.\right. \\ &\quad\left.\left.+\sum_{j=0}^{k}\left\|\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right) - \boldsymbol{b}\right\|^2_{\boldsymbol{B}^{-1}}\right]\right\} \end{aligned} \tag{23}$$

where $C = C_1C_2\left|\boldsymbol{p}_0\right|^{-\frac{1}{2}}p\left[\boldsymbol{a}, \boldsymbol{A}, \boldsymbol{b}, \boldsymbol{B}\right]\exp\left\{-\frac{1}{2}\left\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\right\|^2_{\boldsymbol{P}_0^{-1}}\right\}$.

Applying the natural logarithm to (23) generates

$$\begin{aligned}\ln h = &-\frac{k+1}{2}\ln|\boldsymbol{A}| - \frac{k+1}{2}\ln|\boldsymbol{B}| \\ &-\frac{1}{2}\sum_{j=0}^{k}\left\|\boldsymbol{x}_{j+1} - f_j\left(\boldsymbol{x}_j\right) - \boldsymbol{a}\right\|_{\boldsymbol{A}^{-1}}^2 \\ &-\sum_{j=0}^{k}\left\|\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right) - \boldsymbol{b}\right\|_{\boldsymbol{B}^{-1}}^2 + \ln C \end{aligned} \tag{24}$$

Thus, solving (25) yields (16) (the detailed derivations are described in Appendices 1 and 2).

$$\begin{cases} \dfrac{\partial \ln h}{\partial \boldsymbol{a}}\Big|_{\boldsymbol{a}=\boldsymbol{a}_{k+1}} = 0 \\ \dfrac{\partial \ln h}{\partial \boldsymbol{A}}\Big|_{\boldsymbol{A}=\boldsymbol{A}_{k+1}} = 0 \end{cases} \tag{25}$$

Similarly, solving (26) yields (17) (the detailed derivations are described in Appendices 3 and 4).

$$\begin{cases} \dfrac{\partial \ln h}{\partial \boldsymbol{b}}\Big|_{\boldsymbol{b}=\boldsymbol{b}_{k+1}} = 0 \\ \dfrac{\partial \ln h}{\partial \boldsymbol{B}}\Big|_{\boldsymbol{B}=\boldsymbol{B}_{k+1}} = 0 \end{cases} \tag{26}$$

The proof of Theorem 1 is completed.

The statistics of both process and measurement noises can be online estimated by (16) and (17), and further fed back to the traditional Gaussian filtering procedure for nonlinear state estimation. However, as shown in (16) and (17), the system noise statistics are computed by the arithmetic mean with the common weight $1/k+1$. This means that for each kind of noise statistics, its estimates at all time points have equal contributions to the evaluations of prediction errors. Thus, the calculated prediction errors are not capable of precisely characterizing the actual noise characteristics, resulting in degraded estimation accuracy.

B. RANDOM WEIGHTING MAP ESTIMATION

Based on the principle of random weighting described in Section II, from (16) and (17) the random weighting estimates $\hat{\boldsymbol{a}}^*, \hat{\boldsymbol{A}}^*, \hat{\boldsymbol{b}}^*$ and $\hat{\boldsymbol{B}}^*$ of $\boldsymbol{a}, \boldsymbol{A}, \boldsymbol{b}$ and $\boldsymbol{B}$ at time point $k+1$ for both process noise statistics and measurement noise statistics can be written as

$$\hat{\boldsymbol{a}}_{k+1}^* = \sum_{j=0}^{k}\sigma_j\left[\hat{\boldsymbol{x}}_{j+1} - f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_j}\right] \tag{27}$$

$$\begin{aligned}\hat{\boldsymbol{A}}_{k+1}^* &= \sum_{j=0}^{k}\sigma_j\left\{\left[\hat{\boldsymbol{x}}_{j+1} - f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_j} - \boldsymbol{a}_{k+1}\right]\right. \\ &\quad \left.\left[\hat{\boldsymbol{x}}_{j+1} - f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_j} - \boldsymbol{a}_{k+1}\right]^{\mathrm{T}}\right\} \\ &= \sum_{j=0}^{k}\sigma_j\left\{\left[\hat{\boldsymbol{x}}_{j+1} - \hat{\boldsymbol{x}}_{j+1|j}\right]\left[\hat{\boldsymbol{x}}_{j+1} - \hat{\boldsymbol{x}}_{j+1|j}\right]^{\mathrm{T}}\right\} \end{aligned} \tag{28}$$

and

$$\hat{\boldsymbol{b}}_{k+1}^* = \sum_{j=0}^{k}\sigma_j\left[\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}}\right] \tag{29}$$

$$\begin{aligned}\hat{\boldsymbol{B}}_{k+1}^* &= \sum_{j=0}^{k}\sigma_j\left\{\left[\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}} - \boldsymbol{b}_{k+1}\right]\right. \\ &\quad \left.\left[\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}} - \boldsymbol{b}_{k+1}\right]^{\mathrm{T}}\right\} \\ &= \sum_{j=0}^{k}\sigma_j\left\{\left[\boldsymbol{y}_{j+1} - \hat{\boldsymbol{y}}_{j+1|j}\right]\left[\boldsymbol{y}_{j+1} - \hat{\boldsymbol{y}}_{j+1|j}\right]^{\mathrm{T}}\right\} \end{aligned} \tag{30}$$

where $(\sigma_1, \sigma_2, \cdots, \sigma_n)$ are the random weights subject to Dirichlet distribution $D(1, 1, \cdots, 1)$.

Remark: When implementing the algorithm described by (27)-(30), if the state estimate at time point $k+1$ cannot be obtained, it can be substituted with the state prediction at time point $k+1$ to calculate the system noise statistics.

Theorem 2: The random weighting estimates of the first-order noise statistics, i.e., $\hat{\boldsymbol{a}}_{k+1}^*$ and $\hat{\boldsymbol{b}}_{k+1}^*$ given by (27) and (29), are unbiased. The random weighting estimates of the second-order noise statistics, i.e., $\hat{\boldsymbol{A}}_{k+1}^*$ and $\hat{\boldsymbol{B}}_{k+1}^*$ given by (28) and (30), are sub-optimal unbiased.

Proof of Theorem 2: Define the innovation vector by

$$\boldsymbol{\varepsilon}_{j+1} = \boldsymbol{y}_{j+1} - \hat{\boldsymbol{y}}_{j+1|j} \tag{31}$$

From (31), we have

$$\begin{cases} E\left(\boldsymbol{\varepsilon}_{j+1}\right) = E\left(\boldsymbol{y}_{j+1} - \hat{\boldsymbol{y}}_{j+1|j}\right) = 0 \\ E\left(\boldsymbol{\varepsilon}_{j+1}\boldsymbol{\varepsilon}_{j+1}^{\mathrm{T}}\right) = E\left[\left(\boldsymbol{y}_{j+1} - \hat{\boldsymbol{y}}_{j+1|j}\right)\left(\boldsymbol{y}_{j+1} - \hat{\boldsymbol{y}}_{j+1|j}\right)^{\mathrm{T}}\right] \\ \qquad\qquad = \boldsymbol{P}_{\tilde{\boldsymbol{y}}_{j+1}} \end{cases} \tag{32}$$

From (13), we have

$$\boldsymbol{P}_{j+1|k} - \boldsymbol{P}_{j+1} = \boldsymbol{K}_{j+1}\boldsymbol{P}_{\tilde{\boldsymbol{y}}_{j+1}}\boldsymbol{K}_{j+1}^T \tag{33}$$

and

$$\boldsymbol{P}_{j+1|k} - \boldsymbol{P}_{j+1} = \boldsymbol{K}_{j+1}\boldsymbol{P}_{\tilde{\boldsymbol{y}}_{j+1}}\boldsymbol{K}_{j+1}^T \tag{34}$$

Substituting (18), (13), (31) and (32) into (27), we have

$$\begin{aligned} E\left(\hat{\boldsymbol{a}}_{k+1}^*\right) &= \sum_{j=0}^{k}\sigma_j E\left\{\left[\hat{\boldsymbol{x}}_{j+1} - f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_j}\right]\right\} \\ &= \sum_{j=0}^{k}\sigma_j E\left(\hat{\boldsymbol{x}}_{j+1} - \hat{\boldsymbol{x}}_{j+1|j} + \boldsymbol{a}_{k+1}\right) \\ &= \sum_{j=0}^{k}\sigma_j E\left[\boldsymbol{K}_{j+1}\left(\boldsymbol{y}_{j+1} - \hat{\boldsymbol{y}}_{j+1|j}\right) + \boldsymbol{a}_{k+1}\right] \\ &= \sum_{j=0}^{k}\sigma_j E\left(\boldsymbol{K}_{j+1}\boldsymbol{\varepsilon}_{j+1} + \boldsymbol{a}_{k+1}\right) \\ &= \sum_{j=0}^{k}\sigma_j\boldsymbol{a}_{k+1} \\ &= \boldsymbol{a}_{k+1} \end{aligned} \tag{35}$$

where $\sum_{j=0}^{k}\sigma_j = 1$ is used in the last derivation step.

Similarly, substituting (19), (31) and (32) into (29) yields

$$\begin{aligned} E\left(\hat{\boldsymbol{b}}_{k+1}^{*}\right) &= \sum_{j=0}^{k}\sigma_j E\left\{\left[\boldsymbol{y}_{j+1}-g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}}\right]\right\} \\ &= \sum_{j=0}^{k}\sigma_j E\left(\boldsymbol{y}_{j+1}-\hat{\boldsymbol{y}}_{j+1|j}+\boldsymbol{b}_{k+1}\right) \\ &= \sum_{j=0}^{k}\sigma_j E\left(\varepsilon_{j+1}+\boldsymbol{b}_{k+1}\right) \\ &= \sum_{j=0}^{k}\sigma_j \boldsymbol{b}_{k+1} \\ &= \boldsymbol{b}_{k+1} \end{aligned} \tag{36}$$

where $\sum_{j=0}^{k}\sigma_j = 1$ is used in the last derivation step.

It is known from (35) and (36) that the random weighting estimates $\hat{\boldsymbol{a}}_{k+1}^{*}$ and $\hat{\boldsymbol{b}}_{k+1}^{*}$ of the first-order noise statistics are unbiased.

Define

$$\begin{aligned} \boldsymbol{\rho}_j &= f_j\left(\boldsymbol{x}_j\right)-f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_j} \\ &= f_j\left(\boldsymbol{x}_j\right)-E\left[f_j\left(\boldsymbol{x}_j\right)\big|\boldsymbol{y}_j\right] \end{aligned} \tag{37}$$

Then, we have [23]

$$\boldsymbol{P}_{j+1|j} = E\left(\boldsymbol{\rho}_j\boldsymbol{\rho}_j^{\mathrm{T}}\right)+\boldsymbol{A}_{k+1} \tag{38}$$

Substituting (13), (31)-(34) and (38) into (28) yields

$$\begin{aligned} &E\left(\hat{\boldsymbol{A}}_{k+1}^{*}\right) \\ &= \sum_{j=0}^{k}\sigma_j E\left\{\left[\hat{\boldsymbol{x}}_{j+1}-f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_j}-\boldsymbol{a}_{k+1}\right]\right. \\ &\qquad \left.\times\left[\hat{\boldsymbol{x}}_{j+1}-f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_j}-\boldsymbol{a}_{k+1}\right]^{\mathrm{T}}\right\} \\ &= \sum_{j=0}^{k}\sigma_j E\left\{\left[\hat{\boldsymbol{x}}_{j+1}-\hat{\boldsymbol{x}}_{j+1|j}\right]\left[\hat{\boldsymbol{x}}_{j+1}-\hat{\boldsymbol{x}}_{j+1|j}\right]^{\mathrm{T}}\right\} \\ &= \sum_{j=0}^{k}\sigma_j E\left\{\left[\boldsymbol{K}_{j+1}\left(\boldsymbol{y}_{j+1}-\hat{\boldsymbol{y}}_{j+1|j}\right)\right]\left[\boldsymbol{K}_{j+1}\left(\boldsymbol{y}_{j+1}-\hat{\boldsymbol{y}}_{j+1|j}\right)\right]^{\mathrm{T}}\right\} \\ &= \sum_{j=0}^{k}\sigma_j\left[\boldsymbol{K}_{j+1}E\left(\boldsymbol{\varepsilon}_{j+1}\boldsymbol{\varepsilon}_{j+1}^{\mathrm{T}}\right)\boldsymbol{K}_{j+1}^{\mathrm{T}}\right] \\ &= \sum_{j=0}^{k}\sigma_j E\left[\boldsymbol{K}_{j+1}\boldsymbol{P}_{\tilde{\boldsymbol{y}}_{j+1}}\boldsymbol{K}_{j+1}^{\mathrm{T}}\right] \\ &= \sum_{j=0}^{k}\sigma_j E\left[\boldsymbol{P}_{j+1|j}-\boldsymbol{P}_{j+1}\right] \\ &= \sum_{j=0}^{k}\sigma_j\left[E\left(\boldsymbol{\rho}_j\boldsymbol{\rho}_j^{\mathrm{T}}\right)+\boldsymbol{A}_{k+1}-\boldsymbol{P}_{j+1}\right] \\ &= \sum_{j=0}^{k}\sigma_j\left[E\left(\boldsymbol{\rho}_j\boldsymbol{\rho}_j^{\mathrm{T}}\right)-\boldsymbol{P}_{j+1}\right]+\boldsymbol{A}_{k+1} \\ &\neq \boldsymbol{A}_{k+1} \end{aligned} \tag{39}$$

Equation (39) shows that the random weighting estimate of the second-order process noise statistic, i.e., $\hat{\boldsymbol{A}}_{k+1}^{*}$ described by (28), is biased. However, (28) can be improved as the following unbiased estimate

$$\hat{\boldsymbol{A}}_{k+1}^{*} = \sum_{j=0}^{k}\sigma_j\left[\boldsymbol{K}_{j+1}\boldsymbol{P}_{\tilde{\boldsymbol{y}}_{j+1}}\boldsymbol{K}_{j+1}^{T}+\boldsymbol{P}_{j+1}-E\left(\rho_{j+1}\rho_{j+1}^{T}\right)\right] \tag{40}$$

owing to

$$\begin{aligned} &E\left(\hat{\boldsymbol{A}}_{k+1}^{*}\right) \\ &= E\left(\sum_{j=0}^{k}\sigma_j\left[\boldsymbol{K}_{j+1}\boldsymbol{P}_{\tilde{\boldsymbol{y}}_{j+1}}\boldsymbol{K}_{j+1}^{\mathrm{T}}+\boldsymbol{P}_{j+1}-E\left(\boldsymbol{\rho}_{j+1}\boldsymbol{\rho}_{j+1}^{\mathrm{T}}\right)\right]\right) \\ &= \sum_{j=0}^{k}\sigma_j\left[E\left(\boldsymbol{K}_{j+1}\boldsymbol{P}_{\tilde{\boldsymbol{y}}_{j+1}}\boldsymbol{K}_{j+1}^{\mathrm{T}}\right)+E\left(\boldsymbol{P}_{j+1}\right)-E\left(\boldsymbol{\rho}_{j+1}\boldsymbol{\rho}_{j+1}^{\mathrm{T}}\right)\right] \\ &= \sum_{j=0}^{k}\sigma_j\left[E\left(\boldsymbol{P}_{j+1|j}-\boldsymbol{P}_{j+1}\right)+E\left(\boldsymbol{P}_{j+1}\right)-E\left(\boldsymbol{\rho}_{j+1}\boldsymbol{\rho}_{j+1}^{T}\right)\right] \\ &= \sum_{j=0}^{k}\sigma_j\left[E\left(\boldsymbol{\rho}_j\boldsymbol{\rho}_j^{\mathrm{T}}\right)+\boldsymbol{A}_{k+1}-\boldsymbol{P}_{j+1}+\boldsymbol{P}_{j+1}-E\left(\boldsymbol{\rho}_{j+1}\boldsymbol{\rho}_{j+1}^{\mathrm{T}}\right)\right] \\ &= A_{k+1} \end{aligned} \tag{41}$$

Thus, the random weighting estimate of the second-order process noise statistic, which is described by (28), is sub-optimal unbiased.

Define

$$\begin{aligned} \boldsymbol{\zeta}_j &= g_{j+1}\left(\boldsymbol{x}_{j+1}\right)-g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}} \\ &= g_{j+1}\left(\boldsymbol{x}_{j+1}\right)-E\left[g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\big|\boldsymbol{y}_j\right] \end{aligned} \tag{42}$$

Thus, we have [23]

$$\boldsymbol{P}_{\tilde{\boldsymbol{y}}_{j+1}} = E\left(\boldsymbol{\zeta}_{j+1}\boldsymbol{\zeta}_{j+1}^{\mathrm{T}}\right)+\boldsymbol{B}_{k+1} \tag{43}$$

Substituting (31), (32) and (43) into (30) yields

$$\begin{aligned} E\left(\hat{\boldsymbol{B}}_{k+1}^{*}\right) &= \sum_{j=0}^{k}\sigma_j E\left\{\left[\boldsymbol{y}_{j+1}-g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}}-\boldsymbol{b}_{k+1}\right]\right. \\ &\qquad \left.\left[\boldsymbol{y}_{j+1}-g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}}-\boldsymbol{b}_{k+1}\right]^{\mathrm{T}}\right\} \\ &= \sum_{j=0}^{k}\sigma_j E\left\{\left[\boldsymbol{y}_{j+1}-\hat{\boldsymbol{y}}_{j+1|j}\right]\left[\boldsymbol{y}_{j+1}-\hat{\boldsymbol{y}}_{j+1|j}\right]^{\mathrm{T}}\right\} \\ &= \sum_{j=0}^{k}\sigma_j E\left(\boldsymbol{\varepsilon}_{j+1}\boldsymbol{\varepsilon}_{j+1}^{\mathrm{T}}\right) \\ &= \sum_{j=0}^{k}\sigma_j\boldsymbol{P}_{\tilde{\boldsymbol{y}}_{j+1}} \\ &= \sum_{j=0}^{k}\sigma_j\left[E\left(\boldsymbol{\zeta}_{j+1}\boldsymbol{\zeta}_{j+1}^{\mathrm{T}}\right)+\boldsymbol{B}_{k+1}\right] \\ &\neq \boldsymbol{B}_{k+1} \end{aligned} \tag{44}$$

Equation (44) shows that the random weighting estimate of the second-order measurement noise statistic, i.e., $\hat{\boldsymbol{B}}^*_{k+1}$ described by (30), is biased. However, (30) can be improved as the following unbiased estimate

$$\hat{\boldsymbol{B}}^*_{k+1} = \sum_{j=0}^{k} \sigma_j \left[\boldsymbol{P}_{\tilde{\boldsymbol{y}}_{j+1}} - E\left(\boldsymbol{\zeta}_{j+1}\boldsymbol{\zeta}^T_{j+1}\right)\right] \tag{45}$$

owing to

$$\begin{aligned} E\left(\hat{\boldsymbol{B}}^*_{k+1}\right) &= E\left\{\sum_{j=0}^{k} \sigma_j \left[\boldsymbol{P}_{\tilde{\boldsymbol{y}}_{j+1}} - E\left(\boldsymbol{\zeta}_{j+1}\boldsymbol{\zeta}^{\boldsymbol{T}}_{j+1}\right)\right]\right\} \\ &= E\sum_{j=0}^{k} \sigma_j \left\{ E\left(\boldsymbol{\zeta}_{j+1}\boldsymbol{\zeta}^{\mathrm{T}}_{j+1}\right) + \boldsymbol{B}_{k+1} \right. \\ &\quad \left. - \left[E\left(\boldsymbol{\zeta}_{j+1}\boldsymbol{\zeta}^{\boldsymbol{T}}_{j+1}\right)\right]\right\} \\ &= \boldsymbol{B}_{k+1} \end{aligned} \tag{46}$$

Thus, the random weighting estimate of the second-order process noise statistic, which is described by (30) is sub-optimal unbiased.

The proof of Theorem 2 is completed.

C. DESIGN OF RANDOM WEIGHTS

Define the residual vector of the state prediction by

$$\Delta \boldsymbol{x}_{k-j} = \hat{\boldsymbol{x}}_{k-j} - \hat{\boldsymbol{x}}_{k-j|k-j-1} \quad (j = 1,2,\cdots,n) \tag{47}$$

Define the residual vector of the measurement by

$$\Delta z_{k-j} = \hat{z}_{k-j} - z_{k-j} \quad (j = 1,2,\cdots,n) \tag{48}$$

where $\hat{z}_{k-j} = g_{k-j}\left(\hat{\boldsymbol{x}}_{k-j}\right)$.

In case of a change in the process noise statistics, the contribution of the state prediction $\hat{\boldsymbol{x}}_{k-j|k-j-1}$ to the state estimate will be decreased, leading the prediction to be biased. As a result, the magnitude of the state prediction's residual vector $\Delta \boldsymbol{x}_{k-j}$ will be increased. Similarly, in case of a change in the measurement noise statistics, the measurement's residual Δz_{k-j} will be biased and its magnitude will also be increased.

$$\sigma_j \infty \left\|\Delta \boldsymbol{x}_{k-j}\right\| \times \left\|\Delta z_{k-j}\right\| (j = 1,2,\cdots,n) \tag{49}$$

To characterize the changes of the system noise statistics, consider where $\left\|\Delta \boldsymbol{x}_{k-j}\right\| = \sqrt{\Delta \boldsymbol{x}^{\mathrm{T}}_{k-j}\Delta \boldsymbol{x}_{k-j}}$, $\left\|\Delta z_{k-j}\right\| = \sqrt{\Delta z^{\mathrm{T}}_{k-j}\Delta z_{k-j}}$, and the symbol "$\infty$" denotes the proportional operation.

Equation (49) implies that the larger the value of $\left\|\Delta \boldsymbol{x}_{k-j}\right\| \times \left\|\Delta z_{k-j}\right\|$ is, the larger the weight is. Therefore, the random weights σ_j can be determined as follows.

Let

$$\omega_j = \left\|\Delta \boldsymbol{x}_{k-j}\right\| \times \left\|\Delta z_{k-j}\right\| \quad (j = 1,2,\cdots,n) \tag{50}$$

By normalization, the random weights are obtained as

$$\sigma_j = \frac{\omega_j}{\sum_{j=1}^{n} \omega_j} \tag{51}$$

V. RANDOM WEIGHTED-BASED GAUSSIAN FILTERING

The proposed random weighting-based Gaussian filtering method includes the following steps:

(i) Initialize the estimated state and its associated error covariance

$$\begin{aligned} \hat{\boldsymbol{x}}_0 &= E[\boldsymbol{x}_0] \\ \boldsymbol{P}_0 &= cov\left(\boldsymbol{x}_0, \boldsymbol{x}^{\mathrm{T}}_0\right) = E[(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)^{\mathrm{T}}] \end{aligned} \tag{52}$$

(ii) Calculate the predicted state and its error covariance matrix

$$\hat{\boldsymbol{x}}^*_{k+1|k} = f_k\left(\boldsymbol{x}_k\right)\big|_{\boldsymbol{x}_k \leftarrow \hat{\boldsymbol{x}}_k} + \hat{\boldsymbol{a}}^*_{k+1} \tag{53}$$

$$\boldsymbol{P}^*_{k+1|k} = E(\boldsymbol{\rho}_k \boldsymbol{\rho}^{\mathrm{T}}_k) + \hat{\boldsymbol{A}}^*_{k+1} \tag{54}$$

(iii) Calculate predicted measurement and its error covariance matrix

$$\begin{cases} \hat{\boldsymbol{y}}^*_{k+1|k} = g_{k+1}\left(\boldsymbol{x}_{k+1}\right)\big|_{\boldsymbol{x}_{k+1} \leftarrow \hat{\boldsymbol{x}}_{k+1}} + \hat{\boldsymbol{b}}^*_k \\ \boldsymbol{P}^*_{\tilde{\boldsymbol{y}}_{k+1}} = E\left(\boldsymbol{\zeta}_{k+1}\boldsymbol{\zeta}^T_{k+1}\right) + \hat{\boldsymbol{B}}^*_k \\ \boldsymbol{P}^*_{\tilde{\boldsymbol{x}}_{k+1}\tilde{\boldsymbol{y}}_{k+1}} = E[\tilde{\boldsymbol{x}}_{k+1}\zeta^T_{k+1}] \end{cases} \tag{55}$$

(iv) State update

$$\begin{cases} \hat{\boldsymbol{x}}^*_{k+1} = \hat{\boldsymbol{x}}^*_{k+1|k} + \boldsymbol{K}_{k+1}\left(\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}^*_{k+1|k}\right) \\ \boldsymbol{K}_{k+1} = \boldsymbol{P}^*_{\tilde{\boldsymbol{x}}_{k+1}\tilde{\boldsymbol{y}}_{k+1}} \boldsymbol{P}^{*-1}_{\tilde{\boldsymbol{y}}_{k+1}} \\ \boldsymbol{P}^*_{k+1} = \boldsymbol{P}^*_{k+1|k} - \boldsymbol{K}_{k+1}\boldsymbol{P}^*_{\tilde{\boldsymbol{y}}_{k+1}}\boldsymbol{K}^T_{k+1} \end{cases} \tag{56}$$

Equations (53)-(56) show that the presented Gaussian filtering method based on random weighting can adaptively adjust the weights of the process and measurement noise statistics to restrain system noise interferences on the state estimation, thus leading to improved estimation precision. Fig. 1 illustrates the algorithm of the proposed random weighting-based Gaussian filtering.

VI. PERFORMANCE EVALUATION AND DISCUSSION

A prototype system was implemented using the proposed random weighting-based Gaussian filtering (RWGF) for nonlinear state estimation. Simulations and experimental analysis were conducted to comprehensively evaluate and analyze the performance of the proposed RWGF. Comparison analysis with the traditional Gaussian filtering method such as CKF for system state estimation was also conducted to demonstrate the improved performance of the proposed RWGF.

A. SIMULATION AND ANALYSIS

Simulation analysis was performed to examine the effectiveness of the presented RWGF. Consider the univariate nonstationary growth model [18]

$$\begin{cases} \boldsymbol{x}_k 0.5\boldsymbol{x}_{k-1} + 25\boldsymbol{x}_{k-1}/(1 + \boldsymbol{x}^2_{k-1}) \\ \qquad +8\cos[1.2(k-1)] + \boldsymbol{w}_{k-1} \\ z_k = \boldsymbol{x}^2_k/20 + \boldsymbol{v}_k \end{cases} \tag{57}$$

where both $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are a Gaussian white noise of non-zero mean.

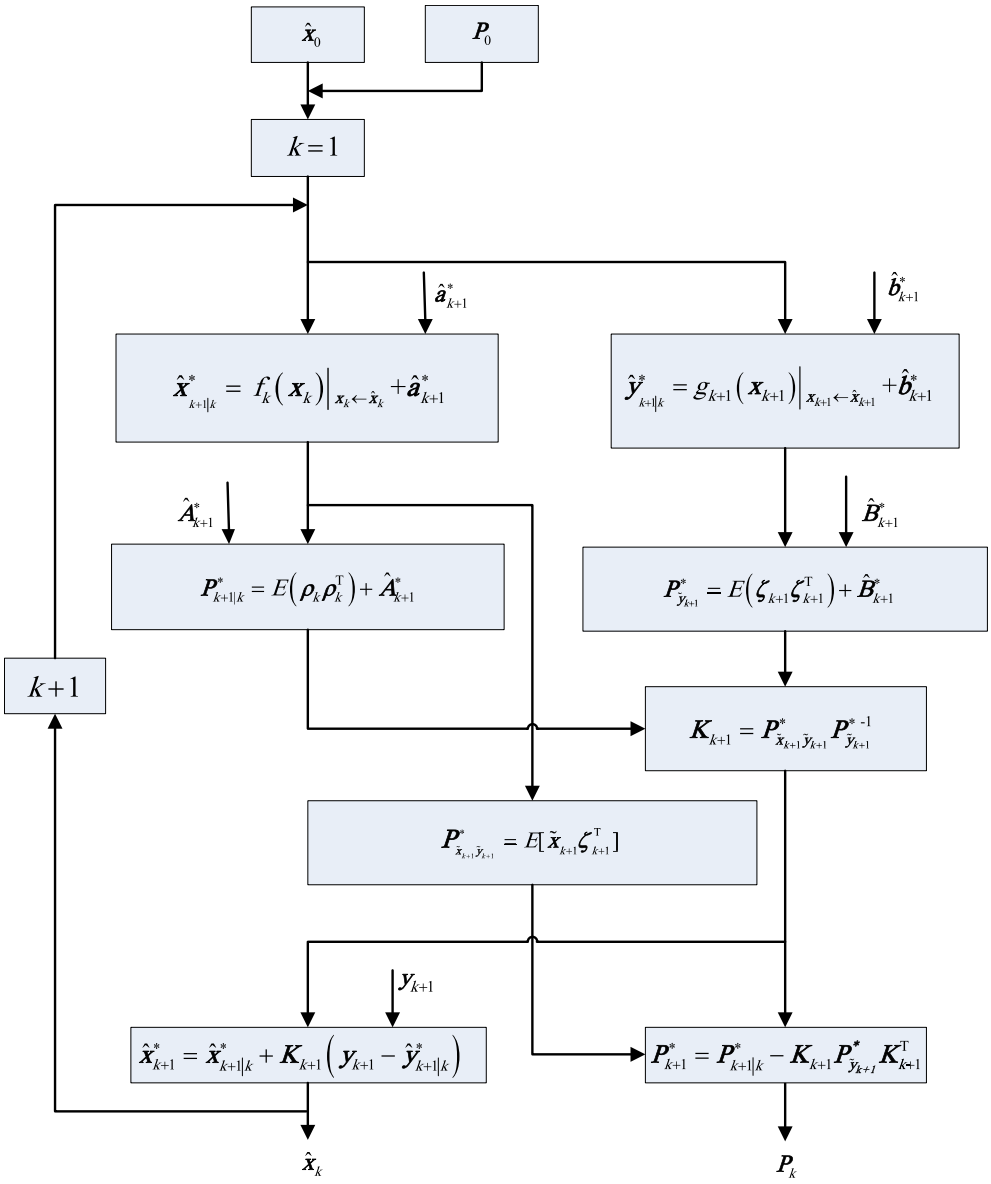


FIGURE 1. The random-weighting Gaussian filtering algorithm.

The initial state was $x_0 = 0.1$, and its estimate was $\hat{x}_0 = 0.1$. The initial estimation error covariance P_0 was set as a unit matrix. Monte Carlo simulations were carried out 150 times.

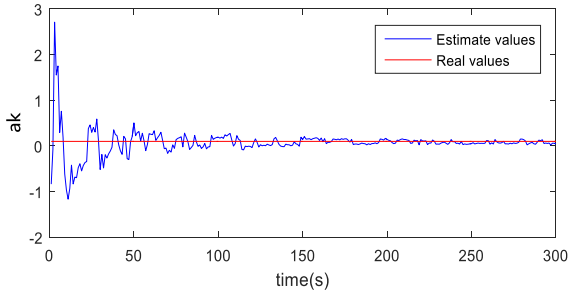
1) ESTIMATION OF PROCESS NOISE STATISTICS

To examine the estimation accuracy of RWGF for the statistics of system process noise, the measurement noise statistics were assumed to be exactly known and they were $b_k = 0$, $\hat{b}_k = 0$, $B_k = 1$ and $\hat{B}_k = 1$. The process noise's mean and covariance were $a_k = 0.1$ and $A_k = 20$ in theory. The initial process noise statistics were $\hat{a}_0 = 0.05$ and $\hat{A}_0 = 4$, which were biased from their theoretical values.

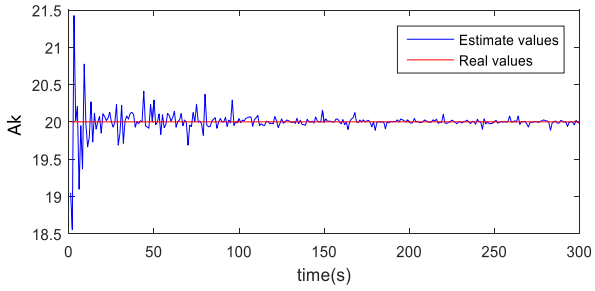
Fig. 2 shows the estimations of process noise's mean and covariance by RWGF, where after the initial oscillations within about 80s, the estimated process noise statistics are gradually converged to their theoretical values, respectively. Fig. 3 shows the state estimation errors by both CKF and RWGF in the presence of the biased process noise statistics. Since RWGF can online estimate the process noise's statistics while CKF lacks this capability, the state estimate error by RWGF is much smaller than that by CKF, in spite of the disturbances of the biased process noise statistics. As shown in Table 1, the mean error and root mean square error (RMSE) are 1.0584 and 1.4443 for CKF, while 0.5351 and 0.6925 for RWGF.

2) MEASUREMENT NOISE STATISTICS

To examine the estimation accuracy of RWGF for measurement noise statistics, the process noise statistics were

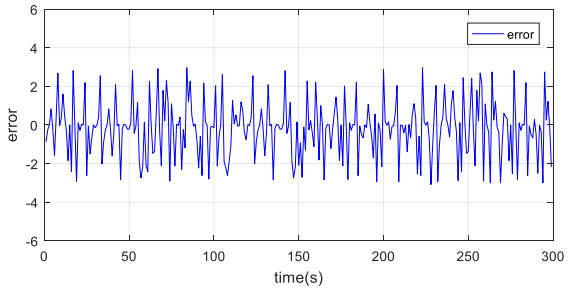


(a) Estimated process noise mean

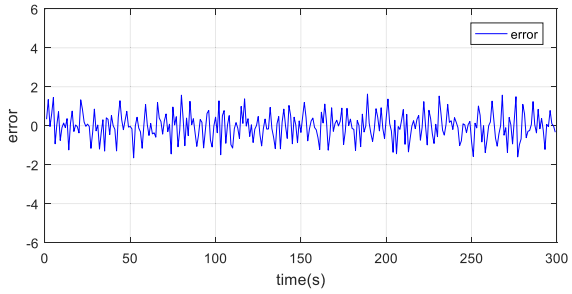


(b) Estimated process noise covariance

FIGURE 2. Estimated process noise statistics by RWGF.



(a) State estimation error by CKF



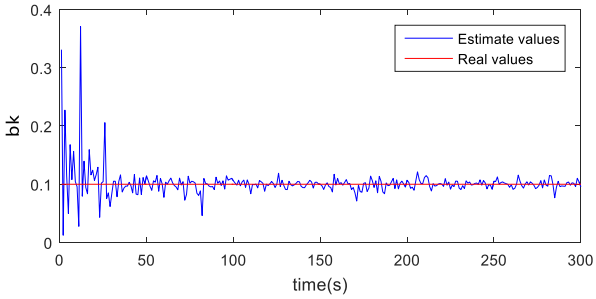
(b) State estimation error by RWGF

FIGURE 3. State estimation errors of both CKF and RWGF under biased process noise statistics.

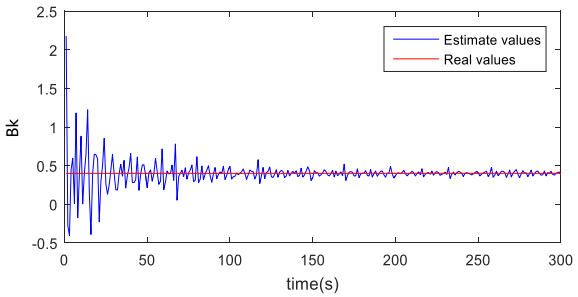
TABLE 1. Statistical errors of state estimation by both CKF and RWGF under biased process noise statistics.

Method	Mean error	RMSE
CKF	1.0584	1.4443
RWGF	0.5351	0.6925

assumed to be exactly known and they were $a_k = 0$, $\hat{a}_k = 0$, $A_k = 5$ and $\hat{A}_k = 1$. The measurement noise's mean and covariance were $b_k = 0.1$ and $B_k = 0.4$ in theory. The initial



(a) Estimated measurement noise mean



(b) Estimated measurement noise covariance

FIGURE 4. Estimated measurement noise statistics by RWGF.

TABLE 2. Statistical errors of state estimation by both CKF and RWGF under biased measurement noise statistics.

Method	Mean error	RMSE
CKF	0.9740	1.1884
RWGF	0.5051	0.6670

measurement noise statistics were $\hat{b}_0 = 0.03$ and $\hat{B}_0 = 10$, which were biased from their theorical values.

Fig. 4 shows the estimations of measurement noise's mean and covariance by RWGF, where after the initial oscillations within about 80s, the estimated measurement noise statistics are gradually converged to their theoretical values, respectively. This demonstrates that the proposed RWGF can effectively estimate the measurement noise statistics.

Fig. 5 shows that the state estimation errors by both CKF and RWGF in the presence of the biased measurement noise statistics. Since RWGF is capable of online estimating measurement noise statistics while CKF does have this capability, the state estimation error achieved by RWGF is much smaller than that by CKF. As shown in Table 2, the mean error and RMSE are 0.9740 and 1.1884 for CKF, while 0.5051 and RMSE 0.6670 for RWGF.

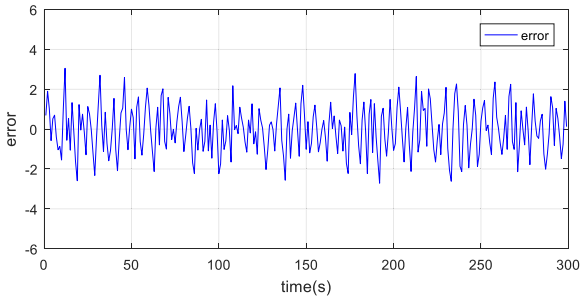
3) SYSTEM STATE ESTIMATION

The means and covariances of both process and measurement noises were assumed to be exactly known, and they were

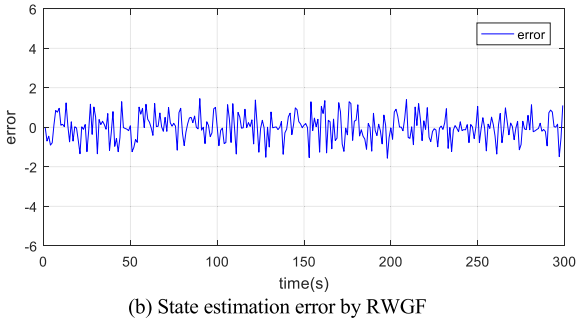
$$a_k = 0, \quad \hat{a}_k = 0, \ b_k = 0 \text{ and } \hat{b}_k = 0 \tag{58}$$

The theoretical covariances of the process and measurement noises, which differed promptly, were represented by

$$A_k = 8 + 2\sin(0.05k) \text{ and } B_k = 1 + 3\cos(0.02k)^2 \tag{59}$$



(a) State estimation error by CKF



(b) State estimation error by RWGF

FIGURE 5. State estimation errors by RWGF and CKF under biased measurement noise statistics.

TABLE 3. Mean errors of the system noise covariances estimated by both MAP and RWGF.

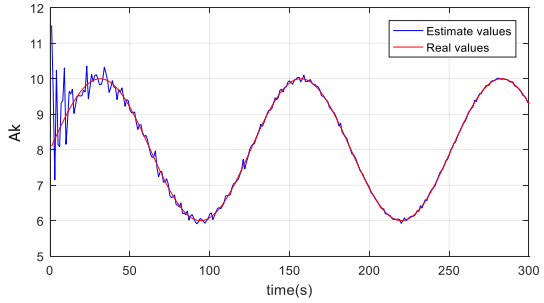
Method	Mean Error (A_k)	Mean Error (B_k)
MAP	0.1631	0.1140
RWGF	0.0703	0.0612

The initial process noise covariance was $\hat{A}_0 = 8$ and the initial measurement noise covariance $\hat{B}_0 = 4$. Obviously, these initial covariances were biased.

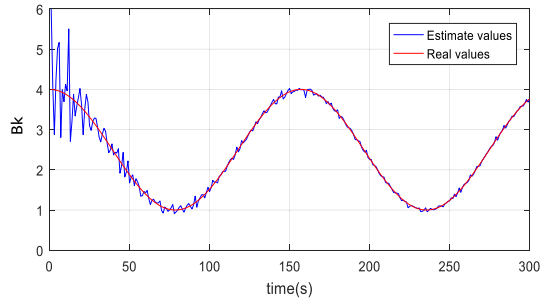
As shown in Fig. 6, the influences on the estimates due to the biases involved in the initial noise covariances are within the initial time interval of ∼30s, where there are significant oscillations in the estimation curves. After this initial time interval, the estimated values of the noise covariances are very close to their theoretical values. These results demonstrate that the proposed RWGF can effectively track the theoretical covariances of both process and measurement noises.

For further evaluation, the noise covariances estimated by RWGF were compared with those by MAP [4, 7] under the same conditions. As shown in Fig. 7, even after the significant oscillations within the initial time period of ∼30s, there are still obvious oscillations remained in the MAP estimation curves. As illustrated in Table 3, the mean errors of the process and measurement noise covariances estimated by RWGF are 0.0703 and 0.0612, and are almost twice smaller than those by MAP, which are 0.1631 and 0.1140.

Given its ability to effectively track the system noise statistics, as shown in Fig. 8, RWGF has a smaller state estimation error than CKF. As illustrated in Table 4, the resultant mean error and RMSE of the state estimation are 0.6138 and

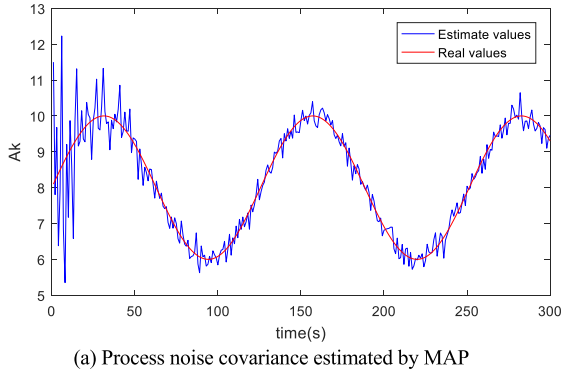


(a) Process noise covariance estimated by RWGF

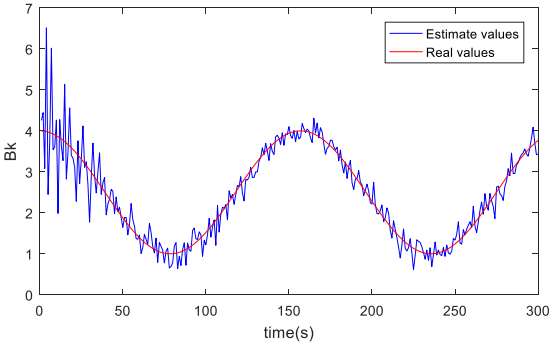


(b)Measurement noise covariance estimated by RWGF

FIGURE 6. System noise covariances estimated by RWGF.



(a) Process noise covariance estimated by MAP

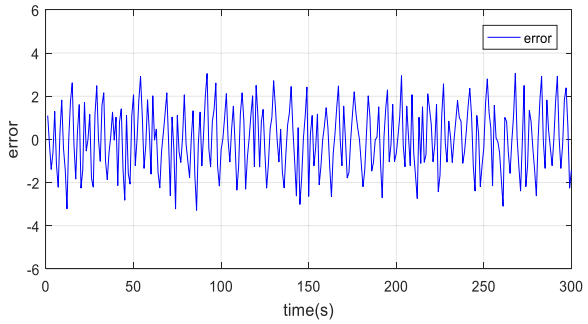


(b)Measurement noise covariance estimated by MAP

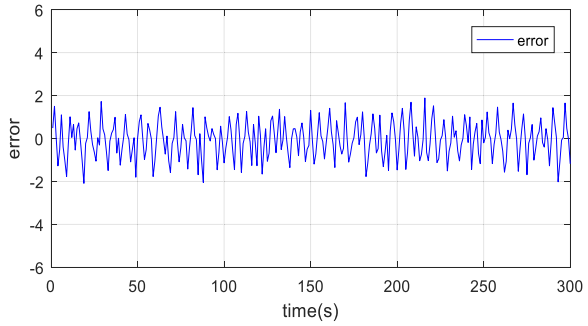
FIGURE 7. System noise covariances estimated by MAP.

0.7402 for RWAGF, while they are 1.6003 and 1.2478 for CKF.

The above simulations and analyses exhibit that the presented RWGF effectively estimates both process and measurement noise statistics, resulting in the increased accuracy for system state estimation comparing to CKF.



(a) State estimation error by CKF



(b) State estimation error by RWGF

FIGURE 8. State estimation errors by both CKF and RWGF under biased system noise statistics.

TABLE 4. Statistical errors of state estimation by CKF and RWGF under biased system noise statistics.

Method	Mean Error	RMSE
CKF	1.6003	1.2478
RWGF	0.6138	0.7402

B. EXPERIMENTS AND ANALYSIS

Experiments of vehicle flight were also performed for the performance evaluation. The vehicle was an ASN 206 reconnaissance UAV (Unmanned Aerial Vehicle). The BDS (Bei Dou System) / MEMS IMU (Micro-Electro-Mechanical System Inertial Measurement Unit) integrated navigation system was mounted on the UAV for positioning and navigation (see Fig. 9). This navigation system includes a self-made MEMS IMU and an AGRMIN GNS430 BDS receiver. Table 5 lists the parameters of the BDS / MEMS IMU integrated navigation system. One additional AGRMIN GNS430 BDS receiver was installed at the local reference station, which was located at a distance of approximate 500m from the initial position of the UAV, to provide the reference data via pseudo-range differential for error analysis.

1) MATHEMATICAL MODEL OF BDS / MEMS IMU INTEGRATED NAVIGATION SYSTEM

a: SYSTEM STATE EQUATION

The BDS / MEMS IMU integrated navigation system adopts the E-N-U (East-North-Up) geography frame as the navigation frame. The state vector of this navigation system is

TABLE 5. The parameters of the BDS / MEMS IMU integrated navigation system.

Sensor	Error source	Value (1σ)
IMU	Gyro constant drift	0.1(°)/h
	Gyro white noise	0.05(°)/h
	Accelerometer zero bias	10-3g
	Accelerometer white noise	10-4g
BDS receiver	Data update rate	1Hz
	positioning accuracy	15m
	Velocity accuracy	0.05m/S

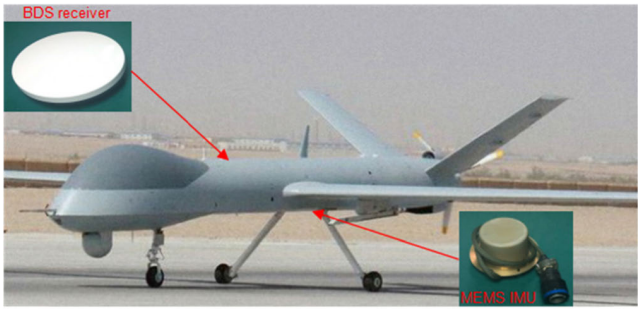


FIGURE 9. The UAV and its BDS / MEMS IMU integrated navigation system.

defined as

$$\boldsymbol{X}(t) = [\delta v_E \quad \delta v_N \quad \delta v_U \quad \delta L \quad \delta\lambda \quad \delta h \quad \phi_E \quad \phi_N \quad \phi_U \quad \varepsilon_x \quad \varepsilon_y \quad \varepsilon_z \quad \nabla_{bx} \quad \nabla_{by} \quad \nabla_{bz}]^{\mathrm{T}}_{15\times 1} \tag{60}$$

where $(\delta v_E, \delta v_N, \delta v_U)$, $(\delta L, \delta\lambda, \delta h)$ and (ϕ_E, ϕ_N, ϕ_U) are the velocity error, position error, and attitude error of the UAV; L, λ and h are the latitude, longitude and altitude of the UAV; $(\varepsilon_x, \varepsilon_y, \varepsilon_z)$ is the constant drift of the gyros; $(\nabla_{bx}, \nabla_{by}, \nabla_{bz})$ is the zero bias of the accelerometers; and

$$\begin{bmatrix} \delta\dot{L} \\ \delta\dot{\lambda} \\ \delta\dot{h} \end{bmatrix} = \begin{bmatrix} \dfrac{\delta v_y}{R+h} + \dfrac{v_y \delta h}{(R+h)^2} \\ \dfrac{\delta v_x}{R+h}\sec L + \dfrac{v_x \delta L}{R+h}\sec L\tan L - \dfrac{v_x \delta h}{(R+h)^2}\sec L \\ \delta v_z \end{bmatrix} \tag{61}$$

where R is the radius of the Earth.

The system state equation of the BDS / MEMS IMU integrated navigation can be represented as

$$\dot{\boldsymbol{X}}(t) = f(\boldsymbol{X}(t)) + \boldsymbol{G}(t)\boldsymbol{w}(t) \tag{62}$$

where $f(\cdot)$ denotes the system nonlinear function, and $\boldsymbol{w}(t)$ the system noise consisting of gyro's Gaussian white noise (w_{gx}, w_{gy}, w_{gz}) and accelerometer's Gaussian white noise (w_{ax}, w_{ay}, w_{az}).

$$\boldsymbol{w}(t) = [w_{gx}, w_{gy}, w_{gz}, w_{ax}, w_{ay}, w_{az}]^{\mathrm{T}}_{6\times 1} \tag{63}$$

$\boldsymbol{G}(t)$ is the coefficient matrix of the system noise and is expressed as

$$\boldsymbol{G}(t) = \begin{bmatrix} C_b^n & 0_{3\times3} \\ 0_{3\times3} & C_b^n \\ 0_{9\times3} & 0_{9\times3} \end{bmatrix}_{15\times6} \tag{64}$$

where C_b^n is the conversion matrix from the body coordinate system to the navigation coordinate system.

b: MEASUREMENT EQUATION

The measurement equation of the BDS / MEMS IMU integrated navigation system is established using velocity error and position error as measurement information.

The measurement equation of position error is described by

$$\boldsymbol{Z}_p(t) = \boldsymbol{H}_p(t)\boldsymbol{X}(t) + \boldsymbol{V}_p(t) = \begin{bmatrix} R\cos L\delta\lambda + n_E \\ R\delta L + n_N \\ \delta h + n_U \end{bmatrix} \tag{65}$$

where $\boldsymbol{H}_p(t)$ is the position measurement matrix, which is expressed as

$$\boldsymbol{H}_p = \left[0_{3\times3} \vdots diag\left[R \;\; R\cos L \;\; 1\right] \vdots 0_{3\times9} \right]_{9\times15} \tag{66}$$

$\boldsymbol{V}_p(t)$ is the position measurement noise, which is expressed as

$$\boldsymbol{V}_p(t) = \begin{bmatrix} n_E & n_N & n_U \end{bmatrix}^T \tag{67}$$

where n_E, n_N and n_U are the position errors of the BDS receiver in the three axes, respectively.

The velocity error measurement equation can be written as

$$\boldsymbol{Z}_v(t) = \boldsymbol{H}_v(t)\boldsymbol{X}(t) + \boldsymbol{V}_v(t) = \begin{bmatrix} \delta v_E + n_{vE} \\ \delta v_N + n_{vN} \\ \delta v_U + n_{vU} \end{bmatrix} \tag{68}$$

where $\boldsymbol{H}_v(t)$ is the velocity measurement matrix, which is expressed as

$$\boldsymbol{H}_v(t) = \left[diag\left[1 \;\; 1 \;\; 1\right] \vdots 0_{3\times12} \right]_{9\times15} \tag{69}$$

$$\boldsymbol{V}_v(t) = \begin{bmatrix} n_{vE} & n_{vN} & n_{vU} \end{bmatrix}^T \tag{70}$$

where $\boldsymbol{V}_v(t)$ is the velocity measurement noise, and n_{vE}, n_{vN} and n_{vU} are the velocity measurement errors of BDS.

According to (65) and (68), the measurement equation of the BDS / MEMS IMU integrated navigation system can be obtained as

$$\boldsymbol{Z}(t) = \begin{bmatrix} \boldsymbol{H}_p(t) \\ \boldsymbol{H}_v(t) \end{bmatrix} \boldsymbol{X}(t) + \begin{bmatrix} \boldsymbol{V}_p(t) \\ \boldsymbol{V}_v(t) \end{bmatrix} = \boldsymbol{H}(t)\boldsymbol{X}(t) + \boldsymbol{V}(t) \tag{71}$$

2) EXPERIMENTAL RESULTS AND ANALYSIS

The flight test was carried out in the City of Xi'an, Shaanxi, China. The UAV lifted off after the ten-minute initialization. The initial position variance was $0.2m^2$ and the initial velocity variance $9.0\times10^{-5}m^2s^{-2}$. The total time of flight was 90min. Fig. 10 illustrates the UAV flight path. The navigation data were selected from a smooth flight of 1800s.

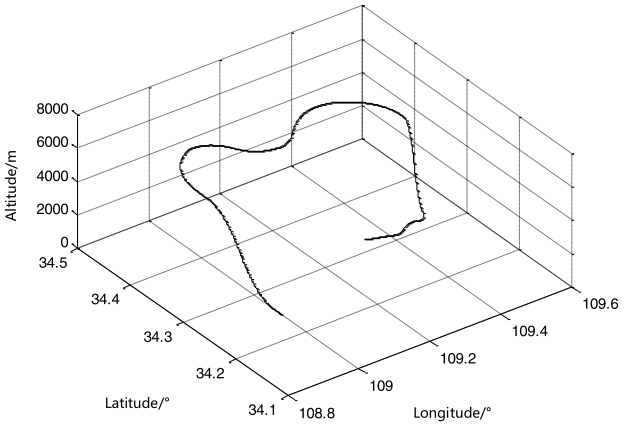


FIGURE 10. UAV flight trajectory.

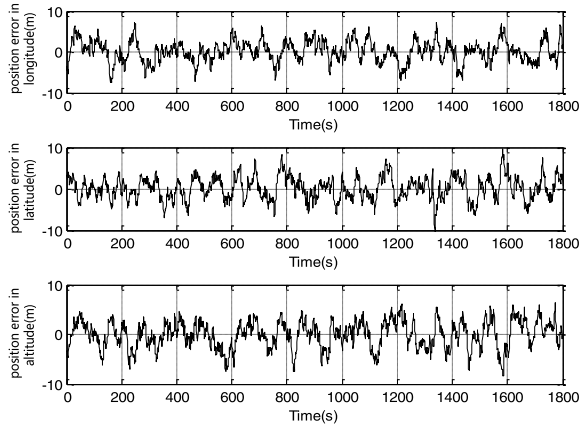


FIGURE 11. Position error by CKF.

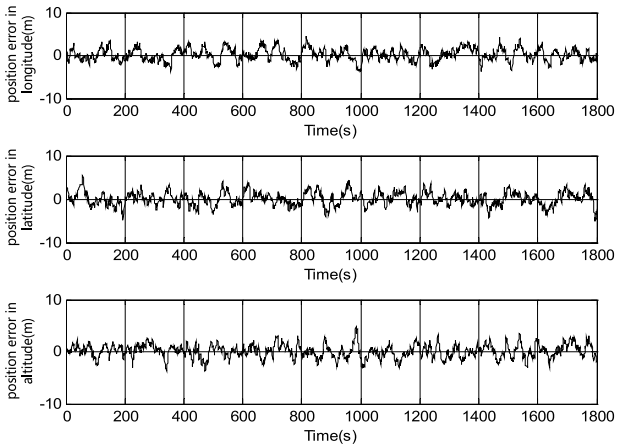


FIGURE 12. Position error by RWGF.

For the purpose of comparison analysis, trials were conducted to process the C/A code observations by both CKF and RWGF. The experimental results show the similar trend as those in the simulation case. As shown in Fig. 11, CKF has significant oscillations in the filtering curve, and its position errors in longitude, latitude and altitude are within (−12m, +12m), (−12m, +12m) and (−10m, +10m), respectively. In contrast, as shown in Fig. 12, the position errors in longitude, latitude and altitude obtained by the proposed RWGF

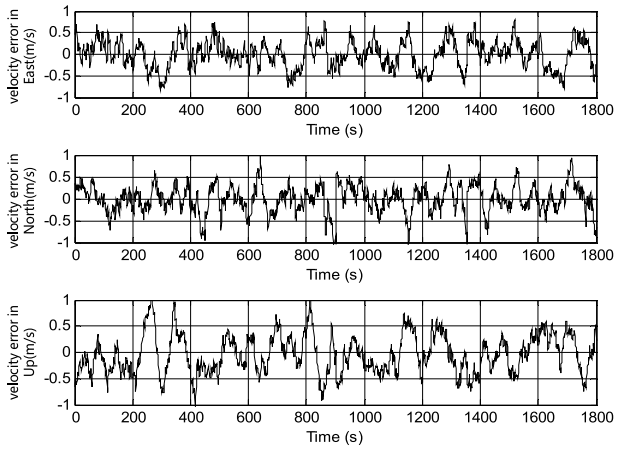


FIGURE 13. Velocity error by CKF.

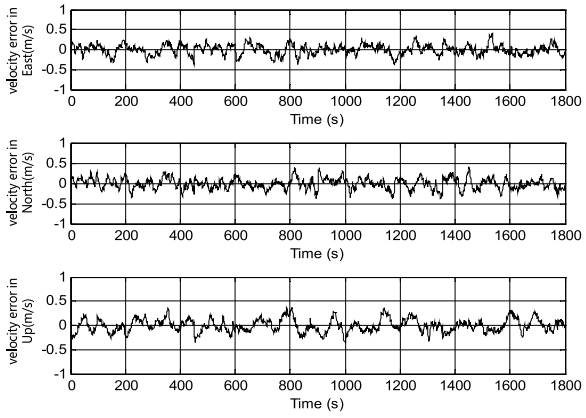


FIGURE 14. Velocity error by RWGF.

TABLE 6. Mean position error and mean velocity error of CKF and RWGF.

Method	Mean Position Error (m)	Mean Velocity Error (m)
CKF	9.6776	0.7850
RWGF	3.4359	0.2830

are within (−3m, +3m), (−3m, +3m) and (−5m, +5m), which are much smaller than those of CKF. Similar to the case of position error, the velocity error by the proposed RWGF is also much smaller than that of CKF. The velocity errors in East, North and Up are within (−0.9m/s, +0.9m/s), (−0.9m/s, +0.9m/s) and (−1m/s, +1m/s) respectively for CKF (see Fig. 13), while (−0.3m/s, +0.3m/s), (−0.3m/s, +0.3m/s) and (−0.4m/s, +0.4m/s) for the proposed RWGF (see Fig. 14). This is because the proposed RWGF has the capability of noise statistics estimation, while CKF does not. As shown in Table 6, the mean errors of position and velocity by RWGF are also much smaller than those of CKF.

VII. CONCLUSION

This paper proposes a new RWGF for estimation of nonlinear system state. This method improves the traditional Gaussian filtering by adopting the random weighting concept to online estimate the statistical characteristics of both process and measurement noises. The theories of random weighting are established on the basis of the MAP theory to online estimate noise statistical characteristics. Upon the random weighting estimations of noise statistical characteristics, the presented RWGF dynamically alters the random factors of noise statistics to restrain system noise interferences on system state estimation, resulting in enhanced estimation accuracy. Simulations, experiments and comparison analysis demonstrate that the presented RWGF has higher accuracy than the traditional Gaussian filtering under unknown or biased noise statistical characteristics.

Future work will focus on improvement of the proposed RWGF. Currently, the proposed RWGF is under the assumption that the system noises are subject to the Gaussian distribution. It is expected to extend the random weighting estimations of Gaussian system noises to non-Gaussian system noises, and thus establishing a new random weighting-based non-Gaussian filter for estimation of nonlinear system state.

APPENDIXES

APPENDIX A
SOLVING (25) FOR $\hat{\boldsymbol{a}}_{k+1}$

By (25),

$$\frac{\partial \ln h}{\partial \boldsymbol{a}} = \frac{1}{2}\sum_{j=0}^{k}\left[\boldsymbol{A}^{-1} + \left(\boldsymbol{A}^{-1}\right)^{\mathrm{T}}\right]\left(\boldsymbol{x}_{j+1} - f_j\left(\boldsymbol{x}_j\right) - \boldsymbol{a}\right) \quad (72)$$

Let

$$\frac{\partial \ln h}{\partial \boldsymbol{a}}\Big|_{\boldsymbol{a}=\boldsymbol{a}_{k+1}} = \frac{1}{2}\left[\boldsymbol{A}^{-1} + \left(\boldsymbol{A}^{-1}\right)^{\boldsymbol{T}}\right]\sum_{j=0}^{k}\left(\boldsymbol{x}_{j+1} - f_j\left(\boldsymbol{x}_j\right) - \boldsymbol{a}\right)\Big|_{\boldsymbol{a}=\boldsymbol{a}_{k+1}} = 0 \quad (73)$$

As $\frac{1}{2}\left[\boldsymbol{A}^{-1} + \left(\boldsymbol{A}^{-1}\right)^{T}\right] \neq 0$, we have

$$\sum_{j=0}^{k}\left(\boldsymbol{x}_{j+1} - f_j\left(\boldsymbol{x}_j\right) - \boldsymbol{a}\right)\Big|_{\boldsymbol{a}=\boldsymbol{a}_{k+1}} = 0 \quad (74)$$

Thus

$$\hat{\boldsymbol{a}}_{k+1} = \frac{1}{k+1}\sum_{j=0}^{k}\left\{\hat{\boldsymbol{x}}_{j+1} - f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j \leftarrow \hat{\boldsymbol{x}}_{j|k+1}}\right\} \quad (75)$$

To simplify the calculation process, substituting state smooth $\hat{\boldsymbol{x}}_{j|k+1}$ with estimated state $\hat{\boldsymbol{x}}_j$ yields

$$\hat{\boldsymbol{a}}_{k+1} = \frac{1}{k+1}\sum_{j=0}^{k}\left\{\hat{\boldsymbol{x}}_{j+1} - f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j \leftarrow \hat{\boldsymbol{x}}_j}\right\} \quad (76)$$

APPENDIX B
SOLVING (25) FOR $\hat{\boldsymbol{A}}_{k+1}$

Consider the conditions as follows:

(i) $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}$ and $\boldsymbol{B}^{\mathrm{T}} = \boldsymbol{B}$

(ii) $\frac{\partial \ln|\boldsymbol{C}|}{\partial \boldsymbol{C}} = \frac{1}{|\boldsymbol{C}|}\cdot\boldsymbol{C}^{*} = \boldsymbol{C}^{-1}\left(\frac{\partial|\boldsymbol{C}|}{\partial \boldsymbol{C}} = \boldsymbol{C}^{*}\right)$

(iii) $\frac{\partial \boldsymbol{C}^{-1}}{\partial \boldsymbol{C}} = -\boldsymbol{C}^{-T}\times\boldsymbol{C}^{-1}$

where $|\boldsymbol{C}|$ is the determinant of matrix $\boldsymbol{C}$.

By (25),

$$\begin{aligned}
&\frac{\partial \ln h}{\partial \boldsymbol{A}}\Big|_{\boldsymbol{A}=\boldsymbol{A}_{k+1}} \\
&= \left[-\frac{k+1}{2}\cdot\frac{\partial \ln|\boldsymbol{A}|}{\partial \boldsymbol{A}}\right. \\
&\qquad \left.-\frac{1}{2}\sum_{j=0}^{k}\left\|\boldsymbol{x}_{j+1}-f_j\left(\boldsymbol{x}_j\right)-\boldsymbol{a}\right\|^2\frac{\partial \boldsymbol{A}^{-1}}{\partial \boldsymbol{A}}\right]\Big|_{\boldsymbol{A}=\boldsymbol{A}_{k+1}} \\
&= \left\{-\frac{k+1}{2}\boldsymbol{A}^{-1}-\frac{1}{2}\sum_{j=0}^{k}\left[\boldsymbol{x}_{j+1}-f_j\left(\boldsymbol{x}_j\right)-\boldsymbol{a}\right]\Big|_{\boldsymbol{a}=\boldsymbol{a}_{k+1}}\right. \\
&\quad \left.\times\left[\boldsymbol{x}_{j+1}-f_j\left(\boldsymbol{x}_j\right)-\boldsymbol{a}\right]^{\mathrm{T}}\times\left(-\boldsymbol{C}^{-T}\times\boldsymbol{C}^{-1}\right)\right\}\Big|_{\boldsymbol{A}=\boldsymbol{A}_{k+1}} \\
&= 0
\end{aligned} \tag{77}$$

Equation (74) can be rewritten as

$$\begin{aligned}
\frac{k+1}{2}\hat{\boldsymbol{A}}_{k+1}^{-1} &= \frac{1}{2}\sum_{j=0}^{k}\left[\boldsymbol{x}_{j+1}-f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_{j|k+1}}-\boldsymbol{a}\right] \\
&\quad\times\left[\boldsymbol{x}_{j+1}-f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_{j|k+1}}-\boldsymbol{a}\right]^{\mathrm{T}}_{\boldsymbol{a}=\boldsymbol{a}_{k+1}}\cdot\left(\hat{\boldsymbol{A}}_{k+1}^{-\mathrm{T}}\cdot\hat{\boldsymbol{A}}_{k+1}^{-1}\right)
\end{aligned} \tag{78}$$

Multiplying both sides of (75) by $\hat{\boldsymbol{A}}_{k+1}$ yields

$$\begin{aligned}
\frac{k+1}{2}\mathrm{I} &= \frac{1}{2}\sum_{j=0}^{k}\left[\boldsymbol{x}_{j+1}-f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_{j|k+1}}-\boldsymbol{a}_{k+1}\right] \\
&\quad\times\left[\boldsymbol{x}_{j+1}-f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_{j|k+1}}-\boldsymbol{a}_{k+1}\right]^{\mathrm{T}}\cdot\hat{\boldsymbol{A}}_{k+1}^{-\mathrm{T}}
\end{aligned} \tag{79}$$

where I is an unit matrix.

Substituting $\hat{\boldsymbol{A}}_{k+1}^{-\mathrm{T}}=\hat{\boldsymbol{A}}_{k+1}^{-1}$ into (76) yields

$$\begin{aligned}
\hat{\boldsymbol{A}}_{k+1} &= \frac{1}{k+1}\sum_{j=0}^{k}\left\{\left[\hat{\boldsymbol{x}}_{j+1}-f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_{j|k+1}}-\boldsymbol{a}_{k+1}\right]\right. \\
&\quad\left.\times\left[\hat{\boldsymbol{x}}_{j+1}-f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_{j|k+1}}-\boldsymbol{a}_{k+1}\right]^{\mathrm{T}}\right\}
\end{aligned} \tag{80}$$

Substituting state smooth $\hat{\boldsymbol{x}}_{j|k+1}$ with state estimate $\hat{\boldsymbol{x}}_j$ yields

$$\begin{aligned}
\hat{\boldsymbol{A}}_{k+1} &= \frac{1}{k+1}\sum_{j=0}^{k}\left\{\left[\hat{\boldsymbol{x}}_{j+1}-f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_j}-\boldsymbol{a}_{k+1}\right]\right. \\
&\quad\left.\times\left[\hat{\boldsymbol{x}}_{j+1}-f_j\left(\boldsymbol{x}_j\right)\Big|_{\boldsymbol{x}_j\leftarrow\hat{\boldsymbol{x}}_j}-\boldsymbol{a}_{k+1}\right]^{\mathrm{T}}\right\}
\end{aligned} \tag{81}$$

APPENDIX C
SOLVING (26) FOR $\hat{\boldsymbol{b}}_{k+1}$

By (26),

$$\frac{\partial \ln h}{\partial \boldsymbol{b}}=\sum_{j=0}^{k}\left[\boldsymbol{B}^{-1}+\left(\boldsymbol{B}^{-1}\right)^{\mathrm{T}}\right]\left(\boldsymbol{y}_{j+1}-g_{j+1}\left(\boldsymbol{x}_{j+1}\right)-\boldsymbol{b}\right) \tag{82}$$

Let

$$\begin{aligned}
\frac{\partial \ln h}{\partial \boldsymbol{b}}\Big|_{\boldsymbol{b}=\boldsymbol{b}_{k+1}} &= \left[\boldsymbol{B}^{-1}+\left(\boldsymbol{B}^{-1}\right)^{\mathrm{T}}\right] \\
&\quad\times\sum_{j=0}^{k}\left(\boldsymbol{y}_{j+1}-g_{j+1}\left(\boldsymbol{x}_{j+1}\right)-\boldsymbol{b}\right)\Big|_{\boldsymbol{b}=\boldsymbol{b}_{k+1}}=0
\end{aligned} \tag{83}$$

Since $\boldsymbol{B}^{-1}+\left(\boldsymbol{B}^{-1}\right)^{\mathrm{T}}\neq 0$, we have

$$\sum_{j=0}^{k}\left(\boldsymbol{y}_{j+1}-g_{j+1}\left(\boldsymbol{x}_{j+1}\right)-\boldsymbol{b}\right)\Big|_{\boldsymbol{b}=\boldsymbol{b}_{k+1}}=0 \tag{84}$$

i.e.,

$$\sum_{j=0}^{k}\left[\boldsymbol{y}_{j+1}-g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\right]\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1|k+1}}-(k+1)\,\boldsymbol{b}\,\Big|_{\boldsymbol{b}=\boldsymbol{b}_{k+1}}=0 \tag{85}$$

Thus

$$\hat{\boldsymbol{b}}_{k+1}=\frac{1}{k+1}\sum_{j=0}^{k}\left\{\boldsymbol{y}_{j+1}-g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1|k+1}}\right\} \tag{86}$$

Substituting state smooth $\hat{\boldsymbol{x}}_{j+1|k+1}$ with state estimate $\hat{\boldsymbol{x}}_{j+1}$ yields

$$\hat{\boldsymbol{b}}_{k+1}=\frac{1}{k+1}\sum_{j=0}^{k}\left\{\boldsymbol{y}_{j+1}-g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}}\right\} \tag{87}$$

APPENDIX D
SOLVING (26) FOR $\hat{\boldsymbol{B}}_{k+1}$

Suppose that the conditions described in Appendix 2 hold. From (26), we readily have

$$\begin{aligned}
&\frac{\partial \ln h}{\partial \boldsymbol{B}}\Big|_{\boldsymbol{B}=\boldsymbol{B}_{k+1}} \\
&= \left[-\frac{k+1}{2}\cdot\frac{\partial \ln|\boldsymbol{B}|}{\partial \boldsymbol{B}}\right. \\
&\qquad \left.-\sum_{j=0}^{k}\left\|\boldsymbol{y}_{j+1}-g_{j+1}\left(\boldsymbol{x}_{j+1}\right)-\boldsymbol{b}\right\|^2\frac{\partial \boldsymbol{B}^{-1}}{\partial \boldsymbol{B}}\right]\Big|_{\boldsymbol{B}=\boldsymbol{B}_{k+1}} \\
&= \left\{-\frac{k+1}{2}\boldsymbol{B}^{-1}-\frac{1}{2}\sum_{j=0}^{k}\left[\boldsymbol{y}_{j+1}-g_{j+1}\left(\boldsymbol{x}_{j+1}\right)-\boldsymbol{b}\right]\Big|_{\boldsymbol{b}=\boldsymbol{b}_{k+1}}\right. \\
&\quad\times\left[\boldsymbol{y}_{j+1}-g_{j+1}\left(\boldsymbol{x}_{j+1}\right)-\boldsymbol{b}\right]\Big|_{\boldsymbol{b}=\boldsymbol{b}_{k+1}}^{\mathrm{T}} \\
&\quad\left.\times\left(-\boldsymbol{B}^{-\mathrm{T}}\times\boldsymbol{B}^{-1}\right)\right\}\Big|_{\boldsymbol{B}=\boldsymbol{B}_{k+1}} \\
&= 0
\end{aligned} \tag{88}$$

Thus

$$\frac{k+1}{2}\hat{\boldsymbol{B}}_{k+1}^{-1} = \frac{1}{2}\sum_{j=0}^{k}\left\{\left[\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\hat{\boldsymbol{x}}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1|k+1}} - \boldsymbol{b}\right]\left[\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\hat{\boldsymbol{x}}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1|k+1}} - \boldsymbol{b}\right]^{\mathrm{T}}\right\}\Big|_{\boldsymbol{b}=\boldsymbol{b}_{k+1}} \times\left(\hat{\boldsymbol{B}}_{k+1}^{-\mathrm{T}}\cdot\hat{\boldsymbol{B}}_{k+1}^{-1}\right) \tag{89}$$

Multiplying both sides of (86) by $\hat{\boldsymbol{B}}_{k+1}$ yields

$$\frac{k+1}{2}\mathrm{I} = \frac{1}{2}\sum_{j=0}^{k}\left[\left[\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1|k+1}} - \boldsymbol{b}\right]\right]\Big|_{\boldsymbol{b}=\boldsymbol{b}_{k+1}} \times\left[\left[\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1|k+1}} - \boldsymbol{b}\right]\right]_{\boldsymbol{b}=\boldsymbol{b}_{k+1}}^{\mathrm{T}}\cdot\hat{\boldsymbol{B}}_{k+1}^{-\mathrm{T}} \tag{90}$$

Substituting $\hat{\boldsymbol{B}}_{k+1}^{-\mathrm{T}} = \hat{\boldsymbol{B}}_{k+1}^{-1}$ into (87) yields

$$\hat{\boldsymbol{B}}_{k+1} = \frac{1}{k+1}\sum_{j=0}^{k}\left\{\left[\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1|k+1}} - \boldsymbol{b}_{k+1}\right]\left[\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1|k+1}} - \boldsymbol{b}_{k+1}\right]^{\mathrm{T}}\right\} \tag{91}$$

Substituting state smooth $\hat{\boldsymbol{x}}_{j+1|k+1}$ with state estimate $\hat{\boldsymbol{x}}_{j+1}$ yields

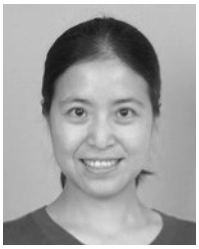
$$\hat{\boldsymbol{B}}_{k+1} = \frac{1}{k+1}\sum_{j=0}^{k}\left\{\left[\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}} - \boldsymbol{b}_{k+1}\right]\left[\boldsymbol{y}_{j+1} - g_{j+1}\left(\boldsymbol{x}_{j+1}\right)\Big|_{\boldsymbol{x}_{j+1}\leftarrow\hat{\boldsymbol{x}}_{j+1}} - \boldsymbol{b}_{k+1}\right]^{\mathrm{T}}\right\} \tag{92}$$

REFERENCES

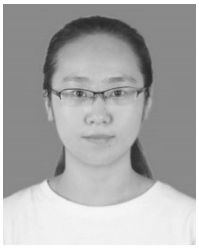
[1] H. Afshari, S. Gadsden, and S. Habibi, "Gaussian filters for parameter and state estimation: A general review of theory and recent trends," *Signal Process.*, vol. 135, pp. 218–238, Jun. 2017.
[2] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Trans. Autom. Control.*, vol. 54, no. 6, pp. 1254–1269, Jun. 2009.
[3] M. Branicki, A. J. Majda, and K. J. H. Law, "Accuracy of some approximate Gaussian filters for the Navier-Stokes equation in the presence of model error," *Multiscale Model. Simul.*, vol. 16, no. 4, pp. 1756–1794, Jan. 2018.
[4] H. Cai, "Bayesian MAP estimation of noise statistics and system adaptive filtering," *J. Nat. Univ. Defense Technol.*, vol. 19, no. 1, pp. 5–8, 1997.
[5] K. P. B. Chandra and D.-W. Gu, *Nonlinear Filtering: Methods and Applications*. Cham, Switzerland: Springer, 2019.
[6] S. Cho and W. Choi, "Robust positioning technique in low-cost DR/GPS for land navigation," *IEEE Trans. Instrum. Meas.*, vol. 55, no. 4, pp. 1132–1142, Aug. 2006.
[7] J. L. Ding and J. Xiao, "Design of adaptive cubature Kalman filter based on maximum a posteriori estimation," *Control Decis.*, vol. 29, pp327-334, Feb. 2014.
[8] B. Gao, S. Gao, G. Hu, Y. Zhong, and C. Gu, "Maximum likelihood principle and moving horizon estimation based adaptive unscented Kalman filter," *Aerosp. Sci. Technol.*, vol. 73, pp. 184–196, Feb. 2018.
[9] B. Gao, S. Gao, Y. Zhong, and C. Gu, "Random weighting estimation of sampling distributions via importance resampling," *Commun. Stat. Simul. Comput.*, vol. 46, no. 1, pp. 640–654, Jan. 2017.
[10] B. Gao, G. Hu, S. Gao, Y. Zhong, and C. Gu, "Multi-sensor optimal data fusion based on the adaptive fading unscented Kalman filter," *Sensors*, vol. 18, no. 2, p. 488, Feb. 2018.
[11] S. Gao, G. Hu, and Y. Zhong, "Windowing and random weighting-based adaptive unscented Kalman filter," *Int. J. Adapt. Control Signal Process.*, vol. 29, no. 2, pp. 201–223, Feb. 2015.
[12] S. Gao, Z. Feng, Y. Zhong, and B. Shirinzadeh, "Random weighting estimation of parameters in generalized Gaussian distribution," *Inf. Sci.*, vol. 178, no. 9, pp. 2275–2281, May 2008.
[13] S. Gao, Y. Zhong, C. Gu, and B. Shirinzadeh, "Weak convergence for random weighting estimation of smoothed Quantile processes," *Inf. Sci.*, vol. 263, pp. 36–42, Apr. 2014.
[14] S. Gao, Y. Zhong, and B. Shirinzadeh, "Random weighting estimation for fusion of multi-dimensional position data," *Inf. Sci.*, vol. 180, no. 24, pp. 4999–5007, Dec. 2010.
[15] S. Gao, Y. Zhong, W. Wei, and C. Gu, "Windowing-based random weighting fitting of systematic model errors for dynamic vehicle navigation," *Inf. Sci.*, vol. 282, pp. 350–362, Oct. 2014.
[16] M. Gautier and P. Poignet, "Extended Kalman filtering and weighted least squares dynamic identification of robot," *IFAC Proc. Vol.*, vol. 33, no. 15, pp. 935–940, Jun. 2000.
[17] D. Del Gobbo, M. Napolitano, P. Famouri, and M. Innocenti, "Experimental application of extended Kalman filtering for sensor validation," *IEEE Trans. Control Syst. Technol.*, vol. 9, no. 2, pp. 376–380, Mar. 2001.
[18] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proc. F Radar Signal Process. UK*, vol. 140, no. 2, p. 107, 1993.
[19] Z. Hu and B. Gallacher, "Extended Kalman filtering based parameter estimation and drift compensation for a MEMS rate integrating gyroscope," *Sens. Actuators A, Phys.*, vol. 250, pp. 96–105, Oct. 2016.
[20] Y. Huang, Y. Zhang, X. Wang, and L. Zhao, "Gaussian filter for nonlinear systems with correlated noises at the same epoch," *Automatica*, vol. 60, pp. 122–126, Oct. 2015.
[21] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Trans. Autom. Control.*, vol. 45, no. 5, pp. 910–927, May 2000.
[22] S. Julier and J. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.
[23] Q. Li and F. Sun, "Adaptive cubature particle filter algorithm," in *Proc. IEEE Conf. Mechatronics Autom.*, Takamatsu, Japan, Aug. 2013, pp. 1356–1360.
[24] M. Narasimhappa, S. L. Sabat, R. Peesapati, and J. Nayak, "An innovation based random weighting estimation mechanism for denoising fiber optic gyro drift signal," *Optik*, vol. 125, no. 3, pp. 1192–1198, Feb. 2014.
[25] A. P. Sage and G. W. Husa, "Adaptive filtering with unknown prior statistics," in *Proc. Joint Amer. Control Conf.*, Boulder, CA, USA, 1969, pp. 769–774, 1969.
[26] H. E. Soken and C. Hajiyev, "Adaptive fading UKF with Q-adaptation: Application to picosatellite attitude estimation," *J. Aerosp. Eng.*, vol. 26, no. 3, pp. 628–636, Jul. 2013.
[27] F. Sun and L. Tang, "Estimation precision comparison of cubature Kalman filter and unscented Kalman filter," *Control Decis.*, vol. 28, no. 2, pp. 303–308, 2013.
[28] G. Wang, Z. Gao, Y. Zhang, and B. Ma, "Adaptive maximum correntropy Gaussian filter based on variational Bayes," *Sensors*, vol. 18, no. 6, p. 1960, Jun. 2018.
[29] H. Wang, H. Li, J. Fang, and H. Wang, "Robust Gaussian Kalman filter with outlier detection," *IEEE Signal Process. Lett.*, vol. 25, no. 8, pp. 1236–1240, Aug. 2018.
[30] M. Wüethrich, S. Trimpe, C. Garcia Cifuentes, D. Kappler, and S. Schaal, "A new perspective and extension of the Gaussian Filter," *Int. J. Robot. Res.*, vol. 35, no. 14, pp. 1731–1749, Dec. 2016.
[31] Y. Yang and T. Xu, "An Adaptive Kalman filter based on sage windowing weights and variance components," *J. Navigat.*, vol. 56, no. 2, pp. 231–240, May 2003.
[32] K. Zhao, P. Li, and S.-M. Song, "Gaussian filter for nonlinear stochastic uncertain systems with correlated noises," *IEEE Sensors J.*, vol. 18, no. 23, pp. 9584–9594, Dec. 2018.
[33] R. Wang, Y. He, C. Huang, X. Wang, and W. Cao, "A novel least-mean kurtosis adaptive filtering algorithm based on geometric algebra," *IEEE Access*, vol. 7, pp. 78298–78310, 2019.

[34] L. Yin, Z. Deng, B. Huo, Y. Xia, and C. Li, "Robust derivative unscented Kalman filter under non-Gaussian noise," *IEEE Access*, vol. 6, pp. 33129–33136, 2018.
[35] X. Lin, Y. Jiao, and D. Zhao, "An improved Gaussian filter for dynamic positioning ships with colored noises and random measurements loss," *IEEE Access*, vol. 6, pp. 6620–6629, 2018.
[36] B. Chen and G. Hu, "Nonlinear state estimation under bounded noises," *Automatica*, vol. 98, pp. 159–168, Dec. 2018.

ZHAOHUI GAO received the Ph.D. degree in control engineering from Northwestern Polytechnical University, China, in 2018. He is currently a Postdoctoral Fellow with the School of Geological Engineering and Surveying and Mapping, Chang'an University, China. His research interests include navigation, guidance and control, and information fusion.



CHENGFAN GU is an Independent Researcher. Her research interests include bio/nano materials characterization and analysis, micro-forming, metallic materials processing, and computational modeling. She was a DECRA Fellow of the Australia Research Council, Australia.



JIAHUI YANG is currently pursuing the Ph.D. degree with the School of Automatics, Northwestern Polytechnical University, China. Her research interests include control theory and engineering, navigation, guidance and control, and information fusion.

SHESHENG GAO is currently a Professor with the School of Automatics, Northwestern Polytechnical University, China. His research interests include control theory and engineering, navigation, guidance and control, optimum estimation and control, integrated inertial navigation systems, and information fusion.

YONGMIN ZHONG is currently an Associate Professor with the School of Aerospace, Mechanical and Manufacturing Engineering, RMIT University, Australia. His research interests include computational engineering, haptics, soft tissue modeling, surgical simulation, aerospace navigation and control, and intelligent systems and robotics.