

## Article

# Hybrid Adaptive Cubature Kalman Filter with Unknown Variance of Measurement Noise

Yuepeng Shi <sup>1</sup>, Xianfeng Tang <sup>2,\*</sup>, Xiaoliang Feng <sup>3,\*</sup> , Dingjun Bian <sup>4</sup> and Xizhao Zhou <sup>5</sup>

<sup>1</sup> College of Intelligent Manufacturing and Automation, Henan University of Animal Husbandry and Economy, Zhengzhou 450011, China; syp@hnuahe.edu.cn

<sup>2</sup> Information Technology Center, Zhejiang University, Hangzhou 310027, China

<sup>3</sup> College of Electrical Engineering, Henan University of Technology, Zhengzhou 450001, China

<sup>4</sup> School of Information Engineering, Zhengzhou University, Zhengzhou 450000, China; bdjbray@163.com

<sup>5</sup> School of Business, University of Shanghai for Science and Technology, Shanghai 200093, China; xizhaozhou@163.com

\* Correspondence: txf1213@zju.edu.cn (X.T.); fxl@haut.edu.cn (X.F.)

Received: 23 September 2018; Accepted: 27 November 2018; Published: 7 December 2018



**Abstract:** This paper is concerned with the filtering problem caused by the inaccuracy variance of measurement noise in real nonlinear systems. A novel weighted fusion estimation method of multiple different variance estimators is presented to estimate the variance of the measurement noise. On this basis, a hybrid adaptive cubature Kalman filtering structure is proposed. Furthermore, the information filter of the hybrid adaptive cubature Kalman filter is also studied, and the stability and filtering accuracy of the filter are theoretically discussed. The final simulation examples verify the validity and effectiveness of the hybrid adaptive cubature Kalman filtering methods proposed in this paper.

**Keywords:** nonlinear system; hybrid adaptive filtering; weighted fusion estimation; square-root cubature Kalman filter; information filter

## 1. Introduction

In recent decades, nonlinear filtering has been widely used in military and civil fields such as target tracking, navigation, positioning, and intelligent manufacturing [1,2]. The theory and method of nonlinear filtering has become one of the most important research issues in the signal processing field, and has attracted increasing attention from researchers.

There are two main kinds of nonlinear filtering methods. The representative of the first kind of nonlinear filter is the extended Kalman filter (EKF), which linearizes the system model by Taylor expansion, holds the first order term, and ignores the second- and higher-order terms. The second kind of nonlinear filtering approximates the statistics of the system state, with examples being the unscented Kalman filter (UKF) and the cubature Kalman filter (CKF). Due to the model error of the linearization of the nonlinear system, the accuracy of the EKF is slightly lower, even leading to filtering divergence. Based on the unscented transformation to approximate the statistics of the system state, UKF was presented in [3]. Further, the cubature Kalman filter algorithm was proposed by Ienkar in [4]. The CKF algorithm uses a third-degree spherical-radial cubature rule based on a Gaussian filtering framework. The algorithm has higher numerical stability and a smaller amount of calculation. Its excellent performance has made it widely used in various nonlinear system scenarios. Many advantages of CKF have attracted scholars to conduct in-depth research on it, considering that the traditional nonlinear filters often need to overcome the filtering divergence caused by high-dimensional operational errors. Drawing on the idea of square root filtering in the Kalman

filter, Ienkaram and Haykin proposed the square-root cubature Kalman filter (SCKF) [5], which further improved the accuracy and stability of filtering. However, the traditional nonlinear filtering method requires knowledge of the mathematical model and the prior statistical information of noise when in practical application. Additionally, the statistical characteristics of noise in actual systems are usually indeterminate, which leads to a decline in the filtering accuracy.

For the problem of unknown statistical characteristics of measurement noise in real applications [6,7], Sage and Husa [8,9] proposed an excellent Sage–Husa suboptimal unbiased maximum a posteriori (MAP) estimator. Many scholars also adopted adaptive filtering techniques to improve the performance of the estimation algorithm, such as sliding window method [10], fading factor adjustment (FFA) [11–13], maximum a posteriori (MAP) estimator [14,15], and the variational Bayesian (VB) method [16–19], etc. Different methods for estimating the statistical characteristics of system noise are usually designed under different estimation criteria. How to use these methods to estimate the statistical characteristics of system noise is still an open issue.

For a class of nonlinear stochastic systems with inaccurate or unknown measurement noise variance (i.e., the priori measurement noise variance is not a precise value), an adaptive filtering algorithm based on SCKF is designed in this paper. Firstly, a novel fusion approach is proposed to estimate the measurement noise variance on the basis of the MAP and VB methods. Then, we use FFA to adjust the part of the variance matrix of the SCKF algorithm to obtain a hybrid adaptive SCKF algorithm (HASCKF) and the corresponding information filter (recorded as HASCIF). This is beneficial to reduce the effects on the adaptive filtering algorithm performance that may be caused by the estimation deviation of noise fusion. At the same time, the performance of the adaptive filtering algorithm is analyzed from two aspects based on the established HASCKF algorithm: the stability of the adaptive filtering algorithm and the filtering accuracy.

This paper is organized as follows: Section 2 formulates the nonlinear stochastic system and describes the problem of the inaccuracy of the the measurement noise variance. In Section 3, a novel noise variance fusion estimation algorithm HASCKF and the corresponding information filter are proposed based on the idea of weighted fusion. In Section 4, two simulation examples are utilized to display and verify the performance of the proposed algorithms. Section 5 provides the conclusions of this work.

## 2. Problem Description

Considering a class of discrete nonlinear stochastic systems, the state space model is described as follows [1]:

$$\mathbf{x}(k+1) = \mathbf{f}(\mathbf{x}(k)) + \mathbf{w}(k), \quad (1)$$

$$\mathbf{z}(k) = \mathbf{h}(\mathbf{x}(k)) + \mathbf{v}(k), \quad (2)$$

where  $\mathbf{x}(k) \in R^n$  is the state of the target,  $\mathbf{z}(k) \in R^m$  is the measurement,  $\mathbf{f} : R^n \rightarrow R^n$  is the evolution process of the nonlinear state, and  $\mathbf{h} : R^n \rightarrow R^m$  is the corresponding nonlinear measurement mapping.

The process noise  $\mathbf{w}(k) \in R^n$  is a Gaussian white noise with zero means and variance  $\mathbf{Q}(k)$ .

The measurement noise  $\mathbf{v}(k) \in R^m$  is a Gaussian white noise with zero means and variance  $\mathbf{R}(k)$ .

**Hypothesis 1.** The process noise  $\mathbf{w}(k)$  and measurement noise  $\mathbf{v}(k)$  in the model are mutually statistically independent.

**Hypothesis 2.** The initial state of the system is  $\mathbf{x}(0)$ , with mean  $\mathbf{x}_0$  and variance  $\mathbf{P}_0$ , and it is uncorrelated with  $\mathbf{v}(k)$  and  $\mathbf{w}(k)$ .

**Hypothesis 3.** The process noise variance  $\mathbf{Q}(k)$  is known, but the measurement noise variance is only with an inaccuracy prior state  $\mathbf{R}_0$ .

For the nonlinear system described in (1) and (2), assuming that at the time  $k$ , we have the optimal estimation  $\hat{\mathbf{x}}(k-1|k-1)$  and the square root matrix of the error variance matrix  $\mathbf{S}(k-1|k-1)$ . Then, the state estimate  $\hat{\mathbf{x}}(k|k)$  and the square root of the variance matrix  $\mathbf{S}(k|k)$  can be calculated according to the standard SCKF algorithm.

For nonlinear systems with determined noise variance, the SCKF algorithm has better estimation performance. However, when the priori value of the measurement noise variance is inaccurate, the final estimation  $\hat{\mathbf{x}}(k|k)$  and  $\mathbf{S}(k|k)$  will have large errors.

### 3. Hybrid Adaptive SCKF Algorithm (HASCKF)

In order to improve the adaptive filtering accuracy for an inaccurate modeling system with unknown measurement noise variance, and to let the estimated noise variance be closer to the true noise variance, in this paper, a novel noise variance fusion estimation algorithm HASCKF is proposed based on the idea of weighted fusion.

#### 3.1. Estimation Method of Measurement Noise

**Theorem 1.** Assume that the measurement noise variance estimated by MAP and VB methods at time  $k$  are denoted as  $\hat{\mathbf{R}}^1(k)$  and  $\hat{\mathbf{R}}^2(k)$ , respectively. Then, the weighted fusion estimation of the measurement noise variance  $\hat{\mathbf{R}}^g(k)$  is

$$\hat{\mathbf{R}}^g(k) = \hat{\mathbf{R}}^1(k) + [T_1(k) - T_{12}(k)][T_1(k) + T_2(k) - 2T_{12}(k)]^{-1}[\hat{\mathbf{R}}^2(k) - \hat{\mathbf{R}}^1(k)], \quad (3)$$

where

$$\begin{cases} T_1(k) = \|\hat{\mathbf{R}}^g(k-1) - \hat{\mathbf{R}}^1(k)\|_F^2, \\ T_{12}(k) = \text{tr}[(\hat{\mathbf{R}}^g(k-1) - \hat{\mathbf{R}}^1(k))(\hat{\mathbf{R}}^g(k-1) - \hat{\mathbf{R}}^2(k))^T], \\ T_2(k) = \|\hat{\mathbf{R}}^g(k-1) - \hat{\mathbf{R}}^2(k)\|_F^2. \end{cases} \quad (4)$$

In the above equation,  $\|\cdot\|_F$  represents the Frobenius norm of the matrix, and the initial value of the fusion estimation  $\hat{\mathbf{R}}^g(0) = \mathbf{R}_0$ .

**Proof.** The weighted fusion estimation of the measurement noise variance  $\hat{\mathbf{R}}^g(k)$  can be expressed as the following linear combination:

$$\hat{\mathbf{R}}^g(k) = a_1(k)\hat{\mathbf{R}}^1(k) + a_2(k)\hat{\mathbf{R}}^2(k). \quad (5)$$

Under the condition  $a_1(k) + a_2(k) = 1$ , we minimize the performance criterion:

$$J(k) = \text{tr} \left\{ [(\mathbf{R}(k) - \hat{\mathbf{R}}^g(k))][(\mathbf{R}(k) - \hat{\mathbf{R}}^g(k))^T] \right\}$$

From  $a_1(k) + a_2(k) = 1$ , we can get  $a_1(k) = (1 - a_2(k))$ , then substituting it into formula (5), we obtain

$$\hat{\mathbf{R}}^g(k) = \hat{\mathbf{R}}^1(k) + a_2(k)[\hat{\mathbf{R}}^2(k) - \hat{\mathbf{R}}^1(k)]. \quad (6)$$

Then, the overall estimation error is

$$\begin{aligned} \tilde{\mathbf{R}}^g(k) &= \mathbf{R}(k) - \hat{\mathbf{R}}^g(k) \\ &= a_1(k)\tilde{\mathbf{R}}^1(k) + a_2(k)\tilde{\mathbf{R}}^2(k) \\ &= [1 - a_2(k)]\tilde{\mathbf{R}}^1(k) + a_2(k)\tilde{\mathbf{R}}^2(k). \end{aligned}$$

So, we can get

$$\begin{aligned} J(k) &= \text{tr}[\tilde{\mathbf{R}}^g(k)(\tilde{\mathbf{R}}^g(k))^T] \\ &= [1 - a_2(k)]^2 \text{tr}[\tilde{\mathbf{R}}^1(k)(\tilde{\mathbf{R}}^1(k))^T] + [1 - a_2(k)] \text{tr}[\tilde{\mathbf{R}}^1(k)(\tilde{\mathbf{R}}^2(k))^T] a_2(k) \\ &\quad + a_2(k) \text{tr}[\tilde{\mathbf{R}}^2(k)(\tilde{\mathbf{R}}^1(k))^T] [1 - a_2(k)] + a_2^2(k) \text{tr}[\tilde{\mathbf{R}}^2(k)(\tilde{\mathbf{R}}^2(k))^T]. \end{aligned} \quad (7)$$

Let  $\frac{\partial J(k)}{\partial a_2(k)} = 0$ , and after simplification, we can get

$$a_2(k) = [T_1(k) - T_{12}(k)][T_1(k) + T_2(k) - T_{12}(k) - T_{21}(k)]^{-1}, \quad (8)$$

where

$$\left\{ \begin{array}{l} T_1(k) = \text{tr}[\tilde{\mathbf{R}}^1(k)(\tilde{\mathbf{R}}^1(k))^T] \\ \quad = \text{tr}[(\mathbf{R}(k) - \hat{\mathbf{R}}^1(k))(\mathbf{R}(k) - \hat{\mathbf{R}}^1(k))^T] \\ \quad = \|\mathbf{R}(k) - \hat{\mathbf{R}}^1(k)\|_F^2, \\ T_{12}(k) = \text{tr}[\tilde{\mathbf{R}}^1(k)(\tilde{\mathbf{R}}^2(k))^T] \\ \quad = \text{tr}[(\mathbf{R}(k) - \hat{\mathbf{R}}^1(k))(\mathbf{R}(k) - \hat{\mathbf{R}}^2(k))^T], \\ T_{21}(k) = \text{tr}[\tilde{\mathbf{R}}^2(k)(\tilde{\mathbf{R}}^1(k))^T] \\ \quad = \text{tr}[(\mathbf{R}(k) - \hat{\mathbf{R}}^2(k))(\mathbf{R}(k) - \hat{\mathbf{R}}^1(k))^T] \\ \quad = T_{12}(k), \\ T_2(k) = \text{tr}[\tilde{\mathbf{R}}^2(k)(\tilde{\mathbf{R}}^2(k))^T] \\ \quad = \text{tr}[(\mathbf{R}(k) - \hat{\mathbf{R}}^2(k))(\mathbf{R}(k) - \hat{\mathbf{R}}^2(k))^T] \\ \quad = \|\mathbf{R}(k) - \hat{\mathbf{R}}^2(k)\|_F^2. \end{array} \right. \quad (9)$$

Since the measurement noise variance  $\mathbf{R}(k)$  is uncertain, Equation (9) cannot be directly calculated. For this reason, we replace  $\mathbf{R}(k)$  with fusion estimated value  $\hat{\mathbf{R}}^g(k-1)$  of the measurement noise variance at time  $(k-1)$ . Obviously, the initial value satisfies  $\hat{\mathbf{R}}^g(0) = \mathbf{R}_0$ , then substituting it to Equations (9), (8), and (6), respectively, we can get Equations (3) and (4).  $\square$

**Note 1** In the above theorem, since the measurement noise variance  $\mathbf{R}(k)$  is uncertain, we replace  $\mathbf{R}(k)$  with fusion estimate  $\hat{\mathbf{R}}^g(k-1)$  of the measurement noise variance. This approximate substitution has certain rationality, especially for the case of constant noise variance and slowly varying noise variance. From Theorem 1, the physical meanings of  $T_1(k)$  and  $T_2(k)$  are equivalent to the estimated error variance of the two noise variance estimation methods, and  $T_{12}(k)$  and  $T_{21}(k)$  are similar to their cross-variance.

**Inference 1.** If we do not consider the correlation between the estimation error of noise variance, that is  $T_{12}(k) = T_{21}(k) = 0$ , then the weighted fusion estimation of measurement noise variance  $\hat{\mathbf{R}}^g(k)$  is

$$\left\{ \begin{array}{l} \hat{\mathbf{R}}^g(k) = [T_1^{-1}(k) + T_2^{-1}(k)]^{-1} T_1^{-1}(k) \hat{\mathbf{R}}^1(k) \\ \quad + [T_1^{-1}(k) + T_2^{-1}(k)]^{-1} T_2^{-1}(k) \hat{\mathbf{R}}^2(k), \\ T^{-1}(k) = T_1^{-1}(k) + T_2^{-1}(k). \end{array} \right. \quad (10)$$

It can be directly derived according to the principle of simple convex combination fusion [1].

**Note 2** Obviously, the result of inference 1 is easy to generalize to the case where the number of noise variance estimators  $N_e \geq 3$ . Assume that the error variance of the  $i$ th estimator  $T_i(k) =$

$\|\hat{\mathbf{R}}^g(k-1) - \hat{\mathbf{R}}^i(k)\|_F^2$ . Assuming that the estimation errors of various estimation methods are not related to each other, the fusion estimation of noise variance is

$$\begin{cases} \hat{\mathbf{R}}^g(k) = \left[ \sum_j^{N_e} T_j^{-1}(k) \right]^{-1} \sum_j^{N_e} T_j^{-1}(k) \hat{\mathbf{R}}^j(k), \\ T_g^{-1}(k) = \sum_j^{N_e} T_j^{-1}(k), \end{cases} \quad (11)$$

where  $T_g(k)$  can be regarded as the error variance of noise variance fusion estimation. Since  $T_i(k) \geq 0$ ,  $T_g(k) \leq T_i(k)$ , which indicates that the estimated noise variance after fusion is superior to that of any single noise variance estimator.

### 3.2. HASCKF Algorithm

Based on Theorem 1, combined with the fading factor adjustment technique [11], we propose the hybrid adaptive SCKF algorithm (HASCKF). The principle block diagram is shown in Figure 1.

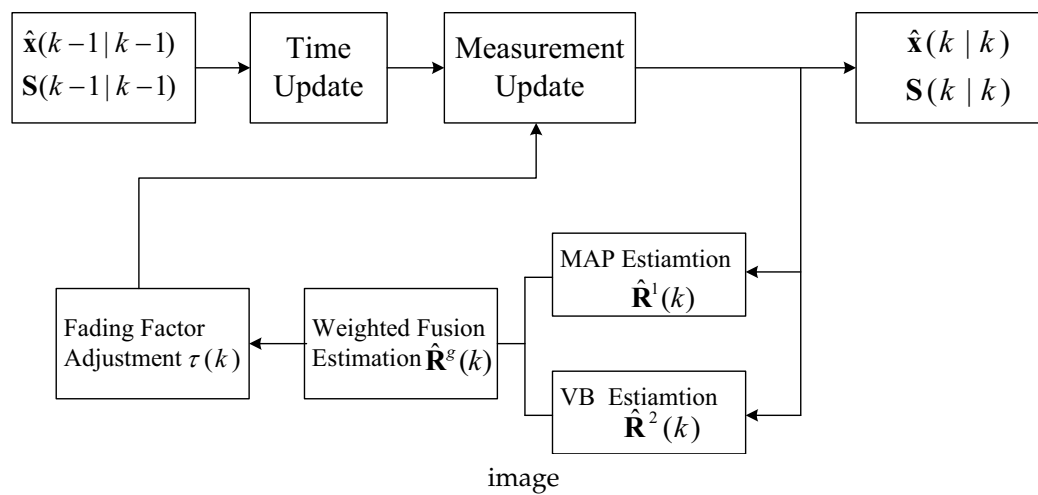


Figure 1. Principle block diagram of the HASCKF algorithm.

Firstly, the variance of measurement noise is estimated by the MAP estimator and VB method respectively. Then, the weighted fusion technique is introduced to fuse the two noise variance estimators. Finally, the variance matrix of the SCKF measurement update is adjusted by the fading factor adjustment technique to obtain the final state estimation and the root-mean square error (RMSE) variance matrix. The detailed implementation of HASCKF is described in the following Theorem 2.

**Theorem 2.** Consider a class of nonlinear system as described in (1) and (2). Under the condition of Hypotheses 1–3, if the optimal estimate  $\hat{\mathbf{x}}(k-1|k-1)$  and the square-root matrix of the estimation error variance  $\mathbf{S}(k-1|k-1)$  have been obtained, the state estimate  $\hat{\mathbf{x}}(k|k)$  and the square-root matrix of the estimation error variance  $\mathbf{S}(k|k)$  can be calculated according to the following steps:

#### Step 1: Time Update

$$\mathbf{x}_i(k-1|k-1) = \mathbf{S}(k-1|k-1)\boldsymbol{\xi}_i + \hat{\mathbf{x}}(k-1|k-1), \quad (12)$$

$$\mathbf{x}_i^*(k|k-1) = f(\mathbf{x}_i(k-1|k-1)), \quad (13)$$

$$\hat{\mathbf{x}}(k|k-1) = \frac{1}{N_x} \sum_{i=1}^{N_x} \mathbf{x}_i^*(k|k-1). \quad (14)$$

In (12)–(14),  $i = 1, 2, \dots, N_x$ ,  $N_x = 2n$ . The parameters  $\xi_i$  are given below ( $\varepsilon_i$  is an  $n$ -order unit vector):

$$\xi_i = \begin{cases} \sqrt{N_x/2} \cdot \varepsilon_i & , i = 1, 2, \dots, n, \\ -\sqrt{N_x/2} \cdot \varepsilon_{i-n_x} & , i = n+1, n+2, \dots, N_x. \end{cases}$$

Calculate the square root of the variance matrix:

$$\mathbf{S}(k|k-1) = \text{Tri}([\mathbf{X}^*(k|k-1) \quad \mathbf{S}_Q(k)]), \quad (15)$$

where  $\text{Tri}(\cdot)$  represents a triangular operation.  $\mathbf{S}_Q(k)$  represents the square-root of the new process noise variance  $\mathbf{Q}(k)$ , that is,  $\mathbf{Q}(k) = \mathbf{S}_Q(k)\mathbf{S}_Q^T(k)$ , and

$$\mathbf{X}^*(k|k-1) = \frac{1}{\sqrt{N_x}} [\mathbf{x}_1^*(k|k-1) - \hat{\mathbf{x}}(k|k-1), \mathbf{x}_2^*(k|k-1) - \hat{\mathbf{x}}(k|k-1), \dots, \mathbf{x}_{N_x}^*(k|k-1) - \hat{\mathbf{x}}(k|k-1)].$$

### Step 2: Measurement Update

(1) According to Equations (16)–(18), we can calculate the predicted value  $\hat{\mathbf{z}}(k|k-1)$ .

$$\mathbf{x}_i(k|k-1) = \mathbf{S}(k|k-1)\xi_i + \hat{\mathbf{x}}(k|k-1), \quad (16)$$

$$\mathbf{z}_i(k|k-1) = \mathbf{h}(\mathbf{x}_i(k|k-1)), \quad (17)$$

$$\hat{\mathbf{z}}(k|k-1) = \frac{1}{N_x} \sum_{i=1}^{N_x} \mathbf{z}_i(k|k-1). \quad (18)$$

(2) The measurement noise variance  $\hat{\mathbf{R}}^1(k)$  of the MAP estimator is calculated by using Lemma 1 (19) or Lemma 2 (20).

**Lemma 1** ([14]). When the measurement noise variance is constant, the suboptimal MAP estimate  $\hat{\mathbf{R}}(k)$  of noise variance  $\mathbf{R}(k)$  can be obtained by the recursive calculation:

$$\begin{aligned} \hat{\mathbf{R}}(k) = \frac{1}{k} [ & (k-1)\hat{\mathbf{R}}(k-1) + \tilde{\mathbf{z}}(k)\tilde{\mathbf{z}}^T(k) \\ & - \mathbf{z}(k|k-1)\mathbf{z}^T(k|k-1)], \end{aligned} \quad (19)$$

where  $\tilde{\mathbf{z}}(k) = \mathbf{z}(k) - \hat{\mathbf{z}}(k|k-1)$  is the residual vector of measurement. The initial value  $\hat{\mathbf{R}}(0) = \mathbf{R}_0$ .

**Lemma 2** ([15]). When the measurement noise variance is time-varying, the suboptimal MAP estimate  $\hat{\mathbf{R}}(k)$  of noise variance  $\mathbf{R}(k)$  can be obtained by the recursive calculation:

$$\begin{aligned} \hat{\mathbf{R}}(k) = [ & 1 - d(k-1)]\hat{\mathbf{R}}(k-1) + d(k-1)[\tilde{\mathbf{z}}(k)\tilde{\mathbf{z}}^T(k) \\ & - \mathbf{z}(k|k-1)\mathbf{z}^T(k|k-1)], \end{aligned} \quad (20)$$

where  $d(k) = (1-b)/(1-b^{k+1})$ .  $b$  is the forgetting factor, and its value range is usually between 0.95 and 0.99.

(3) Using the following Equation (21), combined with Equations (16)–(18), (22), (25), (26), and (23)–(32), we can iteratively compute the measurement noise variance estimation  $\hat{\mathbf{R}}^2(k)$  of the VB method.

Estimation of prediction parameters of measurement noise variance by VB method:

$$\begin{cases} \zeta(k|k-1) = \rho \cdot \zeta(k-1), \\ \eta(k|k-1) = \rho \cdot \eta(k-1), \end{cases} \quad (21)$$

where “.” represents the point operation in Matlab.  $\rho = [\rho_1, \dots, \rho_m]^T$ ,  $\zeta(k) = [\zeta_1(k), \dots, \zeta_m(k)]^T$ ,  $\eta(k) = [\eta_1(k), \dots, \eta_m(k)]^T$ .  $\zeta_i(k)$  and  $\eta_i(k)$  are two parameters of the inverse gamma distribution.  $\rho_i \in (0, 1)$  is the predictive weighting factor. It reflects the degree of correlation between the noise at the last moment and the noise at the current moment. When the difference between the measurement noise variance at the last moment and the measurement noise variance at the current moment is small, a larger  $\rho_i$  value should be used. Conversely,  $\rho_i$  should take a smaller value.

The square-root of the variance matrix:

$$\mathbf{S}_{zz}(k|k-1) = \text{Tri}a([\mathbf{z}(k|k-1) \quad \mathbf{S}_R(k)]), \quad (22)$$

where  $\mathbf{S}_R(k)$  represents the square root of measurement noise variance  $\mathbf{R}(k)$  (namely,  $\mathbf{R}(k) = \mathbf{S}_R(k)\mathbf{S}_R^T(k)$ ) and

$$\begin{aligned} \mathbf{z}(k|k-1) &= \frac{1}{\sqrt{N_x}} [\mathbf{z}_1(k|k-1) - \hat{\mathbf{z}}(k|k-1) \\ &\quad \mathbf{z}_2(k|k-1) - \hat{\mathbf{z}}(k|k-1) \\ &\quad \dots \quad \mathbf{z}_{N_x}(k|k-1) - \hat{\mathbf{z}}(k|k-1)], \\ \mathbf{P}_{xz}(k|k-1) &= \mathbf{x}(k|k-1)\mathbf{z}^T(k|k-1), \end{aligned} \quad (23)$$

where

$$\begin{aligned} \mathbf{x}(k|k-1) &= \frac{1}{\sqrt{N_x}} [\mathbf{x}_1(k|k-1) - \hat{\mathbf{x}}(k|k-1) \\ &\quad \mathbf{x}_2(k|k-1) - \hat{\mathbf{x}}(k|k-1) \\ &\quad \dots \quad \mathbf{x}_{N_x}(k|k-1) - \hat{\mathbf{x}}(k|k-1)]. \end{aligned}$$

Then, the VB method gets  $\hat{\mathbf{R}}(k)$  through  $M$  iterations:

Iterative initialization: let  $t = 1$ , for a given number of iterations  $M$ , we have

$$\begin{cases} \zeta(k) = [1/2, 1/2, \dots, 1/2]^T + \zeta(k|k-1), \\ \hat{\mathbf{x}}^1(k|k-1) = \hat{\mathbf{x}}(k|k-1). \end{cases} \quad (24)$$

Calculating the estimate of the measurement noise variance:

$$\hat{\mathbf{R}}^t(k) = \text{diag}(\eta^t(k) \cdot \zeta(k)), \quad (25)$$

where  $\text{diag}(A)$  represents a diagonal matrix composed of matrix  $A$  diagonal elements.

Use Equations (16) and (32) to calculate the  $t$ -th iteration state estimate  $\hat{\mathbf{x}}^t(k|k)$  and the root of its mean square error matrix  $\mathbf{S}^t(k|k)$ . If  $t < M$ , update the parameter  $\eta^t(k)$ .

$$\eta^t(k) = \eta(k|k-1) + (\mathbf{z}(k) - \hat{\mathbf{z}}^t(k|k-1))^2/2 + \text{diag}\{P_{zz}^t(k|k)\}/2. \quad (26)$$

Let  $t = t + 1$ ,  $\hat{\mathbf{x}}^t(k|k-1) = \hat{\mathbf{x}}^{t-1}(k|k)$ , return to the beginning of the iterative.

When  $t = M$ , end the iteration, we can get

$$\begin{cases} \eta(k) = \eta^M(k) \\ \hat{\mathbf{R}}(k) = \hat{\mathbf{R}}^M(k) \end{cases}, \quad \begin{cases} \hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}^M(k|k) \\ \mathbf{S}(k|k) = \mathbf{S}^M(k|k) \end{cases}. \quad (27)$$

(4) Calculate the fusion estimation  $\hat{\mathbf{R}}^g(k)$  of the measurement noise variance according to Equation (3).

(5) According to Equations (28) and (29),  $\mathbf{S}_{zz}(k|k-1)$  and  $\mathbf{P}_{xz}(k|k-1)$  are adaptively adjusted using the FFA fading factor.

$$\mathbf{S}_{zz}(k|k-1) = \text{Tri}a\left([\mathbf{z}(k|k-1) \Big/ \sqrt{\tau(k)} \quad \mathbf{S}_R(k)]\right) \quad (28)$$

$$\mathbf{P}_{xz}(k|k-1) = \frac{1}{\tau(k)} \mathbf{x}(k|k-1) \mathbf{z}^T(k|k-1), \quad (29)$$

where  $\tau(k)$  is the adaptive factor,  $0 < \tau(k) \leq 1$ , and it is calculated by the following Lemma 3.

**Lemma 3** ([11]). *For nonlinear systems with unknown measurement noise variance, the adaptive fading factor is determined by the following equation:*

$$\tau(k) = \begin{cases} 1, & \text{tr}[\tilde{\mathbf{z}}(k|k-1)\tilde{\mathbf{z}}^T(k|k-1)] \leq \text{tr}[\mathbf{S}_{zz}(k|k-1)\mathbf{S}_{zz}^T(k|k-1)], \\ \frac{\text{tr}[\mathbf{S}_{zz}(k|k-1)\mathbf{S}^T(k|k-1)]}{\text{tr}[\tilde{\mathbf{z}}(k|k-1)\tilde{\mathbf{z}}^T(k|k-1)]}, & \text{tr}[\tilde{\mathbf{z}}(k|k-1)\tilde{\mathbf{z}}^T(k|k-1)] > \text{tr}[\mathbf{S}_{zz}(k|k-1)\mathbf{S}_{zz}^T(k|k-1)]. \end{cases}$$

In the above equation,  $\text{tr}$  denotes the trace of the matrix,  $\tilde{\mathbf{z}}(k|k-1) = \mathbf{z}(k) - \hat{\mathbf{z}}(k|k-1)$  is the measurement residual error vector.

(6) The following Equations (30)–(32) are used to obtain the updated estimation  $\hat{\mathbf{x}}(k|k)$  and the root of its mean square error variance  $\mathbf{S}(k|k)$ .

$$\mathbf{K}(k) = \left( \mathbf{P}_{xz}(k|k-1) / \mathbf{S}_{zz}^T(k|k-1) \right) / \mathbf{S}_{zz}(k|k-1), \quad (30)$$

where the symbol “/” indicates the matrix right divide operation (e.g.,  $A/B = AB^{-1}$ )

$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)[\mathbf{z}(k) - \hat{\mathbf{z}}(k|k-1)], \quad (31)$$

$$\mathbf{S}(k|k) = \text{Tria}([\mathbf{x}(k|k-1) - \mathbf{K}(k)\mathbf{z}(k|k-1) \quad \mathbf{K}(k)\mathbf{S}_R(k)]). \quad (32)$$

Proof: We can directly derive Theorem 1 and Lemma 3 by Lemma 1, omitted here.

**Note 3** Theorem 2 only shows the hybrid adaptive filtering algorithm when the number of noise variance estimators  $N_e$  is 2. Obviously, when  $N_e > 2$ , we first use various noise variance estimators to estimate  $\hat{\mathbf{R}}^j(k)$  and then calculate the fused estimate  $\hat{\mathbf{R}}^g(k)$  of the measurement noise variance according to Equation (11). In other words, only steps (2)–(4) in Theorem 2 need to be adjusted.

### 3.3. HASCIF Information Filter

Compared with traditional filtering, the information filter may not require prior information when it is initialized, and thus has better numerical performance. In addition, the use of an information filter to design a fusion algorithm is also simpler. In the information filter, the state estimate and its estimation error variance matrix are replaced by information vector and information matrix, respectively. Subsequently, we give its corresponding information filtering form (HASCIF) on the basis of HASCKF. According to the literature [20,21], we can obtain the implementation process of HASCIF as follows:

**Step 1:** Time update [20,21]

$$\hat{\mathbf{y}}(k|k-1) = \mathbf{Y}(k|k-1)\hat{\mathbf{x}}(k|k-1), \quad (33)$$

$$\mathbf{Y}(k|k-1) = \mathbf{P}^{-1}(k|k-1) = [\mathbf{S}(k|k-1)\mathbf{S}^T(k|k-1)]^{-1}, \quad (34)$$

where  $\hat{\mathbf{y}}(k|k-1)$  and  $\mathbf{Y}(k|k-1)$  are the predicted information vector and the predicted information matrix, respectively.  $\hat{\mathbf{x}}(k|k-1)$  and  $\mathbf{S}(k|k-1)$  can be calculated according to Equations (14) and (15).



### Step 2: Measurement update

$$\begin{aligned}\hat{\mathbf{y}}(k|k) &= \mathbf{Y}(k|k)\hat{\mathbf{x}}(k|k) \\ &= \hat{\mathbf{y}}(k|k-1) + \mathbf{Y}(k|k-1)\mathbf{P}_{xz}(k|k)[\hat{\mathbf{R}}^g(k)]^{-1}\{\hat{\mathbf{z}}(k|k-1) + \mathbf{P}_{xz}^T(k|k)\mathbf{Y}^T(k|k-1)\hat{\mathbf{x}}(k|k-1)\} \\ &\triangleq \hat{\mathbf{y}}(k|k-1) + \boldsymbol{\theta}(k),\end{aligned}\quad (35)$$

$$\begin{aligned}\mathbf{Y}(k|k) &= \mathbf{Y}(k|k-1) + \mathbf{Y}(k|k-1)\mathbf{P}_{xz}(k|k)[\hat{\mathbf{R}}^g(k)]^{-1}(k)\mathbf{P}_{xz}^T(k|k)\mathbf{Y}^T(k|k-1) \\ &\triangleq \mathbf{Y}(k|k-1) + \boldsymbol{\Theta}(k).\end{aligned}\quad (36)$$

In Equations (35) and (36), the information state vectors  $\hat{\mathbf{y}}(k|k)$  and  $\mathbf{Y}(k|k)$  are the information vector and information matrix of the state estimate, respectively.  $\boldsymbol{\theta}(k)$  and  $\boldsymbol{\Theta}(k)$  are the information contribution vector and the information contribution matrix.  $\hat{\mathbf{R}}^g(k)$  is determined by formula (3), and the cross-variance matrix  $\mathbf{P}_{xz}(k|k)$  can be calculated by Formula (29).

### 3.4. Performance Analysis of HASCKF

References [22–24] proposed the bounded convergence theorem of the UKF algorithm, and Reference [14] extended its theorem to the adaptive cubature Kalman filter (ACKF). In this section, the CKF bounded convergence theorem proposed in Reference [14] and the Cramer–Rao lower bound (CRLB) [25] are used to analyze the convergence of the HASCKF algorithm.

#### 3.4.1. Stability Analysis

Assuming that the noise variances  $\mathbf{Q}(k)$  and  $\mathbf{R}(k)$  are known accurately, and considering the nonlinear systems (1), (2), and the standard CKF algorithm [4], the corresponding state error variance matrix can be written as [14,22,24]:

$$\begin{aligned}\mathbf{P}(k|k-1) &= \mathbf{x}(k|k-1)\mathbf{x}^T(k|k-1) + \mathbf{Q}(k-1) \\ &= \mathbf{A}(k)\mathbf{P}(k-1|k-1)\mathbf{A}^T(k) + \delta\mathbf{P}(k|k-1) + \mathbf{Q}(k-1) \\ &= \mathbf{A}(k)\mathbf{P}(k-1|k-1)\mathbf{A}^T(k) + \boldsymbol{\Xi}(k),\end{aligned}\quad (37)$$

$$\begin{aligned}\mathbf{P}_{zz}(k|k-1) &= \mathbf{z}(k|k-1)\mathbf{z}^T(k|k-1) + \mathbf{R}(k) \\ &= \mathbf{G}(k)\mathbf{P}(k|k-1)\mathbf{G}^T(k) + \delta\mathbf{P}_{zz}(k|k-1) + \mathbf{R}(k) \\ &= \mathbf{G}(k)\mathbf{P}(k|k-1)\mathbf{G}^T(k) + \boldsymbol{\Sigma}(k),\end{aligned}\quad (38)$$

$$\mathbf{K}(k) = \mathbf{P}(k|k-1)\mathbf{G}^T(k)[\mathbf{G}(k)\mathbf{P}(k|k-1)\mathbf{G}^T(k) + \boldsymbol{\Sigma}(k)]^{-1},\quad (39)$$

where

$$\delta\mathbf{P}(k|k-1) = \mathbf{x}(k|k-1)\mathbf{x}^T(k|k-1) - \mathbf{A}(k)\mathbf{P}(k-1|k-1)\mathbf{A}^T(k),\quad (40)$$

$$\delta\mathbf{P}_{zz}(k|k-1) = \mathbf{z}(k|k-1)\mathbf{z}^T(k|k-1) - \mathbf{G}(k)\mathbf{P}(k|k-1)\mathbf{G}^T(k).\quad (41)$$

Each correlation matrix is defined as follows:

$$\begin{cases} \mathbf{A}(k) = \beta(k)\mathbf{F}(k), \\ \mathbf{B}(k) = \alpha(k)\mathbf{H}(k), \\ \mathbf{C}(k) = \mathbf{I} - \mathbf{K}(k)\alpha(k)\mathbf{H}(k), \end{cases}\quad (42)$$

$$\mathbf{G}(k) = \begin{cases} \alpha(k)\mathbf{H}(k)\gamma^T(k), & n \geq m, \\ \gamma^T(k)\alpha(k)\mathbf{H}(k), & n < m, \end{cases}\quad (43)$$

$$\begin{cases} \boldsymbol{\Xi}(k) = \delta\mathbf{P}(k|k-1) + \mathbf{Q}(k-1), \\ \boldsymbol{\Sigma}(k) = \delta\mathbf{P}_{zz}(k|k-1) + \mathbf{R}(k), \end{cases}\quad (44)$$

where

$$\mathbf{F}(k) = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}(k-1|k-1)}, \quad \mathbf{H}(k) = \left. \frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}(k|k-1)}.$$

$\alpha(k)$ ,  $\beta(k)$ , and  $\gamma(k)$  are auxiliary diagonal matrixes. Refer to Reference [22] for specific meanings.

**Lemma 4** ([14]). Consider the nonlinear systems (1), (2), and the standard CKF algorithm. If  $\forall k \geq 0$ , both satisfy the following two assumptions:

(1) There are non-zero real numbers  $a_{\min}$ ,  $a_{\max}$ ,  $b_{\max}$ ,  $c_{\max}$ ,  $g_{\min}$ , and  $g_{\max}$  existing to let the following formulas be established:

$$\begin{cases} a_{\min}^2 \mathbf{I} \leq \mathbf{A}(k) \mathbf{A}^T(k) \leq a_{\max}^2 \mathbf{I}, & \mathbf{B}(k) \mathbf{B}^T(k) \leq b_{\max}^2 \mathbf{I}, \\ g_{\min}^2 \mathbf{I} \leq \mathbf{G}(k) \mathbf{G}^T(k) \leq g_{\max}^2 \mathbf{I}, & \mathbf{C}(k) \mathbf{C}^T(k) \leq c_{\max}^2 \mathbf{I}, \\ [\mathbf{G}(k) - \mathbf{B}(k)][\mathbf{G}(k) - \mathbf{B}(k)]^T \leq (g_{\max} - b_{\max})^2 \mathbf{I}. \end{cases} \quad (45)$$

(2) There are positive real numbers  $p_{\min}$ ,  $p_{\max}$ ,  $q_{\max}$ ,  $r_{\max}$ ,  $\Xi_{\min}$ ,  $\Xi_{\max}$ , and  $\Sigma_{\min}$  existing to let the following forms be established:

$$\begin{cases} p_{\min} \mathbf{I} \leq \mathbf{P}(k|k) \leq p_{\max} \mathbf{I}, & \mathbf{Q}(k) \leq q_{\max} \mathbf{I}, \\ \mathbf{R}(k) \leq r_{\max} \mathbf{I}, & \Xi(k) \leq \Xi_{\max} \mathbf{I}, \\ \Xi(k) > \Xi_{\min} \mathbf{I}, & \Sigma(k) > \Sigma_{\min} \mathbf{I}, \end{cases} \quad (46)$$

where

$$\begin{cases} \Sigma_{\min} = \max(\Sigma_1, \Sigma_2), \\ \Sigma_1 = a_{\max}^2 (g_{\max} - b_{\max})^2 (p_{\max} + p_{\max}^2 a_{\max}^2 \Xi_{\min}^{-1}), \\ \Sigma_2 = b_{\max}^2 (a_{\max}^2 p_{\max} + \Xi_{\max}) - g_{\max}^2 (a_{\min}^2 p_{\min} \Xi_{\min}). \end{cases}$$

Then, the standard CKF state estimation error will be mean-square bounded, that is, the algorithm is stable and convergent.

**Note 4** Lemma 4 shows that the statistical characteristics of noise are closely related to the stability of the CKF algorithm. In addition, if Lemma 4 is established, the SCKF with known noise statistics is also stable and convergent. This is because the theoretical framework of SCKF is consistent with that of CKF. The mean square root matrix  $\mathbf{S}(k|k-1)$  and  $\mathbf{S}(k|k)$  in CKF are only used when transferring the error variance. The root mean square matrix in the SCKF algorithm is obtained through triangulation of the matrix, thus avoiding the filter divergence caused by the non-positive definite variance matrix in the numerical calculation, so it has better stability while ensuring the accuracy of estimation with CKF.

**Theorem 3.** If the standard SCKF algorithm is stable and convergent when the statistical characteristics of noise are known accurately, the introduction of the noise variance fusion estimator and the adaptive fade factor can ensure the stable convergence of the HASCKF algorithm.

**Proof.** When the measured noise variance matrix  $\mathbf{R}(k)$  is inaccurate, other sufficient conditions in the bounded convergence theorem can be satisfied, but the condition (46) will be affected.

(a) First only consider the influence of the weighted fusion noise estimator. Let  $\Delta \mathbf{R}(k) = \hat{\mathbf{R}}^g(k) - \mathbf{R}(k)$ , then formula (40) can be rewritten as follows:

$$\begin{aligned} \tilde{\Sigma}(k) &= \delta \mathbf{P}_{zz}(k|k-1) + \hat{\mathbf{R}}^g(k) \\ &= \delta \mathbf{P}_{zz}(k|k-1) + \mathbf{R}(k) + \Delta \mathbf{R}(k). \end{aligned} \quad (47)$$

When  $\hat{\mathbf{R}}^g(k) \geq \mathbf{R}(k)$ , then  $\Delta \mathbf{R}(k) \geq 0$ , and  $\tilde{\Sigma}(k)$  will become larger. Obviously, condition (46) still holds and the HASCKF algorithm still converges steadily.

When  $\hat{\mathbf{R}}^g(k) < \mathbf{R}(k)$ , then  $\Delta\mathbf{R}(k) < 0$ , and  $\tilde{\Sigma}(k)$  will become smaller. Now, condition (46) may not be satisfied. However, the noise variance weighted fusion estimator  $\hat{\mathbf{R}}^g(k)$  estimates and corrects the inaccurate noise variance matrix in real time so that it gradually tracks the real value  $\mathbf{R}(k)$ , thus making  $\Delta\mathbf{R}(k) \rightarrow 0$ ,  $\tilde{\Sigma}(k) \rightarrow \Sigma(k)$ . In this way,  $\Delta\mathbf{R}(k) \geq 0$  is gradually satisfied to Equation (42) to ensure stable convergence of the HASCKF algorithm.

(b) Further consider the effect of fading factors. According to Theorem 2, the variance matrix for measuring residuals is

$$\begin{aligned}\tilde{\mathbf{P}}_{zz}(k|k-1) &= \mathbf{S}_{zz}(k|k-1)\mathbf{S}_{zz}^T(k|k-1) \\ &= \frac{1}{\tau(k)}\mathbf{z}(k|k-1)\mathbf{z}^T(k|k-1) + \hat{\mathbf{R}}^g(k).\end{aligned}\quad (48)$$

Then, Equation (41) is rewritten as:

$$\begin{aligned}\delta\tilde{\mathbf{P}}_{zz}(k|k-1) &= \frac{1}{\tau(k)}\mathbf{z}(k|k-1)\mathbf{z}^T(k|k-1) - \mathbf{G}(k)\mathbf{P}(k|k-1)\mathbf{G}^T(k) \\ &= \left[\frac{1}{\tau(k)} - 1\right]\mathbf{z}(k|k-1)\mathbf{z}^T(k|k-1) + \mathbf{z}(k|k-1)\mathbf{z}^T(k|k-1) - \mathbf{G}(k)\mathbf{P}(k|k-1)\mathbf{G}^T(k) \\ &= \left[\frac{1}{\tau(k)} - 1\right]\mathbf{z}(k|k-1)\mathbf{z}^T(k|k-1) + \delta\mathbf{P}_{zz}(k|k-1) \\ &\triangleq \Delta\mathbf{P}_{zz}(k|k-1) + \delta\mathbf{P}_{zz}(k|k-1).\end{aligned}\quad (49)$$

Therefore, Equation (44) can be written as

$$\begin{aligned}\tilde{\Sigma}(k) &= \delta\tilde{\mathbf{P}}_{zz}(k|k-1) + \hat{\mathbf{R}}^g(k) \\ &= \delta\mathbf{P}_{zz}(k|k-1) + \mathbf{R}(k) + \Delta\mathbf{P}_{zz}(k|k-1) + \Delta\mathbf{R}(k) \\ &\triangleq \Sigma(k) + \Delta\Sigma(k).\end{aligned}\quad (50)$$

It is known from the definition of fading factor  $0 < \tau(k) \leq 1$ , so we can get

$$\Delta\mathbf{P}_{zz}(k|k-1) = [1/\tau(k) - 1]\mathbf{z}(k|k-1)\mathbf{z}^T(k|k-1) \geq 0. \quad (51)$$

Similar to the analysis in (a), when  $\Delta\mathbf{R}(k) \geq 0$  and  $\Delta\Sigma(k) \geq 0$ , it is obvious that condition (46) holds and the stability of the HASCKF algorithm remains. When  $\Delta\mathbf{R}(k) < 0$ , if  $\Delta\mathbf{P}_{zz}(k|k-1)$  is large enough,  $\Delta\Sigma(k) \geq 0$  can still be satisfied, and condition (46) still holds, so the HASCKF algorithm converges steadily. If  $\Delta\mathbf{P}_{zz}(k|k-1)$  is not enough to guarantee  $\Delta\Sigma(k) \geq 0$ , the introduction of the weighted fusion noise variance estimator can also make  $\Delta\mathbf{R}(k) \rightarrow 0$ , so that  $\Delta\Sigma(k) \geq 0$  is established stepwise to ensure stable convergence of the HASCKF algorithm.

In summary, the introduction of the noise variance fusion estimator and adaptive fading factor in the HASCKF algorithm can ensure the stable convergence of the algorithm.  $\square$

**Note 5** This theorem combined with Note 2 shows that the noise variance estimation based on weighted fusion is superior to the estimate of any single noise variance estimator. Therefore, the hybrid adaptive HASCKF estimation algorithm has better stability than the adaptive SCKF using a single noise variance estimation algorithm.

### 3.4.2. Filtering Accuracy Analysis

There is a lower bound on the minimum variance unbiased estimator of the state of the nonlinear filtering algorithm. It is widely used to evaluate the performance of nonlinear estimation. In practice, the lower limit of Cramer–Rao Lower Bound (CRLB) is commonly used. Denote

$\mathbf{X}_{0:k} = \{\mathbf{x}(0), \mathbf{x}(1), \dots, \mathbf{x}(k)\}$ ,  $\mathbf{Z}_{0:k} = \{\mathbf{z}(0), \mathbf{z}(1), \dots, \mathbf{z}(k)\}$ , and  $p(\mathbf{Z}_{0:k}, \mathbf{X}_{0:k})$  is the joint probability density of  $(\mathbf{Z}_{0:k}, \mathbf{X}_{0:k})$ .  $\hat{\mathbf{x}}(k)$  is the unbiased estimation of  $\mathbf{x}(k)$ . Then, CRLB is defined as [25]:

$$\mathbf{P}(k|k) \triangleq E\{[\mathbf{x}(k) - \hat{\mathbf{x}}(k)][\mathbf{x}(k) - \hat{\mathbf{x}}(k)]^T\} \geq \mathbf{J}^{-1}(k), \quad (52)$$

where  $\mathbf{J}(k)$  is the Fisher information matrix:

$$\mathbf{J}(k) = E \left[ -\frac{\partial^2 \ln p(\mathbf{Z}_{0:k}, \mathbf{X}_{0:k})}{\partial^2 \mathbf{x}^2(k)} \right]. \quad (53)$$

**Theorem 4** Assume that the nonlinear filter is applied to system (1), (2). Then

$$\mathbf{P}(k|k) \geq \mathbf{J}^{-1}(k), \quad (54)$$

where

$$\mathbf{J}(k) = [\mathbf{Q}(k-1) + \mathbf{F}(k)\mathbf{J}^{-1}(k-1)\mathbf{F}^T(k)]^{-1} + \mathbf{H}^T(k)\mathbf{R}^{-1}(k)\mathbf{H}(k). \quad (55)$$

**Proof.** We use  $\mathbf{J}(k)$  to denote the information matrix of state  $\mathbf{x}(k)$ . Then, the information matrix  $\mathbf{J}(k)$  can be recursively calculated according to the following formula [26]:

$$\mathbf{J}(k) = \mathbf{D}_{22}(k) - \mathbf{D}_{21}(k)[\mathbf{J}(k-1) + \mathbf{D}_{11}(k)]^{-1}\mathbf{D}_{12}(k), \quad (56)$$

where

$$\begin{cases} \mathbf{D}_{11}(k) = E\{-\Delta_{\mathbf{x}(k-1)}^{\mathbf{x}(k-1)} \ln p[\mathbf{x}(k)|\mathbf{x}(k-1)]\}, \\ \mathbf{D}_{12}(k) = E\{-\Delta_{\mathbf{x}(k-1)}^{\mathbf{x}(k)} \ln p[\mathbf{x}(k)|\mathbf{x}(k-1)]\}, \\ \mathbf{D}_{21}(k) = E\{-\Delta_{\mathbf{x}(k)}^{\mathbf{x}(k-1)} \ln p[\mathbf{x}(k)|\mathbf{x}(k-1)]\} \\ \quad = \mathbf{D}_{12}^T(k), \\ \mathbf{D}_{22}(k) = E\{-\Delta_{\mathbf{x}(k)}^{\mathbf{x}(k)} \ln p[\mathbf{x}(k)|\mathbf{x}(k-1)]\} \\ \quad + E\{-\Delta_{\mathbf{x}(k)}^{\mathbf{x}(k)} \ln p[\mathbf{z}(k)|\mathbf{x}(k)]\}, \end{cases} \quad (57)$$

where  $\Delta_{\mathbf{x}}^{\mathbf{y}} = \nabla_{\mathbf{x}} \nabla_{\mathbf{y}}^T$  means the second-order operator.  $\nabla_{\mathbf{x}} = \frac{\partial}{\partial \mathbf{x}}$  is the first-order operator. Because process noise and measurement noise are Gaussian white noise, we have

$$\begin{aligned} & -\ln p[\mathbf{x}(k)|\mathbf{x}(k-1)] \\ &= -\ln \left\{ \frac{e^{\{-\frac{1}{2}[\mathbf{x}(k) - \mathbf{f}(\mathbf{x}(k-1))]^T \mathbf{Q}^{-1}(k-1)[\mathbf{x}(k) - \mathbf{f}(\mathbf{x}(k-1))\}]}{\sqrt{2\pi} |\mathbf{Q}(k-1)|} \right\} \\ &= c_1 + \frac{1}{2}[\mathbf{x}(k) - \mathbf{f}(\mathbf{x}(k-1))]^T \mathbf{Q}^{-1}(k-1)[\mathbf{x}(k) - \mathbf{f}(\mathbf{x}(k-1))] \end{aligned} \quad (58)$$

$$\begin{aligned} & -\ln p[\mathbf{z}(k)|\mathbf{x}(k)] \\ &= -\ln \left\{ \frac{e^{\{-\frac{1}{2}[\mathbf{z}(k) - \mathbf{h}(\mathbf{x}(k))]^T \mathbf{R}^{-1}(k)[\mathbf{z}(k) - \mathbf{h}(\mathbf{x}(k))\}]}{\sqrt{2\pi} |\mathbf{R}(k)|} \right\} \\ &= c_2 + \frac{1}{2}[\mathbf{z}(k) - \mathbf{h}(\mathbf{x}(k))]^T \mathbf{R}^{-1}(k)[\mathbf{z}(k) - \mathbf{h}(\mathbf{x}(k))], \end{aligned} \quad (59)$$

where constant  $c_1 = \ln[\sqrt{2\pi} |\mathbf{Q}(k-1)|]$ ,  $c_2 = \ln[\sqrt{2\pi} |\mathbf{R}(k)|]$ .

Substituting Equations (58) and (59) into Equation (57), then we can get

$$\begin{cases} \mathbf{D}_{11}(k) = E\{\nabla_{\mathbf{x}(k-1)} \mathbf{f}^T(\mathbf{x}(k-1))\} \mathbf{Q}^{-1}(k-1) [\nabla_{\mathbf{x}(k-1)} \mathbf{f}^T(\mathbf{x}(k-1))]^T \\ \quad = \mathbf{F}^T(k) \mathbf{Q}^{-1}(k-1) \mathbf{F}(k), \\ \mathbf{D}_{12}(k) = E\{\nabla_{\mathbf{x}(k-1)} \mathbf{f}^T(\mathbf{x}(k-1))\} \mathbf{Q}^{-1}(k-1) \\ \quad = \mathbf{F}^T(k) \mathbf{Q}^{-1}(k-1) = \mathbf{D}_{21}^T(k), \\ \mathbf{D}_{22}(k) = \mathbf{Q}^{-1}(k-1) + E\{\nabla_{\mathbf{x}(k)} \mathbf{h}^T(\mathbf{x}(k))\} \mathbf{R}^{-1}(k) [\nabla_{\mathbf{x}(k)} \mathbf{h}^T(\mathbf{x}(k))]^T \\ \quad = \mathbf{Q}^{-1}(k-1) + \mathbf{H}^T(k) \mathbf{R}^{-1}(k) \mathbf{H}(k). \end{cases} \quad (60)$$

$$\mathbf{J}(k) = \mathbf{Q}^{-1}(k-1) + \mathbf{H}^T(k) \mathbf{R}^{-1}(k) \mathbf{H}(k) - \mathbf{Q}^{-1}(k-1) \mathbf{F}(k) \times [\mathbf{J}(k-1) + \mathbf{F}^T(k) \mathbf{Q}^{-1}(k-1) \mathbf{F}(k)]^{-1} \mathbf{F}^T(k) \mathbf{Q}^{-1}(k-1). \quad (61)$$

Equation (56) can be obtained by applying a matrix inversion lemma to Equation (61).  $\square$

#### 4. Simulation

In this section, two numerical simulation examples are provided to display and verify the performance of the SACKF algorithm proposed in this paper, mainly including the following contents:

(1) For a class of inaccurate modeling with unknown measurement noise variance, we compare the performance difference between the weighted fusion estimator (referred to as WF-NE) and the single noise estimator (e.g., MAP estimation or VB estimation, respectively denoted as MAP-NE and VB-NE).

(2) For a class of inaccurate modeling with unknown measurement noise variance, we study the advantages and disadvantages of the HASCKF algorithm and the standard SCKF algorithm, and the equivalence between the HASCKF and HASCIF algorithms.

**Example 1.** This example is used to evaluate the performance of three kinds of noise estimators, MAP-NE, VB-NE, and WF-NE. Considering the following first-order nonlinear discrete dynamic system:

$$x(k) = 0.5x(k-1) + \frac{0.2x(k-1)}{1+x^2(k-1)} + w(k), \quad (62)$$

$$z(k) = \frac{x^2(k)}{20} + v(k), \quad (63)$$

where  $w(k)$  and  $v(k)$  are mutually independent Gaussian white noise sequences. Assume that the initial value of the system state and the error variance matrix are

$$x_0 = 2, \quad P_0 = 0.01, \quad (64)$$

and the system state initial value  $x_0$  is independent of the two noises. The process noise statistic  $Q(k) = 0.001$ . In the following, simulation experiments are performed for two cases where the measurement noise variance  $R(k)$  is a constant and piecewise continuous function.

(1) Situation 1: If the measurement noise variance is constant and  $R(k) = 0.012$ . Assume that the imprecise measurement noise variance of the initial value  $\hat{R}_0 = 0.04$ . In the simulation, MAP-NE adopted the estimator described in Equation (19). The parameters of VB-NE were selected as follows:  $\rho = 1 - e^{-5}$ ,  $\zeta(0) = 1$ ,  $\eta(0) = 0.04$ ,  $M = 1$ .

In order to compare the performance of various algorithms, we adopted the absolute error (AE) and the mean absolute error (MAE), which are calculated as follows:

$$AE(k) = |y(k) - \hat{y}(k)|, \quad MAE(k) = \frac{1}{N_s} \sum_{k=1}^{N_s} AE(k), \quad (65)$$

where  $y(k)$  and  $\hat{y}(k)$  are the value to be estimated and the estimated value respectively.  $N_s$  is the number of simulation steps. In this case,  $N_s = 1000$ .

The estimated noise variance of the three kinds of noise estimators are shown in Figures 2 and 3, and the estimation error is given in Table 1.

(2) Situation 2: If the measurement noise value is time varying, the true statistical characteristics meet the following formula:

$$R(k) = \begin{cases} R & 0 \leq k \leq 500 \\ 2R & 501 \leq k \leq 1000 \end{cases}, \text{ and } R = 0.012. \quad (66)$$

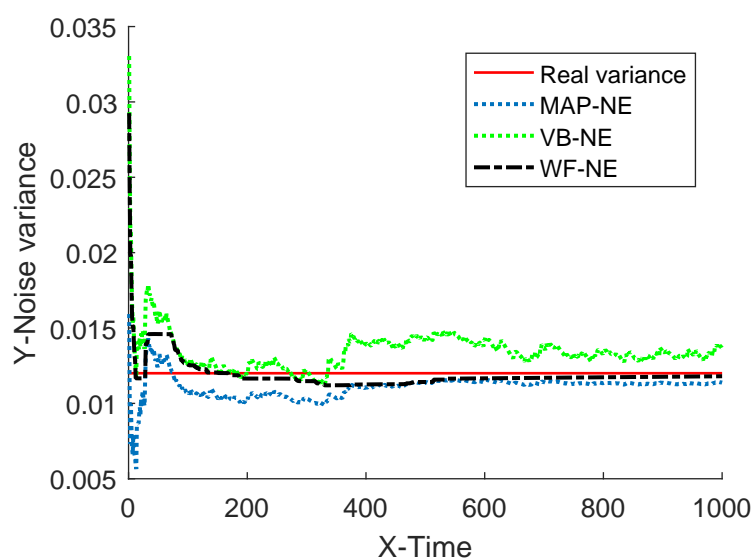
Assume that the known inaccurate initial value of the measurement noise variance is  $\hat{R}_0 = 0.08$ . In the simulation, MAP-NE adopted the estimator described in Equation (20), and the forgetting factor  $b$  was 0.98. The parameters of VB-NE were selected as  $\rho = 1 - e^{-5}$ ,  $\zeta(0) = 1$ ,  $\eta(0) = 0.08$ ,  $M = 1$ . The estimation results of the three kinds of noise estimators on the measured noise variance are shown in Figures 4 and 5, and the mean absolute errors of several algorithms are given in Table 2.

**Table 1.** The mean absolute error of three noise estimation algorithms in Situation 1.

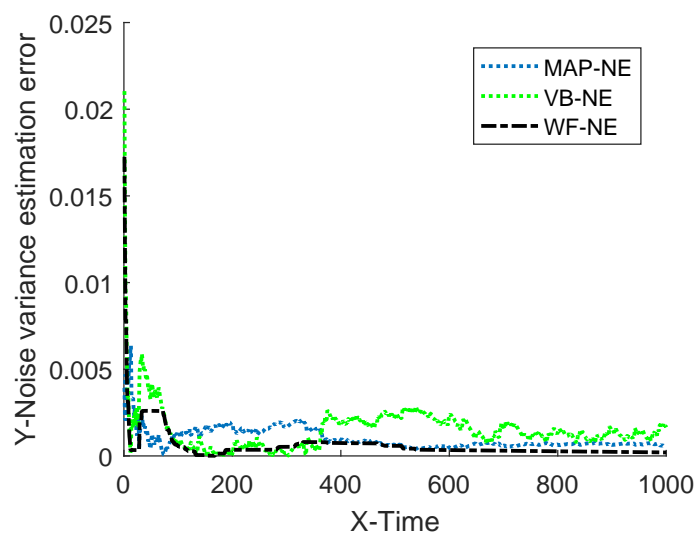
Algorithm	MAP-NE	VB-NE	WF-NE
Mean absolute error	0.0010	0.0015	0.0005
CPU time cost	0.2188	0.2188	0.3725

**Table 2.** The mean absolute error of three noise estimation algorithms in Situation 2.

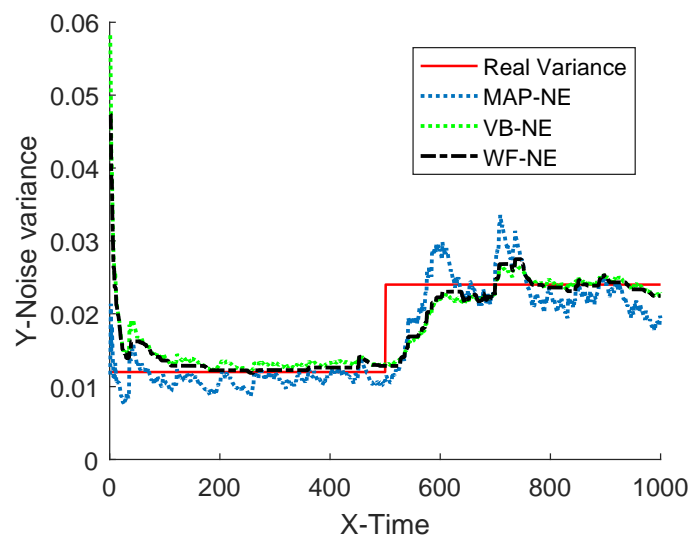
Algorithm	MAP-NE	VB-NE	WF-NE
Mean absolute error	0.0024	0.0022	0.0019



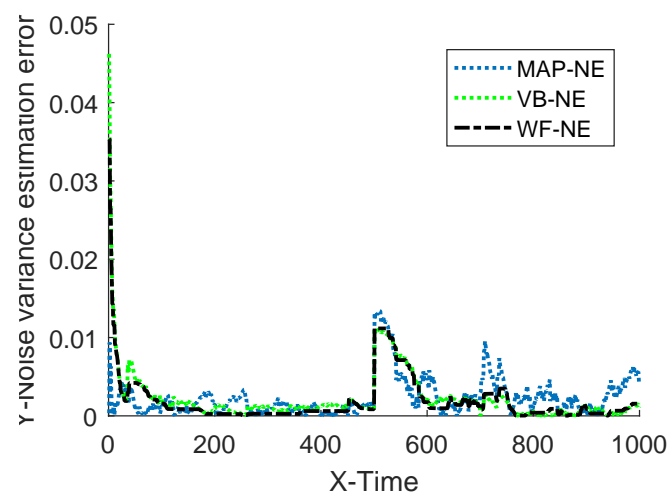
**Figure 2.** Estimation of measurement noise variance.



**Figure 3.** Absolute estimation error of measurement noise variance.



**Figure 4.** Estimation of measurement noise variance.



**Figure 5.** Absolute estimation error of measurement noise variance.

The simulation results of Example 1 show that the fusion estimator WF-NE proposed in Theorem 1 could obtain the best estimated result for either the constant unknown measurement noise variance or for the time-varying one, compared with the MAP-NE and VB-NE. It had the same result with the analysis conclusion in Note 2. In Situation 1, as shown in Figures 2 and 3, compared with the fusion results of MAP-NE and VB-NE, the WF-NE had better estimates in most simulation steps. Due to the randomness of noise, not all the noise variance estimates obtained by WF-NE were better than the ones solved by other methods, especially for the estimation of the time-varying variance in Situation 2. As shown in Figure 3, the MAP-NE method had relatively large estimation errors in the time interval [160–360], and the VB-NE method had relatively large estimation errors in the time interval [390–800]. Nevertheless, the WF-NE could avoid large variation of the noise variance estimation error. Similarly, the general trend of the WF-NE method was best in Figure 4, although the MAP-NE method or the VB-NE method were best in some small time intervals. Although the CPU time cost of WF-NE (0.3725 for 1000 simulation steps) was larger than the other two methods (0.2188 for 1000 simulation steps), it was still acceptable. Meanwhile, the mean absolute errors of WF-NE were smaller than the other two methods, for both the constant noise variance (in Situation 1) and for the time-varying variance (in Situation 2).

**Example 2.** This example is used to verify the pros and cons of the hybrid adaptive SCKF estimation algorithm HASCKF proposed by Theorem 2 and the standard SCKF algorithm described in Reference [4]. Assume that the target is moving in a uniform linear motion on a two-dimensional plane. The system state  $x(k) = [x(k), \dot{x}(k), y(k), \dot{y}(k)]^T$ , where  $x(k)$  and  $y(k)$  are the position components in the east–north coordinate system.  $\dot{x}(k)$  and  $\dot{y}(k)$  are the corresponding velocity components, respectively. Then, the state equation can be described as:

$$x(k) = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot x(k-1) + w(k), \quad (67)$$

where the sampling period  $T = 0.5$  s, the process noise  $w(k)$  is zero mean Gaussian white noise. Its statistical characteristic is  $Q(k) = 0.1 \times \text{diag}(Q1, Q2)$ , and  $Q1 = \begin{bmatrix} T^3/3 & T^2/2 \\ T^2/2 & T \end{bmatrix}$ .

Consider a radar to track the target. The nonlinear measurement equation can be expressed as

$$z(k) = \begin{bmatrix} \sqrt{x^2(k) + y^2(k)} \\ \arctan\left(\frac{y(k)}{x(k)}\right) \end{bmatrix} + v(k). \quad (68)$$

The real statistical characteristic of the measuring noise  $R(k) = \text{diag}\{(4\text{m})^2, (0.1^\circ)^2\}$ . During the simulation, the simulation time was set as  $N_s = 100$  s. Suppose that the inaccuracy initial value of the measurement noise variance  $\hat{R}(0) = \text{diag}\{(81\text{m})^2, (0.3^\circ)^2\}$ . MAP-NE adopts the estimator described in Equation (19), and the parameters of VB-NE were selected as  $\rho = [1 - e^{-5}, 1 - e^{-5}]^T$ ,  $\zeta(0) = [1, 1]^T$ ,  $\eta(0) = [100, 0.02]^T$ ,  $M = 1$ . The initial state estimate and the estimation error variance matrix are

$$x_0 = [10,000 \text{ m}, 150 \text{ m/s}, 15,000 \text{ m}, 200 \text{ m/s}]^T, \\ P_0 = \text{diag}\{(100 \text{ m})^2, (14 \text{ m/s})^2, (100 \text{ m})^2, (15 \text{ m/s})^2\}.$$

The estimated results of SCKF and HASCKF algorithms are shown in Figures 6–10. The estimation error is given in Table 3.

From Figure 6 to Figure 10, it can be clearly seen that the standard SCKF estimation showed a large deviation after 40 s. However, HASCKF could still better estimate the state of the target. This is because the standard SCKF adopts an inaccurate prior noise variance  $\hat{R}(0)$ , while the HASCKF proposed in this



paper estimates and corrects the inaccurate measurement noise variance, thus ensuring the accuracy and stability of the algorithm. As analyzed in Example 1, due to the randomness of noise, not all the noise variance estimates obtained by WF-NE were better. Therefore, the estimates of some components of the state obtained by SCKF were better than the one solved by HASCKF at some time in the interval [10 s, 40 s]. However, the general tend of the proposed HASCKF was better than SCKF, as shown in Figures 6–10.

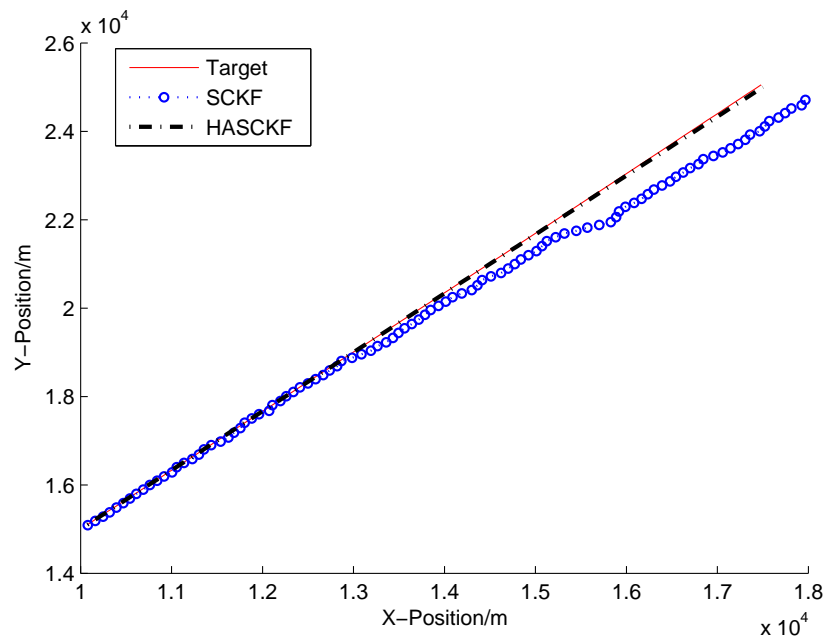


Figure 6. Target's trajectory and tracking result of SCKF and HASCKF.

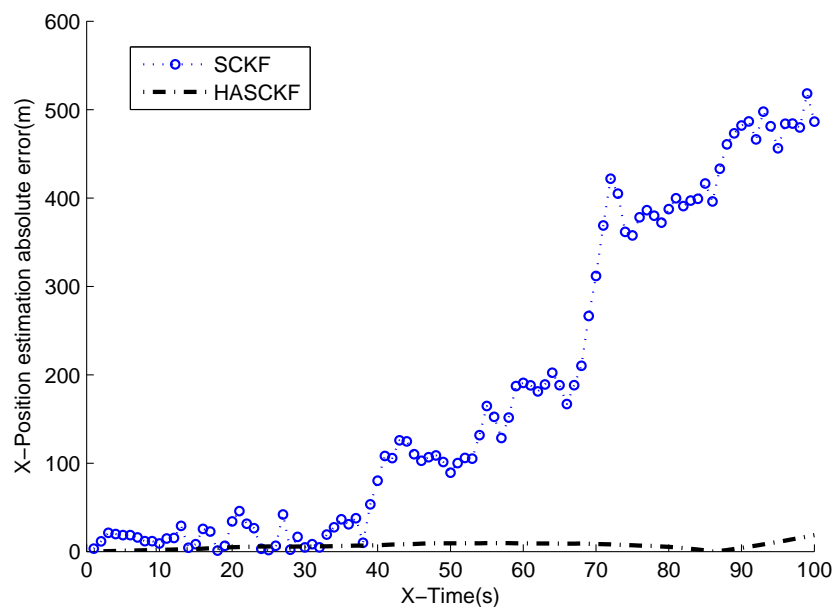
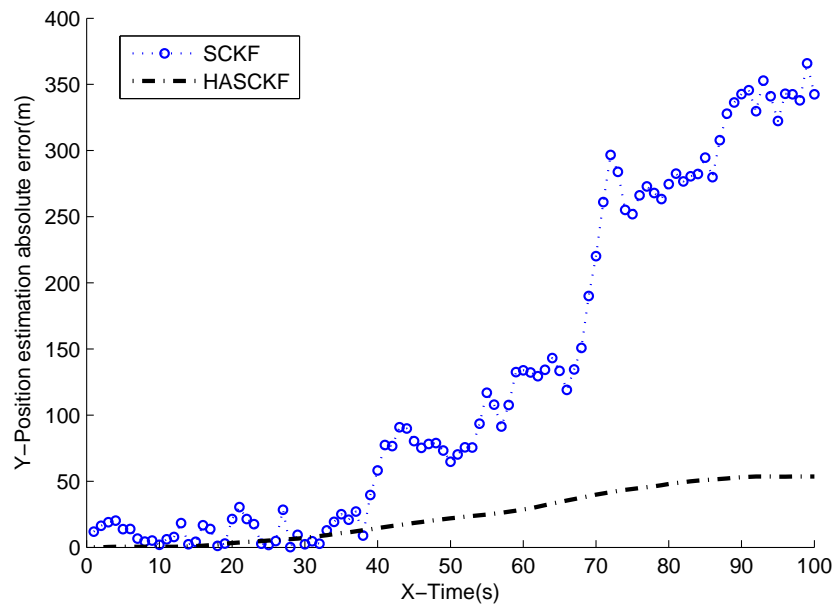
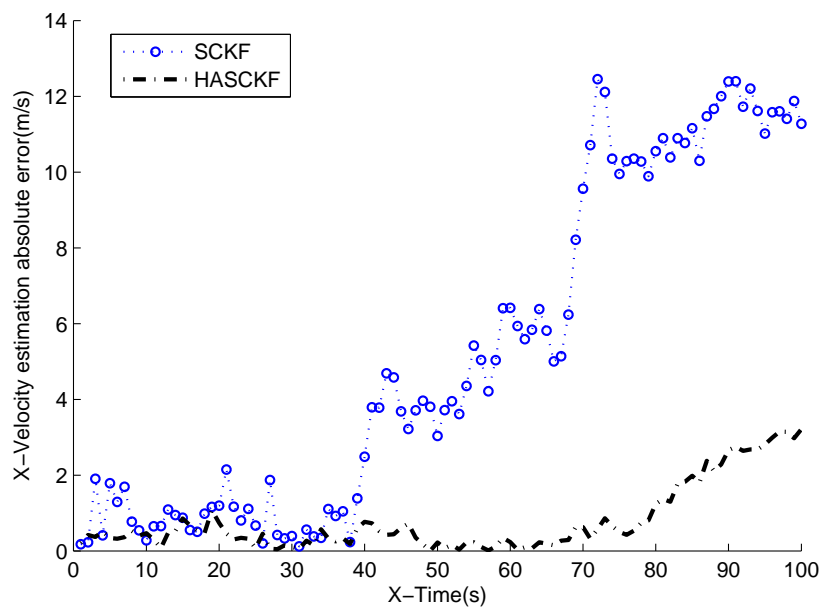


Figure 7. Absolute error curves of X-displacement.

**Table 3.** The mean absolute error of two algorithms.

Mean Absolute Error	Algorithms	
	SCKF	HASCKF
X-Position (m)	182.9477	6.4784
X-Velocity (m/s)	5.2113	0.7898
Y-Position (m)	129.5228	24.6147
Y-Velocity (m/s)	3.6766	1.0834

**Figure 8.** Absolute error curves of Y-displacement.**Figure 9.** Absolute error curves of X-velocity.

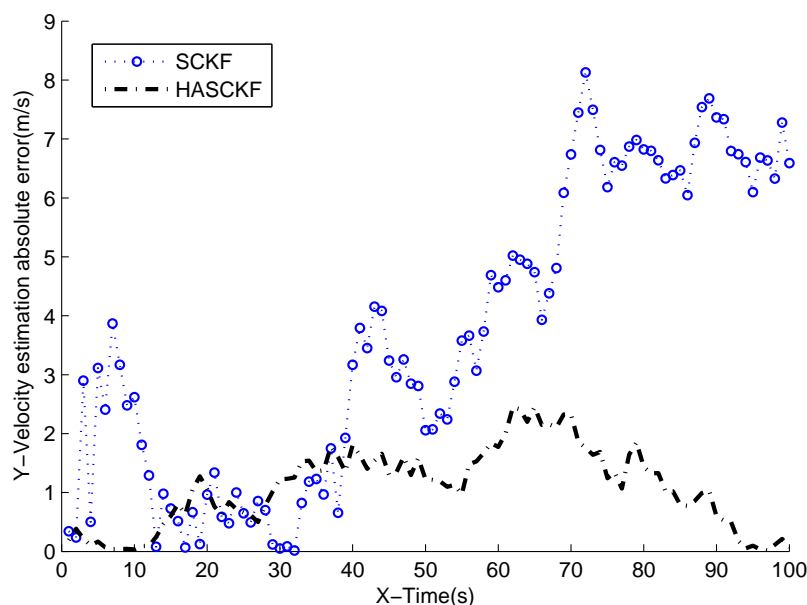


Figure 10. Absolute error curves of Y-velocity.

## 5. Conclusions

In this paper, aiming at the filtering problem caused by the inaccurate measurement noise variance in real systems, a weighted fusion estimation method is designed and a class of hybrid adaptive filtering structures is proposed, based on different measurement noise variance estimators. The stability and the filtering accuracy of the hybrid adaptive filter are analyzed and discussed theoretically. The simulation results showed that the algorithm had excellent accuracy and stability.

The work that can continue to be studied in the future includes further study on the research and design of fusion methods based on hybrid adaptive filtering in nonlinear centralized fusion framework and distributed fusion framework, research on functional equivalence among various fusion algorithms, and so on.

**Author Contributions:** Conceptualization, Y.S.; Methodology, Y.S.; Software, D.B.; Validation, X.Z.; Formal Analysis, X.F. and X.T.; Writing—Original Draft Preparation, Y.S. and X.F.; Writing—Review & Editing, X.Z. and X.T.

**Funding:** This research was supported in part by the National Natural Science Foundation of China (Grant Nos. 61503174, 61503336, 61503136, and 61602141), Henan higher education school young backbone teacher training program (2017GGJS160), Henan University of Animal Husbandry and Economy key scientific research innovation fund (XKYCXJJ2017009).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Han, C.Z.; Zhu, H.Y.; Duan, Z.S. *Multi-source Information Fusion*, 2nd ed.; Tsinghua University Press: Beijing, China, 2010.
2. Ge, Q.; Shao, T.; Yang, Q.; Shen, X.; Wen, C. Multisensor Nonlinear Fusion Methods Based on Adaptive Ensemble Fifth-Degree Iterated Cubature Information Filter for Biomechatronics. *IEEE Trans. Syst. Man Cybern. Syst.* **2017**, *46*, 912–925. [[CrossRef](#)]
3. Julier, S.J.; Uhlmann, J.K. New extension of the Kalman filter to nonlinear systems. In Proceedings of the Signal Processing, Sensor Fusion, and Target Recognition VI, Orlando, FL, USA, 28 July 1997; pp. 182–193.
4. Arasaratnam, I.; Haykin, S. Cubature kalman filter. *IEEE Trans. Autom. Control* **2009**, *54*, 1254–1269. [[CrossRef](#)]

5. Arasaratnam, I.; Haykin, S.; Thomas, R.H. Cubature kalman filtering for continuous-discrete systems: Theory and simulations. *IEEE Trans. Signal Process.* **2010**, *58*, 4977–4993. [[CrossRef](#)]
6. Hernandez, W.; De Vicente, J.; Sergiyenko, O.Y.; Fernández, E. Improving the Response of Accelerometers for Automotive Applications by Using LMS Adaptive Filters: Part II. *Sensors* **2010**, *10*, 952–962. [[CrossRef](#)] [[PubMed](#)]
7. Yang, Y.; Gao, W. A new learning statistic for adaptive filter based on predicted residuals. *Prog. Nat. Sci. Mater. Int.* **2006**, *16*, 833–837.
8. Sage, A.; Husa, G.W. Adaptive filtering with unknown prior statistics. In Proceedings of the Joint Automatic Control Conference, Boulder, CO, USA, 5–7 August 1969; pp. 760–769.
9. Sage, A.P.; Husa, G.W. Algorithms for sequential adaptive estimation of prior statistics. In Proceedings of the 8th Symp on Adaptive Processes Decision and Control, Session 6a, University Park, PA, USA, 17–19 November 1969; pp. 61–70.
10. Gao, S.; Hu, G.; Zhong, Y. Windowing and random weighting-based adaptive unscented Kalman filter. *Int. J. Adapt. Control Signal Process.* **2015**, *29*, 201–223. [[CrossRef](#)]
11. Gao, W.G.; Yang, Y.X.; Cuix, Q. Application of adaptive kalman filtering algorithm in imu/GPS integrated navigation system. *Geo-Spat. Inf. Sci.* **2007**, *10*, 22–26. [[CrossRef](#)]
12. Li, W.; Sun, S.; Jia, Y.; Du, J. Robust unscented Kalman filter with adaptation of process and measurement noise variances. *Digit. Signal Process.* **2016**, *48*, 93–103. [[CrossRef](#)]
13. Zheng, B.; Fu, P.; Li, B.; Yuan, X. A Robust Adaptive Unscented Kalman Filter for Nonlinear Estimation with Uncertain Noise variance. *Sensors* **2018**, *18*, 808. [[CrossRef](#)]
14. Ding, J.L.; Xiao, J. Design of adaptive cubature Kalman filter based on maximum a posteriori estimation. *Control Decis.* **2014**, *29*, 327–334.
15. Zhao, L.; Wang, X.X.; Sun, M.; Ding, J.C.; Yan, C. Adaptive UKF Filtering Algorithm Based on Maximum a Posterior Estimation and Exponential Weighting. *Acta Autom. Sin.* **2010**, *36*, 1007–1019. [[CrossRef](#)]
16. Sarkka, S.; Nummenmaa, A. Recursive noise adaptive kalman filtering by variational bayesian approximations. *IEEE Trans. Autom. Control* **2009**, *54*, 596–600. [[CrossRef](#)]
17. Ge, Q.; Wei, Z.; Liu, M.; Yu, J.; Wen, C. Adaptive Quantized Estimation Fusion Using Strong Tracking Filtering and Variational Bayesian. *IEEE Trans. Syst. Man Cybern. Syst.* **2017**, 1–12. [[CrossRef](#)]
18. Sun, C.; Zhang, Y.; Wang, G.; Gao, W. A New Variational Bayesian Adaptive Extended Kalman Filter for Cooperative Navigation. *Sensors* **2018**, *18*, 2538. [[CrossRef](#)] [[PubMed](#)]
19. Li, J.; Deng, F.; Chen, J. A Fast Distributed Variational Bayesian Filtering for Multisensor LTV System With Non-Gaussian Noise. *IEEE Trans. Cybern.* **2018**, 1–13. [[CrossRef](#)] [[PubMed](#)]
20. Chandra, K.P.B.; Gu, D.W.; Postlethwaite, I. Square root cubature information filter. *IEEE Sens. J.* **2013**, *13*, 750–758. [[CrossRef](#)]
21. Ge, Q.B.; Xu, D.X.; Wen, C.L. Cubature information filters with correlated noises and their applications in decentralized fusion. *Signal Process.* **2014**, *94*, 434–444. [[CrossRef](#)]
22. Xiong, K.; Zhang, H.Y.; Chan, C.W. Performance evaluation of ukf-based nonlinear filtering. *Automatica* **2006**, *42*, 261–270. [[CrossRef](#)]
23. Wu, Y.X.; Hu, D.W.; Hu, X.P. Comments on “performance evaluation of ukf-based nonlinear filtering”. *Automatica* **2007**, *43*, 567–568. [[CrossRef](#)]
24. Xiong, K.; Zhang, H.Y.; Chan, C.W. Author’s reply to “comments on ‘performance evaluation of ukf-based nonlinear filtering’”. *Automatica* **2007**, *43*, 569–570. [[CrossRef](#)]
25. Farina, A.; Ristic, B.; Benvenuti, D. Tracking a ballistic target: Comparison of several nonlinear filters. *IEEE Trans. Aerosp. Electron. Syst.* **2002**, *38*, 854–867. [[CrossRef](#)]
26. Tichavsky, P.; Muravchik, C.H.; Nehorai, A. Posterior cramer-Rao bounds for discrete-time nonlinear filtering. *IEEE Trans. Signal Process.* **1998**, *46*, 1386–1396. [[CrossRef](#)]

