Article

Second-Order EKF White Noise Estimator Design for Hybrid Systems

Yuxiang Liang 1, Huihong Zhao 1,*, Yunlong Shang 2 and Hailong Meng 1

1 School of Mathematics and Statistics, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China; 10431200770@stu.qlu.edu.cn (Y.L.); mhl@qlu.edu.cn (H.M.)
2 School of Control Science and Engineering, Shandong University, Jinan 250061, China; yshang@sdu.edu.cn
* Correspondence: huihongzhao@qlu.edu.cn; Tel.: +86-182-5312-9269

Abstract: The extended Kalman filter (EKF) has a wide range of applications (especially in power battery management systems) with a rapidly increasing market share. It aims to minimize the symmetric loss function (mean square error) and it has high accuracy and efficiency in battery state estimation. This study deals with the second-order extended Kalman filter-based process and the measurement white noise estimation problem for nonlinear continuous-discrete systems. The design of the white noise filter and smoother were, firstly, converted into a linear estimation problem by the second-order Taylor series expansion approximation and the function that makes the second-order term approximately equivalent to the estimation error variance. Secondly, based on the projection formula of the Kalman filtering (KF) theory and the Lemma of expectation for quadratic and quartic product traces of random vectors, the second-order EKF was derived. Then, to generate white noise estimators in the forms of filtering and smoothing, we derived a recursive solution, using an innovation method. Finally, a numerical example is given to show the effectiveness of the proposed method.

Keywords: white noise; KF; continuous-discrete systems; EKF; projection formula

1. Introduction

Process and measurement noise estimation is a task worthy of attention, and it has wide applications in many fields, such as battery condition estimation, multifrequency signal estimation, oil seismic exploration, and image processing [1–5]. Based on error covariance-matrix information, the process white noise estimator was proposed by the KF approach [4]. G. Z. Dai and J. Mendel pioneered the study of white noise estimation with application in oil exploration [6]. For the linear discrete time-varying stochastic system (multi-model and multi-sensor), the optimal weight fusion Kalman estimator and white noise deconvolution were given, respectively [7]. A unified white noise estimate theory based on a modern time series analysis approach was presented, which included a process and measurement white noise estimator design, and proposed a new approach for steady-state optimal state estimation [8]. For linear discrete-time non-Gaussian systems, according to the polynomial filtering theory, a solution to the quadratic estimation problem of non-Gaussian noise was given in [9]. H. Zhao and Z. Li presented a novel Kalman-like nonlinear non-Gaussian noise estimation method based on the packet dropout probability distribution and polynomial filtering technique [10]. W. Liu and Z. Deng solved the design problem of robust white noise deconvolution estimators for a class of uncertain systems with missing measurements, uncertain noise variances, and linearly correlated white noises [11]. The innovation method for the linear least squares estimation problem was extended by T. Kailath to deal with nonstationary continuous time processes on the finite time domain [12]. The nonlinear system is approximated to the linear system by real-time linear Taylor approximation; EKF is designed based on KF. This is the idea of the EKF, which was originally proposed by Stanley Schmidt, so that the Kalman filter could be
applied to nonlinear spacecraft navigation problems [13]. It can be seen from the above discussion that the research on white noise estimation of linear systems has been relatively mature, but there are few studies on the white noise estimation of nonlinear systems; reports on the white noise estimation of continuous discrete hybrid nonlinear systems are even less.

In recent years, state estimator designs for nonlinear system have been actively researched [14,15], and the second-order EKF was better than the first-order EKF in this area [16,17]. D. Simon presented the continuous-discrete system EKF (also called hybrid system EKF) [18]. M. De la Sen and N. Luo dealt with the design of linear observers for a class of linear hybrid systems. Moreover, such systems were composed of continuous-time and digital substrates [19]. In order to solve the estimation problem of continuous-discrete linear systems with parametric uncertainties, V. Shin, D. Y. Kim et al. proposed a novel sub-optimal filter by summing the local KF with weights, depending only on time instants [20]. For nonlinear hybrid stochastic systems, G.Y. Kulikov and M.V. Kulikova Gennady proposed a novel square root algorithm in order to solve the lack of square root implementation within the high-degree cubature KF [21]. In [22], the authors presented the derivation of the dynamical equations of a second-order filter, which estimated the states of the nonlinear system on the base of discrete noisy measurements. For the state of charge estimation of the lithium-ion battery in the linear hybrid systems, the authors of [23] proved that the second-order EKF could improve the estimation effect compared with the first-order EKF. Y. Wang and H. Zhang proposed the accurate Gaussian sum-smoothing method, which was derived by extended-cubature Kalman filters to approximate the non-Gaussian estimation densities as a finite number of weighted sums of Gaussian densities [24].

To the authors’ knowledge, the study of white noise estimation based on second-order EKF for hybrid systems has not been reported. Thus, we discuss the second-order EKF-based white noise estimation problem for a nonlinear hybrid system in this paper. The estimation problem is aimed to minimize a symmetric loss function (mean square error). By the second-order Taylor series expansion approximation, the function that makes the second-order term approximately equivalent to the estimation error variance and projection formula, and the second-order EKF formula, are derived. The Lemmas of expectation for quadratic and quartic product traces of random vectors are proved in detail by using the knowledge of probabilistic property analysis. Then, the continuous process white noise estimator and the discrete measurement white noise estimator are calculated by the Riccati equation, respectively. The main contributions of this paper are as follows: (i) to the best of our knowledge, the process and measurement white noise estimators of second-order EKF for continuous-discrete systems are presented, for the first time. (ii) The results of this paper enrich the traditional theory of white noise estimation and could directly extend to discrete nonlinear systems or continuous nonlinear systems. (iii) The white noise estimation algorithm for hybrid systems, proposed in this paper, is actually a symmetric solution to the fault estimation problem of such systems, under the assumption that the fault signal is white noise. Therefore, the proposed algorithm can be used for fault estimation of hybrid systems under the assumption that the fault signal is white noise.

The rest of this paper is organized as follows. Section 2 introduces some preparations of the second order Taylor expansion approximate for nonlinear hybrid systems, and presents the problem description. Section 3 presents the state estimation of the second-order EKF for systems, with continuous-time system dynamics and discrete-time measurements, and proposes the process and measurement white noise estimator by projection formula. Section 4 compares the performances of white noise estimations for first-order and second-order EKFs, using an example. Finally, we summarize the research results.

**Notation 1.** The superscripts ‘−1’ and ‘T’ stand for the inverse and transpose of a matrix, respectively. \( E[\cdot] \) denotes the expectation operator. \( \delta_{ij} \) is the Kronecker delta function, \( \delta_{ij} = 0 \) for \( i \neq j \) and \( \delta_{ii} = 1 \). \( R^n \) denotes the n-dimensional Euclidean space. For a real matrix, \( P > 0 (P < 0, \text{respectively}) \) means that \( P \) is symmetric and positive (negative, respectively) definite. \( [\cdot] \) is an
integer ceiling function, which is the largest integer not exceeding t. \( f, g \) denotes the covariance of \( f \) and \( g^T \).

2. Materials and Methods

Let us consider the hybrid system with continuous-time system dynamics and discrete-time measurements:

\[
\begin{align*}
\dot{x} &= f(x, u, t) + w(t), \\
y_k &= h_k(x_k) + v_k,
\end{align*}
\]

(1)

where \( f(\cdot) \) and \( h_k(\cdot) \) are nonlinear functions, \( t \in [a, b] \) is the continuous-time index, \([a, b]\) is a finite interval on the real line. \( x \in \mathbb{R}^n \) is the unknown system state, \( y_k \in \mathbb{R}^m \) is the system measurement, the process noise \( w(t) \in \mathbb{R}^n \) is continuous-time white noise, and the measurements noise \( v_k \in \mathbb{R}^m \) is discrete-time white noise.

**Assumption 1.** The variables \( w(t) \in \mathbb{R}^n \) and \( v_k \in \mathbb{R}^m \) are sequences of Gaussian random vectors with zero-means and covariance matrices, as follows

\[
\begin{align*}
E[w(t_1)w^T(t_2)] &= Q(t)\delta_{t_1t_2}, \\
E[v_kv_k^T] &= R_k\delta_{kj}, \\
E[w(t)v_k^T] &= 0, \\
k, j \in N; t, t_1, t_2 \in R,
\end{align*}
\]

it is assumed that \( Q \) and \( R \) are positive definite, and that \( w \) and \( v \) are uncorrelated.

**Assumption 2.** The initial state \( x_0^+ \) is unknown and uncorrelated to \( w(t) \) and \( v_k \) that satisfies

\[
\begin{align*}
E[x_0^+] &= x_0, \\
E[(x_0^+ - x_0)(x_0^+ - x_0)^T] &= P_0.
\end{align*}
\]

Then, we consider only the expansion around a nominal \( x \). The second-order Taylor expansion around \( x = \hat{x} \) versus \( f(x, u, t) \) is:

\[
f(x, u, t) = f(\hat{x}, u_0, t) + \left. \frac{\partial f}{\partial x} \right|_{\hat{x}} (x - \hat{x}) + \frac{1}{2} \sum_{i=1}^n \phi_i (x - \hat{x})^T \frac{\partial^2 f_i}{\partial x^2} \left. \right|_{\hat{x}} (x - \hat{x}),
\]

(2)

where \( n \) is the dimension of the state vector, \( f_i \) is the \( i \)th element of \( f(x, u, t) \), and the \( \phi_i \) vector is defined as an \( n \times 1 \) vector with all zeros, except for a one in the \( i \)th element.

The quadratic term in Education (2) can be written as

\[
(x - \hat{x})^T \frac{\partial^2 f_i}{\partial x^2} \left. \right|_{\hat{x}} (x - \hat{x}) = Tr\left[ \frac{\partial^2 f_i}{\partial x^2} \left. \right|_{\hat{x}} (x - \hat{x})(x - \hat{x})^T \right].
\]

(3)

**Assumption 3.** If we replace the value of \( (x - \hat{x})(x - \hat{x})^T \) in the Equation (3) with its expected value, we obtain

\[
(x - \hat{x})^T \frac{\partial^2 f_i}{\partial x^2} \left. \right|_{\hat{x}} (x - \hat{x}) \approx Tr\left[ \frac{\partial^2 f_i}{\partial x^2} \left. \right|_{\hat{x}} p \right],
\]

(4)

where \( P \) is the variance of the estimation error as \( E[(x - \hat{x})(x - \hat{x})^T] \).

Evaluate Equation (2) at \( x = \hat{x} \), and substitute Equations (4) and (2) into (1), the time-update equation of \( \hat{x} \) is obtained as

\[
\hat{x} = f(\hat{x}, u_0, t) + \frac{1}{2} \sum_{i=1}^n Tr\left[ \frac{\partial^2 f_i}{\partial x^2} \left. \right|_{\hat{x}} p \right] + w(t),
\]

(5)
the time-update equation of $P$ remains the same as in the standard hybrid EKF as shown
the following:
\[
\dot{P} = FP + PF^T + LQL^T,
\]
where $F$ and $L$ are the first partial of $x$ and $w$ at $x = \hat{x}$.

**Problem 1.** In this paper, the problem is to find the linear minimum variance estimation of
the process and measurement noise of a class of continuous-discrete systems. Throughout this
paper, we denote $\hat{\dot{v}}(t|t + 1)$ and $\hat{\dot{v}}(k|k + N)$ as a linear function based on measurement sequence
$\{y(\tau)|_{0 \leq \tau \leq t + 1}\}$ and $\{y_i|_{0 \leq i \leq k + N}\}$ that minimize the mean-squared estimation error.

**Remark 1.** Similar to the KF case, the second-order EKF-based process white noise estimator
$\hat{\dot{v}}(t|t + 1)$ is a filter when $l = 0$ and a smoother when $l > 0$. In the same way, the second-order
EKF based measurement white noise estimator $\hat{\dot{v}}(k|k + N)$ is a filter when $N = 0$ and a smoother
when $N > 0$.

3. **Numerical Analysis Results**

In this section, we design the white noise estimator based on the second-order approx-
imation of the hybrid nonlinear system.

Suppose that the filtering update equation for the state estimate is given as
\[
\hat{x}_k^+ = \hat{x}_k^- + K_k[y_k - h(\hat{x}_k^-, t_k)] - \pi_k,
\]
where $K_k$ is the filtering gain, which is chosen to minimize the trace estimation of variance.
Moreover, $\pi_k$ is a correction term, so that the estimate $\hat{x}_k^+$ unbiased.

We define the state estimation errors as follows
\[
\begin{align*}
\{ e_k^+ &= x_k - \hat{x}_k^+ , \\
\epsilon_k^- &= x_k - \hat{x}_k^- .
\end{align*}
\]

We can see from Equations (1) and (7) that
\[
e_k^+ = e_k^- - K_k[h(x_k, t_k) - h(\hat{x}_k^-, t_k)] - K_kv_k + \pi_k.
\]

Now, the second-order Taylor series expansion of $h(x_k, t_k)$ around the nominal point
$\hat{x}_k^-$ is performed to obtain
\[
h(x_k, t_k) = h(\hat{x}_k^-, t_k) + H_k(x_k - \hat{x}_k^-) + \frac{1}{2} \sum_{i=1}^{m} \varphi_i(x_k - \hat{x}_k^-) \frac{\partial^2 h_i}{\partial x^2} \bigg| \hat{x}_k^- (x_k - \hat{x}_k^-),
\]
where $H_k$ is defined as $\frac{\partial h_i}{\partial x} \bigg| \hat{x}_k^-$, $m$ is the dimension of the measurement vector, $h_i$ is the
$i$th element of $h(x_k, t_k)$, and $\varphi_i$ is defined as an $n \times 1$ matrix, whose elements are all zero,
except for a one in the $i$th element. Substituting Equation (9) into (8), then we have the
filtering estimate error $\epsilon_k^+$ as
\[
\epsilon_k^+ = \epsilon_k^- - K_kH_k\epsilon_k^- - \frac{1}{2} K_k \sum_{i=1}^{m} \varphi_i(\epsilon_k^-)^T H_{k,i}(\epsilon_k^-) - K_kv_k + \pi_k,
\]
where $H_{k,i}$ is defined as
\[
H_{k,i} = \frac{\partial^2 h_i}{\partial x^2} \bigg| \hat{x}_k^- .
\]
Taking the expected value of both sides of Equation (10), and assuming that \( E(e_k^-) = 0 \), we can see that, in order to satisfy \( E(e_k^+)^2 = 0 \), we must let

\[
\pi_k = \frac{1}{2} K_k \sum_{i=1}^{m} \varphi_i \text{Tr}[H_k, P_k^-].
\]

Define the filtering variance matrix \( P_k^+ \) as

\[
P_k^+ = E[e_k^+ (e_k^+)^T],
\]

and using the Equation (8), it can be derived by using the following Lemma 1.

\[
P_k^+ = E[e_k^+ (e_k^+)^T]
= E(((I - K_k H_k) e_k^- - \frac{1}{2} K_k \sum_{i=1}^{m} \varphi_i \text{Tr}[H_k, (e_k^-)^T - P_k^-])) - K_k v_k] [\cdot]^T
= (I - K_k H_k) E[e_k^- (e_k^-)^T] (I - K_k H_k) + K_k E[v_k v_k^T] K_k^T
+ \frac{1}{2} K_k E[(\sum_{i=1}^{m} \varphi_i \text{Tr}[H_k, (e_k^-)^T - P_k^-])] [\sum_{i=1}^{m} \varphi_i \text{Tr}[H_k, (e_k^-)^T - P_k^-]] K_k^T
= (I - K_k H_k) P_k^- (I - K_k H_k) + K_k (R_k + \Lambda_k) K_k^T,
\]

where the matrix \( \Lambda_k \) is defined as

\[
\Lambda_k = \frac{1}{4} E[(\sum_{i=1}^{m} \varphi_i H_k (e_k^-)^T - P_k^-)] [\sum_{i=1}^{m} \varphi_i H_k (e_k^-)^T - P_k^-]] K_k^T.
\]

Let us give a useful probability Lemma.

**Lemma 1.** Suppose we have the \( n \)-element random vector \( x \sim N(0, P) \), then

\[
E[x \text{Tr}(Axx^T)] = 0.
\]

The proof is given in Appendix A.

Further, we define a cost function \( J_k \) that minimize as a weighted sum of estimation errors:

\[
J_k = E[(e_k^+)^T e_k^+] = \text{Tr}[P_k^+].
\]

The \( K_k \) that minimizes this cost function can be computed as

\[
K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k + \Lambda_k)^{-1}.
\]

From the projection theorem and Equation (7), we derive the following formulas:

\[
\begin{align*}
\hat{x}_k^+ & = \sum_{i=0}^{k} \{E[e_k^- e_i^T(i)] [E(e(i)e_i^T(i))]^{-1} e(i) - \pi_i\} \\
& = \hat{x}_k^- + K_k [y_k - h(\hat{x}_k^-, t_k)] - \tau_k \\
& = \hat{x}_k^- + K_k e(k) - \tau_k, \\
K_k & = E[e_k^- e(k)] [E(e(k)e(k))]^{-1} \\
& = P_k^- H_k^T (H_k P_k^- H_k^T + R_k + \Lambda_k)^{-1},
\end{align*}
\]

By substituting (17) into (11), we have

\[
P_k^+ = P_k^- - P_k^- H_k^T (H_k P_k^- H_k^T + R_k + \Lambda_k)^{-1} H_k P_k^-.
\]
We rewrite \( \Lambda_k \) as the double summation
\[
\Lambda_k = \frac{1}{4}E\{\sum_{i,j=1}^{m} \phi_i \phi_j^T \text{Tr}[H_{k,i}(e_{k}^-(e_{k}^-)^T - P_k^-)]\text{Tr}[H_{k,j}(e_{k}^- (e_{k}^-)^T - P_k^-)]\},
\]
where the element in the \( i \)th row and \( j \)th column of \( \Lambda_k \) is given by
\[
\Lambda_k(i,j) = \frac{1}{4}E\{\text{Tr}[H_{k,i}(e_{k}^- (e_{k}^-)^T - P_k^-)]\text{Tr}[H_{k,j}(e_{k}^- (e_{k}^-)^T - P_k^-)]\}.
\]

To calculate (19), we introduce the following Lemma.

**Lemma 2.** Suppose we have the \( n \)-element random vector \( x \sim N(0,P) \), then
\[
E[\text{Tr}(Axx^T Bxx^T)] = E[\text{Tr}(Axx^T)\text{Tr}(Bxx^T)] = 2\text{Tr}(APBP) + \text{Tr}(AP)\text{Tr}(BP).
\]

The proof is given in Appendix B.
According to Lemma 2 and Equation (19), we have
\[
\Lambda_k(i,j) = \frac{1}{2}\text{Tr}(H_{k,i}P_k^- H_{k,j}P_k^-).
\]

### 3.1. The Process White Noise Estimator of Hybrid Nonlinear System

**Theorem 1.** Given system (1) with Assumptions 1–3, and Problem 1, the process white noise estimator \( \hat{v}(t|t+1) \) is calculated by
\[
\hat{v}(t|t+1) = \sum_{k=|t+1|}^{t+1} S_1(t,\tau,k)H_k^T R_k^{-1}e_k d\tau
+ \int_{|t+1|}^{t+1} S_1(t,\tau,k)H_k^T R_k^{-1}e_k d\tau
+ \int_{|t+1|}^{t+1} S_1(t,\tau,[l])H_{[l]}^T R_{[l]}^{-1}e_{[l]} d\tau,
\]
where \([t]\) is the integer ceiling function of \( t \), and \( S_1(t,\tau,k), 0 < \tau \leq 1 \) is given as
\[
\left\{\begin{array}{c}
\frac{dS_1(t,\tau,k)}{dt} = S_1(t,\tau,k)(I - K_k H_k)^T \left( \frac{\eta_k}{\eta_k} \right)^T, 0 < \tau \leq l, \\
S_1(t,\tau,k) = 0, \tau = 0.
\end{array}\right.
\]

**Proof of Theorem 1.** Let \( \hat{v}(t|t+1) \) is the projection of \( v(t) \) onto set \( L\{y(\tau)|_{0 \leq \tau \leq t+1}\} \) that minimizes the mean-square error trace \( \{[v(t) - \hat{v}(t|t+1) ][v(t) - \hat{v}(t|t+1)]^T\} \), where \( y(\tau) \) is the measurement. If we consider the discrete-time measurements, we can replace \( y(\tau), k - 1 < \tau \leq k, k = 1,2,\cdots, N \) by \( y_k(\tau) = y_k \). In order to calculate \( \hat{v}(t|t+1) \) using the projection formula, an innovation sequence is introduced and \( y_k(\tau) \) is given as
\[
\epsilon_k(\tau) \triangleq y_k(\tau) - \hat{y}_k(\tau), \quad 0 < k \leq t + 1, k \in N,
\]
where \( \hat{y}_k(t + \tau) \) is the projection of \( y_k(t + \tau) \) onto the linear space \( L\{y_k(s)|_{t \leq s < t+\tau, i \in N}\} \).

Further, in view of Equations (1) and (9), it follows from (23) that
\[
\epsilon_k(\tau) = H_k e^-(\tau) - \frac{1}{2} \sum_{j=1}^{m} \phi_j \text{Tr}[H_{k,j}(e^- (e^-)^T)] v_k.
\]
Note that \( \hat{w}(t|t+1) \) is the projection of \( w(t) \) onto the linear space \( L\{\varepsilon_i(\tau)|_{\tau,j<i+1,i\in N}\} \). by using projection formula, \( \hat{w}(t|t+1) \) is given by
\[
\hat{w}(t|t+1) = \sum_{k=[t+1]}^{[t]} f_k \langle w(t), \epsilon_k(\tau) \rangle E(\epsilon_k e_k^T)^{-1} \epsilon_k d\tau
\]
\[
+ \int_{[t+1]}^{[t]} (w(t), \epsilon_{[t+1]}(\tau)) E(\epsilon_{[t+1]} e_{[t+1]}^T)^{-1} \epsilon_{[t+1]} d\tau
\]
\[
+ \int_{[t+1]}^{[t+1]} (w(t), \epsilon_{[t]}(\tau)) E(\epsilon_{[t]} e_{[t]}^T)^{-1} \epsilon_{[t]} d\tau.
\]

Let us think about the first part to the right of the equal sign of Equation (25), using Equation (24), we obtain
\[
\sum_{k=[t+1]}^{[t]} f_k \langle w(t), \epsilon_k(\tau) \rangle E(\epsilon_k e_k^T)^{-1} \epsilon_k d\tau
\]
\[
= \sum_{k=1}^{[t+1]} f_{k-1} \langle w(t), H_{k-1} e^{-t}(\tau) \rangle - \frac{1}{2} \sum_{j=1}^{m} \theta_j \text{Tr}[H_{k-1,j} (e^{-t}(\tau))(e^{-t}(\tau))^T] + v_{k-1}
\]
\[
\times R^{-1} e_{k-1} d\tau
\]
\[
= \sum_{k=1}^{[t+1]} f_{k-1} \langle w(t), e^{-t}(\tau) \rangle H_{k-1}^T R_{k-1}^{-1} e_{k-1} d\tau
\]
\[
= \sum_{k=1}^{[t+1]} f_{k-1} S_1(t, \tau, k) H_{k-1}^T R_{k-1}^{-1} e_{k-1} d\tau,
\]
where
\[
S_1(t, \tau, k) = \langle w(t), e^{-t}(\tau) \rangle,
\]
by considering (2), (5), and (10), it follows that
\[
\frac{d\hat{v}(\tau)}{d\tau} = \frac{\partial f}{\partial x} e^\tau(\tau) + w(\tau)
\]
\[
= \frac{\partial f}{\partial x} \left\{ e^{-\tau} - K_k H_k e^{-\tau} - \frac{1}{2} K_k \sum_{i=1}^{m} \varphi_i (e^{-\tau})^T H_{k,i} (e^{-\tau}) - K_k \varkappa_k + \pi_k \right\} + w(\tau)
\]
\[
= \frac{\partial f}{\partial x} \left\{ (I - K_k H_k) e^{-\tau} \right\}
\]
\[
- \frac{\partial f}{\partial x} \left\{ \frac{1}{2} K_k \sum_{i=1}^{m} \varphi_i (e^{-\tau})^T H_{k,i} (e^{-\tau}) + K_k \varkappa_k - \pi_k \right\} + w(\tau).
\]

Noting \( w(t) \) is uncorrelated with \( \varkappa_k \), the term \( \langle w(t), \sum_{i=1}^{m} \varphi_i (e^{-\tau})^T H_{k,i} (e^{-\tau}) \rangle = 0 \) can be derived from Equation (13) in Lemma 1 and taking differential on both sides of (22) with \( \tau \), Equation (22) follows directly from the definition of \( S_1(t, \tau, k) \) and (27).

That is all the proof. \( \square \)

### 3.2. The Measurement White Noise Estimator of Hybrid Nonlinear System

**Theorem 2.** Given system (1) with Assumption 1, Assumption 2, Assumption 3, and Problem 1, the measurement white noise estimator \( \hat{v}(k|N) \) is given by the following formula.

\[
\hat{v}(k|N) = \hat{v}(k|k) + \sum_{i=k+1}^{k+N} S_2(k, i-1) F_{i-1}^T H_i^T (H_i P_i^{-1} H_i^T + R_i + \Lambda_i)^{-1} e_i,
\]

where \( \Lambda_i \) is the difference of white noise estimation by EKF, and \( S_2(k, i) \) is given as

\[
\begin{cases}
S_2(k, i) = \langle \varkappa, e_i^+ \rangle, & i > k, \ i \in N, \\
S_2(k, k) = -R_k K_k^T.
\end{cases}
\]

**Proof of Theorem 2.** \( \hat{v}(k|N) \) is the projection of \( \hat{v}(k) \) onto the linear space \( L\{y_i|_{0 \leq i \leq k+N}\} \), where \( y_i \) is the measurement. In order to calculate \( \hat{v}(k|N) \) using the projection formula, let us calculate \( \langle \varkappa, e_{k+N} \rangle \) first.

\[
\langle \varkappa, e_{k+N} \rangle = \langle \varkappa, H_{k+N} e_{k+N} \rangle - \frac{1}{2} K_{k+N} \sum_{j=1}^{m} \varphi_j \text{Tr}[H_{k+N,j} (e_{k+N}(e_{k+N}^T))^T] + v_{k+N},
\]
since $v_k$ is uncorrelated with the innovation $v_{k+N}$ for $N \geq 1$, and combining Equation (13) in Lemma 1, we obtain
\[
\langle v_k, e_{k+N}\rangle = \langle v_k, H_{k+N}e_{k+N}\rangle = \langle v_k, F_{k+N-1}e_{k+N-1} + w(k + N - 1)\rangle H_{k+N}^T = \langle v_k, F_{k+N-1}^T F_{k+N-1}^T H_{k+N}^T \rangle = S_2(k, k + N - 1) F_{k+N-1}^T H_{k+N}^T.
\]

Note that $\hat{o}(k|k + N)$ is the projection of $v(k)$ onto the linear space $L\{e_{k+N}|0 \leq N, \}$ by using projection formula and Equation (30), $\hat{o}(k|k + N)$ is given by
\[
\hat{o}(k|k + N) = \hat{o}(k|k) + \sum_{i=0}^{k+N} E[\hat{o}(k|k+i)|E(\epsilon_i|e_{k+1})]^{-1} \epsilon_i
\]
\[
= \hat{o}(k|k) + \sum_{i=0}^{k+N} \langle v_k, e_i \rangle (H_i P_i^{-1} H_i^T P_i + R_i + \Lambda_i)^{-1} e_i
\]
\[
= \hat{o}(k|k) + \sum_{i=0}^{k+N} S_2(k, i - 1) F_i^T (H_i P_i^{-1} H_i^T + R_i + \Lambda_i)^{-1} e_i,
\]
where
\[
\hat{o}(k|k) = \sum_{i=0}^{k} E[\hat{o}(k|k+i)|E(\epsilon_i|e_{k+1})]^{-1} \epsilon_i
\]
\[
= E[\hat{o}(k|k)|E(\epsilon_k|e_{k+1})]^{-1} \epsilon_k
\]
\[
= R_k (H_k P_k^{-1} H_k^T + R_k + \Lambda_k)^{-1} \epsilon_k.
\]

Further, the recurrence formula for $S_2(k, k + N)$ is derived
\[
S_2(k, k + N) = \langle v_k, e_{k+N}^+ \rangle
\]
\[
= \langle v_k, e_{k+N}^- - K_{k+N} H_{k+N} e_{k+N}^- - \frac{1}{2} K_{k+N} \sum_{j=1}^{m} \varphi_j (e_{k+N}^-)^T H_{k+N,j} (e_{k+N}^-) - K_{k+N} v_{k+N} + \pi_{k+N} \rangle
\]
\[
= \langle v_k, e_{k+N}^- \rangle (I - K_{k+N} H_{k+N})^T
\]
\[
= \langle v_k, F_{k+N-1} e_{k+N-1}^- \rangle (I - K_{k+N} H_{k+N})^T
\]
\[
= S_2(k, k + N - 1) [(I - K_{k+N} H_{k+N}) F_{k+N-1}]^T
\]
\[
= S_2(k, k) \prod_{i=k+1}^{k+N} [(I - K_{k+i} H_{k+i}) F_{k+i-1}]^T.
\]

That is all the proof. $\Box$

According to the theoretical derivation of Theorem 1 and Theorem 2, we can use a flowchart to represent the white noise estimation algorithm in the following Figure 1, where $k, l, N, and tf$ represent time, smoothing step of process noise, smoothing step of measurement noise, and end time, respectively.
Further, the recurrence formula for $S_{6}(k, k + N)$ is derived

$$S_{6}(k, k + N) = \langle v_{7}, e_{7} \rangle = \langle v_{7}, e_{7} - K_{7} H_{7} e_{7} - \frac{1}{2} K_{7} \sum \phi(\gamma) \theta H_{7}, \phi(\gamma) \phi_{7} \rangle = \langle v_{7}, e_{7} \rangle (I - K_{7} H_{7}) \theta$$

$$= \langle v_{7}, F_{7} e_{7} \rangle (I - K_{7} H_{7}) \theta = S_{6}(k, k) \prod (I - K_{j} H_{j}) F_{j} \theta. \tag{33}$$

That is all the proof. □

According to the theoretical derivation of Theorem 1 and Theorem 2, we can use a flowchart to represent the white noise estimation algorithm in the following Figure 1, where $k, l, N,$ and $t_f$ represent time, smoothing step of process noise, smoothing step of measurement noise, and end time, respectively.

**Figure 1.** Flow chart of second-order EKF white noise estimation algorithm.

**Remark 2.** The algorithm flow in Figure 1 can be described as follows:

**Step 1:** Set $\hat{x}_0^+ = x_0$, $P_0^+ = P_0$, $k = 1$, $l = l_0$, $N = N_0$ and $t_f = t_{f0}$.

**Step 2:** If $k \leq t_f$, go to Step 3; If $k > t_f$, exit.

**Step 3:** Calculate $F$, $L$ by Taylor expansion $f(x, u, k)$, $w(k)$.

**Step 4:** Calculate partial derivatives $H_k, H_k,i, A_k$ after get the measurement.

**Step 5:** Integrate the state estimate and its covariance from time $(k-1)^+$ to time $k^-$ as follows:

$$\dot{x} = f(\hat{x}, u, t) + w(t)$$

$$\dot{P} = AP + P A^T + L Q L^T$$

**Step 6:** Calculate $K_k$ by (17), then compute $\hat{x}^+_k$, $P^+_k$ using (16), (18).

**Step 7:** If $k > l$, calculate $\hat{x}_k^-$ by (21).

**Step 8:** If $k > N$, calculate $\hat{x}(k|k + N)$ by (28).

**Step 9:** Set $k = k + 1$, then go to Step 2.
4. Numerical Simulation

In this section, the first-order EKF and second-order EKF white noise estimators are compared by a numerical simulation; the advantages of the proposed algorithm is verified. Suppose the nonlinear continuous-discrete system (1) is as follows:

\[
\begin{align*}
    \dot{x} &= \cos(x(t)) + w(t), \\
    y_k &= x_k^3 + v_k,
\end{align*}
\]

where the process noise vector \(w(t)\) and measurement noise \(v_k\) are uncorrelated zero-mean Gaussian white noises with variances \(Q = 0.0019, \) and \(R = 0.0088.\)

In this numerical simulation, we compare our algorithm with first-order EKF white noise estimators by setting that \(x(0) = 0.92, P(0) = 0.08\) and the sampling period is 0.01s. We take 50 sampling periods to draw Figures 2–5. According to Theorem 1, the filtering algorithm cannot achieve the process noise estimate. Therefore, in Figure 2, we use one-step smoothing algorithm to estimate the process noise. Figure 2 shows the true value of process white noise \(w(t)\) and its one-step smoothing of two methods. Figure 3 shows the true value of process white noise \(w(t)\) and its three-step smoothing of two methods. Figure 4 shows the true value of measurement white noise \(v_k\) and its filtering of two methods. Figure 5 shows the true value of measurement white noise \(v_k\) and its three-step smoothing of two methods. From Figure 2 to Figure 5, it is shown that the white noise estimators based on the first-order EKF fluctuates greatly, while the estimator based on second-order EKF have higher estimation accuracy. In other words, the second-order EKF white noise estimators are better than the first-order EKF white noise estimators. Moreover, it can be seen that the process noise estimation is affected by the discrete observation equation, and loses some information. Therefore, the estimation effect is not as good as the measurement noise estimation. The measurement noise estimation will gradually approach the real curve with the update of time. Moreover, we take 50 sampling periods to calculate the error loss value of two different methods. Taking the mean square error (MSE) as the evaluation index, the loss value is shown in Table 1.

![Figure 2](image_url)  
Figure 2. True value of process noise, one-step white noise smoothing based on first-order EKF and second-order EKF.
Figure 2. True value of process noise, one-step white noise smoothing based on first-order EKF and second-order EKF.

Figure 3. True value of the process noise, three-step white noise smoothing based on first-order EKF and second-order EKF.

Figure 4. True value of measurement noise, white noise filtering based on first-order EKF and second-order EKF.

Figure 5. True value of measurement noise, three-step white noise smoothing based on the first-order EKF and second-order EKF.
Table 1. MSE of noise estimation.

| MSE of Noise Estimated                        | First-Order EKF | Second-Order EKF |
|-----------------------------------------------|----------------|-----------------|
| One-step smoothing of process noise           | 0.01011        | 0.00733         |
| Three-step smoothing of process noise         | 0.01079        | 0.00884         |
| Measurement noise filtering                   | 0.01932        | 0.01540         |
| Three-step smoothing of measurement noise     | 0.02239        | 0.01642         |

5. Discussion and Conclusions

For the nonlinear hybrid systems with continuous-time dynamic and discrete-time measurements, white noise estimators for process and measurement noises were designed in this paper. Firstly, the nonlinear second-order EKF was derived by using the projection formula after the second-order Taylor expansion approximation and the variance approximation. Moreover, the continuous process noise estimation was processed in segments, and the white noise estimation of the nonlinear hybrid system was performed by combining the innovation of the previous observation point with the second-order EKF theory. The observed white noise estimator was calculated by a difference equation solution of the Riccati equation. Finally, an example was presented to verify the effectiveness of the proposed algorithm.

This paper enriches the second-order EKF theory; the optimal white noise estimation theory is extended to the nonlinear hybrid systems. In the future, we will investigate the color noise estimation of nonlinear hybrid systems.

Author Contributions: Data curation, Y.L.; investigation, Y.L.; methodology, Y.L. and H.Z.; project administration, Y.S.; Software, H.M.; supervision, H.Z.; validation, Y.L.; writing—original draft, Y.L.; writing—review and editing, H.Z., Y.S. and H.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Shandong Province Higher Educational Science and Technology Program (grant no. J18KB157), and the National Natural Science Foundation of P. R. China (grant nos. 61903223, 62122041, 62173211, 62001261, and 61403061).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data presented for reverification of the results are given in the article and the authors have no other data. Software tools for the solution are standard and their use should be known to the scientific community (MATLAB).

Conflicts of Interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

Appendix A

Proof of Lemma 1. Let us look at the general equation:

\[
[(Ax + a)(Bx + b)^T(Cx + c)] = AR_{xx}B^T(Cm_x + c) + AR_{xx}C^T(Bm_x + b) \\
+ Tr(BR_{xx}C^T)(Am_x + a) \\
+ (Am_x + a)(Bm_x + b)^T(Cm_x + c),
\]

(A1)

where \(m_x\) is the expectation of \(x\), and \(R_{xx}\) is the covariance of \(x\).

Equation (8) can be written as:

\[
E[xTr(Axx^T)] = E[xTr(x^TAx)] = E[xx^TAx].
\]

(A2)
According to the random vector $x \sim N(0, P)$ and Equation (A2), we can see that
\[
E[xx^T Ax] = R_{xx}(A + I)m_x + (Tr(R_{xx} A^T) + m_x m_x^T A)m_x = 0. \tag{A3}
\]
This completes the proof. □

Appendix B

Proof of Lemma 2. The first equation of (20) can be obtained directly by the matrix operation, then
\[
E[Tr(Axx^T Bxx^T)] = E[Tr(Axx^T)Tr(Bxx^T)] \\
= E[Tr(x^TAx)Tr(x^TBx)] \\
= E(x^TAxx^TBx). \tag{A4}
\]
Moreover, we have
\[
cov(x^TAx, x^TBx) = E[(x^TAx - E(x^TAx))(x^TBx - E(x^TBx))] \\
= E(x^TAxx^TBx) - E(x^TAx)E(x^TBx), \tag{A5}
\]
combining (A4) and (A5), $E(x^TAxx^TBx)$ can be given by
\[
E(x^TAxx^TBx) = cov(x^TAx, x^TBx) + E(x^TAx)E(x^TBx). \tag{A6}
\]
Furthermore, the associative law of the addition of variance is as follows
\[
Var[x^T(A + B)x] = Var(x^TAx + x^TBx) \\
= Var(x^TAx) + Var(x^TBx) + 2cov(x^TAx, x^TBx), \tag{A7}
\]
by simple deformation of the above equation, we obtain
\[
\text{cov}(x^TAx, x^TBx) = \frac{1}{2} \{\text{Var}[x^T(A + B)x] - \text{Var}(x^TAx) - \text{Var}(x^TBx)\}. \tag{A8}
\]
For $y \sim N(0, I)$ we have
\[
\text{Var}[y^T Ay] = 2Tr(A^2), \tag{A9}
\]
when we consider the $x \sim N(0, P)$, there exists a nonsingular matrix $C$, such that $P = CC^T$. We can see that $y = C^{-1}x \sim N(0, I)$.

Then we can obtain
\[
\text{Var}(x^TAx) = \text{Var}[(Cy)^T A(Cy)] = \text{Var}[y^T (C^T AC)y] \\
= 2Tr(C^T ACC^T AC) = 2Tr(C^T APAC) \\
= 2Tr(ACC^T AP) = 2Tr(APAP), \tag{A10}
\]
\[
\text{Var}(x^T(A + B)x) = 2Tr(APAP + 2APBP + BPBP). \tag{A11}
\]
From (A10) and (A11), we can derive the following formulas:
\[
\text{cov}(x^TAx, x^TBx) = 2Tr(APBP), \tag{A12}
\]
and
\[
E(x^TAx) = E[Tr(x^TAx)] = E[Tr(Axx^T)] \\
= Tr[E(Axx^T)] \\
= Tr[AE(xx^T)] = Tr(AP), \tag{A13}
\]
using Equation (A12), (A13) with (A6), we can see that

\[
E(x^TAx^TBx) = \text{cov}(x^TAx, x^TBx) + E(x^TAx)E(x^TBx) = 2\text{Tr}(APBP) + \text{Tr}(AP)\text{Tr}(BP).
\]

(A14)

At this point, the Lemma 2 is proved. □

References

1. Liu, K.; Shang, Y.; Ouyang, Q.; Widanage, W.D. A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery. IEEE Trans. Ind. Electron. 2020, 68, 3170–3180. [CrossRef]
2. Shang, Y.; Lu, G.; Kang, Y.; Zhou, Z.; Duan, B.; Zhang, C. A multi-fault diagnosis method based on modified Sample Entropy for Lithium-ion battery strings. J. Power Sources 2020, 446, 227–275. [CrossRef]
3. Kania, D.; Borkowski, J. Estimation methods of multifrequency signals with noise and harmonics for PV systems with a DSP processor. In Proceedings of the 40th International Conference on Telecommunications and Signal Processing (TSP), Barcelona, Spain, 5–7 July 2017; pp. 524–527.
4. Mendel, J. White-noise estimators for seismic data processing in oil exploration. IEEE Trans. Autom. Control. 1977, 22, 694–706. [CrossRef]
5. Dong, L.; Zhou, J.; Dai, T. Color Image Noise Covariance Estimation with Cross-Channel Image Noise Modeling. In Proceedings of the 2018 IEEE International Conference on Multimedia and Expo (ICME), San Diego, CA, USA, 23–27 July 2018; pp. 1–6.
6. Dai, G.Z.; Mendel, J. A straightforward and unified approach to the derivation of minimum-variance deconvolution algorithms. IEEE Trans. Autom. Control. 1986, 31, 80–89.
7. Sun, X.J.; Gao, Y.; Deng, Z.; Li, C.; Wang, J.W. Multi-model information fusion Kalman filtering and white noise deconvolution. Inf. Fusion 2010, 11, 163–173. [CrossRef]
8. Deng, Z.L.; Zhang, H.S.; Liu, S.J.; Zhou, L. Optimal and self-tuning white noise estimators with applications to deconvolution and filtering problems. Automatica 1996, 32, 199–216. [CrossRef]
9. Zhao, H.; Zhang, C. Non-Gaussian noise quadratic estimation for linear discrete-time time-varying systems. Neurocomputing 2016, 174, 921–927. [CrossRef]
10. Zhao, H.; Li, Z.; Li, B.; Li, T. A study on input noise second-order filtering and smoothing of linear stochastic discrete systems with packet dropouts. Adv. Differ. Equ. 2020, 2020, 1–14. [CrossRef]
11. Liu, W.; Deng, Z. Robust white noise deconvolution estimators for systems with uncertain variance linearly correlated white noises and missing measurements. In Proceedings of the Chinese Control Conference, Dalian, China, 26–28 July 2017; pp. 5070–5077.
12. Kailath, T. An innovations approach to least-squares estimation–Part I: Linear filtering in additive white noise. IEEE Trans. Autom. Control. 1968, 13, 646–655. [CrossRef]
13. Bellantoni, J.F.; Dodge, K.W. A square root formulation of the Kalman- Schmidt filter. AIAA J. 1967, 5, 1309–1314. [CrossRef]
14. Alessandri, A.; Baglietto, M.; Battistelli, G. Luenberger Observers for Switching Discrete-Time Linear Systems. In Proceedings of the IEEE Conference on Decision and Control and the European Control Conference, Seville, Spain, 15 December 2005.
15. Ibrir, S. Hybrid observer design for a class of nonlinear systems. In Proceedings of the Asian Control Conference, Hong Kong, China, 27–29 August 2009; pp. 142–147.
16. Fang, Y.; Zhang, Q.; Zhang, H.; Xu, W.; Wang, L.; Shen, X.; Yun, F.; Cui, Y.; Wang, L.; Zhang, X. State-of-charge estimation technique for lithium-ion batteries by means of second-order extended Kalman filter and equivalent circuit model: Great temperature robustness state-of-charge estimation. IET Power Electron. 2021, 14, 1515–1528. [CrossRef]
17. Razvan, S.; Jafari, R.; Vargas-Jarillo, C.; Gregov, A.; Forooshani, M. Leakage Detection in Pipeline Based on Second Order Extended Kalman Filter Observer. IFAC-Pap. 2019, 52, 116–121. [CrossRef]
18. Simon, D. Optimal State Estimation; John Wiley & Sons Inc.: Hoboken, NJ, USA, 2006; pp. 413–417.
19. de la Sen, M.; Luo, N. Design of linear observers for a class of linear hybrid systems. In Proceedings of the 38th IEEE Conference on Decision and Control, Phoenix, AZ, USA, 7–10 December 1999; pp. 4903–4908.
20. Shin, Y.; Kim, D.Y.; Shevlyakov, G.; Kim, K. A Suboptimal Filter for Continuous-Discrete Linear Systems with Parametric Uncertainties. In Proceedings of the TENCON 2006-2006 IEEE Region 10 Conference, Hong Kong, China, 14–17 November 2006; pp. 1–4.
21. Kulikov, G.Y.; Kulikova, M.V. Square-root high-degree cubature Kalman filters for state estimation in nonlinear continuous-discrete stochastic systems. Eur. J. Control. 2021, 39, 58–68. [CrossRef]
22. Athans, M.; Wishner, R.; Bertolini, A. Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements. IEEE Trans. Autom. Control. 1968, 13, 504–514. [CrossRef]
23. Huang, M.; Wang, C.; Zhao, J. State of Charge Estimation of Lithium-Ion Battery Based on Second-order Extended Kalman Filter. In Proceedings of the Advanced Information Technology. Electronic and Automation Control Conference, Chengdu, China, 20–22 December 2019; pp. 335–338.

24. Wang, Y.; Zhang, H. Accurate Smoothing for Continuous-Discrete Nonlinear Systems with Non-Gaussian Noise. *IEEE Signal Process. Lett.* 2019, 26, 465–469. [CrossRef]