

Article

A Novel Perspective of the Kalman Filter from the Rényi Entropy

Yarong Luo ¹, Chi Guo ^{1,2,*}, Shengyong You ¹ and Jingnan Liu ^{1,2}

¹ Global Navigation Satellite System Research Center, Wuhan University, Wuhan 430079, China; yarongluo@whu.edu.cn (Y.L.); shengyongyou@whu.edu.cn (S.Y.); jnliu@whu.edu.cn (J.L.)

² Artificial Intelligence Institute, Wuhan University, Wuhan 430079, China

* Correspondence: guochi@whu.edu.cn

Received: 21 July 2020; Accepted: 31 August 2020; Published: 3 September 2020



Abstract: Rényi entropy as a generalization of the Shannon entropy allows for different averaging of probabilities of a control parameter α . This paper gives a new perspective of the Kalman filter from the Rényi entropy. Firstly, the Rényi entropy is employed to measure the uncertainty of the multivariate Gaussian probability density function. Then, we calculate the temporal derivative of the Rényi entropy of the Kalman filter's mean square error matrix, which will be minimized to obtain the Kalman filter's gain. Moreover, the continuous Kalman filter approaches a steady state when the temporal derivative of the Rényi entropy is equal to zero, which means that the Rényi entropy will keep stable. As the temporal derivative of the Rényi entropy is independent of parameter α and is the same as the temporal derivative of the Shannon entropy, the result is the same as for Shannon entropy. Finally, an example of an experiment of falling body tracking by radar using an unscented Kalman filter (UKF) in noisy conditions and a loosely coupled navigation experiment are performed to demonstrate the effectiveness of the conclusion.

Keywords: Rényi entropy; discrete Kalman filter; continuous Kalman filter; algebraic Riccati equation; nonlinear differential Riccati equation

1. Introduction

In the late 1940s, Shannon introduced a logarithmic measure of information [1] and a theory that included information entropy (the literature shows that it is related to Boltzmann entropy in statistical mechanics). The more stochastic and unpredictable a variable is, the larger its entropy is. As a measure of information, entropy has been used in various fields, such as information theory, signal processing, information-theoretic learning [2,3], etc. As a generalization of the Shannon entropy, Rényi entropy, named after Alfréd Rényi [4], allows for different averaging of probabilities through a control parameter α , and is usually used to quantify the diversity, uncertainty, or randomness of random variables. Liang [5] presented the evolutionary entropy equations and the uncertainty estimation for Shannon entropy and relative entropy, which is also called Kullback–Leibler divergence [6], within the framework of dynamical systems. However, higher-order Rényi entropy has some better properties than Shannon entropy by setting the control parameter α in most cases.

The Kalman filter [7] and its variants have been widely used in navigation, control, tracking, etc. Many works focus on combining different entropy and entropy-like quantities with the original Kalman filter to improve the performance. When the state space equation is nonlinear, Rényi entropy can be used to measure the nonlinearity [8,9]. Shannon entropy was used to estimate the weight of each particle from the weights of different measurement models for the fusion algorithm in [10]. Quadratic Rényi entropy [11] of innovation has been used as a minimum entropy criterion under a nonlinear and non-Gaussian circumstance [12] in unscented Kalman filter (UKF) [13] and finite

mixtures [14]. A generalized density evolution equation [15] and polynomial-based non-linear compensation [16] were used to improve the minimum entropy filtering [17]. Relative entropy has been used to measure the similarity between the probabilistic density functions during the recursive processes of the nonlinear filter [18,19]. As for the nonlinear measurement equation with additive Gaussian noise, relative entropy can be deduced to measure the nonlinearity of the measurement [20], and can also be used to measure the approximation error of the i -th measurement element in the partitioned update Kalman filter [21]. When the state variables and the measurement variables do not belong to strict Gaussian distribution, such as in the seamless indoor/outdoor multi-source fusion positioning problem [22], the estimation error can be measured by the relative entropy. Relative entropy can also be used to calculate the number of particles in the unscented particle filter for mobile robot self-localization [23] and to calculate the sample window size in the cubature Kalman filter (CKF) [24] for attitude estimation [25]. Moreover, it has been verified that the original Kalman filter can be derived by maximizing the relative entropy [26]. Meanwhile, the robust maximum correntropy criterion has been adopted as the optimal criterion to derive the maximum correntropy Kalman filter [27,28]. However, there has been no work on the direct connections between the Rényi entropy and the Kalman filter theory until now.

In this paper, we propose a new perspective of the Kalman filter from the Rényi entropy for the first time, which bridges the gap between the Kalman filter and the Rényi entropy. We calculate the temporal derivative of the Rényi entropy for the Kalman filter mean square error matrix, which provides the optimal recursive solution mathematically and will be minimized to obtain the Kalman filter gain. Moreover, from the physical point of view, the continuous Kalman filter approaches a steady state when the temporal derivative of the Rényi entropy is equal to zero, which also means that the Rényi entropy will keep stable. A numerical experiment of falling body tracking in noisy conditions with radar using the UKF and a practical experiment of loosely-coupled integration are provided to demonstrate the effectiveness of the above conclusion.

The structure of this paper is as follows. In Section II, the definitions and properties of Shannon entropy and Rényi entropy are presented. In Section III, the Kalman filter is derived from the perspective of minimizing the temporal derivative of Rényi entropy, and the connection between the Rényi entropy and the algebraic Riccati equation is explained. In Section IV, experimental results and analysis are given by the simulation of the UKF and the real integrated navigation data. We finally conclude this paper and provide an outlook for future work in Section V.

2. The Connection between the Kalman Filter and the Temporal Derivative of the Rényi Entropy

2.1. Rényi Entropy

To calculate the Rényi entropy of the continuous probability density function (PDF), it is necessary to extend the definition of the Rényi entropy to the continuous form. The Rényi entropy of order α for a continuous random variable with a multivariate Gaussian PDF $p(x)$ is defined [4] and calculated [9] as:

$$H_R^\alpha(x) = \frac{1}{1-\alpha} \ln \int_S p^\alpha(x) dx = \frac{N}{2} \ln(2\pi\alpha^{\frac{1}{\alpha-1}}) + \frac{1}{2} \ln(\det \Sigma), \quad (1)$$

where $\alpha > 0, \alpha \neq 1$, and α is a parameter providing a family of entropy functions. N is the dimension of the random variable x . S is the support. Σ is the covariance matrix of $p(x)$.

It is straightforward to show that the temporal derivative of the Rényi entropy is given by [9]:

$$\dot{H}_R^{(\alpha)}(x) = \frac{1}{2} \text{Tr}\{\Sigma^{-1}\dot{\Sigma}\}, \quad (2)$$

where $\dot{\Sigma}$ is the temporal derivative of the covariance matrix and $\text{Tr}(\cdot)$ is the trace operator.

It is easy to get the Shannon entropy for the multivariate Gaussian PDF by taking the limitation of Equation (1) as α approaches 1. This entropy is given as $H(x) = \frac{N}{2} \ln(2\pi e) + \frac{1}{2} \ln(\det \Sigma)$, and the

temporal derivative of the Shannon entropy is given as $\dot{H}(x) = \frac{1}{2} \text{Tr}\{\Sigma^{-1}\dot{\Sigma}\}$. It is obvious the temporal of the Shannon entropy is the same as the temporal of the Rényi entropy. Therefore, we will see later that the conclusion can also be derived from the temporal derivative of the Shannon entropy. However, the Rényi entropy for the multivariate Gaussian PDF instead of the temporal derivative of the Rényi entropy will be used by adjusting the free parameter α for different uncertainty measurements in most cases, as the filtering problem has to account for the nonlinearity and the non-Gaussian noise; we adopt the Rényi entropy as the measurement for uncertainty.

2.2. Kalman Filter

Given the continuous-time linear system [29]:

$$\dot{X}(t) = F(t)X(t) + G(t)w(t) \quad (3)$$

$$Z(t) = H(t)X(t) + v(t), \quad (4)$$

where $X(t)$ is the state vector; $F(t)$ is the state transition matrix; $G(t)$ is the system noise driving matrix; $Z(t)$ is the measurement vector; $H(t)$ is the measurement matrix; and $w(t)$ and $v(t)$ are independent white Gaussian noise with zero mean value; their covariance matrices are $Q(t)$ and $R(t)$, respectively:

$$\mathbb{E}[w(t)] = 0, \mathbb{E}[w(t)w^T(\tau)] = Q(t)\delta(t - \tau) \quad (5)$$

$$\mathbb{E}[v(t)] = 0, \mathbb{E}[v(t)v^T(\tau)] = R(t)\delta(t - \tau) \quad (6)$$

$$\mathbb{E}[w(t)v^T(\tau)] = 0, \quad (7)$$

where $\delta(t)$ is the Dirac impulse function, $Q(t)$ is a symmetric non-negative definite matrix, and $R(t)$ is a symmetric positive matrix.

The continuous Kalman filter can be deduced by taking the limit of the discrete Kalman filter. The discrete-time state-space model is arranged as follows [29]:

$$X_k = \Phi_{k|k-1}X_{k-1} + \Gamma_{k|k-1}W_{k-1} \quad (8)$$

$$Z_k = H_kX_k + V_k \quad (9)$$

where X_k is an n -dimensional state vector; Z_k is an m -dimensional measurement vector; $\Phi_{k|k-1}$, $\Gamma_{k|k-1}$, and H_k are the known system structure parameters, which are called the $n \times n$ dimensional one-step state update matrix, the $n \times l$ dimensional system noise distribution matrix, and the $m \times n$ dimensional measurement matrix, respectively; W_{k-1} is the l -dimensional system noise vector, and V_k is the m -dimensional measurement noise vector. Both of them are Gaussian noise vector sequences with zero mean value, and are independent of each other:

$$\mathbb{E}[W_k] = 0, \mathbb{E}[W_kW_j^T] = Q_k\delta_{kj} \quad (10)$$

$$\mathbb{E}[V_k] = 0, \mathbb{E}[V_kV_j^T] = R_k\delta_{kj} \quad (11)$$

$$\mathbb{E}[W_kV_j^T] = 0. \quad (12)$$

The above equation is the basic assumption for the noise requirement in the Kalman filtering state space model, where Q_k is a symmetric non-negative definite matrix, and R_k is a symmetric positive definite matrix. δ_{kj} is the Kronecker δ function.

The covariance parameters Q_k and R_k play roles similar to those of Q and R in the continuous filter, but they do not have the same numerical values. Next, the relationship between the corresponding continuous and discrete filter parameters will be derived.

To achieve the transformation from the continuous form to the discrete form, the relations between Q and R and the corresponding Q_k and R_k for a small step size T_s are needed. According to the linear system theory, the relation between Q and Q_k from Equation (3) to Equation (8) is as follows:

$$\Phi_{k|k-1} = \Phi(t_k, t_{k-1}) \approx e^{\int_{t_{k-1}}^{t_k} F(\tau) d\tau} \quad (13)$$

$$\Gamma_{k|k-1} W_{k-1} = \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau) G(\tau) w(\tau) d\tau. \quad (14)$$

Denote the discrete-time interval as $T_s = t_k - t_{k-1}$, when $F(t)$ does not change too dramatically within the shorter integral interval $[t_{k-1}, t_k]$. Take the Taylor expansion of $e^{F(t_{k-1})T_s}$ with respect to $F(t_{k-1})T_s$ and set $F(t_{k-1})T_s \ll 1$, so the higher-order terms are negligible and the one-step transition matrix, Equation (13), can be approximated as:

$$\Phi_{k|k-1} \approx e^{F(t_{k-1})T_s} = I + F(t_{k-1})T_s + F^2(t_{k-1})\frac{T_s^2}{2!} + F^3(t_{k-1})\frac{T_s^3}{3!} + \dots \approx I + F(t_{k-1})T_s. \quad (15)$$

Equation (14) shows that $\Gamma_{k|k-1} W_{k-1}$ is the linear transform of the Gaussian white noise $w(\tau)$; the result remains the normal distribution random vector. Therefore, the first- and second-order statistical characteristics can be used to describe and be equivalent to $\Gamma_{k|k-1} W_{k-1}$. Referring to Equation (5), the mean of $\Gamma_{k|k-1} W_{k-1}$ is given as follows:

$$\mathbb{E}[\Gamma_{k|k-1} W_{k-1}] = \mathbb{E}\left[\int_{t_{k-1}}^{t_k} \Phi(t_k, \tau) G(\tau) w(\tau) d\tau\right] = \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau) G(\tau) \mathbb{E}[w(\tau)] d\tau = 0. \quad (16)$$

For the second-order statistical characteristics, when $k \neq j$, the time parameter between the noise $w(\tau_k)$ and $w(\tau_j)$ is independent, so $\Gamma_{k|k-1} W_{k-1}$ and $\Gamma_{j|j-1} W_{j-1}$ are uncorrelated:

$$\mathbb{E}[(\Gamma_{k|k-1} W_{k-1})(\Gamma_{j|j-1} W_{j-1})^T] = 0 \quad (k \neq j). \quad (17)$$

When $k = j$, thus

$$\begin{aligned} \mathbb{E}[(\Gamma_{k|k-1} W_{k-1})(\Gamma_{k|k-1} W_{k-1})^T] &= \mathbb{E}\left\{\left[\int_{t_{k-1}}^{t_k} \Phi(t_k, \tau) G(\tau) w(\tau) d\tau\right]\left[\int_{t_{k-1}}^{t_k} \Phi(t_k, s) G(s) w(s) ds\right]^T\right\} \\ &= \mathbb{E}\left\{\int_{t_{k-1}}^{t_k} \Phi(t_k, \tau) G(\tau) w(\tau) \int_{t_{k-1}}^{t_k} w^T(s) G^T(s) \Phi^T(t_k, s) ds d\tau\right\} \\ &= \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau) G(\tau) \int_{t_{k-1}}^{t_k} \mathbb{E}[w(\tau) w^T(s)] G^T(s) \Phi^T(t_k, s) ds d\tau. \end{aligned} \quad (18)$$

Substituting Equation (5) into the above equation:

$$\begin{aligned} \mathbb{E}[(\Gamma_{k|k-1} W_{k-1})(\Gamma_{k|k-1} W_{k-1})^T] &= \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau) G(\tau) \int_{t_{k-1}}^{t_k} Q(t) \delta(\tau - s) G^T(s) \Phi^T(t_k, s) ds d\tau \\ &= \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau) G(\tau) Q(\tau) G^T(\tau) \Phi^T(t_k, \tau) d\tau. \end{aligned} \quad (19)$$

When the noise control matrix $G(\tau)$ changes slowly during the time interval $[t_{k-1}, t_k]$, Equation (19) becomes:

$$\begin{aligned}
& \mathbb{E}[(\Gamma_{k|k-1} W_{k-1})(\Gamma_{k|k-1} W_{k-1})^T] \\
& \approx \int_{t_{k-1}}^{t_k} [I + F(t_{k-1})(t_k - \tau)] G(t_{k-1}) Q(\tau) G^T(t_{k-1}) [I + F(t_{k-1})(t_k - \tau)]^T d\tau \\
& = [I + F(t_{k-1})T_s] \cdot [G(t_{k-1})Q(t_{k-1})G^T(t_{k-1})T_s] \cdot [I + F(t_{k-1})T_s]^T \\
& \quad + \frac{1}{12} F(t_{k-1})G(t_{k-1})Q(t_{k-1})G^T(t_{k-1})F(t_{k-1})^T T_s^3 \\
& \approx \{[I + F(t_{k-1})T_s]G(t_{k-1})\} \cdot [Q(t_{k-1})T_s] \cdot \{[I + F(t_{k-1})T_s]G(t_{k-1})\}^T.
\end{aligned} \tag{20}$$

When $F(t_{k-1})T_s \ll I$ is satisfied, the above equation can be further approximated:

$$\mathbb{E}[(\Gamma_{k|k-1} W_{k-1})(\Gamma_{k|k-1} W_{k-1})^T] \approx G(t_{k-1}) \cdot [Q(t_{k-1})T_s] \cdot G^T(t_{k-1}). \tag{21}$$

Comparing the result with Equation (10):

$$\Gamma_{k|k-1} \approx [I + F(t_{k-1})T_s]G(t_{k-1}) \approx G(t_{k-1}) \tag{22}$$

$$\mathbb{E}[W_k W_j^T] = Q_k \delta_{kj} = [Q(t_k)T_s] \delta_{kj}. \tag{23}$$

Notice that [29]:

$$Q_k = Q(t_k)T_s. \tag{24}$$

The derivation of the equation relating to R_k and R is more subtle. In the continuous model, $v(t)$ is white, so simple sampling of $Z(t)$ leads to measurement noise with infinite variance. Hence, in the sampling process, we have to imagine averaging the continuous measurement over the T_s interval to get an equivalent discrete sample. This is justified because x is not the Gaussian white noise and can be approximately constant within the interval.

$$Z_k = \frac{1}{T_s} \int_{t_{k-1}}^{t_k} Z(t) dt = \frac{1}{T_s} \int_{t_{k-1}}^{t_k} [H(t)x(t) + v(t)] dt = H(t_k)x_k + \frac{1}{T_s} \int_{t_{k-1}}^{t_k} v(t) dt. \tag{25}$$

Then, the discrete noise matrix and the continuous noise matrix are equivalent:

$$V_k = \frac{1}{T_s} \int_{t_{k-1}}^{t_k} v(t) dt. \tag{26}$$

From Equation (12), we have:

$$\begin{aligned}
\mathbb{E}[V_k V_j^T] &= R_k \delta_{kj} = \frac{1}{T_s^2} \int_{t_{k-1}}^{t_k} \int_{t_{j-1}}^{t_j} \mathbb{E}[v(\tau)v(s)] d\tau ds \\
&= \frac{1}{T_s^2} \int_{t_{k-1}}^{t_k} \int_{t_{j-1}}^{t_j} R(\tau)\delta(s-\tau) d\tau ds = \frac{1}{T_s^2} \int_{t_{k-1}}^{t_k} R(\tau)\delta_{kj} d\tau \approx \frac{R(t_k)}{T_s} \delta_{kj}.
\end{aligned} \tag{27}$$

Comparing it with Equation (6), we have [29]:

$$R_k = \frac{R(t_k)}{T_s}. \tag{28}$$

2.3. Derivation of the Kalman Filter

Assuming that the optimal state estimation at t_{k-1} is \hat{X}_{k-1} , the state estimation error is \tilde{X}_{k-1} , and the state estimation covariance matrix is Σ_{k-1} :

$$\tilde{X}_{k-1} = X_{k-1} - \hat{X}_{k-1} \tag{29}$$

and

$$\Sigma_{k-1} = \mathbb{E}[\tilde{X}_{k-1}\tilde{X}_{k-1}^T] = \mathbb{E}[(X_{k-1} - \hat{X}_{k-1})(X_{k-1} - \hat{X}_{k-1})^T]. \quad (30)$$

If we take the expectation operator of both sides of Equation (8), we obtain the state one-step prediction and the state one-step estimation error:

$$X_{k|k-1}^- = \mathbb{E}[X_k] = \mathbb{E}[\Phi_{k|k-1}X_{k-1} + \Gamma_{k|k-1}W_{k-1}] = \Phi_{k|k-1}\mathbb{E}[X_{k-1}] = \Phi_{k|k-1}\hat{X}_{k-1}, \quad (31)$$

$$\tilde{X}_{k|k-1} = X_k - X_{k|k-1}^-. \quad (32)$$

Substituting Equations (8) and (31) into Equation (32) leads to:

$$\begin{aligned} \tilde{X}_{k|k-1} &= (\Phi_{k|k-1}X_{k-1} + \Gamma_{k|k-1}W_{k-1}) - \Phi_{k|k-1}\hat{X}_{k-1} \\ &= \Phi_{k|k-1}(X_{k-1} - \hat{X}_{k-1}) + \Gamma_{k|k-1}W_{k-1} = \Phi_{k|k-1}\tilde{X}_{k-1} + \Gamma_{k|k-1}W_{k-1}. \end{aligned} \quad (33)$$

Since \tilde{X}_{k-1} is uncorrelated with W_{k-1} , we therefore obtain the covariance of the state one-step estimation error $\tilde{X}_{k|k-1}$ as follows:

$$\begin{aligned} \Sigma_{k|k-1} &= \mathbb{E}[\tilde{X}_{k|k-1}\tilde{X}_{k|k-1}^T] = \mathbb{E}[(\Phi_{k|k-1}\tilde{X}_{k-1} + \Gamma_{k|k-1}W_{k-1})(\Phi_{k|k-1}\tilde{X}_{k-1} + \Gamma_{k|k-1}W_{k-1})^T] \\ &= \Phi_{k|k-1}\mathbb{E}[\tilde{X}_{k-1}\tilde{X}_{k-1}^T]\Phi_{k|k-1}^T + \Gamma_{k|k-1}\mathbb{E}[W_{k-1}W_{k-1}^T]\Gamma_{k|k-1}^T \\ &= \Phi_{k|k-1}\Sigma_{k-1}\Phi_{k|k-1}^T + \Gamma_{k|k-1}Q_{k-1}\Gamma_{k|k-1}^T. \end{aligned} \quad (34)$$

In a similar way, the measurement at t_k can be predicted by the state one-step estimation prediction $X_{k|k-1}^-$ and system measurement Equation (9) as follows:

$$Z_{k|k-1}^- = \mathbb{E}[H_kX_{k|k-1}^- + V_k] = H_kX_{k|k-1}^-. \quad (35)$$

In fact, there is difference between the measurement one-step prediction $Z_{k|k-1}^-$ and the actual measurement Z_k . The difference is denoted as measurement one-step prediction error:

$$\tilde{Z}_{k|k-1} = Z_k - Z_{k|k-1}^-. \quad (36)$$

Substituting the measurement Equations (9) and (35) into Equation (36) yields:

$$\tilde{Z}_{k|k-1} = Z_k - H_kX_{k|k-1}^- = H_kX_k + V_k - H_kX_{k|k-1}^- = H_k\tilde{X}_{k|k-1} + V_k. \quad (37)$$

In general, the measurement one-step prediction error $\tilde{Z}_{k|k-1}$ is called innovation in the classical Kalman filter theory, and it indicates the new information about the state estimate carried by the measurement one-step prediction error.

On the one hand, if the estimation of X_k only includes the state one-step prediction $X_{k|k-1}^-$ of the system state equation, the estimation accuracy will be low, as no information of the measurement equation has been used. On the other hand, according to Equation (37), the measurement one-step prediction error calculated using the system measurement equation contains the information of the state one-step prediction of $X_{k|k-1}^-$. Consequently, it is natural to consider all the state information that comes from the system state equation and the measurement equation, respectively, and correct the state one-step prediction mean $X_{k|k-1}^-$ with the measurement one-step prediction error $\tilde{Z}_{k|k-1}$. Thereby, the optimal estimation of X_k can be calculated by the combination of $X_{k|k-1}^-$ and $\tilde{Z}_{k|k-1}$ as follows:

$$\hat{X}_k = X_{k|k-1}^- + K_k\tilde{Z}_{k|k-1}, \quad (38)$$

where K_k is the undetermined correction factor matrix.

Substituting Equations (31) and (37) into Equation (38) obtains:

$$\begin{aligned}\hat{X}_k &= X_{k|k-1}^- + K_k(Z_k - H_k X_{k|k-1}^-) = (I - K_k H_k) X_{k|k-1}^- + K_k Z_k \\ &= (I - K_k H_k) \Phi_{k|k-1} \hat{X}_{k-1} + K_k Z_k.\end{aligned}\quad (39)$$

From Equation (39), the current state estimation \hat{X}_k is a linear combination of the last state estimation \hat{X}_{k-1} and the current measurement Z_k , which considers the influence of the structural parameters $\Phi_{k|k-1}$ in the state equation and the structure parameters H_k in the measurement equation with different types of construction.

The state estimation error at the current time t_k is denoted as:

$$\tilde{X}_k = X_k - \hat{X}_k, \quad (40)$$

where X_k is the true values and \hat{X}_k is the posterior estimation of X_k .

Substituting Equation (39) into Equation (40) obtains:

$$\begin{aligned}\tilde{X}_k &= X_k - [X_{k|k-1}^- + K_k(Z_k - H_k X_{k|k-1}^-)] = (X_k - X_{k|k-1}^-) - K_k(H_k X_k + V_k - H_k X_{k|k-1}^-) \\ &= \tilde{X}_{k|k-1} - K_k(H_k \tilde{X}_{k|k-1} + V_k) = (I - K_k H_k) \tilde{X}_{k|k-1} - K_k V_k.\end{aligned}\quad (41)$$

Then, the mean square error matrix of state estimation \hat{X}_k is given by:

$$\begin{aligned}\Sigma_k &= \mathbb{E}[\tilde{X}_k \tilde{X}_k^T] = \mathbb{E}\{[(I - K_k H_k) \tilde{X}_{k|k-1} - K_k V_k][(I - K_k H_k) \tilde{X}_{k|k-1} - K_k V_k]^T\} \\ &= (I - K_k H_k) \mathbb{E}[\tilde{X}_{k|k-1} \tilde{X}_{k|k-1}^T] (I - K_k H_k)^T + K_k \mathbb{E}[V_k V_k^T] K_k^T \\ &= (I - K_k H_k) \Sigma_{k|k-1} (I - K_k H_k)^T + K_k R_k K_k^T.\end{aligned}\quad (42)$$

Substituting Equation (34) into Equation (42) obtains:

$$\begin{aligned}\Sigma_k &= (I - K_k H_k) [\Phi_{k|k-1} \Sigma_{k-1} \Phi_{k|k-1}^T + \Gamma_{k|k-1} Q_{k-1} \Gamma_{k|k-1}^T] (I - K_k H_k)^T + K_k R_k K_k^T \\ &= \Phi_{k|k-1} \Sigma_{k-1} \Phi_{k|k-1}^T + K_k H_k \Phi_{k|k-1} \Sigma_{k-1} \Phi_{k|k-1}^T H_k^T K_k^T - \Phi_{k|k-1} \Sigma_{k-1} \Phi_{k|k-1}^T H_k^T K_k^T \\ &\quad - K_k H_k \Phi_{k|k-1} \Sigma_{k-1} \Phi_{k|k-1}^T + \Gamma_{k|k-1} Q_{k-1} \Gamma_{k|k-1}^T - K_k H_k \Gamma_{k|k-1} Q_{k-1} \Gamma_{k|k-1}^T \\ &\quad - \Gamma_{k|k-1} Q_{k-1} \Gamma_{k|k-1}^T H_k^T K_k^T + K_k H_k \Gamma_{k|k-1} Q_{k-1} \Gamma_{k|k-1}^T H_k^T K_k^T + K_k R_k K_k^T.\end{aligned}\quad (43)$$

We now use the approximation $\Phi_{k|k-1} \approx I + F(t_{k-1})T_s$ as Equation (15). From Equation (22) with $\Gamma_{k|k-1} \approx G(t_{k-1})$, we have:

$$\begin{aligned}\Sigma_k &= [I + F(t_{k-1})T_s] \Sigma_{k-1} [I + F(t_{k-1})T_s]^T + K_k H_k [I + F(t_{k-1})T_s] \Sigma_{k-1} [I + F(t_{k-1})T_s]^T H_k^T K_k^T \\ &\quad - [I + F(t_{k-1})T_s] \Sigma_{k-1} [I + F(t_{k-1})T_s]^T H_k^T K_k^T - K_k H_k [I + F(t_{k-1})T_s] \Sigma_{k-1} [I + F(t_{k-1})T_s]^T \\ &\quad + G(t_{k-1}) Q_{k-1} G^T(t_{k-1}) - K_k H_k G(t_{k-1}) Q_{k-1} G^T(t_{k-1}) - G(t_{k-1}) Q_{k-1} G^T(t_{k-1}) H_k^T K_k^T \\ &\quad + K_k H_k G(t_{k-1}) Q_{k-1} G^T(t_{k-1}) H_k^T K_k^T + K_k R_k K_k^T.\end{aligned}\quad (44)$$

Note from Equation (24) that Q_k is of the order of T_s and from Equation (28) that $R_k = \frac{R(t_k)}{T_s}$; then, Equation (44) becomes:

$$\begin{aligned}\Sigma_k &= [I + F(t_{k-1})T_s] \Sigma_{k-1} [I + F(t_{k-1})T_s]^T + K_k H_k [I + F(t_{k-1})T_s] \Sigma_{k-1} [I + F(t_{k-1})T_s]^T H_k^T K_k^T \\ &\quad - [I + F(t_{k-1})T_s] \Sigma_{k-1} [I + F(t_{k-1})T_s]^T H_k^T K_k^T - K_k H_k [I + F(t_{k-1})T_s] \Sigma_{k-1} [I + F(t_{k-1})T_s]^T \\ &\quad + G(t_{k-1}) Q(t_k) T_s G^T(t_{k-1}) - K_k H_k G(t_{k-1}) Q(t_k) T_s G^T(t_{k-1}) \\ &\quad - G(t_{k-1}) Q(t_k) T_s G^T(t_{k-1}) H_k^T K_k^T + K_k H_k G(t_{k-1}) Q(t_k) T_s G^T(t_{k-1}) H_k^T K_k^T + K_k \frac{R(t_k)}{T_s} K_k^T.\end{aligned}\quad (45)$$

2.4. The Temporal Derivative of the Rényi Entropy and the Kalman Filter Gain

To obtain the continuous form of covariance matrix Σ , the limit will be taken. However, the relation between the undetermined correction factor matrix K_k and its continuous form still remains unknown. Therefore, we make the following assumption.

Assumption 1. K_k is of the order of T_s , that is:

$$K(t_k) = \frac{K_k}{T_s}. \quad (46)$$

From the conclusion, we can also derive this assumption conversely. We next draw the conclusion as one theorem under the assumption, as follows:

Theorem 1. The discrete form of the undetermined correction factor matrix is the same as the continuous form when the temporal derivative of Rényi entropy is minimized. This can be presented in a mathematical form as follows:

$$\{K_k = \Sigma_k H_k^T R_k, K = \Sigma H^T R^{-1} | K^* = \arg \min_K \dot{H}_R^{(\alpha)}(K)\}. \quad (47)$$

Proof of Theorem 1. We substitute the expression for K_k into Equation (45) and neglect higher-order terms in T_s ; Equation (45) becomes:

$$\begin{aligned} \Sigma_k &= [I + F(t_{k-1})T_s]\Sigma_{k-1}[I + F(t_{k-1})T_s]^T + T_s K(t_k) H_k \Sigma_{k-1} [I + F(t_{k-1})T_s]\Sigma_{k-1} \\ &\quad [I + F(t_{k-1})T_s]^T H_k^T T_s K^T(t_k) - [I + F(t_{k-1})T_s]\Sigma_{k-1} [I + F(t_{k-1})T_s]^T H_k^T T_s K^T(t_k) \\ &\quad - T_s K(t_k) H_k [I + F(t_{k-1})T_s]\Sigma_{k-1} [I + F(t_{k-1})T_s]^T + G(t_{k-1})Q(t_k)T_s G^T(t_{k-1}) \\ &\quad - T_s K(t_k) H_k G(t_{k-1})Q(t_k)T_s G^T(t_{k-1}) - G(t_{k-1})Q(t_k)T_s G^T(t_{k-1}) H_k^T T_s K^T(t_k) \\ &\quad + T_s K(t_k) H_k G(t_{k-1})Q(t_k)T_s G^T(t_{k-1}) H_k^T T_s K^T(t_k) + T_s K(t_k) \frac{R_k}{T_s} T_s K^T(t_k) \\ &= \Sigma_{k-1} + T_s F(t_{k-1})\Sigma_{k-1} + T_s \Sigma_{k-1} F^T(t_{k-1}) - \Sigma_{k-1} H_k^T T_s K(t_k)^T - T_s K(t_k) H_k \Sigma_{k-1} \\ &\quad + G(t_{k-1})Q(t_k)T_s G^T(t_{k-1}) + T_s K(t_k) \frac{R(t_k)}{T_s} T_s K^T(t_k). \end{aligned} \quad (48)$$

Moving the first term of Equation (48) from right to left and dividing both sides by T_s to form the finite difference expression:

$$\begin{aligned} \frac{\Sigma_k - \Sigma_{k-1}}{T_s} &= F(t_{k-1})\Sigma_{k-1} + \Sigma_{k-1}F^T(t_{k-1}) - \Sigma_{k-1}H_k^T K(t_k)^T - K(t_k)H_k\Sigma_{k-1} \\ &\quad + G(t_{k-1})Q(t_k)G^T(t_{k-1}) + K(t_k)R(t_k)K^T(t_k). \end{aligned} \quad (49)$$

Finally, passing to the limit as $T_s \rightarrow 0$ and dropping of the subscripts lead to the matrix differential equation:

$$\dot{\Sigma} = F\Sigma + \Sigma F^T - \Sigma H^T K^T - K H \Sigma + G Q G^T + K R K^T. \quad (50)$$

Σ is invertible, as it is a positive matrix. Multiplying Σ^{-1} with Equation (50), we can consider the temporal derivative of the Rényi entropy of the mean square error matrix Σ using Equation (2):

$$\begin{aligned}
\dot{H}_R^{(\alpha)} &= \frac{1}{2} \text{Tr}\{\Sigma^{-1} \dot{\Sigma}\} \\
&= \frac{1}{2} \text{Tr}\{\Sigma^{-1} F \Sigma + F^T - H^T K^T - \Sigma^{-1} K H \Sigma + \Sigma^{-1} G Q G^T + \Sigma^{-1} K R K^T\} \\
&= \frac{1}{2} \text{Tr}\{F + F^T - H^T K^T - K H + \Sigma^{-1} G Q G^T + \Sigma^{-1} K R K^T\} \\
&= \frac{1}{2} \text{Tr}\{2F - 2KH + \Sigma^{-1} G Q G^T + \Sigma^{-1} K R K^T\},
\end{aligned} \tag{51}$$

where the invariance under the cyclic permutation property of the trace operator has been used to eliminate Σ^{-1} and Σ , as well as the truth that $\text{Tr}(F) = \text{Tr}(F^T)$ has been used to simplify the formula.

It is obvious that Equation (51) is a quadratic function of the undetermined correction factor matrix K . Thereby, there must be a minimum of $\dot{H}_R^{(\alpha)}(x)$ in a probabilistic sense. Taking the derivative of both sides of Equation (51) with respect to matrix K obtains:

$$\begin{aligned}
\frac{\partial}{\partial K} \dot{H}_R^{(\alpha)} &= -2 \frac{\partial \text{Tr}(KH)}{\partial K} + \frac{\partial \text{Tr}(\Sigma^{-1} K R K^T)}{\partial K} \\
&= -2H^T + \frac{\text{Tr}(\Sigma^{-1} K R (\partial K)^T)}{\partial K} + \frac{\text{Tr}(\Sigma^{-1} (\partial K) R K^T)}{\partial K} \\
&= -2H^T + \Sigma^{-1} K R + (R K^T \Sigma^{-1})^T.
\end{aligned} \tag{52}$$

In addition, since Σ^{-1} and R_k are symmetric matrices, the result is:

$$\frac{\partial}{\partial K} \dot{H}_R^{(\alpha)} = -2H^T + 2\Sigma^{-1} K R. \tag{53}$$

R_k is invertible, as it is a positive matrix. According to the extreme value principle of the function, when the above are equal to zero, then we have:

$$K = \Sigma H^T R^{-1}. \tag{54}$$

So far, we have found the analytic solution to the undetermined correction factor matrix K , which is called the continuous-time Kalman filter gain in the classical Kalman filter. Then, the recursive formulations of the Kalman filter can be established through the Kalman filter gain K . Most importantly, this implies the connection between the temporal derivative of Rényi entropy and the classical Kalman filter: The temporal derivative of the Rényi entropy is minimized when the Kalman filter gain satisfies Equation (54).

Looking back to Assumption 1 and substituting Equation (28) into Equation (54), we obtain:

$$K(t_k) = \frac{K_k}{T_s} = K = \Sigma H^T R^{-1} = \Sigma_k H_k^T R_k(T_s) = \frac{\Sigma_k H_k^T R_k}{T_s}. \tag{55}$$

Therefore, the discrete-time Kalman filter gain can be expressed as follows:

$$K_k = \Sigma_k H_k^T R_k. \tag{56}$$

□

Remark 1. The discrete-time Kalman filter gain has the same form as the continuous-time filter gain, as shown in the Equation (54). In principle, this is consistent with our intuition and proves the correctness and rationality of Assumption 1, in turn.

Remark 2. The Kalman filter gain is equivalent to the minimization of the temporal derivative of the Rényi entropy, although it has the same result as the original Kalman filter, which is deduced under the minimum mean square error criterion.

Substituting Equation (54) into Equation (50), we have:

$$\begin{aligned}\dot{\Sigma} &= F\Sigma + \Sigma F^T - \Sigma H^T K^T - \Sigma H^T R^{-1} H \Sigma + G Q G^T + \Sigma H^T R^{-1} R K^T \\ &= F\Sigma + \Sigma F^T - \Sigma H^T R^{-1} H \Sigma + G Q G^T.\end{aligned}\quad (57)$$

This is a second-order nonlinear differential equation with respect to the mean square error matrix Σ , and it is commonly called the Riccati equation. This is the same result as that of the Bucy–Kalman filter [7].

If the system equation, Equation (3), and the measurement equation, Equation (4), form a linear time-invariant system with constant noise covariance, the mean square error matrix Σ may reach a steady-state value, and $\dot{\Sigma}$ may eventually reach zero. So, we have the continuous algebraic Riccati equation as follows:

$$\dot{\Sigma} = F\Sigma + \Sigma F^T - \Sigma H^T R^{-1} H \Sigma + G Q G^T = 0. \quad (58)$$

As we can see, the time derivative of covariance at the steady state is zero; then, the temporal derivative of the Rényi entropy should also be zero:

$$\dot{H}_R^{(\alpha)} = 0. \quad (59)$$

This implies that when the system approaches a stable state, the Rényi entropy approaches a steady value so that the temporal derivative of the Rényi entropy is zero. This is reasonable when the steady system owns a constant Rényi entropy, as uncertainty is stable, which follows our intuitive understanding. Consequently, it is worth noting that whether the value of the Rényi entropy is stable or not can be a validated indicator of whether the system is approaching the steady state.

3. Simulations and Analysis

In this section, we give two experiments to show that when the nonlinear filter system approaches the steady state, the Rényi entropy of the system approaches stability. The first experiment is a numerical example of a falling body in noisy conditions, tracked by radar [30] using the UKF. The second experiment is a practical experiment of loosely coupled integration [29]. The simulations were carried out on MATLAB 2018a running on a computer with i5-5200U, 2.20 GHz CPU, and the graphs were plotted by MATLAB.

3.1. Falling Body Tracking

In the example of a falling body being tracked by radar, the body falls vertically. The radar is placed at a vertical distance L from the body, and the radar measures the distance y from the radar to the body. The state-space equation of the body is given by:

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= d + g \\ \dot{x}_3 &= 0,\end{aligned}\quad (60)$$

where x_1 is the height, x_2 is the velocity, x_3 is the ballistic coefficient, $g = -9.81 \text{ m/s}^2$ is the gravity acceleration, and d is the air drag, which could be approximated as:

$$d = \frac{\rho x_2^2}{2x_3} = \rho_0 \exp\left(-\frac{x_1}{k}\right) \frac{x_2^2}{2x_3}, \quad (61)$$

where ρ is the air density with an initial value of $\rho_0 = 1.225$; $\rho_0 = 1.225$ and $k = 6705.6$ are constants.

The measurement equation is:

$$y = \sqrt{L^2 + x_1^2}. \quad (62)$$

It is worth noting that the drag and the square root cause severely nonlinearity in the state-space function and measurement function, respectively.

The discrete-time nonlinear system can be given by the Euler discretization method. Combining the additive process with Gaussian white noises for measurement, we can obtain:

$$\begin{aligned}x_1(n+1) &= x_1(n) + x_2(n) \cdot T + w_1(n) \\x_2(n+1) &= x_2(n) + (d + g) \cdot T + w_2(n) \\x_3(n+1) &= x_3(n) + w_3(n)\end{aligned}\quad (63)$$

$$y(n) = \sqrt{L^2 + x_1^2(n)} + v(n). \quad (64)$$

In the UKF numerical experiment, we set the sampling period to $T = 0.4$ s, the horizontal distance to $L = 100$ m, the maximum number of samples to $N = 100$, the process noise to $S_w = \text{diag}(10^5, 10^3, 10^2)$, the measurement noise to $S_v = 10^6$, and the initial state to $x = [10^5; -5000; 400]$. The results are shown as follows:

Figure 1 shows the evolution of covariance matrix Σ . Figures 2 and 3 show the Rényi entropy of covariance matrix Σ and its change in adjacent time, respectively. Notice that the uncertainty increases near the middle of the plots, which is coincident with the drag peak. However, the Rényi entropy fluctuates around 15; even the fourth element of Σ changes dramatically. Of course, the entropy changes are closely accompanied by the drag peak, which means the change of the entropy of covariance reflects the evolution of matrix Σ . Consequently, the Rényi entropy can be viewed as the indicator of whether the system is approaching the steady state or not.

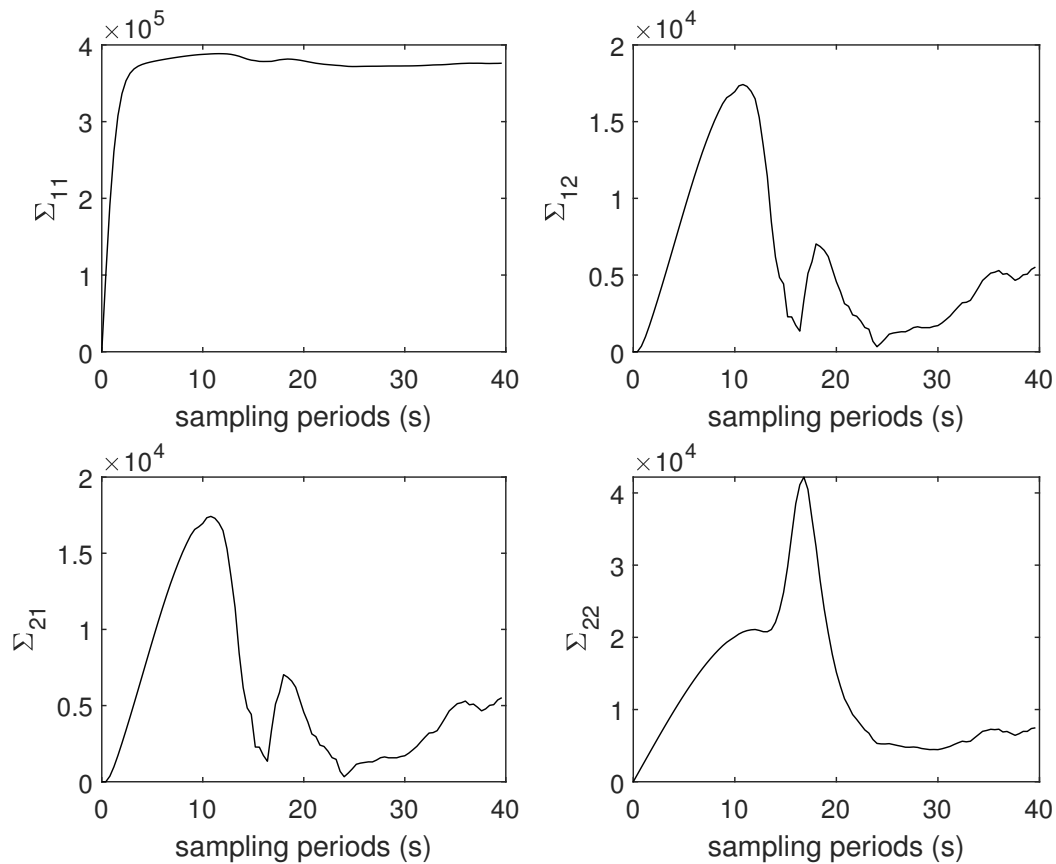


Figure 1. Evolution of matrix Σ .

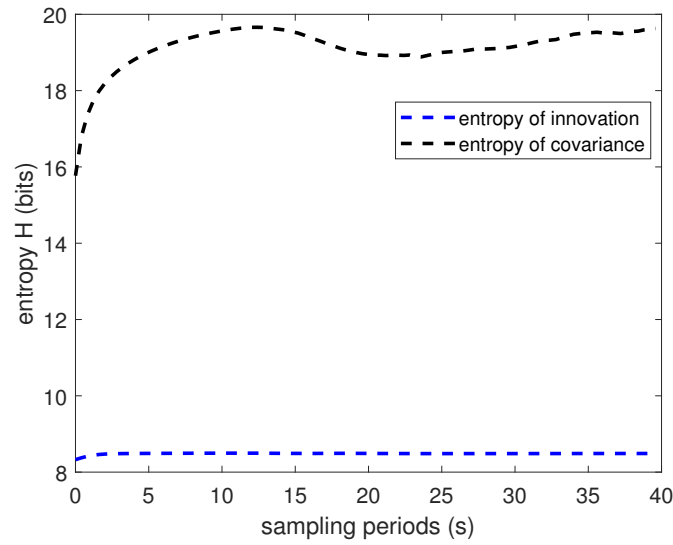


Figure 2. Simulation results for the entropy.

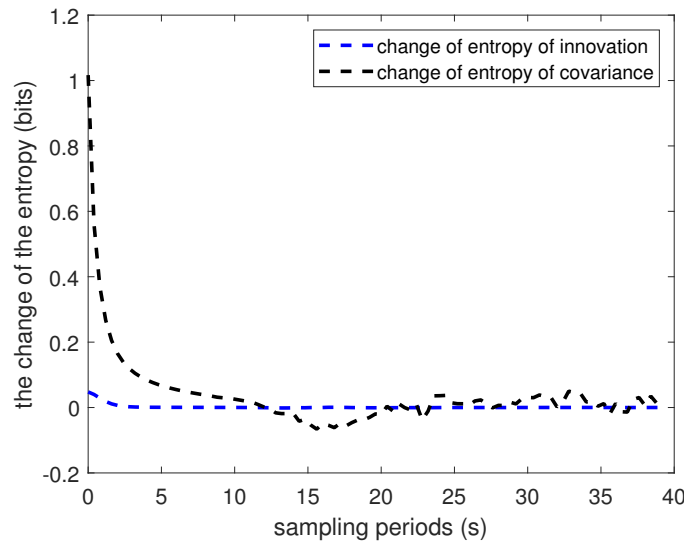


Figure 3. Simulation results for the change of entropy.

3.2. Practical Integrated Navigation

In the loosely integrated navigation system, the system state parameter x is composed of inertial navigation system (INS) error states in the North–East–Down (NED) local-level navigation frame, and can be expressed as follows:

$$x = [(\delta r^n)^T \quad (\delta v^n)^T \quad (\psi)^T \quad (b_g)^T \quad (b_a)^T]^T, \quad (65)$$

where δr^n , δv^n , and ψ represent the position error, the velocity error, and the attitude error, respectively; b_g and b_a are modeled as first-order Gauss–Markov processes, representing the gyroscope bias and the accelerometer bias, respectively.

The discrete-time state update equation is used to update state parameters as follows:

$$x_k = \Phi_{k|k-1} x_{k-1} + G_{k|k-1} w_{k-1}, \quad (66)$$

where $G_{k|k-1}$ is the system noise matrix, w_{k-1} is the system noise, and $\Phi_{k|k-1}$ is the state transition matrix from t_{k-1} to t_k ; this is determined by the dynamic model of the state parameter.

In the loosely coupled integration, the measurement equation can be simply expressed as:

$$\delta z = H_k x_k + v_k, \quad (67)$$

where v_k is the measurement noise, H_k is the measurement matrix, and z_k is the measurement vector calculated by subtracting the global navigation satellite system (GNSS) observation with the inertial navigation system (INS) mechanism.

The experiments reported in this section were carried out by processing the data from an unmanned ground vehicle test. The gyroscope random walk was set to $0.03 \text{ deg}/\sqrt{h}$ and the velocity random walk was set to $0.05 \text{ m/s}/\sqrt{h}$. The sampling rates of the inertial measurement unit (IMU) and the GNSS are 200 Hz and 1 Hz, respectively. The test lasts 48 min.

The position error curve, velocity error curve, and attitude error curve of the loosely coupled integration are shown in Figures 4–6. The root mean squares (RMSs) of the position errors in the north, east, and earth directions are 0.0057 m, 0.0024 m, and 0.0134 m, respectively. The RMS of the velocity errors in the north, east, and earth directions are 0.0023 m/s, 0.0021 m/s, and 0.0038 m/s, respectively. The RMSs of the attitude errors in the roll, pitch, and yaw directions are 0.0034 deg, 0.0030 deg, and 0.0178 deg, respectively.

The Rényi entropy of the covariance P is shown in Figure 7. As we can see, the Rényi entropy fluctuates around -100 once the filter converges, which is consistent with the conclusion from the entropy perspective.

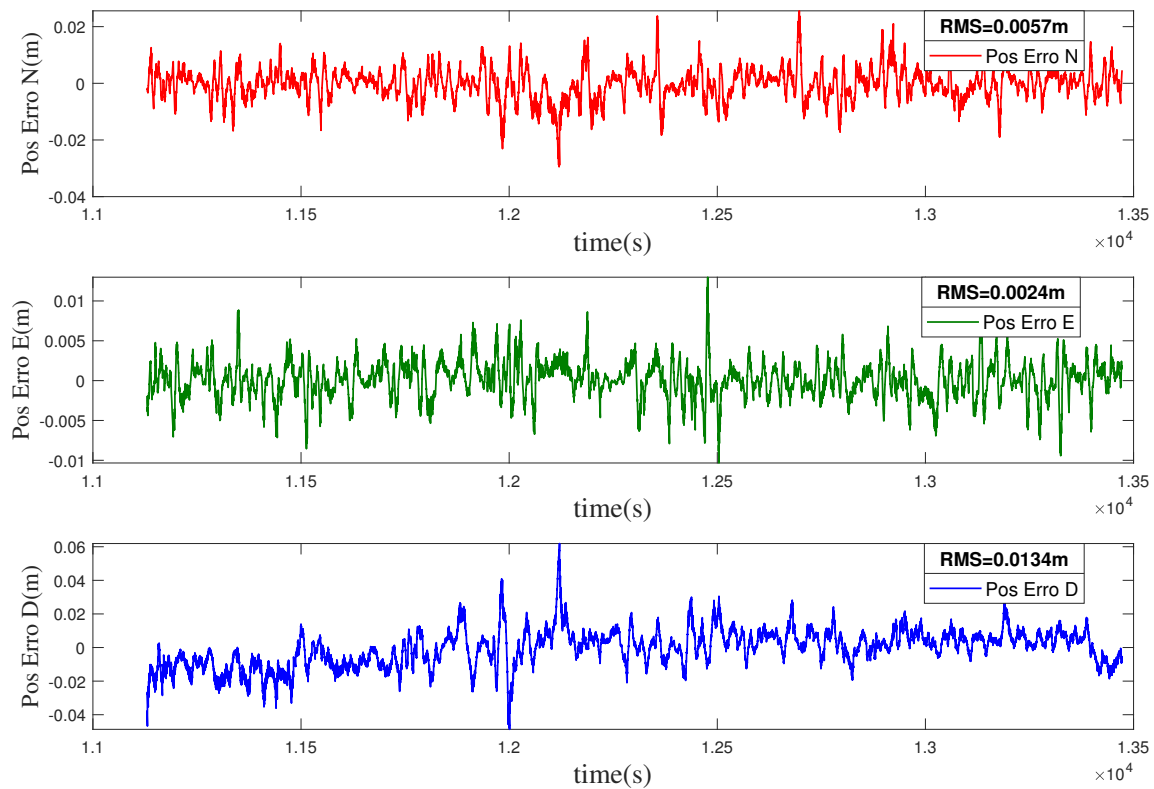


Figure 4. Position error of the loosely coupled integration.

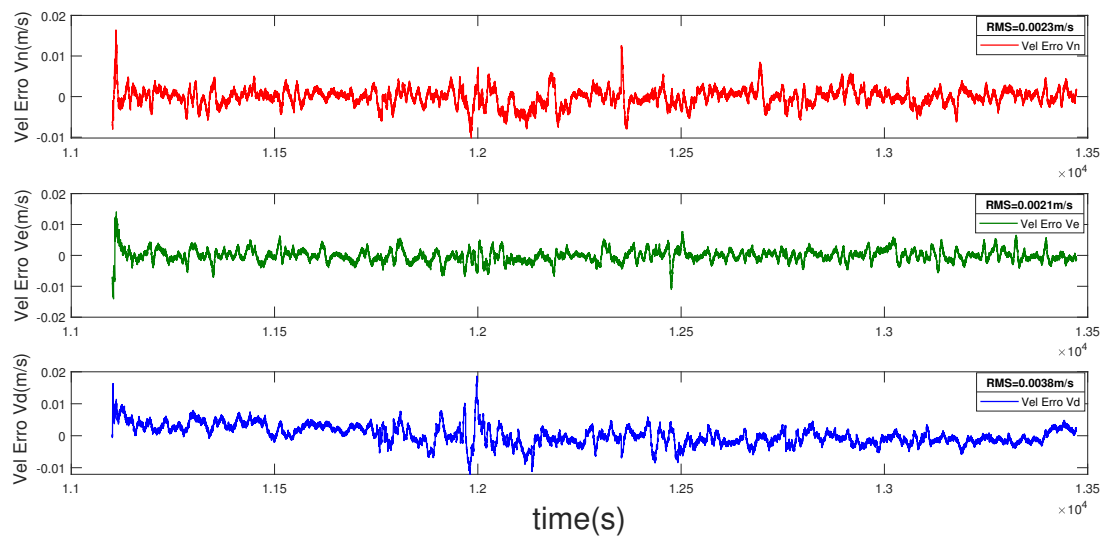


Figure 5. Velocity error of the loosely coupled integration.

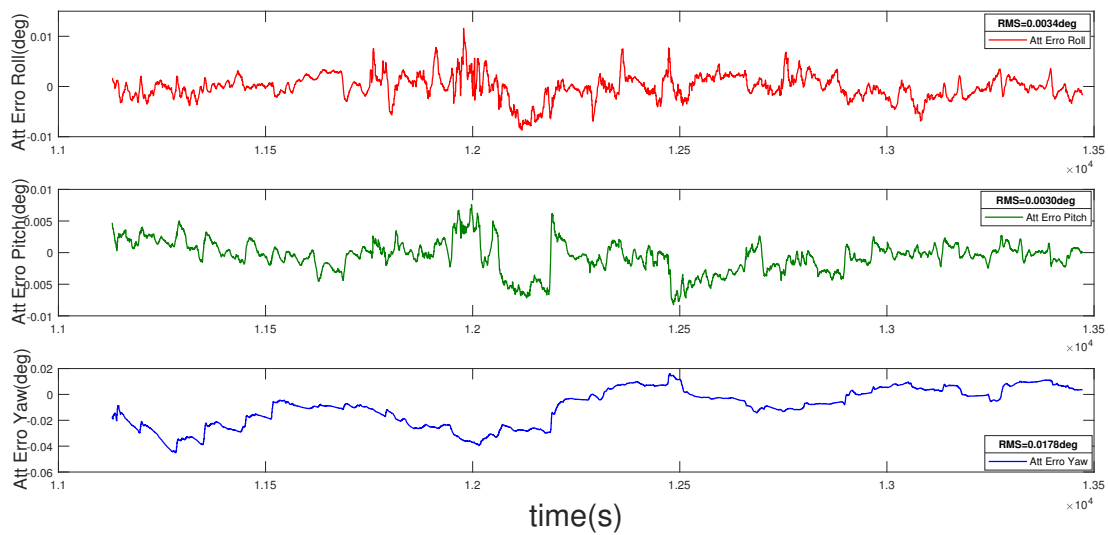


Figure 6. Attitude error of the loosely coupled integration.

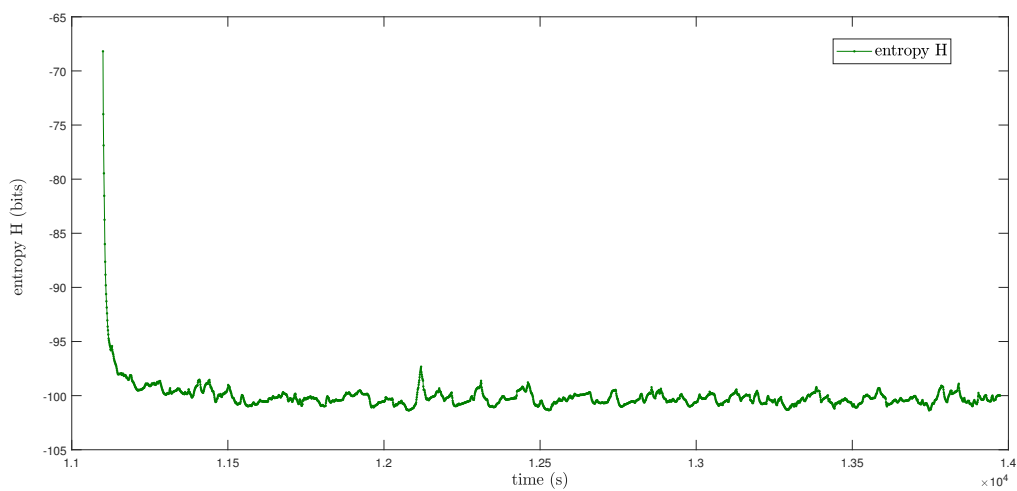


Figure 7. Rényi entropy of the covariance Σ .

4. Conclusions and Final Remarks

We have considered the original Kalman filter by taking the minimization of the temporal derivative of the Rényi entropy. In particular, we show that the temporal derivative of Rényi entropy is equal to zero when the Kalman filter system approaches the steady state, which means that the Rényi entropy approaches a stable value. Finally, simulation experiments and practical experiments show the Rényi entropy truly stays stable when the system becomes steady.

Future work includes calculating the Rényi entropy of the innovation term when the measurements and the noise are non-Gaussian [14] in order to evaluate the effectiveness of measurements and adjust the noise covariance matrix. Meanwhile, we can also calculate the Rényi entropy of the nonlinear dynamical equation to measure the nonlinearity in the propagation step.

Author Contributions: Conceptualization, Y.L. and C.G.; Funding acquisition, C.G. and J.L.; Investigation, Y.L.; Methodology, Y.L., C.G., and S.Y.; Project administration, J.L.; Resources, C.G.; Software, Y.L. and S.Y.; Supervision, J.L.; Validation, S.Y.; Visualization, S.Y.; Writing—original draft, Y.L.; Writing—review and editing, C.G., S.Y., and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by a grant from the National Key Research and Development Program of China (2018YFB1305001).

Acknowledgments: In this section you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

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

References

- Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [\[CrossRef\]](#)
- Principe, J.C. *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives*; Springer Science & Business Media: Berlin, Germany, 2010.
- He, R.; Hu, B.; Yuan, X.; Wang, L. *Robust Recognition via Information Theoretic Learning*; Springer International Publishing: Berlin, Germany, 2014.
- Rényi, A. On measures of entropy and information. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 20 June–30 July 1961.
- Liang, X.S. Entropy evolution and uncertainty estimation with dynamical systems. *Entropy* **2014**, *16*, 3605–3634. [\[CrossRef\]](#)
- Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [\[CrossRef\]](#)
- Kalman, R.E.; Bucy, R.S. New results in linear filtering and prediction theory. *J. Basic Eng.* **1961**, *83*, 95–108. [\[CrossRef\]](#)
- DeMars, K.J. Nonlinear Orbit Uncertainty Prediction and Rectification for Space Situational Awareness. Ph.D. Thesis, The University of Texas at Austin, Austin, TX, USA, 2010.
- DeMars, K.J.; Bishop, R.H.; Jah, M.K. Entropy-based approach for uncertainty propagation of nonlinear dynamical systems. *J. Guid. Control. Dyn.* **2013**, *36*, 1047–1057. [\[CrossRef\]](#)
- Kim, H.; Liu, B.; Goh, C.Y.; Lee, S.; Myung, H. Robust vehicle localization using entropy-weighted particle filter-based data fusion of vertical and road intensity information for a large scale urban area. *IEEE Robot. Autom. Lett.* **2017**, *2*, 1518–1524. [\[CrossRef\]](#)
- Zhang, J.; Du, L.; Ren, M.; Hou, G. Minimum error entropy filter for fault detection of networked control systems. *Entropy* **2012**, *14*, 505–516. [\[CrossRef\]](#)
- Liu, Y.; Wang, H.; Hou, C. UKF based nonlinear filtering using minimum entropy criterion. *IEEE Trans. Signal Process.* **2013**, *61*, 4988–4999. [\[CrossRef\]](#)
- Julier, S.; Uhlmann, J.; Durrant-Whyte, H.F. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Trans. Autom. Control* **2000**, *45*, 477–482. [\[CrossRef\]](#)
- Contreras-Reyes, J.E.; Cortés, D.D. Bounds on rényi and shannon entropies for finite mixtures of multivariate skew-normal distributions: Application to swordfish (*xiphias gladius linnaeus*). *Entropy* **2016**, *18*, 382. [\[CrossRef\]](#)

15. Ren, M.; Zhang, J.; Fang, F.; Hou, G.; Xu, J. Improved minimum entropy filtering for continuous nonlinear non-Gaussian systems using a generalized density evolution equation. *Entropy* **2013**, *15*, 2510–2523. [\[CrossRef\]](#)
16. Zhang, Q. Performance enhanced Kalman filter design for non-Gaussian stochastic systems with data-based minimum entropy optimisation. *AIMS Electron. Electr. Eng.* **2019**, *3*, 382. [\[CrossRef\]](#)
17. Chen, B.; Dang, L.; Gu, Y.; Zheng, N.; Príncipe, J.C. Minimum error entropy Kalman filter. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**. [\[CrossRef\]](#)
18. Gultekin, S.; Paisley, J. Nonlinear Kalman filtering with divergence minimization. *IEEE Trans. Signal Process.* **2017**, *65*, 6319–6331. [\[CrossRef\]](#)
19. Darling, J.E.; DeMars, K.J. Minimization of the kullback–leibler divergence for nonlinear estimation. *J. Guid. Control Dyn.* **2017**, *40*, 1739–1748. [\[CrossRef\]](#)
20. Morelande, M.R.; Garcia-Fernandez, A.F. Analysis of Kalman filter approximations for nonlinear measurements. *IEEE Trans. Signal Process.* **2013**, *61*, 5477–5484. [\[CrossRef\]](#)
21. Raitoharju, M.; García-Fernández, Á.F.; Piché, R. Kullback–Leibler divergence approach to partitioned update Kalman filter. *Signal Process.* **2017**, *130*, 289–298. [\[CrossRef\]](#)
22. Hu, E.; Deng, Z.; Xu, Q.; Yin, L.; Liu, W. Relative entropy-based Kalman filter for seamless indoor/outdoor multi-source fusion positioning with INS/TC-OFDM/GNSS. *Clust. Comput.* **2019**, *22*, 8351–8361. [\[CrossRef\]](#)
23. Yu, W.; Peng, J.; Zhang, X.; Li, S.; Liu, W. An adaptive unscented particle filter algorithm through relative entropy for mobile robot self-localization. *Math. Probl. Eng.* **2013**. [\[CrossRef\]](#)
24. Arasaratnam, I.; Haykin, S. Cubature kalman filters. *IEEE Trans. Autom. Control* **2009**, *54*, 1254–1269. [\[CrossRef\]](#)
25. Kiani, M.; Barzegar, A.; Pourtakdoust, S.H. Entropy-based adaptive attitude estimation. *Acta Astronaut.* **2018**, *144*, 271–282. [\[CrossRef\]](#)
26. Giffin, A.; Urniezius, R. The Kalman filter revisited using maximum relative entropy. *Entropy* **2014**, *16*, 1047–1069. [\[CrossRef\]](#)
27. Chen, B.; Liu, X.; Zhao, H.; Principe, J.C. Maximum correntropy Kalman filter. *Automatica* **2017**, *76*, 70–77. [\[CrossRef\]](#)
28. Chen, B.; Xing, L.; Liang, J.; Zheng, N.; Principe, J.C. Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion. *IEEE Signal Process. Lett.* **2014**, *21*, 880–884.
29. Gongmin, Y.; Jun, W. *Lectures on Strapdown Inertial Navigation Algorithm and Integrated Navigation Principles*; Northwestern Polytechnical University Press: Xi'an, China, 2019.
30. Kumari, L.; Padma Raju, K. Application of Extended Kalman filter for a Free Falling body towards Earth. *IJACSA Ed.* **2011**, *2*, 4. [\[CrossRef\]](#)



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).