A UD factorization-based nonlinear adaptive set-membership filter for ellipsoidal estimation

Bo Zhou\textsuperscript{1,2}, Jianda Han\textsuperscript{1} and Guangjun Liu\textsuperscript{3,*,†}

\textsuperscript{1}Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China
\textsuperscript{2}Graduate School of the Chinese Academy of Sciences, Beijing 100039, China
\textsuperscript{3}Department of Aerospace Engineering, Ryerson University, Toronto, Ont., Canada M5B 2K3

SUMMARY

The extended set-membership filter (ESMF) for nonlinear ellipsoidal estimation suffers from numerical instability, computation complexity as well as the difficulty in filter parameter selection. In this paper, a UD factorization-based adaptive set-membership filter is developed and applied to nonlinear joint estimation of both time-varying states and parameters. As a result of using the proposed UD factorization, combined with a new sequential and selective measurement update strategy, the numerical stability and real-time applicability of conventional ESMF are substantially improved. Furthermore, an adaptive selection scheme of the filter parameters is derived to reduce the computation complexity and achieve sub-optimal estimation. Simulation results have shown the efficiency and robustness of the proposed method. Copyright © 2007 John Wiley & Sons, Ltd.

Received 5 September 2006; Revised 30 September 2007; Accepted 1 October 2007

KEY WORDS: set-membership filter; UD factorization; nonlinear state and parameter estimation; adaptive filter parameter determination

1. INTRODUCTION

The problems of state or parameter estimation are commonly solved via stochastic approaches based on Bayesian theory, with some statistics assumptions such as white noise and known mean and covariance. However, in many cases, it is more practical to assume that the noises are unknown but bounded (UBB), especially when the bounds of noises can be obtained. In view of this, the set-membership filter (SMF), which computes a compact feasible set in which the true state or parameter lies only under the UBB noise assumption, provides an attractive alternative [1, 2]. The bound of state or parameter can be attained by using this guaranteed estimation method, which is required in many robust and optimal control approaches.

\*Correspondence to: Guangjun Liu, Department of Aerospace Engineering, Ryerson University, Toronto, Ont., Canada M5B 2K3.
\†E-mail: gjliu@ryerson.ca

Copyright © 2007 John Wiley & Sons, Ltd.
SMF was first explored by Schweppe and Bertsekas at the end of 1960s [3, 4]. They proposed a bounded ellipsoidal estimation algorithm to contain the true state, but no optimization process was considered. Based on their work, important contributions have been presented by Fogel and Huang, who proposed an optimal-bounded ellipsoid algorithm for the state estimation scheme of linear system, where optimization criterion for parameter bounding is investigated [5, 6]. Maksarov et al. also contributed to the ellipsoidal estimation of state or parameter [7–9]. At present, the theory of guaranteed estimation for linear systems is well developed and has become a mature area of control theory [10, 11].

In the framework of SMF, a classical iterative prediction update method is often used. To describe the uncertain set, some regular geometry shapes, such as ellipsoid, orthotope and paralleltope, are used to approximate the feasible set in order to reduce the complexity and computation load. Among them the ellipsoidal estimation seems to be more popular because of its less demand of information for representing the feasible set, more insightful for analogizing the covariance, invariance with respect to linear transformations in the sense that an ellipsoid remains an ellipsoid after a linear transformation, convenience of optimization, etc.

However, most of the works mentioned above deal with problems where the plant model is assumed to be linear and has no uncertain parameters. This assumption seems unrealistic for many real-life problems. Recently, an extended version of linear SMF is proposed by Scholte and Campbell to implement nonlinear estimation [12]. This extended set-membership filter (ESMF) linearizes the nonlinear dynamics about the current estimate in a manner that is similar to extended Kalman filter [13]. The remaining terms of linearization are then bounded and incorporated into the iterative algorithm as additions to noises. This allows the solution to be obtained for nonlinear systems so long as the boundedness of the nonlinear term is implemented.

Owing to its complex form, the ESMF algorithm has a problem of numerical instability. The direct causes include the fact that the envelop matrix is not positive definite, and the gain computation involves inversion of singular or nearly singular temporary matrix. All of these causes will lead to huge errors or even divergence. Moreover, the large computation load makes it difficult for real-time applications.

Motivated by the above observations, a UD factorization-based SMF, in which the envelope matrices are updated in their UD factorization form, is proposed in this paper to improve the numerical stability. Also, the measurements are updated sequentially and selectively to reinforce the stability and reduce the computation load. Furthermore, the filter parameters are determined by a proposed adaptive strategy to further reduce the computation complexity and achieve sub-optimal estimation.

The rest of the paper is organized as follows. The ESMF algorithm is introduced in Section 2. In Section 3, the UD factorization ESMF combined with the sequential and selective measurement update strategy is derived based on ESMF. Then in Section 4, an adaptive selection of filter parameters for sub-optimal estimation is introduced. In Section 5, comparative simulation results are presented. Finally in Section 6, some conclusions are drawn out.

2. EXTENDED SET-MEMBERSHIP FILTER

The main concept of ESMF is to cast the nonlinear dynamics in a way that allows the algorithm to take advantage of the linear SMF framework. Specifically, the nonlinear dynamics are linearized about the current estimation; the remaining terms are then proven bounded using interval analysis.
and are considered together with the process or measurement noises. Then a linear SMF algorithm is used and an ellipsoid estimation set is achieved. The true value is guaranteed to lie in the ellipsoid.

The nonlinear system considered here is in a form as

\[
x_{k+1} = f(x_k) + w_k \tag{1}
\]
\[
y_{k+1} = h(x_{k+1}) + v_{k+1} \tag{2}
\]

where \(x_k \in \mathbb{R}^n\) and \(y_{k+1} \in \mathbb{R}^m\) are the state and measurement vectors, respectively, \(f(\cdot)\) and \(h(\cdot)\) are nonlinear \(C^2\) functions, \(w_k \in \mathbb{R}^n\) and \(v_{k+1} \in \mathbb{R}^m\) are the process and measurement noise vectors, respectively, and they meet the conditions

\[
w_k \in E(0, Q_k), \quad v_{k+1} \in E(0, R_{k+1}) \tag{3}
\]

where \(E(a, P)\) stands for an ellipsoid set as

\[
E(a, P) = \left\{ x \in \mathbb{R}^n : (x - a)^T P^{-1} (x - a) \leq 1 \right\} \tag{4}
\]

where \(a\) is the center, \(x\) is any point within the ellipsoid, and \(P\) is a positive-definite matrix.

The initial estimation ellipsoid is defined as

\[
x_0 \in E(\hat{x}_{0,0}, P_{0,0}) \tag{5}
\]

where \(\hat{x}_{0,0}\) and \(P_{0,0}\) stand for the initial estimation of ellipsoid center and envelope matrix, respectively. At time \(k\), the estimated ellipsoid is defined as

\[
x_k \in E(\hat{x}_{k,k}, P_{k,k}) \tag{6}
\]

Then, at time \(k+1, k=0, 1, 2, \ldots\), the ESMF algorithm [12] can be summarized as follows:

1. Calculate the state interval based on the ellipsoid extrema:

\[
\tilde{X}_k^i = \left[ \hat{x}_{k,k}^i \pm \sqrt{P_{k,k}^{i,i}} \hat{x}_{k,k}^i \pm \sqrt{P_{k,k}^{i,i}} \right], \quad i = 1, \ldots, n \tag{7}
\]

where the superscript \(i, j\) stands for the \((i, j)\) element of a matrix.

2. Find the maximum interval for the Lagrange remainder using the interval analysis: Expanding the process function, Equation (1) (using one state for example) yields

\[
x_{k+1} = f(\hat{x}_{k,k}) + (\nabla \{ f(x_k) \})_{x_k=\hat{x}_{k,k}} (x_k - \hat{x}_{k,k}) + \frac{1}{2} (x_k - \hat{x}_{k,k})^T \frac{\partial^2 f(\hat{x})}{\partial x^2} (x_k - \hat{x}_{k,k}) \tag{8}
\]

where \(\nabla f(x)\) is the gradient of the function, \(\partial^2 f/\partial x^2\) is the second derivative of the function and \(\hat{\xi}\) can take any value on an interval \(\tilde{X}_k\) where \((x_k - \hat{x}_{k,k})\) is defined. Then the interval of the Lagrange remainder is defined as

\[
R_2(x_k - \hat{x}_{k,k}, \tilde{X}_k) = \frac{1}{2} (x_k - \hat{x}_{k,k})^T \frac{\partial^2 f(\hat{X}_k)}{\partial x^2} (x_k - \hat{x}_{k,k}) \tag{9}
\]
Then the ellipsoid of linearization error is defined as

\[ E \]

where \( q \) is the direct sum of the predicted ellipsoid of the linearized model \( E \) incorporated measurement noise:

\[ \hat{E} \]

which is the intersection of the predicted ellipsoid \( E \) is now defined as \( \bar{X}_k = (\bar{X}^1_k \bar{X}^2_k \ldots \bar{X}^n_k)^T \).

3. Calculate the ellipsoid bounding the linearization error:

\[
[\hat{Q}_k]^i_j = 2(X_{R_k})^2, \quad [\hat{Q}_k]^i_j = 0 \quad (i \neq j)
\]

Then the ellipsoid of linearization error is defined as \( E(0, \hat{Q}_k) \).

4. Calculate the final process/linearization error bound:

\[
\hat{w}_k \in E(0, \hat{Q}_k) \supset E(0, Q_k) \oplus E(0, \hat{Q}_k) = \{ x : x = x_1 + x_2, x_1 \in E(0, Q_k), x_2 \in E(0, \hat{Q}_k) \}
\]

where \( \hat{w}_k \) is the sum of the linearization error and \( w_k \), and

\[
\hat{Q}_k = \frac{\hat{Q}_k}{1-\beta_k}, \quad \beta_k \in (0, 1)
\]

where \( \beta_k \) is a filter parameter to be chosen to minimize the ellipsoid \( E(0, \hat{Q}_k) \). The observation model should be dealt with in the same way as computing \( \hat{w}_k \) in Steps (2)–(4) to attain the incorporated measurement noise:

\[
\hat{v}_{k+1} \in E(0, \hat{R}_{k+1})
\]

5. Calculate the state prediction ellipsoid \( E(\hat{x}_{k+1,k}, P_{k+1,k}) \) using the linear SMF filter, which is the direct sum of the predicted ellipsoid of the linearized model \( E(f(\hat{x}_{k,k}), A_k P_{k,k} A_k^T) \) and the virtual noise ellipsoid \( E(0, \hat{Q}_k) \):

\[
\hat{x}_{k+1,k} = f(\hat{x}_{k,k})
\]

\[
P_{k+1,k} = \Phi_k \frac{P_{k,k}}{1-\beta_k} \Phi_k^T + \frac{\hat{Q}_k}{\beta_k}, \quad \beta_k \in (0, 1)
\]

where \( \beta_k \) is a filter parameter to be chosen to minimize the ellipsoid \( E(\hat{x}_{k+1,k}, P_{k+1,k}) \).

6. Calculate the updated state ellipsoid \( E_{k+1} = E(\hat{x}_{k+1,k+1}, P_{k+1,k+1}) \) using the linear SMF, which is the intersection of the predicted ellipsoid \( E(\hat{x}_{k+1,k}, P_{k+1,k}) \) and the observation set \( S_y = \{ x : (y_{k+1} - h(x))^T \hat{R}_{k+1}(y_{k+1} - h(x)) \leq 1 \} \):

\[
W_k = H_{k+1} \frac{P_{k+1,k}}{1-\rho_k} H_{k+1}^T + \frac{\hat{R}_{k+1}}{\rho_k}, \quad \rho_k \in (0, 1)
\]

\[
K = \frac{P_{k+1,k}}{1-\rho_k} H_{k+1}^T W_k^{-1}
\]
\[
\hat{x}_{k+1,k+1} = \hat{x}_{k+1,k} + K[y_{k+1} - h(\hat{x}_{k+1,k})]
\]
\[
\tilde{P}_{k+1,k+1} = \frac{P_{k+1,k}}{1 - \rho_k} - \frac{P_{k+1,k}H_h^{-1}W_k^{-1}H_h}{1 - \rho_k}
\]
\[
P_{k+1,k+1} = \delta_k \tilde{P}_{k+1,k+1}
\]

where \(\rho_k\) is a filter parameter to be chosen to minimize the ellipsoid \(E(\hat{x}_{k+1,k+1}, P_{k+1,k+1})\), and \(\delta_k\) is defined as
\[
\delta_k = 1 - [y_{k+1} - h(\hat{x}_{k+1,k})]^T W_k^{-1} [y_{k+1} - h(\hat{x}_{k+1,k})]
\]

The linearization of the nonlinear system is
\[
\Phi_k = \left. \frac{\partial f(x_k)}{\partial x} \right|_{x_k = \hat{x}_{k,k}}, \quad H_{k+1} = \left. \frac{\partial h(x_k)}{\partial x} \right|_{x_k = \hat{x}_{k+1,k}}
\]

Special attention should be given to the following two problems. First, the values of the three parameters \(\beta_{Q_k}, \beta_k\) and \(\rho_k\) can be optimized at each time step in an effort to find the smallest bounding ellipsoid. Some criteria such as minimizing the volume of ellipsoid or minimizing the trace of the bounding matrix can be used. Second, when \(\delta_k \leq 0\) in Equation (22), the uncertainty ellipsoid of Equation (21) is not defined. This will occur when the bound assumption of Equation (3) is not satisfied. That is, the practical state or noise is beyond the assumed bound. Thus, it provides an indication of the health of the algorithm.

3. UD FACTORIZATION-BASED ESMF (UD-ESMF)

The ESMF algorithm above suffers from several problems in practical usage. One of the major problems is numerical instability caused by the computer round-off error, which may seriously degrade the performance of the filter or even make it diverge. Another problem is that the selection of three parameters of ESMF, namely, \(\beta_{Q_k}, \beta_k\) and \(\rho_k\), may make the instability problem of the filter even worse in following two ways: (a) the envelope matrix \(P_{k,k}\) is no longer positive definite; and (b) the gain computation of Equation (18) gets involved in inversion of singular or nearly singular temporary matrix \(W_k\).

To solve this numerical instability problem, a UD factorization form is introduced into the original ESMF algorithm to ensure that the envelope matrix \(P\) is symmetrical and positive definite. Using the UD factorization technique [14], the envelope matrix \(P\) is factored such that \(P = UDU^T\), where \(U\) is a unit upper triangular matrix and \(D\) is a diagonal matrix. In each step the factors will be updated instead of the envelope matrix \(P\) itself. As a result, the envelope matrix \(P\) will always be guaranteed to be positive definite, and the round-off errors caused by direct computation of \(P\) is avoided, leading to improved numerical stability.

In addition, a sequential update strategy is adopted for the measurement vector to transform the inversion of the matrix \(W_k\) to the inverse computation of a scalar, which avoids large errors associated with singular matrix. The proposed algorithm is derived below.

Assume that the matrix \(P_{k,k}\) at time \(k\) is positive definite with the UD factorization form of \(P_{k,k} = U_{k,k}D_{k,k}U_{k,k}^T\), where \(U_{k,k}\) is a unit upper triangular matrix and \(D_{k,k}\) is a diagonal matrix. The noises are assumed to be uncorrelated for now, that is, the matrices \(Q_k\) and \(R_{k+1}\) are diagonal.
The situation where $Q_k$ and $R_{k+1}$ are not diagonal is considered later in this section. Thus, the envelope matrices $\hat{Q}_k$ and $\hat{R}_{k+1}$ of the virtual noises $\hat{w}_k$ and $\hat{v}_{k+1}$ are also diagonal from Equations (13) and (14).

For the prediction stage, define

$$S_{k+1,k} = [\Phi_k U_{k,k} \ I]$$

(24)

$$\hat{D}_{k+1,k} = \begin{pmatrix} D_{k,k} & 0 \\ 1 - \beta_k & 0 \\ 0 & \hat{Q}_k / \beta_k \end{pmatrix}$$

(25)

From (15) and (16), we have

$$\hat{x}_{k+1,k} = f(\hat{x}_{k,k})$$

(26)

$$P_{k+1,k} = U_{k+1,k} D_{k+1,k} U_{k+1,k}^T = \Phi_k \frac{P_{k,k}}{1 - \beta_k} \Phi_k^T + \hat{Q}_k / \beta_k$$

$$= \Phi_k \frac{U_{k,k} D_{k,k} U_{k,k}^T}{1 - \beta_k} \Phi_k^T + \hat{Q}_k / \beta_k$$

$$= S_{k+1,k} \hat{D}_{k+1,k} \hat{S}_{k+1,k}^T$$

(27)

where $U_{k+1,k}$ and $D_{k+1,k}$ can be computed by the modified weighted Gram–Schmidt orthogonalization [14].

For the update strategy, the measurement vector is sequentially processed from component to component. First, define

$$H_{k+1} = [H_{k+1(1)}^T \ H_{k+1(2)}^T \ \cdots \ H_{k+1(m)}^T]^T$$

(28)

where $H_{k+1(i)}$, $i = 1, \ldots, m$, are the columns of the Jacobian matrix $H_{k+1}$. $\hat{R}_{k+1(i)}$, $i = 1, \ldots, m$, are the diagonal elements of the matrix $\hat{R}_{k+1}$. $y_{k+1} = (y_{k+1(1)} \ \cdots \ y_{k+1(m)})^T$ is the measurement vector, $h(\cdot) = (h_1(\cdot) \ \cdots \ h_m(\cdot))^T$ is the observation function.

Given the initial condition of the sequential update strategy

$$U_{k+1,k+1}^0 = U_{k+1,k}, \quad D_{k+1,k+1}^0 = D_{k+1,k}, \quad \delta_k^0 = 1$$

(29)

and assuming that the UD factors of $\hat{P}_{k+1,k+1}$ obtained by the $(i - 1)$th measurement component are $\hat{U}_{k+1,k+1}^i$ and $\hat{D}_{k+1,k+1}^i$, we define the following intermediate variables in order to compute the UD factors with the $i$th measurement component:

$$x = H_{k+1(i)} \hat{U}_{k+1,k+1}^i \frac{\hat{D}_{k+1,k+1}^i}{1 - P_{k(i)}} (\hat{U}_{k+1,k+1}^i)^T H_{k+1(i)}^T + \frac{\hat{R}_{k+1(i)}}{P_{k(i)}}$$

(30)

$$s = \frac{D_{k+1,k+1}^i}{1 - P_{k(i)}} (\hat{U}_{k+1,k+1}^i)^T H_{k+1(i)}^T$$

(31)
where \( z \) is a simplified form of \( W_k \) for sequential update, which in general is the envelope matrix of the ellipsoid where each measurement component \( y_{k+1(i)} \) lies. It should be noticed that, since \( y_{k+1(i)} \) is a scalar, the ellipsoid reduces to an interval, and \( z \) is a scalar that defines the bound of the interval.

From (17) to (22), we have

\[
\hat{x}_{k+1,k+1}^i = \hat{x}_{k+1,k+1}^{i-1} + \frac{1}{\alpha} \bar{U}_{k+1,k+1}^{i-1} D_{k+1,k+1}^{i-1}(\bar{U}_{k+1,k+1}^{i-1})^T H_{k+1(i)}^i (y_{k+1(i)} - h_i(\hat{x}_{k+1,k+1}^{i-1}))
\]

\[ (32) \]

\[
\bar{P}_{k+1,k+1}^i = \bar{U}_{k+1,k+1}^{i-1} D_{k+1,k+1}^{i-1}(\bar{U}_{k+1,k+1}^{i-1})^T
\]

\[ (33) \]

\[
\delta_k^i = \frac{1}{\alpha} (y_{k+1(i)} - h_i(\hat{x}_{k+1,k+1}^{i-1}))^2
\]

\[ (34) \]

From Equation (23), the UD factors of \( \bar{P}_{k+1,k+1}^i \) can be obtained:

\[
\bar{U}_{k+1,k+1}^i = \bar{U}_{k+1,k+1}^{i-1} \bar{U}, \quad \bar{D}_{k+1,k+1}^i = \bar{D}
\]

\[ (35) \]

where \( \bar{U} \) and \( \bar{D} \) are the UD factors of the matrix

\[
\frac{D_{k+1,k+1}^{i-1}}{1 - \rho_k(i)} - \frac{1}{\alpha} s s^T
\]

Finally, we have

\[
\hat{x}_{k+1,k+1} = z_{k+1,k+1}^m
\]

\[ (36) \]

\[
U_{k+1,k+1} = \bar{U}_{k+1,k+1}^m, \quad D_{k+1,k+1} = \bar{D}_{k+1,k+1}^m
\]

\[ (37) \]

When \( \hat{Q}_k \) is not diagonal, we can assume that its UD factors are \( \hat{Q}_k = U Q D Q U_Q^T \). Then Equations (24) and (25) will be replaced by

\[
S_{k+1,k} = \left[ \Phi_k U_{k,k} \ U_Q \right]
\]

\[ (38) \]

\[
\tilde{D}_{k+1,k} = \begin{pmatrix} \frac{D_{k,k}}{1 - \beta_k} & 0 \\ 0 & \frac{D_Q}{\beta_k} \end{pmatrix}
\]

\[ (39) \]
When $R_{k+1}$ is not diagonal, we can assume that its UD factors are $R_{k+1} = U_R D_R U_R^T$, and the following transform is necessary:

\[ z_{k+1} = U^{-1}_R y_{k+1} \]
\[ = U^{-1}_R h(x_{k+1}) + U^{-1}_R v_{k+1} \]
\[ = h_a(x_{k+1}) + v_{a,k+1} \]  \hspace{1cm} (40)

where $v_{a,k+1} \in E(0, D_R)$, and $h_a(\cdot)$ is a nonlinear function to replace $h(\cdot)$.

The iterative steps of the UD-ESMF can be summarized as follows:

Steps (1)–(4): Same as those listed in Section 2.

Step (5): Equations (15) and (16) are replaced by Equations (24)–(27).

Step (6): Equations (17)–(22) are replaced by Equations (29)–(35) for each measurement component.

By comparing the above iterative steps with those of normal ESMF outlined in Section 2, we can see that in each step, the envelope matrices $P_{k+1,k}$ and $P_{k+1,k+1}$ are not computed explicitly now but replaced by their UD factors. Because the special configuration of the UD factors can be maintained in each step, that is, $U$ factor is a unit upper triangular matrix and $D$ factor is a diagonal matrix, from the theories of UD factorization, the envelope matrices that can be computed by their UD factors can always be guaranteed to be positive definite according to (27) and (33). Besides, the inversion of the matrix $W_k$ is transformed to the inverse computation of a scalar $\alpha$ by a sequential update strategy, which avoids large computation errors associated with singular matrices. As a result, the numerical stability is therefore greatly improved.

It should also be noticed that the sequential update strategy in step (6) of UD-ESMF can also reduce the computation load, especially when the dimension $m$ of observations is large. A significant portion of the overall computation load is to calculate the inverse of the matrix $W_k$, which is $O(m^3)$. With the sequential update strategy, it is now transformed to the division of a scalar $\alpha$ and the computation load is reduced to $O(m)$. Other matrix computations in the update stage are mostly transformed to scalar computation, which also counteracts the extra computation load produced along with the sequential strategy such as the orthogonalization computation of UD factors. In addition, with the sequential update strategy, the filter can work correctly even when some observations are unavailable due to sensor fault.

Besides, the computation load of ESMF can be further reduced by some other methods such as selective update strategy proposed below. From step (6) in Section 2, the updated state ellipsoid $E_{k+1} = E(\hat{x}_{k+1,k+1}, P_{k+1,k+1})$ is the intersection of the predicted ellipsoid $E(\hat{x}_{k+1,k}, P_{k+1,k})$ and the observation set $S_y$, that is,

\[ E_{k+1} = E(\hat{x}_{k+1,k+1}, P_{k+1,k+1}) = E(\hat{x}_{k+1,k}, P_{k+1,k}) \cap S_y \]  \hspace{1cm} (41)

thus,

\[ E_{k+1} = E(\hat{x}_{k+1,k+1}, P_{k+1,k+1}) \subset E(\hat{x}_{k+1,k}, P_{k+1,k}) \]  \hspace{1cm} (42)

In the case of

\[ E(\hat{x}_{k+1,k}, P_{k+1,k}) \subset S_y \]  \hspace{1cm} (43)
we have
\[ E_{k+1} = E(\hat{x}_{k+1,k}, P_{k+1,k}) \]  (44)

Equation (44) indicates that the update step at time \( k + 1 \) can be saved if this condition is met. However, the condition of Equation (43) cannot be checked easily in practice. In order to give a convenient criterion to realize the selective update, some looser criteria have to be made. There are some studies focusing on the issue of selective update criterion, such as by minimizing the upper bound of the performance index for linear SMF [6] and minimizing the volume of the intersection ellipsoid [15]. Here we propose a simpler criterion:
\[
(y_{k+1} - h(\hat{x}_{k+1,k}))^T \hat{R}_{k+1}^{-1} (y_{k+1} - h(\hat{x}_{k+1,k})) \leq \eta, \quad \eta \in (0, 1]
\]  (45)

where \( \eta \) is a parameter to be selected. It has to be selected carefully according to the initial assumption. Normally, larger \( \eta \) can increase the chance of not updating but will lead to larger estimation bounds. Namely, there is a trade-off between ‘update saving’ and conservative estimation. The simulation in Section 5 shows that, when the over estimation of noise is small, such as below 5%, we can choose \( \eta = 1 \).

Therefore, under the condition of Equation (45), the center of the predicted ellipsoid \( E(\hat{x}_{k+1,k}, P_{k+1,k}) \) lies in the set \( S_y \), and only the prediction step is required. The update step can be saved, while ESMF still guarantees the bound estimation.

4. ADAPTIVE SELECTION OF THE FILTER PARAMETERS

In this section, an adaptive selection strategy of the three parameters \( \beta_Q, \beta_k \) and \( \rho_k \) will be proposed to improve the performance of the filter.

First considering the parameters \( \beta_Q \) and \( \beta_k \), an optimal direct sum of two ellipsoids is proposed below. Without loss of generality, assume that the two ellipsoids are defined as \( E(a_1, P_1) \) and \( E(a_2, P_2) \), while the covered ellipsoid of their direct sum is \( E(a, P) \). Then \( a \) and \( P \) can be selected as
\[
a = \beta a_1 + (1 - \beta) a_2, \quad P = \frac{P_1}{1 - \beta} + \frac{P_2}{\beta}, \quad \beta \in (0, 1)
\]  (46)

The goal of optimization is to choose a proper \( \beta \) to make the covered ellipsoid satisfy some criteria such as minimizing \( \text{det}(P) \) or minimizing the trace \( \text{Tr}(P) \) [15, 16]. Here, the latter is chosen for the simpler form and better robustness. Define
\[
\beta = \arg \min_{\beta \in (0, 1)} \text{Tr}(P)
\]  (47)

According to [16], the optimal \( \beta \) can be selected as
\[
\beta = \frac{\sqrt{\text{Tr}(P_2)}}{\sqrt{\text{Tr}(P_1)} + \sqrt{\text{Tr}(P_2)}}
\]  (48)

The parameters \( \beta_Q \) and \( \beta_k \) can be updated adaptively according to Equation (48). For \( \beta_Q \), we compute the direct sum of two ellipsoids \( E(0, Q_k) \) and \( E(0, \hat{Q}_k) \). And for \( \beta_k \), the direct sum of two ellipsoids \( E(0, \hat{Q}_k) \) and \( E(f(\hat{x}_{k,k}), A_k P_{k,k} A_k^T) \) is considered.
To determine the value of the parameter \( \rho_k \), the matrix \( P_{k+1,k+1} \) is computed as

\[
P_{k+1,k+1} = \delta_k \tilde{P}_{k+1,k+1} = \delta_k \left( \frac{P_{k+1,k+1}}{1-\rho_k} - \frac{P_{k+1,k}}{1-\rho_k} H_{k+1}^T H_{k+1}^{-1} \right)
\]

where

\[
\delta_k = 1 - \|y_{k+1} - h(\hat{x}_{k+1,k})\|_W^{-1} [y_{k+1} - h(\hat{x}_{k+1,k})]
\]

Because of the complex form of the ellipsoid \( S_y \), it is difficult to optimize the parameter \( \rho_k \) by minimizing the volume or trace of the ellipsoid \( \det(P_{k+1,k+1}), \ln \det(P_{k+1,k+1}) \) or \( \text{Tr}(P_{k+1,k+1}) \). Some sub-optimal criterion has to be used. Here the criterion for minimizing the upper bound of the performance index \( \delta_k \) is selected to adaptively update \( \rho_k \) according to Lemma 1:

\[
\rho_k = \arg \min \sup_{\rho_k \in (0,1)} (\delta_k)
\]

**Lemma 1**

By minimizing \( \delta_k \) for computing the optimal observation update factor \( \rho_k \), there exist two positive constants \( c_1 \) and \( c_2 \), which ensure that the following relations hold [6]:

\[
\det(P_{k,k}) \leq c_1 (\delta_k)^n, \quad \text{Tr}(P_{k,k}) \leq c_2 \delta_k
\]

Lemma 1 reveals that minimizing the upper bound of \( \delta_k \) results in the minimization of the upper bounds on interpretable ellipsoidal volume and trace. These results show that, although using \( \delta_k \) minimization, the determinant and trace are not minimized, they are upper bounded by minimized upper bounds in each step. Thus, it is a sub-optimal but efficient method, which can greatly decrease the computation complexity of the original algorithm.

The upper bound of \( \delta_k \) is

\[
\tilde{\delta}_k = 1 - \frac{\|y_{k+1} - h(\hat{x}_{k+1,k})\|^2}{p_m + r_m}
\]

where \( p_m \) and \( r_m \) are maximum singular values of the matrix \( H_{k+1} P_{k+1,k} H_{k+1}^T \) and \( \hat{R}_{k+1} \), respectively. That is, \( p_m = \lambda_{\text{max}}(H_{k+1} P_{k+1,k} H_{k+1}^T) \) and \( r_m = \lambda_{\text{max}}(\hat{R}_{k+1}) \). The minimal upper bound can be computed as

\[
\tilde{\delta}_{k,\text{min}} = 1 - \frac{\|y_{k+1} - h(\hat{x}_{k+1,k})\|^2}{(\sqrt{p_m} + \sqrt{r_m})^2}
\]

when

\[
\rho_k = \rho^*_k = \frac{\sqrt{r_m}}{\sqrt{p_m} + \sqrt{r_m}} \in (0,1)
\]
Equation (55) is an adaptive update algorithm of $\rho_k$. For the sequentially updated UD factorization-based algorithm, a simpler result can be derived by minimizing $\alpha$ in Equation (30) for each component of measurements according to Equation (34). The expression of $\alpha$ can be transformed to

$$\alpha = H_{k+1(i)} \hat{U}_{k+1,k+1}^{-1} \frac{\hat{D}_{k+1,k+1}^{-1}}{1 - \rho_k} (\hat{U}_{k+1,k+1}^{-1})^T H_{k+1(i)}^T + \frac{\hat{R}_{k+1(i)}}{\rho_k} = \frac{d_1}{1 - \rho_k} + \frac{d_2}{\rho_k}$$  

(56)

where

$$d_1 = H_{k+1(i)} \hat{U}_{k+1,k+1}^{-1} \hat{D}_{k+1,k+1}^{-1} (\hat{U}_{k+1,k+1}^{-1})^T H_{k+1(i)}^T$$

and

$$d_2 = \hat{R}_{k+1(i)}$$  

(57)

are positive constants. Minimizing $\alpha$ yields

$$\alpha_{\text{min}} = (\sqrt{d_1} + \sqrt{d_2})^2$$  

(58)

when

$$\rho_k = \rho_k^* = \frac{\sqrt{d_2}}{\sqrt{d_1} + \sqrt{d_2}} \in (0, 1)$$  

(59)

The above equation will replace Equation (55) for sequential update algorithm. Thus, the criteria described in Sections 3 and 4 are incorporated into the final UD factorization-based adaptive filter (AESMF).

5. SIMULATION

The proposed AESMF is tested by simulation to estimate the motion states and slip parameters of a tracked mobile robot as shown in Figure 1. A general kinematics model of the tracked vehicle with consideration of slip is developed in the following.

In order to describe the motion of the tracked vehicle, define a fixed reference frame $F(X_w, Y_w)$ and a moving frame $f(X_m, Y_m)$ attached to the vehicle body, with origin at the geometric center of the vehicle. The motion of the vehicle is composed of the translation velocity $V_e = (V_x, V_y)^T$ and the rotation velocity $\omega_c = \psi/\text{dr}$, where $V_x$ and $V_y$ stand for the projection of $V_e$ on the $x$- and $y$-axes of frame $f$, and $\psi$ is the yaw angle. $V_e$ is termed as the sideslip velocity. Because of the sideslip, the instantaneous center of rotation (ICR) shifts forward of the vehicle centroid by $d$. The angle between the lines of $O_m$-ICR and $O'$-ICR is defined as $\alpha$, where $O'$-ICR is perpendicular to the axis $X_m$.

Then the discrete-time process model is

$$x_{k+1} = \begin{pmatrix} X_{k+1} \\ Y_{k+1} \\ \psi_{k+1} \end{pmatrix} = \begin{pmatrix} X_k + \frac{r \omega_{L,k}(1-i_{L,k}) + r \omega_{R,k}(1-i_{R,k})}{2} \Delta T (\cos \psi_k + \sigma_k \sin \psi_k) \\ Y_k + \frac{r \omega_{L,k}(1-i_{L,k}) + r \omega_{R,k}(1-i_{R,k})}{2} \Delta T (\sin \psi_k - \sigma_k \cos \psi_k) \\ \psi_k + \frac{-r \omega_{L,k}(1-i_{L,k}) + r \omega_{R,k}(1-i_{R,k})}{b} \Delta T \end{pmatrix} + w_k$$  

(60)

where $b$ is the distance between the midpoints of the two tracks; $r$ is the radius of the wheels, which drive the tracks; $\Delta T$ is sampling interval; $\omega_{L,k}$ and $\omega_{R,k}$ are the angular velocities of the
left and right wheels; $i_{L,k}$ and $i_{R,k}$ are the slip ratios of the left and right tracks; and $\sigma_k$ is the sideslip angle coefficient.

In the simulations, $X$, $Y$ and $\psi$ are assumed as measurements. Thus, the observation model is

$$y_{k+1} = x_{k+1} + v_{k+1}$$

The goal of the simulation is to estimate the state $(X, Y, \psi)^T$ and the three slip parameters $i_{L,k}, i_{R,k}$ and $\sigma_k$ simultaneously. To solve this joint nonlinear estimation problem, a new augmented state vector is defined as the combination of state and parameter vector, $x_{a,k} = (X_k, Y_k, \psi_k, i_{L,k}, i_{R,k}, \sigma_k)^T$. The slip parameters are often unknown in practice; hence, they are modeled as follows:

$$p_{k+1} = p_k + w_k$$

where $p_k \in R^p$ is the parameter vector, $w_k \in R^p$ is the additive process noise that drives the model.

The constants in the model of Equation (60) are chosen as $b = 0.65 m$; $r = 0.35 m$; $\Delta T = 0.1s$. The duration of the simulation is 20s. The process and measurement noises are both 5% of nominal value and uniformly distributed. In order to demonstrate the tracking performance, abrupt changes are simulated to occur in the three slip parameters at time $t = 10s$. In the simulation, the normal ESMF and the proposed AESMF are compared in terms of stability, performance and computation load.

The stability improvement of the AESMF over ESMF is demonstrated clearly in Figures 2–4, which are the comparisons between an unstable case of ESMF and the proposed AESMF in parameter estimation, bound estimation and performance index, respectively. In each figure, the results of ESMF are in (a), while those of AESMF are in (b), and the $x$-axis is the simulation time in seconds. The dashed line indicates the true value of states or parameters; the dash-dotted line is the estimation of normal ESMF; and the solid line is the result of the proposed AESMF.
Figure 2. Comparison between unstable ESMF and AESMF on parameter estimation: (a) parameter estimation by normal ESMF and (b) parameter estimation by proposed AESMF.
Figure 3. Comparison between unstable ESMF and AESMF on parameter bound estimation: (a) bound estimation by normal ESMF and (b) bound estimation by proposed AESMF.
Figure 4. Comparison between unstable ESMF and AESMF on performance index: (a) performance index by normal ESMF and (b) performance index by proposed AESMF.
Figure 2(a) shows that, at \( t = 2s \), parameter estimations by the unstable ESMF diverge, and the parameters remain with some values and never change again. This is because the temporal matrix \( W_k \) is nearly singular and the inversion computation causes large errors. Besides, the filter will also diverge if the envelope matrix \( P_{k,k} \) is no longer positive definite or negative elements appear in the diagonal. It can be seen that the index performance at \( t = 2s \) in Figure 4(a) is still positive, which means that the initial noise assumption is still satisfied; hence, in Figure 3(a) the bound estimation does not diverge at \( t = 2s \). Figure 2(b) shows that AESMF remains stable at \( t = 2s \), which indicates that the UD factorization successfully solves the numerical instability problem of the normal ESMF. When abrupt changes occur in the three parameters at \( t = 10s \), the performance index \( \hat{\delta}_k \) in Figure 4(a) becomes and remains negative, which means that the initial noise assumption is no longer satisfied. As a result, the bounds in Figure 3(a) diverge rapidly to \( 10^{41} \) and ESMF cannot generate the correct estimation at all. The proposed AESMF, on the other hand, can handle this problem very well by its UD factorization mechanism. Figure 4(b) shows that the performance index returns to positive value after a short period of adaptive modification, although it does become negative due to the abrupt changes that occurred at \( t = 10s \). After that, the estimation of AESMF can track its true value and the estimated bound covers the true value closely.

Besides the improvement in stability with the proposed AESMF, some other aspects are also compared between the stable case of ESMF and AESMF. The state and parameter estimations, as well as the performance index of a stable ESMF and AESMF, are presented in Figure 5, where the dashed line is the true value, the dash-dotted line is the result of a stable case of ESMF and the solid line is the result of AESMF.

Three aspects are compared between stable ESMF and AESMF: (a) accuracy of both state and parameter estimations; (b) bound estimation of parameters; and (c) CPU time cost. To do this, the prediction accuracies of state and parameter are defined by the following root mean square criteria:

\[
V_1 = \sqrt{\frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N_x} (x_i(t) - \hat{x}_i(t))^2} \tag{63}
\]

\[
V_2 = \sqrt{\frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N_y} (p_i(t) - \hat{p}_i(t))^2} \tag{64}
\]

where \( x_i(t) \) and \( \hat{x}_i(t) \) are the true and estimated states, respectively, \( p_i(t) \) and \( \hat{p}_i(t) \) are true and estimated parameters, respectively, and \( N_x \) and \( N_y \) are the relative dimensions.

The quantified performance of ESMF and AESMF is listed in Table I. All the data in Table I are the average of 20 simulations. From Figures 5(a) and (b) and Table I, we can see that in the case of a stable ESMF, the prediction accuracies of the two filters are very similar. This indicates that the UD factorization combined with the sequential and selective measurement update strategy as well as the adaptive mechanism does not degrade the estimation accuracy of ESMF.

On the other hand, the bounds estimated by AESMF in Figure 5(c) are smaller than those by ESMF, which means that the estimated ellipsoids are smaller in AESMF. This is due to the fact that the three parameters are kept constant in ESMF, while in AESMF they are updated by the adaptive mechanism.

Figure 5. Comparison between stable ESMF and AESMF: (a) state estimation; (b) parameter estimation; (c) parameter bound estimation; and (d) performance index.
Table I. Performance comparison of ESMF and AESMF.

<table>
<thead>
<tr>
<th></th>
<th>State accuracy $V_1$</th>
<th>Parameter accuracy $V_2$</th>
<th>Computation load $t$ (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESMF</td>
<td>1.001e−4</td>
<td>0.031674</td>
<td>42.71</td>
</tr>
<tr>
<td>AESMF</td>
<td>9.7953e−5</td>
<td>0.030627</td>
<td>33.48</td>
</tr>
</tbody>
</table>

As for the CPU time cost, the calculation of each step running by Matlab 6.5 on Pentium-IV PC needs 33.48 ms for AESMF, about 21.6% faster than that of ESMF, which is about 42.71 ms. This is benefited from the sequential and selective measurement update strategy proposed in this paper.

6. CONCLUSION

In this paper, a UD factorization-based adaptive set-membership filter is proposed for the joint state and parameter estimation of nonlinear systems. The normal ESMF is enhanced by UD factorization, sequential and selective measurement updating and adaptive selection of the filter parameters. This improves the numeric stability, reduces computation complexity and realizes the sub-optimal bound-guaranteed performance of normal ESMF. The proposed algorithms are presented and analyzed in detail, and extensive simulations are carried out to perform joint state and parameter estimation of a tracked mobile robot and to make comparisons between the normal ESMF and the proposed AESMF. It has been demonstrated that the proposed AESMF successfully solves the numeric instability problem, realizes the sub-optimal estimation and also improves the real-time applicability without losing estimation accuracy of the normal ESMF.

REFERENCES


