Self-tuning Fusion Kalman Smoother for Multisensor Multi-Channel ARMA Signals and Its Convergence

Guili Tao  
Department of Automation, Heilongjiang University  
Harbin, China  
dengzilihju@yahoo.com.cn

Zili Deng  
Department of Automation, Heilongjiang University  
Harbin, China  
dzl@hlju.edu.cn

Abstract—For the multisensor multi-channel autoregressive moving average (ARMA) signals with white measurement noises and an AR colored measurement noise as common disturbance noises, a multi-stage information fusion identification method is presented when model parameters and noise variances are partially unknown. The local estimators of model parameters and noise variances are obtained by the multi-dimensional recursive instrumental variable (MRIV) algorithm, correlation method, and the Gevers-Wouters algorithm, and the fused estimators are obtained by taking the average of the local estimators. They have the consistency. Substituting them into the optimal fusion Kalman smoother, a self-tuning fusion Kalman smoother for multi-channel ARMA signals is presented. Applying the dynamic error system analysis (DESA) method, it is proved that the proposed self-tuning fusion Kalman smoother converges to the optimal fusion Kalman smoother in a realization, so that it has asymptotic optimality. A simulation example shows its effectiveness.

Keywords—Multisensor information fusion; identification; convergence analysis; self-tuning Kalman smoother

I. INTRODUCTION

In multisensor environment, it is very important how to combine the information from different sensors to obtain the fused estimation for the state of the target when these sensors measure the same target. As we all known, the existing information fusion Kalman filtering is mainly focused on the information fusion Kalman filtering with known model parameters and noise statistics. However, in many applications, the model parameters and/or noise variances are usually unknown. The filtering for the systems with unknown model parameters and/or noise variances is called self-tuning filtering [1]. Several self-tuning weighted fusion filters [2-4] were presented. And [5] presented the self-tuning decoupled fusion Kalman predictor. Their drawbacks are that only the noise variances are assumed to be unknown, while the model parameters are assumed to be known. In [6], self-tuning information fusion Wiener filter was proposed for the single-channel ARMA signals with unknown model parameters and noise variances. However, only a few results were reported for the multisensor multi-channel AR or ARMA signals [7-10]. Their drawbacks are that they can not only solve the fusion filter problem for the systems with noise measurement noises. And in [7], the proposed self-tuning weighted measurement fusion Wiener filter for the multisensor multichannel ARMA signal is based on the online identification of the autoregressive moving average (ARMA). But [11] only suit for the single-channel system with a colored measurement noise. In the existing results, the information fusion is mainly focused on filtering fusion for the multisensor systems with white measurement noises, but the smoothing fusion for the multi-channel signal systems with a colored measurement noise is seldom reported because of their complexity.

The convergence analysis method of the self-tuning Kalman or Wiener fuser has been proved in [2, 3], which is called dynamic error system analysis (DESA) method, and a new concept of convergence in a realization was presented in [2,3], which is weaker than the convergence with probability one. The convergence analysis method of the self-tuning Riccati equation and Lyapunov equation was proved in [5], which is called dynamic variance error system analysis (DVESA) method.

In this paper, for the multisensor multi-channel ARMA signal with a colored measurement noise, when the model parameters and noise variances are partially unknown, based on each sensor, the local estimators of the autoregressive (AR) parameters are obtained by the multi-dimensional recursive instrumental variable (MRIV) algorithm, and the local estimators of the moving average (MA) parameters and noise variances are obtained by the Gevers-Wouters algorithm. Finally, the consistent fused estimators are obtained by taking the average of the local estimators. Using the classical Kalman filtering method, the corresponding self-tuning Kalman smoother is presented. By the DESA method, the convergence of self-tuning information fusion Kalman smoother is proved by the DESA method, i.e., the self-tuning fusion Kalman smoother converges to the optimal fusion Kalman smoother in a realization, so it has asymptotic optimality. Compared with [7-10], the paper presents the information fusion Kalman smoother for multisensor multi-channel ARMA signals with colored measurement noise.
Consider the multisensor multi-channel ARMA signals with L-sensor
\[ A(q^{-1})s(t) = C(q^{-1})w(t) \]
where \( t \) is the discrete time, \( s(t) \in R^n \) is the signal, \( y_i(t) = s(t) + \eta_i(t) + v_i(t) \), \( i = 1, \ldots, L \)
and \( P(q^{-1}) \eta(t) = \xi(t-1) \)
where \( \eta(t) \in R^n \) is a colored noise. \( q^{-1} \) is the backward shift operator, \( a(q^{-1}), C(q^{-1}) \) and \( P(q^{-1}) \) are polynomial matrices have the form as
\[ X(q^{-1}) = X_0 + \ldots + X_n q^{-n} \]
where \( X_0 = I_n, I_n \) is the \( m \times m \) identity matrix, \( n_a \) is the order of polynomial matrix \( A(q^{-1}) \), \( n_s \) is the order of polynomial matrix \( C(q^{-1}) \), and \( n_p \) is the order of polynomial matrix \( P(q^{-1}) \).

**Assumption 1.** \( w(t) \), \( \xi(t) \) and \( v_i(t) \) (\( i = 1, \ldots, L \)) are independent white noises with zero mean and variances \( Q_w, Q_s \) and \( Q_v \), respectively.

**Assumption 2.** \( A(q^{-1}) \) and \( C(q^{-1}) \) are stable polynomials of \( q^{-1} \), and \( (A(q^{-1}),C(q^{-1})) \) are left coprime.

**Assumption 3.** \( A(q^{-1}), C(q^{-1}) \) and \( Q_w \) are unknown, but \( P(q^{-1}), Q_s \) and \( Q_v \) are known.

**Assumption 4** A realization of measurement stochastic process \( y_i(t) \) \((i = 1, 2, \ldots, L)\) is bounded for \( t \).

The problem is to find self-tuning fusion Kalman signal smoother \( \hat{x}_i(t+1|t) \) when the ARMA model parameters and noise variances are partially unknown.

Setting \( w(t) = w(t-1) \), yields that \( w(t) \) has the variance \( Q_w \). The ARMA signal (1) can be rewritten as
\[ A(q^{-1})s(t) = C(q^{-1})w(t) \]
where \( C(q^{-1}) = C_0 + C_1 q^{-1} \ldots + C_{n_c} q^{-n_c}, C_0 = 0, C_i = C_{i-1}, i = 1, 2, \ldots, n_c, n_c = n_c + 1 \)

The signal system (5) has the state space model with the companion form
\[ \alpha(t+1) = A_a \alpha(t) + C_w w(t) \]
\[ s(t) = H_a \alpha(t) \]
with definition \( n_{ac} = \max(n_a, n_c + 1) \), \( A_j = 0(j > n_c) \), \( C_j = 0(j > n_c) \).

Similarly, the system (3) has the state space model with the companion form
\[ \beta(t+1) = P_\beta(t) + R \xi(t) \]
\[ \eta(t) = H_\beta \beta(t) \]
where
\[ \beta = \begin{bmatrix} -P_i & \vdots & I_{\binom{n}{2}} \\ \vdots & \ddots & \vdots \\ -P_{n_s} & \ldots & 0 \end{bmatrix}, R = \begin{bmatrix} I_m \\ 0 \vdots \\ 0 \end{bmatrix}, H_\beta = \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix} \]

Introducing the augmented state \( x(t) \) and augmented input noise \( \tilde{w}(t) \)
\[ x(t) = \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix}, \tilde{w}(t) = \begin{bmatrix} w(t) \\ \xi(t) \end{bmatrix} \]
Then augmented system is given as
\[ x(t+1) = \Phi x(t) + I \tilde{w}(t) \]
\[ y_i(t) = H_x x(t) + v_i(t) \]
\[ s(t) = H_s x(t) \]
where \( \Phi = \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}, \Gamma = \begin{bmatrix} C & 0 \\ 0 & R \end{bmatrix}, H = \begin{bmatrix} H_a & H_\beta \end{bmatrix} \)

**Lemma 1**[12]. For the multi-sensor systems (14)-(16) with known model parameters and noise variances, the \( i \)th sensor subsystem has the local optimal Kalman predictor \( \hat{x}_i(t+1|t) \) of \( x(t) \) as
\[ \hat{x}_i(t+1|t) = \Psi_{yi,i}(t) \tilde{x}_i(t|t-1) + \Sigma_{yi,i}(t) y_i(t) \]
\[ \Psi_{yi,i}(t) = \Phi - K_{yi,i}(t) H \]
\[ K_{yi,i}(t) = \Phi \Sigma_{y(i|t-1)} H^T (H \Sigma_{y(i|t-1)} H^T + Q_{yi,i})^{-1} \]
and the local prediction error variance matrices satisfy the optimal Riccati equations
\[ \Sigma_{y(i|t-1)} = \Phi \Sigma_{y(i|t-1)} \Phi^T + \Sigma_{y(i|t-1)} - \Sigma_{y(i|t-1)} H \Sigma_{y(i|t-1)} H^T + \Sigma_{y(i|t-1)}^{-1} \]
and the local prediction cross-covariances satisfy the Lyapunov equation
\[ \Sigma_{yi}(t+1|t) = \Psi_{yi,i}(t) \Sigma_{yi}(t|t-1) \Psi_{yi,i}^T(t) + \Gamma Q_{yi,i} \Gamma^T, i \neq j \]

**Lemma 2**[12]. For the multi-sensor systems (14)-(16) with assumptions 1 and 2, the \( i \)th sensor subsystem has the local optimal time-varying Kalman smoother as
\[ \hat{x}_i(t+N|t) = \hat{x}_i(t-N|t-N-1) + \frac{1}{\sum_{j=1}^{n_c} (Q_{yi,i})^{-1}} \]

1428
\[
\sum_{j=0}^{N} K_j(t-N \mid t-N+j)\epsilon_j(t-N+j), i=1, \ldots, L
\]  
(23)
where the optimal Kalman predictor \( \hat{x}_i(t-N \mid t-N-1) \) can be computed by (18), and we can obtain
\[
\epsilon_i(t) = y_i(t) - H\hat{x}_i(t \mid t-1)
\]  
(24)
\[
K_i(t \mid t+j) = \Sigma_i(t \mid t-1) \left( \sum_{k=0}^{j-1} \Psi_i^{T}(t+k) \right) H^T \times (H \Sigma_i(t+j \mid t+j-1) H^T + Q_i)^{-1}
\]  
(25)
\[
K_i(t \mid t) = \Sigma_i(t \mid t-1) H^T (H \Sigma_i(t \mid t-1) H^T + Q_i)^{-1}
\]  
(26)
\[
\hat{x}_i(t-N \mid t) = H\hat{x}_i(t-N \mid t) + \hat{\Sigma}_i(t-N \mid t)
\]  
(27)
The error variance matrices and covariance matrices of local optimal Kalman smoother are given as
\[
P_{i}(t-N \mid t) = \Sigma_i(t-N \mid t-N-1) - \sum_{j=0}^{N} K_i(t-N \mid t-N+j) \times (H \Sigma_i(t-N+j \mid t-N+j-1) H^T + Q_i)^{-1}
\]  
(28)
\[
P_{i}(t \mid t) = \Sigma_i(t \mid t-1) H^T (H \Sigma_i(t \mid t-1) H^T + Q_i)^{-1}
\]  
(29)
\[
\Psi_{m_{i}}(t-N) = \sum_{j=0}^{N} [K_{ij}(t-N) - K_{ij}(t-N+k) H \Psi_{m_{j}}(t-N+k, t-N+j)] H \times \Psi_{m_{j}}(t-N+k, t-N+j)
\]  
(30)
\[
K_{ij}(t-N) = - \sum_{k=0}^{N} K_i(t-N \mid t-N+k) H \Psi_{m_{j}}(t-N+k, t-N+j) \times L_j(t-N+p+1) \Gamma, p=0, \ldots, N-1 ; \ K_{ij}(t-N) = 0
\]  
(31)
The cross-covariance among the smoother errors \( \hat{x}_i(t-N \mid t) \) are
\[
P_{ij}(t-N \mid t) = H \Psi_{m_{i}}(t-N \mid t) H^T
\]  
(32)
**Lemma 3** For the multi-sensor systems (14)-(16) with assumptions 1 and 2, we have the optimal information fusion Kalman smoother weighted by scalars
\[
\hat{x}_i(t-N \mid t) = \sum_{j=0}^{L} \omega_j(t-N \mid t) \hat{x}_j(t-N \mid t)
\]  
(33)
where the optimal scalar weighting coefficient vectors \( \omega_j(t-N \mid t) \) are given by
\[
[\omega_1(t-N \mid t), \ldots, \omega_L(t-N \mid t)] = [e^T(P_{i}^{-1}(t-N \mid t)) e]^{-1} e^T \times P_{i}^{-1}(t-N \mid t), i=1, \ldots, n
\]  
(34)
where \( e^T=[1, \ldots, 1] \), and \( L \times L \) matrix \( P_{i}(t-N \mid t) \) is defined as
\[
P_{i}(t-N \mid t) = \text{tr} P_{ij}(t-N \mid t), i, j = 1, \ldots, L
\]  
(35)
whose \( (i, j) \) element \( P_{ij}(t-N \mid t) \) are \( \text{tr} P_{ij}(t-N \mid t) \).

### III. INFORMATION FUSION ESTIMATORS OF MODEL PARAMETERS AND NOISE VARIANCES

For the multisensor systems (1) and (2) with assumptions 1-4, we can apply the information fusion identification method [13,14] to obtain estimators of the model parameters and noise variances. The information fusion estimators can be obtained by the following three stages.

#### A. Information fusion Autoregressive (AR) Parameter Estimator

Substituting (1) and (3) into (2) yields
\[
y_i(t) = A^{-1}(q^{-1}) C(q^{-1}) \omega(t) + P^T(q^{-1}) \xi(t-1) + \nu_i(t)
\]  
(36)
\[
\text{can be rewritten as}
\[
det P(q^{-1}) A(q^{-1}) \nu_i(t) = \text{det} P(q^{-1}) C(q^{-1}) \omega(t) + A(q^{-1}) \text{adj} P(q^{-1}) \xi(t-1) + \text{det} P(q^{-1}) A(q^{-1}) \nu_i(t)
\]  
(37)
\[
\text{Setting det} P(q^{-1}) \nu_i(t) = z_i(t), \text{ so we have}
\[
A(q^{-1}) z_i(t) = \text{det} P(q^{-1}) C(q^{-1}) \omega(t) + A(q^{-1}) \text{adj} P(q^{-1}) \xi(t-1) + \text{det} P(q^{-1}) A(q^{-1}) \nu_i(t)
\]  
(38)

Defining
\[
\tilde{C}(q^{-1}) = \text{det} P(q^{-1}) C(q^{-1}) = I_n + \tilde{C}_1 q^{-1} + \cdots \tilde{C}_n q^{-n},
\]
\[
\tilde{P}(q^{-1}) = A(q^{-1}) \text{adj} P(q^{-1}) = I_n + \tilde{P}_1 q^{-1} + \cdots \tilde{P}_n q^{-n},
\]
\[
\tilde{A}(q^{-1}) = \text{det} P(q^{-1}) A(q^{-1}) = I_n + \tilde{A}_1 q^{-1} + \cdots \tilde{A}_n q^{-n}
\]  
(39)
with \( n = m_{n_1} + n_s \), \( n_s = n_s + n_p, n_s = m_{n_1} + n_s \). \( (38) \) can be rewritten as
\[
A(q^{-1}) z_i(t) = \tilde{C}(q^{-1}) \omega(t) + \tilde{P}(q^{-1}) \xi(t-1) + \tilde{A}(q^{-1}) \nu_i(t)
\]  
(40)

Hence for the \( i \)th sensor, we have the least squares (LS) structure as
\[
z_i(t) = \Theta q(t) + \tau_i(t), i = 1, 2, \ldots, L
\]  
(41)
with the definition
\[
\Theta = \begin{bmatrix} A_1 & \cdots & A_{n_s} \end{bmatrix}
\]  
(42)
\[
\varphi(t) = [-z_1(t-1); \cdots; -z_s(t-n_s)]
\]  
(43)
\[
\tau_i(t) = \tilde{C}(q^{-1}) \omega(t) + \tilde{P}(q^{-1}) \xi(t-1) + \tilde{A}(q^{-1}) \nu_i(t)
\]  
(44)

**Lemma 4** For the \( i \)th sensor subsystem with multi-channel stationary ARMA model (1), the multi-dimensional recursive instrumental variable (MRIV) local estimators \( \hat{\Theta}_i(t) \) of \( \Theta \) are
\[
\tilde{\Theta}_i(t) = \hat{\Theta}_i(t-1) + \frac{[z_i(t) - \hat{\Theta}_i(t-1) \varphi(t)] \hat{\varphi}^T(t) P(t-1)}{1 + \hat{\varphi}^T(t) P(t-1) \varphi(t)}
\]  
(45)
\[\]
\[ P^{\prime}(t) = P^{\prime}(t-1) - \frac{P^{\prime}(t-1)\varphi_{1}(t)\hat{P}^{T}(t)P(t-1)}{1 + \varphi_{1}(t)P(t-1)\varphi_{1}(t)} \]  
\[ \hat{P}^{\prime}(t) = \varphi(t - n_{2}) \]  
(46)
with initial value \( \hat{P}^{\prime}(0) = \hat{P}_{0} \), \( P^{\prime}(0) = \alpha L_{\text{m}} \), \( y_{i}(k) = 0 \) \( k \leq 0 \).

It is proved [13] that multi-dimensional local RIV estimators \( \hat{\Theta}(t) \) are strongly consistent, i.e.
\[ \hat{\Theta}(t) \rightarrow \Theta \quad \text{as} \quad t \rightarrow \infty \quad \text{w.p.1}, \quad i = 1, \ldots, L \]  
(48)
where “w.p.1” denotes “with probability one”. The information fusion estimator \( \hat{\Theta}_{f}(t) \) of \( \Theta \) is defined as
\[ \hat{\Theta}_{f}(t) = \frac{1}{L} \sum_{i=1}^{L} \hat{\Theta}(t) \]  
(57)
so that from (48) it is also strongly consistent, i.e.
\[ \hat{\Theta}_{f}(t) \rightarrow \Theta \quad \text{as} \quad t \rightarrow \infty \quad \text{w.p.1} \]  
(50)
From (49), \( \hat{A}_{i}(t) \) is the \( i \)th element of \( \hat{\Theta}_{f}(t) \). Hence, we have
\[ \hat{A}(q^{-1}) = \hat{A}(q^{-1}) + y_{i}(t) \]  
(52)
It is clearly that \( \hat{A}(q^{-1}) \) is a stable polynomials of \( q^{-1} \), \( y_{i}(t) \) is a stationary stochastic process, so it yields that \( r_{i}(t) \) is also a stationary stochastic process with cross-correlation function as
\[ R_{ij}(k) = E[r_{i}(t)r_{j}^{T}(t-k)] \quad k = 0, \ldots, n_{2} \]  
(53)
with a cut-off as lag \( n_{2} \). At time \( t \), the estimate of the measurement process \( r_{i}(t) \) is defined as
\[ \hat{r}_{i}(t) = \hat{A}(q^{-1})y_{i}(t) \]  
(54)
The on-line sampled correlation function \( \hat{R}_{ij}(k) \) of \( R_{ij}(k) \) is defined as
\[ \hat{R}_{ij}(k) = \frac{1}{L} \sum_{j=1}^{L} \hat{r}_{i}(j)\hat{r}_{j}^{T}(j-k) \]  
(55)
which have the recursive form
\[ \hat{R}_{ij}(k) = \hat{R}_{ij}(k-1) + \frac{1}{t} [\hat{r}_{i}(t)\hat{r}_{j}^{T}(t-k) - \hat{R}_{ij}(k-1)] \quad t = 2, 3, \ldots \]  
(56)
with the initial value \( \hat{R}_{ij}(k) = \frac{1}{t} \sum_{j=1}^{L} r_{i}(t)r_{j}^{T}(1-k) \).

Defining the MA process as
\[ m(t) = \tilde{C}(q^{-1})w(t) \]  
(57)
and \( m(t) \) is a stationary stochastic process, whose correlation function \( R_{m}(k) \) is defined as \( R_{m}(k) = E[m(t)m^{T}(t-k)] \). \( k = 0, \ldots, n_{2} \). From (57), we have that \( R_{m}(k) = \sum_{u=1}^{n_{2}} \tilde{C}_{u}Q_{u}^{-1} \).

In order to yield \( \tilde{C}_{u} \) and \( Q_{u} \), we need to find \( R_{m}(k) \). Computing the correlation function of the two sides of (58) yields that
\[ R_{m}(k) = R_{m}(k) - \sum_{u=1}^{n_{2}} \tilde{C}_{u}Q_{u}^{-1} - \sum_{u=1}^{n_{2}} \tilde{C}_{u}Q_{u}^{-1} \delta_{u}, k = 0, \ldots, n_{2} \]  
(59)
where \( \tilde{C}_{u}(t) = 0(i > n_{2}), \tilde{C}_{u} = 0(i > n_{2}) \).

Substituting the sample estimates \( \hat{R}_{ij}(k) \) and fused estimates \( \hat{\tilde{A}}(l) \) and \( \hat{\tilde{P}}(l) \) into (59) yields the local estimates of \( R_{m}(k) \) as
\[ \hat{R}_{m}(k) = \hat{R}_{m}(k) - \sum_{u=1}^{n_{2}} \hat{C}_{u}(t)Q_{u}^{-1} - \sum_{u=1}^{n_{2}} \hat{C}_{u}(t)Q_{u}^{-1} \delta_{u}, k = 0, \ldots, n_{2} \]  
(60)
Bases on the estimates \( \hat{R}_{m}(k) \), using the Gevers-Wouters algorithm [15] with a dead band, we can obtain the local estimates \( \hat{C}_{m}(u) = 1, \ldots, n_{2}, \hat{C}_{m}(u) = 1, \ldots, L \) and \( \hat{Q}_{m}(u) \) as
\[ \hat{Q}_{m}(u) = \lim_{t \rightarrow \infty} \hat{R}_{m}(u,l) \]  
(61)
\[ \hat{C}_{m}(u) = \lim_{t \rightarrow \infty} \hat{R}_{m}(u,l-k)[\hat{R}_{m}(u,l-k)]^{-1}, \quad u = 1, \ldots, n_{2} \]  
(62)
\[ \hat{R}_{m}(u,l-u) = \hat{R}_{m}(u) - \sum_{r=1}^{n_{2}} \hat{R}_{m}(u,l-r)x \]  
(63)
with the definitions
\[ \hat{R}_{m}(u,l-r) = \hat{R}_{m}(0,l-r) \quad \text{at} \quad l < r \]  
(64)
Then the information fusion estimates \( \hat{C}_{m}(t), \hat{Q}_{m}(t) \) based on all sensors are defined as
\[ \hat{Q}_{m}(t) = \frac{1}{L} \sum_{i=1}^{L} \hat{Q}_{m}(t), \quad \hat{C}_{m}(t) = \frac{1}{L} \sum_{i=1}^{L} \hat{C}_{m}(t), \]  
(65)
\[ \text{where} \quad \text{det} \quad P(t^{-1}) = 1 + g_{1}q^{-1} + \cdots + g_{n_{2}}q^{-n_{2}} \]  
(66)
\[ M = \begin{bmatrix} C_{1} & \cdots & C_{n} \end{bmatrix} \]  
(67)
From (39), we have
\[ \Omega M = Y \]  
(68)
Substituting the estimator \( \hat{C}_i(t) \) \((u = 1, \ldots, n_u)\) into the first formula of (39) yields
\[
\hat{C}(q^{-1}) = \det P(q^{-1}) \hat{C}(q^{-1}) = I_m + \hat{C}_i(t)q^{-1} + \cdots + \hat{C}_u(t)q^{-n_u} \quad (71)
\]

Defining
\[
\hat{M}(t) = \left[ \hat{C}_i(t) \cdots \hat{C}_u(t) \right]^T \quad (73)
\]
Substituting \( \hat{C}_i(t) \) into (68) yields
\[
\Omega \hat{M}(t) = \hat{Y}(t) \quad (74)
\]
where
\[
\hat{Y}(t) = \begin{bmatrix}
\hat{C}_i(t) - g_i I_m \\
\vdots \\
\hat{C}_{mij} - g_{mij} I_m \\
\hat{C}_{me} + g_m I_m \\
\end{bmatrix} \quad (75)
\]
Solving (74) by Pseudo inverse gives that
\[
\hat{M}(t) = (\Omega^T \Omega)^{-1} \Omega^T \hat{Y}(t) \quad (76)
\]

**Theorem 1.** For the multisensor systems (1)-(2) with assumptions 1-3, the fused estimators of MA parameters and noise variances are consistent, i.e.
\[
\hat{A}_i(t) \to \overline{A}_i, \quad i = 1, \ldots, n_\pi, \quad \text{as } t \to \infty, \quad \text{i.a.r} \quad (77)
\]
\[
\hat{\hat{F}}_i(t) \to \overline{F}_i, \quad i = 1, \ldots, n_\pi, \quad \text{as } t \to \infty, \quad \text{i.a.r} \quad (78)
\]
\[
\hat{C}_u(t) \to \overline{C}_u, \quad u = 1, \ldots, n_u, \quad \text{as } t \to \infty, \quad \text{i.a.r} \quad (79)
\]
\[
\hat{Q}_\sigma(t) \to \overline{Q}_\sigma, \quad \text{as } t \to \infty, \quad \text{i.a.r} \quad (80)
\]
\[
\hat{C}_e(t) \to \overline{C}_e, \quad u = 1, \ldots, n_u, \quad \text{as } t \to \infty, \quad \text{i.a.r} \quad (81)
\]
where the notation “i.a.r” denotes the convergence “in a realization”[2].

**Proof.** From (39) and (51), we have that (77) and (78) hold. Applying \( R_m(k) = \sum_{i=1}^{n_\pi} C_{\pi} Q\sigma C_{\sigma}^{-1} \) and the existence theorem of implicit function, \( \overline{C}_\pi \) and \( \overline{Q}_\sigma \) are the continuous functions of the elements of \( R_m(k)(k = 0, \ldots, n_\pi) \) in a sufficiently small neighborhood, i.e.

\[
\overline{C}_\pi = f_\pi (R_m(0), \ldots, R_m(n_\pi)) \quad (82)
\]
\[
\overline{Q}_\sigma = f_\sigma (R_m(0), \ldots, R_m(n_\sigma)) \quad (83)
\]
decking}

where \( f_\pi \) and \( f_\sigma \) are the continuous functions, when \( t \) is sufficiently large, we have relations
\[
\hat{C}_{ij}(t) = f_{ij}(\hat{R}_{mj}(0), \ldots, \hat{R}_{mj}(n_\pi)) \quad (84)
\]
\[
\hat{Q}_{ij}(t) = f_{ij}(\hat{R}_{mj}(0), \ldots, \hat{R}_{mj}(n_\pi)) \quad (85)
\]

According to the ergodicity [13], we have
\[
\hat{R}_{mj}(k) \to \overline{R}_{mj}(k), \quad \text{as } t \to \infty, \quad \text{i.a.r} \quad (86)
\]
and from (59),(60),(77), (78) and (86), we have
\[
\hat{R}_{mj}(k) \to \overline{R}_{mj}(k), \quad \text{as } t \to \infty, \quad \text{i.a.r} \quad (87)
\]
From (82)-(85), (87) and the continuity of \( f_\pi \) and \( f_\sigma \), we have
\[
\hat{C}_{ij}(t) \to \overline{C}_{ij}, \quad \hat{Q}_{ij}(t) \to \overline{Q}_{ij}, \quad \text{as } t \to \infty, \quad \text{i.a.r} \quad (88)
\]
Therefore, from (65) and (88), it follows that (79) and (80) hold. From (70) and (76), the each element of \( \overline{M} \) and \( \hat{M}(t) \) is a continuous function of elements of \( \overline{C}_\pi \) and \( \hat{C}_e(t) \) \((u = 1, \ldots, n_u)\), which yields \( \hat{M}(t) \to \overline{M} \), as \( t \to \infty \), i.a.r i.e., (81) holds.

**IV. SELF-TUNING FUSION KALMAN SMOOTHER**

When model parameters and noise variances are all unknown, substituting their estimators into the optimal fusion Kalman smoother yields the self-tuning fusion Kalman smoother. It consists of the following steps:

Step 1.a) Applying the multi-dimensional recursive instrumental variable (MRIV) algorithm, yields the information fusion parameter estimators \( \hat{A}_i(t) \) \((l = 1, \ldots, n_u)\) at time \( t \). And based on (39), \( \hat{A}_i(t) \) and \( \hat{F}_i(t) \) can be obtained.

b) Based on the above parameters estimators and the sampled correlation function estimators \( \hat{R}_{mj}(k) \), and applying the Gevers-Wouters algorithm with a dead band to (60). The information fusion estimators \( \hat{C}_i(t) \) and \( \hat{Q}_\sigma(t) \) can be obtained. From (76), the fused estimators \( \hat{C}_e(t) \) can be obtained.

Step 2. Substituting all the estimates obtained by step 1 into (9), (12) and (17) yields the estimators \( \hat{\phi}(t) \), \( \hat{\Gamma}(t) \) and \( \hat{\Omega}_{\pi}(t) \) of \( \phi \), \( \Gamma \) and \( Q_{\sigma} \), and we have that
\[
[\hat{\phi}(t) - \phi] \to 0, \quad [\hat{\Gamma}(t) - \Gamma] \to 0, \quad [\hat{\Omega}_{\pi}(t) - Q_{\sigma}] \to 0, \quad \text{as } t \to \infty, \quad \text{i.a.r} \quad (89)
\]

Step 3. In Lemma 1 and Lemma 2, \( \phi \), \( \Gamma \) and \( Q_{\sigma} \) are replaced by \( \hat{\phi}(t) \), \( \hat{\Gamma}(t) \) and \( \hat{\Omega}_{\pi}(t) \) respectively. So, the estimators \( \hat{\psi}_{\pi}(t) \) and \( \hat{\psi}_{\pi}(t) \) are obtained. Substituting the estimates into (18), the self-tuning local Kalman predictor can be given as
\[
\hat{x}_{e}(t+1 | t) = \hat{\psi}_{\pi}(t)\hat{x}_{e}(t | t-1) + \hat{\psi}_{\pi}(t)\hat{y}(t) \quad (90)
\]
And the estimates \( \hat{\mathbf{z}}_i(t) \) satisfy the self-tuning Riccati equations
\[
\dot{\mathbf{z}}_i(t+1) = \mathbf{A}_i(t) \mathbf{z}_i(t) - \mathbf{z}_i(t) \mathbf{B}_i(t) \mathbf{u}_i(t) + Q_i(t) \hat{\mathbf{w}}_i(t)
\]
(91)
and the prediction cross-covariance matrices satisfy the self-tuning Lyapunov equations
\[
\dot{\mathbf{P}}_i(t) = \mathbf{A}_i(t) \mathbf{P}_i(t) + \mathbf{P}_i(t) \mathbf{A}_i(t)^T + \mathbf{Q}_i(t) \mathbf{P}_i(t) \mathbf{B}_i(t) + \mathbf{B}_i(t)^T \mathbf{P}_i(t) \mathbf{Q}_i(t) + \mathbf{R}_i(t)
\]
(92)
Substituting the estimators into (23)-(32), yields the self-tuning local Kalman smoother \( \hat{x}_i(t|T) \) as
\[
\hat{x}_i(t|T) = \mathbf{X}_i(t|T) + \mathbf{K}_i(t|T) \hat{\mathbf{Y}}_i(t|T)
\]
(93)
where \( \mathbf{X}_i(t|T) = \mathbf{F}_i(t|T) \mathbf{X}_i(t|T-1) + \mathbf{K}_i(t|T) \hat{\mathbf{Y}}_i(t|T) \)
(94)
And the self-tuning local Kalman signal smoother is given as
\[
\hat{s}_i(t|T) = \mathbf{H}_i(t|T) \hat{x}_i(t|T)
\]
(95)
Step 4. By (34) and (35), the estimates \( \hat{\mathbf{P}}_i(t|T) \) and \( \hat{\mathbf{Y}}_i(t|T) \) are obtained and the self-tuning fused Kalman signal smoother is given as
\[
\hat{s}_i(t|T) = \sum_{j=1}^{L} \hat{s}_j(t|T) \Psi_i(t|T)
\]
(96)
The above four steps are repeated at each time \( t \).

V. CONVERGENCE ANALYSIS

Now we prove the convergence of the self-tuning local and fused Kalman smoothers by the DESA method.

**Lemma 4** Consider the discrete time error system
\[
\delta(t) = F(t) \delta(t-1) + u(t)
\]
(97)
where \( t \geq 0, \delta(t) \in \mathbb{R}^n \) is the output, \( u(t) \in \mathbb{R}^r \) is the input, and the matrix \( F(t) \in \mathbb{R}^{n \times n} \) is uniformly asymptotically stable[14]. If \( u(t) \) is bounded, then \( \delta(t) \) is bounded. If it goes to zero, \( \delta(t) \) goes to zero, \( u(t) \) goes to zero.

**Theorem 1.** For the multisensor systems (1) and (2) with the assumptions 1-4, the local self-tuning Kalman predictor \( \hat{x}_i(t+1|T) \) converges to the local time-varying optimal Kalman predictor \( \hat{x}_i(t+1|T) \) in the sense that
\[
[\hat{x}_i(t+1|T) - \hat{x}_i(t+1|T)] \rightarrow 0, \text{ as } t \rightarrow \infty, \text{ i.a.r.}
\]
(98)
**Proof.** According to [16], it can be proved similarly that \( \delta(t) \rightarrow 0, \text{ as } t \rightarrow \infty, \text{ i.a.r.} \) Applying the dynamic variance error system analysis (DVES) method [5], it can be proved that \( \delta(t) \rightarrow 0, \text{ as } t \rightarrow \infty, \text{ i.a.r.} \) Hence, according to the stability theory of the Kalman filtering [17], it follows that
\[
[\hat{K}_p(t) - K_p(t)] \rightarrow 0, \text{ as } t \rightarrow \infty, \text{ i.a.r.}
\]
(99)
Setting
\[
\delta(t) = \hat{x}_i(t+1|T) - \hat{x}_i(t+1|T), \hat{Y}_p(t) = \mathbf{Y}_p(t) + \mathbf{R}_p(t)
\]
(100)
Subtracting (18) from (90) yields a dynamic error system
\[
\delta(t) = \mathbf{Y}_{p}(t) \delta(t-1) + u(t)
\]
(101)
where \( u(t) = \delta(t) \). From the assumption 4 and the boundedness of \( \hat{K}_p(t) \), we have that \( \hat{K}_p(t) \) is bounded. From the stability theorem [17] of the Kalman filter, \( \mathbf{Y}_p(t) \) and \( \mathbf{R}_p(t) \) are uniformly asymptotically stable. Hence, applying Lemma 4 to (90) yields that \( \hat{x}_i(t+1|T) \) is bounded. From the boundedness of \( \hat{x}_i(t+1|T) \) and \( \hat{x}_i(t) \), and from (100), we can obtain that \( u(t) \rightarrow 0 \). And applying Lemma 4 to (101) yields that
\[
\delta(t) \rightarrow 0, \text{ as } t \rightarrow \infty, \text{ i.a.r.}
\]
(102)
Hence (98) holds. The proof is completed.

**Theorem 2.** For the multisensor systems (1) and (2) with the assumptions 1-4, the self-tuning fused Kalman signal smoother \( \hat{s}_i(t|T) \) converges to the optimal fused Kalman signal smoother \( \hat{s}_i(t|T) \) in a realization, i.e.
\[
[\hat{s}_i(t|T) - \hat{s}_i(t|T)] \rightarrow 0, \text{ as } t \rightarrow \infty, \text{ i.a.r.}
\]
(103)
**Proof.** From (24), applying (98) yields
\[
[\mathbf{E}_i(t) - \mathbf{E}_i(t)] \rightarrow 0, \text{ as } t \rightarrow \infty, \text{ i.a.r.}
\]
(104)
Similarly, we have that
\[
[\hat{\mathbf{E}}_i(t) - \hat{\mathbf{E}}_i(t)] \rightarrow 0, \text{ as } t \rightarrow \infty, \text{ i.a.r.}
\]
(105)
Setting \( \hat{\mathbf{E}}_i(t) = \mathbf{E}_i(t) + \Delta \hat{\mathbf{E}}_i(t), \hat{\mathbf{E}}_i(t) \rightarrow 0, \text{ as } t \rightarrow \infty, \text{ i.a.r.}
\)
Denoting \( \Delta \hat{\mathbf{E}}_i(t) = \hat{\mathbf{E}}_i(t) - \hat{\mathbf{E}}_i(t) \), and subtracting (23) from (93) yields
\[
\eta_i(t) = \sum_{j=1}^{L} \Delta \hat{E}_j(t) \mathbf{S}_i(t) + \mathbf{E}_i(t)
\]
(106)
From (107) we have
\[
\eta_i(t) = \sum_{j=1}^{L} \Delta \mathbf{E}_j(t) \mathbf{S}_i(t) + \mathbf{E}_i(t)
\]
(108)
Applying the boundedness of \( \mathbf{S}_i(t) \), and \( \Delta \hat{\mathbf{E}}_i(t) \) yields that \( \Delta \hat{\mathbf{E}}_i(t) \) is bounded. From (108), applying (106) yields \( \eta_i(t) \rightarrow 0, \text{ as } t \rightarrow \infty, \text{ i.a.r.}
\) From (98) and \( \eta_i(t) \rightarrow 0, \text{ as } t \rightarrow \infty, \text{ i.a.r.}
\) So from (27) and (95), we have that
\[
[\hat{\mathbf{E}}_i(t) - \mathbf{E}_i(t)] \rightarrow 0, \text{ as } t \rightarrow \infty, \text{ i.a.r.}
\]
\[ \hat{s}_i(t-N|t)-\hat{s}_i(t-N|t) = 0, \text{ as } t \to \infty , \text{ i.a.r.} \] (109)

From Lemma 2, applying (99) and (105), we have
\[ \hat{P}(t-N|t) = \hat{P}(t-N|t), \hat{P}(t-N|t) = \hat{P}(t-N|t) \]
as \( t \to \infty , \text{ i.a.r.} \) (110)
so from (34), we have that
\[ \hat{w}_i(t-N|t) = \hat{w}_i(t-N|t), \text{ as } t \to \infty , \text{ i.a.r.} \] (111)
Setting \( \hat{w}_i(t-N|t) = \hat{w}_i(t-N|t) + \hat{w}_i(t-N|t) \), from (111) we have that \( \hat{w}_i(t-N|t) \to 0 \), subtracting (33) from (96) yields
\[ \hat{s}_i(t-N|t) - \hat{s}_i(t-N|t) = \sum_{j=1}^{I} \hat{d}_i(t-N|t)\hat{s}_j(t-N|t) - \hat{s}_i(t-N|t) + \sum_{j=1}^{I} \hat{d}_i(t-N|t)\hat{s}_j(t-N|t) \] (112)

From (95) and the boundedness of \( \hat{s}_i(t-N|t) \), we have that \( \hat{s}_i(t-N|t) \) is bounded. Applying (109), \( \hat{w}_i(t-N|t) \to 0 \) and the boundedness of \( \hat{s}_i(t-N|t) \) yields that (103) holds.

**VI. SIMULATION EXAMPLE**

Consider the multisensor multi-channel autoregressive moving average (ARMA) signal with white measurement noises and a colored measurement noise
\[ (I_2 + A_1q^{-1} + A_2q^{-2})s(t) = (I_2 + C_1q^{-1})w(t) \] (113)
\[ y_i(t) = s(t) + \eta(t) + v_i(t), i = 1, 2, 3 \] (114)
\[ (I_2 + P_1q^{-1})\eta(t) = \xi(t-1) \] (115)
where the signal \( s(t) = [s_1(t) \ s_2(t)]^T \), \( y_i(t) \in \mathbb{R}^2 \) is the measurement of the \( i \)-th sensor, \( w(t) \), \( \xi(t) \) and \( v_i(t) \) are independent white noises with zero mean and variances \( Q_w \), \( Q_\xi \) and \( Q_{vi} \), respectively. Assume \( A_1, A_2, C_1 \) and \( Q_w \) are unknown. (113)-(115) have the equivalent state model (14)-(17). Hence the problem of finding self-tuning fusion Kalman smoother \( \hat{s}_i(t-N|t) \) of signal \( s(t) \) can be converted into the problem of finding the self-tuning fusion Kalman smoother weighted by scalars of state \( x(t) \)'s first component.

In simulation we take that
\[ A_1 = \begin{bmatrix} 0.8 & -0.5 \\ 0.1 & 0.4 \end{bmatrix}, A_2 = \begin{bmatrix} 0.5 & -0.4 \\ 0.3 & 0.9 \end{bmatrix}, Q_w = \begin{bmatrix} 7 & 0 \\ 0 & 9 \end{bmatrix} \]
\[ C_1 = \begin{bmatrix} -0.1 & -0.3 \\ 0.3 & -0.6 \end{bmatrix}, P_1 = \begin{bmatrix} 0.5 & -0.3 \\ 0.1 & -0.4 \end{bmatrix}, Q_\xi = \begin{bmatrix} 0.15 & 0 \\ 0 & 0.15 \end{bmatrix}, \]
\[ Q_{vi} = \begin{bmatrix} 0.15 & 0 \\ 0 & 0.45 \end{bmatrix}, Q_{v1} = \begin{bmatrix} 0.25 & 0 \\ 0 & 0.6 \end{bmatrix}, Q_{v2} = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.3 \end{bmatrix} \]

In Figs. 1–5, the straight lines denote the true values, and \( M(k,r) \) denotes the \((k,r)\)th element of the matrix \( M \).

Applying Lemma 4, we can obtain the fused estimate \( \hat{A}_i(t) \). The fused curves of \( \hat{A}_1(t) \) and \( \hat{A}_2(t) \) are shown in Figs.1 and 2. Applying the correlation method and Gevers-Wouters algorithm with the dead band \( T_d = 200 \), we can obtain \( \hat{C}_j(t) \) and \( \hat{Q}_n(t) \). The curves of the fused estimates \( \hat{Q}_n(t) \) are shown in Fig.3. The fused estimates of \( \hat{C}_j(t) \) are shown in Fig.4. The error curves between the self-tuning and optimal fused Kalman signal smoothers are presented in Fig.5. We see the errors approximate to zero, which verify the self-tuning fusion Kalman signal smoother converges to the optimal fusion signal smoother.
Kalman smoother converges to the optimal fused Kalman method, it has been proved strictly that the self-tuning fused Gevers-Wouters algorithm, the estimators of the model method. By the MRIV scalars has been presented by the classical Kalman filter tuning information fusion Kalman smoother weighted by unknown model parameters and noise variances, a self-

**ACKNOWLEDGMENT**

This work is supported by National Natural Science Foundation of China under NSFC-60874063, Automatic Control Key Laboratory of Heilongjiang University, and Science and Technology Research Foundation of Heilongjiang Education Department under Grant 12513076.

**REFERENCES**


