Available online at www.sciencedirect.com
ScienceDirect
Signal Processing 87 (2007) 2392–2406
SIGNAL PROCESSING
www.elsevier.com/locate/sigpro

A low-complexity suboptimal filter for continuous–discrete linear systems with parametric uncertainties

Vladimir Shin[a], Du Yong Kim[a,*], Georgy Shevlyakov[b], Kiseon Kim[b]

[a]Department of Mechatronics, Gwangju Institute of Science and Technology, 1 Oryong-Dong Buk-Gu, Gwangju 500-712, South Korea
[b]Department of Information and Communications, Gwangju Institute of Science and Technology, 1 Oryong-Dong Buk-Gu, Gwangju 500-712, South Korea

Received 25 April 2006; received in revised form 7 March 2007; accepted 18 March 2007
Available online 7 April 2007

Abstract

We present a novel suboptimal filtering algorithm addressing estimation problems that arise in mixed continuous–discrete linear time-varying systems with stochastic parametric uncertainties. The suboptimal state estimate is formed by summing of local Kalman estimates with weights depending only on time instants t_k. In contrast to optimal weights, the suboptimal weights do not depend on current measurements, and thus the proposed filter is of a low-complexity and it can easily be implemented in real-time. High accuracy and efficiency of the suboptimal filter are demonstrated on the damper harmonic oscillator motion and the vehicle motion constrained to a plane.

Keywords: Adaptive filtering; Kalman filtering; Continuous–discrete system; Mean-square error; Fusion

1. Introduction

We consider a linear system described by the stochastic differential equation

$$\dot{x}_t = F_t(\theta)x_t + G_t(\theta)v_t, \quad t \geqslant 0, \tag{1}$$

where $x_t \in \mathbf{R}^n$ is the state, $v_t \in \mathbf{R}^q$ is a zero-mean Gaussian white noise with covariance $E(v_t v_s^{\mathrm{T}}) = Q_t\delta(t-s)$, and $F_t \in \mathbf{R}^{n\times n}$, $G_t \in \mathbf{R}^{n\times q}$, and $Q_t \in \mathbf{R}^{q\times q}$.

Discrete linear measurements are taken at time instants t_k:

$$y_{t_k} = H_{t_k}(\theta)x_{t_k} + w_{t_k},$$
$$k = 1,2,\ldots, \quad t_{k+1} > t_k \geqslant t_0 = 0, \tag{2}$$

where $y_{t_k} \in \mathbf{R}^m$ is the measurement $H_{t_k} \in \mathbf{R}^{n\times m}$, and $\{w_{t_k} \in \mathbf{R}^m, k = 1,2,\ldots\}$ is a zero-mean white Gaussian sequence, $w_{t_k} \sim \mathrm{N}(0, R_{t_k}(\theta))$. The distribution of the initial state x_0 is Gaussian, $x_0 \sim \mathrm{N}(\bar{x_0}(\theta), P_0(\theta))$, and x_0, v_t, and $\{w_{t_k}\}$ are assumed independent. The sample times $\{t_k\}$ are scheduled and completely known in advance. In general, the partition of the time instants $\triangle t_k = t_k - t_{k-1}$ is an arbitrary.

In addition, it is assumed that the matrices $F_t(\theta)$, $G_t(\theta)$, $Q_t(\theta)$, $H_{t_k}(\theta)$, $R_{t_k}(\theta)$, $P_0(\theta)$ and the initial mean $\bar{x}_0(\theta)$ include the unknown parameter $\theta \in \mathbf{R}^r$, which takes only a finite set of values

$$\theta \in \{\theta_1, \ldots, \theta_N\}. \tag{3}$$

This finite set might be a result of discretizing a continuous parameter space [1,2]. The parameter θ is time-invariant, that is at the starting point $t = 0$

*Corresponding author. Tel.: +82 62 970 3266.
E-mail address: duyong@gist.ac.kr (D.Y. Kim).

0165-1684/$ - see front matter
doi:10.1016/j.sigpro.2007.03.016

Nomenclature

$\mathbf{R}^n$	Euclidean space of dimension **n**
$\mathbf{R}^{n\times m}$	space of $n \times m$ real matrices
θ	unknown parameter (vector), $\theta \in \mathbf{R}^r$
θ_i	value of parameter θ, $\theta = \theta_i$, $i = 1, \ldots, N$
N	total number of values of parameter $\theta = \{\theta, \ldots, \theta_N\}$
$x_t = x_t(\theta)$	state vector at time instant **t** depending on θ
$x_t^{(i)} = x_t(\theta_i)$	state vector $x_t = x_t(\theta)$ at the value of parameter $\theta = \theta_i$
y_{t_k}	measurement vector at time instant t_k
y_0^t	collection of all measurements up to time **t**
$p(\theta)$	prior probability of θ
$p(\theta \mid y_0^t)$	a posteriori probability of θ given y_0^t
v_t	system noise vector at time instant **t**, $v_t \in \mathbf{R}^q$
w_{t_k}	measurement noise vector at time instant t_k, $w_{t_k} \in \mathbf{R}^m$
$H_{t_k}(\theta)$	measurement matrix at time instant t_k depending on parameter θ, $H_{t_k}(\theta) \in \mathbf{R}^{n\times m}$
$F_t(\theta)$	system matrix at time instant **t** depending on parameter θ, $F_t(\theta) \in \mathbf{R}^{n\times n}$
$Q_t(\theta)$	covariance of system noise depending on parameter θ, $Q_t \in \mathbf{R}^{q\times q}$
$R_{t_k}(\theta)$	covariance of measurement noise depending on parameter θ, $R_{t_k} \in \mathbf{R}^{m\times m}$
$\hat{x}_t^{\mathrm{opt}}, P_t^{\mathrm{opt}}$	optimal mean-square estimate of state x_t given y_0^t and corresponding covariance
$\hat{x}_t^{\mathrm{sub}}, P_t^{\mathrm{sub}}$	suboptimal (fusion) estimate of state x_t given y_0^t and corresponding covariance
$\hat{x}_t^{(i)}, P_t^{(i)}$	optimal local estimate of state $x_t^{(i)}$ and corresponding covariance
$\hat{x}_{t_k}^{(i)-}, P_{t_k}^{(i)-}$	predicted local estimate of state $x_t^{(i)}$ at time t_k and corresponding covariance
$\mathbf{N}(m, P)$	multidimensional normal pdf with mean **m** and P covariance
δ_{ij}	Kronecker delta function
$\delta(t-s)$	Dirac delta function

the parameter is settled, $\theta = \theta_i$, and it cannot change during time progress $t > 0$.

A fundamental problem associated with such systems is estimation of the state x_t from the noisy measurements $y_0^t = \{y_{t_s} : 0 \leqslant t_s \leqslant t\}$.

Many approaches are available for the adaptation of systems. Most identification approaches (see, for example, [1–7]) may be applied to construct an adaptive mechanism. Among existing methods, we are particularly interested in the partitioned adaptive approach that is mathematically based on the Bayesian estimation theory, since it is useful not only for identifying noise statistics but also for estimating unknown system parameters and states, which is sometimes called structure adaptation. In structure adaptation, two methods are primarily used for the system (1)–(3). The first method is based on the extended Kalman filter (EKF) [8–12], and the second one is based on the standard Kalman filter and the Lainiotis partition theorem [1–3]. Note that the second method is often called the adaptive Kalman filtering (AKF). Both filters EKF and AKF are based on the Bayesian approach in which the unknown parameter θ is assumed to be random with a prior known probability $p(\theta)$. The EKF represents a suboptimal nonlinear filtering algorithm to estimate the composite state vector $[x_t\theta]^{\mathrm{T}}$ that contains θ as its component. However, it is difficult to estimate the effect of nonlinear approximations made in the suboptimal realization of EKF [8–12].

The AKF separates the filtering process x_t from identification of the unknown parameter θ [1–5]. In this paper we are interested in such an AKF that constitutes a partitioning of the original nonlinear filter into the bank of much simpler N local Kalman filters where each local filter uses its own system model (1), (2) matched to each possible parameter value $\theta = \theta_i$, $i = 1, \ldots, N$. This AKF is also referred to multiple model adaptive estimation [13–18]. The overall estimate of state of this AKF is given by a weighted sum of local Kalman estimates, thus it can be implemented on a set of parallel processors due to its inherent parallel structure. However, the optimal AKF's weights represent the conditional probabilities of the specific parameter values $p(\theta_i \mid y_0^t)$ which depend on current observations y_0^t and it is rather difficult to implement the AKF in real-time for high-dimension of state vector and large number of local Kalman estimates (filters).

In this paper, discrete filtering of continuous-time linear systems with uncertainties is considered. We extend the well-known optimal discrete and continuous AKFs [1–5] to the mixed continuous–discrete linear systems with parametric uncertainties. But the main objective of the present paper is to give an alternative suboptimal filter (SF) for that kind of systems. Similarly to the optimal AKFs, the SF represents the state estimate as a weighted sum of parameter-conditional estimates (local Kalman estimates) with the weights depending only on time instants and being independent of current measurements y_0^t. It gives an opportunity to design a low-complexity SF that can be easily implemented in real-time, especially in high dimension problems.

This paper is organized as follows. In Section 2, we formulate the adaptive filtering problem and generalize the optimal discrete AKF to mixed continuous-discrete linear dynamic systems. In Section 3, we derive general equations of the SF for continuous–discrete systems. The SF represents a linear combination of local Kalman filters. Each local Kalman filter is fused by the minimum mean-square criterion. In Section 4, the suboptimal filtering algorithm implementation steps are considered. In Section 5, the SF is applied for a special class of linear systems with measurement uncertainties. A solution of the joint detection–estimation problem based on the SF is given. In Section 6, the SF is numerically tested in real-life system models. In Section 7, conclusions are made.

2. The optimal adaptive Kalman filter for continuous–discrete systems

The AKF for continuous and discrete linear systems with uncertainties was firstly proposed by Magill [3] and later generalized by Lainiotis [2,5] to form the framework of partitioned algorithms [1,2]. In this section, we shall present a simple extension of the AKF to the mixed continuous–discrete linear systems. Consider a continuous–discrete linear system with unknown parameters given by Eqs. (1)–(3). According to the Bayesian approach, it is assumed that a prior probability for the parameter θ, $p(\theta)$, is available,

$$p(\theta_i)\geqslant 0, \quad i=1,\dots,N, \qquad p(\theta_1)+\cdots+p(\theta_N)=1. \tag{4}$$

2.1. The optimal estimate and error covariance

According to the Lainiotis partition theorem the optimal mean-square state estimate $\hat{x}_t^{\text{opt}}=E(x_t|y_0^t)$ of x_t and the corresponding estimation error covariance $P_t^{\text{opt}}=E[(x_t-\hat{x}_t^{\text{opt}})(x_t-\hat{x}_t^{\text{opt}})^{\mathrm{T}}|y_0^t]$ are given by the weighted sums

$$\hat{x}_t^{\text{opt}}=\sum_{i=1}^{N}\tilde{c}_t^{(i)}\hat{x}_t^{(i)}, \tag{5}$$

$$P_t^{\text{opt}}=\sum_{i=1}^{N}\tilde{c}_t^{(i)}[P_t^{(i)}+(\hat{x}_t^{(i)}-\hat{x}_t^{\text{opt}})(\hat{x}_t^{(i)}-\hat{x}_t^{\text{opt}})^{\mathrm{T}}], \tag{6}$$

where $\hat{x}_t^{(i)}=E(x_t|y_0^t,\theta_i)$ and $P_t^{(i)}=E[(x_t-\hat{x}_t^{(i)})(x_t-\hat{x}_t^{(i)})^{\mathrm{T}}|y_0^t,\theta_i]$ are the local Kalman estimate and corresponding local error covariance, which are determined by the standard continuous–discrete Kalman filter equations matched to linear system (1), (2) at fixed $\theta=\theta_i$, $i=1,\dots,N$ [10,11].

2.2. The continuous–discrete local Kalman filter for $\hat{x}_t^{(i)}$ and $P_t^{(i)}$

The equations for $\hat{x}_t^{(i)}$ and $P_t^{(i)}$ are given by combining two updates: time update (I) and measurement update (II). In the absence of measurements, the optimal filtering (prediction) is given by performing only the time update portion (I) of the algorithm (by setting $H_{t_k}\equiv 0$). We have

I. *Time update between measurements*:

$$\begin{aligned}
&\dot{\hat{x}}_t^{(i)}=F_t^{(i)}\hat{x}_t^{(i)}, \quad t_{k-1}\leqslant t<t_k, \quad k=1,2,\dots,\\
&\hat{x}_0^{(i)}=\bar{x}_0(\theta_i),\\
&\dot{P}_t^{(i)}=F_t^{(i)}P_t^{(i)}+P_t^{(i)}F_t^{(i)^{\mathrm{T}}}+G_t^{(i)}Q_t^{(i)}G_t^{(i)^{\mathrm{T}}},\\
&P_0^{(i)}=P_0(\theta_i).
\end{aligned} \tag{7}$$

Denote a solution of the "time update" differential equations (7) at $t=t_k$ by

$$\hat{x}_{t_k}^{(i)^-}\triangleq\hat{x}_{t_k}^{(i)}, \quad P_{t_k}^{(i)^-}\triangleq P_{t_k}^{(i)}. \tag{8}$$

Then the previous solutions (8) are updated by adding a new measurement y_{t_k} at time $t=t_k$ and we get

II. *Measurement update* ("*jump*") *at times* t_k:

$$\begin{aligned}
&\hat{x}_{t_k}^{(i)}=\hat{x}_{t_k}^{(i)^-}+K_{t_k}^{(i)}(y_{t_k}-H_{t_k}^{(i)}\hat{x}_{t_k}^{(i)^-}),\\
&K_{t_k}^{(i)}=P_{t_k}^{(i)^-}H_{t_k}^{(i)}[H_{t_k}^{(i)}P_{t_k}^{(i)^-}H_{t_k}^{(i)^{\mathrm{T}}}+R_{t_k}^{(i)}]^{-1},\\
&P_{t_k}^{(i)}=[I_n-K_{t_k}^{(i)}H_{t_k}^{(i)}]P_{t_k}^{(i)^-},
\end{aligned} \tag{9}$$

where I_n is the $n\times n$ identity matrix, and

$$\begin{aligned}
&F_k^{(i)}=F_k(\theta_i), \quad G_k^{(i)}=G_k(\theta_i), \quad Q_k^{(i)}=Q_k(\theta_i),\\
&H_k^{(i)}=H_k(\theta_i), \quad R_k^{(i)}=R_k(\theta_i), \quad i=1,\dots,N.
\end{aligned} \tag{10}$$

Note that local Kalman estimates $\hat{x}_t^{(i)}$ are unbiased, i.e.,

$$E[\hat{x}_t^{(i)}] = E[x_t], \quad i = 1, \ldots, N$$

(see, for example, [8–11]).

2.3. Recursive equations for weights

Given $y_0^{t_k}$, the scalar weights $\tilde{c}_t^{(i)} = p(\theta_i|y_0^t)$, $i = 1, \ldots, N$ in (5) and (6) represent a posteriori probabilities of θ_i, which are described by the following recursive formulas [1,2]:

$$\begin{aligned} p(\theta_i|y_0^{t_k}) &= \frac{L_{t_k}(\theta_i)p(\theta_i|y_0^{t_{k-1}})}{\sum_{j=1}^N L_{t_k}(\theta_j)p(\theta_j|y_0^{t_{k-1}})}, \\ p(\theta_i|y_0^t) &= p(\theta_i), \ i = 1, \ldots, N, \\ L_{t_k}(\theta_i) &= |\tilde{P}_{t_k}^{(i)}|^{-1/2} \exp\left[-\frac{1}{2} z_{t_k}^{(i)^{\mathrm{T}}} (\tilde{P}_{t_k}^{(i)})^{-1} z_{t_k}^{(i)}\right], \\ \tilde{P}_{t_k}^{(i)} &= H_{t_k}^{(i)}[F_{t_k}^{(i)} P_{t_{k-1}}^{(i)} F_{t_k}^{(i)^{\mathrm{T}}} + G_{t_k}^{(i)} Q_{t_k}^{(i)} G_{t_k}^{(i)^{\mathrm{T}}}] H_{t_k}^{(i)^{\mathrm{T}}} + R_{t_k}^{(i)}, \\ z_{t_k}^{(i)} &= y_{t_k}^{(i)} - H_{t_k}^{(i)} F_{t_k}^{(i)} \hat{x}_{t_{k-1}}^{(i)}, \ k = 1, 2, \ldots \end{aligned} \tag{11}$$

and then

$$\tilde{c}_t^{(i)} = \begin{cases} p(\theta_i|y_0^{t_{k-1}}), & t_{k_1} \leqslant t < t_k, \\ p(\theta_i|y_0^{t_k}), & t = t_k, \ k = 1, 2, \ldots. \end{cases} \tag{12}$$

The weights (conditional probabilities) $\tilde{c}_t^{(i)} = p(\theta_i|y_0^t)$ in the weighted sums (5) and (6) are often called the "hypothesis conditional probabilities", which satisfy normalization condition

$$\tilde{c}_t^{(1)} + \cdots + \tilde{c}_t^{(N)} = 1. \tag{13}$$

As it was aforementioned, the resulting optimal continuous–discrete AKF (5)–(12) can be a very costly algorithm to implement, since it requires complex calculations of the a posteriori probabilities $p(\theta_i|y_0^{t_k})$ at each time instant $t_k > 0$.

In this paper we develop an alternative SF for system (1)–(4). This filter does not require calculations of the a posteriori probabilities $p(\theta_i|y_0^{t_k})$ at each time instant t_k, and as a consequence it makes the state estimate computationally feasible for online usage. The obtained suboptimal filtering algorithm reduces the computational burden and online computational requirements.

3. The low-complexity SF for continuous–discrete systems

Similarly to the optimal estimate (6), the suboptimal one is also represented by the weighted sum of the local Kalman estimates $\hat{x}_t^{(i)}$,

$$\hat{x}_t^{\text{sub}} = \sum_{i=1}^N c_t^{(i)} \hat{x}_t^{(i)}, \quad \sum_{i=1}^N c_t^{(i)} = I_n, \tag{14}$$

where in contrast to $\tilde{c}_t^{(i)}$ in (5) the new coefficients $c_t^{(1)}, \ldots, c_t^{(N)}$ represent $n \times n$ weight matrices depending only on time instant and being determined from the mean-square criterion

$$\begin{aligned} \min_{c_t^{(i)}} J_t, \quad J_t &= E\|x_t - \hat{x}_t^{\text{sub}}\|^2 \\ &= \mathrm{tr}\{E[(x_t - \hat{x}_t^{\text{sub}})(x_t - \hat{x}_t^{\text{sub}})^{\mathrm{T}}]\}, \end{aligned} \tag{15}$$

where $\mathrm{tr}(A)$ is the trace of a matrix A.

Remark 1. Weighted sums (5) and (14) are different, since the AKF's weights $\tilde{c}_t^{(i)}$ in (5) depend on current measurements, i.e., $\tilde{c}_t^{(i)} = \tilde{c}_t^{(i)}(y_0^t)$, and thus they can be computed only during the experiment, whereas the SF's weights $c_t^{(i)}$ in (14) can be pre-computed with knowing sample times $\{t_k\}$.

Remark 2. The suboptimal estimate $\hat{x}_t^{\text{sub}}$ is unbiased, i.e., $E(\hat{x}_t^{\text{sub}}) = E(x_t)$. We have

$$\begin{aligned} E(\hat{x}_t^{\text{sub}}) = E\left[\sum_{i=1}^N c_t^{(i)} \hat{x}_t^{(i)}\right] &= \left[\sum_{i=1}^N c_t^{(i)} E(\hat{x}_t^{(i)})\right] \\ &= \left[\sum_{i=1}^N c_t^{(i)}\right] E(x_t) = E(x_t). \end{aligned}$$

Theorem 1. (i) *The weights $c_t^{(1)}, \ldots, c_t^{(N)}$ are given by the following linear algebraic equations*:

$$\begin{aligned} &\sum_{i=1}^N c_t^{(i)}[\tilde{P}_t^{(ij)} - \tilde{P}_t^{(iN)}] = 0, \quad j = 1, \ldots, N-1, \\ &\sum_{i=1}^N c_t^{(i)} = I_n. \end{aligned} \tag{16}$$

(ii) *The overall error covariance $P_t^{\text{sub}} = cov(e_t^{\text{sub}}, e_t^{\text{sub}})$, $e_t^{\text{sub}} = x_t - \hat{x}_t^{\text{sub}}$ is given by*

$$P_t^{\text{sub}} = \sum_{i,j=1}^N c_t^{(i)} \tilde{P}_t^{(ij)} c_t^{(j)^{\mathrm{T}}}, \quad \tilde{P}_t^{(ij)} = \sum_{h=1}^N p(\theta_h) P_t^{(hij)}, \tag{17}$$

where

$$P_t^{(hij)} = cov(e_t^{(hi)}, e_t^{(hj)}), \quad e_t^{(hi)} = x_t(\theta_h) - \hat{x}_t^{(i)} \tag{18}$$

are the local error cross-covariances.

The proof of Theorem 1 is given in Appendix Appendix A.

Equations (16), (17) determining the unknown weights $c_t^{(i)}$ and overall error covariance P_t^{sub} in

Theorem 1 depend on the local cross-covariances $P_t^{(hij)}$, which will be given in Theorem 2.

Theorem 2. *The local error cross-covariances* $P_t^{(hij)}$ *in* (18) *can be represented as*

$$P_t^{(hij)} = L_{xx,t}^{(hh)} - L_{x\hat{x},t}^{(hj)} - L_{x\hat{x},t}^{(hi)^{\mathrm{T}}} + L_{\hat{x}\hat{x},t}^{(ij)},$$
$$i,j,h = 1,\ldots,N, \tag{19}$$

where the second-order moments

$$L_{xx,t}^{(ij)} = E[x_t^{(i)} x_t^{(j)^{\mathrm{T}}}], \quad L_{x\hat{x},t}^{(hj)} = [x_t^{(h)} \hat{x}_t^{(j)^{\mathrm{T}}}],$$
$$L_{\hat{x}\hat{x},t}^{(ij)} = E[\hat{x}_t^{(i)} \hat{x}_t^{(j)^{\mathrm{T}}}],$$

are determined by the following equations:

(A) *Equations for* $L_{xx,t}^{(ij)}$:

$$\dot{L}_{xx,t}^{(ij)} = F_t^{(i)} L_{xx,t}^{(ij)} + L_{xx,t}^{(ij)} F_t^{(j)^{\mathrm{T}}} + G_t^{(i)} Q_t^{(i)^{1/2}} [Q_t^{(j)^{1/2}}]^{\mathrm{T}} G_t^{(j)^{\mathrm{T}}},$$
$$L_{xx,0}^{(ij)} = P_0^{(i)^{1/2}} [P_0^{(j)^{1/2}}]^{\mathrm{T}} + \bar{x}_0^{(i)} \bar{x}_0^{(j)^{\mathrm{T}}},$$
$$t_{k-1} \leqslant t \leqslant t_k, \quad k = 1,2,\ldots. \tag{20}$$

(B) *Equations for* $L_{x\hat{x},t}^{(hj)}$:

(I) *Time update between measurements*:

$$\dot{L}_{x\hat{x},t}^{(hj)^-} = F_t^{(h)} L_{x\hat{x},t}^{(hj)^-} + L_{x\hat{x},t}^{(hj)^-} F_t^{(j)^{\mathrm{T}}},$$
$$L_{x\hat{x},0}^{(hj)^-} = \bar{x}_0^{(h)} \bar{x}_0^{(j)^{\mathrm{T}}}, \quad t_{k-1} \leqslant t \leqslant t_k.$$

(II) *Measurement update* ("*jump*") *at times* $t = t_k$:

$$L_{x\hat{x},t_k}^{(hj)} = A_{t_k}^{(h)} L_{x\hat{x},t_k}^{(hj)^-} + L_{xx,t_k}^{(hj)} B_{t_k}^{(j)^{\mathrm{T}}}. \tag{21}$$

(C) *Equations for* $L_{\hat{x}\hat{x},t}^{(hj)}$:

(I) *Time update between measurements*:

$$\dot{L}_{\hat{x}\hat{x},t}^{(ij)^-} = F_t^{(i)} L_{\hat{x}\hat{x},t}^{(ij)^-} + L_{\hat{x}\hat{x},t}^{(ij)^-} F_t^{(j)^{\mathrm{T}}},$$
$$L_{\hat{x}\hat{x},0}^{(ij)^-} = \bar{x}_0^{(i)} \bar{x}_0^{(j)^{\mathrm{T}}}, \quad t_{k-1} \leqslant t \leqslant t_k.$$

(I-I) *Measurement update* ("*jump*") *at times* $t = t_k$:

$$\begin{aligned} L_{\hat{x}\hat{x},t_k}^{(ij)} = {} & A_{t_k}^{(i)} L_{\hat{x}\hat{x},t_k}^{(ij)^-} A_{t_k}^{(j)^{\mathrm{T}}} + B_{t_k}^{(i)} L_{x\hat{x},t_k}^{(ij)^-} A_{t_k}^{(j)^{\mathrm{T}}} \\ & + A_{t_k}^{(i)^{\mathrm{T}}} L_{x\hat{x},t_k}^{(ji)^-} B_{t_k}^{(j)^{\mathrm{T}}} + B_{t_k}^{(i)} L_{xx,t_k}^{(ij)} B_{t_k}^{(j)^{\mathrm{T}}} \\ & + R_{t_k}^{(i)^{1/2}} [R_{t_k}^{(j)^{1/2}}]^{\mathrm{T}} \delta_{ij}, \end{aligned} \tag{22}$$

where $A_{t_k}^{(i)} = I_n - K_{t_k}^{(i)} H_{t_k}^{(i)}$, $B_{t_k}^{(i)} = K_{t_k}^{(i)} H_{t_k}^{(i)}$, *and* $K_{t_k}^{(i)}$ *stands for the local Kalman gains* (9), (10).

The proof of Theorem 2 is given in Appendix B.

The resulting SF is specified by the set of relations (7)–(10) and (14)–(22). The implementation of SF will be discussed in the following section.

4. A real-time implementation of the SF

Let us consider the SF implementation. All suboptimal filtering equations (7)–(10) and (14)–(22) can be divided into two parts. The first part ("Off-line equations") combines equations which do not depend on real measurements y_{t_k}. This part includes equations for local error covariances $P_t^{(i)} = P_t^{(ii)}$, cross-covariances $P_t^{(hij)}$ and $\tilde{P}_t^{(ij)}$, Kalman gains $K_{t_k}^{(i)}$, and finally for weights $c_t^{(i)}$. These equations depend only on the system matrices and noise statistics, and also on the values of the parameter $\theta = \theta_1,\ldots,\theta_N$ and prior probabilities, which are the part of system model (1)–(4). Therefore they may be pre-computed. The second part ("On-line equations") contains equations for the local Kalman estimates and final suboptimal (fusion) estimate, which depend on the current measurements y_{t_k}. We have

Part 1 ("*Off-line equations*"):

$$\dot{P}_t^{(i)} = F_t^{(i)} P_t^{(i)} + P_t^{(i)} F_t^{(i)^{\mathrm{T}}} + G_t^{(i)} Q_t^{(i)} G_t^{(i)^{\mathrm{T}}}, \quad t_{k-1} \leqslant t < t_k,$$
$$P_{t_k}^{(i)^-} \triangleq P_{t_k}^{(i)}, \quad i = 1,\ldots,N, \tag{23a}$$

$$K_{t_k}^{(i)} = P_{t_k}^{(i)^-} H_{t_k}^{(i)} [H_{t_k}^{(i)} P_{t_k}^{(i)^-} H_{t_k}^{(i)^{\mathrm{T}}} + R_{t_k}^{(i)}]^{-1},$$
$$P_{t_k}^{(i)} = [I_n - K_{t_k}^{(i)} H_{t_k}^{(i)}] P_{t_k}^{(i)^-}, \quad i = 1,\ldots,N, \tag{23b}$$

$$\dot{L}_{xx,t}^{(ij)} = F_t^{(i)} L_{xx,t}^{(ij)} + L_{xx,t}^{(ij)} F_t^{(j)^{\mathrm{T}}} + G_t^{(i)} Q_t^{(i)^{1/2}} (Q_t^{(j)^{1/2}})^{\mathrm{T}} G_t^{(j)^{\mathrm{T}}},$$
$$i,j = 1,\ldots,N, \quad t_{k-1} \leqslant t \leqslant t_k, \tag{23c}$$

$$\dot{L}_{x\hat{x},t}^{(hj)^-} = F_t^{(h)} L_{x\hat{x},t}^{(hj)^-} + L_{x\hat{x},t}^{(hj)^-} F_t^{(j)^{\mathrm{T}}},$$
$$h,j = 1,\ldots,N, \quad t_{k-1} \leqslant t \leqslant t_k.$$
$$L_{x\hat{x},t}^{(hj)} = \begin{cases} L_{x\hat{x},t}^{(hj)^-} & \text{at } t_{k-1} \leqslant t \leqslant t_k, \\ A_{t_k}^{(h)} L_{x\hat{x},t_k}^{(hj)^-} + L_{xx,t_k}^{(hj)^-} B_{t_k}^{(j)^{\mathrm{T}}} & \text{at } t = t_k, \end{cases} \tag{23d}$$

$$\dot{L}_{\hat{x}\hat{x},t}^{(ij)^-} = F_t^{(i)} L_{\hat{x}\hat{x},t}^{(ij)^-} + L_{\hat{x}\hat{x},t}^{(ij)^-} F_t^{(j)^{\mathrm{T}}},$$
$$i,j = 1,\ldots,N, \quad t_{k-1} \leqslant t \leqslant t_k,$$
$$L_{\hat{x}\hat{x},t}^{(ij)} = \begin{cases} L_{\hat{x}\hat{x},t}^{(ij)}, & \text{at } t_{k-1} \leqslant t \leqslant t_k, \\ A_{t_k}^{(i)} L_{\hat{x}\hat{x},t_k}^{(ij)^-} A_{t_k}^{(i)^{\mathrm{T}}} + B_{t_k}^{(i)} L_{x\hat{x},t_k}^{(ij)^-} A_{t_k}^{(j)^{\mathrm{T}}} + A_{t_k}^{(i)} L_{x\hat{x},t_k}^{(ji)^{\mathrm{T}}} B_{t_k}^{(j)^{\mathrm{T}}} & \\ +B_{t_k}^{(i)} L_{xx,t_k}^{(ij)} B_{t_k}^{(j)^{\mathrm{T}}} + R_{t_k}^{(i)^{1/2}} (R_{t_k}^{(j)^{1/2}})^{\mathrm{T}} \delta_{ij} & \text{at } t = t_k, \end{cases}$$
$$\tag{23e}$$

$$P_t^{(hij)} = L_{xx,t}^{(hh)} - L_{x\hat{x},t}^{(hj)} - L_{x\hat{x},t}^{(hi)^{\mathrm{T}}} + L_{\hat{x}\hat{x},t}^{(ij)}, \quad i,j,h = 1,\ldots,N, \tag{23f}$$

$$\sum_{i=1}^{N} c_t^{(i)}[\tilde{P}_t^{(ij)} - \tilde{P}_t^{(iN)}] = 0, \quad j = 1,\ldots,N-1, \sum_{i=1}^{N} c_t^{(i)} = I_n, \tag{23g}$$

$$P_t^{\mathrm{sub}} = \sum_{i,j=1}^{N} c_t^{(i)} \tilde{P}_t^{(ij)} (c_t^{(j)})^{\mathrm{T}}, \quad \tilde{P}_t^{(ij)} = \sum_{h=1}^{N} p(\theta_h) P_t^{(hij)}. \tag{23h}$$

The solution of equations (23a)–(23h) yields the local Kalman gains $K_{t_k}^{(i)}$, optimal weights $c_t^{(1)},\ldots,c_t^{(N)}$ and overall error covariance P_t^{sub} characterizing accuracy of the SF.

Part 2 ("*On-line equations*"):

$$\dot{\hat{x}}_t^{(i)} = F_t^{(i)} \hat{x}_t^{(i)}, \ t_{k-1} \leqslant t < t_k, \ k = 1,2,\ldots; \quad \hat{x}_{t_k}^{(i)^-} \triangleq \hat{x}_{t_k}^{(i)}, \ i = 1,\ldots,N, \tag{24a}$$

$$\hat{x}_{t_k}^{(i)} = \hat{x}_{t_k}^{(i)^-} + K_{t_k}^{(i)}(y_{t_k} - H_{t_k}^{(i)} \hat{x}_{t_k}^{(i)^-}), \quad k = 1,2,\ldots; \quad i = 1,\ldots,N, \tag{24b}$$

$$\hat{x}_t^{\mathrm{sub}} = \sum_{i=1}^{N} c_t^{(i)} \hat{x}_t^{(i)}, \quad \sum_{i=1}^{N} c_t^{(i)} = I_n, \ t \geqslant 0. \tag{24c}$$

Thus, once the measurement schedule has been settled, the real-time implementation of the SF requires only the computation of the local Kalman estimates $\hat{x}_t^{(1)},\ldots,\hat{x}_t^{(N)}$ (24a) and (24b), and the final suboptimal estimate $\hat{x}_t^{\mathrm{sub}}$ in (24c).

Remark 3. Let us consider the existence and uniqueness conditions of solution of linear matrix equations (16). For this purpose we rewrite equations (16) in the following matrix-block form:

$$X_t M_t = D, \tag{25}$$

where

$$X_t = [\, c_t^{(1)} \quad c_t^{(2)} \quad \cdots \quad c_t^{(N)} \,] \in \mathbf{R}^{n \times nN},$$
$$D = [\, O_n \quad \cdots \quad O_n \quad I_n \,] \in \mathbf{R}^{n \times nN}.$$

$O_n \in \mathbf{R}^{n\times n}$ is zero matrix, and $M_t \in \mathbf{R}^{nN\times nN}$ is block-matrix depending on the local cross-covariances $\tilde{P}_t^{(ij)}$, i.e.,

$$M_t = \begin{bmatrix} M_t^{(1,1)} & \cdots & M_t^{(1,N-1)} & I_n \\ M_t^{(2,1)} & \cdots & M_t^{(2,N-1)} & I_n \\ \vdots & \ddots & \vdots & \vdots \\ M_t^{(N,1)} & \cdots & M_t^{(N,N-1)} & I_n \end{bmatrix},$$
$$M_t^{i,j} = \tilde{P}_t^{(ij)} - \tilde{P}_t^{(iN)} \in \mathbf{R}^{n\times n}, \quad i = 1,\ldots,N; \ j = 1,\ldots,N-1. \tag{26}$$

Note that the matrix M_t can be pre-computed, since it depends on unconditional local covariances $P_t^{(i)}$, cross-covariances $\tilde{P}_t^{(ij)}$. It is clear that, in order to determine the unknown weighted matrix X_t, the matrix M_t must be non-singular. The singularity property of the matrix M_t can be analyzed in advance during SF design.

Remark 4. The knowledge of the local error cross-covariances $P^{(hij)}$ is needed in many techniques for distributed fusion. The cross-covariance $P^{(hij)}$ is a key quantity for the best fusion estimation [19]. In real-life problems, the local error covariances $P_t^{(i)} = P_t^{(ii)}$ are usually known. For instance, in linear Kalman filtering, $P_t^{(i)}$ is described by the Riccati equation [10,11]. However, the local cross-covariances $P^{(hij)}$, $i \neq j$ are usually unknown and they should be determined. Theorem 2 presents explicit equations for their computation in the continuous–discrete filtering problem. Fusion of local uncorrelated estimates in static linear regression models is given in [20].

Remark 5. From the fact that the local estimates $\hat{x}_t^{(i)}$ depend on the discrete measurements y_{t_k}, Eqs. (21) and 22 for the second-order moments $L_{x\hat{x},t}^{(hj)}$ and $L_{\hat{x}\hat{x},t}^{(hj)}$ comprise of two portions: time update (I) and measurement update (II). Whereas Eq. (20) for $L_{xx,t}^{(ij)}$ does not depend on measurement noise statistics at time instants t_k, therefore it represents only time update equation. Eq. (20) may be numerically integrated at all times, including the data arrival instants.

Remark 6. Since θ takes a finite number of values $\theta = \theta_1,\ldots,\theta_N$, the local Kalman estimates $\hat{x}_t^{(i)}$ in the weighted sum (14) are separated. Each estimate $\hat{x}_t^{(i)}$ is found independently of other estimates $\hat{x}_t^{(1)},\ldots,\hat{x}_t^{(i-1)}$, $\hat{x}_t^{(i+1)},\ldots,\hat{x}_t^{(N)}$. Therefore, they can be computed in parallel.

Remark 7. If one of the weight $c_t^{(i)}$ tends to zero ($c_t^{(i)} \to 0$), thereby indicating that the corresponding local Kalman estimate $\hat{x}_t^{(i)}$ makes a small contribution to the weighted sum (14) and it can be discarded.

In several applications, there may be a non-zero probability that the measurement model takes N sensor modes. This kind of problem is called a joint classification–estimation problem [1,2,4,5]. One such application is a joint detection–estimation when the tracking of a target trajectory in space is considered where the target may or may not be present, so that target must be detected as well as the trajectory tracked [1,12]. The SF can be used for these applications.

5. Joint classification—estimation problem

In this section, we consider special linear dynamic systems only with measurement uncertainties

$$\dot{x}_t = F_t x_t + G_t v_t, \quad t \geqslant 0,$$
$$y_{t_k} = H_{t_k}(\theta) x_{t_k} + w_{t_k}. \tag{27}$$

Here we assume that the matrices F_t, G_t, Q_t, P_0, and the initial mean $\bar{x}_0$ are completely known, and the matrices $H_{t_k}(\theta)$ and $R_{t_k}(\theta)$ include the unknown parameter $\theta \in \mathbf{R}^r$, which takes only a finite set of values ("different classes of signals") $\theta \in \{\theta_1, \dots, \theta_N\}$.

In this case the state x_t does not depend on the values θ_h unknown parameter θ , hence the local cross-covariances in (18) take the following simple form

$$P_t^{(hij)} = P_t^{(ij)} = cov(e_t^{(i)}, e_t^{(j)}), \quad e_t^{(hi)} = e_t^{(i)} = x_t - \hat{x}_t^{(i)}.$$

And Theorems 1, 2 for system (27) are specified by the following results:

Theorem 3. *Let $\hat{x}_t^{(i)}$, $i = 1, \dots, N$ be the local Kalman estimates in* (14). *Then*

(i) *The optimal weights $c_t^{(1)}, \dots, c_t^{(N)}$ are given by*

$$\sum_{i=1}^{N} c_t^{(i)} [P_t^{(ij)} - P_t^{(iN)}] = 0, \ j = 1, \dots, N-1,$$
$$\sum_{i=1}^{N} c_t^{(i)} = I_n. \tag{28}$$

(ii) *The overall error covariance P_t^{sub} is of the form*

$$P_t^{\text{sub}} = \sum_{i,j=1}^{N} c_t^{(i)} P_t^{(ij)} c_t^{(j)^{\text{T}}}. \tag{29}$$

Here $P_t^{(ii)} \equiv P_t^{(i)}$ is determined by the standard continuous–discrete Kalman filter equations (7)–(10), *and $P_t^{(ij)}$, $i \neq j$ are given by*

Theorem 4. *The local error cross-covariances $P_t^{(ij)}$, $i \neq j$ in* (27) *are determined by the following equations*:

I. *Time update between measurements*:

$$\dot{P}_t^{(ij)^-} = F_t^{(i)} P_t^{(ij)^-} + P_t^{(ij)^-} F_t^{(j)^{\text{T}}}, \quad P_0^{(ij)^-} = P_0,$$
$$t_{k-1} \leqslant t \leqslant t_{k-1}. \tag{30}$$

II. *Measurement update* ("*jump*") *at times $t = t_k$*:

$$P_{t_k}^{(ij)} = A_{t_k}^{(i)} P_{t_k}^{(ij)^-} A_{t_k}^{(i)^{\text{T}}} + K_{t_k}^{(i)} R_{t_k}^{(i)^{1/2}} [R_{t_k}^{(j)^{1/2}}]^{\text{T}} K_{t_k}^{(j)^{\text{T}}} \delta_{ij}.$$

The proofs of Theorems 3 and 4 are the same as for Theorems 1 and 2, respectively.

6. Examples of application of SF

A comparative experimental analysis of the optimal AKF and SF is demonstrated in two practical examples: the damper harmonic oscillator motion and the vehicle motion constrained to a plane.

6.1. Example 1. Joint detection–estimation: the damper harmonic oscillator motion

System model of the harmonic oscillator is considered in [10, p. 104]:

$$\dot{x}_t = \begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\alpha \end{bmatrix} x_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} v_t, \quad 0 \leqslant t \leqslant 1, \tag{31}$$

where $x_t = [x_{1,t} \ x_{1,t}]^{\text{T}}$, $x_{1,t}$ and $x_{2,t}$ represent position and velocity, respectively, v_t is a scalar zero-mean white Gaussian noise with intensity q, $E(v_t v_s) = q\delta(t-s)$, $x_0 \sim \text{N}(\bar{x}_0, P_0)$.

Position is observed with uncertainty. Then the measurement model is written as

$$y_{t_k} = \theta x_{1,t_k} + w_{t_k},$$
$$t_k = k \triangle t, \quad \triangle t = 0.01, \ k = 1, 2, \dots, 100, \tag{32}$$

where $\{w_{t_k}\}$ is zero-mean white Gaussian sequence, $w_{t_k} \sim \text{N}(0, r)$, and the unknown parameter θ takes only two values, i.e.,

$$\theta = \begin{cases} \theta_1 = 1, & p(\theta_1) = 0.5, \\ \theta_2 = 0, & p(\theta_2) = 0.5. \end{cases}$$

This represents the measurement model which takes two sensor modes with $\theta_1 = 1$ (signal-present) and $\theta_2 = 0$ (signal-absent).

We compare two filters: the optimal AKF

$$\hat{x}_t^{\text{opt}} = \tilde{c}_t^{(1)}\hat{x}_t(\theta_1) + \tilde{c}_k^{(2)}\hat{x}_t(\theta_2),$$
$$\tilde{c}_t^{(i)} = p(\theta_i|y_0^t), \quad i = 1, 2$$

and the SF

$$\hat{x}_t^{\text{sub}} = c_t^{(1)}\hat{x}_t(\theta_1) + c_k^{(2)}\hat{x}_t(\theta_2).$$

In this case $N = 2$, and the solution of linear algebraic equations (28) coincides with the Bar-Shalom and Campo formulas for the optimal combination of two correlated estimates [21],

$$c_t^{(1)} = (P_t^{(22)} - P_t^{(21)})(P_t^{(11)} + P_t^{(22)} - P_t^{(12)} - P_t^{(21)})^{-1},$$
$$c_t^{(2)} = (P_t^{(11)} - P_t^{(12)})(P_t^{(11)} + P_t^{(22)} - P_t^{(12)} - P_t^{(21)})^{-1}.$$

The performance of the SF is expressed in the terms of computation load and loss in estimation accuracy with respect to the AKF. The model parameters, noises statistics, and initial conditions are set to

$$\omega_n^2 = 0.64, \quad \alpha = 0.16, \quad q = 1, \quad r = 0.1,$$
$$\bar{x}_0 = [0.0\ 0.0]^{\mathrm{T}}, \quad P_0 = \mathrm{diag}[2.0\ 1.0].$$

Two cases were considered: in the first case, $\theta_1 = 1$ is the true parameter value in (32); in the second case, $\theta_0 = 0$ is the true parameter value. Figs. 1–4 present the time histories of the filter characteristics for the first case. Such time histories are similar to the second case. In Figs. 1 and 2 we show the mean-square errors (MSEs) for the AKF $P_{ii,t}^{\text{opt}} = E[(x_{i,t} - \hat{x}_{i,t}^{\text{opt}})^2]$ and SAF $P_{ii,t}^{\text{sub}} = E[(x_{i,t} - \hat{x}_{i,t}^{\text{sub}})^2]$, respectively using the Monte-Carlo method with 1000 samples.

As it can be seen from Fig. 1, the difference between the optimal and suboptimal MSEs for position is negligible for the steady-state regime. Also the suboptimal MSE for velocity in Fig. 2 is within a few percent of the optimum one. The numerical simulations were performed using a computer with the following specification: Intel Pentium 4 CPU 2.8 GHz 512 MB RAM. The computation time for evaluation of the suboptimal state estimate $\hat{x}_t^{\text{sub}}$ is 3.8 times less than for optimal estimate $\hat{x}_t^{\text{opt}}$. This is due to the fact that the suboptimal weights $c_t^{(i)}$ are pre-computed. This provides the best balance between the computational efficiency and the desired estimation accuracy. Figs. 3 and 4 show one of the possible samples of the optimal and suboptimal position and velocity estimates. Moreover, suboptimal and optimal estimates are very close to each other. This confirms the accuracy of the proposed filter.

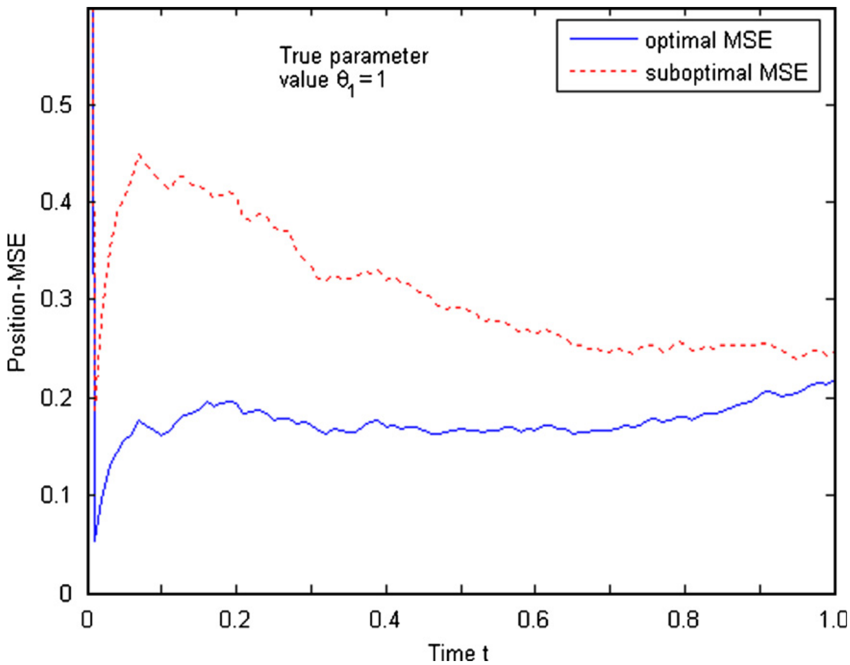

Fig. 1. Comparison of MSEs of the position for 1000 samples: $P_{11,t}^{\text{opt}}$ (solid line); $P_{11,t}^{\text{sub}}$ (dashed line).

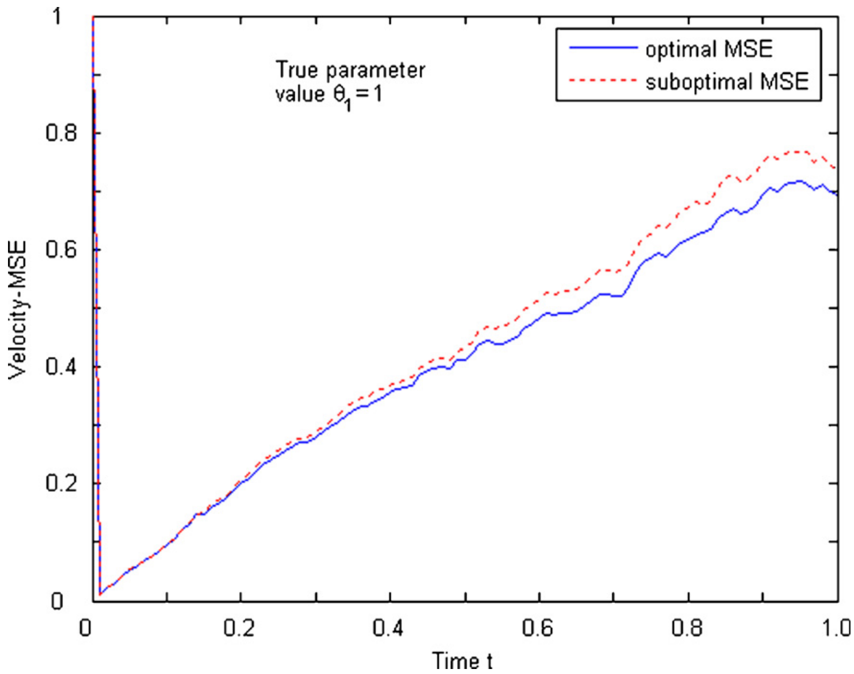

Fig. 2. Comparison of MSEs of the velocity for 1000 samples: $P_{22,t}^{\text{opt}}$ (solid line); $P_{22,t}^{\text{sub}}$ (dashed line).

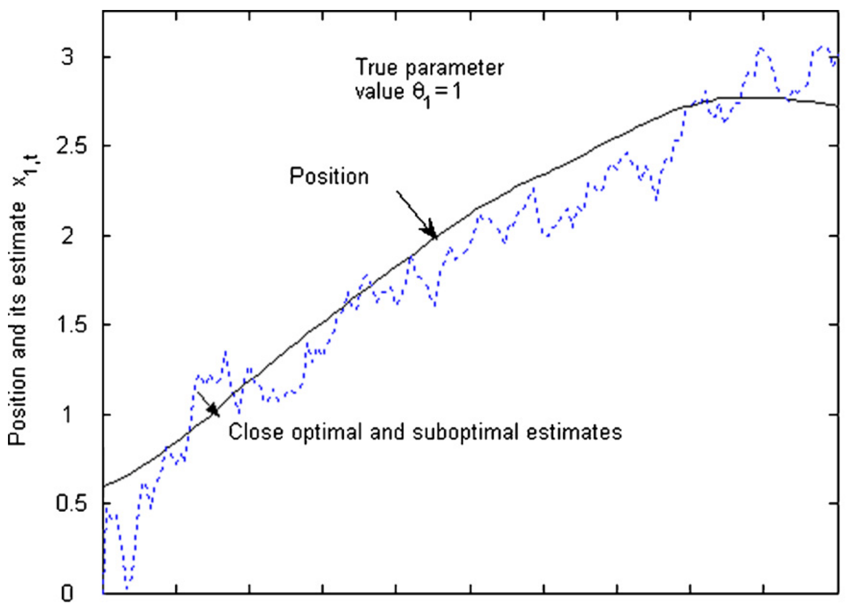

Fig. 3. Optimal and suboptimal position for one sample real position $x_{1,t}$ (solid line); $\hat{x}_{1,t}^{\text{opt}}$ (dotted line); $\hat{x}_{1,t}^{\text{sub}}$ (dashed line).

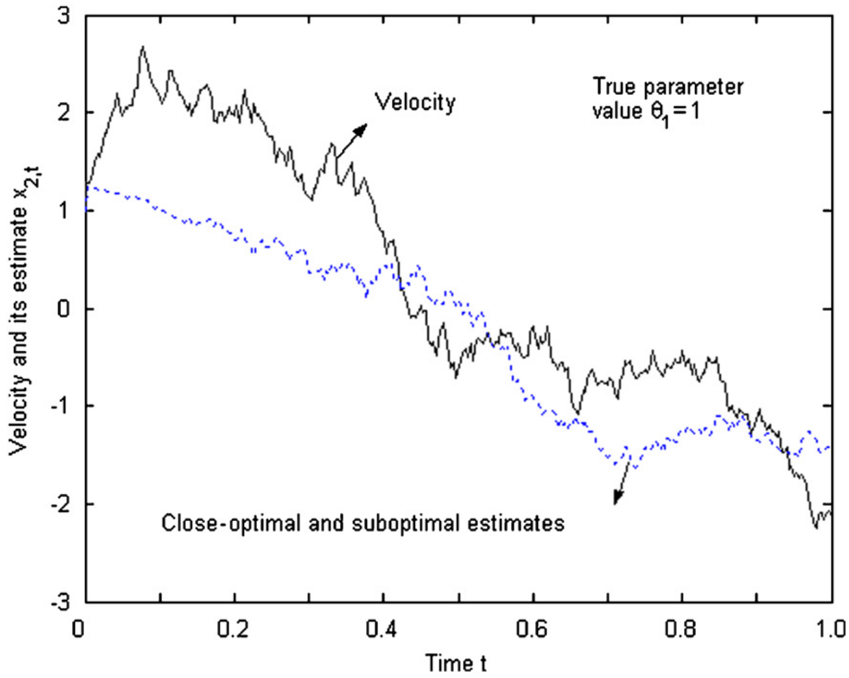

Fig. 4. Optimal and suboptimal velocity for one sample real velocity $x_{2,t}$ (solid line); $\hat{x}_{2,t}^{\text{opt}}$ (dotted line); $\hat{x}_{2,t}^{\text{sub}}$ (dashed line).

6.2. Example 2: the vehicle motion model with unknown initial statistics

Consider a four-dimensional system with state x_t,

$$x_t = \begin{bmatrix} r_x \\ r_y \\ V_x \\ V_y \end{bmatrix} = \begin{bmatrix} x-\text{position} \\ y-\text{position} \\ x-\text{velocity} \\ y-\text{velocity} \end{bmatrix}, \tag{33}$$

which represents a vehicle motion constrained to a plane according to the following equation [1, p. 89]:

$$\dot{r}_x = V_x, \quad \dot{r}_y = V_y, \quad \dot{V}_x = \xi_t, \quad \dot{V}_y = \eta_t,$$

where ξ_t and η_t represent uncorrelated zero-mean white Gaussian noises (random velocities),

$$E(\xi_t\xi_s) = q_\xi\delta(t-s), \;\; E(\eta_t\eta_s) = q_\eta\delta(t-s),$$
$$q_\xi = q_\eta = 10^{-6}.$$

Then state-space description of the model takes the form

$$\dot{x}_t = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} x_t + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} v_t,$$

$$v_t = \begin{bmatrix} \xi_t \\ \eta_t \end{bmatrix}, \quad t \geqslant 0. \tag{34}$$

Assume that the initial state x_0 is zero-mean white Gaussian vector,

$$x_0 \sim \mathbf{N}(\bar{x}_0, P_0). \tag{35}$$

The measurement equation is

$$y_{t_k} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} x_{t_k} + w_{t_k}, \quad k = 1, 2, \ldots, T,$$
$$t_k = k\triangle t, \quad \triangle t = 0.01, \tag{36}$$

where the measurement error $\{w_{t_k} \in \mathbf{R}^2\}$ is zero-mean white Gaussian sequences with covariance $R_{t_k} = \text{diag}(r_1\; r_2)$, and $T = 400$.

We now wish to apply Theorems 1 and 2 to the case where the initial mean $\bar{x}_0 \in \mathbf{R}^4$ is unknown. Let the prior information on $\bar{x}_0$ be given by four hypotheses H_i, $i = 1, 2, 3, 4$,:

$$\bar{x}_0(\text{H}_1) = \bar{x}(\theta_1) = [-14\; 6\; 0.014\; -0.006]^{\text{T}},$$
$$\bar{x}_0(\text{H}_2) = \bar{x}(\theta_2) = [-15\; 7\; 0.018\; -0.005]^{\text{T}},$$
$$\bar{x}_0(\text{H}_3) = \bar{x}(\theta_3) = [-13\; 4\; 0.020\; -0.001]^{\text{T}},$$
$$\bar{x}_0(\text{H}_4) = \bar{x}(\theta_4) = [-12\; 5\; 0.010\; -0.005]^{\text{T}} \tag{37}$$

with the equal prior probabilities $p(\theta_i) = 0.25$ for $i = 1, 2, 3, 4$. The initial covariance P_0 is the same for all hypotheses, i.e., $P_0 = \text{diag}[1\ 1\ 10^{-6}\ 10^{-6}]$.

Here we describe the results of simulations for the two filters in the same manner as example 1: the optimal AKF and SF. Figs. 5 and 6 present the time histories of one sample path of the optimal and suboptimal estimates of the state variables V_x, V_y, whereas Figs. 7 and 8 exhibit the corresponding MSEs for the case with 1000

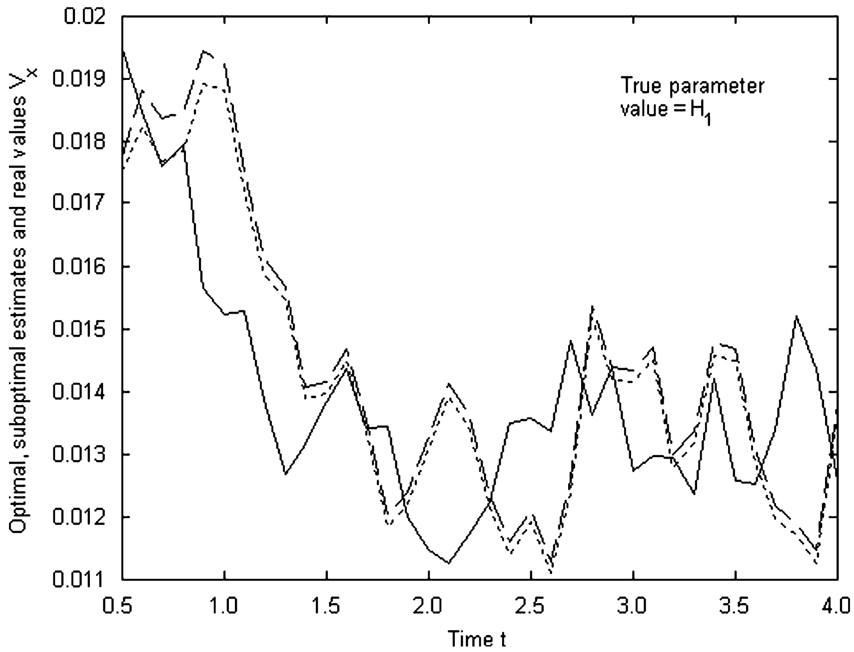

Fig. 5. Optimal and suboptimal velocity for one sample real velocity V_x (solid line); $\hat{V}_x^{\text{opt}}$ (dotted line); $\hat{V}_x^{\text{sub}}$ (dashed line).

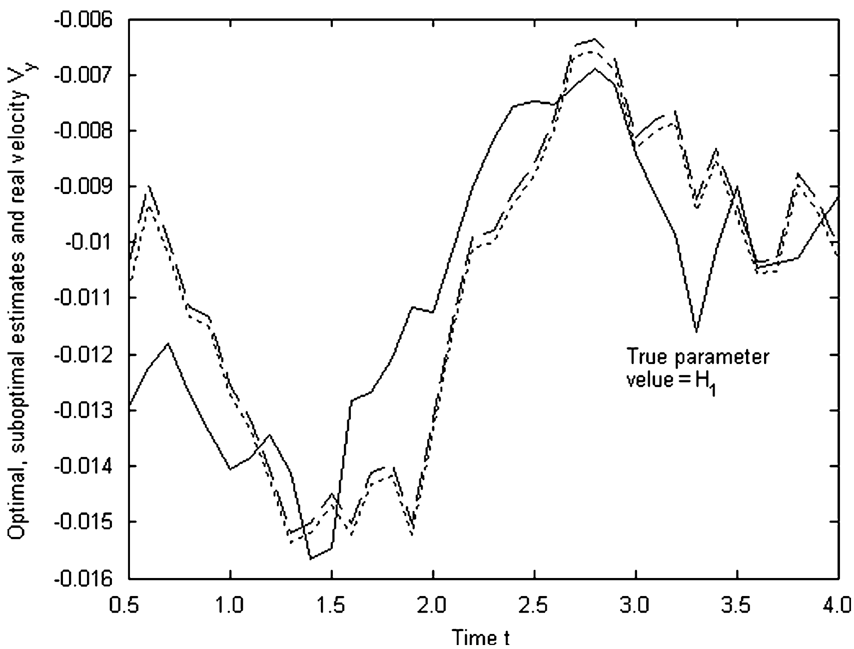

Fig. 6. Optimal and suboptimal velocity for one sample real velocity V_y (solid line); $\hat{V}_y^{\text{opt}}$ (dotted line); $\hat{V}_y^{\text{sub}}$ (dashed line).

samples using the Monte-Carlo method when the true initial mean $\bar{x}_0 = \bar{x}_0(\mathrm{H}_1)$ is assumed to be of hypothesis H_1. Such time histories are perfectly analogous to other hypotheses. The computation time for evaluation of $\hat{x}_t^{\mathrm{sub}}$ is 10.75 times less than for x_t^{opt}.

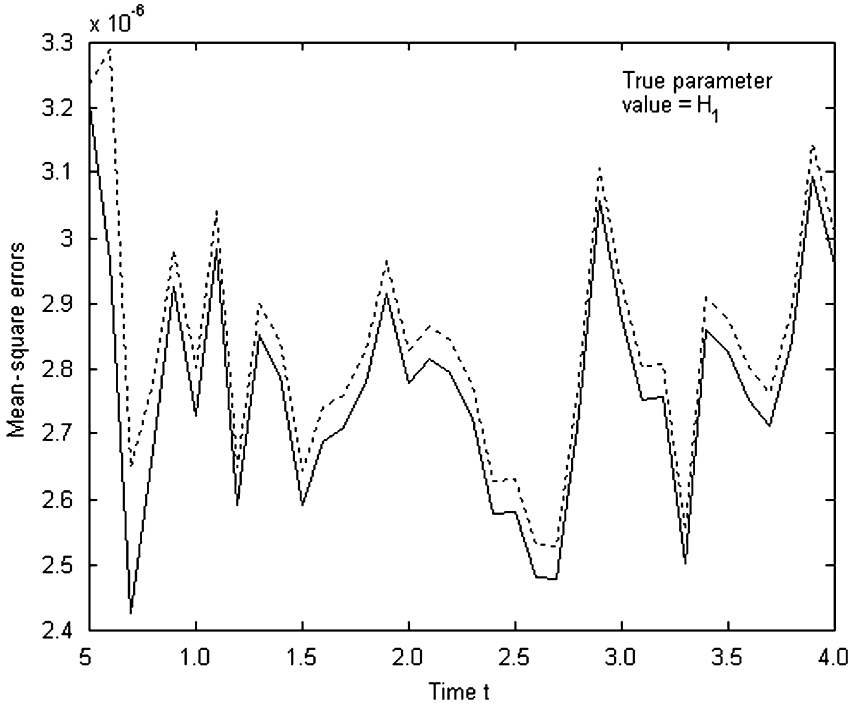

Fig. 7. Comparison of MSEs of the optimal $\hat{V}_x^{\mathrm{opt}}$ (solid line) and suboptimal $\hat{V}_x^{\mathrm{sub}}$ (dotted line) estimates for 1000 samples.

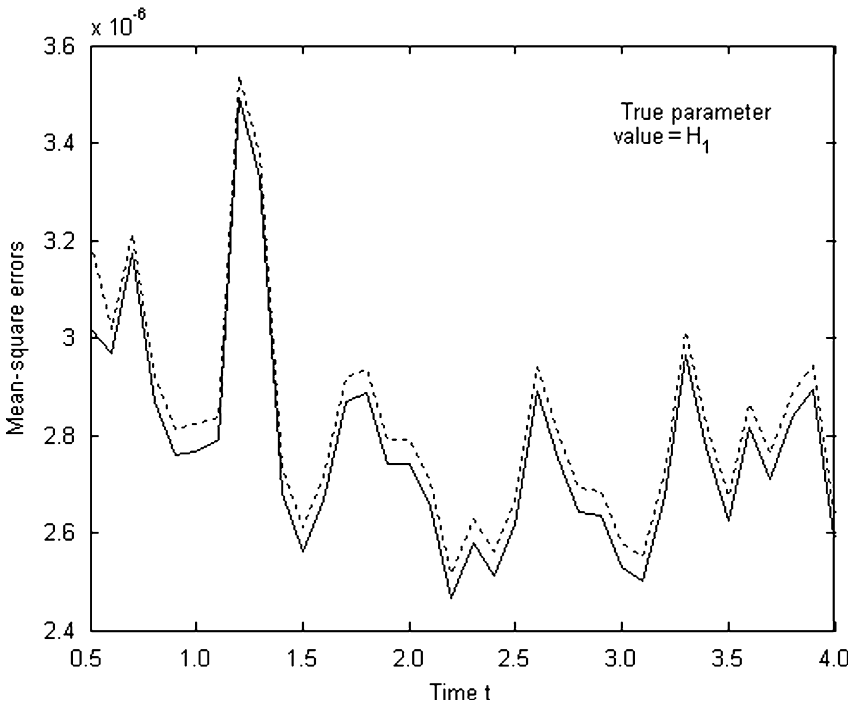

Fig. 8. Comparison of MSEs of the optimal $\hat{V}_y^{\mathrm{opt}}$ (solid line) and suboptimal $\hat{V}_y^{\mathrm{sub}}$ (dotted line) estimates for 1000 samples.

In Figs. 5–8, the comparison of the optimal and suboptimal estimates and the corresponding MSEs show us that performance of the SF is quite similar to the optimal one.

7. Conclusion

In this paper, we have designed a new SF for linear continuous-discrete dynamic systems with uncertainties. This filter represents a linear combination of local Kalman filters with weights depending only on time instances. Each local Kalman filter is fused by the minimum mean-square criterion. The proposed low-complexity filter has a parallel structure and thus it is suitable for parallel processing. Simulation results demonstrate high accuracy of the designed filter.

Appendix A. Proof of Theorem 1

Using (14) and (15), the criterion $J_t = E\|x_t - \hat{x}_t^{\rm sub}\|^2$ can be rewritten as

$$\begin{aligned} J_t &= \sum_{h=1}^{N} p(\theta_h) E\|x_t(\theta_h) - \hat{x}_t^{\rm sub}\|^2 \\ &= \sum_{h=1}^{N} p(\theta_h)\mathrm{tr}\{E[x_t(\theta_h) - \hat{x}_t^{\rm sub}][x_t(\theta_h) - \hat{x}_t^{\rm sub}]^{\rm T}\} \\ &= \sum_{h=1}^{N} p(\theta_h)\mathrm{tr}\left\{E\left[x_t(\theta_h) - \sum_{i=1}^{N} c_t^{(i)}\hat{x}_t^{(i)}\right]\right. \\ &\quad \left.\times\left[x_t(\theta_h) - \sum_{j=1}^{N} c_t^{(j)}\hat{x}_t^{(j)}\right]^{\rm T}\right\}. \qquad (A.1) \end{aligned}$$

Using the normalization condition

$$\sum_{i=1}^{N} c_t^{(i)} = I_n,$$

we can simplify (A.1) to the following form:

$$\begin{aligned} J_t &= \sum_{h=1}^{N} p(\theta_h)\mathrm{tr}\left\{\left[\sum_{i=1}^{N} c_t^{(i)}(x_t(\theta_h) - \hat{x}_t^{(i)})\right]\right. \\ &\quad \left.\times\left[\sum_{j=1}^{N} c_t^{(j)}(x_t(\theta_h) - \hat{x}_t^{(j)})\right]^{\rm T}\right\} \\ &= \sum_{h=1}^{N} p(\theta_h)\mathrm{tr}\left\{\sum_{i=1}^{N}\sum_{j=1}^{N} c_t^{(i)}(x_t(\theta_h) - \hat{x}_t^{(i)})\right. \\ &\quad \left.\times(x_t(\theta_h) - \hat{x}_t^{(j)})^{\rm T} c_t^{(j)^{\rm T}}\right\} \\ &= \sum_{h=1}^{N} p(\theta_h)\mathrm{tr}\left\{\sum_{i,j=1}^{N} c_t^{(i)} E(e_t^{(hi)} e_t^{(hj)^{\rm T}}) c_t^{(j)^{\rm T}}\right\} \\ &= \mathrm{tr}\left\{\sum_{i,j=1}^{N} c_t^{(i)}\left[\sum_{h=1}^{N} p(\theta_h) E(e_t^{(hi)} e_t^{(hj)^{\rm T}})\right] c_t^{(j)^{\rm T}}\right\} \\ &= \mathrm{tr}\left\{\sum_{i,j=1}^{N} c_t^{(i)} \tilde{P}_t^{(ij)} c_t^{(j)^{\rm T}}\right\}, \qquad (A.2) \end{aligned}$$

where

$$\begin{aligned} &\tilde{P}_t^{(ij)} = \sum_{h=1}^{N} p(\theta_h) E(e_t^{(hi)} e_t^{(hj)^{\rm T}}) = \sum_{h=1}^{N} p(\theta_h) P_t^{(hij)}, \\ &P_t^{(hij)} = E(e_t^{(hi)} e_t^{(hj)^{\rm T}}), \\ &e_t^{(hi)} = x_t(\theta_h) - \hat{x}_t^{(i)}, \quad i,j,h = 1,\ldots,N. \qquad (A.3) \end{aligned}$$

Formulas (A.2) and (A.3) give the overall error covariance (17), (18).

Substituting the expression $c_t^{(N)} = I_n - (c_t^{(1)} + \cdots + c_t^{(N-1)})$ into (A.2), we obtain

$$\begin{aligned} J_t = \mathrm{tr}&\left\{\sum_{i,j=1}^{N-1} c_t^{(i)}\tilde{P}_t^{(ij)} c_t^{(j)^{\rm T}} + \sum_{i=1}^{N-1}[c_t^{(i)}\tilde{P}_t^{(iN)} + \tilde{P}_t^{(Ni)} c_t^{(i)^{\rm T}}]\right. \\ &- \sum_{i,j=1}^{N-1}[c_t^{(i)}\tilde{P}_t^{(iN)} c_t^{(j)^{\rm T}} + c_t^{(j)}\tilde{P}_t^{(Ni)} c_t^{(i)^{\rm T}}] \\ &+ \tilde{P}_t^{(NN)} - \left[\sum_{i=1}^{N-1} c_t^{(i)}\right]\tilde{P}_t^{(NN)} \\ &\left. - \tilde{P}_t^{(NN)}\left[\sum_{j=1}^{N-1} c_t^{(j)}\right]^{\rm T} - \sum_{i,j=1}^{N-1} c_t^{(i)}\tilde{P}_t^{(NN)} c_t^{(j)^{\rm T}}\right\}. \qquad (A.4) \end{aligned}$$

Further we differentiate each summand of the criterion J_t in (A.4) with respect to $c_t^{(k)}$, $k = 1,\ldots,N-1$ using the formulas

$$\begin{aligned} &\tilde{P}_t^{(ij)} = \tilde{P}_t^{(ji)^{\rm T}}, \quad \tilde{P}_t^{(ii)} = \tilde{P}_t^{(ii)^{\rm T}}, \\ &\frac{\partial}{\partial c_t^{(k)}}[\mathrm{tr}(c_t^{(k)}\tilde{P}_t^{(ij)})] = \tilde{P}_t^{(ij)^{\rm T}}, \quad \frac{\partial}{\partial c_t^{(k)}}[\mathrm{tr}(\tilde{P}_t^{(ij)} c_t^{(k)})] = \tilde{P}_t^{(ij)}, \\ &\frac{\partial}{\partial c_t^{(k)}}[\mathrm{tr}(c_t^{(k)}\tilde{P}_t^{(ij)} c_t^{(k)^{\rm T}})] = c_t^{(k)}[\tilde{P}_t^{(ij)} + \tilde{P}_t^{(ji)}]. \end{aligned}$$

The notation $\partial J_t/\partial c_t^{(k)}$ represents the operator form of the first derivative of a scalar function J_t

with respect to matrix $c_t^{(i)} = (c_{p,q,t}^{(i)}),\ p,q = 1,\dots,n$, i.e.,

$$\frac{\partial J_t}{\partial c_t^{(i)}} = \left(\frac{\partial J_t}{\partial c_{p,q,t}^{(i)}}\right), \quad p,q = 1,\dots,n.$$

For example,

$$\frac{\partial}{\partial c_t^{(k)}} \operatorname{tr}\left[\sum_{i,j=1}^{N-1} c_t^{(i)} \tilde{P}_t^{(ij)} c_t^{(j)^{\mathrm{T}}}\right] = 2\sum_{i=1}^{N-1} c_t^{(i)} \tilde{P}_t^{(ik)},$$

$$\begin{aligned}\frac{\partial}{\partial c_t^{(k)}} \operatorname{tr}&\left[\sum_{i=1}^{N-1} c_t^{(i)} \tilde{P}_t^{(iN)} + c_t^{(j)} \tilde{P}_t^{(Ni)} c_t^{(i)^{\mathrm{T}}}\right] \\ &= \tilde{P}_t^{(Nk)} + \left(\sum_{i=1}^{N-1} c_t^{(i)}\right)\tilde{P}_t^{(Nk)} + \sum_{i=1}^{N-1} c_t^{(i)} \tilde{P}_t^{(iN)} \\ &= 2P_t^{(Nk)} - c_t^{(N)} \tilde{P}_t^{(Nk)} + \sum_{i=1}^{N-1} c_t^{(i)} \tilde{P}_t^{(iN)},\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial c_t^{(k)}} \operatorname{tr}&\left[\sum_{i,j=1}^{N-1} c_t^{(i)} \tilde{P}_t^{(iN)} c_t^{(j)^{\mathrm{T}}}\right] \\ &= \sum_{i=1}^{N-1} c_t^{(i)} \tilde{P}_t^{(iN)} + \left(\sum_{i=1}^{N-1} c_t^{(i)}\right)\tilde{P}_t^{(Nk)} \\ &= \sum_{i=1}^{N-1} c_t^{(i)} \tilde{P}_t^{(iN)} + \left(I_n - c_t^{(N)}\right)\tilde{P}_t^{(Nk)}.\end{aligned}$$

Setting the result to zero,

$$\frac{\partial J_t}{\partial c_t^{(i)}} = 0, \quad i = 1,\dots,N-1$$

and after simple manipulations we obtain the linear algebraic equations (16) for unknown weights $c_t^{(1)},\dots,c_t^{(N)}$.

This completes the proof of Theorem 1.

Appendix B. Proof of Theorem 2

Representation (19) immediately follows from (18).

The derivation of equations for $L_{xx,t}^{(ij)}$, and "time update" equations for $L_{x\hat{x},t}^{(ij)^-}$ and $L_{\hat{x}\hat{x},t}^{(ij)^-}$ are based on Eqs. (1) and 7 for state $x_t^{(i)} = x_t(\theta_i)$ and local estimate $\hat{x}_t^{(j)^-} = \hat{x}_t(\theta_j)$, respectively,

$$\begin{aligned}\dot{x}_t^{(i)} &= F_t^{(i)} x_t^{(i)} + G_t^{(i)} Q_t^{(i)^{1/2}} \tilde{v}_t, \\ \dot{x}_t^{(j)^-} &= F_t^{(j)} \hat{x}_t^{(j)^-}, \quad t_{k-1} \leqslant t < t_k,\end{aligned} \tag{B.1}$$

where $\tilde{v}_t$ is zero-mean white Gaussian noise with unit intensity, $E(\tilde{v}_t \tilde{v}_s^{\mathrm{T}}) = I_q \delta(t-s)$. Applying the Lyapunov equation for covariance to linear stochastic equations (B.1) [10,11], we obtain the "time update" ordinary differential equations (20), (21) and (22) for the second-order moments $L_{xx,t}^{(ij)}$, $L_{x\hat{x},t}^{(ij)^-}$, and $L_{\hat{x}\hat{x},t}^{(ij)^-}$, respectively.

Substituting the recursive formula (9),

$$\begin{aligned}\hat{x}_{t_k}^{(i)} &= \hat{x}_{t_k}^{(i)^-} + K_{t_k}^{(i)}(y_{t_k} - H_{t_k}^{(i)} \hat{x}_{t_k}^{(i)}) \\ &= (I_n - K_{t_k}^{(i)} H_{t_k}^{(i)})\hat{x}_{t_k}^{(i)^-} + K_{t_k}^{(i)} H_{t_k}^{(i)} x_{t_k} + R_{t_k}^{(i)^{1/2}} \tilde{w}_{t_k} \\ &= A_{t_k}^{(i)} \hat{x}_{t_k}^{(i)^-} + B_{t_k}^{(i)} x_{t_k} + R_{t_k}^{(i)^{1/2}} \tilde{w}_{t_k},\ \tilde{w}_{t_k} \sim \mathrm{N}(0, I_m)\end{aligned} \tag{B.2}$$

into $E(x_{t_k}^{(i)} \hat{x}_{t_k}^{(j)^{\mathrm{T}}})$ and $E(\hat{x}_{t_k}^{(i)} \hat{x}_{t_k}^{(j)^{\mathrm{T}}})$, we obtain the "measurement update" equations (21) and 22 for the second-order moments $L_{x\hat{x},t}^{(ij)}$ and $L_{\hat{x}\hat{x},t}^{(ij)}$, respectively.

This completes the proof of Theorem 2.

References

[1] K. Watanabe, Adaptive Estimation and Control: Partitioning Approach, Prentice-Hall, New York.

[2] D. Lainiotis, Partitioning: a unifying framework for adaptive systems I: estimation, Proc. IEEE 64 (1976) 1126–1142.

[3] D. Magill, Optimal adaptive estimation of sampled stochastic processes, IEEE Trans. Automat. Control 10 (4) (1965) 434–439.

[4] P. Maybeck, Stochastic Models, Estimation, and Control, vol. 2, Navtech, Arlington, VA.

[5] D. Lainiotis, Partitioned linear estimation algorithm: discrete case, IEEE Trans. Automat. Control 20 (3) (1975) 255–257.

[6] H.W.Sorenson, Parameter Estimation, Marcel Dekker, New York.

[7] J.C. Spall (Ed.), Bayesian Analysis of Time Series and Dynamic Models, Marcel Dekker, New York, 1988.

[8] C. Chui, G. Chen, Kalman Filtering with Real-time Applications, Springer, Berlin, Heidelberg.

[9] R. Stengel, Optimal Control and Estimation, Dover Publications, Inc., New York.

[10] F. Lewis, Optimal Estimation with an Introduction to Stochastic Control Theory, Wiley, New York.

[11] A. Gelb, Applied Optimal Estimation, MIT Press, Cambridge, MA.

[12] X.R.L.Y. Bar-Shalom, T. Kirubarajan, Estimation with Applications to Tracking and Navigation, Wiley, New York.

[13] D. Malladi, J. Speyer, A new approach to multiple model adaptive estimation, in Proceedings of the Conference on Decision and Control, San Diego, CA, USA, 1997, pp. 3460–3467.

[14] J. Vasquez, P. Maybeck, Enhanced motion and sizing of bank in moving-bank mmae, IEEE Trans. Aerosp. Electron. Syst. 40 (3) (2004) 770–779.

[15] P. Hanlon, P. Maybeck, Interrelationship of single-filter and multiple-model adaptive algorithms, IEEE Trans. Aerosp. Electron. Syst. 34 (3) (1998) 934–946.

[16] D. Kyger, P. Maybeck, Reducing lag in virtual displays using multiple model adaptive estimation, IEEE Trans. Aerosp. Electron. Syst. 34 (4) (1998) 1237–1248.

[17] F. Wang, V. Balakrishnan, Robust adaptive Kalman filters for linear time-varying systems with stochastic parametric uncertainties, in: Proceedings of the Conference on Decision and Control, vol. 34(4), San Diego, CA, USA, 1999, pp. 440–444.

[18] G. Chen, C. Chui, A modified adaptive Kalman filter for real-time applications, IEEE Trans. Aerosp. Electron. Syst. 27 (1) (1991) 149–154.

[19] S.R.H.R. Hashemipour, A. Laub, Decentralized structures for parallel Kalman filtering, IEEE Trans. Automat. Control 33 (1) (1988) 88–94.

[20] J.W.X. Rong Li, Y. Zhu, C. Han, Optimal linear estimation fusion—part i: unified fusion rules, IEEE Trans. Inf. Theory 49 (9) (2003) 2192–2208.

[21] Y. Bar-Shalom, L. Campo, The effect of the common process noise on the two-sensor fused-track covariance, IEEE Trans. Aerosp. Electron. Syst. 22 (11) (1986) 803–805.