Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Suboptimal Kalman Filter for Dual Estimation under Dynamical Uncertainties

Masaya Murata * **Hidehisa Nagano** * **Kunio Kashino** *,**

* *NTT Communication Science Laboratories, NTT Corporation, 3-1, Morinosato Wakamiya, Atsugi-Shi, Kanagawa 243-0198, Japan (Tel: +81 46-240-3653); e-mail: murata.masaya@lab.ntt.co.jp).*
** *National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430 Japan.*

Abstract: This paper presents a method of designing a suboptimal Kalman filter (SKF) for nonlinear systems, including uncertainties on the system parameter dynamics. In general, governing dynamics behind real-world phenomena has significant associated uncertainties and all we can presume is the approximation to system models. We thus need a robust filter that is capable of rapidly following up the unknown system model parameters while adequately estimating the system states as well. The SKF is suboptimal in terms of standard square errors that the KF minimizes at each time step and this suboptimization results in the enhancement of filter's parameter adaptation ability. The nonlinearity of system models is handled by using an unscented transforming statistical linearization without losing the Gaussianity of states to be estimated. We confirm the effectiveness of the proposed filter by numerical simulations.

1. INTRODUCTION

1.1 Formulation of Dual Estimation Problem

This paper deals with the following nonlinear state-space models with uncertainties on the system parameter dynamics:

$$X_t = f(X_{t-1}) + \mu_{t-1}, \;\; \mu_{t-1} \sim N(0, Q + \Delta Q_{t-1}) \quad (1)$$

$$y_t = HX_t + \omega_t, \;\; \omega_t \sim N(0, r) \quad (2)$$

Here, X_t is an L-dimensional random vector and y_t is a scalar observation at time t. X_t is called a state vector and includes system states and parameters as its elements. $f(X_{t-1})$ expresses the system dynamics and since both system states and parameters are unknown, the function becomes nonlinear in general. Sequentially estimating X_t given the past series of the scalar observations $Y_t = \{y_t, y_{t-1}, \cdots, y_1\}$ is our research problem in this paper. H is an observation matrix and μ_{t-1}, ω_t are system and observation noises following Gaussian distributions both with zero means and variance-covariance matrix $Q + \Delta Q_{t-1}$ and variance r, respectively. ΔQ_{t-1} expresses abrupt time-varying uncertainties on the modeling errors for the system parameter dynamics and the actual (true) values are not known prior to the filter execution.

Since the system parameters are supposed not to greatly vary in time, their dynamics is usually expressed by the following random walk models corrupted by small system noises.

$$X_t^p = X_{t-1}^p + \mu_{t-1}^p, \;\; \mu_{t-1}^p \sim N(0, Q^p) \quad (3)$$

Here, X_t^p corresponds to the system parameter elements in X_t and μ_{t-1}^p is the Gaussian noise brought to the above model. However, this simple model is only valid for cases where the system parameters do not often change with time (stationary cases). For the other cases where the parameter values might abruptly or discontinuously change at random times, an additional technique is required to take the unknown temporal behaviors into consideration.

1.2 Covariance Inflation Techniques

The most widely used technique is to add fictitious system noises to eq. (3). The addition of such noises indicates that the random walk models might include relatively large modeling errors, and it is performed to convey the information for the upcoming observation updating procedure. This kind of technique is called a covariance inflation and due to its simplicity, most practitioners tend to use this fictitious noise approach. The only problem is that we never know how large the noise variances should be to compensate for the modeling errors and so the unknown values remain as parameters that users should tune intensively. Such parameter tuning often troubles the filter's users since the parameter values considerably affect the overall resulting filtered state estimates. Our suboptimal Kalman filter (SKF) approach[1], which is designed to become suboptimal in the framework of square errors of the standard KF[2], addresses this problem and liberates the users from troublesome parameter tunings.

The other approaches are the fading-memory filter[3] and H_∞ filter[4][5][6]. The fading-memory filter increases the past observation noise variances and thus the recent observation update is emphasized. By this procedure, although the filter can retain the covariance inflation effect as the fictitious noise approach does, the inflation parameter remains to be tuned. The H_∞ filter is designed to minimize exponential cumulative square errors and thus speeds up the state convergence to the true value. This effect is also promising for tackling the problem of interest, but the parameter that controls how much the convergence is sped up must be pre-specified and there exists the same problem as in the previous approaches. The inappropriate tuning causes unwanted fluctuations in the state estimates. There is another cost design[7] that might be applicable to our research problem, but the filter equation can not be expressed in closed form and thus a numerical approximation is required. We focus on a closed-form filter approach that is capable of rapidly following up the unknown parameter dynamics.

1.3 Robust Filter to Parameter Uncertainties

Our objective here is to design a filter that rapidly tracks the unknown parameter dynamics. In this area there have been a number of investigations proposing a variety of filters insensitive to the parameter uncertainties[8][9][10][11][12][13]. These filters have been expressed in closed forms but derived under the assumption of uncertain linear state-space models, so that they cannot be applied to dual estimation problem because of coupling nonlinearity. In this paper, we focus on the simultaneous estimations of system states and parameters that might exhibit abrupt changes in time.

1.4 Handling Nonlinear System Models

For handling the nonlinear function in eq. (1), the extended Kalman filter (EKF) or unscented Kalman filter (UKF)[14] is widely used. The EKF linearizes the $f(X_{t-1})$ around the current state estimate and thus the state theoretically retains its Gaussianity during the whole time interval. On the other hand, the UKF approximates expectation and covariance matrix of the nonlinearly transformed state and thus the state generally comes to follow non-Gaussian distributions. Since the Gaussianity is mathematically tractable in designing the optimal filter, the EKF is attractive. However, the UKF, by using unscented transformation techniques, generally provides better state estimates than the EKF. Our approach requires the state Gaussianity, so the UKF is not applicable. For this reason, we propose an unscented transforming statistical linearization (UTSL) technique, which is expected to be superior in handling the nonlinear functions over the Taylor series truncation technique of the EKF. The UTSL is a statistical linearization (SL)[15] used in conjunction with unscented transformation (UT) estimations of its expectations. With this technique, the SL becomes applicable to all kinds of nonlinear functions, while the standard SL has closed-form solutions only for a small class of nonlinear functions.

The paper is organized as follows. Section 2 describes our filter approach, including the cost function that the SKF sequentially minimizes. Section 3 explains the SL and the UTSL techniques for handling nonlinear functions and compares them with the Taylor series first-order truncation technique. We summarize our proposed filter algorithm in Section 4 and show numerical simulation results to confirm its effectiveness in Section 5. We then conclude this paper by mentioning the main contributions and findings.

2. SUBOPTIMAL KALMAN FILTER (SKF)

2.1 Partial Covariance Inflation

The SKF is designed to achieve covariance inflation without adding fictitious system noises or fading the effects of the past state estimate as described in the previous section. In so doing, we consider a suboptimal state estimate in the sense of square errors as follows:

$$\bar{X}_t^{\mathrm{sub}} = \bar{X}_t \pm S_t \tag{4}$$

here, $\bar{X}_t$ is an optimal solution that minimizes the square error of $E[(X_t - f(Y_{t-1}))^T(X_t - f(Y_{t-1}))] = E[||\mathbf{e}_t||^2]$, where $\mathbf{e}_t = X_t - f(Y_{t-1})$ and $||\cdot||$ is an Euclidean norm. S_t is a suboptimal term. Then, the estimation error covariance matrix around $\bar{X}_t^{\mathrm{sub}}$ becomes as follows:

$$\begin{aligned}
\bar{P}_t^{\mathrm{sub}} &= \int \big(X_t - (\bar{X}_t \pm S_t)\big)\big(X_t - (\bar{X}_t \pm S_t)\big)^T \\
&\qquad\qquad \times p(X_t|Y_{t-1})dX_t \\
&= \int (X_t - \bar{X}_t)(X_t - \bar{X}_t)^T p(X_t|Y_{t-1})dX_t \\
&\qquad\qquad + S_t S_t^T \int p(X_t|Y_{t-1})dX_t \\
&= \bar{P}_t + S_t S_t^T
\end{aligned} \tag{5}$$

Here, $\bar{P}_t$ is a state estimation error covariance matrix of the $\bar{X}_t$. Since the $\bar{X}_t$ and $\bar{P}_t$ values correspond to those of the standard KF, we confirm that deliberately making the state estimate suboptimal as the $\bar{X}_t^{\mathrm{sub}}$ yields the partial covariance inflation effect with an amount of StS_t^T (see Subsection 2.3 for the definition). We no longer require any fictitious system noise addition or fading procedures and hereafter focus on this suboptimal state estimate.

2.2 Cost Function Design

The next problem is how to determine the suboptimal term S_t. We do this by adopting a new cost function instead of using the standard square error. The cost function is described as follows:

$$J(\mathbf{e}_t) = ||\mathbf{e}_t||^2_{\bar{P}_t^{-1}} + J_\Delta(\mathbf{e}_t), \tag{6}$$

$$\text{where } \; J_\Delta(\mathbf{e}_t) = \begin{cases} \gamma_t & (||\mathbf{e}_t|| \le \Delta/2) \\ 0 & (||\mathbf{e}_t|| > \Delta/2) \end{cases} \tag{7}$$

Here, the first term is a square error weighted by $\bar{P}_t^{-1}$. This weighting is to balance the two terms and also for the subsequent mathematical tractability. The second term represents the suboptimization effect and it has the form of an inverse uniform error with a height of γ_t and an L dimensional super sphere with the radius $\Delta/2$. These two error terms for the one-dimensional state case are depicted in Fig. 1. The height γ_t is defined as follows:

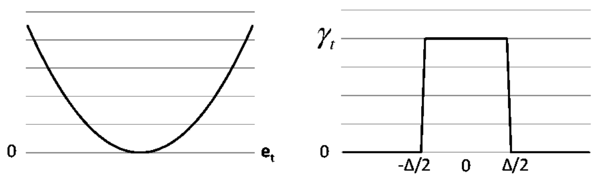

Fig. 1. Cost function: square error (left) and inverse uniform error (right).

$$\gamma_t = \frac{1}{N} \sum_{j=t-N+1}^{j=t} \frac{\nu_j^2}{H\bar{P}_j H^T + r} \tag{8}$$

Here, ν_j is an observation residual at time j ($\nu_j = y_j - H\bar{X}_t$) and $H\bar{P}_j H^T + r$ is its theoretical variance. N is the averaging window. This definition indicates that when large observation residuals are detected, the second term excels and the corresponding suboptimization is exerted. Then, the resulting state estimation error covariance matrix is sufficiently inflated. This filtering behavior can be seen further by investigating the closed-form solution of this cost function.

2.3 Optimal Solution

The optimal solution $\bar{X}_t^{\text{sub}}$ for the cost function in eqs. (6) and (7) satisfies the following equation.

$$\exp\left(-\frac{1}{2}(\bar{X}_t^{\text{sub}} - \bar{X}_t)^T \bar{P}_t^{-1}(\bar{X}_t^{\text{sub}} - \bar{X}_t)\right) = \frac{2}{h_t} \quad (9)$$

h_t is defined as follows:

$$h_t = \frac{\gamma_t D_L(\Delta)}{(\sqrt{2\pi})^L \sqrt{|\bar{P}_t|}}$$

here, $D_L(\Delta) = \frac{\pi^{L/2}}{\Gamma(L/2+1)}(\Delta/2)^L$ (volume of the super sphere with radius of $\Delta/2$). We next restrict our solution as $\bar{X}_t^{\text{sub}} \approx \bar{X}_t$. Then, eq. (9) becomes the following quadratic form.

$$S_t^T \bar{P}_t^{-1} S_t = 2\left(1 - \frac{2}{h_t}\right) = c_t \quad (10)$$

Since the solution for eq. (10) is indefinite, we focus on the solution whose L_2 norm is minimum, that is, whose distance between $\bar{X}_t^{\text{sub}}$ and $\bar{X}_t$ is minimum. We solve eq. (10) by using the method of Lagrange multipliers. The solution is expressed as follows:

$$\begin{aligned} S_t &= \pm\sqrt{c_t \cdot \beta_{t,\min}} M'_t \\ &= \pm\sqrt{2\beta_{t,\min}} \sqrt{\left(1 - \frac{\tau}{\gamma_t}\right)} M'_t \quad (11) \\ \bar{X}_t^{\text{sub}} &= \bar{X}_t \pm S_t \end{aligned}$$

Here, we set $\tau = \frac{2(\sqrt{2\pi})^L \sqrt{|\bar{P}_t|}}{D_L(\Delta)}$. Note that the $\beta_{t,\min}$ and M'_t denote, respectively, the minimum eigenvalue of $\bar{P}_t$ and the corresponding unit eigenvector. Equation (11) indicates that for $\gamma_t \geq \tau$, the solution is suboptimized with an amount of calculated S_t. Besides that case, the solution keep unchanged to be $\bar{X}_t$, that is, the state estimate of the standard KF. Recalling the definition of γ_t in eq. (8), we see that γ_t theoretically obeys a Gaussian distribution $\mathbf{N}(1, 2/N)$ for sufficiently large N by the use of the central limit theorem. Then, the reasonable setting of τ becomes as follows:

$$\tau = 1 + \kappa\sqrt{2/N}, \quad \text{where } \kappa = 2 \sim 3 \quad (12)$$

Equations (11) and (12) express that when the normalized average observation residual γ_t exceeds its two or three sigma range, the suboptimization is performed. This filtering behavior is suited to our research problem. The sign $\pm$ in eq. (11) both leads to the optimal solution; however, we adopt a sign that produces a lower absolute observation residual $|y_t - H\bar{X}_t^{\text{sub}}|$ at each time step.

We summarize the suboptimal state estimate given observations and its estimation error covariance matrix as follows:

$$\bar{X}_t^{\text{sub}} = \begin{cases} \bar{X}_t \pm \sqrt{2\beta_{t,\min}}\sqrt{1 - \dfrac{\tau}{\gamma_t}} M'_t & (\gamma_t \geq \tau) \\ \bar{X}_t & (\text{otherwise}) \end{cases}$$

$$\bar{P}_t^{\text{sub}} = \begin{cases} \bar{P}_t + 2\beta_{t,\min}\left(1 - \dfrac{\tau}{\gamma_t}\right) M'_t (M'_t)^T & (\gamma_t \geq \tau) \\ \bar{P}_t & (\text{otherwise}) \end{cases}$$

A filter that uses these estimates at time t given observations Y_{t-1} is called the SKF. An important feature of the SKF is that the suboptimization direction indicated by M'_t is aligned with the eigenvector associated with the minimum eigenvalue of $\hat{P}_t$. Recalling the random walk model in eq. (3), we see that the estimation error variances for system parameters tend to be very small compared to those of the system states. Therefore, the suboptimization is always performed to the system parameters, and the covariance inflation indicated by $2\beta_{t,\min}\left(1 - \frac{\tau}{\gamma_t}\right) M'_t (M'_t)^T$ enhances the system parameter adaptation ability. We call such enhancement a partial covariance inflation and it significantly differs from a total covariance inflation of the H_∞ filter. This enabling of a rapid follow-up of the uncertain system parameter dynamics is the primary feature of the SKF.

2.4 Gaussianity of State

A strong assumption behind the derivation of the SKF closed-form solution is a Gaussianity of the state. Since the nonlinear function in eq. (1) easily violates this assumption, we need to consider the linearized model to ensure the Gaussianity of state. The next section describes our UTSL technique, which is expected to produce an approximation model superior to that obtainable with the standard linearization technique.

3. UNSCENTED TRANSFORMING STATISTICAL LINEARIZATION (UTSL)

3.1 Statistical Linearization (SL)

The UTSL calculates the expectations in SL models of the target nonlinear systems using the UT technique. The SL model is obtained by solving the following problem[15].

$$\arg\min_{(A,b)} E\Big[||f(X_{t-1}) - \big(A(X_{t-1} - \hat{X}_{t-1}) + b\big)||^2\Big] \quad (13)$$

Here, $\hat{X}_{t-1}$ is a state estimate at time $t-1$ given observations Y_{t-1}. A and b are an inclination matrix and constant vector, respectively, and the optimal values are expressed as follows[16]:

$$A = E[f(X_{t-1})(X_{t-1} - \hat{X}_{t-1})^T]\hat{P}_{t-1}^{-1} \quad (14)$$

$$b = E[f(X_{t-1})] \quad (15)$$

$\hat{P}_{t-1}$ is the estimation error covariance matrix of the $\hat{X}_{t-1}$. Thus, the SL model for eq. (1) becomes

$$\begin{aligned} &f(X_{t-1}) + \mu_{t-1} \approx \\ &\qquad E[f(X_{t-1})\delta^T]\hat{P}_{t-1}^{-1}\delta + E[f(X_{t-1})] + \mu_{t-1} \quad (16) \end{aligned}$$

where $\delta = X_{t-1} - \hat{X}_{t-1}$.

3.2 SL and First-order Taylor Series Truncation Models

It has recently been confirmed that this approximation model corresponds to a first-order truncation of the Fourier-Hermite series expansion[17]. The approximation accuracy of eq. (16) is generally higher than that of a standard linearization (a first-order truncation model of the Taylor series expansion), and this can be confirmed as follows by assuming the 1st-order differentiability as well as the Gaussianity in state:

$$A = \frac{\partial b}{\partial \hat{X}_{t-1}} = E[\frac{\partial}{\partial X_{t-1}} f(X_{t-1})] \tag{17}$$

Equation (17) indicates that the A is an expected inclination using $p(X_{t-1}|Y_{t-1})$, and compared with the standard inclination $\frac{\partial}{\partial X_{t-1}} f(X_{t-1})|_{X_{t-1}=\hat{X}_{t-1}}$, the information of the state distribution is thus taken into account in the SL linearization. Therefore, for a general class of nonlinear functions, it is expected that when the state estimation covariance matrix is small, the SL model becomes closer to the standard linearized model. We summarize both linearized models as follows:

Standard model

$$f(X_{t-1}) + \mu_{t-1} \approx \left.\frac{\partial f}{\partial X_{t-1}}\right|_{X_{t-1}=\hat{X}_{t-1}} \delta + f(\hat{X}_{t-1}) + \mu_{t-1}$$

SL model

$$f(X_{t-1}) + \mu_{t-1} \approx E\left[\frac{\partial f}{\partial X_{t-1}}\right] \delta + E[f(X_{t-1})] + \mu_{t-1}$$

The expectations in eqs. (14) and (15) can not be expressed in the analytical forms in general. We thus apply the UT technique to obtain these values with demonstrated high accuracy. This approach is called here the unscented transforming statistical linearization (UTSL).

3.3 Unscented Transformation (UT) Estimations

The UT estimates an expectation using $2L+1$ number of sigma points χ and the weights w as follows:

$$E[f(X_{t-1})] = w_0 f(\chi_0) + \sum_{n=1}^{2L} w_n f(\chi_n) \tag{18}$$

$$E[(f(X_{t-1})\delta^T)] = \sum_{n=1}^{2L} w_n f(\chi_n)(\chi_n - \hat{X}_{t-1}) \tag{19}$$

where, $\chi_0 = \hat{X}_{t-1}, \chi_n = \hat{X}_{t-1} + (-1)^n \alpha C(\tilde{n})$,

$$\tilde{n} = \lfloor n/2 \rfloor + n \bmod 2$$

Here $C(\tilde{n})$ expresses the $\tilde{n}$th column vector of the Cholesky matrix of $\hat{P}_{t-1}$ ($\hat{P}_{t-1} = \sum_{n=1}^{2L} C(\tilde{n})C(\tilde{n})^T$). $\lfloor\cdot\rfloor$ and $\cdot \bmod 2$ express the floor function and the remainder of the division by 2. By taking the sigma point weights that satisfy $w_0 + 2Lw_1 = 1$, $2w_1\alpha^2 = 1$, and $2w_1\alpha^4 = 3$, eqs. (18) and (19) becomes approximately correct up to the fourth order in terms of the Taylor series expansion. Therefore, the SL model in eq. (16) becomes feasible in conjunction with these UT expectation estimations. The corresponding UKF parameter and sigma point weights are $\alpha = \sqrt{3}$, $w_0 = 1 - L/3$, $w_n = 1/6, (n = 1, 2, \cdots, 2L)$

The UTSL model of eq. (1) enables us to guarantee the Gaussianity of states and to validate the use of the closed-form solution of the SKF explained in the previous section. We next summarize our proposed filter algorithm that sequentially utilizes the UKF and UTSL techniques.

4. PROPOSED FILTER ALGORITHM

Our filter performs a state estimate suboptimization when the observation residuals exceed a specified limit. With this procedure, the system parameter estimation error covariance matrix is inflated and the Kalman gain regains health. Adequate Kalman gain results in the rapid follow-up for the system parameter changes and thus the filter behaves like a robust filter to the parameter dynamics uncertainties, which is our solution addressing the research problem in this paper. Figure 2 plots the flow chart of the SKF algorithm. As shown in this figure, it consists

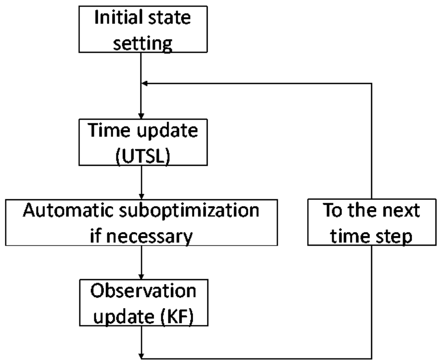

Fig. 2. Flow chart of SKF algorithm.

of an iteration of three main procedures: time update of state, automatic suboptimization of state when required, and observation update of state. The automatic suboptimization yields the partial inflation of the state estimation error covariance matrix, leading to the rapid follow-up for abrupt changes in the system parameters. We call this sequential state estimation algorithm as SKF.

The algorithm equations are summarized as follows:

Time update

$$\chi_0 = \hat{X}_{t-1}$$

$$\chi_n = \hat{X}_{t-1} + (-1)^n \left(\sqrt{3}C(\tilde{n})\right)$$

$$w_n = \begin{cases} 1 - \frac{L}{3}, \text{ for } n = 0 \\ \frac{1}{6}, \text{ for } n \neq 0 \end{cases}$$

$$\bar{X}_t = w_0 f(\chi_0) + \sum_{n=1}^{2L} w_n f(\chi_n)$$

$$F_{\text{UTSL}} = \sum_{n=1}^{2L} w_n f(\chi_n)(\chi_n - \hat{X}_{t-1})^T$$

$$\bar{P}_t = (F_{\text{UTSL}}\hat{P}_{t-1}^{-1})\hat{P}_{t-1}(F_{\text{UTSL}}\hat{P}_{t-1}^{-1})^T + Q$$

Suboptimization

$$\nu_t = y_t - H\bar{X}_t$$

$$\gamma_t = \frac{1}{N}\sum_{j=t-N+1}^{j=t}\frac{\nu_j^2}{H\bar{P}_j H^T + r}$$

$$\tau = 1 + \kappa\sqrt{2/N}$$

$$\bar{X}_t^{\mathrm{sub}} = \begin{cases} \bar{X}_t \pm \sqrt{2\beta_{t,\min}}\sqrt{1-\dfrac{\tau}{\gamma_t}}M_t' & (\gamma_t \geq \tau) \\ \bar{X}_t & (\text{otherwise}) \end{cases}$$

$$\bar{P}_t^{\mathrm{sub}} = \begin{cases} \bar{P}_t + 2\beta_{t,\min}\left(1-\dfrac{\tau}{\gamma_t}\right)M_t'(M_t')^T & (\gamma_t \geq \tau) \\ \bar{P}_t & (\text{otherwise}) \end{cases}$$

Observation update

$$\nu_t^{\mathrm{sub}} = y_t - H\bar{X}_t^{\mathrm{sub}}$$

$$K_t = H\bar{P}_t^{\mathrm{sub}}(H\bar{P}_t^{\mathrm{sub}}H^T + r)^{-1}$$

$$\hat{X}_t = \bar{X}_t^{\mathrm{sub}} + K_t\nu_t^{\mathrm{sub}}$$

$$\hat{P}_t = \bar{P}_t^{\mathrm{sub}} - K_t H\bar{P}_t^{\mathrm{sub}}$$

Here, $\beta_{t,\min}$ and M_t' are the minimum eigenvalue of $\bar{P}_t$ and the unit eigenvector. We adopt a sign from $\pm$ which produces a lower absolute observation residual $|y_t - H\bar{X}_t^{\mathrm{sub}}|$. κ and N are the algorithm design parameters, and considering that κ is a threshold of unusual observation residuals and that N is related to the validation of using the central limit theorem, their recommendation settings are $\kappa = 2 \sim 3, N = 10$. Since these parameter values do not depend on the problem to be solved, our proposed filter is liberated from troublesome parameter tunings. The other approaches mentioned in Section 1 require intensive parameter tunings and since the parameter values considerably affect the overall state estimation results. Our approach is clearly superior to them from the practical viewpoint. We numerically confirm the effectiveness of the proposed filter algorithm in the next section.

5. NUMERICAL SIMULATIONS

5.1 Performance of SKF

This section presents the result of a numerical investigation of our proposed approach. Our filter algorithm is designed to rapidly follow up the uncertain dynamics of system parameters. The example problem taken is as follows:

$$X_t = f(X_{t-1}) + \mu_{t-1}, \quad \mu_{t-1} \sim \mathbf{N}(0, Q)$$

$$y_t = HX_t + \omega_t, \quad \omega_t \sim \mathbf{N}(0, r),$$

$$\text{where } X_t \equiv [x_{1,t}, x_{2,t}, a_t, b_t]^T, \quad Q = \mathrm{diag}(0.2, 0, 0, 0),$$

$$r = 0.1, \quad H = [1, 0, 0, 0]$$

here, the nonlinear function $f(\cdot)$ is expressed as follows:

$$f(X_{t-1}) = \begin{bmatrix} a_{t-1}\sin(x_{1,t-1}) + b_{t-1}x_{2,t-1} \\ x_{1,t-1} \\ a_{t-1} \\ b_{t-1} \end{bmatrix},$$

$$\text{where } \begin{cases} a_t = 1.5, \quad b_t = -0.9 & (t \leq 2000) \\ a_t = 1.0, \quad b_t = -0.5 & (2000 < t \leq 4000) \\ a_t = 1.5, \quad b_t = -0.9 & (4000 < t \leq 6000) \end{cases}$$

The system state dynamics is a second-order Markovian process with known modeling error magnitude and it is expressed as $Q_{11} = 0.2$. On the other hand, the system parameter dynamics is a random walk model with unknown modeling error magnitude. The two parameters significantly change after $t = 2000$ and $t = 4000$, and these changing points of time are also unknown prior to the filter design phase. Therefore, some strategies must be pursued in order to follow up the uncertain parameter dynamics. We compare the system parameter estimation performance of the proposed filter with that obtained by the UKF with the fictitious system noise approach.

Figure 3 plots the observations y_t and the true evolution of system parameters. The observation time series

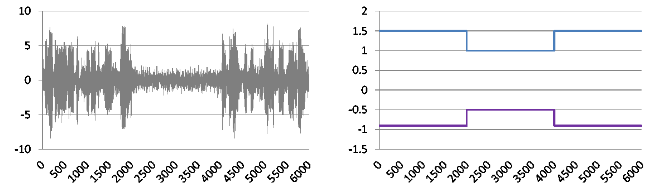

Fig. 3. y_t (left). True parameter dynamics (right).

exhibit two changing patters as the parameters vary. The sequential state estimations were simulated using these observations and the results were averaged over 100 times Monte Carlo runs.

Figure 4 plots the parameter estimation results for the proposed filter and UKF with fictitious system noises. The first three graphs (upper left, upper right, and lower left) are the results for the UKF with tunings of Q_{33}, Q_{44} that are brought to the system parameter dynamics. The last graph (lower right) is the result for the proposed filter. The middle lines at each graph are the filtered estimates

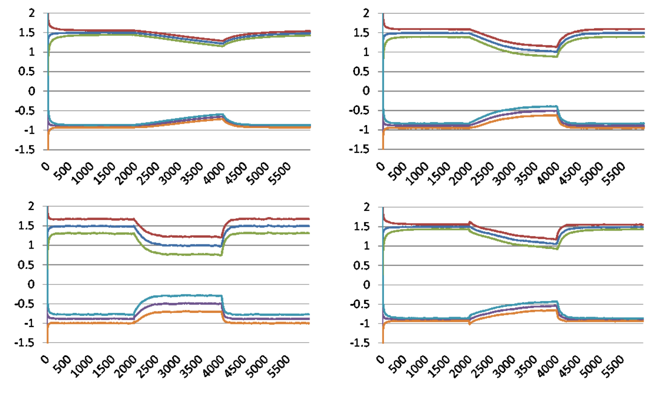

Fig. 4. Results for UKF with fictitious system noises: $Q_{33} = Q_{44} = 10^{-6}$ (upper left): $Q_{33} = Q_{44} = 10^{-5}$ (upper right): $Q_{33} = Q_{44} = 10^{-4}$ (lower left). Results for proposed filter: $Q_{33} = Q_{44} = 0$ (lower right).

and the upper and lower lines indicate the estimation error bars (± 2 sigma). From Fig. 4, the UKF enhances its parameter adaptation ability by adding fictitious system noises to the unknown dynamics. However, this approach easily leads to surplus covariance inflation and prevents the estimation errors from gradually becoming small. This makes the filter always sensitive to the new observations and causes fluctuations in estimation results. Besides that, the estimation results considerably depend on the settings of the fictitious noise variances and appropriate tuning of the settings is essential.

On the other hand, our proposed filter automatically detects significant changes in system parameters and deals with them at the cost function level. Therefore, fictitious noises are not necessary and when the observation residuals are within the normal range, it is expected that the estimation errors becomes sufficiently lowered. The algorithm parameters are set $\kappa = 2, N = 10$ by following the recommendation values in Section 2.

5.2 κ and N dependency

N is the average number in eq. (8) and must be large to theoretically validate the use of the central limit theorem. To investigate the parameter dependency, we show the estimation results with $\kappa = 2, N = 5$ and $\kappa = 2, N = 20$ in Fig. 5. From these figures, we can observe that the

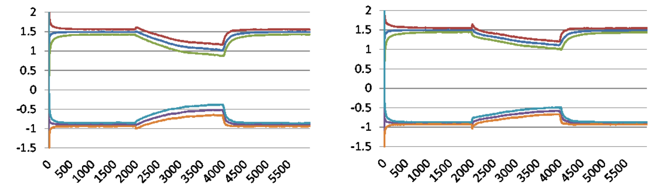

Fig. 5. Results for proposed filter. $N = 5$ (left). $N = 20$ (right).

results do not considerably depend on the setting of N. The recommendation value is thus $N = 10$. However, κ is relatively important parameter since it controls the balance of two cost terms in eqs. (6)-(7). Recall that κ relates to Δ through the following equation.

$$\tau = \frac{2(\sqrt{2\pi})^L \sqrt{|\bar{P}_t|}}{D_L(\Delta)} = 1 + \kappa\sqrt{2/N}$$

Larger κ indicates smaller Δ and smaller Δ contributes less effect of the second cost term (suboptimal term) in eqs. (6)-(7). In other words, when κ is wrongly set small, since the effect of the suboptimal term becomes excessive, suboptimization will be performed frequently, leading to the unstable state estimation results. To investigate the dependency of κ, we show the estimation results with $\kappa = 1, N = 10$ and $\kappa = 3, N = 10$ in Fig. 6. The left

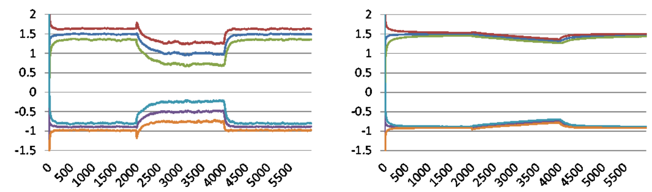

Fig. 6. Results of proposed filter. (Left) $\kappa = 1$. (Right) $\kappa = 3$.

figure shows that the estimation error bars do not become sufficiently small, indicating that the suboptimization was excessively performed even during the observation residuals are within the normal range. To the contrary, the right figure clearly shows that the suboptimization was not performed when needed and thus, the estimation results do not exhibit the rapid parameter follow-up that can be seen in Fig. 3 (lower right). From these results, we need a special attention to set the value κ and when the state dimension becomes small, larger κ (smaller Δ) would be more appropriate since it reduces the effect of the suboptimization term. We thus recommended $\kappa = 2 \sim 3$ in Section 2.

6. CONCLUSION

Governing dynamics behind real-world phenomena has significant uncertainties and all we can presume is approximation models of the uncertainties. The approximation models are composed of unknown system states and parameters, and the system parameters do not generally stay constant over time. We therefore need a robust filter that is capable of rapidly following up the unknown system parameter dynamics and thus yielding the accurate state estimates as well. Our proposed filter is based on a SKF algorithm that requires state Gaussianity and thus the filter uses UTSL models for the nonlinear system dynamics. The UTSL calculates expectations in the standard SL solutions using UT and it is thus applicable to a general class of nonlinear functions. We have confirmed the effectiveness of the proposed approach by numerical simulations.

REFERENCES

[1] M. Murata, H. Nagano and K. Kunio. Robustifying Kalman Filter to Rapidly Adapt to Significant Changes in System Model Parameters of State-Space Models. In Proc. *IEEE Conf. Decision and Control*, to appear, 2013.

[2] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Trans. ASME, J. Basic Eng.*, Vol. 82, pages. 33–45, 1960.

[3] S. L. Fagin. Recursive linear regression theory, optimal filter theory and error analysis of optimal system. *IEEE Conv. Rec.*, vol. 12, pages. 216–245, 1964.

[4] K. M. Nagpal and P. P. Khargonekar. Filtering and Smoothing in an H_∞ Setting. *IEEE Trans. on Automatic Control*, vol.36, no.2, 1991.

[5] R. N. Banavar. A game theoretic approach to linear dynamic estimation. Doctoral Dissertation, University of Texas at Austin, 1992.

[6] J. L. Speyer, C. H. Fan and R. N. Banavar. Optimal Stochastic Estimation with Exponential Cost Criteria. In Proc. *IEEE Conf. Decision and Control*, pages. 2293–2298, 1992.

[7] D. Simon. Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches. *John Wiley & Sons, Inc.*, Hoboken, New Jersey, 2006.

[8] L. Xie, Y. C. Soh and C. E. Souza. Robust Kalman Filtering for Uncertain Discrete-Time Systems. *IEEE Trans. on Automatic Control*, vol.39, no.6, 1994.

[9] Y. Theodor and U. Shaked. Robust Discrete-Time Minimum-Variance Filter. *IEEE Trans. on Automatic Control*, vol.44, no.2, 1996.

[10] A. H. Sayed. A Framework for State-Space Estimation with Uncertain Models. *IEEE Trans. on Automatic Control*, vol.46, no.7, 2001.

[11] A. Subramanian and A. H. Sayed. Regularized Robust Filters for Time-Varying Uncertain Discrete-Time Systems. *IEEE Trans. on Automatic Control*, vol.49, no.6, 2004.

[12] H. Xu and S. Mannor. A Kalman Filter Design Based on the Performance/Robustness Tradeoff. *IEEE Trans. on Automatic Control*, vol.54, no.5, 2009.

[13] S. M. K. Mohamed and S. Nahavandi. Robust Finite-Horizon Kalman Filtering for Uncertain Discrete-Time Systems. *IEEE Trans. on Automatic Control*, vol.57, no.6, 2012.

[14] S. J. Julier and J. K. Uhlmann and H. F. Durrant-Whyte. A New Approach for Filtering Nonlinear Systems. In Proc. *American Control Conference*, pages. 1628–1632, 1995.

[15] Y. Sunahara. An Approximate Method of State Estimation for Nonlinear Dynamical Systems. *Trans. ASME, J. Basic Eng.*, Vol. 92D, No.2, pages. 385–393, 1970.

[16] A. Gelb. Applied Optimal Estimation. *MIT Press.*, 1974.

[17] J. Sarmavuori and S. Sarkka. Fourier-Hermite Kalman Filter. *IEEE Trans. on Automatic Control*, vol.AC-57, no.6, 2012.