Study of Nonlinear Parameter Identification Using UKF and Maximum Likelihood Method

Sun, Zhen; Yang, Zhenyu

Published in:
I E E E Conference on Control Applications. Proceedings

DOI (link to publication from Publisher):
10.1109/CCA.2010.5611170

Publication date:
2010

Document Version
Early version, also known as pre-print

2010 IEEE International Conference on Control Applications
Part of 2010 IEEE Multi-Conference on Systems and Control
Yokohama, Japan, September 8-10, 2010

Study of Nonlinear Parameter Identification Using UKF and Maximum Likelihood Method

Zhen Sun and Zhenyu Yang

Abstract— **The nonlinear parameter identification is studied using UKF and Maximun Likelihood (ML) method. The proposed scheme consists of two sequential stages. The first stage conducts the state estimation using UKF, where the estimated state is a function of unknown parameters. A likelihood function is constructed in the second stage based on the estimated state. Thereby, the parameter identification problem becomes an optimization of the parameterized likelihood function. The proposed method is further compared with EKF based approach. Several case studies show a clear benefit using UKF instead of EKF based approach for a class of nonlinear identification in terms of precision and fast convergence.**

I. INTRODUCTION

Modern systems and equipments are often subjected to significant uncertain factors. Stochastic Differential Equation (SDE) is a natural way to describe the time evolution of dynamic phenomena [6]. The SDE is now widely used in system modeling, analysis and design, for example, in biology, medicine and finance areas. Normally, a SDE model consists of a deterministic part, usually referred to as *drift term*, and a nondeterministic part, usually referred to as *diffusion term*. In general system situation, besides a set of SDE describing the dynamics of the system in continuous time, a set of discrete time measurement equations, is also used to describe a considered system. This kind of hybrid framework provides a general platform for modeling dynamic systems [14].

It is no doubt that the parameter identification for SDE described systems plays a crucial rule in modeling and analysis. The parameter identification of SDE was first studied by Arato, Kolmogorov and Sinai in [8]. Extensive research and results can be found in recent decades, such as method of moments, filter based techniques, statistic methods, and so on [4], [11]. In the last two decades, the Kalman Filter (KF) technique has been more and more used for parameter identification [7]. Generally the approaches using KF can be classified into two different categories. One category is referred to as direct approaches. This kind of approach takes both the state variable and the unknown parameter(s) into an augmented system state. Then, KF or Extended Kalman Filter (EKF) is used to estimate the new state and thereby the estimation of unknown parameter(s) [7]. However, if the diffusion term of the SDE contains unknown parameters, this kind of approach could not be directly used.

The other category is to combine KF technique with some statistic methods. The scheme consists of two sequential stages. The first stage conducts the state estimation using KF, where the estimated state is a function of unknown parameters. Then, a statistic criterion, such as Maximum Likelihood (ML) and Mean Least Square (MLS), is set up in the second stage based on the estimated state. Thereby, the parameter identification problem becomes an optimization of a parameterized statistic problem. This approach can be directly applied to linear systems and an explicit solution may be found [1], [10], [12]. Nevertheless, this kind of approach needs to be extended in order to handle nonlinear cases. Recently, a ML/Prediction Error Decomposition (PED) method for direct estimation of embedded parameters in SDE is proposed in [3] based on the EKF. [9] set up the scheme of parameter identification based on the EKF and ML as well as Maximum A Posteriori (MAP) estimation with software implementation. Both of the two methods can handle parameter identification for cases that the diffusion item consists of the unknown parameter(s).

In 1997, Julier and Uhlman [5] proposed a new type of nonlinear filter, named Unscented Kalman Filter (UKF). It is based on the nonlinear Unscented Transformation (UT). Thereby, UKF does not use the linearization like EKF. The UKF produces a set of selecting points (called sigma-points) in such a way that they together capture the full mean and covariance of the state and make the estimation based on these points. The accuracy of using UKF for nonlinear estimation is investigated in [15], [16]. Normally, the UKF can provide a better estimation than EKF for a wide class of nonlinear systems [13].

In this paper we focus on using UKF and ML method for nonlinear parameter identification. In order to evaluate the proposed method, the approach using EKF and ML method is also carried out for comparison purpose. Several case studies show a clear benefit using UKF instead of EKF based approach for a class of nonlinear identification in terms of precision and fast convergence. The remainder of the paper is organized as follows: The considered problem is formulated in Section II; The method to solve this problem is given in Section III; Section IV illustrates the proposed algorithm via several case studies; Finally, we conclude the paper in Section V.

II. PROBLEM FORMULATION

A. The System Model

The considered system is described by the following SDE:

$$dX(t) = g_1(X(t), u(t), t, \theta)dt + g_2(t, \theta)dB_t, \qquad (1)$$

where $t \in \mathbb{R}$ is the time variable, $X(t) \in \mathscr{X} \subset \mathbb{R}^n$ is a vector of state variables, $u(t) \in \mathscr{U} \subset \mathbb{R}^m$ is a vector of input

Zhen Sun and Zhenyu Yang are with Automation & control, Department of Electronic Systems, Aalborg University, Esbjerg Campus, Niels Bohrs Vej 8, 6700 Esbjerg, Denmark. zhen, yang@aaue.dk

variables, $g_1(\cdot) \in \mathbb{R}^n$, $g_2(\cdot) \in \mathbb{R}^{n\times n}$ are nonlinear or linear functions and $\{B_t\}$ is an n-dimensional standard Wiener process. $\theta \in \Theta$ is the unknown parameter. For simplicity purpose, $X(t), u(t)$ are denoted as X, u respectively.

The measurement of the considered system is described by

$$Y_k = h(X_k, t_k) + \varepsilon_k, \tag{2}$$

where $Y_k \in \mathscr{Y} \subset \mathbb{R}^l$ is a vector of output variables, $h(\cdot) \in \mathbb{R}^l$, t_k, $k = 0, 1, \ldots, N$ are sampling instants, $\{\varepsilon_k\}$ is an l-dimensional white noise process with $\varepsilon_k \sim \mathscr{N}(\mathbf{0}, \mathbf{R})$ ($\mathbf{R}$ is an $l \times l$ matrix) and X_k is the state value at time t_k.

SDE may be interpreted both in the sense of Stratonovich and in the sense of Itô, but since the Stratonovich interpretation is less suitable for parameter identification [6], [10], the Itô interpretation is adapted here. Furthermore, the diffusion term is assumed to be independent of the state variables at this stage.

B. Problem Description

The considered parameter identification problem could be described as:

(P): Estimate the unknown parameter θ in the system (1) based on a set of data which consists of some measured output points Y_k generated by (2) and input signal $u(t_k)$.

III. PROBLEM SOLUTION

The scheme and techniques used to solve the problem **(P)** is given in the following.

A. Discretization of Continuous Time System

The stochastic continuous time model is discretized using the Euler method [10]:

$$\begin{aligned} X_j = X_{j-1} &+ g_1(X_{j-1}, u_{j-1}, t_{j-1}, \theta)(t_j - t_{j-1}) \\ &+ g_2(t_{j-1}, \theta)(B_{t_j} - B_{t_{j-1}}), \end{aligned} \tag{3}$$

where $t_j, j = 1, 2, \cdots$ are a discretization of the given time interval, the subscript j stands for the value of the corresponding variable at t_j.

B. Kalman Filter Theory

To solve the problem **(P)**, the KF technique is applied to make the state estimation. Here the KF technique used in the paper is summarized in the following.

Extended Kalman Filter (EKF):

Initialization with: X_0 and P_0.

Time-updated (Prediction):

$$\begin{aligned} &\hat{X}_{k|k-1} = \hat{X}_{k-1|k-1} + g_1(\hat{X}_{k-1|k-1}, u_{k-1}, t_{k-1}, \theta)(t_k - t_{k-1}), \\ &P_{k|k-1} = \Phi_{k-1} P_{k-1|k-1} \Phi_{k-1}^T + 2 g_2 g_2^T(t_k, \theta), \\ &S_k = H_k P_{k|k-1} H_k^T + R, \\ &K_k = P_{k|k-1} H_k^T S_k^{-1}, \end{aligned}$$

Measurement-updated (Update):

$$\begin{aligned} &r_k = Y_k - h(\hat{X}_{k|k-1}, t_k), \\ &\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k r_k, \\ &P_{k|k} = (I - K_k H_k) P_{k|k-1}, \end{aligned}$$

where $k \geq 1$, $g_2^T(\cdot)$ stands for the transpose of $g_2(\cdot)$ and

$$\begin{aligned} &\Phi_k = \frac{\partial (X_k + g_1(X_k, u_k, t_k, \theta)(t_k - t_{k-1}))}{\partial X_k}\Big|_{X_k = \hat{X}_{k|k}}, \\ &H_k = \frac{\partial (h(X_k, t_k))}{\partial X_k}\Big|_{X_k = \hat{X}_{k|k-1}} \end{aligned}$$

Unscented Kalman Filter (UKF):

Initialization with: X_0 and P_0.

The first step consists of creating $2n+1$ sigma-points in such a way that they together captures the full mean and covariance of the augmented state. The χ matrix is chosen to contain these points, and its columns are calculated as follows:

$$\begin{aligned} &\chi_{i,k-1} = X_{k-1}, && i = 0 \\ &\chi_{i,k-1} = X_{k-1} + (\sqrt{(n+\lambda) P_{k-1}})_i, && i = 1, \ldots, n \\ &\chi_{i,k-1} = X_{k-1} - (\sqrt{(n+\lambda) P_{k-1}})_{i-n}, && i = n+1, \ldots, 2n \end{aligned}$$

where subscript i means the i-th column of the square root of the covariance matrix, $k \geq 1$, $\lambda = \alpha^2(n + \kappa) - n$ is a scaling parameter, α determines the spread of the sigma points around X_{k-1} and is usually set to a small positive value to avoid non-local effects (in the examples, α is set to 0.001), κ is a secondary scaling parameter which is usually set to 0.

Each sigma-point is assigned a weight. These weight are derived by comparing the moments of the sigma-points with a Taylor series expansion of the models. The resulting weights for mean and covariance estimates are given:

$$\begin{aligned} &W_0^{(m)} = \tfrac{\lambda}{(n+\lambda)}, \\ &W_0^{(c)} = \tfrac{\lambda}{(n+\lambda)} + (1 - \alpha^2 + \beta), \\ &W_i^{(m)} = W_i^{(c)} = \tfrac{1}{2(n+\lambda)}, && i = 1, \ldots, 2n \end{aligned}$$

where β is used to incorporate prior knowledge of the distribution of X, generally for Gaussian distributions $\beta = 2$ is optimal. The superscript m and c mean that the corresponding weights are used to calculate the mean and covariance of the state.

The filter then predicts next state by propagating the sigma-points through the state and measurement models, and then calculating weighted averages and covariance matrices of the states:

$$\begin{aligned} &\chi_{i,k|k-1} = \chi_{i,k-1} + g_1(\chi_{i,k-1}, u_{k-1}, t_{k-1}, \theta)(t_k - t_{k-1}) \\ &\hat{X}_{k|k-1} = \sum_{i=0}^{2n} W_i^{(m)} \chi_{i,k|k-1} \\ &P_{k|k-1} = \sum_{i=0}^{2n} W_i^{(c)} [\chi_{i,k|k-1} - \hat{X}_{k|k-1}][\chi_{i,k|k-1} - \hat{X}_{k|k-1}]^T \\ &Y_{k|k-1} = h(\chi_{k|k-1}, t_{k-1}) \\ &\hat{Y}_{k|k-1} = \sum_{i=0}^{2n} W_i^{(m)} Y_{i,k|k-1} \end{aligned}$$

The predictions are then updated by: first, calculating the measurement covariance and state-measurement cross correlation matrices, and then, determining the Kalman gain:

$$
\begin{aligned}
P_{YY} &= \sum_{i=0}^{2n} W_i^{(c)}[Y_{i,k|k-1} - \hat{Y}_{k|k-1}][Y_{i,k|k-1} - \hat{Y}_{k|k-1}]^T + R \\
P_{XY} &= \sum_{i=0}^{2n} W_i^{(c)}[\chi_{i,k|k-1} - \hat{X}_{k|k-1}][Y_{i,k|k-1} - \hat{Y}_{k|k-1}]^T \\
K_k &= P_{XY}P_{YY}^{-1} \\
r_k &= (Y_k - \hat{Y}_{k|k-1}) \\
\hat{X}_{k|k} &= \hat{X}_{k|k-1} + K_k r_k \\
P_{k|k} &= P_{k|k-1} - K_k P_{YY} K_k^T
\end{aligned}
$$

The first stage to solve the problem **(P)** is to use the above KF techniques to estimate the state. In the scheme, the UKF is adopted. As a result, the estimated mean and covariance of the state are obtained.

C. Maximum Likelihood Estimation

The second stage is to make the ML estimation of the parameter. Given the model structure in (1) and (2), the ML estimation of the unknown parameter can be determined by finding the parameter θ that maximize the likelihood function of a given sequence of measurements $Y_0, Y_1, \ldots, Y_k, \ldots, Y_N$. Introducing the notation

$$\mathscr{Y}_k = [Y_k, Y_{k-1}, \ldots, Y_1, Y_0],$$

then, the likelihood function becomes the joint probability density, i.e.,

$$L(\theta; \mathscr{Y}_N) = p(\mathscr{Y}_N \mid \theta), \tag{4}$$

or equivalently

$$L(\theta; \mathscr{Y}_N) = \left(\prod_{k=1}^{N} p(Y_k \mid \mathscr{Y}_{k-1}, \theta) \right) p(Y_0 \mid \theta). \tag{5}$$

In order to carry out the optimization of the likelihood function, the state estimation needs to be solved beforehand. For the SDE in (1) is driven by a Brown Motion that is a special Wiener process, and since increments of a Wiener process are Gaussian, it is reasonable to assume the conditional densities can be well approximated by Gaussian densities, which need the means and covariances. Based on the state estimation results, the parameterized likelihood function can be rewritten as

$$L(\theta; \mathscr{Y}_N) = \left(\prod_{k=1}^{N} \frac{\exp(-\frac{1}{2} r_k^T P_{YY}^{-1} r_k)}{\sqrt{\det(P_{YY})}(\sqrt{2\pi})^n} p(Y_0 \mid \theta) \right), \tag{6}$$

where P_{YY} is the covariance matrix of Y, while the same matrix is represented as S_k in EKF, and superscript -1 stands for the inversion of the corresponding matrix.

Then, the considered problem **(P)** converts to an optimization problem which could be described as:

Given a set of measured output $\mathscr{Y}_k$ and input $u(t_k) \in \mathscr{U}$, find θ by solving the optimization problem defined in the following

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \{ -\ln(L(\theta; \mathscr{Y}_N \mid Y_0)) \}. \tag{7}$$

D. Optimization Computing

To solve the nonlinear optimization problem (7), the quasi-Newton method [15] is used in the paper.

E. Scheme

The scheme to solve the problem **(P)** is given:

- Initialization with X_0 and P_0,
- Using UKF to estimate the state,
- Form the Maximum Likelihood function of the parameter based on the result of the state estimation,
- Solve the optimization problem of the Maximum Likelihood function, then get the result of the parameter identification.

IV. ILLUSTRATIVE CASE STUDIES

In the following, the illustration of the proposed method and comparison with EKF based method are conducted through a number of numerical examples.

A. Example 1

The first example we use is like the example 1 in the [9]. The system is described as

$$
\begin{aligned}
d\begin{pmatrix} X^1 \\ X^2 \\ X^3 \end{pmatrix} &= \begin{pmatrix} VX^1 - \frac{UX^1}{X^3} \\ -\frac{VX^1}{Y} + \frac{U(10-X^2)}{X^3} \\ U \end{pmatrix} dt \\
&+ \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} dB_t,
\end{aligned}
$$

where $(X^1, X^2, X^3)^T$ is the state of the system, and

$$V = \theta \frac{X^1}{0.5X^2X^2 + X^2 + 0.03},$$

θ is the system parameter in the drift term of the SDE, U is the input variable. $\sigma_1, \sigma_2, \sigma_3$ are unknown parameters in the diffusion term.

The measurement equation is given as

$$\begin{pmatrix} Y^1 \\ Y^2 \\ Y^3 \end{pmatrix}_k = \begin{pmatrix} X^1 \\ X^2 \\ X^3 \end{pmatrix}_k + \varepsilon_k,$$

where $(Y^1, Y^2, Y^3)^T$ is the measurement of the state, and $\varepsilon_k \sim \mathscr{N}(\mathbf{0}, \mathbf{S})$ with

$$\mathbf{S} = \begin{pmatrix} S_{11} & 0 & 0 \\ 0 & S_{22} & 0 \\ 0 & 0 & S_{33} \end{pmatrix}$$

and $S_{11} = 0.01, S_{22} = 0.001, S_{33} = 0.01$. Note that superscript $i, i = 1,2,3$ stands for the i-th state variable. And the square of X is noted as XX to avoid potential confusion.

The true parameters are assumed as $\theta = 1$, $\sigma_1 = \sigma_2 = \sigma_3 = \sigma = 0.1$, and the initial state is $(1, 0.24495, 1)^T$. The U is a kind of sweeping signal which is plotted in the Fig.1. A set of outputs (100 samples) is generated by simulating the predefined system and the data is plotted in Fig. 2.

Both the EKF and UKF plus ML methods are examined and compared in the following two scenarios.

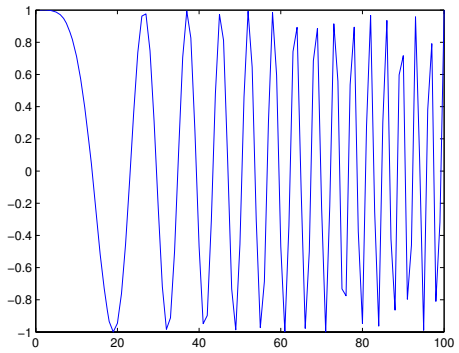

Fig. 1. The input U.

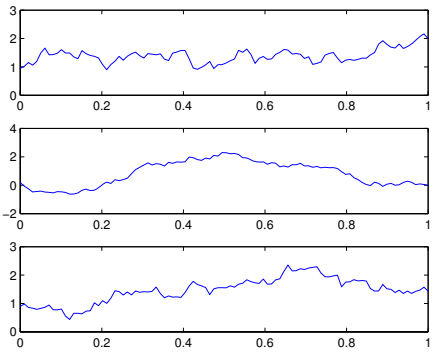

Fig. 2. The real value of the measurement $(Y^1, Y^2, Y^3)^T$.

1) Nominal test, i.e., the data used for identification is generated from the true system.
   - Precision:
     The estimation results are shown in Table I. It can be observed that the parameter estimated using UKF based method is closer to the real value than the situation using EKF based method. Since the model is a nonlinear one, using UKF could provide a more accurate state estimation [16].
   - Convergence issue.
     The convergence properties of the two methods can be judged according to the number of iterations required to reach same tolerant criteria. The tolerant level is selected as $1.0000e-004$ in our concern, and Table II shows iteration numbers of these two approaches. It can be noticed that the UKF plus ML method converges faster than EKF based method does for this example.
   - Computation load.
     The two approaches are implemented under the same computational condition (cpu: Intel Core2 Duo CPU T5900. Memory: 3GB.). The EKF based method needs 4.272164 seconds while UKF based needs 8.672853 seconds. From the computational time point of view, it is clear that UKF based method needs more calculation power than EKF based method does. Since UKF uses $2n+1$ sigma-points and the Cholesky decomposition of the covariance matrix needs to be carried out as well.
2) Robustness test, i.e., the data are generated from the system in which has the modeling error.
   Here the modeling error concerned only happens in variable V. The data is generated according to the new V.
   $$V = \theta \frac{X^1}{0.55X^2X^2 + X^2 + 0.03}.$$
   However, the following estimation still uses the original system model. The convergent values are listed in Table III. It can be observed both results have some deviations compared with "true" identification. Here the criterion to evaluate the robustness is made as:
   $$l_a = \frac{|\,\hat{a} - \hat{a}_e\,|}{\hat{a}},$$
   where $\hat{a}$ is the nominal result of the identification while $\hat{a}_e$ is the result based on the modeling error data (assume a is an unknown parameter of the system). According to this criterion, $l_\theta = 0.0866$, $l_\sigma = 0.1572$ for the EKF based method while $l_\theta = 0.2599$, $l_\sigma = 0.0853$ for the UKF based method. The results show UKF based method has larger deviations than EKF based method. This means that the UKF based method is more sensitive than EKF based method in the parameter identification regarding the modeling error.

TABLE I

THE ESTIMATION RESULTS FOR EXAMPLE 1

Approach	EKF MLE	UKF MLE
θ	1.0422	0.9983
σ	0.0935	0.0984

TABLE II

THE NUMBER OF REQUIRED ITERATIONS FOR EXAMPLE 1

Approach	EKF MLE	UKF MLE
The number of the iteration	73	53

TABLE III

THE ESTIMATED PARAMETER FOR ROBUSTNESS TEST: EXAMPLE 1

Approach	EKF+ML	UKF+ML
θ	1.1325	1.2578
σ	0.1082	0.1115

B. Example 2

In example 2, two senarios are investigated: nonlinear systems described as a polynomial format and a division format. For simplicity, all the systems are simulated in one time unit and the parameter identification is based on 50 sampling points with uniform time intervals of 0.01.

TABLE IV

THE ESTIMATION RESULT FOR EXAMPLE 2-1

Approach	EKF+ML	UKF+ML
θ	0.7729	0.8012
σ	0.1056	0.1045

TABLE V

THE NUMBER OF REQUIRED ITERATIONS EXAMPLE 2-1

Approach	EKF MLE	UKF MLE
The number of the iteration	53	71

TABLE VI

THE ESTIMATION RESULT FOR EXAMPLE 2-2

method	EKF	UKF
θ	0.7881	0.4883
σ	0.0950	0.0973

TABLE VII

THE NUMBER OF REQUIRED ITERATIONS FOR EXAMPLE 2-2

Approach	EKF MLE	UKF MLE
The number of the iteration	82	49

The system is generally described as:

$$\begin{aligned} dX &= f(X,U,\theta)dt + \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} dB_t \\ Y_k &= h(X_k) + \varepsilon_k \end{aligned}$$

where X is the system state, and it is rewritten as $(X^1, X^2, X^3)^T$, U is the input variable. θ is the system unknown parameter, and there is $\sigma_1 = \sigma_2 = \sigma_3 = \sigma$. $Y_k = (Y^1, Y^2, Y^3)_k^T$ is the measurement, $f(\cdot) \in \mathbb{R}^3$, $h(\cdot) \in \mathbb{R}^l, l \leq 3$ are some specific nonlinear or linear functions.

1: The function $f(\cdot)$ is a nonlinear polynomial:

$$f(X,U,\theta) = \begin{pmatrix} X^2X^2X^1 + UX^1 \\ X^3 + UX^2 \\ \theta X^1(X^2 + X^3) + U \end{pmatrix}$$

and the measurement equation is

$$Y_k = X_k^1 + \varepsilon_k$$

with $\varepsilon_k \sim \mathcal{N}(0, 0.1)$. Here the real values are that $\theta = 0.8$, $\sigma = 0.1$, and the initial state is $(1,0,1)^T$. It should be remarked that the system states become partially measurable, i.e., only X^1 is measured, while in the example 1 all system states are directly measured. Similarly as what we do for example 1, the estimated and computing results are listed in Table IV and Table V.

- Parameters estimation (Table IV)
- Number of iteration (Table V)
- Computation load.
  Here the condition of the computation is as the same to the example 1. The EKF method need 2.376364 seconds while UKF need 6.419088 seconds.

2: The function $f(\cdot)$ has simple divisions. The only change to the example 2-1 is that the function $f(\cdot)$ converts to the following function which has simple divisions.

$$f(X,U,\theta) = \begin{pmatrix} X^2X^2/X^3 + UX^1/X^3 \\ \theta X^3/X^2 \\ X^1 + U \end{pmatrix}$$

Here the real value of θ is 0.5, initial state is $(1,1,1)^T$ and other variables are just the same to **1**. Repeat the same process and the results are shown in the below tables (Table VI and VII).

- Parameters estimation (Table VI)
- Number of iteration (Table VII)
- Computation load.
  The EKF based method needs 3.666436 seconds while UKF based method needs 7.084774 seconds under the same computing condition.

In the example 2, the robustness test is not listed because the results have the same conclusion with example 1.

The two case studies of example 2 show almost same results as situations with example 1. In the polynomial case, the two estimation results illustrate UKF based method has better performance than EKF based method, if the computation load won't be a concern. Regarding to the converging property, the EKF based method is a slightly better than UKF based method. However, regarding to the division case, the UKF based method is obviously better than EKF based method without concerning of computational loads. But EKF based method converges much faster than UKF based method.

C. Discussion

Through the above studies, the characteristics of both EKF and UKF based methods are illustrated. In general, the UKF based method can provide more accurate result than EKF based method. Meanwhile, the UKF based method also provides faster converging rate than EKF based method although there are a few special cases. Since the EKF just picks the first order term through linearization of the nonlinear system and drops all items higher than the first order. If the influence of the higher order items can not be ignored in the system, the EKF may provide a poor performance. In contrast, the UKF uses sigma-points that are dedicately chosen. [13] indicated that UKF yields results comparable to a third order approximation of Taylor expansion. As a result, it provide a better estimation to the state of the system. That could be the reason why UKF based method is generally better in parameter estimation. The payoff for better performance of the UKF based method is the computational load. The UKF needs to handle Cholesky decomposition and calculation based on double-sized sigma-points. Moreover, it has been found that the UKF based method is more sensitive than EKF based method regarding to potential modeling errors.

V. CONCLUSIONS AND FUTURE WORKS

A nonlinear parameter identification approach is proposed by combining the UKF and ML method. The comparison of the proposed method and EKF plus ML method is also conducted through a number of case studies. In general,

the proposed method shows a better performance than EKF based approach in terms of precision and converging rate. However, the proposed method requires more computational power and it is more sensitive to potential modeling errors. To extend this method to handle more general cases, e.g., the diffusion term of SDE is state dependent, will be part of our future work.

REFERENCES

[1] H. Madsen, Statistically determined dynamical models for climate processes, PhD thesis, IMSOR, DTH, Lyngby, 1985.

[2] H. Madsen, PhD course materials: System Identification, AAU PhD school, AAU, Aalborg, Denmark, 2009.

[3] H. Melgaard and H. Madsen, CTLSM version 2.6-a program for parameter estimation in stochastic differential equations, Technical report, IMSOR, No. 1, 1993.

[4] L. P. Hansen, Large sample properties of generalized method of moments, *Econometrica*, 50, pp. 1029-1054, 1982.

[5] J. Julier and J. K. Uhlmann, A new extension of the Kalman Filter to nonlinear systems, *Proc. of AeroSense: The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Control*,1997.

[6] K. J. Åström. Introduction to stochastic controltheory. Academic Press, New York, USA, 1970.

[7] L. Ljung, Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems, *Automatic Control, IEEE Transactions*, Vol. 24, 1979, pp 36-50.

[8] M. Arato, A. N. Kolmogorov and Y.G. Sinai, On estimation of parameters of complex stationary Gaussian processes, *Doklady of Academy of Sciences*, 1962.

[9] N. R. Kristensen, H. Madsen and S. B. Jørgensen, Parameter estimation in stochastic grey-box models, *Automatica*, 40, 2004, pp 225-237.

[10] P. E. Kloeden and E. Platen, Numerical solution of stochastic differential equations, Springer, Berlin, Germany, 1992.

[11] P. Young, Parameter estimation for continuous-time models-a survey, *Automatica*, 17, pp. 23-29, 1981.

[12] R. Kalman and R. Bucy, New results in linear prediction and filtering theory, *Journal of Basic Engineering (ASME)*, 83D, 1961.

[13] S. Julier, J. Uhlmann and H. D. Whyte, A new approach for filering nonlinear systems, *Proceedings of the 1995 American Control Conference*, IEEE Press, June 1995.

[14] T. Bohlin and S. F. Graebe, Issues in nonlinear stochastic grey-box identification, *Intrernational Journal of Adaptive Control and Signal Processing*, No. 9, 1995, pp. 465-490.

[15] T. Bohlin, A grey-box process identification tool: Theory and practice, *Technical Report IR-S3-REG-0103*, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, 2001.

[16] W. Eric and V. D. M. Rudolph, The unscented kalman filter for nonlinear estimation, *Proc. of IEEE Symposium 2000 (AS-SPCC)*, Lake Louise, Alberta, Canada, October 2000.