



# Nonlinear filters based on taylor series expansions

Tanizaki, Hisashi  
Roberto S. Mariano

---

(Citation)

Communications in Statistics: Theory and Methods, 25(6):1261-1282

(Issue Date)

1996

(Resource Type)

journal article

(Version)

Accepted Manuscript

(URL)

<https://hdl.handle.net/20.500.14094/90000567>



# NONLINEAR FILTERS BASED ON TAYLOR SERIES EXPANSIONS\*

Hisashi TANIZAKI  
Faculty of Economics  
Kobe University  
Nadaku, Kobe 657, JAPAN

Roberto S. MARIANO  
Department of Economics  
University of Pennsylvania  
Philadelphia, PA 19104, USA

**ABSTRACT:** The nonlinear filters based on Taylor series approximation are broadly used for computational simplicity, even though their filtering estimates are clearly biased. In this paper, first, we analyze what is approximated when we apply the expanded nonlinear functions to the standard linear recursive Kalman filter algorithm. Next, since the state variables  $\alpha_t$  and  $\alpha_{t-1}$  are approximated as a conditional normal distribution given information up to time  $t - 1$  (i.e.,  $I_{t-1}$ ) in approximation of the Taylor series expansion, it might be appropriate to evaluate each expectation by generating normal random numbers of  $\alpha_t$  and  $\alpha_{t-1}$  given  $I_{t-1}$  and those of the error terms  $\epsilon_t$  and  $\eta_t$ . Thus, we propose the Monte-Carlo simulation filter using normal random draws. Finally we perform two Monte-Carlo experiments, where we obtain the result that the Monte-Carlo simulation filter has a superior performance over the nonlinear filters such as the extended Kalman filter and the second-order nonlinear filter.

**KEY WORDS:** Taylor Series Expansion, Extended Kalman Filter, Second-Order Nonlinear Filter, Monte-Carlo Simulation Filter, Random Draws.

## 1 INTRODUCTION

The purpose of this paper is to reconsider the traditional nonlinear filters based on the Taylor series expansion and propose a nonlinear filter generating random numbers.

In the case where the measurement and transition equations depend on the past information, the standard linear recursive Kalman filter algorithm can be derived, when we have the assumptions that both equations in the state-space model are linear and that the error terms are normally distributed (see Harvey (1989)). Unless the distributions of error terms are normal and the measurement and transition equations are linear, we cannot derive the explicit expression for the filtering algorithm. Therefore, some approximation is necessary for estimation. There are

---

\*We are grateful to Francis X. Diebold and Marc Nerlove for valuable suggestions and comments. However, responsibility for any errors remains entirely with us.

two approaches to obtain nonlinear filtering algorithms. One is approximating the nonlinear measurement and transition equations by the Taylor series expansion and applying the linearized nonlinear functions directly to the conventional linear recursive Kalman filter algorithm (see, for example, Anderson and Moore (1979), Gelb (1974), Harvey (1989), Wishner, Tabaczynski and Athans (1969)). Another approach is approximating the underlying density functions of the state vector, where a recursive algorithm on the densities is derived from Bayes' formula (see Alspach and Sorenson (1972), Anderson and Moore (1979), Harvey (1989), Kitagawa (1987), Kramer and Sorenson (1988), Sorenson and Alspach (1971) and Tanizaki (1991)). However, the estimators based on the densities require a great amount of computational burden, compared with those based on the Taylor series approximation of the nonlinear functions, although the density approximation approach gives us asymptotically unbiased filtering estimates. Therefore, in this paper, we consider only the nonlinear filters using the Taylor series expansion.

In the first part of the paper, we discuss what we have to approximate when we apply the Taylor series expansion to the nonlinear functions. There we will see that we encounter some problems with the approximated error terms (which are regarded as the residuals): (i) the expectation of the error terms is not necessarily zero, (ii) the state vector is correlated with the error terms, (iii) the error term in the measurement equation is correlated with the error term in the transition equation, and (iv) the error terms are not normal. Thus, the approximated error terms are not well-behaved. If we approximate the error terms to be well-behaved and apply the linearized functions to the linear recursive algorithm, we have the algorithms based on the Taylor series expansion. Clearly, ignoring these problems of the error terms implies that the filtering estimates of the state vector are biased because  $E(g(x)) \neq g(E(x))$  for a nonlinear function  $g(\cdot)$  and a random variable  $x$ . From the computational point of view, however, the approach based on the Taylor series approximation is broadly used and attractive.

Next, a new nonlinear filtering estimation is proposed, which is based on the Monte-Carlo simulation. We call this nonlinear filter the Monte-Carlo simulation filter. As pointed out in Mariano and Brown (1983, 1989) and Brown and Mariano (1984, 1989), approximating the expectations of nonlinear functions by the Taylor series expansion gives us the biased estimates, and accordingly, we might have a better approximation of the expectations if random numbers are generated for the error terms. Thus, generating normal random numbers to evaluate expectations, we propose the new estimator which is easy and simple in programming and computational time and gives less biased estimates.

Moreover, we sometimes take a functional form such that either the measurement equation or the transition equation is linear in either the state variable or the error term. We show that, under this functional form of the measurement equation or the transition equation, the error terms (or residuals) are uncorrelated with each other. Therefore, in this case, we can have a better approximation.

Finally, three Monte-Carlo experiments are performed to examine the extended Kalman filter, the second-order nonlinear filter and the Monte-Carlo simulation filter. The experiments give the result that the Monte-Carlo simulation filter is

a less biased estimator with respect to the state variable and estimation of unknown parameter, compared with the other two nonlinear filters.

## 2 RECONSIDERATION OF TAYLOR SERIES EXPANSION APPROACH

In this section, we consider what is approximated when we apply the Taylor series expansion to the nonlinear measurement and transition equations. When a nonlinear function is linearized, we have a new approximated error term, which is different from an original one and referred to as the residual. The approximated error is not normally distributed even if the original one follows a normal density. Therefore, one of our purposes in the section is to resolve properties of the approximated error terms.

We treat the nonlinear filtering problem, where the model is specified as follows:

$$\text{(Measurement Equation)} \quad y_t = h_t(\alpha_t, \epsilon_t), \quad (1)$$

$$\text{(Transition Equation)} \quad \alpha_t = g_t(\alpha_{t-1}, \eta_t), \quad (2)$$

$$y_t : g \times 1, \quad \alpha_t : k \times 1, \quad \epsilon_t : g \times 1, \quad \eta_t : k \times 1,$$

where the nonlinear functions  $h_t(\cdot, \cdot)$  and  $g_t(\cdot, \cdot)$  may depend on the other exogenous variables (here, we omit them for simplicity). Assume that  $\epsilon_t$  and  $\eta_t$  are independently distributed as normal random vectors, i.e.,

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} H_t & 0 \\ 0 & Q_t \end{pmatrix} \right).$$

Let  $I_t$  define the information available at time  $t$ , i.e.,  $I_t = \{y_t, y_{t-1}, \dots, y_1\}$ . Consider estimating the state vector  $\alpha_t$  given  $I_t$  or  $I_{t-1}$ . That is, we estimate  $a_{t|s} = E(\alpha_t | I_s)$  and  $\Sigma_{t|s} = \text{Cov}(\alpha_t | I_s)$  for  $s = t-1, t$ .

In the linear case, if the measurement and transition equations depend on the past information  $I_{t-1}$  (i.e., the lagged observed variables), we have to assume normality for the error terms in order to obtain the standard Kalman filter algorithm (see, for example, Anderson and Moore (1979), Chow (1983), Gelb (1974) and Harvey (1989)).

The approach using the Taylor series expansion is common in the literature. In order to apply the standard linear recursive Kalman filter algorithm, the nonlinear functions have to be linearized and the approximated errors have to be assumed as normal random variables. Therefore, from equations (1) and (2), we consider the following expressions which represent the exact relationships, not approximations.

$$\begin{aligned} \text{(Measurement Equation)} \\ y_t &= h_t(\alpha_t, \epsilon_t) \end{aligned} \quad (3)$$

$$= c_{t|t-1} + Z_{t|t-1}(\alpha_t - a_{t|t-1}) + u_t,$$

$$\text{(Transition Equation)}$$

$$\begin{aligned}\alpha_t &= g_t(\alpha_{t-1}, \eta_t) \\ &= d_{t|t-1} + T_{t|t-1}(\alpha_{t-1} - a_{t-1|t-1}) + v_t,\end{aligned}\tag{4}$$

where the measurement equation (3) is transformed to a linear function of  $\alpha_t$ , depending on the information  $I_{t-1}$ . The transition equation (4) is also linear in  $\alpha_{t-1}$ . Here, we can take any functions for  $c_{t|t-1}$ ,  $Z_{t|t-1}$ ,  $d_{t|t-1}$  and  $T_{t|t-1}$ , which depend on the information  $I_{t-1}$ . The expectations of  $u_t$  and  $v_t$  are not necessarily zero. The somewhat *ad hoc* but more general linearization is given in equations (3) and (4). We take such a linearization in order to analyze what has to be approximated in the case of applying the Taylor series expansion to the nonlinear state-space model.

Note that the linear recursive algorithm can be obtained if  $u_t$  and  $v_t$  are normal with zero means. Clearly, however, both error terms  $u_t$  and  $v_t$  are not normally distributed with zero means, because the error terms are represented as the residuals, i.e.,

$$u_t = y_t - c_{t|t-1} - Z_{t|t-1}(\alpha_t - a_{t|t-1})\tag{5}$$

$$= h_t(\alpha_t, \epsilon_t) - c_{t|t-1} - Z_{t|t-1}(\alpha_t - a_{t|t-1}),$$

$$v_t = \alpha_t - d_{t|t-1} - T_{t|t-1}(\alpha_{t-1} - a_{t-1|t-1})\tag{6}$$

$$= g_t(\alpha_{t-1}, \eta_t) - d_{t|t-1} - T_{t|t-1}(\alpha_{t-1} - a_{t-1|t-1}).$$

In the model (3) and (4), there are some problems. The basic assumptions on the error terms are that the state-vector in the measurement equation is uncorrelated with the error term, and that the lagged state-vector in the transition equation is not correlated with the error term (see Harvey (1981, 1989)). This is not the case, however, because  $u_t$  is a function of  $(\alpha_t - a_{t|t-1})$  and also  $v_t$  is a function of  $(\alpha_{t-1} - a_{t-1|t-1})$ . Thus,  $u_t$  is correlated with  $(\alpha_t - a_{t|t-1})$ , and in addition,  $v_t$  is also correlated with  $(\alpha_{t-1} - a_{t-1|t-1})$ . Furthermore,  $u_t$  is correlated with  $v_t$ . Summarizing the above facts, we have the following inequalities:

$$(i) \ E(u_t|I_{t-1}) \neq 0 \text{ and } E(v_t|I_{t-1}) \neq 0,$$

$$(ii) \ E(u_t(\alpha_t - a_{t|t-1})'|I_{t-1}) \neq 0 \text{ and } E(v_t(\alpha_{t-1} - a_{t-1|t-1})'|I_{t-1}) \neq 0,$$

$$(iii) \ E(u_t v_t'|I_{t-1}) \neq 0,$$

Also, we have the problem: (iv)  $u_t$  and  $v_t$  are not normal. Thus, all of the assumptions required for the recursive algorithm are violated in the equations (3) and (4).

Here we consider eliminating the correlation between  $u_t$  and  $(\alpha_t - a_{t|t-1})$  and the correlation between  $v_t$  and  $(\alpha_{t-1} - a_{t-1|t-1})$ , transforming the error terms  $u_t$  and  $v_t$ . The following theorem found in a usual textbook is useful for the transformation.

**Theorem:** Let  $x$  and  $y$  be vectors of random variables, which are distributed as

$$E \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \quad \text{Cov} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma'_{xy} & \Sigma_{yy} \end{pmatrix}.$$

Define  $\tilde{x} = x - \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y)$ . Then,  $(\tilde{x} - \mu_x)$  is uncorrelated with  $(y - \mu_y)$ , i.e.,  $E((\tilde{x} - \mu_x)(y - \mu_y)') = 0$ .

For the measurement equation, to exclude the correlation between  $u_t$  and  $\alpha_t$ , we transform  $u_t$  using the above theorem. The transformed error term  $\tilde{u}_t$  can be represented as:

$$\begin{aligned}\tilde{u}_t &= u_t - y_{t|t-1} + c_{t|t-1} - (M_{t|t-1} - Z_{t|t-1}\Sigma_{t|t-1})\Sigma_{t|t-1}^{-1}(\alpha_t - a_{t|t-1}) \\ &= y_t - y_{t|t-1} - M_{t|t-1}\Sigma_{t|t-1}^{-1}(\alpha_t - a_{t|t-1}),\end{aligned}$$

where

$$\begin{aligned}M_{t|t-1} &= E((y_t - y_{t|t-1})(\alpha_t - a_{t|t-1})'|I_{t-1}), \\ y_{t|t-1} &= E(y_t|I_{t-1}) = E(h_t(\alpha_t, \epsilon_t)|I_{t-1}).\end{aligned}$$

Now, by performing the above transformation, clearly it is shown that we have  $E(\tilde{u}_t|I_{t-1}) = 0$  and  $\text{Cov}(\tilde{u}_t, \alpha_t|I_{t-1}) = 0$ .

Similarly, for the transition equation, we eliminate the correlation between  $v_t$  and  $\alpha_{t-1}$ . The transformed error term  $\tilde{v}_t$  is represented as:

$$\begin{aligned}\tilde{v}_t &= v_t - a_{t|t-1} + d_{t|t-1} - (N_{t|t-1} - T_{t|t-1}\Sigma_{t-1|t-1})\Sigma_{t-1|t-1}^{-1}(\alpha_{t-1} - a_{t-1|t-1}) \\ &= \alpha_t - a_{t|t-1} - N_{t|t-1}\Sigma_{t-1|t-1}^{-1}(\alpha_{t-1} - a_{t-1|t-1}),\end{aligned}$$

where

$$N_{t|t-1} = E((\alpha_t - a_{t|t-1})(\alpha_{t-1} - a_{t-1|t-1})'|I_{t-1}).$$

From the transformation shown above,  $\tilde{u}_t$  is not correlated with  $(\alpha_t - a_{t|t-1})$  and at the same time  $(\alpha_{t-1} - a_{t-1|t-1})$  is uncorrelated with  $\tilde{v}_t$ . Moreover,  $\tilde{u}_t$  and  $\tilde{v}_t$  have zero-mean, i.e.,  $E(\tilde{u}_t|I_{t-1}) = E(\tilde{v}_t|I_{t-1}) = 0$ . The measurement equation and the transition equation in (3) and (4) can then be modified as:

$$\begin{aligned}& \text{(Measurement Equation)} \\ y_t &= y_{t|t-1} + M_{t|t-1}\Sigma_{t|t-1}^{-1}(\alpha_t - a_{t|t-1}) + \tilde{u}_t,\end{aligned}\tag{7}$$

$$\begin{aligned}& \text{(Transition Equation)} \\ \alpha_t &= a_{t|t-1} + N_{t|t-1}\Sigma_{t-1|t-1}^{-1}(\alpha_{t-1} - a_{t-1|t-1}) + \tilde{v}_t.\end{aligned}\tag{8}$$

In equations (7) and (8), the new error terms  $\tilde{u}_t$  and  $\tilde{v}_t$  have less problems: (iii)  $\text{Cov}(\tilde{u}_t, \tilde{v}_t|I_{t-1}) \neq 0$ , and (iv)  $\tilde{u}_t$  and  $\tilde{v}_t$  are nonnormal. That is, the problems (i) and (ii) are avoided, i.e.,  $E(\tilde{u}_t|I_{t-1}) = E(\tilde{v}_t|I_{t-1}) = 0$  and  $E(\tilde{u}_t\alpha_t'|I_{t-1}) = E(\tilde{v}_t\alpha_{t-1}'|I_{t-1}) = 0$ . Hereafter, we call  $\tilde{u}_t$  and  $\tilde{v}_t$  the residuals to distinguish them from the error terms  $\epsilon_t$  and  $\eta_t$ .

In general, however,  $\tilde{u}_t$  is still correlated with  $\tilde{v}_t$  in equations (7) and (8). Furthermore, the residuals  $\tilde{u}_t$  and  $\tilde{v}_t$  are not normally distributed. Ignoring these non-normal error terms and their correlation, we apply equations (7) and (8) to the standard Kalman filter algorithm, and the following algorithm can be obtained.

$$a_{t|t-1} = E(g_t(\alpha_{t-1}, \eta_t)|I_{t-1}),\tag{9}$$

$$\Sigma_{t|t-1} = E((\alpha_t - a_{t|t-1})(\alpha_t - a_{t|t-1})'|I_{t-1}),\tag{10}$$

$$y_{t|t-1} = E(h_t(\alpha_t, \epsilon_t)|I_{t-1}), \quad (11)$$

$$F_{t|t-1} = E((y_t - y_{t|t-1})(y_t - y_{t|t-1})'|I_{t-1}), \quad (12)$$

$$M_{t|t-1} = E((y_t - y_{t|t-1})(\alpha_t - a_{t|t-1})'|I_{t-1}), \quad (13)$$

$$k_t = M_{t|t-1}'F_{t|t-1}^{-1}, \quad (14)$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - k_t F_{t|t-1} k_t', \quad (15)$$

$$a_{t|t} = a_{t|t-1} + k_t(y_t - y_{t|t-1}), \quad (16)$$

for  $t = 1, \dots, T$ .

Note that we have the standard Kalman filter algorithm if the nonlinear functions  $g_t(\cdot, \cdot)$  and  $h_t(\cdot, \cdot)$  are linear in the state variables and the errors. Approximating  $\tilde{u}_t$  and  $\tilde{v}_t$  to be normal is equivalent to approximating  $\alpha_t$  given  $I_t$  or  $I_{t-1}$  to be normal.

The above algorithm is not operational, because the expectations of the nonlinear functions are included in the algorithm. Therefore, next we need to approximate the nonlinear measurement and transition functions in order to evaluate the expectations.

Here, before evaluating the expectations in the above equations (9) – (16), consider the estimation method in the case where unknown parameters (say,  $\theta$  and  $\tau$ ) are included in  $h_t(\cdot, \cdot)$  and  $g_t(\cdot, \cdot)$ , i.e.,  $h_t = h_t(\alpha_t, \epsilon_t; \theta)$  and  $g_t = g_t(\alpha_{t-1}, \eta_t; \tau)$ . In equations (7) and (8), since the residuals  $\tilde{u}_t$  and  $\tilde{v}_t$  are distributed as approximately normal random variables,  $y_t$  is also approximately normal with mean  $y_{t|t-1}$  and variance  $F_{t|t-1}$ , which are given in equations (11) and (12). We have the following approximated innovation form of the likelihood function:

$$\begin{aligned} & P(y_T, y_{T-1}, \dots, y_1) \\ &= P(y_T|I_{T-1})P(y_{T-1}|I_{T-2}) \cdots P(y_2|y_1)P(y_1) \\ &= \prod_{t=1}^T P(y_t|I_{t-1}) \\ &= \prod_{t=1}^T (2\pi)^{-g/2} |F_{t|t-1}|^{-1/2} \exp\left(-\frac{1}{2}(y_t - y_{t|t-1})' F_{t|t-1}^{-1} (y_t - y_{t|t-1})\right). \end{aligned} \quad (17)$$

The unknown parameter vectors  $\theta$  and  $\tau$  are included in  $y_{t|t-1}$  and  $F_{t|t-1}$ . Thus, the likelihood function (17) is maximized with respect to the unknown parameter vectors  $\theta$  and  $\tau$ .

Now, in the above algorithm (9) – (16) and the likelihood function (17), we need to evaluate the expectations. We have made some approximations on  $\tilde{u}_t$  and  $\tilde{v}_t$ . One is assuming that the residuals  $\tilde{u}_t$  and  $\tilde{v}_t$  are approximately normally distributed, and another approximation is no correlation between them. Moreover, we will perform one additional approximation of the above expectations included in equations (9) – (13). The following nonlinear filters are derived by approximating the expectations.

**Extended Kalman Filter and Higher-Order Nonlinear Filters:** In the case where we approximate nonlinear functions  $h_t(\alpha_t, \epsilon_t)$  and  $g_t(\alpha_{t-1}, \eta_t)$  around  $(\alpha_t, \epsilon_t) = (a_{t|t-1}, 0)$  and  $(\alpha_{t-1}, \eta_t) = (a_{t-1|t-1}, 0)$  by the first-order Taylor series expansion, the extended Kalman filter can be obtained by applying the linearized

nonlinear functions to the algorithm (9) – (16) in order to evaluate the expectations. See Anderson and Moore (1979), Gelb (1974), Harvey (1989), and Wishner, Tabaczynski and Athans (1969) for the extended Kalman filter.

It is known that the estimate of  $\alpha_t$  given  $I_t$  is biased because the nonlinearity gives  $E(g(x)) \neq g(E(x))$  for a random variable  $x$  and a function  $g(x)$ . To reduce this problem and obtain the more precise filtered estimates, the second-order approximation can be applied to the nonlinear functions.

When we evaluate each expectation in equations (9) – (16) by using the nonlinear measurement and transition equations approximated by the second-order Taylor series expansion, the second-order nonlinear filtering algorithm can be obtained (see Gelb (1974), Harvey (1989), and Wishner, Tabaczynski and Athans (1969)). In the derivation procedure, the third- and the fourth-moments of  $(\alpha_{t-1} - a_{t-1|t-1})$  and  $(\alpha_t - a_{t|t-1})$  are required. Note that the third-moment of a normal distribution is zero and the fourth-moment is three times the second-moment.

The second-order nonlinear filter might be a less biased estimator than the extended Kalman filter because the bias correction terms (i.e., the second-order terms) are included in the approximation of the nonlinear measurement and transition equations.

As discussed above, we can derive the extended Kalman filter and the second-order nonlinear filter, approximating the nonlinear measurement and transition equations by the first- and the second-order Taylor series expansions with respect to the state vectors and the error terms. It is possible to consider higher-order nonlinear filters, i.e., the third-order nonlinear filter, the fourth-order nonlinear filter, and so on (see Gelb (1974)). In the exactly same fashion, higher-order nonlinear filters can be derived.

Even if the higher-order nonlinear filters give us less biased filtering estimates than the extended Kalman filter, the filtering estimates by higher-order nonlinear filters are still biased because the nonlinear functions are approximated.

Note the following important point: when we approximate the expectations included in the algorithm (9) – (16) by the Taylor series expansions (for example, the first-order approximation, the second-order expansion and so on), the residuals  $\tilde{u}_t$  and  $\tilde{v}_t$  in equations (7) and (8) turn out to have non-zero means and be correlated with each other (see Appendix). Therefore, the problems of the nonlinear filters such as the extended Kalman filter and the second-order nonlinear filters are:

$$E(\tilde{u}_t|I_{t-1}) \neq 0 \text{ and } E(\tilde{v}_t|I_{t-1}) \neq 0, \quad (18)$$

$$E\left(\tilde{u}_t(\alpha_t - a_{t|t-1})'|I_{t-1}\right) \neq 0 \text{ and } E\left(\tilde{v}_t(\alpha_{t-1} - a_{t-1|t-1})'|I_{t-1}\right) \neq 0, \quad (19)$$

$$E(\tilde{u}_t\tilde{v}_t'|I_{t-1}) \neq 0, \quad (20)$$

$$\tilde{u}_t \text{ and } \tilde{v}_t \text{ are not normal.} \quad (21)$$

Accordingly, in the case of the extended Kalman filter and the other higher-order nonlinear filters,  $\tilde{u}_t$  and  $\tilde{v}_t$  have to be assumed as well-behaved error terms in order to apply the usual linear recursive algorithm, even if they are not well-behaved.



**Monte-Carlo Simulation Filter:** Loosening the approximation of expanding the functions  $h_t(\alpha_t, \epsilon_t)$  and  $g_t(\alpha_{t-1}, \eta_t)$  with respect to  $(\alpha_t, \epsilon_t)$  and  $(\alpha_{t-1}, \eta_t)$ , here we propose a new algorithm based on the Monte-Carlo stochastic simulation, where the random numbers are generated for the state variables  $\alpha_t$  and  $\alpha_{t-1}$  and the error terms  $\epsilon_t$  and  $\eta_t$  to evaluate the expectations (9) – (13) more correctly. Since we have assumed  $\epsilon_t$  and  $\eta_t$  to be normal and approximated  $\tilde{u}_t$  and  $\tilde{v}_t$  to be normal, it might be possible to generate the random numbers of  $\alpha_t$  given  $I_t$  and  $I_{t-1}$ ,  $\epsilon_t$  and  $\eta_t$ .

First, to obtain the algorithm of the nonlinear filter based on equations (9) – (16), consider approximating the nonlinear functions  $h_t(\alpha_t, \epsilon_t)$  and  $g_t(\alpha_t, \eta_t)$  by random numbers. Equations (9) – (13) are approximated by:

$$a_{t|t-1} = \frac{1}{n} \sum_{i=1}^n \alpha_{i,t|t-1}, \quad (22)$$

$$\begin{aligned} \Sigma_{t|t-1} &= \frac{1}{n} \sum_{i=1}^n (\alpha_{i,t|t-1} - a_{t|t-1})(\alpha_{i,t|t-1} - a_{t|t-1})' \\ &= \frac{1}{n} \sum_{i=1}^n \alpha_{i,t|t-1} \alpha_{i,t|t-1}' - a_{t|t-1} a_{t|t-1}', \end{aligned} \quad (23)$$

$$y_{t|t-1} = \frac{1}{n} \sum_{i=1}^n y_{i,t|t-1}, \quad (24)$$

$$\begin{aligned} F_{t|t-1} &= \frac{1}{n} \sum_{i=1}^n (y_{i,t|t-1} - y_{t|t-1})(y_{i,t|t-1} - y_{t|t-1})' \\ &= \frac{1}{n} \sum_{i=1}^n y_{i,t|t-1} y_{i,t|t-1}' - y_{t|t-1} y_{t|t-1}', \end{aligned} \quad (25)$$

$$\begin{aligned} M_{t|t-1} &= \frac{1}{n} \sum_{i=1}^n (y_{i,t|t-1} - y_{t|t-1})(\alpha_{i,t|t-1} - a_{t|t-1})' \\ &= \frac{1}{n} \sum_{i=1}^n y_{i,t|t-1} \alpha_{i,t|t-1}' - y_{t|t-1} a_{t|t-1}', \end{aligned} \quad (26)$$

where

$$\begin{aligned} \alpha_{i,t|t-1} &= g_t(\alpha_{i,t-1|t-1}, \eta_{i,t}), \\ y_{i,t|t-1} &= h_t(\alpha_{i,t|t-1}, \epsilon_{i,t}). \end{aligned}$$

The random numbers  $\alpha_{i,t-1|t-1}$ ,  $\eta_{i,t}$ ,  $\alpha_{i,t|t-1}$  and  $\epsilon_{i,t}$  for  $i = 1, \dots, n$  are approximately generated from the following normal distributions:

$$\begin{aligned} \begin{pmatrix} \alpha_{i,t-1|t-1} \\ \eta_{i,t} \end{pmatrix} &\sim N \left( \begin{pmatrix} a_{t-1|t-1} \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{t-1|t-1} & 0 \\ 0 & Q_t \end{pmatrix} \right), \\ \begin{pmatrix} \alpha_{i,t|t-1} \\ \epsilon_{i,t} \end{pmatrix} &\sim N \left( \begin{pmatrix} a_{t|t-1} \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{t|t-1} & 0 \\ 0 & H_t \end{pmatrix} \right). \end{aligned}$$

Approximating  $\tilde{u}_t$  and  $\tilde{v}_t$  to be normal is equivalent to approximating  $\alpha_t$  and  $\alpha_{t-1}$  given  $I_{t-1}$  to be normal. Therefore, in the algorithm above, we generate the normal random numbers for  $\alpha_t$  and  $\alpha_{t-1}$  in order to evaluate the expectations.

According to this estimator, asymptotically, the problems (18) and (19) on the errors  $\tilde{u}_t$  and  $\tilde{v}_t$  are improved. However, we still have the problems (20) and (21).

The approximations for all of the algorithms introduced in this section are that we have assumed the distributions of  $\tilde{u}_t$  and  $\tilde{v}_t$  in equations (7) and (8) to be approximately normally distributed, and that we have ignored the correlation between  $\tilde{u}_t$  and  $\tilde{v}_t$ . In order to apply the model (7) and (8) directly to the Kalman filter algorithm, the residuals  $\tilde{u}_t$  and  $\tilde{v}_t$  must be conditionally normally distributed given  $I_{t-1}$ . They are not normally distributed, however, because of the nonlinearity of the functions  $h_t(\cdot, \cdot)$  and  $g_t(\cdot, \cdot)$ . Moreover,  $\tilde{u}_t$  is correlated with  $\tilde{v}_t$ . Approximation of the expectations in equations (9) – (13) might be more appropriate for the Monte-Carlo simulation filter than the extended Kalman filter and the second-order nonlinear filter.

In the following section, we will examine the condition under which we can relax one of the approximations for the algorithm (9) – (16). Recall that the problems in the algorithm (9) – (16) are that  $\tilde{u}_t$  is correlated with  $\tilde{v}_t$  and that  $\tilde{u}_t$  and  $\tilde{v}_t$  are nonnormal. Under a certain functional form of either  $g_t(\cdot, \cdot)$  or  $h_t(\cdot, \cdot)$ , it will be shown that  $\tilde{u}_t$  is uncorrelated with  $\tilde{v}_t$ .

### 3 Theorems

Under a certain functional form of either  $h_t(\cdot, \cdot)$  or  $g_t(\cdot, \cdot)$ , the correlation between the residuals  $\tilde{u}_t$  and  $\tilde{v}_t$  in equations (7) and (8) disappears. This implies that we do not need to take into account one of the approximations made clear in the previous section in the case where we choose the functional form. In this section, we propose two theorems. These are useful for the Monte-Carlo simulation filter proposed in this paper, because the Monte-Carlo simulation filter has the fewer problems, i.e., zero-mean errors uncorrelated with the state vectors.

To show no correlation between  $\tilde{u}_t$  and  $\tilde{v}_t$  under a certain functional form, first, consider the following measurement and transition equations:

$$\text{(Measurement equation)} \quad y_t = h_{1t}(\alpha_t)h_{2t}(\epsilon_t), \quad (27)$$

$$\text{(Transition equation)} \quad \alpha_t = g_{1t}(\alpha_{t-1})g_{2t}(\eta_t). \quad (28)$$

Transforming the error terms in the same way,  $\tilde{u}_t$  and  $\tilde{v}_t$  are represented by:

$$\begin{aligned} \tilde{u}_t &= h_{1t}(h_{2t} - \bar{h}_{2t}) + \bar{h}_{2t}(h_{1t} - \bar{h}_{1t}) \\ &\quad - \bar{h}_{2t}E((h_{1t} - \bar{h}_{1t})(\alpha_t - a_{t|t-1})'|I_{t-1})\Sigma_{t|t-1}^{-1}(\alpha_t - a_{t|t-1}), \end{aligned} \quad (29)$$

$$\begin{aligned} \tilde{v}_t &= g_{1t}(g_{2t} - \bar{g}_{2t}) + \bar{g}_{2t}(g_{1t} - \bar{g}_{1t}) \\ &\quad - \bar{g}_{2t}E((g_{1t} - \bar{g}_{1t})(\alpha_{t-1} - a_{t-1|t-1})'|I_{t-1})\Sigma_{t-1|t-1}^{-1}(\alpha_{t-1} - a_{t-1|t-1}), \end{aligned} \quad (30)$$

where we define as follows:  $h_{1t} \equiv h_{1t}(\alpha_t)$ ,  $h_{2t} \equiv h_{2t}(\epsilon_t)$ ,  $g_{1t} \equiv g_{1t}(\alpha_{t-1})$  and  $g_{2t} \equiv g_{2t}(\eta_t)$ .  $\bar{h}_{2t}$  and  $\bar{g}_{2t}$  denote the expectations with respect to the error terms  $\epsilon_t$  and  $\eta_t$ .  $\bar{h}_{1t}$  and  $\bar{g}_{1t}$  represent the conditional expectations with respect to  $\alpha_t$  and  $\alpha_{t-1}$ , given information up to time  $t - 1$ , i.e.,  $I_{t-1}$ . With this transformation, we have no correlation between  $\tilde{u}_t$  and  $(\alpha_t - a_{t|t-1})$  and no correlation between  $\tilde{v}_t$  and

$(\alpha_{t-1} - a_{t-1|t-1})$ .  $\tilde{u}_t$  is, however, correlated with  $\tilde{v}_t$ , because the third term in equation (30) is correlated with both the second and the third terms in equation (29). From the structure of the residuals  $\tilde{u}_t$  and  $\tilde{v}_t$ , the following theorem can be proposed.

**Theorem 1:** In the case of the state-space model given by equations (27) and (28), if one of the functions  $h_{1t}(\alpha_t)$ ,  $h_{2t}(\epsilon_t)$ ,  $g_{1t}(\alpha_{t-1})$  and  $g_{2t}(\eta_t)$  is linear in its argument, i.e., if  $h_{1t}(\alpha_t) = Z_t\alpha_t$ ,  $h_{2t}(\epsilon_t) = S_t\epsilon_t$ ,  $g_{1t}(\alpha_{t-1}) = T_t\alpha_{t-1}$  and  $g_{2t}(\eta_t) = R_t\eta_t$ , then  $\tilde{u}_t$  is not correlated with  $\tilde{v}_t$ . Note that  $Z_t$ ,  $S_t$ ,  $T_t$  and  $R_t$  may depend on  $I_{t-1}$ .

**Proof:** First, consider the case where either  $h_{1t}(\alpha_t)$  or  $h_{2t}(\epsilon_t)$  is linear. Equation (29) reduces to:

$$\tilde{u}_t = h_{1t}(h_{2t} - \bar{h}_{2t}). \quad (31)$$

Since  $(h_{2t} - \bar{h}_{2t})$  is a function of  $\epsilon_t$ , it is uncorrelated with  $\eta_t$ ,  $\alpha_t$  and  $\alpha_{t-1}$ . Therefore, we have  $E(\tilde{u}_t\tilde{v}_t') = 0$ .

Next, consider the case where either  $g_{1t}(\alpha_{t-1})$  or  $g_{2t}(\eta_t)$  is linear. Similarly, equation (30) reduces to:

$$\tilde{v}_t = g_{1t}(g_{2t} - \bar{g}_{2t}), \quad (32)$$

which is not correlated with  $\epsilon_t$ ,  $\alpha_t$  and  $\alpha_{t-1}$ . In this case, we also have  $E(\tilde{u}_t\tilde{v}_t') = 0$ .

Thus, in the case where one of the functions  $h_{1t}(\alpha_t)$ ,  $h_{2t}(\epsilon_t)$ ,  $g_{1t}(\alpha_{t-1})$  and  $g_{2t}(\eta_t)$  is linear, we do not have to take into account the correlation between the error terms (the residuals). In this case, we might expect a better approximation. This theorem is useful for the Monte-Carlo simulation filter, because the expectations are more correctly evaluated than the other filters.

Next, consider the following nonlinear state-space model:

$$\text{(Measurement equation)} \quad y_t = h_{1t}(\alpha_t) + h_{2t}(\epsilon_t), \quad (33)$$

$$\text{(Transition equation)} \quad \alpha_t = g_{1t}(\alpha_{t-1}) + g_{2t}(\eta_t). \quad (34)$$

In this case,  $M_{t|t-1}$  is computed as:

$$\begin{aligned} M_{t|t-1} &= E((y_t - y_{t|t-1})(\alpha_t - a_{t|t-1})' | I_{t-1}) \\ &= E((h_{1t} - \bar{h}_{1t})(\alpha_t - a_{t|t-1})' | I_{t-1}), \end{aligned}$$

because  $(y_t - y_{t|t-1})$  is represented as:

$$y_t - y_{t|t-1} = (h_{1t} - \bar{h}_{1t}) + (h_{2t} - \bar{h}_{2t}),$$

and  $(h_{2t} - \bar{h}_{2t})$  is not correlated with  $(\alpha_t - a_{t|t-1})$ .

Thus,  $\tilde{u}_t$  is written as follows:

$$\begin{aligned}\tilde{u}_t &= (h_{1t} - \bar{h}_{1t}) + (h_{2t} - \bar{h}_{2t}) \\ &\quad - E((h_{1t} - \bar{h}_{1t})(\alpha_t - a_{t|t-1})' | I_{t-1}) \Sigma_{t|t-1}^{-1} (\alpha_t - a_{t|t-1}).\end{aligned}$$

Similarly,  $\tilde{v}_t$  is given by:

$$\begin{aligned}\tilde{v}_t &= (g_{1t} - \bar{g}_{1t}) + (g_{2t} - \bar{g}_{2t}) \\ &\quad - E((g_{1t} - \bar{g}_{1t})(\alpha_{t-1} - a_{t-1|t-1})' | I_{t-1}) \Sigma_{t-1|t-1}^{-1} (\alpha_{t-1} - a_{t-1|t-1}).\end{aligned}$$

Here, the following theorem can be proposed.

**Theorem 2:** In the case of the state-space model given by equations (33) and (34), if either  $h_{1t}(\alpha_t)$  or  $g_{1t}(\alpha_{t-1})$  is linear in its argument, i.e., if either  $h_{1t}(\alpha_t) = Z_t \alpha_t$  or  $g_{1t} = T_t \alpha_{t-1}$ , then  $\tilde{u}_t$  is not correlated with  $\tilde{v}_t$ . Note that  $Z_t$  and  $T_t$  may depend on  $I_{t-1}$ .

**Proof:** When we take  $h_{1t}(\alpha_t) = Z_t \alpha_t$ ,  $\tilde{u}_t$  is rewritten as follows:

$$\tilde{u}_t = h_{2t} - \bar{h}_{2t},$$

because

$$\begin{aligned}h_{1t} - \bar{h}_{1t} &= Z_t(\alpha_t - a_{t|t-1}), \\ E((h_{1t} - \bar{h}_{1t})(\alpha_t - a_{t|t-1})' | I_{t-1}) \Sigma_{t|t-1}^{-1} (\alpha_t - a_{t|t-1}) &= Z_t(\alpha_t - a_{t|t-1}).\end{aligned}$$

Clearly,  $\tilde{u}_t = h_{2t} - \bar{h}_{2t}$  is independent of  $\tilde{v}_t$ , because  $\tilde{u}_t$  depends on the error term  $\epsilon_t$  only.

Similarly, if  $g_{1t}(\alpha_{t-1}) = T_t \alpha_{t-1}$ ,  $\tilde{v}_t$  is given by:

$$\tilde{v}_t = g_{2t} - \bar{g}_{2t},$$

which is independent of  $\tilde{u}_t$ , because  $\tilde{v}_t$  is a function of the error term  $\eta_t$  only.

Thus, as shown in this section, under a certain functional form of either  $h_t(\alpha_t, \epsilon_t)$  or  $g_t(\alpha_{t-1}, \eta_t)$ , the correlation between the residuals  $\tilde{u}_t$  and  $\tilde{v}_t$  in equations (7) and (8) disappears.

## 4 Monte-Carlo Experiments

In this section, based on the criteria of BIAS (bias) and RMSE (root mean squared error), we examine the extended Kalman filter, the second-order nonlinear filter and the Monte-Carlo simulation filter. Three experiments are performed in this section.

For one experiment, the filtering estimates are compared using a logistic type of nonlinear functions for the three nonlinear filters. 1000 simulations are performed, and BIAS and RMSE between the estimated state variables and the simulated ones are computed for each time  $t$ .

For another experiment, we take an ARCH(1) model as an example. An attempt is made to obtain the filtering estimates in the case where an unknown parameter is included in the system, and the parameter estimates are compared for the three filters.

For the third experiment, the functional forms investigated in Kitagawa (1987) and Carlin, Polson and Stoffer (1992) are taken.

**Experiment I:** Consider the following logistic measurement and transition equations:

$$\text{(Measurement equation)} \quad y_t = \frac{\exp(\alpha_t)}{\exp(\alpha_t) + \exp(\epsilon_t)}, \quad (35)$$

$$\text{(Transition equation)} \quad \alpha_t = \frac{\exp(\alpha_{t-1})}{\exp(\alpha_{t-1}) + \exp(\eta_t)}, \quad (36)$$

where  $\epsilon_t$  and  $\eta_t$  are assumed to be normally distributed with:

$$E \begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and

$$\text{Cov} \begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Note that both  $y_t$  and  $\alpha_t$  lie on the interval between zero and one.

The experiment procedure is as follows:

- (i) Generate normal random numbers for  $\epsilon_t$  and  $\eta_t$  with mean  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  and variance  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  given  $\alpha_0 = 0.5$ , and we obtain the artificial data for  $y_t$  and  $\alpha_t$ ,  $t = 1, \dots, T$ , from equations (35) and (36), where  $T = 100$  is taken.
- (ii) Given the variances of  $\epsilon_t$  and  $\eta_t$ , and the data  $y_t$  obtained in procedure (i), we can compute the filtering estimates (i.e.,  $a_{t|t}$ ) using the algorithms introduced in Section 3. Here, the artificially generated state variable at time  $t = 1$ ,  $\alpha_1$ , is taken as the initial value. The initial variance is given by  $\Sigma_{1|1} = 0$ .
- (iii) Repeat the procedures (i) – (ii)  $m$  times (i.e., we perform  $m$  simulations). Let  $\alpha_t^{(i)}$ ,  $t = 1, \dots, T$  and  $i = 1, \dots, n$ , be the  $i$ -th set of artificial data (the  $i$ -th true state variable), and  $a_{t|t}^{(i)}$  be the  $i$ -th simulation run (the  $i$ -th estimated state variable) in a series of  $m$  simulation runs, where  $m = 1000$  is set.
- (iv) For each nonlinear filter, compute the bias (BIAS) as:

$$\text{BIAS} = \frac{1}{T-1} \sum_{t=2}^T \text{BIAS}_t,$$

Table 1: A Comparison of the Three Filters (Experiment I)

	$n$	BIAS	RMSE
E		-.0228	.1971
S		.0100	.1960
M	5	.0043	.2731
M	20	.0005	.2146
M	50	.0003	.2058
M	100	.0000	.2026
M	500	.0001	.2019

where  $\text{BIAS}_t$  is defined as:

$$\text{BIAS}_t = \frac{1}{m} \sum_{i=1}^m (\alpha_t^{(i)} - a_{t|t}^{(i)}),$$

and the root mean squared error (RMSE) as,

$$\text{RMSE} = \frac{1}{T-1} \sum_{t=2}^T \text{RMSE}_t,$$

where  $\text{RMSE}_t$  is defined as:

$$\text{RMSE}_t = \left( \frac{1}{m} \sum_{i=1}^m (\alpha_t^{(i)} - a_{t|t}^{(i)})^2 \right)^{1/2}.$$

All of the values in Table 1 are represented by average of 1000 simulations. E, S and M denotes the extended Kalman filter, the second-order nonlinear filters and the Monte-Carlo simulation filter. Recall that  $n$  denotes the number of the random draws used in the Monte-Carlo simulation filter. In Figures 1 and 2, the case  $n = 500$  is displayed for the Monte-Carlo simulation filter.

Table 1 and Figures 1 and 2 show the results of BIAS and RMSE, defined above. The filtering estimates are overestimated for the extended Kalman filter and underestimated for the second-order nonlinear filter. Both estimates are clearly biased (see Figure 1). The Monte-Carlo simulation filter gives us the most unbiased estimates even if  $n$  (the number of random draws) is small (see Table 1 and Figure 1). Under the RMSE criterion, the second-order nonlinear filter show the smallest values, i.e., 0.1960 (see Table 1 and Figure 2). The Monte-Carlo simulation filter takes the largest RMSE, although the RMSE decreases as  $n$  increases (see Table 1).

Summarizing the results, we have the following: (i) the filtering estimates obtained from the extended Kalman filter and the second-order nonlinear filter are biased but small RMSE, (ii) the Monte-Carlo simulation filter gives asymptotically unbiased filtering estimates but large RMSE, and (iii) for the Monte-Carlo simulation filter, precision of the estimates is improved in the sense of RMSE as  $n$  increases.

Figure 1: A Comparison of the Three Filters (Experiment I)  
(BIAS<sub>t</sub>)

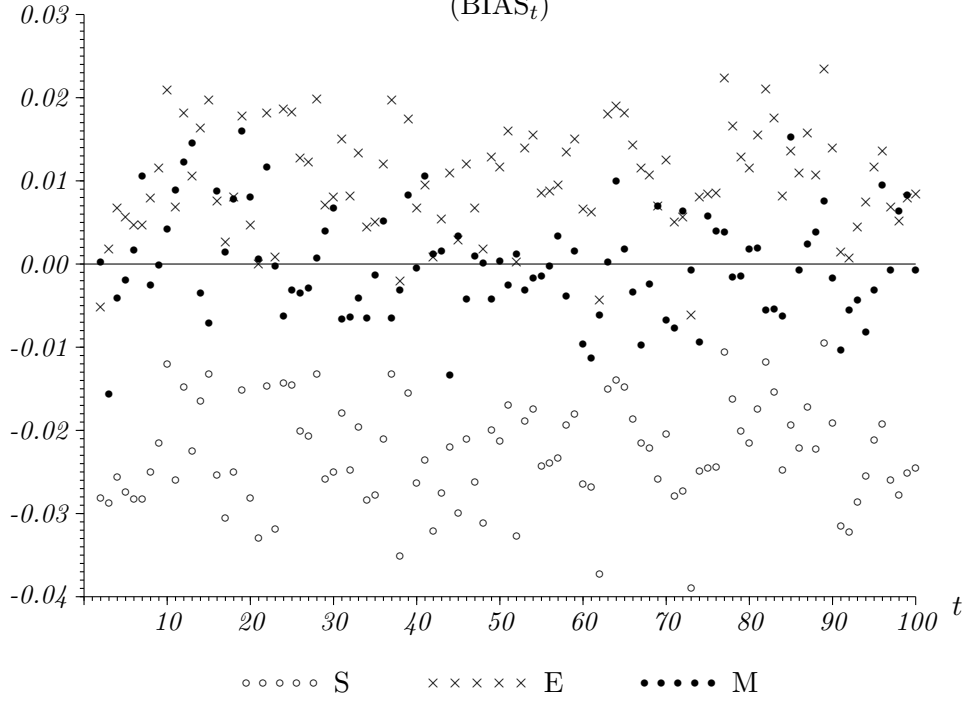


Figure 2: A Comparison of the Three Filters (Experiment I)  
(RMSE<sub>t</sub>)

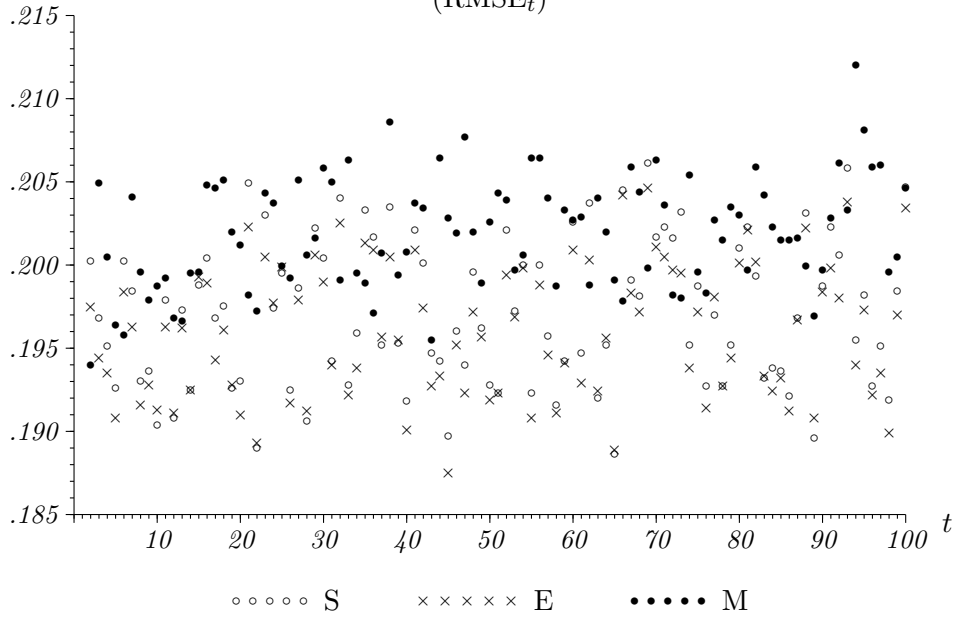


Table 2: A Comparison of the Three Filters (Experiment II)

	$n$	$b$				
		0.0	0.3	0.6	0.9	0.95
BIAS	E	-.0024	-.0023	-.0023	-.0012	-.0008
	S	-.0024	-.0023	-.0023	-.0012	-.0006
	M 50	.0021	.0024	.0024	.0019	.0013
	M 100	-.0025	-.0023	-.0020	-.0010	-.0009
	M 500	.0020	.0023	.0022	.0019	.0013
RMSE	E	.7091	.7051	.6810	.5404	.4525
	S	.7091	.7050	.6797	.5381	.4503
	M 50	.7159	.7126	.6897	.5638	.4916
	M 100	.7096	.7051	.6803	.5501	.4765
	M 500	.7136	.7093	.6844	.5487	.4694

**Experiment II:** In the second example, estimation of an unknown parameter is considered using an ARCH(1) model, which is written as  $\alpha_t = \sigma_t \eta_t$ , where  $\sigma_t = (a + b\sigma_{t-1}^2 \eta_{t-1}^2)^{1/2}$ ,  $a > 0$  and  $0 \leq b < 1$ .  $\eta_t$  is mutually independently, identically and normally distributed and normalized to the normal random variable with mean zero and variance one. Eliminating  $\sigma_t$  from the above two equations, it is rewritten as  $\alpha_t = (a + b\alpha_{t-1}^2)^{1/2} \eta_t$ , which corresponds to the transition equation in Experiment II. Also, normalizing the unconditional variance of  $\alpha_t$  to one,  $a = 1 - b$  is taken. For the measurement equation, we assume that the observed variable  $y_t$  consists of the ARCH(1) process  $\alpha_t$  and a random shock  $\epsilon_t$ . Therefore, we have the following state-space form:

$$\text{(Measurement equation)} \quad y_t = \alpha_t + \epsilon_t, \quad (37)$$

$$\text{(Transition equation)} \quad \alpha_t = (1 - b + b\alpha_{t-1}^2)^{1/2} \eta_t, \quad (38)$$

where  $\alpha_t$  is the state variable and we assume  $\epsilon_t$  and  $\eta_t$  as mutually independent standard normal random variables. In this model, the expectations given by equations (9) and (11) – (13) are explicitly computed without any approximation. Approximation of equation (10) is performed by the first- and the second-order Taylor series expansions or the Monte-Carlo stochastic simulations. In this experiment, we consider estimating  $\alpha_t$  and  $b$  simultaneously.

The experiment procedure is as follows. Given the parameter  $b$ , generate normal random draws of  $\epsilon_t$  and  $\eta_t$ , and compute  $y_t$  and  $\alpha_t$  for  $t = 1, \dots, T$  ( $T = 100$ ). For the initial value  $\alpha_0$ , we take the normal distribution with mean zero and variance one. Then, based on  $y_t$ ,  $t = 1, \dots, T$ , we estimate  $\alpha_t$  given  $I_t$ , and the unknown parameter  $b$  using the likelihood function (8). Here,  $m$  sets of data  $\{y\}_{t=1}^T$  are generated and  $m$  estimates of  $b$  are obtained where  $m = 1000$ . We choose the true parameter as  $b = 0.0, 0.3, 0.6, 0.9, 0.95$ . For the Monte-Carlo Simulation filter,  $n = 50, 100, 500$  is taken.

The maximization of the likelihood function is performed by a simple grid search,



Table 3: True Parameters and Estimates (Experiment II)

$b$		$n$	AVE	VAR	MSE	10%	25%	50%	75%	90%
.0	E		.14697	.03622	.05753	.0	.0	.04	.26	.45
	S		.14719	.03797	.05963	.0	.0	.04	.25	.46
	M	50	.13113	.05202	.06921	.0	.0	.0	.19	.45
	M	100	.14712	.06014	.08179	.0	.0	.0	.21	.53
	M	500	.16628	.06957	.09722	.0	.0	.0	.24	.62
.3	E		.22383	.04161	.04741	.0	.0	.19	.38	.52
	S		.23528	.04585	.05004	.0	.0	.20	.40	.54
	M	50	.24262	.07609	.07938	.0	.0	.16	.40	.67
	M	100	.27581	.09004	.09062	.0	.0	.18	.47	.78
	M	500	.31049	.09541	.09552	.0	.0	.25	.52	.81
.6	E		.36752	.04769	.10174	.0	.22	.39	.53	.64
	S		.39117	.05340	.09701	.0	.22	.41	.57	.67
	M	50	.41584	.10628	.14019	.0	.08	.42	.68	.89
	M	100	.45506	.12150	.14251	.0	.09	.46	.78	.95
	M	500	.52120	.11188	.11809	.0	.24	.57	.81	.95
.9	E		.67703	.03761	.08733	.43	.56	.69	.83	.92
	S		.70462	.03742	.07560	.46	.60	.73	.85	.93
	M	50	.77452	.08422	.09997	.29	.68	.90	.99	.99
	M	100	.83002	.06936	.07425	.44	.80	.95	.99	.99
	M	500	.88362	.03401	.03428	.71	.86	.96	.99	.99
.95	E		.78423	.02905	.05653	.55	.68	.81	.92	.99
	S		.80724	.02643	.04681	.60	.73	.84	.93	.99
	M	50	.84681	.06929	.07994	.53	.85	.97	.99	.99
	M	100	.89626	.04477	.04766	.74	.91	.99	.99	.99
	M	500	.93710	.01631	.01648	.84	.94	.99	.99	.99

Table 4: A Comparison of the Three Filters (Experiment III)

	$n$	BIAS	RMSE
E		.6416	23.7023
S		.1124	82.1384
M	5	-.0073	16.0883
M	20	-.1381	10.2972
M	50	-.0867	9.5633
M	100	-.0920	9.3411
M	500	-.0885	9.1979

in which the function is maximized by changing the parameter value of  $b$  by 0.01 in the interval between 0 and 0.99.

Table 2 represents BIAS and RMSE between the artificially simulated state variable (i.e.,  $\alpha_t^{(i)}$ ) and the estimated state variable (i.e.,  $a_{t|t}^{(i)}$ ). Table 3 shows true parameter values ( $b = 0.0, 0.3, 0.6, 0.9, 0.95$ ) and the corresponding estimates for each estimator (i.e., E, S and M). E, S and M denote the extended Kalman filter, the second-order nonlinear filter and the Monte-Carlo simulation filter. AVE, VAR and MSE represent the average, variance and mean squared error of the estimates obtained in 1000 simulations. Also, 10%, 25%, 50%, 75% and 90% indicate each percent value from the 1000 estimates.

As precision of approximation of the expectations is improved, the average of the estimates is closer to the true parameter value (see AVE in Table 3). In this sense, the Monte-Carlo simulation filter is the best estimator but gives the largest variance and relatively large MSE.

**Experiment III:** The following functions are taken in Kitagawa (1987) and Carlin, Polson and Stoffer (1992).

$$\text{(Measurement equation)} \quad y_t = \frac{\alpha_t^2}{20} + \epsilon_t, \quad (39)$$

$$\begin{aligned} \text{(Transition equation)} \quad \alpha_t = & \frac{1}{2}\alpha_{t-1} + \frac{25\alpha_{t-1}}{1 + \alpha_{t-1}^2} \\ & + 8 \cos(1.2(t-1)) + \eta_t, \end{aligned} \quad (40)$$

The error terms are mutually independently distributed, i.e.,

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 10 \end{pmatrix} \right).$$

The distribution of the initial value is assumed to be normal, i.e.,

$$\alpha_0 \sim N(0, 10).$$

The results are in Table 4. In Experiment I, differences among E, S and M are small. However, in this experiment, M performs much better than E and S from

Table 4. The Monte-Carlo simulation filter is extremely better than the extended Kalman filter and the second-order nonlinear filter in the sense of both BIAS and RMSE criteria.

## 5 Summary

In this paper, the three nonlinear filters are derived from the same theoretical framework (i.e., the extended Kalman filter, the second-order nonlinear filter and the Monte-Carlo simulation filter). There, we have discussed the following: first, we must impose some approximations on the approximated error terms (or the residuals)  $\tilde{u}_t$  and  $\tilde{v}_t$  and also linearize the nonlinear functions in order to evaluate the expectations in equations (9) – (13). Some of these approximations imply that the approximated error terms  $\tilde{u}_t$  and  $\tilde{v}_t$  are mutually, independently and normally distributed with zero means. Summarizing the problems of the nonlinear filters such as the extended Kalman filter and the second-order nonlinear filter, we have the problems (18) – (21) on the approximated error terms, which are not well-behaved. In order to use the Taylor series expansion approach, we have to approximate the error terms  $\tilde{u}_t$  and  $\tilde{v}_t$  to be well-behaved.

Second, the nonlinear filter was proposed, where the expectations are approximated by generating the normal random numbers for the state variable  $\alpha_t$  given  $I_{t-1}$  and the error terms  $\epsilon_t$  and  $\eta_t$ . The approximation based on the normal random numbers might be appropriate because  $\tilde{u}_t$  and  $\tilde{v}_t$  are approximated to be normal in deriving the algorithm (9) – (16) and accordingly the normality approximation of  $\tilde{u}_t$  and  $\tilde{v}_t$  implies the normality approximation of the state variables  $\alpha_t$  and  $\alpha_{t-1}$  given  $I_{t-1}$ . For the Monte-Carlo simulation filter proposed in this paper, however, we still have the problems (20) and (21) while (18) and (19) are improved.

Third, in the case where one of the functions  $h_{1t}(\alpha_t)$ ,  $h_{2t}(\epsilon_t)$ ,  $g_{1t}(\alpha_{t-1})$  and  $g_{2t}(\eta_t)$  is linear, we do not need to consider the correlation between the error terms (the residuals). We might expect a better approximation if we take such a functional form for the measurement equation or the transition equation. The theorems proposed here are useful in the case where  $\tilde{u}_t$  and  $\tilde{v}_t$  have zero-means, and accordingly useful for the Monte-Carlo simulation filter.

The Monte-Carlo experiments indicated that the Monte-Carlo simulation filter gives asymptotically unbiased filtering estimates and small RMSE, depending on the functional form of the state-space model, and that, in simultaneous estimation of the state variables and the unknown parameter, the average of the estimated parameter is closer to the true parameter value as the precision of approximation of the expectations is improved.

The approximation by a Taylor series expansion is broadly used in the case of nonlinear estimation. When approximating nonlinear functions by the Taylor series expansion, however, we need to recognize that the structure of the original function is completely different from that of the approximated function. Especially the approximated error terms consist of two parts; the stochastic part and the deterministic part. According to the Taylor series expansion approach, we have to assume the approximated error terms to be normal, which implies that we must assume

that both the stochastic part and the deterministic part are normally distributed. Clearly, it is not appropriate to assume the deterministic part to be normal. Therefore, we have to justify the approximation when we use the Taylor series expansion approach. In this paper, this discussion was not given.

Finally, note as follows. We have discussed about the nonlinear filters based on Taylor series expansions. It has been shown in this paper that in the class of the Taylor series expansion approaches the Monte-Carlo simulation filter is optimal at least in the sense of BIAS. Moreover, because of less computational burden, the Monte-Carlo simulation filter might be recommended. Recently, the density-based nonlinear filters have been developed, where the underlying density functions, rather than the nonlinear measurement and transition equations, are approximated. Kitagawa (1987) and Kramer and Sorenson (1988) evaluated the integration included in the density-based filtering algorithm by numerical integration. Carlin, Polson and Stoffer (1992) derived the algorithm by the Gibbs sampler in a Bayesian framework. Tanizaki (1991), Tanizaki and Mariano (1994) and Mariano and Tanizaki (1995) applied Monte-Carlo integration with importance sampling to the density approximation. Moreover, Mariano and Tanizaki (1995) and Tanizaki and Mariano (1995) proposed the nonlinear filtering algorithm without evaluating any density function, where the random draws are generated directly from the filtering density at each time. Thus, recently, numerous papers have been devoted to the density-based nonlinear filters, rather than the Taylor series expansion approaches, because it is well known that the density-based nonlinear filters are better than the Taylor series expansion approaches from both BIAS and RMSE criteria. However, we have a serious problem on the density-based nonlinear filters, which is the computational disadvantage although computer progresses day by day. Therefore, in this paper, we have proposed the Monte-Carlo simulation filter, which has much less computational burden compared with the density-based approaches.

## Appendix

In the case of approximation based on the first-order Taylor series expansion, equations (7) and (8) reduces to:

$$\begin{aligned} & \text{(Measurement Equation)} \\ y_t &= c_{t|t-1} + Z_{t|t-1}(\alpha_t - a_{t|t-1}) + \tilde{u}_t, \end{aligned} \tag{41}$$

$$\begin{aligned} & \text{(Transition Equation)} \\ \alpha_t &= d_{t|t-1} + T_{t|t-1}(\alpha_{t-1} - a_{t-1|t-1}) + \tilde{v}_t, \end{aligned} \tag{42}$$

where  $c_{t|t-1}$  and  $d_{t|t-1}$  are  $h_t(\alpha_t, \epsilon_t)$  and  $g_t(\alpha_{t-1}, \eta_t)$  evaluated at  $(\alpha_t, \epsilon_t) = (a_{t|t-1}, 0)$  and  $(\alpha_{t-1}, \eta_t) = (a_{t-1|t-1}, 0)$ .  $Z_{t|t-1}$  and  $T_{t|t-1}$  are the first- derivatives with respect to  $\alpha_t$  and  $\alpha_{t-1}$  evaluated at  $(\alpha_t, \epsilon_t) = (a_{t|t-1}, 0)$  and  $(\alpha_{t-1}, \eta_t) = (a_{t-1|t-1}, 0)$ .  $\tilde{u}_t$  and  $\tilde{v}_t$  above are clearly nonnormal with nonzero means, correlated with each other, and correlated with the corresponding state variable.

For the higher-order nonlinear filters, the measurement and transition equations are similarly approximated as equations (41) and (42).

## REFERENCES

- Alspach, D.L. and Sorenson, H.W. (1972). "Nonlinear Bayesian Estimation Using Gaussian Sum Approximations," *IEEE Transactions on Automatic Control*, Vol.AC-17, pp.439-448.
- Anderson, B.D.O. and Moore, J.B. (1979). *Optimal Filtering*, Prentice-Hall, New York.
- Brown, B.W. and Mariano, R.S. (1984). "Residual-based Procedures for Prediction and Estimation in a Nonlinear Simultaneous System," *Econometrica*, Vol.52, pp.321-343.
- Brown, B.W. and Mariano, R.S. (1989). "Measures of Deterministic Prediction Bias in Nonlinear Models," *International Economic Review*, Vol.30, pp.667-684.
- Carlin, B.P., Polson, N.G. and Stoffer, D.S. (1992). "A Monte Carlo Approach to Nonnormal and Nonlinear State Space Modeling," *Journal of the American Statistical Association*, Vol.87, pp.493 – 500.
- Chow, G.C. (1983). *Econometrics*, McGraw-Hill Book Company.
- Gelb, A. (1974). *Applied Optimal Estimation*, MIT Press.
- Harvey, A.C. (1981). *Time Series Models*, Philip Allen Publishers Limited, Oxford.
- Harvey, A.C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press.
- Kitagawa, G. (1987). "Non-Gaussian State-Space Modeling of Nonstationary Time Series," *Journal of the American Statistical Association*, Vol.82, pp.1032-1063 (with discussion).
- Kramer, S.C. and Sorenson, H.W. (1988). "Recursive Bayesian Estimation Using Piece-wise Constant Approximations," *Automatica*, Vol.24, pp.789-801.
- Mariano, R.S. and Brown, B.W. (1983). "Asymptotic Behavior of Predictors in a Nonlinear Simultaneous System," *International Economic Review*, Vol.24, pp.523-536.
- Mariano, R.S. and Brown, B.W. (1989). "Stochastic Simulation, Prediction and Validation of Nonlinear Models," in *Economics in Theory and Practice: An Eclectic Approach* (L.R. Klein and J. Marquez, Eds.), Kluwer Academic Publishers.

- Mariano, R.S. and Tanizaki, H. (1995). "Prediction of Final Data with Use of Preliminary and/or Revised Data," *Journal of Forecasting*, Vol.14, pp.351-380.
- Mariano, R.S. and Tanizaki, H. (1996). "Nonlinear and Nonnormal State-Space Modeling with Monte-Carlo Simulations: Theory and Applications," forthcoming in *Simulation-Based Inference in Econometrics: Methods and Applications* (R.S. Mariano, M.Weeks and T. Schuermann, Eds.), Cambridge University Press.
- Sorenson, H.W. and Alspach, D.L. (1971). "Recursive Bayesian Estimation Using Gaussian Sums," *Automatica*, Vol.7, pp.465-479.
- Tanizaki, H. (1991). "Nonlinear Filters: Estimation and Applications," Dissertation, University of Pennsylvania.
- Tanizaki, H. and Mariano, R.S. (1994). "Prediction, Filtering and Smoothing in Nonlinear and Nonnormal Cases Using Monte-Carlo Integration," *Journal of Applied Econometrics*, Vol.9, pp.163-179.
- Tanizaki, H. and Mariano, R.S. (1995). "Nonlinear and Nonnormal State-Space Modeling with Monte-Carlo Stochastic Simulations," forthcoming in *Journal of Econometrics*.
- Wishner, R.P., Tabaczynski, J.A. and Athans, M. (1969). "A Comparison of Three Non-Linear Filters," *Automatica*, Vol.5, pp.487-496.