A NOTE ON TRUNCATED SECOND-ORDER NONLINEAR FILTERS

Rolf Henriksen

Econometric Research Program Research Memorandum No. 242

April 1979

Abstract

The truncated second-order nonlinear filter is rederived both for discrete-time and continuous systems. The rederivation reveals that a significant error appears in previous derivations of this filter. What has previously been termed the modified truncated second-order filter will be shown to be, provided a small correction is made in the discrete-time case, the correct form of the truncated second-order filter.

Econometric Research Program
Princeton University
207 Dickinson Hall
Princeton, New Jersey

1. INTRODUCTION

Since the introduction of the Kalman filter in 1960 [1], similar techniques have been developed for nonlinear dynamical systems. Most of these works were carried out and published in the late sixties, and a thorough presentation of the different techniques was made in the book by Jazwinski in 1970 [2].

of all the nonlinear filters which have appeared in the literature, especially the so-called second-order filters, i.e., the truncated second-order filter and the Gaussian second-order filter seem to have attracted much attention. Nonlinearities are in both these filters carried to second order only, but while third and higher order central moments are neglected in the truncated second-order filter, fourth-order moments are taken into account in the Gaussian second-order filter by approximating the conditional pdf. which is involved by a normal or Gaussian pdf. Simulation or experimental results with these filters clearly show that they may improve the estimates compared with the extended Kalman filter, but the improvement may depend on the system nonlinearities and the magnitude of the plant and measurement noises, see Schwartz et al. 131, Jazwinski 141, Carney et al. 151, and Henriksen et al. 161.

Andrade Netto et al. [7] have compared several nonlinear filters for discrete-time systems and concluded that the truncated second-order filter should not be used because of a term in the covariance equation which tends to decrease the covariance matrix. Jazwinski [2] also makes some comments on the apparent discrepancy between the truncated and the Gaussian second-order filters where a term enters the covariance equations of the two filters with opposite signs. Obviously feeling this to be disquieting, Jazwinski suggests a third type of second-order filter, the so-called

modified second-order filter, where the aforementioned term has been removed.

The truncated second-order nonlinear filter was first developed by Zazwinski [8] and independently by Bass et al. [9]. What we are about to show in this paper is that the truncated second-order filter, as it appears in the literature, actually is wrong since the implications of the assumptions on which it is based have not been fully recognized. In fact, it will be shown that the modified second-order filter is the correct form of what has been termed the truncated second-order filter.

The paper is organized as follows. In Section 2 we prove the Proposition which is used to correct the truncated second-order filter, and derive the correct form of this filter for discrete-time systems. The correct form of the truncated second-order filter for continuous-time systems is then derived in Section 3.

2. DISCRETE-TIME TRUNCATED SECOND-ORDER FILTER

Consider a discrete-time nonlinear system described by the equations

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{t}_k) + \mathbf{G}(\mathbf{x}_k, \mathbf{t}_k) \mathbf{v}_k \tag{1}$$

$$y_k = h(x_k, t_k) + w_k$$
 (2)

where \mathbf{x}_k and \mathbf{y}_k are the state and observation vectors, respectively. The noise processes $\{\mathbf{v}_k\}$ and $\{\mathbf{w}_k\}$ are assumed to be mutually independent zero-mean white processes, both assumed to be independent of \mathbf{x}_0 . The covariance matrices of \mathbf{v}_k and \mathbf{w}_k are denoted by, respectively, \mathbf{v}_k and \mathbf{w}_k .

Define \mathbf{Y}_k to be the sequence of observations \mathbf{y}_0 , \mathbf{y}_1 ,..., \mathbf{y}_k up to time k , viz.

$$Y_k = (y_0, y_1, \dots, y_k)$$
,

and define $\hat{x_i}_{|k|}$ to be the expectation of x_i with respect to $\sigma(Y_k)$, the σ -algebra generated by Y_k , viz.

$$\hat{\mathbf{x}}_{\mathbf{i}|\mathbf{k}} = \mathbf{E}(\mathbf{x}_{\mathbf{i}}|\sigma(\mathbf{Y}_{\mathbf{k}})) \stackrel{\triangle}{=} \mathbf{E}(\mathbf{x}_{\mathbf{i}}|\mathbf{Y}_{\mathbf{k}})$$
(3)

Similarly, define

$$\mathbf{X}_{\mathbf{i}|\mathbf{k}} = \mathbf{E}((\mathbf{x}_{\mathbf{i}} - \hat{\mathbf{x}}_{\mathbf{i}|\mathbf{k}}) (\mathbf{x}_{\mathbf{i}} - \hat{\mathbf{x}}_{\mathbf{i}|\mathbf{k}})^{\mathrm{T}} | \sigma(\mathbf{Y}_{\mathbf{k}})) \stackrel{\triangle}{=} \mathbf{E}((\mathbf{x}_{\mathbf{i}} - \hat{\mathbf{x}}_{\mathbf{i}|\mathbf{k}}) (\mathbf{x}_{\mathbf{i}} - \hat{\mathbf{x}}_{\mathbf{i}|\mathbf{k}})^{\mathrm{T}} | \mathbf{Y}_{\mathbf{k}})$$
(4)

We also introduce the notation

$$\mathbf{E}_{\mathbf{k}}(.) \stackrel{\triangle}{=} \mathbf{E}(. \mid \sigma(\mathbf{Y}_{\mathbf{k}})) \tag{5}$$

The exact equations for $\hat{x}_{k+1|k}$ and $x_{k+1|k}$ are from Equations (1) - (2) found to be

$$\hat{x}_{k+1|k} = E_k(f(x_k, t_k))$$
(6)

$$X_{k+1|k} = E_{k}(f(x_{k}, t_{k})f^{T}(x_{k}, t_{k})) + E_{k}(G(x_{k}, t_{k})V_{k}G^{T}(x_{k}, t_{k}))$$

$$- \hat{x}_{k+1|k}\hat{x}_{k+1|k}$$
(7)

Now, carrying the nonlinearities to second order only and neglecting third and higher order central moments, we find

$$\hat{x}_{k+1|k} = f(\hat{x}_{k|k}, t_k) + \frac{1}{2}(x_{k|k}f_{xx}(\hat{x}_{k|k}, t_k))$$
(8)

where the i'th component of the vector $(x_k|_k f_{xx}(\hat{x}_k|_k, t_k))$ is given by

$$(\mathbf{x}_{\mathbf{k}|\mathbf{k}^{\mathbf{f}}\mathbf{x}\mathbf{x}}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}}^{\mathbf{f}},\mathbf{t}_{\mathbf{k}}))_{\mathbf{i}} = \sum_{\mathbf{j},\ell=1}^{n} \mathbf{x}_{\mathbf{k}|\mathbf{k}}^{\mathbf{j}\ell} \frac{\partial^{2} \mathbf{f}_{\mathbf{i}}}{\partial \mathbf{x}_{\mathbf{k}}^{\mathbf{j}} \partial \mathbf{x}_{\mathbf{k}}^{\ell}} (\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}}^{\mathbf{f}},\mathbf{t}_{\mathbf{k}})$$
(9)

where $x_{k\,|\,k}^{j\,\ell}$ is element (j, ℓ) of $x_{k\,|\,k}$, f, is the i'th component of f while x_k^j is the j'th component of x_k .

Before we proceed, let us make a note about the magnitude of fourth-order central moments compared to the square of second-order central moments. If ξ is a normal random variable with variance σ^2 , we know that the fourth-order central moment of ξ is $3\sigma^4$ which certainly is greater than σ^4 . This can easily be shown to hold for any random variable.

Proposition

Let ξ be a random zero-mean vector, and let $F(\xi\xi^T)$ be a vector function of $\xi\xi^T$. Then $^{1)}$

$$E(F(\xi\xi^{T})F^{T}(\xi\xi^{T})) \geq E(F(\xi\xi^{T}))E(F^{T}(\xi\xi^{T}))$$
(10)

Proof. Absolutely trivial. We have

$$\begin{split} \mathbf{E} \left[(\mathbf{F} (\boldsymbol{\xi} \boldsymbol{\xi}^{\mathbf{T}}) - \mathbf{E} (\mathbf{F} (\boldsymbol{\xi} \boldsymbol{\xi}^{\mathbf{T}}))) (\mathbf{F} (\boldsymbol{\xi} \boldsymbol{\xi}^{\mathbf{T}}) - \mathbf{E} (\mathbf{F} (\boldsymbol{\xi} \boldsymbol{\xi}^{\mathbf{T}})))^{\mathbf{T}} \right] \\ &= \mathbf{E} (\mathbf{F} (\boldsymbol{\xi} \boldsymbol{\xi}^{\mathbf{T}}) \mathbf{F}^{\mathbf{T}} (\boldsymbol{\xi} \boldsymbol{\xi}^{\mathbf{T}})) - \mathbf{E} (\mathbf{F} (\boldsymbol{\xi} \boldsymbol{\xi}^{\mathbf{T}})) \mathbf{E} (\mathbf{F}^{\mathbf{T}} (\boldsymbol{\xi} \boldsymbol{\xi}^{\mathbf{T}})) \\ &\geq \mathbf{0} \end{split}$$

Expanding the right-hand terms in Equation (7) in Taylor series about $\hat{x}_{k|k}$, taking expectations with respect to $\sigma(Y_k)$, and neglecting third-order central moments, we find

For two symmetric matrices A and B, $A \ge B$ or $A - B \ge 0$ means that A - B is positive semidefinite.

$$\begin{split} \mathbf{x}_{\mathbf{k}+\mathbf{1}|\mathbf{k}} &= \mathbf{f}_{\mathbf{x}}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}})\mathbf{x}_{\mathbf{k}|\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{x}}^{\dagger}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}}) \\ &+ \frac{1}{4}\mathbf{E}_{\mathbf{k}}^{\dagger}\{[(\mathbf{x}_{\mathbf{k}}^{\dagger}-\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}})(\mathbf{x}_{\mathbf{k}}^{\dagger}-\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}})^{\mathsf{T}}\mathbf{f}_{\mathbf{x}\mathbf{x}}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}})][((\mathbf{x}_{\mathbf{k}}^{\dagger}-\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}})(\mathbf{x}_{\mathbf{k}}^{\dagger}-\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}})^{\mathsf{T}}\mathbf{f}_{\mathbf{x}\mathbf{x}}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}})][((\mathbf{x}_{\mathbf{k}}^{\dagger}-\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}})(\mathbf{x}_{\mathbf{k}}^{\dagger}-\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}})^{\mathsf{T}}\mathbf{f}_{\mathbf{x}\mathbf{x}}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}})]^{\mathsf{T}}] \\ &+ \mathbf{G}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}})\mathbf{V}_{\mathbf{k}}^{\mathsf{G}}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}})^{\mathsf{T}} + ((\mathbf{x}_{\mathbf{k}|\mathbf{k}}^{\dagger}\mathbf{V}_{\mathbf{k}}^{\mathsf{G}}\mathbf{x}^{2}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}}))) \\ &+ \frac{1}{2}((\mathbf{x}_{\mathbf{k}|\mathbf{k}}^{\dagger}\mathbf{V}_{\mathbf{k}}^{\mathsf{G}}\mathbf{x}\mathbf{x}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}}))\mathbf{G}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}})) \\ &+ \frac{1}{2}((\mathbf{x}_{\mathbf{k}|\mathbf{k}}^{\dagger}\mathbf{V}_{\mathbf{k}}^{\mathsf{G}}\mathbf{x}\mathbf{x}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}}))\mathbf{G}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}}))^{\mathsf{T}} \\ &- \frac{1}{4}((\mathbf{x}_{\mathbf{k}|\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{x}\mathbf{x}}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}}))(((\mathbf{x}_{\mathbf{k}|\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{x}\mathbf{x}}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}}))^{\mathsf{T}} \\ &- \frac{1}{4}((\mathbf{x}_{\mathbf{k}|\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{x}\mathbf{x}}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}}))(((\mathbf{x}_{\mathbf{k}|\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{x}\mathbf{x}}(\hat{\mathbf{x}}_{\mathbf{k}|\mathbf{k}'}^{\dagger}\mathbf{t}_{\mathbf{k}})))^{\mathsf{T}} \\ &- \frac{1}{4}((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}}\mathbf{x})((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{x}})((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}}\mathbf{x}))((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}}\mathbf{f}_{\mathbf{k}}))^{\mathsf{T}} \\ &+ (\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}})((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}})((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}}))((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}}))^{\mathsf{T}} \\ &+ ((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}})((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}})((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}}))((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}})((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}}))((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}})((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}}))((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}})((\mathbf{x}_{\mathbf{k}}^{\dagger}\mathbf{f}_{\mathbf{k}})))((\mathbf$$

where the components of the vector $[(\mathbf{x}_k - \hat{\mathbf{x}}_k|_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k|_k)\mathbf{f}_{\mathbf{x}\mathbf{x}}(\hat{\mathbf{x}}_k|_k,\mathbf{t}_k)]$, and the elements of the matrices $(\mathbf{x}_k|_k\mathbf{v}_k\mathbf{G}_{\mathbf{x}}^2(\hat{\mathbf{x}}_k|_k,\mathbf{t}_k))$ and $(\mathbf{x}_k|_k\mathbf{v}_k\mathbf{G}_{\mathbf{x}\mathbf{x}}^2(\hat{\mathbf{x}}_k|_k,\mathbf{t}_k))$ are given by, respectively

$$\{(\mathbf{x}_{k}^{-\hat{\mathbf{x}}_{k}}|\mathbf{k})(\mathbf{x}_{k}^{-\hat{\mathbf{x}}_{k}}|\mathbf{k})^{\mathrm{T}}\mathbf{f}_{\mathbf{x}\mathbf{x}}(\hat{\mathbf{x}}_{k}|\mathbf{k},\mathbf{t}_{k})\}_{\mathbf{i}} = \sum_{j,\ell=1}^{n} (\mathbf{x}_{k}^{j-\hat{\mathbf{x}}_{k}^{j}}|\mathbf{k})(\mathbf{x}_{k}^{\ell}-\hat{\mathbf{x}}_{k}^{\ell}|\mathbf{k}) \frac{\partial^{2}\mathbf{f}_{\mathbf{i}}}{\partial \mathbf{x}_{k}^{j}\partial \mathbf{x}_{k}^{\ell}}(\hat{\mathbf{x}}_{k}|\mathbf{k},\mathbf{t}_{k})$$
(12)

$$(x_{k|k}v_{k}G_{x}^{2}(\hat{x}_{k|k},t_{k}))_{iq} = \sum_{j,\ell=1}^{s} \sum_{p,t=1}^{n} x_{k|k}^{pt}v_{k}^{j\ell} \frac{\partial G_{ij}}{\partial x_{k}^{p}}(\hat{x}_{k|k},t_{k}) \frac{\partial G_{q\ell}}{\partial x_{k}^{t}}(\hat{x}_{k|k},t_{k})$$
(13)

$$(\mathbf{x}_{k}|\mathbf{k}^{\mathbf{V}_{k}}\mathbf{G}_{\mathbf{x}\mathbf{x}}(\hat{\mathbf{x}}_{k}|\mathbf{k}^{\mathsf{T}_{k}})\mathbf{G}(\hat{\mathbf{x}}_{k}|\mathbf{k}^{\mathsf{T}_{k}}))_{\mathbf{i}\mathbf{q}} = \sum_{\mathbf{j}, \ell=1}^{\mathbf{S}} \sum_{\mathbf{p}, t=1}^{\mathbf{n}} \mathbf{x}_{k}^{\mathbf{p}t}\mathbf{V}_{k}^{\mathbf{j}\ell} \frac{\partial^{2}\mathbf{G}_{\mathbf{i}\mathbf{j}}}{\partial \mathbf{x}_{k}^{\mathbf{p}}\partial \mathbf{x}_{k}^{\mathbf{t}}}(\hat{\mathbf{x}}_{k}|\mathbf{k}^{\mathsf{T}_{k}})\mathbf{G}_{\mathbf{q}\ell}(\hat{\mathbf{x}}_{k}|\mathbf{k}^{\mathsf{T}_{k}})$$

$$(14)$$

We have here assumed x_k and v_k to be n and s vectors, respectively. Elements of the matrices $x_{k|k}$, v_k , and G are denoted by, respectively, $x_{k|k}^{pt}$, $v_k^{j\ell}$, and G_{ij} .

Since we have assumed fourth-order central moments to be negligible, the second right-hand term in Equation (11) can be dropped. The remainder of that equation then constitutes the covariance equation for the one-stage

predicted estimate as it appears in, say, Jazwinski [1]. However, by the previous proposition, the order of magnitude of the second right-hand term in Equation (11) is at least as great as the last term. So if we assume that the second right-hand term is negligible, it would be not only illogical but even wrong to retain the last term since the sum of these two terms is nonnegative definite. Neglecting these two terms in Equation (11) yields

$$X_{k+1|k} = f_{x}(\hat{x}_{k|k}, t_{k}) X_{k|k} f_{x}^{T}(\hat{x}_{k|k}, t_{k}) + G(\hat{x}_{k|k}, t_{k}) V_{k} G^{T}(\hat{x}_{k|k}, t_{k})$$

$$+ (X_{k|k} V_{k} G_{x}^{2}(\hat{x}_{k|k}, t_{k})) + \frac{1}{2} (X_{k|k} V_{k} G_{xx}(\hat{x}_{k|k}, t_{k})) G(\hat{x}_{k|k}, t_{k})$$

$$+ \frac{1}{2} (X_{k|k} V_{k} G_{xx}(\hat{x}_{k|k}, t_{k})) G(\hat{x}_{k|k}, t_{k})$$

$$+ \frac{1}{2} (X_{k|k} V_{k} G_{xx}(\hat{x}_{k|k}, t_{k})) G(\hat{x}_{k|k}, t_{k})$$
(15)

Andrade Netto et al. [7] conclude that the truncated filter should not be used because of the last term in Equation (11) which has a tendency to decrease the covariance matrix and eventually may force some of the diagonal terms to be negative. Since we now know that this term is not properly present and that Equation (15) is the correct covariance equation, their conclusion can therefore be discarded.

In order to derive the equations for the filtered estimates $\hat{x}_{k+1|k+1}$ and $x_{k+1|k+1}$, the following form usually appears in the literature

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + A(y_{k+1} - \hat{y}_{k+1|k})$$
(16)

$$X_{k+1|k+1} = B + C(y_{k+1} - \hat{y}_{k+1|k})$$
 (17)

Furthermore, if C=0 is assumed, the form is usually termed the modified truncated second-order filter.

Filtering equations of the form given by Equations (16) - (17) have been derived both for the truncated and Gaussian second-order filters, ending up with similar expressions for the tensor C in the two cases

but with opposite signs. Obviously disquieted by this, Jazwinski [1] also suggests forms with C=0 , the so-called modified second-order filters.

A careful computation, applying the previous proposition, will actually reveal that C=0 1) in the truncated filter when the structure given by Equations (16) - (17) is assumed. Therefore, the term modified is quite superfluous and should not be used in connection with the truncated second-order filter. Assuming that third and fourth-order central moments are negligible will in fact imply that the truncated second-order filter becomes "modified" if we still may use that term. The structure of the filtering equations is given in Equations (20) - (22).

Summing up, the true and only truncated second-order filter is of the following form:

Prediction

$$\hat{x}_{k+1|k+1} = f(\hat{x}_{k|k}, t_{k}) + \frac{1}{2}(x_{k|k}f_{xx}(\hat{x}_{k|k}, t_{k}))$$

$$x_{k+1|k} = f_{x}(\hat{x}_{k|k}, t_{k})x_{k|k}f_{x}^{T}(\hat{x}_{k|k}, t_{k}) + G(\hat{x}_{k|k}, t_{k})v_{k}G_{x}^{T}(\hat{x}_{k|k}, t_{k})$$

$$+ (x_{k|k}v_{k}G_{x}^{2}(\hat{x}_{k|k}, t_{k})) + \frac{1}{2}(x_{k|k}v_{k}G_{xx}(\hat{x}_{k|k}, t_{k}))G(\hat{x}_{k|k}, t_{k}))$$

$$+ \frac{1}{2}(x_{k|k}v_{k}G_{xx}(\hat{x}_{k|k}, t_{k}))G(\hat{x}_{k|k}, t_{k}))^{T}$$

$$(18)$$

$$+ \frac{1}{2}(x_{k|k}v_{k}G_{xx}(\hat{x}_{k|k}, t_{k})) + \frac{1}{2}(x_{k|k}v_{k}G_{xx}(\hat{x}_{k|k}, t_{k}))^{T}$$

$$(19)$$

Filtering

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}[y_{k+1} - h(\hat{x}_{k+1|k} + k_{k+1}) - \frac{1}{2}(X_{k+1|k} + x_{k+1|k} + k_{k+1})]$$
(20)

The elaborate computation of the matrices A and B, and the tensor

C is dropped since the filtering equations, once C=0 has been verified,

are quite easily obtained, see the end of this section. Also see the next

section where the result C=0 is shown to hold in the continuous-time case.

$$X_{k+1|k+1} = [I-K_{k+1}^{h} x(\hat{x}_{k+1|k}, t_{k+1})]X_{k+1|k}$$
 (21)

where

$$K_{k+1} = X_{k+1} |_{k}^{T} (\hat{x}_{k+1}|_{k}, t_{k+1}) [h_{x}(\hat{x}_{k+1}|_{k}, t_{k+1}) X_{k+1} |_{k}^{T} (\hat{x}_{k+1}|_{k}, t_{k+1}) + W_{k+1}]^{-1}$$
(22)

The i'th component of the vector $(x_{k+1|k}^h x_x (x_{k+1|k}^t t_{k+1}))$ in Equation (20) is of the form

$$(x_{k+1}|_{k}^{h}_{xx}(\hat{x}_{k+1}|_{k},t_{k+1}))_{i} = \sum_{j,\ell=1}^{m} x_{k+1}^{j\ell}|_{k} \frac{\partial^{2}h_{i}}{\partial x_{k+1}^{j}\partial x_{k+1}^{\ell}}(\hat{x}_{k+1}|_{k},t_{k+1})$$
(23)

The terms which erroneously have been retained in previous papers on truncated second-order filters will appear in Equations (19) and (22). The filter above is almost identical to what has been termed the modified truncated second-order filter in Jazwinski [1]. However, another term, $-\frac{1}{4}(\mathbf{X}_{k+1}\mathbf{h}_{xx}(\mathbf{x}_{k+1}|\mathbf{k},\mathbf{t}_{k+1}))(\mathbf{X}_{k+1}|\mathbf{k}^{\mathbf{h}}_{xx}(\mathbf{x}_{k+1}|\mathbf{k}'\mathbf{t}_{k+1}))^{T}$ appears inside the brackets in Equation (22) in that book. Using the previous proposition we again find that this term should be dropped.

The filtering equations, Equations (20) - (22), can easily be obtained by assuming Gaussian distributions and simply computing the first and second order moments of the posterior pdf. $p(x_{k+1}|Y_{k+1})$.

3. CONTINUOUS-TIME TRUNCATED SECOND ORDER FILTER

Consider a system governed by the following set of Ito stochastic differential equations:

$$dx(t) = f(x(t),t)dt + G(x(t),t)d\beta(t)$$
(25)

$$dz(t) = h(x(t),t)dt + d\eta(t)$$
(26)

where $\{\beta(t),t\geq t_0\}$ and $\{\eta(t),t\geq t_0\}$ are mutually independent Wiener processes, both assumed to be independent of $x(t_0)$. Furthermore, assume $E(d\beta(t)d\beta^T(t))=V(t)dt$ and $E(d\eta(t)d\eta^T(t))=W(t)dt$.

Define $\hat{x}(t|t)$ and X(t|t) to be, respectively, the mean and covariance matrix of x(t) given $Z(t) = \{z(\tau), t_0 \le \tau \le t\}$. Furthermore, let E_t denote the expectation operator with respect to $\sigma(Z(t))$. Equations for the evaluation of $\hat{x}(t|t)$ and X(t|t) are derived in Jazwinski [1], pp.182-184. For the conditional mean $\hat{x}(t|t)$ we have

$$d\hat{x}(t|t) = E_{t}(f(x(t),t))dt + [E_{t}(x(t)h^{T}(x(t),t))$$

$$- \hat{x}(t|t)E_{t}(h^{T}(x(t),t))]W^{-1}(t)[dz(t) - E_{t}(h(x(t),t))dt]$$
(27)

Taking nonlinearities to second order and neglecting third and fourth-order central moments, we end up with

$$d\hat{x}(t|t) = [f(\hat{x}(t|t),t) + \frac{1}{2}(X(t|t)f_{xx}(\hat{x}(t|t),t))]dt + K(t)\{dz(t) - [h(\hat{x}(t|t),t) + \frac{1}{2}(X(t|t)h_{xx}(\hat{x}(t|t),t))]dt\}$$
(28)

where

$$K(t) = X(t|t)h_{X}^{T}(\hat{x}(t|t),t)W^{-1}(t)$$
 (29)

$$(\mathbf{X}(\mathsf{t}\,|\mathsf{t})\,\mathbf{f}_{\mathbf{x}\mathbf{x}}(\hat{\mathbf{x}}(\mathsf{t}\,|\mathsf{t})\,\mathsf{,t}))_{\mathbf{i}} = \sum_{\mathbf{j},\mathbf{k}=1}^{n} \mathbf{X}_{\mathbf{j}\mathbf{k}}(\mathsf{t}\,|\mathsf{t})\,\frac{\partial^{2}\mathbf{f}_{\mathbf{i}}}{\partial\mathbf{x}_{\mathbf{j}}\partial\mathbf{x}_{\mathbf{k}}}(\hat{\mathbf{x}}(\mathsf{t}\,|\mathsf{t})\,\mathsf{,t}) \tag{30}$$

$$(\forall (t|t) h_{xx}(\hat{x}(t|t),t))_{i} = \sum_{j,k=1}^{n} x_{jk}(t|t) \frac{\partial^{2} h_{i}}{\partial x_{j} \partial x_{k}} (\hat{x}(t|t),t)$$
(31)

The (i,t) th component of X(t|t) will satisfy the equation

$$dx_{ij} = \{ [E_{t}(x_{i}f_{j}) - \hat{x}_{i}E_{t}(f_{j})] + [E_{t}(x_{j}f_{i}) - \hat{x}_{j}E_{t}(f_{i})] + E_{t}(GVG^{T})_{ij}$$

$$- [E_{t}(x_{i}h) - \hat{x}_{i}E_{t}(h)]^{T}W^{-1}[E_{t}(x_{j}h) - \hat{x}_{j}E_{t}(h)] \}dt$$

$$+ [E_{t}(x_{i}x_{j}h) - E_{t}(x_{i}x_{j})E_{t}(h) - \hat{x}_{i}E_{t}(x_{j}h)$$

$$- \hat{x}_{i}E_{t}(x_{i}h) + 2\hat{x}_{i}\hat{x}_{i}E_{t}(h)]^{T}W^{-1}[dz-E_{t}(h)dt]$$
(32)

where all arguments of the functions have been dropped for the sake of simplicity.

First, consider the term in front of $[dz-E_t(h)dt]$. Taking the non-linearities to second order, we find

$$[E_{t}(x_{i}x_{j}h) - E_{t}(x_{i}x_{j})E_{t}(h) - \hat{x}_{i}E_{t}(x_{j}h) - \hat{x}_{j}E_{t}(x_{i}h) + 2\hat{x}_{i}\hat{x}_{j}E_{t}(h)]$$

$$= \frac{1}{2}E_{t}\{(x_{i}-\hat{x}_{i})(x_{j}-\hat{x}_{j})[(x-\hat{x})(x-\hat{x})^{T}h_{xx}]\} - \frac{1}{2}X_{ij}(xh_{xx})$$
(33)

where the term $[(x-\hat{x})(x-\hat{x})^Th_{XX}]$ has a similar interpretation as the term (Xh_{XX}) which is defined in Equation (31). Neglecting fourth-order central moments in Equation (33) and not being aware of the previous proposition, the term $-\frac{1}{2}X_{ij}(Xh_{XX})$ has been retained in previous papers on truncated second-order filters, thus obtaining a random forcing term in the covariance equation. However, the proper approximation of the right-hand side of Equation (3.3) is actually 0 which completely eliminates the random forcing term from the covariance equation.

Taking the nonlinearities of the remainder of Equation (31) to second order, we finally end up with the covariance equation being

$$\frac{d}{dt}X(t|t) = f_{\mathbf{x}}(\hat{\mathbf{x}}(t|t),t)X(t|t) + X(t|t)f_{\mathbf{x}}^{\mathbf{T}}(\hat{\mathbf{x}}(t|t),t)$$

$$+ G(\hat{\mathbf{x}}(t|t),t)V(t)G^{\mathbf{T}}(\hat{\mathbf{x}}(t|t),t)$$

$$+ (X(t|t)V(t)G_{\mathbf{x}}^{2}(\hat{\mathbf{x}}(t|t),t))$$

$$+ \frac{1}{2}(X(t|t)V(t)G_{\mathbf{xx}}(\hat{\mathbf{x}}(t|t),t)G(\hat{\mathbf{x}}(t|t),t))$$

$$+ \frac{1}{2}(X(t|t)V(t)G_{\mathbf{xx}}(\hat{\mathbf{x}}(t|t),t)G(\hat{\mathbf{x}}(t|t),t))^{\mathrm{T}}$$

$$- X(t|t)h_{\mathbf{x}}^{\mathbf{T}}(\hat{\mathbf{x}}(t|t),t)W^{-1}(t)h_{\mathbf{x}}(\hat{\mathbf{x}}(t|t),t)X(t|t)$$
(34)

where the matrices $(X(t|t)V(t)G_X^2(\hat{x}(t|t),t))$ and $(X(t|t)V(t)G_{XX}(\hat{x}(t|t),t)G(\hat{x}(t|t),t))$ have a similar interpretation as in the previous section.

Summing up, the correct form of the truncated second-order nonlinear filter for continuous-time systems consists of the following equations

$$d\hat{x}(t|t) = [f(\hat{x}(t|t),t) + \frac{1}{2}(X(t|t)f_{XX}(\hat{x}(t|t),t))]dt$$

$$+ K(t)\{dz(t) - [h(\hat{x}(t|t),t) + \frac{1}{2}(X(t|t)h_{XX}(\hat{x}(t|t),t))]dt\}$$

$$+ K(t)\{dz(t) - [h(\hat{x}(t|t),t) + \frac{1}{2}(X(t|t)h_{XX}(\hat{x}(t|t),t))]dt\}$$

$$+ K(t)\{dz(t) - [h(\hat{x}(t|t),t) + \frac{1}{2}(X(t|t),t)]dt\}$$

$$+ K(t)\{dz(t) - [h(\hat{x}(t|t),t) + \frac{1}{2}(X(t|t),t)](\hat{x}(t|t),t)]dt\}$$

$$+ K(t)\{dz(t) + \frac{1}{2}(X(t|t))X(t|t) + \frac{1}{2}(X(t|t),t)X(t|t) + \frac{1}{2}(X(t|t))X(t)G^{T}(\hat{x}(t|t),t)G(\hat{x}(t|t),t))$$

$$+ \frac{1}{2}(X(t|t))X(t)G_{XX}(\hat{x}(t|t),t)G(\hat{x}(t|t),t))^{T} - K(t)W(t)K^{T}(t)$$

$$+ \frac{1}{2}(X(t|t))X(t)G_{XX}(\hat{x}(t|t),t)G(\hat{x}(t|t),t))^{T} - K(t)W(t)K^{T}(t)$$

$$+ \frac{1}{2}(X(t|t))X(t)G_{XX}(\hat{x}(t|t),t)G(\hat{x}(t|t),t))^{T} - K(t)W(t)K^{T}(t)$$

$$+ \frac{1}{2}(X(t|t))X(t)G_{XX}(\hat{x}(t|t),t)G(\hat{x}(t|t),t)$$

where

$$K(t) = X(t|t)h_{\mathbf{v}}^{\mathbf{T}}(\hat{\mathbf{x}}(t|t),t)\mathbf{w}^{-1}(t)$$
(37)

A random forcing term appears in the covariance equation in previous papers on the truncated second-order filter, see, say, Jazwinski [1], but we have shown that this term is not properly present in this filter. A similar term is, however, properly present in the Gaussian second-order filter, but it enters with the opposite sign of what it seemed to do in the truncated filter. Obviously disquieted by this, Jazwinski also suggests a modified second-order filter where the random forcing term is removed from the covariance equation. That filter is identical to the filter which has been derived here, but, as we have already pointed out in the discrete-time case, there is no need to use the term modified.

4. CONCLUSION

We have rederived the truncated second-order nonlinear filter for both discrete-time and continuous-time systems and have shown that previous derivations contain a significant error. This is due to the fact that the full implications of the assumptions have not been recognized in previous papers. The derivations in this paper reveal that what has previously been termed the modified truncated second-order filter is the correct form of the truncated filter provided a small correction is made in the discrete-time case. The term modified can therefore be dropped in connection with the truncated second-order filter.

A truncated second-order filter for discrete-time implicit systems has been derived by Henriksen [10,11]. However, in order to obtain a computationally feasible solution, it is shown in those papers that a stronger set of assumptions than appears in this paper will have to be made.

REFERENCES

- [1] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," Trans. ASME, J. Basic Eng., Vol. 82 (1960), 35-45.
- [2] A. H. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, 1970.
- [3] L. Schwartz, E. B. Stear, "A Computational Comparison of Several Non-linear Minimal Variance Filters," IEEE Trans. Automatic Control, Vol. AC-13 (1968), 83-86.
- [4] A. H. Jazwinski, "Nonlinear Filtering-Numerical Examples," NASA-Goddard Astrodynamics Conference, Greenbelt, Maryland (April 1966).
- [5] T. M. Carney, R. M. Goldwyn, "Numerical Experiments with Various Optimal Estimators," J. Optimization Theory Appl., Vol. 1 (1967), 113-130.
- [6] R. Henriksen, T. O. Olsen, "Comparison of Different Parameter Estimation Methods in Flotation Processes," 4th IFAC Symp. Multivariable Technological Systems, Fredericton, N.B., Canada (July 1977).
- [7] M. L. Andrade Netto, L. Gimeno, M. J. Mendes, "Nonlinear Filtering of Discrete-Time Systems," 4th IFAC Symp. Identification and System Parameters Estimation, Tbilisi, USSR (September 1976).
- [8] A. H. Jazwinski, "Filtering for Nonlinear Dynamical Systems," IEEE Trans. Automatic Control, Vol. AC-11 (1966), 765-766.
- [9] R. D. Bass, V. D. Norum, L. Schwartz, "Optimal Multichannel Nonlinear Filtering," J. Math. Anal. Appl., Vol. 16 (1966), 152-164.
- [10] R. Henriksen, "Nonlinear Filtering in Econometric Models," Princeton University, Econometric Research Program, Princeton, New Jersey, Research Memorandum No. 240 (April 1979).
- [11] R. Henriksen, "Nonlinear Filtering in Discrete-Time Implicit Models,"
 Princeton University, Econometric Research Program, Princeton, New
 Jersey, paper (April 1979).