2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

A COMBINED VARIABLE STRUCTURE AND KALMAN FILTERING APPROACH

Saeid Habibi, Member, ASME, IEEE
McMaster University, Hamilton, Canada

ABSTRACT

The Extended Kalman Filter (EKF) can be used for parameter as well as for state estimation. However, when used for parameter estimation, EKF is sensitive to modeling uncertainties and prone to instability. In this paper, the Variable Structure filtering concept is introduced and used for modifying the EKF into a robust form. The proposed method is model-based and has a guaranteed stability given a stable internal model. The application of the new filter for parameter estimation in a system that is described by a smooth nonlinear function is presented. The performance of the new filter is discussed by its comparison to an equivalent Extended Kalman Filter.

1. INTRODUCTION

Unscheduled preventative maintenance is a concept that is of significant economic value to companies concerned with safety or involved in mass production and distribution (e.g. the aerospace and the process industries). In this concept, the inception of a fault is detected in real-time, diagnosed, and used to trigger preventative maintenance. There are various forms of fault detection strategies that commonly involve vibrations analysis, neural networks, or model-based prediction. Model-based strategies are increasingly used in some niche applications such as in the fluid power as they are more conducive to fault diagnosis.

In model-based fault detection, an internal physical model is used to predict and estimate the operation of the system and its outputs. If there is an error between the predicted and the actual outputs of the system, then this error is used to correct the estimation model. This correction is achieved through changes in some pre-selected parameters that have physical significance and can be specifically linked to fault conditions. Changes in these parameters are monitored in real-time and compared to predetermined thresholds that would signal gradual or abrupt deteriorations in the plant operation as a result of fault conditions. As such, parameter estimation not only detects faults but also provides considerable information for the fault diagnosis. The Extended Kalman Filter (EKF) is a model-based strategy that is commonly used for parameter estimation and fault detection. However, the EKF formulation for this application is overly sensitive to modeling uncertainties that occur after the tuning of the filter as a result of a fault condition. In this paper, a new parameter estimation concept that is robust to modeling uncertainties and referred to as the Variable Structure Filter (VSF) is considered, [1]. The VSF and EKF concepts are then combined into a new form that takes advantage of the robustness of the VSF and of the performance of the EKF. The Nomenclature used in this paper is provided in Section 2. The use of the Extended Kalman Filter (EKF) for parameter estimation is briefly reviewed in sections 3 and 4. A description of the VSF is provided in section 5. A new EKF/VSF estimation strategy is presented in Section 6 and compared to an equivalent EKF design. The concluding remarks are contained in Section 7.

2. NOMENCLATURE

Matrices and vectors are denoted by using bold letters. Their elements are denoted by italic lower case letters with subscripts *i* and/or *j*. *k* denotes calculation step. Subscripts *k*|*k* and *k*|*k-1* are used to identify *a posteriori* and *a priori* estimates. The symbols ^ and ~ are used to identify an estimated value and the error in the estimated value. The subscript **max** is used to identify an upper bound. The symbols *E*, ' and + denote expectation, transpose, and pseudo inverse. The symbol $\left| \; \right|_{\mathbf{ABS}}$ is used to denote a matrix or vector made up of absolute values such that $\left|\mathbf{e}_{z_{k|k-1}}\right|_{\mathbf{ABS}}$ has elements that have the absolute value of the corresponding elements of $\mathbf{e}_{z_{k|k-1}}$. The nomenclature is provided in the following table. Note that the elements of vectors and matrices are not listed and can be readily identified by the above-mentioned notation.

Symbol	Comments	
B	Input multiplier.	1 x 1
$\mathbf{e}_{z_{k\|k}}, \mathbf{e}_{z_{k\|k-1}}$	Output estimation error calculated (a posteriori and a priori)	*m* x 1
G	Input matrix.	*n* x *p*
$\mathcal{F}, \mathcal{H}$	Nonlinear functions.	
H	Output matrix.	*m* x *n*
I	Identity matrix.	
i,j	Subscripts used to identify elements of matrices and vectors.	1 x 1
K	Calculation step index.	1 x 1
$\mathbf{K}_k$	VSF gain.	*n* x 1
$\mathbf{K}_{\mathbf{Kalman}}$	Kalman gain.	*n* x 1
M	Number of measurements.	1 x 1
n	Number of states.	1 x 1

p	Number of inputs.	1 x 1
$\mathbf{P}$	Error covariance matrix.	m x m
$\mathbf{Q}$	System noise covariance.	n x n
$\mathbf{R}$	Measurement noise cov.	m x m
Sat	Saturation function.	
$\mathbf{s}_{\mathbf{gn}}$	$\mathrm{s_{gn}}\left(\mathbf{e}_{z_{k\mid k-1}}\right)=[sign(e_{z_{1_{k\mid k-1}}})\cdots]'$	
$\mathbf{u}$	Input.	m x 1
Vec	Any vector	
$\mathbf{v}, \mathcal{V}, \mathbf{V}_{\max}$	Meas. noise and its bound.	m x 1
$\mathbf{w}, \mathcal{w}, \mathbf{W}_{\max}$	System noise and its bound.	n x 1
$\mathbf{x}, \mathbf{x}_o$	States and initial conditions.	n x 1
$\hat{\mathbf{x}}_{k\mid k}, \hat{\mathbf{x}}_{k\mid k-1}$	A posteriori, a priori est.	m x 1
$\mathbf{z}$	Measured output.	m x 1
$\hat{\mathbf{z}}_{k\mid k}, \hat{\mathbf{z}}_{k\mid k-1}$	A posteriori/priori output est.	m x 1
$\boldsymbol{\alpha}$	Diag. matrix of rate of decay.	n x n
β	Constant.	m x 1
ν	Lyapunov function.	
$\boldsymbol{\Phi}$	System matrix.	n x n
$\boldsymbol{\gamma}$	Diag.matrix with $\gamma_{ii} \geq 1$.	n x n
τ	Sampling time.	1 x 1
$\boldsymbol{\Psi}$	Boundary layer	n x 1
ω	Natural frequency.	1 x 1
ς	Damping ratio.	1 x 1

3. Review Of The Extended Kalman Filter

Many practical applications of estimation pertain to systems that are modeled by smooth nonlinear functions as

$$\mathbf{x}_{k+1} = \mathcal{F}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{w}_k) \tag{1.}$$

$$\mathbf{z}_k = \mathcal{H}(\mathbf{x}_k, \mathbf{v}_k) \tag{2.}$$

As such, both the states and measurements are assumed to be subject to white noise.

The Extended Kalman Filter (EKF) is one of the most widely used tools for state and parameter estimation, [2-5]. The EKF is a predictor-corrector method that uses a nonlinear model (equations (1) and (2)) to generate an initial or *a priori* set of estimates denoted by $\hat{\mathbf{x}}_{k\mid k-1}$. These *a priori* estimates are then corrected into an *a posteriori* or refined form, $\hat{\mathbf{x}}_{k\mid k}$. Further to initial assumption of white noise, the EKF correction is optimal in magnitude and has a direction that is orthogonal to the true trajectory of the states or the parameters that are being estimated, [2]. The magnitude of this correction is obtained by using a linearized model of the problem along a nominal trajectory. This linearized model is a truncated Taylor series expansion where the higher order terms are neglected. It is valid if the nominal and actual state trajectories are sufficiently close; otherwise, the neglected higher order terms of the Taylor series expansion become significant thus potentially causing instability. An effective strategy for dealing with this problem is to use the last best estimate of the trajectory for deriving the linearized model. Hence the nominal operating point for linearization may be defined as: $\hat{\mathbf{x}}_{k\mid k}$ for the state equation (1) and $\hat{\mathbf{x}}_{k\mid k-1}$ for the output equation (2). The linearization in effect implies that the state trajectory is divided into successive piece-wise linear regions, and as the trajectory moves from one region to the next, the linearized model is updated accordingly. Within a piece-wise linear region, the model of equations (1) and (2) can be expressed as:

$$\begin{aligned}
\mathbf{x}_{k+1} &= \mathcal{F}(\hat{\mathbf{x}}_{k\mid k}, \mathbf{u}_k, \mathrm{w}_k) + \left.\frac{\partial \mathcal{F}(\mathbf{x}, \mathbf{u}, \mathrm{w})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}_{k\mid k}, \mathrm{w}_k=0} \left(\mathbf{x}_k - \hat{\mathbf{x}}_{k\mid k}\right) + \mathcal{w}_k \\
&= \mathcal{F}(\hat{\mathbf{x}}_{k\mid k}, \mathbf{u}_k, \mathrm{w}_k) + \Phi_k\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k\mid k}\right) + \mathcal{w}_k \\
\mathbf{z}_k &= \mathbf{z}_k^{nom} + \left.\frac{\partial \mathcal{H}(\mathbf{x}, \mathrm{v})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}=\hat{\mathbf{x}}_{k\mid k-1}, \mathrm{v}_k=0} \left(\mathbf{x}_k - \hat{\mathbf{x}}_{k\mid k-1}\right) + \mathcal{V}_k \\
&= \hat{\mathbf{z}}_{k\mid k} + \mathbf{H}_k\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k\mid k-1}\right) + \mathcal{V}_k
\end{aligned} \tag{4.}$$

where:

$$\mathbf{\Phi}_k = \left.\frac{\partial \mathcal{F}(\mathbf{x}, \mathbf{u}, \mathbf{w})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}=\hat{\mathbf{x}}_{k\mid k}, \mathbf{u}=\mathbf{u}_k, \mathbf{w}=0}, \quad \mathcal{W}_k = \left.\frac{\partial \mathcal{F}(\mathbf{x}, \mathbf{u}, \mathbf{w})}{\partial \mathbf{w}}\right|_{\mathbf{x}=\mathbf{x}=\hat{\mathbf{x}}_{k\mid k}, \mathbf{u}=\mathbf{u}_k, \mathbf{w}=0} \mathbf{w}_k,$$

$$\mathbf{H}_k = \left.\frac{\partial \mathcal{H}(\mathbf{x}, \mathbf{v})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}=\hat{\mathbf{x}}_{k\mid k-1}, \mathbf{v}=0} \text{,and}$$

$$\mathcal{V}_k = \left.\frac{\partial \mathcal{H}(\mathbf{x}, \mathbf{v})}{\partial \mathbf{v}}\right|_{\mathbf{x}=\mathbf{x}=\hat{\mathbf{x}}_{k\mid k-1}, \mathcal{V}=0} \mathbf{v}_k.$$

The EKF caters for the presence of noise in the system and its measurement. The characterization of process and measurement noise is made through covariance matrices $\mathbf{R}_k$ and $\mathbf{Q}_k$. Further to the model of equations (3) and (4), the EKF process may be summarized as follows:

1. An *a priori* error covariance matrix $\mathbf{P}_{k\mid k-1}$ is obtained and used in the calculation of an optimal EKF gain vector $\mathbf{K}_{\mathbf{Kalman}\,k}$ such that:

$$\mathbf{P}_{k\mid k-1} = \mathbf{\Phi}_{k-1}\mathbf{P}_{k-1\mid k-1}\mathbf{\Phi}_{k-1}^T + \mathbf{Q}_{k-1} \tag{5.}$$

$$\mathbf{K}_{\mathbf{Kalman}\,k} = \mathbf{P}_{k\mid k-1}\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_{k\mid k-1}\mathbf{H}_k^T + \mathbf{R}_k)^{-1} \tag{6.}$$

2. The *a priori* state estimate is refined into *a posteriori* state estimate $\hat{\mathbf{x}}_{k\mid k}$ such that:

$$\hat{\mathbf{x}}_{k\mid k} = \hat{\mathbf{x}}_{k\mid k-1} + \mathbf{K}_{\mathbf{Kalman}\,k}(\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_{k\mid k-1}) \tag{7.}$$

3. The next *a priori* state estimate and the *a posteriori* error covariance matrix are predicted such that:

$$\hat{\mathbf{x}}_{k+1\mid k} = \hat{\mathcal{F}}(\hat{\mathbf{x}}_{k\mid k}, \mathbf{u}_k) \tag{8.}$$

$$\mathbf{P}_{k\mid k} = (\mathbf{I} - \mathbf{K}_{\mathbf{Kalman}\,k}\mathbf{H}_k)\mathbf{P}_{k\mid k-1} \tag{9.}$$

4. Steps 1 to 3 are repeated for each time step.

In calculating the optimal correction vector, $\mathbf{K}_{\mathbf{Kalman}\,k}(\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_{k\mid k-1})$, the EKF does not explicitly account for modeling uncertainties. The definition of the noise covariance matrices $\mathbf{R}_k$ and $\mathbf{Q}_k$ are used to implicitly capture modeling uncertainties and have a strong influence on the stability and performance of the filter. The initial conditions for the Kalman gain pertaining to $\mathbf{P}_{0|0}$ and $\hat{\mathbf{x}}_{0|0}$ also strongly influence the performance. Generally, an excellent estimation performance is achieved by tuning the covariance matrices and the initial conditions by trial and error.

4. Parameter Estimation Using The Extended Kalman Filter

In parameter estimation, the internal model of the filter is formulated such that the parameter that is being estimated is treated as a state. For example, consider the problem of estimating the damping ratio of a second order system, (taken from [2]), where:

$$\ddot{x} + 2\zeta\omega\dot{x} + \omega^2 x = bu + w(t) \tag{10.}$$

$$z(t) = x(t) + v(t) \tag{11.}$$

The system of equations (10) and (11) has two states $x_1 = x$ and $x_2 = \dot{x}$. If the damping ratio ζ is to be estimated, then it is defined as a third state $x_3 = \zeta$, thus transforming the problem into a nonlinear form. Using a simplified discrete representation, the system equations become:

$$\begin{bmatrix} x_{1_{k+1}} \\ x_{2_{k+1}} \\ x_{3_{k+1}} \end{bmatrix} = \begin{bmatrix} x_{1_k} + \tau x_{2_k} \\ \left(1 - 2\omega\tau x_{3_k}\right)x_{2_k} - \omega^2 \tau x_{1_k} \\ x_{3_k} \end{bmatrix} + \begin{bmatrix} w_{1_k} \\ w_{2_k} \\ w_{3_k} \end{bmatrix} + \begin{bmatrix} 0 \\ b \\ 0 \end{bmatrix} u \tag{12.}$$

$$z_k = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{1_k} \\ x_{2_k} \\ x_{3_k} \end{bmatrix} + v_k \tag{13.}$$

In applying the EKF strategy to this estimation problem, the internal model of the filter may be defined from equations (12) and (13), as:

$$\hat{\mathbf{x}}_{k+1|k} = \hat{\mathcal{F}}(\hat{\mathbf{x}}_{k|k}, \mathbf{u}_k) = \begin{bmatrix} \hat{x}_{1_{k|k}} + \tau\hat{x}_{2_{k|k}} \\ \left(1 - 2\hat{\omega}\tau\hat{x}_{3_{k|k}}\right)\hat{x}_{2_{k|k}} - \hat{\omega}^2 \tau\hat{x}_{1_{k|k}} \\ \hat{x}_{3_{k|k}} \end{bmatrix} + \begin{bmatrix} 0 \\ \hat{b} \\ 0 \end{bmatrix} u \tag{14.}$$

$$\hat{\mathbf{z}}_{k|k} = \hat{\mathbf{H}}\hat{\mathbf{x}}_{k|k} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_{1_{k|k}} \\ \hat{x}_{2_{k|k}} \\ \hat{x}_{3_{k|k}} \end{bmatrix} \tag{15.}$$

The corresponding estimated system and output matrices of the linearized model become:

$$\hat{\mathbf{\Phi}}_k = \begin{bmatrix} 1 & \tau & 0 \\ -\hat{\omega}^2\tau & \left(1 - 2\hat{\omega}\tau\hat{x}_{3_{k|k}}\right) & -2\hat{\omega}\tau\hat{x}_{2_{k|k}} \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \hat{\mathbf{H}} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}.$$

In this example, let the model be exactly known such that: $\hat{\omega} = \omega = 5rad/s$ and $\hat{b} = b = 12$. The simulation time is 30 secs and, during simulation, the value of ζ is changed from 0.2 to 0.5 to 0.9. The sampling interval τ is 0.01secs and, the initial value of the actual and estimated states are specified as $\begin{bmatrix} x_{1_0} \\ x_{2_0} \\ x_{3_0} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0.2 \end{bmatrix}$ and $\begin{bmatrix} \hat{x}_{1_{0|0}} \\ \hat{x}_{2_{0|0}} \\ \hat{x}_{3_{0|0}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0.1 \end{bmatrix}$. The input to the system is a random signal in the range of -1 to 1, and the single measurement from the system is subjected to white noise of maximum amplitude of $\mathbf{V}_{\max} = 0.1$, corresponding to 10% of range. All system states are subjected to white noise of amplitude, $\mathbf{W}_{\max} = \begin{bmatrix} 0.001 \\ 0.001 \\ 0.001 \end{bmatrix}$.

Details for the design of an EKF to this estimation problem is provided in [2]. The EKF here is tuned by trial error such that the initial error covariance matrix $\mathbf{P}_{0|0}$ and, its noise covariant matrices are set to: $\mathbf{P}_{0|0} = \begin{bmatrix} 20 & & \\ & 20 & \\ & & 20 \end{bmatrix}$,

$\mathbf{Q}_k = \tau \begin{bmatrix} 0.1 & & \\ & 0.1 & \\ & & 0.1 \end{bmatrix}$ and $\mathbf{R}_k = 0.1$.

In a real practical application, the system states are not available and therefore the exact value of the estimation error is unknown. In this example however, the study is conducted by using computer simulation and, therefore the states and the exact estimation errors are available and can be used for evaluating the effectiveness of the filter. Figures 1a to 1c illustrate the tracking performance of the EKF filter. Note that the change in the damping ratio is detected very well by the EKF; however its accuracy is poor due to the low sensitivity of the model to this parameter. Notwithstanding the accuracy of the estimation process, the objective of this paper is to improve the stability problem of the EKF due to modeling uncertainties. The performance of the EKF is severely degraded when modeling uncertainties are introduced. The estimation process becomes unstable when the estimated natural frequency $\hat{\omega}$ is changed from its exact value of $\hat{\omega} = \omega = 5rad/s$ to any value beyond $\hat{\omega} = 10rad/s$. The stability of the EKF when used for parameter estimation is much dependent on modeling uncertainties; more so than when applied to state estimation. In the parameter estimation problem formulation, the associated state for the parameter under consideration (e.g. x_3 for the damping factor in this example) is artificial and loosely related to the other state variables. As such, referring to equation (25), $\hat{x}_{3_{k+1|k}} = \hat{x}_{3_{k|k}}$ and its change is solely dependent on the EKF corrective action:

$$\hat{x}_{3_{k+1|k+1}} = \hat{x}_{3_{k+1|k}} + \sum_{i=1}^{m} K_{Kalman_{i_k}} (z_{i_k} - \sum_{j=1}^{n} H_{ij}\hat{x}_{j_{k|k-1}}).$$

Inaccuracies due to the linearization process or modeling uncertainties in the derivation of $\mathbf{K}_{\mathbf{Kalman}}$ therefore have a strong influence on the estimated parameter and hence the overall estimation model. Furthermore, a requirement for stability is that the estimation model that is used in the *a* *priori* stage of the process be stable. In state estimation, this internal model is unaffected by the estimation process and remains as predefined. In parameter estimation, the internal model is updated continuously and as such inaccuracies in the EKF corrective term could cause deviations that would result in an unstable internal model. In the following section, the EKF strategy is modified in order to improve its stability by using the Variable Structure Filtering Concept, initially introduced in a different form in [1].

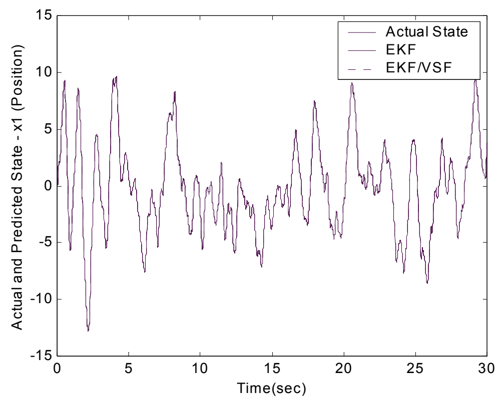

Figure 1a: Estimated and Actual States Associated with State 1 (Position)

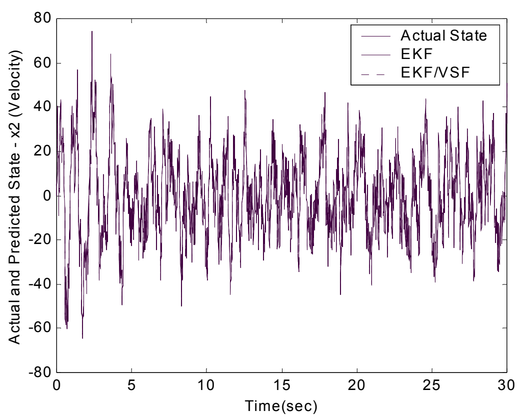

Figure 1b: Estimated and Actual States Associated with State 2 (Velocity)

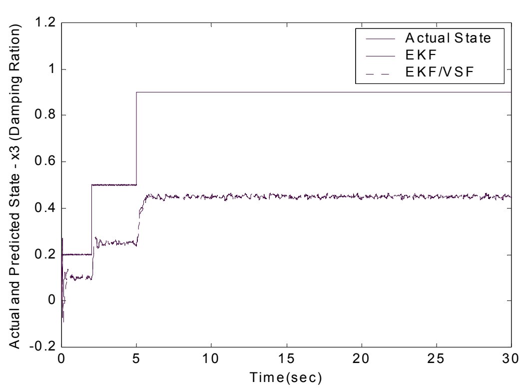

Figure 1c: Estimated and Actual States Associated with State 3 (Damping ratio)

Figure 1: Estimated and Actual States

5. THE VARIABLE STRUCTURE FILTER CONCEPT

The state estimation concept presented in this paper is referred to as the Variable Structure Filter (VSF), [1]. The VSF is model-based and can be used to track a state trajectory in time. It is a predictor corrector method that uses an internal uncertain model of the system to produce an initial estimate of the states and the system outputs. The measured and estimated outputs are compared, and a corrective term proportional to this error is then applied to the states in order to improve their accuracy. The VSF can be applied to systems that are modeled by a smooth function as specified in equation (1). It is assumed however that the relationship between the measurement signals and the states is linear or at the very least piece-wise linear. This assumption is valid for most systems as the sensors that are used for measurement are in general, specifically designed to be linear over their operating range and are well calibrated and characterized. As such, let:

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k \tag{16.}$$

Furthermore in the majority of applications, the output matrix $\mathbf{H}$ is positive and pseudo-diagonal.

Further to the model of equations (1) and (16), the VSF estimation process can be summarized as follows.

1 An *a priori* state estimate is predicted by using the estimated model of the system such that:

$$\hat{\mathbf{x}}_{k+1|k} = \hat{\mathcal{F}}(\hat{\mathbf{x}}_{k|k}, \mathbf{u}_k) \tag{17.}$$

This estimate is obtained by using the previous *a posteriori* state estimate $\hat{\mathbf{x}}_{k|k}$ or, at the inception of the process, by using the initial conditions, $\mathbf{x}_o$. The estimated states are then used for predicting the *a priori* estimates of measurements such that:

$$\hat{\mathbf{z}}_{k+1|k} = \hat{\mathbf{H}}\hat{\mathbf{x}}_{k+1|k} \tag{18.}$$

2 A corrective gain $\mathbf{K}_{k+1} \in \mathbf{R}^{n\times 1}$ is calculated as a function of the error in the *a priori* predicted output.

3 The *a priori* state estimate is refined into an *a posteriori* form such that:

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1} \tag{19.}$$

4 Steps 1 to 3 are repeated for each time step.

The corrective VSF term $\mathbf{K}_k$ is discontinuous and switches between two extreme values in order to force the estimated states towards the actual system states. The strategy is formulated such that with every iteration, the magnitude of the estimation error is reduced as shown in Figure 1. This effect may be stated as:

$$E\left[\left|\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k+1}\right|_{\mathbf{ABS}}\right] \le E\left[\left|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right|_{\mathbf{ABS}}\right] \tag{20.}$$

If condition (20) is satisfied, the system is evidently stable and convergent such that:

$\lim_{k\to\infty}\left(\left|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right|_{\mathbf{ABS}}\right) \le \beta$ where β is constant.

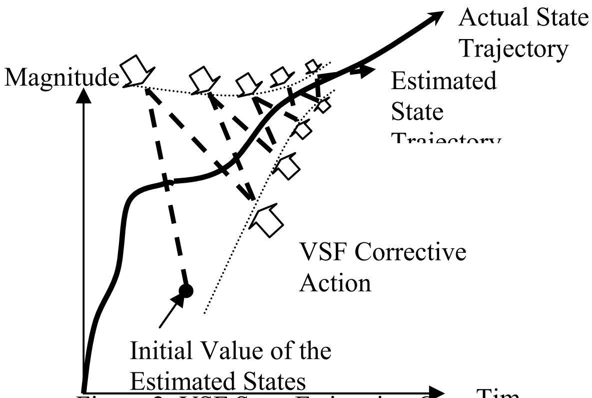

Figure 2: VSF State Estimation Concept

The VSF strategy achieves stability and is derived by using concepts closely related to the Variable Structure Systems theory, [6-16]. In the VSF, the corrective action switches direction every time that the estimated state trajectory crosses the true state trajectory as shown in Figure 2. As such the true state trajectory can be considered as a discontinuity surface or hyperplane. The true state trajectory is not available and cannot in practice be used as a switching hyperplane. However, if the relationship between the system states and measurements is linear and if the system is observable, then the trajectory of the output can be used as the switching hyperplane. This implies that the more accurate the estimated value of the measured output, the more accurate the estimated states. As such, the error between the estimated and actual value of the measured signals is used as an indicator of the error in the estimated value of the state variables. Two variables that are critical to the VSF process are the *a priori* and the *a posteriori* output error estimates:

$$\mathbf{e}_{\mathbf{z}_{k|k}} = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k} \tag{21.}$$

$$\mathbf{e}_{\mathbf{z}_{k|k-1}} = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1} \tag{22.}$$

These are used as indicators of the error in the *a priori* and the *a posteriori* state estimates. As such, convergence is obtained if the magnitude of the *a posteriori* error is reduced in time. Condition (20) can now be restated as:

$$\left|\mathbf{e}_{z_{k|k}}\right|_{\mathbf{ABS}} < \left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}} \tag{23.}$$

Condition (23) leads to a stable VSF estimation process. This can be verified by defining a discrete Lyapunov function such that $\nu_k = \mathbf{e}_{z_{k|k}} \circ \mathbf{e}_{z_{k|k}}$. The process is stable if $\left(\Delta\nu_k = \mathbf{e}_{z_{k|k}} \circ \mathbf{e}_{z_{k|k}} - \mathbf{e}_{z_{k-1|k-1}} \circ \mathbf{e}_{z_{k-1|k-1}}\right) < 0$, which is a condition satisfied by equation (23). The VSF corrective term can be derived by using condition (23) as stated in the following theorem.

Theorem 1: Further to the VSF estimation process, the stability condition of equation (23) is satisfied if the corrective action $\mathbf{K}_k$ in turn satisfies the following conditions:

$$\left|\mathbf{e}_{z_{k|k-1}}\right|_{\mathbf{ABS}} < \left|\hat{\mathbf{H}}\mathbf{K}_k\right|_{\mathbf{ABS}} < \left|\mathbf{e}_{z_{k|k-1}}\right|_{\mathbf{ABS}} + \left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}} \tag{24.}$$

$$\mathbf{s}_{\mathbf{gn}}(\hat{\mathbf{H}}\mathbf{K}_k) = \mathbf{s}_{\mathbf{gn}}(\mathbf{e}_{z_{k|k-1}}) \tag{25.}$$

Proof: From condition (23), the estimation process is stable if: $\left|\mathbf{e}_{z_{k|k}}\right|_{\mathbf{ABS}} < \left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}}$.

From (24), if $\left|\hat{\mathbf{H}}\mathbf{K}_k\right|_{\mathbf{ABS}} < \left|\mathbf{e}_{z_{k|k-1}}\right|_{\mathbf{ABS}} + \left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}}$ then:

$$\left|\hat{\mathbf{H}}\mathbf{K}_k\right|_{\mathbf{ABS}} - \left|\mathbf{e}_{z_{k|k-1}}\right|_{\mathbf{ABS}} < \left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}} \tag{26.}$$

From (25), if $\mathbf{s}_{\mathbf{gn}}(\hat{\mathbf{H}}\mathbf{K}_k) = \mathbf{s}_{\mathbf{gn}}(\mathbf{e}_{z_{k|k-1}})$, then from (26):

$$\left|\mathbf{e}_{z_{k|k-1}} - \hat{\mathbf{H}}\mathbf{K}_k\right|_{\mathbf{ABS}} < \left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}} \tag{27.}$$

From equations (18),(19), (21) and (22):

$$\mathbf{e}_{z_{k|k}} = \mathbf{e}_{z_{k|k-1}} - \hat{\mathbf{H}}\mathbf{K}_k \tag{28.}$$

Substituting (28) into (27), then $\left|\mathbf{e}_{z_{k|k}}\right|_{\mathbf{ABS}} < \left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}}$ and the stability condition of condition (23) is satisfied. Further to condition (23), for a constant output matrix and white measurement noise, condition (20) is satisfied for completely observable system.

A stable VSF corrective gain that satisfies Theorem 1 may be stated in a simple form as:

$$\mathbf{K}_k = \hat{\mathbf{H}}^{+}\left|\left(\left|\mathbf{e}_{z_{k|k-1}}\right|_{\mathbf{ABS}} + \boldsymbol{\gamma}\left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}}\right)\right|_{\mathbf{ABS}} \circ \mathbf{s}_{\mathbf{gn}}(\mathbf{e}_{z_{k|k-1}}) \tag{29.}$$

where $\boldsymbol{\gamma} \in \mathbf{R}^{m\times m}$ is a diagonal matrix with elements such that $0 < \gamma_{ii} < 1$. The rate of convergence may be obtained from the error equation of the filter. The error equation may be derived from (28), or $\mathbf{e}_{z_{k|k}} = \mathbf{e}_{z_{k|k-1}} - \hat{\mathbf{H}}\mathbf{K}_k$.

Substituting $\hat{\mathbf{H}}\mathbf{K}_k$ from (29) into (28) and rearranging:

$$\mathbf{e}_{z_{k|k}} = \mathbf{e}_{z_{k|k-1}} - \hat{\mathbf{H}}\left(\left(\hat{\mathbf{H}}^{+}\left|\left(\left|\mathbf{e}_{z_{k|k-1}}\right|_{\mathrm{ABS}} + \boldsymbol{\gamma}\left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathrm{ABS}}\right)\right|_{\mathrm{ABS}}\right) \circ \mathrm{s}_{\mathrm{gn}}(\mathbf{e}_{z_{k|k-1}})\right) \tag{30.}$$

For $\hat{\mathbf{H}}$ pseudo-diagonal and positive, equation (30) simplifies to the following:

$$\mathbf{e}_{z_{k|k}} = -\boldsymbol{\gamma}\left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}} \circ \mathbf{s}_{\mathbf{gn}}(\mathbf{e}_{z_{k|k-1}}) \tag{31.}$$

For the elements of the diagonal matrix $\boldsymbol{\gamma}$ defined such that $\gamma_{ii} < 1$, equation (31) leads to a series with a decreasing magnitude. From (31):

$$\left|\mathbf{e}_{z_{k|k}}\right|_{\mathbf{ABS}} = \boldsymbol{\gamma}\left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}} \tag{32.}$$

$\boldsymbol{\gamma}$ is related to the rate of decay $\boldsymbol{\alpha}$ in (32), such that:

$$\boldsymbol{\alpha} = \frac{1}{\tau}\ln(\boldsymbol{\gamma}) \tag{33.}$$

where τ is the sampling time. It should be noted that the stability of the VSF is confirmed by equation (33): for $\gamma_{ii} < 1$, $\boldsymbol{\alpha}$ is negative and signifies decay.

$\mathbf{K}_k$ results in a high frequency of switching that would limit the applications of the VSF and introduce chattering in the estimated states. The chattering may be filtered out by using a smoothing function with a boundary layer $\mathbf{\Psi}$ around the switching surface as shown in Figure 3. Outside of this smoothing boundary layer, the sign function is maintained to ensure stability while inside this layer, $\mathbf{K}_k$ is interpolated to obtain a smooth function. This is achieved by replacing the function $\mathbf{s_{gn}}(\mathbf{vec})$ for any vector **vec**, by the function $\mathbf{sat}(\mathbf{vec},\Psi)$ with elements defined as follows:

$$sat_i(\mathbf{vec},\Psi)=\begin{cases}(vec_i/\psi_i) & \text{for } |vec_i/\psi_i|\le 1\\ sign((vec_i/\psi_i)) & \text{for } |vec_i/\psi_i|>1\end{cases} \tag{34.}$$

6. A COMBINED ESTIMATION STRATEGY USING THE EKF AND SVSF CONCEPT

The VSF provides an estimation process that is sub-optimal albeit stable. It is hence beneficial to be able to combine the optimal performance of the EKF with the stability of the VSF. As such, the VSF strategy can be used to force the estimated states to within a boundary – in this case the smoothing boundary – and then on, the corrective action being principally transferred to the EKF. For ensuring continuity during this transfer, a combined strategy that requires a modification to the VSF is needed. This modified VSF takes advantage and requires the orthogonality of the Kalman correction with respect to the actual system states, [2].

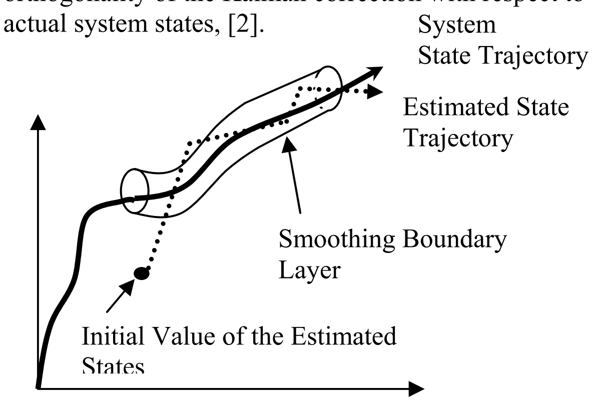

Figure 3: Convergence of the Estimated State Trajectory due to VSF Action

The VSF gain of equation (29) is augmented with a constant positive vector $\mathbf{\Pi}$ and forced in a direction determined by the vector $\mathbf{d}_x$ such that:

$$\mathbf{K_k}=\left(\hat{\mathbf{H}}^{+}\left|\left(\left|\mathbf{e}_{z_{k|k-1}}\right|_{\mathbf{ABS}}+\gamma\left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}}\right)\right|_{\mathbf{ABS}}+\mathbf{\Pi}\right)\circ\mathbf{S_{gn}}(\mathbf{d}_{x_k}) \tag{35.}$$

If $\mathbf{d}_{x_k}$ is a vector that originates from the *a priori* state estimates and is orthogonal to the trajectory traced by the actual system states, then $\mathbf{d}_x$ points towards $\mathbf{x}_k$ and thus the switching surface. The direction of $\mathbf{d}_x$ therefore coincides with the direction of the *a priori* output error vector such that for $\mathbf{\Pi}$ positive, $\mathbf{S_{gn}}(\mathbf{e}_{z_{k|k-1}})=\mathbf{S_{gn}}\left(\hat{\mathbf{H}}_k^{+}\mathbf{\Pi}\circ\mathbf{S_{gn}}(\mathbf{d}_x)\right)$ and $\left(\hat{\mathbf{H}}^{+}\left|\mathbf{e}_{z_{k|k-1}}\right|_{\mathbf{ABS}}\right)\circ\mathbf{S_{gn}}(\mathbf{e}_{z_{k|k-1}})=\left(\hat{\mathbf{H}}^{+}\left|\mathbf{e}_{z_{k|k-1}}\right|_{\mathbf{ABS}}\right)\circ\mathbf{S_{gn}}(\mathbf{d}_{x_k})$.

Substituting $\hat{\mathbf{H}}\mathbf{K}_k$ from (35) into (28) provides the error equation:

$$\mathbf{e}_{z_{k|k}}=-\gamma\left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}}\circ\mathbf{S_{gn}}(\mathbf{d}_x)-\hat{\mathbf{H}}_k\left(\mathbf{\Pi}\circ\mathbf{S_{gn}}(\mathbf{d}_x)\right) \tag{36.}$$

From (36), the maximum and minimum amplitudes of $\left(\left|\mathbf{e}_{z_{k|k}}\right|_{\mathbf{ABS}}\right)_{\mathbf{MAX}}$ are obtained as:

$$\left(\left|\mathbf{e}_{z_{k|k}}\right|_{\mathbf{ABS}}\right)_{\mathbf{MAX}}=\gamma\left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}}+\hat{\mathbf{H}}\mathbf{\Pi} \tag{37.}$$

$$\left(\left|\mathbf{e}_{z_{k|k}}\right|_{\mathbf{ABS}}\right)_{\mathbf{MAX}}=\gamma\left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}}-\hat{\mathbf{H}}\mathbf{\Pi} \tag{38.}$$

Equations (36), (37) and (38) lead to a set of closed loop discrete transfer functions of the form:

$$-\frac{\sum_{j=1}^{m}H_{ij}\pi_j}{z+(\gamma_{ii}\pm\sum_{j=1}^{m}H_{ij}\pi_j)} \tag{39.}$$

Where in the context of equation (39), z is the z-transform variable, and γ_{ii}, H_i, and π_i denote elements of γ, $\hat{\mathbf{H}}$ and $\mathbf{\Pi}$. From (39), the process remains stable if the root of $z+(\gamma_{ii}+\sum_{j=1}^{m}H_{ij}\pi_j)$ and $z+(\gamma_{ii}-\sum_{j=1}^{m}H_{ij}\pi_j)$ are within the unit circle or:

$$\left|(\gamma_{ii}-\sum_{j=1}^{m}H_{ij}\pi_j)\right|<1 \tag{40.}$$

$$\left|(\gamma_{ii}+\sum_{j=1}^{m}H_{ij}\pi_j)\right|<1 \tag{41.}$$

Subject to the conditions of equations (40) and (41), the stability of the estimation process is preserved. The VSF and the EKF strategies may now be combined by setting:

- $\mathbf{d}_{x_k}=\mathbf{K}_{\mathbf{Kalman}_k}\mathbf{e}_{z_{k|k-1}}$ that is orthogonal (by the virtue of EKF derivation) to the actual system states $\mathbf{x}_k$;
- the additive term $\mathbf{\Pi}$ to be relatively large with respect to γ in order to emphasize EKF action within the boundary layer, e.g. $\Pi=0.8\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}$, yet conditions (40) and (41);
- the gain γ set such that conditions (40) and (41) are satisfied e.g. $\gamma=\begin{bmatrix}0.1 & & \\ & 0.1 & \\ & & 0.1\end{bmatrix}$; and
- the boundary layer $\mathbf{\Psi}$ is initially set to $\mathbf{\Psi}=\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}$ in order to as nearly as possible preserve the EKF action.

The gain of the combined EKF/VSF strategy may be stated in a matrix form as:

$$\mathbf{K_k}=\left(\hat{\mathbf{H}}^{+}\left|\left(\left|\mathbf{e}_{z_{k|k-1}}\right|_{\mathbf{ABS}}+\gamma\left|\mathbf{e}_{z_{k-1|k-1}}\right|_{\mathbf{ABS}}\right)+\mathbf{\Pi}\right|_{\mathbf{ABS}}\right)\circ\mathbf{sat}(\mathbf{K}_{\mathbf{Kalman}_k}\mathbf{e}_{z_{k|k-1}},1) \tag{42.}$$

Assuming a pseudo-diagonal output matrix $\hat{\mathbf{H}}$, the elements of the corrective term of equation (42) may be stated as:

$$K_{i_k} = \begin{cases} \left(\hat{H}_{ii}^{+}\left(\left|e_{z_{ik|k-1}}\right| + \gamma_{ii}\left|e_{z_{ik-1|k-1}}\right|\right) + \pi_i\right)\left(K_{Kalman_k} e_{z_{ik|k-1}}\right) & \text{for} \quad \left|K_{Kalman_k} e_{z_{ik|k-1}} / 1\right| \le 1 \\ \left(\hat{H}_{ii}^{+}\left(\left|e_{z_{ik|k-1}}\right| + \gamma_{ii}\left|e_{z_{ik-1|k-1}}\right|\right) + \pi_i\right) & \text{for} \quad \left|K_{Kalman_k} e_{z_{ik|k-1}} / 1\right| > 1 \end{cases} \tag{43.}$$

From equation (43), it can be observed that outside of the smoothing boundary layer, the VSF gain is applied to ensure convergence and stability. Inside the boundary layer, the corrective gain is in a combined form involving both EKF and VSF. The larger the relative magnitude of the term $\mathbf{\Pi}$ with respect to $\boldsymbol{\gamma}$, the more dominant the EKF component of the estimation strategy. Further to the example provided in Section 3, the EKF is augmented into the combined EKF/SVSF form with the gain of equation (41), $\boldsymbol{\gamma} = 0.1\mathbf{I_3}$, $\boldsymbol{\Psi} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{\Pi} = \begin{bmatrix} 0.8 \\ 0.8 \\ 0.8 \end{bmatrix}$. For a known model where $\hat{\omega} = \omega = 5 rad/s$, the estimation results from the combined EKF/VSF strategy are nearly identical to the EKF as shown in Figures 1a to 1c. However, the new strategy is considerably more stable. Although, the performance of the new filter is degraded without any adjustments to the filter settings, the process is stable with introduced errors that move the internal model close to marginal stability. The filter remains stable for large errors such as $\hat{\omega} = 500 rad/s$ or nearly 10000% error in the estimated value of the natural frequency! The estimation results for a 1000% error in this value, i.e. $\hat{\omega} = 55 rad/s$ are shown in Figures 4a to 4c. The results confirm the theoretical expectations form the combined strategy in that the performance of EKF is retained while the stability characteristics are considerably improved. Some degradation of performance can be observed in the estimation results in the form of chattering in Figures 4b and 4c. Intuitively, if the estimation model of equations (17) and (18) are exact, and if there is no noise in the system, then the *a priori* estimates and the real states coincide. The inaccuracy in estimation is the result of uncertainties in the initial conditions, the model, and the system and the measurement noise. The magnitude of the corrective action is proportional to the level of uncertainties and noise. The level of chattering is in turn proportional to the magnitude of the corrective term. Hence, the larger the uncertainty and noise, the larger the chattering effect and the larger the width of the smoothing boundary layer that is needed to alleviate chattering. A time varying boundary layer may be used as suggested for sliding mode control by [12]. In this example, some improvement in performance is obtained by increasing the width of the boundary layer to $\boldsymbol{\Psi} = \begin{bmatrix} 1 \\ 1 \\ 100 \end{bmatrix}$ for the case involving modeling error i.e. $\hat{\omega} = 55 rad/s$. The simulation results with the increased boundary layer are provided in Figures 5a to 5c.

7. Conclusions

In this paper, the use of the Extended Kalman Filter for parameter estimation is briefly reviewed. In parameter estimation, the EKF strategy is very sensitive to modeling uncertainties and is suceptable to instability. The concept of Variable Structure Filtering is reviewed and combined with the EKF. This combined strategy retains the near optimal performance of the EKF when applied to an uncertain system. It has the added benefit of presenting a considerable improvement in the robustness of the estimation process. A simple example is provided to demonstrate the implementation and advantages of the combined method.

REFERENCES

[1] S.R. Habibi and R. Burton, "The Variable Structure Filter", *ASME J.Dyn.Syst. Meas.,&Control*, Vol. 125.

[2] M.S. Grewal and A.P. Andews, "Kalman filtering theory and practice", 2nd edition, *Wiley,* 2001.

[3] S. Haykin, "Introduction to adaptive Filters" *Macmillam Publishing Co*., 1984.

[4] T. Kailath, "A review of 3 decades of linear filtering theory", *IEEE T.Inf.Theory,* Vol.20,No.2, 1974

[5] H. W. Sorenson, "Least-squares estimation: from Gauss to Kalman", *IEEE Spectrum*, pp.63-68, 1970.

[6] V.I. Utkin, "Variable structure systems with sliding mode"*, IEEE T.Aut.Contr.*,Vol.AC-22, 1977.

[7] V.I. Utkin, "Sliding mode and their application in variable structure systems", *Mir Pub.,* 1978.

[8] E.P. Ryan and M. Corless, "Ultimate boundedness and asymptotic stability of a class of uncertain dynamical systems via continuous and discontinuous feedback control", *J. Math. Control and Information*, Vol. 1.

[9] R.G. Morgan and U. Ozguner, "A decentralised variable structure control algorithm for robotic manipulators," *IEEE J.Rob.&Aut.*,Vol.1, No. 1, 1985.

[10] K.S. Yeung and V.P. Chen, "A new controller design for manipulators using the theory of VSS", *IEEE T.Aut.Conr.*,Vol.33,No.2, 1988.

[11] S.R. Habibi and R.J. Richards,"Sliding mode control of an electrically powered industrial robot", *IEE Proc. D: Cont.Theory& appl.,*Vol.139,No.2, 1992.

[12] H. Asada and J.J.E. Slotine, "Robot analysis and control", *John Wiley Publications*, 1986.

[13] E. Bailey and A. Arapostathis, "Simple SMC scheme applied to robots", *Int. J. Cont.*, Vol. 45, No. 4, 1987.

[14] J.J.E. Slotine, "Sliding controller design for nonlinear systems," *Int. J. Control*, Vol. 40, No. 2, 1984.

[15] J.J.E. Slotine, "The robust control of robot manipulators", *Int. J. of Rob. Res.*, Vol. 4,No.2, 1985.

[16] H. Buhler, "Reglage par mode de glissement", *Press Polytechnique Romande*, 1986.

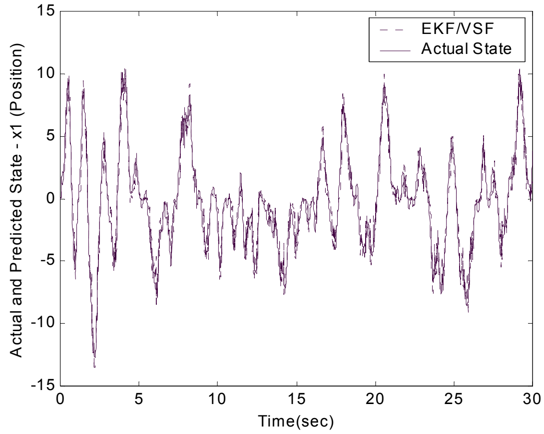

Figure 4a: Estimated and Actual States Associated with Modeling Error (Position)

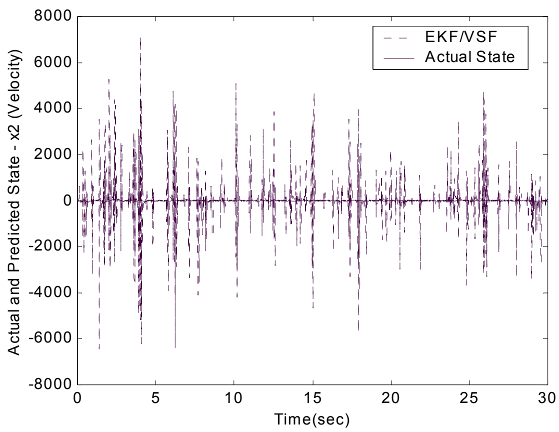

Figure 4b: Estimated and Actual States Associated with Modeling Error (Velocity)

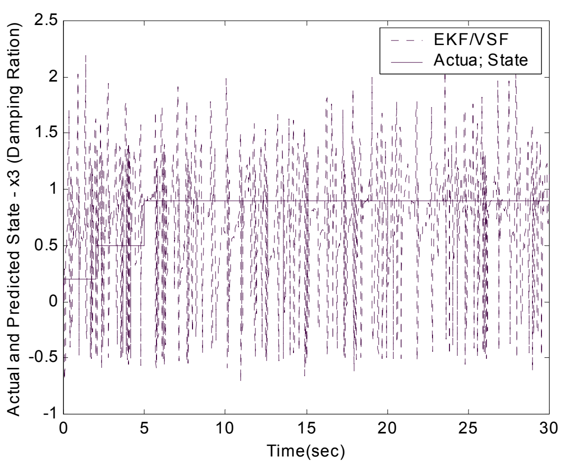

Figure 4c: Estimated and Actual States Associated with Modeling Error (Damping ratio)

Figure 4: Estimated and Actual States with 1000% Error in Natural Frequency

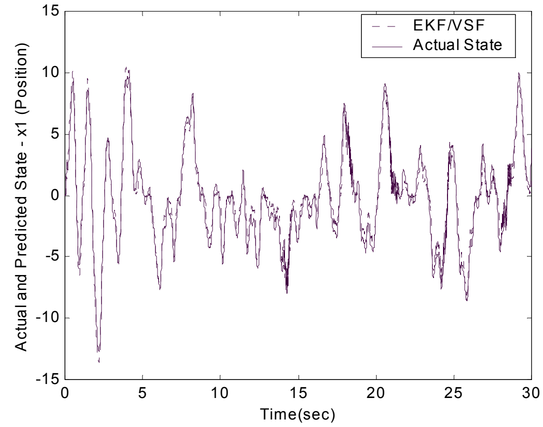

Figure 5a: Estimated and Actual States Associated with Modeling Error and increased boundary layer (x1 - Position)

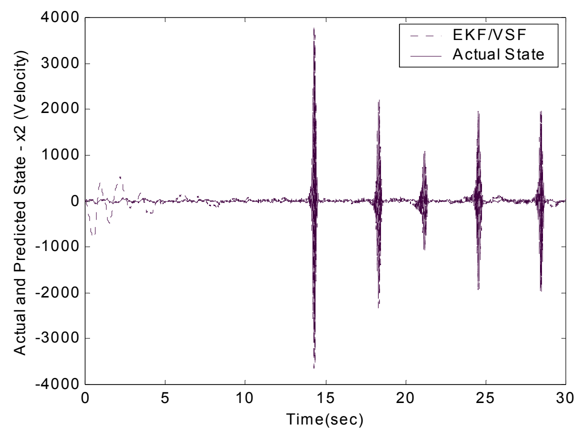

Figure 5b: Estimated and Actual States Associated with Modeling Error and increased boundary layer (x2- Velocity)

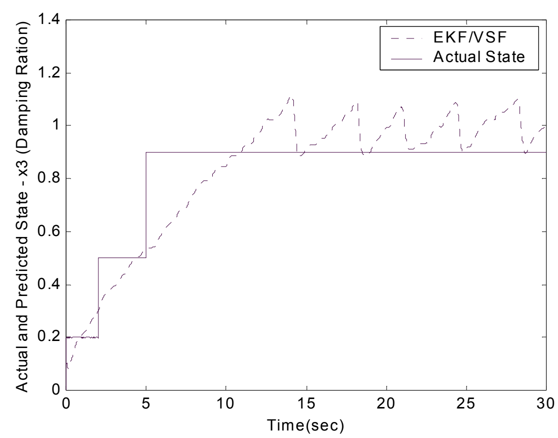

Figure 5b: Estimated and Actual States Associated with Modeling Error and increased boundary layer (x3 - Damping ratio)

Figure 5: Estimated and Actual States Associated with Modeling Error and increased boundary layer