Set-Valued State Observers for Nonlinear Systems *

Jeff S. Shamma Kuang-Yang Tu
Department of Aerospace Engineering and Engineering Mechanics
The University of Texas at Austin
Austin, TX 78712
[shamma,edwardtu]@jaws.ae.utexas.edu

Abstract

A set-valued observer (SVO) produces a set of possible states based on output measurements and *a priori* models of exogenous disturbances and noises. Previous work considered linear time-varying systems and unknown-but-bounded exogenous signals. In this case, the sets of possible state vectors take the form of polytopes whose centers are optimal state estimates. These polytopic sets can be computed by solving several small linear programs. A SVO can be constructed conceptually for nonlinear systems, however the set of possible state vectors no longer takes the form of polytopes which in turn inhibits their explicit computation. This paper considers an "extended SVO". As in the extended Kalman filter, the state equations are linearized about the state estimate, and a linear SVO is designed along the linearization trajectory. Under appropriate observability assumptions, it is shown that the extended SVO provides an exponentially convergent state estimate in the case of sufficiently small initial condition uncertainty, and provides a non-divergent state estimate in the case of sufficiently small exogenous signals.

1 Introduction

Constructions of observers for nonlinear systems often rely on some form of underlying linear dynamics. The extended Kalman filter (EKF) [5] linearizes the state trajectory about the current state estimate, resulting in approximate linear time-varying error dynamics. Output injection methods [7] employ a state-transformation to obtain exact linear time-invariant error dynamics. Similarly, reference [4] employs a state-transformation to obtain exact linear time-varying error dynamics where the "time-variations" actually are due to a measured endogenous signal. See references [10, 11, 15] and references therein for a further overview of nonlinear observers.

A guaranteed state estimator, alternatively called set-valued observer (SVO), assumes *a priori* bounds on exogenous disturbances and noises and constructs *sets* of possible states which are consistent with the *a priori* bounds and current measurements. The survey article [9] presents a historical account of such methods. See also the text [2] and collection [8]. References [13, 14] also consider the construction of SVO's for linear time-varying systems. The authors present a recursive method to construct these sets of possible states and show that the centers of these sets represent optimal state estimates in an ℓ^∞-induced norm sense.

In the linear case, sets of possible states generally take the form of (convex) polytopes. While it is possible to define conceptually a SVO for nonlinear systems [2], an explicit construction of the set of possible states is essentially prevented by the generality of (possibly disconnected) shapes.

In this paper, we mimic the EKF and construct an extended SVO for nonlinear systems. As in the EKF, the extended SVO linearizes the state equations about the current state estimate. Unlike the EKF, the extended SVO does not neglect the linearization errors. Rather, the linearization errors are considered as exogenous disturbances and are used to bound the set of possible states. An attractive feature of this approach is that the linear SVO optimally minimizes the effect of exogenous disturbances, and hence possibly the effect of linearization errors, on the estimation error. The main shortcoming of the extended SVO is the real-time computational burden of solving several small linear programs.

This paper is an abridged version of
Jeff S. Shamma and K.-Y. Tu, "Approximate set-valued observers for nonlinear systems", to appear, *IEEE Transactions on Automatic Control*, 1997.
which contains all proofs.

2 Notation

For $x \in \mathcal{R}^n$, let x_i denote the i^{th} component of x, and define $|x| = \max_i |x_i|$. The closed unit box in

*Supported by NSF grant #ECS–9258005, EPRI grant #8030–23, and Ford Motor Co.

$\mathcal{R}^n$ centered at x_o is denoted $\mathcal{B}(x_o)$. Define $\mathcal{K}_{\mathcal{R}^n}$ to be the set of all (nonempty) compact subsets of $\mathcal{R}^n$. Then $\mathcal{K}_{\mathcal{R}^n}$ is a metric space when equipped with the Hausdorff metric [12, p. 279].

Let $\mathcal{Z}^+$ denote the set of non-negative integers. For a sequence $\{x(0), x(1), x(2), \ldots\} \subset \mathcal{R}^n$, define

$$\|x\|_{\ell^\infty} = \sup_{k \in \mathcal{Z}^+} |x(k)|,$$

and

$$x_{[k_1,k_2]} = \begin{pmatrix} x(k_1) \\ \vdots \\ x(k_2) \end{pmatrix},$$

where $k_2 = \infty$ is possible.

For $f : \mathcal{R}^n \to \mathcal{R}^n$, Df denotes the Jacobian matrix. For notational simplicity, Df also denotes the Jacobian matrix of $f(\cdot, k)$ when $f : \mathcal{R}^n \times \mathcal{Z}^+ \to \mathcal{R}^n$. For $f : \mathcal{R}^n \times \mathcal{R}^p \times \mathcal{Z}^+ \to \mathcal{R}^q$, $D_1 f$ denotes the $q \times n$ Jacobian matrix with respect to the first variable.

3 Nonlinear SVO

Consider a nonlinear system of the form

$$\begin{aligned} x(k+1) &= f(x(k), d(k), k), \quad x(0) \in \mathcal{M}_0 \\ y(k) &= h(x(k), n(k), k), \end{aligned} \tag{1}$$

where $d(k) \in \mathcal{R}^{n_d}$ is unknown process noise, $y(k) \in \mathcal{R}^{n_y}$ is the measured output, and $n(k) \in \mathcal{R}^{n_n}$ is measurement noise. The effects of a known input can be incorporated as time-variations. Initial conditions are restricted to the set $M_0 \subset \mathcal{R}^{n_x}$.

We make the following *a priori* assumptions on (1).

Assumption 3.1 *There exist bounding functions $d_{\max}, n_{\max} : \mathcal{Z}^+ \to \mathcal{R}^+$ such that for all $k \in \mathcal{Z}^+$ the signals d and n in (1) satisfy*

$$|d(k)| \le d_{\max}(k), \quad |n(k)| \le n_{\max}(k).$$

We are interested in constructing the *set* of possible states, denoted $X(k)$, which are consistent with the *current* measurement trajectory and *a priori* Assumption 3.1.

First define the set $\tilde{X}(k)$ by

$$\begin{aligned} \tilde{X}(k) = \{&x \in \mathcal{R}^{n_x} : \\ &y(k) = h(x, n, k) \text{ for some } |n| \le n_{\max}(k)\}. \end{aligned}$$

In words, $\tilde{X}(k)$ represents the set of possible states at time k based on the single measurement $y(k)$ only. Similarly, define

$$\begin{aligned} X_{pre}(k+1) = \{&x : x = f(\tilde{x}, d, k), \\ &\text{for some } \tilde{x} \in X(k), |d| \le d_{\max}(k)\}. \end{aligned}$$

In words, $X_{pre}(k+1)$ denotes the *anticipated* set of possible states at time $k+1$ based on measurements up to time k. Note that X, $\tilde{X}$, and X_{pre} all depend on the current measurement trajectory. However, this dependence is not explicitly expressed for the sake of notational simplicity.

Algorithm 3.1 (SVO) *Let $y = \{y(0), y(1), y(2), \ldots\}$ be a measurement trajectory of the system (1) under Assumption 3.1. Suppose $x(0) \in X_o \in \mathcal{K}_{\mathcal{R}^{n_x}}$.*

Initialization

$$X_{pre}(0) = X_o,$$
$$X(0) = X_o \bigcap \tilde{X}(0).$$

Propagation

$$\begin{aligned} X(k) &= X_{pre}(k) \bigcap \tilde{X}(k) \\ &\{x : x = f(\tilde{x}, d, k), \text{ for some } \tilde{x} \in X(k-1), \\ &|d| \le d_{\max}(k-1)\} \bigcap \tilde{X}(k). \end{aligned}$$

■

Note that the SVO algorithm is causally dependent on the measurement trajectory.

Associated with the set $X(k)$ is the *central estimate* $\hat{x}_c(k)$ defined as follows. For each component, $x_i(k)$, with $i = 1, \ldots, n_x$, define

$$\overline{x}_i(k) = \max\{x_i(k) : x(k) \in X(k)\},$$

$$\underline{x}_i(k) = \min\{x_i(k) : x(k) \in X(k)\},$$

Then the central estimate is defined as

$$\hat{x}_c(k) = \frac{1}{2}\overline{x}(k) + \frac{1}{2}\underline{x}(k).$$

Optimality properties of central estimates are considered in [9, 13, 14].

4 Extended SVO

In the nonlinear case, the SVO algorithm must propagate general sets in $\mathcal{R}^{n_x}$. This essentially prevents any computational implementation of the algorithm. In this section, we mimic the EKF and construct an extended SVO for nonlinear systems.

We will consider the simplified nonlinear system

$$\begin{aligned} x(k+1) &= f(x(k), k) + d(k) \\ y(k) &= Cx(k) + n(k). \end{aligned} \tag{2}$$

Having the disturbances, d, enter linearly can always be satisfied at the cost of higher order dynamics by augmenting the system with a delay. The linear output assumption is made with some loss of generality. In some cases, the output can be part of the state vector after an appropriate transformation.

Assumption 4.1 *For any* $k \in \mathcal{Z}^+$, *the function* $f(\cdot, k) : \mathcal{R}^{n_x} \to \mathcal{R}^{n_x}$ *in (2) is continuously differentiable, and for all* $x_o, x \in \mathcal{R}^{n_x}$,

$$f(x, k) = f(x_o, k) + Df(x_o, k)(x - x_o) + R(x, x_o, k),$$

where

$$|R(x, x_o, k)| \leq \gamma |x - x_o|^2 .$$

Assumption 4.1, as stated, requires that the linearization residuals are uniformly quadratically bounded. In fact, the forthcoming extended SVO only requires that these residuals are uniformly bounded over all state/estimate trajectories. This essentially reflects that the system evolves over a (not necessarily small) compact set.

The forthcoming extended SVO will produce sets of states, $X_e(k)$, which bound the actual sets of possible states, i.e.,

$$X(k) \subset X_e(k).$$

Let $\hat{x}_e(k)$ denote the central estimate based on $X_e(k)$. Linearizing (2) about $\hat{x}_e(k)$ leads to

$$\begin{aligned} x(k+1) &= f(\hat{x}_e(k), k) + Df(\hat{x}_e(k), k)(x(k) - \hat{x}_e(k)) \\ &+ R(x(k), \hat{x}_e(k), k) + d(k), \end{aligned} \tag{3}$$

Based on this linearization, the bounding sets $X_e(k)$ can be computed as follows. Define $\tilde{X}(k)$ as before. It will be convenient to express the sets $X_e(k)$ as deviations from their centers. Towards this end, define

$$X_e'(k) = \{v : \hat{x}_e(k) + v \in X_e(k)\}$$

and

$$\rho(k) = \sup_{v \in X_e'(k)} |v| .$$

Algorithm 4.1 (Extended SVO) *Let* $y = \{y(0), y(1), y(2), \ldots\}$ *be a measurement trajectory of the system (2) under Assumptions 3.1 and 4.1. Suppose* $x(0) \in X_o \in \mathcal{K}_{\mathcal{R}^{n_x}}$.

Initialization

$$X_e(0) = \tilde{X}(0) \bigcap X_o.$$

Propagation

$$\begin{aligned} X_e(k) = \tilde{X}(k) \bigcap \{&x : x = f(\hat{x}_e(k-1), k-1) \\ &+ Df(\hat{x}_e(k-1), k-1)v + r + d, \\ &\text{for some } v \in X_e'(k-1), |d| \leq d_{\max}(k-1), \\ &|r| \leq \gamma\rho^2(k-1)\} . \end{aligned}$$

■

We see that the extended SVO bounds the sets $X(k)$ by considering the linearization residuals, $R(x(k), \hat{x}_e(k), k)$ as exogenous disturbances. This is unlike the traditional EKF which simply ignores the linearization residuals (although it is possible to include "expected" residuals in a "second order" EKF). If available, tighter bounds may be used in place of $\gamma\rho^2$. In fact, the residual bounds can be a function of the current set-valued state estimate (at the cost of increased computational burden). As with the linear SVO, the sets $X_e(k)$ are polytopes and can be computed by solving several linear programs.

It can be shown that the extended SVO estimate, $\hat{x}_e$, has the following convergence and non-divergence properties under appropriate observability assumptions:

- In case $d, n = 0$,

  $$\hat{x}_e(k) \to x(k)$$

  for sufficiently small initial condition uncertainty.

- In case $d, n \neq 0$,

  $$\|x - \hat{x}_e\|_{\ell^\infty} \leq \gamma \|(d, n)\|_{\ell^\infty} ,$$

  for sufficiently small initial condition uncertainty and sufficiently small $\|(d, n)\|_{\ell^\infty}$.

Details may be found in the full journal version of this paper.

5 Special Cases

Two special classes of systems previously considered for nonlinear observers are nonlinear systems whose dynamics after state transformations take the form

$$x(k+1) = \phi(y(k)) + Ax(k),$$

$$y(k) = Cx(k),$$

or more generally

$$x(k+1) = \phi(y(k)) + A(y(k))x(k),$$

$$y(k) = Cx(k).$$

The first system, considered in [3, 7], represents a nonlinear system which is state equivalent to a linear system with output injection. The second system, considered in [4], represents a special structure which resembles a linear time-varying system whose "time-variations" are actually due to the measured output.

For either structure, it is possible to generate error dynamics which are either linear time-invariant or resemble linear time-varying dynamics via the observer

$$\hat{x}(k+1) = \phi(y(k)) + A(y(k))x(k) + H(k)(y(k) - C\hat{x}(k)).$$

Let $e(k) = x(k) - \hat{x}(k)$. Then

$$e(k+1) = (A(y(k)) - H(k)C)e(k).$$

The underlying linear error dynamics then greatly simplify observer design.

It turns out that for these classes of systems, the extended SVO actually generates the exact set of possible states. In particular, consider

$$\begin{aligned} y(k+1) &= \phi_y(x(k), k) + d_y(k) \\ z(k+1) &= \phi_z(y(k), k) + A(y(k), k)z(k) + d_z(k) \\ y(k) &= (\, I_{n_y} \quad 0\,)\, x(k), \end{aligned} \tag{4}$$

where the state vector is partitioned $x(k) = \begin{pmatrix} y(k) \\ z(k) \end{pmatrix}$, and there is no measurement noise. Note that equation (4) includes both previous structures, but *after* appropriate state transformations have been made.

Proposition 5.1 *Consider Algorithm 4.1 applied to the nonlinear system (4) with $\gamma = 0$. The sets $X_e(k)$ exactly represent the nonlinear SVO of Algorithm 3.1.*

Example 5.1 Consider the dynamics of a freely rotating rigid body

$$\begin{aligned} \dot{\omega}_1 &= \alpha\omega_2\omega_3 \\ \dot{\omega}_2 &= \beta\omega_1\omega_3 \\ \dot{\omega}_3 &= \gamma\omega_1\omega_2 \\ y &= \omega_1. \end{aligned}$$

These dynamics were considered in [6] as well as [4]. Both references derived state transformations which linearize the error dynamics, however the transformation considered in [6] relies on *a priori* knowledge of maximal and minimal values of y.

A discretized version of these dynamics yields

$$\begin{aligned} \omega_1(k+1) &= \omega_1(k) + h\alpha\omega_2(k)\omega_3(k) + d_1(k) \\ \omega_2(k+1) &= \omega_2(k) + h\alpha\omega_1(k)\omega_3(k) + d_2(k) \\ \omega_3(k+1) &= \omega_3(k) + h\alpha\omega_1(k)\omega_2(k) + d_3(k) \end{aligned}$$

which are already in the form of (4) without any state transformations. The quantities d_i can be used to reflect discretization errors.

■

6 Simulation Example

We will consider state estimation for a discretized Van der Pol equation

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -9x_1 + \mu(1 - x_1^2)x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} d_c$$

$$y = x_1 + n_c,$$

which was also considered in [1]. Performing a discretization of step size h leads to

$$\begin{aligned} \begin{pmatrix} x_1(k+1) \\ x_2(k+1) \end{pmatrix} &= f(x) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} d(k) \\ &= \begin{pmatrix} x_1(k) + hx_2(k) \\ x_2(k) + h\Big(-9x_1(k) + \mu(1 - x_1^2(k))x_2(k)\Big) \end{pmatrix} \\ &\quad + \begin{pmatrix} 0 \\ d(k) \end{pmatrix}, \end{aligned}$$

$$y(k) = x_1(k) + n(k),$$

where d and n denote discrete-time noises. Note that these equations are in the form of (4) in case $n = 0$.

Let $e = x - x_o$. Linearizing the above right-hand-side about x_o leads to

$$\begin{aligned} f(x) &= f(x_o) + \\ &\begin{pmatrix} 1 & h \\ -9h - 2h\mu x_{o,1}x_{o,2} & 1 + \mu h - \mu h x_{o,1}^2 \end{pmatrix} e \\ &+ R(x, x_o), \end{aligned}$$

where

$$R(x, x_o) = -\mu h \left(\begin{matrix} 0 \\ e^T \begin{pmatrix} x_{o,2} & x_{o,1} \\ x_{o,1} & 0 \end{pmatrix} e + e_1^2 e_2 \end{matrix} \right).$$

The simulated extended SVO followed Algorithm 4.1 except that it exploited maximal values of $x(k) - \hat{x}_e(k)$ to bound on $R(x(k), x_e(k))$. An EKF as described in [15] also was included in the simulations.

The following simulation parameters were used for Simulations 1–4:

- System parameters: $h = 0.02$, $\mu = 2$, $x(0) = \begin{pmatrix} 5 \\ 0 \end{pmatrix}$.
- Noise bounds: $d_{\max} = n_{\max} = 1$.
- SVO initial condition: $X(0) = \begin{pmatrix} 4 \le x_1 \le 8 \\ 0 \le x_2 \le 2 \end{pmatrix}$.
- Initial *a priori* covariance matrix: $\overline{P}(0) = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}$.
- Initial *a priori* state estimate: $\overline{x}(0) = \begin{pmatrix} 6 \\ 2 \end{pmatrix}$.

The particular simulations are described as follows:

1. Constant Noise: $E[d^2(k)] = E[n^2(k)] = 1/3$; $d(k) = n(k) = 0$.
2. Constant Noise: $E[d^2(k)] = E[n^2(k)] = 1/3$; $d(k) = -1$; $n(k) = 1$.
3. Uniform Random Noise: $E[d^2(k)] = E[n^2(k)] = 1/3$; $d(k), n(k) \in [-1, 1]$.

Simulation	x_1		x_2	
	SVO	EKF	SVO	EKF
#1	0	4.40×10^{-3}	1.74×10^{-5}	4.18×10^{-4}
#2	1.00	3.09	4.41	7.96
#3	5.87×10^{-2}	2.25×10^{-2}	3.75	1.51
#4	5.16×10^{-3}	1.21×10^{-1}	2.13	6.10

Table 1: Mean Square Estimation Error

4. Bang-Bang Random Noise: $E[d^2(k)] = E[n^2(k)] = 1$; $d(k), n(k) \in \{-1, +1\}$.

Both the EKF and extended SVO generally follow the state trajectory, however the SVO state bounds are very conservative. Figure 1 illustrates these bounds for Simulation #1. Figure 2 of Simulation #2 shows that the true state initially follows the SVO lower bound. Figures 3–4 show the various time responses in Simulation #2.

Table 1 summarizes the mean square estimation errors starting *after* time $k = 0$. The extended SVO seems to outperform the EKF whenever the simulation significantly departs from the stochastic structure for which the linear Kalman filter is optimal. This includes Simulation #4, in which the extended Kalman filter was provided with the correct variances.

The following simulation parameters exhibiting large initial condition uncertainty led to divergence of the EKF while the extended SVO locked on to the true state within 2 time steps:

- Noise Bounds: $d_{\max} = 0.0001$, $n_{\max} = 1$.
- SVO initial condition: $X(0) = \begin{pmatrix} 0 \leq x_1 \leq 40 \\ 0 \leq x_2 \leq 20 \end{pmatrix}$.
- Initial *a priori* covariance matrix: $\overline{P}(0) = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}$.
- Initial *a priori* state estimate: $\overline{x}(0) = \begin{pmatrix} 20 \\ 10 \end{pmatrix}$.
- Uniform random noise: $d(k) \in [-d_{\max}, d_{\max}]$; $n(k) \in [-n_{\max}, n_{\max}]$; $E[d^2(k)] = d_{\max}/3$; $E[n^2(k)] = n_{\max}/3$.

Despite these results, it is unclear whether either observer generally exhibits superior convergence and performance. The extended SVO does have a significantly larger computational burden.

7 Concluding Remarks

In this paper, we have considered an extended SVO for nonlinear systems and derived guaranteed convergence and non-divergence properties. The main shortcoming of the extended SVO is the significant computational burden of solving several linear programs. The number of variables for these linear programs approximately equals the number of state variables and exogenous disturbances. The number of constraints depends on the complexity of the resulting sets of possible states. Theoretically, this number could increase with the number of measurements. In the simulation example, however, the number of constraints was rarely larger than 20 and usually less than 10. Given this computational burden, real-time application seems unlikely for systems with fast dynamics.

References

[1] M.S. Ahmed. An innovation representation for nonlinear systems with application to parameter and state estimation. *Automatica*, **30**(12):1967–1974, 1994.

[2] F.L. Chernousko. *State Estimation for Dynamic Systems.* CRC Press, Inc., Boca Raton, FL, 1994.

[3] S.-T. Chung and J.W. Grizzle. Sampled-data observer error linearization. *Automatica*, **26**(6):997–1008, 1990.

[4] M. Jankovic. A new observer for a class of nonlinear systems. *Journal of Mathematics, Estimation, and Control*, **3**(2):225–246, 1993.

[5] A.H. Jazwinski. *Stochastic Processes and Filtering Theory.* Academic Press, New York, 1970.

[6] W. Kang and A.J. Krener. Observation of a rigid body from measurements of a principal axis. *Journal of Mathematical Systems, Estimation, and Control*, **1**:197–208, 1991.

[7] A.J. Krener and A. Isidori. Linearization by output injection and nonlinear observers. *Systems & Control Letters*, **3**:47–52, 1983.

[8] A.B. Kurzhanski and V.M Veliov, editors. *Modeling Techniques for Uncertain Systems.* Birkhäuser, Boston, 1993.

[9] M. Milanese and V. Vicino. Optimal estimation theory for dynamic systems with set membership uncertainty: An overview. *Automatica*, **27**:997–1009, 1991.

[10] E.A. Misawa and J.K. Hedrick. Nonlinear observers—A state-of-the-art survey. *ASME Journal of Dynamic Systems, Measurement, and Control*, **111**:344–352, 1989.

[11] P.E. Moraal and J.W. Grizzle. Observer design for nonlinear systems with discrete-time measurements. *IEEE Transactions on Automatic Control*, **AC–40**(3):395–404, 1995.

[12] J.R. Munkres. *Topology: A First Course.* Prentice-Hall, Inc., Englewood Cliffs, NJ, 1975.

[13] J.S. Shamma and K.-Y. Tu. Optimality of set-valued observers for linear systems. In *Proceedings of the 34th IEEE Conference on Decision and Control*, New Orleans, LA, December 1995.

[14] J.S. Shamma and K.-Y. Tu. Set-valued observers and optimal disturbance rejection. submitted to *IEEE Transactions on Automatic Control*, 1995.

[15] Y. Song and J.W. Grizzle. The extended Kalman filter as a local asymptotic observer for discrete-time nonlinear systems. *Journal of Mathematics, Estimation, and Control*, **5**(1):59–78, 1995.

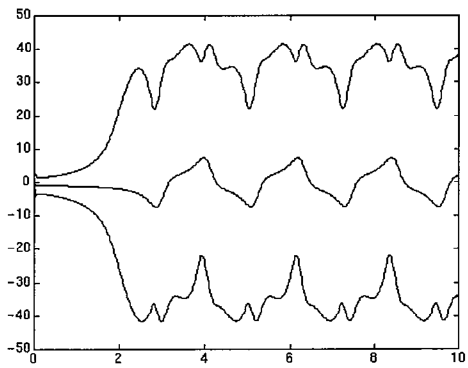

Figure 1: State trajectory $x_2(k)$ and extended SVO error bounds (Simulation #1)

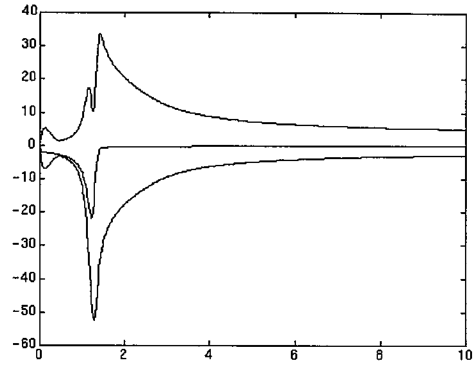

Figure 2: State trajectory $x_2(k)$ and extended SVO error bounds (Simulation #2)

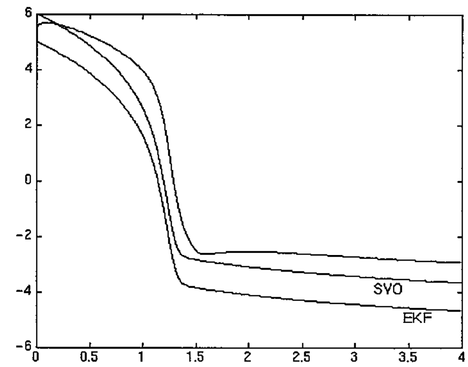

Figure 3: State trajectory $x_1(k)$ with extended SVO and EKF estimates (Simulation #2)

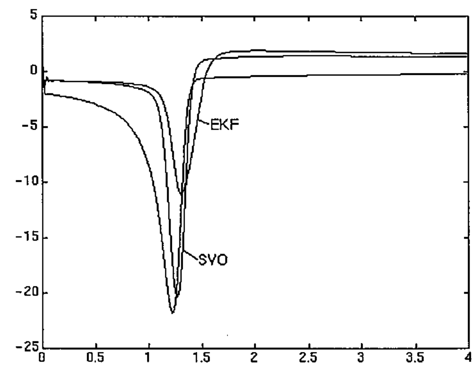

Figure 4: State trajectory $x_2(k)$ with extended SVO and EKF estimates (Simulation #2)