Proceedings of 25th Conference
on Decision and Control
Athens, Greece • December 1986

NONLINEAR STATE ESTIMATION USING SLIDING OBSERVERS

J.-J. E. Slotine[1], J. K. Hedrick, and E. A. Misawa [2]

Nonlinear Systems Laboratory
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

Sliding controllers have recently been shown to feature excellent robustness and performance properties for specific classes of nonlinear tracking problems. This paper examines the potential use of sliding surfaces for observer design. A particular observer structure including switching terms is shown to have promising properties in the presence of modelling errors and sensor noise.

1. INTRODUCTION

The notion of a sliding surface (Filippov, 1960) has been investigated mostly in the Soviet literature (see Utkin, 1977 for a review), where it has been used to stabilize a class of non-linear systems. Although it theoretically features excellent robustness properties in the face of *parametric* uncertainty, classical sliding mode control presents several important drawbacks that severely limit its practical applicability. In particular, it involves large control authority and control *chattering*. Chattering is in general highly undesirable in practice (with exceptions, such as the control of electric motors using pulse width modulation), since it implies extremely high control activity, and further may excite high-frequency dynamics neglected in the course of modeling, such as resonant structural modes, neglected actuator time-delays, or sampling effects. These problems can be remedied by replacing the chattering control by a smooth control interpolation in a boundary layer neighboring a time-varying sliding surface (Slotine and Sastry, 1983) and monitoring the boundary layer width so as not to excite the high-frequency unmodeled dynamics (Slotine, 1984).

In this paper, we consider the dual problem of designing state *observers* using sliding surfaces. We show that, as can be expected, sliding observers potentially offer advantages similar to those of sliding controllers, in particular inherent robustness to parametric uncertainty and easy application to important classes of nonlinear systems. Further , contrary to the case of controller design, chattering issues in sliding observer design are only linked to numerical implementation rather than 'hard' mechanical limitations.

Basic concepts on implicit dynamics using sliding surfaces are introduced in Section 2. Section 3 applies the development to the design of sliding observers for nonlinear systems in 'companion form', i.e. of the form

$$x^{(n)} = f$$

where f is a nonlinear, uncertain function of the system state:

$$\mathbf{x} = [x, \dot{x}, ..., x^{(n-1)}]^T$$

Section 4 extends the methodology to general nonlinear systems, and discusses observability requirements and their relationship to sliding observers. Section 5 shows some possible applications and concluding remarks are offered in Section 6.

2. BASIC CONCEPTS

2.1 Sliding Surfaces

Let us first briefly summarize the basic idea of a sliding mode, linked to the potential advantages of using discontinuous (switching) control laws. Consider the dynamic system:

$$x^{(n)}(t) = f(\mathbf{x};t) + b(\mathbf{x};t)u(t) + d(t) \tag{1}$$

where u(t) is scalar control input, x is the scalar output of interest, and $\mathbf{x} = [x, \dot{x}, ..., x^{(n-1)}]^T$ is the state. In equation (1) the function f(x;t) (in general nonlinear) is not exactly known, but the *extent of the imprecision* $|\Delta f|$ *on f(x;t) is upper bounded by a known continuous function of x and t*; similarly control gain b(x;t) is not exactly known, but is of known sign, and is bounded by known, continuous functions of x and t. Both f(x;t) and b(x;t) are assumed to be continuous in $\mathbf{x}$. The disturbance *d(t)* is unknown but bounded in absolute value by a known continuous function of time. The control problem is to get the state $\mathbf{x}$ to *track a specific state* $\mathbf{x}_d = [x_d, \dot{x}_d, ..., x_d^{(n-1)}]^T$ *in the presence of model imprecision* on f(x;t) and b(x;t), *and of disturbances d(t)*. To guarantee that this is achievable using a finite control u, we must assume :

$$\tilde{\mathbf{x}}|_{t=0} = 0 \tag{2}$$

where $\tilde{\mathbf{x}} := \mathbf{x} - \mathbf{x}_d = [\tilde{x}, \dot{\tilde{x}}, ..., \tilde{x}^{(n-1)}]^T$ is the tracking error vector; the relaxation of this assumption shall be further discussed later. We define a *time-varying sliding surface s(t)* in the state-space R^n as $s(\mathbf{x};t) = 0$ with

$$s(\mathbf{x};t) := \left(\frac{d}{dt} + \lambda\right)^{n-1} \tilde{x}, \quad \lambda > 0 \tag{3}$$

where λ is a positive constant. Given initial condition (2), *the problem of tracking* $\mathbf{x} \equiv \mathbf{x}_d$ is equivalent to that of remaining on the surface s(t) for all $t > 0-$ indeed $s \equiv 0$ represents

[1] supported in part by NSF grant 86-11640-ECS

[2] On leave from Escola Politécnica da Universidade de São Paulo (Dept. of Mechanical Eng.) and supported by Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, grant 3495/82-5.

CH2344-0/86/0000-0332 $1.00 © 1986 IEEE

Authorized licensed use limited to: UNIVERSIDADE DE SAO PAULO. Downloaded on November 15, 2008 at 14:13 from IEEE Xplore. Restrictions apply.

a linear differential equation whose unique solution is $\tilde{x} \equiv 0$ given initial conditions (2). Now a sufficient condition for such positive invariance of s(t) is to choose the control law u of (1) such that outside of s(t)

$$\frac{1}{2}\frac{d}{dt} s^2(\mathbf{x};t) \leq -\eta|s| \tag{4}$$

where η is a positive constant. Inequality (4) constrains trajectories to point towards the surface s(t) (Figure 1), and is referred to as the *sliding condition.*

The idea behind equations (3, 4) is to pick-up a well-behaved function of the tracking error, s, according to (3), and then select the feedback control law u in (1) such that s^2 remains a Liapunov function of the closed-loop system *despite the presence of model imprecision and of disturbances.* Further, satisfying (3) guarantees that if condition (2) is not exactly verified, i.e. if $\mathbf{x}|_{t=0}$ is actually off $\mathbf{x}_d|_{t=0}$, the surface S(t) will none the less be reached in a finite time inferior to $|s(\mathbf{x}(0);0)|/\eta$, while definition (3) then guarantees that $\mathbf{x} \to 0$ as $t \to \infty$. Control laws that satisfy (4), however, have to be *discontinuous* across the sliding surface, thus leading in practice to control chattering.

The obvious problem in similarly exploiting sliding behavior in the design of *observers*, rather than controllers, is precisely that the full state is not available for measurement, and thus that a sliding surface definition analog to (3) is not adequate. Some intuition can be developed for addressing this difficulty by considering simple second-order dynamics.

2.2 Shearing Effect and Sliding Patches

Let us consider the generation of sliding behavior in a second-order system through input switching according to the value of a *single* component of the state, rather than a linear combination of both components, as in (3). The system

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -k_2 sgn(x_1)$$

where k_2 is a positive constant and *sgn* is the sign function, clearly exhibits no sliding behavior (Figure 2). Instead, let us consider the system

$$\dot{x}_1 = x_2 - k_1 sgn(x_1)$$
$$\dot{x}_2 = -k_2 sgn(x_1)$$

where k_1 and k_2 are positive constants. The corresponding phase-plane trajectories are illustrated in Figure 3, which can be constructed from Figure 2 by shifting the trajectories on the right half-plane upwards, by the quantity k_1, and similarly shifting the left half-plane trajectories by $-k_1$. This shearing effect generates sliding behavior in the region

$$|x_2| \leq k_1 \tag{5}$$

which we shall refer to as the *sliding patch.*

Let us detail the analysis. The condition

$$\frac{d}{dt}(x_1)^2 < 0$$

is satisfied if condition (5) holds, which defines the sliding patch. The dynamics on the sliding patch itself can be derived from Filippov's solution concept (Filippov, 1960), which formalizes engineering intuition: the dynamics on the patch can only be a convex combination (i.e., an *average* of the dynamics on each side of the discontinuity surface

$$\dot{x}_1 = \gamma(x_2 + k_1) + (1-\gamma)(x_2 - k_1)$$
$$\dot{x}_2 = \gamma\, k_2 + (1-\gamma)(-k_2)$$

The value of γ , and therefore the resulting dynamics, are then formally determined by the invariance of the patch itself:

$$\dot{x}_1 = 0 \quad => \quad \dot{x}_2 = -(k_2/k_1)\, x_2$$

Thus, x_2 exponentially decreases to 0 after reaching the sliding patch, with a time-constant k_1/k_2 . Further, one can easily show that all trajectories starting on the x_2 axis reach the patch in a time smaller than $|x_2(t=0)|/(k_1 k_2)$ Actually, sliding can be guaranteed from time t=0 by making $k_1$andk_2

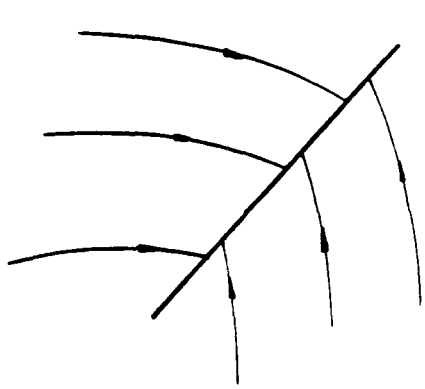

Figure 1: *The sliding condition*

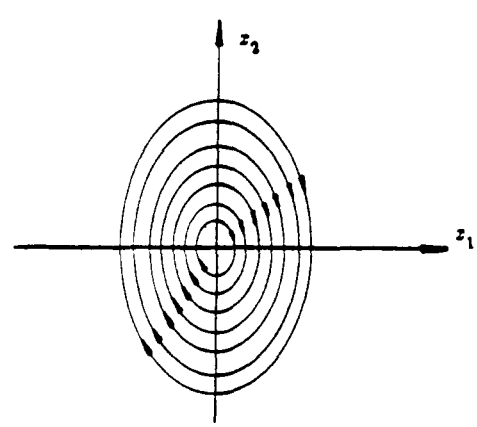

Figure 2: *Second order system with single input switching*

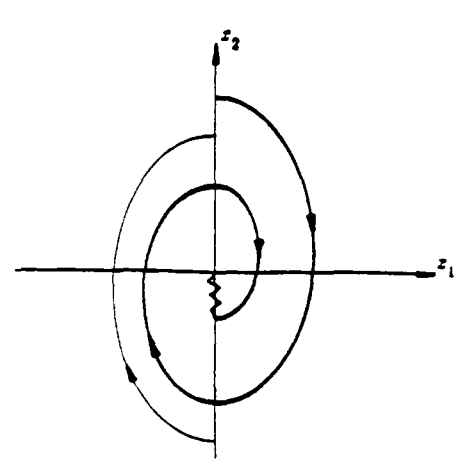

Figure 3: *Shearing effect*

Authorized licensed use limited to: UNIVERSIDADE DE SAO PAULO. Downloaded on November 15, 2008 at 14:13 from IEEE Xplore. Restrictions apply.

time-varying, with

$$k_2/k_1 \geq a$$
$$k_1 > |x_2(t=0)|e^{-at}$$

where $a = a(t)$ is any positive function of time .

2.3 System Damping

Consider now the system

$$\dot{x}_1 = -\alpha_1 x_1 + x_2 - k_1 sgn(x_1)$$
$$\dot{x}_2 = -\alpha_2 x_1 - k_2 sgn(x_1)$$

Repeating the previous analysis, the sliding condition is verified in the extended region

$$x_2 \leq k_1 + \alpha_1 x_1 \quad \text{if } x_1 > 0$$
$$x_2 \geq -k_1 + \alpha_1 x_1 \quad \text{if } x_1 < 0$$

as illustrated in Figure 4. Thus, the addition of the damping term in α_1 increases the region of direct attraction. Further, the value of α_2 only affects the capture phase but not the dynamics on the patch itself, which remains unchanged:

$$\dot{x}_2 = -(k_2/k_1)\, x_2$$

3. IMPLICIT REDUCED-ORDER OBSERVERS FOR NONLINEAR SYSTEMS IN COMPANION FORM

3.1 Systems with a Single Measurement

Let us now consider the system

$$\ddot{x}_1 = f$$

where f is a nonlinear, uncertain function of the state $\mathbf{x} = [x_1, x_2 = \dot{x}_1]^T$, and let us exploit the preceding development to design an observer for this system, based on the measurement of x_1 alone. From the previous dicussion, we use an observer structure of the form

$$\begin{aligned} \dot{\hat{x}}_1 &= -\alpha_1 \tilde{x}_1 + \hat{x}_2 - k_1 sgn(\tilde{x}_1) \\ \dot{\hat{x}}_2 &= -\alpha_2 \tilde{x}_1 + \hat{f} - k_2 sgn(\tilde{x}_1) \end{aligned} \tag{6}$$

where $\tilde{x}_1 = \hat{x}_1 - x_1$, $\hat{f}$ is the estimated value of f , and the constants α_i are chosen as in a Luenberger observer (which would correspond to $k_1 = 0, k_2 = 0$) so as to place the poles of the linearized system at desired locations $-\varsigma_j$. The resulting error dynamics can be written:

$$\dot{\tilde{x}}_1 = -\alpha_1 \tilde{x}_1 + \tilde{x}_2 - k_1 sgn(\tilde{x}_1)$$
$$\dot{\tilde{x}}_2 = -\alpha_2 \tilde{x}_1 + \Delta f - k_2 sgn(\tilde{x}_1)$$

The value of $\Delta f = \hat{f} - f$ depends both on the modelling effort and of the computational complexity allowable in the observer itself. In this paper, we assume that dynamic uncertainty Δf is explicitely bounded. Known nonlinear terms may also, for simplicity, be treated as bounded error (using known bounds on the actual system state) and included in Δf. The effect of Δf is compensated by exploiting this knowledge of its (generally time-varying) bound, as we shall later illustrate.

The methodology can be directly extended to n^{th}-order systems in companion form:

$$x_1^{(n)} = f$$

where x_1 is the single measurement available. The observer

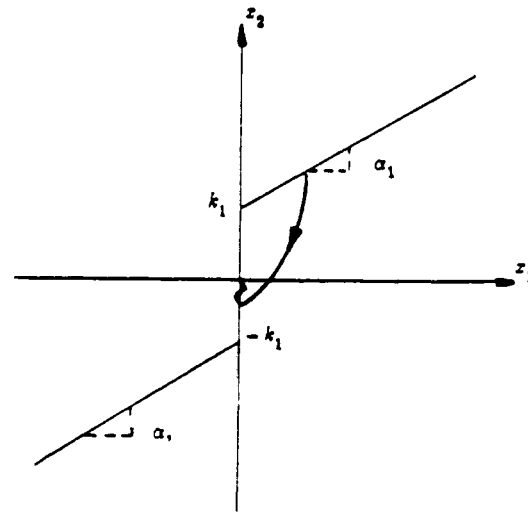

Figure 4: *Effect of damping on reachability*

structure is then of the form

$$\begin{aligned} \dot{\hat{x}}_1 &= -\alpha_1 \tilde{x}_1 + \hat{x}_2 - k_1 sgn(\tilde{x}_1) \\ \dot{\hat{x}}_2 &= -\alpha_2 \tilde{x}_1 + \hat{x}_3 - k_2 sgn(\tilde{x}_1) \\ &\cdots \\ \dot{\hat{x}}_n &= -\alpha_n \tilde{x}_1 + \hat{f} - k_n sgn(\tilde{x}_1) \end{aligned}$$

The $n-1$ poles associated with the implicit dynamics on the patch are defined by

$$det(p\mathbf{I}_{n-1} - \begin{bmatrix} -k_2/k_1 & 1 & 0 & \dots & 0 \\ -k_3/k_1 & 0 & 1 & \dots & 0 \\ & & & \ddots & \\ & & & & \dots 1 \\ -k_n/k_1 & 0 & 0 & \dots & 0 \end{bmatrix}) = 0 \tag{7}$$

Thus, the poles on the patch can be placed arbitrarily by proper selection of the ratios $k_i/k_1, [i = 2, ..., n]$. A possible choice is to define k_1 as the desired precision in $\tilde{x}_2$, let

$$k_n \geq |\Delta f|$$

and in a constant ratio with k_1, and finally define the remaining poles $k_i, [i = 2, ..., n-1]$ so that the implicit dynamics associated with the patch be critically damped, i.e., have all poles real and equal to a positive constant λ. One can then easily show that trajectories starting on the $\tilde{x}_1$ axis and in the sliding patch remain in the patch, and verify

$$|\tilde{x}_2^{(i)}| \leq (2\lambda)^i k_1 \qquad i = 0, \dots, n-2$$

from which the precision on $\tilde{\mathbf{x}}$ can be derived.

Remark that there is no reason for the implicit bandwidth on the patch to be identical to the bandwidth of the Luenberger part . In particular, reducing k_1 *increases* the bandwidth on the patch, yet may potentially *reduce* sensitivity to measurement noise by reducing the amplitude of the discontinuity in $\tilde{x}_1$. It is this unusual nonlinear effect that we must try and exploit to make sliding observers superior to Luenberger observers or extended Kalman filters under certain noise conditions, as we now discuss.

Authorized licensed use limited to: UNIVERSIDADE DE SAO PAULO. Downloaded on November 15, 2008 at 14:13 from IEEE Xplore. Restrictions apply.

3.2 Effects of Measurement Noise

Consider again a second order system with a single measurement, now corrupted by noise $v = v(t)$

$$\begin{aligned} \dot{\tilde{x}}_1 &= -\alpha_1(\tilde{x}_1 + v) + \tilde{x}_2 - k_1 sgn(\tilde{x}_1 + v) \\ \dot{\tilde{x}}_2 &= -\alpha_2(\tilde{x}_1 + v) + \Delta f - k_2 sgn(\tilde{x}_1 + v) \end{aligned} \tag{8}$$

Although the presence of the terms in $sgn(\tilde{x}_1 + v)$ makes an exact stochastic analysis fairly involved, useful insight can be obtained by using appropriate simplifying approximations.

Assume, first, that v is a deterministic C^1 signal of bounded spectrum:

$$0 \le \omega < \omega_- \quad \text{or} \quad \omega > \omega_+ \qquad => \qquad F_v(\omega) = 0$$

where F_v is the Fourier transform of v. Sliding behavior, if any, can then only occur on the surface

$$\tilde{x}_1 + v = 0$$

Repeating the analysis of Section 2.2, the sliding region is then defined by

$$|\tilde{x}_2 + \dot{v}| \le k_1 \tag{9}$$

and the equivalent dynamics are given by

$$\begin{aligned} \tilde{x}_1 &= -v \\ \dot{\tilde{x}}_2 + (k_2/k_1)\,\tilde{x}_2 &= -(k_2/k_1)\,\dot{v} + \Delta f \end{aligned}$$

Two limiting cases deserve particular attention:

a) $\omega_+ << (k_2/k_1)$. We then have, if $\Delta f = 0$

$$\tilde{x}_2 \approx -\dot{v} << (k_2/k_1)\, v_{max}$$

In particular, the estimate of x_2 is exact if the measurement error in x_1 is constant.

b) $\omega_- >> (k_2/k_1)$. We then have, if $\Delta f = 0$

$$\tilde{x}_2 \approx 0$$

However, the bound on k_2/k_1 also implies that the observer's robustness to model uncertainty is directly limited by the value of ω_-. The corresponding precision in $\tilde{x}_2$ is then

$$|\tilde{x}_2| \le r\,F\,/\,\omega_-$$

where $r \approx 3$ is the desired ratio between ω_- and k_2/k_1, and F is the available (in general time-varying) bound on $|\Delta f|$. It is obtained by choosing k_1 and k_2 according to

$$\begin{aligned} k_1 &\ge |\dot{v}| + rF/\omega_- \\ k_2 &= k_1\omega_-/r \end{aligned}$$

so as to satisfy (9) while maintaining k_2 larger or equal to F.

The above discussion implies that, as could be expected, the system cannot remain in a pure sliding mode in the presence of arbitrary measurement noise, since this could require using an infinitely large $\mathbf{k}$. Instead, assuming that the measurement noise is bounded by some constant v_0 , the system will remain in a vicinity of the x_2 axis of width v_0. The major potential advantage of the proposed sliding observer, over e.g. an extended Kalman filter, is that the sliding observer can still be made considerably more robust to parametric uncertainty. This can be easily understood by considering the 'average' error of the observer, $\tilde{\mathbf{x}}_a$, whose dynamics can be approximated as

$$\begin{aligned} \dot{\tilde{x}}_{1a} &= -\alpha_1\tilde{x}_{1a} + \tilde{x}_{2a} - k_1 \text{ Average}[\text{sgn}(\tilde{x}_{1a} + v)] \\ \dot{\tilde{x}}_{2a} &= -\alpha_2\tilde{x}_{1a} - k_2 \text{ Average}[\text{sgn}(\tilde{x}_{1a} + v)] + \Delta f_a \end{aligned}$$

where Average $[\tilde{x}_{1a} + v]$ is computed over 'short' time periods during which $\tilde{x}_a$ is treated as a constant, and Δf_a is defined as Average $[\Delta f]$. If we assume for simplicity that $v(t)$ is white noise, then

$$\begin{aligned} \text{Average}[\text{sgn}(\tilde{x}_1 + v)] &= \text{Expectation}[\text{sgn}(\tilde{x}_1 + v)] \\ &= \int_{-\infty}^{+\infty} \text{sgn}(\tilde{x}_1 + v)p(v)dv = 2\int_0^{\tilde{x}_1} p(v)dv \end{aligned}$$

where the last equality assumes that the probability density $p(v)$ is symmetric. Thus, the average value is an odd continuous function of $\tilde{x}_1$. For instance, if v is uniformly distributed on the interval $[-v_0, v_0]$, we get

$$\text{Expectation}[\text{sgn}(\tilde{x}_1 + v)] = \tilde{x}_1/v_0$$

so that the average dynamics can be written

$$\begin{aligned} \dot{\tilde{x}}_{1a} &= -(\alpha_1 + k_1/v_0)\tilde{x}_{1a} + \tilde{x}_{2a} \\ \dot{\tilde{x}}_{2a} &= -(\alpha_2 + k_2/v_0)\tilde{x}_{1a} + \Delta f_a \end{aligned} \tag{10}$$

Thus, *the effect of the switching terms is to modify the effective bandwidth of the average dynamics according to the actual level of the measurement noise.* In particular, we recover Fillipov's equivalent sliding dynamics as the noise level v_0 tends to zero. Indeed,

$$1/v_0 \to \infty \quad => \quad \tilde{x}_{1a} \to 0 \quad => \quad \dot{\tilde{x}}_{1a} \to 0$$

so that

$$\begin{aligned} \dot{\tilde{x}}_{2a} &\to -[(\alpha_2 + k_2/v_0)/(\alpha_1 + k_1/v_0)]\tilde{x}_{2a} + \Delta f_a \\ &\to -(k_2/k_1)\tilde{x}_{2a} + \Delta f_a \end{aligned}$$

The above simplified analysis can be used to guide the choice of the switching gains k_j . Consider, for instance, the average error dynamics of a third-order system

$$\begin{aligned} \dot{\tilde{x}}_{1a} &= -(\alpha_1 + k_1/v_0)\tilde{x}_{1a} + \tilde{x}_{2a} \\ \dot{\tilde{x}}_{2a} &= -(\alpha_2 + k_2/v_0)\tilde{x}_{1a} + \tilde{x}_{3a} \\ \dot{\tilde{x}}_{3a} &= -(\alpha_3 + k_3/v_0)\tilde{x}_{1a} + \Delta f_a \end{aligned}$$

still with uniform bounded white noise of amplitude v_0 , and choose the α_i as in a standard Kalman filter. It is then reasonable to select the k_j so that the average dynamics be critically damped, e.g.:

$$\begin{aligned} \alpha_1 + k_1/v_0 &= 3\lambda \\ \alpha_2 + k_2/v_0 &= 3\lambda^2 \\ \alpha_3 + k_3/v_0 &= \lambda^3 \end{aligned}$$

Further, the minimum acceptable value of λ is determined by the condition

$$|\tilde{x}_{1a}| \le v_0$$

which can be written

$$v_0 \ge F/\lambda^3$$

where F is a constant (or 'slowly' varying, as compared to bandwidth λ) upper bound on Δf. The value of λ that yields the smallest k_j's is then

$$\lambda = (F/v_0)^{1/3} \tag{11}$$

which represents a reasonable choice as long as the corresponding k_j' s remain positive. The bounds on $\tilde{x}_{2a}$ and $\tilde{x}_{3a}$ can be computed accordingly. In particular, they can be easily analyzed in the frequency domain: letting p be the Laplace variable, we have

Authorized licensed use limited to: UNIVERSIDADE DE SAO PAULO. Downloaded on November 15, 2008 at 14:13 from IEEE Xplore. Restrictions apply.

$$\tilde{x}_{2_a} = [1 - (3\lambda^2 p + \lambda^3)/(p + \lambda)^3]\,(\Delta f_a/p^2)$$

$$\tilde{x}_{3_a} = [1 - \lambda^3/(p + \lambda)^3]\,(\Delta f_a/p)$$

Thus, for the observer to be fully effective, λ must be larger than the frequency content of Δf_a , which, if tuning (11) is used, imposes in turn an upper bound on the noise level v_0. Alternatively, the condition can be satisfied by increasing λ to a value larger than (11), thus also increasing the k_j and thereby the noise content of the state estimates.

3.3 Noise Sensitivity and Implementation Aspects

Based on the previous averaging analysis, the algorithm may be modified as follows: the magnitude of the switching terms can be modulated proportionally to the average value of $\tilde{x}_1$ over a short period of past history; in particular, the gains can be made small when this average is close to zero, which reflects small values of Δf_a and therefore adequacy of the Kalman filter part alone. A practical way to do so is to multiply the vector **k** by a smoothing factor of the form $sat(|\tilde{x}_{1f}|/\phi)$, where $\tilde{x}_{1f}$ is a lowpass filtered value of $\tilde{x}_1$, and ϕ is a suitably chosen boundary layer width. This procedure will be illustrated in Section 3.5. Note that the multiplication of **k** by the smoothing factor leaves the ratios k_i/k_1 unchanged.

Further, in practice, the ratios k_i/k_1 are limited by the sampling rate, i.e., the bandwidth of the equivalent dynamics on the sliding patch has to be smaller than the sampling frequency by a factor 2 or 3 . Moreover, in order for the boundary layer interpolation to be effective, the boundary layer width must be modulated, so that the time the estimation error $\tilde{x}_1$ would take to cross the boundary layer, in the absence of measurement noise, be at least two sampling periods.

3.4 Systems with Multiple Measurements

The case of systems in companion form with multiple measurements is a particular instance of a more general class of systems, discussed in Section 4.

3.5 Numerical Example

Let us consider a second order nonlinear system, consisting of a mass connected to a nonlinear spring in the presence of dry friction and stiction , in companion form :

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = -\kappa x_1^3 - f(x_2) + u$$

$$z = x_1 + v$$

where v is the measurement noise, κ is a constant nonlinear spring coefficient, and $f(x_2)$ represents dry friction with stiction. For this system, the sliding mode observer can be written :

$$\dot{\hat{x}}_1 = \alpha_1 \tilde{z} + \hat{x}_2 + k_1 sgn(\tilde{z})$$

$$\dot{\hat{x}}_2 = \alpha_2 \tilde{z} - \hat{\kappa}\hat{x}_1^3 - \hat{f}(\hat{x}_2) + u + k_2 sgn(\tilde{z}).$$

$$\tilde{z} = z - \hat{x}_1$$

The numerical values used in the simulations are :

$$\kappa = 1.0$$

$$F_s = \text{static friction} = 1.0$$

$$F_d = \text{dynamic friction} = 0.75$$

while the estimated values used in the observer are :

$$\hat{\kappa} = 0.75$$

$$\hat{F}_s = 1.25$$

$$\hat{F}_d = 1.00$$

Dry friction with stiction represents a multivalued function at the origin, and therefore, no linearization technique can be applied. Assume a white measurement noise with standard deviation .1 .Modelling the effect of parametric uncertainty as a white process noise acting on state x_2 with standard deviation 1, approximating (incorrectly) dry friction by viscous friction ,and linearizing, the Kalman Filter gains α_1 and α_2 can be obtained :

$$\alpha_1 = 3.785$$

$$\alpha_2 = 7.162$$

The switching gain k_1 is chosen as a bound on the steady state estimation error on x_2, and k_2 is chosen to be larger than modelling errors .

$$k_1 = 0.1$$

$$k_2 = 2.0$$

The true system is excited by a sinusoidal input :

$$u = sin(t)$$

and the initial conditions are :

$$x_1(0) = 0.0 \text{ and } x_2(0) = 0.5$$

with the estimated initial conditions :

$$\hat{x}_1(0) = 0.0 \text{ and } \hat{x}_2(0) = 0.2$$

The simulation results, in absence of measurement noise, are shown in Figures 5 . The results with gaussian measurement noise of standard deviation .1, are shown in Figure 6 . These simulations show that, in spite of the parameter mismatch, the sliding mode observer provides adequate performance .

The reader may be curious about how a simple "Kalman filter" would perform in this problem. This would correspond to using the linearized model alone and to letting k_1 and k_2 equal zero in the sliding mode observer. The corresponding comparative results are shown in Figure 7 . As one can see, the sliding observer yields, with a minor increase in complexity, a considerable increase in performance.

We now consider, for further illustration, a purely linear problem, to which the Kalman Filter applies directly. As-

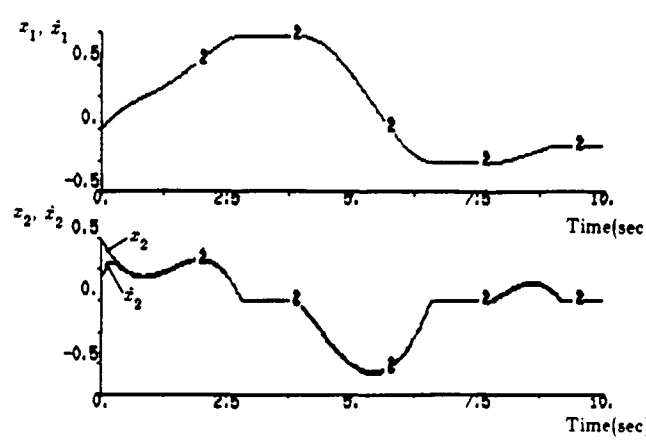

Figure 5 : Nonlinear deterministic case:true and estimated states. One can see that the estimation is essentially perfect.

Authorized licensed use limited to: UNIVERSIDADE DE SAO PAULO. Downloaded on November 15, 2008 at 14:13 from IEEE Xplore. Restrictions apply.

sume that the true system is a harmonic oscillator, and has been modelled (poorly) as a pure integrator . Introducing measurement noise, with standard deviation .1, we obtain the plots shown in Figure 8. Comparing the estimation errors resulted from the Kalman Filter and the Sliding Mode Observer, one can see that Sliding Mode Observer provided an estimation error for x_2 which is 62% of the one from the Kalman Filter. This significative improvement is due to the fact that the Sliding Mode Observer explicitly accounts for modelling errors . Note, however, that further tuning of the Kalman Filter may allow it to handle modelling errors more effectively.

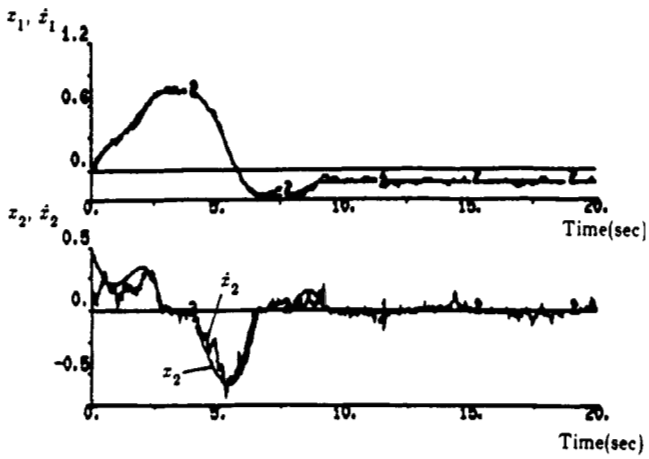

Figure 6 : Nonlinear stochastic case:true and estimated states.

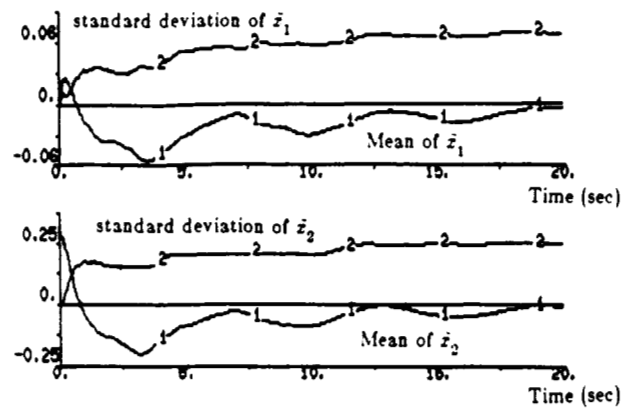

Figure 7a : Nonlinear stochastic case: statistics of estimation errors (Kalman Filter) .

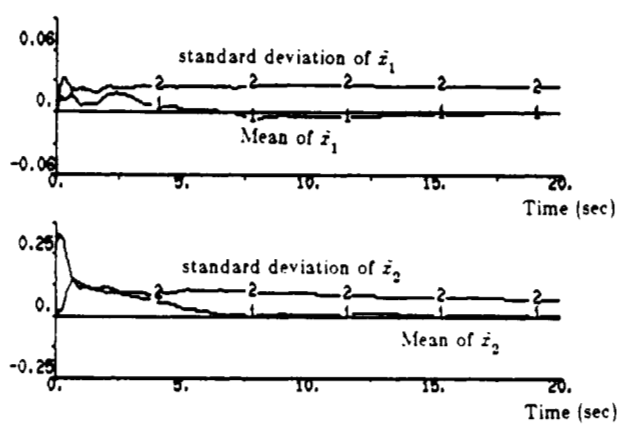

Figure 7b : Nonlinear stochastic case: statistics of estimation errors (Sliding Mode Observer) .

4. EXTENSION TO GENERAL NONLINEAR OBSERVABLE SYSTEMS

4.1 Observer Structure

Consider the nth order nonlinear system :

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},t) \quad , \quad \mathbf{x} \in R^n$$

and, for convenience, consider a vector of measurements that are linearly related to the state vector :

$$\mathbf{z} = \mathbf{C}\mathbf{x} \quad , \quad \mathbf{z} \in R^p$$

We define an observer with the following structure :

$$\dot{\hat{\mathbf{x}}} = \hat{\mathbf{f}}(\mathbf{x},t) - \mathbf{H}\tilde{\mathbf{z}} - \mathbf{K}\mathbf{1}_s$$

where $\hat{\mathbf{x}} \in R^n$, $\hat{\mathbf{f}}$ is our *model* of $\mathbf{f}$, $\mathbf{H}$ and $\mathbf{K}$ are $n \times p$ gain matrices to be specified, and $\mathbf{1}_s$ is the $p \times 1$ vector

$$\mathbf{1}_s = [sgn(\tilde{z}_1) \;\; sgn(\tilde{z}_2) \;\; sgn(\tilde{z}_3) \ldots sgn(\tilde{z}_p)]^T$$

where

$$\tilde{z}_i := \mathbf{c}_i\hat{\mathbf{x}} - z_i \tag{12}$$

and $\mathbf{c}_i$ is the i-th row of the $p \times n$ $\mathbf{C}$ matrix. We also define the error vectors :

$$\mathbf{s} := \tilde{\mathbf{z}} = \mathbf{C}(\hat{\mathbf{x}} - \mathbf{x}) \tag{13}$$

$$\tilde{\mathbf{x}} := \hat{\mathbf{x}} - \mathbf{x} \tag{14}$$

Using equations (8) and (9) we have :

$$\dot{\tilde{\mathbf{x}}} = \Delta\mathbf{f} - \mathbf{H}\tilde{\mathbf{z}} - \mathbf{K}\mathbf{1}_s \tag{15}$$

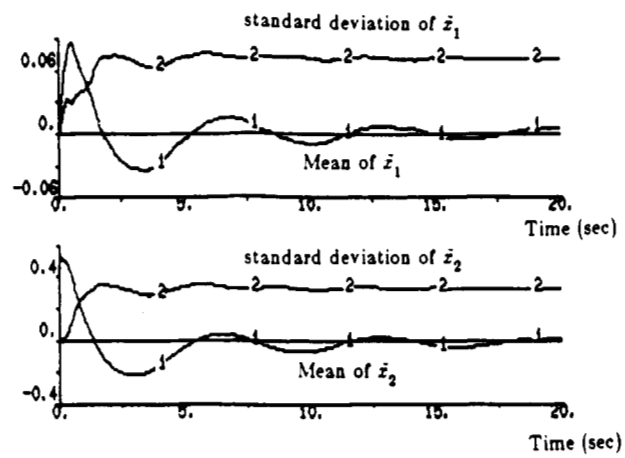

Figure 8a : Linear stochastic case: statistics of estimation errors (Kalman Filter) .

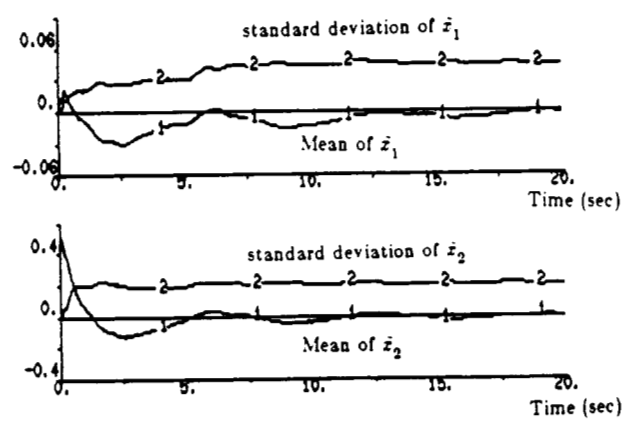

Figure 8b : Linear stochastic case: statistics of estimation errors (Sliding Mode Observer) .

Authorized licensed use limited to: UNIVERSIDADE DE SAO PAULO. Downloaded on November 15, 2008 at 14:13 from IEEE Xplore. Restrictions apply.

where

$$\Delta \mathbf{f} := \hat{\mathbf{f}}(\hat{\mathbf{x}},t) - \mathbf{f}(\mathbf{x},t)$$

For convenience we can rewrite (15) as ,

$$\dot{\tilde{\mathbf{x}}} = \tilde{\mathbf{f}} \ , \ \tilde{\mathbf{f}} = \Delta \mathbf{f} - \mathbf{H}\tilde{\mathbf{s}} - \mathbf{K}\mathbf{1}_s \tag{16}$$

The p dimensional surface, $\mathbf{s} = \mathbf{0}$,will be *attractive* if ,

$$s_i \dot{s}_i < 0 \ , \quad i=1,...p \tag{17}$$

Sliding will occur on the surface if in an arbitrarily small vicinity of s_i ,

$$s_i \dot{s}_i \leq -\eta |s_i| \ , \quad i=1,...p \tag{18}$$

Equations (18) define the *sliding region* , i.e. the multivariable extension of the *sliding patch* defined by (5). During sliding the system dynamics are effectively reduced from n^{th} order system to a $n-p$ equivalent or *reduced order* system .

The approximate dynamics on this reduced order manifold can be formally derived using the so-called 'equivalent control' method (Utkin,1977), which is equivalent to Fillipov's solution concept in the case of linear input switching . During sliding, the switching term in (16) is acting to keep $\mathbf{s} \equiv \mathbf{0}$, hence, formally, $\dot{\mathbf{s}} \equiv \mathbf{0}$. We can express the second condition as

$$\text{grad}(\mathbf{s}).\tilde{\mathbf{f}}(\tilde{\mathbf{x}},\tilde{\mathbf{1}}_s) = \mathbf{0} \tag{19}$$

where

$$\tilde{\mathbf{f}} := \Delta \mathbf{f} - \mathbf{H}\tilde{\mathbf{s}} - \mathbf{K}\tilde{\mathbf{1}}_s \tag{20}$$

and $\mathbf{1}_s$ is the equivalent switching vector , which can be obtained from (13), (19) and (20) :

$$\mathbf{C}(\Delta \mathbf{f} - \mathbf{H}\tilde{\mathbf{s}} - \mathbf{K}\tilde{\mathbf{1}}_s) = \mathbf{0}$$

so that

$$\tilde{\mathbf{1}}_s = (\mathbf{CK})^{-1}\mathbf{C}\Delta \mathbf{f} \tag{21}$$

Thus , the equivalent dynamics on the reduced order manifold is given by :

$$\dot{\tilde{\mathbf{x}}} = (\mathbf{I} - \mathbf{K}(\mathbf{CK})^{-1}\mathbf{C})\Delta \mathbf{f}$$

$$\mathbf{C}\tilde{\mathbf{x}} = \mathbf{0} \tag{22}$$

Example 4.1 Companion form (single measurement)

With $\mathbf{C} = [1\ 0\ 0\ ...\ 0]$

$\mathbf{k} = [k_1\ ....\ k_n]^T$

$\Delta \mathbf{f} = [\tilde{x}_2, \tilde{x}_3, ..., \tilde{x}_n, \hat{f}(\hat{\mathbf{x}}) - f(\mathbf{x})]^T$

expression (22) yields the reduced order dynamics previously obtained in Section 3.1.

Example 4.2 Non-companion form (single measurement)

Consider the second order nonlinear system

$$\dot{x}_1 = f_1(x_1,x_2)$$
$$\dot{x}_2 = f_2(x_1,x_2)$$
$$z = x_1$$

The observer dynamics can be written as

$$\dot{\hat{x}}_1 = \hat{f}_1(\hat{x}_1,\hat{x}_2) - \alpha_1 \tilde{x}_1 - k_1 sgn(\tilde{x}_1)$$
$$\dot{\hat{x}}_2 = \hat{f}_2(\hat{x}_1,\hat{x}_2) - \alpha_2 \tilde{x}_1 - k_2 sgn(\tilde{x}_1)$$

In order to use equivalent dynamics (22) we identify :

$$\mathbf{C} = [1\ 0] \ , \ \mathbf{K} = [k_1\ k_2]^T \ , \ \Delta \mathbf{f} = [\Delta f_1\ \Delta f_2]^T$$

where $\Delta f_i = \hat{f}_i - f_i$. Condition (17) becomes

$$\tilde{x}_1(\Delta f_1 - \alpha_1 \tilde{x}_1 - k_1 sgn(\tilde{x}_1)) < 0$$

Thus, sliding occurs when $\alpha_1 \tilde{x}_1 + k_1 > \Delta f_1$ for $\tilde{x}_1 > 0$, and $\alpha_1 \tilde{x}_1 - k_1 < \Delta f_1$ for $\tilde{x}_1 < 0$. Equation (22) yields the sliding dynamics

$$\dot{\tilde{x}}_1 = 0$$
$$\dot{\tilde{x}}_2 = -(k_2/k_1)\Delta f_1 + \Delta f_2 \tag{23}$$

The structure of Δf_1 and Δf_2 must be known before any further analysis can be done, as we now discuss .

4.2 Observability Requirements

Consider the system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},t) \ , \ \mathbf{x} \in R^n$$

$$\mathbf{z} = \mathbf{g}(\mathbf{x},t) \ , \ \mathbf{z} \in R^p$$

This system must be observable in order for any observer structure to be succesfull in reconstructing the state $\mathbf{x}$ from the measurement $\mathbf{z}$. Convenient algebraic observability conditions on $\mathbf{f}$ and $\mathbf{g}$ are not nearly so easy to find as in the linear case. (Hermann and Krener,1977) discuss the use of Lie derivatives to develop local conditions. Intuitively, in order for the system to be observable one must be able to perform successive differential operations on $\mathbf{g}(\mathbf{x})$ until an implicit inversion can be performed to obtain $\mathbf{x}$.

For instance, consider again the order nonlinear system of Example 4.2 :

$$\dot{x}_1 = f_1(x_1,x_2)$$
$$\dot{x}_2 = f_2(x_1,x_2)$$
$$z = x_1$$

In order for this system to be observable, f_1 must be a single valued function of x_2. One can see from equation (23) that Δf_1 must be a function of $\tilde{x}_2$ in order for the control term $-(k_2/k_1)\Delta f_1$ to have any influence on the error dynamics .

In general, the observability condition is strongly linked to equation (22) through the structure of the $\Delta \mathbf{f}$ vector, and an unobservable system will result in uncontrollable error dynamics .

5. APPLICATIONS

* Although high-precision ($\sim$ 16 bits) measurements of joint displacements are typically available in robot manipulators, tachometers are generally needed to generate adequate velocity feedback signals. In high performance robots, such as direct-drive arms, the cost of the tachometers may be comparable to the cost of the motors themselves. Sliding observers should provide the capability of generating clean velocity signals directly from joint displacement measurements, using the nonlinear robot model, therefore reducing cost and furthermore potentially yielding improved performance.

* The performance of force control schemes is critically dependent on the quality of the derivative of the contact force signal (Slotine, 1986). However, force signals are typically corrupted by large amounts of mechanical noise generated by the robot and the environment (from gears, stiction, etc.). Furthermore, force observer design requires a model of the generally poorly known environment stiffness. This is again a domain where robust observer design using sliding behavior could potentially be very effective.

Authorized licensed use limited to: UNIVERSIDADE DE SAO PAULO. Downloaded on November 15, 2008 at 14:13 from IEEE Xplore. Restrictions apply.

* In electric motors, pure sliding mode (switching) control may be used in place of pulse-width modulation. This requires, however, to use third-order models, which involves estimating joint velocities and accelerations from encoder signals (as well as, perhaps, tachometer signals and armature current measurements) in order to compute the value of s . Using each actuator in a pure saturated switched mode

$$u = -u_{max} sgn(s)$$

in conjunction with a sliding observer, at switching frequencies similar to those of PWMs (10-40 KHz), would in principle allow extremely high performance using simple design. In a robot manipulator, it would also permit each motor to be controlled independently (i.e., without accounting explicitely for the coupling effects and nonlinear dynamic forces). This would allow modularity in the robot design, with afferent costs and maintenance savings.

* Along the same lines, pure sliding mode can be applied directly in end-effector space: the expression of the vector **s** (of components s_i) is then based on errors in end-effector position rather than errors in joint displacements. The resulting end-effector forces are then expressed in terms of joint torques θ to be applied, using the manipulator Jacobian **J** (Asada and Slotine, 1986):

$$\theta = \mathbf{J}^T \mathbf{F}_{sliding}$$

where the components of $\mathbf{F}_{sliding}$ are of the form

$$F_i = -F_{imax} sgn(s_i)$$

Accurate computation of **s** involves again the use of a nonlinear observer.

* Another potentially useful application of sliding mode observer is to estimate shaft torques in automotive powertrain applications. Force sensors are available , but (again) several orders of magnitude too expensive for practical implementation . The extremely nonlinear nature of the engine/transmission environment prohibits the use of linear estimations.

6. CONCLUDING REMARKS

But is it Art ? Clearly, this study is only a step in developing a complete and systematic methodology of sliding observer design for nonlinear systems, and the reader may not want to throw away his Kalman filters yet. However, sliding observers have intriguing properties in the presence of measurement noise, and have predictably robust properties in the presence of modelling errors.

By a simple but very nonlinear example it was shown how a simple sliding observer could be designed, and that its switching part provided robustness to modelling errors. A simple linear example with modelling errors also illustrated that the sliding observer significantly reduced the estimation error over its Kalman Filter counterpart. These properties should make sliding observers worthy of extensive further research.

REFERENCES

Asada, H., and Slotine, J.-J. E., 1986. Robot Analysis and Contr ol, *Wiley.*

Filippov, A.F., 1960. *Am. Math. Soc. Trans.*, 62, 199.

Hermann,R. and Krener,A.J.,1977 Nonlinear Controllability and Observability,*I.E.E.E. Trans.Autom.Control* 22,798.

Hunt, L.R., Su, R., and Meyer, G., 1983. Design for Multi Input Nonlinear Systems, *Differential Geometry and Control Theory Conf.*,Birkhauser, Boston, 268

Itkis,V.,1976. Control Systems of Variable Structure, *Wiley.*

Krener,A.J. and Respondek,W., 1985. Nonlinear Observers with Linearizable Error Dynamics,*S.I.A.M. J.Contr Optim.* 23, 2

Slotine,J.J.E.. 1984 Sliding Controller Design for Nonlinear Systems, *Int. J. Control*, 40, 421.

Slotine, J.J.E., and Sastry, S.S., 1983. Tracking Control of Nonlinear Systems Using Sliding Surfaces, with Application to Robot Manipulators, *Int. J. Control*, 38, 465.

Slotine, J.-J. E., 1986. Robustness and Adaptation in Robot Cont rol, *Japan-USA Symposium on Flexible Automation, Osaka.*

Utkin, V.I., 1977. Variable Structure Systems with Sliding Mode: A Survey, *I.E.E.E. Trans. Autom. Control* 22, 212.

Authorized licensed use limited to: UNIVERSIDADE DE SAO PAULO. Downloaded on November 15, 2008 at 14:13 from IEEE Xplore. Restrictions apply.