Design and Analysis of Incipient Fault Diagnosis Schemes Using On-Line Approximators

M. M. Polycarpou[†] and **M. A. Demetriou**[‡]

[†]Department of Electrical and Computer Engineering
University of Cincinnati, Cincinnati, OH 45221-0030, USA

[‡]Mechanical Engineering Department
Worcester Polytechnic Institute, Worcester, MA 01609-2280, USA

Abstract

Detection of incipient (slowly developing) faults is crucial in automated maintenance problems where early detection of worn equipment is required. In this paper, a general framework for model-based fault detection and diagnosis of a class of incipient faults is developed. The changes in the system dynamics due to the fault are modeled as nonlinear functions of the state and input variables, while the time profile of the failure is assumed to be exponentially developing. An automated fault diagnosis architecture using nonlinear on-line approximators with an adaptation scheme is designed and analyzed.

1 Introduction

Increased productivity requirements and stringent performance specifications lead to more demanding operating conditions of many modern engineering systems. Such conditions increase the possibility of system failures which are characterized by critical, unpredictable changes in the system dynamics. In general, feedback control algorithms which are designed to handle small system perturbations that may arise under "normal" operating conditions (typically, in the linear regime), cannot accommodate abnormal behavior due to faults. Automated maintenance for early detection of worn equipment is becoming a crucial problem is many practical applications. Therefore, the development of new design and analysis methods for health monitoring and fault diagnosis is a key component in the safe operation of advanced engineering systems.

The process of fault diagnosis consists of three steps: (i) *detection* deals with determining if a malfunction has occured in the supervised system; (ii) *diagnosis* considers the problem of identifying the location, type, and characteristics of a failure; and (iii) *accommodation* attempts to self-correct a particular failure, typically through reconfiguration of the control decision policy.

The design of fault diagnosis algorithms using the model-based analytical redundancy approach has received a lot of attention during the last two decades (see, for example, survey papers by Frank [4], Gertler [5], Isserman [8], and Willsky [15]). Various quantitative models (such as state-space models, parametric models, parity relations), as well as qualitative models (such as expert systems) have been used to generate a residual vector that provides a measure of the deviation between estimated and measured signals. In general, a fault is declared if the "size" of the residual vector exceeds a certain threshold value.

The nature of possible failure situations may be classified as *abrupt* (sudden) failures, which are typically modeled as step-like deviations, and *incipient* (slowly developing) failures, which are represented by drift-type changes. In abrupt type failures, it is crucial that the fault diagnosis scheme is able to detect the changes quickly so as to avoid catastrophic consequences. In such cases, early detection and accommodation are the key objectives of fault diagnosis. On the other hand, incipient failures are more important in maintenance activities where it is required that slowly developing problems are detected early enough to avoid more serious consequences. Therefore, the development of effective fault diagnosis schemes for incipient faults plays a key role in the automation of inspection procedures and minimization of maintenance activities and costs. One of the main difficulties in dealing with incipient faults is the compensating effect of feedback control, which tends to diminish the effect of small incipient faults on the tracking performance.

In this paper, a fault diagnosis methodology for incipient faults is developed. We consider nonlinear dynamical systems whose dynamics change at some unknown time due to a failure. This change is modeled as an unknown nonlinear function of the state and input variables with a time-varying failure profile. In order to capture the nonlinear characteristics of faults, we design a nonlinear estimator using the on-line approximation approach [12] with an adaptive scheme for the adjustable parameters or weights.

The stability and performance properties of the fault diagnosis scheme are rigorously established under the assumption of full state measurement. These results are obtained in the presence of *approximation errors*, that is, errors arising as a result of imperfect modeling of the system deviations due to faults by the on-line approximator.

From an adaptive theory viewpoint, the objective of this paper is to develop a *learning* methodology for incipient failure detection. In this framework, on-line approximators (such as neural networks, spline functions, wavelets, etc.) are used to monitor the system for any deviations due to faults. By using the adaptivity capabilities of on-line approximators, they can be used not only to detect the occurrence of system failures but also to provide an on-line estimate of the fault characteristics (diagnosis).

The paper is organized as follows: In Section 2 we outline the class of dynamical systems under study and describe the general structure of the nonlinear estimator. The synthesis of the fault diagnosis scheme is presented in Section 3. In Section 4 we investigate the robustness, stability and performance properties of the incipient fault diagnosis scheme.

2 General Formulation

Most fault diagnosis schemes developed so far have dealt exclusively with linear models subject to faults that are represented as external additive input signals (of time). Although such linear techniques allow the derivation of many analytical results, in real engineering applications linear-based methods may lead to degraded performance of the fault diagnosis scheme. To capture some of the characteristics of practical failure situations, in this section we present a *nonlinear* modeling framework for representing failures and developing estimation schemes.

2.1 Representation of Failures

The class of dynamical systems under study is described by

$$\dot{x}(t) = \xi(x(t), u(t)) + \mathcal{B}(t-T) f(x(t), u(t)) \qquad (1)$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^m$ is the input vector, $\xi, f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ are smooth vector fields, $T \geq 0$ is the beginning time of the failure, and $\mathcal{B}$ is a square $n \times n$ matrix function representing the time profiles of failures. We consider incipient faults that are modeled by

$$\mathcal{B}(t-T) = \operatorname{diag}\left(\beta_1(t-T), \beta_2(t-T), \ldots, \beta_n(t-T)\right),$$

where for each $i = 1, 2, \ldots, n$,

$$\beta_i(\tau) = \begin{cases} 0 & \text{if } \tau < 0 \\ 1 - e^{-\rho_i \tau} & \text{if } \tau \geq 0 \end{cases} \qquad (2)$$

and $\rho_i > 0$ is an unknown constant that represents the rate at which the failure in state x_i evolves. For large values of ρ_i, the time profile function β_i approaches a step function, which models abrupt failures.

The objective is to design a fault diagnosis scheme that processes input and state information to determine the presence and characteristics of any incipient faults. Since this paper does not address fault accommodation, below we make the standard assumption that the control input u and the state vector x remain bounded prior and after the occurrence of a fault:
(A1) There exist compact sets $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{U} \subset \mathbb{R}^m$ such that $x(t) \in \mathcal{X}$ and $u(t) \in \mathcal{U}$ for all $t \geq 0$.

The "healthy" system in the absence of any faults is described by

$$\dot{x}_h(t) = \xi(x_h(t), u(t)) := \xi^*(x_h(t), u(t)) + \tilde{\xi}(x_h(t), u(t)),$$

where ξ^* represents the nominal dynamics (known) and $\tilde{\xi}$ characterizes any discrepancy between the actual plant and nominal model that may occur due to modeling errors. It is well known in the fault diagnosis literature that the presence of modeling errors, in general, increases the probability of false alarms. During the last few years the design of so-called *robust fault diagnosis* schemes have resulted in a variety of tools for dealing with such modeling uncertainties [10, 13]. An intuitive approach is to use a small threshold in the residual error to account for modeling uncertainties; in this case, a fault is declared if the residual error is greater than the selected threshold. Another approach attempts to decouple the effects of faults and modeling errors as a way of improving robustness. In this work, we first consider the ideal case where $\tilde{\xi} \equiv 0$ and then the case where $|\tilde{\xi}(x, u)| \leq \xi_0$ for all $(x, u) \in (\mathcal{X} \times \mathcal{U})$, where $\xi_0 \geq 0$ is a known constant. In general, the design and analysis of robust fault diagnosis architectures based on nonlinear modeling techniques requires further investigation.

2.2 Nonlinear Estimator

The failure representation described by (1) provides a framework for characterizing a wide class of faults. In general, the magnitude of faults in practical applications depends on the state of the system as well as the system input. The nonlinear fault representation (1) captures these dependencies of f on the state x and the input u. The price that one has to pay for the potential to model a larger class of failures is the need to approximate unknown nonlinear functions, which leads to nonlinear fault diagnosis techniques. This can be realized by the utilization of parameterized on-line approximation structures with adjustable parameters. Such an adaptive nonlinear estimator is given by

$$\hat{x} = W(s)[z] \quad (3)$$
$$z = \zeta(x, u, \hat{\theta}) \quad (4)$$
$$\dot{\hat{\theta}} = \eta(x, u, \hat{x}, \hat{\theta}), \quad (5)$$

where $W(s)$ is an $n \times n$ stable filter matrix, (3), (4) represents an observer-based nonlinear estimation scheme, and (5) is the adaptive law of the adjustable parameters. Next, we proceed to the design of $W(s)$, ζ and η.

3 Fault Diagnosis Scheme

Following the formulation of [12], we consider a nonlinear model-based estimator given by

$$\dot{\hat{x}}(t) = A\hat{x}(t) + \xi^*(x(t), u(t)) + \hat{f}(x(t), u(t); \hat{\theta}(t)) - Ax(t) \quad (6)$$

where $\hat{x} \in \mathbb{R}^n$ is the *estimated state vector*, $\hat{f}$ is the *on-line approximation model*, $\hat{\theta} \in \mathbb{R}^q$ is a vector of *adjustable parameters*, and A is a constant $n \times n$ matrix that satisfies the Lyapunov equation $A^T\Pi + \Pi A = -Q$, with $\Pi = \Pi^T > 0$ and $Q > 0$. In the framework of the general estimation scheme (3), (4), $W(s) = (sI - A)^{-1}$ and $\zeta(x, u, \hat{\theta}) = \xi^*(x, u) + \hat{f}(x, u; \hat{\theta}) - Ax$. The construction of an accurate nonlinear model based estimator, able to follow any variations in the physical system, is a crucial component of the overall learning scheme. The nonlinear estimator described by (6) is based on a series-parallel, error filtering scheme [12], which is shown in Section 4 to have some desirable stability and performance properties. An alternative approach, pursued in [3] for fault accommodation, uses the estimated state $\hat{x}$ instead of the measured state x in the nonlinear estimation scheme.

The initial condition for the estimated model (6) is $\hat{x}(0) = x(0)$ and $\hat{\theta}(0) = \hat{\theta}^0$ is selected such that $\hat{f}(x, u; \hat{\theta}^0) = 0$ for all x and u. Starting from these initial conditions, the objective is to develop a parameter adaptive law for $\hat{\theta}(t)$ so that the on-line approximator $\hat{f}(x, u; \hat{\theta})$ approximates the function $\mathcal{B}(t-T)f(x, u)$. Once this is achieved, then the on-line approximator $\hat{f}$ may be used not only to detect failures but also to diagnose these failures in the sense of identifying their magnitude and dependency on x and u. Where appropriate, the on-line approximator may be used for failure accommodation. Note that in this paper it is assumed that the failure modes, described by f, are unknown. In the special case that all possible failure modes are known apriori then a multiple-model estimation scheme that takes into consideration the knowledge of each failure mode can be used to improve performance [1].

During the last few years several on-line approximation models have been studied in the context of intelligent and learning control [14]. In addition to conventional approximation models like polynomials, rational functions, spline functions etc., various neural network topologies such as sigmoidal neural networks, radial basis function networks, CMAC networks etc, and new network structures such as wavelet networks have emerged. In the framework of adaptive networks, (x, u) is the input vector to the network, $\hat{\theta}$ is a vector of adjustable parameters or weights, and $y = \hat{f}(x, u; \hat{\theta})$ is the output of the network. In this paper we consider a general class of sufficiently smooth on-line approximators; that is, $\hat{f} \in C^\infty$.

First we consider the ideal case of no modeling errors; i.e., $\tilde{\xi} = 0$. Later, we modify the adaptive law to handle modeling uncertainty. Using Lyapunov redesign methods [9], we obtain the following adaptive law

$$\dot{\hat{\theta}} = \mathcal{P}\left\{\Gamma Z \Pi e\right\}, \quad (7)$$

where $e = x - \hat{x}$ is the state estimation error, $\Gamma = \Gamma^T > 0$ is the learning rate, Z is a $q \times n$ matrix given by

$$Z = \left[\frac{\partial \hat{f}(x, u; \hat{\theta})}{\partial \hat{\theta}}\right]^T, \quad (8)$$

and $\mathcal{P}$ is the projection operator, which constrains the parameter $\hat{\theta}$ to some selected compact, convex region $\mathcal{M}_{\hat{\theta}}$ of the parameter space $\mathbb{R}^q$.

The projection operator in the adaptive law is used to prevent parameter drift of the adjustable weights, a phenomenon that may occur with standard adaptive laws in the presence of modeling uncertainty and approximation error [7]. As shown in the analysis given below, the selection of the compact region $\mathcal{M}_{\hat{\theta}}$ does not require knowledge of an upper bound on the optimal parameter θ^*. In the special case that $\mathcal{M}_{\hat{\theta}}$ is chosen to be a hypersphere of size M (i.e., $\mathcal{M}_{\hat{\theta}} = \left\{\hat{\theta} \in \mathbb{R}^q \,:\, |\hat{\theta}| \leq M\right\}$), then the above adaptive law can be expressed as

$$\dot{\hat{\theta}} = \Gamma Z \Pi e - \chi^* \Gamma \frac{\hat{\theta}\hat{\theta}^T}{|\hat{\theta}|^2} \Gamma Z \Pi e, \quad (9)$$

where χ^* denotes the *indicator* function given by

$$\chi^* = \begin{cases} 0 & \text{if } \ (|\hat{\theta}| < M) \text{ or } (|\hat{\theta}| = M \ \& \ \hat{\theta}^T \Gamma Z \Pi e \leq 0) \\ 1 & \text{if } \ (|\hat{\theta}| = M \ \& \ \hat{\theta}^T \Gamma Z \Pi e > 0). \end{cases}$$

3.1 Robust Fault Diagnosis

Under ideal conditions of no modeling errors a fault is declared whenever the output of the on-line approximator $y = \hat{f}(x, u, \hat{\theta})$ becomes non-zero. A straightforward and practical way of improving the robustness of the algorithm with respect to modeling uncertainties is to declare a fault whenever $|y| \geq \delta$, where $\delta \geq 0$ is a design parameter that depends on the magnitude of the modeling uncertainties. This approach to improving robustness is incorporated into the learning methodology developed above by modifying the adaptive law (7) as follows:

$$\dot{\hat{\theta}} = \mathcal{P}\left\{\Gamma Z \Pi D[e]\right\}, \tag{10}$$

where $D[\cdot]$ is the dead-zone operator [11], defined as

$$D[e] := \begin{cases} 0 & \text{if } |e| \leq \epsilon \\ e & \text{if } |e| > \epsilon \end{cases},$$

where $\epsilon > 0$ is a design constant. The selection of the dead-zone size ϵ clearly induces a trade-off between reducing the possibility of false alarms (robustness) and improving the sensitivity to faults. In the next section we derive a value for the dead-zone size ϵ (in terms of the modeling uncertainty bound ξ_0) that guarantees robustness in the presence of any modeling uncertainty satisfying the given bound.

4 Analysis

The fault diagnosis scheme described in Section 3 has some desirable stability, performance and robustness properties, which are presented in this section. These results are obtained for the case of incipient failures that occur at some unknown time T and develop with unknown rates ρ_i. The incipient failure changes the dynamics of the system but is assumed to retain the boundedness of the state and input variables (Assumption A1).

First consider the ideal case of no modeling errors. In the time interval $t \in [0, T)$ (i.e., prior to the occurrence of a fault), the state estimation error $e = x - \hat{x}$ and parameter estimate $\hat{\theta}$ satisfy

$$\begin{aligned} \dot{e} &= Ae - \hat{f}(x,u,\hat{\theta}) & e(0) &= 0, \quad (11) \\ \dot{\hat{\theta}} &= \mathcal{P}\left\{\Gamma Z \Pi e\right\} & \hat{\theta}(0) &= \hat{\theta}^0. \quad (12) \end{aligned}$$

Since $\hat{\theta}^0$ is chosen such that $\hat{f}(x,u,\hat{\theta}^0) = 0$ for all x and u, the vector $\left(e,\ \hat{\theta}\right) = \left(0,\ \hat{\theta}^0\right)$ is an equilibrium of the above system. Therefore, $e(t) = 0$ and $\hat{\theta}(t) = \hat{\theta}^0$ for $t \in [0, T)$.

In the presence of modeling errors, (11) becomes

$$\dot{e} = Ae + \tilde{\xi}(x,u) - \hat{f}(x,u,\hat{\theta}). \tag{13}$$

According to the robust adaptive law (10), the output of the on-line approximator remains zero a long as $|e(t)| \leq \epsilon$. To determine an appropriate value for ϵ, we derive an upper bound for $e(t)$ in the case $\hat{f}(x(t), u(t), \hat{\theta}(t)) = 0$. From (13), we have

$$e(t) = \int_0^t e^{A(t-\tau)}\tilde{\xi}(x(\tau), u(\tau))\, d\tau.$$

Since A is a stability matrix, there exist positive constants ρ and α such that $\|e^{At}\| \leq \rho e^{-\alpha t}$. Therefore

$$\begin{aligned} |e(t)| &\leq \rho \int_0^t e^{-\alpha(t-\tau)}\xi_0\, d\tau \\ &= \frac{\rho\xi_0}{\alpha}(1 - e^{-\alpha t}). \end{aligned}$$

This implies that if the dead-zone size is chosen as $\epsilon = \frac{\rho}{\alpha}\xi_0$, then $e(t)$ remains within the dead-zone for all $t \leq T$ and the output of the on-line approximator remains zero; in other words, the set $\{(e,\ \hat{\theta})\ :\ |e| < \epsilon,\ \hat{\theta} = \hat{\theta}_0\}$ is a positively invariant set [9].

Therefore, the learning algorithm given by (10) is robust in the sense that it is not affected by modeling uncertainties that satisfy $|\tilde{\xi}(x,u)| \leq \xi_0$. Furthermore, by letting $\tilde{\xi}(x(t), u(t)) = \xi_0$ for all t, it is easy to verify that the selected bound for the dead-zone size ϵ is not conservative.

Next we consider the time interval $t \geq T$, after the occurrence of a fault. Using (1), (6), the state estimation error satisfies

$$\begin{aligned} \dot{e} &= Ae + \tilde{\xi}(x,u) + \mathcal{B}(t-T)f(x,u) - \hat{f}(x,u,\hat{\theta}) \\ &= Ae + \tilde{\xi}(x,u) + \mathcal{B}(t-T)\hat{f}(x,u,\theta^*) - \hat{f}(x,u,\hat{\theta}) \\ &\quad + \nu(t) \end{aligned} \tag{14}$$

where $\nu(t)$ is the *approximation error* given by

$$\nu(t) = \mathcal{B}(t-T)\left[f(x(t),u(t)) - \hat{f}(x(t),u(t);\theta^*)\right]. \tag{15}$$

The "optimal" parameter θ^* is chosen as the value of $\hat{\theta}$ that minimizes the L_2-norm (energy-norm) distance between $f(x,u)$ and $\hat{f}(x,u;\hat{\theta})$ over all $(x,u) \in \mathcal{X} \times \mathcal{U}$ subject to the constraint that $\hat{\theta} \in \mathcal{M}_{\hat{\theta}}$ [12]. It is noted that θ^* is an "artificial" quantity required only for analysis purposes.

Under smoothness assumptions on $\hat{f}$, (14) can be expressed as

$$\begin{aligned} \dot{e} &= Ae + \tilde{\xi}(x,u) - [I - \mathcal{B}(t-T)]\hat{f}(x,u,\theta^*) \\ &\quad - \frac{\partial \hat{f}(x,u;\hat{\theta})}{\partial \hat{\theta}}\left(\hat{\theta} - \theta^*\right) - \Delta + \nu \end{aligned} \tag{16}$$

where Δ is given by

$$\begin{aligned} \Delta(x,u;\hat{\theta},\theta^*) &= \hat{f}(x,u;\hat{\theta}) - \hat{f}(x,u,\theta^*) \\ &\quad - \frac{\partial \hat{f}(x,u;\hat{\theta})}{\partial \hat{\theta}}\left(\hat{\theta} - \theta^*\right) \end{aligned}$$

Intuitively, Δ represents the higher-order terms of the Taylor series expansion of $\hat{f}(x,u,\hat{\theta})$ with respect to $\hat{\theta}$. Indeed, it can be readily shown using the Mean Value Theorem [2] that

$$\left|\Delta(x,u;\hat{\theta},\theta^*)\right| \leq p(x,u,\hat{\theta},\theta^*)|\hat{\theta} - \theta^*|,$$

where $\lim_{\hat{\theta}\to\theta^*} p(x,u,\hat{\theta},\theta^*) = 0$ for all $(x,u) \in \mathcal{X}\times\mathcal{U}$. In the special case of a linearly parameterized approximator (i.e., $\hat{f}(x,u;\hat{\theta}) = \Omega(x,u)^T\hat{\theta}$), the higher-order term component Δ is identically equal to zero. Examples of linearly parameterized approximators include polynomial functions and radial basis function networks with fixed centers and widths.

By letting $\tilde{\theta}(t) = \hat{\theta} - \theta^*$, $\omega(t) = -\Delta(x(t), u(t); \hat{\theta}(t), \theta^*) + \nu(t)$, and using (8), the error equation (16) becomes

$$\dot{e} = Ae + \tilde{\xi}(x,u) - \Phi\hat{f}(x,u,\theta^*) - Z^T\tilde{\theta} + \omega, \quad (17)$$

where $\Phi(t) = I - \mathcal{B}(t-T)$ is a diagonal matrix. Clearly, For $t \geq T$ the matrix Φ satisfies

$$\dot{\Phi}(t) = -P\Phi(t) \qquad\qquad \Phi(T) = I,$$

where P is a constant positive definite matrix given by

$$P = \mathrm{diag}(\rho_1, \rho_2, \ldots, \rho_n).$$

If the norm of the state estimation error is within the dead-zone (i.e., $|e| \leq \epsilon$) then $\dot{\hat{\theta}} = 0$ and hence stability follows trivially. In order to analyze the stability and performance properties of the fault diagnosis scheme in the case $|e| > \epsilon$, we consider the Lyapunov function candidate

$$\begin{aligned} V(t) &= \frac{1}{2}e^T(t)\Pi e(t) + \frac{1}{2}\tilde{\theta}^T(t)\Gamma^{-1}\tilde{\theta}(t) \\ &\quad + \frac{1}{2}\,\mathrm{trace}\left\{\lambda\Phi(t)P^{-1}\Phi(t)\right\}, \end{aligned}$$

where $\lambda > 0$ is a constant scalar to be selected later. The time derivative of $V(t)$ evaluated along the trajectories of (17), (7), yields

$$\begin{aligned} \dot{V} &= \tfrac{1}{2}e^T\left(A^T\Pi + \Pi A\right)e - e^T\Pi\Phi\hat{f}(x,u,\theta^*) \\ &\quad - e^T\Pi Z^T\tilde{\theta} + e^T\Pi\omega + e^T\tilde{\xi}(x,u) \\ &\quad + \tilde{\theta}^T\Gamma^{-1}\mathcal{P}\left\{\Gamma Z\Pi e\right\} - \mathrm{trace}\left\{\lambda\Phi\Phi\right\} \\ &= -\tfrac{1}{2}e^TQe + e^T\tilde{\xi}(x,u) - e^T\Pi\Phi\hat{f}(x,u,\theta^*) \\ &\quad - \chi^*\tilde{\theta}^T\tfrac{\hat{\theta}\hat{\theta}^T}{|\hat{\theta}|^2}\Gamma Z\Pi e + e^T\Pi\omega - \mathrm{trace}\left\{\lambda\Phi\Phi\right\}. \end{aligned} \quad (18)$$

Using standard techniques from adaptive control [7, 12] it can be shown that the projection term can only make the derivative of the Lyapunov function more negative; i.e.,

$$\chi^*\tilde{\theta}^T\frac{\hat{\theta}\hat{\theta}^T}{|\hat{\theta}|^2}\Gamma Z\Pi e \geq 0.$$

Now, using the smoothness assumption on $\hat{f}$ and the uniform boundedness of $x(t)$ and $u(t)$, there exists a finite constant c_1 such that

$$c_1 = \sup_{t\geq T}\left\{\hat{f}(x(t), u(t), \theta^*)\right\}.$$

The Frobenius matrix norm, defined as $\|A\|_F^2 = \sum_{ij}|a_{ij}|^2 = \mathrm{trace}\{AA^T\}$, satisfies $\|A\|_2 \leq \|A\|_F$ (see [6]). Therefore, (18) becomes

$$\begin{aligned} \dot{V} &\leq -\frac{\lambda_{min}(Q)}{2}|e|^2 - \lambda\|\Phi\|_F^2 + c_1\|\Pi\|_2\|\Phi\|_F|e| \\ &\quad + e^T\Pi\omega + e^T\tilde{\xi}(x,u). \end{aligned}$$

By completing the squares and setting $\lambda := \frac{4(c_1\|\Pi\|_2)^2}{\lambda_{min}(Q)}$ it can be readily shown that

$$\begin{aligned} \dot{V} &\leq -\frac{\lambda_{min}(Q)}{4}|e|^2 - \frac{\lambda}{2}\|\Phi\|_F^2 \\ &\quad - \left[\frac{\lambda_{min}(Q)}{8}|e|^2 - e^T\left(\Pi\omega + \tilde{\xi}(x,u)\right)\right] \\ &\quad - \left[\frac{\lambda_{min}(Q)}{8}|e|^2 - c_1\|\Pi\|_2\|\Phi\|_F|e| + \frac{\lambda}{2}\|\Phi\|_F^2\right] \\ &\leq -\frac{\lambda_{min}(Q)}{4}|e|^2 - \frac{\lambda}{2}\|\Phi\|_F^2 + k_1|\omega|^2 \\ &\quad + k_2|\tilde{\xi}(x,u)|^2, \end{aligned} \quad (19)$$

where the constants k_1, k_2 are given by $k_1 = \frac{2(\lambda_{max}(\Pi))^2}{\lambda_{min}(Q)}$ and $k_2 = \frac{2}{\lambda_{min}(Q)}$.

When $\frac{\lambda_{min}(Q)}{2}|e(t)|^2 + \frac{\lambda}{2}\|\Phi\|_F^2 > k_1|\omega(t)|^2 + k_2|\tilde{\xi}(x,u)|^2$ we have that $\dot{V}(t) \leq 0$. Since $\omega(t)$ and $\tilde{\xi}(x(t), u(t))$ are uniformly bounded, the above inequality implies that there exists a constant k_3 such that if $|e(t)| > k_3$ then $\dot{V}(t) \leq 0$. Furthermore, the projection operator guarantees that $\hat{\theta}$ is uniformly bounded. Thus we can infer that $V, e, \tilde{\theta}$ are also uniformly bounded. By integrating (19) over any finite interval $[T,\ T+\tau]$ and re-arranging terms we obtain

$$\begin{aligned} \int_T^{T+\tau}|e(t)|^2\,dt &\leq \frac{4}{\lambda_{min}(Q)}\Big[V(T) - V(T+\tau) \\ &\quad + k_1\int_T^{T+\tau}|\omega(t)|^2\,dt \\ &\quad + k_2\int_T^{T+\tau}|\tilde{\xi}(x(t), u(t))|^2\,dt\Big] \\ &\leq \lambda_1 + \lambda_2\int_T^{T+\tau}|\omega(t)|^2\,dt \\ &\quad + \lambda_2\int_T^{T+\tau}|\tilde{\xi}(x(t), u(t))|^2\,dt, \end{aligned} \quad (20)$$

where $\lambda_1 = \dfrac{4}{\lambda_{min}(Q)}\sup_{\tau\geq 0}\left[V(T) - V(T+\tau)\right]$ and $\lambda_2 = \dfrac{8}{\lambda_{min}^2(Q)}\max\{1,\ \lambda_{max}^2(\Pi)\}$.

The above analysis guarantees the uniform boundedness of the nonlinear estimator, including the adjustable parameters of the on-line approximator. Furthermore, the extended L_2-norm of the state estimation error $e(t)$ over any finite time interval is, at most, of the same order as the extended L_2-norm of $\omega(t)$ and $\tilde{\xi}(x(t), u(t))$. Hence the inequality (20) gives a qualitative relationship between the performance of the learning scheme and $\omega(t)$, as well as the modeling uncertainty $\tilde{\xi}(x(t), u(t))$.

According to the definition of $\Phi(t)$ the time profile of the failure is required to satisfy $\dot{\Phi}(t) = -P\Phi(t)$, where P is a positive-definite matrix. It is important to note that the stability analysis presented above is still valid as long as $\dot{\Phi}(t) \leq -P\Phi(t)$. Therefore, a wider class of failure situations can be treated in a similar framework.

In the special case of linearly parameterized approximators, we have $\Delta \equiv 0$; hence $\omega(t) = \nu(t)$. If, in addition, $\nu(t) = 0$ (that is, the on-line approximator can match the fault exactly) or $\nu \in L_2$, and the modeling uncertainty $\tilde{\xi}(x(t), u(t)) = 0$ then clearly $e(t)$ is square integrable. Therefore, in this case Barbălat's Lemma [7] can be used to show that $\lim_{t\to\infty} e(t) = 0$.

A nonlinear estimator (adaptive diagnostic estimator) was designed via Lyapunov redesign methods and shown to actually provide a means of both detecting and diagnosing the characteristics of the system failures described by nonlinear functions of the state. Modifications to the adaptation rules were added to account for modeling errors in the plant dynamics, thus leading to a robust (with respect to false alarms) diagnostic scheme. Since the aforementioned diagnostic estimator required knowledge of the full state, a natural extension is to extend the above results in the case of partial state information. This would require that the nominal input-output transfer function be strictly positive real in order to utilize Lyapunov redesign methods. This is currently pursued by the authors.

References

[1] A. Alessandri and T. Parisini, *Model-based fault diagnosis using nonlinear estimators: a neural approach*, in Proc. of the 1997 Amer. Control Conf. (to appear), June 1997.

[2] J. D. DePree and C. W. Swartz, *Introduction to Real Analysis*, Wiley, New York, 1988.

[3] J. A. Farrell, T. Berger, and B. D. Appleby, *Using learning techniques to accommodate unanticipated faults*, IEEE Contr. Sys. Mag., 13 (1993), pp. 40–49.

[4] P. M. Frank, *Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy - a survey and some new results*, Automatica, 26 (1990), pp. 459–474.

[5] J. Gertler, *Survey of model-based failure detection and isolation in complex plants*, IEEE Control System Magazine, 8 (1988), pp. 3–11.

[6] G. H. Golub and C. F. V. Loan, *Matrix Computations*, The John Hopkins University Press, Baltimore, MD, 2 ed., 1989.

[7] P. A. Ioannou and J. Sun, *Stable and Robust Adaptive Systems*, Prentice Hall, Englewood Cliffs, NJ, 1995.

[8] R. Isermann, *Process fault detection based on modeling and estimation methods: a survey*, Automatica, 20 (1984), pp. 387–404.

[9] H. K. Khalil, *Nonlinear Systems*, Macmillan, New York, 1992.

[10] R. J. Patton, *Robust model based fault diagnosis: the state-of-the-art*, in Proc. IFAC Symposium on Fault Detection, supervision and Safety for Processes (SAFEPROCESS), Espoo, Finland, June 1994, pp. 1–24.

[11] B. B. Peterson and K. S. Narendra, *Bounded error adaptive control*, IEEE Transactions on Automatic Control, 27 (1982), pp. 1161–1168.

[12] M. M. Polycarpou and A. J. Helmicki, *Automated fault detection and accomodation: A learning systems approach*, IEEE Trans. on Systems, Man and Cybernetics, 25 (1995), pp. 1447–1458.

[13] A. T. Vemuri and M. M. Polycarpou, *On-line approximation methods for robust fault detection*, in Proceedings of the 13th World Congress of Inter. Fed. Autom. Control, vol. K, July 1996, pp. 319–324.

[14] D. A. White and D. A. Sofge, eds., *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches*, Van Nostrand and Reinhold, 1993.

[15] A. Willsky, *A survey of design methods for failure detection in dynamic systems*, Automatica, 12 (1976), pp. 601–611.