Proceedings of the 7th Mediterranean Conference on Control and Automation (MED99) Haifa, Israel - June 28-30, 1999

Residual-Sensitive Fault Detection Filter

Robert H. Chen and Jason L. Speyer *
Mechanical and Aerospace Engineering Department
University of California, Los Angeles, California 90095-1597 †

Abstract

A fault detection and identification algorithm, called the residual-sensitive fault detection filter, is presented. The objective of the filter is to monitor certain faults called target faults and block other faults which are called nuisance faults. This filter is derived from solving a min-max problem which makes the residual sensitive to the target fault, but insensitive to the nuisance faults. It is shown that this filter approximates the properties of the classical fault detection filter such that in the limit where the weighting on the nuisance faults is zero, the residual-sensitive fault detection filter is equivalent to the unknown input observer and there exists a reduced-order filter. Fault detection filter designs can be obtained for both linear time-invariant and time-varying systems.

1 Introduction

Any system under automatic control demands a high degree of system reliability and this requires a health monitoring system capable of detecting any system, actuator and sensor fault as it occurs and identifying the faulty component. One approach, analytical redundancy which reduces the need for hardware redundancy, uses a modeled dynamic relationship between system inputs and measured system outputs to form a residual process used for detecting and identifying faults. Nominally, the residual is nonzero only when a fault has occurred and is zero at other times.

A popular approach to analytical redundancy is the detection filter which was first introduced by (Beard, 1971) and refined by (Jones, 1973). It is also known as the Beard-Jones fault detection filter. A geometric interpretation of this filter is given in (Massoumnia, 1986). Design algorithms have been developed (White and Speyer, 1987; Douglas and Speyer, 1996, 1999) which improved detection filter robustness. The idea of a detection filter is to put the reachable subspace of each fault into invariant subspaces which do not overlap with each other. Then, when a nonzero residual is detected, a fault can be announced and identified by projecting the residual onto each of the invariant subspaces. Therefore, multiple faults can be monitored in one filter.

Another related approach, the unknown input observer (Massoumnia *et al.*, 1989), simplifies the detection filter problem by dividing the faults into a target fault and nuisance fault group where the nuisance faults are placed into one unobservable subspace. Although only one fault can be detected in each unknown input observer, additional flexibility in fault detection filter design for robustness and time-varying system is obtained by using an approximate fault detection filter (Chung and Speyer, 1998; Lee, 1994; Brinsmead *et al.*, 1997; Chen and Speyer, 1999).

*Email: chrobert@talus.seas.ucla.edu, speyer@seas.ucla.edu

†This work was sponsored by the California Department of Transportation, Agreement No. 65A0013, MOU 315 and Air Force Office of Scientific Research, Award No. F49620-97-1-0272

In this paper, the residual-sensitive fault detection filter, motivated by (Chen and Speyer, 1999; Chung and Speyer, 1998) is presented. In (Chen and Speyer, 1999), it is shown that the optimal stochastic fault detection filter has the same properties as the unknown input observer in the limit where the disturbance attenuation bound is zero. However, the limiting filter can not be determined directly from this formulation because as the sensor noise variance goes to zero, the filter gain depends on its inverse. In contrast, the game-theoretic fault detection filter in (Chung and Speyer, 1998) can be derived in the limit by using the result from singular optimal control (Bell and Jacobsen, 1975). Also, these two approaches yield the same estimator equations with similar but not identical Riccati equations when it is not in the limit. Therefore, a problem, which is similar to the game-theoretic fault detection filter, but has the same result as optimal stochastic fault detection filter, is formulated and solved. The result is called the residual-sensitive fault detection filter. This filter, which is derived similarly to (Chung and Speyer, 1998), is an approximation of the unknown input observer. Many properties obtained in (Chung and Speyer, 1998) also apply to this filter. However, some new properties are given. For example, the target fault direction is now explicitly in the filter gain calculations. This provides a mechanism for enhancing the sensitivity of the filter to the target fault. Furthermore, the projector, which annihilates the residual direction associated with the nuisance faults and was assumed in the problem formulation of (Chung and Speyer, 1998), is not required to determine the filter gain. Finally, the nuisance faults were generalized to include the invariant zero directions for the time-invariant system. It is also shown that this filter has a minimal (C, A)-unobservability subspace for the nuisance faults and is equivalent to the unknown input observer in the limit. However, the filter gains in the limit are the result of solving a worst-case-design detection filter.

The problem is formulated in Section 2 and the solution is derived in Section 3. In Section 4, some conditions for this problem have been derived by using a linear matrix inequality (Chung and Speyer, 1998). In Section 5, the filter is derived for the limiting case (Bell and Jacobsen, 1975; Chung and Speyer, 1998). In Section 6, it is shown that the nuisance faults have been put into an invariant subspace in the limit. For the time-invariant system, this subspace is the minimal (C, A)-unobservability subspace, just like the unknown input observer. In Section 7, for the time-invariant system, a reduced-order filter is derived in the limit (Chung and Speyer, 1998). In Section 8, numerical examples are given.

2 Problem Formulation

Consider a linear system,

$$\dot{x} = Ax + Bu \tag{1a}$$
$$y = Cx \tag{1b}$$

where system matrices A, B and C can be time-varying and all system variables belong to real vector spaces $x \in \mathcal{X}$, $u \in \mathcal{U}$ and $y \in \mathcal{Y}$ with $n = \dim \mathcal{X}$ and $m = \dim \mathcal{Y}$. From (Beard, 1971; Massoumnia, 1986; White and Speyer, 1987; Chung and Speyer, 1998), any plant, actuator and sensor fault can be modeled as an additive term in the state equation (1a). Therefore, a linear system with q failure modes can be modeled by

$$\dot{x} = Ax + Bu + \sum_{i=1}^{q} \bar{F}_i \bar{\mu}_i \tag{2a}$$
$$y = Cx \tag{2b}$$

where $\bar{F}_i$ can be time-varying and $\bar{\mu}_i$ belong to real vector spaces $\bar{\mu}_i \in \bar{\mathcal{M}}_i$ with $\bar{q}_i = \dim \bar{\mathcal{M}}_i$. The failure modes $\bar{\mu}_i$ are vectors that are unknown and arbitrary functions of time and are zero when there is no failure. The failure signatures $\bar{F}_i : \bar{\mathcal{M}}_i \to \bar{\mathcal{F}}_i \subseteq \mathcal{X}$ are maps that are known. A failure mode $\bar{\mu}_i$ models the time-varying amplitude of a failure while a failure signature $\bar{F}_i$ models the directional characteristics of a failure. Assume the $\bar{F}_i$ are monic so that $\bar{\mu}_i \neq 0$ implies $\bar{F}_i\bar{\mu}_i \neq 0$. In this paper, the residual-sensitive fault detection filter is designed to detect only one fault and not to be affected by other faults. Therefore, let $\mu_1 = \bar{\mu}_1$ be the target fault and $\mu_2 = [\bar{\mu}_2^T, \cdots, \bar{\mu}_q^T]^T$ be the nuisance fault with $q_1 = \bar{q}_1$ and $q_2 = \sum_{i=2}^{q} \bar{q}_i$. Then, (2a) can be rewritten as

$$\dot{x} = Ax + Bu + F_1\mu_1 + F_2\mu_2 \tag{3a}$$

where $F_1 = \bar{F}_1$ and $F_2 = [\bar{F}_2, \cdots, \bar{F}_q]$. Since the state x is unknown, its determination from the min-max problem below produces the desired estimator. Therefore, the relation between the data sequence y and the state (2b) should be modified as

$$y = Cx + v \tag{3b}$$

where v can be considered a contrivance which will be made large to detect the occurrence of μ_1 and small to ensure that μ_2 is not observed. In fact, it will turn out to be the residual process. The mechanism for explicitly doing this is given by the formulation of the differential game and the associated cost criterion given below.

There are two assumptions about the system (3). The first one ensures the separation of faults μ_1 and μ_2 (Massoumnia, 1986; Chung and Speyer, 1998). The second one ensures a nonzero residual in steady state when target fault μ_1 occurs (Chen and Speyer, 1999).

Assumption 2.1. F_1 and F_2 are output separable.

Assumption 2.2. (C, A, F_1) does not have transmission zero at origin.

The objective of blocking the nuisance fault while detecting target fault can be achieved by solving the following min-max problem,

$$\min_{\mu_1} \max_{\mu_2} \max_{x(t_0)} J \tag{4}$$

where

$$J = \int_{t_0}^{t_1} (\| \mu_1 \|^2_{Q_1^{-1}} - \| \mu_2 \|^2_{\bar{Q}_2^{-1}} - \| y - Cx \|^2_{\bar{V}^{-1}})dt - \| x(t_0) - \hat{x}_0 \|^2_{\Pi_0} \tag{5}$$

subject to (3a). The current time is t_1 and y, the measurement, is assumed given. Q_1, $\bar{Q}_2$, $\bar{V}$ and Π_0 are positive definite weightings. Let $\bar{Q}_2^{-1} = \gamma Q_2^{-1}$ where γ is a small positive scalar and Q_2^{-1} is positive definite because the interest will be on small $\bar{Q}_2^{-1}$. The interpretation of (5) is that μ_1 tries to make the residual, $y - Cx$, large and μ_2, $x(t_0)$ try to make $y - Cx$ small. Therefore, the fault detection and identification is achieved by blocking the nuisance fault μ_2 from the residual while retaining the transmission from the target fault μ_1. As discussed in Sections 5 and 6, when γ becomes smaller, the residual is affected less by the nuisance fault and in fact when γ goes to zero, the problem becomes singular and the nuisance fault is completely blocked from the residual. In fact, the cost criterion could be viewed as being derived from a disturbance attenuation problem with a bound γ if $\bar{V}^{-1} = \gamma V^{-1}$ and $\Pi_0 = \gamma P_0^{-1}$ where V^{-1}

and P_0^{-1} would be the assumed weighting on measurement and initial condition uncertainties in the disturbance attenuation problem, respectively. In the limit, the weightings V and P_0 go to zero such that $\bar{V}$ and Π_0 remain finite and positive definite. This again illustrates that v is a construct, designed to produce a robust fault detection filter with limits consistent with previous developments (Massoumnia *et al.*, 1989). Also, note that the filter derived from worst case design should still work well when the faults μ_1 and μ_2 are different from their optimal strategies because the fault detection filter construction is based essentially on the direction instead of the magnitude of the faults. In Section 6, it is shown that, for time-invariant system, the nuisance fault μ_2 has been put into its minimal (C, A)-unobservability subspace $\mathcal{T}_2$ in the limit and similarly for the time-varying system. Therefore, it does not matter what μ_2 is, the residual will not be affected by μ_2 at all. Also, since μ_2 is in the unobservability subspace $\mathcal{T}_2$, the residual would naturally be

$$r = \hat{H}(y - C\hat{x}) \tag{6}$$

where $\hat{x}$ is the state estimate of the fault detection filter and

$$\hat{H} : \mathcal{Y} \to \mathcal{Y} \ \ , \ \ \text{Ker } \hat{H} = C\mathcal{T}_2 \ \ , \ \ \hat{H} = I - C\mathcal{T}_2[(C\mathcal{T}_2)^T C\mathcal{T}_2]^{-1}(C\mathcal{T}_2)^T$$

Remark 1. The differential game solved for the game-theoretic fault detection filter in (Chung and Speyer, 1998) is

$$\min_{\hat{x}} \max_{\mu_2} \max_{y} \max_{x(t_0)} J$$

where

$$J = \int_{t_0}^{t_1} (\| \hat{H}C(x - \hat{x}) \|_Q^2 - \| \mu_2 \|^2_{\gamma M^{-1}} - \| y - Cx \|^2_{\gamma V^{-1}}) dt - \| x(t_0) - \hat{x}_0 \|^2_{\gamma P_0^{-1}}$$

subject to

$$\dot{x} = Ax + Bu + F_2\mu_2$$

Note that the target fault μ_1 is not included in the differential game and the system. Also, the projector $\hat{H}$ is defined apriori.

3 Solution

In this section, the min-max problem (4) will be solved. The variational Hamitonian of the problem is

$$\mathcal{H} = \| \mu_1 \|^2_{Q_1^{-1}} - \| \mu_2 \|^2_{\gamma Q_2^{-1}} - \| y - Cx \|^2_{\bar{V}^{-1}} + \lambda^T (Ax + Bu + F_1\mu_1 + F_2\mu_2)$$

where $\lambda(t) \in \mathcal{R}^n$ is a continuously differentiable Lagrange multiplier. Then, take the first-order variation with respect to μ_1, μ_2 and x, respectively. The first-order necessary conditions imply that (Bryson and Ho, 1975) optimal strategies and dynamics for the Lagrange multiplier are

$$\mu_1^* = -Q_1 F_1^T \lambda \tag{7a}$$

$$\mu_2^* = \frac{1}{\gamma} Q_2 F_2^T \lambda \tag{7b}$$

$$\dot{\lambda} = -A^T\lambda - C^T\bar{V}^{-1}(y - Cx) \tag{7c}$$

with boundary conditions,

$$\lambda(t_0) = \Pi_0[x^*(t_0) - \hat{x}_0] \tag{7d}$$
$$\lambda(t_1) = 0 \tag{7e}$$

By substituting (7a) and (7b) into (3a) and combining with (7c), the two-point boundary value problem requires the solution to

$$\begin{bmatrix} \dot{x}^* \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} A & \frac{1}{\gamma}F_2Q_2F_2^T - F_1Q_1F_1^T \\ C^T\bar{V}^{-1}C & -A^T \end{bmatrix} \begin{bmatrix} x^* \\ \lambda \end{bmatrix} + \begin{bmatrix} Bu \\ -C^T\bar{V}^{-1}y \end{bmatrix} \tag{8}$$

with boundary conditions (7d) and (7e). Note that x^* is now the state using the optimal strategies (7a) and (7b). The form of (7d) suggests that

$$\lambda = \Pi(x^* - \hat{x}) \tag{9}$$

where $\Pi(t_0) = \Pi_0, \hat{x}(t_0) = \hat{x}_0$ and $\hat{x}$ is an intermediate state. Differentiate (9) and by using (8) and (9),

$$\begin{aligned} 0 = {} & [\dot{\Pi} + \Pi A + A^T\Pi + \Pi(\frac{1}{\gamma}F_2Q_2F_2^T - F_1Q_1F_1^T)\Pi - C^T\bar{V}^{-1}C]x^* \\ & - [\dot{\Pi} + A^T\Pi + \Pi(\frac{1}{\gamma}F_2Q_2F_2^T - F_1Q_1F_1^T)\Pi]\hat{x} - \Pi\dot{\hat{x}} + \Pi Bu + C^T\bar{V}^{-1}y \end{aligned}$$

Add and subtract $\Pi A\hat{x}$ and $C^T\bar{V}^{-1}C\hat{x}$,

$$\begin{aligned} 0 = {} & [\dot{\Pi} + \Pi A + A^T\Pi + \Pi(\frac{1}{\gamma}F_2Q_2F_2^T - F_1Q_1F_1^T)\Pi - C^T\bar{V}^{-1}C](x^* - \hat{x}) \\ & - \Pi\dot{\hat{x}} + \Pi A\hat{x} + \Pi Bu + C^T\bar{V}^{-1}(y - C\hat{x}) \end{aligned}$$

Therefore, (9) is a solution to (8) if

$$-\dot{\Pi} = \Pi A + A^T\Pi + \Pi(\frac{1}{\gamma}F_2Q_2F_2^T - F_1Q_1F_1^T)\Pi - C^T\bar{V}^{-1}C \ \ , \ \ \Pi(t_0) = \Pi_0 \tag{10}$$
$$\Pi\dot{\hat{x}} = \Pi A\hat{x} + \Pi Bu + C^T\bar{V}^{-1}(y - C\hat{x}) \ \ , \ \ \hat{x}(t_0) = \hat{x}_0 \tag{11}$$

Substitute the μ_1^* (7a), μ_2^* (7b) and (9) into the cost J (5),

$$J^* = \int_{t_0}^{t_1} (- \| x^* - \hat{x} \|^2_{\Pi(\frac{1}{\gamma}F_2Q_2F_2^T - F_1Q_1F_1^T)\Pi} - \| y - Cx^* \|^2_{\bar{V}^{-1}})dt - \| x^*(t_0) - \hat{x}_0 \|^2_{\Pi_0}$$

Add the identical zero term $\| x^*(t_0) - \hat{x}_0 \|^2_{\Pi(t_0)} - \| x^*(t_1) - \hat{x}(t_1) \|^2_{\Pi(t_1)} + \int_{t_0}^{t_1} \frac{d}{dt} \| x^* - \hat{x} \|^2_{\Pi} \, dt = 0$ to J^*,

$$\begin{aligned} J^* = \int_{t_0}^{t_1} & [- \| x^* - \hat{x} \|^2_{\Pi(\frac{1}{\gamma}F_2Q_2F_2^T - F_1Q_1F_1^T)\Pi} - \| y - Cx^* \|^2_{\bar{V}^{-1}} \\ & + (\Pi\dot{x}^* - \Pi\dot{\hat{x}})^T(x^* - \hat{x}) + (x^* - \hat{x})^T\dot{\Pi}(x^* - \hat{x}) + (x^* - \hat{x})^T(\Pi\dot{x}^* - \Pi\dot{\hat{x}})]dt \end{aligned}$$

Note that $\| x^*(t_1) - \hat{x}(t_1) \|^2_{\Pi(t_1)} = 0$ because of the boundary condition (7e). Substitute $\dot{x}^*$ from (8), (10) and (11) into J^*, J^* can be reduced to

$$J^* = \int_{t_0}^{t_1} - \| y - C\hat{x} \|^2_{\bar{V}^{-1}} \, dt$$

Therefore, the optimal filter is (11) subject to (10).

Remark 2. The Riccati equation for the game-theoretic fault detection filter (Chung and Speyer, 1998) is

$$-\dot{\Pi} = \Pi A + A^T \Pi + \frac{1}{\gamma} \Pi F_2 M F_2^T \Pi + C^T H^T Q H C - C^T \bar{V}^{-1} C$$

and the filter is the same as (11).

Remark 3. The filter (11) and the Riccati equation (10) are the same as the optimal stochastic fault detection filter (Chen and Speyer, 1999).

The sufficient condition for (11) to be stable is given in theorem 3.1.

Theorem 3.1. The optimal filter (11) is stable if

$$\frac{1}{\gamma} F_2 Q_2 F_2^T - F_1 Q_1 F_1^T \geq 0$$

Proof. The stability of (11) depends on the eigenvalues of the closed-loop A matrix,

$$A_{cl} = A - \Pi^{-1} C^T \bar{V}^{-1} C$$

Substitute this into (10),

$$\dot{\Pi} + \Pi A_{cl} + A_{cl}^T \Pi = -\Pi(\frac{1}{\gamma} F_2 Q_2 F_2^T - F_1 Q_1 F_1^T)\Pi - C^T \bar{V}^{-1} C$$

If $\dot{\Pi} + \Pi A_{cl} + A_{cl}^T \Pi < 0$, Lyapunov's Stability Theorem implies the eigenvalues of A_{cl} are in the open left-half plane. Since Π and $C^T \bar{V}^{-1} C$ are positive definite, the sufficient condition for $\dot{\Pi} + \Pi A_{cl} + A_{cl}^T \Pi < 0$ is $\frac{1}{\gamma} F_2 Q_2 F_2^T - F_1 Q_1 F_1^T \geq 0$.

Note that Q_1 is a free parameter which can be tuned to meet this sufficient condition.

Remark 4. The stability of the residual-sensitive fault detection filter depends on the target fault's weighting Q_1 because the optimization problem is trying to make the residual sensitive to the target fault μ_1 and sometimes μ_1 might destabilize the filter. However, the game-theoretic fault detection filter (Chung and Speyer, 1998) does not have this concern because the target fault is not in the problem formulation.

4 Conditions for the Nonpositivity of the Cost

In this section, the cost (5) is converted into an equivalent linear matrix inequality. Sufficient conditions for optimality for the singular control problem can be derived from the linear matrix inequality for the limiting case. The linear matrix inequality, associated with the solution optimality, is just the left half of the saddle point inequality,

$$J(\mu_1^*, \mu_2, x(t_0), v) \leq J(\mu_1^*, \mu_2^*, x(t_0)^*, v^*) = 0 \leq J(\mu_1, \mu_2^*, x(t_0)^*, v^*)$$

The asterisk indicates that the optimal strategy is being used for that element.

Substitute the optimal μ_1^* (7a) into the cost J (5) and adjoin with the constraint (3a) by a Lagrange multiplier, $(x-\hat{x})^T\Pi$,

$$J = \int_{t_0}^{t_1} [\| x-\hat{x} \|^2_{\Pi F_1 Q_1 F_1^T \Pi} - \| \mu_2 \|^2_{\gamma Q_2^{-1}} - \| y - Cx \|^2_{\bar{V}^{-1}} + (x-\hat{x})^T \Pi (Ax + Bu + F_1\mu_1 + F_2\mu_2 - \dot{x})]dt - \| x(t_0) - \hat{x}_0 \|^2_{\Pi_0}$$

Add and subtract $(x-\hat{x})^T\Pi A\hat{x}$ and $(x-\hat{x})^T\Pi\dot{\hat{x}}$ to J,

$$J = \int_{t_0}^{t_1} [\| x-\hat{x} \|^2_{\Pi A} - \| \mu_2 \|^2_{\gamma Q_2^{-1}} - \| y - Cx \|^2_{\bar{V}^{-1}} + (x-\hat{x})^T \Pi F_2 \mu_2 - (x-\hat{x})^T\Pi(\dot{x}-\dot{\hat{x}}) + (x-\hat{x})^T(\Pi A\hat{x} + \Pi Bu - \Pi\dot{\hat{x}})]dt - \| x(t_0) - \hat{x}_0 \|^2_{\Pi_0}$$

Integrate $(x-\hat{x})^T\Pi(\dot{x}-\dot{\hat{x}})$ by parts and substitute (3a) into J. Then, add and subtract $\hat{x}^T A^T \Pi (x-\hat{x})$ to J,

$$\begin{aligned} J = \int_{t_0}^{t_1} [&\| x-\hat{x} \|^2_{\dot{\Pi}+\Pi A + A^T\Pi - \Pi F_1 Q_1 F_1^T \Pi} - \| \mu_2 \|^2_{\gamma Q_2^{-1}} - \| y - Cx \|^2_{\bar{V}^{-1}} + (x-\hat{x})^T \Pi F_2 \mu_2 \\ &+ \mu_2^T F_2^T \Pi (x-\hat{x}) + (x-\hat{x})^T(\Pi A\hat{x} + \Pi Bu - \Pi\dot{\hat{x}}) + (\Pi A\hat{x} + \Pi Bu - \Pi\dot{\hat{x}})^T(x-\hat{x})]dt \\ &- \| x(t_0) - \hat{x}_0 \|^2_{\Pi_0 - \Pi(t_0)} - \| x(t_1) - \hat{x}(t_1) \|^2_{\Pi(t_1)} \end{aligned}$$

By expanding $\| y - Cx \|^2_{\bar{V}^{-1}}$ into $\| (y - C\hat{x}) - C(x-\hat{x}) \|^2_{\bar{V}^{-1}}$,

$$\begin{aligned} J = \int_{t_0}^{t_1} [&\| x-\hat{x} \|^2_{\dot{\Pi}+\Pi A + A^T\Pi - \Pi F_1 Q_1 F_1^T \Pi - C^T\bar{V}^{-1}C} - \| \mu_2 \|^2_{\gamma Q_2^{-1}} - \| y - C\hat{x} \|^2_{\bar{V}^{-1}} \\ &+ (x-\hat{x})^T \Pi F_2 \mu_2 + \mu_2^T F_2^T \Pi (x-\hat{x}) \\ &+ (x-\hat{x})^T(-\Pi\dot{\hat{x}} + \Pi A\hat{x} + \Pi Bu + C^T\bar{V}^{-1}(y - C\hat{x})) \\ &+ (-\Pi\dot{\hat{x}} + \Pi A\hat{x} + \Pi Bu + C^T\bar{V}^{-1}(y - C\hat{x}))^T(x-\hat{x})]dt \\ &- \| x(t_0) - \hat{x}_0 \|^2_{\Pi_0 - \Pi(t_0)} - \| x(t_1) - \hat{x}(t_1) \|^2_{\Pi(t_1)} \end{aligned}$$

Substitute the optimal filter (11),

$$\begin{aligned} J = \int_{t_0}^{t_1} \{ &\begin{bmatrix} (x-\hat{x})^T & \mu_2^T \end{bmatrix} \begin{bmatrix} \dot{\Pi}+\Pi A + A^T\Pi - \Pi F_1 Q_1 F_1^T \Pi - C^T\bar{V}^{-1}C & \Pi F_2 \\ F_2^T \Pi & -\gamma Q_2^{-1} \end{bmatrix} \begin{bmatrix} x-\hat{x} \\ \mu_2 \end{bmatrix} \\ &- \| y - C\hat{x} \|^2_{\bar{V}^{-1}} \} dt - \| x(t_0) - \hat{x}_0 \|^2_{\Pi_0 - \Pi(t_0)} - \| x(t_1) - \hat{x}(t_1) \|^2_{\Pi(t_1)} \end{aligned}$$

Therefore, the sufficient conditions for $J \leq 0$ are

$$\begin{aligned} &\begin{bmatrix} \dot{\Pi}+\Pi A + A^T\Pi - \Pi F_1 Q_1 F_1^T \Pi - C^T\bar{V}^{-1}C & \Pi F_2 \\ F_2^T \Pi & -\gamma Q_2^{-1} \end{bmatrix} \leq 0 \\ &\Pi_0 - \Pi(t_0) \geq 0 \\ &\Pi(t_1) \geq 0 \end{aligned}$$

The sufficient conditions for $J \leq 0$ in the limit where $\gamma \to 0$ are

$$\Pi F_2 = 0 \tag{12a}$$

$$\dot{\Pi}+\Pi A + A^T\Pi - \Pi F_1 Q_1 F_1^T \Pi - C^T\bar{V}^{-1}C \leq 0 \tag{12b}$$

More details about the limit will be discussed in next section.

5 Limiting Case

In this section, the min-max problem (4) is solved in the limit where $\gamma \to 0$. In the limit, the cost J (5) becomes

$$J = \int_{t_0}^{t_1} (\| \mu_1 \|^2_{Q_1^{-1}} - \| y - Cx \|^2_{\bar{V}^{-1}}) dt - \| x(t_0) - \hat{x}_0 \|^2_{\Pi_0} \tag{13}$$

This is a singular problem with respect to μ_2. Therefore, the Goh transformation has been used to make it become nonsingular. Let

$$\phi_1(t) = \int_{t_0}^{t} \mu_2(\tau) d\tau \tag{14a}$$

$$\alpha_1 = x - F_2 \phi_1 \tag{14b}$$

Differentiate (14b) and by using (3a),

$$\dot{\alpha}_1 = A\alpha_1 + Bu + F_1 \mu_1 + B_1 \phi_1 \tag{15}$$

where $B_1 = AF_2 - \dot{F}_2$. Substitute (14b) into the limiting cost (13),

$$\begin{aligned} J = \int_{t_0}^{t_1} [\| \mu_1 \|^2_{Q_1^{-1}} - \| \phi_1 \|^2_{F_2^T C^T \bar{V}^{-1} C F_2} - \| y - C\alpha_1 \|^2_{\bar{V}^{-1}} + (y - C\alpha_1)^T \bar{V}^{-1} C F_2 \phi_1 \\ + \phi_1^T F_2^T C^T \bar{V}^{-1} (y - C\alpha_1)] dt - \| \alpha_1(t_0^+) + F_2 \phi_1(t_0^+) - \hat{x}_0 \|^2_{\Pi_0} \end{aligned} \tag{16}$$

Then, the new min-max problem is

$$\min_{\mu_1} \max_{\phi_1} \max_{\alpha_1(t_0^+)} J \tag{17}$$

subject to (15).

The variational Hamitonian of the problem is

$$\begin{aligned} \mathcal{H} = \| \mu_1 \|^2_{Q_1^{-1}} - \| \phi_1 \|^2_{F_2^T C^T \bar{V}^{-1} C F_2} - \| y - C\alpha_1 \|^2_{\bar{V}^{-1}} + (y - C\alpha_1)^T \bar{V}^{-1} C F_2 \phi_1 \\ + \phi_1^T F_2^T C^T \bar{V}^{-1} (y - C\alpha_1) + \lambda^T (A\alpha_1 + Bu + F_1 \mu_1 + B_1 \phi_1) \end{aligned}$$

where $\lambda(t) \in \mathcal{R}^n$ is a continuously differentiable Lagrange multiplier. Then, take the first-order variation with respect to μ_1, ϕ_1 and α_1, respectively.

$$\mu_1^* = -Q_1 F_1^T \lambda \tag{18a}$$

$$\phi_1^* = (F_2^T C^T \bar{V}^{-1} C F_2)^{-1} [B_1^T \lambda + F_2^T C^T \bar{V}^{-1} (y - C\alpha_1)] \tag{18b}$$

$$\dot{\lambda} = -A^T \lambda - C^T \bar{V}^{-1} (y - C\alpha_1)^T + C^T \bar{V}^{-1} C F_2 \phi_1^* \tag{18c}$$

and

$$\phi_1^*(t_0^+) = -(F_2^T \Pi_0 F_2)^{-1} F_2^T \Pi_0 [\alpha_1^*(t_0^+) - \hat{x}_0] \tag{18d}$$

$$\lambda(t_0^+) = \Pi_0 [\alpha_1^*(t_0^+) + F_2 \phi_1^*(t_0^+) - \hat{x}_0] \tag{18e}$$

$$\lambda(t_1) = 0 \tag{18f}$$

Substitute (18d) into (18e),

$$\lambda(t_0^+) = [\Pi_0 - \Pi_0 F_2 (F_2^T \Pi_0 F_2)^{-1} F_2^T \Pi_0][\alpha_1^*(t_0^+) - \hat{x}_0] \tag{19}$$

By substituting (18a), (18b) into (15) and (18b) into (18c), a two-point boundary value problem with boundary conditions (19) and (18f) results for satisfying

$$\begin{aligned}\begin{bmatrix} \dot{\alpha}_1^* \\ \dot{\lambda} \end{bmatrix} &= \begin{bmatrix} \bar{A} & -F_1 Q_1 F_1^T + B_1 (F_2^T C^T \bar{V}^{-1} C F_2)^{-1} B_1^T \\ C^T \bar{H}^T \bar{V}^{-1} \bar{H} C & -\bar{A}^T \end{bmatrix} \begin{bmatrix} \alpha_1^* \\ \lambda \end{bmatrix} \\ &+ \begin{bmatrix} Bu + B_1 (F_2^T C^T \bar{V}^{-1} C F_2)^{-1} F_2^T C^T \bar{V}^{-1} y \\ -C^T \bar{H}^T \bar{V}^{-1} \bar{H} y \end{bmatrix}\end{aligned} \tag{20}$$

where $\bar{A} = A - B_1 (F_2^T C^T \bar{V}^{-1} C F_2)^{-1} F_2^T C^T \bar{V}^{-1} C$ and $\bar{H} = I - C F_2 (F_2^T C^T \bar{V}^{-1} C F_2)^{-1} F_2^T C^T \bar{V}^{-1}$. Note that α_1^* is now the state using the optimal strategies (18a) and (18b). The form of (19) suggests that

$$\lambda = S(\alpha_1^* - \hat{x}) \tag{21a}$$

$$S(t_0^+) = \Pi_0 - \Pi_0 F_2 (F_2^T \Pi_0 F_2)^{-1} F_2^T \Pi_0 \tag{21b}$$

$$\hat{x}(t_0^+) = \hat{x}_0 \tag{21c}$$

where $\hat{x}$ is an intermediate state. Differentiate (21a) and by using (20) and (21a), the following dynamic filter structure results

$$\begin{aligned}-\dot{S} &= S[A - B_1 (F_2^T C^T \bar{V}^{-1} C F_2)^{-1} F_2^T C^T \bar{V}^{-1} C] + [A - B_1 (F_2^T C^T \bar{V}^{-1} C F_2)^{-1} F_2^T C^T \bar{V}^{-1} C]^T S \\ &+ S[-F_1 Q_1 F_1^T + B_1 (F_2^T C^T \bar{V}^{-1} C F_2)^{-1} B_1^T] S - C^T \bar{H}^T \bar{V}^{-1} \bar{H} C\end{aligned} \tag{22}$$

$$S\dot{\hat{x}} = SA\hat{x} + SBu + [SB_1 (F_2^T C^T \bar{V}^{-1} C F_2)^{-1} F_2^T C^T \bar{V}^{-1} + C^T \bar{H}^T \bar{V}^{-1} \bar{H}](y - C\hat{x}) \tag{23}$$

subject to (21b) and (21c).

By substituting the optimal μ_1^* (18a), ϕ_1^* (18b), $\phi_1^*(t_0^+)$ (18d) and (21a) into the cost J (16), the cost becomes

$$\begin{aligned}J^* &= \int_{t_0}^{t_1} [\| \alpha_1^* - \hat{x} \|^2_{S[F_1 Q_1 F_1^T - B_1 (F_2^T C^T \bar{V}^{-1} C F_2)^{-1} B_1^T] S} - \| y - C\alpha_1^* \|^2_{\bar{H}^T \bar{V}^{-1} \bar{H}}] dt \\ &- \| \alpha_1^*(t_0^+) - \hat{x}_0 \|^2_{\Pi_0 - \Pi_0 F_2 (F_2^T \Pi_0 F_2)^{-1} F_2^T \Pi_0}\end{aligned}$$

Add the identical zero term $\| \alpha_1^*(t_0^+) - \hat{x}_0 \|^2_{S(t_0^+)} - \| \alpha_1^*(t_1) - \hat{x}(t_1) \|^2_{S(t_1)} + \int_{t_0}^{t_1} \frac{d}{dt} \| \alpha_1^* - \hat{x} \|^2_S$ $dt = 0$ to J^*,

$$\begin{aligned}J^* = \int_{t_0}^{t_1} & [\| \alpha_1^* - \hat{x} \|^2_{S[F_1 Q_1 F_1^T - B_1 (F_2^T C^T \bar{V}^{-1} C F_2)^{-1} B_1^T] S} - \| y - C\alpha_1^* \|^2_{\bar{H}^T \bar{V}^{-1} \bar{H}} \\ & + (S\dot{\alpha}_1^* - S\dot{\hat{x}})^T (\alpha_1^* - \hat{x}) + (\alpha_1^* - \hat{x})^T \dot{S} (\alpha_1^* - \hat{x}) + (\alpha_1^* - \hat{x})^T (S\dot{\alpha}_1^* - S\dot{\hat{x}})] dt\end{aligned}$$

Note that $\| \alpha_1^*(t_1) - \hat{x}(t_1) \|^2_{S(t_1)} = 0$ because of the boundary condition (18f). Substitute $\dot{\alpha}_1^*$ from (20), (22) and (23) into J^*, J^* can be reduced to

$$J^* = \int_{t_0}^{t_1} - \| y - C\hat{x} \|^2_{\bar{H}^T \bar{V}^{-1} \bar{H}} \, dt$$

Therefore, the optimal filter is (23) subject to (21c), (22) and (21b). Note that the second part of the optimization problem is solved differently as (Chung and Speyer, 1998) but is consistent with the derivation of the optimization problem solved in Section 3. Also, this makes the derivation clearer and more compact.

If $F_2^T C^T \bar{V}^{-1} C F_2$ fails to be positive definite, (17) is still a singular problem because it needs the inverse, $(F_2^T C^T \bar{V}^{-1} C F_2)^{-1}$. Then, the Goh transformation has to be used until it is nonsingular. There are two types of singularity.

1. If $F_2^T C^T \bar{V}^{-1} C F_2 = 0$, repeat the Goh transformation,

$$\begin{aligned} \phi_2(t) &= \int_{t_0}^{t} \phi_1(\tau) d\tau \\ \alpha_2 &= \alpha_1 - B_1 \phi_2 \\ \dot{\alpha}_2 &= A\alpha_2 + Bu + F_1\mu_1 + B_2\phi_2 \\ B_2 &= AB_1 - \dot{B}_1 \end{aligned}$$

2. If $F_2^T C^T \bar{V}^{-1} C F_2 \geq 0$, the Goh transformation is applied only on the singular part.

The transformation process is stopped if $B_1^T C^T \bar{V}^{-1} C B_1$ is positive definite. Otherwise, continue the transformation as above until there exists B_k such that $B_{k-1}^T C^T \bar{V}^{-1} C B_{k-1}$ is positive definite and the min-max problem (4) becomes

$$\min_{\mu_1} \max_{\phi_k} \max_{\alpha_k(t_0^+)} J$$

where

$$\begin{aligned} J = \int_{t_0}^{t_1} [\| \mu_1 \|^2_{Q_1^{-1}} - \| \phi_k \|^2_{B_{k-1}^T C^T \bar{V}^{-1} C B_{k-1}} - \| y - C\alpha_k \|^2_{\bar{V}^{-1}} + (y - C\alpha_k)^T \bar{V}^{-1} C B_{k-1} \phi_k \\ + \phi_k^T B_{k-1}^T C^T \bar{V}^{-1} (y - C\alpha_k)] dt - \| \alpha_k(t_0^+) + \bar{B}\bar{\phi}(t_0^+) - \hat{x}_0 \|^2_{\Pi_0} \end{aligned}$$

and $\bar{B} = [F_2 \;\; B_1 \;\; B_2 \;\; \cdots \;\; B_{k-1}]$, $\bar{\phi} = [\phi_1^T \;\; \phi_2^T \;\; \cdots \;\; \phi_k^T]^T$ and subject to

$$\dot{\alpha}_k = A\alpha_k + Bu + F_1\mu_1 + B_k\phi_k$$

Following the same procedure, the solution to this problem is

$$S\dot{\hat{x}} = SA\hat{x} + SBu + [SB_k(B_{k-1}^T C^T \bar{V}^{-1} C B_{k-1})^{-1} B_{k-1}^T C^T \bar{V}^{-1} + C^T \bar{H}^T \bar{V}^{-1} \bar{H}](y - C\hat{x}) \tag{24a}$$

$$\hat{x}(t_0^+) = \hat{x}_0$$

$$\begin{aligned} -\dot{S} = {} & S[A - B_k(B_{k-1}^T C^T \bar{V}^{-1} C B_{k-1})^{-1} B_{k-1}^T C^T \bar{V}^{-1} C] \\ & + [A - B_k(B_{k-1}^T C^T \bar{V}^{-1} C B_{k-1})^{-1} B_{k-1}^T C^T \bar{V}^{-1} C]^T S \\ & + S[-F_1 Q_1 F_1^T + B_k(B_{k-1}^T C^T \bar{V}^{-1} C B_{k-1})^{-1} B_k^T] S - C^T \bar{H}^T \bar{V}^{-1} \bar{H} C \end{aligned} \tag{24b}$$

$$S(t_0^+) = \Pi_0 - \Pi_0 \bar{B}(\bar{B}^T \Pi_0 \bar{B})^{-1} \bar{B}^T \Pi_0$$

where $\bar{H} = I - CB_{k-1}(B_{k-1}^T C^T \bar{V}^{-1} C B_{k-1})^{-1} B_{k-1}^T C^T \bar{V}^{-1}$.

Theorem 5.1 shows that the solution (24b) satisfies the sufficient conditions (12) in Section 4 and this implies that S is the limit of Π.

Theorem 5.1.

$$\begin{gathered} S\begin{bmatrix} B_{k-1} & B_{k-2} & \cdots & B_1 & F_2 \end{bmatrix} = 0 \\ \dot{S} + SA + A^T S - SF_1 Q_1 F_1^T S - C^T \bar{V}^{-1} C \leq 0 \end{gathered}$$

Proof. Multiply (24b) by B_{k-1} from the right and subtract $S\dot{B}_{k-1}$ from both sides,

$$\frac{d}{dt}(SB_{k-1}) = -[A^T - SF_1Q_1F_1^T + (SB_k - C^T\bar{V}^{-1}CB_{k-1})(B_{k-1}^TC^T\bar{V}^{-1}CB_{k-1})^{-1}B_k^T]SB_{k-1}$$

It is a homogeneous differential equation and its boundary condition is zero because $S(t_0^+)$ is a projector that maps $\bar{B}$ to zero and B_{k-1} is contained in $\bar{B}$. Therefore $SB_{k-1} = 0$. Similarly, multiply (24b) by B_{k-2} from the right and subtract $S\dot{B}_{k-2}$ from both sides,

$$\frac{d}{dt}(SB_{k-2}) = -[A^T - SF_1Q_1F_1^T + (SB_k - C^T\bar{V}^{-1}CB_{k-1})(B_{k-1}^TC^T\bar{V}^{-1}CB_{k-1})^{-1}B_k^T]SB_{k-2}$$

Therefore, $SB_{k-2} = 0$ because its boundary condition is zero. Iterate this procedure by using $B_{k-3}, \cdots, B_1, F_2$,

$$S\begin{bmatrix} B_{k-1} & B_{k-2} & \cdots & B_1 & F_2 \end{bmatrix} = 0$$

To prove the second part of this theorem, (24b) can be rewritten as

$$\begin{aligned} &\dot{S} + SA + A^TS - SF_1Q_1F_1^TS - C^T\bar{V}^{-1}C \\ &= -(B_k^TS - B_{k-1}^TC^T\bar{V}^{-1}C)^T(B_{k-1}^TC^T\bar{V}^{-1}CB_{k-1})^{-1}(B_k^TS - B_{k-1}^TC^T\bar{V}^{-1}C) \end{aligned}$$

and it is non-positive definite.

Remark 5. The filter (24a) and Riccati equation (24b) will also be the limiting solution for the optimal stochastic fault detection filter.

6 Properties of the null space of S

In this section, some properties of the null space of S are given. Theorem 6.1 shows that the null space of S is a $(C,\ A)$-invariant subspace. Theorem 6.2 shows that the null space of S contains the minimal $(C,\ A)$-invariant subspace of F_2. For the time-invariant system, after modifying the nuisance fault direction, the invariant zero directions of (C, A, F_2) are also in the null space of S. Therefore, $\operatorname{Ker} S = \mathcal{T}_2$ by the definition of minimal $(C,\ A)$-unobservability subspace (Massoumnia, 1986). The unknown input observer (Massoumnia *et al.*, 1989) and Beard-Jones fault detection filter also put the fault into a minimal $(C,\ A)$-unobservability subspace.

Theorem 6.1. $\operatorname{Ker} S$ is a (C, A)-invariant subspace.

Proof. Consider the 'state matrix' of the estimator, (24a),

$$A_{cl} = S[A - B_k(B_{k-1}^TC^T\bar{V}^{-1}CB_{k-1})^{-1}B_{k-1}^TC^T\bar{V}^{-1}C - C^T\bar{H}^T\bar{V}^{-1}\bar{H}C]$$

Since S is positive semi-definite with a nontrivial null space, there exists a state transformation Γ such that

$$\Gamma^TS(t)\Gamma = \begin{bmatrix} \bar{S}(t) & 0 \\ 0 & 0 \end{bmatrix} \tag{25}$$

where $\bar{S}(t)$ is symmetric and positive definite. Multiply A_{cl} by Γ^T from the left and Γ from the right.

$$\begin{aligned}\Gamma^T A_{cl}\Gamma &= \begin{bmatrix} \bar{S} & 0 \\ 0 & 0 \end{bmatrix} \Gamma^{-1}[A - B_k(B_{k-1}^T C^T \bar{V}^{-1} C B_{k-1})^{-1} B_{k-1}^T C^T \bar{V}^{-1} C - C^T \bar{H}^T \bar{V}^{-1} \bar{H} C]\Gamma \\ &\triangleq \begin{bmatrix} \bar{S} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix} = \begin{bmatrix} \bar{S} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \bar{A}_{11} & 0 \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix}\end{aligned} \tag{26}$$

where $\bar{A}_{12} = 0$ is shown in (31) in Section 7. If the error e initially lies in $\operatorname{Ker} S$,

$$\Gamma^{-1}e = \begin{bmatrix} 0 \\ \bar{e} \end{bmatrix} \in \operatorname{Ker} \begin{bmatrix} \bar{S} & 0 \\ 0 & 0 \end{bmatrix}$$

In the absence of exogenous inputs, (26) implies that $\Gamma^{-1}e$ will be propagated by way of

$$\frac{d}{dt}(\Gamma^{-1}e) = \begin{bmatrix} \bar{S} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \bar{A}_{22}\bar{e} \end{bmatrix}$$

This shows the error will never leave $\operatorname{Ker} S$. Therefore, $\operatorname{Ker} S$ is a (C, A)-invariant subspace. ◀

Theorem 6.2. $\operatorname{Ker} S$ contains the minimal (C, A)-invariant subspace of F_2.

Proof. The minimal (C, A)-invariant subspace of F_2 is $B_{k-1}, B_{k-2} \cdots B_1, F_2$ (Massoumnia, 1986; Chung and Speyer, 1998). Therefore, the following is needed to be shown.

$$\begin{aligned} S \begin{bmatrix} B_{k-1} & B_{k-2} & \cdots & B_1 & F_2 \end{bmatrix} &= 0 \\ SB_k &\neq 0 \end{aligned}$$

Theorem 5.1 already shows the first part. To prove the second part, multiply (24b) by B_k^T from the left and B_k from the right,

$$\begin{aligned} -\frac{d}{dt}(B_k^T S B_k) &= -\dot{B}_k^T S B_k - B_k^T S \dot{B}_k + B_k^T S[A - B_k(B_{k-1}^T C^T \bar{V}^{-1} C B_{k-1})^{-1} B_{k-1}^T C^T \bar{V}^{-1} C] B_k \\ &+ B_k^T [A - B_k(B_{k-1}^T C^T \bar{V}^{-1} C B_{k-1})^{-1} B_{k-1}^T C^T \bar{V}^{-1} C]^T S B_k \\ &+ B_k^T S[-F_1 Q_1 F_1^T + B_k(B_{k-1}^T C^T \bar{V}^{-1} C B_{k-1})^{-1} B_k^T] S B_k - B_k^T C^T \bar{H}^T \bar{V}^{-1} \bar{H} C B_k \end{aligned}$$

If $SB_k = 0$, $\bar{H}CB_k = 0$ which is not true. Therefore, $SB_k \neq 0$ and $\operatorname{Ker} S$ contains the minimal (C, A)-invariant subspace of F_2. ◀

For time-invariant system, it is important to discuss the invariant zero directions of (C, A, F_2) when designing the fault detection filter. It is shown that the invariant zeros will be included with the eigenvalues of the fault detection filter if associated invariant zero directions are not included in the invariant subspace of F_2 (Massoumnia, 1986; Massoumnia *et al.*, 1989). Therefore, the null space of S includes at least the invariant zero directions associated with the invariant zeros on the right-half plane and $j\omega$-axis because the filter is stable which is shown in Section 7. However, the invariant zeros on the left-half plane might be part of the filter eigenvalues since there is no guarantee that associated invariant zero directions are in the null space of S. It is important that the filter can assign its eigenvalues freely because the invariant zeros might be ill-conditioned. This could be done by modifying the nuisance fault direction to enforce the null space of S to include the invariant zero directions. The invariant zero of (C, A, F_2) is defined by

$$\begin{bmatrix} zI - A & F_2 \\ C & 0 \end{bmatrix} \begin{bmatrix} \nu \\ w \end{bmatrix} = 0 \tag{27}$$

where z is the invariant zero and ν is the invariant zero direction. If the nuisance fault direction is changed to $[F_2 \;\; \nu]$, Assumption 2.1 in Section 2 is not satisfied. However, if the nuisance fault direction is changed to ν, the null space of S will be span$\{\nu \;\; A\nu \;\; A^2\nu \;\; \cdots \;\; A^k\nu\}$ which is equivalent to span$\{F_2 \;\; AF_2 \;\; \cdots \;\; A^{k-1}F_2 \;\; \nu\}$ according to (27). This guarantees that every invariant zero directions are in the null space of S. If (C, A, ν) has invariant zero, the same procedure above will be repeated until there is no invariant zero. From theorems 6.1, 6.2 and using modified nuisance fault direction, the null space of S is the minimal (C, A)-unobservability subspace. Note that the null space of S can be obtained apriori and it is fixed for the time-invariant system. This is very important for deriving the reduced-order filter in Section 7.

Remark 6. If the nuisance fault direction is not modified, the invariant zero directions associated with invariant zeros of (C, A, F_2) on the right-half plane and $j\omega$-axis should be included in the null space of $S(t_0^+)$ in (21b). This could be done by making the null space of Π_0 contain these invariant zero directions.

Remark 7. This modification of the nuisance fault direction can also apply to the game-theoretic fault detection filter and it makes the game-theoretic fault detection filter also equivalent to an unknown input observer.

7 Reduced-Order Filter

In this section, a reduced-order filter is derived for the limiting residual-sensitive fault detection filter (24a) for the time-invariant system. It is shown that this reduced-order filter can completely block the nuisance fault. Also, the sufficient condition for this filter to be stable is given.

Consider the state transformation Γ in (25), it can be computed apriori because $\operatorname{Ker} S$ can be obtained apriori in terms of the system matrices A, C and F_2. Note that Γ is not unique. Apply the transformation Γ to the system matrices,

$$\Gamma^{-1}A\Gamma = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \Gamma^{-1}B = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix} \quad C\Gamma = \begin{bmatrix} C_1 & C_2 \end{bmatrix}$$
$$\Gamma^{-1}F_1 = \begin{bmatrix} N_1 \\ N_2 \end{bmatrix} \qquad \Gamma^{-1}B_k = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \quad \Gamma^{-1}B_{k-1} = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix}$$

Since $SB_{k-1} = 0$, $\Gamma^T S\Gamma\Gamma^{-1}B_{k-1} = \bar{S}D_1 = 0$ implies $D_1 = 0$. Then,

$$\Gamma^{-1}B_{k-1} = \begin{bmatrix} 0 \\ D_2 \end{bmatrix}$$

Also, apply this transformation to the estimator state,

$$\hat{\eta} = \begin{bmatrix} \hat{\eta}_1 \\ \hat{\eta}_2 \end{bmatrix} = \Gamma^{-1}\hat{x}$$

to transform (24a) into

$$\begin{bmatrix} \bar{S} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \dot{\hat{\eta}}_1 \\ \dot{\hat{\eta}}_2 \end{bmatrix} = \begin{bmatrix} \bar{S} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} \hat{\eta}_1 \\ \hat{\eta}_2 \end{bmatrix} + \begin{bmatrix} \bar{S} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} M_1 \\ M_2 \end{bmatrix} u$$
$$+ \{ \begin{bmatrix} C_1^T \\ C_2^T \end{bmatrix} \bar{H}^T \bar{V}^{-1} \bar{H} + \begin{bmatrix} \bar{S} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} (\begin{bmatrix} 0 & D_2^T \end{bmatrix} \begin{bmatrix} C_1^T \\ C_2^T \end{bmatrix} \bar{V}^{-1} \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} 0 \\ D_2 \end{bmatrix})^{-1}$$
$$\begin{bmatrix} 0 & D_2^T \end{bmatrix} \begin{bmatrix} C_1^T \\ C_2^T \end{bmatrix} \bar{V}^{-1} \} (y - \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} \hat{\eta}_1 \\ \hat{\eta}_2 \end{bmatrix})$$

Then, it can be transformed into two equations,

$$\begin{aligned} \bar{S}\dot{\hat{\eta}}_1 &= \bar{S}A_{11}\hat{\eta}_1 + \bar{S}A_{12}\hat{\eta}_2 + \bar{S}M_1 u \\ &\quad + [\bar{S}G_1(D_2^T C_2^T \bar{V}^{-1} C_2 D_2)^{-1} D_2^T C_2^T \bar{V}^{-1} + C_1^T \bar{H}^T \bar{V}^{-1} \bar{H}](y - C_1\hat{\eta}_1 - C_2\hat{\eta}_2) && \text{(28a)} \\ 0 &= C_2^T \bar{H}^T \bar{V}^{-1} \bar{H}(y - C_1\hat{\eta}_1 - C_2\hat{\eta}_2) && \text{(28b)} \end{aligned}$$

From (28b),

$$\bar{H}C_2 = 0 \tag{29}$$

because $y - C_1\hat{\eta}_1 - C_2\hat{\eta}_2$ is arbitrary. Also, apply the transformation Γ on (24b),

$$\begin{aligned} 0 &= \bar{S}[A_{12} - G_1(D_2^T C_2^T \bar{V}^{-1} C_2 D_2)^{-1} D_2^T C_2^T \bar{V}^{-1} C_2] && \text{(30a)} \\ -\dot{\bar{S}} &= \bar{S}[A_{11} - G_1(D_2^T C_2^T \bar{V}^{-1} C_2 D_2)^{-1} D_2^T C_2^T \bar{V}^{-1} C_1] \\ &\quad + [A_{11} - G_1(D_2^T C_2^T \bar{V}^{-1} C_2 D_2)^{-1} D_2^T C_2^T \bar{V}^{-1} C_1]^T \bar{S} \\ &\quad + \bar{S}[-N_1 Q_1 N_1^T + G_1(D_2^T C_2^T \bar{V}^{-1} C_2 D_2)^{-1} G_1^T]\bar{S} - C_1^T \bar{H}^T \bar{V}^{-1} \bar{H} C_1 && \text{(30b)} \end{aligned}$$

From (30a),

$$A_{12} - G_1(D_2^T C_2^T \bar{V}^{-1} C_2 D_2)^{-1} D_2^T C_2^T \bar{V}^{-1} C_2 = 0 \tag{31}$$

because $\bar{S}$ is positive definite. Put (29) and (30a) into (28a), the reduced-order filter is

$$\dot{\hat{\eta}}_1 = A_{11}\hat{\eta}_1 + M_1 u + [G_1(D_2^T C_2^T \bar{V}^{-1} C_2 D_2)^{-1} D_2^T C_2^T \bar{V}^{-1} + \bar{S}^{-1} C_1^T \bar{H}^T \bar{V}^{-1} \bar{H}](y - C_1\hat{\eta}_1) \tag{32}$$

subject to (30b). The dimension of this reduced-order filter is $n - \dim(\operatorname{Ker} S) = n - \dim \mathcal{T}_2$ where $\mathcal{T}_2$ is the minimum (C, A)-unobservability subspace of F_2. The residual (6) becomes

$$r = \hat{H}(y - C_1\hat{\eta}_1)$$

Theorem 7.1 shows that the reduced-order filter (32) can completely block the nuisance fault. The sufficient condition for (32) to be stable is given in theorem 7.2 and the stability of the reduced-order residual-sensitive fault detection filter depends on the target fault's weighting Q_1.

Theorem 7.1. The error between the plant states and estimator states is not affected by nuisance fault μ_2.

Proof. Define

$$\eta = \Gamma^{-1} x = \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$$
$$e_1 = \eta_1 - \hat{\eta}_1$$
$$e_2 = \eta_2 - \hat{\eta}_2$$

Apply the transformation Γ on (3),

$$\dot{\eta}_1 = A_{11}\eta_1 + A_{12}\eta_2 + N_1\mu_1 \tag{33a}$$
$$\dot{\eta}_2 = A_{21}\eta_1 + A_{22}\eta_2 + N_2\mu_1 + D_2\mu_2$$
$$y = C_1\eta_1 + C_2\eta_2 + v \tag{33b}$$

Using (32), (33) and (31), the error equation for the reduced-order filter is

$$\begin{aligned}\dot{e}_1 = {} & [A_{11} - G_1(D_2^T C_2^T \bar{V}^{-1} C_2 D_2)^{-1} D_2^T C_2^T \bar{V}^{-1} C_1 - \bar{S}^{-1} C_1^T \bar{H}^T \bar{V}^{-1} \bar{H} C_1] e_1 \\ & + N_1\mu_1 - [G_1(D_2^T C_2^T \bar{V}^{-1} C_2 D_2)^{-1} D_2^T C_2^T \bar{V}^{-1} + \bar{S}^{-1} C_1^T \bar{H}^T \bar{V}^{-1} \bar{H}] v\end{aligned}$$

This shows that the error is not affected by the nuisance fault μ_2.

Theorem 7.2. The reduced-order filter (32) is stable if

$$-N_1 Q_1 N_1^T + G_1(D_2^T C_2^T \bar{V}^{-1} C_2 D_2)^{-1} G_1^T \geq 0$$

Proof. The stability of (32) depends on the eigenvalues of the closed-loop A matrix,

$$\bar{A}_{cl} = A_{11} - G_1(D_2^T C_2^T \bar{V}^{-1} C_2 D_2)^{-1} D_2^T C_2^T \bar{V}^{-1} C_1 - \bar{S}^{-1} C_1^T \bar{H}^T \bar{V}^{-1} \bar{H} C_1$$

Substitute this into (30b),

$$\dot{\bar{S}} + \bar{S}\bar{A}_{cl} + \bar{A}_{cl}^T \bar{S} = -\bar{S}[-N_1 Q_1 N_1^T + G_1(D_2^T C_2^T \bar{V}^{-1} C_2 D_2)^{-1} G_1^T] \bar{S} - C_1^T \bar{H}^T \bar{V}^{-1} \bar{H} C_1$$

If $\dot{\bar{S}} + \bar{S}\bar{A}_{cl} + \bar{A}_{cl}^T \bar{S} < 0$, Lyapunov's Stability Theorem implies the eigenvalues of $\bar{A}_{cl}$ are in the open left-half plane. Since $\bar{S}$ and $C_1^T \bar{H}^T \bar{V}^{-1} \bar{H} C_1$ are positive definite, the sufficient condition for $\dot{\bar{S}} + \bar{S}\bar{A}_{cl} + \bar{A}_{cl}^T \bar{S} < 0$ is $-N_1 Q_1 N_1^T + G_1(D_2^T C_2^T \bar{V}^{-1} C_2 D_2)^{-1} G_1^T \geq 0$.

Note that Q_1 is a free parameter which can be tuned to meet this sufficient condition.

Remark 8. The reduced-order filter (32) and reduced-order Riccati equation (30b) will also be the limiting solution for the reduced-order optimal stochastic fault detection filter.

8 Example

In this section, two numerical examples are used to demonstrate the properties of the limiting Riccati matrix S. In Section 8.1, a numerical example from (White and Speyer, 1987) shows that the null space of S includes the nuisance fault direction and the invariant zero directions associated with right-half plane invariant zeros. In Section 8.2, the previous example is modified to have left-half plane invariant zeros and it shows that the null space of S includes only the nuisance fault direction, but not the invariant zero directions associated with left-half plane invariant zeros. It also shows that the left-half plane invariant zero directions will be in the null space of S after the nuisance fault direction is modified.

8.1 Example 1

This example is from (White and Speyer, 1987). The system matrices are

$$A = \begin{bmatrix} 0 & 3 & 4 \\ 1 & 2 & 3 \\ 0 & 2 & 5 \end{bmatrix} \quad , \quad C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad F_1 = \begin{bmatrix} 1 \\ -0.5 \\ 0.5 \end{bmatrix} \quad , \quad F_2 = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}$$

where F_1 is the target fault direction and F_2 is the nuisance fault direction. (C, A, F_2) has an invariant zero at 3 and the invariant zero direction is $[1 \ \ 0 \ \ 0]^T$. The weightings are $Q_1 = 0.5$, $Q_2 = 1$ and $\bar{V}^{-1} = I$. The steady-state solutions to the Riccati equation (10) when $\gamma = 10^{-6}$ and limiting Riccati equation (22) are

$$\Pi = \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0010 & -0.0002 \\ 0.0000 & -0.0002 & 0.0964 \end{bmatrix} \quad , \quad S = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.0964 \end{bmatrix}$$

This example demonstrates that the nuisance fault direction F_2 and the invariant zero direction ν associated with right-half plane invariant zero are in the null space of S.

8.2 Example 2

This example is modified from previous example such that the invariant zero is in the left-half plane instead of the right-half plane. The system matrices are the same except

$$A = \begin{bmatrix} 0 & 3 & 4 \\ -1 & 2 & 3 \\ 0 & 2 & 5 \end{bmatrix}$$

(C, A, F_2) has an invariant zero at -3 and the invariant zero direction ν is $[1 \ \ 0 \ \ 0]^T$. The weightings are the same. The steady-state solutions to the Riccati equation (10) when $\gamma = 10^{-6}$ and limiting Riccati equation (22) are

$$\Pi = \begin{bmatrix} 0.0307 & 0.0916 & -0.0603 \\ 0.0916 & 0.2747 & -0.1803 \\ -0.0603 & -0.1803 & 0.2150 \end{bmatrix} \quad , \quad S = \begin{bmatrix} 0.0305 & 0.0915 & -0.0601 \\ 0.0915 & 0.2746 & -0.1802 \\ -0.0601 & -0.1802 & 0.2146 \end{bmatrix}$$

This shows that the null space of S includes only the nuisance fault direction F_2, but not the invariant zero direction ν. Also, the filter has an eigenvalue at -3. If the nuisance fault direction is changed to ν and the weightings are the same, the steady-state solutions to the Riccati equation (10) when $\gamma = 10^{-12}$ and limiting Riccati equation (22) are

$$\Pi = \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0014 & -0.0003 \\ 0.0000 & -0.0003 & 0.0965 \end{bmatrix} \quad , \quad S = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.0964 \end{bmatrix}$$

This shows that the null space of S includes both the nuisance fault direction F_2 and the invariant zero direction ν. Note that $\text{span}\{F_2 \ \ \nu\} = \text{span}\{\nu \ \ A\nu\}$

9 Conclusion

The residual-sensitive fault detection filter is derived from solving a min-max problem which makes the residual sensitive to the target fault, but not to the nuisance fault. In the limit as the nuisance fault weighting goes to zero, this filter is equivalent to an unknown input observer which puts the nuisance fault into an unobservability subspace. Furthermore, there exists a reduced-order filter in the limit. Since the target fault is explicit in this derivation, the reduced-order filter is found with respect to the target fault direction and weighting. This filter also extends the unknown input observer to a time-varying system.

References

Beard, R. V. (1971). *Failure Accomodation in Linear Systems through Self-Reorganization*, Ph.D. thesis, Massachusetts Institute of Technology.

Bell, D. J. and D. H. Jacobsen (1975). *Singular Optimal Control Problems*, vol. 117 of *Mathematics in Science and Engineering*, Academic Press.

Brinsmead, T. S., J. S. Gibson, G. C. Goodwin, G. H. Lee, and D. L. Mingori (1997). "Fault detection: A quadratic optimization approach," in *Proceedings of the 36th Conference on Decision and Control*, pp. 3970–3975.

Bryson, A. E. and Y.-C. Ho (1975). *Applied Optimal Control: Optimization, Estimation and Control*, Hemisphere.

Chen, R. H. and J. L. Speyer (1999). "Optimal stochastic fault detection filter," To be published in the *Proceedings of the American Control Conference.*

Chung, W. H. and J. L. Speyer (1998). "A game theoretic fault detection filter," *IEEE Transactions on Automatic Control*, **AC-43**, no. 2, pp. 143–161.

Douglas, R. K. and J. L. Speyer (1996). "Robust fault detection filter design," *AIAA Journal of Guidance, Control, and Dynamics*, **19**, no. 1, pp. 214–218.

Douglas, R. K. and J. L. Speyer (1999). "$\mathcal{H}_\infty$ bounded fault detection filter," *AIAA Journal of Guidance, Control, and Dynamics*, **22**, no. 1, pp. 129–138.

Jones, H. L. (1973). *Failure Detection in Linear Systems*, Ph.D. thesis, Massachusetts Institute of Technology.

Lee, G. H. (1994). *Least-Squares and Minimax Methods for Filtering, Identification and Detection*, Ph.D. thesis, University of California at Los Angeles.

Massoumnia, M.-A. (1986). "A geometric approach to the synthesis of failure detection filters," *IEEE Transactions on Automatic Control*, **AC-31**, no. 9, pp. 839–846.

Massoumnia, M.-A., G. C. Verghese, and A. S. Willsky (1989). "Failure detection and identification," *IEEE Transactions on Automatic Control*, **AC-34**, no. 3, pp. 316–321.

White, J. E. and J. L. Speyer (1987). "Detection filter design: Spectral theory and algorithms," *IEEE Transactions on Automatic Control*, **AC-32**, no. 7, pp. 593–603.