A Game Theoretic Multiple-Fault Detection Filter

Emmanuell A. Murray and Jason L. Speyer

Abstract—In this paper, the game-theoretic fault detection filter is extended to the multiple-fault case. It approximates detection filters based on spectral and geometric theories with a set of disturbance attenuation problems, and extends these detection filters to time-varying systems. The new detection filter is derived for a nonzero disturbance attenuation bound and evaluated in an example.

I. INTRODUCTION

The detection filter using analytical redundancy for fault detection and isolation was first introduced by Beard [1] and Jones [2], now called the Beard-Jones detection filter (BJDF). The idea of the BJDF is to place the reachable subspace of each fault into non-overlapping invariant subspaces called detection spaces. A slight generalization of the BJDF, called the restricted diagonal detection filter (RDDF), was derived using geometric theory in [3]. Some faults may not need to be detected, but simply blocked from the residual output. In that case, the RDDF places these faults into the unobservable subspace of the residual. Then, when a fault occurs, the residual can be projected onto the orthogonal complement of each detection space so that the fault can be identified. Each detection space includes both the reachable subspace of the associated fault and the directions associated with the transmission of the fault to the residual. Constraints on the locations of the invariant zeros are imposed to guarantee the desired properties of the detection filter. Geometric and spectral analyses were given in [3] and [4], respectively. Design algorithms were developed based on spectral theory in [4], [5], [6] and geometric theory in [7].

The unknown input observer (UIO), originally used for fault-tolerant estimation, was applied to the single-fault detection filter problem in [8]. The UIO simplifies the detection filter problem by requiring that all but one fault be placed in an unobservable detection space, which can be annihilated via model reduction. It was shown that constraints on the fault directions and the location of invariant zeros are relaxed compared to the multiple-fault filters. Thus, a bank of UIOs can be used to detect multiple faults, trading relaxed constraints for possibly increased computational complexity. Recently, banks of UIOs were applied to multi-vehicle actuator fault detection in [9] and to fault detection of Markovian jump linear systems in [10].

The main drawback of the spectral and geometric methods is the rigidity of their structure and sensitivity to noise.

To increase the flexibility of the detection filter problem, it has been approximated by relaxing the requirement of strict blocking of undesirable faults and noises. Recently, much attention has been devoted to approximating UIOs for fault detection using methods based on $H_\infty$ estimation, which seeks to minimize the $H_\infty$ norm of the disturbances’ transfer function. Enhanced sensitivity to the detected faults was examined in [11], [12], [13] and extended to time-varying systems [14] and multiple-fault detection filters [15], [16], though the latter is limited to certain special classes of faults.

Another approximation method is based on disturbance attenuation, and is the focus of this paper. This method generally applies to more complex systems and disturbances than the $H_\infty$-based methods. In [17], a time-varying approximation of the UIO, called the game-theoretic fault detection filter (GTFDF), was obtained by optimizing a disturbance attenuation problem (DAP) with respect to the disturbance inputs and estimate via a differential game. The structure of the UIO is largely recovered in the limit as the disturbance attenuation bound goes to zero. However, it was shown in [18] that the invariant zero directions are not automatically included in the detection spaces created by the optimization, though they can be included artificially by modifying the fault directions. The GTFDF was applied to decentralized fault detection in [19] and a similar detection filter with enhanced sensitivity to the detected fault was derived in [20].

A stochastic method called the optimal stochastic fault detection filter (OSFDF) was examined in [18], [21] and applied to the multiple-fault case in [22]. The OSFDF uses a stochastic description of the estimation error covariance to construct a cost function. The optimization chooses the filter gain to minimize the transmissions of disturbances and maximize transmission of the detected fault to the projected output error. In the multiple-fault case, multiple error covariances are constructed so that detected faults can be isolated by their respective projected residuals. An example showed that this method automatically included the faults’ invariant zero directions in their proper detection spaces. Thus, in the limit as the weight on the nuisance fault transmission goes to infinity, the structure of the RDDF is recovered. However, this approach assumes that all fault magnitudes are white noise and very little could be stated about solution optimality.

In this paper, the game-theoretic multiple-fault detection filter (GTMFDF) is derived. The GTMFDF extends the GTFDF to the multiple-fault case by modeling the detection filter problem as a set of DAPs to be optimized via a single differential game. However, since the globally optimal solution for the filter gain is difficult to obtain, sufficient conditions for satisfying the DAPs are derived instead. It is
shown that the result is similar to the multiple-fault OSFDF with a more general description of the fault magnitudes and simpler solution requirements. Thus, the flexibility, simplicity, and robustness of the GTDFD problem for single-fault detection is combined with the multiple-fault optimization of the OSFDF, resulting in a robust multiple-fault detection filter with relatively few assumptions on the system and fault structure compared to the current literature.

This paper is organized as follows. To establish a simple basis for understanding the multiple-fault problem, the RDDF structure is discussed and approximated by a set of DAPs in Section II. Then, the implied differential game problem is simplified into a feasibility problem to find a filter gain that satisfies the DAPs. Sufficient conditions for satisfying the DAPs are derived in Section III when the disturbance attenuation bound is nonzero. It is shown that these conditions require certain Riccati differential inequalities to be satisfied with nonnegative solutions. Then, the Riccati inequalities are used as the constraints of a secondary optimization problem to determine the filter gain. Finally, a numerical example is discussed in Section IV.

II. DIFFERENTIAL GAME PROBLEM FORMULATION

In this section, the detection filter problem is formulated as a set of DAPs that can be optimized via a differential game problem. First, an overview of the RDDF is given in Section II-A to examine the invariant subspace structure of the detection filter. Then, the RDDF is extended to the time-varying case and approximated with a set of DAPs in Section II-B. Finally, the DAPs are converted into a differential game feasibility problem and the required assumptions are discussed in Section II-C.

A. Restricted-Diagonal Detection Filter Background

Consider an observable, LTI system with $q$ faults

$$
\dot{x}(t) = Ax(t) + Bu(t) + \sum_{i=1}^{q} F_i \mu_i(t),
$$

$$
y(t) = Cx(t) + Du(t),
$$

with state $x(t) \in \mathbb{R}^n$, control input $u(t) \in \mathbb{R}^l$, measurement $y(t) \in \mathbb{R}^m$, unknown, arbitrary fault magnitude $\mu_i(t) \in \mathbb{R}$, and a priori known fault direction $F_i$. Assume that the measurements are linearly independent, and so $C$ is full rank ($m \leq n$). The detection filter is a linear observer of the form

$$
\hat{x}(t) = A\hat{x}(t) + Bu(t) + L(y(t) - C\hat{x}(t) - Du(t)),
$$

$$
r(t) = y(t) - C\hat{x}(t) - Du(t),
$$

with state estimate $\hat{x}(t) \in \mathbb{R}^n$, filter gain $L$, and residual $r(t) \in \mathbb{R}^m$. Using (1) and (2), the estimation error $e(t) \triangleq x(t) - \hat{x}(t)$ is subject to the dynamic system

$$
\dot{e}(t) = (A - LC)e(t) + \sum_{i=1}^{q} F_i \mu_i(t),
$$

$$
r(t) = Ce(t).
$$

A detection filter is required to detect one or more target faults. However, some faults, called nuisance faults [7], [6], are disturbances that may not need to be detected explicitly, simply blocked from the residuals. Without loss of generality, assume that the first $s \leq q$ faults are target faults and the remaining $q - s$ faults are nuisance faults. Several residuals can be generated, each sensitive to only one target fault, by multiplying $r(t)$ by a residual projector $H_i$. Therefore, the projected residual associated with $\hat{\mu}_i(t)$ is written as

$$
r_i(t) = H_iCe(t).
$$

The RDDF problem is to choose $L$ so that the detection filter satisfies the following objectives [3], [7], [23], [6]:

- When the target fault $\mu_i$ occurs, the residual $r(t)$ lies in a fixed subspace that is linearly independent from the subspace associated with the remaining faults.
- The projected residual $r_i(t)$ is nonzero if and only if the target fault $\mu_i(t)$, $i = 1, \ldots, s$ occurs.
- The eigenvalues of the filter can be chosen arbitrarily.
- The steady-state residual response to a constant bias fault is nonzero.

These objectives lead to a description of the state space structure of the detection filter problem. The remainder of this section is focused on deriving the state space geometry.

The invariant subspace structure of the RDDF is formulated around the occurrence of the complementary fault $\hat{\mu}_i \triangleq [\mu_1 \ldots \mu_{i-1} \mu_{i+1} \ldots \mu_q]$ with direction

$$
\hat{F}_i \triangleq [F_1 \ldots F_{i-1} F_{i+1} \ldots F_q].
$$

When it occurs, the estimation error lies in $W_i \triangleq W_1 + \ldots + W_{i-1} + W_{i+1} + \ldots + W_q$ and the residual lies along $CW_i$ where for $j = 1, \ldots, q$

$$
W_j \triangleq \text{Im} \left[ F_j \ (A - LC) F_j \ldots (A - LC)^{n-1} F_j \right].
$$

Let $\delta_i$ be the smallest nonnegative integer such that $C(A - LC)^{\delta_i} F_i \neq 0$ for $i = 1, \ldots, s$. Therefore, $C(A - LC)^k F_i = 0$ for $k = 0, \ldots, \delta_i - 1$, which implies $CA^{\delta_i} F_i = 0$ and $C(A - LC)^{\delta_i} F_i = CA^{\delta_i} F_i$. Let the filter gain $L$ be chosen such that $\delta_i + 1$ of the eigenvectors of $A - LC$ span

$$
W_i^* = \text{Im} \left[ F_i \ A F_i \ldots A^{\delta_i} F_i \right], i = 1, \ldots, s.
$$

Thus, $W_i^*$ is the smallest reachable, observable subspace associated with $F_i$ [3], [4] that satisfies

$$
(A - LC) W_i^* \subseteq W_i^* \quad \text{(6a)}
$$

$$
\text{Im} \ F_i \subseteq W_i^* \quad \text{(6b)}
$$

By (6), $W_i^*$ is invariant under $(A - LC)$, and is known as the minimal $(C, A)$-invariant subspace of $F_i$. When the target fault $\hat{\mu}_i$ occurs, the residual will lie along $CW_i^*$. Further, let other eigenvectors of $A - LC$ span the minimal $(C, A)$-invariant subspace $W_i^*$ of $F_i \triangleq [F_{s+1} \ldots F_q]$ (see [25] for the algorithm to calculate this subspace) and define $W_i^* = W_i^* \cup \ldots + W_{i-1}^* + W_{i+1}^* + \ldots + W_s^* + \hat{W}^*$ where $\hat{W}^*$ and $W_i^*$ satisfy (6) for $F$ and $F_i$, respectively. When the complementary fault $\hat{\mu}_i$ occurs, the residual will lie along
To satisfy the first objective, \( CW^*_i \) and \( \hat{CW}^*_i \) must be linearly independent for \( i = 1, \ldots, s \), i.e. the target faults must be \((C,A)\) output separable from their complementary faults. Note that output separability implies that \( m \geq q \).

To satisfy the second objective, the RDDF must generate a set of residuals that are each insensitive to all but a given target fault [3]. To isolate a target fault direction \( F_i \), the detection filter is designed so that \( \hat{F}_i \) is placed in the unobservable subspace of the projected residual \( r_i(t) \) from (4), where the residual projector \( \hat{H}_i \) is defined as [23]

\[
\hat{H}_i = I_m - CW^*_i \left( (CW^*_i)^T CW^*_i \right)^{-1} (CW^*_i)^T.
\]

Thus, when the faults are output separable, \( \hat{H}_i CW^*_i \neq 0 \) and so the target fault is observable. Further, \( \text{Ker} \, \hat{H}_i = CW^*_i q \) and so the complementary fault is unobservable, thus satisfying the second objective. Since \( \hat{H}_i \) is an orthogonal projector, \( \hat{H}_i = H_i^T = H_i^2 \).

To satisfy the third objective, it was shown in [3] that the eigenvectors of \( A - LC \) must span both \( W^*_i \) and the invariant zero directions of \( (C,A,F_i) \) for \( i = 1, \ldots, q \). Otherwise, some of the eigenvalues will be located at the associated invariant zeros. Denote the subspace of invariant zero directions associated with \( (C,A,F_i) \) as \( V_i \). To satisfy the third objective, choose the filter gain such that some of the eigenvectors span \( T_i = W^*_i \oplus V_i \), called the minimal \((C,A)\)-unobservability subspace of \( F_i \) [3] or the detection space of \( F_i \) [2]. Thus, \( T_i \) contains all and only the directions that satisfy (6) for \( F_i \), where \( C T_i = CW^*_i \) [3]. Also, by using basic linear systems theory, if there exists an invariant zero at the origin, the steady state residual due to a constant input from the associated fault is zero [24]. Therefore, in order to satisfy the fourth objective, there can be no invariant zeros at the origin that are associated with a target fault.

Invariant zeros of \( (C,A,[F_1 \ldots F_q]) \) that are not associated with \( (C,A,F_i) \), \( i = 1, \ldots, q \) also become eigenvalues of the detection filter. However, since these extra invariant zeros are not associated with a single fault, the resulting eigenvalues cannot be moved without altering the state space (one method is discussed in [4]). Denote the subspace of extra invariant zero directions as \( V_{\text{ext}} \). If \( (C,A,[F_1 \ldots F_q]) \) contains no more invariant zeros than \( (C,A,F_i) \), \( i = 1, \ldots, q \), then every eigenvalue of the detection filter can be specified arbitrarily and the faults are mutually detectable.

The remainder of the state space is called the complementary subspace \( C \). Thus, the state space composition is

\[
T_1 \oplus \ldots \oplus T_q \oplus V_{\text{ext}} \oplus C = \mathbb{R}^n.
\]

To summarize, in order to generate an arbitrarily stable detection filter, the faults must be observable, output separable, mutually detectable, and have no invariant zeros at the origin.

**Remark I:** To simplify the above analysis, this section has considered only scalar faults. However, the GTMFDF derivation applies to the vector fault case as well. For a detailed derivation of the state space geometry in the vector fault case, see [25].

### B. Extension to the Time-Varying and Approximate Cases

Redefine the system in (1) to be time-varying where

\[
\dot{x}(t) = A(t)x(t) + B(t)u(t) + \sum_{i=1}^q F_i(t)\mu_i(t)
\]

\[
y(t) = C(t)x(t) + v(t)
\]

is defined from initial time \( t_0 \) to final time \( t_1 < \infty \). A measurement noise vector \( v(t) \) has also been included. The time-varying extension of \( W^*_i(t) \) in (5) is obtained by requiring that the detection filter dynamics place \( F_i(t) \) in the observable subspace of \( (C(t),A(t) - L(t)C(t)) \) for the entire time interval from \( t_0 \) to \( t_1 \). Therefore, define

\[
W^*_i(t) \triangleq \text{Im} \left[ B_i^0(t) \ B_i^1(t) \ldots B_i^{\beta_i}(t) \right]
\]

where the columns of \( W^*_i(t) \) are constructed by [17]

\[
B_i^0(t) = F_i(t)
\]

\[
B_i^1(t) = A(t)B_i^{\beta_i-1}(t) - \hat{B}_i^{\beta_i-1}.
\]

and \( \beta_i \) is the smallest nonnegative integer such that \( C(t)B_i^{\beta_i}(t) \neq 0 \) \( \forall t \in [t_0,t_1] \).

To approximate a multiple-fault detection filter for linear time-varying systems, a set of DAPs are formulated by requiring that the target faults be observable to the projected residual and relaxing the requirement on strict blocking implied by the first two RDDF objectives. Instead, the transmissions of the complementary fault, sensor noise, and initial condition error, henceforth referred to collectively as the *disturbance parameters*, are bounded above by a preset level. So that the target faults are observable, \( F_i(t) \) must be \((C(t),A(t))\) output separable from \( \hat{F}_i(t) \) \( \forall t \in [t_0,t_1] \) [17]. Further, the third RDDF objective is relaxed to requiring only that the detection filter dynamics be stable. To ensure that a stabilizing solution exists, \((C(t),A(t))\) is assumed to be detectable \( \forall t \in [t_0,t_1] \). The fourth RDDF objective only applies to the infinite-time case.

For the \( i \)th DAP, the transmissions of the complementary faults and disturbances are separated from the transmission of the target fault into their own state \( x_i(t) \) where

\[
\dot{x}_i(t) = A(t)x_i(t) + B(t)u(t) + \hat{F}_i(t)\hat{\mu}_i(t)
\]

\[
y_i(t) = C(t)x_i(t) + v_i(t).
\]

Note that to simplify the derivation each measurement \( y_i(t) \) has been given its own noise \( v_i(t) \). By substituting \( y_i(t) \) for \( y(t) \) in (2) and using (11), the dynamics of the state error \( e_i(t) \) \( \triangleq x_i(t) - \hat{x}(t) \) and the projected residual \( r_i(t) \) are

\[
\dot{e}_i(t) = (A(t)-L(t)C(t))e_i(t) + \hat{F}_i(t)\hat{\mu}_i(t)
\]

\[
r_i(t) = \hat{H}_i(t)C(t)e_i(t) + \hat{H}_i(t)v_i(t)
\]

for \( i = 1, \ldots, s \). Since the projected residual contains a direct feedthrough term from the sensor noise, the projected output error \( \hat{H}_i(t)C(t)e_i(t) \) is used instead of \( r_i(t) \) to represent the transmission of the fault to the output. Thus, the \( i \)th DAP is
written for \( i = 1, \ldots, s \) as

\[
\int_{t_0}^{t_1} \left[ \| \hat{H}_i C e_i \|_{Q_i}^2 - \| \hat{\mu}_i(t) \|_{M_i}^2 + \| v_i(t) \|_{V_i}^2 \right] dt + \| e_i(t_0) \|_{P_0}^2 \leq \gamma, \tag{13}
\]

subject to the dynamic system (12) for any \( \hat{\mu}_i(t), v(t), \) and \( e_i(t_0) \) that satisfy \( \int_{t_0}^{t_1} \| \hat{\mu}_i(t) \|_{M_i}^2 dt < \infty \) and \( \int_{t_0}^{t_1} \| v_i(t) \|_{V_i}^2 dt < \infty \). \( \gamma > 0 \) is the arbitrary DAP bound and \( Q_i \geq 0, \) \( M_i > 0, \) \( V_i > 0, \) and \( P_0 > 0 \) are arbitrary design weightings. Typically, \( V_i \) is chosen as the covariance of the measurement noise. Further, when the design weightings \( M_i, V_i, \) and \( P_0 \) are chosen to be larger, the projected residual becomes less sensitive to the complementary fault, sensor noise, and initial condition error, respectively, which can also be achieved simultaneously by choosing \( Q_i \) to be larger.

C. Problem Formulation

By multiplying both sides of (13) by the denominator of the left-hand side, subtracting the right-hand side from the left, and setting the left-hand side equal to \( J_i \), (13) is converted into the nonconvex cost function

\[
J_i = \frac{1}{2} \int_{t_0}^{t_1} \left[ \| \hat{H}_i C e_i(t) \|_{Q_i}^2 - \| \hat{\mu}_i(t) \|_{M_i}^2 \right] dt - \| v_i(t_0) \|_{P_0}^2 \tag{14}
\]

for \( i = 1, \ldots, s \) where \( P_0 \triangleq \gamma P_0^{-1} \) and \( V_i \triangleq \gamma^{-1}V_i \). The detection filter problem is modeled as a differential game optimization by summing (14) over \( i = 1, \ldots, s \), minimizing the sum with respect to the filter gain, and maximizing the sum with respect to the disturbance parameters. Therefore, the differential game problem is

\[
\min_{L(t)} \max_{\psi(t)|_{\Delta t}} \max_{\phi(t)} \max_{\Pi_i} \sum_{i=1}^{s} J_i \tag{15}
\]

subject to (12a) for \( i = 1, \ldots, s \). Note that \( \mu_1, \ldots, \mu_s \) are not independent since they share elements with each other.

Since the detection filter gain \( L(t) \) does not appear in the game cost (14) and enters linearly into the constraint (12a), (15) is singular with respect to \( L(t) \) [26]. This makes the process of finding a globally optimal solution for \( L(t) \) that will generate the desired fault detection properties very complex. However, in order to satisfy the DAPs (13), it is only required that (14), \( i = 1, \ldots, s \), is nonpositive for any value of \( \hat{\mu}_i(t), v_i(t), \) and \( e_i(t_0) \) \( \forall t \in [t_0, t_1] \). Thus, we only require a solution to a feasibility problem in \( L(t) \) such that (14) is nonpositive. To further simplify the problem statement, assume that all of the disturbance parameters are independent. Note that this assumption will only affect the problem statement, not how we eventually solve for the filter gain. Therefore, to determine a filter gain sufficient to satisfy (13), we will solve the following simplified problem:

**Problem 1:** Find \( L(t) \) such that

\[
\max_{v_i(t)} \max_{\hat{\mu}_i(t)} \max_{e_i(t_0)} J_i \leq 0 \forall i = 1, \ldots, s,
\]

subject to (12a) and (14).

III. DETECTION FILTER PROBLEM SOLUTION

In this section, a sub-optimal solution for the GTMFDF gain in Problem 1 is determined for the general case where \( \gamma > 0 \). First, the conditions under which (14) is nonpositive are determined in Section III-A. After appending the dynamics, it is shown that (14) can be rewritten such that the disturbances parameters enter as nonpositive quadratic terms. The problem requires that the filter gain be chosen such that a certain set of nonpositive Riccati differential inequalities have positive solutions. In Section III-B, it is shown that this derivation generalizes and clarifies the GTFDF [17] and OSFDF [22]. Solutions for \( L(t) \) given an arbitrary secondary cost function are derived in Section III-C. Finally, results for the infinite-time case are discussed in Section III-D.

A. Conditions for Nonpositivity of the Game Cost

We are now ready to consider the conditions under which the game cost (14) is nonpositive. First, the estimation error dynamics (12a) are appended to (14) using the Lagrange multiplier \( e_i^T \Pi_i \), which yields

\[
J_i = \frac{1}{2} \int_{t_0}^{t_1} \left[ \| \hat{H}_i C e_i \|_{Q_i}^2 - \| \hat{\mu}_i \|_{M_i}^2 - \| v_i \|_{V_i}^2 \right] dt - \frac{1}{2} \| e_i(t_0) \|_{P_0}^2.
\]

Note that for compactness the variables’ time dependence is no longer shown. By integrating \( \int_{t_0}^{t_1} e_i^T \Pi_i e_i dt \) by parts, substituting (12a), adding and subtracting \( \int_{t_0}^{t_1} e_i^T \Pi_i e_i dt \) by parts, substituting (12a), adding and subtracting \( \int_{t_0}^{t_1} e_i^T \Pi_i \hat{e}_i dt \), and collecting terms,

\[
J_i = \frac{1}{2} \int_{t_0}^{t_1} \left[ -\| \hat{\mu}_i - 1 \gamma M_i \hat{F}_i \Pi_i e_i \|_{M_i}^2 - \| v_i + V L^T \Pi_i e_i \|_{V_i}^2 \right] dt - \frac{1}{2} \| e_i(t_0) \|_{P_0 - \Pi_i(t_0)}^2 - \frac{1}{2} \| e_i(t_1) \|_{\Pi_i(t_1)}^2 \tag{16}
\]

where

\[
\Psi_i(\Pi_i, L) = \Pi_i + \Pi_i (A - LC) + (A - LC)^T \Pi_i + \Pi_i \left( \frac{1}{\gamma} F_i M_i \hat{F}_i^T + LV L^T \right) \Pi_i + C^T H_i Q_i H_i C.
\]

Clearly, (16) is nonpositive if there exists some \( L \) such that

\[
0 \geq \Psi_i(\Pi_i, L, t) \tag{18}
\]

\[
0 \leq \Pi_i - \Pi_i(t_0) \tag{19}
\]

\[
0 \leq \Pi_i(t_1), \tag{20}
\]

which implies that the faults are placed into approximate detection spaces to be isolated by each projected residual. Therefore, Problem 1 requires a solution to the coupled Riccati inequalities (18) given (17) with boundary conditions (19) and (20), though for implementation purposes we restrict \( \Pi_i \) to be positive definite. Since the degree of fault blocking can be changed by adjusting \( \gamma \), the structure of the GTMFDF is less constrained than the detection filters based on spectral [4], [7] and geometric [3], [8] theories.
B. Comparison to Previous Robust Detection Filters

In this section, the GTMFDF is compared to the previous Riccati-based robust fault detection methods. First, it is shown that the above problem formulation generalizes the solution to the single-fault problem by Chung and Speyer in [17]. Further, though the constraint equations of the current problem are similar to those of the multiple-fault OSFDF by Chen and Speyer in [22], the GTMFDF problem is shown to be clearer and more general.

In [17], it was proven that the single-fault DAP (assumes \( s = 1 \)) is satisfied when the Riccati variable \( \Gamma \) is propagated by the differential equation

\[
0 = \dot{\Gamma} + \Gamma A + A^T \Gamma + \frac{1}{\gamma} \dot{F}_i M_i \dot{F}_i^T + C^T \left( \dot{H}_i Q_i \dot{H}_i - V^{-1} \right) C
\]

where \( \Gamma(t_0) = \Gamma_0 \) and the solution for the filter gain \( L \) is

\[
L = \Gamma^{-1} C^T V^{-1}.
\]

However, by adding and subtracting \( C^T V^{-1} C \) to (17),

\[
\Psi_i = \dot{\Pi}_i + \Pi_i A + A^T \Pi_i + \Pi_i (L - C^T V^{-1}) \left( V (L - C^T V^{-1}) \right)^T
\]

\[
+ \frac{1}{\gamma} \dot{F}_i M_i \dot{F}_i^T \Pi_i + C^T \left( \dot{H}_i Q_i \dot{H}_i - V^{-1} \right) C.
\]

Clearly, (23) is simply (21) with an added quadratic term to account for the difference between \( L \) and \( \Pi_i^{-1} C^T V^{-1} \). Thus, (21) and (22) are a special case of the solution to (18) using (23).

The advantage of (21) and (22) is that the Riccati solution can be computed independently of \( L \), simplifying its calculation. However, by generalizing the constraint as in (23), the filter gain can be chosen achieve secondary objectives. For example, if we assume that

\[
L = \Pi_i^{-1} C^T V^{-1} + L_0
\]

where \( L_0 \) is an arbitrary \( m \times n \) matrix, then (23) becomes

\[
\Psi_i = \dot{\Pi}_i + \Pi_i A + A^T \Pi_i + \Pi_i \left[ \frac{1}{\gamma} \dot{F}_i M_i \dot{F}_i^T + L_0 V L_0^T \right] \Pi_i
\]

\[
+ C^T \left( \dot{H}_i Q_i \dot{H}_i - V^{-1} \right) C
\]

and it may be possible to choose or optimize \( L_0 \) to decrease the eigenvalues of \( A - LC \). Such optimization of the filter gain is the subject of Section III-C.

Next, to compare the GTMFDF derivation to the multiple-fault OSFDF derivation in [22], the Riccati differential inequality (18) is rewritten in terms of \( P_i = \Pi_i^{-1} \). Assuming that (17) equals zero, multiplying on the left and right by \( P_i \), and substituting \( P_i = \Pi_i^{-1} \) over \( i = 1, \ldots, s \),

\[
\dot{P}_i = (A - LC) P_i + P_i (A - LC)^T + \frac{1}{\gamma} \dot{F}_i M_i \dot{F}_i^T + L V L^T
\]

\[
+ P_i C^T \dot{H}_i Q_i \dot{H}_i C P_i
\]

where \( P_i(t_0) = \gamma P_0 \). In [22], the filter gain of the multiple-fault OSFDF is optimized subject to

\[
\dot{W}_i = (A - LC) W_i + W_i (A - LC)^T + \frac{1}{\gamma} \dot{F}_i M_i \dot{F}_i^T + L V L^T - F_i N_i F_i^T
\]

for \( i = 1, \ldots, s \) where \( N_i \) is a design weighting on the transmission of the target fault to the residual.

Though the two constraint equations (24) and (25) share some similarities, the GTMFDF improves upon the multiple-fault OSFDF of [22] in three ways. First, since the GTMFDF compares outputs to disturbances instead of target faults to disturbances, the GTMFDF gives a clearer representation of how the disturbance parameters are blocked. Second, the OSFDF requires the solution to a specific cost function. On the other hand, the GTMFDF only requires a solution to a feasibility problem for the constraints (24) and \( P_i \geq 0 \) for \( i = 1, \ldots, s \). If desired, a secondary cost function can be used to achieve other objectives, such as enhancing sensitivity to the target fault, thereby increasing the flexibility of the detection filter problem. Finally, the multiple-fault OSFDF derivation assumes that the disturbances are modeled as white noise processes. The GTMFDF requires no such assumption, proving that constraints like (24) and (25) are applicable to more general systems.

Remark 2: It is possible to introduce target fault sensitivity into the DAPs by adding \( \int_{t_0}^{t_f} \lVert \mu_i(t) \rVert_{N_i^{-1}}^2 \, dt \) to the numerator of (13), allowing \( \mu_i(t) \neq 0 \), and minimizing with respect to \( \mu_i(t) \). In the single-fault case, this resembles the \( H_\infty \) controller synthesis problem where \( \mu_i(t) \) is the control [27]. However, since the user does not have control over \( \mu_i(t) \), \( H_\infty \) results cannot be guaranteed and it is unclear how target fault detection is affected. When \( Q_i = 0 \), as in [20], [22], a constraint identical to (25) can be obtained without the assumption of white noise processes for all disturbances. However, this introduces a positive quadratic term for the target fault into the game cost (16). Thus, we will be unable to claim that there exists some \( L \) such that (16) will be nonnegative independent of all fault magnitudes.

C. Filter Gain Optimization

In this section, the filter gain \( L \) is optimized with respect to a new cost function. Since any solution \( \Pi_i \geq 0 \) to (18) automatically implies that (14) is nonpositive, the specific cost function used at this stage is arbitrary. Thus, let the optimal filter gain minimize the cost function \( J \), stated as

\[
\min_L \ J = \min_L \sum_{i=1}^{s} \int_{t_0}^{t_f} \text{tr} \ \Omega_i \ dt
\]

where the integrand \( \Omega_i \) is an \( n \times n \) real, symmetric, differentiable function chosen by the user. For convenience at this stage of the problem, substitute \( P_i = \Pi_i^{-1} \) subject to (24). Also for convenience, assume that \( \Omega_i \) is a function of \( P_i \), \( L \),

\footnote{In [22], the optimization is actually subject to two differential equations, one Riccati and one Lyapunov, for each target fault. In (25), the two have been summed into a single equation with a single variable.}
and $t$ only. At the end of the section, suggestions on choosing $\Omega_i$ such that (26) has a non-trivial solution are discussed and an example is presented.

To determine the first-order necessary conditions for optimality of (26), use $\Lambda_i$ to append (24) to $J$ to obtain

$$J = \sum_{i=1}^{s} \int_{t_0}^{t_1} \left\{ \Omega_i \left( P_i, L, t \right) + \Lambda_i \left[ (A - LC) P_i + P_i (A - LC)^T \right] + \frac{1}{\gamma} \dot{F}_i M_i \dot{F}_i^T + LV L^T + P_i C^T \hat{H}_i Q_i \hat{H}_i C P_i - \dot{P}_i \right\} dt.$$ 

Integrating $\sum_{i=1}^{s} \int_{t_0}^{t_1} \text{tr}(\Lambda_i \dot{P}_i) \ dt$ by parts and taking the first-order variation with respect to $L$ and $P_i$, the first-order necessary conditions for optimality are

$$0 = \sum_{i=1}^{s} \left[ 2 \left( V L^T - CP_i^* \right) \Lambda_i + \frac{\partial \Omega_i (P_i^*, L^*, t)}{\partial L^*} \right]$$

(27)

$$-\dot{\Lambda}_i = \Lambda_i \left( A - L^* C + P_i^* C^T \hat{H}_i Q_i \hat{H}_i C \right)$$

$$+ \left( A - L^* C + P_i^* C^T \hat{H}_i Q_i \hat{H}_i C \right)^T \Lambda_i + \frac{\partial \Omega_i (P_i^*, L^*, t)}{\partial P_i^*}$$

(28)

$$0 = \Lambda_i(t_1)$$

(29)

for $i = 1, \ldots, s$ where $L^*$ is the optimal strategy for the filter gain and $P_i^*$ is the Riccati variable using $L^*$. Therefore, since $\Omega_i$ is symmetric by assumption, $\Lambda_i$ is the solution of a Lyapunov differential equation. In order to obtain a non-trivial solution for $L^*$, $\Omega_i$ should be chosen such that $\frac{\partial \Omega_i}{\partial L^*}$ is nonzero. Otherwise, from (28) and (29), (24) will not constrain the filter gain solution since $\Lambda_i$ will equal zero. The optimal filter gain is therefore determined by solving a two-point boundary value problem which includes a set of Riccati equations (24) and Lyapunov equations (28) coupled by (27) with boundary conditions $P_i(t_0) = P_0$ and (29).

Finally, as an example, the simplified cost function of the multiple-fault OSFDF in [22] is minimized with respect to $L$. It was proven in [17] that as $\gamma \rightarrow 0$, $\Pi_i$ obtains a nullspace that contains $W_{\Pi_i}^n$, implying that the range space of $P_i$ is dominated by $W_{\Pi_i}^n$. Thus, the optimization should attempt to minimize the transmission of $\hat{F}_i$ by placing $P_i$ approximately in the nullspace of $\hat{H}_i C$. So, choose $\Omega_i$ as

$$\text{tr} \ \Omega_i = \text{tr} \ K_i \hat{H}_i C P_i^* C^T \hat{H}_i$$

(30)

where $K_1, \ldots, K_s$, are design weightings on the complementary fault transmissions. Thus, the optimization problem (26) attempts to choose $L$ such that $P_i$ has the aforementioned geometric structure. When $K_j$, $j \in \{1, \ldots, s\}$, is large, the transmission from $\hat{F}_j$ to the residual $r_j$ is smaller. By differentiating $\Omega_i$ with respect to $L$ and $P_i$ and substituting into (27) and (28), the optimal filter gain is

$$L^* = \left( \sum_{i=1}^{s} \Lambda_i \right)^{-1} \left[ \sum_{i=1}^{s} \Lambda_i P_i^* C^T V^{-1} \right]$$

(31)

where $\Lambda_i$ is subject to the matrix differential equation

$$-\dot{\Lambda}_i = \Lambda_i \left( A - L^* C + P_i^* C^T \hat{H}_i Q_i \hat{H}_i C \right)$$

$$+ \left( A - L^* C + P_i^* C^T \hat{H}_i Q_i \hat{H}_i C \right)^T \Lambda_i + C^T \hat{H}_i K_i \hat{H}_i C$$

(32)

with boundary condition (29). A numerical example using the preceding cost function is given in Section IV. Alternative cost functions can include terms to enhance sensitivity to the target fault or minimize the eigenvalues of $A - LC$.

D. Steady-State Detection Filter

Finally, it is important to discuss the steady-state (infinite time) results for the GTMFDF. In the previous sections, we assumed that the design parameters $Q_i$, $M_i$, $V$, and $\Pi_0$ are chosen so that there exists a real, symmetric, positive-definite solution to (17) and (18). Several conditions for the existence of such Riccati solutions for the steady-state case can be found in [28]. However, many of them cannot be used because (17) resembles Riccati equations used for $H_\infty$ control. In this section, conditions on the existence of nonnegative solutions to (17) and (18) are linked to the stability of $A - LC$. Further, it is shown that $\Pi_i$ is full rank when the eigenvectors of $A - LC$ do not line up perfectly with the nullspaces of $\hat{H}_i C$, $i = 1, \ldots, s$.

Assume that $\Pi_i \rightarrow 0$ as $t_1 \rightarrow \infty$. In that case, Problem 1 requires a solution $\Pi_i > 0$ to

$$0 \geq \Pi_i (A - LC) + (A - LC)^T \Pi_i + C^T \hat{H}_i Q_i \hat{H}_i C$$

$$+ \Pi_i \left( \frac{1}{\gamma} \dot{F}_i M_i \dot{F}_i^T + LV L^T \right) \Pi_i$$

(33)

for $i = 1, \ldots, s$. The following theorem states some results when $A - LC$ is asymptotically stable. The main result is that when there exists a nonnegative solution for $\Pi_i$, the solution is positive-definite when the complementary faults are allowed to be nearly but not completely unobservable in $(\hat{H}_i C, A - LC)$, $i = 1, \ldots, s$, which is generally the case when $\gamma > 0$. Therefore, $L$ should be chosen so that $A - LC$ is stable, the Hamiltonians $\hat{H}_i$, $i = 1, \ldots, s$ have no purely imaginary eigenvalues, and $(\hat{H}_i C, A - LC)$ is observable.

Theorem 1: Assume that (33) satisfies equality and that the associated Hamiltonian has no purely imaginary eigenvalues. If $A - LC$ is asymptotically stable, then there exists a real, symmetric, nonnegative solution $\Pi_i$. Further, $\text{Ker} \ \Pi_i = 0$ if $(\hat{H}_i C, A - LC)$ is observable.

Proof: The Hamiltonian of (33) is defined as [28]

$$H_i = \begin{bmatrix} A - LC & \bar{R}_i \\ -Q_i & -(A - LC)^T \end{bmatrix}$$

where

$$\bar{R}_i = \frac{1}{\gamma} \dot{F}_i M_i \dot{F}_i^T + LV L^T$$

$$\bar{Q}_i = C^T \hat{H}_i Q_i \hat{H}_i C.$$

From Theorem 13.6 of [28], when $H_i$ has no imaginary eigenvalues and $\bar{R}_i$ is a nonnegative symmetric matrix, there exists a real, symmetric, stabilizing solution $\Pi_i$ if and only
if \((A - LC, \bar{R}_i)\) is stabilizable. Since \(A - LC\) is stable, such a solution exists. Further, by rewriting (33) as
\[
0 \geq \Pi_i (A - LC) + (A - LC)^T \Pi_i
\]
\[
= -\Pi_i \left( \frac{1}{\gamma} F_i M_i F_i^T + LV L^T \right) \Pi_i - C^T \hat{H}_i Q_i \hat{H}_i C,
\]
it is clear that \(A - LC\) satisfies a nonpositive Lyapunov inequality. Since \(A - LC\) is stable, \(\Pi_i \geq 0\).

Finally, we show that the nullspace of \(\Pi_i\) is non-trivial only if \((\hat{H}_i C, A - LC)\) is unobservable by using part of the proof of Theorem 13.7 in [28]. Assume that \(\text{Ker} \Pi_i \neq 0\). Then there exists \(0 \neq x \in \text{Ker} \Pi_i\). Multiply (33) on the left by \(x^T\) and on the right by \(x\) to get
\[
\hat{H}_i C x = 0.
\]
Now multiply (33) on the right by \(x\) to get
\[
\Pi_i (A - LC) x = 0.
\]
Thus, \(\text{Ker} \Pi_i\) is an \((A - LC)\)-invariant subspace, and so there exists a \(\lambda\) such that \((A - LC) x = \lambda x\). By combining this with (34), \(x\) is an unobservable mode of \((\hat{H}_i C, A - LC)\).

IV. Numerical Example

In this section, a linear time-invariant numerical example for the F16XL aircraft [7], [22] is used to demonstrate the performance of the GTMFDF. The system has four states (longitudinal velocity \(x_u\), normal velocity \(x_w\), pitch rate \(x_q\), and pitch angle \(x_\phi\)), one control input (elevon deflection angle \(u_\delta\)), four measurements (longitudinal velocity \(y_u\), normal velocity \(y_w\), pitch rate \(y_q\), and pitch angle \(y_\phi\)), one disturbance input (wind gust \(\mu_{wg}\)), and sensor noise \(v\). The system matrices are
\[
A = \begin{bmatrix}
-0.0674 & 0.0430 & -0.8886 & -0.5587 \\
0.0205 & -1.4666 & 16.5800 & -0.0299 \\
0.1377 & -1.6788 & -0.6819 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}
\]
\[
B_\delta = \begin{bmatrix}
-0.1672 \\
-1.5179 \\
0 \\
0.0430 \\
-1.4666 \\
-1.6788 \\
0
\end{bmatrix}^T
\]
\[
B_{wg} = \begin{bmatrix}
0.0430 \\
-1.4666 \\
-1.6788 \\
0
\end{bmatrix}^T
\]
\[
C = I.
\]

Three faults in the pitch angle sensor \(y_\phi\), elevon deflector \(u_\delta\), and wind gust \(u_{wg}\) are considered, with fault directions
\[
F_\theta = \begin{bmatrix}
0 & 0 & 0 & 1.0000 \\
-0.5587 & -0.0299 & 0 & 0
\end{bmatrix}^T,
\]
\[
F_\delta = B_\delta, \quad F_{wg} = B_{wg}.
\]

It is desired to detect faults in the pitch angle sensor and elevon deflector in the presence of a wind gust disturbance and sensor noise. Thus, \(F_\theta\) and \(F_\delta\) are the target fault directions and \(F_{wg}\) is the nuisance fault direction. It can be verified that the faults are output separable, mutually detectable, and have no invariant zeros at the origin. Thus, two residual projectors are calculated using (7) to isolate the two target faults. The design weightings for the problem are chosen as
\[
\gamma = 10^{-4}, \quad \bar{V} = 10^{-6} I,
\]
\[
Q_i = M_i = K_i = I.
\]

To determine a filter gain that satisfies the DAPs (13), the MATLAB function "fminunc" is used to obtain a numerical solution to the optimization problem (26) for the integrand (30) given the Riccati constraint (24). The resulting detection filter dynamics are stable with eigenvalues at \(-4810, -11.15, -0.5346, -0.8598\). Then, assuming that \(u_\delta\) is nominally zero, the dynamic system (1) and detection filter (2) are integrated simultaneously for a unit bias fault at \(t = 2s\) in \(y_\phi\), \(u_\delta\), and \(u_{wg}\), sequentially. Fig. 1 shows response magnitudes of the two projected residuals to the three faults. The first projected residual is sensitive to the pitch angle sensor fault, shown as a dashed line. Further, it is insensitive to the elevon deflector and wind gust faults, shown as dotted and solid lines near zero, respectively. The performance is similar for the second projected, which is sensitive to the elevon deflector fault but insensitive to the pitch angle sensor and wind gust faults. Therefore, the GTMFDF sufficiently isolates the target faults and blocks the nuisance fault.

V. Conclusions and Future Work

To extend a previous robust fault detection filter to the multiple-fault case, the detection filter problem has been modeled as a set of disturbance attenuation problems. The game-theoretic multiple-fault detection filter has been derived and evaluated via a numerical example. When a solution is available, the detection filter meets the disturbance attenuation bound, implying that the faults are restricted to approximate invariant subspaces in the general case so that they can be isolated in the filter residual. The derivation above assumes scalar faults since the equations are more transparent. However, the results apply to both the scalar and vector fault cases.
In the future, the detection filter will be evaluated in the limit as the disturbance attenuation bound goes to zero. To implement the limiting case, the invariant subspace structure of the limiting detection filter must be examined. This will include an analysis of the invariant zeros in the time-invariant case. Further, it was shown that the detection filter problem obtains a non-trivial nullspace in the limit containing the invariant subspaces of the nuisances faults. Thus, a reduced-order model of the detection filter will be generated by truncating these invariant subspaces. Given the analysis above, the size of the reduced-order limiting GTMFDF is expected to be the same as reduced-order detection filters based on the spectral and geometric theories.

**References**


