Automatica 56 (2015) 70–77

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Parametrization of optimal fault detection filters[☆]

Xiaobo Li[a], Hugh H.T. Liu[b,1], Bin Jiang[c]

[a] General Stim Inc., 184 Technology Dr. Ste 105, Irvine, CA 92618, USA

[b] University of Toronto, Institute for Aerospace Studies, Toronto, Ontario, M3H 5T6, Canada

[c] College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, 210016, China

ARTICLE INFO

Article history:
Received 16 April 2013
Received in revised form
18 August 2014
Accepted 9 December 2014
Available online 11 April 2015

Keywords:
Fault detection
Multi-objective optimization
$\mathscr{H}_-$ index
Filter design
Robustness

ABSTRACT

This paper is concerned with the fault detection filter design problem for linear time-invariant systems subject to disturbance and possible faults. To make a tradeoff between fault sensitivity and disturbance sensitivity, frameworks such as $\mathscr{H}_-/\mathscr{H}_\infty$, $\mathscr{H}_2/\mathscr{H}_\infty$, and $\mathscr{H}_\infty/\mathscr{H}_\infty$ are considered. It is shown that for $\mathscr{H}_-/\mathscr{H}_\infty$ and $\mathscr{H}_\infty/\mathscr{H}_\infty$ frameworks, all optimal fault detection filters can be characterized by an envelope besieged by the special solutions serving as upper and lower bounds. Furthermore, the explicit parametrization is given in terms of a free contractive system/parameter. This free parameter provides freedom to emphasize fault and disturbance sensitivities at different frequency ranges, which can be used to construct new optimal fault detection filter. The results are also extended to the optimization over finite frequency range, for which the upper and/or lower bound of the optimal solution set may not be rational. On the other hand, it is shown that the solution set of $\mathscr{H}_2/\mathscr{H}_\infty$ problem is unique up to square inner system. Finally, examples are given to illustrate the results.

1. Introduction

Model-based fault diagnosis has attracted a great of interest in the last several decades (Chen & Patton, 1999; Ding, 2008; Frank & Ding, 1997). As the most critical part of fault diagnosis, fault detection is concerned with designing a filter called residual generator that generates residual signal to predict the occurrence of faults (Chen & Patton, 1999). The fault detection designs based on optimization have been proposed in the last twenty years (Casavola, Famularo, & Franze, 2005; Henry & Zolghadri, 2005; Khosrowjerdi, Nikoukhah, & Safari-Shad, 2005; Rank & Niemann, 1999). Specifically, in order to make this trade-off of two objectives: robustness to disturbance and sensitivity to fault, many design criteria and the corresponding techniques have been proposed (Ding, Jeinsch, Frank, & Ding, 2000; Hou & Patton, 1996; Jaimoukha, Li, & Papakos, 2006; Li & Liu, 2013; Li, Mazars, & Jaimoukha, 2006; Li & Zhou, 2009b; Liu, Wang, & Yang, 2005; Wang, Lam, Ding, & Zhong, 2005; Wang, Yang, & Liu, 2007).

Several criteria such as $\mathscr{H}_-/\mathscr{H}_\infty$, $\mathscr{H}_2/\mathscr{H}_\infty$ and $\mathscr{H}_\infty/\mathscr{H}_\infty$ problems have been employed in fault detection design recently. In Liu and Zhou (2007, 2008), a unified optimal filter of an observer form is obtained. The same unified solution has also been derived in Ding et al. (2000) for the different objective that maximizes the ratio of fault sensitivity and disturbance sensitivity, such as $\frac{\mathscr{H}_-}{\mathscr{H}_\infty}$ problem formulation. Furthermore, in Li and Zhou (2009a) and Li, Mou, and Zhou (2010). The result has been extended the result to linear time-varying systems based on the definitions and problem formulations in time domain. With the matrix factorization technique, a solution for $\mathscr{H}_\infty/\mathscr{H}_-$ framework was proposed in Jaimoukha et al. (2006), where the disturbance effect is minimized in terms of $\mathscr{H}_\infty$ norm, while the minimal fault sensitivity is constrained in terms of $\mathscr{H}_-$ index. However, several limitations and unanswered questions exist in those works. In N. Liu and K. Zhou's work (Liu & Zhou, 2007), only one 'unique' solution is given to the fault detection filter for all three problems. Similarly, the same 'unique' solution is given in S. X. Ding, T. Jeinsch and et al.'s work (Ding et al., 2000) for different frameworks such as $\frac{\mathscr{H}_-}{\mathscr{H}_\infty}$ (Jaimoukha et al., 2006). Most importantly, it is not easy to apply these filter designs to strictly proper systems. It also fails to consider fault and disturbance sensitivities at different frequency ranges. On the other hand, the fault detection designs via linear matrix inequality (LMI), always give merely one solution (Wang et al., 2007).

☆ This work is supported by the National Sciences and Engineering Research Council (NSERC) Discovery Grant of Canada (RGPIN 227674) and Natural Science Foundation of China (61273171, 61428303). The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Delin Chu under the direction of Editor Ian R. Petersen.

E-mail addresses: lixb@generalstim.com (X. Li), liu@utias.utoronto.ca (H.H.T. Liu), binjiang@nuaa.edu.cn (B. Jiang).

[1] Tel.: +1 416 667 7928.

http://dx.doi.org/10.1016/j.automatica.2014.12.043
0005-1098/

This work is to investigate the solution sets of the above problem frameworks and explore a method to construct new solutions. Based on a novel comparison relation for stable transfer matrices, we show that the optimal solutions of $\mathscr{H}_-/\mathscr{H}_\infty$ can be characterized by the upper bound and lower bound. In particular, the special solution given in Ding et al. (2000) and Liu and Zhou (2008) is only the upper bound. Furthermore, we also find that all solutions can be parametrized by a free contractive parameter/system in terms of combination of lower bound and upper bound. This free parameter provides freedom to construct other optimal solutions from the above two special solutions. By designing U properly, i.e., of a low-pass filter form, we can emphasize fault and disturbance sensitivities at different frequency ranges. The results have also been extended to finite frequency range, for which the upper bound and/or lower bound are not necessarily rational. The same parametrization is derived for $\mathscr{H}_\infty/\mathscr{H}_\infty$ problem which the lower bound may not be rational. In addition, we also show that the solutions of $\mathscr{H}_2/\mathscr{H}_\infty$ problem are unique up to square inner systems.

2. Notations and definitions

The set of $m \times n$ real (complex) matrices is denoted as $\mathscr{R}^{m\times n}$ ($\mathscr{C}^{m\times n}$). For a matrix $A \in \mathscr{C}^{m\times n}$, we use A^* to denote its conjugate transpose and A' to denote its transpose. For a matrix $A \in \mathscr{C}^{m\times n}$, we use $\bar{\sigma}(A)$ to denote the largest singular value of A and $\underline{\sigma}(A)$ to denote the smallest singular value of A if $m \geq n$. $A \in \mathscr{R}^{m\times n}$ is called tall (wide, or square) matrix if $m > n$ ($m < n$, or $m = n$). We denote $\mathscr{L}_2^n$ the set of all real energy bounded signals with dimension n. We use $\mathscr{RL}_\infty^{m\times n}$ to denote the set of all $m \times n$ real rational proper transfer matrices with no poles on the imaginary axis. The superscripts for dimensions will usually be dropped when they are either not important or clear from context. $\mathscr{RH}_\infty$ is a subset of $\mathscr{RL}_\infty$ with all stable transfer matrices. Similarly $\mathscr{RH}_2$ is the set of all real rational strictly proper stable transfer matrices. A state space realization of a transfer matrix $G(s)$ is denoted as $G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ such that $G(s) = C(sI - A)^{-1}B + D$. $G^\sim(s) := G'(-s)$ is the para-Hermitian complex conjugate transpose of G. G is called inner if $G \in \mathscr{RH}_\infty$ and $G^\sim G = I$. G is called co-inner if $G \in \mathscr{RH}_\infty$ and $GG^\sim = I$. $G \in \mathscr{RH}_\infty$ is called outer if all its transmission zeros are stable (Zhou, Doyle, & Glover, 1996). For $G \in \mathscr{RH}_2$ we define the $\mathscr{H}$ norm of G as (Zhou et al., 1996) $\|G\|_2 := \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty}\text{Trace}\{G^\sim(j\omega)G(j\omega)\}d\omega}$. For $G \in \mathscr{RH}_\infty$ we define the $\mathscr{H}_\infty$ norm of G as (Zhou et al., 1996) $\|G\|_\infty := \sup_{\omega\in\mathscr{R}}\bar{\sigma}(G(j\omega))$. For $G \in \mathscr{RH}_\infty$ the $\mathscr{H}_-$ index of G is defined as (Frank & Ding, 1997) $\|G\|_- := \inf_{\omega\in\mathscr{R}}\underline{\sigma}(G(j\omega))$. Two transfer matrices G and H are called equivalent up to square inner system if there exists a square inner system U such that $UG = H$. It is denoted by $G \equiv H$.

Definition 1. For two transfer matrices G and H in $\mathscr{RH}_\infty^{m\times n}$, we denote $G \trianglelefteq H$ ($G \trianglerighteq H$) if $G^\sim(j\omega)G(j\omega) \leq H^\sim(j\omega)H(j\omega)$ ($G^\sim(j\omega)G(j\omega) \geq H^\sim(j\omega)H(j\omega)$) holds for all $\omega \in \mathscr{R}\cup\infty$.

3. Plant and problem formulations

Consider plant:

$$y = G_u u + G_d d + G_f f \quad (1)$$

where $x(t) \in \mathscr{R}^n$ is the state vector, $y(t) \in \mathscr{R}^{n_y}$ is the output measurement, $u(t) \in \mathscr{R}^{n_u}$ represents the control input, $d(t) \in \mathscr{R}^{n_d}$ represents the unknown/uncertain disturbance and measurement noise, and $f(t) \in \mathscr{R}^{n_f}$ denotes the process, sensor or actuator fault vector. G_u, G_d, and G_f are $n_y \times n_u$, $n_y \times n_d$ and $n_y \times n_f$ transfer matrices respectively and their state-space realizations are $\begin{bmatrix} G_u & G_d & G_f \end{bmatrix} = \left[\begin{array}{c|ccc} A & B & B_d & B_f \\ \hline C & D & D_d & D_f \end{array}\right]$. We make the following assumptions. 1. (C, A) is detectable. 2. $n_d \geq n_y = n_f$. 3. $\begin{bmatrix} A - j\omega I & B_d \\ C & D_d \end{bmatrix}$ has full row rank for all $\omega \in \mathscr{R}$. Or, equivalently, the transfer matrix G_d has no transmission zero on the imaginary axis. 4. D_d has full row rank. 5. $\begin{bmatrix} A - j\omega I & B_f \\ C & D_f \end{bmatrix}$ has full row rank for all $\omega \in \mathscr{R}$. Or, equivalently, the transfer matrix G_f has no transmission zero on the imaginary axis. 6. D_f has full column rank.

Assumption 1 is a standard assumption for all fault detection problems. It will be seen later that Assumptions 3 and 4 are used to derive the stable and rational upper bound of the optimal solution set. With the removal of the two assumptions, the upper bound may be not stable as it may have pole on imaginary axis. However, this upper bound is a strictly tight upper bound, since we can use optimal stable transfer matrix to approximate it with arbitrary accuracy. In addition, our derived solution set still works and can be used to derive other solutions.

It will be seen later that Assumptions 5 and 6 are necessary to derive the stable and rational lower bound of optimal solution set. Actually, with removal of the two assumptions, the $\mathscr{H}_-$ index in infinite frequency range representing the weakest sensitivity is always zero. In other words, there always exist un-detectable faults. However, fault always happens in finite frequency range, and thus the two assumptions can be partially removed. Specifically, Assumptions 5 and 6 can be relaxed to be that $\begin{bmatrix} A - j\omega I & B_f \\ C & D_f \end{bmatrix}$ has full row rank for all ω in the frequency range that fault occurs, i.e., $[0, \omega_f]$. Obviously, it cannot be completely removed, otherwise, some fault may be undetectable.

Assumption 2 is the case we are investigated in this paper in terms of system dimensions. First of all, it is necessary to have $n_y = n_f$ in sensor fault detection as faults could happen in each output channel. Even if not, we have two cases, $n_y > n_f$ and $n_y < n_f$. For the former we can always choose proper outputs to make $n_y = n_f$. For the latter, theoretically, some fault cannot be detected, since it always holds that $\|QN_f\|_- = 0$. Thus, it makes no sense to characterize all solutions.

Second, we argue that $n_y \leq n_d$ is reasonable since noises always exist in every output channels in practice and thus we can treat them as extra disturbance input (see Liu & Zhou, 2007). Even for the case $n_d \leq n_f$, it is still possible to construct a square system $\hat{G}_f = \begin{bmatrix} G_f & G_c \end{bmatrix}$ where G_c is an extra system so that $\|\hat{G}_f\|_- \approx \|G_f\|_-$. Thus, we can use $\hat{G}_f$ to replace G_f in the derivation of lower bound. The essential of this augment is to introduce a fictitious fault f_c corresponding to G_c. To set G_c as large as possible, f_c may have very significant sensitivity and thus makes negligible effect to $\mathscr{H}_-$ index.

With Assumption 1, we have a left coprime factorization as $\begin{bmatrix} G_u & G_d & G_f \end{bmatrix} = M^{-1}\begin{bmatrix} N_u & N_d & N_f \end{bmatrix}$. As in Frank and Ding (1997) that the fault detection filter F can take the following general form $r = Q(My - N_u u) =: F\begin{bmatrix} y \\ u \end{bmatrix}$ where r is the residual vector for detection, $Q \in \mathscr{RH}_\infty^{n_y\times n_y}$ is a free stable transfer matrix to be designed. By computation, we have $r = QN_d d + QN_f f$. It can be seen that fault detection filter design becomes to design Q. We will call Q instead of F as optimal filter or solution.

To make a tradeoff between fault sensitivity and disturbance sensitivity, we consider the following problem formulations (Liu & Zhou, 2007) ($\|\cdot\|_i$ represents $\|\cdot\|_-$, $\|\cdot\|_2$ or $\|\cdot\|_\infty$).

Problem 1 ($\mathscr{H}_i/\mathscr{H}_\infty$ *Problem*). Let a system be described by Eq. (1) and let $\beta > 0$ be a given disturbance rejection level. Find a stable transfer matrix $Q \in \mathscr{RH}_\infty^{n_y\times n_y}$ such that $\|G_{rd}\|_\infty \leq \beta$ and $\|G_{rf}\|_-$ is maximized, i.e., $\max_{Q\in\mathscr{RH}_\infty^{n_y\times n_y}}\left\{\|QN_f\|_i : \|QN_d\|_\infty \leq \beta\right\}$.

4. Parametrization for $\mathscr{H}_-/\mathscr{H}_\infty$ framework

By Zhou et al. (1996), with Assumptions 1, 2, 3 and 4, we have the co-inner–outer factorization for N_d, i.e., $N_d = V_d I_d$ where I_d is a co-inner. With Assumptions 1, 2, 5 and 6, we have the co-