Fault Diagnosis for Non-minimum Phase Systems using $\mathcal{H}_\infty$ Optimization.

Henrik Niemann

Ørsted•DTU, Automation
Technical University of Denmark
Building 326
DK-2800 Lyngby, Denmark.
E-mail: hhn@oersted.dtu.dk

Jakob Stoustrup

Dept. of Control Engineering
Aalborg University
Fr. Bajers Vej 7C
DK-9220 Aalborg Ø, Denmark
E-mail: jakob@control.auc.dk
URL: www.control.auc.dk/~jakob

Abstract

Analysis and design algorithms for residual generators for non minimum phase systems are given. It is shown that $\mathcal{H}_\infty$ optimization of residual generators applied directly to systems including non minimum phase zeros can be very conservative. To remove this conservatism in the $\mathcal{H}_\infty$ optimization of the residual generators, a factorization of the non minimum phase system into a minimum phase part and an all-pass factor including the non minimum phase zeros can be applied. The optimization of the residual generator can then be done with respect to the minimum phase part of the system only. It is shown that the effect from the all-pass factor will not affect the 2-norm of the residual vector.

1 Introduction

In the area of model-based fault diagnosis, the design of residual generators can be derived in a large number of different ways, as e.g. design methods based on $\mathcal{H}_2$ and $\mathcal{H}_\infty$ optimization, eigenstructure assignment, state/output estimation, parameterization, parameter identification, parity equations, and a host of others. An introduction to the area of model-based fault diagnosis can be found in the books by Gertler [4], Chen and Patton [1] and by Mangoubi [5].

Design of residual generators using different $\mathcal{H}_\infty$ based optimization methods has been very attractive, see e.g. [3, 8, 12, 14]. One of the reasons for this, is the attractive properties $\mathcal{H}_\infty$ optimization methods have with respect to robustness. This has turned out to be very important in connection with feedback control of uncertain systems, [15]. It has been shown in the papers mentioned above, that $\mathcal{H}_\infty$ optimization of residual generators can be applied with advantages. Using $\mathcal{H}_\infty$ optimization, it is possible to give upper bounds on the norm of the residual vector/estimation error. These bounds are very useful in connection with the selection of the threshold values for the residual vector. However, there is a number of limitations in connection with using $\mathcal{H}_\infty$ optimization of the residual generator.

One of these cases is when the system includes non minimum phase zeros, [13]. In many cases, this limitation will result in unacceptable bounds on the residual vector/estimation error, depending on the location of the non minimum phase zeros. The main problem is that design based on an $\mathcal{H}_\infty$ method will optimize the worst case situation only. For non minimum phase systems, the worst case will be in the input/output directions for the non minimum phase zeros and at those specific frequencies, where there exist some interpolation constraints on the closed loop transfer functions, [13]. These interpolation constraints will in general spoil an $\mathcal{H}_\infty$ optimization, if nothing is done to remove/minimize this. The most direct way to handle interpolation constraints from non minimum phase zeros is to include weighting matrices in the design problem. Another approach, as we will suggest in this paper, is to apply a factorization of the non minimum phase systems into a minimum phase part and an all-pass part that include the non minimum phase zeros. The key result in this paper is that it is possible to apply only the minimum phase part of the system for the $\mathcal{H}_\infty$ optimization of residual generators without affecting the 2-norm of the residual vector/estimation error, when the residual generator is applied on the real system.

The rest of this paper is organized as follows. In Section 2, the system setup is given along with a number of definitions and a preliminary analysis of the use of the $\mathcal{H}_\infty$ norm for non minimum phase systems. An analysis of the effect of non minimum phase zeros on the sensitivity and the complementary sensitivity functions for the (fault diagnosis) filtering problem is considered in Section 3. In Section 4 a method is presented for designing residual generators using $\mathcal{H}_\infty$ optimization for non minimum phase systems, where the effect from the non minimum phase zeros are removed/minimized by using factorization of the non minimum phase systems. A conclusion is given in Section 5.
2 System Setup and Problem Formulation

Consider the following state space description for a plant or a system given by

\[
\Sigma : \begin{cases}
\sigma x = Ax + Ed + L_1 f \\
y = Cx + D_d f + D_1 f
\end{cases}
\]  

(1)

\(\sigma\) is an operator indicating the time derivation \(\frac{df}{dt}\) for continuous-time systems and a forward unit time shift for discrete-time systems, \(x \in \mathbb{R}^n\) is the state vector, \(d \in \mathbb{R}^m\) is a disturbance signal vector, and \(y \in \mathbb{R}^p\) is the measurement vector. The fault signal vector \(f \in \mathbb{R}^k\) is a collection of fault signals \(f_i, i = 1, 2, \ldots, k\), into a vector. Further, the coefficient matrices \(L_1\) and \(D_1\) are referred to in the literature as failure signatures associated with the fault vector \(f\).

The system setup given in (1) can be rewritten in a transfer function form given by:

\[y(\alpha) = G_f(\alpha)f(\alpha) + G_d(\alpha)d(\alpha)\]

where \(\alpha\) should be interpreted as a complex variable introduced by either Laplace or Z-transform. We now proceed to formulate certain fault estimation (detection and/or isolation) problems.

Let the residual signal \(r\) be given by

\[r(\alpha) = H(\alpha)y(\alpha) = \Psi(d, f)(\alpha)\]

(2)

where \(r\) is a time function that takes values in \(\mathbb{R}^q\). In general, we might have to take \(H\) to be a nonlinear bounded-input, bounded-output stable operator which makes \(\Psi\) also a nonlinear operator mapping disturbances and faults to a residual signal \(r\). Of course, if \(H\) is linear then there exist transfer matrices \(G_{rf}\) and \(G_{rd}\) such that

\[r(\alpha) = G_{rf}(\alpha)f(\alpha) + G_{rd}(\alpha)d(\alpha)\]

where \(G_{rf} = HG_f\) and \(G_{rd} = HG_d\).

One of the basic issues that concerns fault detection, isolation and estimation in general system setup is whether one can achieve such a detection, isolation or an estimation when the disturbance \(d\) affects the system. This points out a need to have a residual generator which is insensitive to the external disturbance \(d\). That is, we need that

\[\Psi(d, f) = \Psi(0, f)\]

for all disturbances \(d\) and all fault signals \(f\) or at least that the dependence of \(r\) on \(d\) can be made arbitrarily/sufficiently small with respect to some specified norm. If \(H\) is linear then this implies that we impose that the transfer matrix \(G_{rd}\) is zero or arbitrarily small in some specific norm.

Before we continue, let us give the definition of fault detection and fault isolation, [7].

**Definition 1** Given the residual generator \(H \in \mathcal{RH}_\infty\), the residual \(r\) is said to achieve fault detection (FD) without disturbance if any non-zero fault vector \(f\) and \(d \equiv 0\) results in a non-zero residual \(r\).

**Definition 2** Given the residual generator \(H \in \mathcal{RH}_\infty\), the residual \(r\) is said to achieve fault detection and isolation (FDI) without disturbance if for any two different fault vectors \(f_1\) and \(f_2\) and all different fault vectors \(d\), the corresponding residuals \(r_1\) and \(r_2\) are different.

**Definition 3** Given the residual generator \(H \in \mathcal{RH}_\infty\), the residual \(r\) is said to achieve robust fault detection with respect to some fault set \(F\) and some disturbance set \(D\) if there exists a threshold \(\tau\) such that

\[f \in F \Rightarrow \|Hy\|_2 > \tau\]

\[f_1 \equiv 0 \Rightarrow \|Hy\|_2 < \tau\]

for any \(d \in D\).

**Definition 4** Given the residual generator \(H \in \mathcal{RH}_\infty\), the residual \(r\) is said to achieve robust fault detection with respect to some fault set \(F\) and some disturbance set \(D\) if there exists a threshold \(\tau\) such that

\[f \in F \Rightarrow \|Hy\|_2 > \tau\]

\[f_1 \equiv 0 \Rightarrow \|Hy\|_2 < \tau\]

for any \(d \in D\) and \(H_i\) is the operator from \(y\) to \(r_i\).

It will be assumed in the rest of this paper that only a single fault can appear at any time. In general, the results presented in the rest of this paper can be generalized to allow faults occurring simultaneously without further conditions.

Let us consider the standard estimation approach considered in e.g. [10] or in [15]. Using the general system setup from [10] given by:

\[
\sum_g : \begin{cases}
\sigma z = Ax + Bu \\
y = Cy + Du \nu \\
z = C_2 z + Fu \nu
\end{cases}
\]

(3)

where \(\nu \in \mathbb{R}^m\) indicate an input vector to the system, and \(z \in \mathbb{R}^p\) is the desired output vector to be estimated. The estimation problem is then to estimate the external output \(\hat{z}\) by using the filter \(H\) given by:

\[\hat{z} = Hy\]

such that the difference between \(z\) and \(\hat{z}\) is minimized in a suitable way. Let the estimation error \(e_z\) be defined as the difference between the external output \(z\) and the estimated output \(\hat{z}\), i.e.

\[e_z = z - \hat{z}\]

To be more precise, let us consider an \(H_\infty\) problem formulation of the estimation problem. The \(H_\infty\) problem formulation is given by, [16]:

\[4433\]
Problem 1 $\mathcal{H}_\infty$ Filtering. Given a $\gamma > 0$, find a causal filter $H \in \mathcal{R} \mathcal{H}_\infty$, if it exists, such that the $\mathcal{H}_\infty$ norm of the transfer function matrix from $v$ to $e$ is smaller than or equal to $\gamma$.

The reason is that the non-minimum phase zeros give a useful residual generator. However, in the case where $H$ or $G_f$ or both are strictly proper, the $\mathcal{H}_\infty$ filtering problem given in Problem 1 is not restricted to include the fault estimation case. Instead of using $F_e$ as given in (4), we can use $F_e = [V\ 0]$ where the design problem depends on the selection of $V$. (5) then takes the following form:

$$\| [I - HG_f - HG_d] \|_\infty < \gamma$$

(6)

The fault estimation problem is obtained by using $V = I$. Both fault detection as well as fault isolation can be obtained by the selection of $V$.

In both the fault detection case as well as in fault isolation case, the $V$ matrix is a free design matrix, [9]. $V$ can also be a dynamical matrix. In the following, it will be assumed that $V$ is fixed (static or dynamical).

In the following, the term fault diagnosis will be used for the above design problem, where both fault detection, fault isolation and fault estimation can be obtained depending on the structure of $V$.

Due to the direct term in the $\mathcal{H}_\infty$ fault diagnosis problem in (6), a weight matrix should be included for the solution of the problem to be meaningful, [6, 8, 14]. Premultiplying (6) with a weight matrix $W$ gives

$$\| [W(V-HG_f) - WHG_d] \|_\infty < \gamma$$

(7)

If $W$ is selected as a strictly proper transfer matrix, the direct term in (6) (i.e. the $D_{11}$ term in the standard setup) has been removed.

Using $\mathcal{H}_\infty$ optimization for the design of the residual generator $H$ in (6) or in (7) is an attractive method, [2, 3, 8, 12]. However, in the case where $G_f$ include non-minimum phase zeros, the $\mathcal{H}_\infty$ optimization method will not in general result in a useful residual generator. The reason is that the non-minimum phase zeros give interpolation constraints which in turn imply a lower bound on $\gamma$, [15]. Let $q$ be a non-minimum phase zero of $G_f$. Then a lower bound on $\gamma$ is given by

$$\gamma > \| [V - HG_f - HG_d] \|_\infty$$

$$= \sup_{s \in C^+} \| [V - H(s)G_f(a) - H(s)G_d(a)] \|$$

$$\geq \| [V - H(q)G_f(q) - H(q)G_d(q)] \|$$

for the discrete time case and equivalently for the continuous time case by replacing the half plane with the unit circle. Similarly, by using (7) we obtain

$$\gamma > \| [W(V-HG_f) - WHG_d] \|_\infty$$

$$= \sup_{s \in C^+} \| [W(s) [V - H(s)G_f(s) - H(s)G_d(s)] \|$$

$$\geq \| [W(q)(V - H(q)G_f(q)) - W(q)H(q)G_d(q)] \|$$

(8)

From (8), it can be seen that $\gamma$ will be larger than 1 if $H$ or $G_f$ or both are strictly proper as they would usually be (zeros at infinity), which imply that the estimation error can be more than 100% (it is assumed that $\|V\|_\infty = 1$. From (9) $\gamma$ has to be larger than $\|W(q)V\|$. If $\|W(q)V\|$ is not small, $\gamma$ will also in this case be unacceptable large.

Non-minimum phase zeros in a MIMO system will have both input and output directions. The result of this is that the effect from a non-minimum phase zero can be seen in some directions and not in others. With respect to fault diagnosis, a non-minimum phase zero will not affect the diagnosis of all faults, in general only some of them will fail to be diagnosed. However, using a standard $\mathcal{H}_\infty$ optimization method, the worst case will be optimized. Therefore, the effect from a non-minimum phase zero will indirectly affect the diagnosis of all faults. This is in general not acceptable. There is a number of ways to overcome this. One way is to select the weighting matrix $W$ in (7) to include the same non-minimum phase zeros as $G_f$. This requires that both input and output directions for the zeros are identical with the directions of $G_f$. The other way to overcome the problem is to make a factorization of $G_f$ in a minimum phase part and a non-minimum phase part as will be shown in the sequel.

3 Analysis of Fault Diagnosis

An analysis of the effect from non minimum phase zeros in $G_f$ on the fault diagnosis problem will shortly be given in the following.

In the same line as for feedback control, sensitivity and complementary sensitivity functions can be designed, [13]. The sensitivity function $S$ and the complementary sensitivity function $T$ for the fault diagnosis problem are given by:

$$S(a) = (VHG_f(a))V^{-1},\ T(a) = HG_f(a)V^{-1}$$

(10)

The sensitivity function is important, because the sensitivity function is included in the transfer function for the error, given by

$$e_s = S(a)f - HG_d(a)d$$
Let \( q \) be a non minimum phase zero of \( G_f \). Then, there exist a non-zero vector \( \phi \) such that
\[
S(q)\eta = \phi, \quad T(q)\phi = 0
\]
The relation between the zero input direction \( \eta \) and \( \phi \) is given by:
\[
\eta = V^{-1}\phi
\]
It can be seen directly from the above equation, that a non minimum phase zero will not affect all input directions in the transfer function from the fault vector \( f \) to the fault error \( e_d \). Together with the result on the lower bound on the \( H_\infty \) norm given in (8), it is clear that non minimum phase zeros results in a limitation in the performance for the derived residual generators, see also \([13, 15]\). Further, from (8), we have that an \( H_\infty \) optimization directly of the fault diagnosis problem when \( G_f \) include non minimum phase zeros is not useful. At least a weighting matrix as shown in (9) needs to be included.

4 Design of Residual Generators

The analysis of non minimum phase systems in Section 3 will be applied in this section in connection with formulation of an \( H_\infty \) design problem for residual generators. As pointed out in Section 2, using \( H_\infty \) optimization methods directly on design problems involving non minimum phase zeros can be very conservative. However, using a factorization of \( G_f \), it is possible to overcome/reduce the effect from non minimum phase zeros in the \( H_\infty \) optimization of the residual generators without affecting the optimality of the residual generator significantly.

Before we continue, we need to give the following definition concerning the weighting matrix \( W(\alpha) \) to be selected.

**Definition 5** A filter \( W(\alpha) \) is said to be a \((\mathcal{F}, \beta, \delta)\)-compatible weighting if for any fault signal \( f \in \mathcal{F} \) the following inequality is satisfied:
\[
\|WB_m f\|_2 > \beta \|f\|_2 + \delta
\]
where \( B_m \) is a right all-pass factor of \( G_f \):
\[
G_f(\alpha) = G_{m,f}(\alpha)B_m(\alpha)
\]
chosen such that \( G_{m,f} \) is minimum phase.

A fault set \( \mathcal{F} \) that allows a \((\mathcal{F}, \beta, \delta)\)-compatible weighting is called \((\beta, \delta)\)-separable. A fault set \( \mathcal{F} \) that allows a non-trivial \((\mathcal{F}, \beta, \delta)\)-compatible weighting, i.e. a \((\mathcal{F}, \beta, \delta)\)-compatible weighting for some \( \beta > 0 \) and \( \delta > 0 \) is called separable.

We consider the system in Figure 1, where \( d \) is assumed to be a norm bounded disturbance,
\[
\|d\|_2 < d_{\text{max}} \tag{11}
\]

**Figure 1:** FDI system with factorization

Assume that \( H \) satisfies:
\[
\|W(V - HG_{m,f})\|_{\infty} < \varepsilon_f \tag{12}
\]
and
\[
\|WHG_d\|_{\infty} < \varepsilon_d \tag{13}
\]

Note, that if it exists, such a filter can be found by \( H_\infty \) optimization by solving a standard \( H_\infty \) filtering problem (see e.g. \([15]\) with the following data:
\[
\begin{bmatrix}
G_{zw} & G_{zr} \\
G_{yw} & G_{yr}
\end{bmatrix} =
\begin{bmatrix}
-\frac{1}{\varepsilon_f}WV & 0 \\
\frac{1}{\varepsilon_f}G_{m,f} & \frac{1}{\varepsilon_d}G_d
\end{bmatrix} W
\]
where \( w = [f^T \ d^T]^T \). Clearly, from (11) and (13), we can see that an observation of the output of the filter satisfying:
\[
\|Wd\|_2 > \varepsilon_d d_{\text{max}}
\]
leads to the conclusion that a fault has occurred.

From (12) and (13) we have:
\[
\|Wr\|_2 = \|WVf + W(HG_{m,f} - V)f + WHG_dd\|_2
\geq \|WVF\|_2 - \|W(HG_{m,f} - V)f\|_2 - \|WHG_dd\|_2
\geq \|WVF\|_2 - \varepsilon_f \|f\|_2 - \varepsilon_d \|d\|_2
= \|WB_m f\|_2 - \varepsilon_f \|f\|_2 - \varepsilon_d \|d\|_2
\]
from the triangular inequality, where it has been exploited that \( B_m \) is inner (all-pass).

Now, assuming that any nonzero fault \( f \) satisfies
\[
\|WB_m f\|_2 > \varepsilon_f \|f\|_2 + 2\varepsilon_d \|d\|_2
\]
it follows that
\[
\|Wr\|_2 > \varepsilon_d d_{\text{max}}
\]
which allow us to conclude from the measurement that a fault has occurred.

From this we obtain the following:

**Theorem 1** A system \( \Sigma \) is robustly detectable with respect to some fault set \( \mathcal{F} \) and some fault set \( D \), if and only if \( \mathcal{F} \) is separable, and there exists a \( W \) and a solution to the \( H_\infty \) standard problem (14) for which \( W \) is a \((\mathcal{F}, \varepsilon_f, 2\varepsilon_d d_{\text{max}})\)-compatible weighting.
In practice, this result should be used by fixing $W$ in an 'optimal' way relative to $F$, or alternatively to determine a suitable $W$ by iteration.

The cases where $\varepsilon_f$ and/or $\varepsilon_d$ are zero or arbitrary small are special cases of the above general case. These cases have been considered explicitly in [7, 11], where a detailed analysis are given along with solvability conditions for a number of fault detection, fault isolation and fault estimation problems. In the case of fault estimation, ($V = I$), it is possible to obtain exact fault estimation, ($\varepsilon_f = \varepsilon_d = 0$) and almost exact fault estimation, ($\varepsilon_f$ and $\varepsilon_d$ are arbitrary small) under different restricted solvability conditions, [7]. In contrast to this, in the fault detection case and the fault isolation case, there is no difference between the solvability conditions for obtaining e.g. exact fault detection and almost exact fault detection, [11]. These solvability conditions can be used in connection with the optimization of the residual generator for the system given by $(G_{m,f}, G_d)$, especially in the fault estimation case.

It was assumed in Section 2 that $V$ is a fixed matrix in the optimization of the residual generator $H$. Except in the fault estimation case, where $V$ is given by $V = I$, the selection/design of $V$ should be included in the design of the residual generator. In the fault detection case and in the fault isolation case, only the structure of $V$ is fixed. It is not possible to include $V$ directly in the design problem. Instead, the design of $V$ in connection with the design of the residual generator $H$ can be done by iteration. In [9], two different iterative approaches are given for the design/selection of $V$ in connection with the residual generator $H$.

5 Conclusion

The problem of designing residual generators for non minimum phase systems using $H_\infty$ optimization has been considered. It has been shown that non minimum phase zeros in the transfer function from fault vector to measurement vector will give unnecessarily hard limitations in the performance for an $H_\infty$ optimised residual generator. These limitations can be removed by considering only the minimum phase part of the transfer function from the fault vector to the measurement vector. The only price for this might be an increased detection time for faults appearing in the direction of the non-minimum phase zeros. However, applying a minimum phase factorization on this transfer function in connection with an $H_\infty$ optimization of a residual generator will not affect the 2-norm of the (weighted) residual vector.

References


