Disturbance decoupled fault reconstruction using sliding mode observers


Published in:
Asian Journal of Control

Publication Status:
Accepted/In press: 01/07/2009

DOI:
10.1002/asjc.231

Document Version
Publisher's PDF, also known as Version of record

General rights
Copyright for the publications made accessible via Ulster University’s Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Ulster University’s institutional repository that provides access to Ulster’s research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person’s rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact pure-support@ulster.ac.uk.
DISTURBANCE DECOUPLED FAULT RECONSTRUCTION USING SLIDING MODE OBSERVERS

Kok Yew Ng, Chee Pin Tan, Rini Akmeliawati, and Christopher Edwards

ABSTRACT

This paper investigates and presents conditions that guarantee disturbance decoupled fault reconstruction using sliding mode observers, which are less stringent than those of previous work, and show that disturbance reconstruction is not necessary. An aircraft model validates the ideas proposed in this paper.

Key Words: Decoupling, fault reconstruction, sliding mode observer.

I. INTRODUCTION

Disturbance decoupled fault reconstruction (DDFR) is the ability to generate an accurate reconstruction of a fault that is totally insensitive to disturbances. Edwards et al. [1] used a sliding mode observer [2] to reconstruct faults, but there was no explicit consideration of the disturbances. Tan and Edwards [3] built on their work and designed an algorithm for the observer to minimize the $L_2$ gain from the disturbances to the fault reconstruction. However, total decoupling from the disturbances was not guaranteed. Saif and Guan [4] combined the faults and disturbances to form a new augmented ‘fault’ vector and used an unknown input observer to reconstruct the new ‘fault’, including the disturbances. Although this successfully decouples the disturbances from the fault reconstruction, it requires very stringent conditions to be fulfilled, and is conservative because the disturbances do not need to be reconstructed, only rejected/decoupled. Edwards and Tan [5] later compared the fault reconstruction performances of [1] and [4], and found that disturbance reconstruction was not necessary to achieve DDFR. A counter example was presented in [5], but the conditions for disturbance decoupling were not formally investigated.

This paper extends the work in [5], and its main contribution is the investigation of conditions that guarantee DDFR. It is found that the conditions that guarantee DDFR are less stringent than those in [4] which prove that disturbance reconstruction is not necessary for DDFR. In addition, the conditions in this paper are easily testable on the original system matrices, making it possible to determine immediately whether DDFR is feasible. An aircraft system will be used to validate the results in this paper. The paper is organized as follows: Section II introduces the system and sets up the framework for the investigation of the existence conditions; Section III investigates the conditions for DDFR to be achieved; an example to validate the conditions is given in Section IV and finally Section V makes some conclusions.

II. PRELIMINARIES AND PROBLEM STATEMENT

Consider systems having the following form:

$$\dot{x} = Ax + Mf + Q\zeta, \quad y = Cx$$

(1)

where $x \in \mathbb{R}^n$, $y \in \mathbb{R}^p$, $f \in \mathbb{R}^q$ and $\zeta \in \mathbb{R}^h$ are the states, outputs, faults and disturbances respectively. The signal $\zeta$ represents the mismatches between the model and the actual plant. Assume $p > q$. 

© 2010 John Wiley and Sons Asia Pte Ltd and Chinese Automatic Control Society
Tan and Edwards [3] proposed a scheme which minimized the effect of $\xi$ on the fault reconstruction if and only if the following are satisfied:

A1. $\text{rank}(CM) = \text{rank}(M)$
A2. $(A, M, C)$ is minimum phase.

The method in [3] however, does not totally reject $\xi$. Saif and Guan [4] totally reject $\xi$ by combining $\xi$ and $f$ to form $\hat{f} = [\xi^T, f^T]^T$ and then design a fault reconstruction scheme to reconstruct $\hat{f}$. One of the necessary and sufficient conditions for their scheme is:

B1. $\text{rank}(CMQ) = \text{rank}(CMQ) - \text{rank}(CQ)$

Since only the reconstruction of $f$ is required, reconstructing $\hat{f}$ is conservative. This paper investigates conditions that guarantee DDFR with less stringent conditions than [4]. To this end, define $k := \text{rank}(CQ)$ where $k \leq h$. Assume A1 holds as well as

N1. $\text{rank}(CMQ) = \text{rank}(CMQ) + \text{rank}(CQ)$

**Proposition 1.** If A1 and N1 hold then there exist transformations $z \mapsto z_2x$ and $\xi \mapsto x_1^{-1}\xi$ such that $C = [0 C_2]$ and $(A, M, Q)$ have the following structure:

$$
A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}, \quad M = \begin{bmatrix} 0 \\ M_2 \end{bmatrix}, \quad Q = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}
$$

(2)

where $A_4 \in \mathbb{R}^{n \times p}$, $M_2 = [0 M_2^T]^T$ and

$$
Q_1 = \begin{bmatrix} \bar{Q}_1 & 0 \\ 0 & 0 \end{bmatrix} \oplus_{h-k} p-q-k, \quad Q_2 = \begin{bmatrix} 0 \\ \bar{Q}_2 \end{bmatrix} \oplus_q k
$$

(3)

where $M_o, C_2, \bar{Q}_1$ and $\bar{Q}_2$ are square and invertible.

**Proof.** See Proposition 1 of [6].

### 2.1 Sliding mode observer for fault reconstruction

The scheme described in this paper system (1) will use a sliding mode observer [2] of the form

$$
\dot{\hat{x}} = A\hat{x} - G_I e_y + G_n v, \quad \hat{y} = C\hat{x}
$$

(4)

where $e_y = \hat{y} - y$. The term $v$ is a nonlinear discontinuous term defined by $v = -\rho \frac{e_y}{\|e_y\|}$ where $\rho \in \mathbb{R}_+$. The matrices $G_I, G_n \in \mathbb{R}^{n \times p}$ are the observer gains to be designed. In the coordinates of (2), $G_n$ is assumed to have the following structure:

$$
G_n = \begin{bmatrix} -L \\ I_p \end{bmatrix} (P_oC_2)^{-1}, \quad L = \begin{bmatrix} L_o & 0 \\ p-q & q \end{bmatrix} \oplus_{n-p} q
$$

(5)

where $P_o = P_o^T > 0$. By analysing the error arising from (4) and (1), performing the coordinate transformation in (4) from [3], and assuming that the gains $G_I, G_n$ have been well designed to attain a sliding motion, the error equations become (see §2.2 of [3])

$$
\dot{e}_1 = (A_1 + L A_3) e_1 - (Q_1 + L Q_2) \xi
$$

(6)

$$
0 = C_2 A_3 e_1 + P^{-1}_o v_{eq} - C_2 M_2 f - C_2 Q_2 \xi
$$

(7)

where $v_{eq}$ is the equivalent output error injection to maintain the sliding motion, which can be approximated to any degree of accuracy by replacing the denominator of $v$ with $\|e_y\| + \delta$ where $\delta$ is a small positive scalar.

Define a fault reconstruction signal to be $\hat{f} = C \dot{\hat{f}} - f$. From (7) it follows

$$
e_f = -W A_3 e_1 + W Q_2 \xi
$$

(8)

Ideally, $e_f = 0$ (i.e. $f = \hat{f}$), but from (6) and (8), it is clear that $\xi$ excites $e_f$ and that the design freedom is represented by $L_o$ and $W_1$. The objective is to decouple $e_f$ from $\xi$ by choice of $L_o$ and $W_1$. Define $Q_o$ to be the left $h-k$ columns of $Q$ in (2). The following theorem states the main result of the paper:

**Theorem 1.** Suppose assumptions A1 and N1 hold. Then $e_f$ will be decoupled from $\xi$ by choice of $L_o$ and $W_1$ if the following conditions are satisfied:

C1. $\text{rank}(\Phi) - \text{rank}(M) = \text{rank}(CMQ) = \text{rank}(\Omega) - \text{rank}(Q)$
C2. $(A, [M Q], C)$ is minimum phase.

where $\Phi = [CAQ_o CM CQ]$ and $\Omega = [AQ_o Q]$. Note that C1 and C2 are easily testable conditions in terms of the original system matrices. In addition C1 is not as stringent as B1. The following section provides a constructive proof of Theorem 1.

### III. PROOF OF THEOREM 1

From (6) and (8), $e_f$ is the sum of a filtered and scaled version of $\xi$. Therefore, there is a need for these two parameters to be zero in order to achieve $e_f = 0$. The term $W Q_2$ is the direct feedthrough in (6) and (8). Partition $Q_2$ from (3) as

$$
Q_2 = \begin{bmatrix} 0 & 0 \\ 0 & \bar{Q}_2 \\ 0 & 0 \end{bmatrix} \oplus_{p-q} \oplus_q
$$

(9)
It is obvious that \( W_1 Q_{21} = W Q_2 = 0 \) is satisfied by
\[
W_1 = [W_{12} \ 0], \quad W_{12} \in \mathbb{R}^{q \times (p-q)} \tag{10}
\]

**Remark 1.** Recall that \( \text{rank}(CQ) = k \) by definition. If \( N_1 \) is not satisfied, then \( \text{rank}(C[M \ 
Q]) < \text{rank}(CM) + \text{rank}(CQ) \) which then results in \( \text{rank}(Q_{21}) < \text{rank}(CQ) = k \). From the expression for \( \hat{f} \), it follows that to satisfy \( W Q_2 = 0 \) requires \( W_1 Q_{21} = -M_o^{-1} Q_{22} \) which in turn requires
\[
\text{rank}(Q_{21}) = \text{rank} \begin{bmatrix} Q_{21} \\ Q_{22} \end{bmatrix} = \text{rank}(CQ) = k \tag{11}
\]
which is not satisfied if \( N_1 \) is not satisfied. Therefore, \( N_1 \) is necessary for \( W Q_2 = 0 \). Furthermore, \( N_1 \) is sufficient for \( W Q_2 = 0 \) as it enables the transformation in Proposition 1 to be attained.

To decouple the filtered version of \( \xi \) through to \( e_f \), partition \( A_1, A_3 \) from (2) as
\[
A_1 = \begin{bmatrix}
A_{11} & A_{12} & \hat{h} \\
A_{13} & A_{14} & \hat{h} \\
\end{bmatrix} \quad A_3 = \begin{bmatrix}
A_{31} & A_{32} & \hat{h} \\
A_{33} & A_{34} & \hat{h} \\
A_{35} & A_{36} & \hat{h} \\
\end{bmatrix} \tag{12}
\]

From (10) and (12), \( W A_3 \) can be written as
\[
\tilde{C} = [-W_1 A_{31} - M_o^{-1} A_{35} - W_1 A_{32} - M_o^{-1} A_{36}] \tag{13}
\]

Partition \( L_0 = \begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix} \). Then \((A_1 + L A_3)\) and \((Q_1 + L Q_2)\) in (6), when expressed using (12)–(13), will respectively produce
\[
\begin{aligned}
\tilde{A} := \begin{bmatrix}
A_{11} + L_{11} A_{31} + L_{12} A_{33} & A_{12} + L_{11} A_{32} + L_{12} A_{34} \\
A_{13} + L_{21} A_{31} + L_{22} A_{33} & A_{14} + L_{21} A_{32} + L_{22} A_{34} \\
\end{bmatrix} \\
\tilde{B} := \begin{bmatrix}
\hat{b}_1 \\
\hat{b}_2 \\
\hat{b}_3 \\
\end{bmatrix} = \begin{bmatrix}
\hat{Q}_1 \\
L_{12} \hat{Q}_2 \\
L_{22} \hat{Q}_2 \\
\end{bmatrix} \tag{14}
\end{aligned}
\]

Partition \( e_1 = [e_{11}^T \ e_{12}^T]^T \) where \( e_{11} \in \mathbb{R}^{h-k} \) and \( \xi = [\xi_{e1}^T \ \xi_{e2}^T]^T \) where \( \xi_{e1} \in \mathbb{R}^{h-k} \). Then partition the matrices \( \tilde{A}, \tilde{C} \) conformally with (13)–(14) as
\[
\tilde{A} = \begin{bmatrix}
\tilde{A}_1 \\
\tilde{A}_2 \\
\end{bmatrix}, \quad [\tilde{C}_1 \ \tilde{C}_2].
\]
It can be observed that \( e_{11} \) will always be affected by \( \xi_{e1} \) because \( \hat{Q}_1 \neq 0 \). However, \( e_{12} \) can be decoupled from \( \xi \) by setting \( L_{22} = 0 \) and \( \tilde{A}_3 = 0 \). In order for \( e_f \) not to be influenced by \( e_{11} \), \( \tilde{C}_1 = 0 \) is necessary. With \( L_{22} = 0 \), to make \( \tilde{A}_3 = 0 \) and \( \tilde{C}_1 = 0 \), requires \( \text{rank}(A_{31}) = \text{rank}[A_{13}^T A_{31}^T]^T \) and \( \text{rank}(A_{31}) = \text{rank}[A_{31}^T A_{33}^T]^T \) which in turn requires
\[
E1. \quad \text{rank}(A_{31}) = \text{rank}[A_{13}^T A_{31}^T A_{33}^T]^T \tag{15}
\]
which corresponds to E1, and the proof is complete.

**Lemma 1.** Condition E1 is satisfied if and only if
\[
\text{rank}(\Phi) - \text{rank}(CM) - \text{rank}(CQ) = \text{rank}(\Omega) - \text{rank}(Q) \tag{17}
\]

**Proof.** From (2)–(3) and (12)–(13) and the structure of \( Q_o \) in Theorem 1, it can be shown that
\[
\text{rank}(\Phi) - \text{rank}(CM) - \text{rank}(CQ) = \text{rank}(A_{31}) \tag{18}
\]
since \( C_2, M_o, \) and \( \tilde{Q}_1 \) are full rank and invertible. Also, it can be shown from (12)–(13) and (2)–(3) that
\[
\text{rank}(\Omega) - \text{rank}(Q) = \text{rank}[A_{13}^T A_{31}^T A_{33}^T]^T \tilde{Q}_1 \tag{19}
\]
Thus, (17) in terms of (18) and (19) is equivalent to
\[
\text{rank}(A_{31}) = \text{rank}[A_{13}^T A_{31}^T A_{33}^T]^T \tilde{Q}_1 \tag{20}
\]
which corresponds to E1, and the proof is complete.

Substituting \( A_1 \) into \( N_1 \), and \( N_1 \) into (17) yields
\[
\text{rank}(\Phi) - \text{rank}(M) - \text{rank}(CQ) = \text{rank}(\Omega) - \text{rank}(Q) \tag{21}
\]
which corresponds to C1. Substituting \( L_{22} = 0 \) into (14) yields \( \tilde{A}_3 = A_{13} + L_{21} A_{31} \). Choosing \( L_{21} \) as follows:
\[
L_{21} = -A_{13} + L_{21}(I - A_{31} A_{31}^T) \tag{22}
\]
where \( L_{211} \) is full rank and \( \tilde{A}_3 = 0 \) as a result, from (14), in order for \( \tilde{A}_3 \) to be stable, both \( \tilde{A}_1 \) and \( \tilde{A}_4 \) have to be stable. From \( L_{21} \) in (21), \( \tilde{A}_4 \) from (14) becomes \( \tilde{A}_4 + L_{211} \tilde{C}_4 \) where \( \tilde{A}_4 = A_{14} - A_{13} A_{31} A_{32}, \tilde{C}_4 = (I - A_{31} A_{31}^T) A_{32} \). So, for \( \tilde{A}_4 \) to be stable, \( \tilde{A}_4, \tilde{C}_4 \) must be detectable. Likewise \( A_{11} \) can be written as
\[
\tilde{A}_1 = A_{11} + [L_{11} \ L_{12}][A_{31}^T A_{33}^T]^T \tag{23}
\]
which implies that \( (A_{11}, [A_{31}^T A_{33}^T]^T) \) has to be detectable for \( \tilde{A}_1 \) to be stable.

**Proposition 2.** \( (\tilde{A}_4, \tilde{C}_4) \) and \( (A_{11}, [A_{31}^T A_{33}^T]^T) \) are detectable if \( (A, [M \ Q], C) \) is minimum phase.

**Proof.** By forming the Rosenbrock matrix [2] from \( (A, [M \ Q], C) \), it can be found that its zeros are the zeros of \( (A_{14}, A_{13}, A_{32}, A_{31}) \). The matrix \( A_{31} \)
can be decomposed to $A_{31} = R_1 \times \text{diag}(0, A_{312}) \times R_2$ where $R_1$ and $R_2$ are orthogonal and $A_{312} \in \mathbb{R}^{r \times r}$ is invertible. Then an appropriate choice of $A_{31}^\dagger$ will be $R_1^T \times \text{diag}(0, A_{312}^{-1}) \times R_2^T$. Since $\text{rank}(A_{31}) = \text{rank}(A_{13}^T A_{31}^T)$ by assumption, $A_{13}$ will become $A_{13} = [0 \quad A_{132}^T R_2^T]_2$ where $A_{132} \in \mathbb{R}^{(n-p-k) \times r}$. Partition $A_{32} = R_1 \begin{bmatrix} A_{321} \\ A_{322} \end{bmatrix}$ and then pre-multiply the Rosenbrock matrix of $(A_{14}, A_{13}, A_{32}, A_{31})$ with $\begin{bmatrix} I \\ 0 \\ Y \end{bmatrix}$ where $Y = [0 \quad -A_{132} A_{312}^{-1}]$. Using the Popov-Hautus-Rosenbrock (PHR) [2] rank test, the zeros of $(A, [M \quad Q], C)$ are the unobservable modes of $(A_{14} - A_{132} A_{312}^{-1} A_{322}, A_{321})$.

Now evaluate $(\tilde{A}_4, \tilde{C}_4)$ using the structures of $A_{31}, A_{13}$ and $A_{32}$ shown earlier. It is easy to verify that $A_4 = A_{14} - A_{132} A_{312}^{-1} A_{322}$, and $\tilde{C}_4 = R_1[A_{321}^T \quad 0]^T$.

Therefore, if $(A, [M \quad Q], C)$ is minimum phase, then $(\tilde{A}_4, \tilde{C}_4)$ is detectable and the first part of the proposition is proved.

If $(A, [M \quad Q], C)$ is minimum phase, then $(A, C)$ is detectable. From the PHR rank test, the detectability of $(A, C)$ implies the detectability of $(A_1, A_3)$, which by using (12)–(13) further implies the detectability of $(A_{11}, [A_{13}^T \quad A_{31}^T \quad A_{33}^T \quad A_{35}^T])^T$.

However, since by assumption condition E1 holds, the detectability of $(A_{11}, [A_{13}^T \quad A_{31}^T \quad A_{33}^T \quad A_{35}^T])^T$ implies that $(A_{11}, [A_{31}^T \quad A_{33}^T \quad A_{35}^T])^T$ is detectable.

Hence, if $(A, [M \quad Q], C)$ is minimum phase, then $(A_1, [A_{31}^T \quad A_{33}^T \quad A_{35}^T])^T$ is detectable and the second part of the proposition is proved. □

Proposition 2 corresponds to C2, guaranteeing a stable sliding motion, and Theorem 1 is proven. □

Remark 2. In Remark 3.6 of [5], Edwards and Tan provided an example where B1 is not satisfied but it is still possible to reconstruct the fault robustly. It can be verified that C1 and C2 are satisfied for the example.

Remark 3. Note that C1 is less conservative than B1, and hence the scheme in this paper can be applied to a wider class of systems. Condition B1 implies that $Q_d = \emptyset$ (empty matrix), which will satisfy C1. However, the converse is not necessarily true.

Remark 4. There have been efforts to generate disturbance decoupled fault detection residuals using linear observers: for example, [7, 8] which uses eigenstructure assignment and the ‘special coordinate basis’. However, from detailed examination (see [7] and equation (13) of [8]), it can be seen that certain elements of the matrix A are required to be zero. From the analysis in this paper, no such condition is required; the only requirement on the matrix A is that E1 is satisfied (which is guaranteed if $A_{31}$ has full column rank). Hence, this paper has shown how the conditions for DDFR using sliding mode observers are less stringent than ones involving linear observers.

IV. AN EXAMPLE

The theory in this paper will be demonstrated using an aircraft system [9]. Due to space constraints, the numerical matrices are not given here, and the reader is referred to [9]. The states are the bank angle, yaw rate, roll rate, sideslip angle, wash-out filter state, rudder deflection, aileron deflection and yaw angle. The inputs are the rudder command and the aileron command. Assume that the bank angle, sideslip angle, rudder deflection, aileron deflection and yaw angle are measurable. Assume both inputs are potentially faulty and therefore $M = B$. Suppose $A$ is imprecisely known. The state equation becomes

$$\dot{x} = (A + \Delta A)x + Bu + Mf$$  \hspace{1cm} (23)

where $\Delta A$ is the discrepancy between the known A and its actual value. By inspection, only rows 2 to 4 in A have uncertainties due to the nature of the state equations. Writing (23) in the framework of (1) yields $\Delta Ax = Q\xi$ where $Q = [0_{1 \times 3} \quad I_3 \quad 0_{4 \times 3}]^T$ and $\xi = \Delta Ax$ where $\Delta A$ are rows 2–4 of $A$. Appropriate choices for $T_1, T_2$ to attain the structure in (2) are

$$T_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad T_2 = \begin{bmatrix} I_3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & I_2 \end{bmatrix}$$

It can be shown that C1 and C2 are satisfied, hence it is possible to obtain a fault reconstruction that is decoupled from $\xi$. Since $\text{rank}(Q) = 3$, $\text{rank}(M) = 2$ and $\text{rank}(CQ) < \text{rank}(Q)$, B1 is not satisfied and it is not possible to reconstruct $\xi$. After attaining the structure in (2), it can be shown that $A_{35} = 0$ and $I - A_{31} A_{31}^\dagger = 0$, which from (16) leads to $W_{12} = 0$. Substituting $W_{12} = 0$ into (10), yields $W_1 = 0$.

Since $I - A_{31} A_{31}^\dagger = 0$, the matrix $\tilde{A}_4$ is independent of $L$ and $\dot{\lambda}(A_{14} - A_{13} A_{31} A_{32}) = -4$. Choosing $L_{11} = -I_2$ means $\dot{\lambda}(A_{11} + L_{11} A_{31} + L_{12} A_{33}) = (-2.5430 - 0.6110)$. Once $L$ is obtained, a suitable choice for $G_l$ and a $P_0$ that satisfies §2.2 of [3] and $G_n$ can then be obtained.
In the following simulation, the gain $\rho$ is set to $\rho=800$ while $\delta=0.001$. Faults have been induced in both actuators, hence $x \neq 0 \Rightarrow \Delta Ax = Q \xi \neq 0 \Rightarrow \xi \neq 0$. Figure 1 shows the faults and their reconstructions. It can be seen that $\hat{f}$ provides accurate estimates of $f$ that are independent of $\xi$ without reconstructing $\xi$.

V. CONCLUSION

This paper has investigated and presented conditions that guarantee DDFR using sliding mode observers, that are easily testable in terms of the original system matrices. The conditions investigated in this paper have shown that reconstructing the disturbances is not necessary for robust fault reconstruction. An aircraft model has been used to validate the results.

REFERENCES