Fault Diagnosis of Linear Singly Perturbed Systems

F. Gong and K. Khorasani

Department of Electrical and Computer Engineering
Concordia University
Montreal, QC H3G 1M8
Email: kash@ece.concordia.ca

Abstract—In this paper, the singular perturbation approach is used to study the diagnosability of linear two-time scale systems. Based on a power series expansion of the slow manifold around $\epsilon = 0$, higher order corrected models for both slow and fast subsystems are obtained. The diagnosability of the decomposed subsystems vary with respect to different corrected-order models. It is shown that if the original singularly perturbed system is fault diagnosable, a composite observer-based residual generator for the original system with actuator faults can be synthesized from the observers of the two separate subsystems. An example of a linear singularly perturbed system is worked out to illustrate the methodology.

I. INTRODUCTION

Because of the increasing demand on the reliability and high efficiency in most of the industrial processes, fault diagnosis has become an important field of research. Several model-based approaches for fault detection and isolation (FDI) have been investigated, including observer-based techniques [1], parity relation methods [2, 3], and parameter estimation methods [4, 5]. The core element of model-based fault diagnosis methods is to generate residual as a fault indicating signal. A residual generator uses available input and output information of the system and generates a residual vector, which should be normally zero or close to zero when no fault is present, but is distinguishably different from zero when a fault occurs.

For a two time-scale system, the slow and fast dynamics in the system respond to external stimuli of different frequencies. Consequently, for the problem of fault detection and isolation, it is necessary to investigate the observability of both slow and fast modes. Approximate models of slow and fast subsystems can be constructed by either equating terms in a power series expansion about $\epsilon$ or through coordinate transformations [6]. Consequently, based on singular perturbation theory [7, 8, 10], analysis and design problems can be solved for both slow and fast modes in two separate time-scales to approximate the original system’s behavior. The fault diagnosis problem of linear singularly perturbed systems was also investigated in [15] which considered the fast subsystem as a modeling error and isolation (FDI) have been investigated, including observer-based techniques [1], parity relation methods [2, 3], and parameter estimation methods [4, 5].

In this paper, the problem of fault diagnosis of the full-order model is addressed by designing decomposed subsystems with utilizing higher order corrected slow and fast subsystems. That is, a full-order observer that generates residuals for the original system is decomposed based on the observers designed separately for the two subsystems. A geometric approach [9] is also applied to guarantee the residual vector has the required isolability property to accomplish fault detection and isolation.

II. SLOW AND FAST SUBSYSTEMS

We consider a linear singularly perturbed system [6]

$$x_1 = A_{11}x_1 + A_{12}x_2 + B_1u, \quad x_1 \in \mathbb{R}^n$$  \hspace{1cm} (1)

$$\epsilon x_2 = A_{21}x_1 + A_{22}x_2 + B_2u, \quad x_2 \in \mathbb{R}^n, u \in \mathbb{R}, \epsilon > 0$$  \hspace{1cm} (2)

$$y = C_1x_1 + C_2x_2.$$  \hspace{1cm} (3)

An $n$-dimensional slow manifold $M_s$ may be defined for the above system according to [16,17]

$$M_s : x_2 = \phi(x_1, u, \epsilon)$$

$= \phi_0(x_1, u) + \epsilon \phi_1(x_1, u) + \epsilon^2 \phi_2(x_1, u) + ...$$  \hspace{1cm} (4)

The above manifold satisfies the following condition which is determined by differentiating equation (4) and substituting it into (1)-(2), that is

$$\epsilon \frac{\partial}{\partial x_1}(A_{11}x_1 + A_{12}x_2 + B_1u) = A_{11}x_1 + A_{12}\phi(x_1, u, \epsilon) + B_1u.$$  \hspace{1cm} (5)

The uncorrected slow manifold is now obtained at $\epsilon = 0$,

$$M_0 : x_2 = \phi_0(x_1, u) = -(A_{22})^{-1}(A_{21}x_1 + B_2u).$$  \hspace{1cm} (6)

With $\phi_0(x_1, u)$ known, we equate the terms of power one $\epsilon$ and obtain

$$\phi_1(x_1, u) = -A_{22}^{-1}A_{21}\phi_0(x_1, u) = -A_{22}^{-1}A_{21}\phi_{n-1}(x_1, u) + B_2u.$$  \hspace{1cm} (7)

The above process can be continued to define higher order corrections, $\epsilon^2 \phi_2(x_1, u)$, etc. At the $n$th stage, we obtain

$$\phi_n = A_{22}^{-1} \phi_{n-1}.$$  \hspace{1cm} (9)

and a manifold $O(\epsilon^n)$ closer to the exact manifold $M_s$ compared to the $(n-1)$th stage. Equipped with the above characterization at $\epsilon = 0$, the uncorrected slow and fast subsystems are obtained as follows

$$x_0^s = A_{11}^s x_0^s + B_1^su, \quad \tau$$

$$x_0^f = C_1^sx + D_1^su,$$  \hspace{1cm} (10)

$$y_0^s = C_1^s x_0^s + D_1^su,$$  \hspace{1cm} (11)

and

$$\frac{dx_0^f}{d\tau} = A_{12}^s x_0^s + B_2^su,$$  \hspace{1cm} (12)

$$y_0^f = C_2^f x_0^f,$$  \hspace{1cm} (13)
Following the same procedure, the first-order corrected slow and fast subsystems are obtained as follows
\[ x_i' = A^0_i x_i' + B^0_i u, \quad y_i' = C^0_i x_i' + D^0_i u, \]
and
\[ \frac{dx_i'}{dt} = A^i_i x_i' + B^i_i u, \]
where
\[ A^0_i = A_{12} - A_{12} A_{22} A_{21}, \quad B^0_i = B_1 - A_{12} A_{22} B_2, \]
\[ C^0_i = C_1 - C_1 A_{12} A_{22}, \quad D^0_i = D_1 - C_1 A_{12} A_{22}, \]
\[ A^i_i = A_{22} + A_{22} A_{12} A_{11}, \quad B^i_i = B_2 + A_{22} A_{12} B_1, \]
\[ C^i_i = C_2 + C_2 A_{22} A_{12}. \]

III. DESIGN OF AND COMPOSITE OBSERVERS

Corresponding to the original system (1)-(3), a full-order observer may be constructed as follows
\[ \hat{x} = (A - GC) \hat{x} + Gy + Bu, \]
\[ \hat{y} = C \hat{x}, \]
where
\[ x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad G = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix}, \quad C = [C_1, C_2]. \]
That is,
\[ \begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{bmatrix} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} y + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u, \]
where
\[ \hat{A}_{11} = A_{11} - G_1 C_1, \quad \hat{A}_{12} = A_{12} - G_1 C_2, \]
\[ \hat{A}_{21} = A_{21} - G_2 C_1, \quad \hat{A}_{22} = A_{22} - G_2 C_2. \]
The state error may be defined by
\[ e = \hat{x} - x, \]
that satisfies
\[ e = (A - GC)e. \]
In [11] a composite observer was designed for the original system (1)-(3) based on the observers for the uncorrected slow and fast models (10)-(13). Let us define \( G^0_i \) and \( G^i_i \) to be the gain matrices of the observers for the uncorrected slow and fast subsystems (10)-(13) respectively. When the full-order observer (18)-(19) is applied to the original system (1)-(3) with
\[ G_1 = A_{12} A_{22}^{-1} G^0_f + G^i_i (I_n - C_2 A_{12} A_{22}^{-1} G^0_i), \]
\[ G_2 = G^i_i, \]
this composite observer is uniformly stable for any \( e \in (0, e^*], e^* > 0 \) if \( A^0_i - G^0_i C^0_i \) and \( A^i_i - G^i_i C^i_i \) are uniformly stable.

The assumption for the existence of the full-order observer designed based on \( G^0_i \) and \( G^i_i \) is that both the uncorrected slow and fast models (10)-(13) are observable. Therefore, when the uncorrected models are not observable, the observability of the higher order corrected models should be investigated in order to construct a full-order observer. In other words, in this work we further the results in [11] by considering the first-order corrected models of the slow and fast subsystems.

A full-order observer for the first-order-corrected slow subsystem (14-15) is given by
\[ \dot{\hat{x}}_1 = (A^i_i - G^i_i C^i_i) \hat{x}_1 + G^i_i y_1 + (Bu_{i}), \]
\[ \dot{\hat{y}}_1 = C^i_i \hat{x}_1 + D^i_i u_1 - C^i_i A_{12} B^i_i u, \]
where \( G^i_i \) is the gain matrix. The state error is defined by
\[ e_1 = \hat{x}_1 - x_1, \]
which satisfies
\[ e_1 = (A^0_i - G^0_i C^0_i) e_1, \]
where
\[ A^0_i - G^0_i C^0_i = (A^0_i - G^0_i C^0_i) - \epsilon (A_{12} A_{22} A_{21} A_{11}) \]
\[ + \epsilon C_2 A_{12} A_{22} A_{21} A_{11} \]
\[ A^i_i - G^i_i C^i_i = (A^i_i - G^i_i C^i_i) + \epsilon (A_{12} A_{22} A_{21} A_{11}) \]
\[ + \epsilon C_2 A_{12} A_{22} A_{21} A_{11}. \]

In order to design a composite observer based on the observers for the slow and fast subsystems, the error dynamics (22) is first decomposed into slow and fast subsystems by utilizing the coordinate transformation [6],
\[ \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = \begin{bmatrix} I_n & -\epsilon \hat{H} \\ -\hat{L} & I_n \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}, \]
That is,
\[ \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = \begin{bmatrix} I_n & -\epsilon \hat{H} \\ -\hat{L} & I_n \end{bmatrix} \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} I_n - \epsilon \hat{H} \end{bmatrix} E_1, \]
where \( \hat{L}(e) \) and \( \hat{H}(e) \) satisfy the algebraic equations
\[ R(\hat{L}, e) = A_{21} - A_{22} \hat{L} + \epsilon \hat{L} A_{11} - \epsilon \hat{L} A_{12} \hat{L} = 0, \]
\[ S(\hat{H}, e) = \epsilon (A_{11} - A_{12} \hat{L}) \hat{L} - \hat{H} (A_{21} + \epsilon \hat{L} A_{12}) + A_{22} = 0. \]
An approximation to \( \hat{L} \) and \( \hat{H} \) up to \( o(\varepsilon^2) \) may be obtained as:

\[
L = \hat{L}_0 + \varepsilon \hat{L}_1 = \hat{A}_{11} \hat{A}_{21} + \varepsilon \hat{A}_{21} \hat{L}_0 (\hat{A}_{11} - \hat{A}_{12} \hat{A}_{22} \hat{A}_{11}) ,
\]

\[
H = \hat{H}_0 + \varepsilon \hat{H}_1 = \hat{A}_{12} \hat{A}_{22} + \varepsilon (\hat{A}_{11} - \hat{A}_{12} \hat{A}_{22} \hat{A}_{11}) \hat{H}_0 - \hat{H}_0 \hat{A}_{12} \hat{A}_{22} \hat{A}_{11} ,
\]

where

\[
\hat{L}_0 = \hat{A}_{12} \hat{A}_{22}, \quad \hat{H}_0 = \hat{A}_{12} \hat{A}_{22}
\]

Substituting (46)-(47) into (35) we obtain

\[
E_i = \begin{bmatrix} I_s - e\hat{H}_\varepsilon \hat{L}_0 (\hat{A}_{11} - \hat{A}_{12} \hat{A}_{22} \hat{A}_{11}) & 0 \\
E_i & \hat{A}_{22} + e\hat{A}_{22} \hat{A}_{22} \hat{A}_{12} \end{bmatrix} E_i
\]

where

\[
(\text{I}_s - e\hat{H}_\varepsilon \hat{L}_0 (\hat{A}_{11} - \hat{A}_{12} \hat{A}_{22} \hat{A}_{11}))
\]

Now, the first-order corrected model of (36) is given by

\[
E_i = \begin{bmatrix} \text{I}_s - e\hat{H}_\varepsilon \hat{L}_0 & 0 \\
E_i & \hat{A}_{22} + e\hat{A}_{22} \hat{A}_{22} \hat{A}_{12} \end{bmatrix} E_i
\]

Therefore, from the error dynamics (22) and using (28), (33) and (35), we can conclude that

\[
e_i = e_i + \varepsilon [M + \hat{H}_\varepsilon E_i] + o(\varepsilon^2)
\]

\[
e_i = \begin{bmatrix} I_s - e\hat{H}_\varepsilon \hat{L}_0 (\hat{A}_{11} - \hat{A}_{12} \hat{A}_{22} \hat{A}_{11}) & 0 \\
E_i & \hat{A}_{22} + e\hat{A}_{22} \hat{A}_{22} \hat{A}_{12} \end{bmatrix} E_i + o(\varepsilon^2)
\]

Therefore, for the observations (18)-(19), (25)-(26) and (30)-(31) applied to systems (1)-(3) (14)-(15) and (16)-(17), respectively, where

\[
G_i = A_{21} A_{22} G_j^0 + G_j^0 (I - C_{21} A_{22} G_j^0),
\]

will guarantee that the error of the full-order system (1)-(3) satisfies the equations

\[
e_i = e_i + \varepsilon [M + \hat{H}_\varepsilon e_i (\tau)] + o(\varepsilon^2)
\]

\[
e_i = \begin{bmatrix} I_s - e\hat{H}_\varepsilon \hat{L}_0 & 0 \\
E_i & \hat{A}_{22} + e\hat{A}_{22} \hat{A}_{22} \hat{A}_{12} \end{bmatrix} E_i + o(\varepsilon^2)
\]

IV. A GEOMETRIC METHOD FOR FAULT DETECTION

Consider the linear time-invariant model:

\[
x = Ax + Bu + \sum_{i=1}^{n} L_i f_i,
\]

\[
y = Cx
\]

where \( L_i = B_i \), \( B_i \) is the \( i \)th column of \( B \), and \( f_i(t) \) is the fault in the \( i \)th actuator. When the actuator has failed, then \( f_i(t) = -u_i(t) \) where \( u_i(t) \) is the \( i \)th component of \( u(t) \); and when there is a bias fault on the same actuator, their \( f_i(t) \) is some nonzero constant.

The geometric approach presented in [9] is applied here based on the so-called Beard-Jones approach [12, 13] to diagnose the actuator faults in the system (58). A Beard-Jones detection filter is a full-order observer that can be described as:

\[
\dot{x} = (A - KC)x + Ky + Bu
\]

\[
r = R(\bar{y} - y)
\]

with \( K \) and \( R \) in the error dynamics:

\[
e = (A - KC)e - \sum_{i=1}^{n} L_i f_i ,
\]

\[
r = R C e
\]

such that \( R \) is a constant matrix and \( R C e \) is the projection of the residual onto the output space. The observer-based residual generator (60) will result in a directional residual vector \( r(t) \) which lies in a fixed and fault-specified direction in the output space. Moreover, to identify the fault, only the magnitude of the residual and not its functional behavior is used. Therefore, the failure can be identified by finding the projection of \( r(t) \) onto the output space and comparing the magnitude of this projection to a threshold.

V. NUMERICAL EXAMPLE

In order to illustrate and demonstrate our proposed methodology, we design a detection filter for a system with slow and fast modes in the form (1)-(3) [14] with \( \varepsilon = 0.01 \), having the following specifications:

\[
A_{11} = \begin{bmatrix} 0 & 0 \\
0 & 0.345 \end{bmatrix}, A_{12} = \begin{bmatrix} 0 & 0 \\
0 & 0 \end{bmatrix}, A_{21} = \begin{bmatrix} 0 & -0.524 \\
0 & 0 \end{bmatrix}, A_{22} = \begin{bmatrix} 0 & 0.465 \\
0 & -1 \end{bmatrix}
\]

\[
B_i = \begin{bmatrix} 0 \\
1 \end{bmatrix}, C_1 = \begin{bmatrix} 1 \\
0 \end{bmatrix}, C_2 = \begin{bmatrix} 0 & 0 \\
0 & 1 \end{bmatrix}
\]

For the slow subsystem, the uncorrected slow model is characterized by

\[
A_0 = \begin{bmatrix} 0 & 0.4 \\
0 & -0.3888 \end{bmatrix}, B_0 = \begin{bmatrix} 0 \\
1 \end{bmatrix}, C_0 = \begin{bmatrix} 1 \\
0 \end{bmatrix}, D_0 = \begin{bmatrix} 0 \\
0 \end{bmatrix}
\]

We choose \( G_0 \) as
\( G^o_j = \begin{bmatrix} 1 & -0.355 \\ 0 & -1.4298 \end{bmatrix} \) \hspace{1cm} (63)

to place the poles of \( A^i_j - G^o_j C^o_j \) at \(-1, -2\). Then we design a fault detection filter for the system (62) based on the observer gain (63) as \( K_j = G^o_j \), and choose \( R_j = 1.49 \). Here, the residual is generated by the second output of the system model and its observer.

Fig. 1 shows a simulation result corresponding to the following scenario: the actuator \( u_s \) is supposed to provide constant step input equal to 1, \( u_s = 1 \). A fault \( \Delta f \) on the actuator \( u_s \) occurs at time \( t = 20 \) and ends at \( t = 30 \) [Fig. 1(a)]. Fig. 1(b) depicts the output of the residual generator of the uncorrected model, which clearly shows the occurrence of the fault \( \Delta f \) and identification of its actual value.

For the fast subsystem, the uncorrected model is given by

\[
\begin{bmatrix} A_j^0 \\ B_j^0 \\ C_j^0 \end{bmatrix} = \begin{bmatrix} -0.465 & 0.262 \\ 0 & -1 \\ 1 & 0 \end{bmatrix} \]

which places the poles of \( A^K - G^K C^K \) at \(-1, -2, -211, -282\). For the purpose of fault diagnosis, we chose \( R = 22.6 \). For comparison, we also design an observer directly for the full-order system (61) as

\[
\begin{bmatrix} A_j^0 \\ B_j^0 \\ C_j^0 \end{bmatrix} = \begin{bmatrix} -0.4611 & 0.262 \\ 0 & -1 \\ 1.0084 & 0.0022 \end{bmatrix} \]

We choose

\[
G_j^f = \begin{bmatrix} 3.493 \\ 0.0113 \\ 0 \end{bmatrix} \hspace{1cm} (66)
\]
to place the poles of \( A^f_j - G^f_j C^j \) at \(-2, -3\). Then a detection filter is designed based on the observer gain (66) as \( K_j = G_j^f \) and \( R = 525 \). Fig. 2(a) shows an actuator fault \( \Delta f \) in the fast subsystem whose actuator \( u_f \) is supposed to provide a constant step input equal to \( u_f = 1 \). Fig. 2(b) is the output of the residual generator.

Based on the observer gains (63) and (66) designed for the slow and fast subsystems, we construct a composite observer for the full-order system (61) by following the equations (54)-(55) that are obtained as

\[
G = \begin{bmatrix} 1 & -4.5422 \\ 0 & -22.3662 \\ 0 & 349.3026 \\ 0 & 760.4272 \end{bmatrix} \hspace{1cm} (67)
\]

which places the poles of \( A^o - G C \) at \(-1, -1.99, -211.51, -2822.9\). For comparison, we also design an observer directly for the full-order system (61) as

\[
K = \begin{bmatrix} 3.1098 & -5.4282 \\ -2.3960 & -11.0242 \\ -52.4009 & 346.3902 \\ -259.8070 & 757.5931 \end{bmatrix} \hspace{1cm} (68)
\]

which places the poles of \( A^K - K C^K \) at \(-1, -2, -211, -282\), and for the purpose of fault diagnosis we choose \( R = 10.4 \).

Fig. 3(a) simulates an actuator fault that has occurred in the full-order system (61) with \( \varepsilon = 0.01 \). Fig. 3(b) shows the output of the residual generator designed based on the composite observer, whereas Fig. 3(c) depicts the residual generated based on the observer directly designed for the full-order system. By comparing the error between the fault
and the residual in Fig. 3(b) and between the fault and the residual in Fig. 3(c), we can observe from Fig. 3(d) that the detection filter based on the composite observer has better capability in fault diagnosis.

To investigate the impact of $\varepsilon$ on fault diagnosis of the full-order system (61), we set $\varepsilon = 0.01$: (a) the actuator fault occurred in the original system; (b) the residual generated based on the composite observer; (c) the residual generated based on the observer directly designed for the full-order system; (d) solid line is the error between the fault and the residual in (b), dashed line is the error between the fault and the residual in (c).

To investigate the impact of $\varepsilon$ on fault diagnosis of the full-order system (61), we set $\varepsilon = 0.5$ and repeat our proposed design procedures. When $\varepsilon = 0.5$, the uncorrected slow model is the same as (62), and we choose the same observer $G_1^s$ as in (63). For the fast subsystem, the first-order corrected model is now given by

$$A_j^f = \begin{bmatrix} -0.2706 & 0.262 \\ 0 & -1 \end{bmatrix}, B_j^f = \begin{bmatrix} 0.5634 \\ 0 \end{bmatrix}. \quad (69)$$

$$C_j^f = \begin{bmatrix} 0 & 0 \\ 1.4180 & 0.1092 \end{bmatrix}. \quad (70)$$

We choose $G_j^f = \begin{bmatrix} 0 & 2.1003 \\ 0 & 6.8578 \end{bmatrix}$ to place the poles of $A_j^f - G_j^f C_j^f$ at $\{-2, -3\}$. In order to identify the fault, we chose $R'_f = 1.49$ and $R'_f = -7.5$. Fig. 4a shows the simulation results for an actuator fault that has occurred in the slow and fast subsystem with $\varepsilon = 0.5$, Fig. (4b) and Fig. (4c) depict the output of the residual generators based on the observers $G_1^s$ and $G_j^f$, respectively.

Based on the observer gains $G_1^s$ and $G_j^f$, we designed a composite observer for the original system at $\varepsilon = 0.5$ as

$$G = \begin{bmatrix} 1 & -3.3298 \\ 0 & -16.3040 \\ 0 & 4.2003 \\ 0 & 13.7156 \end{bmatrix}. \quad (71)$$
which places the poles of at \( A - G'C \{-1, -1.27, -3 \pm 4.34i \} \). For fault diagnosis, we choose \( K' = G' \) and \( R' = 16.7 \). We also designed an observer directly for the full-order system (61) as

\[
K' = \begin{bmatrix}
3.6046 & -0.1072 \\
4.5348 & -0.4592 \\
-2.6760 & 1.7654 \\
-2.0836 & -1.1175 \\
\end{bmatrix}
\] (72)

which places the poles of \( A - K'C \{-1,-1.2,-3,-3 \} \), and for the purpose of fault diagnosis we choose \( R' = 2.35 \).

![Fig. 5 Detection and identification of the actuator fault in the original system with \( \varepsilon = 0.5 \): (a) the actuator fault occurred in the full-order system; (b) the residual generated based on the composite observer; (c) the residual generated based on the observer directly designed for the full-order system; (d) solid line is the error between the fault and the residual in (b), dashed line is the error between the fault and the residual in (c).](image)

Fig. 5 shows the simulations for an actuator fault that has occurred in the full-order system (61) with \( \varepsilon = 0.5 \). Fig. 5(b) is the output of the residual generator designed based on the composite observer, while Fig. 5(c) is the residual generated based on the observer directly designed for the full-order system. By comparing the errors between the fault and the residual in Fig. 5(b) and between the fault and the residual in Fig. 5(c), we can observe from Fig. 5(d) that the detection filter based on the composite observer has better capability in fault diagnosis.

REFERENCES