Fault Estimation Methods for Semi-Active Suspension Systems

Diana Hernandez-Alcantara, Ruben Morales-Menendez, Luis Amezquita-Brooks, Olivier Sename, Luc Dugard

To cite this version:

HAL Id: hal-01233700
https://hal.science/hal-01233700
Submitted on 25 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Fault Estimation Methods for Semi-Active Suspension Systems

Diana Hernandez-Alcantara*, Ruben Morales-Menendez*, Luis Amezquita-Brooks‡, Olivier Sename ‡ and Luc Dugard‡

*Tecnologico de Monterrey, Escuela de Ingenieria y Ciencias. Av. Garza Sada 2501. Monterrey N.L., Mexico
Email: {A00469139, rmm}@itesm.mx

‡Universidad Autonoma de Nuevo Leon. Av. Universidad s/n. San Nicolas de los Garza, N.L., Mexico
Email: amezquita-brooks@ieee.org

‡ GIPSA-Lab, Grenoble INP. Domaine Universitaire BP46, Saint Martin d’Hères, France.
Email: {olivier.sename, luc.dugard}@gipsa-lab.grenoble-inp.fr

Abstract—Semi-Active (SA) suspension systems aim to improve the stability and comfort of vehicles. Although they offer better performance than passive suspensions, the actuators such as magneto-rheological dampers are more susceptible to failure. Oil leakage is the most common fault, and its effect is a reduction of the damping force. The estimation of suspension faults can be used with a Fault Tolerant Control system to prevent handling and comfort deterioration. However, fault estimation schemes introduce additional challenges due to the damper non-linear dynamics and the strong influence of the disturbances (i.e. the road profile). One of the first obstacles for appropriate damper fault detection is the modeling of the fault, which has been shown to be of multiplicative nature. However, many of the most widespread fault detection schemes consider additive faults due to mathematical convenience. Two complementary model-based fault estimation schemes for semi-active dampers are proposed: an observer-based approach, which is intended to estimate additive faults; and a parameter identification approach, which is intended to estimate multiplicative faults. The performance of these schemes is validated and compared through simulations using a pickup truck model. Early results shows that a parameter identification approach is more accurate in fault estimation, whereas an observer-based approach is less sensible to parametric uncertainty.

I. INTRODUCTION

The goal of a suspension system is to isolate the vehicle body from the natural road disturbances, while providing an adequate road-holding. In a passive suspension, there exists a trade-off between these objectives. Semi-Active (SA) suspensions systems are an option to cope with this. These systems consist of a SA shock absorber, which is able to online change its damping characteristics, improving simultaneously both road disturbance rejection and road-holding. These actuators have also the benefit of being lower power consuming and less expensive and bulky in comparison with active dampers. Among the SA technologies the Magneto and Electro-Rheological are the most mature and widely used because of their wide force range and fast time response.

While the SA suspensions have advantages over the passive ones, the use of more complex actuators increases the necessity of fault detection. In particular, it has been noted that SA dampers are more susceptible to failures than their traditional counterpart. This is specially relevant since the effect of damper faults could vary from the increase of chassis vibrations to the loss of stability. Weak shock absorbers may deteriorate the vehicle handling and increase the braking distances up to 20 %, [1].

There are few papers dealing with Fault Detection and Isolation (FDI) for SA dampers, and most of them are based on methods that inherently handle additive faults. These methods have the advantage of low computational requirements and do not need permanent excitation. In [2], [3] the parity space method is used to detect the damper fault. However, the SA damper nonlinearities are not considered nor model uncertainties. Also, in [4] a robust fault estimator based on the parity space method is proposed. The residual is designed to minimize the sensitivity to uncertainties, while maximizing the sensitivity to damper faults. In [5] a bank of unknown input observers is used to detect and isolate actuator failures in an active suspension system. In [6] a bank of Lyapunov observers is proposed to detect faults in a building SA suspension system.

Some FDI approaches based on parametric identification have been also proposed. These approaches have the advantage of providing deeper insight into the process; however, most online estimation methods need excitation and other operating requirements which may be restrictive when the road profile is unknown. In [7], [8], the parameters of a faulty Quarter of Vehicle (QoV) suspension model are estimated using a recursive algorithm; these parameters are then used to generate residuals and fault signatures. In [9], [1] the FDI system is extended by applying Artificial Intelligence methods to isolate the fault.

Two different fault detection approaches are applied to SA suspension systems to estimate actuator faults are proposed: the observer-based approach and the parameter estimation approach. The outline of the paper is as follows: Section II introduces the QoV model and defines the fault estimation problem. Section III details the design of both fault estimation approaches. Section IV describes the used tests to validate the fault estimation approaches. Conclusions are presented in section V.
The dynamic behavior of a $QoV$ model with a SA suspension is described by:

$$
\begin{align*}
    m_i z_i &= -F_D - k_i (z_i - z_{as}) \\
    m_{as} z_{as} &= k_i (z_i - z_{as}) + F_D - k_i (z_{as} - z_r)
\end{align*}
$$

where $z_i$, $z_{as}$ and $z_r$ are the sprung mass, unsprung mass and road profile vertical positions, respectively; $m_i$ is the sprung mass which represents the chassis, $m_{as}$ is the unsprung mass which represents the wheel, tire, etc.; $k_i$ and $k_c$ are the suspension and the tire stiffness, and $F_D$ represents the damping force. The SA damper behavior may be represented according to this model, [10]:

$$
F_D = k_p z_{def} + c_p z_{def} + f_c u \tanh(a_c z_{def} + a_d z_{def})
$$

where $z_{def} = z_i - z_{as}$ is the suspension deflection, and $u$ is the damper control input (i.e. electric current for a MR damper), $f_c$ is the damping force due to the input, $c_p$ is a linear viscous damping coefficient, $k_p$ is a linear stiffness coefficient, $a_c$ and $a_d$ are coefficients used to represent the hysteretic behavior.

### Table I. Typical pickup truck model parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Units</th>
<th>Parameter</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_i$</td>
<td>315</td>
<td>Kg</td>
<td>$m_{as}$</td>
<td>37.5</td>
<td>Kg</td>
</tr>
<tr>
<td>$k_i$</td>
<td>29,300</td>
<td>N/m</td>
<td>$k_c$</td>
<td>230,000</td>
<td>N/m</td>
</tr>
<tr>
<td>$c_p$</td>
<td>1,500</td>
<td>N/m</td>
<td>$a_c$</td>
<td>7.89</td>
<td>m/s</td>
</tr>
<tr>
<td>$a_d$</td>
<td>-13.8</td>
<td>1/m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$f_c$</td>
<td>441</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The considered model for multiplicative fault is:

$$
F_D = \alpha \cdot [f_c u \tanh(a_c z_{def} + a_d z_{def}) + c_p \cdot z_{def} + k_p \cdot z_{def}]
$$

where $\alpha$ represents the effectiveness of the damper, i.e. $\alpha = 1$ is a healthy damper, $\alpha = 0$ represents a 100% damper failure. The aim is to estimate $\alpha$ using the masses accelerations and the suspension deflection, while the road profile $z_r$ is an unknown perturbation input, because it is expensive to measure.

### III. Fault Estimation Methods

#### A. Observer-based Fault Estimation

A FDI system based on an Unknown Input Observer (UIO) and the $QoV$ model (1) is proposed. The idea is to synthesize a residual, which is sensitive to the damper fault and insensitive to disturbances (i.e. road profile). The damper fault is first modelled as an additive fault:

$$
F_D = F_{\Delta} - F_{\delta}
$$

where $F_{\Delta}$ is the nominal force (i.e. the force of a healthy damper) and $F_{\delta}$ is the loss of force due to the fault. By using (1), (2) and (4), a state space system is:

$\dot{x}(t) = Ax(t) + Bu(t) + B_2 z_r(t) + B_4 F_\delta(t)$

$y(t) = Cx(t) + Du(t) + D_2 z_r(t) + B_4 F_\delta(t)$

where $x = [z_s \ z_s \ z_{as} \ z_{def}]^T$, $y = [z_{as} \ z_s \ z_{def}]^T$

$A = \begin{bmatrix} -k & 1 & 0 & 0 \\ 0 & -m_i & -c_p & 0 \\ -k & 0 & -k_c & k_c \\ 0 & m_i & -m_{as} & -c_p \\ 0 & 0 & 0 & -1 \end{bmatrix}$

$B_a = \begin{bmatrix} f_c \\ 0 \\ 0 \\ f_c \end{bmatrix}$

$B_f = \begin{bmatrix} \frac{k_p}{m_{as}} \\ 0 \\ 0 \end{bmatrix}$

$D_a = \begin{bmatrix} \frac{f_c}{m_{as}} \end{bmatrix}$

$D_f = \begin{bmatrix} \frac{1}{m_{as}} \end{bmatrix}$

$D_z = \begin{bmatrix} \frac{1}{m_{as}} \\ 0 \\ 0 \end{bmatrix}$

$D_s = \begin{bmatrix} \frac{k_p}{m_{as}} \end{bmatrix}$

$D_{ss} = \begin{bmatrix} \frac{k_p}{m_{as}} \end{bmatrix}$

$D_{ss} = \begin{bmatrix} \frac{k_p}{m_{as}} \end{bmatrix}$

$D_{ss} = \begin{bmatrix} \frac{k_p}{m_{as}} \end{bmatrix}$

$D_{ss} = \begin{bmatrix} \frac{k_p}{m_{as}} \end{bmatrix}$

with

$$
\rho(t) = \tanh(a_c z_{def} + a_d z_{def})
$$

An important result of applying the (UIO) theory to this domain is that, without road information, the system is non-observable and non-detectable [11]. This indicates that there are deeply rooted problems which will affect the state estimations.

To deal with this problem, an approximated UIO observer can be obtained by decoupling the system (5) from the road surface using the procedure reported in [12]. The decoupled system can be written as:

$$
\dot{x}(t) = \bar{A}x(t) + \bar{B}_a pu(t) + \bar{B}_4 F_\delta(t) + G \bar{y}(t)
$$

$y_2(t) = \tilde{C}x(t) + \tilde{D}_a pu(t) + \tilde{B}_4 F_\delta(t)

where $\bar{A}$ is a $(4 \times 4)$ matrix whose first two rows are:

$$
\bar{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -k & -m_i & -c_p & 0 \\ -k & 0 & -k_c & k_c \\ 0 & m_i & -m_{as} & -c_p \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad \bar{B}_a = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
$$

$\bar{B}_f = \begin{bmatrix} \frac{1}{m_{as}} \end{bmatrix}, \quad \bar{B}_\delta = \begin{bmatrix} \frac{k_p}{m_{as}} \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\tilde{C} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ -k & -m_i & -c_p & 0 \\ -k & 0 & -k_c & k_c \\ 0 & m_i & -m_{as} & -c_p \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad \tilde{D}_a = \begin{bmatrix} D_a \\ 0 \\ 0 \end{bmatrix}

D_f = \begin{bmatrix} D_f \end{bmatrix}, \quad y_1 = z_{as}, y_2 = z_{def}, y_3 = z_s, \tilde{y} = \begin{bmatrix} y_2 \\ y_3 \end{bmatrix}$

For the state estimation the decoupled system (7) with the output $y_2$ will be used since this output does not depend on the fault. On the other hand, for the residual generation, $y_3$, which is affected by the fault, will be used.

If $\bar{A}, \bar{C}_1$ is observable or detectable an observer can be designed for the decoupled system to estimate the states. However, the observability matrix of system (7), has rank 2; thus, two of the observer poles cannot be located arbitrarily. Moreover, for the system (7) the non-observable subset is related to two poles in $s = 0$; therefore, the system is non-observable and non-detectable. A possible solution consists on generating a stable system which approximates the frequency response of the UIO observer, [13].

An observer is a dynamical system whose inputs are the process input and output vectors. For the decoupled system (7), the observer poles in $s = 0$ are related to the input contribution to the state estimation. Since these poles are not
stable regardless the observer gain, the best estimation effort is the open-loop simulation of these states. In general the state estimation based only in the simulation of the process (which is equivalent to an observer with observer gain equal to zero) is not advised. However, the input related with these integrators is the unsprung-mass acceleration, which is indeed a measured output, but it was considered as an input to achieve the decoupling of the system from the road profile. Thus, the non-observable states are estimated via the output, like it is done by using high-gain observers.

The observer gain matrix \( L \) is designed to place the two observable poles at \((-20.38 \pm 1.09j)\). The resulting observer combines the closed-loop estimation dynamics and the integration of acceleration measurements. For the states which are estimated by numerical integration, it is well known that low frequency noise in the measurements induces a divergence called drift [14]. Without additional measurements it is not possible to eliminate the drift by modifying the observer gain, [11], [15]. Since the suspension deflection is limited and the masses velocities converge to zero, it is possible to use high-pass filters to eliminate the drift. Thus, the observer is complemented by adding a pair of high-pass filters \( g_r(s) = s/(s+\omega_f) \) for each of the integrators. In this way, only the non-detectable states are filtered, while the observable ones are estimated by a closed-loop observer. The cut-frequency \( \omega_f \) of the filter is set at 0.14 Hz.

By using the estimated states with the U/O observer and the measurement of the sprung mass acceleration, this residual can be computed to detect a loss of damping force:

\[
r(t) = y_3 - [\hat{C}_2 \dot{x}(t) + \hat{D}_a u(t)] = D_f \dot{z}_d(t) \quad (8)
\]

then the magnitude of the fault can be estimated as:

\[
\dot{\hat{z}}_d = [\hat{D}_f]^{\top} r(t) \quad (9)
\]

where \( \dagger \) stands for the Moore Penrose pseudo inverse.

B. Parameter Estimation Approach

By considering the multiplicative model of the fault (3), the system can be represented in state-space form as:

\[
\begin{align*}
\dot{x}(t) &= A_n x(t) + g(x,u) + B_s z_s(t) \\
y(t) &= C_n(\alpha)x(t) + h(x,u) + D_s z_s(t)
\end{align*} \quad (10)
\]

with

\[
A_n = \begin{bmatrix}
0 & 1 & 0 & 0 \\
-\frac{K_s}{m_s} & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
\frac{K_s}{m_s} & 0 & 0 & 0
\end{bmatrix},
B_s = \begin{bmatrix}
0 \\
0 \\
0 \\
-\frac{K_s}{m_s}
\end{bmatrix},
C_n(\alpha) = \begin{bmatrix}
0 \\
0 \\
0 \\
\frac{K_s}{m_s}
\end{bmatrix},
D_s = \begin{bmatrix}
0 \\
0 \\
0 \\
-\frac{K_s}{m_s}
\end{bmatrix}
\]

Although the fault affects both the sprung and unsprung mass dynamics, for this estimator only the sprung mass dynamic will be considered since it is independent of the unmeasured road profile. Two measurements are used: \( \dot{z}_s \) and \( z_{def} \). From (10), the discrete-time representation of the sprung mass dynamic can be written in the regression form as:

\[
\zeta = \Psi^T \theta \quad (11)
\]

where \( \zeta = -k_s(z_{def}) - m_s(\dot{z}_s), \quad \Psi^T = -f_p u + c_p(z_{def}) + k_p(z_{def}) \) and \( \theta = \alpha \). The damper fault can be online estimated by using the Recursive Least Squares (RLS) with forgetting factor:

\[
\begin{align*}
e(k) &= \zeta - \Psi^T(\hat{\theta}(k-1)) \\
\gamma(k) &= \frac{P(k-1)\psi(k)}{\lambda + \psi^T(k)P(k-1)\psi(k)} \\
P(k) &= \frac{1}{\gamma(k)}(I - \gamma(k)\Psi^T(k))P(k-1) \\
\hat{\theta}(k) &= \hat{\theta}(k-1) + \gamma(k)e(k)
\end{align*} \quad (12)
\]

where \( \hat{\theta} \) is the estimation of \( \theta \) and \( \lambda \) is the forgetting factor. The measurement of \( \dot{z}_{def} \) is not available, therefore an estimation of this signal is required in order to implement the parameter estimation. \( \dot{z}_{def} \) may be estimated either by the numerical derivative approximation of \( \dot{z}_{def} \) or by an state-observer. In this case, an approximation based on the bandwidth-limited derivative is used.

IV. RESULTS

Four tests were used to evaluate the performance of the FDI proposals. These tests were an ISO Type D as the road profile at a vehicle speed of 30km/h, which represents a normal driving condition. Table II summarizes the tests characteristics.

<table>
<thead>
<tr>
<th>No.</th>
<th>Description</th>
<th>Specific objective to analyze</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>Fault ( \alpha \in [0.5, 1] ), ( u = 2\alpha )</td>
<td>Estimation of abrupt fault at constant damper input</td>
</tr>
<tr>
<td>#2</td>
<td>Fault ( \alpha \in [0.1, 1] ), ( u = 2\alpha )</td>
<td>Estimation of gradual fault at constant damper input</td>
</tr>
<tr>
<td>#3</td>
<td>Fault ( \alpha \in [0.1, 1] ), ( u = [0.5, 3]A )</td>
<td>Estimation of gradual fault at varying damper input</td>
</tr>
<tr>
<td>#4</td>
<td>Fault ( \alpha \in [0.5, 1] ), and steps in ( z ).</td>
<td>Sensitivity to road disturbances.</td>
</tr>
<tr>
<td>#5</td>
<td>Fault ( \alpha \in [0.5, 1] ), and ( 1.2m_s )</td>
<td>Robustness to uncertainty in ( m_s )</td>
</tr>
</tbody>
</table>

Although both strategies are used to estimate faults of different nature (i.e. additive or multiplicative), a comparison can be made by calculating the equivalence between them. For the multiplicative fault the equivalent additive fault is:

\[
F_\delta = F_D \left( \frac{1}{\alpha} - 1 \right) \quad (13)
\]

Figure 1 shows the estimation of the loss of force \( F_\delta \) in the Test #1. An abrupt damper fault is modelled as a step change in \( \alpha \); however, the equivalent additive fault also depends on \( F_D \), therefore even if \( \alpha \) is constant, \( F_\delta \) is not. The result, is that for a step change in \( \alpha \), the Least Squares Estimator (LEE) error converges to zero while the Observer-Based Estimator (OBE) error does not, see Fig. 1.

Figure 2 shows the multiplicative equivalent fault of the Test #1, this confirms the last discussion. The LEE is better posed for constant multiplicative faults than the OBE. Since real damper faults are more similar to multiplicative faults and due to lack of space, for the remaining tests only the responses of the equivalent multiplicative fault \( \alpha \) will be presented.

Figure 3 shows that both approaches are able to estimate gradual variations in the multiplicative fault level, Test #2. In particular, when the multiplicative fault \( \alpha \) has a constant slope, neither of the estimator errors converge to zero. Nonetheless, the LEE has a lower overall error. In Fig. 4 the estimation of \( \alpha \) is compared in the same fault conditions as in Test #2, but, in this case a varying damper input is considered, Fig. 5. It can be observed that both approaches can discriminate
between the damper input and the fault signal. This is an important observation because this characteristic, essential for SA suspension systems, has not been widely explored with these algorithms. This figure also shows that the LEE has lower sensitivity to damper input variation than the OBE.

Test #4 is used to assess the sensitivity of the fault estimators to sudden road profile changes, which can represent road abnormalities such as potholes. The road considered in this test is shown in Fig. 6. The road decoupling of the OBE is not perfect since a perfect UIO was not possible to design due to observability issues. This is more evident when the road presents sudden changes, inducing a greater level of error. On the other hand, the LEE is not visibly affected by these road characteristics.

The last test is used to assess the effect of sprung mass uncertainty on the estimated fault. A sprung mass 20% heavier was considered. The results show, Fig. 8), that the mass uncertainty induces steady state error in both schemes. In particular, the LEE has a greater error. In addition, it was observed that the LEE tends to always underestimate the fault level while the OBE can present underestimation or overestimation depending
on the fault level.

Finally, a further quantitative comparison of the estimation approaches was made using the Root Mean Square Error (RMSE) index. Table IV summarizes the results. It can be observed that the LEE presents an overall lower level of error; for instance, in Test #2 the LEE has 493% lower RMSE in the estimation of the loss of force $F_\delta$. If the fault is represented in the multiplicative form, the LEE also yields a lower RMSE, but the differences are smaller. For example, in test #2 the LEE is also lower but 57%. Only when sprung-mass uncertainty is considered the OBE has a lower RMSE than the LEE, in this case 55% lower for $F_\delta$.

V. CONCLUSIONS

A comparison of two fault detection proposals for the estimation of 5A damper faults is presented. One method is based on classical least squares parametric identification to estimate a multiplicative fault. The second method is based on an unknown input observer, which is used to calculate a residual and to estimate the loss of force due to an additive fault.

In the literature, it has been reported that damper faults are normally of multiplicative nature; nonetheless additive fault estimation methods are more widely reported. Since it is possible to derive the multiplicative equivalent from the additive estimation, both methods were compared in terms of multiplicative damper faults.

The comparison was made through simulations devised to investigate different scenarios, including mass uncertainty, damper input sensitivity and road profile sensitivity. The results show that the least squares estimator achieves better qualitative and quantitative performance levels in all cases with the exception of mass uncertainty. This also suggests that, for this system, identification algorithms based on regression models are highly sensitive to plant parameters changes. It can be concluded that the two estimators present complementary characteristics which could be exploited depending on the operating conditions in order to improve the overall system performance.

The results also show that for this application a simple fault estimation algorithm, such as least squares, is enough as long as a multiplicative fault structure is used and a low level of parameter uncertainty is present.

ACKNOWLEDGMENT

Authors thank Tecnol\'ogico de Monterrey and CONACyT (PCP project 06/13 M\'exico-France) for their partial support.

REFERENCES