FDI WITH ADAPTIVE RESIDUAL GENERATION APPLIED TO A DC-SERVO

Erik Frisk[*]**, Mattias Nyberg**[**] **and Lars Nielsen**[***]

Vehicular Systems, Linköping University,
S-581 83 Linköping, Sweden. Fax: +46-13-282035
[*]*e-mail: frisk@isy.liu.se,* [**]*e-mail:matny@isy.liu.se,* [***]*e-mail:lars@isy.liu.se*

Abstract: When parity equations are designed ad hoc., the low frequency gain of the fault-to-residual transfer function can become low or even zero which is highly undesirable. As a solution, a systematic design procedure, utilizing the freedom available in residual generator design, is proposed. The method is called *residual shaping* and guarantees reasonable performance. It is used together with RLS to form an adaptive residual generator, which is important to handle e.g. aging. The RLS algorithm also makes it possible to differentiate between additive and multiplicative faults. The proposed FDI system is evaluated on a DC-servo where the strengths of the system is illustrated.

Keywords: fault detection, fault isolation, fault identification, adaptive, DC-servo, fault diagnosis

1. INTRODUCTION

Fault Detection and Isolation (FDI) systems are becoming more common in many industrially produced products. An important example are automotive vehicles. Among typical requirements for an FDI system is that it should 1) work well for all individuals of the product, 2) be robust against aging and different operating conditions, 3) be capable of detecting and isolating a variety of specified faults, and 4) be able to estimate fault sizes. The design of the FDI system should be based partly on these requirements and also on other considerations. The objective of this paper is to demonstrate a design of an FDI system, where the listed requirements are taken into account. The whole structure of the

Fig. 1. The FDI system.

proposed FDI system, is illustrated in Figure 1. To meet the first two requirements an adaptive FDI system is desirable. Therefore is an adaptive residual generator developed, which is discussed in Section 4.

The third requirement implies that different kinds of faults may need to be detected and diagnosed. Different techniques are suitable for different kinds of faults. As can be seen in the figure, the proposed FDI system contains parity function based residuals for sensor and actuator faults and parameter estimation for detecting changes in parameters, i.e process faults. Techniques, for achieving the fourth requirement, are *residual shaping*, and a *don't care* option in the residual structure which also increases the robustness provided by the residual structure. These topics are discussed in Section 3.

The process chosen to exemplify the design is an experimental setup of a DC-servo. The set of faults and the process choice is inspired both by a practical application of a throttle in an automotive engine (Nyberg and Nielsen, 1997b) and by the benchmark proposed in (Blanke *et al.*, 1995). The process and the faults are described in Section 2. An FDI system with the proposed structure is developed for the DC-servo and the resulting system is tested experimentally and evaluated in Section 6.
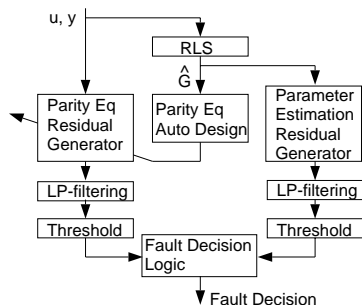
2. PROCESS DESCRIPTION

The experimental setup used to validate the FDI system is a standard laboratory DC-servo equipped with both angle sensor and tachometer. It is assumed that any load on the DC-servo can be estimated. The dynamics are well captured by a 2nd-order linear model.

In this paper a discrete model is used. However the results are applicable in the continuous case as well. In this investigation, model performance is not of primary concern. Comparisons of model output and validation data show that the model captures process dynamics well up to approximately 13 rad/s ≈ 2.5 times system bandwidth.

2.1. Fault modelling

The faults considered are: angle-sensor fault f_ϕ, tacho-meter fault f_ω, actuator fault f_a, and a parametric change of the input gain of the DC-servo. The sensor and actuator faults are modeled as *additive* faults whilst the parameter change is modeled as a multiplicative fault as

$$\omega(t+1) = a\omega(t) + (b + \Delta b)(u(t) + f_a(t)) \quad (1)$$

The choice of investigating a change in Δb rather than Δa (or both) has no physical motivation. This work is a principle study and extending the algorithm with Δa-monitoring is straightforward, therefore only Δb-changes are monitored.

Further on, three different fault characteristics are considered, step-faults with constant amplitude, incipient faults, and complete failures, e.g. shortcuts of sensor signals or motor seizure.

To illustrate the complications with introducing both additive and multiplicative faults, (1) is rewritten

$$\omega(t+1) = a\omega(t) + bu(t) +$$
$$+(b + \Delta b)(\frac{\Delta b}{b + \Delta b}u(t) + f_a(t)) \quad (2)$$

It is clear from the equation above that if no further assumptions are made, $\frac{\Delta b}{b+\Delta b}u(t)$ will be misinterpreted as an actuator fault. The assumption made in this paper is that $f_a(t)$ is uncorrelated with $u(t)$, thus it is possible to separate multiplicative and additive faults by utilizing the correlation between $\frac{\Delta b}{b+\Delta b}u(t)$ and the known signal $u(t)$.

When considering a real application and parametric faults are to be considered, a continuous model might be more suitable as physical parameters then are directly coupled to model parameters. In a discrete model, the relationships between model and physical parameters are much more involved.

Sensor faults are simulated by adding the desired fault directly onto the measured signal. For actuator faults and parametric changes the commanded control-signal is recalculated such that with the desired faults the resulting control-signal is the actual commanded control-signal. Denote the commanded control-signal with $u(t)$ and the recalculated signal

$u'(t)$ which is the one fed to the FDI system. The calculations are performed as follows.

$$\omega(t+1) = a\omega(t) + bu(t) =$$
$$= a\omega(t) + (b + \Delta b)(u'(t) + f_a(t))$$

From the equation above it is clear that

$$u'(t) = \frac{b}{b + \Delta b}u(t) - f_a(t)$$

All this is possible since the DC-servo does not operate in closed-loop, which would complicate the fault simulations.

3. RESIDUAL GENERATION

Many design methods for linear residual generation exists. All results in a filter for which the computational form, i.e. how the residual is computed, is

$$r = \frac{A_1 y_1 + \ldots + A_m y_m + B_1 u_1 + \ldots + B_k u_k}{C} \quad (3)$$

where A_i, B_i, and C are polynomials in q if discrete time is considered. This results in a transfer function

$$G_{rf_j} = \frac{D_j(q)}{C(q)}$$

from fault j to the residual. The objective of residual generation is to create a signal that is affected by a specific subset of faults but not by any other signals. This is equivalent to finding a filter which fulfills the following two requirements: the transfer functions from the faults, in the specific subset, to the residual must be non-zero, and the transfer functions from all other signals to the residual must be zero, i.e. *decoupling*. These two requirements introduces a constraint on the numerator polynomial of (3) only. Thus the denominator polynomial C can be chosen freely. The numerator polynomial is often denoted *parity function* (Chow and Willsky, 1984). Therefore, residuals designed by primarily studying the numerator polynomial are, in this work, denoted *parity function based residuals*.

A simple and straight-forward method for deriving numerator polynomials which fulfills these requirements is the *Universal Linear Parity Equation* (ULPE) scheme (Nyberg, 1997). Also this method is universal in the sense that all possible numerator polynomials, satisfying the requirements, can be found.

In the DC-servo example, three sets of numerator polynomials with different decoupling properties exists. For polynomials within these three sets, f_ω, f_ϕ, and f_u are decoupled respectively.

Another kind of residuals are based on parameter estimation. A typical residual in this case is

$$r = \| \hat{\theta} - \theta_0 \|$$

where $\hat{\theta}$ is an estimation of model parameters and θ_0 is the initial estimate. In the DC-servo example a residual $r_{\Delta b} = \hat{b} - b_0$ is used, where $\hat{b}$ is obtained from the RLS algorithm.

2

3.1. Residual Shaping

As noted above, the C polynomial can be chosen freely. Also for a certain polynomial degree there may exist several linear independent numerator polynomials with the same decoupling properties. This adds to the freedom available. It is important to consider this freedom in the design of the residual generator, so that the residuals respond to the fault in the frequency range of interest. Otherwise the frequency response can become highly undesirable. For example, a fault that is strongly detectable (Chen and Patton, 1994) can in the design result in a residual in which the fault is not strongly detectable.

The freedom can be utilized to *shape* the residuals in the frequency domain so that the transfer functions from faults to residuals get appropriate properties in the frequency range of interest. One desirable property is that no fault frequencies should be weighted higher than other, i.e. the transfer function $G_{rf}(z)$ should be constant for all frequencies. For some faults this is not possible without using an integration or unstable residual generator. For these cases it is only possible to obtain a $G_{rf}(q)$ that is approximately constant for some frequencies. In the DC-servo example, this is the case for the angle sensor fault, which is not strongly detectable. Note that strong detectability is a *system* property and not a property of the residual generator. It is possible to show that there exists *no* linear residual generators where f_ϕ is strongly detectable (Nyberg, 1997).

A simple method to achieve this is to first utilize the freedom to select A_i and B_i polynomials so that, if possible, the polynomial D_j have no zeros outside the unit circle. Then the C-polynomial can be chosen to be equal to D_j. If this is not possible, one can choose the C-polynomial such that each zero equals or mirrors the zeros of D_j. For zeros of D_j on the unit circle, the zeros of the C-polynomial should be placed close to the zeros of D_j and within the unit circle. This method is systematic and possible to automatize which is the basis of the adaptive residual generators.

If a pair of polynomials A_i and B_i is considered, then for different faults, the corresponding polynomials D_j are generally not equal. This implies that different C polynomials may be needed for different faults. In other words, for each pair of polynomials A_i and B_i, one residual is needed for each fault that is not decoupled.

When this procedure is applied to the DC-servo example, the transfer functions from monitored faults become $G_{rf}(q) = 1$ for u-faults and w-faults. The transfer functions from the ϕ-fault get a frequency response according to Figure 2. As said above, this is a consequence of that this fault is not strongly detectable.
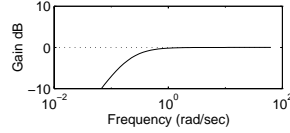

Fig. 2. Transfer functions from ϕ-fault to residual

3.2. Residual Structure

To decide what fault has occured, the residuals need to match a column in the residual structure (also denoted coding set, incidence matrix etc.). Because each residual is designed to detect and diagnose one specific fault, it is desirable that when this fault occurs, residuals that responds to the fault, but is not designed specifically for the fault, should not be considered in the fault decision. A residual structure with only ones and zeros provides no support for this. A natural solution is to introduce *don't care*, see (Nyberg and Nielsen, 1997a), in the residual structure. If the residuals are required to match the residual structure to signal alarm, the residual structure provides robustness against false alarm. Introducing don't care will increase this robustness, i.e. the false alarm rate is decreased. Another advantage is that individual thresholds can be used for each fault. The idea to have a specific residual dedicated for each fault does not affect the isolation properties, e.g. strong isolability, of the residual structure.

The residual structure of the parity function based residuals, for the DC-servo, is shown in Figure 3. Also included is the parameter estimation residual. It can be seen that this residual structure is strongly isolating and that residual shaping can be performed individually for each fault, without affecting the ability to detect other faults. The meaning 0^* is that before the RLS has converged, 0^* should be considered an X but when the RLS estimator has converged, this is a true 0.

	u	ϕ	ω	Δb
r_1	X	1	0	0^*
r_2	1	X	0	0^*
r_3	X	0	1	0^*
r_4	1	0	X	0^*
r_5	0	X	1	0^*
r_6	0	1	X	0^*
$r_{\Delta b}$	0^*	0	0^*	1

Fig. 3. The residual structure of the FDI system.

3.3. Evaluation of Residual Shaping

The importance of residual shaping can for example be seen by studying incipient faults in combination with faults that are not strongly detectable. Figure 4 illustrates the importance of shaping in this case. The left plot shows a non-shaped residual response to an incipient, non strongly detectable fault, e.g. a ϕ-fault. The right shows a shaped residual. In this case the shaping consists of plac-

ing a discrete pole in 0.99. The non shaped residual does not reach the threshold, but the shaped does. It is clear that although the lack of strong detectability problem is not solved, the shaping makes the fault easier to detect.
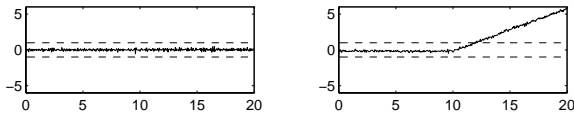

Fig. 4. The importance of residual shaping in the case of an incipient fault in combination with a not strongly detectable fault.

To further study the effect of residual shaping, faults are applied to the DC-servo. The resulting residuals are shown in Figures 5 and 6 for ω-, and ϕ-fault respectively. The faults are step faults of sizes 0.2 V, and 2 V, applied at time 10 s. It can be seen that the residuals respond in correspondence with the residual structure in Figure 3. Also in Figure 5, r_3 and r_5 have a response in well correspondence with the fault, i.e. are estimations of the fault-size. A similar behavior is can be seen for u-faults. This is not the case for ϕ-fault as it is not strongly detectable. For the the ω- (and u-)fault, it can be
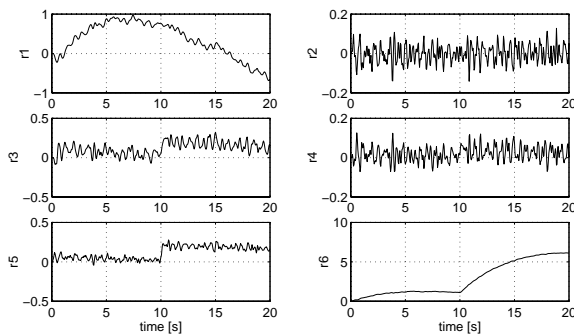

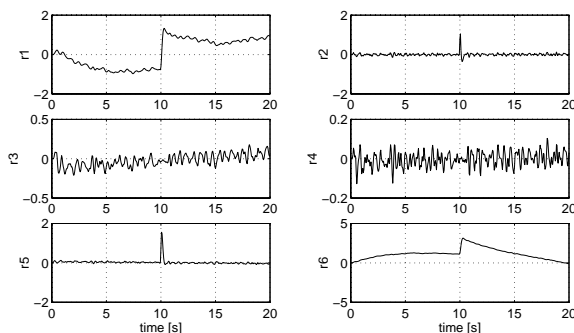Fig. 5. The residuals when a ω-fault is applied.

Fig. 6. The residuals when a ϕ-fault is applied.

seen in Figure 5 that the shaping has an overall advantage, in the sense that it shortens detection delay or increases the effect of the fault onto the residual. This is especially obvious if r_3 and r_4 are compared where r_4 is shaped especially for ω-faults.

In Figure 6 it can be seen that the residuals designed to detect ϕ-faults, i.e. r_1 and r_6, drifts away

from zero before the fault is applied. The reason is a non-desirable side-effect of the residual shaping, which is close to a pure integration. However if the plot of r_1 and r_6, are compared to r_2 and r_5 which are not shaped for this fault, it can be seen that the shaping increases the effect of the fault on the residuals. Thus for faults that are not strongly detectable, there is almost always a compromise between a non-desirable drift and offset in the fault-free case, and a large effect of the fault onto the residual.

4. ADAPTIVITY

A model based FDI system relies on a model, and this model should be as good as possible. However for mass produced products, it is impossible to identify an optimal model for each individual; a universal model that captures the mean behavior of all individuals must be used. This together with ability to handle aging and varying operating conditions is a motivation for including some kind of adaptability in the FDI system.

The proposed FDI system structure, shown in Figure 1, has adaptive residual generation. Another possibility, not utilized in this work, is adaptive thresholds as in (Höfling and Isermann, 1996). The principle of the adaptive residual generator is to on-line estimate a model of the process and redesign the residual generator with a certain time interval. Adaptivity have also been studied in e.g. (Ding et al., 1989, Höfling and Isermann, 1996).

The on-line model estimator used is a straightforward RLS algorithm. From the estimated model, the residual generator including residual shaping is automatically designed according to the procedure described in Section 3.1. Polynomials A_i and B_i are derived with the ULPE scheme for each residual. These polynomials are chosen such that the zeros are placed within the unit circle if possible. Then an appropriate C polynomial is chosen.

For the DC-servo only parametric changes Δb are studied and therefore other model parameters do not need to be estimated. This can be accomplished by using (2) and calculating the signal used in the RLS algorithm as

$$
\begin{aligned}
e(t) &= \omega_s(t) - a\omega_s(t-1) - bu(t-1) = \\
&= (f_\omega(t) - af_\omega(t-1) + (b+\Delta b)f_a(t-1)) + \\
&\quad + \Delta bu(t-1) = f_1(t) + \Delta bu(t-1) = \\
&= \begin{bmatrix} 1 & u(t-1) \end{bmatrix} \begin{bmatrix} f_1 \\ \Delta b \end{bmatrix} = \varphi^T(t)\theta(t)
\end{aligned}
$$

where $\omega_s(t)$ is the sensor signal, i.e. $\omega_s(t) = \omega(t) + f_\omega(t)$. Assuming that θ is uncorrelated with $u(t)$, θ can be estimated, and thereby Δb. Note that this estimation can be performed even if other faults are present. The only requirement is an exciting $u(t)$. It is however important to suspend the estimation algorithm when $u(t)$ is non-exciting, e.g. constant, due to the risk of diverging estimations. The forgetting factor in RLS has been chosen $\lambda = 0.96$

to get a good balance between parameter tracking accuracy and speed. The residual generator is re-designed with a 1 second interval.

5. RESIDUAL EVALUATION

The residual evaluation consists of two steps, 1) detection and isolation of a fault and 2) estimation of fault size. The first step is a simple thresholding test of the residuals and comparing the residuals with the columns in Figure 3. If there exists a match, the fault is detected and isolated. The thresholds are set so that as small faults as possible can be detected and isolated without generating any false-alarms on the data gathered during this work. Before thresholding the residuals are LP-filtered with cut-off frequency 13 rad/s because, as we noted in Section 2, our model is valid up to about this frequency. The following identification of fault size is one of the shaped residuals, e.g. for estimation of a u-fault, r_2 or r_4 are both, according to Figure 3, estimations of f_a. Note that the size of ϕ-faults cannot be estimated with the algorithm described here as the fault is not strongly detectable.

6. EXPERIMENTS WITH THE COMPLETE FDI SYSTEM

The minimum size of faults that can be detected depends partly on the algorithm used, but first and foremost on the quality of the model. In this work, where emphasis has been on the algorithms not the model, it is possible to detect 0.2 V-faults in the actuator and ω-sensor, 2 V-faults in the ϕ-sensor. This corresponds to 2% and 20% respectively of total signal-span. When, and how accurate Δb changes can be detected strongly depends on how exciting the control-signal is.

6.1. Adaptive performance

The adaptive residual generators improves performance significantly when system dynamics change. This is illustrated by Figure 7. At $t = 5$, a change
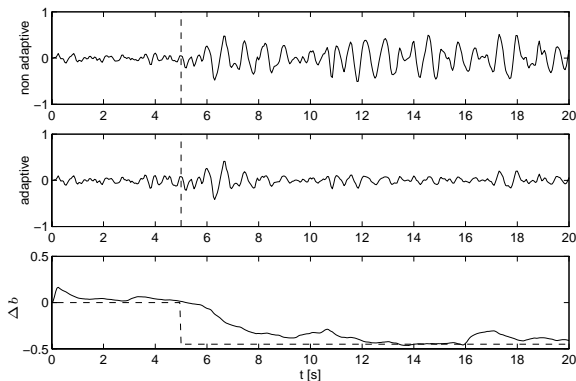

Fig. 7. Illustration of how adaptability improves system performance

in parameter Δb is introduced. No other faults are present, and thus the residuals should be 0. The lower plot shows how well the parameter change is estimated where the solid line is the estimation and the dotted line the real value of the tracked parameter. The upper plot show one of the residuals without adaptability and the middle shows the same with adaptability. It is clear that when the parameter estimation has found its stationary value the adaptive residual has distinctly lower variance than the non-adaptive thus giving a safer behavior.

6.2. Discrimination between multiplicative and additive faults

Adaptability also leads to more accurate fault estimation as can be seen in Figure 8, where a change in Δb is introduced at $t = 5$, and an actuator fault is introduced at $t = 10$. In the upper plots the binary decision for actuator faults is shown and in the middle, an estimation of the fault size $\hat{f}_a$, calculated as $\hat{f}_a = d_a(t)(r_2(t) + r_4(t))/2$. $d_a(t)$ is the binary decision for actuator faults seen in the upper plots. The lowest plots shows the Δb-estimations. It is
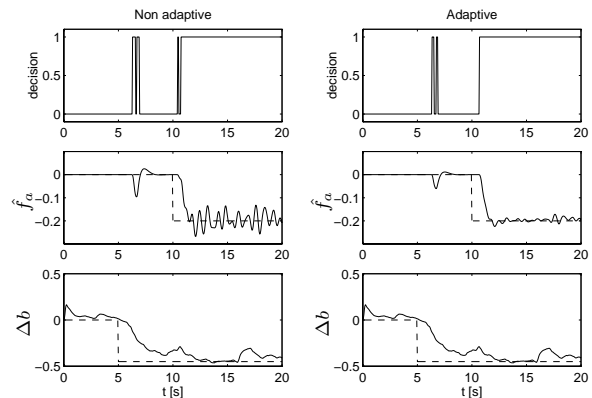

Fig. 8. Discrimination between Δb and f_a

seen that a false-alarm is present, at $t \approx 6$, both in the adaptive and the non-adaptive system. This is due to that the Δb estimation has not yet converged to the correct value. When the Δb-estimation has converged at $t \approx 11$, the fault-size estimation $\hat{f}_a$ is distinctly better when adaptation is used. These plots also show how the RLS filter makes it possible to differentiate between the multiplicative change Δb and the additive fault f_a which is not possible with a straightforward parity equation approach as seen in Equation (2).

6.3. Complete failure

A common fault in a real-world application is a short circuit between e.g. earth and sensor output thus making sensor output zero. A short-circuit in the ω-sensor, and with the additive fault model assumed, $f_\omega(t) = -\omega(t)$. If $w(t)$ is constant, an abrupt short-circuit will be seen as a step fault. However in a real situation $f_\omega(t)$ might look like the top plot in Figure 9. The middle plots shows the residual response to the short-circuit where the dotted lines are the thresholds. Only four residuals are

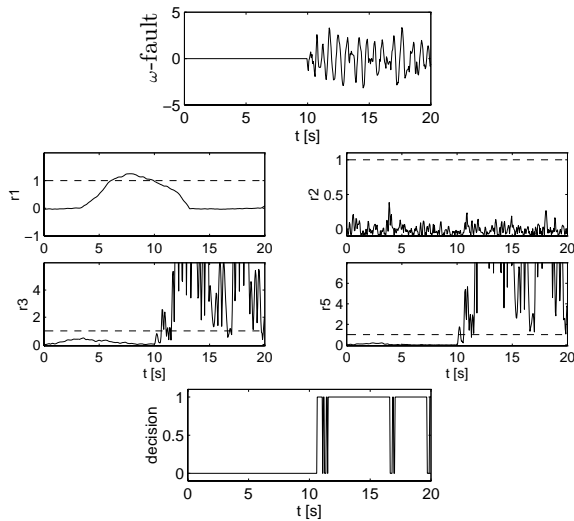plotted since r_4 and r_6 has a don't care in the corresponding positions in the residual structure. The

Fig. 9. Residuals and decision plots when ω-sensor has a short-circuit to earth

lowest plot show the binary fault decision. From Figure 9 it is clear that the model-based approach works well also for this fault type. However, the performance for short-circuits is strongly dependent on the sensor signal. If it is possible to choose a sensor where 0 V is not in the operating range, a more traditional fault detection approach would be more suitable for this fault type. Another realistic fault is a complete motor seizure, a fault that is not explicitly modeled here. A seizure can here naturally be seen as an u-fault with $f_a(t) = -u(t)$. The residuals will behave similarly as in the short-circuit example, which has been validated on the experimental setup. Further details on the seizure experiment and experiments with incipient faults can be found in (Frisk *et al.*, 1997).

6.4. ϕ-sensor fault

A fault in the ϕ-sensor is hard to detect as they are not strongly detectable. This, and the fact that even small model faults can result in considerable effects on the ϕ estimations due to the integration in the model lead to that only rather large and abrupt faults can be detected with the algorithm presented here. Thus it is possible that also here, a traditional range/rate-test would be more effective.

7. CONCLUSIONS

In this paper, the structure and design of an adaptive FDI system is proposed. When the freedom available in parity equation design is utilized ad hoc., the low frequency gain of the fault-to-residual transfer function can become low or even zero which is highly undesirable. As a solution, a systematic design procedure, utilizing the freedom available, is proposed. The method is called *residual shaping* and guarantees reasonable performance. The design method has been automatized and together

with an on-line RLS estimator, this forms the basis for an adaptive residual generator scheme. With an adaptive approach it is possible to handle individual differences, aging, and different operating conditions in mass-produced products. The use of the RLS algorithm also makes it possible to differentiate between additive and multiplicative faults which was not possible without the RLS filter.

The proposed system is evaluated on an experimental DC-servo setup, where the angle-sensor and tacho-meter, the actuator, and the input gain of the DC-servo is monitored. Experiments show that all strongly detectable faults, i.e. ω-, and u-faults as well as the parametric change is well monitored. Also the angle sensor is monitored, but because the fault is not strongly detectable, this problem is more challenging. When input gain is changed, the adaptive residual generator is shown to distinctly improve system performance.

8. REFERENCES

Blanke, M., S.A. Bøgh, R.B. Jørgenssen, and R.J. Patton (1995). Fault detection for a diesel engine actuator - a benchmark for FDI. *Control Eng. Practice*, **3**(12), 1731–1740.

Chen, J. and R.J. Patton (1994). A re-examination of fault detectability and isolability in linear dynamic systems. Fault Detection, Supervision and Safety for Technical Processes, pp. 567–573. IFAC, Espoo, Finland.

Chow, E.Y. and A.S. Willsky (1984). Analytical redundancy and the design of robust failure detection systems. *IEEE Trans. on Automatic Control*, **29**(7), 603–614.

Ding, X., P.M. Frank, and L. Guo (1989). Fault detection via adaptive observers based on orthogonal functions. Advanced Information Processing in Automatic Control, pp. 95–100. IFAC, Nancy, France.

Frisk, E., M. Nyberg, and L. Nielsen (1997). FDI with adaptive residual generation applied to a DC-servo. Technical Report LiTH-R-1952, Department of Electrical Engineering, Linköping University, S-581 83 Linköping, Sweden.

Höfling, T. and R. Isermann (1996). Fault detection based on adaptive parity equations and single-parameter tracking. *Control Eng. Practice*, **4**(10), 1361–1369.

Nyberg, Mattias (1997). *Model Based Diagnosis with Application to Automotive Engines*. Licentiate Thesis, Linköping University, Sweden.

Nyberg, M. and L. Nielsen (1997a). Design of a complete FDI system based on a performance index with application to an automotive engine. Fault Detection, Supervision and Safety for Technical Processes. IFAC, Hull, United Kingdom.

Nyberg, M. and L. Nielsen (1997b). Model based diagnosis for the air intake system of the SI-engine. *SAE Paper* (970209).