

# Design of a Complete FDI System based on a Performance Index with Application to an Automotive Engine

Mattias Nyberg  
Vehicular Systems, ISY, Linköping University  
S-581 83 Linköping, Sweden.  
e-mail: matny@isy.liu.se

September 8, 1997

## Abstract

Assuming residual generators are already available, there are still several choices to be made when a complete FDI system is to be designed. This is a time-consuming engineering work so for this purpose, a systematic procedure is proposed. The procedure is phrased as an optimization problem. The goal is to minimize a new probability based performance index, which is derived from measurements on the real process. To increase the robustness of the FDI system, a *don't care* option is introduced in the residual structure. The procedure is successfully applied to the problem of FDI design for the air intake system of an SI-engine.

## 1 Introduction

Model based fault detection and isolation (FDI) has received much attention during the last decade. The reasons are increased requirements in some key areas, e.g. automotive diagnosis, and the increased availability of microprocessors, which are needed for the realization of any model based FDI system. Important performance issues are false alarm rate, fault detection delay, probability of missed detection of faults, and minimal fault sizes that can be detected. All these properties need to be minimized and the major limiting factor is the model accuracy. However to build accurate models requires a lot of work and in most applications, there are practical limits of how good models that can be obtained. It is therefore of high importance that, based on a model, that may not be very good, at least the FDI system is optimal in the sense of the properties listed above.

Another important issue is to minimize the time-consuming engineering work, that is frequently needed for the design of FDI systems. Even though residual generators are already available, there are still several choices to be made when a complete FDI system is to be designed. For instance, in the automotive engine example provided in this work, there are  $2^{15}$  different residual structures to choose between. The search for a relevant engineering method to handle this problem of designing a complete FDI system, that is also optimal, is the key idea of this paper.

The design of the FDI system is set up as an optimization problem. The goal is to minimize a probability based performance index, described in Section 2, aiming at measuring the performance of the whole system. The optimization results in a selected set of residuals, threshold values, and a residual structure including a *don't care* option. The *don't care* option increases the robustness of the system and is discussed in Section 3.3. The procedure for solving the optimization problem is provided in Section 4. In Section 5, the procedure is applied to a real application, FDI of the air-intake system of an SI-engine.

## 2 Performance Indices

To design an optimal FDI system, some kind of performance index (PI) is needed. Some PI's for linear residual generators, have been proposed in the literature. Typically these are based on some relative gain from faults to residual. Some examples are

$$\begin{aligned}\pi_1 &= \frac{|G_{rf}(q)|}{J_{thresh}} \\ \pi_2 &= \frac{|G_{rf}(0)|}{\sigma^2} \theta \\ \pi_3 &= \frac{\|G_{rf}(q)\|}{\|G_{rd}(q)\|}\end{aligned}$$

$\pi_1$  from (Gertler and Costin, 1994) is frequency dependent and normalized with the threshold value.  $\pi_2$  from (Chow and Willsky, 1984) is a DC-gain normalized with  $\theta$  being the magnitude of the coefficients, and the variance of the residual in the fault free case.  $\pi_3$  from (Frank, 1993) also takes signals that need to be decoupled, into account. Such signals are disturbances in the form of structured uncertainties and faults that need to be decoupled because of isolation reasons.

The purpose of these PI's is to optimize the residual generators. In for example (Frank, 1993),  $\pi_3$  is used to maximize the sensitivity of a residual to faults while minimizing the sensitivity from disturbances and non monitored faults. However these PI's are not aimed at guiding the design of the complete FDI system. Also they are not generally applicable to non-linear systems.

Another problem with the previously proposed PI's, more exactly  $\pi_1$  and  $\pi_2$ , is that residuals affected by unstructured model uncertainties are not always

correctly evaluated. For  $\pi_3$ , unstructured uncertainties are not taken in account at all.

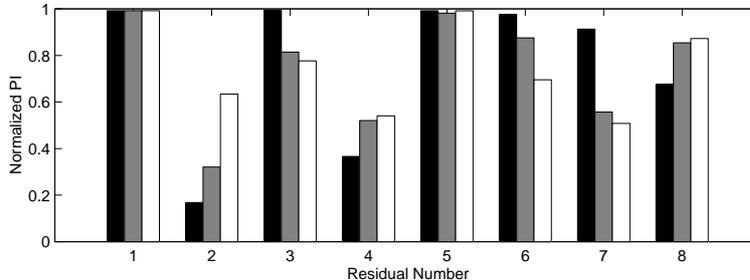


Figure 1: Comparison of three PI's for eight residuals.

To see how different PI's performs when used to compare different residuals, a test was carried out. Eight residuals, affected by model uncertainties, were evaluated by three PI's. Note that these PI's are *not* the PI's  $\pi_1$ ,  $\pi_2$ , and  $\pi_3$ . Figure 1 shows the three PI's, which are represented by black, gray, and white bars respectively, when applied to the eight residuals. The data are based on measurements from a real process in a test case when a constant fault is present all the time. The residuals were affected by a significant amount of unstructured model uncertainties. Two of the PI's require that thresholds are fixed. Therefore each threshold has been chosen as the maximum value of the residual in the fault free case. The black bars are the amount of time the residuals are above the thresholds. This is similar to the PI proposed in Section 2.1. The gray bars are the residual mean over time divided by the threshold, and thus equals  $\pi_1$  for constant faults. The white bars are the residual mean over time divided by the standard deviation of the residual in the fault free case. Except for the coefficient  $\theta$ , this is similar to  $\pi_2$  for constant faults. The grey and white bars have been normalized so that they equals the black bar for residual number 1.

Now if the PI's are to be used to compare residual performance, there is a strong indication that residual 6 is better than residual 7, since each bar of residual 6 is higher than the corresponding bar for residual 7. However for most residuals it is not easy to conclude which ones that are the best. For example if the white bars are considered, it is seen that residual 2 performs slightly better than residual 7. However according to the black bars, residual 2 correctly detects the fault during only 17% of the time during this test, while residual 7 detects the fault during 91% of the time. Therefore residual 7 should be the better one in this case. A similar discrepancy between the different PI's is seen if residual 7 and 8 are compared. From this study, it can be concluded that the PI's represented by the gray and white bars and closely related to the previously proposed  $\pi_1$  and  $\pi_2$ , do not correctly evaluate the ability of the

residuals to detect faults when unstructured model uncertainties are present.

## 2.1 A probability based PI

As an alternative to the previously proposed PI's  $\pi_1$ ,  $\pi_2$ , and  $\pi_3$ , a probability based PI is proposed. The idea is to study the probabilities of the following naturally defined events:

- “not detect a fault given that no fault has occurred”
- “detect fault  $j$  given that fault  $j$  has occurred”
- “detect a fault given that no fault has occurred”
- “not detect fault  $j$  given that fault  $j$  has occurred”

The first and the third event are the complement of each other. The same is true for the second and fourth event. Therefore without loss of generality, the proposed PI is the probability of the third and fourth event. These probabilities can be interpreted as the *probability of false alarm* and *probability of missed detection* respectively. In other words, the PI is defined as the probability of malfunction of the FDI system. The PI should be minimized for several different fault situations. This requires special care, which is discussed in Section 4. Note that in this work, only single faults are considered.

The proposed PI have the advantage that it can be used both in the context of residuals and also in the context of a complete FDI system. Also it is not restricted to linear systems and correctly evaluates the impact of unstructured model uncertainties.

## 2.2 The PI for Residuals

In the context of residuals, the PI measures residual performance and can for example be used to select the threshold level. Note that the PI measures performance of the residual in combination with a threshold. Let  $B_i$  denote the event that the  $i$ :th residual signals *false alarm*. Also in the case fault  $j$  is present, let  $A_{ij}$  denote the event that residual  $i$  does not respond in correspondence with what it was designed to do. In other words,  $A_{ij}$  is the event *missed detection* if residual  $i$  is designed to respond to fault  $j$ , and the event *false alarm* if residual  $i$  is designed to not respond to fault  $j$ . The actual PI is in this case denoted  $P(B_i)$  and  $P(A_{ij})$  respectively. The goal is to minimize these probabilities but in most applications, the probabilities are not known. Therefore estimates, derived from measurements on the real process, must be used. An example of how this can be done is provided in Section 4.

## 2.3 The PI for the Complete FDI System

To guide the design of what residuals to include in the complete FDI system, a PI related to the total performance is needed. Thus for the complete FDI system, let  $\bar{B}$  denote the event *false alarm* and let  $\bar{A}_j$  denote the event *missed*

detection of fault  $j$ . The corresponding probabilities, i.e. the PI, are denoted  $P(\bar{B})$  and  $P(\bar{A}_j)$  respectively. These events and probabilities are more thoroughly discussed in the following section.

### 3 The Residual Structure

The residual structure (also called coding set or incidence matrix) is the basis for the design of an isolation scheme. It can be derived from the equations describing the residuals. Often it is desirable that the residual structure is *strongly isolating*. A strongly isolating structure protects small faults from being incorrectly isolated in the case when some residuals are not excited enough to reach the thresholds. Two examples of strongly isolating residual structures are shown in Figure 2.

	$f_1$	$f_2$	$f_3$		$f_1$	$f_2$	$f_3$
$r_1$	1	0	0	$r_1$	0	1	1
$r_2$	0	1	0	$r_2$	1	0	1
$r_3$	0	0	1	$r_3$	1	1	0

Figure 2: Two examples of residual structures.

#### 3.1 Selecting Residual Structure

The residual structure can to some extent be determined already in the design of the residual generators. In addition, selecting different sets of already designed residual generators results in different residual structures. Thus a possible large freedom exists in the design of the residual structure. However except from isolation requirements, it is not obvious how this freedom should be utilized. Following are two examples aiming at illustrating this problem.

Consider the residual structures in Figure 2. An advantage with the left residual structure compared to the right, is that it can isolate multiple faults. However it is less robust against disturbances in the fault free case. Suppose disturbances are present, then to get a false alarm, it is in the left structure sufficient that only one residual change from zero to nonzero. In the right structure, it is necessary that two residuals change from zero to nonzero. In this way the right residual structure provides robustness against false alarms. If also the probability of missed detection is considered and without more knowledge, it is however impossible to conclude what residual structure is the best.

To further illustrate the problem, consider the task of designing or selecting residual generators to be included in the FDI system. It is important to know what set of residuals that constitutes the optimal structure. For example, is it an advantage to include an extra residual? For the moment, consider only fault 1 in the right residual structure in Figure 2. Obviously at least one of

$r_2$  and  $r_3$  must be included. Also obvious is that many residuals increase the robustness against false alarm and few residuals increase the robustness against missed detection of faults. However without more knowledge, that is about everything that can be concluded so even in this simple case there is no straight answer to what residuals to be included to get the optimal residual structure. To answer the question, it is first necessary to specify what is meant by optimal residual structure; a PI is needed.

### 3.2 The PI for Selection of Residual Structure

A natural PI is the one defined in the previous section, i.e. the probability of false alarm  $\bar{B}$  and the probability of missed detection  $\bar{A}_j$ . The idea is best illustrated by an example. Consider the right residual structure in Figure 2. In the following it is assumed that fault detection and isolation is the same task, i.e. the FDI system contains only isolation, and detection is an implication of isolation. All residual related probabilities, associated with this residual structure, are summarized in Table 3.2.

	no fault	$f_1$	$f_2$	$f_3$
$r_1$	$P(B_1)$	$P(A_{11})$	$P(A_{12})$	$P(A_{13})$
$r_2$	$P(B_2)$	$P(A_{21})$	$P(A_{22})$	$P(A_{23})$
$r_3$	$P(B_3)$	$P(A_{31})$	$P(A_{32})$	$P(A_{33})$

Table 1: All residual related probabilities associated with the right residual structure in Figure 3.2.

Let  $\bar{B}_j$  denote the event that in the fault free case, the FDI system signals that fault  $j$  has occurred. This means that in the example,  $\bar{B}_1 = B_1^C \cap B_2 \cap B_3$  etc. Note that the index  $i$  of  $B_i$  denotes residual number while the index  $j$  of  $\bar{B}_j$  denotes fault number. The event of any false alarm is

$$\bar{B} = \bar{B}_1 \cup \bar{B}_2 \cup \bar{B}_3$$

If we assume that the residual structure is at least weakly isolating, then the events  $\bar{B}_j$  are disjoint, which implies

$$P(\bar{B}) = P(\bar{B}_1) + P(\bar{B}_2) + P(\bar{B}_3)$$

This relationship is used in the procedure in Section 4, in which the minimization of  $P(\bar{B})$  is approached by trying to minimize all  $P(\bar{B}_j)$ . The other probabilities that are used as PI are the probabilities of missed detection of fault  $j$ , that is

$$P(\bar{A}_j) = P(A_{1j} \cup A_{2j} \cup A_{3j}) \quad j = 1 \dots 3$$

To summarize, the global goal is to minimize  $P(\bar{B})$  and all  $P(\bar{A}_j)$ . The optimization must be performed under the constraint, that the residual structure

must be weakly isolating and possibly also strongly isolating. Details on how this optimization can be performed are provided in Section 4.

### 3.3 Don't Care in the Residual Structure

*Don't care*, X, is an alternative to *ones* and *zeros* in the residual structure. The meaning of an X, in the position for the  $i$ :th residual and the  $j$ :th fault, is that the fault decision should not take any notice about the value of residual  $i$  when deciding if fault  $j$  has occurred. In most situations a residual is affected differently by different faults. This is the underlying reason why there is a need for *don't care*, or X, in the residual structure. Explicit examples of such cases are:

- Model uncertainty can force the residual threshold to be set high. This can have the effect that faults with low fault-to-residual gain are missed although faults with high gain are still detected. For detection / isolation of a fault with low gain, it is not a good idea to require some specific behavior of this residual, i.e. a *one* or *zero* in the residual structure. Introducing *don't care* solves this problem.
- Non linear physics in the process can make the fault-to-residual gain for a specific fault, decrease or even get zero in some operating areas. This makes the residual useless for detecting this fault but it can still be useful for other faults. Again *don't care* solves the problem.

An additional reason to use *don't care* is that a residual structure can be augmented so that each fault get its own residuals dedicated to only this fault. In this way no residuals must be a compromise between detecting different faults. Isolation properties are not affected by *don't care* used in this way. There are several applications for this technique. One is individual thresholds for each fault. This is useful to maximize the robustness of the FDI system and is used to some extent in the procedure described in Section 4.

Another application is residual shaping (filtering) (Frisk *et al.*, 1997) which is a technique to shape the frequency response, from fault to residual, of the residual generator. Different shaping is probably needed for the different faults, and the use of residual structure augmentation in combination with *don't care* makes this possible.

	$f_1$	$f_2$	$f_3$
$r_1$	X	1	0
$r_2$	X	X	1
$r_3$	1	0	X
$r_4$	0	1	1

Figure 3: An example of a residual structure including *don't care*.

An example of a residual structure including *don't care* is given in Figure 3. For anything but small residual structures, it is difficult to see if a residual structure containing X's is strongly, weakly or not isolating. Therefore the following Theorems are useful.

**Theorem 1.** *A residual structure is strongly isolating if and only if there is a one in each fault column and for all pairs of faults, it holds that the faults are strongly isolated from each other when rows containing X's, in any of the two fault columns, are neglected.*

**Theorem 2.** *A residual structure is weakly isolating if and only if there is a one in each fault column and for all pairs of faults, it holds that the faults have distinct columns when rows containing X's, in any of the two fault columns, are neglected.*

For example the residual structure in Figure 3 is only weakly isolating because fault 2 and 3 are not strongly isolated from each other.

When introducing X's, different strategies are possible. One that makes sense, and therefore used in the procedure described in the next section, is to only allow *ones* to be replaced by X's. In that case, for a residual sensitive to  $p$  faults, at a maximum  $p - 1$  X's can be introduced, in  $2^p - 1$  different combinations.

Introduction of X's will increase residual robustness and also the robustness of the complete system if done properly. When and how these X's should be introduced, can because of model uncertainties, generally not be concluded by studying the equations describing how the residuals are computed. Instead measurements on the real process must be used. This is in analogy with the strategy adopted in the present work, because X's are introduced based on the PI defined in the previous sections, and the PI is derived from measurements.

## 4 FDI Design based on the PI

Many methods for residual generation design exist, especially for linear systems. By using these methods a number of residual generators can be designed. It is not sure that all these residual generators perform satisfactory. Also the number of residual generators that can be designed is generally larger than what is needed to be able to detect and isolate the faults. In addition to using specific design methods, many residual generators can be designed with an ad hoc approach. Ad hoc design is often necessary for non-linear systems because of the scarcity of design methods. All this results in a possibly large set of residuals that are candidates to be included in the final FDI system. These residuals are denoted *base residuals*.

When residual generators, generating base residuals, have been designed, the remaining tasks of the FDI design are the selection between these residuals,

selection of thresholds, and introduction of X's in the residual structure. The idea is to perform this design by optimizing the probability based PI described in Section 2. The procedure for this optimization is described in the next section.

To formally describe the optimization problem, let the base residuals be denoted  $r_k$ ,  $k = 1, \dots, n$  and let  $p_k$  denote the number of faults that base residual  $r_k$  is sensitive to. By introducing X's in different combinations, each base residual  $r_k$  generates  $m_k \leq 2^{p_k} - 1$  residual candidates  $r_{kl}$ ,  $l = 1, \dots, m_k$ .

The global optimization problem can now be set up as shown in Table 2. The first column represents all residual candidates  $r_{kl}$  obtained from base residuals denoted  $r_k$ ,  $k = 1, \dots, n$ . The optimization is with respect to the variables in the other two columns. Each variable  $b_{kl}$  in the second column determines whether residual candidate  $r_{kl}$  is included in the FDI system or not. These variables take values *yes* or *no*. If  $b_{kl} = \text{yes}$ , then also the threshold level  $J_{kl}$  needs to be optimized.

residual candidate	included	threshold level
$r_{11}$	$b_{11}$	$J_{11}$
$r_{12}$	$b_{12}$	$J_{12}$
$\vdots$	$\vdots$	$\vdots$
$r_{1m_1}$	$b_{1m_1}$	$J_{1m_1}$
$r_{21}$	$b_{21}$	$J_{21}$
$\vdots$	$\vdots$	$\vdots$
$r_{2m_2}$	$b_{2m_2}$	$J_{2m_2}$
$\vdots$	$\vdots$	$\vdots$
$r_{nm_n}$	$b_{nm_n}$	$J_{nm_n}$

Table 2: The variables in the optimization problem.

The present optimization problem is nonlinear and contains both discrete and continuous variables, and is therefore generally hard to solve. The approach in the proposed procedure is to divide the global optimization problem into two subproblems. First, residual performance is optimized by introducing X's and selecting optimal threshold values. Based on the result of this optimization, the performance of the whole FDI system is optimized by selecting residuals to be included.

#### 4.1 The Procedure

The whole design procedure aims at solving the two subproblems described above. It contains four steps. The first subproblem is solved in steps 1A, 1B, and 1C, and the second in the single step 2. The flowchart of the procedure

is shown in Figure 4. Following is a detailed description of each step in the procedure.

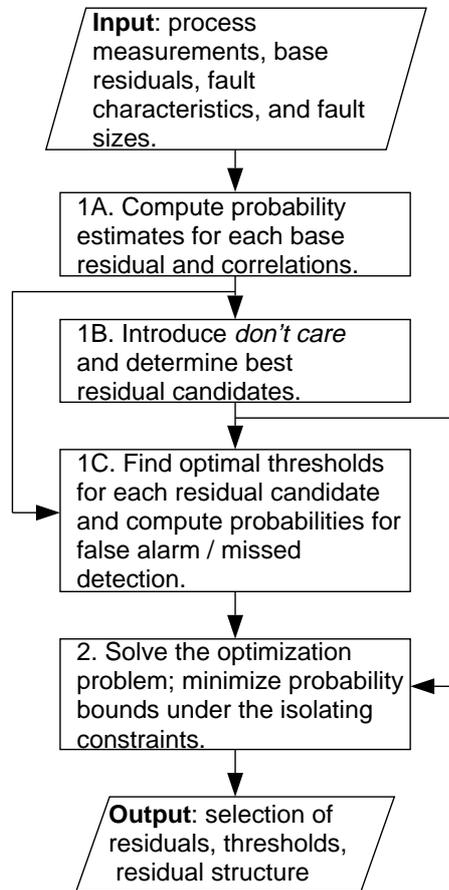


Figure 4: The flow chart of the residual evaluation design procedure. The arrows represent flow of information.

### Input

The inputs to the procedure are the already designed residual generators, generating base residuals, fault-free measurements on the real process, faults characteristics, and fault sizes. If constant bias or gain (additive or multiplicative) faults are chosen, the computations in step 1A are straightforward.

If measurements, in which the faults are present, are available, they can also be used. In that case, faults characteristics and sizes are not needed explicitly as an input.

### Step 1A

For the PI, probability estimates of  $B_i$  and  $A_{ij}$  are needed. These will be denoted  $\hat{P}(B_i)$  and  $\hat{P}(A_{ij})$  respectively. The estimates are of course dependent on the threshold levels. Therefore estimates are computed for each threshold level. Principally this is the same as estimating the probability density function. The estimation can be performed by measuring on the real process for the different fault cases and then calculating the amount of time that the residual is less than the threshold compared to the total time. This estimate converges asymptotically to the true probability, when measurement time is increased. If fault free measurement data are used, the estimates can be computed by introducing the faults in the measurements.

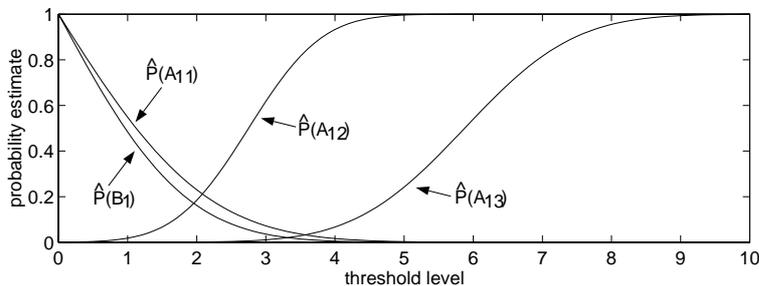


Figure 5: Probability estimates for a residual as a function of the threshold level.

An example of probability estimates can be seen in Figure 5. The example corresponds to residual number 1 in the right residual structure in Figure 2. As can be seen,  $\hat{P}(A_{11})$  is fundamentally different from  $\hat{P}(A_{12})$  and  $\hat{P}(A_{13})$ . This is because the residual structure has a *zero* for fault 1 and *ones* for fault 2 and 3. When perfect fault decoupling is achieved,  $\hat{P}(A_{ij})$  corresponding to *zeros* in the residual structure, coincide with  $\hat{P}(B_i)$ . This is not the case in the figure.

For each threshold level, the plots show what the probability is that the residual does not respond in accordance with the residual structure. For example if the threshold  $J_{thresh} = 2$ , the probability that the residual would respond to fault 1, even if it should not, is about 0.2. In the fault free case, the probability that the residual would reach the threshold is also about 0.2.

The result of step 1A is that probabilities as function of thresholds, like the ones plotted in Figure 5, are computed for each base residual. Also the correlation coefficients between all base residuals are computed.

### Step 1B

Figure 5 shows that whatever threshold is chosen, the probability that residual 1 would respond the way it shouldn't, never becomes less than about 0.2. This is an example of when *don't care* would increase the robustness of the performance. If the *one* corresponding to fault 2 in the residual structure is replaced with an X, the threshold can be chosen such that the probability that residual 1 would respond the way it shouldn't, becomes lower than 0.05, i.e.  $J_{thresh} \approx 3.5$ . By gradually replacing the limiting *ones* in this way, each base residual can generate  $p$  different residual candidates, where  $p$  is the number of *ones* in the base residual. This strategy generates less than the maximum number of residual candidates, which was discussed in Section 3.3. The motivation is that it is desirable to have few residual candidates and the ones selected with this strategy is probably the best ones in most cases.

The reason why the number of residual candidates should be kept low, is to lower the computational load in step 2. Further actions to reduce the number of residual candidates are that residual candidates for which the probability that they respond the way they shouldn't, is high for all threshold levels, are neglected. Also by using the correlation coefficients computed in the previous step, if two residual candidates are highly correlated, only the best one is kept. The correlation coefficient is sensitive to time delays between the residuals so to be sure all correlated residuals are detected, the correlation coefficient should be calculated for several different time delays.

### Step 1C

By using the probability estimates, computed in step 1A, optimal thresholds are chosen for all residual candidates individually. The optimal threshold for the  $i$ :th residual is defined as

$$J_{thresh,i} = \arg \min_{J_{thresh}} (\max(\hat{P}(B_i), \hat{P}(A_{i1}), \hat{P}(A_{i2}), \dots))$$

where the maximization is taken over  $\hat{P}(B_i)$  and all  $\hat{P}(A_{ij})$  corresponding to faults  $j$  for which there is *not* an X in the residual structure. If more than one threshold minimizes the maximum probability estimate, the threshold  $J_{thresh,i}$  can be chosen as the mean value of all minimizing thresholds.

When the thresholds are determined, all probabilities needed for the selection of residual candidates, can be computed from the probability estimates, obtained in step 1A. This results in a set of probabilities like the ones in Table 3.2.

### Step 2

As described in Section 2.3, the goal is to minimize the probabilities  $P(\bar{B}_j)$  and  $P(\bar{A}_j)$ . To calculate the true probabilities, the joint multi-dimensional density function would be needed. Unfortunately this density function can generally

not be derived from the input data provided to the procedure. However, it is possible to calculate an upper bound of  $P(\bar{B}_j)$  and both lower and upper bounds of  $P(\bar{A}_j)$ . For the example of the right residual structure in Figure 2, the bound of  $P(\bar{B}_1) = P(B_1^C \cap B_2 \cap B_3)$  is

$$P(B_1^C \cap B_2 \cap B_3) \leq \min(P(B_1^C), P(B_2), P(B_3))$$

and the bounds of  $P(\bar{A}_1) = P(A_{11} \cup A_{21} \cup A_{31})$  are

$$\begin{aligned} \max(P(A_{11}), P(A_{21}), P(A_{31})) &\leq P(A_{11} \cup A_{21} \cup A_{31}) \\ &\leq 1 - P(A_{11}^C)P(A_{21}^C)P(A_{31}^C) \end{aligned}$$

The second bound of  $P(\bar{A}_1)$  follows from the assumption that for example

$$P(A_{31}^C | A_{11}^C \cap A_{21}^C) \geq P(A_{31}^C) \quad (1)$$

and

$$P(A_{21}^C | A_{11}^C) \geq P(A_{21}^C) \quad (2)$$

These relations says that it is more probable that a residual responds in correspondence with the residual structure if it is given that other residuals respond in correspondence with the residual structure. From the relation (2) and the definition of conditional probability, it follows that

$$P(A_{11}^C \cap A_{21}^C) \geq P(A_{11}^C)P(A_{21}^C)$$

which implies

$$P(A_{11} \cup A_{21}) \leq 1 - P(A_{11}^C)P(A_{21}^C)$$

If also the relation (1) is used, the second bound of  $P(\bar{A}_1)$  will follow.

The idea is to minimize the bounds of  $P(\bar{B}_j)$  and  $P(\bar{A}_j)$ , but with estimated probabilities instead of true probabilities. Minimizing the bounds actually gives the same result as minimizing the true probabilities in some special cases. The number of bounds for each residual structure is three times the number of faults.

The optimization goal is to, under some isolation constraints, find the residual structure for which the bounds are minimized. All bounds are generally not minimized by the same residual structure. This problem can be solved by replacing the minimization with a constraint that for a residual structure, all bounds must be “almost minimized”. Generally this does not give one residual structure but a set of residual structures. Formally this set is

$$\{\Phi | \text{Bound}_i(\Phi) - \min_{\phi} \text{Bound}_i(\phi) \leq \epsilon, i = 1 \dots 3n_f\}$$

where  $\Phi$  and  $\phi$  denotes residual structures,  $\text{Bound}_i(\Phi)$  the  $i$ :th bound for residual structure  $\Phi$ , and  $n_f$  the number of faults. The value of  $\epsilon$  determines how conservative the minimization is and the number of residual structures obtained is dependent on this value. For a value  $\epsilon = 0$ , true minimization is achieved but this may result in that no residual structures are obtained.

## Output

The output from the procedure is one, or possibly several, complete FDI systems. Each consists of a set of selected residuals, one threshold for each residual, and a residual structure including X's.

## 4.2 Remarks

The optimization problem is divided into two subproblems. When the first subproblem has been solved, then thresholds have been fixed. These thresholds are also the ones that minimize the bounds in step 2. In spite of this, it is not assured that these thresholds are globally optimal. The reason is that minimizing the bounds may not necessarily result in that  $P(\bar{B}_j)$  and  $P(\bar{A}_j)$  are minimized.

Another reason for non-optimality is the way X's are introduced. For instance only *ones* are allowed to be replaced with X's. Further, one may also obtain a non-optimal result if only strongly isolating structures are considered. This constraint may however be relaxed if sufficient computer power is available.

One potential problem with using the procedure is requirements of low probabilities of false alarm and missed detection. If the estimations of the probabilities  $A_{ij}$  and  $B_i$  are performed as depicted in step 1A, then in most realistic applications, the requirements are such that the probability estimates  $\hat{P}(B_i)$  and  $\hat{P}(A_{ij})$  would be zero for a broad range of thresholds. This is not a problem when choosing the thresholds in step 1C, but becomes a serious problem in step 2 because many probabilities from step 1C will be zero. This means that the probabilities provide no or little guidance in comparing different residual structures.

To overcome this problem there are at least four possible solutions. One is to use longer measurement sequences. However practical limitations can make this difficult. Another is to select fault sizes that are smaller than are actually required to be detected. Suppose the fault sizes are smaller by some factor, then the thresholds obtained from the procedure can be raised with the same factor. The third solution is to estimate the probabilities by using another technique, that works with realistic fault sizes. The final solution is to keep the small fault sizes used in the design, and take care of the false alarms / missed detections by adding some passive robustness in the residual evaluation. For example the time, for which the original FDI system signals alarm, can be summed up and the alarm can be suppressed until the time-sum reaches a threshold.

The whole procedure is automatic, i.e. when input data are provided, all steps can be performed without any human involvement. The procedure has been implemented as a Matlab command.

## 5 FDI of the Air-Intake System of an SI-Engine

In the field of automotive engines, environmentally based legislative regulations such as OBDII (CARB, 1993) and EOBD specify hard requirements on the performance of the FDI system. This makes the area a challenging application for model based FDI. In (Nyberg and Nielsen, 1997), model based FDI is applied to the FDI of the air-intake system of a spark-ignition engine. In this section it is shown how the previously described procedure is applied to the same problem.

The task is to diagnose three sensors: the throttle angle sensor  $\alpha$ , the air mass-flow sensor  $\dot{m}$ , and the manifold pressure sensor  $p$ . Also included among the inputs to the residual generators is the engine speed  $n$  and the throttle control signal  $u$ . The process model is non-linear and contains one state, which is the manifold pressure. For a more detailed description of the system, and the experimental setup, the reader is referred to (Nyberg and Nielsen, 1997).

### 5.1 Application of the Procedure

The system belongs to a class of non-linear systems for which there exists no general design method for residual generation. However by using primarily non-linear parity equations and non-linear observers in different configurations, 12 different residual generators were constructed. The computational form of these 12 base residuals are

$$\begin{aligned}
 r_1 &= m_s - \hat{m}_1(a_s, p_s) \\
 r_2 &= m_s - \hat{m}_2(n, p_s) \\
 r_3 &= p_s - \hat{p}_1(a_s, n, p_s) \\
 r_4 &= m_s - \hat{m}_3(a_s, n, \dot{m}_s) \\
 r_5 &= p_s - \hat{p}_2(a_s, \dot{m}_s, n) \\
 r_6 &= a_s - \hat{a}_1(u, a_s, \dot{m}_s, p_s) \\
 r_7 &= m_s - \hat{m}_4(a_s, n, p_s) \\
 r_8 &= r_2 - r_1 = \hat{m}_1(a_s, p_s) - \hat{m}_2(n, p_s) \\
 r_9 &= r_4 - r_2 = \hat{m}_2(n, p_s) - \hat{m}_3(a_s, n, \dot{m}_s) \\
 r_{10} &= r_4 - r_1 = \hat{m}_1(a_s, p_s) - \hat{m}_3(a_s, n, \dot{m}_s) \\
 r_{11} &= r_3 - r_5 = \hat{p}_2(a_s, \dot{m}_s, n) - \hat{p}_1(a_s, n, p_s) \\
 r_{12} &= a_s - \hat{a}_2(\dot{m}_s, p_s)
 \end{aligned}$$

where  $\hat{m}_i$ ,  $\hat{a}_i$ , and  $\hat{p}_i$  are different estimates of the output signals. The index  $s$  indicates that the signal is obtained from a sensor. For more details on these residuals, see (Nyberg and Nielsen, 1997). The residual structure of the 12 base residuals can be derived from the equations and is shown in Figure 6. Also the influence of a possible fault in the engine speed signal  $n$  is included.

	$f_\alpha$	$f_m$	$f_p$	$f_n$
$r_1$	1	1	1	0
$r_2$	0	1	1	1
$r_3$	1	0	1	1
$r_4$	1	1	0	1
$r_5$	1	1	1	1
$r_6$	1	1	1	0
$r_7$	1	1	1	1
$r_8$	1	0	1	1
$r_9$	1	1	1	1
$r_{10}$	1	1	1	1
$r_{11}$	1	1	1	1
$r_{12}$	1	1	1	0

Figure 6: The residual structure of the base residuals.

### Input

To generate input data to the procedure, data were collected during a one minute fault free test cycle, see (Nyberg and Nielsen, 1997). Constant bias faults were chosen. The fault sizes were 4% for the  $\alpha$ -fault, 5% for the  $m$ -fault, and 4% for the  $p$ -fault. These fault sizes were chosen as a result of some trial and error to get suitable sized probability estimates in step 1A. Thus the input to the procedure is 12 base residuals, one minute fault free measurements, the choice of constant bias fault characteristics, and fault sizes 4%, 5% and 4% respectively.

### Step 1A

Probability estimates are computed for thresholds in the range 0 to 10. The faults are introduced in the measurement data. For each fault, two probabilities estimates are used, one for negative faults and one for positive faults. The final estimate is then the mean. The probability estimates for all 12 base residuals are shown in Figure 7. It can be seen that residual 1 and residual 7 have very similar probability curves. This is also true for residual 3 and 8, and 5 and 11. This similarity is confirmed when the correlation coefficient is calculated:  $C(r_1, r_7) = 0.99$ ,  $C(r_3, r_8) = 0.94$ , and  $C(r_5, r_{11}) = 0.99$ . Note that even though these residuals may seem to be identical, they are not.

### Step 1B

To limit the computational load in step 2, residual 1 and residual 11, are because of similarity, not included in the set of residual candidates. As can be seen in Figure 7, the use of X's will for most residuals, increase the robustness. X's are introduced and residuals for which the lowest probability that they will

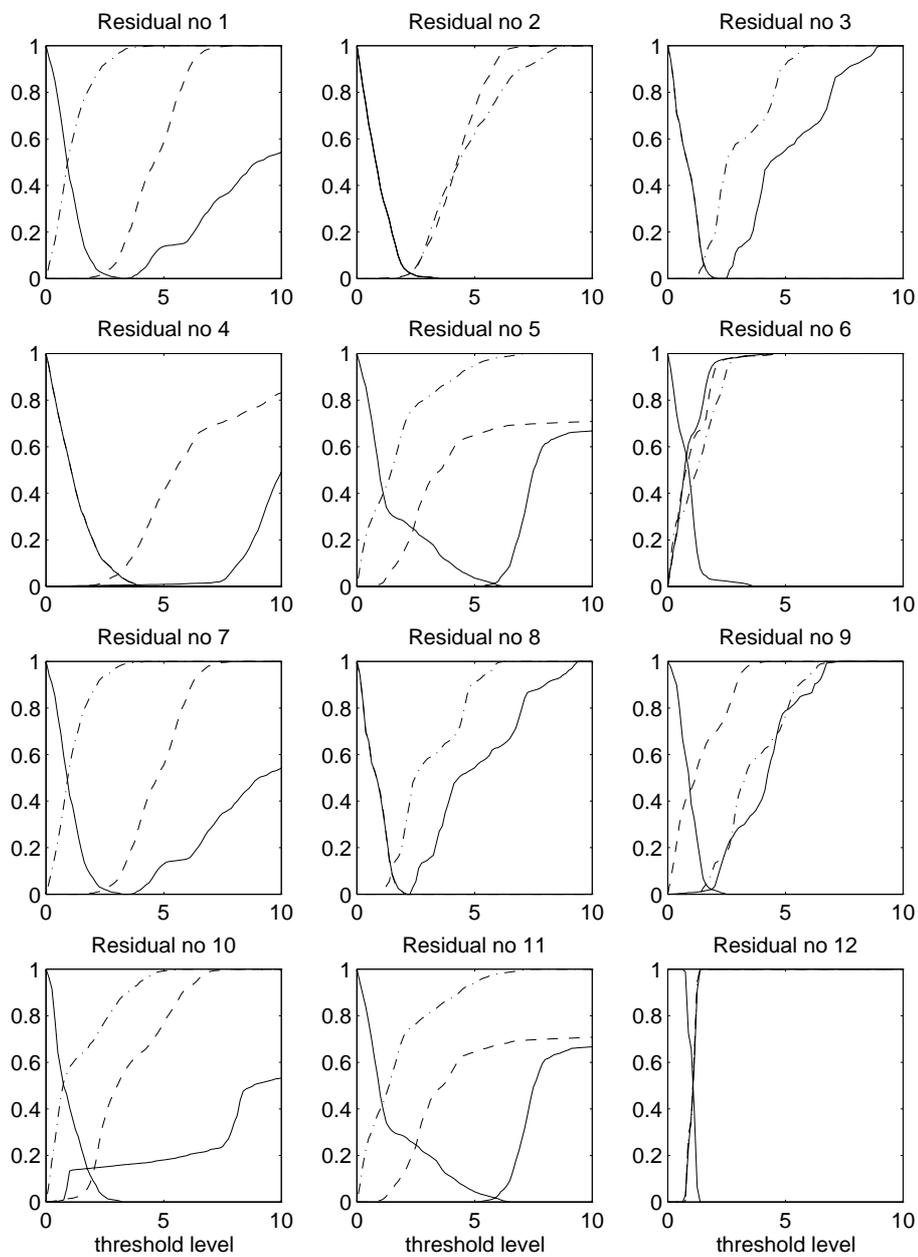


Figure 7: Probability estimates for each residual as a function of the threshold: no fault (solid),  $\alpha$ -fault (solid),  $m$ -fault (dashed), and  $p$ -fault (dash-dotted).

signal false alarm or miss a fault, is higher than 0.2, are omitted. After this, there are 15 residual candidates left.

$r$	no fault	$\alpha$ -fault	$\dot{m}$ -fault	$p$ -fault	$J_{thresh}$
2	.0240	0 .0241	1 .0208	1 .0251	2.250
2	.0240	0 .0241	1 .0208	X	2.250
3	.0732	1 0.0	0 .0724	1 .0548	1.500
3	0	1 0.0	0 0	X	2.250
4	.0584	1 .0058	1 .0484	0 .0593	3.000
4	.0053	1 .0072	X	0 .0051	3.875
5	.0102	1 .0144	X	X	5.875
7	.0198	1 0.0	1 .0171	X	2.500
7	0	1 0.0	X	X	3.375
8	.0854	1 0.0	0 .0841	1 .1182	1.500
8	0	1 0.0	0 0	X	2.250
9	.0405	1 .0160	X	1 .0414	1.625
9	.0204	1 .0206	X	X	1.875
10	.1428	1 .1438	1 .0568	X	1.750
10	.1153	X	1 .0816	X	1.875

Table 3: All 15 residual candidates, with probabilities of false alarm and missed detection. The rightmost column is the threshold.

### Step 1C

Thresholds are calculated for all 15 residual candidates. All 15 residual candidates are shown in Table 3. For each row, the first column contains the number of the base residual. The second column is the estimated probability of false alarm  $\hat{P}(B_i)$ . Columns three to five correspond to the different faults  $\alpha$ ,  $\dot{m}$ , and  $p$ , and contain the residual structure together with estimated probabilities of false alarm or missed detection  $\hat{P}(A_{ij})$ . The last column is the threshold. When a position in the residual structure contains an X, there is of course no probability associated with it.

### Step 2

There are  $2^{15}$  possible residual structures. Of these only 6144 are strongly isolating which is concluded with the help of Theorem 1. These numbers are sufficiently low for performing a complete search. The result is that for totally 7 residual structures, all 9 bounds are almost minimized, i.e.  $\epsilon = 0.005$ . These have 4, 4, 5, 4, 5, 5, and 6 residuals respectively. Four of these residual structures are shown in Figure 8.

	$f_\alpha$	$f_{\dot{m}}$	$f_p$
$r_2$	0	1	1
$r_3$	1	0	1
$r_4$	1	1	0
$r_8$	1	0	X

	$f_\alpha$	$f_{\dot{m}}$	$f_p$
$r_2$	0	1	1
$r_3$	1	0	1
$r_4$	1	1	0
$r_7$	1	X	X

	$f_\alpha$	$f_{\dot{m}}$	$f_p$
$r_2$	0	1	1
$r_3$	1	0	1
$r_4$	1	1	0
$r_7$	1	X	X
$r_8$	1	0	X

	$f_\alpha$	$f_{\dot{m}}$	$f_p$
$r_2$	0	1	1
$r_3$	1	0	1
$r_3$	1	0	X
$r_4$	1	1	0

Figure 8: Four of seven residual structures obtained from the procedure.

## 5.2 Confirmation of the Design

For the final selection in a real application, a residual structure with 4 residuals should be the best choice. This is to minimize the FDI system complexity and computational load.

To confirm the design, the upper left residual structure in Figure 8 is chosen. In Figure 9, each column contains plots for the case of no fault,  $\alpha$ -,  $\dot{m}$ -, and  $p$ -fault respectively. The upper four rows are plots for residuals  $r_2$ ,  $r_3$ ,  $r_4$ , and  $r_8$  respectively. The last row contains the fault decision. It is clear that the procedure successfully manage to generate an FDI system, that can detect and isolate these faults. However it can be seen that there are two false alarms. Both are  $p$ -fault false alarms. Also for the cases with faults, there are some missed detections. Thus the performance is not perfect but as said in Section 4.2, the reason is the small fault sizes used to make it possible to apply the procedure. However, compared to the study in (Nyberg and Nielsen, 1997), the faults are less than half the size and the performance is still fairly good.

## 6 Conclusions

Assuming residual generators are already available, there are still several choices to be made when a complete FDI system is to be designed. The key issue has been to find a systematic procedure to support this time-consuming engineering work. To the author's knowledge, this is the first attempt, even though there should be several other possibilities. For this purpose, a new probability based PI is proposed. This PI is applicable to measuring both residual performance and also complete FDI system performance. To increase the robustness of the FDI system, a *don't care* option X in the residual structure is introduced.

The procedure is phrased as an optimization problem consisting of selecting

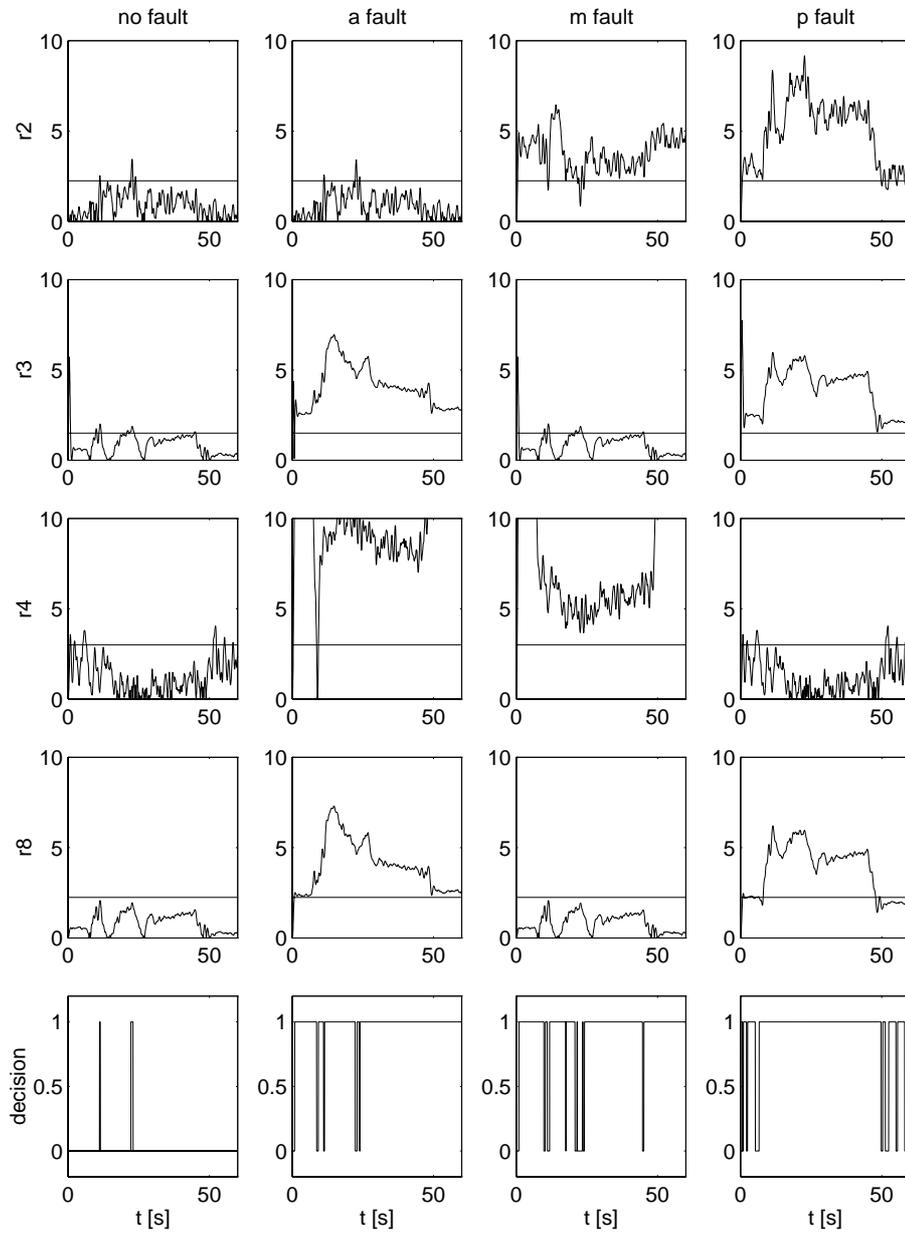


Figure 9: Confirmation of the FDI system for the cases no fault  $\alpha$ -,  $m$ -, and  $p$ -fault respectively.

the set of residuals to be included, introducing X's in the residual structure, and selecting thresholds. The procedure is successfully applied to the problem of FDI design for the air-intake system of an SI-engine. The resulting FDI system is able to detect and isolate faults of less than half the size, compared to the former system (Nyberg and Nielsen, 1997), for which considerably more time was spent on engineering and tuning.

## 7 Acknowledgments

This research is supported by NUTEK (Swedish National Board for Industrial and Technical Development).

## References

- CARB (1993). California's OBD-II regulation (section 1968.1, title 13, California code of regulations), resolution 93-40, July 9. pp. 220.7 – 220.12(h).
- 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.
- Frank, P.M. (1993). Advances in observer-based fault diagnosis. Proc. TOOLDIAG'93, pp. 817–836. CERT, Toulouse, France.
- Frisk, E., M. Nyberg, and L. Nielsen (1997). FDI with adaptive residual generation applied to a DC-servo. Fault Detection, Supervision and Safety for Technical Processes. IFAC, Hull, United Kingdom.
- Gertler, J. and M. Costin (1994). Model-based diagnosis of automotive engines. IFAC Fault Detection, Supervision and Safety for Technical Processes, pp. 393–402. Espoo, Finland.
- Nyberg, M. and L. Nielsen (1997). Model based diagnosis for the air intake system of the SI-engine. *SAE Paper* (970209).