Off-line Robust Residual Selection Using Sensitivity Analysis

Hamed Khorasgani[1] **Daniel Eriksson**[2] **Gautam Biswas**[1] **Erik Frisk**[2] and **Mattias Krysander**[2]
[1]Institute of Software Integrated Systems, Vanderbilt University, USA
e-mail: {hamed.g.khorasgani,gautam.biswas}@vanderbilt.edu
[2]Dept. of Electrical Engineering, Linkoping University, Sweden
e-mail: {daner,frisk,matkr}@isy.liu.se

Abstract

Model-based approaches to fault detection and isolation (FDI) rely on accurate models of the plant and a sufficient number of reliable measurements for residual generation and analysis. However, in realistic situations, there can be uncertainties in the plant models and measurements, which have a negative impact on the diagnosability performance that depends on the system state. In other words, the impact of the uncertainties can be larger in some operating regions as compared to others. To achieve acceptable performance in practice, it is necessary to find a set of residuals that are sufficiently sensitive to faults but robust to uncertainties across all operating conditions. In this paper, a quantitative measure, called detectability ratio, is used to evaluate and quantify detectability performance of different residuals in different operating regions. This measure is used to find a minimal residual set that fulfills a set of desired diagnosability performance requirements. The proposed method is demonstrated and validated through a case study.

1 Introduction

In model-based diagnosis, finding a suitable set of residuals which fulfils required detectability and isolability is an important part of the diagnosis system design process [Chen and Patton, 2012; Blanke and Schröder, 2003]. A common approach to design FDI systems utilizes the concept of structured residuals, see for example [Gertler, 1991; Cordier *et al.*, 2004], which utilizes a set of residuals, where each residual is sensitive to a subset of faults, and the set of residuals can detect and isolate the set of faults. Several algorithms have been proposed to find sets of residuals which fulfil a required detection and isolation performance (see for example [Armengol *et al.*, 2009; Svärd *et al.*, 2013; Bregon *et al.*, 2014]).

Depending on the system size and the number of potential sensors that may be used, the number of residual candidates can become very large and search for the optimal set is computationally infeasible. In [Svärd *et al.*, 2013], a greedy search strategy is used to select a minimal set of residuals that can detect and isolate a set of faults. The problem of finding sets of residuals for fault diagnosis is also covered in [Travé-Massuyes *et al.*, 2006; Nejjari *et al.*, 2010] as a component of the sensor placement problem. On-line residual selection strategies have been proposed to reduce the computational cost [Krysander *et al.*, 2010]. In [Eriksson *et al.*, 2012], an on-line sequential test selection strategy is proposed where the detectability performance of each residual is quantified using a measure called distinguishability [Eriksson *et al.*, 2013]. In [Nyberg, 1999], a loss function is used as a predefined cost for different situations that produce incorrect results such as false alarms and incorrect isolations hypotheses, then the expected value of the loss function is used as a measure to evaluate effectiveness of different diagnosis system candidates.

Even if faults are detectable and isolable in theory, parameter uncertainties can complicate the FDI problem producing false and missed alarms, and degrade the detection and isolation performance. Given a model of a dynamic system, and a set of measurements that can be made on the system, there are several residual candidates that can be derived using ARR approaches. In this paper, we study the problem of choosing a subset of residuals that are the most sensitive to faults and at the same time are robust to the uncertainties. Sensitivity of a residual with respect to a fault or an uncertainty represent the residual's response to the fault or uncertainty respectively. Our primary goal in this paper is to maximize detectability of faults, and then extend the analysis to establish isolability results. However, choosing the best set of residuals for fault diagnosis is not trivial. It is a challenge to evaluate the effects of uncertainties and faults on each residual. In [Djeziri *et al.*, 2007], the effects of uncertainties are compensated for by the use of adaptive thresholds. In [Eriksson and Sundström, 2014], residual performance is quantified using the Kullback-Leibler divergence. In relation to previous work, a contribution here is to develop an approach to residual selection that achieves sufficiently high detectability, thus reducing the chances of false alarms over different operating regions of system behavior. For fault isolation, it reduces the chances for incorrect diagnoses.

Moreover, as the system behavior evolves, the performance of each residual may vary over time and the detectability performance will depend on the operating regions of the system. A residual that is more effective in detecting a fault in one operating region, may be quite ineffective in another. Therefore, different residual sets may be necessary to fulfill the required performance over multiple operating regions of the system. However, since the set of residual candidates can be large, it is important to keep the search process constrained to avoid large computational expenses in the selection process.

In previous work [Khorasgani *et al.*, 2014] used sensitiv-

ity analysis (see for example [Ascher and Petzold, 1997]) to define a *detectability ratio* measure to quantify the performance of a residual. By comparing the sensitivity of the residual value to the fault versus the sensitivity of the residual value to uncertainty in model parameters, the detectability ratio quantifies the separation of the effect of the fault from uncertainty effects. In the previous work, the detectability ratio was used for on-line residual selection when the system switched from one operating region to the next. In this paper we use the detectability ratio for off-line residual selection. The objective is to find subsets of residuals that fulfil a specified detectability performance for the operating regions of the system. The selected residual sets can then be invoked on-line for the corresponding operating range. This reduces the overall computational complexities for on-line detection.

The outline of this paper is as follows. Section 2 formulates the problem and briefly reviews dynamic system models, residual selection and the fault detection problem. Section 3 presents a running example to describe the residual selection method, the approach for residual generation given a system model and measurements, and the residual candidates for the example problem. A quantitative measure to evaluate residual performance for fault detection is presented in Section 4. Two algorithms for choosing an optimal set of residuals for the different operating ranges of system behavior are presented in Section 5. Section 6 extends the work to discuss residual selection for fault isolation using the detectability ratio criterion. Section 7 discusses the conclusions of the paper.

2 Problem formulation

In this work, we study the problem of off-line residual selection for linear dynamic systems.

2.1 System representation

The general model of linear dynamic systems takes the form

$$\begin{aligned} \dot{x} &= A(\theta_n, \delta, f)x + B(\theta_n, \delta, f)v \\ y &= C(\theta_n, \delta, f)x + D(\theta_n, \delta, f)v \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^{l_x}$ represents state variables, $v \in \mathbb{R}^{l_v}$ actuator signals, $y \in \mathbb{R}^{l_y}$ sensor signals, $\theta_n \in \mathbb{R}^{l_{\theta_n}}$ nominal system parameters, $\delta \in \mathbb{R}^{l_\delta}$ parameter uncertainties, and $f \in \mathbb{R}^{l_f}$ faults. It is assumed that the true values of parameters of the system model may not be known, and the uncertainty in the parameters is characterized by a multiplicative term. Therefore, each parameter θ with nominal value θ_n is represented as: $\theta = \theta_n(1 + \delta_\theta)$, where δ_θ models the uncertainty in the parameter. It is also assumed that each parameter uncertainty δ_θ has a known upper bound Δ_θ:

$$|\delta_\theta| \leq \Delta_\theta \tag{2}$$

2.2 Fault detection

A residual r is an analytical redundancy relation (ARR) between parameters of the system, process measurements, and inputs. Once a persistent fault occurs the computed value from the ARR will most likely not match the measured value. To detect a fault f we need a residual sensitive to the fault and, at the same time, invariant or at least robust to uncertainties [Frank and Ding, 1994]. In this paper we model residuals for the linear dynamic system (1) as a non linear dynamic relationship between nominal parameters of the system, process measurements, and inputs in the following manner:

$$\begin{aligned} \dot{\hat{x}} &= f(\hat{x}, \theta_n, u, y) \\ r(\hat{x}, \theta_n, u, y) &= 0, \end{aligned} \tag{3}$$

where $\hat{x}$ represents the internal dynamics of the residual and u represents measured inputs. When the inputs are known or we have ideal sensors with no faults or uncertainties, u is equal to v.

2.3 Residual selection

The total number of possible residual candidates for fault detection grows exponentially as the number of measurements increases [Armengol *et al.*, 2009; Svärd *et al.*, 2013]. In (3) the total number of redundancies introduced into the system model is equal to the number of measurements l_y. Theoretically, each ARR can include from one to l_y measurements. Therefore, the total number of residuals N_{ARR} is proportional to the number of all the possible combinations of the measurements :

$$N_{ARR} \propto \sum_{i=1}^{l_y} \binom{l_y}{i} = 2^{l_y} \tag{4}$$

Therefore in general, there are many different residual subsets that we can use to detect the faults.

When there are uncertainties in the system, the sensitivity of these residuals to each uncertainty have to be computed to determine the residuals that provide the best performance. The sensitivity of the residuals to the faults are not the same. Therefore, different sets of residuals can have different performances for fault detection. Moreover, for a given behavior trajectory, different residuals can have different performance in different operating regions of the system. As a result, the best performance over an entire range of operations of the system can not be achieved by considering a single set of residuals. To achieve the best possible performance we need to find the best set of residuals in each operating region and switch to that set as the system behavior evolves and crosses from one region to another. To minimize the computational costs at the same time, we want each residual set to have minimal cardinality. As discussed earlier, solving the general problem of finding optimal residual sets for a set of faults over different operating regions is a computationally intractable problem. Therefore, we simplify the problem by finding the best subset of residuals to detect each possible fault and then we consider the minimal size sets formed by the union of the individual sets as the final answer. This simplification still gives us the best performance but the computational cost is not globally optimum.

In this paper we define detectability ratio of fault f_i for a residual r, $D(f_i|r)$, as a measure that captures the performance of residual r in detecting fault f_i. Using this measure we can choose the best residual to detect f_i within each operating region. To achieve the maximum performance with minimum on-line required computation we have developed an off-line residual selection algorithm that can choose a subset of residuals to detect f_i that guarantee maximum performance for a given trajectory. The problem is presented as follows.

Let $\mathcal{R} = \{r_1, r_2, \ldots, r_n\}$ denote a set of residual candidates. We develop an algorithm to find a subset of residuals

R_i which guarantees a maximum quantitative detectability ratio for each fault f_i over the time interval T, and has minimum number of residuals.

$$\begin{aligned} \min \quad & R_i \subseteq \mathcal{R} \\ \text{s.t.} \quad & \forall t \in T, \exists r \in R_i : \\ & D(f_i|r(t)) = \max(D(f_i|r_j(t))|\forall r_j \in \mathcal{R}), \\ & R = \bigcup_{i=1}^{l_f} R_i \end{aligned} \tag{5}$$

where t represents time and R is the final subset of residuals for fault detection.

3 Running example

As the running example for this paper, consider a simple DC motor with a mechanical load shown in Figure 1.

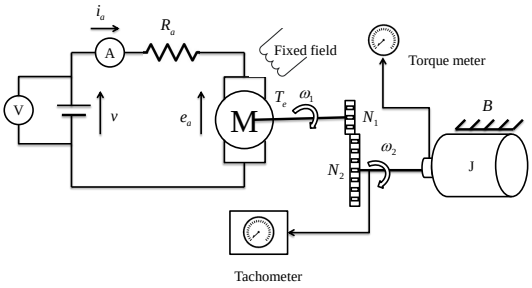

Figure 1: DC motor diagram.

The rotational speed of the mechanical load is considered as the only state variable in the system. The corresponding state equation is:

$$\begin{aligned} \dot{\omega}_2 = -(\frac{B(1+\delta_B)}{J(1+\delta_J)} + \frac{K_a(1+\delta_{K_a})K_f(1+\delta_{K_f})}{J(1+\delta_J)R_a(1+\delta_{R_a})})\omega_2 \\ + \frac{N_2}{N_1}(1+f_g)\frac{K_a(1+\delta_{K_a})}{J(1+\delta_J)R_a(1+\delta_{R_a})}v, \end{aligned} \tag{6}$$

where v is the DC voltage applied to the motor, R_a is the nominal value of armature resistance. The parameter K_a is the unit rate of the proportional relationship between motor torque T_e and current i_a. The parameter K_f represents the rate of change of back electromotive force (EMF) voltage e with respect to electric motor rotational speed ω_1. The gear helps in increasing the torque and reducing the speed of mechanical load. The mechanical torque is given by $T_m = \frac{N_2}{N_1}T_e$ and mechanical rotational speed by $\omega_2 = \frac{N_1}{N_2}\omega_1$. A fault in the gear is denoted f_g. Mechanical load has nominal inertia J and nominal friction coefficient B.

The system has four known variables, the input u, and three measurements, y_1, y_2, and y_3 with associated sensors. The equations for the measured input and the three output measurements are given by:

$$\begin{aligned} u =& (1+\delta_u)v, \\ y_1 =& (1+\delta_{y_1})i_a \\ =& (1+\delta_{y_1})\left(-\frac{K_f(1+\delta_{K_f})}{R_a(1+\delta_{R_a})}\frac{N_1}{N_2(1+f_g)}\omega_2 \right. \\ & \left. + \frac{1}{R_a(1+\delta_{R_a})}v\right), \\ y_2 =& (1+\delta_{y_2})T_m \\ =& (1+\delta_{y_2})\left(-\frac{K_a(1+\delta_{K_a})K_f(1+\delta_{K_f})}{R_a(1+\delta_{R_a})}\omega_2 \right. \\ & \left. + \frac{N_2}{N_1}(1+f_g)\frac{K_a(1+\delta_{K_a})}{R_a(1+\delta_{R_a})}v\right), \\ y_3 =& (1+\delta_{y_3})\omega_2, \end{aligned} \tag{7}$$

where u represents the output of the first sensor: a voltmeter which measures the battery voltage, v. The signal y_1 is the ammeter output that measures the motor current. y_2 represents the output associated with the torque transducer and, i.e., the mechanical torque T_m. y_3 is the output measurement recorded by the tachometer, ω_2. The uncertainty in each parameter θ is denoted δ_θ. The block diagram of the system, derived from the state and output equations is shown in Figure 2.

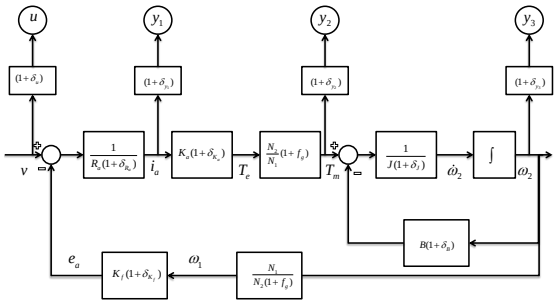

Figure 2: Block diagram model of DC motor.

Several methods have been proposed for designing residuals [Frank, 1990]. Given u and y_3 the following residual can be designed to detect f_g by comparing y_3 and its predicted value $\hat{\omega}_2$ as:

$$\begin{aligned} \dot{\hat{\omega}}_2 = -(\frac{B}{J} + \frac{K_aK_f}{JR_a})\hat{\omega}_2 + \frac{N_2}{N_1}\frac{K_a}{JR_a}u \\ r_1 = y_3 - \hat{\omega}_2. \end{aligned} \tag{8}$$

Another residual that can be used to detect f_g is computed as the difference between a measurement y_2 and its predicted value $\hat{T}_m$.

$$\begin{aligned} \dot{\hat{\omega}}_2 &= -(\frac{B}{J} + \frac{K_aK_f}{JR_a})\hat{\omega}_2 + \frac{N_2}{N_1}\frac{K_a}{JR_a}u \\ \hat{T}_m &= -\frac{K_aK_f}{R_a}\hat{\omega}_2 + \frac{N_2}{N_1}\frac{K_a}{R_a}u \\ r_2 &= y_2 - \hat{T}_m. \end{aligned} \tag{9}$$

Usually an ARR is considered as a relationship between inputs and outputs, however it can also represent analytical redundancy between two or more outputs. r_3 is derived us-

ing y_1 to predict y_3.

$$\begin{aligned}\dot{\hat{\omega}}_2 &= -\frac{B}{J}\hat{\omega}_2 + \frac{K_a}{J}\frac{N_2}{N_1}y_1 \\ r_3 &= y_3 - \hat{\omega}_2.\end{aligned} \tag{10}$$

Considering y_1 and y_2 we have:

$$r_4 = y_2 - K_a\frac{N_2}{N_1}y_1. \tag{11}$$

So far the residuals were a function of two known variables. The following residuals represent an analytical relationship between three known variables i.e., using u and y_3 to predict y_2:

$$r_5 = y_1 - \frac{1}{R_a}(u - K_f\frac{N_1}{N_2}y_3). \tag{12}$$

The next residual uses u and y_2 to predict y_1.

$$\begin{aligned}\dot{\hat{\omega}}_2 &= -\frac{B}{J}\hat{\omega}_2 + y_2 \\ r_6 &= y_1 - \frac{1}{R_a}(u - K_f\frac{N_1}{N_2}\hat{\omega}_2).\end{aligned} \tag{13}$$

Similarly, the following residual uses u and y_3 to predict y_2.

$$r_7 = y_2 - \frac{1}{R_a}K_a\frac{N_2}{N_1}u + K_f\frac{1}{R_a}K_a y_3. \tag{14}$$

We can also estimate the input by $\hat{v} = R_a y_1 + K_f\frac{N_1}{N_2}y_3$ and substitute it in (9). This results in a new residual, which is a function of y_1, y_2 and y_3.

$$\begin{aligned}\hat{v} &= R_a y_1 + K_f\frac{N_1}{N_2}y_3 \\ \dot{\hat{\omega}}_2 &= -(\frac{B}{J} + \frac{K_a K_f}{JR_a})\hat{\omega}_2 + \frac{N_2}{N_1}\frac{K_a}{JR_a}\hat{v} \\ \hat{T}_m &= -\frac{K_a K_f}{R_a}\hat{\omega}_2 + \frac{N_2}{N_1}\frac{K_a}{R_a}\hat{v} \\ r_8 &= y_2 - \hat{T}_m.\end{aligned} \tag{15}$$

The number of possible residuals increase as the number of analytical redundancies in the system increase. It should also be mentioned here that in general any linear or even non-linear combination of residuals can be used to generate a new residual. In this paper, we propose a methodology to choose an optimal subset of residuals for a given system trajectory.

4 Quantifying the performance of residuals

To select the best set of residuals for fault detection we need to quantify their performance in the face of model uncertainties. As discussed, we use sensitivity analysis as the basis evaluating the effect of model uncertainties on the residuals. In this section we present the sensitivity analysis methodology, and then use this method to define a quantitative measure for fault detection performance of a given residual.

4.1 Sensitivity analysis methodology

Sensitivity analysis helps evaluate how model behaviors are affected by variations in model parameters [Serban and Hindmarsh, 2005]. We start by computing the derivative of each system variable with respect to each parameter using the chain rule. Consider a general linear dynamic system model (1). The chain rule is applied to the state and output equations to obtain [Ascher and Petzold, 1997]:

$$\begin{aligned}\dot{p}_\psi &= Ap_\psi + \frac{\partial A}{\partial \psi}x + \frac{\partial B}{\partial \psi}v \\ q_\psi &= Cp_\psi + \frac{\partial C}{\partial \psi}x + \frac{\partial D}{\partial \psi}v,\end{aligned} \tag{16}$$

where $p_\psi = \frac{\partial x}{\partial \psi}$ and $q_\psi = \frac{\partial y}{\partial \psi}$ represent the sensitivity of state variables and measurements to the parameter ψ, which can be a fault, f_i or the uncertainty associated with a parameter, δ_j. It is assumed that v is an external control signal which does not depend on ψ, implying $\frac{\partial v}{\partial \psi} = 0$.

Consider the DC motor state dynamic equations (6). We can present the dynamics associated with $p_{\delta_B} = \frac{\partial \omega_2}{\partial \delta_B}$ for nominal behavior as:

$$\dot{p}_{\delta_B} = -(\frac{B}{J} + \frac{K_a K_f}{JR_a})p_{\delta_B} - \frac{B}{J}\omega_2. \tag{17}$$

Using equation (7) and the chain rule we can derive the sensitivity of known variables of the system to δ_B as a function of p_{δ_B}:

$$\begin{aligned}\frac{\partial u}{\partial \delta_B} &= 0 \\ \frac{\partial y_1}{\partial \delta_B} &= -\frac{K_f N_1}{R_a N_2}p_{\delta_B} \\ \frac{\partial y_2}{\partial \delta_B} &= -\frac{K_a K_f}{R_a}p_{\delta_B} \\ \frac{\partial y_3}{\partial \delta_B} &= p_{\delta_B}.\end{aligned} \tag{18}$$

Sensitivity of known variables to other uncertainties δ_{R_a}, δ_{K_a}, δ_{K_f}, δ_J, δ_u, δ_{y_1}, δ_{y_2},δ_{y_3} and fault f_g are derived in a similar manner but due to the lack of space we do not list all of them in this paper.

4.2 Sensitivity analysis of residuals

Consider the residual model given by equation (3). We can use the sensitivity of known variables to uncertainties and faults derived in the previous section, to compute the sensitivity of the residual to faults and uncertainties, represented by ψ, as:

$$\begin{aligned}\dot{\hat{p}}_\psi &= \frac{\partial f}{\partial \hat{x}}\hat{p}_\psi + \frac{\partial f}{\partial u}\frac{\partial u}{\partial \psi} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \psi} \\ \frac{\partial r}{\partial \psi} &= \frac{\partial r}{\partial \hat{x}}\hat{p}_\psi + \frac{\partial r}{\partial u}\frac{\partial u}{\partial \psi} + \frac{\partial r}{\partial y}\frac{\partial y}{\partial \psi},\end{aligned} \tag{19}$$

where $\hat{p}_\psi = \frac{\partial \hat{x}}{\partial \psi}$. As an example, the sensitivity of the first residual r_1 to δ_B is derived as:

$$\begin{aligned}\dot{\hat{p}}_{\delta_B} &= -(\frac{B}{J} + \frac{K_a K_f}{JR_a})\hat{p}_{\delta_B} \\ \frac{\partial r_1}{\partial \delta_B} &= \frac{\partial y_3}{\partial \delta_B} - \hat{p}_{\delta_B}.\end{aligned} \tag{20}$$

where $\hat{p}_{\delta_B} = \frac{\partial \hat{\omega}_2}{\partial \delta_B}$.

4.3 Detectability ratio for fault detection

We use sensitivity analysis to compute the effects of uncertainties and faults in the residuals. It is assumed that the fault magnitude and uncertainties are small and approximately constant over time. A first order linear approximation of

a residual with respect to the faults f and uncertainties δ is given by:

$$r(y,u) \approx \sum_{i=1}^{l_f} \frac{\partial r}{\partial f_i} f_i + \sum_{j=1}^{l_\delta} \frac{\partial r}{\partial \delta_j} \delta_j. \quad (21)$$

The partial derivatives $\frac{\partial r}{\partial f_i}$ and $\frac{\partial r}{\partial \delta_j}$ are computed using (19). Note that if a residual r is not sensitive to a fault or uncertainty then the corresponding partial derivative is zero.

When quantifying the detectability performance of a residual r with respect to a fault f_i the relative effect of the fault is compared to the total effect of the uncertainties δ. Since the actual magnitudes of the fault and uncertainties are unknown, the maximum values of the magnitude of uncertainties and minimum magnitudes of a fault f_i are used for the calculations. This gives us the worst case scenario of the difficulty in detecting a fault. We use a quantitative measure of detectability performance defined in previous work [Khorasgani *et al.*, 2014] as follows.

Definition 1. *(Detectability Ratio) Given a dynamic system (1) the detectability ratio of a fault f_i for a residual r (3) is defined as:*

$$D(f_i|r) = \frac{\left|\frac{\partial r}{\partial f_i} F_{min}\right|}{\left|\frac{\partial r}{\partial f_i} F_{min}\right| + \sum_{j=1}^{l_\delta} \left|\frac{\partial r}{\partial \delta_j} \Delta_j\right|}, \quad (22)$$

where Δ_j is the upper bound of uncertainty δ_j and F_{min} is an absolute minimum fault magnitude of f_i to detect. If $\left|\frac{\partial r}{\partial f_i} F_{min}\right| = 0$ then $D(f_i|r) = 0$.

The detectability ratio has a value in the interval $[0, 1]$, where 0 corresponds to the situation where the residual is not sensitive to the fault f_i and 1 if there are no uncertainties affecting the residual's ability to detect the fault. If the effect of a fault is larger than the total effect of all the uncertainties then $D(f|r) > 0.5$, which means that the fault is detectable.

Table 1 lists parameters for the running example. Considering the parameters, $v = 10V$, 1% uncertainty in each parameter and the sensors, and $F_{min} = 25\%$, the detectability ratio of $r_1, r_2, r_3, r_4, r_5, r_6, r_7$, and r_8 are plotted in Figure 3. We can see in this figure that r_4 has the highest detectability ratio and is the best choice to detect fault f_g. This is not surprising because r_4 is a function of only one parameter and two known variables and when all the elements have the same percentage of uncertainty, the residuals with fewer number of parameters and known variables provide the better performance. However, in the general case each element can have different level of uncertainty and we need to perform sensitivity analysis to find the residual that is most sensitive to the fault and has the greatest robustness to the uncertainties. The other interesting observation is that all of the eight residuals have detectability ratio higher than 0.5, so they can be used to detect a fault with magnitude, $F_{min} = 25\%$. But if we wish to detect smaller faults we need to use residuals that have a higher detectability ratio.

In this simple example we only considered the system operation in steady state, but in general a dynamic system can have different trajectories which lead to different detectability ratio for each residual over different intervals of time. In the next section, we will discuss the problem of residual selection when we have more complicated trajectories.

Table 1: Running Example Parameters

Parameter	Value and unit
R_a	1Ω
K_a	0.5 N.m/Amp
K_f	0.68 Volt/(rad/sec)
$\frac{N_1}{N_2}$	0.5
B	0.6 N.m.s/rad
J	0.4 kg.m^2

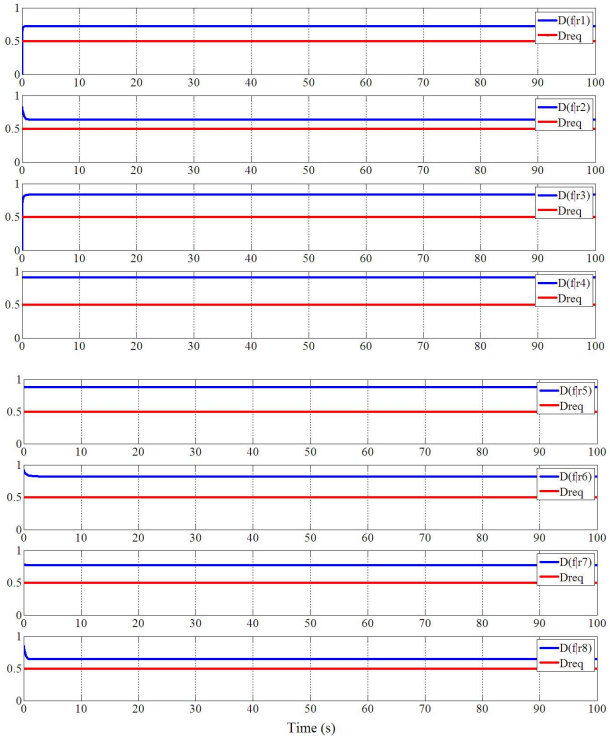

Figure 3: Detectabilty ratio of fault f_g for different residuals in the running example.

5 Algorithm for residual selection

In this section we present two methods for residual selection. The first one is the solution to the problem we presented in (5). The second approach adds a constraint to the problem but instead of the best performance it only satisfies a minimum required performance.

5.1 Best performance with minimum number of residuals

To address problem (5) we have to choose a subset of residuals provide the best performance in different working regions and switch between them accordingly as the system behavior transitions from one region to the next. In other words, we compute the detectability ratio of all the residual candidates off-line and choose a subset of residuals that contains the residuals with the best detectability ratio in each working range. When operating on-line we only need to compute the detectability ratio of this subset of residuals and switch between them. This method is straight forward and guarantees the best performance in on-line fault detection. However, there is no limitation on the number of required switchings between the residuals in on-line scenarios

to achieve this performance.

5.2 Required performance with minimum number of switching between residuals

Switching between residuals can be computationally expensive and in some applications where the system has a fast dynamic or the performance of residuals are similar ensuring minimum amount of switching between the residuals is a reasonable criterion to impose on the choice of residuals. In this section we propose an algorithm to choose a set of residuals that guarantee the required detectability ratio with minimum number of residual switching for a given trajectory.

Note that to achieve global minimum switching between the residuals, we have to consider all the possible faults and all the residuals sensitive to them simultaneously as one optimization problem. However, to simplify the problem, in this paper we consider the problem for each individual fault separately. Therefore, there is no guarantee that the final answer gives us the global minimum switching for the system but if the number of possible faults are small in the system the answer is close to the global optimal.

Let $\mathcal{R}$ denote a set of residual candidates for fault detection. The purpose is to find a subset of residuals R_i which fulfils a minimum required detectability ratio D^{req} for detecting fault f_i, and requires the minimum number of switching between the residuals for a given time interval T which in general represents a set of different operating ranges. For example, consider the case that we have a residual that gives us acceptable performance in the current system working range and then the system behavior changes to another region. We prefer to make no change if the current residual still provides sufficient performance. If the performance falls below a threshold, then we prefer to change to a new residual such that its performance will stay above the threshold for the longest possible period of time . This should minimize subsequent residual switching in the future. The problem is presented as follows.

$$\begin{aligned} \min \quad & (|r_j \rightarrow r_k| \,|r_j, r_k \in R_i) \\ \text{s.t.} \quad & \forall t \in T \exists r \in R_i : D(f_i|r(t)) \geq D^{\text{req}} \\ & R_i \subseteq \mathcal{R} \\ & R = \bigcup_{i=1}^{l_f} R_i, \end{aligned} \tag{23}$$

where $r_j \rightarrow r_k$ represents switching from residual r_j to r_k and we want to minimize the total number of required switching, denoted $|r_j \rightarrow r_k|$, between the residuals $r \in R_i$ while keeping $D(f_i|r)$ above D^{req} for the given trajectory. We can easily change this problem to the problem of finding shortest path in graphs [Cherkassky *et al.*, 1996] and use one of the well developed polynomial algorithms for the shortest path problem to solve it. The process is summarized as follows.

1- Consider a source and a terminal node for the graph.

2- For each region of operation, consider a layer in the graph and for each residual that has detectability ratio above the threshold in that region of operation consider a node in that layer.

3- Connect the source node to the nodes of layer one with weight zero.

4- Connect each node in each layer to all of the nodes in the next layer. If start and end node of a connection represent the same residual assign weight zero to that connection, otherwise assign weight one to the connection.

5- Connect the nodes in the last layer to the terminal node with weight zero.

We elaborate on this process using an example in the following section.

5.3 Example for residual selection

To increase DC motor rotational speed we can reduce the field flux, this is equivalent with decreasing K_f in (6). To show how different trajectories can change the detectability ratio of the residuals in this simulation, we decrease the flux field at a constant rate from $40s$ to $50s$ to 10% of its original value. All the parameters have the same values reported in Table 1. The maximum uncertainty in each sensor and parameter considered is 1%, $F_{min} = 25\%$ and $D^{\text{req}} = 0.8$. Figure 4 shown the detectability ratios and the required detectability ratio for each residual.

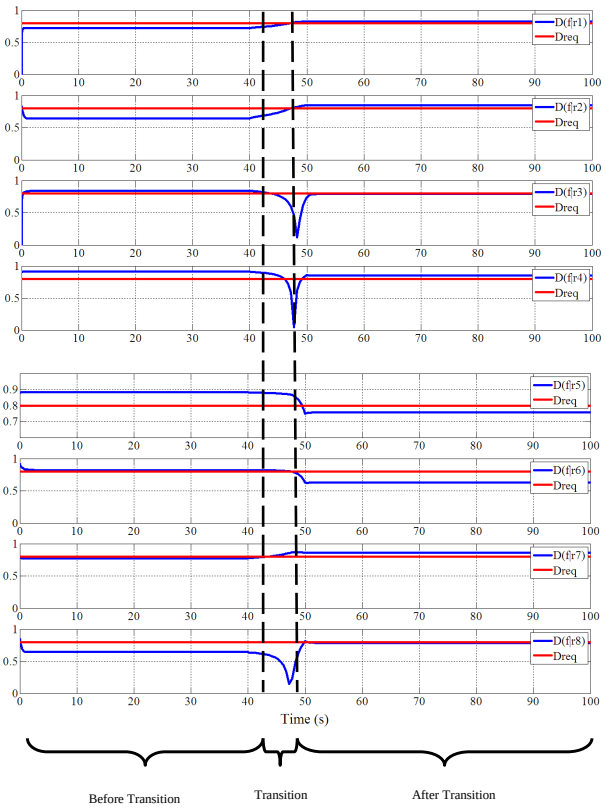

Figure 4: Detectability ratio and minimum required detectability ratio of fault f_g for different residuals in the running example.

Residual selection for best performance with minimum number of residuals

For the given trajectory the system has three main working ranges: before transition, transition, and after transition. To choose the best performance with minimum number of residuals we consider these three time intervals as the three

operating regions. Before the transition r_4 has the highest detectability ratio, during the transition r_5 has the highest detectability ratio and after transitions r_7 is the best choice. Therefore, the subset (r_4, r_5, r_7) is enough to achieve the best possible performance. The total number of switching with these residuals for the given trajectory is two, the first one from r_4 to r_5 and the second one from r_5 to r_7.

Residual selection for required performance with minimum number of switching between residuals

To minimize the number of switching between residuals we consider the second problem. Before transitioning to a new region r_3, r_4, r_5 and r_6 have detectability ratios higher than the minimum requirement so we need one of them in our final set. In the second region, r_5 and r_7 are our choices and finally for the last interval we need to choose one from r_2, r_4, r_7. As we mentioned earlier, we can look at the problem as finding the shortest path in a graph. As is shown in Figure 5, we consider a source and a terminal node and then for each residual switching we consider an edge of weight 1. If we find the shortest path from the initial region of operation to the final region of operation we have found the residual set with minimum required switching.

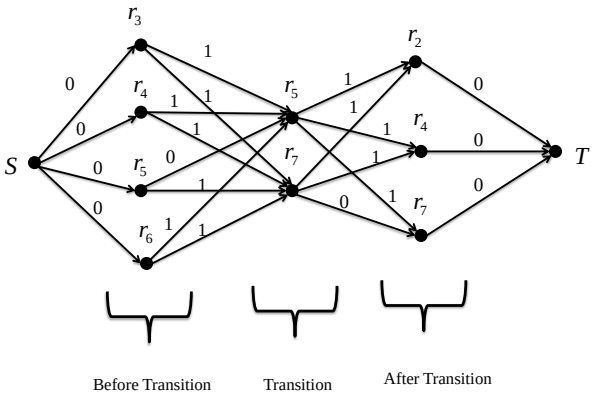

Figure 5: Residual selection as a shortest path problem.

In this example, the shortest path gives us one of the following residual pairs:

$$(r_3, r_7), (r_4, r_7), (r_5, r_7), (r_6, r_7), (r_5, r_2), (r_5, r_4). \quad (24)$$

Note that the total weight of each path is equal to 1 which means only one residual switching is required for the given trajectory. In this example, we only have one possible fault, therefore, the answer is the global optimal answer.

6 Isolability problem

In diagnosis systems, fault detection is typically followed by fault isolation. To isolate a fault f_i from the other possible faults in the system, we need a residual that is sensitive to fault f_i but invariant or robust to others. However, the effects of faults and uncertainties in the residuals are unknown, and we have to, like before, estimate them using sensitivity analysis.

Making the single fault assumption, it is assumed that only one of the possible faults occurs, and we want to quantify the performance of each residual to isolate that specific fault from the others. The magnitudes of the possible faults and uncertainties are unknown, therefore, to quantify the performance of residual r to isolate fault f_i from another fault f_j, the minimum magnitude of f_i and the maximum magnitude of f_j and uncertainties are considered. That is, the other fault f_j is considered as an uncertainty. Then, isolability ratio is defined, using equation (21), as a quantitative measure of isolation performance as follows.

Definition 2. *(Isolability Ratio) Given the dynamic system (1) the isolability ratio of a fault f_i from another fault f_j, given residual r is defined as:*

$$I(f_i|r, f_j) = \frac{\left|\frac{\partial r}{\partial f_i} F_{min}\right|}{\left|\frac{\partial r}{\partial f_i} F_{min}\right| + \left|\frac{\partial r}{\partial f_j} F_j\right| + \sum_{k=1}^{l_\delta} \left|\frac{\partial r}{\partial \delta_k} \Delta_k\right|}, \quad (25)$$

where Δ_k and F_j are the upper bound of the uncertainty δ_k and the fault f_j respectively, and F_{min} represents an absolute minimum fault magnitude of f_i.

If $I(f_i|r, f_j) > 0.5$ means that the effect of a fault f_i on r is always larger than the total effects of the uncertainties and the fault f_j and we can use r to isolate f_i from f_j. Note that if r is sensitive to f_j and we do not know the maximum magnitude of f_j, i.e. "$F_j = \infty$" , then r can not be used to isolate f_i from f_j and $I(f_i|r, f_j) = 0$.

To demonstrate the performance of the isolability ratio consider the running example with possible fault f_u in the volt-meter.

$$u = (1 + f_u)(1 + \delta_u)v \quad (26)$$

Figure 6 represents the isolability ratio of gear fault f_g from volt meter fault f_u when the minimum magnitude of f_g is 25%, maximum magnitude of f_u is 100%, $v = 10$ and system parameters are represented in Table 1. Residuals r_3, r_4

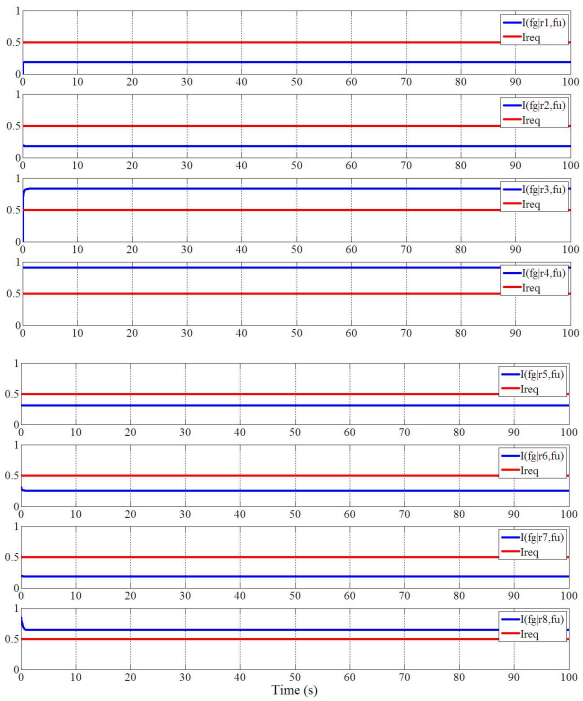

Figure 6: Isolability ratio of fault f_g from fault f_u.

and r_8 have isolability ratio higher than 0.5 and can be used to isolate fault f_g from fault f_u. Among them r_4 has the

highest isolability ratio and is the most reliable residual to isolate f_g from f_u in the presence of uncertainties.

Note that isolability ratio is simply a generalization of detectability ratio. In fact, the detectability ratio of fault f_g is the isolability ratio of f_g from the no fault case. Therefore, we can extend the algorithms developed for residual selection, to address the problem of residual selection with required isolability performance of fault f_g from fault f_u by considering $I(f_g|r, f_u)$ as the quantitative measure.

7 Conclusion

There are several methods for fault detection and isolation. These methods suggest automated algorithms to determine a set of minimal residuals to maximize diagnosability of the system. However it is often not clear how the performance of each method is affected in the presence of uncertainties in the system. In this paper, we used the detectability ratio to select a set of residuals that guarantee a required performance for a given trajectory. Two approaches, one based on maximizing performance, and a second based on minimizing switching between residuals, but still maintaining detectability performance above a pre-specified threshold are presented. We also extended the measure for fault detectability to a fault isolability measure to determine the ability to isolate one fault from another. The algorithms developed in this paper are for off-line analysis and selection. Therefore, the computational problem of the sensitivity analysis was not considered to be a central issue.

The detectability and isolability ratios give useful information of how diagnosability performance varies in dynamic systems. This can be used to design diagnosis systems which adapts to different operating conditions as discussed in this work but also other applications such as optimizing input signals for active fault diagnosis. Using sensitivity analysis, makes the method capable to address additive and multiplicative faults and uncertainties, however, it reduces the accuracy of the method dealing with high frequency signals. In future work, we will investigate other possible methods to take noise and time-varying faults and uncertainties into consideration in the diagnosability analysis.

References

[Armengol *et al.*, 2009] Joaquim Armengol, Anibal Bregón, Teresa Escobet, E Gelso, Mattias Krysander, Mattias Nyberg, Xavier Olive, Belarmino Pulido, and Louise Travé-Massuyès. Minimal structurally overdetermined sets for residual generation: A comparison of alternative approaches. In *Proceedings of the 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, SAFEPROCESS09*, pages 1480–1485, 2009.

[Ascher and Petzold, 1997] Uri M Ascher and Linda R Petzold. Computer methods for ordinary differential equations and differential-algebraic equations. *Society for Industrial and Applied Mathematics*, 1997.

[Blanke and Schröder, 2003] Mogens Blanke and Jochen Schröder. *Diagnosis and fault-tolerant control*, volume 115. Springer, 2003.

[Bregon *et al.*, 2014] Anibal Bregon, Matthew Daigle, Indranil Roychoudhury, Gautam Biswas, Xenofon Koutsoukos, and Belarmino Pulido. An event-based distributed diagnosis framework using structural model decomposition. *Artificial Intelligence*, 210:1–35, 2014.

[Chen and Patton, 2012] Jie Chen and Ron J Patton. *Robust model-based fault diagnosis for dynamic systems*. Springer Publishing Company, Incorporated, 2012.

[Cherkassky *et al.*, 1996] Boris V Cherkassky, Andrew V Goldberg, and Tomasz Radzik. Shortest paths algorithms: Theory and experimental evaluation. *Mathematical programming*, 73(2):129–174, 1996.

[Cordier *et al.*, 2004] M-O Cordier, Philippe Dague, François Lévy, Jacky Montmain, Marcel Staroswiecki, and Louise Travé-Massuyès. Conflicts versus analytical redundancy relations: a comparative analysis of the model based diagnosis approach from the artificial intelligence and automatic control perspectives. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 34(5):2163–2177, 2004.

[Djeziri *et al.*, 2007] Mohand Arab Djeziri, Rochdi Merzouki, B Ould Bouamama, and Genevieve Dauphin-Tanguy. Robust fault diagnosis by using bond graph approach. *Mechatronics, IEEE/ASME Transactions on*, 12(6):599–611, 2007.

[Eriksson and Sundström, 2014] Daniel Eriksson and Christofer Sundström. Sequential residual generator selection for fault detection. 2014.

[Eriksson *et al.*, 2012] Daniel Eriksson, Erik Frisk, and Mattias Krysander. A sequential test selection algorithm for fault isolation. In *Proceedings of the 10th European Workshop on Advanced Control and Diagnosis, ACD 2012, Copenhagen, Denmark*. Citeseer, 2012.

[Eriksson *et al.*, 2013] Daniel Eriksson, Erik Frisk, and Mattias Krysander. A method for quantitative fault diagnosability analysis of stochastic linear descriptor models. *Automatica*, 49(6):1591–1600, 2013.

[Frank and Ding, 1994] Paul M Frank and Xianchun Ding. Frequency domain approach to optimally robust residual generation and evaluation for model-based fault diagnosis. *Automatica*, 30(5):789–804, 1994.

[Frank, 1990] Paul M Frank. Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: A survey and some new results. *Automatica*, 26(3):459–474, 1990.

[Gertler, 1991] Janos Gertler. Analytical redundancy methods in fault detection and isolation. In *Proceedings of IFAC/IAMCS symposium on safe process*, volume 1, pages 9–21, 1991.

[Khorasgani *et al.*, 2014] H Khorasgani, D Eriksson, G Biswas, Frisk E, and Krysander M. Robust residual selection for fault detection. In *53rd IEEE Conference on Decision and Control*, 2014.

[Krysander *et al.*, 2010] Mattias Krysander, Fredrik Heintz, Jacob Roll, and Erik Frisk. Flexdx: A reconfigurable diagnosis framework. *Engineering applications of artificial intelligence*, 23(8):1303–1313, 2010.

[Nejjari *et al.*, 2010] Fatiha Nejjari, Ramon Sarrate, and Albert Rosich. Optimal sensor placement for fuel cell system diagnosis using bilp formulation. In *Control & Automation (MED), 2010 18th Mediterranean Conference on*, pages 1296–1301. IEEE, 2010.

[Nyberg, 1999] Mattias Nyberg. Automatic design of diagnosis systems with application to an automotive engine. *Control Engineering Practice*, 7(8), 1999.

[Serban and Hindmarsh, 2005] Radu Serban and Alan C Hindmarsh. CVODES, the sensitivity-enabled ODE solver in SUNDIALS. In *Proceedings of the 5th International Conference on Multibody Systems, Nonlinear Dynamics and Control, Long Beach, CA*, 2005.

[Svärd *et al.*, 2013] Carl Svärd, Mattias Nyberg, Erik Frisk, and Mattias Krysander. Automotive engine fdi by application of an automated model-based and data-driven design methodology. *Control Engineering Practice*, 21(4):455–472, 2013.

[Travé-Massuyes *et al.*, 2006] Louise Travé-Massuyes, Teresa Escobet, and Xavier Olive. Diagnosability analysis based on component-supported analytical redundancy relations. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 36(6):1146–1160, 2006.