15th Triennial World Congress, Barcelona, Spain

FDI DECISION MAKING USING A GAME FORMULATION

Boris Darkhovski * and **Marcel Staroswiecki** **,[1]

* *ISA RAS, 117312 Moscow, Russia*
** *LAIL-CNRS-UPRESA 8021, USTL, EUDIL, F-59655 Villeneuve d'Ascq cedex, France*

Abstract: This paper proposes a game theoretic formulation of the fault detection problem, in a Bayesian setting, in the presence of nuisance parameters. Solutions are characterized and an illustrative example is provided in the linear frame.

Keywords: Fault Detection, Minmax, Game Theory

1. INTRODUCTION

Real-time fault detection (FD) is a decision problem in which the healthy or faulty state of a system has to be inferred from the observation of the available data. Two kinds of decision problems, namely hypotheses testing or change point detection problems can be stated. Fault isolation extends the decision problem to the consideration of more than two hypotheses. This paper addresses the FD problem, by means of the two hypotheses testing problem.

The main difficulty in FD is that the observed data depends on some nuisances : unknown parameters, inputs, initial conditions. In control (the so-called geometric approach), the system model is used to eliminate those parameters (Cocquempot *et al.*, 1991), (Seliger and Frank, 1991). This produces a *residual* vector, which is created using identification, parity space or observer based approaches, which are related in some sense (Gertler, 2000), (Cocquempot and Christophe, 2000), (Magni and Mouyon, 1994). The decision procedure applies to this residual. However, there are many cases in which this approach does not work, either because all the nuisances cannot be eliminated (heuristic approximate decoupling is then used, (Gertler and Kunwer, 1993), (Staroswiecki *et al.*, 1993)), or because nuisances and faults act in the same spaces (see (Staroswiecki and Darkhovski, 2001) for example).

Statistical decision approaches directly consider the (stochastic) available observation, by setting some optimisation problem in which false alarms and missed detections are minimized (Basseville and Nikiforov, 1993), (Brodsky and Darkhovsky, 2000). Relations between control and statistical decision approaches have been exhibited (Basseville and Nikiforov, 1993), (Nikiforov *et al.*, 1996). Again, when nuisances are present, this problem receives only heuristic solutions, whose *asymptotic* optimality (i.e., when the number of observations tends to infinity) has been proven only under restrictive assumptions on the faults (like the assumption about the bringing hypotheses - see (Borovkov, 1984)).

This paper proposes an approach to the FD problem which is not purely geometric and not purely statistical. The decision problem is based on a game theoretic formulation, which does not need the decoupling of nuisances, provides non-asymptotical results, and can be interpreted in terms of some rational behaviour of the decision maker. Section 2 presents the system normal and faulty operation models, and gives a Bayesian

[1] The corresponding author:
Email: marcel.staroswiecki@univ-lille1.fr,
Tel: (33) 3 20 33 7190, Fax: (33) 3 20 33 7189.

setting of a unified decision problem, which is developed in a game theoretic frame in section 3. Section 4 gives three results which characterize the solutions of the game, and Section 5 provides an illustrative application for linear systems. Concluding remarks are given in Section 6.

2. NORMAL AND FAULTY OPERATION

Consider a dynamic system in discrete time:

$$\begin{aligned} x_{k+1} &= F(x_k, u_k, v_k, \varphi_k),\ x_0 \in X_0 \\ y_k &= G(x_k, u_k, v_k, \varphi_k, \xi_k) \end{aligned} \tag{1}$$

Here $x_k \in \mathbb{R}^n$ is the state vector, $X_0 \subseteq \mathbb{R}^n$ is a given set, and $y_k \in \mathbb{R}^p$ is the output vector. The measurements are corrupted by some stochastic vector $\xi_k \in \mathbb{R}^r$. The system inputs are the control vector $u_k \in \mathbb{R}^m$ – *a known function of time*, and two unknown inputs, namely $v_k \in \Theta \subseteq \mathbb{R}^q$ which represents external disturbances or model uncertainties which are not to be detected (Θ is a given set), and $\varphi_k \in \mathbb{R}^r$ which represents the fault vector. The system operation is considered on a given time window of finite lenghth $\{0, 1, ...N\}$, and the problem is to decide whether the system is in normal or in faulty operation on that time window.

Using the notation

$$\begin{aligned} \bar{\varphi} &= (\varphi_0^\tau, \varphi_1^\tau, \ldots, \varphi_N^\tau)^\tau, \\ \bar{\xi} &= (\xi_0^\tau, \xi_1^\tau, \ldots, \xi_N^\tau)^\tau \end{aligned}$$

where the superscript τ stands for transposition, and supposing that joint density function (d.f.) $f(\cdot)$ of vector $\bar{\xi}$ is known, normal and faulty operation can be defined as follows.

Problem 1

$$\begin{aligned} H_0 &: \text{normal operation, } \bar{\varphi} \in \Phi_0,\ \bar{\xi} \sim f_0(\cdot) \\ H_1 &: \text{faulty operation, } \bar{\varphi} \in \Phi_1,\ \bar{\xi} \sim f_0(\cdot) \end{aligned} \tag{2}$$

where Φ_0 and Φ_1 are two given sets, $\Phi_0 \cap \Phi_1 = \emptyset$ for consistency, and f_0 is a given d.f.

Problem 2

$$\begin{aligned} H_0 &: \text{normal operation, } \bar{\varphi} \in \Phi_0,\ \bar{\xi} \sim f_0(\cdot) \\ H_1 &: \text{faulty operation, } \bar{\varphi} \in \Phi_0,\ \bar{\xi} \sim f_1(\cdot) \end{aligned} \tag{3}$$

2.1 *Decision problem 1*

Define

$$\begin{aligned} z &= (x_0^\tau, \bar{v}^\tau, \bar{\varphi}^\tau)^\tau \\ Z_0 &= X_0 \times V \times \Phi_0 \\ Z_1 &= X_0 \times V \times \Phi_1 \end{aligned}$$

The (off-line) fault detection problem consists of testing if the fault vector $\bar{\varphi}$ belongs to Φ_0 or to Φ_1. For any given control sequence $\{u_0, u_1, ...u_{N-1}\}$, the observation $\bar{y} = (y_0^\tau, y_1^\tau, ...y_N^\tau)^\tau$ depends on the initial state $x_0 \in X_0$, on the unknown vectors $\bar{v} = \left(v_0^\tau, v_1^\tau, ...v_{N-1}^\tau\right)^\tau$ (which belongs to the known set $V = \Theta^N$) and $\bar{\varphi}$ (which belongs either to Φ_0 or to Φ_1 according to the system operation mode) and on the stochastic vector $\bar{\xi}$. Let g be the d.f. of the observation vector, which is generated by (1) and the d.f. f_0. It obviously depends on the parameter z. Therefore, for a given f_0, there is a *collection* of density functions $g(\cdot, z)$. Let

$$\mathcal{G}_0 = \{g(., z),\ z \in Z_0\},\ \mathcal{G}_1 = \{g(., z),\ z \in Z_1\}$$

then the fault detection problem can be set as a problem of statistical hypotheses testing:

$$\begin{aligned} &\mathbf{H}_0 : \text{the d.f. of the vector } \bar{y} \text{ belongs to } \mathcal{G}_0 \\ &\mathbf{H}_1 : \text{the d.f. of the vector } \bar{y} \text{ belongs to } \mathcal{G}_1 \end{aligned} \tag{4}$$

This problem is called the *two parametric sets* problem (Z_0 and Z_1). A *necessary and sufficient* condition for the faults (2) to be detectable is:

$$\mathcal{G}_0 \cap \mathcal{G}_1 = \emptyset \tag{5}$$

2.2 *Decision problem 2*

In decision problem 2, φ_k plays the same role as v_k, so considering only the latter in the parameter vector $z = (x_0^\tau, \bar{v}^\tau)^\tau$ is enough. Under the fault (3), the d.f. of the observation vector changes for any z from $g_0(., z)$ to $g_1(., z)$ (due to the change in the d.f. of $\bar{\xi}$ from f_0 to f_1), and the hypotheses testing problem is still described by (4), where $\mathcal{G}_0$ and $\mathcal{G}_1$ are now

$$\begin{aligned} \mathcal{G}_0 &= \{g_0(., z),\ z \in X_0 \times V\} \\ \mathcal{G}_1 &= \{g_1(., z),\ z \in X_0 \times V\} \end{aligned} \tag{6}$$

This problem is called the *one parametric set* problem ($Z = X_0 \times V$). The *necessary and sufficient* condition for such faults to be detectable is still given by (5).

2.3 *Unified decision problem*

The two above decision problems can be stated in a unified frame. Indeed, denote by $\zeta \in \mathcal{Z} = Z_0 \times Z_1$ a new unknown parameter and consider the following density functions

$$\gamma_0(\bar{y}, \zeta) = g(\bar{y}, \mathrm{Pr}_{Z_0}\zeta), \quad \gamma_1(\bar{y}, \zeta) = g(\bar{y}, \mathrm{Pr}_{Z_1}\zeta),$$

where Pr_A is the projection operator on the set A. Obviously, one has $\gamma_0(\bar{y}, \zeta) = g(\bar{y}, z_0)$ for some $z_0 \in Z_0$ and $\gamma_1(\bar{y}, \zeta) = g(\bar{y}, z_1)$ for some $z_1 \in Z_1$. It follows that for the functions γ_0, γ_1 and the parameter $\zeta \in \mathcal{Z}$ the problem is reduced to a one parametric set one.

Now, let $\bar{y} \in \mathbb{R}^{\mathcal{N}}$, where $\mathcal{N} = (N+1)p$, be the data, and $z \in Z \subset \mathbb{R}^{\mathcal{M}}$, where $\mathcal{M} = n \times Nq$, be the unknown parameter vector, Z is a given set. The problem is to test the two hypotheses :

$$\begin{array}{l} \mathbf{H}_0 : \text{d.f. of } \bar{y} \text{ is } g_0(\bar{y}, z) \\ \mathbf{H}_1 : \text{d.f. of } \bar{y} \text{ is } g_1(\bar{y}, z) \end{array}, \quad z \in Z$$

Let $\rho(\bar{y})$ be a decision function, i.e., a measurable map from $\mathbb{R}^{\mathcal{N}}$ to $\{0,1\}$. If $\rho(\bar{y}) = 0$, $\mathbf{H}_0$ is accepted, otherwise $\mathbf{H}_1$ is accepted. Let $\mathbf{P}_{z,0}$, $\mathbf{P}_{z,1}$ be the probabilistic measures associated with the hypotheses $\mathbf{H}_0$, $\mathbf{H}_1$ and the parameter value z respectively.

Hypotheses testing problems can be set using either the Bayesian or the Neyman-Pearson (NP) approach. In this paper, only the Bayesian approach is presented (the NP approach can be developed using similar lines), i.e. the criterion is:

$$J\big(\rho(\cdot), z\big) = \alpha \mathbf{P}_{z,1}\big(\rho(\bar{y}) = 0\big) + \beta \mathbf{P}_{z,0}\big(\rho(\bar{y}) = 1\big), \tag{7}$$

where $\alpha > 0$, $\beta > 0$ are given real numbers, for *any* $z \in Z$.

Some remarks and comments are in order here. First, the problem with one parametric set is not usually considered in the literature. Second, when strict decoupling can be achieved, geometric approaches can be seen to solve problem (4) under the simple situation where

$$Z_0 = \Phi_0, \quad Z_1 = \Phi_1$$

As already noticed, strict decoupling is not possible in many cases. Approximate decoupling introduces heuristic approaches, based on geometric manipulations, in order to "maximise the influence" of $\bar{\varphi}$ while "minimizing the influence" of x_0 and of $\bar{v}$ on the observations $\bar{y}$. However, such approaches allow for different means of defining the "influences", which cannot be easily related with the optimality criterion (7).

On another hand, problem (4) is well known in mathematical statistics (see (Borovkov, 1984) for example). Considering the Bayesian statement (7), the following decision rule is used in practice

$$\begin{array}{l} \text{if} \quad \dfrac{\max\limits_{z \in \mathcal{Z}_0} g(\bar{y}, z)}{\max\limits_{z \in \mathcal{Z}_1} g(\bar{y}, z)} > \eta, \text{ then accept} \quad \mathbf{H}_0, \\ \text{in the opposite case accept} \quad \mathbf{H}_1 \end{array} \tag{8}$$

where η is some constant. However, it is known ((Borovkov, 1984), p.349-350) that this rule is *not the optimal one in the general case*, and that it is only possible to prove *asymptotic optimality* (i.e., when $N \to \infty$) under special assumptions on the faults (for example, the assumption about the bringing hypotheses).

3. THE GAME PROBLEM SETTING

In this section, a game theoretic based approach to the solution of the one parametric set decision problem, with criterion (7), is proposed. This formulation provides a rational scheme from the decider's point of view, which needs neither any decoupling property, nor any statistical consideration.

For *any given* value of the parameter, the well known optimal rule for problem (7) is:

$$\begin{array}{l} \bar{y} \in \mathbf{D}(z) = \big\{\bar{y} : \dfrac{\alpha g_1(\bar{y}, z)}{\beta g_0(\bar{y}, z)} > 1\big\} \Longrightarrow \rho(\bar{y}) = 1 \\ \bar{y} \in \mathbb{R}^{\mathcal{N}} \backslash \mathbf{D}(z) \Longrightarrow \rho(\bar{y}) = 0 \end{array} \tag{9}$$

Denote by $\rho_z(\bar{y}) \triangleq \rho_z(\cdot)$ the optimal solution when the parameter is equal to z. Obviously, this rule cannot be implemented, since it depends on the unknown parameter. In the proposed approach, the decision is made *based on the guess* that the parameter is equal to $\hat{z} \in Z$ while its true value is equal to $z \in Z$.

Since the unknown parameter is guessed to be $\hat{z}$, it is natural to make the decision $\rho_{\hat{z}}(\cdot)$. Thus, the decision will be $\rho_{\hat{z}}(\bar{y}) = 1$ on the set $\mathbf{D}(\hat{z})$ and $\rho_{\hat{z}}(\bar{y}) = 0$ on the supplementary set. The cost functional under such decision will be

$$J\big(\rho_{\hat{z}}(\cdot), z\big) = \beta + \int_{\mathbb{R}^{\mathcal{N}} \backslash \mathbf{D}(\hat{z})} \Big(\alpha g_1\big(\bar{y}, z\big) - \beta g_0\big(\bar{y}, z\big)\Big) d\bar{y}$$

The minimum cost (which one could get if the true value of the parameter would be known) is

$$J\big(\rho_z(\cdot), z\big) = \beta + \int_{\mathbb{R}^{\mathcal{N}} \backslash \mathbf{D}(z)} \Big(\alpha g_1\big(\bar{y}, z\big) - \beta g_0\big(\bar{y}, z\big)\Big) d\bar{y}$$

Now the following game between the decision maker and the Nature is considered :

- the decision maker chooses the vector $\hat{z} \in Z$
- the Nature chooses the vector $z \in Z$
- the loss function of the game is

$$L_B(\hat{z}, z) \triangleq J\big(\rho_{\hat{z}}(\cdot), z\big) - J\big(\rho_z(\cdot), z\big) \tag{10}$$

Therefore, the initial hypotheses testing problem with unknown parameter can be reduced to some *finite-dimensional* problem, whose setting depends on the assumption which is made about Nature's behaviour. First, suppose that Nature plays randomly, according to some probabilistic measure, with d.f. $p(.)$, on the parametric set. Such measure can be considered as some *a priori* available knowledge about the parameter. Then it is natural to consider the following problem :

$$\lambda(\hat{z}) \triangleq \int_Z L_B(\hat{z}, z) p(z) dz \longrightarrow \min_{\hat{z} \in Z} \tag{11}$$

When no *a priori* knowledge is available, or when guaranteed results are needed, the worst case situation is considered for Nature's decision. Then, the following problem arises:

$$\lambda(\hat{z}) \triangleq \max_{z\in Z} L_B(\hat{z}, z) \longrightarrow \min_{\hat{z}\in Z} \qquad (12)$$

Let $z^* \in Z$ be the solution of problem (11) or (12). Then the final solution of the initial problem is to accept hypothesis $\mathbf{H}_1$ on the set $\mathbf{D}(z^*)$.

It can be underlined that the direct setting of the min-max approach to the initial problem, i.e., considering a problem of the type

$$\max_{z\in Z} J\big(\rho(\cdot), z\big) \longrightarrow \inf_{\rho(\cdot)}$$

where the inf operation is carried out over all measurable maps from $\bar{y}$ into $\{0, 1\}$, is very difficult. Note also that instead of minimizing the initial cost functional, one might consider minimizing its upper estimate, i.e., the problem

$$\tilde{J}\big(\rho(\cdot)\big) = \int \rho(\bar{y}) \max_{z\in Z} \big(\alpha g_1(\bar{y}, z) - g_0(\bar{y}, z)\big) d\bar{y} \longrightarrow \inf_{\rho(\cdot)} \qquad (13)$$

whose optimal solution would be

$$\tilde{\rho}(\bar{y}) = \begin{cases} 0 \text{ if } \max_{z\in Z} \big(\alpha g_1(\bar{y}, z) - \beta g_0(\bar{y}, z)\big) > 0 \\ 1 \text{ in opposite case,} \end{cases}$$

However, this solution seems unsuitable because in general it is impossible to construct any simple region for the acceptation of the hypothese in the $\bar{y}$-space.

Remark 1. Note that under the proposed approach, the detectability conditions are not used. This is because the decision is not a *statistical* one in the usual meaning. The detectability conditions guarantee that under any rational statistical decision, the cost functional tends to zero as the size of the sample tends to infinity (due to the law of large numbers). In the proposed approach, one is not interested in the asymptotic features of the solution, but one would only like to find a *rational solution* for finite sample sizes. The above mentioned solution is just such one.

4. THREE RESULTS FOR THE GAME PROBLEM

Let $Z \subset \mathbb{R}^{\mathcal{M}}$ be the given parametric set, and $\mathcal{L}(\hat{z}, z)$ be a given function, for $\hat{z} \in Z$, $z \in Z$. Everywhere below (to avoid trivial complexity) it is assumed that Z is compact.

Depending on the assumption which is made about Nature's behaviour, two problem settings are considered, namely

$$\lambda(\hat{z}) \stackrel{\triangle}{=} \max_{z\in Z} \mathcal{L}(\hat{z}, z) \longrightarrow \min_{\hat{z}\in Z} \qquad (14)$$

when Nature is supposed to play in the most aggressive way, and

$$\lambda(\hat{z}) \stackrel{\triangle}{=} \int_Z \mathcal{L}(\hat{z}, z) p(z) dz \longrightarrow \min_{\hat{z}\in Z} \qquad (15)$$

when Nature is supposed to play randomly by choosing the unknown parameter vector z under some known probabilistic measure $p(\cdot)$.

Theorem 4.1. (The minmax problem). Let the following assumptions hold:

1) the set Z has a center of symmetry,

2) the function $\mathcal{L}$ depends only of the difference $\hat{z} - z$ and it is symmetric with respect to zero,

3) $\mathcal{L}$ is a continuous and quasi-convex function.

Then the center of symmetry of the set Z is a solution of problem (14).

Proof. Without any loss of generality, assume that $0 \in Z$ and 0 is the center of symmetry. It has to be proven that $0 \in \arg\min \lambda(\hat{z})$.

Recall that a function f is called quasi-convex iff its Lebesgue set

$$\{x : f(x) \le b\}$$

is convex for any $b \in \mathbb{R}$. It is easy to see that if f is a continuous quasi-convex function then for any x there exists at least one halfspace $\mathfrak{R}(x)$ such that

$$f(x + z) \ge f(x) \ \forall z \in \mathfrak{R}(x).$$

From the central symmetry of $Z\,(-Z = Z)$ it follows :

$$\begin{aligned} \lambda(\hat{z}) &= \max_{z\in Z} \mathcal{L}(\hat{z} - z) = \max_{z\in Z} \mathcal{L}(\hat{z} + z) \\ &= \max_{z\in Z} \big(\mathcal{L}(z) + \mathcal{L}(\hat{z} + z) - \mathcal{L}(z)\big) \\ &\ge \mathcal{L}(z^*) + \mathcal{L}(z^* + \hat{z}) - \mathcal{L}(z^*), \end{aligned}$$

where $z^* \in \arg\max_{z\in Z} \mathcal{L}(z)$.

If $\hat{z} \in \mathfrak{R}(z^*)$, then from the above inequality and the quasi-convexity of $\mathcal{L}$ it follows that $\lambda(\hat{z}) \ge \mathcal{L}(z^*)$.

Suppose that $\hat{z} \not\in \mathfrak{R}(z^*)$. Therefore, $\mathcal{L}(z^* + \hat{z}) < \mathcal{L}(z^*)$. Using the symmetry of $\mathcal{L}$ one has

$$\begin{aligned} \mathcal{L}(-z^* - \hat{z}) &= \mathcal{L}\big(-(-z^* - \hat{z})\big) \\ &< \mathcal{L}(z^*) = \mathcal{L}(-z^*) \end{aligned}$$

and therefore $(-\hat{z}) \not\in \mathfrak{R}(-z^*)$, i.e., $\hat{z} \in \mathfrak{R}(-z^*)$.

But due to the symmetry, one has $(-z^*) \in \arg\max_{z\in Z} \mathcal{L}(z)$ from which it follows that

$$\lambda(\hat{z}) \ge \mathcal{L}(-z^* + \hat{z}) - \mathcal{L}(-z^*) + \mathcal{L}(-z^*) \ge \mathcal{L}(-z^*) = \mathcal{L}(z^*)$$

Therefore, for any $\hat{z} \in Z$, $\lambda(\hat{z}) \geq \mathcal{L}(z^*) = \max_{z \in Z} \mathcal{L}(z)$ holds. But on another hand

$$\begin{aligned}\min_{\hat{z} \in Z} \lambda(\hat{z}) &\leq \lambda(0) \\ &= \max_{z \in Z} \mathcal{L}(-z) = \mathcal{L}(z^*)\end{aligned}$$

Hence

$$\min_{\hat{z} \in Z} \max_{z \in Z} \mathcal{L}(\hat{z} - z) = \max_{z \in Z} \mathcal{L}(z)$$

and so $\hat{z} = 0$ is an optimal guess. □

Theorem 4.2. (The probabilistic problem). Let the following assumptions hold:

1) the set Z has a center of symmetry,

2) the function $\mathcal{L}$ depends only of the difference $\hat{z} - z$ and it is symmetric with respect to zero,

3) the function $\mathcal{L}$ is continuous and convex,

4) the probability measure $p(\cdot)$ is symmetric with respect to the center of symmetry.

Then

$$\min_{\hat{z} \in Z} \lambda(\hat{z}) = \lambda(0) = \int_Z \mathcal{L}(z) p(z) dz$$

Proof. Without any loss of generality, assume that 0 is the center of symmetry of Z. Due to the symmetry and convexity 0 belongs to the set of minima $\mathcal{L}$. Therefore it is possible to assume that $\mathcal{L} \geq 0$. From assumption 2, one has

$$\lambda(\hat{z}) = \int_Z \mathcal{L}(\hat{z} - z) p(z) dz$$

Fix vector $\hat{z} \in Z$ and denote

$$\begin{aligned}R(\hat{z}) &= \{z \in Z : \mathcal{L}(\hat{z} + z) \geq \mathcal{L}'\hat{z})\}, \\ R(-\hat{z}) &= \{z \in Z : \mathcal{L}(-\hat{z} + z) \geq \mathcal{L}(-\hat{z})\}\end{aligned}$$

Obviously, due to the symmetry of Z it follows that $z \in R(\hat{z}) \Leftrightarrow (-z) \in R(-\hat{z})$ and $Z = R(\hat{z}) \bigcup R(-\hat{z})$. Therefore one has

$$\lambda(\hat{z}) = \int_{R(\hat{z})} \big(\mathcal{L}(z + \hat{z}) + \mathcal{L}(z - \hat{z})\big) p(z) dz \quad (16)$$

As the function $\mathcal{L}$ is convex and continuous, it has at any point the subdifferential $\partial\mathcal{L}$. Therefore, for any $z \in Z, h \in Z$ there exists a vector $u(z) \in \partial\mathcal{L}(z)$ such that

$$\mathcal{L}(z + h) \geq \mathcal{L}(z) + \langle u(z), h \rangle$$

In particular, if $\mathcal{L}(\cdot)$ is a differentiable function, $u(z) = \frac{d\mathcal{L}}{dz}(z)$. Therefore,

$$\begin{aligned}\mathcal{L}(z + \hat{z}) &\geq \mathcal{L}(z) + \langle u(z), \hat{z} \rangle, \\ \mathcal{L}(z - \hat{z}) &\geq \mathcal{L}(z) + \langle u(z), -\hat{z} \rangle.\end{aligned} \quad (17)$$

From (16) and (17) and due to the assumptions one has

$$\lambda(\hat{z}) \geq 2 \int_{R(\hat{z})} \mathcal{L}(z) p(z) dz = \int_Z \mathcal{L}(z) p(z) dz = \lambda(0)$$

which is the result. □

Another result concerning the probabilistic problem is as follows.

Theorem 4.3. Let the following assumptions hold:

1) the set Z has a center of symmetry,

2) the function $\mathcal{L}$ depends only of the difference $\hat{z} - z$, symmetric with respect to zero, and nonnegative,

3) the probability measure $p(\cdot)$ is symmetric with respect to the center of symmetry, convex and twice differentiable.

Then

$$\min_{\hat{z} \in Z} \lambda(\hat{z}) = \lambda(0) = \int_Z \mathcal{L}(z) p(z) dz$$

Proof. Without loss of generality assume that 0 is the center of symmetry of Z. We can assume that the function $p(\cdot)$ is defined on the whole space but has the support Z. Then, taking the integral over the whole space, we have

$$\lambda(\hat{z}) = \int \mathcal{L}(u) p(u - \hat{z}) du$$

Calculating the derivative of λ at 0, we obtain

$$\lambda'(0) = \int \mathcal{L}(u) p'(u) du = 0$$

due to the symmetry of $\mathcal{L}$ and p. Further,

$$\lambda''(0) = \int \mathcal{L}(u) p''(u) du$$

and this matrix is positive semi-definite due to nonnegativity of $\mathcal{L}$ and convexity of p. Therefore, the point $\hat{z} = 0$ satisfies a weak local minimum condition. The function λ being convex (by assumptions 1 and 3), therefore $0 \in \arg\min \lambda(\hat{z})$. □

Note that if the matrix $\int \mathcal{L}(u) p''(u) du$ is definite positive, the point 0 is at least a local minimum of $\lambda(\cdot)$ without the convexity assumption of $p(\cdot)$.

5. AN ILLUSTRATIVE APPLICATION

Consider the LTI version of system (1) :

$$\begin{aligned}x_{k+1} &= A x_k + B u_k + E v_k \\ \zeta_k &= C x_k + \varepsilon_k\end{aligned} \quad (18)$$

where $x_0 \in X_0 \subset \mathbb{R}^n$, $v_k \in \Theta \subset \mathbb{R}^q$, and the considered fault is the change of the d.f. of ε

from $g_0(.)$ in normal operation to $g_1(.)$ in faulty operation. From (18), it follows that

$$\bar{\zeta} = OBSx_0 + COM_B\bar{u} + COM_E\bar{v} + \bar{\varepsilon} \quad (19)$$

where OBS, COM_B, COM_E are respectively the observability matrix, and the influence matrices associated with the known and unknown inputs. (19) obviously writes under the form

$$\bar{y} = Hz + \epsilon,$$

where $\bar{y} = \bar{\zeta} - COM_B\bar{u} \in \mathbb{R}^{\mathcal{N}}$ is the observation vector, $\epsilon = \bar{\varepsilon}$ is the random vector, $z = \begin{pmatrix} x_0 \\ \bar{v} \end{pmatrix} \in Z \subset \mathbb{R}^{\mathcal{M}}$ is the parameter, and $H = \begin{pmatrix} OBS, & COM_E \end{pmatrix}$.

First, note that the applicability of the parity space approach depends on the rank of E. Second, when parity space works, the decision procedure leads to detect a change in the means of the residual $r = W\bar{y}$ (whose evaluation form is $W\epsilon$), where W is the parity space matrix, such that $WH = 0$.

Assuming that Z is a given central symmetric convex compact set, with center of symmetry 0, and under some mild hypothesis about the d.f. $g_0(\cdot)$ and $g_1(\cdot)$ (which are satisfied in the classical setting, where $g_0(.)$ and $g_1(.)$ are gaussian, have the same covariance matrix, but different means), it can be shown that the optimal guess for the parameter is $\hat{z} = 0$. Thus, the resulting decision problem resumes in detecting a change in the means of the observation $\bar{y}$ (instead of a change in the means of the residual r).

6. CONCLUSION

Fault Detection and Isolation in dynamic systems rises decision problems in the presence of nuisance parameters. In this paper, the problem has been set in a game theoretic framework, in which the decider guesses the nuisance parameters and Nature chooses the actual ones. The loss function expresses the penalty endured by the decider for not being able to guess the actual parameter values. Different problem formulations have been proposed, and solutions have been characterized under some reasonable assumptions. It has to be noted that this approach does not need any pre-treatment of the data (e.g. producing first a residual vector independent on the initial state), since it directly takes into account the known sets to which the unknown parameters belong.

7. REFERENCES

Basseville, M. and I. V. Nikiforov (1993). *Detection of Abrupt Changes–Theory and Applications.* Prentice Hall, Information and System Sciences Series.

Borovkov, A.A. (1984). *Mathematical Statistics.* Nauka, Moscow.

Brodsky, B.E. and B.S. Darkhovsky (2000). *Nonparametric Statistical Diagnosis. Problems and Methods.* Kluwer, Dordrecht.

Cocquempot, V. and C. Christophe (2000). On the equivalence between observer-based and parity space approaches for fdi in nonlinear systems. Proc. of IFAC Safeprocess. Budapest. pp. 232–237.

Cocquempot, V., J. Ph. Cassar and M. Staroswiecki (1991). Generation of robust analytic redundancy relations. Proc. of the European Control Conference (ECC'91). Grenoble, France.

Gertler, J. J. and M. M. Kunwer (1993). Optimal residual decoupling for robust fault diagnosis. Tooldiag'93 International Conference on Fault Diagnosis. Toulouse, France.

Gertler, J.J. (2000). All linear methods are equal – and extendable to nonlinearities. Plenary lecture, Proceeding IFAC Safeprocess'2000. Hungary, Budapest.

Magni, J. F. and Ph. Mouyon (1994). On residual generation by observer and parity space approaches. *IEEE Trans. on Automatic Control* **39**, 441–447.

Nikiforov, I., M. Staroswiecki and B. Vozel (1996). Duality of analytical redundancy and statistical approach in fault diagnosis. Proceeding of 13th IFAC World Congress. San Francisco, CA.

Seliger, R. and P. M. Frank (1991). Fault diagnosis by disturbance decoupled non-linear observers. Proc. of 30th IEEE Conf. on Decision and Control. Brighton. pp. 2248–2253.

Staroswiecki, M. and B. Darkhovski (2001). On structural and parity space detectability. Proceeding of ECC'01. Portugal, Porto.

Staroswiecki, M., J. Ph. Cassar and V. Cocquempot (1993). Generation of optimal structured residuals in the parity space. Proc. of 12 th IFAC Word Congress. Sydney. pp. 535–542.