Robust Fault Detection using Set-based Approaches

Shuang Zhang[1] and Vicenç Puig[1]

Abstract— **This paper presents the performance of zonotopic fault detection (FD) for additive and multiplicative fault using direct test and inverse test. Zonotopic set-based approaches use the zonotope to describe the uncertain state, parameter and noise which are assumed unknown but bounded to reduce their influences on FD. These FD test methods aim at checking the consistency between the measured and estimated behaviour obtained from estimator in the parameter or output space. When an inconsistency is detected between these two, a fault can be indicated. At last, a motor model will be used to compare the performance of direct test and inverse test for additive and multiplicative faults.**

I. INTRODUCTION

Fault Detection (FD) plays an important role in improving the safety and reliability of automatic control systems. Model-based fault detection checks the consistency based on generating the residual between the estimated behaviors and the observations obtained from sensors. Ideally, the residuals should only be zero when there's not fault. However, due to the existence of modelling uncertainty, unknown noise and disturbance, it leads to mismatch between actual process and estimated process. Therefore, it's important to deal with the uncertainties. Besides, due to the different locations where the faults happen in, there are two types of fault, i.e. additive faults and multiplicative faults. So, the other important issue is to make sure which test method is suitable for certain type of fault through comparing the detection performances of different test methods on these two-types faults.

In this paper, the uncertainty is assumed to be unknown but bounded. So it's better to use zonotopes to describe the uncertainties that is related to measurement noise, disturbances and the modelling error. The advantage is that basic set operations can be reduced to simple matrix calculations [1]. Then the fault can be detected by comparing the residual with a threshold value derived from zonotope [2]. If the residual is larger than such threshold, the existence of the fault can be proved [3], [4],[5]. Otherwise, the system is assumed to be still healthy.

As for fault detection methods, the direct test is based on verifying if the residual is inside the interval of possible values. The inverse test is based on verifying if there exists a value that can explain the measured output of the system, inside the set of possible parameters. If there isn't a value satisfying the actual output, then a fault has been detected [6]. Based on [3]and [7], the direct test is closely related to state estimation methods and more suitable for additive faults, while the inverse test is related to parameter estimation methods better suited for multiplicative (parametric) faults.

Based on the above research, this paper will make a contrast of the performance of direct test in output space and inverse test in parameter space by computing the sensitivity of the residual. The main contribution of this paper is to study the performance of zonotopic fault detection of additive and multiplicative fault when using direct test and inverse tests.

This paper is organized as follows. Section II introduces the considered uncertain dynamic model for the system which is able to deal with uncertainties. Section III reviews two ways to detect faults. Section IV presents and compares Interval Observer Approach (IOA) and Set-Membership Approach (SMA) using zonotopes for state estimation. Section V proposes zonotopic Recursive Least Square (ZRLS) for parameter estimation. Section VI presents the performance of direct test and inverse test for additive faults and multiplicative faults mathematically. In Section VII, an actual motor model with additive fault and multiplicative fault for fault detection using direct test and inverse test are presented. Finally, conclusions are summarised in Section VIII.

II. PROBLEM FORMULATION

The system considered in this paper can be described by uncertain parameter-dependent discrete-time dynamic model which is shown as below.

$$x_{k+1} = A(\theta)x_k + B(\theta)u_k + w_k \quad (1a)$$

$$y_k = C(\theta)x_k + D(\theta)u_k + v_k \quad (1b)$$

where $u_k \in \mathbb{R}^{n_u}$, $y_k \in \mathbb{R}^{n_y}$, $x_k \in \mathbb{R}^{n_x}$, $w_k \in \mathbb{R}^{n_x}$, $v_k \in \mathbb{R}^{n_y}$ are the input, output, state, process noise and measurement noise vectors, respectively. Moreover, the process noise and measurement noise are assumed to be unknown but bounded. Besides, $A(\theta) \in \mathbb{R}^{n_x \times n_x}$, $B(\theta) \in \mathbb{R}^{n_u \times n_u}$, $C(\theta) \in \mathbb{R}^{n_y \times n_x}$, $D(\theta) \in \mathbb{R}^{n_y \times n_u}$ are the state-space matrices, where $\theta \in \mathbb{R}^{n_\theta}$ is the vector of parameters.

To simplify the notations, the index $k+1$ will be replaced by $+$, $k-1$ will be replaced by $-$ and k will be omitted. Then, the dynamical model (1) can be rewritten as

$$x_+ = A(\theta)x + B(\theta)u + w \quad (2a)$$

$$y = C(\theta)x + D(\theta)u + v \quad (2b)$$

*This work was supported by Chinese Scholarship Council (CSC)

[1]S. Zhang and V. Puig are with Automatic Control Department, Universitat Politecnica de Catalunya, Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Supervision, Safety and Automatic Control Research Center (CS2AC), C/. Llorens i Artigas 4-6, 08028 Barcelona, Spain (shuang.zhang,vicenc.puig)@upc.edu

III. DIRECT VS INVERSE TESTS

The main idea of detection methods is to compare the estimated behavior obtained with the system model against the behavior measured with sensors from physical system. The inconsistencies between them are called residuals, which are calculated in the following way under ideal conditions.

$$\mathbf{r}(k) = \mathbf{y}(k) - \hat{\mathbf{y}}(k) \tag{3}$$

where $\mathbf{r}(k)$ is the residuals vector, $\mathbf{y}(k)$ is the system output measurement, and $\hat{\mathbf{y}}(k)$ is the estimated output.

Ideally, the detection test is based on checking if $\mathbf{r}(k) = 0$. If this condition is satisfied then there is no inconsistency between the model and the system and therefore it is assumed that there is no fault. Otherwise if $\mathbf{r}(k) \neq 0$, a fault can be indicated. However under actual conditions, it is necessary to consider the influence of the different sources of uncertainty. Furthermore, considering the faults type, this paper utilizes two methods, i.e. direct test in output space and inverse test in parameter space respectively, for better recognizing the faults.

A. Output Space

Considering noise, the fault detection test in output space yields to check if

$$\mathbf{y}(k) \in [\hat{\mathbf{y}}(k) - \sigma, \bar{\hat{\mathbf{y}}}(k) + \sigma] \tag{4}$$

where σ is the noise bound. In output space, the test can be implemented by zonotopic sate estimation in Section IV. In this case, the estimated output is generated from the estimated state bounded by a zonotope. Then, the direct test involves checking whether the measured output is contained in the estimated output interval. If not, a fault can be indicated.

B. Parameter Space

The inverse test tends to check if there exists a parameter in the set of nominal parameters set that enables the model to be consistent with the measurement.

$$\exists \theta \in \mathsf{Q} \mid \hat{\mathbf{y}}(k, \theta) \in [\mathbf{y}(k) - \sigma, \mathbf{y}(k) + \sigma] \tag{5}$$

This test can be implemented by parameter estimation procedure in Section V checking the intersection between parameter zonotope and strip at each instant.

$$\mathsf{Q}_{k+1} = \mathsf{Q}_k \cap \mathbb{F}_k \tag{6}$$

where

$$\mathbb{F}_k = \left\{ \theta \in \mathbb{R}^{n_\theta} \mid \mathbf{y}(k) - \sigma \leq c(k)^T \theta(k) \leq \mathbf{y}(k) + \sigma \right\} \tag{7}$$

is the strip of parameters consistent with the current measurements. In the inverse test if $\mathsf{Q}_{k+1} = \emptyset$, a fault is indicated .

IV. DIRECT TEST IMPLEMENTATION USING ZONOTOPES

The direct test implementation methods use state estimation methods to obtain the nominal estimation plus the uncertainty interval based on measured data from the system. Then, they check whether the measurement behavior is consistent with the estimated behavior in output space.

A. Interval Observer

To estimate the interval dynamic model (1), IOA designs an observer of the form

$$\hat{x}_+ = A\hat{x} + Bu + L(y - C\hat{x} - Du - v) \tag{11}$$

where $\hat{x}_+$ is the estimated state, L is the observer gain, which changes the modeling of uncertainty propagation by weighting the relative influences of v and w.

Assume that the initial state $\hat{x}_0$ belongs to the set $X_0^{io} = \left\langle c_{x,0}^{io}, R_{x,0}^{io} \right\rangle$, where $c_{x,0}^{io} \in \mathbb{R}^{n_x}$ denotes the center and $R_{x,0}^{io} \in \mathbb{R}^{n_x \times n_{R_{x,0}^{io}}}$ is a shape matrix of the initial state bounding zonotope. According to [8], [9], the state $\hat{x}$ at each instant belongs to such state bounding zonotope $\hat{X}_k^{io}$, i.e.,

$$\hat{X}_k^{io} = \left\langle c_x^{io}, R_x^{io} \right\rangle$$

which can be recursively computed using

$$c_{x,+}^{io} = (A - LC)\, c_x^{io} + Bu + Ly \tag{12a}$$

$$R_{x,+}^{io} = \left[(A - LC)\, \bar{R}_x^{io} \quad w \quad -Lv \right] \tag{12b}$$

where $\bar{R}_x^{io} = \downarrow_q \left(R_x^{io} \right)$ and L provides degrees of freedom to tune the system with respect to some aim. This approach can be regarded as the deterministic counterpart of the stochastic Kalman Filter (KF) when uncertainties are assumed to be unknown but bounded [1].

Thus, this IOA is equal to Zonotopic KF in [9]. And the optimal observer gain L^* can be obtained according to [9], with the aim of min $\left\| R_x^{io} \right\|_{F,W}^2$

$$L^* = \frac{A R_x^{io} {R_x^{io}}^T C^T}{C R_x^{io} {R_x^{io}}^T C^T + v v^T} \tag{13}$$

B. Set-Membership

The Set-Membership Approach (SMA) is an alternative approach for estimating the state of the system. Considering the dynamical model (2), the center c_x^{sm} and the shape matrix R_x^{sm} of the state bounding zonotope $\hat{X}^{sm}$ corrected by the i^{th} output [10], [1],

$$\hat{X}_i^{sm} = \langle c_x^{sm}, R_x^{sm} \rangle \tag{14}$$

can be obtained by the intersection of the prediction state set $P^{sm} = \left\langle c_p^{sm}, R_p^{sm} \right\rangle$, where c_p^{sm} and R_p^{sm} denote the center and shape of the zonotope P_k^{sm}, respectively, and the set of states consistent with each output strip S_y as

$$c_x^{sm} = c_p^{sm} + \lambda \left(y - C c_p^{sm} \right) \tag{15a}$$

$$R_x^{sm} = \left[(I - \lambda C) R_p^{sm} \quad -\lambda v \right] \tag{15b}$$

with

$$c_p^{sm} = A c_{x,-}^{sm} + Bu_- \tag{16a}$$

$$R_p^{sm} = \left[A R_{x,-}^{sm} \quad w_- \right] \tag{16a}$$

where the optimal λ can be obtained with the aim of $\min \|R_x^{sm}\|_F^2$.

$$\lambda^* = \frac{R_p^{sm} {R_p^{sm}}^T C}{C R_p^{sm} {R_p^{sm}}^T C^T + v v^T} \tag{17}$$

C. Comparison of IOA and SMA

From theoretical aspect, IOA obtains the estimated state by the propagation of state and output at the previous time instant. SMA allows to estimate the state set by means of the intersection between the predicted state set at the previous time instant with the strip obtained by the current measurement [10]. So, the main difference between IOA and SMA is the instant that the used measurements to obtain the state estimation.

Following the work of [11] and [9], Pourasghar[12] proposed a current IOA (CIOA), which is the deterministic case of stochastic Current estimation-type KF, to relate IOA to SMA given by

$$c_x^{cio} = c_p^{cio} + L \ \ y - C c_p^{cio} \tag{18a}$$

$$R_x^{cio} = \ \ (I - LC) R_p^{cio} \quad - Lv \tag{18b}$$

with

$$c_p^{cio} = A c_{x;-}^{cio} + B u_- \tag{19a}$$

$$R_p^{cio} = \ \ A R_{x;-}^{cio} \quad w_- \tag{19a}$$

$$L^* = \frac{R_p^{cio} {R_p^{cio}}^T C}{C R_p^{cio} {R_p^{cio}}^T C^T + v v^T} \tag{20}$$

where c_x^{cio} and R_x^{cio} denote the center and the shape of zonotope bounding the set of estimated states. It is worth noted that the CIOA is performed based on the information of the measurement given at the current time instant. And also the observer gain L^* of CIOA and parameter λ of SMA are identical with the same aim of minimizing the zonotope size. Therefore, SMA and IOA are related through introducing CIOA.

V. INVERSE TEST IMPLEMENTATION USING ZONOTOPES

The inverse test implementation methods related to parameter estimation are only based on a set of inputs and measurement outputs in a non-faulty scenario to obtain estimated parameters in the form of zonotopes, and then check whether there is an intersection between parameter zonotope and strip at each instant.

Based on Ogata's book [11], current estimation-type KF (for state estimation) after applying the substitution (21) and Recursive Least Square (RLS) (for parameter estimation) are of identical form. Hence, we can establish a conversion from state estimation to parameter estimation by using (21). Thus, zonotopic RLS (ZRLS) can be obtained from zonotopic current estimation-type KF (CIOA) with the substitution (21).

$$A = I; B = 0; C = c(k)^T; D = 0; w = 0 \tag{21}$$

After using (21), the system can be expressed in the regressor form as follows:

$$y(k) = c(k)^T \theta + v(k) = \hat{y}(k) + v(k) \tag{22}$$

where $c(k)$ is the regressor vector with dimension of n_θ which contains functions of inputs $u(k)$ and outputs $y(k)$; $v(k)$ is additive noise where $|v(k)| \le \sigma$. Then, regard parameters as the states: $x(k) = \theta(k)$. The regressor model can be rewrited in the sate space form:

$$\theta(k+1) = \theta(k) \tag{23a}$$

$$y(k) = c(k)^T \theta(k) + v(k) \tag{23b}$$

where $\theta \in \mathsf{Q}_k$ is the parameter vector of dimension n_θ and Q_k is the set that bounds parameter values, which can be described by a zonotope as follows [13]:

$$\mathsf{Q}_k = P \oplus \mathbf{H} B^n = \{P + \mathbf{H}\mathbf{z} : \mathbf{z} \in B^n\} \tag{24}$$

Thus, substituting (18)-(20) by (21) and combining (24), the inverse test implementation method ZRLS is as follows:

$$P = P_- + L \ \ y - c^T P_- \tag{25a}$$

$$H = \ \ I - L c^T \ \ H_-; -Lv \tag{25b}$$

$$L^* = \frac{H_- {H_-}^T c}{c^T H_- {H_-}^T c + v v^T} \tag{26}$$

It is worth to note that after applying the substitution (21) to SMA, we can obtain SMA for parameter estimation, which has the same formula as ZRLS.

VI. COMPARISON IN FAULT DETECTION

In order to compare the performance of proposed fault detection methods, the sensitivity of residual to a fault can be used [14][7]:

$$S_f = \frac{\partial r}{\partial f} \tag{27}$$

which describes how sensitive the residual r is to a given fault f. Normally, $S_f \in [-1; 1]$. If S_f is equal to a non-zero constant, it means the residual is sensitive to the fault. And the closer the constant to 1, the greater the sensitivity. on the contrary, if S_f is varying or equal to 0, it means the residual is not sensitive to the fault.

In fault detection, two kinds of faults are typically considered [15]:

- Additive faults: input sensor fault f_u, output sensor fault f_y and actuator sensor fault f_a.
- Multiplicative faults: parametric fault f_θ

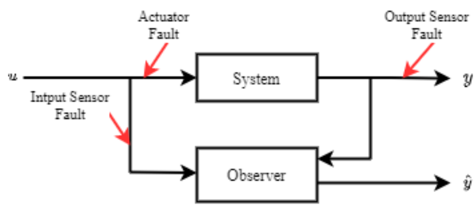

Fig. 1. Different faults

A. Direct Test in Output Space

In case of considering these faults in direct test, then the dynamical model can be rewritten as:

$$x_+ = A(\theta)x + B(\theta)(u + f_a) + w \quad (28a)$$

$$y = C(\theta)x + D(\theta)u + v + f_y \quad (28b)$$

where $\theta = \theta_{nf} + f_\theta$, and θ_{nf} denotes θ in non-faulty scenario. Besides, f_u will affect the observer:

$$\hat{x}_+ = A(\theta)\hat{x} + B(\theta)(u + f_u) + L(y - \hat{y}) \quad (29)$$

Hence, y is affected by f_a, f_θ and f_y, $\hat{y}$ is affected by f_u. In output space the residual

$$r = y - \hat{y} \quad (30)$$

Then, fault sensitivity can be particularized, in the case of:

- sensor output faults as

$$S_{fy} = \frac{\partial r}{\partial f_y} = \frac{\partial (y - \hat{y})}{\partial f_y} = 1 \quad (31)$$

- sensor input faults as

$$\begin{aligned} S_{fu} &= \frac{\partial(-\hat{y})}{\partial f_u} = -\frac{\partial (C(\theta)(A(\theta)\hat{x} + B(\theta)(u + f_u)) + v)}{\partial f_u} \\ &= -C(\theta)\frac{\partial B(u + f_u)}{\partial f_u} = -C(\theta)B(\theta) \end{aligned} \quad (32)$$

- actuator faults as

$$\begin{aligned} S_{fa} &= \frac{\partial r}{\partial f_a} = \frac{\partial y}{\partial f_a} = C(\theta)\frac{\partial (A(\theta)x + B(\theta)(u + f_a) + w)}{\partial f_a} \\ &= C(\theta)B(\theta) \end{aligned} \quad (33)$$

- parametric faults

$$\begin{aligned} S_{f_\theta} &= \frac{\partial r}{\partial f_a} = \frac{\partial y}{\partial f_a} = \frac{\partial (C(\theta)x + D(\theta)u + v)}{\partial f_a} \\ &= \frac{\partial C(\theta)}{\partial f_a}x + \frac{\partial D(\theta)}{\partial f_a}u \end{aligned} \quad (34)$$

Based on (31)-(34), direct test for sensor output faults, input sensor faults and actuator faults are non-zero constant, so it is sensitive to additive faults. While for parametric faults, the sensitivity is varying. Therefore, direct test is not permanently sensitive to multiplicative faults.

B. Inverse Test in Parameter Space

For inverse test, the dynamic model considering faults is as follows:

$$\theta(k+1) = \theta(k) \quad (35a)$$

$$y(k) = c(k)^T \theta(k) + v(k) + f_y \quad (35b)$$

and the observer is:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + L(y(k) - \hat{y}(k)) \quad (36)$$

So, parametric faults are $f_\theta = \theta - \hat{\theta}$. Then considering residual with

- $y(k)$ affected by f_y and actuator faults f_a.
- regressor vector $c(k)$ affected, depending on the regressor structure, by f_u if u is present in the regressor and by f_y and f_a if y is present in the regressor [15].
- $\theta(k)$ affected by multiplicative faults f_θ.

Furthermore, the residual in parameter space is

$$r = \theta - \hat{\theta} \quad (37)$$

Then, fault sensitivity can be obtained separately in parameter space:

- sensor output faults as

$$\begin{aligned} S_{fy} &= \frac{\partial r}{\partial f_y} = \frac{\partial (\theta - \hat{\theta})}{\partial f_y} = -L\frac{\partial (c^T\theta + v + f_y)}{\partial f_y} \\ &= -L(\frac{\partial c^T}{\partial f_y}\theta + 1) \end{aligned} \quad (38)$$

- sensor input faults as

$$S_{fu} = \frac{\partial (\theta - \hat{\theta})}{\partial f_u} = -\frac{\partial (L(y - \hat{y}))}{\partial f_y} = L\frac{\partial c^T}{\partial f_y}\theta \quad (39)$$

- actuator faults as

$$S_{fa} = \frac{\partial (\theta - \hat{\theta})}{\partial f_a} = 0 \quad (40)$$

- parametric faults

$$S_{f\theta} = \frac{\partial (\theta - \hat{\theta})}{\partial f_u} = 1 \quad (41)$$

Based on (38)-(41), inverse test for sensor output faults, input sensor faults and actuator faults are varying or zero, so it is not sensitive to additive faults. While for parameter faults, the sensitivity is 1. Therefore, inverse test is sensitive to multiplicative faults.

VII. EXAMPLE

This section introduces a BLDC motor model to test the above approaches, namely, to estimate the states and parameters of this model in no-faulty situation using IOA and SMA, and then apply inverse test and direct test using both state estimation and parameter estimation in each different faulty situation to compare the test performance.

A. BLDC Motor Model

A BLDC motor bears a resemblance to a permanent magnet synchronous machine (PMSM): surface-mounted permanent magnets in the rotor and a 3-phase connected to winding in the stator. The state space model is as follows, which is from [16]

$$\frac{d}{dt}\begin{bmatrix} i \\ \omega_r \end{bmatrix} = \begin{bmatrix} -\frac{R_{eq}}{L_{eq}} & -\frac{k_e}{L_{eq}} \\ \frac{k_T}{L_{eq}} & -\frac{B_r}{J} \end{bmatrix}\begin{bmatrix} i \\ \omega_r \end{bmatrix} + \begin{bmatrix} \frac{1}{L_{eq}} & 0 \\ 0 & -\frac{1}{J} \end{bmatrix}\begin{bmatrix} V_{dc} \\ T_L \end{bmatrix}$$

where the state variables are the current i and rotor speed ω_r; the input variables are DC voltage V_dc and resisting (or load) torque T_L; the parameters in state transmission matrix are the equivalent resistance of the simplified circuit R_{eq}, the back emf coefficient k_e, the equivalent inductance of the simplified circuit L_{eq}, the torque coefficient k_t, damping (or viscous friction) coefficient B_r and the moment of inertia of the rotational system J, respectively.

B. BLDC Motor Faults

As discussed in Section III, faults can be classified into additive faults and multiplicative faults. Usually, additive faults happen in input or output sensors, while multiplicative faults happen in parameters. Here, introduce both types of faults into this motor model:

- Additive faults: stuck sensor fault (f_θ), current sensor fault (f_i).
- Multiplicative faults: resistance fault (f_R), inductance fault(f_L), friction fault (f_B) and inertia fault (f_J).

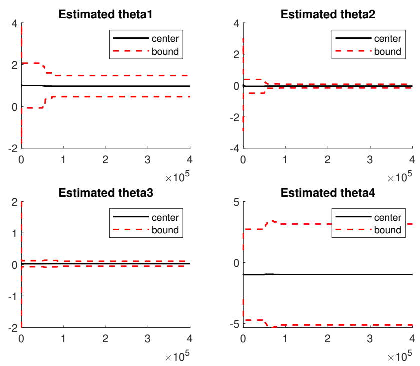

Fig. 2. Parameter Estimation Using ZRLS

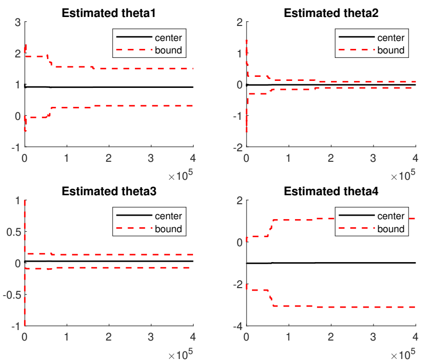

Fig. 3. Parameter Estimation Using SMA

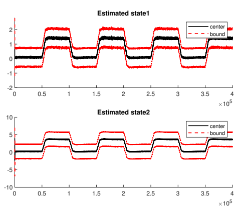

Fig. 4. State Estimation-IOA

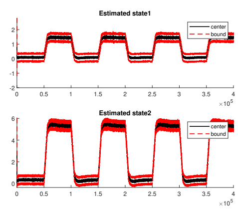

Fig. 5. State Estimation-SMA

C. Consistency Test

For implementation the consistency test, state estimation and parameter estimation are conducted firstly based on the discrete-time motor model in no-faults scenario. And the parameter and state estimation results are shown as Fig.(2,3) and Fig.(4,5) respectively.

To compare the performance of zonotopic fault detection of additive and multiplicative fault direct test and inverse test, this paper introduces 6 faults using the implementation methods in Section IV and V. The detection results are as Fig.(6-11).

It's worth remarking that if there is no intersection between the zonotope and the strip in inverse test in Figs.6-9. It means that this system contains a fault. For direct test, if the measured output (yellow line) is not in the estimated output bound, it indicates a fault, for example Fig.10. Especially, faults can also be detected during state estimation stage using SMA, to check whether the intersection (15) is empty or not. The fault will be detected if the intersection is empty, like Fig.(11), the intersection at instant 100001 is empty.

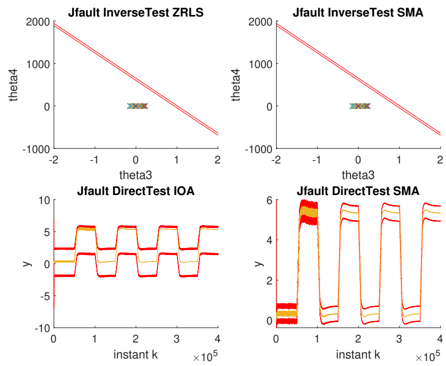

Fig. 6. J Fault Detection

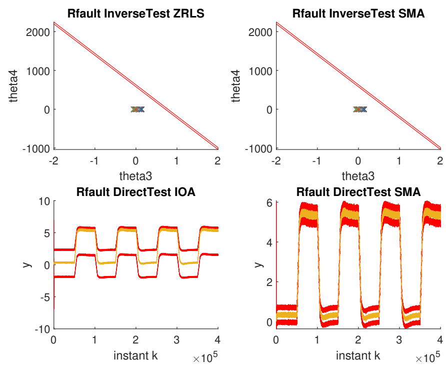

Fig. 7. R Fault Detection

According to these detection results, inverse test can detect all the multiplicative faults, while for some additive faults it not works. At the same time, direct test works well in the detection of additive faults, but is not suitable for multiplicative faults. Hence, we can conclude that inverse and direct test have a better detection performance for multiplicative faults and additive faults respectively.

VIII. CONCLUSIONS

This paper has mathematically and experimentally proved that direct test and inverse test have a better performance