A Tractable Nonlinear Fault Detection and Isolation Technique with Application to the Cyber-Physical Security of Power Systems

P. Mohajerin Esfahani, M. Vrakopoulou, G. Andersson, and J. Lygeros

Abstract—This article consists of two parts: a theoretical part concerned with fault detection schemes, and an application part dealing with cyber security of power systems. In the first part, we develop a tractable approach to design a robust residual generator to detect and isolate faults in high dimensional nonlinear systems. Previous approaches on fault detection and isolation problems are either confined to linear systems or they are only applicable to low dimensional dynamics with more specific structures. In contrast, we propose a novel methodology to robustify a linear residual generator for a nonlinear system in the presence of certain disturbance signatures. To this end, we formulate the problem into the framework of quadratic programming which enables us to solve relatively high dimensional systems. In the second part, the application is motivated by the emerging problem of cyber security in power networks. We provide description of a multi-machine power system that represents a two-area power system, and we model a cyber-physical attack emanating from the vulnerabilities introduced by the interaction between IT infrastructure and power system. The algorithm developed in the first part is finally used to diagnose such an intrusion before the functionality of the power system is disrupted.

I. INTRODUCTION

The task of fault detection and isolation (FDI) in dynamical systems is the problem of generating a diagnostic signal sensitive to the occurrence of specific faults. This problem essentially has the connotation of designing a filter with all available information as inputs which leads to a non-interactive map from faults to residual. Therefore the concept of residual plays a central role for the FDI problem. Earlier works on residual generators for linear systems are commonly concerned with more specific classes of models such as transfer function [14], state-space model [8], and descriptive model [15]. Roughly speaking, for the aforementioned models the residual generators are classified into two categories: observer-based and parity-space-like approaches.

In the observer-based approach, Beard [7] and Jones [16] are the first pioneers where the filter is a Luenberger observer such that failures of different system components affect the residuals in linearly independent directions. Some inherent limitation of Beard-Jones approach was improved in

Massoumnia et al. [22]. Later, this approach was extended to more general classes of systems by Seliger and Frank [26] surveyed in [13], and was comprehensively investigated by Speyer and coauthors in the presence of measurement noise, see [9] and [6].

Parity-space-like approaches have been studied in the framework of descriptor models in several papers e.g. [20] and [23]. In more recent work, Nyberg and Frisk extended the class of systems as well as the notion of detectability [23]. In this article the class of functions is extended to linear differential-algebraic equations (DAEs) that covers all the previous classes, and fault detectability is rather defined as a system property. DAE models appear in many applications such as electrical systems, robotic manipulators, and mechanical systems. For instance, a motion of robot constrained to a certain geometrical area can be modelled in this class.

In the context of nonlinear systems, De Persis and Isidori [24] have proposed a differential geometric approach to extend the approach of [21] and design the FDI filter for the observation of certain states in the presence of unknown disturbances. The problem of fault detection and isolation has been characterized in terms of the properties of certain distribution, which can be considered as the nonlinear analog of the unobservability subspaces first introduced in [21, Section IV]. However, as one is required to verify the conditioned invariant property of certain distributions, the algorithm is rather intractable for relatively large and sophisticated dynamics.

Motivated by this fact, in this article we aim to find a compromise between theoretical soundness and practical feasibility. For this purpose, we restrict the class of FDI filters to linear residual generators which allows us to explicitly track the contributions of linear and nonlinear dynamic terms as well as fault signals into the residual. Our main goal is then to control the nonlinear term contributions in the presence of certain disturbance patterns. In another word, in contrast to existing literature toward FDI methods, we impose constraints on disturbance signals rather than nonlinearity structure of the system dynamic.

In more details, we first review some results developed for linear systems of [23] in §III-A. We then propose a linear programming (LP) formulation, Lemma 3.1, as an alternative characterization of residual generators. This is in fact an LP counterpart of matrix polynomial formulation in the literature. In §III-B, we generalize the DAE model to contain nonlinear terms as well, see (6). In order to extend the scheme to the nonlinear model, we propose two
approaches where each approach may be viewed from a
certain class of applications. The first approach in §III is a
straightforward extension of the LP formulation developed in
the preceding subsection, thank to the fact that the FDI filter
can be designed up to a scaler. The idea may be justified in
applications that the system works during normal operation
in a neighbourhood of an equilibrium. This suggests to
neglect the contribution of nonlinear terms, and mainly focus
on the mapping from fault signal into the residual. The
second approach in §III is the main contribution of the article.
Given some particular signatures of possible disturbances,
we approximate the contributions of the nonlinear terms
into the residual. The approximation scheme is based on
the projection of these contributions into a finite-dimensional
function space which is closed under differentiation operator.
In the following we formulate the $L_2$ norm of errors in terms
of a family of quadratic programming (QP) problems where
the number of QP problems is the degree of the FDI filter.

In the second part of the article, §IV, we first describe
the mathematical model of the IEEE 118-bus power network
equipped with primary and secondary frequency control.
The latter is also referred as Automatic Generation Control
(AGC) and is one of the few control loops that are closed
over the SCADA system without human operator interven-
tion. This interaction with the IT infrastructure may give rise
to cyber security issues which are investigated in our earlier
work [11], [12]. It was shown that if an attacker gains access
to the AGC signal, unacceptable frequency deviations and
power oscillations may occur. This can trigger out-of-step,
der under frequency and generator frequency protection relays,
and hence lead to load shedding and generation tripping. If
the intrusion is detected on time, one may prevent further
damage by disconnecting the AGC. Therefore, it is crucial
to utilize available measurements to diagnose the AGC
intrusion sufficiently fast, even in the presence of unknown
load deviations. By invoking the proposed FDI scheme, a
protection layer is constructed that permits us to address the
aforementioned security concern.

The article is organized as follows. In §II a general class
of linear model is described, and basic definitions of residual
generators and detectability notion are introduced. §III pro-
vides two algorithms to tackle nonlinear FDI problems. We
then explain a multi-machine power system equipped with
AGC in §IV, and in §V apply our technique developed in
the preceding sections to diagnose the AGC intrusion. We
conclude with some remarks and directions for future work
in §VI.

II. MODEL DESCRIPTION AND BASIC DEFINITIONS

In this section we introduce the class of linear models
proposed in [23], and follow the basic definitions in this
article. The model is considered as

$$
H(p)x + L(p)z + F(p)f = 0,
$$

where $p$ is the distributional derivative operator [2, Section
I], and $H, L, F$ are polynomial matrices in the operator $p$.
We assume that $x, y, z$ are piece-wise continuous functions
from $\mathbb{R}_+ \times \mathbb{R}^n_z, \mathbb{R}^m_z, \mathbb{R}^m_f$ respectively. We denote these
sets by $\mathcal{W}_z, \mathcal{W}_z, \mathcal{W}_f$. In the model (1), $x$ represents all
unknowns signals, e.g. internal system states and unknown
exogenous disturbances. $z$ contains all the known signals,
for instance control signals and state measurements, and $f$
stands for the signals to be detected such as faults or intrusion
signals.

One may extend the space of functions $x, z, f$ to Sobolev
spaces, but an elaborate discussion regarding this issue is out-
side the scope of our study. On the other hand, if these spaces
are restricted to the smooth functions, then the operator $p$
can be understood as the classical differentiation operator.
Throughout this article we will focus on continuous-time
models, but one can obtain similar results for discrete-time
models by changing the operator $p$ to the time-shift operator
$q$. In the rest of the article, we use $p$ when the matrices
involved are viewed as an operator, e.g. $H(p)$, and if they
are dealt as a polynomial matrices, we shall use the complex
variable $s$ instead of $p$, e.g. $H(s)$.

The following example indicates that an ordinary state-
space description is indeed a particular case of the linear
model (1). Consider the model

$$
\begin{align*}
E\dot{x}(t) &= AX(t) + Bu(t) + B_d d(t) + B_f f(t) \\
Y(t) &= CX(t) + Du(t) + D_d d(t) + D_f f(t)
\end{align*}
$$

(2)

where $u(\cdot)$ is the input signal, $d(\cdot)$ unknown disturbance,
$Y(\cdot)$ state measurements, and $f(\cdot)$ possible faults/attack
signal to be detected. Therefore, defining $x := [X' \quad d']'$
and $x := [Y' \quad u']'$ and matrices

$$
H(p) := \begin{bmatrix} -pE + A & B_d \\ C & D_d \end{bmatrix}, \quad L(p) := \begin{bmatrix} 0 & B_u \\ -I & D_u \end{bmatrix},
$$

$$
F(p) := \begin{bmatrix} B_f \\ D_f \end{bmatrix},
$$

directly fits the model (2) to (1).

Note that the model (1) affords an appropriate framework
deal with the algebraic constraints. Moreover, we do not
assume any condition on initial values of the signals $x, z, f$.
The only assumption one may impose on the model matrices
(1) is that there is no linear dependency in the model when
$f \equiv 0$. This condition is satisfied when $[H(s) L(s)]$ has full
row rank.

Let us proceed with some basic definitions and clarify
what we mean by sensitivity and residual generator. To this
end, let us formally characterize all possible observations
of the model (1) in the absence of the fault signal $f$ as
$\mathcal{M} := \{z \in \mathcal{W}_z \mid \exists x \in \mathcal{W}_x : H(p)x + L(p)z = 0\}$.
This set of observation is called behavior of the system used
in the behavioral approach to systems theory, see [25, Section
2.4] for more details.

**Definition 2.1 (Residual Generator):** A proper linear time
invariant filter $r := R(p)z$ is a residual generator for (1) if
for all $z \in \mathcal{M}$, it holds that $\lim_{t \to \infty} r(t) = 0$.

The following Definition provides a notion of sensitivity
for the above residual generators with respect to a specific
fault:
Definition 2.2 (Fault Sensitivity): The residual generator introduced in Definition 2.1 is sensitive to fault $f_i$ if the transfer function from $f_i$ to $r$ is nonzero, where $f_i$ is the $i^{th}$ element of the signal $f$.

III. DESIGN OF RESIDUAL GENERATOR

The main objective of this section is to establish a tractable approach, possibly for nonlinear systems, to design a sensitive linear residual generator in the sense of Definitions 2.1 and 2.2. For this purpose we first characterize the residual generator as a polynomial matrix equation and then make a link from the polynomial matrix formulation to an LP problem. In the sequel we extend the approach to a class of nonlinear systems. To that end, we propose a new framework, in a QP formulation, so as to minimize the contributions of nonlinear terms into the residual of the designed filter.

A. Linear Models

Consider a linear model as defined in (1). Along the same vein as [25, Section 2.5.2], one may observe that the behavior set $\mathcal{M}$ can alternatively be characterized as $\mathcal{M} = \{ z \in \mathbb{C}^m : N_H(p)L(p)z = 0 \}$, where the collection of the rows of $N_H(s)$ forms an irreducible polynomial basis for the left null-space of the matrix $H(s)$. This representation is the basic idea to design a residual generator of model (1). Namely, by picking a linear combination of $N_H(p)$ rows and adding stable dynamic $d(p)$ of sufficiently order, we arrive at a residual generator in the sense of Definition 2.1 with transfer operator

$$ R(p) = d^{-1}(p)\gamma(p)N_H(p)L(p) := d^{-1}(p)N(p)L(p) \quad (3) $$

The above filter can easily be realized by an explicit state-space description (2) with the input $z$ and output $r$. Hence, a sensitive residual generator can be characterized as the polynomial matrix equations

$$ N(s)H(s) = 0, \quad (4a) $$
$$ N(s)F(s) \neq 0, \quad (4b) $$

where the polynomial vector $N(s)$ is to be chosen. Let us recall that equations (4a) and (4b) in fact address the required conditions of Definition 2.1 and Definition 2.2 respectively.

The following lemma expresses the nontrivial matrix polynomial equations (4) in a linear programming framework.

Lemma 3.1: Let $N(s)$ be the solution of (4), where

$$ H(s) := \sum_{i=0}^{d_H} H_is^i, \quad F(s) := \sum_{i=0}^{d_F} F_is^i, \quad N(s) := \sum_{i=0}^{d_N} N_is^i. $$

Then the conditions in (4) can equivalently be written as

$$ \bar{N}H = 0, \quad (5a) $$
$$ \| \bar{N}F \|_{\infty} \geq 1, \quad (5b) $$

where $\| \cdot \|_{\infty}$ is infinite vector norm, and

$$ \bar{N} := [N_0 \quad N_1 \quad \cdots \quad N_{d_N}] $$

Proof: By definitions, it is easy to observe that

$$ N(s)H(s) = \bar{N}H[I \quad sI \quad \cdots \quad s^kI]', \quad k := d_N + d_H. $$

Moreover, in light of the linear structure of equations (4), one can simply scale the equations and arrive at the assertion of the Lemma.

Remark 3.2: It is straightforward to inspect that if $\bar{N}$ is a solution to (5), then so is $-\bar{N}$. Hence, the inequality (5b) can be understood as an $m$ different LP problems where $m = n_f(d_F + d_N + 1)$ is the number of $\bar{F}$ columns, and $n_f$ is the dimension of signal $f$ in the model (1). That is, in each true LP problem, one can only focus on a component of the vector $\bar{N}\bar{F}$ and replace the inequality (5b) with

$$ \bar{N}\bar{F}v \geq 1, \quad v := [0, \cdots, 1, \cdots, 0]' $$

B. Nonlinear Models

In this section we extend the model of (1) with a nonlinear term $E(\cdot)$ as a function of unknown signal $x$:

$$ E(x) + H(p)x + L(p)z + F(p)f = 0. \quad (6) $$

It is straightforward to see that the residual of the filter (3) is obtained as

$$ r := R(p)z = -d^{-1}(p)N(p)(F(p)f + E(x)). \quad (7) $$

Therefore, the main objective, roughly speaking, is to reduce the contribution of $E(x)$ into residual (7) while increase the residual sensitivity with respect to the fault $f$. For this purpose, we propose two approaches focusing on different terms of the residual (7). In both approaches we assume that the denominator of the FDI filter is fixed and the main target to design is the numerator coefficients, i.e., $N(p)$ in (7).

Approach (I): The main objective of this approach is to somehow increase the sensitivity of the residual (7) with respect to the fault $f$. Without loss of generality, one may extract the linear part of $E(x)$, and assume that the contribution of the nonlinear term $E(x)$ can be neglected providing that the system (6) normally works around an equilibrium point $x_E$. From practical perspective this could be a reasonable assumption since in many applications a system dynamic is considerably deviated from a nominal operating point if there exists a fault/attack signal. In these cases it is essentially important to only detect the fault/attack signal on time. However, since the signal $E(x(\cdot))$ passes through the FDI filter, and consequently the derivatives of the contribution signals are also involved, it is not clear any more if the assumption holds in view of the residual.
issue will be addressed in the next approach. Hence, as a first approach, one can slightly modify the formulation in (5) and arrive at

$$\max_{\mathcal{N}} \| \tilde{N} \tilde{F} \|_{\infty} \quad (8a)$$

s.t. \[ \tilde{N} \tilde{H} = 0 \]

$$\| \tilde{N} \|_{\infty} \leq 1 \quad (8b)$$

where the objective function (8a) targets the contribution of signal $f$ into the residual. Let us recall that $\tilde{N} \tilde{F}$ is the vector containing all numerator coefficients of the transfer function from signal $f$ to residual $r$, i.e., $\tilde{N}(s)$ in (4a). Moreover, in a similar fashion as Remark 3.2, it is immediate to consider the objective function (8a) essentially as an $m$ different LP formulations, where $m$ is the number of $\tilde{F}$ columns. Regarding the optimization constraints, we add the second constraint in (8b) to avoid unbounded solutions. Obviously this constraint does not lose generality as the filter $R(p)$ in (6) can be computed up to a scalar. It is also a classical result that the second constraint in (8b) is indeed an LP constraint in an augmented state space, see for instance [19, Section 5.4.3].

**Approach (III):** This approach is the main theoretical contribution of the article. In contrast to existing literature on nonlinear FDI methods, here we impose constraints on disturbance signals rather than nonlinearity structure of system dynamics. Namely, we assume that some rough information about the disturbances pattern is available, i.e., we restrict the disturbances to a certain family of signatures. We then aim to control the contribution of the nonlinear term $E(x)$ into the residual in the presence of these disturbances. In essence, the main objective is to train the FDI filter in order to identify the normal behavior of the system while such disturbances appear. For this purpose, let us fix a certain pattern for the signal $x$. We approximate the mapping $t \mapsto E(x(t))$ in the presence of this disturbance over a given time horizon $[0,T]$. The approximation step is in fact the projection of the function $E_k(x(\cdot))$, $k^{th}$ component of $E(x(\cdot))$, into the linear vector space $\mathcal{N} := \text{span}\{b_0, b_1, \ldots, b_n\}$ where $\{b_i(\cdot)\}_{i=0}^n$ is a basis of smooth functions for $\mathcal{N}$. Let formally introduce this step as

$$e(t) := E(x(t)) \approx \sum_{i=0}^n \beta_i b_i(t) = \beta B(t), \quad t \in [0,T] \quad (9)$$

where $\beta := [\beta_0, \ldots, \beta_n]$ is a constant matrix, and $B := [b_0, \ldots, b_n]^T$ is a vector of smooth functions. Further, we assume that the subspace $\mathcal{N}$ is closed under differentiation operator $p$. This requirement, for instance, is satisfied for the polynomial or Fourier basis. The aforementioned assumption gives rise to translate the linear operator $p$ as a matrix operator, i.e.,

$$pB(t) := \frac{d}{dt} B(t) = DB(t). \quad (10)$$

Let us define $r_e(t) := N(p)e(t)$. In accordance to approximation (9) and operator (10), and in view of projection into the subspace $\mathcal{N}$, one can also approximate the error of residual as follows:

$$r_e(t) \approx \tilde{N} \tilde{D}B(t), \quad \tilde{D} := [\beta' \ D'\beta' \ \ldots \ D'^n \beta'] \quad (11)$$

where $\tilde{N}$ is defined as in Lemma 3.1, and $d_N$ is the degree of FDI filter. Hence, it is now straightforward to formulate the $L_2$ norm of $r_e$ as a quadratic function of the FDI filter coefficients $\tilde{N}$. Namely

$$\| r_e \|_{L_2}^2 \approx \tilde{N} \tilde{D} \tilde{G} \tilde{D}' \tilde{N}', \quad G_{ij} := \int_0^T b_{i-1} b_{j-1} dt, \quad (12)$$

where the matrix $\tilde{D}$ is defined as in (11) and $G$ is a symmetric matrix with dimension $(d_N + 1)$. Note that $G$ is indeed the Gram matrix of the subspace $\mathcal{N}$ contained in a Hilbert space endowed with the inner product $(f, g) := \int_0^T f g dt$ [18, Section 3.6]. Now we are at a place to modify the formulation (8) in order to control the nonlinear term contribution into the residual. To this end, we suggest the following QP type formulation:

$$\min_{\mathcal{N}} NQ \tilde{N}', \quad Q := \tilde{D} \tilde{G} \tilde{D}' \quad (13a)$$

s.t. \[ \tilde{N} \tilde{H} = 0 \]

$$\| \tilde{N} \tilde{F} \|_{\infty} \geq 1 \quad (13b)$$

where $\tilde{D}$ and $G$ are defined in (11) and (12), respectively. Let us recall once again that in light of Remark 3.2 the formulation (13) can be viewed as $m$ different true QP problem where $m = n_f (d_P + d_N + 1)$.

**Remark 3.3:** In practice it may be required to robustify the FDI filter to more than one disturbance pattern, say $x_i(\cdot)$ for $i = 1, \ldots, m$. For this purpose it suffices to first compute the matrices $Q_i$ corresponding to each of $x_i(\cdot)$ and then solve the QP problem in (13) with $Q := \sum_{i=1}^m Q_i$.

**IV. CASE STUDY: MULTI-MACHINE TWO-AREA POWER NETWORK**

In this section a multi-machine power system [5], is described. The system is arbitrarily divided into two control areas. The generators are equipped with primary frequency control and each area is under the so called Automatic Generation Control (AGC) which adjusts the generating setpoints of specific generators so as to regulate frequency and maintain the power exchange between the two areas to its scheduled value. Each generator is also equipped with a turbine that is represented by a first order transfer function.

We consider a system comprising of $n$ buses and $g$ number of generators. Let $G = \{i\}^g_{i=1}$ denote the set of generator indices and $A_1 = \{i \in G \mid i \text{ in Area 1}\}$, $A_2 = \{i \in G \mid i \text{ in Area 2}\}$ the sets of generators that belong to Area 1 and Area 2, respectively. Let also $L_{tie} = \{(i,j) | i,j \text{ edges of a tie line from area } k \text{ to the other areas}\}$ where a tie line is a line connecting the two independently controlled areas and let also $K = \{1, 2\}$ be the set of the indices of the control areas in the system.

Using the classical generator model every synchronous machine is modelled as constant voltage source behind its transient reactance. The dynamic states of the system are
the rotor angle $\delta_i$ (rad), the rotor electrical frequency $f_i$ (Hz) and the mechanical power (output of the turbine) $P_{m,i}$ (MW) for each generator $i \in G$. We also have one more state that represents the output of the AGC $\Delta P_{agc,k}$ for each control area $k \in K$.

We denote by $E_{G_i} \in \mathbb{C}^n$ a vector consisting of the generator internal node voltages $E_{G_i} = |E_{G_i, i}|/\delta_i$ for $i \in G$. The phase angle of the generator voltage node is assumed to coincide with the rotor angle $\delta_i$ and $|E_{G_i, i}|$ is a constant. The voltages of the rest of the nodes are included in $V_N \in \mathbb{C}^n$, whose entries are $V_{N,i} = |V_{N,i}|/\delta_i$ for $i = 1, \ldots, n$. To remove the algebraic constraints that appear due to the Kirchhoff’s first law for each node, we retain the internal nodes (behind the transient reactance) of the generators and eliminate the rest of the nodes. This could be achieved only under the assumption of constant impedance loads since in that way they can be included in the network admittance matrix.

The node voltages can then be linearly connected to the internal node voltages and hence to the dynamic state $\delta_i$. Moreover, this results in a reduced admittance matrix that corresponds only to the internal nodes of the generators. The power flows, which are a function of the node voltages, can be now expressed directly by the dynamic states of the system. The resulting model of the two area power system is described by the following set of equations.

\[
\delta_i = 2\pi(f_i - f_0),
\]
\[
\dot{f}_i = \frac{f_0}{2\pi S_{B_i}}(P_{a_i} - P_{e_i}(\delta) - \frac{1}{D_i}(f_i - f_0) - \Delta P_{load,i}),
\]
\[
\dot{P}_{m,a_k} = \frac{1}{T_{ch,a_k}}(P_{m,a_k} + v_{a_k}\Delta P_{sat,a} + w_{a_k}\Delta P_{sat,a_k} - P_{m,a_k}),
\]
\[
\dot{\Delta}P_{agc,k} = \sum_{j \in A_k} e_{kj}(f_j - f_0)
\]
\[
+ \sum_{j \in A_k} b_{kj}(P_{m,j} - P_{e,j}(\delta) - \Delta P_{load,j}) - \frac{1}{T_{N_k}}g_k(\delta, f)
\]
\[
- C_{p,i}h_k(\delta, f) - \frac{K_k}{T_{N_k}}(\Delta P_{agc,k} - \Delta P_{sat,a_k}).
\]

where $i \in G$, $a_k \in A_k$ for $k \in K$. Superscript $sat$ on the AGC output $P_{agc,k}$. and on the primary frequency control signal $\Delta P_{sat,a}$ highlights the saturation that the signals are subjected to. The primary frequency control is given by $\Delta P_{sat,a} = -(f_i - f_0)/S_i$ . Based on the reduced admittance matrix, the generator electric power output is given by

\[
P_{ei} = \sum_{j=1}^{n} E_{G_j}E_{G_j}(G_{e,j}^T \cos(\delta_j - \delta_j) + B_{e,j}^T \sin(\delta_i - \delta_j)).
\]

Moreover, $g_k = \sum_{(i,j) \in E} (P_{i,j} - P_{G_i})$ and $h_k = d g_k / dt$, where the power flow $P_{i,j}$, based on the initial admittance matrix of the system, is given by

\[
P_{i,j} = |V_{N,i}||V_{N,j}|(G_{ij} \cos(\theta_i - \theta_j) + B_{ij} \sin(\theta_i - \theta_j))
\]

All undefined variables are constants. Details on the derivation of the models can be found in [10].

The demonstration of the FDI methodology will be based on an undesirable signal $U$, additive to the AGC signal. For instance, if the attack signal is imposed in Area 1, the mechanical power dynamics of Area 1 will be modified as

\[
\dot{P}_{m,a_1} = \frac{1}{T_{ch,a_1}}(P_{m,a_1}^0 + v_{a_1}\Delta P_{sat,a_1} + w_{a_1}(\Delta P_{agc,a_1} + U) - P_{m,a_1}).
\]

The described model can be compactly written as

\[
\begin{align*}
\dot{X}(t) &= h(X(t)) + Bdd(t) + Bff(t) \\
Y(t) &= CX(t)
\end{align*}
\]

where $X := \{\{\delta_i\}^T, \{f_i\}^T, \{P_{m,a_1}\}_{a_1 \in A_1}, \{P_{m,a_2}\}_{a_2 \in A_1}, \Delta P_{agc_1}, \Delta P_{agc_2}\}^T$, $d := \{\Delta P_{load}(t)\}^T$ is the unknown load disturbance, and $f(t) := U$ corresponds to the fault signal we want to detect. The measurement state $Y(\cdot)$ is assumed to be $Y := \{\{f_i\}^T, \{P_{m,a_1}\}_{a_1 \in A_1}, \{P_{m,a_2}\}_{a_1 \in A_1}\}^T$. The nonlinear function $h(\cdot)$ and the constant matrices $B_d$, $B_f$ and $C$ can be easily obtained by the mapping between the analytical model and (14).

The model can be then written in the form of (6) by defining $x := [X' - X_e' d']^T$, $z := Y - CX_e$ and

\[
E(x) := \begin{bmatrix} h(X) - A(X - X_e) \\ 0 \end{bmatrix}, \quad L(p) := \begin{bmatrix} 0 \\ -I \end{bmatrix}, \quad H(p) := \begin{bmatrix} -p + A & B_d \\ C & 0 \end{bmatrix}, \quad F(p) := \begin{bmatrix} B_f \\ 0 \end{bmatrix},
\]

where $X_e$ is the equilibrium of (14), i.e., $h(X_e) = 0$, and $A := \frac{\partial h}{\partial X}|_{X=X_e}$. The following section will highlight via simulations the security and reliability of the filter.

V. SIMULATION RESULTS

A. Test System

To illustrate the FDI methodology we employed the IEEE 118-bus system. The data of the model are retrieved from a snapshot available at [1]. It includes 19 generators, 177 lines, 99 load buses and 7 transmission level transformers. Since there were no dynamic data available, typical values provided by [3] were used for the simulations. The network was arbitrarily divided into two control areas and the nonlinear frequency model of the network was developed according to [4] so as to be the test case for the filter described in [3].

B. Diagnosis filters

In this part we apply the FDI schemes proposed in (8) and (13) to detect the cyber attack on the AGC of the first area, in the presence of load deviations $\Delta P_{load}$ in all nodes. The filter must be insensitive to the normal situation of the network operating conditions (including acceptable load deviations), and highly reflects any undesirable intrusion in the AGC command. In the following simulations we fix the degree of filters as $d_N = 7$, and solve the equations introduced in (8) and (13) using YALMIP toolbox [17].

Figure 1 illustrates the results of the FDI filter obtained from the LP formulation in (8). Fig.1.a depicts a load deviation in node 5 ($\Delta P_{load_5}$) at time $t = 1$, and an attack signal in the first area AGC at time $t = 10$. As demonstrated in Fig.1.b, the FDI filter works very well while the inputs are measurements from an ideal linearized model. However, as shown in Fig.1.c, the filter is highly sensitive to nonlinearities and immediately reacts to the load deviation at node 5.

In the second simulation, we aim to overcome the nonlinearities contributions into the residual with the aid of QP formulation of (13). To this end, we choose the polynomial
functions up to degree $n = 40$ as the basis of approximation scheme. Namely, $b_i(t) := t^i$ for $i$ in $\{0, 1, \ldots, n\}$ and $B(t) := [1, t, \ldots, t^n]^\top$. We select $T = 10$ as the approximation horizon, and refer to [10] for details on the differentiation matrix $D$ and Gram matrix $G$ as introduced in (10) and (12), respectively. We further assume step functions as particular signatures of load deviations which individually appear at each node. Therefore, each load deviation results in a certain pattern of $e(t)$ introduced as in (9). For the approximation step in (9) and computation of matrix $\beta$, we refer the reader to [18, Chapter 3].

Figure 2 illustrates the results of the FDI filter obtained from the QP formulation in (8). As demonstrated in Fig. 1.a-c, not only is the residual sensitive to the attack signal in the first area, but also the contribution of nonlinear terms in the presence of load deviation is significantly decoupled.

VI. CONCLUDING REMARKS AND FUTURE DIRECTIONS

We proposed a tractable algorithm to design an FDI residual generator for nonlinear systems. The technique has been formulated as a family of QP problems that the number of problems is linear with respect to the degree of FDI filter. To illustrate the performance of our theoretical results, we applied the proposed diagnosis filter to a two-area power system so as to detect a cyber intrusion in the AGC signal. In future work, we plan to test the effectiveness of the proposed approach on a large scale power network, including also voltage dynamics. Moreover, we aim to extend the framework to address a larger class of disturbances in a probabilistic fashion.

REFERENCES

[5] [1], Power System Analysis, Power System Laboratory, ETH Zurich.