

New Hypothesis Testing-Based Methods for Fault Detection for Smart Grid Systems

Qian He and Rick S. Blum
ECE Department, Lehigh University
19 Memorial Drive West, Bethlehem, PA 18015
{qih207; rblum}@lehigh.edu

Abstract—Fault detection plays an indispensable role in ensuring the security of smart grid systems. Based on the dynamics of the generators, we show the time evolution of the smart grid system can be modeled by a discrete-time linear state space model. We focus on faults that can be described by changes in system matrices of the state space model. Newly developed locally optimum tests are discussed and utilized to improve the performance for detecting small changes. Numerical results are provided which demonstrate the superiority of the new approaches when compared to existing methods, especially in the detection of small changes.

Index Terms—Fault detection, hypothesis testing, smart grid, state space model.

I. INTRODUCTION

The vulnerability of smart grid systems is a growing concern. To demonstrate the validity of this concern, we tested the IEEE 9-bus power system in Fig. 1 using the MATPOWER simulation package under the default settings specified by the IEEE document, and we obtain a set of solutions which include the amount of power flow at each branch. Suppose the message sent from bus No. 9 to the control center (i.e., SCADA) containing the load information is intercepted and tampered with by a malicious attacker, who induces a false change in the demanded power at bus No. 9 from the original 125 MW to 560 MW while the overall power demand of this system remains under its maximum generation capacity. We again simulate this new 9-bus system and we find that as a result (see Fig. 2), significant power overflows will occur at branches a, b, and i. These overflows would likely burn out lines and damage equipment. Thus the health of this power system is in great jeopardy due to this malicious data attack.

Copyright 2011 IEEE. Published in the 45th Conference on Information Sciences and Systems (CISS) in Mar 2011 in Baltimore, MD, USA. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works, must be obtained from the IEEE. Contact: Manager, Copyrights and Permissions / IEEE Service Center / 445 Hoes Lane / P.O. Box 1331 / Piscataway, NJ 08855-1331, USA. Telephone: + Intl. 908-562-3966.

This material is based on research supported by the U.S. Army Research Office under grant No. W911NF-08-1-0449, the Air Force Research Laboratory under agreement No. FA9550-09-1-0576, by the National Science Foundation under Grant No. CCR-0112501, and by the Fundamental Research Funds for the Central Universities under grant No. ZYGX2009J019. Q. He is also with EE Department, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731 China.

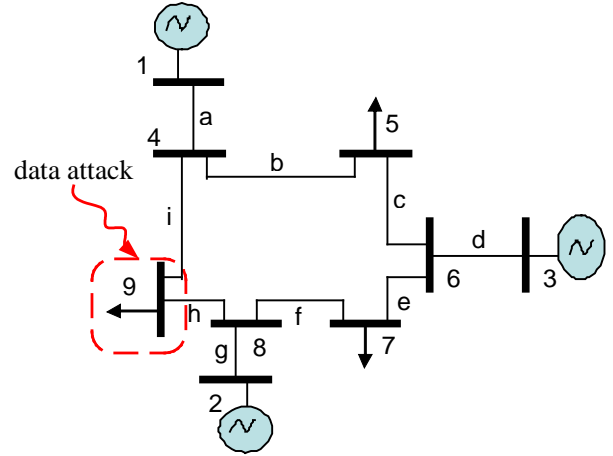


Fig. 1. IEEE 9-bus system with data attack on bus No. 9.

Based on this example, it is clear that the ability to recognize even a single intrusion is very desirable.

The use of a state space model to describe the dynamics of a power system has been justified in the past [1], [2]. In order to allow this paper to be self contained, we will also provide explicit justification in Section II of this paper. Here we propose to recognize changes in a power system by employing hypothesis testing to detect changes in the state space model for the system, a topic which has received little study to date in the power systems community. The time for such studies appears right since phasor measurement devices [5], which allow simultaneous measurement of both phase and magnitude of any voltage or current, are now available to provide significant gains in monitoring capability. It is worth noting that the theory discussed in this paper is directly applicable to detecting changes in systems for other applications beyond power systems and these ideas fall in the general rapidly developing area called cybersystems.

In the course of our study of employing hypothesis testing to detect changes in power systems, as our results will indicate, we found that if the change in the matrices describing the power system are very large, then estimating the change and employing the estimate in the resulting hypothesis test tends to work well. This is the so-called generalized likelihood ratio (GLR) test. On the other hand this approach may not work well when the changes are small. Here we developed some

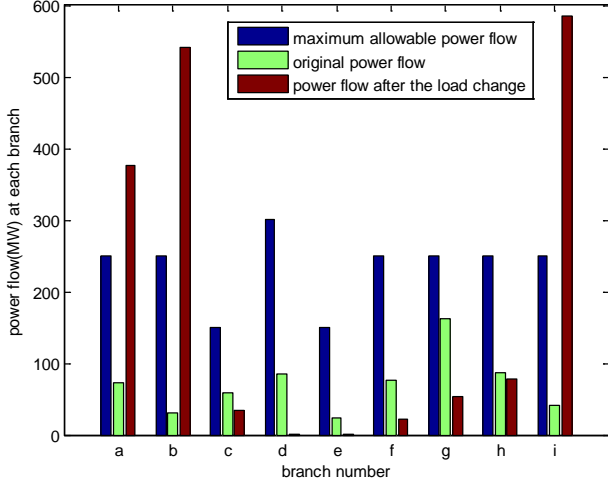


Fig. 2. Power flow at each branch before and after attack.

new locally optimum (LO) tests which tend to outperform the GLR test for cases with small changes. Previous study of LO tests were limited to cases where only a scalar parameter is different under the different hypotheses. In fact, the general results apply only for positive scalar parameters or for other very constrained cases. Here we develop new LO tests for cases with non-scalar parameters which need not be positive.

II. STATE SPACE MODEL FOR SMART GRID SYSTEMS

In this section, we show that a smart grid system can be modeled using a linear state space model. Consider a smart grid system consisting of n generators. The mechanical and electrical dynamics of the i th ($i = 1, \dots, \mathcal{M}$) generator can be described by [1]

$$\frac{d}{dt}\delta_i(t) = \omega_i(t) \quad (1)$$

$$\frac{d}{dt}\omega_i(t) = -\frac{D_i}{J_i}\omega_i(t) + \frac{1}{J_i}[P_{mi}(t) - P_{ei}(t)] \quad (2)$$

and

$$\frac{d}{dt}E'_{qi}(t) = \frac{1}{T'_{doi}}[E_{fi}(t) - E_{qi}(t)] \quad (3)$$

where $\delta_i(t)$ denotes the power angle of the i th generator, $\omega_i(t)$ the angular velocity of the i th generator, D_i in (2) the damping constant of the i th generator, and J_i the rotor inertia of the i th generator. In (2), $P_{mi}(t)$ denotes the mechanical input power of the i th generator and

$$P_{ei}(t) = I_{qi}(t)E'_{qi}(t) = \sum_{j=1}^{\mathcal{M}} E'_{qj}(t)[B_{ij} \sin(\delta_i(t) - \delta_j(t)) + G_{ij} \cos(\delta_i(t) - \delta_j(t))]E'_{qi}(t) \quad (4)$$

denotes the active power delivered by the i th generator, where $G_{ij} + jB_{ij}$ represents the ij th element of the nodal transient admittance matrix after eliminating all physical busbars, $I_{qi}(t)$ represents the quadrature axis stator currents and $E'_{qi}(t)$ in

(3) denotes the transient electromotive force (EMF) in the quadrature axis for the i th generator. In (3), T'_{doi} denotes the transient time constant in open circuit for the i th generator, $E_{fi}(t)$ the equivalent EMF in the excitation coil of the i th generator, and

$$E_{qi}(t) = E'_{qi}(t) + (x_{di} - x'_{di})I_{di}(t) \quad (5)$$

denotes the EMF in the quadrature axis for the i th generator, where x_{di} and x'_{di} represent the direct axis reactance and the direct axis transient reactance, $I_{di}(t)$ the direct axis stator currents of the i th generator, and

$$Q_{ei}(t) = I_{di}(t)E'_{qi}(t) = \sum_{j=1}^{\mathcal{M}} E'_{qj}(t)[G_{ij} \sin(\delta_i(t) - \delta_j(t)) - B_{ij} \cos(\delta_i(t) - \delta_j(t))]E'_{qi}(t) \quad (6)$$

represents the reactive power of the i th generator.

Let the variables on the left hand side of (1)-(3) be the continuous-time state variables. The state vector of the i th generator can be written as

$$\mathbf{x}_i(t) = \begin{bmatrix} \frac{d}{dt}\delta_i(t) \\ \frac{d}{dt}\omega_i(t) \\ \frac{d}{dt}E'_{qi}(t) \end{bmatrix} \equiv \begin{bmatrix} x_{1i}(t) \\ x_{2i}(t) \\ x_{3i}(t) \end{bmatrix}. \quad (7)$$

From (1)-(3), the time derivatives of the state variables are

$$\frac{d}{dt}\mathbf{x}_i(t) = \begin{bmatrix} -\frac{D_i}{J_i}\frac{d}{dt}\omega_i(t) + \frac{1}{J_i}\left[\frac{d}{dt}P_{mi}(t) - \frac{d}{dt}P_{ei}(t)\right] \\ \frac{1}{T'_{doi}}\left[\frac{d}{dt}E_{fi}(t) - \frac{d}{dt}E_{qi}(t)\right] \end{bmatrix}. \quad (8)$$

Let the input vector to the system be

$$\mathbf{u}_i(t) = \begin{bmatrix} \frac{d}{dt}E_{fi}(t) \\ \frac{d}{dt}I_{fi}(t) \end{bmatrix} \equiv \begin{bmatrix} u_{1i}(t) \\ u_{2i}(t) \end{bmatrix}. \quad (9)$$

Assume the mechanical torque is constant, so that $P_{mi}(t) = P_{mi}$ does not vary with time. Then in the second component of the vector in (8), $\frac{d}{dt}P_{mi}(t)$, becomes zero and can be dropped. Plugging (4)-(7) into (8), after simplification, approximations, and manipulation [1], we obtain the continuous-time noise-free state equation

$$\frac{d}{dt}\mathbf{x}_i(t) = \mathbf{A}_i(t)\mathbf{x}_i(t) + \mathbf{B}_i\mathbf{u}_i(t) + \mathbf{g}_i(t) \quad (10)$$

where

$$\mathbf{A}_i(t) = \begin{bmatrix} 0 & 1 & 0 \\ p_{1i}(t) & p_{2i} & p_{3i}(t) \\ p_{4i}(t) & 0 & p_{5i}(t) \end{bmatrix} \quad (11)$$

with

$$p_{1i}(t) = \frac{1}{J_i} \sum_{j=1, j \neq i}^{\mathcal{M}} E'_{qi}(t)E'_{qj}(t)[G_{ij} \sin(\delta_{ij0}) - B_{ij} \cos(\delta_{ij0})]$$

where δ_{ij0} is a constant approximation of $(\delta_i(t) - \delta_j(t))$ [1],

$$p_{2i}(t) = -\frac{D_i}{J_i}$$

$$p_{3i}(t) = -\frac{2G_{ii}E'_{qi}(t)}{J_i} - \frac{1}{J_i} \sum_{j=1, j \neq i}^M E'_{qj}(t) \times [B_{ij} \sin(\delta_{ijo}) + G_{ij} \cos(\delta_{ijo})]$$

$$p_{4i}(t) = -\frac{(x_{di} - x'_{di})}{T'_{doi}} \sum_{j=1, j \neq i}^M E'_{qj}(t) \times [B_{ij} \sin(\delta_{ijo}) + G_{ij} \cos(\delta_{ijo})]$$

and

$$p_{5i}(t) = -\frac{1}{T'_{doi}} + \frac{(x_{di} - x'_{di})}{T'_{doi}} B_{ii}.$$

The \mathbf{B}_i in (10) is

$$\mathbf{B}_i = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \frac{1}{T'_{doi}} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ p_{6i} & 0 \end{bmatrix} \quad (12)$$

where

$$p_{6i} = \frac{1}{T'_{doi}}.$$

The $\mathbf{g}_i(t)$ in (10) is

$$\mathbf{g}_i(t) = \sum_{j=1, j \neq i}^M \mathbf{R}_{ij}(t) \mathbf{x}_j(t) \equiv \begin{bmatrix} 0 \\ g_{1i}(t) \\ g_{2i}(t) \end{bmatrix} \quad (13)$$

where

$$\mathbf{R}_{ij}(t) = \begin{bmatrix} 0 & 0 & 0 \\ r_{1ij}(t) & 0 & r_{2ij}(t) \\ r_{3ij}(t) & 0 & r_{4ij}(t) \end{bmatrix} \quad (14)$$

with

$$r_{1ij}(t) = -\frac{1}{J_i} E'_{qi}(t) E'_{qj}(t) [G_{ij} \sin(\delta_{ijo}) - B_{ij} \cos(\delta_{ijo})]$$

$$r_{2ij}(t) = -\frac{1}{J_i} E'_{qi}(t) [B_{ij} \sin(\delta_{ijo}) + G_{ij} \cos(\delta_{ijo})]$$

$$r_{3ij}(t) = \frac{(x_{di} - x'_{di})}{T'_{doi}} E'_{qj}(t) [B_{ij} \sin(\delta_{ijo}) + G_{ij} \cos(\delta_{ijo})]$$

and

$$r_{4ij}(t) = -\frac{(x_{di} - x'_{di})}{T'_{doi}} [G_{ij} \sin(\delta_{ijo}) - B_{ij} \cos(\delta_{ijo})].$$

The state equation in (10) can be further simplified using similar approximations [1], and then converted to a discrete-time version using standard manipulations to get

$$\mathbf{x}_{k+1} = \mathbf{F}(k+1) \mathbf{x}_k + \mathbf{G} \mathbf{u}_k + \mathbf{w}_k.$$

The output equation can be obtained in a similar way as

$$\mathbf{z}_k = \mathbf{H} \mathbf{x}_k + \mathbf{J} \mathbf{u}_k + \mathbf{e}_k.$$

III. DISCRETE-TIME LINEAR STATE SPACE MODEL

As we have just shown, a smart grid system can be modeled using a linear state space model using approximations frequently made in the power system literature [1], [2]. Let \mathbf{x}_k be an N dimensional vector which denotes the state of the system at time k [1], [2]. Denote the available measurements by the output vector \mathbf{z}_k . The state and output time evolutions are

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{F}(k+1) \mathbf{x}_k + \mathbf{G} \mathbf{u}_k + \mathbf{w}_k \\ \mathbf{z}_k &= \mathbf{H} \mathbf{x}_k + \mathbf{J} \mathbf{u}_k + \mathbf{e}_k \end{aligned} \quad (15)$$

where \mathbf{u}_k denotes the Gaussian or deterministic input to the system, \mathbf{w}_k the zero mean Gaussian state disturbance, and \mathbf{e}_k the zero mean Gaussian observation noise, which we assume are each independent and identically distributed (iid) sequences mutually independent of each other. We further assume the initial state \mathbf{x}_0 is deterministic or Gaussian, so \mathbf{x}_k and \mathbf{z}_k are Gaussian sequences. Let Σ_u , Σ_w and Σ_e denote the covariance matrices of \mathbf{u}_k , \mathbf{w}_k and \mathbf{e}_k , respectively. Methods for finding the parameters of such models ($\mathbf{F}(k+1)$, \mathbf{G} , ...) have been extensively studied, see for example [6].

Here we propose to recognize a failure or intrusion by recognizing a change in the system. If we want to detect a change in the internals of the system, not the input, then a focus on a change in $\mathbf{F}(k+1)$ appears very relevant. Since it may be impossible to detect the difference between a change in \mathbf{G} and a change in the input, we do not focus on changes in \mathbf{G} . A change in \mathbf{H} implies some change in the part of the system that couples the state to the output. This is interesting and could be handled in a similar way. Here we focus on changes in $\mathbf{F}(k+1)$ for simplicity. Note that in (15) we use the notation $\mathbf{F}(k+1)$ to remind the reader that $\mathbf{F}(k+1)$ in (15) will impact the state and output at the next time $k+1$. In normal operation the matrix $\mathbf{F}(k+1)$ will be constant, but if a failure or intrusion occurs it can change at a given time, say $k+1 = n$. We want to recognize this change.

IV. HYPOTHESIS TESTING

One way to detect changes in such a system is to continuously estimate the parameters ($\mathbf{F}(k+1)$, \mathbf{G} , ...) or some related quantities (for example spectral characteristic, [7]). One can also attack the problem more directly by employing hypothesis testing. For example, if we believe the system is operating for a long time with $\mathbf{F}(k+1) = \mathbf{F}_o$ and desire to judge if a change occurred at time $k+1 = n$ we can formulate the hypothesis testing problem as¹

H_0 : we observe $\mathbf{z} = (\mathbf{z}_1, \dots, \mathbf{z}_n)^T$ from (15) with

$$\mathbf{F}(k+1) = \mathbf{F}_o \text{ for } k = 0, \dots, n-1$$

H_1 : we observe $\mathbf{z} = (\mathbf{z}_1, \dots, \mathbf{z}_n)^T$ from (15) with

$$\mathbf{F}(k+1) = \begin{cases} \mathbf{F}_o & \text{for } k = 0, \dots, n-2 \\ \mathbf{F}_c \neq \mathbf{F}_o & \text{for } k = n-1 \end{cases} \quad (16)$$

¹More general problems where the change might occur over a window of possible times will be discussed in future work.

for a known n . If \mathbf{F}_o and \mathbf{F}_c are both known², then the Neyman-Pearson optimum test, maximizing detection probability for fixed false alarm probability, compares the likelihood ratio (LR) of the observations $\mathbf{z} = (\mathbf{z}_1, \dots, \mathbf{z}_n)^T$ to a threshold [8] where the LR is

$$\begin{aligned} & \frac{p_{\mathbf{z}}(\mathbf{z}|H_1)}{p_{\mathbf{z}}(\mathbf{z}|H_0)} \\ &= \frac{p(\mathbf{z}_n|\mathbf{z}_{n-1}; \mathbf{F}_c)p(\mathbf{z}_{n-1}|\mathbf{z}_{n-2}; \mathbf{F}_o) \cdots p(\mathbf{z}_1|\mathbf{z}_0; \mathbf{F}_o)p(\mathbf{z}_0; \mathbf{F}_o)}{p(\mathbf{z}_n|\mathbf{z}_{n-1}; \mathbf{F}_o)p(\mathbf{z}_{n-1}|\mathbf{z}_{n-2}; \mathbf{F}_o) \cdots p(\mathbf{z}_1|\mathbf{z}_0; \mathbf{F}_o)p(\mathbf{z}_0; \mathbf{F}_o)} \\ &= \frac{p(\mathbf{z}_n|\mathbf{z}_{n-1}; \mathbf{F}_c)}{p(\mathbf{z}_n|\mathbf{z}_{n-1}; \mathbf{F}_o)}. \end{aligned}$$

Here $p_{\mathbf{z}}(\mathbf{z}|H_1)$ denotes the probability density function (pdf) of the observation vector \mathbf{z} and $p(\mathbf{z}_{k+1}|\mathbf{z}_k; \mathbf{F}(k+1))$ denotes the conditional pdf of \mathbf{z}_{k+1} given \mathbf{z}_k and $\mathbf{F}(k+1)$.

Under the assumptions made concerning the model in (1), the conditional pdf of \mathbf{z}_{k+1} given \mathbf{z}_k becomes

$$p(\mathbf{z}_{k+1}|\mathbf{z}_k; \mathbf{F}(k+1)) = \frac{e^{-\frac{1}{2}(\mathbf{z}_{k+1}-\boldsymbol{\mu}_{k+1})^T \boldsymbol{\Sigma}_{k+1}^{-1}(\mathbf{z}_{k+1}-\boldsymbol{\mu}_{k+1})}}{(2\pi)^{\frac{N}{2}} |\boldsymbol{\Sigma}_{k+1}|^{\frac{1}{2}}}$$

where

$$\begin{aligned} \boldsymbol{\mu}_{k+1} &= \mathbf{H}\mathbf{F}(k+1)\mathbf{H}^{-1}\mathbf{z}_k \\ &+ [\mathbf{H}\mathbf{G} - \mathbf{H}\mathbf{F}(k+1)\mathbf{H}^{-1}\mathbf{J} + \mathbf{J}]\mathbb{E}\{\mathbf{u}_k\} \end{aligned}$$

and

$$\begin{aligned} \boldsymbol{\Sigma}_{k+1} &= \mathbf{H}\mathbf{F}(k+1)\mathbf{H}^{-1}\boldsymbol{\Sigma}_e(\mathbf{H}\mathbf{F}(k+1)\mathbf{H}^{-1})^T \\ &+ [\mathbf{H}\mathbf{G} - \mathbf{H}\mathbf{F}(k+1)\mathbf{H}^{-1}\mathbf{J}]\boldsymbol{\Sigma}_u[\mathbf{H}\mathbf{G} - \mathbf{H}\mathbf{F}(k+1)\mathbf{H}^{-1}\mathbf{J}]^T \\ &+ \mathbf{J}\boldsymbol{\Sigma}_u\mathbf{J}^T + \mathbf{H}\boldsymbol{\Sigma}_w\mathbf{H}^T + \boldsymbol{\Sigma}_e. \end{aligned}$$

The case where \mathbf{F}_c is unknown is generally more complicated. One approach is to choose the maximum likelihood estimate of \mathbf{F}_c , the GLR test [9], which works well if the estimate is very accurate. We give some alternative approaches later in his paper where we first develop an LO test for a case with an unknown vector parameter. We have not seen LO tests for this case provided in the literature to date.

V. NEW LOCALLY OPTIMUM TESTS

A. LOUD Test and LOED Test

Now we develop an LO test for a general detection problem with a deterministic unknown N -dimensional vector parameter (like $\text{vec}(\mathbf{F}(n))$), which we call the LO unknown direction (LOUD) test. Consider the case where we observe a random vector \mathbf{y} whose pdf $p_{\mathbf{y}}(\mathbf{y}; \boldsymbol{\theta})$ is dependent on a deterministic parameter $\boldsymbol{\theta}$. Consider a binary hypothesis testing problem where $\boldsymbol{\theta}$ takes on values from two disjoint sets Θ_o and Θ_c , which distinguishes the two hypotheses

$$\begin{aligned} H_0: \mathbf{y} \text{ has pdf } p_{\mathbf{y}}(\mathbf{y}; \boldsymbol{\theta}) \text{ with } \boldsymbol{\theta} \in \Theta_o \\ H_1: \mathbf{y} \text{ has pdf } p_{\mathbf{y}}(\mathbf{y}; \boldsymbol{\theta}) \text{ with } \boldsymbol{\theta} \in \Theta_c \end{aligned} \quad (17)$$

²We only consider $\mathbf{F}_o, \mathbf{F}_c$ that are full rank (invertible) so that we will not have a singular detection problem. Thus $\mathbf{F}_c\mathbf{x}_k \neq \mathbf{F}_o\mathbf{x}_k$. To avoid similar problems, we also assume \mathbf{H} is full rank (invertible).

We assume $\Theta_o = \{\boldsymbol{\theta}_o\}$ and Θ_c includes all $\boldsymbol{\theta}$ in a very small ball around $\boldsymbol{\theta}_o$ and we focus on obtaining good performance for all such $\boldsymbol{\theta}$ on average. By employing the N -dimensional spherical coordinate system, we write a length r vector in N -dimensional space as $r\boldsymbol{\Omega}$ with

$$\boldsymbol{\Omega} = \begin{pmatrix} \cos(\phi_1) \\ \sin(\phi_1)\cos(\phi_2) \\ \vdots \\ \sin(\phi_1)\sin(\phi_2)\cdots\cos(\phi_{N-1}) \\ \sin(\phi_1)\sin(\phi_2)\cdots\sin(\phi_{N-1}) \end{pmatrix} \quad (18)$$

where ϕ_i is in $[0, \pi]$ for $i < N-1$ or $[0, 2\pi]$ for $i = N-1$, such that $(\boldsymbol{\theta}_o + r\boldsymbol{\Omega})$ describes the points on the surface of an N -sphere centered at $\boldsymbol{\theta}_o$. This sphere has a radius r and a surface area of $D = 2\pi^{N/2}r^{N-1}/\Gamma(N/2)$. Denote $\boldsymbol{\Phi} = (\phi_1, \phi_2, \dots, \phi_{N-1})^T$ and $d\phi_1 d\phi_2 \cdots d\phi_{N-1} = d\boldsymbol{\Phi}$. We define the differential surface area element as

$$ds = r^{N-1} \sin^{N-2}(\phi_1) \sin^{N-3}(\phi_2) \cdots \sin(\phi_{N-2}) d\boldsymbol{\Phi}.$$

Let $P_d(\delta; \boldsymbol{\theta})$ denote the detection probability of the test with test function δ when the parameter is $\boldsymbol{\theta}$. We define the LOUD test as the test, with test function δ (which describes the conditional probability that we decide for H_1 conditioned on our observing \mathbf{y}), that maximizes

$$\frac{\partial^m}{\partial r^m} \left[\int_{\boldsymbol{\Phi}} P_d(\delta; \boldsymbol{\theta}_o + r\boldsymbol{\Omega}) g(\boldsymbol{\Phi}) d\boldsymbol{\Phi} \right] \Big|_{r=0} \quad (19)$$

where m is the lowest order derivative for which (19) is nonzero, and

$$g(\boldsymbol{\Phi}) \equiv \frac{r^{N-1} \sin^{N-2}(\phi_1) \sin^{N-3}(\phi_2) \cdots \sin(\phi_{N-2})}{D} \quad (20)$$

is the fractional differential area of a small surface element centered at $\boldsymbol{\theta}_o + r\boldsymbol{\Omega}$. The integrations span over the relevant angles so that $\boldsymbol{\theta}_o + r\boldsymbol{\Omega}$ ranges over all points in a radius r N -dimensional sphere around $\boldsymbol{\theta}_o$. The LOUD test criterion searches for a test that produces the largest average $P_d(\delta; \boldsymbol{\theta})$ when we average over all possible small changes in the parameter $\boldsymbol{\theta}$. This is a reasonable extension of the LO test which is relevant for cases when we have no information concerning the direction of change from $\boldsymbol{\theta}_o$ to $\boldsymbol{\theta}$. If some directions are known to be more likely, this is easy to incorporate.

Since a LOUD test maximizes the lowest order term in a Taylor series of the quantity differentiated in (19) then its test function δ will ensure

$$\int_{\boldsymbol{\Phi}} P_d(\delta; \boldsymbol{\theta}_o + r\boldsymbol{\Omega}) g(\boldsymbol{\Phi}) d\boldsymbol{\Phi} \geq \int_{\boldsymbol{\Phi}} P_d(\delta'; \boldsymbol{\theta}_o + r\boldsymbol{\Omega}) g(\boldsymbol{\Phi}) d\boldsymbol{\Phi}$$

when compared to any other test δ' for $0 < r < r_{\max}$ and some r_{\max} . More details about the LOUD test can be found in [10].

Instead of attempting to optimize performance averaged over all the possible directions, an alternative approach is to estimate the change direction from the observations and to use this estimate in the LO test. With this idea, we define the LO estimated direction (LOED) test as the test, with test

function δ , that maximizes $P_d(\delta; \theta_c)$, among all possible tests, for changes $\theta_c - \theta_o = r\hat{\Omega}$ along the estimated direction $\hat{\Omega}$ for $0 < r < r_{\max}$ and some r_{\max} . See [10] for more details.

B. LOUD-GLR Test and LOED-GLR Test

The LOUD and LOED tests are designed for small changes $(\theta_c - \theta_o)$. On the other hand, if the change is sufficiently large such that the unknown parameter can be estimated accurately enough, the performance of the GLR test is known to be good. Accordingly, we propose to combine the LOUD (or LOED) test with the GLR test to obtain a new test called the LOUD-GLR (or LOED-GLR test). The combined test employs the LOUD (or LOED) test when the change looks small, and switches to the GLR test when the change looks large. An intuitive way to assess the size of the change is to measure the distance between the observed \mathbf{y} and the mean of the observation under H_0 hypothesis, $\mu_0 = E\{\mathbf{y}; \theta_o\}$. Thus, we compute $\epsilon = \sqrt{(\mathbf{y} - \mu_0)^T \Sigma_0^{-1} (\mathbf{y} - \mu_0)}$, where $\Sigma_0 = E\{(\mathbf{y} - \mu_0)(\mathbf{y} - \mu_0)^T; \theta_o\}$. If ϵ is small then this observation \mathbf{y} would occur under H_0 with high probability, so we select the LOUD or LOED test. If ϵ is large we select the GLR test. The switch between these two tests can be done gradually or abruptly.

For simplicity, we present an example for a case where \mathbf{y} is a scalar. Let the test statistic of the combined tests be

$$T_{\text{LO-GLR}}(y) = (1 - \lambda)T_{\text{LO}}(y) + \lambda\hat{T}_{\text{GLR}}(y) \quad (21)$$

where

$$\begin{aligned} \hat{T}_{\text{GLR}}(y) = & [T_{\text{GLR}}(y) - \tau_{s_1}]\nu(y - \mu_0 - \Delta) \\ & + [T_{\text{GLR}}(y) - \tau_{s_2}]\nu(\mu_0 - y - \Delta) \end{aligned}$$

with $\nu(t) = 1$ if $t \geq 0$ or 0 if $t < 0$. In (21), τ_{s_1} and τ_{s_2} are employed to make the test statistic in (21) continuous in y , and $T_{\text{LO}}(y)$ and $T_{\text{GLR}}(y)$ denote the test statistics of the LO and the GLR tests, where the subscript ‘‘LO’’ represents LOUD or LOED. The factor $0 \leq \lambda \leq 1$ will switch gradually between the test statistics used in (21). For example, we can choose

$$\lambda = \mathcal{P}(\xi, (\epsilon - \Delta)\nu(\epsilon - \Delta)/s)$$

where $\mathcal{P}(a, b) = \int_0^b t^{a-1} e^{-t} dt / (\int_0^\infty t^{a-1} e^{-t} dt)$ denotes the regularized Gamma function, $\xi > 0$ and $s > 0$ are the shape and scale parameters that set the range where we have transition, and the predetermined scalar value Δ indicates a distance between y and μ_0 within which the LO test will be used. A more general transform (even nonlinear) on $T_{\text{GLR}}(\mathbf{y})$ could be employed in non-scalar cases. We now discuss one way to choose Δ . Let y_c be the nearest (with respect to μ_0) critical point of the LOUD or LOED test statistic which satisfies $\frac{\partial T_{\text{LO}}(y)}{\partial y}|_{y=y_c} = 0$ and $\frac{\partial^2 T_{\text{LO}}(y)}{\partial y^2}|_{y=y_c} < 0$. Then $\Delta \equiv |y_c - \mu_0|$.

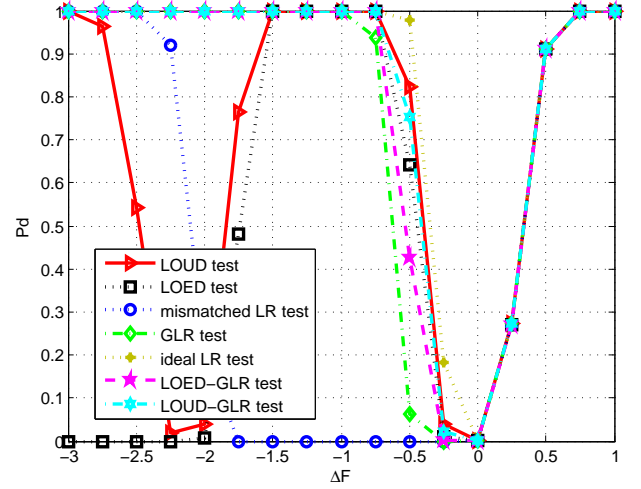


Fig. 3. P_d versus $r\Omega$ curves for various tests for $P_f = 10^{-3}$.

VI. NUMERICAL RESULTS

In this section, we apply the proposed LO tests to a simplified version of the model in (15) and illustrate the change detection performance of the new tests and compare them with the conventional approaches via numerical investigations. Here, for simplicity, we consider the case where the variables and parameters involved are scalars. Assume $\Sigma_w = 0.001$, $\Sigma_e = 0.005$, $H = 1$ and $G = 0$ and the hypothesis testing problem in (16). In particular, we assume that prior to times $k + 1 = n$ we have $F(k + 1) = 1$. However, if an abnormal event occurs then $F(n) = F_c = F_o + \Delta F = 1 + r\Omega$, where $r = |\Delta F|$ and Ω , indicating the change direction, is either -1 or 1 . Of course if the abnormal event does not occur then $F(n) = 1$. To employ the LO tests we denote $\theta = \text{vec}(F(n)) = F(n)$ and $p_y(y; \theta_o) = p_{z_n|z_{n-1}}(z_n|z_{n-1}; F_o)$. The false alarm probability is fixed at $P_f = 10^{-3}$. The simulation results are obtained using 5000 Monte Carlo runs. The LOUD, LOED, mismatched LR, GLR, ideal LR, LOED-GLR, and LOUD-GLR tests are considered, where the mismatched LR detector assumes $F(n) = F_c = 1.5$ ($\Delta F = 0.5$) and is always matched to this value no matter what the actual change is. Alternatively, the ideal LR detector is perfectly matched to the actual F_c (thus the actual ΔF). For the LOUD-GLR or LOED-GLR test, we use the test statistic

$$T_{\text{LO-GLR}}(y) = \begin{cases} T_{\text{LO}}(y), & \text{if } |y - \mu_0| \leq \Delta \\ \hat{T}_{\text{GLR}}(y), & \text{if } |y - \mu_0| > \Delta \end{cases}$$

to switch between the two tests abruptly, where $T_{\text{LO}}(y)$ and $\hat{T}_{\text{GLR}}(y)$ are defined after (21), $\Delta = |y_c - \mu_0|$ and y_c is the nearest (with respect to μ_0) critical point of the LOUD or LOED test.

In Fig. 3, the detection probability P_d is plotted versus ΔF for various tests, assuming the last observation $z_{n-1} = 1$. It is observed that when the change is positive, all the curves are almost on top of each other, indicating similar performances. When the change becomes negative, these tests exhibit very different behaviors. We see from the curves that the ideal

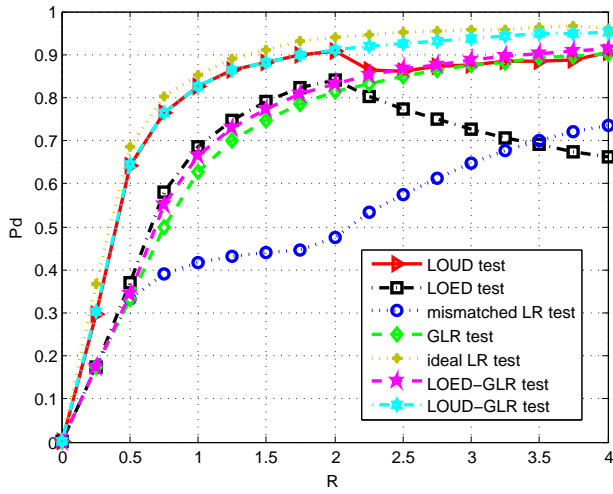


Fig. 4. P_d versus R curves for various tests for $P_f = 10^{-3}$.

LR test (a genie type approach which perfectly sees the unknown change) has the best performance, which provides a benchmark for evaluating different tests. The mismatched LR is significantly affected by the change directions when the amplitude of the change is small. The LOUD and LOED tests outperform the GLR test for small changes, as expected, but their performance is not acceptable in some regime for negative large changes since they are designed only for small changes. The performance of the LOUD-GLR and LOED-GLR tests are close to those of the LOUD and LOED tests when the change is small, and are good as well when the change is large due to the switching to the GLR test.

In the second example, we look at the case where ΔF is uniformly distributed in $[-R, R]$. Assume the last observation $z_{n-1} = 2$. We plot P_d versus R in Fig. 4 for the tests considered in the previous example. It is seen that the LOUD and LOED tests have superior performance compared with the traditional GLR test or mismatched LR test for small changes. The performance the LOUD and LOED test degrade when R becomes large. The LOUD-GLR and LOED-GLR have similar performance as the LOUD and LOED for small R , and maintain a good performance for large R . Again, these results illustrate the superiority of the proposed LO tests for small changes. They also show that the LOUD-GLR and LOED-GLR tests are preferable, since they perform nearly as good as the LOUD and LOED tests when the change is small and also have satisfactory performance when the change is large. We studied many other examples and obtain a similar conclusion.

VII. CONCLUSIONS

The importance of failure and intrusion detection of smart grid systems was illustrated. We showed that a smart grid system can be modeled by a linear state space model using some frequently made approximations. We introduced two new locally optimum tests, the LOUD test and the LOED test. Then we discussed combining the LO tests with the GLR test. It was shown that the new LO tests are favorable for detecting

small changes with unknown directions, while the LOUD-GLR and LOED-GLR tests have satisfactory performance for the detection of both small and large changes.

VIII. ACKNOWLEDGEMENT

The authors would like to thank Yang Yang and Liang Zhao for providing the motivating simulation result given in the Introduction section. We would also thank Prof. Mooi Choo Chuah from Lehigh University for useful discussions on this same topic.

REFERENCES

- [1] M. Dehghani and S.K.Y. Nikravesh, "State-space model parameter identification in large-scale power systems," in *IEEE Trans. on Power Systems*, vol. 23, no. 3, pp. 1449-1457, 2008.
- [2] Y. Yu, *Electric Power System Dynamics*. New York: Academic, 1983.
- [3] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*, Marcel Dekker, 2004.
- [4] A. Jain and N. R. Shivakumar, "Power system tracking and dynamic state estimation," *IEEE PSCE Conf.*, 2009.
- [5] H. Xua, Q. Jia, et al. "A dynamic state estimation method with PMU and SCADA measurement for power systems," in *Proc. of IPEC 2007*, pp. 848-853, 2007.
- [6] L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification*. MIT Press, Cambridge, MA, 1983.
- [7] M. Basseville and I.V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Information and system science series. Prentice Hall, NJ, 1993.
- [8] H. Vincent Poor, *An Introduction to Signal Detection and Estimation*. Springer-Verlag, New York, 1988.
- [9] P. J. Bickel and K. A. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*, Holden-Day, CA, 1977.
- [10] Q. He and R. S. Blum, "Smart grid monitoring for intrusion and fault detection with new locally optimum testing procedures," in *Proc. of the 36th International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011.
- [11] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*. Springer Verlag, New York, 1988.