From Static to Dynamic Anomaly Detection with Application to Power System Cyber Security

KAIKAI PAN, PETER PALENSKY, AND PEYMAN MOHAJERIN ESFAHANI

Abstract. Developing advanced diagnosis tools to detect cyber attacks is the key to security of power systems. It has been shown that multivariate data injection attacks can bypass bad data detection schemes typically built on static behavior of the systems, which misleads operators to disruptive decisions. In this article, we depart from the existing static viewpoint to develop a diagnosis filter that captures the dynamics signatures of such a multivariate intrusion. To this end, we introduce a dynamic residual generator approach formulated as robust optimization programs in order to detect a class of disruptive multivariate attacks that potentially remain stealthy in view of a static bad data detector. We investigate two possible desired features: (i) a non-zero transient and (ii) a non-zero steady-state behavior of the residual generator in the presence of an attack. In case (i), the problem is reformulated as a finite, but possibly non-convex, optimization program. We further develop a linear programming relaxation that improves the scalability, and as such practicality, of the diagnosis filter design. In case (ii), it turns out that the resulting robust program admits an exact convex reformulation, yielding a Nash equilibrium between the attacker and the residual generator. This assertion has an interesting implication: the proposed approach is not conservative in the sense that the additional knowledge of the worst-case attack does not improve the diagnosis performance. To illustrate our theoretical results, we implement the proposed diagnosis filter to detect multivariate attacks on the system measurements deployed to generate the so-called Automatic Generation Control signals in a three-area IEEE 39-bus system.

1. Introduction

The digital transformation of our power system does not only lead to better observability, flexibility and efficiency, but also introduces a phenomenon that is new to power system controls: cyber security threats. NIST [7] defines five functions for protecting Information and Communication Technology (ICT): (i) Identify, (ii) Protect, (iii) Detect, (iv) Respond, (v) Recover. It would be naive to think an ICT system can be perfectly protected in order to address the issues raised by (iii)-(v). This paper focuses on (iii) Detection for supervisory control and data acquisition (SCADA) systems, which are in charge of transmitting measurement and control signals between power system substations and control centers. Such SCADA systems are notorious for being based on legacy ICT, and are a popular target for adversaries [13, 6] nowadays. The consequences of a successful attack on SCADA systems can be catastrophic to an economy and society in general [24, 17]. In this light, it is of utmost importance to detect these attacks and respond accordingly. Notably, if the malicious attacks can be detected sufficiently fast, the corrupted signals can be disconnected or corrected by resilient controls, preventing further severe damage [34].

Literature on anomaly detection. Traditionally, SCADA systems deploy bad data detection (BDD) to filter out possible erroneous measurements due to sensor failures or anomalies [33]. The BDD process captures only a snapshot of the steady states of system trajectories, and thus only exploits possible *static* impact of

Date: September 22, 2019.

The authors are with the Delft University of Technology, The Netherlands (email: {K.Pan, P.Palensky, P.MohajerinEsfahani}@tudelft.nl).

intrusions. Although this method can perform successfully in detecting basic attacks, it may fail in the presence of the so-called *stealthy multivariate attacks* that carefully launch synthesized false data injections given full knowledge of the system model [15].

It was first explored in [20] that such an attack can perturb the state estimation function without triggering alarms in BDD. Since then vulnerability and impact analysis of stealthy attacks on power systems have been a prominent subject in the literature. A typical notion to quantify the vulnerability to stealthy attacks is directly concerned with the level of efforts required to alter specific measurements [12, 27]. Without advanced diagnosis tools, tampering measurements remains undetected, causing state deviations, equipment damages or even cascading failures [18]. Techniques proposed to deal with stealthy attacks include statistical methods such as sequential detection using Cumulative Sum (CUSUM)-type algorithms [16], and measurements consistency assessment under certain observability assumptions [35]. A detection method that leverages online information is described in [3], which is applicable by ensuring the availability and accuracy of load forecasts and generation schedules. In [19], a mechanism is introduced to formulate the detection scheme as a matrix separation problem, but it only recovers intrusions among corrupted measurements over a particular period of time.

These techniques are essentially static detection methods that may be confined by certain prior assumptions on the distribution of measurement errors. Despite an extensive and ongoing literature focusing on the static part of BDD mechanism, the following question remains largely unexplored:

Would it be possible to detect stealthy multivariate attacks in a real-time operation by exploiting the attack impact on the dynamics of system trajectories during the transient?

The importance of an appropriate answer to this question has been reinforced thanks to recent advances in sensing technology in the modern power systems. Our main objective in this article is to address this question.

Related work. Detection methods concerning system dynamics have primarily emerged under the topic of *fault detection and isolation filters*. A subclass of these schemes is the observer-based approach applied initially to linear models [21]; see also [9] for a comprehensive summary of the large body of literature. The authors in [25] further extend the modeling framework to general linear differential-algebraic equations (DAEs), enhancing the applicability of such methods particularly for power system applications due to the common governing physical laws in this setting. Recently, a variant of observer-based methods is also investigated in [1] so as to deal with unknown natural exogenous inputs.

An inherent shortcoming of many observer-based approaches is that the degree of the resulting diagnosis filter is effectively the same as the system dynamics, which may yield an unnecessarily complex filter in large-scale power systems. To our best of knowledge, there are relatively much fewer studies in the literature on the design of the reduced-order observers where the conditions for a minimum order existence need to be satisfied [9, 10]. The closest approach in the literature is [23] where a scalable optimization-based filter design is developed for high-dimensional nonlinear control systems. However, the proposed method opts for mainly dealing with a single fault scenario, and may not be as effective in case of smart multivariate adversarial inputs.

An effective approach toward security and modeling the interaction between attackers and detectors builds on the rich framework of game theory. Recently, the authors in [32] propose a two player mixed strategy game to address a dynamic resource-planning problem between an attacker targeting the communication equipment and a defender protecting the control network. Similar frameworks have also been deployed to model the dynamics of information flow between an advanced persistent threat and a detector [30, 29].

Our contributions. The main objective of this article is to develop a *diagnosis filter* to detect *multivariate data injection attacks* in a real-time operation. For this purpose, considering a class of disruptive multivariate attack scenarios (Definition 2.5), we first characterize the attack impact on power system dynamics through a set of differential equations. Having transferred the dynamics into the discrete-time domain, we further restrict the diagnosis filter to a family of dynamic residual generators that entirely decouples the contributions of the attacks from the system states and natural disturbances. In order to identify an admissible multivariate attack scenario, we propose an optimization-based framework to robustify the diagnosis filter with respect to such attacks, i.e., aiming to design a filter whose residual (output) is sensitive to any plausible disruptive multivariate attacks. The main contributions of this article are as follows:

(i) Unlike the existing literature, we go beyond a static viewpoint of anomaly detection to capture the attack impact on the dynamics of system trajectories. To this end, we characterize the diagnosis filter design approach as a robust optimization program. It is guaranteed that while the filter residual is decoupled from system states and disturbances, it still remains sensitive to all admissible disruptive multivariate attacks even if the attacker has full knowledge about the diagnosis filter architecture (Definition 4.1 and the program (18)).

(ii) To detect attacks during the transient behavior, we reformulate the resulting robust program as a finite, possibly non-convex, optimization program (Theorem 4.3). To improve the scalability of the proposed solution, we further propose a linear programming relaxation which is highly tractable for large scale systems (Corollary 4.4). It is guaranteed that if the optimal value of the relaxed program is positive, the resulting diagnosis filter is able to detect any admissible disruptive attack scenarios, which may remain stealthy through the lens of a static detector.

(iii) We further explore the steady-state behavior of the diagnosis filter in the presence of a plausible attack scenario (Lemma 4.6). In this case, we develop an exact convex reformulation of the resulting robust program. As a byproduct, we show that the proposed solution is indeed a Nash equilibrium (saddle point) between the attacker and the residual generator (Theorem 4.7). An interesting implication of such a Nash equilibrium is that the information of the attack signal may not necessarily improve the performance of the diagnosis filter. In other words, if the proposed convex optimization fails to have a desirable feasible solution, it then implies that there exists a disruptive stealthy attack where the exact knowledge of the attack signal still does not help design a successful residual generator.

In addition to the above theoretical results, we validate the performance and effectiveness of the proposed diagnosis filter on a multi-area IEEE 39-bus system. Numerical results illustrate that the diagnosis filter successfully generates a residual "alert" in the presence of multivariate attacks that are stealthy in a static viewpoint, even in a noisy environment with imprecise measurements.

Section 2 introduces the problem of power system cyber security, and the challenges posed by multivariate attacks are highlighted. Section 3 discusses a model instance of power system dynamics under attacks on measurements. Our diagnosis filter design is proposed in Section 4 where an optimization framework is introduced, and numerical simulations are reported in Section 5.

Notation. The symbols $\mathbb{R}$, $\mathbb{N}$ represent the set of real numbers and integers, respectively. Given a matrix $A \in \mathbb{R}^{m\times n}$, $A^\top$ denotes its transpose, and the space $\mathrm{Im}(A)$ represents its range space. Throughout the paper, the matrix I is the identity matrix with an appropriate dimension. Given a column vector $a \in \mathbb{R}^m$, $\mathrm{diag}(a)$ denotes an $m \times m$ diagonal matrix with the elements of vector a sitting on the main diagonal and the rest of the elements being zero. We also denote by $\mathrm{diag}[A_1,\ A_2,\ \ldots,\ A_k]$ a block matrix whose main diagonal elements are the matrices A_1, A_2, $\ldots$, A_k. Given a vector $a \in \mathbb{R}^m$, the associated $\ell_\infty-$norm is denoted by $\|a\|_\infty = \max_{i\leq m} |a_i|$.

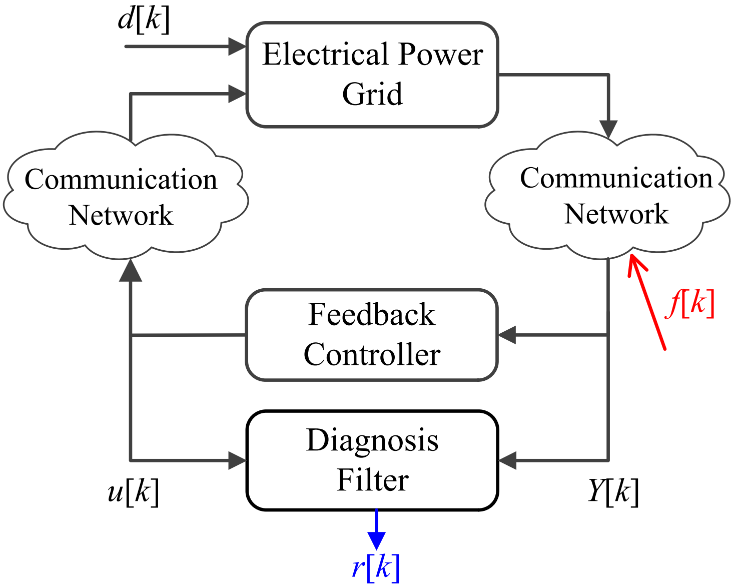


Figure 1. Schematic block diagram of the system model.

2. Problem Statement

2.1. Static detection and system modeling

For a power grid, measurements are collected by remote sensors and transmitted through a SCADA network. The typical BDD is conducted to detect the erroneous measurements at each time instance. We can see this as a static process: it only concerns the system states $X[k] \in \mathbb{R}^{n_X}$ and measurements $Y[k] \in \mathbb{R}^{n_Y}$ at time step $k \in \mathbb{N}$, which can be described by

$$Y[k] = CX[k] + D_f f[k], \tag{1}$$

where $C \in \mathbb{R}^{n_Y \times n_X}$ is the measurement matrix, and $f[\cdot] \in \mathbb{R}^{n_f}$ represents the data injection attacks on measurements. Note that the matrix D_f characterizes which measurement is vulnerable to attacks. It is customary to define a *residual signal* for a static detector, $r_S[k] := Y[k] - \hat{Y}[k]$, where $\hat{Y}[\cdot]$ denotes the estimated measurements. In the traditional weighted least squares estimation, the estimate of state is $(C^\top C)^{-1}C^\top Y[k]$, assuming that C has full column rank with high measurement redundancy. Then the measurements estimate is $C(C^\top C)^{-1}C^\top Y[k]$, and the residual signal can be further expressed as

$$r_S[k] = \big(I - C(C^\top C)^{-1}C^\top\big)Y[k]. \tag{2}$$

Such an anomaly detector has shown a good effectiveness in detecting erroneous data and basic attacks [8]. However, in the face of coordinated attacks on multiple measurements, this static detector can fail. In this article, motivated by this shortcoming, we take a dynamic design perspective where we shift the emphasis on an attack as a static process to its effects on power system dynamics. In particular, we opt for differentiating the attack impact on the systems trajectories from natural disturbances such as load deviations.

To model its impact on the dynamics, let us consider a more general modeling framework in Figure 1. The electrical grid is operated by a digital controller that receives measurements as inputs and sends control signals to the actuators through communication networks. These transmitted data are applied in discrete-time samples. On the power grid side, the input $d[k] \in \mathbb{R}^{n_d}$ represents natural disturbances. On the controller side, a control signal $u[k] \in \mathbb{R}^{n_u}$ is computed given the measurements $Y[k]$. Note that with the closed-loop control, the corruptions $f[k]$ on the measurements would affect the system dynamics. The dynamics of the closed-loop system is

$$\begin{cases} X[k+1] = A_x X[k] + B_d d[k] + B_u u[k], \\ Y[k] = CX[k] + D_f f[k], \end{cases} \tag{3}$$

where A_x, B_d and B_u are constant matrices. Let us highlight the difference between the dynamical system (3) and the respective static counterpart (1). In fact, the time independence of the first equation in (3) describes the dynamics of the system, while the algebraic equation (1) represents the relation on each time instance and describes a static relation between the states and outputs. The aim of this study is to exploit such dynamics information in (3) in order to design a diagnosis filter to detect stealthy multivariate attacks. To illustrate the attack impact on the system dynamics, we can simply consider the feedback controller as a linear operator such that $u[k] = GY[k]$ where $G \in \mathbb{R}^{n_u \times n_Y}$ is a matrix gain. By defining the closed-loop system matrices $A_{cl} := A + B_u GC$ and $B_f := B_u G D_f$, we can reformulate (3) into

$$\begin{cases} X[k+1] = A_{cl} X[k] + B_d d[k] + B_f f[k], \\ Y[k] = CX[k] + D_f f[k]. \end{cases} \tag{4}$$

Remark 2.1 (Dynamic feedback controller)**.** *The restriction to only a static feedback controller $u[k] = GY[k]$ to transfer from (3) to (4) is without loss of generality. Namely, the proposed framework is rich enough to subsume a dynamic controller architecture as well. Indeed, when the controller has certain dynamics, it suffices to augment the system dynamics (3) with the controller states and outputs. We refer to Appendix 2.1, for such a detailed analysis.*

Remark 2.2 (Attacks impact on the dynamics of system trajectories)**.** *In light of (4), matrices B_f, D_f capture the attack impact on the power system dynamics, mapping attacks $f[\cdot]$ to the system states and measurements respectively.*

In the following, we show that the state-space description (4) is a particular case of DAE model. By introducing a time-shift operator $q : qX[k] \to X[k+1]$, one can fit (4) into

$$H(q)x[k] + L(q)y[k] + F(q)f[k] = 0, \tag{5}$$

where $x := [X^\top \ d^\top]^\top$ represents the unknown signals of system states and disturbances; $y := Y$ contains all the available data for the operator. Let n_x and n_y be the dimensions of $x[\cdot]$, $y[\cdot]$. We denote n_r as the number of rows in (5). Then $H,\ L,\ F$ are polynomial matrices in terms of the time-shift operator q with n_r rows and n_x, n_y, n_f columns separately, by defining,

$$H(q) := \begin{bmatrix} -qI + A_{cl} & B_d \\ C & 0 \end{bmatrix}, \quad L(q) := \begin{bmatrix} 0 \\ -I \end{bmatrix}, \quad F(q) := \begin{bmatrix} B_f \\ D_f \end{bmatrix}.$$

2.2. Challenge: multivariate attacks

We start this subsection with an existing result characterizing the set of stealthy multivariate attacks that can bypass the static detector.

Lemma 2.3 (Stealthy attack values [20, Theorem 1])**.** *Consider the measurement equation (1) and the static detector with the respective residual function (2). Then, an attack $f[\cdot]$ remains stealthy, i.e., it does not cause any additional residue to (2), if it takes values from the set*

$$\mathcal{F} := \big\{ f[k] \in \mathbb{R}^{n_f} : \ D_f f[k] \in \mathit{Im}(C), \quad k \in \mathbb{N} \big\}, \tag{6}$$

One can observe that a stealthy attack $D_f f[\,\cdot\,]$ described in (6) has the knowledge of the system model (1) through the range space of C. That is, it represents a tampered value $D_f f[k] = C\Delta X$ where $\Delta X \in \mathbb{R}^{n_X}$ can be any injected bias influencing certain sensor measurements. Such multivariate attacks would also challenge the detector design as they may neutralize the diagnosis filter outputs.

Assumption 2.4 (Stationary attacks)**.** *Throughout this article, we consider attacks $f[\,\cdot\,]$ that are time-invariant, i.e., $f[k] = 0$ for all $k \leq k_{\min}$; $f[k] = f \in \mathcal{F}$ for all $k > k_{\min}$. Namely, the attack occurs as a constant bias injection f on measurements during the system operations at a specific unknown time instance $k_{\min}$, and it remains unchanged since then.*

Advanced attacks also pursue a maximized impact on the system dynamics. Thus, an adversary would try to inject "*smart*" false data, possibly with large magnitudes, in such a way that it causes the maximum damage. The next definition opts to formalize this class of attacks.

Definition 2.5 (Disruptive stealthy attack)**.** *Consider a set of vectors $F_{\mathrm{b}} := [f_1, f_2, \ldots, f_d]$ representing a finite basis for the set of stealthy attacks (6), i.e., the set $\mathcal{F}$ defined in (6) can equivalently be represented by*

$$\mathcal{F} = \left\{ F_{\mathrm{b}}^{\top}\alpha = \sum_{i=1}^{d} \alpha_i f_i \;\middle|\; \alpha = [\alpha_1, \alpha_2, \cdots, \alpha_d]^{\top} \in \mathbb{R}^d \right\}.$$

We call a signal $f \in \mathcal{F}$ disruptive stealthy attack *if its corresponding coefficients α is a polytopic set, i.e., it belongs to*

$$\mathcal{A} := \left\{\alpha \in \mathbb{R}^d \mid A\alpha \geq b\right\}, \tag{7}$$

where $A \in \mathbb{R}^{n_b \times d}$ and $b \in \mathbb{R}^{n_b}$ are given matrices. We emphasize that the subsequent analysis and the proposed diagnosis filter design only rely on the convexity of the set $\mathcal{A}$. Namely, the choice (7) may be adjusted according to the application at hand, as long as the convexity of the set is respected.

3. Cyber Security of Power Systems: AGC modeling

In this section, we first go through a modeling instance of power system dynamics in the form of (4): Automatic Generation Control (AGC) closed-loop system under attacks. This model will be used to validate our diagnosis filter. Figure 2 depicts the diagram of a three-area IEEE 39-bus system. AGC is a feedback controller that tunes the setpoints of participated generators (e.g., G11 of Area 1) to maintain the frequency as its nominal value and the tie-line (e.g., L1-2 between Area 1 and 2) power as the scheduled one.

In the work of AGC, a linearized model is commonly used for the load-generation dynamics [28]. For a three-area system, the frequency dynamics in Area i can be written as

$$\Delta\dot{\omega}_i = \frac{1}{2H_i}(\Delta P_{m_i} - \Delta P_{tie_i} - \Delta P_{l_i} - D_i\Delta\omega_i), \tag{8a}$$

where H_i is the equivalent inertia constant; D_i is the damping coefficient and ΔP_{l_i} denotes load deviations. Here ΔP_{tie_i}, ΔP_{m_i} represent the total tie-line power exchanges from Area i and the total generated power in Area i, i.e., $\Delta P_{tie_i} = \sum_{j\in\mathcal{E}_i} \Delta P_{tie_{i,j}}$ where $\mathcal{E}_i$ denotes the set of areas that connect to Area i, and $\Delta P_{m_i} = \sum_{g=1}^{G_i} \Delta P_{m_{i,g}}$ where G_i denotes the number of participated generators in Area i, and we have

$$\Delta\dot{P}_{m_{i,g}} = -\frac{1}{T_{ch_{i,g}}}(\Delta P_{m_{i,g}} + \frac{1}{S_{i,g}}\Delta\omega_i - \phi_{i,g}\Delta P_{agc_i}), \tag{8b}$$

$$\Delta\dot{P}_{tie_{i,j}} = T_{ij}(\Delta\omega_i - \Delta\omega_j), \tag{8c}$$

where $T_{ch_{i,g}}$ is the governor-turbine's time constant; $S_{i,g}$ denote the droop coefficient; T_{ij} is the synchronizing parameter between Area i and j. Note that ΔP_{agc_i} is the signal from AGC for the participated generators to

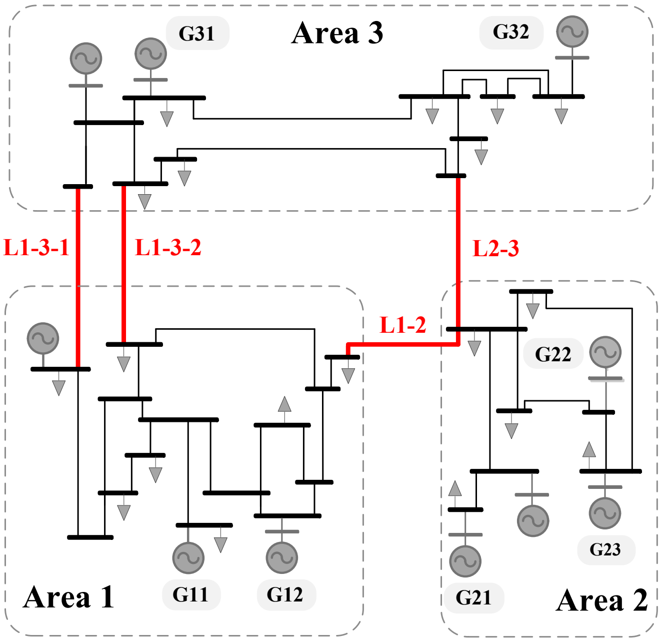


FIGURE 2. Three-area 39-bus system: the measurements of the tie-lines (in red) L1-3, L1-2, L2-3 are attacked.

track the load changes, and $\phi_{i,g}$ is the participating factor, i.e., $\sum_{g=1}^{G_i} \phi_{i,g} = 1$. After receiving the frequency and tie-line power measurements, the *area control error* (ACE) is computed for an integral action,

$$ACE_i = B_i \Delta\omega_i + \sum_{j \in \mathcal{E}_i} \Delta P_{tie_{i,j}}, \tag{8d}$$

$$\Delta \dot{P}_{agc_i} = -K_{I_i} ACE_i, \tag{8e}$$

where B_i is the frequency bias and K_{I_i} represents the integral gain. Based on the equations (8), the linearized model of Area i can be presented as the state equation

$$\dot{X}_i(t) = A_{ii} X(t) + B_{i,d} d_i(t) + \sum_{j \in \mathcal{E}_i} A_{ij} X_j(t), \tag{9}$$

where X_i is the state vector; $d_i := \Delta P_{l_i}$ denotes load deviations. Recall Remark 2.1 that (9) is an augmented model for the closed-loop AGC system that X_i consists of not only the electrical grid states (e.g., frequency, generator output and tie-line power) but also the controller state ΔP_{agc_i}, i.e.,

$$X_i := \begin{bmatrix} \{\Delta P_{tie_{i,j}}\}_{j \in \mathcal{E}_i} & \Delta\omega_i & \{\Delta P_{m_{i,g}}\}_{1:G_i} & \Delta P_{agc_i} \end{bmatrix}^\top .$$

Besides in (9), A_{ii} is the system matrix of Area i; A_{ij} is a matrix whose only non-zero element is $-T_{ij}$ in row 1 or 2 and column 3; $B_{i,d}$ is the matrix for load deviations.

In addition to (9), we assume a measurement model with high redundancy that the measurements of each tie-line power ($\Delta P_{tie_{i,j}}$) and the total tie-lines' power (ΔP_{tie_i}), the frequency ($\Delta\omega_i$), each generator output ($\Delta P_{m_{i,g}}$) and the total generated power (ΔP_{m_i}), and the AGC controller output (ΔP_{agc_i}) are all available. Besides, vulnerabilities within SCADA networks may allow cyber intrusions. Thus the output equation is

$$Y_i(t) = C_i X(t) + D_{i,f} f_i(t), \tag{10}$$

where Y_i is the system output and C_i is the output tall-matrix with full column rank. Here f_i denotes multivariate attacks and the matrix $D_{i,f}$ quantifies which output is attacked. In the aforementioned section, due to the feedback loop, attacks on the measurements would also affect the frequency dynamics. Hence the state equation (9) during attacks becomes

$$\dot{X}_i(t) = A_{ii}X(t) + B_{i,d}d_i(t) + B_{i,f}f_i(t) + \sum_{j\in\mathcal{E}_i} A_{ij}X_j(t),$$

where $B_{i,f}$ is the matrix that relates attacks to system states.

Using the state equations of each area, the continuous-time model of the three-area system can be obtained,

$$\dot{X}(t) = \tilde{A}_{cl}X(t) + \tilde{B}_d d(t) + \tilde{B}_f f(t), \tag{11}$$

where X is the vector consisting of groups of dynamic states in each area; d is the vector for all areas' load deviations; f denotes all the attack signals in the three-area, namely,

$$X = \begin{bmatrix} X_1^\top & X_2^\top & X_3^\top \end{bmatrix}^\top, \quad d = \begin{bmatrix} \Delta P_{l_1} & \Delta P_{l_2} & \Delta P_{l_3} \end{bmatrix}^\top, \quad f = \begin{bmatrix} f_l^\top & f_2^\top & f_3^\top \end{bmatrix}^\top.$$

In (11), $\tilde{A}_{cl}$ is the closed-loop system matrix; $\tilde{B}_d$, $\tilde{B}_f$ are constant matrices that relate load deviations and attacks to system states. For the three-area system, these matrices are

$$\tilde{A}_{cl} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, \quad \tilde{B}_d = \operatorname{diag}\begin{bmatrix} B_{1,d}, & B_{2,d}, & B_{3,d} \end{bmatrix}, \quad \tilde{B}_f = \operatorname{diag}\begin{bmatrix} B_{1,f}, & B_{2,f}, & B_{3,f} \end{bmatrix}.$$

We can also obtain the output equation of the system,

$$Y(t) = CX(t) + D_f f(t), \tag{12}$$

where Y is the system output vector containing all the three areas' outputs; C is the output matrix; D_f quantifies all the vulnerable signals. Similarly, these matrices are

$$Y = \begin{bmatrix} Y_1^\top & Y_2^\top & Y_3^\top \end{bmatrix}^\top, \quad C = \operatorname{diag}\begin{bmatrix} C_1, & C_2, & C_3 \end{bmatrix}, \quad D_f = \operatorname{diag}\begin{bmatrix} D_{1,f}, & D_{2,f}, & D_{3,f} \end{bmatrix}.$$

To obtain the sampled discrete-time model as (4), (11) and (12) must be discretized. We deploy a zero-order hold (ZOH)[1] discretization for a given sampling period T_s [26],

$$A_{cl} = e^{\tilde{A}_{cl}T_s}, \quad B_d = \int_0^{T_s} e^{\tilde{A}_{cl}(T_s - t)}\tilde{B}_d \mathrm{d}t. \tag{13}$$

Note that the attack matrix $\tilde{B}_f$ has the same matrix transformation as $\tilde{B}_d$, resulting B_f. The above approximation is exact for a ZOH and (13) corresponds to the analytical solution of the discretization. Therefore, the above model can be described in the form of (4) which again can be fitted into the DAE (5). In Appendix 2.2, we provide the detailed description of the involved parameters of the three-area 39-bus system as well as the attack scenarios on the AGC measurements.

4. Robust Dynamic Detection

4.1. Preliminaries for diagnosis filter construction

An ideal detection aims to implement a non-zero mapping from the attack to the diagnostic signal while decoupled from system states and disturbances, given the available data $y[\,\cdot\,]$ in the control center. In the power system dynamics described via a set of DAE, we restrict the diagnosis filter to a type of dynamic residual generator in the form of linear transfer functions, i.e., $r_D[k] := R(q)y[k]$ where r_D is the residual

[1] The inputs signals $d(\,\cdot\,)$ and $f(\,\cdot\,)$ in (11) are assumed to be piecewise constant within the sampling periods.

signal of the diagnosis filter and $R(q)$ is a transfer operator. Note that $y[\,\cdot\,]$ is associated with the polynomial matrix $L(q)$ in (5). We propose a formulation of transform operator $R(q)$ as

$$R(q) := a(q)^{-1}N(q)L(q),$$

where $N(q)$ is a polynomial vector with the dimension of n_r and a predefined order d_N. To make $R(q)$ physically realizable, stable dynamics $a(q)$ with sufficient order need to be added as the denominator where all the roots are strictly contained in the unit circle. Note that, unlike the observer-based methods, here d_N can be much less than the dimension of system dynamics. Then $N(q)$ and $a(q)$ are the two variables for a diagnosis filter design. By multiplying $a(q)^{-1}N(q)$ in the left of (5), we have

$$r_D[k] = a(q)^{-1}N(q)L(q)y[k] = -\underbrace{a(q)^{-1}N(q)H(q)x[k]}_{\text{(I)}} - \underbrace{a(q)^{-1}N(q)F(q)f[k]}_{\text{(II)}}, \tag{14}$$

where term (I) in (14) is due to $x[\,\cdot\,]$ of system states and natural disturbances. Term (II) is the desired contribution from the attacks $f[\,\cdot\,]$. In view of this diagnosis filter description, we introduce a class of residual generator which is sensitive to disruptive stealthy attacks as defined in Definition 2.5.

Definition 4.1 (Robust residual generator). *Consider a linear residual generator represented via a polynomial vector $N(q)$. This residual generator is robust with respect to disruptive stealthy attacks introduced in Definition 2.5 if*

$$\begin{cases} (I) & N(q)H(q) = 0, \\ (II) & N(q)F(q)F_{\mathrm{b}}\alpha \neq 0, \quad \forall \alpha \in \mathcal{A}, \end{cases} \tag{15}$$

where the basis matrix F_{b} and the set $\mathcal{A}$ are the same as the ones in Definition 2.5.

In the next step, we show that the polynomial equations (15) in Definition 4.1 can be characterized as a feasibility problem of a finite robust program.

Lemma 4.2 (Linear program characterization). *Consider the polynomial matrices $H(q) = \sum_{i=0}^{1} H_i q^i$, $N(q) := \sum_{i=0}^{d_N} N_i q^i$ and $F(q) = F$, where $H_i \in \mathbb{R}^{n_r\times n_x}$, $N_i \in \mathbb{R}^{n_r}$, and $F \in \mathbb{R}^{n_r\times n_f}$ are constant matrices. Then, the family of robust residual generators in (15) is characterized by*

$$\begin{cases} (I) & \bar{N}\bar{H} = 0, \\ (II) & \left\|\bar{N}V(\alpha)\right\|_\infty > 0, \quad \forall \alpha \in \mathcal{A}, \end{cases} \tag{16}$$

where $\|\cdot\|_\infty$ denotes the infinite vector norm, and

$$\bar{N} := \begin{bmatrix} N_0 & N_1 & \cdots & N_{d_N} \end{bmatrix}, \quad \bar{H} := \begin{bmatrix} H_0 & H_1 & 0 & \cdots & 0 \\ 0 & H_0 & H_1 & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & H_0 & H_1 \end{bmatrix}, \quad V(\alpha) := \begin{bmatrix} FF_{\mathrm{b}}\alpha & 0 & \cdots & 0 \\ 0 & FF_{\mathrm{b}}\alpha & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & FF_{\mathrm{b}}\alpha \end{bmatrix}.$$

Proof. The proof follows a similar line of arguments as [23, Lemma 4.2]. The key step is to observe that $N(q)H(q) = \bar{N}\bar{H}[I,\ qI,\ \cdots,\ q^{d_N+1}I]^\top$, and $N(q)FF_{\mathrm{b}}\alpha = \bar{N}V(\alpha)[I,\ qI,\ \cdots,\ q^{d_N}I]^\top$. The rest of the proof follows rather straightforwardly, and we omit the details for brevity. □

4.2. Robust diagnosis filter: transient behavior

In light of (16), we can define a symmetric set for the design variable $\bar{N}$ of the dynamic residual generator,

$$\mathcal{N} := \{\bar{N} \in \mathbb{R}^{(d_N+1)n_r} \mid \bar{N}\bar{H} = 0,\ \|\bar{N}\|_\infty \leq \eta\}. \tag{17}$$

The second constraint in the set is added to avoid possible unbounded solutions. To design a robust residual generator, we aim to find an $\bar{N} \in \mathcal{N}$ that for all $\alpha \in \mathcal{A}$, (16) can be satisfied. To this end, a natural reformulation of the residual synthesis is to consider an objective function as the second quantity in (16) influenced by the parameters $\mathcal{N}$ and the attacker action α, i.e., $\mathcal{J}(\bar{N}, \alpha) := \|\bar{N}V(\alpha)\|_\infty$. A successful scenario from an attacker viewpoint is to minimize this objective function given a residual generator. Therefore, we take a rather conservative viewpoint where the attacker may have complete knowledge of the system model and even the residual generator parameters, and exploits it so as to synthesize a stealthy attack. We then reformulate the diagnosis filter design as the robust optimization program,

$$\gamma^\star := \max_{\bar{N} \in \mathcal{N}} \min_{\alpha \in \mathcal{A}} \Big\{ \mathcal{J}(\bar{N}, \alpha) := \|\bar{N}V(\alpha)\|_\infty \Big\}. \tag{18}$$

The optimal value $\gamma^\star$ of the robust reformulation (18) is indeed an indication whether the attack still remains stealthy in the dynamic setting, i.e., if $\gamma^\star > 0$ then the optimal solution $\bar{N}^\star$ yields a diagnosis filter in the form of (14) which detects all the admissible attacks introduced in Definition 2.5. However, if $\gamma^\star = 0$, then it implies that for any possible detectors (static or dynamic) there exists a stationary disruptive attack that remains stealthy. In the next step, we show that the robust program (18) can be equivalently reformulated as a finite (non-convex) optimization problem.

Theorem 4.3 (Finite reformulation of (18))**.** *The robust optimization* (18) *can be equivalently described via the finite optimization program*

$$\begin{aligned} \gamma^\star = \quad & \max_{\bar{N},\ \beta,\ \lambda} && b^\top \lambda \\ & \quad s.t. && \sum_{i=0}^{d_N} (\beta_{2i} - \beta_{2i+1}) N_i F F_{\mathrm{b}} = \lambda^\top A, \\ & && \mathbf{1}^\top \beta = 1,\ \beta \geq 0, \\ & && \bar{N} \in \mathcal{N},\ \lambda \geq 0, \end{aligned} \tag{19}$$

where $\beta = [\beta_0,\ \beta_1,\ \cdots,\ \beta_{2d_N+1}]^\top$ is an $\mathbb{R}^{2d_N+2}$-valued auxiliary variable.

Proof. See Appendix 1.1. □

The exact reformulation program (19) for (18) is unfortunately non-convex due to the bilinearity between the variables β and N_i in the first constraint. In the following corollary, we suggest a convex relaxation of the program by restricting the feasible set of the variable β to a $2d_N + 2$ finite possibilities where $\beta = [0,\ \cdots,\ 1,\ \cdots,\ 0]^\top$ in which the only non-zero element of the vector is the i-th element.

Corollary 4.4 (Linear program relaxation)**.** *Given $i \in \{1,\ \ldots,\ 2d_N + 2\}$, consider the linear program*

$$\begin{aligned} \gamma_i^\star := \quad & \max_{\bar{N},\lambda} && b^\top \lambda \\ & \quad s.t. && (-1)^i N_{\lfloor i/2 \rfloor} F F_{\mathrm{b}} = \lambda^\top A, \\ & && \bar{N} \in \mathcal{N},\ \lambda \geq 0, \end{aligned} \tag{LP$_i$}$$

where $\lfloor \cdot \rfloor$ is the ceiling function that maps the argument to the least integer. Then, the solution to the program (LP$_i$) *is a feasible solution to the exact robust design reformulation* (19)*, and* $\max_{\{i \leq 2d_N+2\}} \gamma_i^\star \leq \gamma^\star$*. In particular, if for any $i \in \{1,\ \ldots,\ 2d_N+2\}$ we have $\gamma_i^\star > 0$, then the solution to* LP$_i$ *offers a robust residual generator detecting all admissible disruptive attacks introduced by Definition 2.5.*

Corollary 4.4 suggests that the maximum optimal value of $\{\gamma_0^\star,\ \gamma_1^\star,\ \cdots,\ \gamma_{2d_N+2}^\star\}$ and its corresponding $\bar{N}^\star$ provide a suboptimal solution to the original robust design (18).

We note that the focus of this article is on stationary (time-invariant) attacks. It is also important to highlight that the robust design perspective (18) allows the attacker to know the system model and filter parameters. In such a setting, the detection procedure could be much more difficult if the attacker would be able to dynamically adapt the attacks over the time, i.e., the attack signal is time-varying. In fact, in a multivariate attack scenario, one can construct a disruptive time-varying attack bypassing any linear residual generators. The next remark alludes more to this situation.

Remark 4.5 (Time-varying stealthy attacks)**.** *Consider a multivariate attack $f = [f_1 \ f_2 \ \cdots \ f_{n_f}]^\top$ where each element is a time-varying signal $f_i = f_i[k]$. Then, the residual (14) can be rewritten as*

$$a(q)r_D[k] = -\sum_{i=1}^{n_f} \Big(N(q)F_i f_i[\,\cdot\,]\Big)[k], \tag{20}$$

where $F = [F_1 \ F_2 \ \cdots \ F_{n_f}]$ represents the attack dynamics matrix. One can inspect that when the time-varying relation $\sum_{i=1}^{n_f} \big(N(q)F_i f_i[\,\cdot\,]\big)[k] = 0$ holds for every k, for instance when

$$f_{n_f}[k] = -\big(N(q)F_{n_f}\big)^{-1} \sum_{i=1}^{n_f-1} \Big(N(q)F_i f_i[\,\cdot\,]\Big)[k],$$

then the residual outcome (20) stays zero for all k, and as such, the attack remains undetected.

The proposed robust design in (18) does not necessarily enforce a non-zero steady-state residual of the diagnosis filter under multivariate attacks. Namely, the design perspective of (18) focuses on detection of attacks during the transient behavior without any requirements on long-term behavior of the residual. Indeed, the residual signal r_D may return to zero value after a successful reaction to the attack occurrence. A more stringent perspective is to require a non-zero steady-state behavior under any admissible attack scenario in $\alpha \in \mathcal{A}$. This extension is addressed in the next subsection.

4.3. **Robust diagnosis filter: steady-state behavior**

In order to design a diagnosis filter with non-zero steady-state residual "alert" when a multivariate attack occurs, the robust optimization (18) can be modified by a more conservative (smaller) objective function $\mathcal{J}(\bar{N}, \alpha) := |\bar{N}\bar{F}\alpha|$ where

$$\bar{F} := \Big[FF_\mathrm{b} \quad FF_\mathrm{b} \quad \cdots \quad FF_\mathrm{b}\Big]^\top. \tag{21}$$

A similar treatment as the preceding subsection can establish a framework for computational purposes. The next lemma follows similar objective as in Lemma 4.2 with a more demanding requirement of the non-zero long-term residual behavior.

Lemma 4.6 (Non-zero steady-state residual characterization)**.** *For the polynomial matrices $H(q)$, $N(q)$ and $F(q)$ as defined in Lemma 4.2, the family of dynamic residual generators with non-zero steady-state residual under multivariate attacks can be characterized by the algebraic relations*

$$\begin{cases} (I) & \bar{N}\bar{H} = 0, \\ (II) & |\bar{N}\bar{F}\alpha| > 0, \quad \forall \alpha \in \mathcal{A}, \end{cases} \tag{22}$$

where $\bar{F}$ is defined in (21), and the matrices $\bar{N}, \bar{H}$ are as defined in Lemma 4.2.

Proof. Recall that $N(q)H(q) = \bar{N}\bar{H}[I, \ qI, \ \cdots, \ q^{d_N+1}I]^\top$. Thus if $\bar{N}\bar{H} = 0$, the diagnosis filter becomes $r_D[k] = -a(q)^{-1}N(q)f[k]$. Note the steady-state value of the filter residual under attacks would be $-a(q)^{-1}N(q)F(q)f|_{q=1}$. Thus for the multivariate attack with α, the steady-state value of the filter residual is $-a(1)^{-1}N(1)F(1)F_\mathrm{b}\alpha$. The proof concludes by noting that $N(1)F(1)F_\mathrm{b}\alpha = \bar{N}\bar{F}\alpha$. □

In a similar fashion, the robust design perspective in (18) can be modified accordingly as

$$\mu^\star := \max_{\bar{N}\in\mathcal{N}} \min_{\alpha\in\mathcal{A}} \Big\{\mathcal{J}(\bar{N},\alpha) := |\bar{N}\bar{F}\alpha|\Big\}. \tag{23}$$

Notice the relation between the new objective function with the absolute value and the one in (18) with the infinity-norm. As it appears in the next result, the new setting is in fact a restricted case of the finite reformulation in Theorem 4.3.

Theorem 4.7 (Residual long-term behavior: exact convex reformulation and Nash equilibrium)**.** *Consider the minimax counterpart of the program* (18) *as defined*

$$\varphi^\star := \min_{\alpha\in\mathcal{A}} \max_{\bar{N}\in\mathcal{N}} \Big\{\mathcal{J}(\bar{N},\alpha) := |\bar{N}\bar{F}\alpha|\Big\}. \tag{24}$$

Each of the program (23) *and* (24) *can be equivalently reformulated through the linear programs*

$$\begin{aligned} \mu^\star = \quad & \max_{\bar{N},\ \lambda} \quad b^\top\lambda \\ & \quad s.t. \quad\ \bar{N}\bar{F} = \lambda^\top A \\ & \qquad\qquad \bar{N}\in\mathcal{N},\ \ \lambda\geq 0\,, \end{aligned} \tag{25a}$$

$$\begin{aligned} \varphi^\star = \quad & \min_{v_1,\,v_2,\,w,\,\alpha} \quad \mathbf{1}^\top v_1 + \mathbf{1}^\top v_2 \\ & \quad s.t. \qquad \bar{H}w + v_1 - v_2 = \bar{F}\alpha \\ & \qquad\qquad\quad v_1\geq 0,\ \ v_2\geq 0, \\ & \qquad\qquad\quad A\alpha\geq b\,. \end{aligned} \tag{25b}$$

Moreover, the value of each of these two programs coincide, i.e., $\mu^\star = \varphi^\star$*.*

Proof. See Appendix 1.2. □

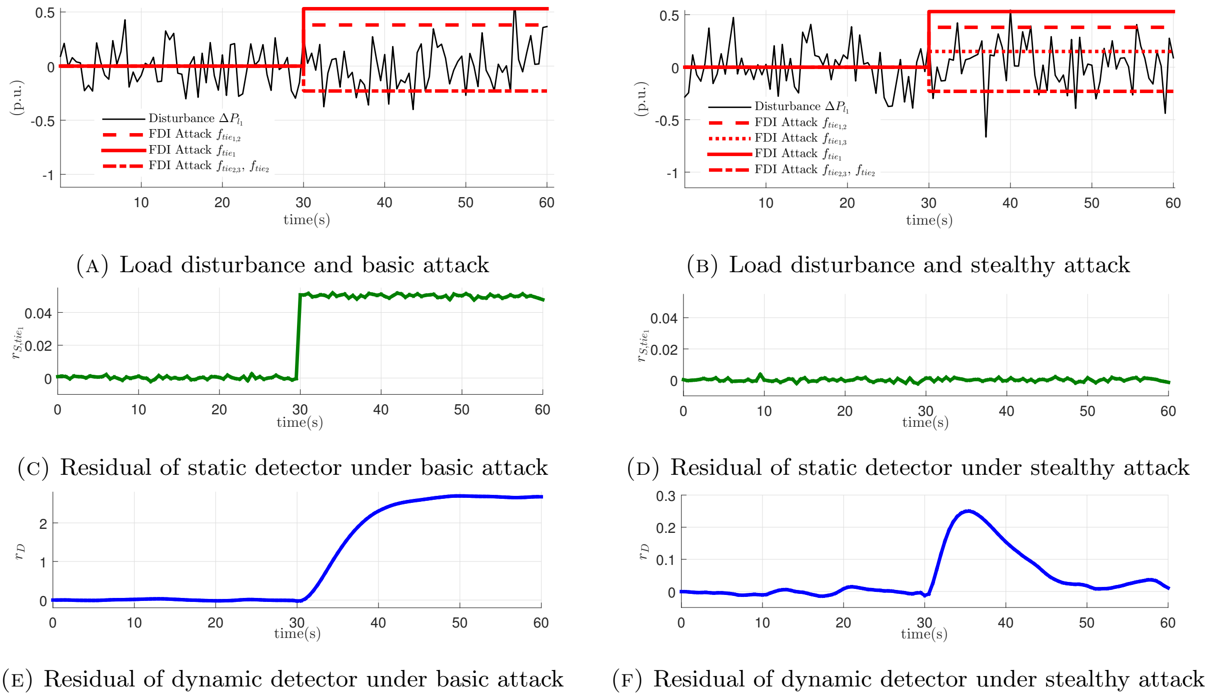
It is worth noting the difference between the robust perspective of (23) versus the minimax program (24). While in the design perspective of (23) the filter is oblivious to the possible attack scenarios, in the perspective of (24) the filter is aware of the attack signal and opts to detect that particular signal in the presence of natural disturbances. Obviously, the former setting is the one closer to the reality and, in general, the knowledge of the attack signal should help the detection significantly. This observation can indeed be translated through the usual weak inequality of $\mu^\star \leq \varphi^\star$. However, Theorem 4.7 indicates that the filter performance, in view of the long-term behavior of the worst-case attack scenario, indeed does not depend on the exact knowledge of the attacker signal and the inequality holds as the equality. We summarize this discussion in the following remark.

Remark 4.8 (Nash equilibrium interpretation)**.** *If the linear programs* (25a) (25b) *admit a positive optimal value* $\varphi^\star = \mu^\star > 0$*, then the resulting filter can detect all the admissible multivariate attacks described by Definition 2.5 along with a non-zero steady-state residual level. On the other hand, if the optimal values coincide with* $\varphi^\star = \mu^\star = 0$*, it then implies that there is no linear filter being able to decouple the admissible attack with* $\alpha^\star$*, the solution to* (25b)*, from the natural disturbances in a long-term horizon.*

5. Numerical Results

5.1. Test system and diagnosis filter description

In order to validate the effectiveness of the diagnosis filter with application to power system cyber security, we employed the IEEE 39-bus system which is well-known as a standard system for testing of new power system analysis. As shown in Figure 2, this system consists of 3 areas and 10 generators where 7 of them



(A) Load disturbance and basic attack

(B) Load disturbance and stealthy attack

(C) Residual of static detector under basic attack

(D) Residual of static detector under stealthy attack

(E) Residual of dynamic detector under basic attack

(F) Residual of dynamic detector under stealthy attack

FIGURE 3. Static detector in (2) versus dynamic detector (diagnosis filter) from Corollary 4.4 under basic and stealthy attacks.

are equipped with AGC for frequency control. All the participating generators in each area are with equal participation factors. The total load of the three-area system is 5.483 GW for the base of 100 MVA and 60 Hz. The generator specifications and AGC parameters of each area are referred to [4], and the linear frequency dynamics model has been developed in the preceding Section 3. Thus we result in a 19-order model in the form of (4).

We apply the diagnosis filter proposed in Section 4 to detect multivariate disruptive attacks on the measurements of AGC system. In the following simulations, we set the degree of the dynamic residual generator $d_N = 3$ which is much less than the order of the dynamics model, the sampling time $T_s = 0.5$ sec and the finite time horizon 60 sec. To design the filter, we set the denominator in the form $a(q) = (q-p)^{d_N}/(1-p)^{d_N}$ where p is a user-defined variable acting as the *pole* of the transfer operator $R(q)$, and it is normalized in steady-state value for all feasible poles. The pole is set to be $p = 0.8$ for a stable dynamic behavior at the beginning, and we have deployed the solver CPLEX to solve the corresponding optimization problems.

5.2. **Simulation results**

To evaluate the performance of the diagnosis filter, the disturbances $d_i = \Delta P_{l_i}$ are modeled as stochastic load patterns. To capture its uncertainty, as shown in Figure 3a and Figure 3b, we mainly model ΔP_{l_1} in Area 1 as random zero-mean Gaussian signals. It should be noted that tie-line power flow measurements are much more vulnerable to cyber attacks, comparing with frequency measurements (e.g., the anomalies in frequency can be easily detected by comparing the corrupted reading with the normal one.) [5]. Therefore as indicated in Figure 2 we mainly focus on the scenario that there are 5 vulnerable tie-line power measurements, namely $\Delta P_{tie_{1,2}}$, $\Delta P_{tie_{1,3}}$, ΔP_{tie_1}, $\Delta P_{tie_{2,3}}$ and ΔP_{tie_2}. Recalling Definition 2.5 for stealthy attack basis, thus there exist 3 basis vectors in the spanning set and we model them as follows: $f_1 = [0.1\ 0\ 0.1\ 0\ 0]^T$, $f_2 = [0.1\ 0.15\ 0.25\ 0\ 0]^T$, $f_3 = [0\ 0\ 0\ 0.1\ 0.1]^T$ (all in p.u.). Here each basis vector lies in the range space of the output matrix that the corrupted measurements still align with an actual physical state, bypassing the static detector $r_S[\cdot]$. Furthermore, without loss of generality we set $A = \mathbf{1}^\top$ and $b = 1.5$ in the set $\mathcal{A}$ and $\eta = 10$ in the set $\mathcal{N}$. The design variable $\bar{N}$ of the robust residual generator is first derived by solving (18) through (LP$_i$). The optimal value achieves maximum for $i = 2$ that $\gamma_2^\star = 300$, which implies a robust detection during the transient behavior as Corollary 4.4. For the given $\bar{N}$, the multivariate attack coordinates $\alpha = [2.8 \; 1 \; -2.3]^\top$ are obtained by solving the inner minimization of (18). Next, we look into the steady-state behavior of the filter with the above sets $\mathcal{N}$ and $\mathcal{A}$. For this, following Theorem 4.7 we solve (23) and (24) through the programs (25a) and (25b). It turns out that the derived optimal values satisfy the equality $\varphi^\star = \mu^\star = 0$, indicating that the optimal multivariate attack with $\alpha^\star$, the optimizer of the program (25b) and an optimal solution to (24), is a stealthy attack in the long-term horizon. We highlight that, thanks to the fact that the optimal values of the programs (25a) (25b) form a Nash equilibrium, even with the exact information of the stealthy attack coefficients $\alpha^\star$, we still cannot decouple the long-term behavior of the residual from the natural disturbances; see Remark 4.8.

In the first simulation, we begin with a general scenario where the multivariate attack is not carefully coordinated, i.e., basic attack. Thus as shown in Figure 3a, only 4 of 5 vulnerable measurements are compromised that $f_{tie_{1,2}} = 0.38 p.u.$, $f_{tie_1} = 0.53 p.u.$, $f_{tie_{2,3}} = -0.23 p.u.$ and $f_{tie_2} = -0.23 p.u.$. Note that since the injected data on $\Delta P_{tie_{1,2}}$ and ΔP_{tie_1} are inconsistent, the static detector is also expected to be triggered. To test the detectors in a more realistic setup, we also consider the presence of process and measurements noises. The process noise term added to the state equation of Area 1 is zero-mean Gaussian noises with the covariance matrix $R_{X_1} = 0.03 \times \text{diag}([1 \; 1 \; 0.03 \; 1 \; 1 \; 1 \; 1]^\top)$, i.e., the covariance of the noise to the frequency is 0.009 and the covariance of other states' noise is 0.03 [1]. Similarly, the measurement noise term added to the measurements of Area 1 is with the covariance matrix $R_{Y_1} = 0.03 \times \text{diag}([1 \; 1 \; 1 \; 0.03 \; 1 \; 1 \; 1 \; 1 \; 1]^\top)$, i.e., the covariance of the frequency measurement is 0.009 and the covariance of other measurements' noise is 0.03 [1]. Note the residue r_S of BDD in (2) becomes $r_S[k] = (I - C(C^\top R_Y^{-1} C)^{-1} C^\top R_Y^{-1}) Y[k]$ under the noisy system. The attacks are launched at $k_{\min} = 30$ sec. In Figure 3c and Figure 3e, results of the static detector in (2) and the proposed dynamic detector (diagnosis filter) are presented. Both detectors have succeeded to generate a diagnostic signal when attacks occurred, and the diagnosis filter residual r_D is significantly decoupled from stochastic load disturbances, and keeps sensitive to the multivariate attacks for a successful detection under noisy system settings.

In the second simulation, to challenge the detectors, now the multivariate attacks have been launched on all the 5 vulnerable measurements and the derived attack coefficient α from the optimization results has been used for a more intelligent adversary. Thus in Figure 3b, the corruptions become $f_{tie_{1,2}} = 0.38 p.u.$, $f_{tie_{1,3}} = 0.15 p.u.$, $f_{tie_1} = 0.53 p.u.$, $f_{tie_{2,3}} = -0.23 p.u.$ and $f_{tie_2} = -0.23 p.u.$. This corresponds to the worst case for the diagnosis filter that the adversary is given the knowledge of the residual generator's parameter $\bar{N}$ that it tries to minimize the payoff function over $\mathcal{A}$. Besides, the noisy system settings have been considered. Figure 3d and Figure 3f demonstrate all the simulation results. In Figure 3d, the static detector becomes totally blind to the occurrence of such an intelligent attack. However, as we can see in Figure 3f, even in the worst case, the diagnosis filter works perfectly well under the noisy system, generate a residual "alert" for the presence of multivariate attacks. We can also see that the residual output becomes close to zero value again after a successful detection during the transient behavior in Figure 3f, which is consistent to the aforementioned result $\varphi^\star = \mu^\star = 0$ and Remark 4.8. These simulations also prove the effectiveness and robustness of the proposed diagnosis filter design.

5.3. **Further discussions**

In this section we elaborate several practical aspects of the proposed filter in the preceding section.

5.3.A. *Diagnosis sensitivity to filter poles*

While the denominator of the filter $a(q)$ in (14) is chosen rather arbitrarily, up to a stability condition, the poles however has a significant impact on the residual sensitivity. As a general rule, the smaller the poles, the faster the residual responds, and the more sensitive the residual responds to model imprecision and noises. Simulation results in Figure 4 in Appendix 2.3 numerically illustrate this relation when the filter poles vary.

5.3.B. *Other types of attacks*

In addition to a smart multivariate measurement attacks, the main focus of this study, there are several other types of attacks that we briefly discuss in the following:

- *Denial-of-service (DoS) attack*: A type of availability attack where the attacker aims to prevent some specific data from being delivered to the respective destinations.
- *Replay attack*: A two-stage attack where the adversary gathers a sequence of data packets at stage 1, and then replays the recorded data afterwards at stage 2.

From a detection point of view, DoS attacks are trivially detectable without any sophisticated mechanisms as the absence of data is not stealthy. In the typical DoS attack modeling, the missing data is typically replaced with the last received ones [31]. In such a mechanism, the DoS can be treated as an "injection" attack. We investigate the performance of our filter in the presence of this class of attacks in Figure 5 in Appendix 2.3. Numerical results confirm that the proposed filter can successfully detect the DoS attacks. In regard with the replay attack, the articles [22, 14] offer sufficient conditions under which plausible attacks may remain stealthy irrespective of the detection mechanism providing that the attacker has access all the necessary data channels and excite attack of stage 2 at a suitable time.

5.3.C. *Observer-based diagnosis filters*

Another major technique for anomaly detection builds on observer-based techniques. In this view, the estimate of the system states, or in more general setting *output observer*, is a reference to alert the abnormality [11]. We close this section by a brief summary of the differences between these approaches and the one proposed in this study.

- The observer-based approaches typically yield diagnosis filters with higher dynamical system degrees than the approach proposed in this study. A low-order diagnosis filter is often more desired due to practical aspects of online implementation particularly for large-scale power systems.
- Observer-based diagnosis filters usually rely on a precondition of system observability. An extended version of such filters relaxes this condition to the so-called Luenberger-type conditions [2]. Our diagnosis filter, however, requires a weaker condition reflected through the feasibility condition of the resulting optimization programs, e.g., when the program (16) in Lemma 4.2 is feasible.
- Thanks to the optimization-based framework, unlike the observer-based approaches, we have a systematic approach to incorporate a multivariate attack scenario into the framework.

6. Conclusion

In this article, we investigated the problem of anomaly detection in the power system cyber security with a particular focus on exploiting the dynamics information where tempering multiple measurements data may be possible. Our study showed that a dynamical perspective to the detection task indeed offers powerful diagnosis tools to encounter attack scenarios that may remain stealthy from a static point of view. The effectiveness of this result was validated by simulations in the IEEE 39-bus system. Future research directions that we envision include an extension to nonlinear systems, as well as a setting exposed to the "dynamic" (time-variant) attacks in Remark 4.5, as opposed to the linear models and stationary attack scenarios studied in this article.

Appendix I: Technical Proofs

1.1. Proof of Theorem 4.3

Let us recall that $\bar{N}V(\alpha) = \Big[N_0 F F_{\mathrm{b}}\alpha \quad N_1 F F_{\mathrm{b}}\alpha \quad \cdots \quad N_{d_N} F F_{\mathrm{b}}\alpha\Big]$, and as such, the payoff function of the robust reformulation (18) is $\mathcal{J}(\bar{N},\alpha) = \max_i |N_i F F_{\mathrm{b}}\alpha|$ where $i \in \{0,\cdots,d_N\}$. By introducing an auxiliary variable β in the simplex set $\mathcal{B} \coloneqq \{\beta \in \mathbb{R}^{2d_N+2} \mid \beta \ge 0,\ \mathbf{1}^\top\beta = 1\}$, one can rewrite $\mathcal{J}$ as

$$\mathcal{J}(\bar{N},\alpha) = \max_{\beta\in\mathcal{B}} \sum_{i=0}^{d_N} (\beta_{2i} - \beta_{2i+1}) N_i F F_{\mathrm{b}}.$$

In this light, the original robust strategy (18) can be equivalently described via

$$\max_{\bar{N}\in\mathcal{N}} \min_{\alpha\in\mathcal{A}} \max_{\beta\in\mathcal{B}} \left\{ \sum_{i=0}^{d_N} (\beta_{2i} - \beta_{2i+1}) N_i F F_{\mathrm{b}}\alpha \right\}.$$

Note that given a fixed $\bar{N}$ the inner minimax optimization is indeed a bilinear objective in the decision variables and the respective feasible sets $\mathcal{A}$ and $\mathcal{B}$ are convex. Since one of the sets, $\mathcal{B}$, is also compact, then the zero-duality gap holds. Therefore, interchanging the optimization over $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$ yields

$$\gamma^\star = \max_{\bar{N}\in\mathcal{N},\ \beta\in\mathcal{B}} \left\{ \min_{\alpha\in\mathcal{A}} \sum_{i=0}^{d_N} (\beta_{2i} - \beta_{2i+1}) N_i F F_{\mathrm{b}}\alpha \right\}. \tag{26}$$

The inner minimization of (26) is a (feasible) linear program. We can use the duality again. To this end, let us assume that the decision variables $\bar{N}$ and β are fixed and consider the Lagrangian function

$$\mathcal{L}(\alpha;\lambda) = b^\top\lambda + \Big(\sum_{i=0}^{d_N} (\beta_{2i} - \beta_{2i+1}) N_i F F_{\mathrm{b}} - \lambda^\top A\Big)\alpha,$$

where optimizing over an unconstrained variable α becomes

$$\min_{\alpha} \mathcal{L}(\alpha;\lambda) = \begin{cases} b^\top\lambda & \text{if } \begin{cases} \sum_{i=0}^{d_N} (\beta_{2i} - \beta_{2i+1}) N_i F F_{\mathrm{b}} = \lambda^\top A \\ \lambda \ge 0 \end{cases} \\ -\infty & \text{otherwise,} \end{cases}$$

Using the above characterization as the most inner optimization program in (26) leads to

$$\begin{aligned} \max_{\lambda} \quad & b^\top\lambda \\ \text{s.t.} \quad & \sum_{i=0}^{d_N} (\beta_{2i} - \beta_{2i+1}) N_i F F_{\mathrm{b}} = \lambda^\top A, \\ & \lambda \ge 0. \end{aligned} \tag{27}$$

It then suffices to combine maximizing over the auxiliary variable λ together with the variables $\bar{N}$ and β to arrive at the main result in (19).

1.2. Proof of Theorem 4.7

We first prove the convex reformulation. For a given $\bar{N} \in \mathcal{N}$, the inner minimization of (23) can be translated as

$$\begin{array}{cl} \min\limits_{\alpha \in \mathcal{A},\, r} & r \\ \text{s.t.} & \bar{N}\bar{F}\alpha - r \le 0, \\ & -\bar{N}\bar{F}\alpha - r \le 0. \end{array}$$

The Lagrangian of the inner minimization reads as

$$\mathcal{L}(\alpha,\, r;\, \beta,\, \lambda) = b^\top \lambda + \big((\beta_0 - \beta_1)\bar{N}\bar{F} - \lambda^\top A\big)\alpha + (1 - \beta_0 - \beta_1)r.$$

Optimizing over the variables $\alpha,\ r$ yields

$$\min_{\alpha,\ r} \mathcal{L}(\alpha,\, r;\, \beta,\, \lambda) = \begin{cases} b^\top \lambda & \text{if} \begin{cases} (\beta_0 - \beta_1)\bar{N}\bar{F} = \lambda^\top A \\ \beta_0 + \beta_1 \le 1 \\ \beta_0 \ge 0,\ \beta_1 \ge 0,\ \lambda \ge 0 \end{cases} \\ -\infty & \text{otherwise.} \end{cases}$$

Then, combining maximization over the auxiliary variables $\lambda,\ \beta_0,\ \beta_1$ together with the variable $\bar{N}$ arrives at the optimization program,

$$\begin{array}{rcl} \mu^\star = & \max\limits_{\bar{N},\, \beta_0,\, \beta_1,\, \lambda} & b^\top \lambda \\ & \text{s.t.} & (\beta_0 - \beta_1)\bar{N}\bar{F} = \lambda^\top A, \\ & & \beta_0 + \beta_1 \le 1,\ \beta_0 \ge 0,\ \beta_1 \ge 0, \\ & & \bar{N} \in \mathcal{N},\ \lambda \in \mathbb{R}^{n_b},\ \lambda \ge 0. \end{array} \tag{28}$$

Note that the actual program (25a) is a restriction of (28) where the variables β_0 and β_1 are restricted to $\beta_0 = 1$ and $\beta_1 = 0$. Next, we show that this restriction is indeed without loss of generality. To this end, suppose the tuple $(\beta_0^\star,\, \beta_1^\star,\, \bar{N}^\star,\, \lambda^\star)$ is an optimal solution to the program (28). Note that the optimal variables $\beta_0^\star$ and $\beta_1^\star$ may satisfy one of the following three properties:

(i) $\beta_0^\star = \beta_1^\star$: In this case, $\lambda^\star = 0$, and therefore the optimal value $\mu^\star = 0$. This optimal solution can be trivially achieved in the program (25a) by setting $\bar{N} = 0$.

(ii) $\beta_0^\star > \beta_1^\star$: Observe that the tuple $\big(\beta_0^{'} = 1,\, \beta_1^{'} = 0,\, \bar{N}^{'} = \bar{N}^\star,\, \lambda^{'} = \lambda^\star/(\beta_0^\star - \beta_1^\star)\big)$ is a feasible solution with the objective value $b^\top \lambda^\star/(\beta_0^\star - \beta_1^\star)$. Since $b^\top \lambda^\star \ge 0$ by optimality assumption and $\beta_0^\star - \beta_1^\star \in (0, 1]$, then this feasible solution has a possibly higher optimal value, and therefore $\beta_0^\star - \beta_1^\star = 1$. That is, $\beta_0^\star = 1$ and $\beta_1^\star = 0$.

(iii) $\beta_0^\star < \beta_1^\star$: Following similar steps as the previous case together with the symmetric property of the feasible set $\mathcal{N}$, one can show that the optimal value of the program (28) also coincides with the restricted version in (25a).

This concludes the proof of the convex reformulation from (23) to (25a). In regard with the minimax problem (24), let us recall the symmetric property of the feasible set $\mathcal{N}$ in the variable $\bar{N}$. With a fixed α, the inner maximization can be directly formed as $\max_{\bar{N} \in \mathcal{N}} \bar{N}\bar{F}\alpha$ whose Lagrangian becomes

$$\mathcal{L}(\bar{N}; v, w) = -(\mathbf{1}^\top v_1 + \mathbf{1}^\top v_2) + \big(w^\top \bar{H}^\top + v_1^\top - v_2^\top - (\bar{F}\alpha)^\top\big)\bar{N}^\top,$$

Optimizing over the variable $\bar{N}$ leads to

$$\min_{\bar{N}} \mathcal{L}(\bar{N}; v, w) = \begin{cases} -\mathbf{1}^\top v_1 - \mathbf{1}^\top v_2 & \text{if } \begin{cases} \bar{H}w + v_1 - v_2 = \bar{F}\alpha \\ v_1 \geq 0, \ v_2 \geq 0 \end{cases} \\ -\infty & \text{otherwise.} \end{cases}$$

Thus, combining minimization over the auxiliary variables v_1, v_2, w together with the variable α, the minimax optimization (24) can be reformulated as the linear program (25b).

Finally, we show that the solution to programs (25) indeed forms a Nash equilibrium between the programs (23) and (24). Thus far, we have reformulated maximin and minimax problems as linear programs (25). The idea is to show that these programs have the same optimal values. In fact, we show that the programs are dual of each other, and that the strong duality holds when both programs are feasible. To this end, we resort to the duality of (25a) with the Lagrangian

$$\mathcal{L}(\bar{N}, \lambda; \alpha, v, w) = \big(w^\top \bar{H}^\top + v_1^\top - v_2^\top - (\bar{F}\alpha)^\top\big)\bar{N}^\top + (\alpha^\top A^\top - b^\top)\lambda - (\mathbf{1}^\top v_1 + \mathbf{1}^\top v_2).$$

Optimizing over the variables $\bar{N}$, λ yields

$$\min_{\bar{N}, \lambda} \mathcal{L}(\bar{N}, \lambda; \alpha, v, w) = \begin{cases} -\mathbf{1}^\top v_1 - \mathbf{1}^\top v_2 & \text{if} \begin{cases} \bar{H}w + v_1 - v_2 = \bar{F}\alpha \\ A\alpha \geq b \\ v_1 \geq 0, \ v_2 \geq 0 \end{cases} \\ -\infty & \text{otherwise.} \end{cases}$$

It is not difficult to see that the above program coincides with the program (25b); this concludes the proof.

Appendix II: System Parameters & Added Simulation Results

2.1. Dynamic Feedback Controller Modeling

Consider a dynamical system (e.g., the electrical power system studied in Section 3). Suppose the control signal is implemented as a *dynamic* feedback controller described by the discrete-time dynamics

$$\begin{cases} X_c[k+1] = A_c X_c[k] + B_c Y[k], \\ u[k] = C_c X_c[k] + D_c Y[k], \end{cases}$$

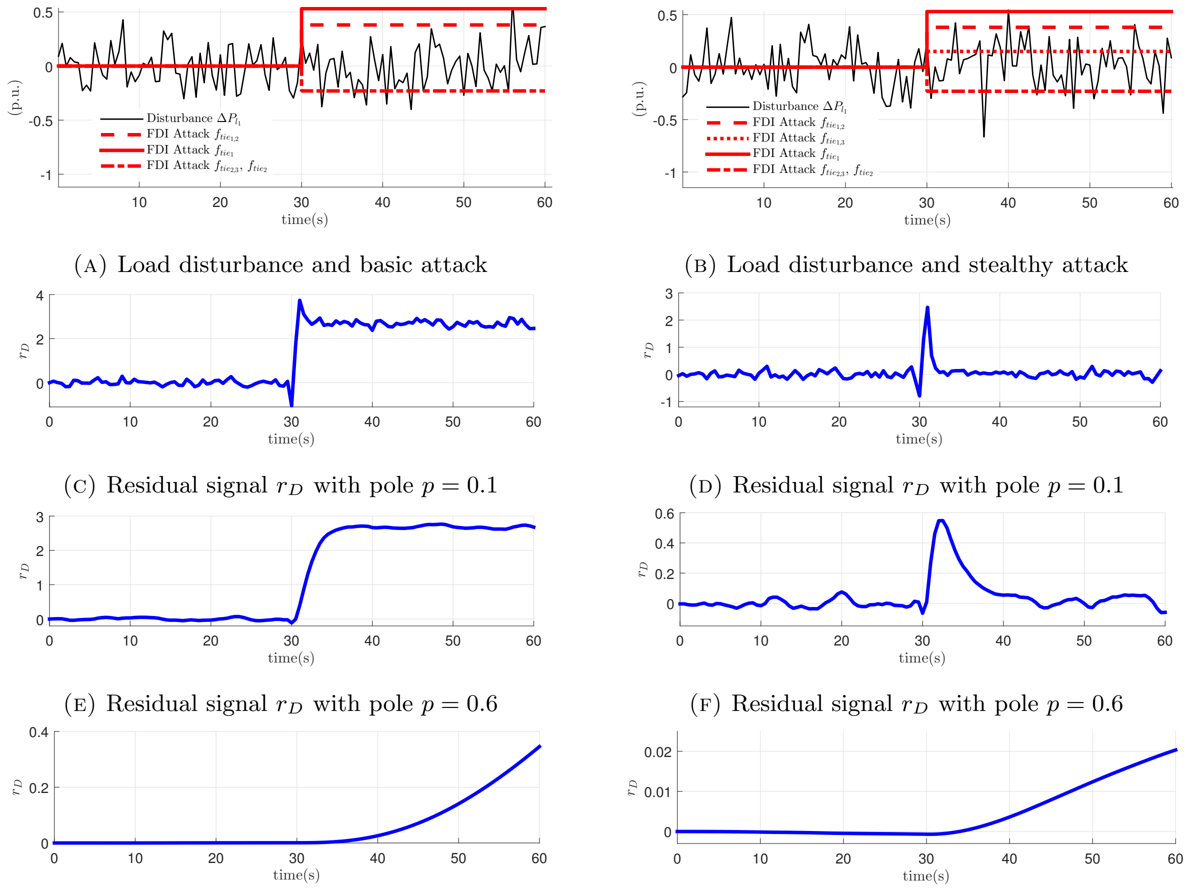
where the input is the dynamical system measurements $Y[\,\cdot\,]$, the output the control signal $u[\,\cdot\,]$, and the internal state of the controller is denoted by $X_c \in \mathbb{R}^{n_c}$. When an attack occurs on the measurements, it affects the dynamics of the controller and consequently the involved physical system. To study the control dynamics together with the original dynamical system, one can augment the states of the system (3) together with the controller's as $\hat{X} := [X^\top \ X_c^\top]^\top$. Assuming that the control signal can also be measured, one can also introduce an augmented measurement signals as $\hat{Y} = [Y^\top \ u^\top]^\top$. Following this procedure, the dynamics of the closed-loop system is described by

$$\begin{cases} \hat{X}[k+1] = \hat{A}_{cl}\hat{X}[k] + \hat{B}_d d[k] + \hat{B}_f f[k], \\ \hat{Y}[k] = \hat{C}\hat{X}[k] + \hat{D}_f f[k]. \end{cases} \tag{29}$$

where the involved matrices are defined as

$$\hat{A}_{cl} := \begin{bmatrix} A_x + B_u D_c C & B_u C_c \\ B_c C & A_c \end{bmatrix}, \quad \hat{B}_d := \begin{bmatrix} B_d \\ 0 \end{bmatrix}, \quad \hat{B}_f := \begin{bmatrix} B_u D_c D_f \\ B_c D_f \end{bmatrix},$$

$$\hat{C} := \begin{bmatrix} C & 0 \\ D_c C & C_c \end{bmatrix}, \quad \hat{D}_f := \begin{bmatrix} D_f \\ D_c D_f \end{bmatrix}.$$



(A) Load disturbance and basic attack

(B) Load disturbance and stealthy attack

(C) Residual signal r_D with pole $p = 0.1$

(D) Residual signal r_D with pole $p = 0.1$

(E) Residual signal r_D with pole $p = 0.6$

(F) Residual signal r_D with pole $p = 0.6$

(G) Residual signal r_D with pole $p = 0.98$

(H) Residual signal r_D with pole $p = 0.98$

FIGURE 4. Results of dynamic detector (diagnosis filter) with different *poles* ($p = 0.1,\ 0.6,\ 0.98$) under basic and stealthy attacks.

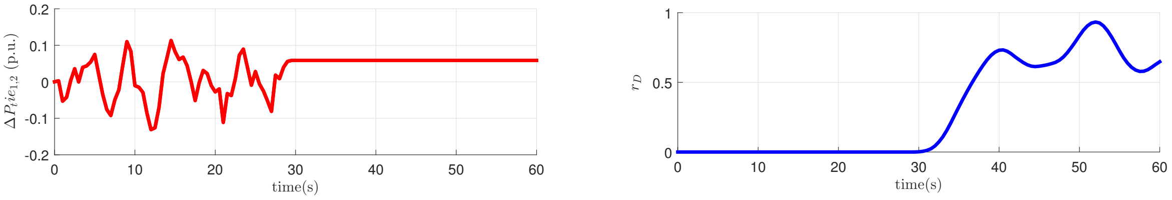
In this view, the augmented system (29) shares the same structure as (4) studied in the main part of the article for the case of static feedback controller.

2.2. AGC Parameters of the three-area 39-bus system

In this subsection we provide the involved matrices and parameters of the three-area 39 system. We take the model description of Area 1 in the three-area system in Figure 2 of Section 3 as an instance,

$$B_{1,d} = \begin{bmatrix} 0 & 0 & -\frac{1}{2H_1} & 0 & 0 & 0 \end{bmatrix}^\top,$$

$$A_{11} = \begin{bmatrix} 0 & 0 & T_{12} & 0 & 0 & 0 \\ 0 & 0 & T_{13} & 0 & 0 & 0 \\ -\frac{1}{2H_1} & -\frac{1}{2H_1} & -\frac{D_1}{2H_1} & \frac{1}{2H_1} & \frac{1}{2H_1} & 0 \\ 0 & 0 & -\frac{1}{T_{ch_{1,1}} S_{1,1}} & -\frac{1}{T_{ch_{1,1}}} & 0 & \frac{\phi_{1,1}}{T_{ch_{1,1}}} \\ 0 & 0 & -\frac{1}{T_{ch_{1,2}} S_{1,2}} & 0 & -\frac{1}{T_{ch_{1,2}}} & \frac{\phi_{1,2}}{T_{ch_{1,2}}} \\ -K_{I_1} & -K_{I_1} & -K_{I_1} B_1 & 0 & 0 & 0 \end{bmatrix}.$$



(A) $\Delta P_{tie_{1,2}}$ under DoS attacks from $k_{dos} = 30$

(B) Residual signal r_D under DoS attacks

FIGURE 5. Results of dynamic detector (diagnosis filter) under DoS attacks on $\Delta P_{tie_{1,2}}$ ($p = 0.8$).

As we have assumed a measurement model with high redundancy, the matrix C_i for Area 1 becomes

$$C_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}^\top .$$

In Area 1, the vulnerable measurements to cyber attacks are the ones of tie-line power flows $\Delta P_{tie_{1,2}}$, $\Delta P_{tie_{1,3}}$ and ΔP_{tie_1}. Thus the AGC signal ΔP_{agc_1} would be corrupted into

$$\Delta \dot{P}_{agc_1} = -k_1(B_1 \Delta\omega_1 + \Delta P_{tie_{1,2}} + f_{tie_{1,2}} + \Delta P_{tie_{1,3}} + f_{tie_{1,3}}) .$$

Then the parameters regarding multivariate attacks are

$$f_1 = \begin{bmatrix} f_{tie_{1,2}} & f_{tie_{1,3}} & f_{tie_1} \end{bmatrix}^\top ,$$

$$D_{1,f} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}^\top , \quad B_{1,f} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -k_1 \\ 0 & 0 & 0 & 0 & 0 & -k_1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^\top .$$

2.3. Additional simulation results

In Figure 4 we present the simulation results of the residual signal r_D from the proposed diagnosis filter under different *poles* ($p = 0.1,\ 0.6,\ 0.98$, respectively). We also show the simulation results of the residual signal r_D from the proposed diagnosis filter under DoS attacks in Figure 5.

REFERENCES

[1] A. AMELI, A. HOOSHYAR, E. EL-SAADANY, AND A. YOUSSEF, *Attack detection and identification for automatic generation control systems*, IEEE Transactions on Power Systems, (2018), p. 1.

[2] V. ANDRIEU AND L. PRALY, *On the existence of a kazantzis–kravaris/luenberger observer*, SIAM Journal on Control and Optimization, 45 (2006), pp. 432–456.

[3] A. ASHOK, M. GOVINDARASU, AND V. AJJARAPU, *Online detection of stealthy false data injection attacks in power system state estimation*, IEEE Transactions on Smart Grid, 9 (2018), pp. 1636–1646.

[4] H. BEVRANI, *Robust Power System Frequency Control*, Power Electronics and Power Systems, Springer, 2008.

[5] C. CHEN, K. ZHANG, K. YUAN, L. ZHU, AND M. QIAN, *Novel detection scheme design considering cyber attacks on load frequency control*, IEEE Transactions on Industrial Informatics, 14 (2018), pp. 1932–1941.

[6] T. M. CHEN AND S. ABU-NIMEH, *Lessons from stuxnet*, Computer, 44 (2011), pp. 91–93.

[7] C. Cybersecurity, *Framework for improving critical infrastructure cybersecurity version 1.1*, tech. report, National Institute of Standards and Technology, Apr. 2018.

[8] R. Deng and H. Liang, *False data injection attacks with limited susceptance information and new countermeasures in smart grid*, IEEE Transactions on Industrial Informatics, (2018), p. 1.

[9] S. X. Ding, *Model-based fault diagnosis techniques: design schemes, algorithms, and tools*, Springer Science & Business Media, 2008.

[10] X. Gao, X. Liu, and J. Han, *Reduced order unknown input observer based distributed fault detection for multi-agent systems*, Journal of the Franklin Institute, 354 (2017), pp. 1464–1483.

[11] W. Ge and C.-Z. FANG, *Detection of faulty components via robust observation*, International Journal of Control, 47 (1988), pp. 581–599.

[12] A. Giani, E. Bitar, M. Garcia, M. McQueen, P. Khargonekar, and K. Poolla, *Smart grid data integrity attacks*, IEEE Transactions on Smart Grid, 4 (2013), pp. 1244–1253.

[13] S. Gorman, *Electricity grid in US penetrated by spies*, The Wall Street Journal, 8 (2009).

[14] A. Hoehn and P. Zhang, *Detection of replay attacks in cyber-physical systems*, in American Control Conference, 2016, pp. 290–295.

[15] G. Hug and J. A. Giampapa, *Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks*, IEEE Transactions on Smart Grid, 3 (2012), pp. 1362–1370.

[16] S. Li, Y. Yilmaz, and X. Wang, *Quickest detection of false data injection attack in wide-area smart grids*, IEEE Transactions on Smart Grid, 6 (2015), pp. 2725–2735.

[17] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, *The 2015 Ukraine blackout: Implications for false data injection attacks*, IEEE Transactions on Power Systems, 32 (2017), pp. 3317–3318.

[18] J. Liang, L. Sankar, and O. Kosut, *Vulnerability analysis and consequences of false data injection attack on power system state estimation*, IEEE Transactions on Power Systems, 31 (2016), pp. 3864–3872.

[19] L. Liu, M. Esmalifalak, Q. Ding, V. A. Emesih, and Z. Han, *Detecting false data injection attacks on power grid by sparse optimization*, IEEE Transactions on Smart Grid, 5 (2014), pp. 612–621.

[20] Y. Liu, P. Ning, and M. K. Reiter, *False data injection attacks against state estimation in electric power grids*, in 16th ACM Conference on Computer and Communication Security, New York, 2009, pp. 21–32.

[21] M. A. Massoumnia, G. C. Verghese, and A. S. Willsky, *Failure detection and identification*, IEEE Transactions on Automatic Control, 34 (1989), pp. 316–321.

[22] Y. Mo and B. Sinopoli, *Secure control against replay attacks*, in 47th Annual Allerton Conference on Communication, Control, and Computing, 2009, pp. 911–918.

[23] P. Mohajerin Esfahani and J. Lygeros, *A tractable fault detection and isolation approach for nonlinear systems with probabilistic performance*, IEEE Transactions on Automatic Control, 61 (2016), pp. 633–647.

[24] P. Mohajerin Esfahani, M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, *Cyber attack in a two-area power system: Impact identification using reachability*, in American Control Conference, 2010, pp. 962–967.

[25] M. Nyberg and E. Frisk, *Residual generation for fault diagnosis of systems described by linear differential-algebraic equations*, IEEE Transactions on Automatic Control, 51 (2006), pp. 1995–2000.

[26] K. Ogata, *Discrete-time Control Systems (2Nd Ed.)*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1995.

[27] K. Pan, A. Teixeira, M. Cvetkovic, and P. Palensky, *Cyber risk analysis of combined data attacks against power system state estimation*, IEEE Transactions on Smart Grid, (2018), p. 1.

[28] E. Rakhshani, D. Remon, A. M. Cantarellas, J. M. Garcia, and P. Rodriguez, *Virtual synchronous power strategy for multiple hvdc interconnections of multi-area agc power systems*, IEEE Transactions on Power Systems, 32 (2017), pp. 1665–1677.

[29] D. Sahabandu, S. Moothedath, L. Bushnell, R. Poovendran, J. Aller, W. Lee, and A. Clark, *A game theoretic approach for dynamic information flow tracking with conditional branching*, in American Control Conference, 2019, pp. 2289–2296.

[30] D. Sahabandu, B. Xiao, A. Clark, S. Lee, W. Lee, and R. Poovendran, *Dift games: Dynamic information flow tracking games for advanced persistent threats*, in IEEE Conference on Decision and Control, Dec. 2018, pp. 1136–1143.

[31] L. Schenato, *To zero or to hold control inputs with lossy links?*, IEEE Transactions on Automatic Control, 54 (2009), pp. 1093–1099.

[32] P. Shukla, A. Chakrabortty, and A. Duel-Hallen, *A cyber-security investment game for networked control systems*, in American Control Conference, 2019, pp. 2297–2302.

[33] A. Teixeira, S. Amin, H. Sandberg, K. H. Johansson, and S. S. Sastry, *Cyber security analysis of state estimators in electric power systems*, in IEEE Conference on Decision and Control, 2010.

[34] E. E. Tiniou, P. Mohajerin Esfahani, and J. Lygeros, *Fault detection with discrete-time measurements: An application for the cyber security of power networks*, in IEEE Conference on Decision and Control, 2013.

[35] J. Zhao, L. Mili, and M. Wang, *A generalized false data injection attacks against power system nonlinear state estimator and countermeasures*, IEEE Transactions on Power Systems, (2018), p. 1.