Online Analytical Characterization of Outliers in Synchrophasor Measurements: A Singular Value Perturbation Viewpoint

Kaveri Mahapatra, *Student Member, IEEE*, Nilanjan Ray Chaudhuri, *Senior Member, IEEE*, Rajesh Kavasseri, *Senior Member, IEEE*, and Sukumar Brahma, *Senior Member, IEEE*

***Abstract*—This paper presents a Principal Component Analysis (PCA)-based method for online characterization of outliers in synchrophasor measurements. To that end, a linearized framework is established to analyze dynamical response from a system under nominal and off-nominal (e.g. faulted) conditions, which are contained in the same window of synchrophasor data. Inspired by the singular value perturbation theory, a bound on the change in the norm of the Principal Component (PC) scores as a function of system state matrices is presented. It is shown that in presence of bad data outliers these bounds for higher dimensional PC scores will be significantly larger compared to lower-dimensions. The effect of the number of samples in the data window on the results of the analysis is established. Case studies on a simulated test system and on field data collected from a US utility are presented to support the analytical results. Finally, an online classifier for characterization of outliers is developed to illustrate the usefulness of the proposed framework for machine learning-based methods.**

***Index Terms*—PCA, Bad data, Outlier, SVD, Perturbation theory**

NOMENCLATURE

x, u, z	State, input, and algebraic variables, respectively of the nonlinear model.
$\Delta x, \Delta u$ $\Delta y, \Delta z$	State, input, output, and algebraic variables, respectively of the linearized model.
A_n, B_n C_n	State, input, and output matrices, respectively of the linear model under nominal condition.
A_f, B_f C_f	State, input, and output matrices, respectively of the linear model under faulted condition.
θ	Variance of the measurement noise.
T_s	Sampling time of the synchrophasor output.
Y_{bus}	Bus admittance matrix.
$[Y]$	Synchrophasor measurements of 'p' signals with nominal system response in a matrix.
$\left[\tilde{Y}\right]$	Synchrophasor measurements of 'p' signals with nominal system response and outlier samples in a matrix.
U, Σ V	Matrices containing left singular vectors, singular values, and right singular vectors, respectively obtained from Singular Value Decomposition.
$\vartheta[n]$	Noise at the n^{th} sample in measurement matrix $[\tilde{Y}]$.
$\vec{\tilde{\varsigma}}_i$	PC score in the i^{th} principal direction.
$[\xi]$	Perturbation matrix of the same size as $[Y]$.
$\underline{\sigma}(\cdot)$	Minimum singular values.
$\overline{\sigma}(\cdot)$	Maximum singular values.
$\wp$	Orthogonal projection onto the column space of the unperturbed matrix $[Y]$.

Kaveri Mahapatra and Nilanjan Ray Chaudhuri are with Department of Electrical Engineering, The Pennsylvania State University, State College, PA 16802, USA (e-mail: *kzm221@psu.edu, nuc88@engr.psu.edu*).

Rajesh G. Kavasseri is with the Department of ECE, North Dakota State University, Fargo, ND 58108 USA (e-mail:*rajesh.kavasseri@ndsu.edu*).

Sukumar M. Brahma is with Department of ECE, New Mexico State University, Las Cruces, NM 88003, USA (e-mail:*sbrahma@nmsu.edu*)

Financial support from NSF under grant awards CNS 1544621 and CNS 1739206 is gratefully acknowledged.

I. INTRODUCTION

PHasor Measurement Units (PMUs) play a major role in wide-area monitoring, protection, and control systems (WAMPAC) [1]–[3], and power system security assessment [4]. Errors in PMU measurement are direct consequences of limited measurement precision, telecommunication equipment noise, two-way communication systems [5], interference from devices, and cyber-attacks such as eavesdropping, GPS spoofing, and data tampering [6]. Events such as faults and line tripping cause the system to deviate from its quasi-steady state operating condition and appear as outliers in PMU measurements. Bad data can have similar appearance as of fault outliers which can jeopardize decision-making and determining the true state of the network. As an example, Fig. 1 shows outliers in frequency measured by PMUs installed at a US utility for two cases: a) Case 1 - spurious bad data outlier, and b) Case 2 - disturbance outlier. In this paper we aim to address the broad research question: how to develop an analytical framework for online characterization of synchrophasor data outliers?

There is a significant amount of work done on bad data detection in the literature - pertaining to state estimation [7]–[19]. Literature on detection of bad data originated by cyber attacks in PMU dynamic data samples include a common path algorithm [20], a hybrid intrusion detection system [21], and a Bayesian-based approximation filter proposed in [22]. Researchers have also focused on event detection algorithms that have proposed clustering methods [23], ellipsoid method [24] combined with decision tree [25], an event unmixing method [26], and Principal Component Analysis (PCA) [27]–[29].

In reference [27] lower dimensional signatures of the states of the power system are obtained by PCA. It was shown that

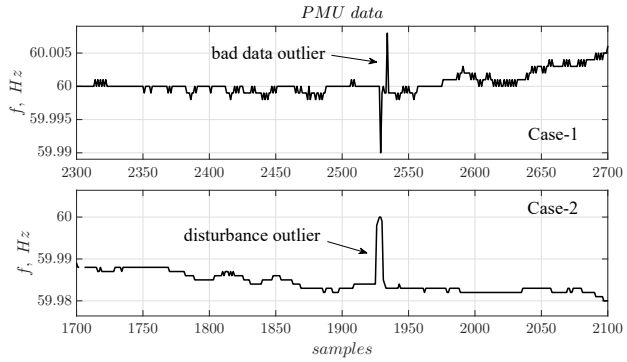

Fig. 1. Field PMU data obtained from a US utility for two cases: 1) bad data and 2) disturbance outlier. Outliers in frequency measurement are highlighted.

the occurrence of a system event will alter the core subspaces of the PMU data, which can be used as an early event detector. In contrast to [27]–[29] and prior works focused on event detection, the contributions of this paper are:

1) First, in section II, we establish a linearized framework that relates the system's intrinsic properties (via its state matrices) to synchrophasor measurements on a sliding window which includes mixed data, i.e., pre-disturbance (nominal) data as well as new (incoming) data - triggered by genuine faults or disturbances.
2) Second, in section III, using singular value perturbation analysis, we establish bounds on deviations in the Principal Component (PC) scores in terms of the system matrices. These bounds explicitly show the effect of the system matrices associated with the nominal and faulted states on these deviations, thereby serving as a basis to distinguish between data induced by faults versus measurement anomalies. These results also allow one to quantify the effect of bad data outliers on PCs in the higher versus lower dimensional subspaces.
3) Finally, in section IV, we validate the analytical results with dynamic simulations on: a) the 16-machine New England-New York system and b) actual field-recorded PMU data originating from a utility-owned transmission network in the U.S. The applicability of these concepts is shown through a nonlinear classifier using PC scores for discriminating outliers in synchrophasor data streams.

II. Proposed Framework for Analyzing Sampled Dynamic Response from Different Systems in Same Synchrophasor Data Window

First, let us consider the problem of online characterization of outliers caused by system disturbances (e.g. faults) using a moving-window PCA-based algorithm. For that, a linearized framework is established in this section, which will be able to analyze pre-disturbance data samples along with the outliers that just entered the window. In general, from a '*systems*' standpoint this framework will be capable of analyzing responses from different system models contained in the same window.

The model of the power system can be expressed in the form of nonlinear differential and algebraic equations as

$$\begin{aligned} \dot{x} &= f(x,u,z) \\ 0 &= g(x,u,z) \end{aligned} \tag{1}$$

where, x, u and z are the state variables, input variables, and algebraic variables, respectively. Linearizing the model around a nominal operating condition (x_0, u_0, z_0), and considering the measurement noise $\vartheta(t)$ as zero-mean independent Gaussian noise, we can write

$$\begin{aligned} \Delta\dot{x}(t) &= A_n \Delta x(t) + B_n \Delta u(t) \\ \Delta y(t) &= C_n \Delta x(t) + \vartheta(t), \vartheta(t) \sim \mathrm{N}(0,\theta) \end{aligned} \tag{2}$$

where, A_n, B_n, and C_n are the state matrix, input matrix, and output matrix, respectively of the nominal model; and θ is the variance of the measurement noise. Let us now consider a faulted condition, which is modeled as a high shunt conductance connected to a node that modifies the bus admittance matrix Y_{bus}. With this modified Y_{bus}, we can run loadflow, initialize the dynamic model and linearize it around the faulted condition (x_f, u_0, z_f). The following points should be noted for such linearization -

- The loadflow needs to converge, which may not be guaranteed even after relaxing the upper and lower limits of voltages. As an example, we were able to obtain convergence with fault conductance of $100.0pu$ on a $100MVA$ base in certain buses in the 16-machine 5-area New England-New York system, described later.
- Linearization will only make sense if the control limits are not hit under the operating condition.

Assuming the measurement noise remains unchanged, the linearized model under faulted condition is

$$\begin{aligned} \Delta\dot{x}(t) &= A_f \Delta x(t) + B_f \Delta u(t) \\ \Delta y(t) &= C_f \Delta x(t) + \vartheta(t) \end{aligned} \tag{3}$$

where A_f, B_f, and C_f are the state matrix, input matrix, and output matrix of the faulted model.

Let us consider that the initial state of the nominal system is Δx_0 and the fault takes place at $t = t_0$. The system states at fault instant are given by $\Delta x(t_0)$, which cannot change instantaneously. Therefore, the expression for output signals can be written as:

$$\begin{aligned} \Delta y(t) = {} & C_n \left[\mathrm{H}(t) - \mathrm{H}(t-t_0)\right] \left\{ e^{A_n t} \Delta x_0 + \int_0^t e^{A_n (t-\tau)} B_n \Delta u(\tau)\, d\tau \right\} \\ & + C_f \mathrm{H}(t-t_0)\, e^{A_f t} \left\{ e^{A_n t_0} \Delta x_0 + \int_0^{t_0} e^{A_n (t_0-\tau)} B_n \Delta u(\tau)\, d\tau \right\} \\ & + C_f \mathrm{H}(t-t_0) \int_0^t e^{A_f (t-\tau)} B_f \Delta u(\tau)\, d\tau + \vartheta(t) \end{aligned} \tag{4}$$

where, $\mathrm{H}(\cdot)$ is the Heaviside step function. Considering $\Delta y(t)$ as the measured signal, this equation expresses the signal as a function of two different system models. In practice, these measured signals will come from synchrophasors in the form of discrete samples.

Now, let us consider a data window of the measured signals from $t = [0, t_1]$ in which the fault took place at $t = t_0$.

Discretizing the linear systems in (2) and (3), we get

$$\begin{aligned} t=[0,t_0]&\begin{cases}\Delta x\left[k+1\right]=F_n\Delta x\left[k\right]+G_n\Delta u\left[k\right]\\ \Delta y\left[k\right]=C_n\Delta x\left[k\right]+\vartheta\left[k\right]\end{cases}\\ t=[t_0,t_1]&\begin{cases}\Delta x\left[k+1\right]=F_f\Delta x\left[k\right]+G_f\Delta u\left[k\right]\\ \Delta y\left[k\right]=C_f\Delta x\left[k\right]+\vartheta\left[k\right]\end{cases}\\ &F_n=e^{A_nT_s},G_n=A_n^{-1}\left(e^{A_nT_s}-I\right)B_n,\\ &F_f=e^{A_fT_s},G_f=A_f^{-1}\left(e^{A_fT_s}-I\right)B_f,\\ &\vartheta\left[k\right]\sim\mathrm{N}\left(0,\theta_d\right),\theta_d=\theta\end{aligned} \tag{5}$$

where, T_s denotes the sampling time of synchrophasor output.

We will assume that control input $u(t)$ (e.g. exciter voltage reference of generators) have not changed in the interval $[0, t_1]$ and that the fault instant t_0 is an integral multiple of sampling time T_s, which leads to

$$\Delta u\left(t\right)=0, r=\frac{t_0}{T_s}, r\in N \tag{6}$$

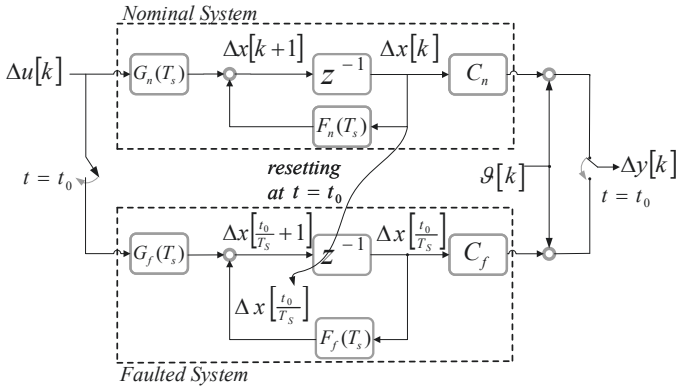

Fig. 2. Schematic representation of the proposed linearized framework for analyzing data samples from nominal and faulted systems in the same window. The state variable in the faulted system is reset to the corresponding value from the nominal system at the r^{th} sample, where $r=\frac{t_0}{T_s}$.

We are now in a position to formulate the proposed linearized framework to analyze the sampled synchrophasor data in the time window $[0, t_1]$. The following points should be noted in this context -

- When the fault occurs at $t = t_0$, the dynamic state $\Delta x[\frac{t_0}{T_s}]$ of the nominal state-space model *cannot change* instantly.
- The measured output *changes* instantly at $t = t_0$.

Figure 2 shows the schematic representation of the proposed linearized framework. The state variable in the faulted system is reset to the corresponding value from the nominal system at the r^{th} sample, where $r=\frac{t_0}{T_s}$. The input and the output signals are also switched to the faulted discrete model at the same instant. Considering p signals in synchrophasor measurements, the window of measured data can be represented by a matrix $\left[\tilde{Y}\right]$ as follows:

$$\left[\tilde{Y}\right]^T=\left[\underbrace{\Delta y\left[1\right]\Delta y\left[2\right]\cdots\Delta y\left[r-1\right]}_{Nominal}\underbrace{\Delta y\left[r\right]\cdots\Delta y\left[n\right]}_{Faulted}\right]_{p\times n} \tag{7}$$

where, $n=\frac{t_1}{T_s}$ is the number of samples from each signal present in the window. Noting that $\Delta y[r] = C_f\Delta x[\frac{t_0}{T_s}] + \vartheta\left[\frac{t_0}{T_s}\right]$ we can rewrite the data matrix as:

$$\begin{aligned}\left[\tilde{Y}\right]^T=&\Bigg[\underbrace{C_n\quad C_nF_n\quad\cdots\quad C_nF_n^{r-2}}_{Nominal}\\ &\underbrace{C_fF_n^{r-1}\quad C_fF_fF_n^{r-1}\quad\cdots\quad C_fF_f^{n-r}F_n^{r-1}}_{Faulted}\Bigg]\Delta x\left[1\right]\\ &+\left[\begin{array}{cccc}\vartheta\left[1\right]&\vartheta\left[2\right]&\cdots&\vartheta\left[n\right]\end{array}\right]\end{aligned} \tag{8}$$

The data matrix can be reconstructed after full PCA decomposition of $\left[\tilde{Y}\right]$ as sum of p orthonormal basis functions $\vec{\tilde{v}}_1$ to $\vec{\tilde{v}}_p$ as shown in (9)

$$\left[\tilde{Y}\right]=\begin{bmatrix}\tilde{\varsigma}_{1,1}\\ \\ \tilde{\varsigma}_{n,1}\end{bmatrix}\vec{\tilde{v}}_1+\begin{bmatrix}\tilde{\varsigma}_{1,2}\\ \\ \tilde{\varsigma}_{n,2}\end{bmatrix}\vec{\tilde{v}}_2+\cdots+\begin{bmatrix}\tilde{\varsigma}_{1,p}\\ \\ \tilde{\varsigma}_{n,p}\end{bmatrix}\vec{\tilde{v}}_p \tag{9}$$

The PC scores [30] for each observation in the i^{th} principal direction can be expressed as

$$\vec{\tilde{\varsigma}}_i=\left[\begin{array}{cccc}\tilde{\varsigma}_{1,i}&\tilde{\varsigma}_{2,i}&\cdots&\tilde{\varsigma}_{n,i}\end{array}\right]^T \tag{10}$$

These are calculated by applying Singular Value Decomposition (SVD) [31] on the data matrix $\left[\tilde{Y}\right]$ of size (n, p) given by

$$\left[\tilde{Y}\right]=U\Sigma V^T \tag{11}$$

where the principal vectors as columns of U_1 can be obtained by selecting first p columns of left singular vectors, U shown below.

$$\begin{aligned}U&=\left[\begin{array}{ccccccc}u_1&u_2&\cdots&u_p|&u_{p+1}&\cdots&u_n\end{array}\right]\\ &=\left[\begin{array}{cc}U_1&U_2\end{array}\right]_{n\times n}\end{aligned} \tag{12}$$

Also, the matrix Σ can be written as

$$\Sigma=\begin{bmatrix}\Sigma_1\\ 0\end{bmatrix}_{n\times p} \tag{13}$$

where, Σ_1 is a diagonal matrix of size (p, p) that contains p singular values for the matrix $\left[\tilde{Y}\right]$ in its diagonal elements. The right singular vectors are given by

$$V=\left[\begin{array}{ccccc}\vec{\tilde{v}}_1^T&\vec{\tilde{v}}_2^T&\cdots&\vec{\tilde{v}}_p^T\end{array}\right]_{p\times p} \tag{14}$$

Finally, the coefficients of the PCs are calculated as:

$$\tilde{\Xi}=\left[\begin{array}{ccccc}\vec{\tilde{\varsigma}}_1&\vec{\tilde{\varsigma}}_2&\cdots&\vec{\tilde{\varsigma}}_p\end{array}\right]_{n\times p}=U_1\Sigma_1 \tag{15}$$

As an orthogonal transformation, PCA maps the higher dimensional data to a lower dimensional subspace while preserving correlation between the original variables and maintaining maximum variance of the original data in its lower dimensional representation.

III. Bound on the Change in Norm of PC Scores: A Singular Value Perturbation Viewpoint

A. Presence of Disturbance Outliers

Let us now consider the case where a sliding window of n samples over p PMU signals are taken into account and the

fault outlier enters the window as the last sample. Therefore, (8) can be written as

$$\left[\tilde{Y}\right]^T = \left[\underbrace{C_n \quad C_nF_n \quad \cdots \quad C_nF_n^{n-2}}_{Nominal} \quad \underbrace{C_fF_n^{n-1}}_{Fault} \right] \Delta x\,[1] + \left[\vartheta\,[1] \quad \vartheta\,[2] \quad \cdots \quad \vartheta\,[n] \right] \tag{16}$$

The challenge is to quantify the alteration in the PC scores when the fault outliers enter the window. We propose to approach this problem from a singular value perturbation viewpoint.

To that end, let us consider the system response in absence of any fault, which is denoted by a data matrix $[Y]$ given by

$$[Y]^T = \left[\underbrace{C_n \quad C_nF_n \quad \cdots \quad C_nF_n^{n-2}}_{Nominal} \quad \underbrace{C_nF_n^{n-1}}_{Nominal} \right] \Delta x\,[1] + \left[\vartheta\,[1] \quad \vartheta\,[2] \quad \cdots \quad \vartheta\,[n] \right] \tag{17}$$

We view the faulted data matrix $\left[\tilde{Y}\right]$ as a perturbed matrix of $[Y]$, where the perturbation is denoted by $[\xi]$

$$\left[\tilde{Y}\right] = [Y] + [\xi] \tag{18}$$

The perturbation matrix can be written as

$$[\xi]^T = \left[0 \quad 0 \quad \cdots \quad 0 \quad (C_f - C_n)\,F_n^{n-1} \right] \Delta x\,[1] \tag{19}$$

Applying the classical perturbation bound derived by Weyl [32] we obtain $\forall i = 1(1)p$:

$$\left| \left\| \vec{\tilde{\varsigma}}_i \right\| - \left\| \vec{\varsigma}_i \right\| \right| \leq \left\| (\Delta x\,[1])^T \left(F_n^{n-1}\right)^T \left(C_f^T - C_n^T\right) \right\|_2 \tag{20}$$

where, $\left\| \vec{\varsigma}_i \right\|$ is the spectral radius of the i^{th} PC score of the unperturbed matrix. From now on, the 2-norm ($\|\bullet\|_2$) will be considered everywhere unless indicated otherwise. Using well-known norm inequalities we can write

$$\left| \Delta \left\| \vec{\tilde{\varsigma}}_i \right\| \right| \leq \left\| (\Delta x\,[1])^T \right\| \left\| \left(F_n^{n-1}\right)^T \left(C_f^T - C_n^T\right) \right\| \tag{21}$$

$$\boxed{\left| \Delta \left\| \vec{\tilde{\varsigma}}_i \right\| \right|_{\max} \propto \left\| \left(F_n^{n-1}\right)^T \left(C_f^T - C_n^T\right) \right\|} \tag{22}$$

■ *Remarks:* Equation (22) presents the bound on the alteration of the spectral radius of the i^{th} PC score as a function of the system matrices. Since the disturbance occurred at the last sample in the window and the system states do not change instantaneously, this bound is *not* a function of F_f. Notably, the above framework is generic and one can find such a bound on change in PC scores if more than one sample under faulted condition is considered where F_f will also be a part of the expression, as observed in equation (8).

Although (22) presents a bound on the change in the norm of PC scores in general, it does not give any indication on the variation of these bounds across different dimensions. We will apply a variant of the classical perturbation theorem proposed by Stewart [33] to obtain a clearer picture on these bounds across different dimensions:

$$\boxed{\left\| \vec{\tilde{\varsigma}}_i \right\|^2 = \left(\left\| \vec{\varsigma}_i \right\| + \gamma_i \right)^2 + \eta_i^2} \tag{23}$$

$$|\gamma_i| \leq \|\wp\xi\|, \underline{\sigma}\left(\wp^{\perp}\xi\right) \leq \eta_i \leq \overline{\sigma}\left(\wp^{\perp}\xi\right), \wp^{\perp} = (I - \wp) \tag{24}$$

where, $\wp$ is the orthogonal projection onto the column space of the unperturbed matrix $[Y]$, and $\underline{\sigma}(\cdot)$ and $\overline{\sigma}(\cdot)$ are the minimum and maximum singular values, respectively.

Therefore, we can write the following equations:

$$\wp = Y\left(Y^TY\right)^{-1}Y^T = \begin{bmatrix} \wp_{11} & \wp_{12} & \cdots & \wp_{1n} \\ \wp_{21} & \wp_{22} & \cdots & \wp_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \wp_{n1} & \wp_{n2} & \cdots & \wp_{nn} \end{bmatrix} \tag{25}$$

$$|\gamma_i| \leq \|\wp\xi\| = \left\| \begin{bmatrix} \wp_{1n} & \wp_{2n} & \cdots & \wp_{nn} \end{bmatrix} \right\| \|\beta\| \tag{26}$$

$$0 \leq \eta_i \leq \left\| \begin{bmatrix} -\wp_{1n} & -\wp_{2n} & \cdots & (1-\wp_{nn}) \end{bmatrix} \right\| \|\beta\| \tag{27}$$

$$\Delta\left\| \vec{\tilde{\varsigma}}_i \right\| \leq -\left\| \vec{\varsigma}_i \right\| + \left[\left\| \begin{bmatrix} -\wp_{1n} & -\wp_{2n} & \cdots & (1-\wp_{nn}) \end{bmatrix} \right\|^2 \times \|\beta\|^2 + \left(\left\| \vec{\varsigma}_i \right\| + \left\| \begin{bmatrix} \wp_{1n} & \wp_{2n} & \cdots & \wp_{nn} \end{bmatrix} \right\| \|\beta\| \right)^2 \right]^{\frac{1}{2}} \tag{28}$$

where, $\beta = (\Delta x\,[1])^T \left(F_n^{n-1}\right)^T \left(C_f^T - C_n^T\right)$. Equation (28) expresses the bound on the change in the norm of individual PC scores as a function of system state matrices.

B. Presence of Bad Data Outliers

Let us consider a case where bad data outliers exist in some of the PMU signals and enters the window at the last sample. We present the following arguments to quantify the bound on change in scores in lower and higher dimensional subspace.

■ *Lower dimensional subspace:* As described before, outliers originating from a system disturbance follow its dynamical characteristics. In contrast, the spurious outlier will not have this property. Let us express the data matrix $\left[\tilde{Y}\right]$ with signature anomaly that just entered the window as:

$$\left[\tilde{Y}\right]^T = \left[\underbrace{C_n \quad C_nF_n \quad \cdots \quad C_nF_n^{n-2}}_{No\,min\,al} \quad \underbrace{\Gamma F_n^{n-1}}_{Anomaly} \right] \times \Delta x\,[1] + \left[\vartheta\,[1] \quad \vartheta\,[2] \quad \cdots \quad \vartheta\,[n] \right] \tag{29}$$

Here, we have expressed the anomaly as a modification in the output matrix C_n of the system with matrix Γ. As before, viewing the data matrix $\left[\tilde{Y}\right]$ as a perturbed matrix of the nominal system response $[Y]$, the norm of the perturbation matrix $[\xi]$ is given by:

$$\|\xi\| = \left\| (\Delta x\,[1])^T \left(F_n^{n-1}\right)^T \left(\Gamma^T - C_n^T\right) \right\| \tag{30}$$

We assume that such modifications due to spurious outliers will take place only in a limited number of signals, thereby altering only the corresponding rows in C_n. Therefore, most of the entries in $\left(\Gamma^T - C_n^T\right)$ will be zeros. Hence the value of $\|\xi\|$ in this case can be expected to be smaller than the faulted scenario. Note that with an increasing number of bad

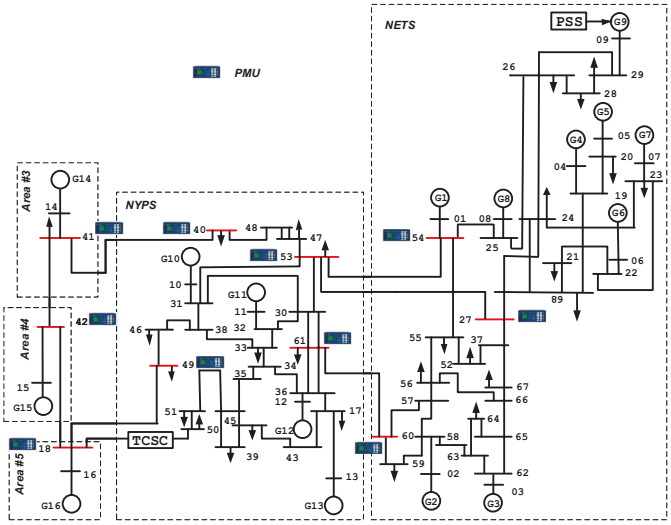

Fig. 3. Single-line diagram of 16-machine, 5-area New England-New York system with PMUs installed at major inter-tie buses highlighted in red.

data outliers, this will not be true, as will be evident from the studies in section IV-C.

Since the norm of PC scores $\left\|\vec{\varsigma}_i\right\|$ in the lower dimensions represent the system's dynamical information, those PC scores will have the most significant magnitude. So, it is reasonable to assume $\left\|\vec{\varsigma}_i\right\| >> \|\xi\|$. Therefore, the first term in equation (23) dominates, which leads to

$$\begin{aligned} \left\|\vec{\tilde{\varsigma}}_i\right\| &\cong \left\|\vec{\varsigma}_i\right\| + \gamma_i \\ \Rightarrow |\Delta \left\|\vec{\tilde{\varsigma}}_i\right\|| &\leq \|\wp\xi\| \end{aligned} \tag{31}$$

For $p << n$, i.e. for reasonably large number of time-samples in the window compared to the number of synchrophasor channels, $\|\wp\xi\| << \|\xi\|$. Therefore, as the window size increases, the bound on $|\Delta \left\|\vec{\tilde{\varsigma}}_i\right\||$ will be tighter in the lower dimensional subspace. This will be demonstrated in section IV-B Case 1.

■ *Higher dimensional subspace:* In the higher dimensional subspace for the cases where $\left\|\vec{\varsigma}_i\right\| \sim \|\xi\|$, the second term η_i will dominate and for $p << n$, the value of $\left\|\wp^{\perp}\xi\right\|$ will be significant. This will result in substantially higher bound on $|\Delta \left\|\vec{\tilde{\varsigma}}_i\right\||$ in the higher dimension.

IV. Case Study

A. Simulated Case

We have considered the 16-machine, 5-area New England-New York system [34] with PMUs installed at major inter-tie buses highlighted in red, which are shown in Fig. 3. Detrending was performed on all signals, and each angle signal was normalized with respect to its initial value before applying PCA.

1) Linear Analysis Results: To validate the concepts laid out in sections II and III, we linearized the test system (Fig. 3) model around 4 faulted conditions - faults at buses 53, 54, 60, and 61, respectively. The linearization was possible when a fault conductance of $100.0pu$ on a $100MVA$ base was assumed. The model was discretized using a sampling time

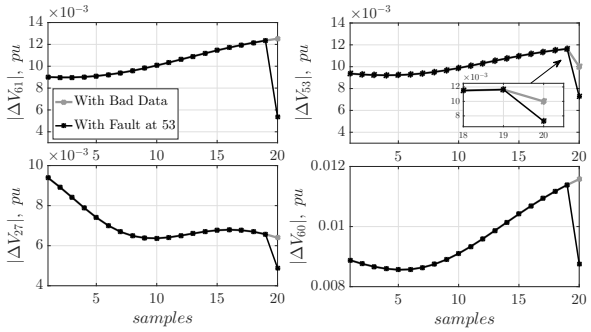

Fig. 4. Linear analysis results: Trajectories of 4 voltage magnitudes (out of 9) with two conditions - fault at bus 53 (black trace) and bad data outlier injected in bus voltage 53 (grey trace). Both outliers occur at the 20^{th} sample.

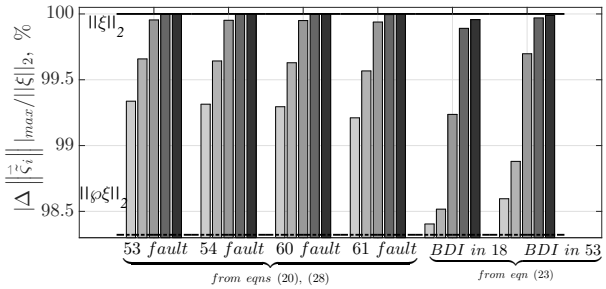

Fig. 5. Linear analysis results: Calculation of bounds on the change in the norm of PC scores in the first 5 dimensions obtained from equations (20) and (28) for faults. These bounds were also computed from (23) for bad data outliers deliberately injected in voltages. The bounds were computed using a 20−sample window and are expressed as a percentage of $\|\xi\|_2$. Color code: Lightest for 1^{st} dimension, Darkest for 5^{th} dimension.

$T_s = 0.0167s$ to consider the PMU output data rate of $60Hz$. We considered $\Delta x[1]$ to be 1.0% of the nominal steady-state values of the system states. Voltage magnitudes from buses $18, 27, 41, 42, 49, 53, 54, 60, 61$ were used as output signals and a 20-sample window was selected in which fault was simulated at the 20^{th} sample as described in section II. Separately, two bad data outliers were injected in different simulations at 20^{th} sample in two individual voltage measurements at buses 18 and 53, respectively. Figure 4 shows the trajectories of 4 voltage magnitudes under the two conditions mentioned above.

Figure 5 shows the calculation of bounds on the change in the norm of PC scores in the first 5 dimensions obtained from equations (20) and (28) for faults and using equation (23) for bad data outliers. The bounds are expressed as a percentage of $\|\xi\|_2$. It can be seen that spurious outliers coming from bad data injection (BDI) results in smaller relative bounds for change in the norm of PC scores in the 1^{st} and the 2^{nd} dimensions as compared to the faulted cases, which is in agreement with discussions in section III-B. The relative bounds for changes in the higher dimensions is comparable for both cases.

2) Nonlinear Simulation Results: Nonlinear time-domain simulation was conducted using the test system with the initial values of dynamic states deviated from their respective nominal conditions by 1.0% of steady state values. Uncorrelated zero mean gaussian noise is injected at different load buses to emulate quasi-static operation of power system. For this study the voltage magnitudes and angles of buses highlighted in Fig. 3 were measured and a 300-sample data window was considered. A sub-cycle fault was emulated by a shunt

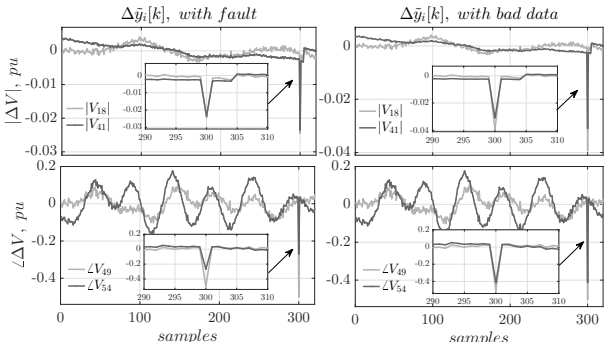

Fig. 6. Nonlinear simulation results: Trajectories of 2 voltage magnitudes and 2 separate angles after normalization and de-trending. Left subplots consider a sub-cycle fault at bus 53 in the 300^{th} sample. Right subplots consider BDI at the 300^{th} sample.

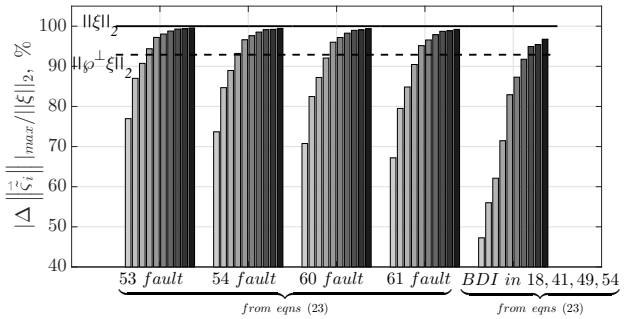

Fig. 7. Nonlinear simulation results: Calculation of bounds on the change in the norm of PC scores in the first 10 dimensions obtained from equations (23) for different faults and BDI. The bounds are expressed as a percentage of $\|\xi\|_2$. Color code: Lightest for 1^{st} dimension, Darkest for 10^{th} dimension.

conductance of $100.0pu$ at the same four buses as was done for the linear simulation (section IV-A1). Simultaneous BDI in four randomly-chosen variables - $|V_{18}|$, $|V_{41}|$, $\angle V_{49}$, $\angle V_{54}$ were simulated. Figure 6 right column shows these variables (after normalization and de-trending) with bad data and the left column shows the corresponding variables following the fault at bus 53. The fault and BDI was simulated at the 300^{th} sample.

Figure 7 shows the calculation of bounds on the change in the norm of PC scores in the first 10 dimensions obtained from equation (23) for different faults and BDI. The bounds are expressed as a percentage of $\|\xi\|_2$ and are calculated using a 300-sample window. Figure 8 shows the actual changes of the PC norms from $[Y]$ to $\left[\tilde{Y}\right]$ as a percentage of $\|\xi\|_2$. These results are consistent with the linearized analysis in suggesting that:

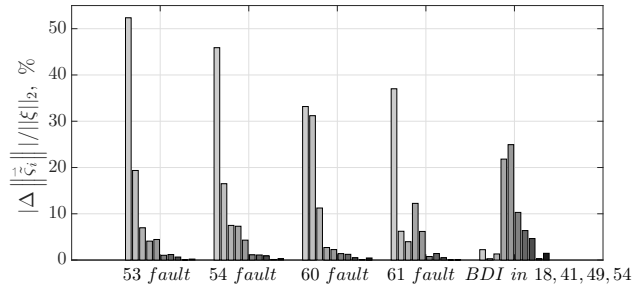

Fig. 8. Nonlinear simulation results: Calculation of change in the norm of PC scores in the first 10 dimensions for different faults and BDI using 300-sample window, the outliers being at the last sample. The changes are expressed as a percentage of $\|\xi\|_2$. Color code: Lightest for 1^{st} dimension, Darkest for 10^{th} dimension.

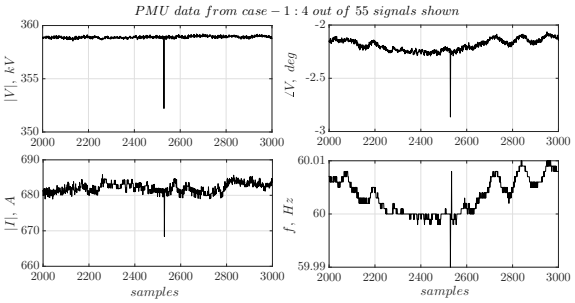

Fig. 9. Case 1 - Spurious Outliers: Four signals out of fifty five with outliers at the 2529^{th} sample.

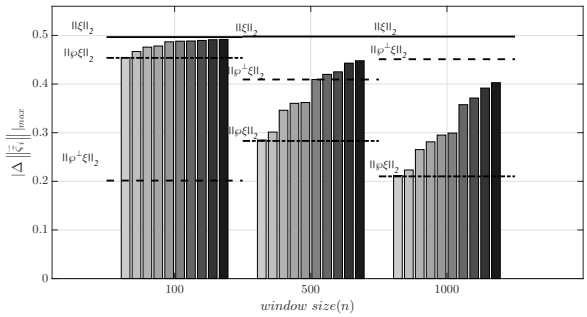

Fig. 10. Case 1: Calculation of bounds on the change in the norm of PC scores in the first 10 dimensions obtained from equation (23) for different window sizes. For each window the last sample is the outlier. Color code: Lightest for 1^{st} dimension, Darkest for 10^{th} dimension.

- For spurious data not triggered by physical system events, the deviations in PC norms are significantly greater in the higher dimensions compared to lower dimensions.
- For outliers that are triggered by genuine physical disturbances (i.e. faults), the changes in PC norms are most apparent in the lower dimensions.

B. Field PMU Data from a US Utility

We have considered synchrophasor data from the transmission network of a US utility. Fifty five signals from different PMUs containing measurement of voltage magnitudes, angles, current magnitudes, and frequencies were analyzed for two outlier case studies. As described in section IV-A, all signals are normalized and de-trended before applying PCA.

1) Case 1 - Spurious Outliers: Figure 9 shows four signals with outliers at the 2529^{th} sample.

■ *Bound on alteration of PC norms & effect of window size:* Based on our previous convention, field PMU data is denoted by the notation $\left[\tilde{Y}\right]$. To obtain the unperturbed matrix $[Y]$, the samples at the outlier location were replaced with the preceding samples, i.e. non-outliers. We have applied concepts described in section III and calculated the bounds on the change in the norm of PC scores in the first 10 dimensions using equation (23). Figure 10 shows the variation of these bounds with changing window size. For each case the last sample is considered to be the outlier. As the window size increases the following can be observed:

- $\left\|\wp^{\perp}\xi\right\|_2$ becomes significantly higher compared to $\|\wp\xi\|_2$ and determines the bound on higher dimensions.
- Alteration of PC norm in the lowest dimension is bounded consistently by $\|\wp\xi\|_2$ as per equation (31). Clearly,

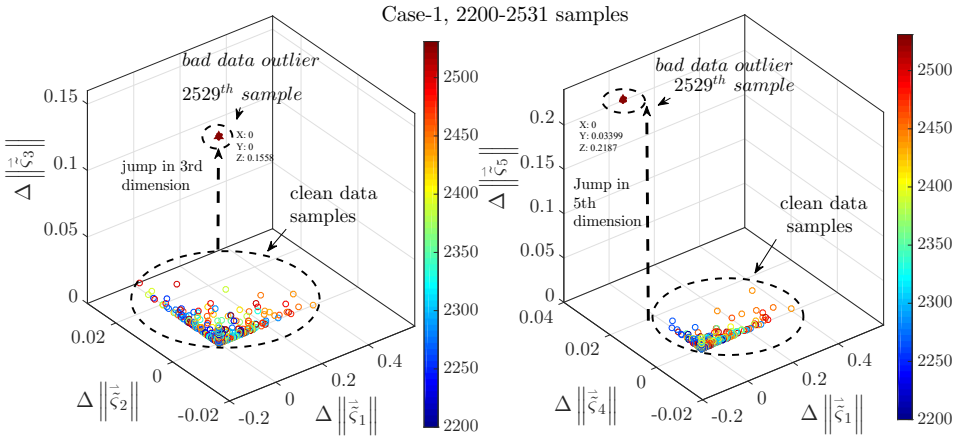

Fig. 11. Case 1: Calculation of change in the norm of PC scores in consecutive windows in the first 5 dimensions in two separate 3-D plots for a 100-sample sliding window PCA starting at the 2200^{th} sample, see in Fig. 9. The PC score samples from each window are color-coded in the side bar. Significant alteration is observed only in the higher (3^{rd} and 5^{th}) dimensional subspace when the 2529^{th} sample enters the window.

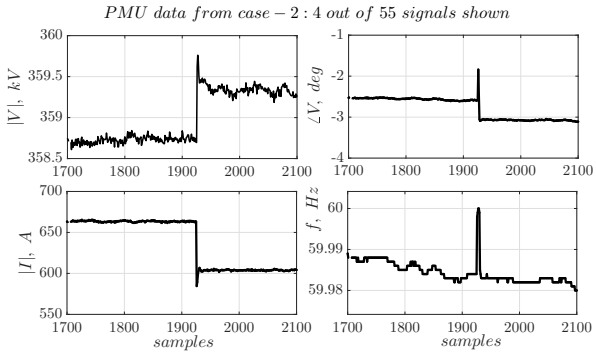

Fig. 12. Case 2 - Disturbance Outliers: Four signals out of fifty five with multiple outliers at $1926 - 1928^{th}$ sample.

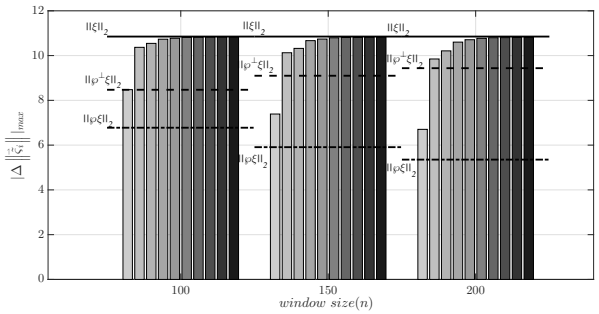

Fig. 13. Case 2: Calculation of bounds on the change in the norm of PC scores in the first 10 dimensions obtained from equation (23) for different window sizes. For each window the last sample is the outlier. Color code: Lightest for 1^{st} dimension, Darkest for 10^{th} dimension.

bound in the lower dimension gets progressively tighter as the window size increases.

- These observations suggest that these outliers can be classified as bad data. This will be shown more conclusively by looking at the actual changes in PC scores using a sliding window PCA.

■ *Sliding window PCA for online outlier characterization:* For an online characterization of the outlier, PCA was performed using a sliding window. The PC scores obtained from the previous window is subtracted from the current window to obtain the change in PC scores. Figure 11 shows the alteration in the norm of PC scores in the first 5 dimensions in two separate 3-D plots for a 100-sample sliding window PCA starting at the 2200^{th} sample of PMU data (Fig. 9). The PC score samples from each window are color-coded in the side bar. Significant alteration is observed only in the higher (3^{rd} and 5^{th}) dimensional subspace when the 2529^{th} sample enters the window. There is hardly any change in the lower dimensions, which implies that these outliers are bad data.

2) Case 2 - Disturbance Outliers: Figure 12 shows four signals with multiple outliers in $1926^{th} - 1928^{th}$ sample.

■ *Bound on alteration of PC norms & effect of window size:* Similar to Case 1 we applied the singular value perturbation theory on data from Case 2 and observe the effect of window size. Unlike Case 1 we do not observe the alteration in the lowest dimension to be bounded by $\|\wp\xi\|_2$. Also, the none of the alterations in higher dimensions are bounded by $\|\wp^{\perp}\xi\|_2$. Therefore, it is likely that these outliers are disturbance outliers. We will confirm this by a sliding window PCA described next.

■ *Sliding window PCA for online outlier characterization:* As before, PCA was performed using a sliding window. The PC scores obtained from the previous window are subtracted from the latest window to obtain the change in PC scores. Figure 14 shows the alteration in the norm of PC scores in the first 5 dimensions in two separate 3-D plots for a 100-sample sliding window PCA starting at the 1700^{th} sample. The PC score samples from each window are color-coded in the side bar. Significant alteration is observed only in the lower (1^{st} and 2^{nd}) dimensional subspace when $1926^{th} - 1928^{th}$ sample enters the window consecutively. This implies that these outliers are induced from system disturbance.

C. Case 3 : Effect of Number of Bad Data Outliers on Sliding Window PCA for Online Outlier Characterization

In order to study the effect of number of bad data outliers on sliding window PCA for online outlier characterization, we have examined the presence of bad data outlier in multiple (1 to 6) signals simultaneously out of 20 signals obtained from the nonlinear simulation case. Corresponding results with the increase in number of corrupted channels at the 300^{th} sample are shown in Fig. 15 in 6 subplots. The alteration in the norm of PC scores in the 1^{st}, 3^{rd}, and the 5^{th} dimensions

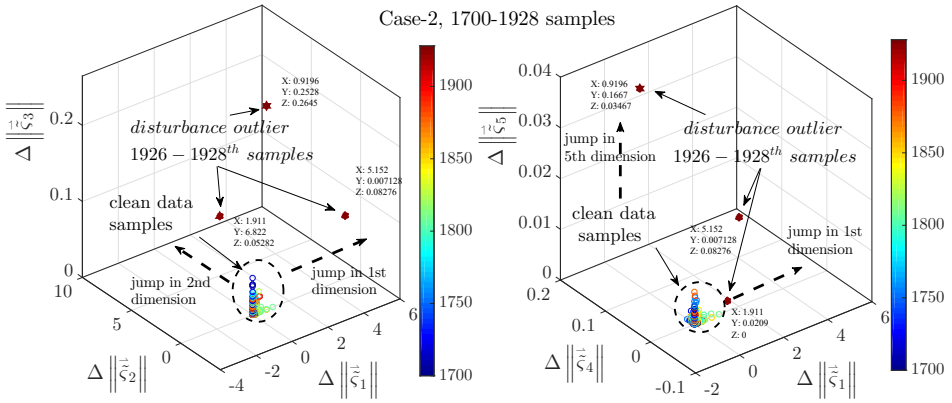

Fig. 14. Case 2: Calculation of change in the norm of PC scores in the first 5 dimensions in two separate 3-D plots for a 100-sample sliding window PCA starting at the 1700^{th} sample. The PC score samples from each window are color-coded in the side bar. Significant alteration is observed only in the lower (1^{st} and 2^{nd}) dimensional subspace when $1926-1928^{th}$ samples enter the window consecutively.

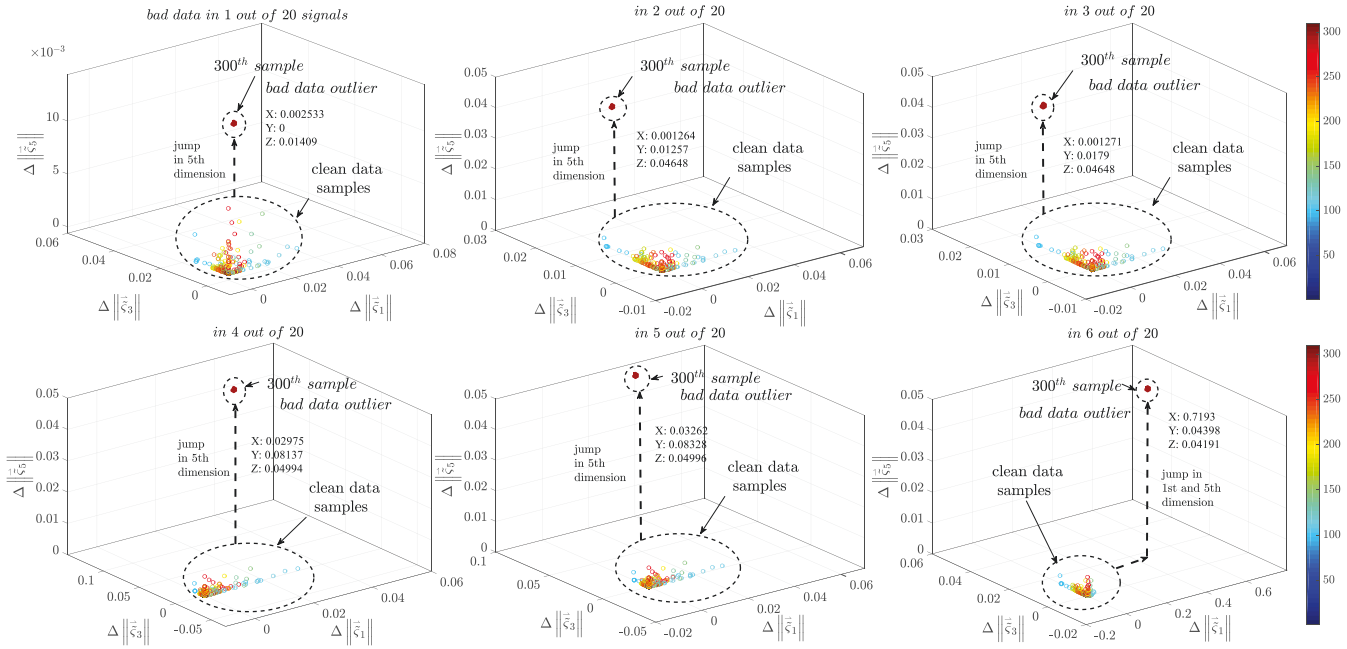

Fig. 15. Case 3: Comparison of change in the norm of PC scores in the 1^{st}, 3^{rd}, and 5^{th} dimensions with a 100-sample sliding window PCA starting at the 1^{st} sample for the nonlinear simulation. The six subplots include bad data outliers injected in a specific number of signals out of 20 simultaneously at the 300^{th} sample. Significant alteration is observed only in the higher (5^{th}) dimensional subspace around the 300^{th} sample for all the cases but also in lower (1^{st}) dimensional subspace when 6 variables are corrupted with bad data outlier.

are shown in the 3-D plots for a 100-sample sliding window PCA. The results are shown with the moving window PCA starting at the 1^{st} sample on nonlinear simulation data up to the 310^{th} sample. For up to 5 signals out of 20 being corrupted simultaneously, significant alteration in change in norm of PC scores is observed only in the higher (5^{th}) dimensional subspace when bad data outliers at the 300^{th} sample enter the window.

However, significant deviation is observed in the 1^{st} dimension as well when more than 5 signals are affected simultaneously by bad data. As the number of bad data outliers increases up to 6 or more than 6 out of 20, lower dimensional subspace gets affected. This leads to the following concluding remarks on this experiment:

- Beyond a certain amount of bad data outliers injected at any instant, the lower dimensional subspaces get affected. This is inline with the explanation provided in section III-B following equation (30).
- Since the distinguishing characteristics between the bad data and fault outliers rely on a significant deviation in higher dimensional subspace and minimal deviation in lower dimensional subspace, simultaneous injection of bad data can be tolerated in about 25% of the signals.

D. Case 4: PCA Feature-Driven Nonlinear Classifier

The results in this section are only meant to illustrate the utility of the proposed framework for machine learning-based methods. Here, we show for example, how an Artificial Neural Network (ANN)-based classifier capable of providing online bad data detection using the PC scores across different dimen-

sions can be developed. The performance of the classifier is evaluated for several operating conditions on the 16 machine test system (Fig. 3). A Monte Carlo (MC) simulation with 550 runs was performed with zero mean white Gaussian noise added to the loads with measurements sampled at 60 Hz.

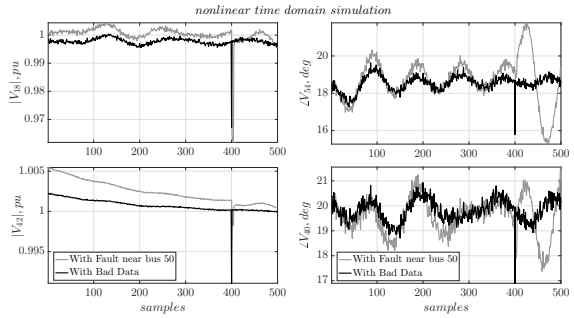

Fig. 16. Case 4: Four signals out of twenty from test system during fault near bus 50 at the 400^{th} sample and /or injected bad data in signals from buses 18, 40, 42, 54 obtained from two different MC runs.

Each run includes a $1-3\%$ perturbation in the initial system states and 400 samples are collected. Five hundred runs are obtained with bad data injected in 4 random measurements and the rest with fault near a randomly-selected bus at the 400^{th} sample. One typical MC run with bus 50 fault and another with bad data for buses 18, 42, 40 and, 54 are shown in Fig. 16. The norm of PC scores are extracted from the 1^{st} to the 4^{th} dimension at any instant obtained from a 300 sample moving window data and these features in 4 dimensions are utilized as input to ANN classifier for training and testing. A 5-fold cross-validation methodology is employed for examining the classifier performance for the proposed features. ANN with 4 input neurons, 1 output neuron, and 2 hidden layer neurons is trained with 70% MC trials of bad data and 70% MC trials of fault cases for each of the 5 subsamples. Table I highlights the results of confusion matrix obtained from trained ANN when the remaining 30% of MC trials for one subsample were used as 'testing data'. The average percentage of correct detection for good data is 99.99% and that for bad data is 96%. This shows that the proposed features are sufficient and robust enough for distinguishing outliers caused due to disturbance as well as due to bad data from normal operating condition. The average CPU time taken for detection of the data type after receiving measurement samples for the current instance is $0.0096s$ for this study.

TABLE I
CONFUSION MATRIX RESULTS ON THE PREDICTIVE PERFORMANCE FOR A SMALL TEST DATA SET FROM THE SYSTEM SHOWN IN FIG. 3

Data Type	True # Good Data Samples	True # Bad Data Samples	Correct Prediction
Output # Good Data Samples	13167 (99.75%)	2 (0.02%)	(13167/13169) (99.98%)
Output # Bad Data Samples	3 (0.02%)	28 (0.21%)	(28/31) (90.32%)
Correct Detection	(13167/13170) (99.99%)	(28/30) (93.33%)	(13195/13200) (99.96%)

E. Case 5: Effect of Window Size on Proposed Classifier

In Case 4, a 300-sample moving window was used. To analyze the impact of window size on the confusion matrix shown in Table I, three window sizes of samples 100, 200, and 350 were considered. The confusion matrices in Tables II, III, and IV summarize the results corresponding to different window sizes.

TABLE II
WITH WINDOW SIZE OF 100 SAMPLES : CONFUSION MATRIX RESULTS ON THE PREDICTIVE PERFORMANCE FOR A SMALL TEST DATA SET FROM THE SYSTEM SHOWN IN FIG. 3

Data Type	True # Good Data Samples	True # Bad Data Samples	Correct Prediction
Output # Good Data Samples	13160 (99.7%)	6 (0.08%)	(13160/13166) (99.95%)
Output # Bad Data Samples	10 (0.08%)	24 (0.18%)	(24/34) (70.59%)
Correct Detection	(13160/13170) (99.92%)	(24/30) (80%)	(13184/13200) (99.88%)

TABLE III
WITH WINDOW SIZE OF 200 SAMPLES : CONFUSION MATRIX RESULTS

Data Type	True # Good Data Samples	True # Bad Data Samples	Correct Prediction
Output # Good Data Samples	13161 (99.7%)	1 (0.07%)	(13161/13162) (99.99%)
Output # Bad Data Samples	9 (0.07%)	29 (0.22%)	(29/38) (76.32%)
Correct Detection	(13161/13170) (99.93%)	(29/30) (96.67%)	(13190/13200) (99.92%)

TABLE IV
WITH WINDOW SIZE OF 350 SAMPLES: CONFUSION MATRIX RESULTS

Data Type	True # Good Data Samples	True # Bad Data Samples	Correct Prediction
Output # Good Data Samples	13170 (99.77%)	1 (0.001%)	(13170/13171) (99.99%)
Output # Bad Data Samples	0 (0.0%)	29 (0.22%)	(29/29) (100%)
Correct Detection	(13170/13170) (100%)	(29/30) (96.67%)	(13199/13200) (99.99%)

Further, the mean efficiency for correct detection of True Bad data samples were calculated using a 5-fold cross validation for each of these window sizes. The results are summarized in the bar graph in Fig. 17.

Based upon the above analysis, we conclude that a typical window size of 300 samples is adequate in this particular case.

We would like to offer the following guidelines for selecting the window size –

- The larger the window size, the better the performance of the classifier would be.
- Increase in the window size comes at the cost of increased computational burden.
- Beyond a certain window size the performance gain will taper off. This window size will vary depending on the number of signals considered and can be determined from extensive offline studies.
- For a practical situation, the above guidelines should lead to a suitable window size based on the accuracy vs computational burden tradeoff.

F. Case 6: Comparison with an Outlier Detection Technique

Finally, we have compared the proposed method for characterizing the outliers with a popular technique used for

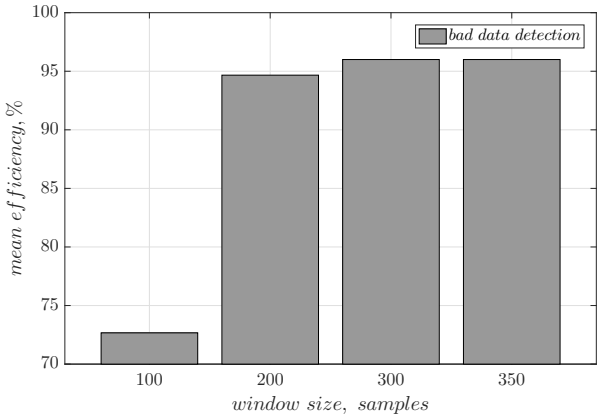

Fig. 17. Nonlinear simulation results: Mean efficiency of bad data detection obtained from a 5-fold cross validation performed on the ANN classifier.

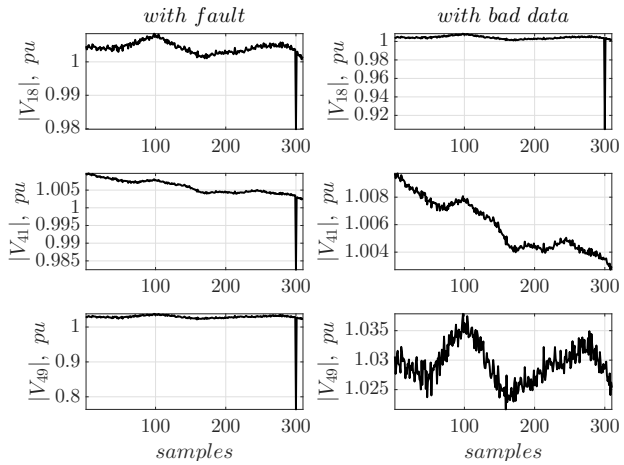

Fig. 18. Nonlinear simulation results: Trajectories of 3 selected voltage magnitudes used for MVEE experiment. Left subplots consider a sub-cycle fault at bus 53 in the 300^{th} sample. Right subplots consider BDI at the 300^{th} sample of bus 18 voltage magnitude measurement during ambient condition.

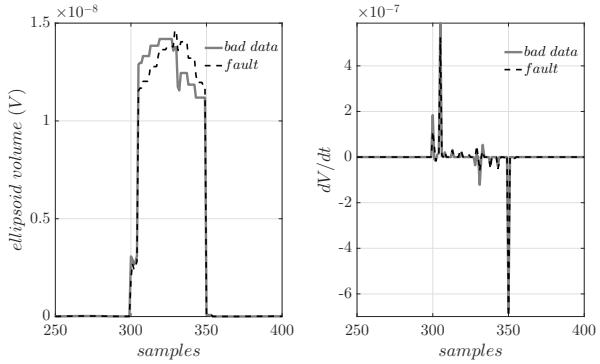

Fig. 19. Case 6: Comparison of ellipsoid volumes and rate of change of ellipsoid volume plots of a 50-sample sliding window-based minimum volume enclosing ellipsoid (MVEE) for fault case and bad data case.

event detection known as minimum volume enclosing ellipsoid (MVEE) [24]. MVEE uses a certain number of data samples of selected signals to construct a minimum volume enclosing second order surface by solving an optimization problem. The event detection capability of the method depends upon changes observed in the geometrical properties such as volume, rate of change in volume, etc. of the constructed ellipsoid.

In this work, we have compared the ellipsoids formed at the instances of bad data sample and fault sample in the measured data collected from selected PMUs of the nonlinear simulation case. The selected signals are the voltage magnitudes of the buses 18, 41, and 49, and their trajectories during a disturbance and bad data outliers upto the 310^{th} samples are shown in Fig. 18. MVEE technique was applied on a sliding window of 50 samples as used for event detection in [24]. The plots of volume and the rate of change of volume of the ellipsoids shown in Fig. 19 have similar appearances and overlap very closely for both the disturbance and bad data cases.

Similarly, the change in norm of PC scores in the 1^{st}, 3^{rd}, and 5^{th} dimensions from the sliding window analysis used in the proposed method are plotted in Fig. 20. The most significant deviation due to fault outliers is observed in the 1^{st} dimension in Fig. 20, however that due to bad data is observed only in the 5^{th} dimension. This trait can be very

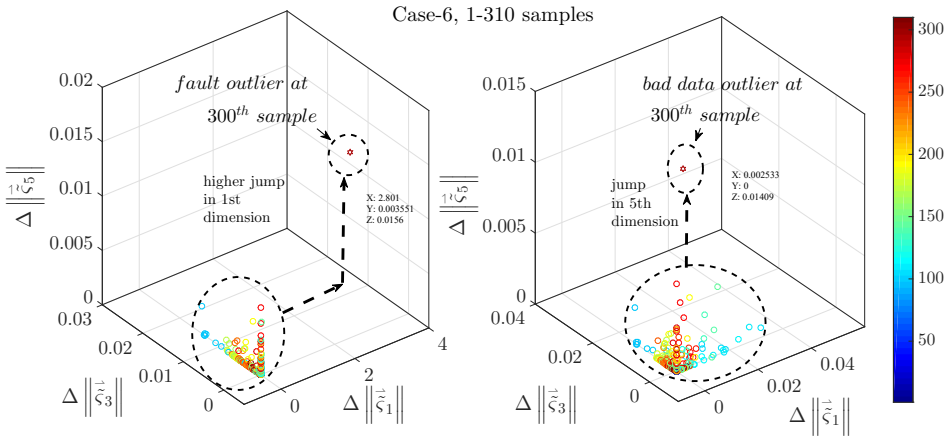

Fig. 20. Case 6: Comparison of change in the norm of PC scores in the 1^{st}, 3^{rd}, and the 5^{th} dimensions in two separate 3-D plots for fault case and bad data injection case with a 100-sample sliding window PCA starting at the 1^{st} sample for the nonlinear simulations. The PC score samples from each window are color-coded in the side bar. Significant alteration is observed only in the lower (1^{st}) dimensional subspace around the 300^{th} sample for fault case and in higher (5^{th}) dimensional subspace for bad data case.

useful for distinguishing bad data from disturbance outlier. In summary,

- Popular event detection methods such as MVEE are capable of detecting outliers, but fail to distinguish between genuine outliers due to disturbance and those due to bad data.
- The proposed method of analysis with change in norm of PC scores successfully characterizes disturbance and bad data outliers with the distinguishing features observed in higher and lower dimensions.

V. Conclusion

This paper presents a PCA-based method for online detection of "bad data"outliers in synchrophasor measurements which may closely resemble genuine disturbance data. A linearized framework to analyze dynamical response with mixed data (nominal and faulted or outliers) is established. Exploiting concepts from singular value perturbation theory, explicit bounds are derived to quantify the deviations in PC scores in terms of the system's state matrices. Using both simulated and field data, it is shown that in the presence of spurious data, the deviations in these bounds are relatively greater in the higher dimensions compared to the lower dimensions, in sharp contrast to data induced by genuine physical system disturbances where deviations are strongest in the lower dimensions. The proposed framework also allows one to quantify how these bounds are affected in the lower and higher dimensions, and by the window size. The results presented in this paper provide both theoretical and numerical insight into using PC-score based metrics for online classification as demonstrated by a nonlinear classifier.

References

[1] A. G. Phadke, "The Wide World of Wide-area measurement," *IEEE Power and Energy Magazine*, vol. 6, no. 5, pp. 52–65, Sept. 2008.
[2] V. Terzija, G. Valverde, D. Cai, P. Regulski, V. Madani, J. Fitch, S. Skok, M. M. Begovic, and A. Phadke, "Wide-Area Monitoring, Protection, and Control of Future Electric Power Networks," in *Proc. of the IEEE*, vol. 99, no. 1, Jan. 2011, pp. 80–93.
[3] L. Xie, Y. Chen, and H. Liao, "Distributed Online Monitoring of Quasi-Static Voltage Collapse in Multi-Area Power Systems," *IEEE Trans. Power Systems*, vol. 27, no. 4, pp. 2271–2279, Nov. 2012.
[4] Y. V. Makarov, P. Du, S. Lu, T. B. Nguyen, X. Guo, J. W. Burns, J. F. Gronquist, and M. A. Pai, "PMU-based Wide-Area Security Assessment: Concept, Method, and Implementation," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1325–1332, Sept. 2012.
[5] O. Vuković and G. Dán, "Security of Fully Distributed Power System State Estimation: Detection and Mitigation of Data Integrity Attacks," *IEEE J. on Selected Areas in Communications*, vol. 32, no. 7, pp. 1500–1508, Jul. 2014.
[6] S. Cui, Z. Han, S. Kar, T. T. Kim, H. V. Poor, and A. Tajer, "Coordinated Data-injection attack and Detection in the Smart Grid: A Detailed look at enriching Detection Solutions," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 106–115, Sept. 2012.
[7] A. Abur and A. G. Exposito, *Power system state estimation: theory and implementation*. CRC press, 2004.
[8] A. Monticelli, *State estimation in electric power systems: a generalized approach*. Springer Science & Business Media, 1999, vol. 507.
[9] Y. Liu, P. Ning, and M. K. Reiter, "False Data Injection Attacks Against State Estimation in Electric Power Grids," in *Proc. of the 16th ACM Conference on Computer and Communications Security*, ser. CCS '09, Nov. 2009, pp. 21–32.
[10] A. Tajer, "Energy grid state estimation under random and structured bad data," in *2014 IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM)*. IEEE, 2014, pp. 65–68.
[11] Z.-H. Yu and W.-L. Chin, "Blind false data injection attack using pca approximation method in smart grid," *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1219–1226, 2015.
[12] A. S. Dobakhshari and A. M. Ranjbar, "A wide-area scheme for power system fault location incorporating bad data detection," *IEEE Transactions on Power Delivery*, vol. 30, no. 2, pp. 800–808, April 2015.
[13] Z. H. Yu and W. L. Chin, "Blind false data injection attack using pca approximation method in smart grid," *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1219–1226, May 2015.
[14] D. Shi, D. Tylavsky, and N. Logic, "An adaptive method for detection and correction of errors in pmu measurements," in *2013 IEEE Power Energy Society General Meeting*, July 2013, pp. 1–1.
[15] I. Esnaola, S. M. Perlaza, H. V. Poor, and O. Kosut, "Maximum distortion attacks in electricity grids," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 2007–2015, July 2016.
[16] A. Tajer, "Energy grid state estimation under random and structured bad data," in *2014 IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, June 2014, pp. 65–68.
[17] J. Zhao, G. Zhang, K. Das, G. N. Korres, N. M. Manousakis, A. K. Sinha, and Z. He, "Power system real-time monitoring by using pmu-based robust state estimation method," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 300–309, Jan 2016.
[18] K. Nishiya, J. Hasegawa, and T. Koike, "Dynamic state estimation including anomaly detection and identification for power systems," *IEE Proceedings C - Generation, Transmission and Distribution*, vol. 129, no. 5, pp. 192–198, September 1982.
[19] A. M. L. da Silva, M. B. D. C. Filho, and J. M. C. Cantera, "An efficient dynamic state estimation algorithm including bad data processing," *IEEE Power Engineering Review*, vol. PER-7, no. 11, pp. 49–49, Nov 1987.
[20] S. Pan, T. Morris, and U. Adhikari, "Classification of disturbances and cyber-attacks in power systems using heterogeneous time-synchronized data," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 3, pp. 650–662, June 2015.
[21] ——, "Developing a hybrid intrusion detection system using data mining for power systems," *IEEE Transactions on Smart Grid*, vol. 6, no. 6, pp. 3104–3113, Nov 2015.
[22] H. M. Khalid and J. C. H. Peng, "A bayesian algorithm to enhance the resilience of wams applications against cyber attacks," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 2026–2037, July 2016.
[23] O. P. Dahal, S. M. Brahma, and H. Cao, "Comprehensive clustering of disturbance events recorded by phasor measurement units," *IEEE Transactions on Power Delivery*, vol. 29, no. 3, pp. 1390–1397, June 2014.
[24] J. Ma, Y. Makarov, C. Miller, and T. Nguyen, "Use of multi-dimensional ellipsoid to monitor dynamic behavior of power systems based on PMU measurement," in *2008 IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century*.
[25] J. Ma, Y. V. Makarov, R. Diao, P. V. Etingov, J. E. Dagle, and E. D. Tuglie, "The characteristic ellipsoid methodology and its application in power systems," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2206–2214, Nov 2012.
[26] W. Wang, L. He, P. Markham, H. Qi, Y. Liu, Q. C. Cao, and L. M. Tolbert, "Multiple event detection and recognition through sparse unmixing for high-resolution situational awareness in power grid," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1654–1664, July 2014.
[27] L. Xie, Y. Chen, and P. R. Kumar, "Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2784–2794, Nov 2014.
[28] Y. Chen, L. Xie, and P. R. Kumar, "Dimensionality reduction and early event detection using online synchrophasor data," in *2013 IEEE Power Energy Society General Meeting*, July 2013, pp. 1–5.
[29] ——, "Power system event classification via dimensionality reduction of synchrophasor data," in *2014 IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, June 2014, pp. 57–60.
[30] J. E. Jackson, *A user's guide to principal components*. John Wiley & Sons, 2005.
[31] I. Jolliffe, *Principal Component Analysis*. Wiley Online Library, 2002.
[32] H. Weyl, "Das asymptotische verteilungsgesetz der eigenwerte linearer partieller differentialgleichungen (mit einer anwendung auf die theorie der hohlraumstrahlung)," *Mathematische Annalen*, vol. 71, no. 4, pp. 441–479, 1912.
[33] "A note on the perturbation of singular values," *Linear Algebra and its Applications*, vol. 28, pp. 213 – 216, 1979.
[34] B. Pal and B. Chaudhuri, *Robust control in power systems*, ser. Power electronics and power systems. New York: Springer, 2005.

Kaveri Mahapatra (S'16) received the M. Tech. degree from Siksha 'O' Anusandhan University, India in 2013. She is currently pursuing her Ph.D. degree in the School of Electrical Engineering and Computer Science at the Pennsylvania State University, USA. Her current research interests include wide-area monitoring, protection and control, cybersecurity, soft computing and optimization, and power system dynamics.

Nilanjan Ray Chaudhuri (S'08-M'09-SM'16) received his Ph.D. degree from Imperial College London, London, UK in 2011 in Power Systems. From 2005-2007, he worked in General Electric (GE) John F. Welch Technology Center. He came back to GE and worked in GE Global Research Center, NY, USA as a Lead Engineer during 2011-2014. Presently, he is an Assistant Professor with the School of Electrical Engineering and Computer Science at Penn State, University Park, PA. He was an Assistant Professor with North Dakota State University, Fargo, ND, USA during 2014-2016. He is a member of the *IEEE* and *IEEE PES*. Dr. Ray Chaudhuri is the lead author of the book *Multi-terminal Direct Current Grids: Modeling, Analysis, and Control* (Wiley/IEEE Press, 2014), and an Associate Editor of the IEEE TRANSACTIONS ON POWER DELIVERY. Dr. Ray Chaudhuri is the recipient of the National Science Foundation Early Faculty CAREER Award in 2016.

Rajesh G. Kavasseri (M'02) received the B.E. degree from Visvesvaraya Regional College of Engineering, Nagpur, India, in 1995, the M.S. degree from the Indian Institute of Science, Bangalore, India, in 1998, and the Ph.D. degree from Washington State University, Pullman, WA, USA, in 2002, all in electrical engineering. His research interests focus on power systems analysis and computing. Since 2002, he has been with the Department of Electrical and Computer Engineering,North Dakota State University, Fargo, ND, USA, where he is currently an Associate Professor.

Sukumar Brahma received his Bachelor of Engineering from Gujarat University in 1989, Master of Technology from Indian Institute of Technology, Bombay in 1997, and PhD in from Clemson University, Clemson, USA in 2003; all in Electrical Engineering. He is presently William Kersting Endowed Chair Associate Professor and Associate Director of the Electric Utility Management Program (EUMP) at New Mexico State University, USA. Dr. Brahma is the past Chair of IEEE Power and Energy Society's Life Long Learning Subcommittee, past Chair of Distribution System Analysis Subcommittee, current Chair of Power and Energy Education Committee, and a member of Power System Relaying and Control Committee (PSRCC). He is an editor for IEEE Transactions on Power Delivery, and served as Guest Editor in Chief for the Special Issue on Frontiers of Power System Protection for the journal.