Received December 18, 2017, accepted February 9, 2018, date of current version March 28, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2805815

A Novel Data-Driven Situation Awareness Approach for Future Grids–Using Large Random Matrices for Big Data Modeling

XING HE[1], LEI CHU[1], ROBERT CAIMING QIU[1,2], (Fellow, IEEE), QIAN AI[3], (Senior Member, IEEE), AND ZENAN LING[1]

[1]Department of Electrical Engineering, Research Center for Big Data and Artificial Intelligence Engineering and Technologies, State Energy Smart Grid R&D Center, Shanghai Jiaotong University, Shanghai 200240, China
[2]Department of Electrical and Computer Engineering, Tennessee Technological University, Cookeville, TN 38505, USA
[3]Department of Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

Corresponding author: Robert Caiming Qiu (rqiu@ieee.org)

This work was supported in part by the NSF of China under Grant 61571296 and in part by the NSF, U.S., under Grant CNS-1619250.

ABSTRACT Data-driven approaches, when tasked with situation awareness, are suitable for complex grids with *massive datasets*. It is a challenge, however, to efficiently turn these massive datasets into useful big data analytics. To address such a challenge, this paper, based on random matrix theory, proposes a data-driven approach. The approach models massive datasets as large random matrices; it is model-free and requires no knowledge about physical model parameters. In particular, the large data dimension N and the large time span T, from the spatial aspect and the temporal aspect, respectively, lead to favorable results. The beautiful thing lies in that these linear eigenvalue statistics (LESs) are built from data matrices to follow Gaussian distributions for very general conditions, due to the *latest breakthroughs* in probability on the central limit theorems of those LESs. Numerous case studies, with both simulated data and field data, are given to validate the proposed new algorithms.

INDEX TERMS Big data analytics, linear eigenvalue statistics, random matrix theory, situation awareness, statistical indicator.

I. INTRODUCTION

Situation awareness (SA) is of great significance for power system operation, and a reconsideration of SA is essential for future grids [1]. These future grids are always huge in size and complex in topology. Operating under a novel regulation, their management mode is much different [2]. Data are more and more easily accessible, on the other hand, and data-driven approaches become natural for future grids. Towards this vision, following problems need to be solved urgently:

- There are massive data in power grids. The so-called curse of dimensionality [3] occurs inevitably.
- The resource cost (time, hardware, human, etc.) for extracting big data analytics should be tolerable.
- For a massive data source, there often exist realistic ''bad'' data, e.g. the incomplete, the inaccurate, the asynchronous, and the unavailable. For system operations, decisions such as relay actions, should be highly reliable.

This paper is built upon our previous work in the last several years. See Section I-B for details. Motivated for data mining, our line of research is based on the high-dimensional statistics. By high-dimensionality, the datasets are represented in terms of large random matrices. These data matrices can be viewed as data points in high-dimensional vector space—each vector is very long.

Data-driven approaches and data utilization for smart grids are current stressing topics, as evidenced in the special issue of ''Big Data Analytics for Grid Modernization'' [1]. This special issue is most relevant to this paper in spirit. Several SA topics are discussed. We highlight the anomaly detection and classification [4], [5], the estimation of active ingredients such as PV installations [6], [7], and the online transient stability evaluation using real-time data [8].

In addition, some researches are concerned with the improvement in wide-area monitoring, protection and control (WAMPAC) and the utilization of PMU data [9]–[11], together with the fault detection and location [12], [13]. Xie *et al.* [14] based on principal component analysis (PCA), propose an online application for early event detection by introducing a reduced dimensionality. Lim *et al.* [15] based on singular value decomposition (SVD), study the quasi-steady-state operational problem relevant to the voltage instability phenomenon. Their work has a special connection to this paper.

A. CONTRIBUTIONS OF THIS PAPER

Randomness is critical to future grids since rapid fluctuations in voltages and currents are ubiquitous. Often, these fluctuations exhibit Gaussian statistical properties [15]. The central interest in this paper is to model these rapid fluctuations using the framework of random matrix theory (RMT). This new algorithms are made possible due to the *latest breakthroughs* in probability on the central limit theorems of the linear eigenvalue statistics (LESs) [16, Ch. 7]. See [17] for a recent review.

1) Starting from fundamental formulas of power systems, a theoretical justification is given for the validity of modeling complex grids as large random matrices. This data modeling framework ties together the RMT and the power system analysis. This part is basic in nature.
2) This paper studies numerous basic problems including the technical route and applied framework, data-processing and relevant procedures, evaluation system and indicator sets, and the advantages over classical methodologies.
3) This paper makes a comparison between RMT-based approach and PCA-based one.
4) On the basis of big data analytics, this paper studies some power system applications: anomaly detection and location, empirical spectral density test, sensitivity analysis, statistical indicator system and its visualization, and, finally, robustness against asynchronous data.

B. RELATIONSHIP TO OUR PREVIOUS WORK

The work [2] is the first attempt to introduce the mathematical tool of RMT into power systems. Later, numerous papers demonstrate the power of this tool. Ring Law and Marchenko-Pastur (M-P) Law are regarded as the statistical foundation, and Mean Spectral Radius (MSR) is proposed as the high-dimensional indicator. Then we move forward to the second stage—paper [18] studies the correlation analysis under the above framework. The concatenated matrix $\mathbf{A}_i$ is the object of interest. It consists of the basic matrix $\mathbf{B}$ and a factor matrix $\mathbf{C}_i$, i.e., $\mathbf{A}_i = [\mathbf{B}; \mathbf{C}_i]$. In order to seek the sensitive factors, we compute the advanced indicators that are based on the LESs of these concatenated matrices $\mathbf{A}_i$. This study contributes to fault detection and location, line-loss reduction, and power-stealing prevention. Based on the same theoretical foundation, analysis for power transmission equipment is also conducted [19]. Paper [20] is the third step in which the LES set is studied. Based on the LES set, a statistical and data-driven indicator system, rather than its deterministic and model-based counterpart, is built to describe the system from a high-dimensional perspective. The robustness against spatial data error, precisely, data losses in the core area, is emphasized.

C. ADVANTAGES OF RMT-BASED APPROACH

The data-driven approach conducts analysis requiring no prior knowledge of the system topology, the unit operation/control mechanism, the causal relationship, etc. Comparing with classical data-driven methodologies such as PCA-based method, the RMT-based counterpart has some unique advantages:

1) The massive dataset of power systems are in a high-dimensional vector space; the temporal variations (T sampling instants) are simultaneously observed together with spatial variations (N grid nodes). The extraction of information from the above temporal-spatial variations is a challenge that does not meet the prerequisites of most classical statistical algorithms. Unifying time and space through their ratio $c = T/N$, RMT deal with such kind of data mathematically rigorously.

2) The statistical indicator is generated from all the data in the form of matrix entries. This is not true to principal components; the rank of the covariance matrix is unknown. The large size of the data enhances the robustness of the final decision against the bad data (inaccuracy, losses, or asynchronization), as well as those inevitable challenges in classical data-driven methods, such as error accumulations and spurious correlations [18].

3) For the statistical indicator, a theoretical or empirical value is obtained in advance. The statistical indicator such as LES follows a Gaussian distribution, and its variance is bounded [21] and decays very fast in the order of $O(N^{-2})$ for a given data dimension N, say $N = 118$.

4) The proposed approach can flexibly handle heterogenous data to realize data fusion via matrix operations, such as the blocking [2], the sum [22], the product [22], and the concatenation [18] of the matrices. Data fusion is guided by the latest mathematical research [16, Ch. 7].

5) Only eigenvalues are used for further analyses, while the eigenvectors are omitted. This leads to a much faster data-processing speed and less required memory space. Although some information is lost, there is still rich information contained in the eigenvalues [23], especially those outliers [24], [25].

6) Particularly, for a certain RMM, various forms of LES, in the form of $\tau_{\mathrm{F}} = \sum_{i=1}^{N} \varphi_F \left(\lambda_{\mathbf{M},i}\right)$, can be constructed by designing test functions $\varphi_F(\cdot)$ without introducing any system error. Each LES, similar to a filter, provides a unique view-angle. As a result, the system is understood piece by piece. Besides, some specific signal can be detected and tracked using LES technologies.

Section II gives the mathematical background and theoretical foundation. Spectrum test is introduced as a novel tool. Section III studies the details about the RMT-based method. Section IV and Section V, using the simulated data and field data respectively, study the function designing based on the proposed method. Section VI concludes this paper.

II. MATHEMATICAL BACKGROUND AND THEORETICAL FOUNDATION

A. RANDOM MATRIX MODELING

Operating in a balance situation, power grids obey

$$\begin{cases} \Delta P_i = P_{is} - P_i(\mathbf{V}, \boldsymbol{\theta}) \\ \Delta Q_i = Q_{is} - Q_i(\mathbf{V}, \boldsymbol{\theta}), \end{cases} \tag{1}$$

where P_{is} and Q_{is} are the power injections of node i, and $P_i(\mathbf{V}, \boldsymbol{\theta})$ and $Q_i(\mathbf{V}, \boldsymbol{\theta})$ are the power injections of the network, satisfying

$$\begin{cases} P_i = V_i \sum_{j=1}^{n} V_j \left(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}\right) \\ Q_i = V_i \sum_{j=1}^{n} V_j \left(G_{ij}\sin\theta_{ij} - B_{ij}\cos\theta_{ij}\right). \end{cases} \tag{2}$$

Combining (1) and (2), we obtain

$$\mathbf{w}_0 = f(\mathbf{x}_0, \mathbf{y}_0), \tag{3}$$

where $\mathbf{w}_0$ is the vector of nodes' power injections depending on P_{is}, Q_{is}, $\mathbf{x}_0$ is the system status variables depending on V_i, θ_i, and $\mathbf{y}_0$ is the network topology parameters depending on B_{ij}, G_{ij}.

Then, the system fluctuations, thus randomness in datasets, are formulated as

$$\mathbf{w}_0 + \Delta\mathbf{w} = f(\mathbf{x}_0 + \Delta\mathbf{x}, \mathbf{y}_0 + \Delta\mathbf{y}). \tag{4}$$

With a Taylor expansion, (4) is rewitten as

$$\begin{aligned} \mathbf{w}_0+\Delta\mathbf{w} &= f(\mathbf{x}_0, \mathbf{y}_0) + f'_{\mathbf{x}}(\mathbf{x}_0, \mathbf{y}_0)\Delta\mathbf{x} + f'_{\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)\Delta\mathbf{y} \\ &\quad + \frac{1}{2}f''_{\mathbf{xx}}(\mathbf{x}_0, \mathbf{y}_0)(\Delta\mathbf{x})^2 + \frac{1}{2}f''_{\mathbf{yy}}(\mathbf{x}_0, \mathbf{y}_0)(\Delta\mathbf{y})^2 \\ &\quad + f''_{\mathbf{xy}}(\mathbf{x}_0, \mathbf{y}_0)\Delta\mathbf{x}\Delta\mathbf{y} + \cdots. \end{aligned} \tag{5}$$

The value of system status variables $\mathbf{x}$ are relatively stable, which means that the second-order term $(\Delta\mathbf{x})^2$ and higher-order terms are ignorable. Besides, (2) shows that $f''_{\mathbf{yy}}(\mathbf{x}, \mathbf{y}) = 0$. As a result, (5) is turned into

$$\begin{aligned} \Delta\mathbf{w} &= f'_{\mathbf{x}}(\mathbf{x}_0, \mathbf{y}_0)\Delta\mathbf{x} + f'_{\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)\Delta\mathbf{y} \\ &\quad + f''_{\mathbf{xy}}(\mathbf{x}_0, \mathbf{y}_0)\Delta\mathbf{x}\Delta\mathbf{y}. \end{aligned} \tag{6}$$

Suppose the network topology is unchanged, i.e., $\Delta\mathbf{y} = 0$. From (6), it is deduced that

$$\Delta\mathbf{x} = \left(f'_{\mathbf{x}}(\mathbf{x}_0, \mathbf{y}_0)\right)^{-1}(\Delta\mathbf{w}) = \mathbf{S}_0\Delta\mathbf{w}. \tag{7}$$

On the other hand, suppose the power demands is unchanged, i.e., $\Delta\mathbf{w} = 0$. From (6), it is deduced that

$$\Delta\mathbf{x} = \mathbf{S}_0\Delta\mathbf{w}_y, \tag{8}$$

where $\mathbf{w}_y = [\mathbf{I} + f''_{\mathbf{xy}}(\mathbf{x}_0, \mathbf{y}_0)\Delta\mathbf{y}\mathbf{s}_0]^{-1}[f'_{\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)]$.

Note that $\mathbf{S}_0 = \left(f'_{\mathbf{x}}(\mathbf{x}_0, \mathbf{y}_0)\right)^{-1}$, i.e., the inversion of the Jacobian matrix $\mathbf{J}_0$.

Thus, the power system operation is modeled in the form of random matrices. If there exists an unexpected active power change or short circuit, the corresponding change of system status variables $\mathbf{x}_0$, i.e. V_i, θ_i, will obey (7) or (8) respectively.

For a practical system without dramatic changes, rich statistical empirical evidence indicates that the Jacobian matrix $\mathbf{J}$ keeps nearly constant, so does $\mathbf{s}_0$. Considering T random vectors observed at time instants $i = 1, \cdots, T$, the relationship is built in the form of $\Delta\mathbf{X_s} = \mathbf{S}_0\Delta\mathbf{W}$ with a similar procedure as (3) to (8), where $\Delta\mathbf{X_s}$ denotes the variation of state $[\Delta\mathbf{x}_1, \cdots, \Delta\mathbf{x}_T]$, and $\Delta\mathbf{W}$ denotes the variation of power injections or topology parameters accordingly.

Taking the case in [20] as an example, for an equilibrium operation system (the topology is unchanged, the reactive power is almost constant or changes much more slowly than the active one), the relationship model between voltage magnitude and active power is just like the Multiple Input Multiple Output (MIMO) model in wireless communication [16], [22]. We write $\mathbb{V} = \Xi\mathbb{P}$. Note that most variables of vector $\mathbb{V}$ are random due to the ubiquitous noises, e.g., small random fluctuations in $\mathbb{P}$. Furthermore, with the normalization, the standard random matrix model (RMM) is built in the form of $\tilde{\mathbb{V}} = \tilde{\Xi}\mathbf{R}$, where $\mathbf{R}$ is a standard Gaussian random matrix.

B. ANOMALY DETECTION BASED ON ASYMPTOTIC EMPIRICAL SPECTRAL DISTRIBUTION

Often, these rapid fluctuations exhibit Gaussian statistical properties [15], as pointed out above. In practice, Gaussian unitary ensemble (GUE) and Laguerre unitary ensemble (LUE) are used in the proposed models:

$$\mathbf{A} = \begin{cases} \frac{1}{2}\left(\mathbf{X} + \mathbf{X}^H\right), & \mathbf{X} \in \mathbb{X}^{N\times N}, \quad \text{GUE;} \\ \frac{1}{N}\mathbf{X}\mathbf{X}^H, & \mathbf{X} \in \mathbb{X}^{N\times T}, \quad \text{LUE.}, \end{cases} \tag{9}$$

where $\mathbf{X}$ is the standard Gaussian random matrix whose entries are independent identically distributed (i.i.d.) complex Gaussian random variables.

Let $f_A(x)$ be the empirical density of A, and define its empirical spectral distribution (ESD) $F_A(x)$:

$$F_A(x) = \frac{1}{N}\sum_{i=1}^{N} I_{\{\lambda_i \le x\}}, \tag{10}$$

where A is GUE or LUE matrix, and $I(\cdot)$ represents the event indicator function. We investigate the rate of convergence of the expected ESD $\mathbb{E}\{F_A(x)\}$ to the Wigner's Semicircle Law or Wishart's M-P Law.

Let $g_A(x)$ and $G_A(x)$ denote the true eigenvalue density and the true spectral distribution of A, and the Wigner's Semicircle Law and Wishart's M-P Law say:

$$g_A(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4 - x^2}, & x \in [-2, 2], \quad \text{GUE;} \\ \frac{1}{2\pi c x}\sqrt{(x-a)(b-x)}, & x \in [a, b], \quad \text{LUE;}, \end{cases} \tag{11}$$

where $a = \left(1-\sqrt{c}\right)^2$, $b = \left(1+\sqrt{c}\right)^2$.

$$G_A(x) = \int_{-\infty}^{x} g_A(u)\, du. \tag{12}$$

Then, we denote the Kolmogorov distance between $\mathbb{E}\{F_A(x)\}$ and $G_A(x)$ as Δ:

$$\Delta = \sup_x \left| \mathbb{E}\{F_A(x)\} - G_A(x) \right|. \tag{13}$$

Gotze and Tikhomirov [26], in their work, prove an optimal bound for Δ of order $O\left(N^{-1}\right)$.

Lemma 2.1: There exists a positive constant C such that, for any $N \geq 1$,

$$\Delta \leq CN^{-1}. \tag{14}$$

They also prove that the convergence of the density of standard Semicircle Law and M-P Law to the expected spectral density $f_A(x)$ satisfies following lemmas.

Lemma 2.2: For GUE matrix, there exists a positive constant ε *and* C *such that, for any* $x \in \left[-2+N^{-\frac{1}{3}}\varepsilon, 2-N^{-\frac{1}{3}}\varepsilon\right]$,

$$\left| f_A(x) - g(x) \right| \leq \frac{C}{N\left(4-x^2\right)}. \tag{15}$$

Lemma 2.3: For LUE matrix, let $\beta = N/T$, *there exists some positive constant* β_1 *and* β_2 *such that* $0 < \beta_1 \leq \beta \leq \beta_2 < 1$, *for all* $N \geq 1$. *Then there exists a positive constant* C *and* ε *depending on* β_1 *and* β_2 *and for any* $N \geq 1$ *and* $x \in \left[a+N^{-\frac{2}{3}}\varepsilon, b-N^{-\frac{2}{3}}\varepsilon\right]$,

$$\left| f_A(x) - h(x) \right| \leq \frac{C}{N(x-a)(b-x)}. \tag{16}$$

Lemma 2.2 and 2.3 also describe how fast the population distribution functions converge to the asymptotic ESD limit. This ESD-based test is interesting for anomaly detection about a complex grid; the effectiveness is validated in Section IV. We exploit the mathematical knowledge that the ESD converges to its limit with a optimal convergence rate of N^{-1}.

III. THE METHOD OF SITUATION AWARENESS

A. TECHNICAL ROUTE AND PRACTICAL PROCEDURES

The proposed RMT-based method consists of three procedures as illustrated in Fig. 1: 1) big data model—to model the system using experimental data for the RMM; 2) big data analysis—to conduct big data anlytics for the indicator system; 3) engineering interpretation—to visualize and interpret the statistical results to operators for decision-making.

This method is universal. Numerous successful attempts have already be made in the field of anomaly detection and diagnosis for both the grid network [2], [18] and the transmission equipment [19]. In addition, [27] and [28] based on RMT, study the steady stability and transient stability.

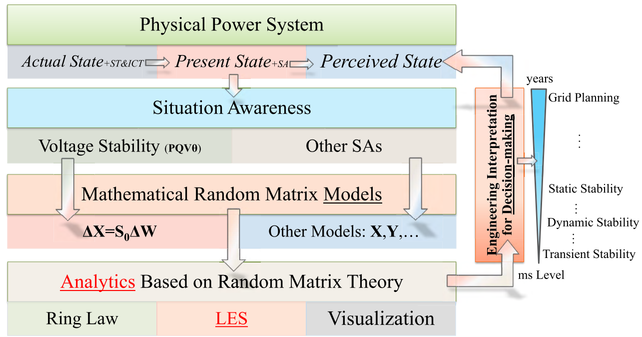


FIGURE 1. RMT-based Method for SA.

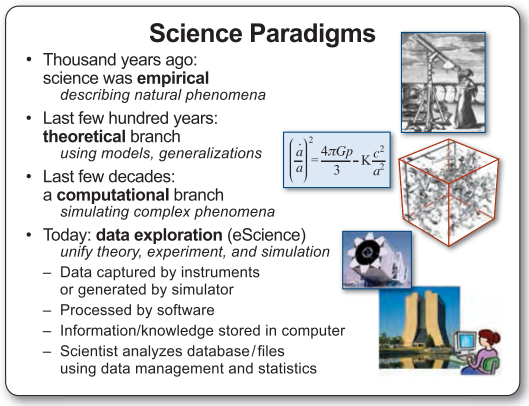


FIGURE 2. Science paradigms [30].

B. PARADIGMS AND METHOD

Fig. 2 in [29] is referred as a clue. It is the age of 4th-paradigm—data-intensive scientific discovery. Besides, the summaries for the classical decision-making approaches and for the proposed ones, obtained initially in [2], are improved as Fig. 3.

The second and third paradigms are typically model-based—they use equations, formulas, and simulations to describe the system. The blue line in Fig. 3 depicts the general procedure and corresponding tools. These tools cannot deal with massive data due to the essence of mechanism models—the models are in low dimensions, leading to deterministic results which are fully dependent upon only a few parameters.[1] It may cause inefficient or even incorrect big data analytics. For instance, only under ideal conditions, is the wind power proportional to the cube of wind speed. Moreover, some physical parameters, e.g., admittance matrixes, will introduce system error due to the ubiquitous randomness and uncertainty.

Under classical statistical framework, only two typical data matrices in the form of $\mathbf{X} \in \mathbb{R}^{N\times T}$ are at our disposal: 1) N, T are small, and 2) N is small, T is very large (compare with N). This prerequisite greatly restricts the utilization of the

[1] E.g., $y = ax^2+bx+c$ is a 3-dimensional model; the relationship between x and y fully depends on a, b, and c.

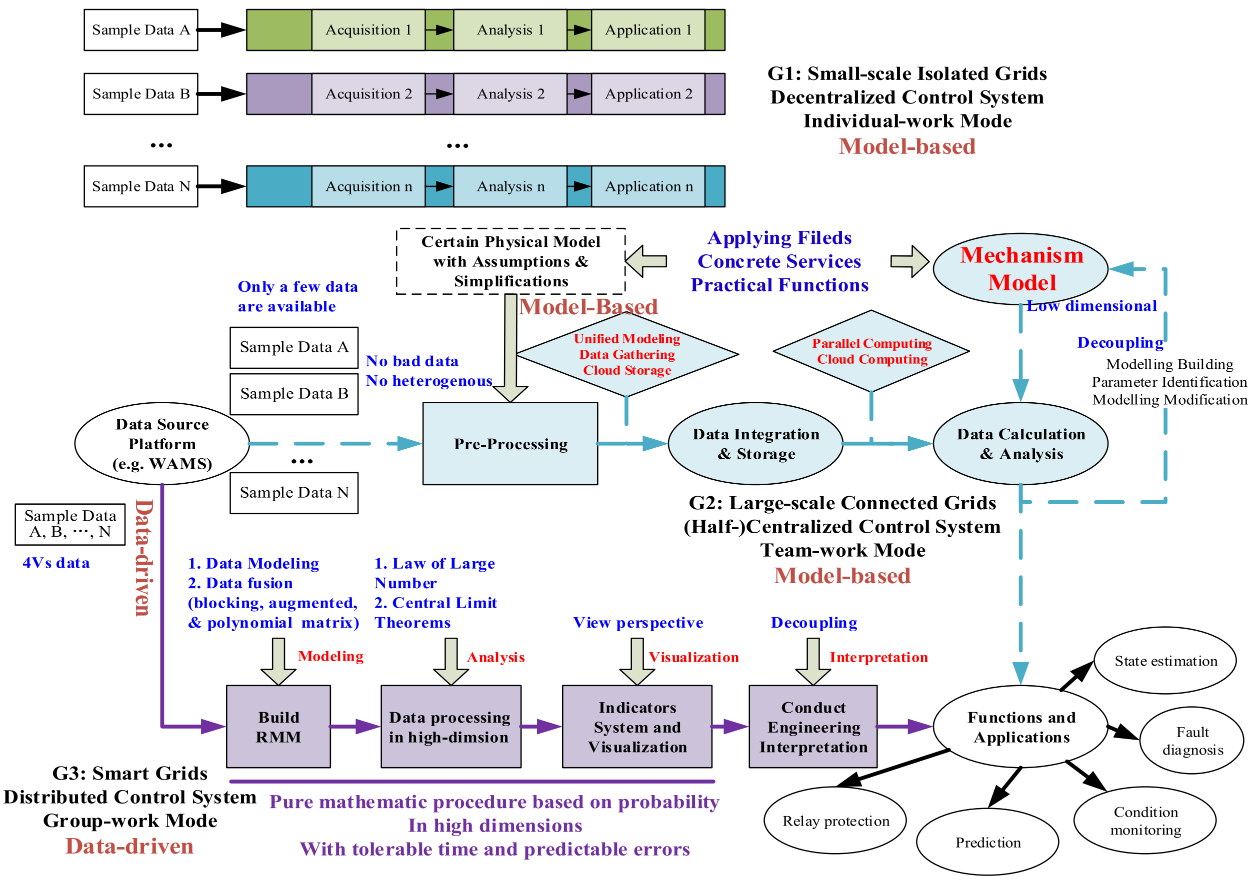


FIGURE 3. Data utilization method for power systems. The above, middle, and below parts indicate the data processing procedures and the work modes for G1, G2, and G3, respectively.

massive data; we should enable more data to speak for themselves [31]. In other words, model-based framework is not able to turn massive data into useful big data analytics. Although these massive data can contribute to model improvement and parameters correction, we can hardly conduct analysis more precisely with extremely large data volumes. Even worse, more data mean more errors; if those bad data are taken into the fixed model, poor results are obtained almost surely. Besides, the bias, caused by challenges such as error accumulations and spurious correlations, will not be alleviated via a low-dimensional procedure [18]—the dimensions of the procedure are limited by the dimensions of the model. The belief that data-driven mode is adapted to the future grid's analysis agrees with the core viewpoint of the 4th-paradigm. The classical data utilization methodology needs be revisited.

C. CLASSICAL DIMENSIONALITY REDUCTION ALGORITHM–PCA

Data-driven methodology is an alternative; it is model-free and able to process massive data in a holistic way. Principal component analysis (PCA) is one of the classical data processing algorithms which are sensitive to relative scaling original variables. It uses an orthogonal transformation to convert a set of possibly correlated raw variables into a set of linearly uncorrelated variables called principal components. The number of principal components is often much less than the number of original variables. In [14], PCA is used for dimensionality reduction from 14 PMU datasets to extract the event indicators. For PCA, the procedure consists of three parts: 1) Singular Value Decomposition (SVD) [15], 2) Projection, and 3) Indicators.

This procedure is applied to conduct early event detection; details can be found in [14]. The comparison between the PCA-based approach and the RMT-based approach, and the advantages of the later are summarized in I-C.

D. DATA-DRIVEN APPROACH BASED ON RANDOM MATRIX THEORY

The framework of RMT-based approach starts with the use of sample covariance matrix to replace the true covariance matrix. It is well known that this replacement is far from optimal. The almost optimal estimation of large covariance matrices using tools from RMT [32] can be used, instead. The procedure based on RMT is outlined below.

1) RING LAW AND MSR

Ring Law Analysis conducts SA as follows:

Steps of Ring Law Analysis

1) Select arbitrary raw data (or all available data) as data source Ω.
2) At a certain time t_i, form $\hat{\mathbf{X}}$ as random matrix.
3) Obtain $\tilde{\mathbf{Z}}$ by matrix transformations ($\hat{\mathbf{X}} \rightarrow \tilde{\mathbf{X}} \rightarrow \mathbf{X}_u \rightarrow \mathbf{Z} \rightarrow \tilde{\mathbf{Z}}$ [2]).
4) Calculate eigenvalues $\lambda_{\tilde{\mathbf{Z}}}$ and plot the Ring on the complex plane.
5) Conduct high-dimensional analysis.
 5a) Observe the experimental ring and compare it with the reference.
 5b) Calculate $\tau_{\text{MSR}} = \sum_{i=1}^{N} \left|\lambda_{\mathbf{Z},i}\right| / N$ as the *statistical indicators*.
 5c) Compare τ_{MSR} with the theoretical value $\mathbb{E}(\tau_{\text{MSR}})$.
6) Repeat 2)-5) at the next time point ($t_i = t_i + 1$).
7) Visualize τ on the time series, i.e. draw τ–t curve.
8) Make engineering explanations.

In *Steps 2–7*, with a high-dimensional procedure, one conducts SA without any prior knowledge, assumption, or simplification. In *step 2*, arbitrary raw data, even those from distributed nodes or intermittent time periods, are at our disposal. The size of $\hat{\mathbf{X}}$ is controllable, and as a result the dimensionality curse is relieved in some ways.

2) M-P LAW AND LES

For the M-P Law Analysis, the steps are very similar, except for the following differences:

Partial Steps of M-P Law Analysis

3] Obtain $\mathbf{M}$ by matrix transformations ($\mathbf{M} = \frac{1}{N}\tilde{\mathbf{X}}\tilde{\mathbf{X}}^H$).
4] Calculate eigenvalues $\lambda_{\mathbf{M}}$.
 5b] Calculate $\tau = \sum_{i=1}^{N} \varphi\left(\lambda_{\mathbf{M},i}\right)$ as the *statistical indicators*.
 5c] Compare τ with the theoretical value $\mathbb{E}(\tau)$.

Notice that Ring Law maps the information from datasets to the complex plane ($\mathbb{C}^{N\times T} \mapsto \mathbb{C}$), while M-P law does this to the right half real-axis ($\mathbb{C}^{N\times T} \mapsto \mathbb{R}^{+}$). This fundamental difference plays a critical role in data visualization.

IV. CASE STUDIES USING SIMULATED DATA

A. BACKGROUND AND ASSUMPTION OF THE CASE

A standard IEEE 118-node system is adopted as Fig. 16, shown in Appendix A, and the events are assumed as Table 2, shown in Appendix B. Thus, the power demand on each node is obtained as the system injections (Fig. 4a), while the voltage is accessible as the operation status (Fig. 4b). Suppose that the power demand data is *unknown* or unqualified for SA due to the low sampling frequency or the bad quality. For further analysis, we just start with data source $\mathbf{\Omega_V}$: $\hat{v}_{i,j} \in \mathbb{R}^{118\times 2500}$ and assign the analysis matrix as $\mathbf{X} \in \mathbb{R}^{118\times 240}$ (4 minutes' time span). Firstly, category is conducted for the system operation status; the results are given

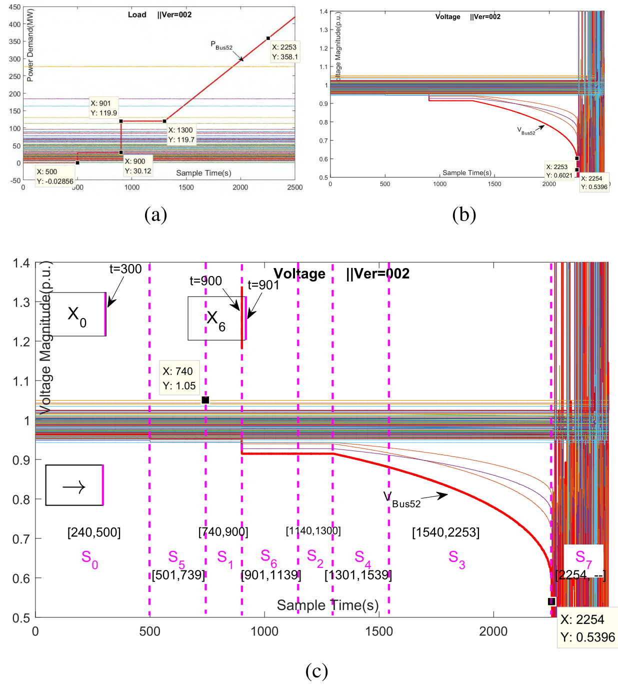


FIGURE 4. Background of Case 1. (a) Assumed event, unavailable. (b) Raw voltage, Ω_V for analysis. (c) Category for operation status.

in Fig. 4c. In general, according to the data feature (events on time-series) and the matrix length (time span, i.e., T), the operation status of the system is divided into 8 stages. Note that $S4$, $S5$, and $S6$ are transition stages, and their time span is right equal to the analysis matrix length minus one, i.e, $T - 1 = 239$. These stages are described as follows:

- For $S0, S1, S2$, white noises play a dominant part. $P_{\text{Node-52}}$ is rising in turn.
- For $S3$, $P_{\text{Node-52}}$ keeps a sustained and stable growth.
- $S4$, transition stage. Ramping signal exists.
- $S5$, **S**6, transition stages. Step signal exists.
- For $S7$, voltage collapse.

Two typical data sections, at stage $S0$ and $S6$ respectively, are selected: $\mathbf{X}_0 \in \mathbb{R}^{118\times 240}$, covering period $t = [61 : 300]$ and at sampling time $t_{\text{end}} = 300$, and 2) $\mathbf{X}_6 \in \mathbb{R}^{118\times 240}$, covering period $t = [662 : 901]$ and at sampling time $t_{\text{end}} = 901$.

B. ANOMALY DETECTION

1) BASED ON RING LAW AND M-P LAW

According [2], RMM $\tilde{\mathbb{V}}$ is build from the raw voltage data. Then, τ_{MSR} is employed as a statistical indicator to conduct anomaly detection. For the selected data section $\mathbf{X}_0$ and $\mathbf{X}_6$, their M-P Law and Ring Law Analysis are shown as Fig. 5a, 5b, 5c and 5d. With sliding-window, the τ_{MSR}-t curve is obtained as Fig. 5e.

Fig. 5 shows that when there is no signal in the system, the experimental RMM well matches Ring Law and M-P Law, and the experimental value of LES is approximately equal to the theoretical value. This validates the theoretical justification for modeling rapid fluctuation of

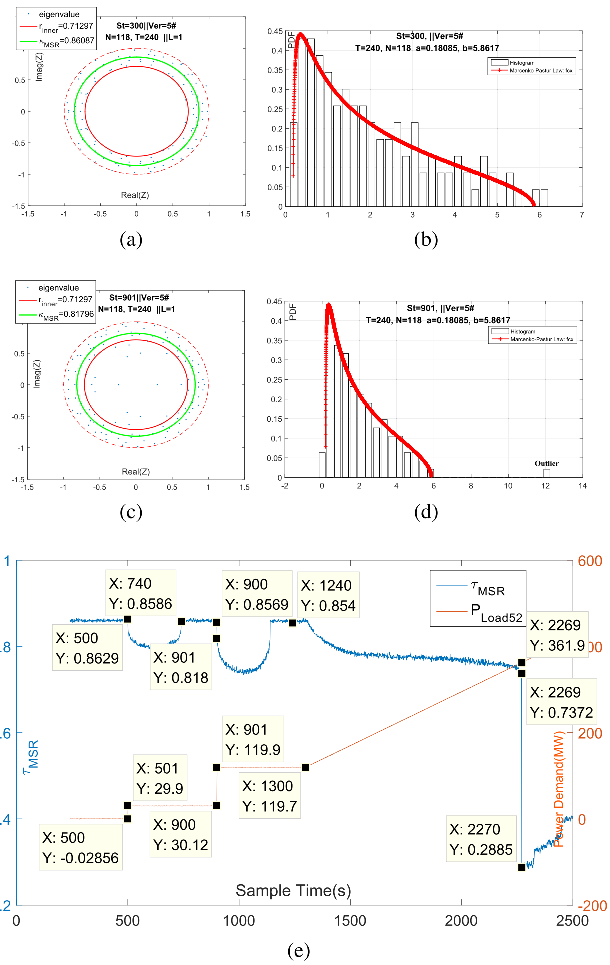


FIGURE 5. **Anomaly detection results. (a) Ring Law for X_0. (b) M-P Law for X_0. (c) Ring Law for X_6. (d) M-P Law for X_6. (e) τ_{MSR}-t curve using MSW method on time series.**

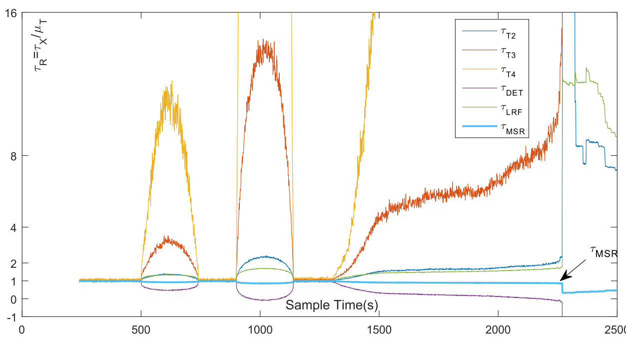


FIGURE 6. **Illustration of various LES indicators.**

each node using white Gaussian noises, as the description in Section II-A. On the other hand, Ring Law and M-P Law are violated at the very beginning ($t_{\text{end}} = 901$) of the step signal. Besides, the proposed high-dimensional indicator τ_{MSR}, is extremely sensitive to the anomaly—at $t_{\text{end}} = 901$, the τ_{MSR} starts the dramatic change (Fig. 5e, τ_{MSR}-t curve), while the raw voltage magnitudes are still in the normal range (Fig. 4c). Moreover, following [20], we design numerous kinds of LES τ and define $\mu_0 = \tau/\mathbb{E}(\tau)$. The detection results using results τ are shown in Fig. 6, proving that different

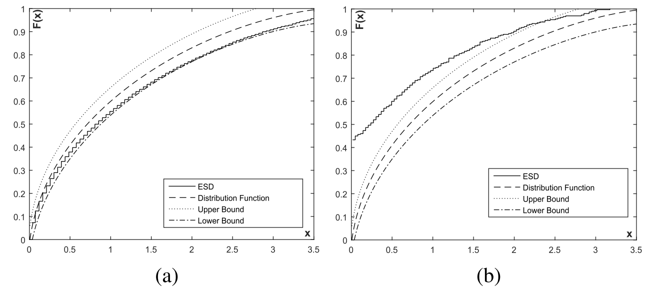


FIGURE 7. **Anomaly Detection Using LUE matrices. (a) ESD of $\mathbf{Y}_0$ (Normal). (b) ESD of $\mathbf{Y}_6$ (Abnormal).**

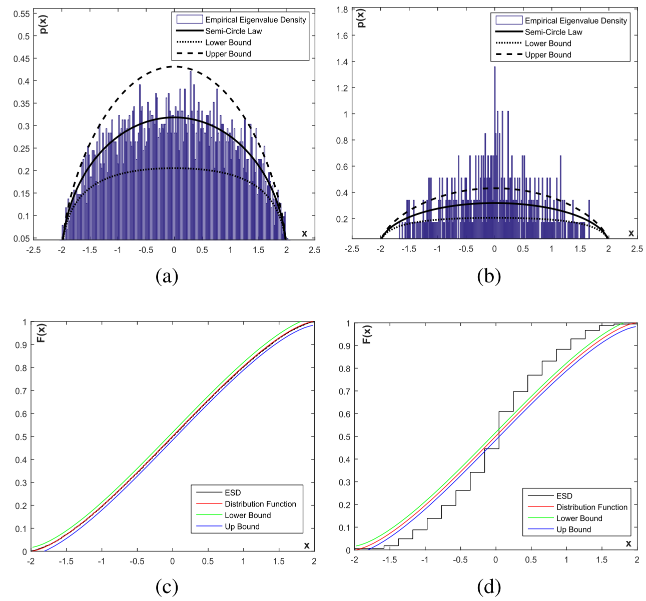


FIGURE 8. **Anomaly detection using GUE matrices. (a) Density of $\mathbf{Z}_0$ (Normal). (b) Density of $\mathbf{Z}_6$ (Abnormal). (c) ESD of $\mathbf{Z}_0$ (Normal). (d) ESD of $\mathbf{Z}_6$ (Abnormal).**

indicators have different effectiveness; this suggests another topic to explore in the future.

2) BASED ON SPECTRUM TEST

The sampling time is still set at $t_{\text{end}} = 300$ and $t_{\text{end}} = 901$. Following Lemma 2.2 and Lemma 2.3, $\mathbf{Y}_0, \mathbf{Y}_6 \in \mathbb{R}^{118\times 240}$ (span $t = [61 : 300]$ and $t = [662 : 901]$), and $\mathbf{Z}_0, \mathbf{Z}_6 \in \mathbb{R}^{118\times 118}$ (span $t = [183 : 300]$ and $t = [784 : 901]$) are selected. The analysis results are shown in Fig. 7 and Fig. 8. These results validate that empirical spectral density test is competent to conduct anomaly detection—when the power grid is under a normal condition, the empirical spectral density $f_A(x)$ and the ESD function $F_A(x)$ are almost strictly bounded between the upper bound and the lower bound of their asymptotic limits. On the other hand, these results also validate that GUE and LUE are proper mathematical tools to model the power grid operation.

C. STEADY STABILITY EVALUATION

The $V - P$ curve (also called nose curve) and the smallest eigenvalue of the Jacobian Matrix [15] are two clues

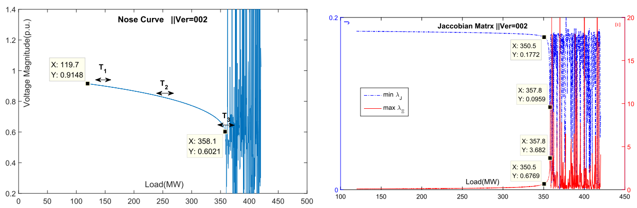


FIGURE 9. The $V-P$ curve and $\lambda-P$ curve. (a) $V-P$ Curve. (b) $\lambda-P$ Curve.

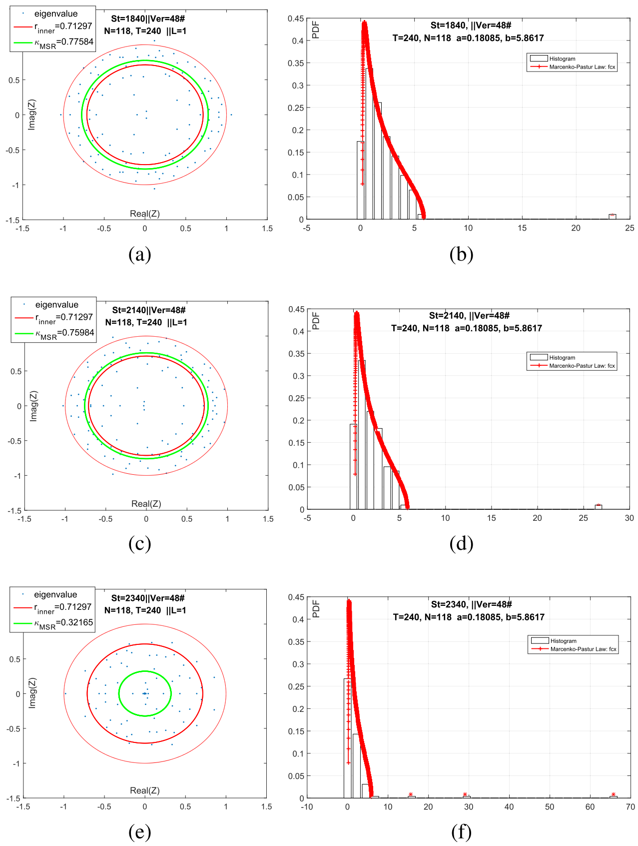


FIGURE 10. RMT-based results for voltage stability evaluation. (a) Ring Law for $\mathbf{T_1}$. (b) M-P Law for $\mathbf{T_1}$. (c) Ring Law for $\mathbf{T_2}$. (d) M-P Law for $\mathbf{T_2}$. (e) Ring Law for $\mathbf{T_3}$. (f) M-P Law for $\mathbf{T_3}$.

for steady stability evaluation. In this case, we focus on **E**4 stage during which $P_{\text{Node-52}}$ keeps increasing until the system exceeds its steady stability limit. The $V-P$ curve and $\lambda-P$ curve, respectively, are given in Fig. 9a and Fig. 9b. Furthermore, some data section are chosen, $\mathbf{T}_1$: [1601 : 1840]; $\mathbf{T}_2$: [1901 : 2140]; $\mathbf{T}_3$: [2101 : 2340], shown as Fig. 9a. The RMT-based results are shown as Fig. 10. The outliers become more evident as the stability degree decreases. The statistics of the outliers, in some sense, are similar to the smallest eigenvalue of Jacobian Matrix, Lyapunov Exponent or the entropy.

For further analysis, the signal and stage division are taken into account. In general, sorted by the stability degree, the stages are ordered as $S0 > S1 > S2 \gg \max(S3, S4, S5) > \min(S3, S4, S5) \gg S6 \gg S7$.

TABLE 1. Indicator of various LESs at each stage.

	MSR	T_2	T_3	T_4	DET	LRF
E_0: Theoretical Value						
$\mathbb{E}(\tau)$	0.8645	1338.3	10069	8.35E4	48.322	73.678
$S0$ [0240:0500, 261]: Small fluctuations around 0 MW ①						
$\overline{\tau_{\mathbf{X}}}_{\mathrm{R}}$	0.995	1.010	1.040	1.080	0.959	1.014
$S5$ [0501:0739, 239]: A step signal (0 MW ↑ 30 MW) is included ④						
$\overline{\tau_{\mathbf{X}}}_{\mathrm{R}}$	0.9331	1.280	2.565	7.661	0.5453	1.284
$S1$ [0740:0900, 161]: Small fluctuations around 30 MW ②						
$\overline{\tau_{\mathbf{X}}}_{\mathrm{R}}$	0.9943	1.010	1.039	1.084	0.9568	1.015
$S6$ [0901:1139, 239]: A step signal (30 MW ↑ 120 MW) is included ⑦						
$\overline{\tau_{\mathbf{X}}}_{\mathrm{R}}$	0.8742	2.054	1.06E1	7.22E1	7E−2	1.597
$S2$ [1140:1300, 161]: Small fluctuations around 120 MW ③						
$\overline{\tau_{\mathbf{X}}}_{\mathrm{R}}$	0.9930	1.019	1.067	1.135	0.9488	1.021
$S4$ [1301:1539, 239]: A ramp signal (119.7 MW ↗) is included ④						
$\overline{\tau_{\mathbf{X}}}_{\mathrm{R}}$	0.9337	1.295	2.787	9.615	0.5316	1.294
$S3$ [1540:2253, 714]: Steady increase (↗ 358.1 MW) ④						
$\overline{\tau_{\mathbf{X}}}_{\mathrm{R}}$	0.8906	1.717	6.530	3.48E1	0.1483	1.545
$S7$ [2254:2500, 247]: Static voltage collapse (361.9 MW ↗) ⑧						
$\overline{\tau_{\mathbf{X}}}_{\mathrm{R}}$	0.4259	1.02E1	2.11E2	4.65E3	−1.4E1	1.08E1

$^*\overline{\tau_{\mathbf{X}}}_{\mathrm{R}} = \overline{\tau_{\mathbf{X}}}/\mathbb{E}(\tau)$.

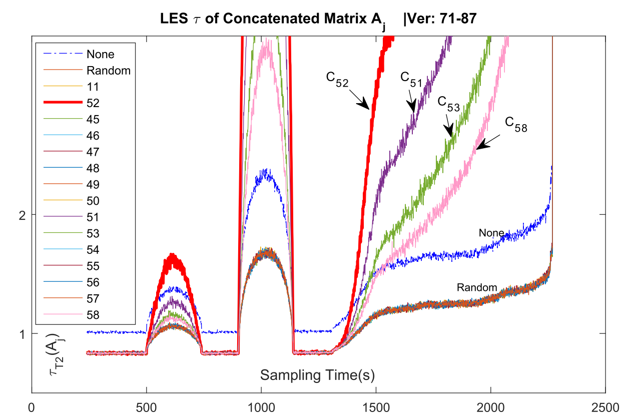


FIGURE 11. Sensitivity analysis based on concatenated matrix.

According to Fig. 6, Table 1 is obtained. The high-dimensional indicators $\overline{\tau_{\mathbf{X}}}_{\mathrm{R}}$ has the same trend as the stability degree order. These statistics have the potential for data-driven stability evaluation. Moreover, based on the Gaussian property of LES indicators, hypothesis tests are designed for the anomaly detection; see [33] for details.

D. CORRELATION ANALYSIS

The key for correlation analysis is the concatenated matrix $\mathbf{A}_i$, which consist of two part—the basic matrix $\mathbf{B}$ and a certain factor matrix $\mathbf{C}_i$, i.e., $\mathbf{A}_i = [\mathbf{B}; \mathbf{C}_i]$. For details, see [18]. The LES of each $\mathbf{A}_i$ is computed in parallel, and Fig. 11 shows the results.

In Fig. 11, the blue dot line (marked with None) shows the LES of basic matrix $\mathbf{B}$, and the orange line (marked with Random) shows the LES of the concatenated matrix $[\mathbf{B}; \mathbf{R}]$ ($\mathbf{R}$ is the standard Gaussian Random Matrix). Fig. 11 demonstrates that: 1) node 52 is the causing factor of the anomaly;

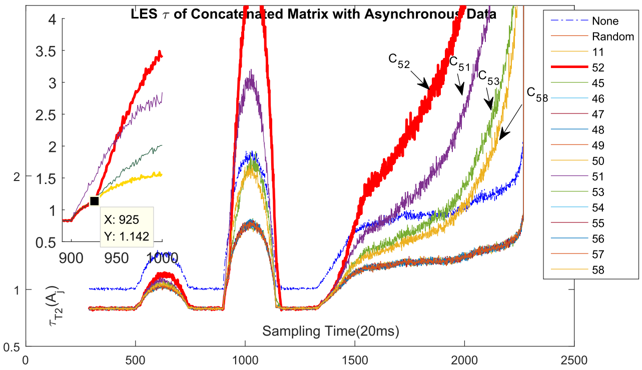


FIGURE 12. Situation awareness with asynchronous data.

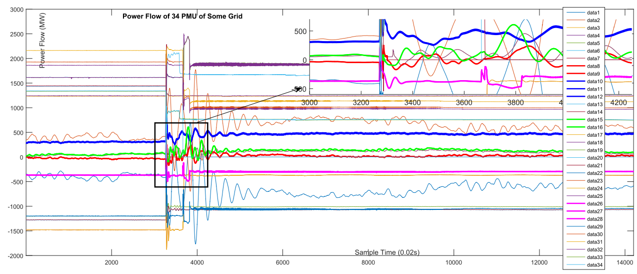


FIGURE 13. Raw power flow data of 34 PMUs.

2) sensitive nodes are 51, 53, and 58; and 3) nodes 11, 45, 46, etc, are not affected by the anomaly. Based on this algorithm, it is able to conduct behavior analysis, e.g., detection and estimation of residential PV installations [6]. It is another hot topic which is expanded in our research [33].

E. SA WITH ASYNCHRONOUS DATA

The proposed data-driven method is robust against bad data both in space and in time. He *et al.* [20] have successfully conducted SA with data loss in the core area. This paper stuides SA with asynchronous data. It is common that asynchronous data exists in the data platforms such as SCADA and WAMS. The problem is mainly caused by erroneous time-tags or communication delays. Sometimes, for a certain signal, the proper delay protection or interaction/response mechanism may also lead to asynchronous data. It is hard to measure or even to detect the time delay via traditional methods. The proposed approach has a special meaning here.

Using the simulated data, we make an artificial delay of 25 sampling points for 7 nodes—11, 14, 50, 52, 53, 77, and 81. With the concatenation operation introduced above, the results is obtained as shown in Fig. 12. It is an interesting discovery that the approach is robust against asynchronous data: 1) the anomalies are detected at $t = 501$ and $t = 901$; 2) Node 52 is the most sensitive node; 3) with more detailed observation, it is even able to quantitatively draw the conclusion that there exists a 25 sampling points delay $(925 - 900)$ for Node 52. It is surprising that the exact delay value can be recovered for the particular node! The power of the proposed approach is vividly exhibited here.

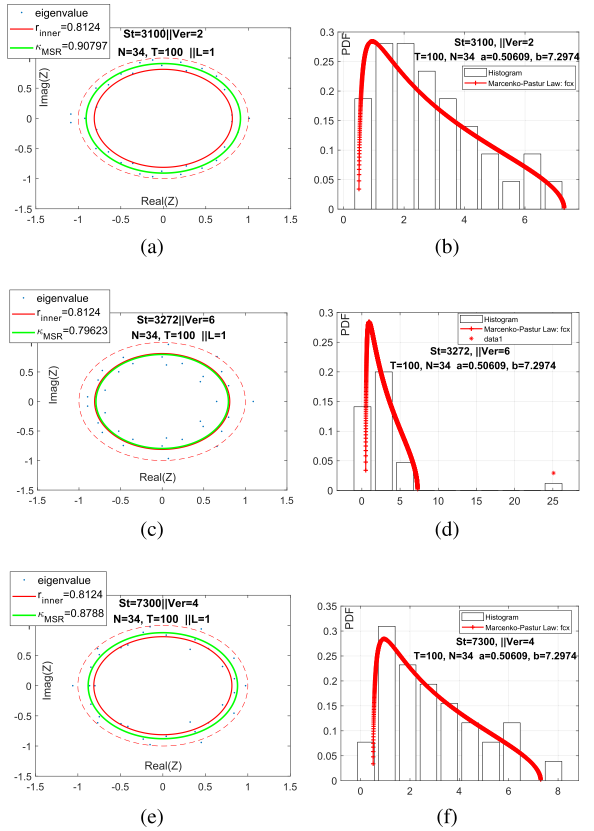


FIGURE 14. Ring Law and M-P Law for the fault. (a) Pre-fault: Ring Law. (b) Pre-fault: M-P Law. (c) During fault: Ring Law. (d) During fault: M-P Law. (e) Post-fault: Ring Law. (f) Post-fault: M-P Law.

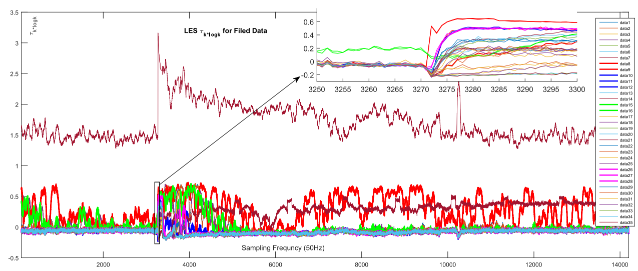


FIGURE 15. LES $t - \tau$ curves.

V. CASE STUDIES USING FIELD DATA

Some power grid of China is selected, with 34 PMUs collecting power flow data. The raw data are shown as Fig. 13; it is quite obvious that the fault begins at sampling time $t_s = 3271$. The ring distribution and M-P law pre-fault $(3101 - 3100)$, during fault $(3173 - 3272)$, and post-fault $(7201 - 7300)$ are given as Fig 14. This implies that the real-world data do follow Ring Law and M-P Law under normal condition, and they violate these laws when the fault is occurring. Moreover, the LES $t - \tau$ curves of basic matrix $\mathbf{B}$ and concatenated matrix $\mathbf{C}_i$ are obtain as Fig. 15. It shows that Node 8, 9, 26, 27, 28, 10, 11, and 12 are most relevant to this fault; while Node $1 - 7$ are not so sensitive.

VI. CONCLUSION

This paper has made significant progress on the basis of the previous work in the context of big data analytics for future grids. Randomness and uncertainty are at the heart of this data modeling and analysis. The approach exploits the massive spatial-temporal datasets of power systems. Random matrix theory (RMT) appears very natural for the problem at hands; in a random matrix of $\mathbb{C}^{N\times T}$, we use N nodes to represent the spatial nodes and T data samples to represent the temporal samples. When the number of nodes N is large, very unique mathematical phenomenon occurs, such as concentration of measure. Phase transition as a function of data size N is a result of this deep mathematical phenomenon. This is the very reason why the proposed algorithms are so powerful in practice.

Explicitly expressed in forms of linear eigenvalue statistics (LESs), the proposed RMT-based algorithms have numerous unique advantages. In the form of a large random matrix, they handle massive data that are in high dimensions and within a wide time span all at once. The trick is to treat these data *as a whole* at the disposal of RMT. In this way, highly reliable decisions are still attainable with some imperfect data, e.g., the asynchronous data. Moreover, with the statistical processing such as test function setting, the proposed data-driven approach has the potential to balance the perspectives of the speed, the sensitivity, and the reliability in practice.

The stability evaluation and behavior analysis are two big topics along this direction. Besides, the statistical indicators are good starting points for artificial intelligence and machine learning. For example, we can extract the linear eigenvalue statistics as features; those extracted features are used for further data processing in the pipeline using algorithms such as random forest, decision trees, and support vector machine.

APPENDIX A

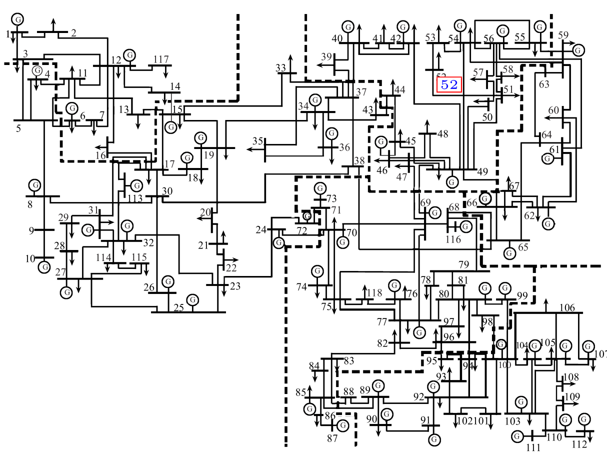


FIGURE 16. Partitioning network for the IEEE 118-node system.

APPENDIX B

The power demand of other nodes are assigned as

$$\tilde{y}_{\text{load}\_nt} = y_{\text{load}\_nt} \times (1 + \gamma_{\text{Mul}} \times r_1) + \gamma_{\text{Acc}} \times r_2, \quad (17)$$

TABLE 2. Series of events.

Stage	$E1$	$E2$	$E3$	$E4$
Time (s)	1–500	501–900	901–1300	1301–2500
$P_{\text{Node-52}}$ (MW)	0	30	120	$t/4 - 205$

P_{52} is the power demand of node 52.

where r_1 and r_2 are the element of standard Gaussian random matrix; $\gamma_{\text{Acc}} = 0.1$, $\gamma_{\text{Mul}} = 0.001$.

REFERENCES

[1] T. Hong *et al.*, "Guest editorial big data analytics for grid modernization," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2395–2396, Sep. 2016.

[2] X. He *et al.*, "A big data architecture design for smart grids based on random matrix theory," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 674–686, Mar. 2017.

[3] L. S. Moulin, A. P. Alves da Silva, M. A. El-Sharkawi, and R. J. Marks, II, "Support vector machines for transient stability analysis of large-scale power systems," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 818–825, May 2004.

[4] M. Rafferty, X. Liu, D. M. Laverty, and S. McLoone, "Real-time multiple event detection and classification using moving window PCA," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2537–2548, Sep. 2016.

[5] H. Jiang, X. Dai, D. W. Gao, J. J. Zhang, Y. Zhang, and E. Muljadi, "Spatial-temporal synchrophasor data characterization and analytics in smart grid fault detection, identification, and impact causal analysis," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2525–2536, Sep. 2016.

[6] X. Zhang and S. Grijalva, "A data-driven approach for detection and estimation of residential PV installations," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2477–2485, Sep. 2016.

[7] H. Shaker, H. Zareipour, and D. Wood, "A data-driven approach for estimating the power generation of invisible solar sites," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2466–2476, Sep. 2016.

[8] B. Wang, B. Fang, Y. Wang, H. Liu, and Y. Liu, "Power system transient stability assessment based on big data and the core vector machine," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2561–2570, Sep. 2016.

[9] A. G. Phadke, "The wide world of wide-area measurement," *IEEE Power Energy Mag.*, vol. 6, no. 5, pp. 52–65, Sep./Oct. 2008.

[10] V. Terzija *et al.*, "Wide-area monitoring, protection, and control of future electric power networks," *Proc. IEEE*, vol. 99, no. 1, pp. 80–93, Jan. 2011.

[11] L. Xie, Y. Chen, and H. Liao, "Distributed online monitoring of quasi-static voltage collapse in multi-area power systems," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2271–2279, Nov. 2012.

[12] Q. Jiang, X. Li, B. Wang, and H. Wang, "PMU-based fault location using voltage measurements in large transmission networks," *IEEE Trans. Power Del.*, vol. 27, no. 3, pp. 1644–1652, Jul. 2012.

[13] A. H. Al-Mohammed and M. A. Abido, "A fully adaptive PMU-based fault location algorithm for series-compensated lines," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2129–2137, Sep. 2014.

[14] L. Xie, Y. Chen, and P. R. Kumar, "Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2784–2794, Nov. 2014.

[15] J. M. Lim and C. L. DeMarco, "SVD-based voltage stability assessment from phasor measurement unit data," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2557–2565, Jul. 2016.

[16] R. Qiu and P. Antonik, *Smart Grid and Big Data*. Hoboken, NJ, USA: Wiley, 2015.

[17] R. C. Qiu, "Large random matrices and big data analytics," in *Big Data for Complex Network*. Boca Raton, FL, USA: CRC Press, 2016.

[18] X. Xu, X. He, Q. Ai, and R. C. Qiu, "A correlation analysis method for power systems based on random matrix theory," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1811–1820, Jul. 2017.

[19] Y. Yan *et al.*, "The key state assessment method of power transmission equipment using big data analyzing model based on large dimensional random matrix," in *Proc. CSEE*, vol. 36. Jan. 2016, pp. 435–445.

[20] X. He, R. C. Qiu, Q. Ai, L. Chu, and X. Xu. (Dec. 2015). "Linear eigenvalue statistics: An indicator ensemble design for situation awareness of power systems." [Online]. Available: https://arxiv.org/pdf/1512.07082.pdf

[21] M. Shcherbina. (Jan. 2011). "Central limit theorem for linear eigenvalue statistics of the wigner and sample covariance random matrices." [Online]. Available: https://arxiv.org/abs/1101.3249
[22] C. Zhang and R. C. Qiu, "Massive MIMO as a big data system: Random matrix models and testbed," *IEEE Access*, vol. 3, no. 4, pp. 837–851, Apr. 2015.
[23] J. R. Ipsen and M. Kieburg, "Weak commutation relations and eigenvalue statistics for products of rectangular random matrices," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 89, no. 3, 2014, Art. no. 032106.
[24] F. Benaych-Georges and J. Rochet, "Outliers in the single ring theorem," *Probab. Theory Rel. Fields*, vol. 165, pp. 313–363, Jun. 2016. [Online]. Available: http://dx.doi.org/10.1007/s00440-015-0632-x
[25] T. Tao, "Outliers in the spectrum of iid matrices with bounded rank perturbations," *Probab. Theory Rel. Fields*, vol. 155, nos. 1–2, pp. 231–263, 2013.
[26] F. Götze and A. Tikhomirov, "The rate of convergence for spectra of gue and lue matrix ensembles," *Open Math.*, vol. 3, no. 4, pp. 666–704, 2005.
[27] Q. Wu, D. Zhang, D. Liu, W. Liu, and C. Deng, "A method for power system steady stability situation assessment based on random matrix theory," in *Proc. CSEE*, 2016, pp. 5414–5420.
[28] W. Liu, D. Zhang, X. Wang, D. Liu, and Q. Wu, "Power system transient stability analysis based on random matrix theory," in *Proc. CSEE*, 2016, pp. 4854–4863.
[29] T. Hey, S. Tansley, and K. Tolle, *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Redmond, WA, USA: Microsoft Research, Oct. 2009. [Online]. Available: https://www.microsoft.com/en-us/research/publication/fourth-paradigm-data-intensive-scientific-discovery/
[30] J. Gray, "Jim Gray on eScience: A transformed scientific method," in *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Mountain View, CA, USA: Microsoft Research, 2009, pp. 17–31.
[31] R. Kitchin, "Big data and human geography opportunities, challenges and risks," *Dialogues Hum. Geogr.*, vol. 3, no. 3, pp. 262–267, 2013.
[32] J. Bun, J.-P. Bouchaud, and M. Potters. (Oct. 2016). "Cleaning large correlation matrices: Tools from random matrix theory." [Online]. Available: https://arxiv.org/abs/1610.08104
[33] X. He, R. Qiu, L. Chu, Q. Ai, Z. Ling, and J. Zhan. (Oct. 2017). "Detection and estimation of the invisible units using utility data based on random matrix theory." [Online]. Available: https://arxiv.org/abs/1710.10745

XING HE was born in Chuzhou, China, in 1986. He received the B.S. degree in electrical engineering from Southeast University, Nanjing, in 2008, and the M.S. and Ph.D. degrees in electrical engineering from Shanghai Jiaotong University, Shanghai, in 2012 and 2017, respectively.

Since 2017, he has been holding a post-doctoral position with the Research Center for Big Data and Artificial Intelligence Engineering and Technologies, Shanghai Jiaotong University. His research interests include situation awareness, smart grid, intelligent algorithms, and high-dimensional statistics.

Dr. He was a recipient of the National Scholarship for Doctoral Students in 2016.

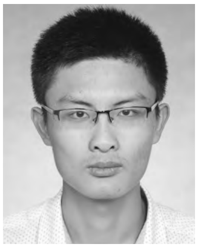
LEI CHU is currently pursuing the Ph.D. degree with Shanghai Jiaotong University. He is also a Research Assistant with the Research Center for Big Data and Artificial Intelligence Engineering and Technologies, Shanghai Jiaotong University, Shanghai. He has authored or co-authored one book chapter and several journal papers and holds patents. His research interests are in the theoretical and algorithmic studies in signal processing, multivariate statistics, statistical learning for high-dimensional data, random matrix theory, deep learning, and their applications in communications, localization, and smart grid.

ROBERT CAIMING QIU (S'93–M'96–SM'01–F'14) received the Ph.D. degree in electrical engineering from New York University.

He joined the Department of Electrical and Computer Engineering, Center for Manufacturing Research, Tennessee Technological University, Cookeville, Tennessee, as an Associate Professor in 2003, where he has been a Professor since 2008. He has also been with the Department of Electrical Engineering, Research Center for Big Data and Artificial Intelligence Engineering and Technologies, State Energy Smart Grid R&D Center, Shanghai Jiaotong University, since 2015. His current interest is in high-dimensional statistics, machine learning, wireless communication and networking, and smart grid technologies.

Dr. Qiu was a Founder CEO and a President of Wiscom Technologies, Inc., manufacturing and marketing wide band code division multiple access chipsets. Wiscom was sold to Intel in 2003. He was with Verizon, Waltham, MA, USA, and also with Bell Labs, Lucent, Whippany, NJ, USA. He focused on wireless communications and network, machine learning, smart grid, digital signal processing, electromagnetic scattering, composite absorbing materials, RF microelectronics, ultrawideband (UWB), underwater acoustics, and fiber optics. In 1998, he developed the first three courses on 3G for Bell Labs researchers. He was an Adjunct Professor with Polytechnic University, Brooklyn, NY, USA. He has authored over 100 journal papers/book chapters and 120 conference papers and holds over six patents. He has 15 contributions to 3GPP and the IEEE standards bodies. He has co-authored *Cognitive Radio Communication and Networking: Principles and Practice* (John Wiley, 2012) and *Cognitive Networked Sensing: A Big Data Way* (Springer, 2013), and authored *Big Data and Smart Grid* (John Wiley, 2015). He is a Guest Book Editor of *Ultra Wideband Wireless Communications* (New York: Wiley, 2005), and three special issues on UWB including the IEEE Journal on Selected Areas in Communications, the IEEE Transactions on Vehicular Technology, and the IEEE Transactions on Smart Grid. He serves as a TPC Member for GLOBECOM, ICC, WCNC, MILCOM, and ICUWB. In addition, he served on the Advisory Board of the New Jersey Center for Wireless Telecommunications. He is included in Marquis Who's Who in America. He serves as an Associate Editor for the IEEE Transactions on Vehicular Technology and other international journals.

QIAN AI (M'03–SM'16) received the bachelor's degree from Shanghai Jiaotong University, Shanghai, China, in 1991, the master's degree from Wuhan University in 1994, and the Ph.D. degree from Tsinghua University 1999, all in electrical engineering. After spending one year at Nanyang Technological University, Singapore, and two years at the University of Bath, U.K., he returned to Shanghai Jiaotong University. He is currently a Professor with the School of Electronic Information and Electrical Engineering, Shanghai Jiaotong University. His main interests include power quality, load modeling, smart grids, Microgrid and intelligent algorithms.



ZENAN LING received the bachelor's degree from the Department of Mathematics, Nanjing University, in 2015. He is currently pursuing the Ph.D. degree with Shanghai Jiaotong University. He is also a Research Assistant with the Research Center for Big Data and Artificial Intelligence Engineering and Technologies, Shanghai Jiaotong University, Shanghai. His research interests are in the machine learning, random matrix and free probability, and their applications in smart grid, stock market, and image processing.