Fast Event Identification through Subspace Characterization of PMU Data in Power Systems

Wenting Li, *Student Member, IEEE,* Meng Wang, *Member, IEEE*, Joe H. Chow, *Fellow, IEEE*
Department of Electrical, Computer, and Systems Engineering
Rensselaer Polytechnic Institute, Troy, NY
Email: {liw14,wangm7,chowj}@rpi.edu

***Abstract*—This paper proposes a data-driven fast event identification method based on the measurements of Phasor Measurement Units (PMUs). The central idea is to characterize an event by the low-dimensional subspace spanned by the dominant singular vectors of a data matrix that contains spatial-temporal blocks of PMU data. A dictionary of subspaces that correspond to different events are established off-line, and an event is identified online with the most similar subspace in the dictionary. Both theoretical and numerical analysis demonstrate that the subspace of PMU data is characteristic of system events and is robust to system initial conditions. This compact representation reduces the dictionary size and the computational time of the event identification method. The proposed method is evaluated numerically on simulated events in an IEEE 68-bus power systems.**

***Index Terms*—disturbance identification, synchrophasor measurements, low rank matrices.**

I. Introduction

Power system faults and events happen irregularly and potentially lead to wide-area oscillations and even cascading failures. The Northeast blackout of 2003 in the U.S. started with a local fault but escalated to a widespread blackout partially because the operator was not aware of the fault and took no action at the early stage. Therefore, fast event detection, identification, and location are of critical importance in the wide-area situational awareness of power systems.

Quite a few methods have been developed to locate events such as generator trips [1], line outages [2], [3], and oscillations [4]–[6]. Model-based identification methods (see e.g., [2], [3], [7]) usually require the modeling of the power system, and the identification performance depends critically on the accuracy of estimated parameters. Data-driven methods (e.g., [8], [9]) analyze the measurements directly without power system models. These methods extract features (including direct features like a frequency [10] or its derivative [11], as well as indirect features like wavelet coefficients [12]) from various datasets and classify those with similar features as cases resulting from the same event type.

Data-driven methods are receiving increasing attentions due to the recent wealth of data provided by modern devices such as Phasor Measurement Units [13], which measure voltage and current phasors directly at a rate of thirty samples per second or above. The computational complexity of a data-driven method generally increases significantly with the dataset size. That limits its practical applications in large-scale power systems with an increasing denser coverage of PMUs. Thus, it is important to develop a compact representation of data features to keep the computation tractable.

This paper proposes a novel data-driven method to identify and locate events without modeling the power system. The critical innovation is to characterize an event with a compact representation of system dynamics, computed from the spatial temporal blocks of measurements collected by multiple PMUs. Specifically, we demonstrate that the system dynamics can be characterized by a low-dimensional row subspace spanned by the dominant singular vectors of the PMU data matrix. This subspace characterization is robust to system initial conditions. Then event identification can be achieved by comparing obtained data with a pre-computed event dictionary with each dictionary atom corresponding to a row subspace of an event.

This proposed method has the following distinctive features: (1). The low-dimensional subspace can be easily computed through Singular Value Decomposition (SVD) or its faster variants like [14]. No modeling or parameter estimation is required. (2). The dictionary size is much smaller compared with a dictionary built directly from time series measurements, due to the dimensionality reduction in subspace characterization and the robustness of subspaces to system initial conditions. That in turn reduces the computational complexity of both the offline dictionary construction and the online event identification. (3). The method can identify events shortly after the event starts (e.g., within one second) and can be implemented in real-time, while existing methods are mostly designed for past event analysis.

The rest of the paper is organized as follows. The motivation of subspaces representation is described in Section II. The event identification method is introduced in Section III. Section IV records the numerical experiments in the IEEE 68-bus test system. Section V concludes the paper.

II. Technical Motivation

A. Low-rank Property of PMU Data Matrices

If PMU measurements at multiple buses across time are collected into an $m \times T$ matrix, where m is the number of measurements at one time instant, and T is the number of time steps, then the matrix can be approximated by a low-rank matrix with rank r much less than m and T. This approximate low-rank property of PMU data matrices have been observed

[15]–[17] and exploited to recover missing PMU data [16], detect system events [16], and identify cyber data attacks [18].

Let matrix M_r denote the rank-r approximation to the data matrix, and let $M_r = U_r \Sigma_r V_r^\dagger$ denote its r non-zero singular value decomposition, where $\Sigma_r \in \mathbb{C}^{r\times r}$ contains the singular values as its diagonal entries, and $U_r, V_r \in \mathbb{C}^{m\times r}$ contain r left and right singular vectors, respectively. $V_r^\dagger$ denotes the conjugate transpose of V_r. U_r's characterize the correlations of measurements in different PMUs. V_r's characterize the dominating dynamics in the power system during an event. Columns of V_r form a unitary basis of the row subspace of M_r. This paper proposes to represent an event by the row subspace of M_r, represented by V_r.

B. Event Characterization through Row Subspaces of Data Matrices

We first use a simplified model to motivate the low rankness of data matrices and the subspace characterization. If the power system is linearized around one equilibrium point, the resulting discrete-time model is

$$\mathbf{x}(\mathbf{t+1}) = \mathbf{A}\mathbf{x}(\mathbf{t}) + \mathbf{B}\mathbf{u}(\mathbf{t}), \tag{1}$$

$$\mathbf{y}(\mathbf{t}) = \mathbf{C}\mathbf{x}(\mathbf{t}) + \mathbf{D}\mathbf{u}(\mathbf{t}) + \boldsymbol{\epsilon}(\mathbf{t}), \tag{2}$$

where vectors $\mathbf{x}(\mathbf{t}) \in \mathbb{C}^{\mathbf{n}}$ and $\mathbf{y}(\mathbf{t}) \in \mathbb{C}^{\mathbf{m}}$ represent deviations at time t in state variables and measurements from the equilibrium point. $\mathbf{u}(\mathbf{t})$ is the control input, and $\boldsymbol{\epsilon}(t)$ is noise.

Let λ_k, $\mathbf{l}_k$, and $\mathbf{r}_k$ denote the kth largest eigenvalue of A and the corresponding left and right eigenvectors. Suppose the system matrix A is approximately low-rank, i.e., $|\lambda_{r+1}|$ is very close to zero for some r much less than n. Assume that the control input $\mathbf{u}(\mathbf{t})$ and measurement noise $\boldsymbol{\epsilon}(t)$ are zero. Let $\mathbf{x_0}$ denote the system state after an event starts. The measurements from time 1 to T are collected into matrix $M = [\mathbf{y}(\mathbf{1}), \mathbf{y}(\mathbf{2}), \cdots, \mathbf{y}(\mathbf{T})]$. One can verify that M can be approximated by a rank-r matrix, i.e.,

$$M \approx \sum_{k=1}^{r} \mathbf{l}_k^\dagger \mathbf{x_0} \cdot \mathbf{r_k} \cdot \boldsymbol{\beta}_\mathbf{k}^\dagger, \tag{3}$$

where

$$\boldsymbol{\beta}_k^\dagger = [1, \lambda_k, \lambda_k^2, \cdots, \lambda_k^T]. \tag{4}$$

Moreover, *the row subspace of M is spanned by $\boldsymbol{\beta}_k$'s, and independent of the system initial condition $\mathbf{x_0}$. The row space is shift-invariant when the measurement window is shifted*. The row subspace of $M_\kappa = [\mathbf{y}(\mathbf{1}+\kappa), \mathbf{y}(\mathbf{2}+\kappa), \cdots, \mathbf{y}(\mathbf{T}+\kappa)]$ is also spanned by $\boldsymbol{\beta}_k$' for any positive integer κ.

We emphasize that although the motivating analysis is based on a simplified model and does not consider noise, the low-rank property exists for complex power systems. For instance, even when A is not low-rank for a large-scale power system, its reduce model under one event may have a low-rank system matrix. Moreover, we observe numerically that the row subspace is indeed robust to system initial conditions and shift-invariant to a certain degree (see Section IV-A for details). Therefore, the row subspace of a data matrix is a compact and robust representation of system dynamics.

The proposed event identification method here is to build a dictionary of events off-line based on historical data and then identify an event online by comparing the obtained measurements with the dictionary. The key innovation here is to characterize an event by a low-dimensional row subspace of a spatial-temporal PMU data matrix. Compared with the dictionary composed of time series of PMU measurements (e.g., [9], [19]), using row subspace as dictionary atoms can significantly reduce the dictionary size without scarifying the identification accuracy. Moreover, the subspaces can be obtained easily through SVD or its fast variants like [14].

III. Online Event Identification

This paper focuses on subspace-based event identification, while the proposed method includes event detection, identification, and location for a practical implementation. The method is outlined in Fig. 1. A dictionary of events are constructed offline from sample PMU historical datasets during a window size T. Once an event is detected (through either the detection method here or any other event detector), all the PMU data in the next T time steps are collected to estimate the row subspace and column subspace. The event is identified through comparing the row subspace with the dictionary computed offline and located by comparing the significance of changes at different locations.

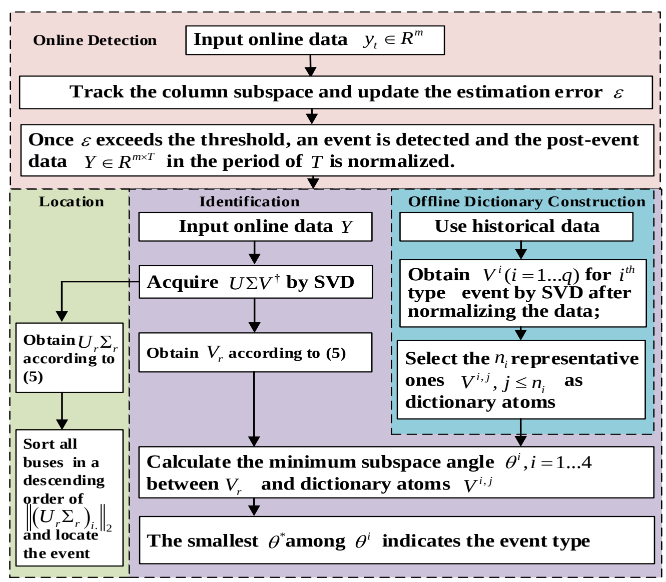

Fig. 1: The flow diagram of the online detection, identification and location method

The event identification requires subspace estimation and comparison, summarized as follows. Let σ_i denote its ith largest singular value of matrix $M \in \mathbb{C}^{m\times T}$. Given a predetermined $\tau \in (0,1)$, $\bar{r}$ is the smallest integer such that

$$(\Sigma_{i=1}^{\bar{r}}\sigma_i)/(\Sigma_{j=1}^{m}\sigma_j) \ge \tau \tag{5}$$

holds. The approximate rank r is defined as

$$r = \arg\max_{1\le i\le \bar{r}} \frac{\sigma_i}{\sigma_{i+1}}. \tag{6}$$

τ in (5) controls the approximation ratio, and only dominating features are maintained in the approximation through (6). The rank-r approximation is $M_r = U_r \Sigma_r V_r^\dagger$. Columns in V_r form a basis of the r-dimensional subspace in $\mathbb{C}^T$.

One can measure the affinity of two subspaces in $\mathbb{C}^T$ through the average subspace angle [20]. Specifically, the angle between subspaces spanned by unitary bases V_l and V_k respectively can be computed as

$$\theta(V_l, V_k) = \arccos(\sqrt{\|V_k^\dagger V_l\|_F^2 / \min\{k, l\}}). \quad (7)$$

θ equals $90°$ if two subspaces are orthogonal to other. θ equals $0°$ if two subspaces are the same ($l = k$) or one is embedded in the other ($l \neq k$). Thus, a smaller θ indicates higher affinity of two subspaces.

The proposed method includes the following components.

1. Offline construction of an event dictionary

The dictionary is constructed from historical event datasets, as shown in Fig. 2. Given each dataset, we compute a row subspace from PMU data within a window size T (e.g. thirty samples for one second of PMU data) and a given approximation ratio τ. These row subspaces are candidate dictionary atoms. Since a row subspace is robust to system initial conditions as described in Section II-B, row subspaces computed from events with the same type, at nearly location may have a small subspace angle. This property is exploited to reduce the dictionary size.

We refine the dictionary as follows. After selecting a subspace as an atom, all subspaces with angles less than a predefined threshold ϑ are removed from the dictionary. Let $\bar{V}^{i,j}$ denote the ith dictionary atom that corresponds to event type j. The total number of atoms in a dictionary is $\Sigma_{i=1}^q n_i$ for q types of events like line trip, load change, and three-phase short circuit events.

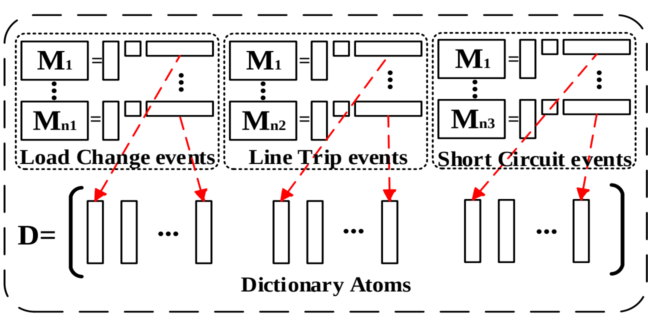

Fig. 2: Dictionary construction from historical datasets

2. Online event detection

We detect an event when the column subspace changes, similar to [16]. At time step t, the obtained measurement vector $\mathbf{y}(t)$ is projected onto the current estimate of the column subspace, represented by a basis $\hat{U}_r$. If the projection error, defined as $\varepsilon = \|\mathbf{y}_t - \hat{U}_r \hat{U}_r^\dagger \mathbf{y}(t)\|_2$, exceeds a predetermined threshold, then $\mathbf{y}(t)$ does not belong to the span of $\hat{U}_r$, and an event is declared to have happened. The column subspace can be estimated by SVD of past data in a fixed window or computed by subspace tracking methods like SPIRIT [21]. The event detection functionality is independent of others and can be achieved with other event detectors.

3. Online event identification based on the row subspace

Once an event is detected, all the PMU data in T time steps after an event happens are collected into a matrix. Given τ, we determine the approximate rank r and compute the row subspace spanned by V. Define the minimum subspace angle θ^* with the dictionary as

$$\theta^* = \min_{j=1}^{q} \min_{i=1}^{n_j} \theta(V, \bar{V}^{i,j}). \quad (8)$$

If the minimum is achieved with some dictionary atom $\bar{V}^{i^*,j^*}$, the event is identified to be type i^*.

4. Online event location

Here we use bus voltage magnitude measurements for event location. Given the rank r-approximation $M_r = U_r \Sigma_r V_r^\dagger$ of a data matrix, the ith row of $U_r \Sigma_r$, denoted by $(U_r \Sigma_r)_{i\cdot}$, contains coordinates of $(M_r)_{i\cdot}$, which contains voltage magnitudes of bus i across time, with respect to the basis V_r. Then $\|(U_r \Sigma_r)_{i\cdot}\|_2 = \|(M_r)_{i\cdot}\|_2$.

With the intuition that the bus closest to the origination of the event might have the largest value of $\|(M_r)_{i\cdot}\|_2$, we sort the buses in a descending order of $\|(U_r \Sigma_r)_{i\cdot}\|_2$ and locate the event by indicating that the event could happen at any of the top k buses with a predetermined integer k.

IV. Numerical Results

The method is tested on an IEEE 16-generator 68-bus test system as shown in Fig. 3. We generate synthetic PMU measurements by Power System Toolbox (PST) [22] for dictionary construction and method validation. Three types of events are considered: load change (the load changes from 0.5 p.u. to 1.5 p.u.), line trip and three phase short circuit (cleared after 0.2 second) events. The data rate is set to be 30 samples per second. Voltage magnitudes measurements are reported here, while we obtained similar results for phasor angle measurements. One second of data after an event starts are selected for analysis, i.e., $T = 30$.

Event datasets are generated with different pre-event system conditions. The difference between pre-event conditions are measured by the average relative difference of active power flow over all lines, denoted by $\bar{\eta}$. Specifically,

$$\bar{\eta}_a^b = (\Sigma_{j=1}^L |P_j^b - P_j^a| / |P_j^a|)/L, \quad (9)$$

where P_j^a and P_j^b are active power injections in the line j under condition a and b, respectively. L is the total number of lines.

A. Robustness of subspaces to pre-event systems conditions

We first verify that the row subspace is robust to pre-event system conditions. Fig. 4 shows the voltage magnitudes of four buses when the line connecting bus 1 and bus 2 is tripped under two different system conditions a and b. The

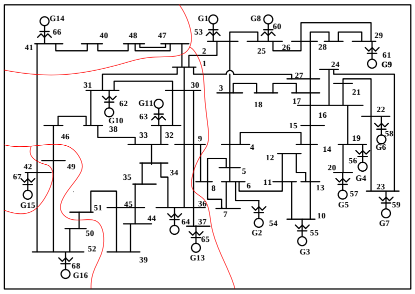

Fig. 3: Single line diagram of an IEEE 16-machine 68 bus power system. Five areas are decided based on the coherency of generators.

average relative difference of two conditions is $\bar{\eta}_a^b = 93.45\%$. Although these two datasets correspond to the same event at the same location, the differences in pre-event conditions lead to different time series measurements. Directly comparing the time series measurements may miss the similarities.

We pick one second of data in both datasets starting from $t_1 = 0.55$ second. With $\tau = 0.99$, the approximate rank defined in (5)-(6) is 6 for both data matrices. Fig. 5 shows the corresponding six principal right singular vectors, which form a basis of the row subspace. Note that the same subspace can be represented by infinitely possible basis selections. The angle of these two subspaces (defined in (7)) is $1.13°$, indicating that these two subspaces are very close to each other. Thus, the row subspace is indeed robust to pre-event conditions.

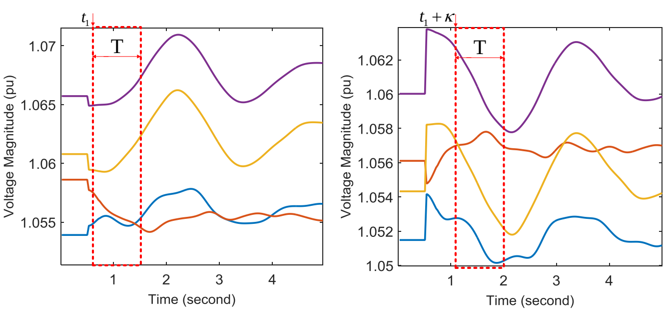

Fig. 4: The voltage magnitude of bus 1,25,40,48 when line 1-2 is tripped under pre-event condition a (left) ; under pre-event condition b with $\eta_a^b = 93.45\%$ (right)

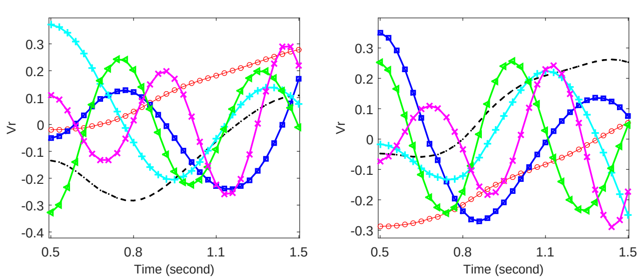

Fig. 5: The six principal singular vectors of V_r when line 1-2 was tripped under pre-event condition a (left); under pre-event condition b with $\eta_a^b = 93.45\%$ (right)

We then verify the shift-invariance property of the row subspace. We fix the one-second of data under condition a and its rank-6 approximation as before, while shifting the observation window in the other dataset to start from $t_1 + \kappa$. We set τ to be 0.95 when approximating the data matrix under condition b to maintain dominating features. As shown in Table I, even though the observation window in two datasets are not synchronized, the subspace angle is still relatively small. Therefore, minor differences in the observation windows do not affect the subspace comparison.

Table I: The subspace angles between two datasets in Fig. 4 when two observation windows are separated by κ

κ (second)	0.03	0.15	0.3	0.6	0.9	1.5
Subspace angle (°)	0.29	1.62	1.31	0.87	1.05	0.29

B. Dictionary construction

We simulate 190 cases, including 80 line trip events, 50 load change events, and 60 short circuit events at different locations with the same pre-event system condition a. A event dictionary is constructed following Step 1 in Section III. We set $\tau = 0.99$, $\vartheta = 5°$. After the refinement, the dictionary has 33 atoms, listed in Table II. Line trip 1-2 indicates tripping the line between bus 1 and 2. Load change (or short circuit) 32 means a load change (or short circuit) event at bus 32. Each atom corresponds to a row-subspace, represented by a $30 \times r$ matrix $\bar{V}$ (r is usually $3 \sim 6$ here).

Table II: The labels of different types of dictionary atoms

Dictionary Types	The Label of Dictionary Atoms
Line Trip	1-2, 4-14, 5-8, 12-11, 15-16, 30-2, 32-33, 33-34, 35-45, 36-37, 42-41, 44-45
Load Change	15, 16, 24, 32, 41, 42, 47, 48
Short Circuit	4, 12, 24, 29, 53, 59, 61, 63, 66, 28, 32, 50, 65

C. Event identification and location

Another 380 cases under two different pre-event conditions b and c are generated to test the identification performance. $\bar{\eta}_a^b = 93.4\%$, and $\bar{\eta}_a^c = 48.7\%$. Under each pre-event condition, 80 line trip events, 50 load change events (the load changes from 0.5 p.u. to 1.5 p.u.), and 60 three phase short circuits[1] events at different locations are generated. $\tau = 0.99$.

Table III: Minimum subspace angles between a test case with a pre-event condition c and dictionary atoms with the same event type with pre-event condition a. $\bar{\eta}_a^c = 48.7\%$.

DT / Events	Load Change	Line Trip	Short Circuit
Load Change 1	**1.03**	10.63	18.50
Load Change 7	**1.77**	10.94	20.55
Line Trip 1-30	11.80	**0.67**	14.32
Line Trip 16-19	10.21	**1.39**	13.24
Short Circuit 5	21.18	11.75	**1.97**
Short Circuit 15	22.80	18.32	**5.46**

Table III records the identification results of six sample cases under condition c. The subspace angle between each test case and each dictionary atom is computed. The minimum

[1]The one-second observation window includes the short circuit event before and after the clearance. Since the system topology changes after clearing the fault, the corresponding system model changes. Thus, the clearance time should be aligned when comparing two short circuit datasets.

subspace angle between a test case and dictionary atoms with the same event type are listed in Table III. As highlighted in bold, the minimum subspace angle is achieved when a dictionary atom has the same event type with the test case. Moreover, the subspace angle between a test case with dictionary atoms corresponding to different event types are generally much larger. Thus, the events can be correctly identified.

Table IV records the identification and location results under three criteria:

Identification Accuracy Rate (IAR): The ratio of the number of accurately identified events to the total number of events;

Approximate Location Accuracy Rate (ALAR): the ratio the number of events with actual locations[2] among the top k buses selected by Step 4 to the total number of events.

Exact Location Accuracy Rate (ELAR): a special case of ALAR when $k = 1$, i.e., the event location is exact.

Table IV: Statistical results of 380 *cases with pre-event conditions b and c with* $\bar{\eta}_a^b = 93.4\%$ *or* $\bar{\eta}_a^c = 48.7\%$ *respectively*

Type of event	IAR %	ELAR %	ALAR %
Line Trip	100	85	94(among 3 buses)
Short Circuit	100	77	90 (among top 3 buses)
Load Change	100	46	90 (among top 5 buses)

All the event types of 380 cases are correctly identified. The exact location can be identified with 85% and 77% accuracy for serious events like line trips and short circuit events. The exact location rate is smaller for minor events like load changes. The location rates all increase to 90% to 95% if we allow locating the events among 3-5 buses.

V. Conclusion and Future Work

This paper develops a data-driven event identification method by comparing the measurements with a dictionary of events constructed offline. It proposes to characterize system dynamics by the row subspace of a spatial-temporal PMU data matrix without estimating the power system model. This compact representation reduces the dictionary size while maintaining the identification accuracy. The method is verified on simulated datasets in an IEEE 68-bus test system. We are currently testing this method on historical disturbance data provided by ISO-NE. The results are not included here due to the page limit. The other ongoing work is to identify events outside a control region based on its local measurements.

Acknowledgement

We thank Frankie Qiang Zhang, Xiaochuan Luo, and Slava Maslennikov at ISO-NE for helpful discussions. This research is supported in part by the ERC Program of NSF and DoE under the supplement to NSF Award EEC-1041877 and the CURENT Industry Partnership Program, and in part by NSF Grant 1508875.

[2]The location of line trip events are considered as successful if one of the two related buses are correctly identified.

References

[1] J. N. Bank, R. M. Gardner, J. K. Wang, A. J. Arana, and Y. Liu, "Generator trip identification using wide-area measurements and historical data analysis," in *2006 IEEE PES Power Systems Conference and Exposition*, 2006, pp. 1677–1681.

[2] J. E. Tate and T. J. Overbye, "Line outage detection using phasor angle measurements," *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1644–1652, 2008.

[3] H. Zhu and G. B. Giannakis, "Sparse overcomplete representations for efficient identification of power line outages," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2215–2224, 2012.

[4] M. L. Crow and A. Singh, "The matrix pencil for power system modal extraction," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 501–502, 2005.

[5] G. Liu, J. Quintero, and V. Venkatasubramanian, "Oscillation monitoring system based on wide area synchrophasors in power systems," in *Bulk Power System Dynamics and Control-VII. Revitalizing Operational Reliability, 2007 iREP Symposium*, 2007, pp. 1–13.

[6] D. Trudnowski, J. Johnson, and J. Hauer, "Making prony analysis more accurate using multiple signals," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 226–231, 1999.

[7] M. J. Smith and K. Wedeward, "Event detection and location in electric power systems using constrained optimization," in *Proc. IEEE Power Energy Society General Meeting*, 2009, pp. 1–6.

[8] A. K. Ghosh and D. L. Lubkeman, "The classification of power system disturbance waveforms using a neural network approach," *IEEE Trans. Power Del.*, vol. 10, no. 1, pp. 109–115, 1995.

[9] Y. Song, W. Wang, Z. Zhang, H. Qi, and Y. Liu, "Multiple event analysis for large-scale power systems through cluster-based sparse coding," in *Proc. IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2015, pp. 301–306.

[10] O. P. Dahal, S. M. Brahma, and H. Cao, "Comprehensive clustering of disturbance events recorded by phasor measurement units," *IEEE Trans. Power Del.*, vol. 29, no. 3, pp. 1390–1397, 2014.

[11] A. Bykhovsky and J. H. Chow, "Power system disturbance identification from recorded dynamic data at the northfield substation," *International Journal of Electrical Power & Energy Systems*, vol. 25, no. 10, pp. 787 – 795, 2003.

[12] H. Jiang, J. J. Zhang, and D. W. Gao, "Fault localization in smart grid using wavelet analysis and unsupervised learning," in *Proc. Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, 2012, pp. 386–390.

[13] A. Phadke and J. Thorp, *Synchronized phasor measurements and their applications*. Springer, 2008.

[14] C. G. Baker, K. A. Gallivan, and P. Van Dooren, "Low-rank incremental methods for computing dominant singular subspaces," *Linear Algebra and its Applications*, vol. 436, no. 8, pp. 2866–2888, 2012.

[15] N. Dahal, R. L. King, and V. Madani, "Online dimension reduction of synchrophasor data," in *Proc. IEEE PES Transmission and Distribution Conference and Exposition (T&D)*, 2012, pp. 1–7.

[16] P. Gao, M. Wang, S. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stefopoulos, "Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements," *IEEE Trans. Power Systems*, vol. 31, no. 2, pp. 1006 – 1013, 2016.

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

[18] P. Gao, M. Wang, J. H. Chow, S. Ghiocel, B. Fardanesh, G. Stefopoulos, and M. P. Razanousky, "Identification of successive "unobservable" cyber data attacks in power systems." *IEEE Trans. Signal Process.*, vol. 64, no. 21, pp. 5557–5570, 2016.

[19] 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 Trans. Smart Grid*, vol. 5, no. 4, pp. 1654–1664, 2014.

[20] M. Soltanolkotabi, E. Elhamifar, E. J. Candes *et al.*, "Robust subspace clustering," *The Annals of Statistics*, vol. 42, no. 2, pp. 669–699, 2014.

[21] S. Papadimitriou, J. Sun, and C. Faloutsos, "Streaming pattern discovery in multiple time-series," in *Proceedings of the 31st international conference on Very large data bases*. VLDB Endowment, 2005, pp. 697–708.

[22] J. H. Chow and K. W. Cheung, "A toolbox for power system dynamics and control engineering education and research," *IEEE transactions on Power Systems*, vol. 7, no. 4, pp. 1559–1564, 1992.