Power Grid Faults Classification via Low-Rank Tensor Modeling

Matthew Repasky, Yao Xie
H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology
Atlanta, GA
mwrepasky@gatech.edu, yao.xie@isye.gatech.edu

Yichen Zhang
Department of Electrical Engineering, University of Texas at Arlington
Arlington, TX
yichen.zhang@uta.edu

Feng Qiu
Energy Systems and Infrastructure Analysis Division, Argonne National Laboratory
Lemont, IL
fqiu@anl.gov

2023 57th Asilomar Conference on Signals, Systems, and Computers | 979-8-3503-2574-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/IEEECONF59524.2023.10476907

***Abstract*—With the increased extent of synchrophasor-based wide area monitoring systems in modern large-scale power systems, more data-limited classification of fault events in smart grids becomes possible. Recent works have identified and exploited the low-rank property of synchrophasor data and power system dynamics for fault type classification. We develop a new methodology for modeling particular types of power grid faults from a tensor perspective, taking advantage of this well-known low-rank property, and classifying observed faults into categories by comparing to these low-rank representations. Using simulated data from the IEEE 39-bus system, we demonstrate that this approach is robust to limited temporal observation and sparse embedding of measurement units on the grid.**

***Index Terms*—Synchrophasor data, fault classification, low-rank modeling, tensor models.**

I. Introduction

To monitor and diagnose patterns observed in large-scale power systems, many smart grid technologies have been implemented in electric power systems, for example in the case of grids equipped with phasor measurement units (PMUs) [1], [2]. PMUs measure voltage, current, and frequency at embedded locations across a power grid to produce time-synchronized *synchrophasors*. This high-frequency but time-synchronized property of PMU data has become a critical component in the advent of data-driven techniques for smart grid diagnostics [3].

A key area of interest is the identification of faults in power systems, which may lead to cascading failures and blackouts without timely intervention [4], [5].However, the intervention time frames for such fault-induced electrical instabilities are usually in the order of milliseconds to tens of seconds [6]. Therefore, it is necessary for classification to require little data to make accurate characterizations. PMUs enable such classification, but the size and shape of PMU data requires particular consideration. It is important to carefully consider the three-way nature of PMU data when conducting analyses: the sensor in the grid, the sensor channel, and time.

Modern sensor data can often be naturally represented as multi-way arrays, or *tensors*, and decomposition is central to their analysis [7]. A common approach for matrices is the singular value decomposition (SVD), where two component matrices contain the singular vectors weighted by the singular values. SVD allows one to identify *rank*, or the number of modes required to capture (essentially) all information in the data. Many types of data are low rank [8]–[10]; see [11] for more applications. PMU data are low rank [12], and many methods use this to recover missing data [13]–[15] or for bad PMU data detection [16], [17].

The low-rank modeling and classification of fault types is discussed in Section II. This is followed by a series of experiments described in Section III, including a consideration of the time window used for fault type classification (Section III-A), an ablation on the individual PMU channels (Section III-B), and a study of the effect of a partial observation of the grid (Section III-C). We conclude with a discussion in Section IV.

A. Related Works

Typical classification approaches are often applied for fault identification [18], [19], including unsupervised cluster-based approaches [20]. Many approaches use explicit feature modeling, such as those extracted by intrinsic mode functions from signals in [21]. Others use low-rank models; for example, [22] apply principal component analysis (PCA) to PMU data, detecting changes in these components. PCA is also used by [23] for classification feature compression.

Training models requires a large amount of data covering different configurations and initial conditions [12]. Low-rank methods address this by suggesting that while a particular type of fault in the grid may yield different measurements at different times, the principal modes may share similar characteristics. For example, [24] estimate subspaces of PMU data for comparison to newly observed data. Others consider low-rank representation of multiple PMU channels jointly, such as [25], which are then used for feature construction in convolutional neural network classification. Our approach utilizes multiple channels of PMU data to construct low-rank representations of each fault type directly, comparing new observations to these models as in [24].

II. Methods

In our approach, the typical behavior of the power grid under a given fault type is modeled and new observations are classified according to the type of fault by comparing

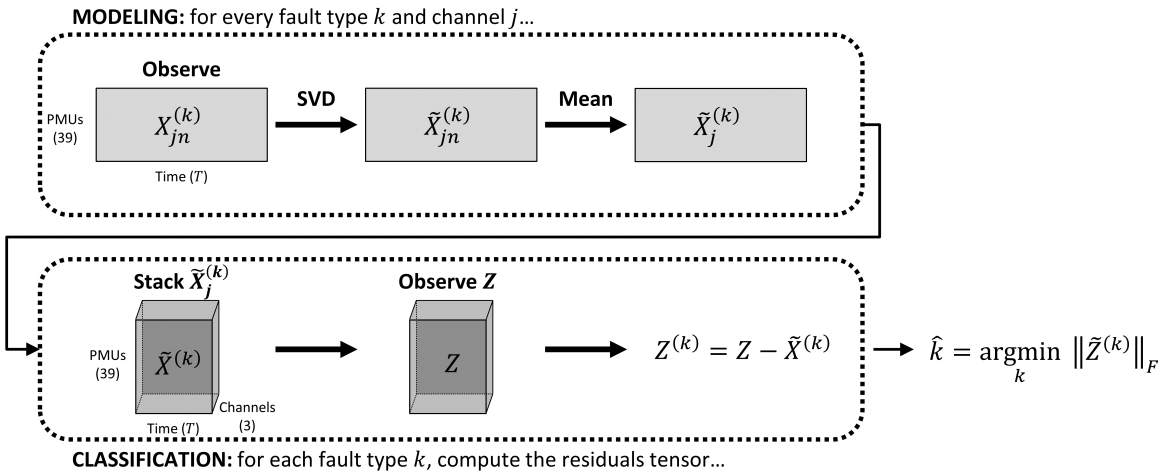

Fig. 1. Flowchart of the modeling and classification approach used for fault type classification. First, as outlined in the top panel, the readings from previously-observed faults are used to construct an approximation for standard power grid response to each type of fault. Then, as outlined in the bottom panel, when a new fault is observed, the approximation is used to classify the fault type which is most similar to the observed readings.

to these models. The problem is formalized in Section II-A. The truncated SVD for event type modeling is described in Section II-B, with the comparison of new observations outlined in Sections II-C and II-D. The outline of our approach is visualized as a flowchart in Figure 1.

A. Problem Formulation and Notation

In a simulated grid with 39 PMUs, each sensor reads voltage magnitude (VOLT), phase angle (ANGLE), and frequency (FREQ) with a period of 5 ms. Line trip, load increase, and temporary line-to-ground faults are simulated. We have a set of observations measuring 10 seconds of readings. In each observation, the fault takes effect at the 1 second mark. Define:

- $k \in \{1, 2, 3\}$: Fault type index
- $i \in \{1, \ldots, 39\}$: PMU sensor index
- $j \in \{1, 2, 3\}$: "Channel" index (indicator for VOLT, ANGLE, or FREQ)
- $n \in \{1, \ldots, N^{(k)}\}$: Power grid sample/observation index
- $t \in \{1, \ldots, T\}$: Time index

$N^{(k)}$ is the number of observations of fault type k. The data are tensors $X^{(k)}$ of shape $(N^{(k)} \times 39 \times 3 \times T)$. Each component is denoted $x_{ijnt}^{(k)}$. Under fault type k, for PMU i, channel j, and time t, the data can be constructed using an approximation:

$$x_{ijnt}^{(k)} = \mu_{ijt}^{(k)} + \epsilon_{ijnt}^{(k)}. \tag{1}$$

The *residuals* of observation $x_{ijnt}^{(k)}$ with respect to the reconstruction $\mu_{ijt}^{(k)}$ are denoted $\epsilon_{ijnt}^{(k)}$.

B. Low-Rank Representation

Matrix $X_{jn}^{(k)}$ of shape $(39 \times T)$ contains T readings of 39 PMUs at channel j and observation n. The SVD is:

$$X_{jn}^{(k)} = U_{jn}^{(k)} \Sigma_{jn}^{(k)} \left(V_{jn}^{(k)}\right)^{\mathrm{T}}, \tag{2}$$

where $\Sigma_{jn}^{(k)}$ is a diagonal matrix containing the (descending-order) singular values of $X_{jn}^{(k)}$, and the columns of $U_{jn}^{(k)}$ and $V_{jn}^{(k)}$ contain the singular vectors. A rank d representation of $X_{jn}^{(k)}$ is constructed by the first d principal components:

$$\tilde{X}_{jn}^{(k)} = U_{jn}^{(k)} \tilde{\Sigma}_{jn}^{(k)} \left(V_{jn}^{(k)}\right)^{\mathrm{T}}, \tag{3}$$

where $\tilde{\Sigma}_{jn}^{(k)}$ contains only the first d singular values. We take the component-wise mean of each $\tilde{X}_{jn}^{(k)}$ for all $N^{(k)}$ observations to obtain the low-rank, average representation $\tilde{X}_{j}^{(k)}$. Each row, $\hat{\mu}_{ij}^{(k)}$, is the approximation of channel j in PMU i during fault type k through time. There are 3 matrices for each fault type (one per channel). This is visualized in the top of Figure 1.

C. Residuals of the Approximation

Individual components of $X^{(k)}$ can be constructed:

$$x_{ijnt}^{(k)} = \hat{\mu}_{ijt}^{(k)} + \epsilon_{ijnt}^{(k)}, \tag{4}$$

with residual $\epsilon_{ijnt}^{(k)}$. Given $\hat{\mu}_{ij}^{(k)}$, residuals may be computed for every component:

$$\epsilon_{ijnt}^{(k)} = x_{ijnt}^{(k)} - \hat{\mu}_{ijt}^{(k)}. \tag{5}$$

Since each channel may have different scale, residuals may also be scaled differently. To address this, we define the sets

$$\mathcal{E}_{j'}^{(k')} = \left\{\epsilon_{ijnt}^{(k)} : j = j', k = k'\right\}, \tag{6}$$

containing all residuals for fault type k' and channel j'. Let $\bar{\mathcal{E}}_{j'}^{(k')}$ be the mean and $\sigma\left(\mathcal{E}_{j'}^{(k')}\right)$ be the standard deviation of these sets, respectively. Given new residual $z_j^{(k)}$ from channel j comparing to fault type k, define the standardized residual:

$$\tilde{z}_j^{(k)} = \frac{z_j^{(k)} - \bar{\mathcal{E}}_j^{(k)}}{\sigma\left(\mathcal{E}_j^{(k)}\right)}. \tag{7}$$

This ensures residuals are of comparable magnitude.

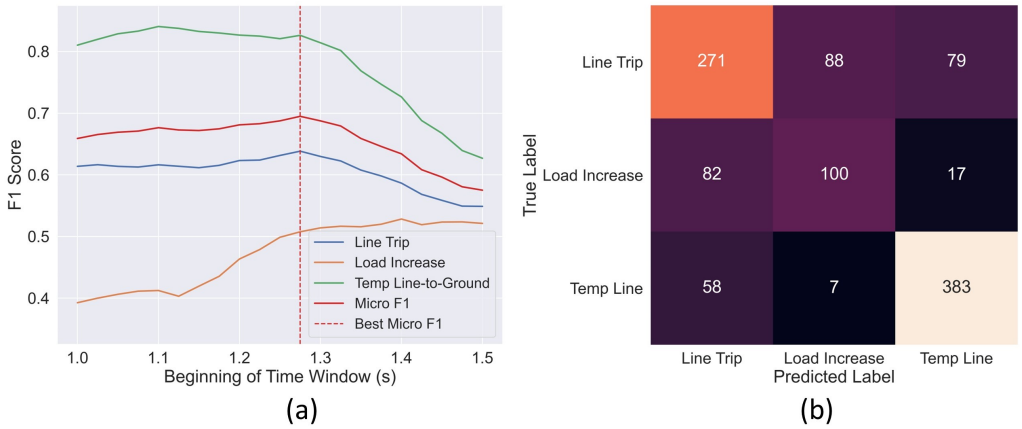

Fig. 2. In (a), the F_1 score of classification for each of the three classes of fault type is considered. The x-axis displays the beginning of the 0.25 second window used for classification, where the time is relative to the beginning of the simulation in which the fault is applied at the one second mark. The F_1 score is displayed on the y-axis. The red curve plots the micro-averaged F_1 score across all classes, which experiences its maximal value for the window beginning at the 1.275 second mark. In (b), the confusion matrix for the corresponding "best" micro-F_1 window is visualized.

TABLE I
PMU CHANNEL SUBSET CLASSIFICATION PERFORMANCE

	FULL	VOLT	ANGLE	FREQ	VOLT+ANGLE	ANGLE+FREQ	VOLT+FREQ
F_1	0.69	0.68	0.42	0.33	**0.72**	0.34	0.68
Line Trip Accuracy	0.62	0.52	0.98	0.54	0.77	0.62	0.53
Load Increase Accuracy	0.50	0.49	0.01	0.55	0.31	0.48	0.51
Temp Line	0.85	0.92	0.06	0.03	0.85	0.00	0.90

D. Fault Type Classification

Suppose a fault is detected and an observation tensor Z_τ of shape $(39 \times 3 \times \tau)$ contains τ time-steps of readings for the 39 PMUs and 3 channels per PMU. This is compared to each $\tilde{X}_j^{(k)}$ to classify Z_τ. First, define $\tilde{X}_\tau^{(k)}$ as the $(39 \times 3 \times \tau)$ tensor which "stacks" $\tilde{X}_j^{(k)}$ and truncates to the appropriate time window. Then,

$$Z_\tau^{(k)} = Z_\tau - \tilde{X}_\tau^{(k)}. \tag{8}$$

Each $Z_\tau^{(k)}$ contains the residuals (5), which can be standardized by (7) to produce $\tilde{Z}_\tau^{(k)}$. The classification of the type of fault can be determined according to which $\tilde{Z}_\tau^{(k)}$ has smallest squared Frobenius norm. Effectively, the classification selects the fault type whose approximation according to Section II-B most closely matches the observed fault. The observation of a new fault and subsequent classification are visualized in the bottom half of the flowchart in Figure 1.

III. EXPERIMENTS

Consider labeled simulated data containing 2,194 line trip faults, 998 load increase faults, and 2,240 temporary line-to-ground faults in a 39-bus power grid. Faults occur at one second followed by nine more seconds of measurement (period of 5 milliseconds). 80% of the data is used to develop $\tilde{X}_j^{(k)}$.

The remaining 20% of the data are used for classification. We demonstrate the importance of the time window considered when conducting classification in Section III-A. Then, the effect of each channel is examined in an ablation study (Section III-B). Finally, we study the effect on classification of partial observations of the power grid in Section III-C.

A. Windowed Classification

To evaluate the performance of classification through time, we consider a moving window classification approach. At any point in time, we consider a window of 0.25 seconds of readings (in the case of our data, this is equivalent to $\tau = 50$ time steps). Beginning with the point in time at which the fault takes place (the one second mark in our simulated data), we consider 21 such windows, where the start time is shifted by 0.025 seconds each time. That is, we consider windows from 1.000-1.250 seconds, 1.025-1.275 seconds, ..., 1.500-1.750 seconds. For each of these time windows, we perform classification according to Section II-D on the classification split of the data.

The F_1 scores for each class corresponding to window choice are visualized in Figure 2(a). The line trip score increases for earlier windows while the score of the load increase fault increases more drastically. The score of temporary line-to-ground faults plateaus at early times followed by a sharp decline. The micro-averaged F_1 gradually increases through time, peaking at the 1.275-1.525 second window. The confusion matrix for this window is visualized in Figure 2(b), indicating that performance for line trip and temporary line-to-ground faults is better than that of load increase faults.

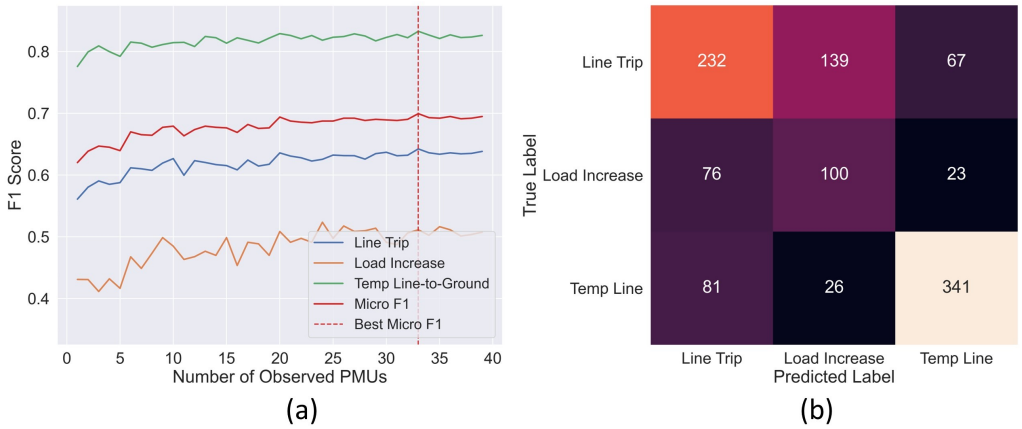

Fig. 3. In (a), the F_1 score of classification for each of the three classes of fault type is considered. The x-axis displays the number of PMUs (n_{PMU}) used for classification, and the observation window is 1.275-1.525 seconds measured from the beginning of the simulation (recall, the fault occurs at one second). The F_1 score is again plotted on the vertical axis. The red curve plots the micro-averaged F_1 score across all classes. In (b), the confusion matrix for the observation of a single PMU is visualized.

B. Channel Ablation

Next, we conduct an ablation study to determine the effect of each PMU channel (voltage, phase angle, and frequency) on the classification performance. To do so, rather than using the entire residuals tensor $Z_\tau^{(k)}$ in the classification, (channel dimension) sub-matrices are used. That is, we perform classification using only voltage, phase angle, or frequency, or a combination of two channels. Using the result from Section III-A, this section uses the time window from 1.275-1.525 seconds. These $\tau = 50$ time readings are used to perform the classification, which is repeated for all possible combinations of the three channels.

The performance in the ablation study is displayed in Table I. Voltage alone has the highest F_1 score for single-channel classification, followed by angle and frequency. Considering pairs of channels, voltage combined with angle achieves the highest score followed closely by combining voltage with frequency. Considering angle and frequency together achieves very low score. These results indicate that voltage is necessary for high performance. The VOLT + ANGLE classification exhibits the best F_1, but compared to Figure 2(b) using all channels, we observe worse performance in classifying load increase type faults.

C. Partial Observation

In many cases, it is unrealistic to assume that all buses in the power grid will have an associated PMU. To study the effectiveness of our approach in these more realistic cases, we aim to determine how the performance of classification is affected by only considering a subset of the PMU readings. In this case, we consider sub-matrices corresponding to subsets of the entire set of 39 PMUs in the power grid. In each trial, we fix the number of PMUs to be used, denoted n_{PMU}. That is, n_{PMU} sub-matrices from tensor $Z_\tau^{(k)}$ are considered to conduct the classification. We choose a random subset n_{PMU} number of PMUs to consider for each observation. We again use a window from 1.275-1.525 seconds.

Figure 3(a) shows the F_1 score for each fault type over a range of n_{PMU}. Temporary line-to-ground faults achieve the best performance and load increase faults achieve the worst. The micro-averaged F_1 score increases only slightly with n_{PMU}, from just over 0.6 when a single PMU is observed to just under 0.7 when more than 30 are observed. The confusion matrix for the observation of a single PMU is displayed in Figure 3(b). Line trip and temporary line-to-ground faults experience a slight decrease in performance compared to Figure 2(b), but the overall performance is only slightly diminished.

IV. Conclusion

We have demonstrated the utility of low-rank tensor-based modeling of faults. Our approach is interpretable and efficient, constructing simple models and using difference-based classification. We have demonstrated that, given a detected fault in simulated data, our approach can effectively classify faults according to the three types of observed disturbances. Furthermore, an ablation study indicates that using a subset of the PMU channels can yield similar performance. Finally, we have found that this approach works well given only an observation of a subset of the grid.

Next steps can combine our approach with rapid detection, potentially providing a method for both quick detection and classification of fault type. This combination of temporal and categorical information could lead to effective and informed response to failures in the grid.

Acknowledgments

The work is partially supported by the U.S. Department of Energy Office of Electricity – Advanced Grid Modeling Program. Repasky and Xie are partially supported by an NSF CAREER CCF-1650913, NSF DMS-2134037, CMMI-2015787, CMMI-2112533, DMS-1938106, and DMS-1830210.

REFERENCES

[1] D. Novosel, K. Vu, V. Centeno, S. Skok, and M. Begovic, "Benefits of synchronized-measurement technology for power-grid applications," in *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*. IEEE, 2007, pp. 118–118.

[2] "2020 smart grid system report," Tech. Rep., 2022.

[3] E. Elbouchikhi, M. F. Zia, M. Benbouzid, and S. El Hani, "Overview of signal processing and machine learning for smart grid condition monitoring," *Electronics*, vol. 10, no. 21, p. 2725, 2021.

[4] A. Muir and J. Lopatto, "Final report on the august 14, 2003 blackout in the united states and canada: causes and recommendations," 2004.

[5] G. A. Maas, M. Bial, and J. Fijalkowski, "Final report-system disturbance on 4 november 2006," *Union for the Coordination of Transmission of Electricity in Europe, Tech. Rep*, p. 1, 2007.

[6] P. Henneaux, P.-E. Labeau, and J.-C. Maun, "A level-1 probabilistic risk assessment to blackout hazard in transmission power systems," *Reliability Engineering & System Safety*, vol. 102, pp. 41–52, 2012.

[7] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. Phan, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE signal processing magazine*, vol. 32, no. 2, pp. 145–163, 2015.

[8] A. Ahmed and J. Romberg, "Compressive multiplexing of correlated signals," *IEEE Transactions on Information Theory*, vol. 61, no. 1, pp. 479–498, 2014.

[9] A. Singer, "A remark on global positioning from local distances," *Proceedings of the National Academy of Sciences*, vol. 105, no. 28, pp. 9507–9511, 2008.

[10] A. Ramlatchan, M. Yang, Q. Liu, M. Li, J. Wang, and Y. Li, "A survey of matrix completion methods for recommendation systems," *Big Data Mining and Analytics*, vol. 1, no. 4, pp. 308–323, 2018.

[11] M. A. Davenport and J. Romberg, "An overview of low-rank matrix recovery from incomplete observations," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 4, pp. 608–622, 2016.

[12] M. Wang, J. H. Chow, D. Osipov, S. Konstantinopoulos, S. Zhang, E. Farantatos, and M. Patel, "Review of low-rank data-driven methods applied to synchrophasor measurement," *IEEE Open Access Journal of Power and Energy*, vol. 8, pp. 532–542, 2021.

[13] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stefopoulos, "Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1006–1013, 2015.

[14] D. Osipov and J. H. Chow, "Pmu missing data recovery using tensor decomposition," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4554–4563, 2020.

[15] S. Konstantinopoulos, G. M. De Mijolla, J. H. Chow, H. Lev-Ari, and M. Wang, "Synchrophasor missing data recovery via data-driven filtering," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4321–4330, 2020.

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

[17] S. Zhang and M. Wang, "Correction of corrupted columns through fast robust hankel matrix completion," *IEEE Transactions on Signal Processing*, vol. 67, no. 10, pp. 2580–2594, 2019.

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

[19] P. G. Axelberg, I. Y.-H. Gu, and M. H. Bollen, "Support vector machine for classification of voltage disturbances," *IEEE Transactions on power delivery*, vol. 22, no. 3, pp. 1297–1303, 2007.

[20] Y. Song, W. Wang, Z. Zhang, H. Qi, and Y. Liu, "Multiple event detection and recognition for large-scale power systems through cluster-based sparse coding," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4199–4210, 2017.

[21] M. Mishra and P. K. Rout, "Detection and classification of micro-grid faults based on hht and machine learning techniques," *IET Generation, Transmission & Distribution*, vol. 12, no. 2, pp. 388–397, 2018.

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

[23] Y. Wang, M. Liu, Z. Bao, and S. Zhang, "Stacked sparse autoencoder with pca and svm for data-based line trip fault diagnosis in power systems," *Neural computing and applications*, vol. 31, no. 10, pp. 6719–6731, 2019.

[24] W. Li, M. Wang, and J. H. Chow, "Fast event identification through subspace characterization of pmu data in power systems," in *2017 IEEE Power & Energy Society General Meeting*. IEEE, 2017, pp. 1–5.

[25] W. Li and M. Wang, "Identifying overlapping successive events using a shallow convolutional neural network," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4762–4772, 2019.