

Review of Low-Rank Data-Driven Methods Applied to Synchronphasor Measurement

Meng Wang, *Member, IEEE*, Joe H. Chow, *Life Fellow, IEEE*, Denis Osipov, *Member, IEEE*,
Stavros Konstantinopoulos, *Student Member, IEEE*, Shuai Zhang, *Student Member, IEEE*,
Evangelos Farantatos, *Senior Member, IEEE*, Mahendra Patel, *Fellow, IEEE*

There is a growing acceptance of using synchronphasor data collected over large power systems in control centers to enhance the reliability of power system operations. The spatial and temporal nature of power system ambient and disturbance response allows the analysis of large amount of synchronphasor data by low-rank methods. This paper provides an overview of several applications of synchronphasor data utilizing the low-rank property. The tools to capitalize on the low-rank property include matrix completion methods, tensor analysis, adaptive filtering, and machine learning. The applications include missing data recovery, bad data correction, and disturbance recognition.

Index Terms—Synchronphasor data, low rankness, matrix completion, tensor analysis, adaptive filtering.

I. INTRODUCTION

WIDE-AREA monitoring systems (WAMS) built on synchronphasor networks consisting of phasor measurement units (PMUs) interconnected through communication networks have seen a growing acceptance by power system control centers to enhance the visibility of events and dynamics propagation throughout the system [1]. At its core, PMU data in a WAMS show the dependence of system dynamics on the underlying power network and its operating condition. As such, synchronphasor networks offer opportunities to develop data-driven methods to improve the reliability of power system operation, such as detection of malfunctioning equipment [2] and wide-area damping control [3]. The purpose of this paper is to address the PMU data quality based on its low-rank property and applications that are based on this specified property.

The low-rank property of data from a synchronphasor network reflects the spatial and temporal variations of voltage and current phasors measured at nearby high-voltage substations. Absent of disturbances, the data show the slowly varying adjustments of generator setpoints to meet the daily load cycles. The almost simultaneity of the daily load cycles further preserves the low rankness of PMU data over longer time spans.

Data of many networks, whether man-made or based on natural phenomena, will exhibit low rankness at some level. For example, forecasts of temperature or precipitation for points within a few tens of kilometers of each other tend to be highly correlated [4]. The Netflix user preference system developed many years ago can recommend movie choices to

users with very high accuracy [5]. The main idea of that algorithm is to cluster users into affinity groups and movies into genres, resulting in a mostly low-rank linkage of user groups and movie genres. As a result, many good numerical algorithms have been developed to capitalize on data with low rank such as [6], [7], [8], [9], [10]. Such development has benefited many fields, including PMU data analysis. However, it should be pointed out that these algorithms have been developed for sparse data structure, whereas PMU analysis requires algorithms dealing with non-sparse data structure.

This paper provides a review of low-rank methods for enhancing the quality of PMU data as well as other related applications. Although it will focus on the work by the authors, other relevant research results will also be included to serve as a literature survey. One of the aims of this paper is to illustrate the richness of the low-rank property for research and development.

The remainder of this paper is organized as follows. Section II provides an example of disturbance event PMU data to illustrate low rankness. Section III presents three methods to reliably recover missing PMU using matrix, tensor, and adaptive filtering techniques. Section IV applies the low-rank property to detect bad data. Section V extends the low-rank approach treating the PMU data as responses from a nonlinear system. Section VI describes the use of PMU data to classify disturbances.

II. A LOW-RANK PMU DATA EXAMPLE

As PMU data are time-synchronized measurements at many different locations spread over a large expanse of a power system, they capture the variations of voltages and currents in ambient and disturbance conditions. From a data analytics point of view, it is useful to organize the data into matrices with time on one axis and measured data values on the other axis. The measured data is a result of the power network reacting to the generations and load changes during normal operating conditions and control systems reacting to large and

The work described in this paper was supported in part by the Engineering Research Center Program of the NSF and the DOE under the supplement to NSF Award Number EEC1041877 and the CURENT Industry Partnership Program, NSF award #1932196, and Electric Power Research Institute.

M. Wang, J. H. Chow, D. Osipov, S. Konstantinopoulos, and S. Zhang are with the Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, NY 12180, USA (e-mail: {wangm7, chowj, osipod, konst, zhangs21}@rpi.edu).

E. Farantatos and M. Patel are with the Grid Operations and Planning Department, Electric Power Research Institute (EPRI), Knoxville, TN, USA (e-mail: {efarantatos, mpatel}@epri.com).

small disturbances. As the current injection phasors $\tilde{I}_{inj}$ and voltage phasors $\tilde{V}$ in a power network are related by

$$Y\tilde{V} = \tilde{I}_{inj}, \quad (1)$$

where $\tilde{V} \in \mathbb{C}^N$ and $\tilde{I}_{inj} \in \mathbb{C}^N$ are in the vector form, N is the number of buses in the network, and $Y \in \mathbb{C}^{N \times N}$ is the matrix of network admittances [11]. In the formulation (1), it is expected that strongly connected nearby buses would exhibit similar voltage responses to sufficiently far away disturbances.

One of the earliest observations of the low-rank matrix of PMU data was made in [12], which used principal component analysis to reduce the dimensionality of the PMU data. In this section, this low-rank property is illustrated by PMU data measured during a disturbance event in central New York. Fig. 1 shows the locations of six PMUs in central New York. These multi-channel PMUs provide 11 bus voltage and 26 line current positive-sequence phasors. The voltage magnitude response to a disturbance is shown in Fig. 2, with the onset of the disturbance at about 2.3 s.

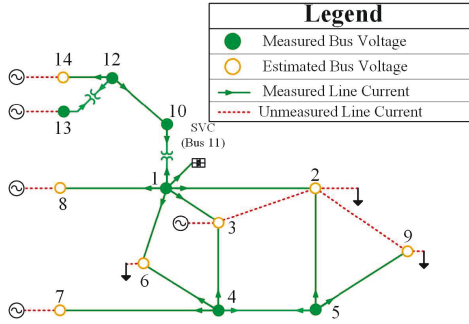


Fig. 1: Six PMUs in Central New York Power System [13].

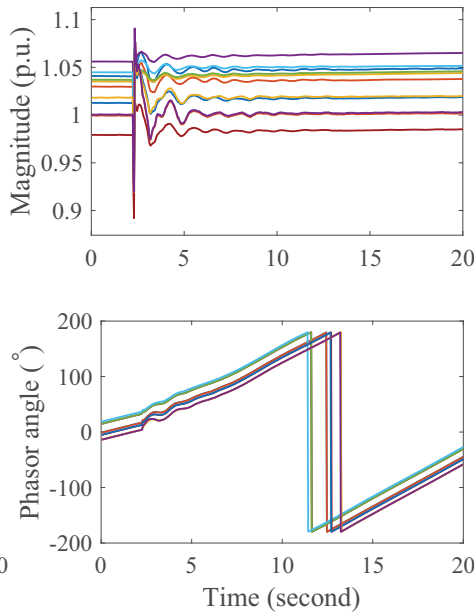


Fig. 2: Voltage Phasor Response to a Disturbance [14].

Assembling all the measured phasor data into a single matrix over a 20-second data window with 30 points per

second, a complex matrix X^* of dimension 37 by 600 can be obtained. Here a single phasor quantity is split into two channels, one for its magnitude and one for its phase. A plot of the singular values of the matrix is shown in Fig. 3. The ten largest singular values of X^* are 894.59, 36.83, 20.72, 8.34, 3.08, 2.48, 1.97, 1.35, 0.59 and 0.25. The largest singular value is due to the nonzero steady-state values of the measurements, and as such, it is much larger than the other singular values. Thus, the data matrix X^* can be approximated by the eight largest singular values and their left and right singular vectors with a small error.

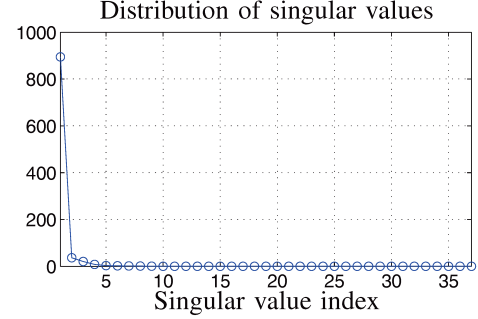


Fig. 3: Singular Values of the PMU Data Matrix from a Disturbance Event [13].

The low-rank condition of the PMU data matrix is maintained if additional rows are time shifted and included in X^* , forming a Hankel matrix. Given an $m \times n$ matrix X with the j th column denoted by $\mathbf{x}_j$, for any integer $\kappa > 1$, a Hankel matrix $\mathcal{H}_\kappa(X)$ is defined as

$$\mathcal{H}_\kappa(X) = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_3 & \cdots & \mathbf{x}_{n+1-\kappa} \\ \mathbf{x}_2 & \mathbf{x}_3 & \cdots & \mathbf{x}_{n+2-\kappa} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_\kappa & \mathbf{x}_{\kappa+1} & \cdots & \mathbf{x}_n \end{bmatrix}. \quad (2)$$

Fig. 4 shows the relative approximation errors of rank- r matrix to the $\mathcal{H}_\kappa(X^*)$ with different r and κ . All the matrices $\mathcal{H}_\kappa(X^*)$ can be approximated by a rank-2 or rank-3 matrix with a negligible error.

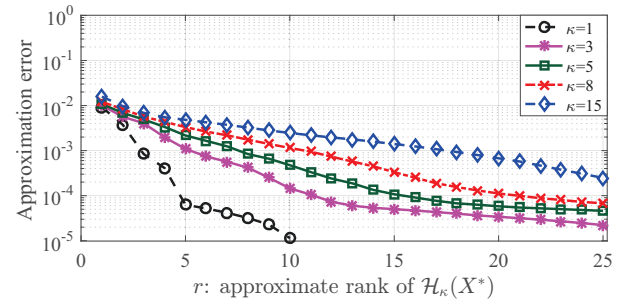


Fig. 4: The Approximation Errors of Low-rank Matrices to Hankel Matrices from X^* [14].

As shown in [15], the power system can be approximated by a linear dynamical system under minor disturbances. In theory, the total number of modes n depends on the order of the dynamical system. In practice, a mode might be highly

damped, or not excited by a disturbance input, or not directly measured in the measurements. Thus, if only r ($r \ll n$) dominant modes are present in the measurements, the Hankel matrix of the measurements is approximately rank r . When the system is under a severe event such that the linear approximation does not hold, one can apply the Koopman theory [16], [17] to lift the nonlinear dynamical system to a possibly infinite-dimensional linear system. Then if the number of excited modes in this lifted system is smaller than the ambient dimension, the low-rank Hankel property still holds [18].

III. MISSING DATA RECOVERY

Synchrophasor data collected by system operators often suffer from quality issues such as data losses, bad data, and possibly cyber data attacks due to communication delays, device malfunction, and attacks from malicious intruders. To ensure reliable outcomes from PMU-data-based applications, we first need to address these data issues before sending them to any application. In this section, we will discuss various methods to recover missing data by exploiting the low-rank property, while the discussions on correcting bad data and cyber data attacks are deferred to Section IV.

A. Matrix Method

Let the matrix $X^* \in \mathbb{C}^{m \times n}$ contain synchrophasor data from m PMU channels in n time instants. We consider the general setup that X^* can contain complex values to characterize phasors, and all the results apply to real-valued matrices as well.

Because X^* is low-rank, as illustrated in Section II, the missing data recovery problem can be formulated as a low-rank matrix completion problem [19], [13], which has been extensively studied in the past twelve year or so due to its wide applications [9]. It has been established theoretically that if an $m \times n$ matrix is rank r with r much smaller than m and n , then as long as the number of observed entries of X^* is in the order of $rn \log^2 n$ (assuming $m \leq n$), all the remaining missing entries can be corrected recovered. Thus, directly applying existing low-rank matrix completion methods such as Singular Value Thresholding (SVT) [7] and Fast Iterative Hard Thresholding (FIHT) [20] on PMU data can provide accurate estimates of the missing points.

Despite the encouraging results [19], [13], a direct application of low-rank methods has two limitations. First, the above theoretical guarantee is based on the assumption that the locations of the missing points are randomly distributed in the matrix, while PMU data losses are often consecutive in time or simultaneous across channels, resulting from communication congestion. The theoretical guarantee of recovering missing points at a non-uniform location is established in [13], but the required number of observations is much larger than $rn \log^2 n$, leaving room for future theoretical improvement.

The other limitation is that the low-rank property does not characterize the temporal correlations in PMU data fully. To see this, consider arbitrarily swapping some columns in X^* , then the resulting matrix X' still has the same rank, but each row of X' does not correspond to a legitimate time series. One

way to characterize the correlations is to exploit the low-rank property of the Hankel matrix of the PMU data.

Let Ω denote the set of indices of the entries that are observed. Then, X_{ij}^* is observed if (ij) is in Ω , and X_{ij}^* is missing otherwise. From the property of the low-rank Hankel matrix, the data recovery problem is formulated as

$$\begin{aligned} \min_{X \in \mathbb{C}^{m \times n}} \quad & \sum_{(ij) \in \Omega} (X_{ij}^* - X_{ij})^2 \\ \text{s.t.} \quad & \text{rank}(\mathcal{H}_\kappa(X)) \leq r, \end{aligned} \quad (3)$$

in [15]. The solution to (3), denoted by $\hat{X}$, is used to estimate the missing points in X^* . The optimization (3) is nonconvex due to the rank constraint.

An Accelerated Multi-channel Fast Iterative Hard Thresholding (AM-FIHT) algorithm is developed in [15], [21] to solve (3). At initialization, AM-FIHT sets all the missing entries to zero and finds the best rank- r approximation of the resulting $\kappa m \times n$ Hankel matrix based on the observations. Then in each iteration, it first updates through accelerated gradient descent, then projects to a rank- r $\kappa m \times n$ matrix, and converts back to an $m \times n$ matrix.

If the locations of the missing points are randomly distributed, AM-FIHT is guaranteed to recover the missing entries accurately, provided that the number of observations is in the order of $r^2 \log^2 n$ [15]. This bound is significantly smaller than $rn \log^2 n$ by the conventional low-rank matrix completion methods, demonstrating the effectiveness of exploiting the dynamics in the time series using the low-rank Hankel property. Moreover, AM-FIHT has a low computational complexity. It converges linearly, meaning the number of iterations needed to return a solution within error ϵ is only in the order of $\log(1/\epsilon)$, and the per-iteration complexity is at most in the order of $r^2 mn + rmn \log n + r^3$. The result can also be generalized to the case that the obtained measurements contain noise [15].

The recovery performance of AM-FIHT is compared with Singular Value Thresholding (SVT) [7] and Fast Iterative Hard Thresholding (FIHT) [20] on the dataset visualized in Fig. 2. SVT solves the convex relaxation of the low-rank matrix completion problem. We apply it to both the data matrix directly and the resulting Hankel matrix. FIHT is a non-convex algorithm that recovers the missing data of single-channel Hankel matrices, and it is implemented to recover data in each PMU channel separately. The x -axis is the data loss percentage. The y -axis is the relative recovery error, defined as $\|\hat{X}_{\Omega^c} - X_{\Omega^c}^*\|_F / \|X_{\Omega^c}^*\|_F$, where $\hat{X}_{\Omega^c}$ and $X_{\Omega^c}^*$ denote the recovered and ground-truth data of the missing points, respectively. In Case 1, measurements are lost randomly and independently across channels and time. In Case 2, measurements in all PMU channels are lost simultaneously at some randomly selected time instants. In Case 3, all measurements at some fixed channels are lost simultaneously and consecutively for some time. Directly applying SVT on the data matrix fails to recover fully lost columns in Case 2. AM-FIHT achieves the best performance among all the methods in all three cases.

B. Tensor Method

A phasor data matrix $X^* \in \mathbb{C}^{m \times n}$ can be represented as a tensor $\mathcal{X}^* \in \mathbb{R}^{m \times n \times 2}$ by separating the phasor magnitudes

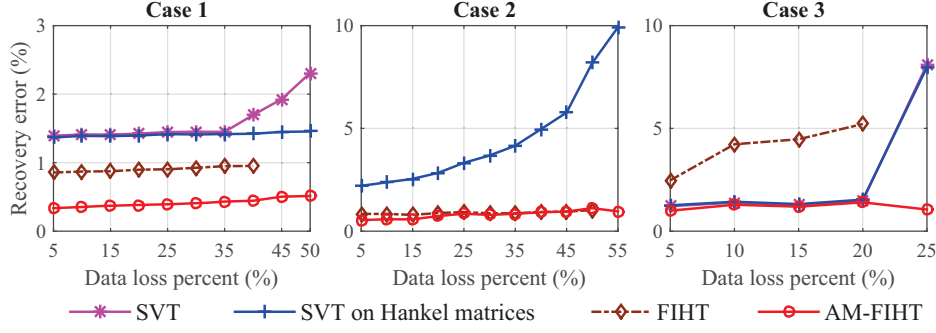


Fig. 5: Relative Recovery Error of SVT, FIHT and AM-FIHT on the Disturbance Data (visualized in Fig. 2) [21].

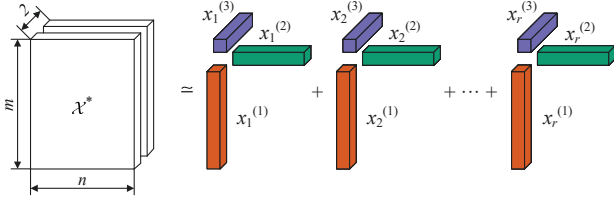


Fig. 6: Polyadic Decomposition of Tensor $\mathcal{X}^*$.

and angles into two matrices (slices of the tensor). A tensor is a multidimensional array which can be decomposed into a set of low-dimension structures using a tensor decomposition method [22]. For example, in the polyadic tensor decomposition (PD), the tensor is approximated by the sum of a finite number of rank-one tensors:

$$\mathcal{X}^* \approx \sum_{i=1}^r x_i^{(1)} \circ x_i^{(2)} \circ x_i^{(3)} \quad (4)$$

where r is the rank of the decomposed tensor, $x_i^{(k)}$ is a factor vector corresponding to i -th rank and k -th dimension, and $\circ$ denotes the outer product. Expression (4) is illustrated in Fig. 6. The factor vectors in (4), which have the same dimension, can be grouped into a factor matrix:

$$\mathbf{X}^{(k)} = [x_1^{(k)}, x_2^{(k)}, \dots, x_r^{(k)}] \quad (5)$$

The set of all factor matrices represents the decomposed tensor: $\{\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \mathbf{X}^{(3)}\}$

The missing entries in $\mathcal{X}^*$ can be recovered by solving the optimization problem:

$$\min_{\mathbf{X}^{(k)}} \frac{1}{2} \|\mathcal{W} \odot (\mathcal{X}^* - \llbracket \mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \mathbf{X}^{(3)} \rrbracket)\|_F^2 \quad (6)$$

where $\mathcal{W} \in \{0, 1\}^{m \times n \times 2}$ is the observation tensor whose entries are equal to 0 for the corresponding missing entries of tensor $\mathcal{X}^*$ and 1 otherwise, $\odot$ denotes the Hadamard product, and $\|\cdot\|_F$ is the Frobenius norm. The problem (6) is solved using a nonlinear least-squares formulation and the Gauss-Newton method [23], [24]. Reconstructed from factor matrices $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \mathbf{X}^{(3)}$, the tensor $\mathcal{X}$ will not have missing entries. Detailed description of the PD-based method for PMU missing data recovery can be found in [25].

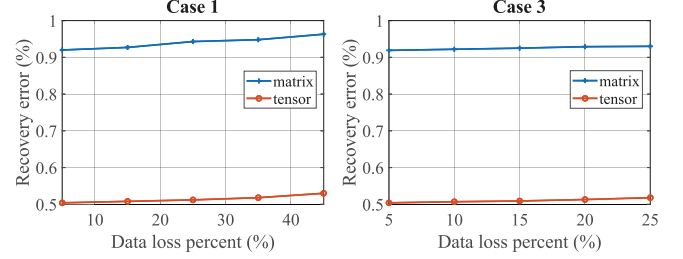


Fig. 7: Relative Recovery Error of the Disturbance Data.

The performance of the PD-based method is tested by applying it to matrix X^* and tensor $\mathcal{X}^*$ for Cases 1 and 3 of the data used in Section III-A. The results are shown in Fig. 7. The results demonstrate that conversion of the data into higher-dimensional objects can reduce the recovery error.

C. Adaptive Filtering Method

When significant temporal behaviors like periodic time series are present in the data, an adaptive filtering approach, called the temporal OnLine Algorithm for PMU data processing (OLAP-t), can yield good data recovery results. Let a measurement matrix $\Psi = X_k^T \in \mathbb{R}^{n \times m}$ be a sliding window matrix with n time samples and m channels. Assuming missing points only in the last row comprising the newest measurements, we partition the left singular vector matrix U of Ψ as

$$U = \begin{bmatrix} U_1 & U_0 \end{bmatrix} \quad (7)$$

where $U_0 \in \mathbb{R}^{n \times (m-r)}$ denotes the less significant singular vectors, and $U_2 \in \mathbb{R}^{1 \times r}$ is the last row of the first r left singular vectors. $U_1 \in \mathbb{R}^{(n-1) \times r}$ captures the one-step back temporal singular vectors (corresponding to known measurements.) Then, we predict the missing measurements by optimizing

$$\min_{c \in \mathbb{C}^{r \times n}} \|\Psi_{n-1} - U_1 c\|_2^2 \quad (8)$$

where Ψ_{n-1} includes the first $n-1$ rows of Ψ_k . The last row containing missing measurements is predicted by,

$$\hat{\Psi}_n = U_2 c \quad (9)$$

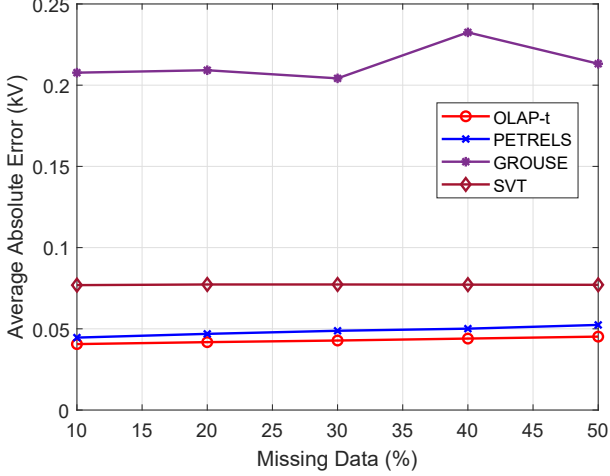


Fig. 8: Switching Event Comparison of Mean Absolute Error in kV for Increasing Frequency 5-Point Outages [26].

where Ψ_n corresponds to the last row of Ψ_k . The above process can be reformulated in an adaptive filtering form. If the $n - 1$ past measurements in step k are

$$\Psi_{n-1} = \begin{bmatrix} \psi_{k-n+1} \\ \vdots \\ \psi_{k-1} \end{bmatrix} \quad (10)$$

then

$$\hat{\Psi}_n = U_2(U_1^T U_1)^{-1} U_1^T \Psi_{n-1} \quad (11)$$

Thus, OLAP-t solves a least-squares minimization problem with a closed-form solution (11). By defining and calculating the coefficients c as

$$c = U_2(U_1^T U_1)^{-1} U_1^T = [c_1 \ c_2 \ \dots \ c_{n-1}] \quad (12)$$

we can obtain the missing data recovery (for a single channel) from the prediction equation

$$\hat{\psi}_k = c_1 \psi_{k-n+1} + \dots + c_{n-1} \psi_{k-1} \quad (13)$$

The predictions will be utilized to fill in potential missing data in the PMU data. For a detailed discussion on the filter formulation, stability and performance on various event types, the reader is referred to [26]. Fig. 8 shows the performance of OLAP-t compared to three other methods for recovering missing data in an oscillation event [26]. The error metric utilized is the average absolute error of the estimated measurements, defined as $\sum_{i=1}^N |\hat{\Psi}_i - \Psi_i|/N$, where $\hat{\Psi}_i$ implies an estimated value, and Ψ_i implies the ground-truth value of the i th missing data point. N is the number of missing values in the examined dataset. PETRELS [27] and GROUSE [28] are subspace tracking and missing point recovery methods of steaming data. SVT [7], on the other hand, solves the matrix recovery problem via a nuclear norm minimization, where the observed entries of the original matrix are preserved through the introduction of equality constraints in the optimization.

IV. BAD DATA DETECTION

Bad data is another type of data quality issue that can result from device malfunction, communication error, and cyber data attacks. Conventional bad data detectors usually require redundancy in the measurements and employ circuit laws to identify outliers. Exploiting the low-rank property, one can locate and correct bad data without any information about the system topology and line impedance.

Let $X^* \in \mathbb{C}^{m \times n}$ denote the ground-truth data from m channels in n time instant. Let S^* denote arbitrary additive errors in the measurements. The only assumption about S^* is that it contains a small percentage of non-zero entries, while most entries are zero. That indicates the percentage of bad data is low, but the error values can be large. The measurements with bad data can be represented by $M = X^* + S^*$. When there is no missing data, the problem of recovering the low-rank X^* from M is studied under the terminology Robust Principal Component Analysis (RPCA). When M further contains missing data, the problem of estimating X^* is called Robust Matrix Completion (RMC). Various convex and non-convex methods have been developed for both RPCA [8], [29], [30], [31], [32], [33] and RMC [34], [35], [36], [37].

The location of the bad data cannot be concentrated in one row or column for the successful correction by RPCA and RMC. For example, if the fraction of bad data is at most in the order of $1/r$ in each column and row, then methods such as [34], [38] are guaranteed to correct bad data and recover X^* accurately. However, conventional RPCA and RMC methods cannot recover X^* if all PMU channels have quality issues (missing data or bad data) simultaneously. Similarly, they also fail if all the data in one PMU channel have quality issues continuously.

One approach to address this issue is to exploit the low-rank property of the Hankel matrix to characterize the temporal correlations in the data governed by the underlying dynamical system. The data recovery problem is formulated in [39], [40] as

$$\begin{aligned} \min_{X, S} \quad & \sum_{(ij) \in \Omega} (M_{ij} - X_{ij} - S_{ij})^2 \\ \text{s.t.} \quad & \text{rank}(\mathcal{H}_k(X)) \leq r \quad \text{and} \quad \|S\|_0 \leq s, \end{aligned} \quad (14)$$

where Ω denotes the location of missing data, and $\|\cdot\|_0$ measures the number of nonzero entries. The parameters r and s are predetermined constants.

An alternating minimization algorithm named structured alternating projections (SAP) with recovery guarantee is provided in [39], [40] to solve (14). SAP iteratively removes bad data by hard thresholding using a threshold that decreases across iterations and computes a low-rank Hankel matrix that best approximates the remaining values in the current iteration. As long as the number of observations is at least in the order of $r^3 \log^2(n)$ (assuming $m \leq n$), the fraction of bad data in each channel is at most in the order of $1/r$, and the Hankel matrix of X^* has rank at most r , then SAP identifies all the bad data and recovers X^* accurately. Compared with the conventional RPCA and RMC methods, SAP has no constraint on the number of bad data at a given time instant

and can recover simultaneous bad data across all channels. The required number of observations is significantly smaller than conventional RMC methods, indicating that SAP can tolerate a much higher percentage of missing data. SAP converges linearly, and the per-iteration computational complexity is in the order of $rmn \log(n)$.

Fig. 9 compares the performance of SAP with R-RMC [38], a non-convex RMC algorithm, as well as using Alternating Direction Method of Multipliers (ADMM) to solve the convex relaxation of RMC problem. The dataset is shown in Fig. 2 with the time window from $t = 2.5$ s to 20s. In Fig. 9, given the percentage of bad data and missing data, the location of these data losses and outliers are selected uniformly at random in each independent trial. The results are averaged over 50 trials. A white block means accurately recovering X^* in all these trails. A black block corresponds to failures in all trails. SAP achieves the best performance among these methods. When data losses and bad data happen simultaneously across all the channels, SAP shows a similar performance as those in Fig. 9, while the other two methods fail because they cannot handle fully corrupted or lost columns.

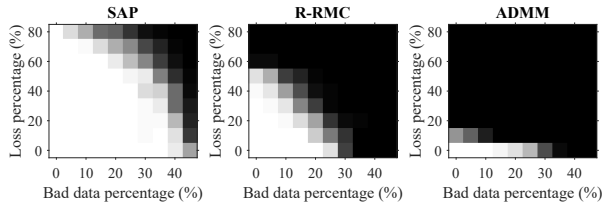


Fig. 9: Performance of SAP, R-RMC, and ADMM on the Disturbance Data (visualized in Fig. 2) [41] (©Cambridge Univ. Press, Reprinted with permission).

One advantage of low-rank methods is that no modeling of the power systems is required. However, if power system topology and line parameters are available, they can be easily incorporated into these methods to enhance performance. Because bus voltage phasors and line current phasors are related by circuit laws, their dependence can be modeled by additional equations in the constraints of (14). For example, [42] incorporates the topology information to identify cyber data attacks and proves that with these additional constraints, the developed method can correct a higher percentage of bad data resulting from cyber data attacks.

Ref. [14] develops a streaming PMU data recovery method to correct bad data and fill in missing entries in an online fashion. Assuming all the data before time t have already been corrected and recovered, the objective is to correct bad data and fill in missing entries in the observations obtained at time t . The idea is to construct a Hankel matrix using clean data from $t-L+1$ to $t-1$ ($L > 1$) and estimate the low-dimensional subspace by singular value decomposition. Then, based on the estimated subspace and the cleaned data from $t-\kappa+1$ to $t-1$, the data at time t can be estimated, where κ is the number of time instants in each column of the Hankel matrix. Then the method compares the estimation with the actual observations at time t to correct bad data and fill in missing data. The advantage of the online method is that the window length L

can be much smaller than the window length n in the block processing method in (14), because one only needs a small Hankel matrix of clean data to estimate the subspace. The computational complexity of the online method in [14] at time t is in the order of $\kappa m L^2$.

V. NONLINEAR SYSTEM APPROACH

The matrix, tensor, and filtering methods in Sections III-IV essentially assume the power system in the observation window can be approximated by a linear dynamical system. When the power system is undergoing a large disturbance, a nonlinear approach based on low-rankness can offer further advantages.

From the Koopman theory, for every nonlinear dynamical system, there exists a Koopman operator that lifts the system to an infinite-dimensional linear system with the same observations [16], [17]. That means there exists a linear dynamical system with order N (N can be infinite) described by

$$\mathbf{z}_{t+1} = \mathbf{A}\mathbf{z}_t, \quad \mathbf{x}_t = \mathbf{C}\mathbf{z}_t. \quad (15)$$

where $\mathbf{z}_t \in \mathbb{C}^N$ is the state variables in the lifted space at time t , the state matrix $\mathbf{A} \in \mathbb{C}^{N \times N}$, and the output matrix $\mathbf{C} \in \mathbb{C}^{m \times N}$.

Let $\mathbf{Z}$ contain the data in a window of length L , i.e.,

$$\mathbf{Z} = [\mathbf{z}_{t-L+1} \quad \mathbf{z}_{t-L+2} \quad \cdots \quad \mathbf{z}_t] \in \mathbb{C}^{N \times L}, \quad (16)$$

and

$$\mathcal{H}_\kappa(\mathbf{Z}) = \begin{bmatrix} \mathbf{z}_{t-L+1} & \mathbf{z}_{t-L+2} & \cdots & \mathbf{z}_{t-\kappa+1} \\ \mathbf{z}_{t-L+2} & \mathbf{z}_{t-L+3} & \cdots & \mathbf{z}_{t-\kappa+2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{z}_{t-L+\kappa} & \mathbf{z}_{t-L+\kappa+1} & \cdots & \mathbf{z}_t \end{bmatrix}. \quad (17)$$

If the lifted linear dynamical system can be approximated by a reduced-order system in the lifted space $\mathbb{C}^N$, then the rank of $\mathcal{H}_\kappa(\mathbf{Z})$ is much smaller than its ambient dimension. The low-rank property thus holds for $\mathcal{H}_\kappa(\mathbf{Z})$.

Ref. [18] considers the missing data recovery problem in the online setting, where the objective is to fill in the missing entries in $\mathbf{x}_t$ at time t using historical information in a window of length L . Suppose $\mathbf{z}_j$ are all known for $j < t$ and $\mathbf{C}$ is known. One can first use the historical data to estimate the r -dimensional column subspace in $\mathbb{C}^N$ of $\mathcal{H}_\kappa(\mathbf{Z})$, and then use $\mathbf{z}_{t-\kappa+1}, \dots, \mathbf{z}_{t-1}$ to determine the location of the last column vector of $\mathcal{H}_\kappa(\mathbf{Z})$ in (17). Then $\mathbf{z}_t$ can be estimated, followed by the estimation of $\mathbf{x}_t$. One can compare the estimation with the obtained observations to correct bad data and fill in missing entries.

Unfortunately, both $\mathbf{z}$ and $\mathbf{C}$ are unknown. However, the above described recovery method only requires $\langle \mathbf{z}_i, \mathbf{z}_j \rangle$ for any $1 \leq i, j \leq n$ rather than $\mathbf{z}_i$ directly. Therefore, one can employ a kernel function Φ [43] such that

$$\Phi(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{z}_i, \mathbf{z}_j \rangle. \quad (18)$$

Then, $\langle \mathbf{z}_i, \mathbf{z}_j \rangle$ can be computed directly from $\mathbf{x}_i$ and $\mathbf{x}_j$ without knowing $\mathbf{z}$. The kernel function is pre-defined, and popular choices include Gaussian kernel, polynomial kernel,

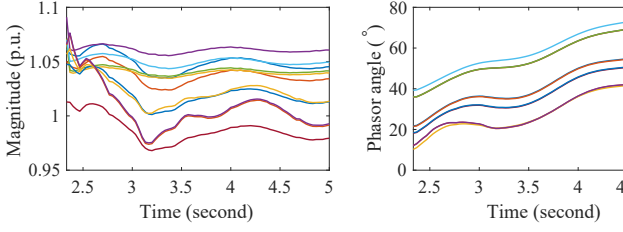


Fig. 10: Voltage Phasors Immediately after an Event [1]

and sigmoid kernel [43]. Overall, the computational complexity of correcting data issues for observations at time t is in order of $\kappa m L^2$.

This method is evaluated on the event data in Fig. 2 during the time interval 2.3s to 5s because the system is nonlinear immediately after a line trip. The voltage phasors in the interval are shown in Fig. 10. Given a fixed data loss percentage, we tested four cases of data loss patterns. Case 1 is missing data at random locations and times. Case 2 is simultaneous data losses across all channels at random time instants. Case 3 is simultaneous data losses across all channels for consecutive time instants. Case 4 is block-wise data losses occur across some (but not all) randomly selected channels and some consecutive instants repeatedly.

Fig. 11 shows the comparison of five different methods. Duplication simply replaces a missing entry with the most recent observation in the same channel. The low-rank Hankel method is the online method that exploits the low-rank-Hankel property of previous data samples to estimate the missing points in the current time instant [14]. Kalman filter [44] is a recursive filter that estimates the states of a dynamical system from noisy measurements and is implemented for each channel separately. Two versions of the lifted Hankel method are implemented. One method only recovers missing data (without bad data correction), while the other one also detects and corrects bad data (with bad data correction). A Gaussian kernel is selected in the lifted Hankel methods. The estimation error is measured by $\|\hat{X} - X^*\|_F / \|X^*\|_F$, where $\hat{X}$ and X^* denote the recovered and ground-truth data, respectively.

The Kalman filter method performs well when the data loss percentage is not high (not over 20%), and this is consistent with the conclusion in [44]. In contrast, the lifted Hankel methods can achieve a small estimation error in a wide range of data loss percentages and clearly outperform the other ones when the data loss percentage is high.

VI. DISTURBANCE CLASSIFICATION

Cascading failures often start from a preliminary phase that lasts from tens of seconds to hours. Once the power system stability is violated, a rapid phase is triggered, and events occur successively within milliseconds to tens of seconds [45]. Early identification and mitigation of events can prevent cascading failures [46], [47]. Data-driven methods extract features such as frequency [48], the rate of change of frequency (ROCOF) [49], and wavelet coefficients [50] from measurements and identify events based on the representative features. In addition to conventional event identification methods using SCADA

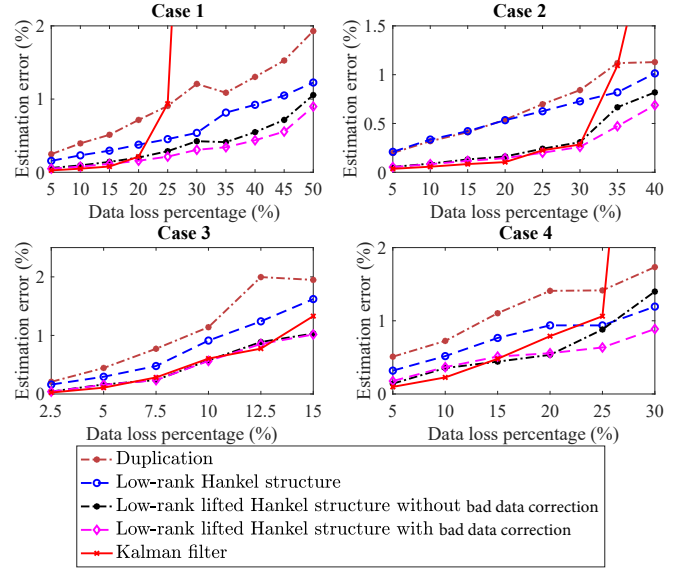


Fig. 11: Estimation Errors of Missing Voltage Phasors [18].

data, some data-driven methods using PMU data have been developed recently [51], [52], [53], [54], [55], [56], [57].

One idea is to use recorded event data with the corresponding labels to train a classifier, usually one variant of Neural Networks, such that when taking the PMU measurements as the input, the learned classifier outputs the label that corresponds to the event type. Despite the superior empirical performance of neural networks for image classification, natural language processing, and computer vision, its application on event identification has two limitations. The first one is the requirement of a large training set of various events at different system topologies with diverse initial conditions to train a reasonable classifier. The neural network architecture is thus complex and requires many computational resources to train. The second limitation is that training data usually contain individual events, while the actual observations may be affected by multiple successive events like cascading failures. For example, if a second event occurs when the system is still under the disturbance resulting from a previous event, directly applying the trained classifier may not identify the second event accurately.

To address the first issue, refs. [58], [59] propose to train on low-dimensional features extracted from PMU data rather than training on time series measurements directly. Given the measurements from m PMU channels in n time instants, [58] chooses the r -dimensional row subspace of the $m \times n$ data matrix as the feature because the row subspace is only determined by the r ($r \leq n$) dominant modes of the linear dynamical system and independent of the pre-event condition. Ref. [59] chooses r dominant singular values of the PMU data matrix and r dominant modes of the linear dynamical system as the low-dimensional representation of the data. The singular values of the data matrix are affected by the significant dynamics in the event. The modes of the linear dynamical system can be computed from Dynamic Mode Decomposition [60] and are affected by the characteristics of events because different events excite different sets of modes. The feature

size is thus reduced, and the features can compactly represent event data. For example, a line trip event at the same line but with different pre-event conditions leads to different time series measurements, but the excited modes might be the same. Then, different time series of event data may have similar features, and the required number of training cases can be significantly reduced, which in turn indicates that the neural network architecture can be greatly simplified (a two-layer convolutional neural network is used in [59]). Moreover, these features can be computed using data in a short time window (like one second) after an event happens rather than using a long window of data, and thus, an event can be identified immediately after it happens. Ref. [58] uses the low-rank subspaces of voltage magnitude, current magnitude, and frequency, respectively, to identify events. The identification performance using these three types of data are comparable. Ref. [59] computes r dominant modes in voltage magnitude, active power, and frequency, respectively, and uses all these $3r$ features to identify events.

To address the second issue, [58] assumes that the effects of different events on the measurements are additive and subtracts the effect of the first event before applying the classifier to identify the second event. Suppose the first event happens at time t_1 and the second event happens at time t_2 ($t_2 > t_1$). To predict the impact of the first event, one can use measurements during time $t_2 - L$ ($L < t_2 - t_1$) to $t_2 - 1$ and apply the online Hankel matrix completion method in [14] to estimate the observations after time t_2 if the second event did not happen. Then, subtracting this prediction from the actual observations after time t_2 provides the net effect of the second event.

Ref. [59] compares the proposed method, denoted by CNN-F, with training directly on time series measurements using convolutional neural networks, denoted by CNN-T, on simulated events that include line trips (LTs), generator trips (GTs), three-phase short circuit (TPs), and generator reference-voltage changes (GRCs). The number of parameters of CNN-T is eight times larger than that of CNN-F. However, CNN-F performs significantly better than CNN-T. The identification accuracy rate (IAR) of CNN-F is about 15% higher than that of CNN-T. Moreover, when the second event happens within one second of the first event, subtracting the impact of the first event increases the IAR of the second event by about 10%.

Fig. 12 compares the identification performance of CNN-F and CNN-T when the number of training samples decreases. The IAR of CNN-F is much higher than that of CNN-T and robust with respect to the size of the training data, while CNN-T degrades significantly when the number of events for training is reduced. Because the practical recorded events might not include all the possible event locations and pre-event conditions, CNN-F is a more practical option.

The methods for missing data recovery and bad data correction described in Sections III and IV essentially first estimate the low-rank subspaces based on the reliable data points and then use the estimated subspaces to fill in missing entries and correct bad data. If the disturbance data contain missing or bad data, one can simply apply the methods in Sections III and IV to estimate the low-rank subspaces and then apply the identification methods. Because the data recovery

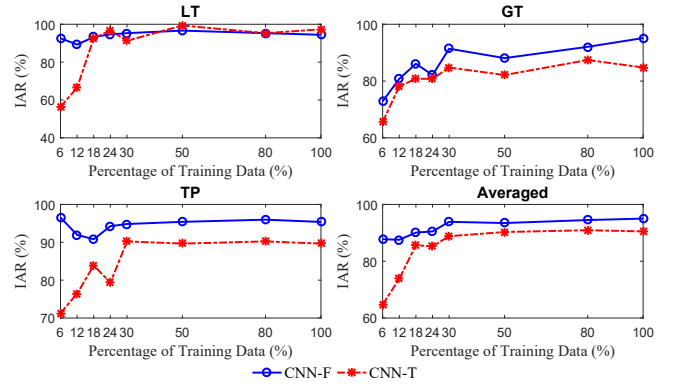


Fig. 12: Performances of CNN-F and CNN-T when Partial Training Datasets are Available [59].

methods can handle a significant percentage of missing or bad data, the estimated low-rank subspaces are accurate, and the identification performance will not be affected.

VII. CONCLUSION

This paper summarizes a recent line of work on PMU data processing by exploiting the low-rank property, which results from the strong dynamical corrections in the PMU measurements. The low-rank property enables reliable data recovery from severe data quality issues and simplifies disturbance identification without sacrificing the identification accuracy. These low-rank-based methods only provide the best estimate, while one future direction is to provide an uncertainty index to quantify the reliability of the returned solution. Other applications of low-rank methods may include data compression, quantifying the significance of disturbance events, and reducing the complexity in PMU-data machine learning methods.

VIII. ACKNOWLEDGMENT

We thank Prof. Hanoch Lev-Ari, Drs. Pengzhi Gao, Ying-shuai Hao, and Wenting Li for the contributions to the work discussed here. We also thank New York Power Authority for providing recorded PMU datasets.

REFERENCES

- [1] B. J. Walker, "Smart grid system report," *technical report*, November 2018.
- [2] S. Maslennikov, B. Wang, and E. Litvinov, "Dissipating energy flow method for locating the source of sustained oscillations," *International Journal of Electrical Power & Energy Systems*, vol. 88, pp. 55–62, June 2017.
- [3] B. J. Pierre, F. Wilches-Bernal, D. A. Schoenwald, R. T. Elliott, D. J. Trudnowski, R. H. Byrne, and J. C. Neely, "Design of the pacific dc intertie wide area damping controller," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3594–3604, March 2019.
- [4] N. B. Ingleby, "The statistical structure of forecast errors and its representation in the met. office global 3-d variational data assimilation scheme," *Quarterly Jour. Royal Meteorological Society*, vol. 29, pp. 127–231, January 2001.
- [5] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, August 2009.
- [6] L. Balzano, R. Nowak, and B. Recht, "Online identification and tracking of subspaces from highly incomplete information," in *Proc. Allerton Conf. Communication, Control, and Computing*, September 2010, pp. 704–711.

- [7] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optimiz.*, vol. 20, no. 4, pp. 1956–1982, January 2010.
- [8] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM (JACM)*, vol. 58, no. 3, p. 11, June 2011.
- [9] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Comput. Math.*, vol. 9, no. 6, pp. 717–772, December 2009.
- [10] P. Jain, R. Meka, and I. S. Dhillon, "Guaranteed rank minimization via singular value projection," in *Adv. Neural Inf. Process. Syst.*, September 2010, pp. 937–945.
- [11] J. H. Chow and J. Sanchez-Gasca, *Power System Modeling, Computation, and Control*. Wiley-IEEE Press, 2020.
- [12] L. Xie, C. Yang, and P. R. Kumar, "Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis," *IEEE Trans. Power Systems*, vol. 29, no. 6, pp. 2784–2794, April 2014.
- [13] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stofopoulos, "Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1006–1013, March 2016.
- [14] Y. Hao, M. Wang, J. H. Chow, E. Farrantos, and M. Patel, "Model-less data quality improvement of streaming synchrophasor measurements by exploiting the low-rank Hankel structure," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6966–6977, June 2018.
- [15] S. Zhang, Y. Hao, M. Wang, and J. H. Chow, "Multi-channel missing data recovery by exploiting the low-rank hankel structures," in *Proc. Int. Workshop Comput. Adv. Multi-Sensor Adaptive Process. (CAMSAP)*, December 2017, pp. 1–5.
- [16] Y. Susuki and I. Mezić, "Nonlinear koopman modes and power system stability assessment without models," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 899–907, November 2014.
- [17] E. Barocio, B. C. Pal, N. F. Thornhill, and A. R. Messina, "A dynamic mode decomposition framework for global power system oscillation analysis," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 2902–2912, December 2015.
- [18] Y. Hao, M. Wang, and J. H. Chow, "Likelihood analysis of cyber data attacks to power systems with markov decision processes," *IEEE Trans. Smart Grid*, pp. 3191–3202, November 2018.
- [19] P. Gao, M. Wang, S. G. Ghiocel, and J. H. Chow, "Modelless reconstruction of missing synchrophasor measurements," in *Proc. IEEE Power and Energy Society (PES) General Meeting (one of the Best Papers)*, July 2014, pp. 1–5.
- [20] J.-F. Cai, T. Wang, and K. Wei, "Fast and provable algorithms for spectrally sparse signal reconstruction via low-rank hankel matrix completion," *Applied and Computational Harmonic Analysis*, vol. 46, no. 1, pp. 94–121, January 2019.
- [21] S. Zhang, Y. Hao, M. Wang, and J. H. Chow, "Multi-channel Hankel matrix completion through nonconvex optimization," *IEEE J. Sel. Topics Signal Process., Special Issue on Signal and Information Processing for Critical Infrastructures*, vol. 12, no. 4, pp. 617–632, April 2018.
- [22] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, August 2009.
- [23] L. Sorber, M. V. Barel, and L. D. Lathauwer, "Unconstrained optimization of real functions in complex variables," *SIAM Journal on Optimization*, vol. 22, no. 3, pp. 879–898, July 2012.
- [24] L. Sorber, M. Van Barel, and L. De Lathauwer, "Optimization-based algorithms for tensor decompositions: Canonical polyadic decomposition, decomposition in rank- $(\mathbf{L}_r, \mathbf{L}_r, 1)$ terms, and a new generalization," *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 695–720, April 2013.
- [25] D. Osipov and J. H. Chow, "PMU missing data recovery using tensor decomposition," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4554–4563, May 2020.
- [26] 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, April 2020.
- [27] Y. Chi, Y. C. Eldar, and R. Calderbank, "Petrels: Parallel subspace estimation and tracking by recursive least squares from partial observations," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5947–5959, September 2013.
- [28] L. Balzano, R. Nowak, and B. Recht, "Online identification and tracking of subspaces from highly incomplete information," in *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, September 2010, pp. 704–711.
- [29] Y. Chen and M. J. Wainwright, "Fast low-rank estimation by projected gradient descent: General statistical and algorithmic guarantees," *arXiv:1509.03025*, September 2015.
- [30] Y. Chen, H. Xu, C. Caramanis, and S. Sanghavi, "Robust matrix completion with corrupted columns," in *Proc. International Conference on Machine Learning*, June 2011.
- [31] D. Hsu, S. M. Kakade, and T. Zhang, "Robust matrix decomposition with sparse corruptions," *IEEE Transactions on Information Theory*, vol. 57, no. 11, pp. 7221–7234, June 2011.
- [32] P. Netrapalli, U. Niranjan, S. Sanghavi, A. Anandkumar, and P. Jain, "Non-convex robust PCA," in *Adv. Neural Inf. Process. Syst.*, October 2014, pp. 1107–1115.
- [33] X. Yi, D. Park, Y. Chen, and C. Caramanis, "Fast algorithms for robust PCA via gradient descent," in *Advances in Neural Information Processing Systems* 29, May 2016, pp. 4152–4160.
- [34] Y. Chen, A. Jalali, S. Sanghavi, and C. Caramanis, "Low-rank matrix recovery from errors and erasures," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4324–4337, March 2013.
- [35] Q. Gu, Z. Wang, and H. Liu, "Low-rank and sparse structure pursuit via alternating minimization," in *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, vol. 51, May 2016, pp. 600–609.
- [36] O. Klopp, K. Lounici, and A. B. Tsybakov, "Robust matrix completion," *Probability Theory and Related Fields*, vol. 169, no. 1, pp. 523–564, October 2017.
- [37] A. Kyrillidis and V. Cevher, "Matrix alps: Accelerated low rank and sparse matrix reconstruction," in *2012 IEEE Statistical Signal Processing Workshop (SSP)*, August 2012, pp. 185–188.
- [38] Y. Cherapanamjeri, K. Gupta, and P. Jain, "Nearly optimal robust matrix completion," in *International Conference on Machine Learning*, July 2017, pp. 797–805.
- [39] S. Zhang and M. Wang, "Correction of simultaneous bad measurements by exploiting the low-rank hankel structure," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, June 2018, pp. 646–650.
- [40] S. Zhang and M. Wang, "Correction of corrupted columns through fast robust hankel matrix completion," *IEEE Trans. Signal Process.*, vol. 67, no. 10, pp. 2580–2594, March 2019.
- [41] M. Wang and J. Chow, "Data quality and privacy enhancement," in *Advanced Data Analytics for Power Systems*, A. Tajer, S. Perlaza, and H. Poor, Eds. Cambridge, UK: Cambridge University Press, 2021, pp. 261–282.
- [42] P. Gao, M. Wang, J. H. Chow, S. G. Ghiocel, B. Fardanesh, G. Stofopoulos, and M. P. Ranzanousky, "Identification of successive "unobservable" cyber data attacks in power systems," *IEEE Trans. Signal Process.*, vol. 64, no. 21, pp. 5557–5570, November 2016.
- [43] T. Hofmann, B. Schölkopf, and A. J. Smola, "Kernel methods in machine learning," *Ann. Stat.*, pp. 1171–1220, June 2008.
- [44] K. D. Jones, A. Pal, and J. S. Thorp, "Methodology for performing synchrophasor data conditioning and validation," *IEEE Trans. Power Syst.*, vol. 30, no. 3, pp. 1121–1130, May 2015.
- [45] P. Henneaux, P.-E. Labeau, and J.-C. Maun, "A level-1 probabilistic risk assessment to blackout hazard in transmission power systems," *Reliab. Eng. Syst. Safe.*, vol. 102, pp. 41–52, June 2012.
- [46] D. Novosel et al., "IEEE PSRC report on performance of relaying during wide-area stressed conditions," *IEEE Trans. Power Del.*, vol. 25, no. 1, pp. 3–16, January 2010.
- [47] H. H. Alhelou et al., "A survey on power system blackout and cascading events: Research motivations and challenges," *Energies*, vol. 12, no. 4, pp. 1–28, February 2019.
- [48] 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, October 2014.
- [49] A. Bykhovsky and J. H. Chow, "Power system disturbance identification from recorded dynamic data at the northfield substation," *Int. J. of Elec. Power*, vol. 25, no. 10, pp. 787 – 795, December 2003.
- [50] H. Jiang, J. J. Zhang, and D. W. Gao, "Fault localization in smart grid using wavelet analysis and unsupervised learning," in *Rec. 46th Asilomar Conf. Signals, Systems and Computers*, November 2012, pp. 386–390.
- [51] P. G. Axelberg, I. Y.-H. Gu, and M. H. Bollen, "Support vector machine for classification of voltage disturbances," *IEEE Trans. Power Del.*, vol. 22, no. 3, pp. 1297–1303, July 2007.
- [52] M. Garcia, T. Catanach, S. Vander Wiel, R. Bent, and E. Lawrence, "Line outage localization using phasor measurement data in transient state," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 3019–3027, September 2016.
- [53] 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, April 1995.

- [54] M. Mishra and P. K. Rout, "Detection and classification of micro-grid faults based on HHT and machine learning techniques," *IET Gener. Transm. Dis.*, vol. 12, pp. 388–397, February 2018.
- [55] M. Rafferty, X. Liu, D. M. Lavery, 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, April 2016.
- [56] 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 Trans. Power Syst.*, vol. 32, no. 6, pp. 4199–4210, January 2017.
- [57] 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 Comput. and Appl.*, vol. 31, no. 10, pp. 6719–6731, October 2019.
- [58] W. Li, M. Wang, and J. H. Chow, "Fast event identification through subspace characterization of PMU data in power systems," in *Proc. IEEE Power and Energy Society (PES) General Meeting*, July 2017, pp. 1–5.
- [59] W. Li and M. Wang, "Identifying overlapping successive events using a shallow convolutional neural network," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4762–4772, Nov. 2019.
- [60] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, "On dynamic mode decomposition: Theory and applications," *J. Comput. Dyn.*, vol. 1, no. 2, pp. 391–421, Nov. 2014.



Stavros Konstantinopoulos received the Diploma in Electrical and Computer Engineering from the National Technical University of Athens, Greece, in 2015 and the M.S. degree in Electrical Engineering from Rensselaer Polytechnic Institute, Troy, NY, USA in 2018.

He is currently a Ph.D. candidate at Rensselaer Polytechnic Institute. His research interests include integration and control of renewable generation, transient stability analysis and control and power system monitoring and PMU applications.



Shuai Zhang received the B.E. degree from University of Science and Technology of China, Hefei, China, in 2016.

He is pursuing the Ph.D. degree in electrical engineering at Rensselaer Polytechnic Institute, Troy, NY. His research interests include signal processing and high dimensional data analysis.



Meng Wang (M'12) received the Ph.D. degree from Cornell University, Ithaca, NY, USA, in 2012.

She is an Associate Professor in the department of Electrical, Computer, and Systems Engineering at Rensselaer Polytechnic Institute. Her research interests include high dimensional data analysis and their applications in power systems monitoring and network inference.



Joe H. Chow (F'92) received the M.S. and Ph.D. degrees from the University of Illinois, Urbana-Champaign, Urbana, IL, USA.

After working in the General Electric power system business in Schenectady, NY, USA, he joined Rensselaer Polytechnic Institute, Troy, NY, USA, in 1987, where he is Institute Professor of Electrical, Computer, and Systems Engineering. His research interests include power system dynamics and control, FACTS controllers, and synchronized phasor data. He is a member of the US National Academy

of Engineering. He is a past recipient of the IEEE PES Charles Concordia Power Engineering Award.



Evangelos Farantatos (S'06-M'13-SM'18) received the Diploma in electrical and computer engineering from the National Technical University of Athens, Greece, in 2006, and the M.S. and Ph.D. degrees from the Georgia Institute of Technology, Atlanta, GA, USA, in 2009 and 2012, respectively.

He is currently a Senior Project Manager with the Grid Operations and Planning Research and Development Group, EPRI, Palo Alto, CA, USA. He is managing and leading the technical work of various research and development projects related to synchrophasor technology, power systems monitoring and control, power systems stability and dynamics, renewable energy resources modeling, grid operation, and protection with high levels of inverter-based resources. In summer 2009, he was an intern with MISO.



Mahendra Patel (M'74-SM'81-F'16) received the B.E. from Sardar Patel University, Vallabh Vidyanagar, India, the M.S.E.E. degree from West Virginia University, Morgantown, WV, USA, and M.B.A. degree from the University of Pittsburgh, Pittsburgh, PA, USA.

He is a Technical Executive with the Grid Operations and Planning Department, Electric Power Research Institute (EPRI), Knoxville, TN, USA. He worked at PJM before joining EPRI. He has more than 40 years of experience in the Electric

Power Industry in transmission planning and system reliability, transmission technologies, system dynamics, synchrophasor technology, switching surges and insulation coordination, voltage stability, system protection and power quality.



Denis Osipov (S'14, M'19) received his B.S. and M.S. degrees in electrical engineering from Donetsk National Technical University, Donetsk, Ukraine in 2004 and 2005, respectively. He received his Ph.D. degree in electrical engineering from University of Tennessee, Knoxville, TN in 2018.

He is currently a postdoctoral research associate at the Department of Electrical, Computer, and Systems Engineering in Rensselaer Polytechnic Institute, Troy, NY. His research interests include power system stability, modeling, and monitoring.