Identification of "Unobservable" Cyber Data Attacks on Power Grids

Meng Wang[1], *Member, IEEE,* Pengzhi Gao,[1] *Student Member, IEEE*, Scott G. Ghiocel[1], *Member, IEEE,* Joe H. Chow[1], *Fellow, IEEE,* Bruce Fardanesh[2], *Fellow, IEEE,* George Stefopoulos[2], Michael P. Razanousky[3]
[1]Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, 12180, USA
[2]New York Power Authority, White Plains, NY, 10601, USA
[3]New York State Energy Research and Development Authority, Albany, NY, 12203.

***Abstract*—This paper presents a new framework of identifying cyber data attacks on synchrophasor measurements. We focus on detecting "unobservable" cyber data attacks that cannot be detected by any existing detection method that purely relies on measurements received at one time instant. Leveraging the approximate low-rank property of phasor measurement unit (PMU) data, we formulate the unobservable cyber attack identification problem as a matrix decomposition problem where the observed data matrix is the sum of a low-rank matrix plus a linear projection of a column-sparse matrix. We propose a convex-optimization-based decomposition method and provide its theoretical guarantee in the attack identification. Numerical experiments on actual PMU data and synthetic data are conducted to verify the effectiveness of the proposed method.**

I. Introduction

The monitoring, dispatch, and scheduling of power systems can benefit a lot from the integration of cyber infrastructures into future smart grids. Such integration, however, makes the power systems more susceptible to cyber attacks. It is reported that cyber spies have penetrated U.S. electrical grid [24]. Researchers have also launched an experimental cyber attack that caused a generator to self-destruct [15].

State estimation [1] is a critical component of power system monitoring. It continuously updates the system operator about the operating state based on measurements over various locations of the system. Errors in the measurement can affect the state estimations and mislead the system operator. Therefore, many efforts have been devoted to develop methods that can identify the bad data, see e.g., [6], [14], [23], [25], [32].

Cyber data attacks are firstly studied in [21] and can be viewed as "the worst interacting bad data injected by an adversary"[17]. Malicious intruders can simultaneously manipulate multiple measurements so that these attacks cannot be detected by any bad data detector. Because the removal of affected measurements would make the system unobservable, these attacks are termed as "unobservable attacks"[1] in [17].

State estimation in the presence of cyber data attacks has attracted much research attention recently [3], [9], [17], [21], [27], [28]. Existing approaches include protecting a small number of key measurement units such that the intruders cannot inject unobservable attacks without hacking protected units [3], [9], [16], as well as detectors designed for attacks in the observable regime [17]. Only one recent paper [28] considered the detection of unobservable attacks in Supervisory Control and Data Acquisition (SCADA) system and proposed a detection method based on statistical learning. The method in [28] has no theoretical guarantee and relies critically on the assumption that the measurements at different time instants are i.i.d. samples of random variables. This assumption might not hold when the system is experiencing some disturbances.

We propose a new method that can identify the unobservable cyber data attacks to PMUs. It has the theoretical guarantee of attack detection even when the system is under disturbance, provided that the attacker only. The intuition is that although these attacks are undetectable to detectors that rely only on instantaneous measurements, they can be identified by examining the temporal correlations in a sequence of measurements, as long as the intruders do not know the system dynamics.

Low-dimensional structure of PMU data matrix is recently observed in [7], [8], [12]. We formulate the identification problem as a decomposition problem of a low-rank matrix plus a linear projection of a column-sparse matrix. The matrix decomposition problem has attracted much research attention recently, see e.g., [4], [5], [26], [31], and have wide applications in areas like Internet monitoring [18], [22], [29], medical imaging [10], [11], image processing [2], etc. The situation that one component is a projection of a sparse matrix, however, has not been much addressed, except for a recent paper [22] that has a different problem formulation from ours.

The contributions of this paper are threefold. (1) We propose the idea of exploiting spatial-temporal correlations in measurements to identify unobservable data attacks. This differentiates our work with most existing methods that only explore spatial correlations. (2) We formulate the identification problem into a matrix decomposition problem and propose a computationally efficient identification method that does not involve the modeling of power system dynamics. (3) We provide the theoretical guarantee of our method in attack detection, as well as the general matrix decomposition problem.

The rest of the paper is organized as follows. We formulate the unobservable attack problem and point out its connection to other applications in Section II. We propose an identification method and provide its theoretical guarantee in Section III. Section IV records our numerical experiments. We conclude the paper in Section V.

[1]The term "unobservable" is used in this sense throughout the paper.

978-1-4799-4934-2/14/$31.00 ©2014 IEEE

II. Problem Formulation and Related Work

A. PMU measurements

Consider a power grid that contains n buses with PMUs installed on some buses. Let p denote the total number of PMU channels that measure voltage and current phasors. Matrix $M \in \mathbb{C}^{t\times p}$ contains the collected PMU measurements in t synchronized time instants. $\bar{\mathcal{J}} \in [\![p]\!]$ denotes the set of PMU channels that are under data attacks. We model the measurements under attacks as actual phasors plus additive errors. The observed measurement matrix can be presented as

$$M = \bar{L} + \bar{D}, \tag{1}$$

where $\bar{L} \in \mathbb{C}^{t\times p}$ represents the actual phasors without data attacks, and $\bar{D} \in \mathbb{C}^{t\times p}$ corresponds to injections by data attacks. Note column $\bar{D}_i$ is a zero vector for every $i \notin \mathcal{J}$.

As observed in [7], [8], [12], the PMU matrix $\bar{L}$ exhibits low-dimensional structure. In Section II of [12], we analyzed actual PMU data from the Central New York Power System (Fig. 1). Six PMUs measure 37 voltage and current phasors, and the data are reported at 30 samples per second. Fig. 2 shows the current magnitudes of PMU data for a 20-s period. An event occurs around $2.5s$. The singular values of the data matrix are plotted in Fig. 3. The ninth largest singular value is 0.5930, while the largest one is 894.6. Therefore we can approximate the PMU data matrix by a low-rank matrix.

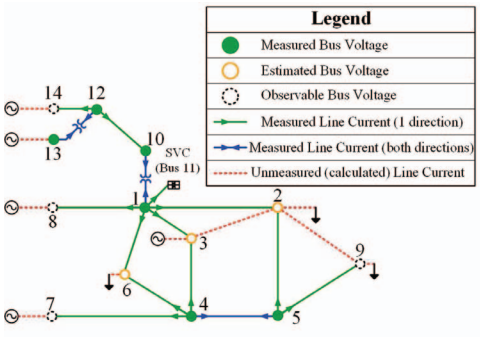

Fig. 1: Six PMUs in the Central NY Power System

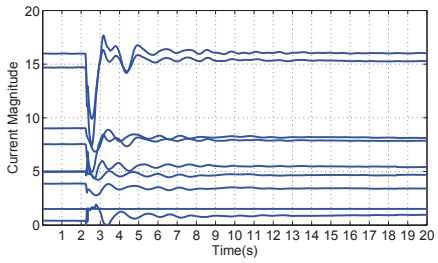

Fig. 2: Current magnitudes of PMU data (9 current phasors out of 37 phasors)

The Singular Value Decomposition (SVD) of $\bar{L}$ is

$$\bar{L} = \bar{U}\bar{\Sigma}\bar{V}^{\dagger},$$

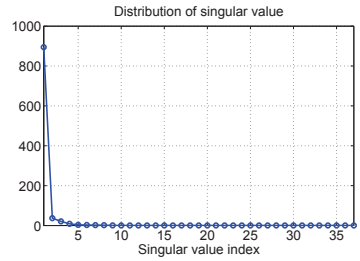

Fig. 3: Singular values of PMU data matrix in decreasing order

where $\bar{U} \in \mathbb{C}^{t\times r}$, $\bar{\Sigma} \in \mathbb{C}^{r\times r}$, $\bar{V} \in \mathbb{C}^{p\times r}$ $(r \ll t, p)$[2], and $\bar{V}^{\dagger}$ is the conjugate transpose of $\bar{V}$. We assume that every nonzero column of $\bar{D}$ does not lie in the column space of $\bar{L}$, which is legitimate when the intruders do not have full information about the system dynamics.

The notations are summarized in Table I. Matrix A is *column-sparse* if it contains a small fraction of non-zero columns. We call the set of indices of nonzero columns the *column support* of A.

TABLE I: Notations

A_i	the ith column of matrix A.
$A_{i,*}$	the ith row of matrix A.
$A_{\mathcal{I}}$	the submatrix of A with column indices in $\mathcal{I}$.
$\mathcal{P}_{\mathcal{I}}(A)$	matrix obtained from A by setting A_i to zero for all $i \notin \mathcal{I}$.
$A \in \mathcal{P}_{\mathcal{I}}$	if and only if $\mathcal{P}_{\mathcal{I}}(A) = A$.
$\|A\|$	the spectral norm.
$\|A\|_*$	the nuclear norm of A, which is the sum of singular values.
$\|A\|_F$	the Frobenius norm.
$\|A\|_{1,2}$	the sum of ℓ_2 norms of the columns A_i.
$\|A\|_{\infty,2}$	the largest ℓ_2 norm of the columns.
$\mathcal{P}_U(A)$	$:= UU^{\dagger}A$, the projection of A onto the column space of L.
$\mathcal{P}_V(A)$	$:= AVV^{\dagger}$, the projection of A onto the row space.
$\mathcal{P}_T(\cdot)$	$:= \mathcal{P}_U(\cdot) + \mathcal{P}_V(\cdot) - \mathcal{P}_U\mathcal{P}_V(\cdot)$.
$\mathcal{P}_{U^{\perp}}(A)$	$:= (I - UU^{\dagger})A$.
$\mathcal{P}_{V^{\perp}}(A)$	$:= A(I - VV^{\dagger})$.
$\mathcal{P}_{T^{\perp}}(A)$	$:= \mathcal{P}_{U^{\perp}}\mathcal{P}_{V^{\perp}}(A)$.
$A \in \mathcal{P}_T$	if and only if $\mathcal{P}_T(A) = A$.
$\mathcal{I}^c$	the complimentary set of set $\mathcal{I}$.

B. Unobservable cyber data attacks

We use voltage phasors as state variables, and let $X \in \mathbb{C}^{t\times n}$ contain the state variables at t instants. We use the π equivalent model to represent a transmission line (Fig. 4). Z^{ij} and Y^{ij} denote the impedance and admittance of the transmission line between bus i and bus j. Current I^{ij} from bus i to bus j is related to bus voltage V^i and V^j by

$$I^{ij} = \frac{V^i - V^j}{Z^{ij}} + V^i\frac{Y^{ij}}{2}.$$

We define $\bar{W} \in \mathbb{C}^{p\times n}$ as follows. If the kth PMU channel measures the voltage phasor of bus j, $\bar{W}_{kj} = 1$; if it measures the current phasor from bus i to bus j, then $\bar{W}_{ki} = 1/Z^{ij} +$

[2]Strictly speaking, $\bar{L}$ is approximately low-rank and can be viewed as the summation of a low-rank matrix plus noise. We assume $\bar{L}$ is strictly low-rank for notational simplicity, and the results can be extended to approximate low-rank matrices with simple modifications.

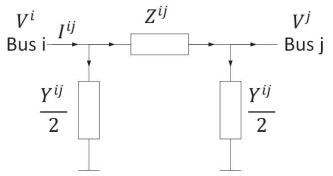

Fig. 4: π model of a transmission line

$Y^{ij}/2$, $\bar{W}_{kj} = -1/Z^{ij}$; $\bar{W}_{kj} = 0$ otherwise. The actual PMU measurements and the state variables are related by

$$\bar{L} = X\bar{W}^T. \tag{2}$$

The attack at time i is called *unobservable*[3] if and only if

$$M_{i,*} = \bar{L}_{i,*} + \bar{D}_{i,*} = X_{i,*}\bar{W}^T + \bar{D}_{i,*} = (X_{i,*} + \alpha^i)\bar{W}^T$$

holds for some nonzero row vector $\alpha^i \in \mathbb{C}^{1\times n}$. The attack, denoted by data injection $\bar{D}_{i,*}$, is unobservable since no detector can differentiate $X_{i,*}$ and $X_{i,*} + \alpha^i$ based on $M_{i,*}$. We consider the scenario that the attacks at all time instants are unobservable. The attack matrix can be represented by

$$\bar{D} = \begin{bmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^t \end{bmatrix} \bar{W}^T := \bar{C}W^T \tag{3}$$

where $W_j = \bar{W}_j/\|\bar{W}_j\|$. $\bar{C}$ represents the additive error (up to a scaling factor) to bus voltages due to data attacks, i.e., $\|\bar{W}_j\|\bar{C}_j$ is the error to bus voltage V^j. Let $\bar{\mathcal{I}} \in [\![n]\!]$ denote the column support of $\bar{C}$. We assume $\bar{C}$ is column-sparse because intruders might only alter some of the state variables due to resource constraints.

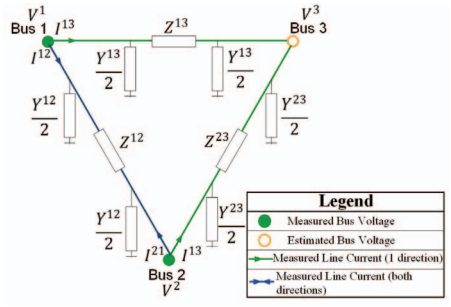

Fig. 5: Three-bus example

We use a three-bus network (Fig. 5) to illustrate the notations. Let $\mathbf{V}^1$, $\mathbf{V}^2$, $\mathbf{V}^3$, $\mathbf{I}^{12}$, $\mathbf{I}^{13}$, $\mathbf{I}^{21}$, $\mathbf{I}^{23} \in \mathbb{C}^{t\times 1}$ denote the bus voltages and line currents in t instants. Assume two PMUs are installed at bus 1 and bus 2. The PMU measurements without attacks are

$$\bar{L} = [\mathbf{V}^1, \mathbf{I}^{12}, \mathbf{I}^{13}, \mathbf{V}^2, \mathbf{I}^{21}, \mathbf{I}^{23}] = [\mathbf{V}^1, \mathbf{V}^2, \mathbf{V}^3]\bar{W}^T \tag{4}$$

[3][21] is centered on DC model where power measurements and state variables are approximately related by linear equations. Here PMU measurements and state variables are accurately related by linear equation (2).

where $\bar{W}^T$ is shown as follows.

$$\begin{bmatrix} 1 & \frac{1}{Z^{12}} + \frac{Y^{12}}{2} & \frac{1}{Z^{13}} + \frac{Y^{13}}{2} & 0 & -\frac{1}{Z^{12}} & 0 \\ 0 & -\frac{1}{Z^{12}} & 0 & 1 & \frac{1}{Z^{12}} + \frac{Y^{12}}{2} & \frac{1}{Z^{23}} + \frac{Y^{23}}{2} \\ 0 & 0 & -\frac{1}{Z^{13}} & 0 & 0 & -\frac{1}{Z^{23}} \end{bmatrix}$$

Suppose the intruder attacks all channels of PMU 1 and the channel of PMU 2 that measures $\mathbf{I}^{13}$ and manipulates these measurements so that the system operator would have the wrong estimation that the system states are $[\mathbf{V}^1 + \boldsymbol{\beta}^1, \mathbf{V}^2 + \boldsymbol{\beta}^2, \mathbf{V}^3]$ for any nonzero $\boldsymbol{\beta}^1$, $\boldsymbol{\beta}^2 \in \mathbb{C}^{t\times 1}$. In this case, the measurements under attacks should be

$$\begin{aligned} M &= [\mathbf{V}^1 + \boldsymbol{\beta}^1, \mathbf{V}^2, \mathbf{V}^3 + \boldsymbol{\beta}^2]\bar{W}^T \\ &= [\mathbf{V}^1 + \boldsymbol{\beta}^1, \mathbf{I}^{12} + \frac{\boldsymbol{\beta}^1}{Z^{12}} + \frac{\boldsymbol{\beta}^1 Y^{12}}{2}, \mathbf{I}^{13} + \frac{\boldsymbol{\beta}^1 - \boldsymbol{\beta}^2}{Z^{13}} + \frac{\boldsymbol{\beta}^1 Y^{12}}{2}, \\ &\quad \mathbf{V}^2, \mathbf{I}^{21} - \boldsymbol{\beta}^1/Z^{12}, \mathbf{I}^{13} - \boldsymbol{\beta}^2/Z^{13}]. \end{aligned}$$

The additive errors due to attacks are

$$\bar{D} = M - \bar{L} = [\boldsymbol{\beta}^1, \boldsymbol{\beta}^2, \mathbf{0}]\bar{W}^T = [\|\bar{W}_1\|\boldsymbol{\beta}^1, \|\bar{W}_2\|\boldsymbol{\beta}^2, \mathbf{0}]W^T.$$

Therefore, we have

$$M = \bar{L} + \bar{C}W^T. \tag{5}$$

With the increasing installation of PMUs in a smart grid, we anticipate that the total number of PMU channels p will be larger than the number of buses n. Thus, the transform in (3) reduces the dimension of the unknowns in data attacks.

C. Applications beyond power system monitoring

Our problem formulation in (5) is closely related to those in [31] and [22]. We borrow the ideas from [31] for our identification method and analysis; however, we address a more general framework of matrix decomposition than the one in [31], and the analysis here is more involved. $\bar{C}$ in [22] is assumed to be a sparse matrix where the locations of nonzero entries in different rows are different and selected independently of other rows, while we consider the scenario that the locations of the nonzero entries are the same for each row of $\bar{C}$. The corresponding decomposition methods and analysis are also different.

Our proposed method can be applied to other domains that involve decomposing a matrix as in (5). As discussed in [22], these applications include unveiling network traffic anomalies [18], [29], dynamic magnetic resonance imaging [10], [11], face recognition [2], and separation of singing voices from music accompanies [19], [20]. For the detection of network traffic anomalies, each column of M corresponds to the traffic volume on one link in the Internet across t time instants. $\bar{L}$ represents the normal traffic on the links, and $\bar{C}$ represents the anomalies in origin-to-destination (OD) flows. W represents the routing matrix, which is usually a flat matrix. [22] considers the scenario that the OD flows that contain anomalies are different in different sampling instants, while we consider the scenario that the abnormal OD flows do not change in a short period of time.

Method 1 Unobservable cyber attack identification method

Input: PMU measurements in t instants, represented by M

Find (L^*, C^*), the optimum solution to the following optimization problem

$$\min_{L\in\mathbb{C}^{t\times p},C\in\mathbb{C}^{t\times n}} \|L\|_* + \lambda\|C\|_{1,2} \quad \text{s.t.} \quad M = L + CW^T \quad (6)$$

Compute the SVD of $L^* = U^*\Sigma^*V^{*\dagger}$.

Find column support of $D^* = C^*W^T$, denoted by $\mathcal{J}^*$.

Return: $L^*_{\mathcal{J}^{*c}}$, U^* and $\mathcal{J}^*$.

III. Identification of Data attacks

A. Identification method and guarantee

We say a method can *identify* an unobservable attack if it successfully determines the set $\mathcal{J}$ of PMU channels that are under attack and reconstructs measurements that are not attacked. Our proposed method is summarized in Method 1. Note that (6) is a convex program and can be solved efficiently by genetic solvers such as CVX[13]. Given $\bar{L} = \bar{U}\bar{\Sigma}\bar{V}^\dagger$ and W, we define

$$\epsilon := \|\bar{V}^\dagger W\|_{\infty,2},$$

$$\mu := \max_{i\neq j} \|W_i^\dagger W_j\|,$$

$$\sigma_k := \max_{\mathcal{I}:|\mathcal{I}|\leq k} \|(W_\mathcal{I}^\dagger W_\mathcal{I})^{-1}\|.$$

Note that $\sigma_1 = 1$ as W has unit-norm columns. ϵ depends on the rank r of $\bar{L}$, since $\|\bar{V}\|_F^2 = r$.

Given positive k, define the function

$$f_k(x) := 1/(2-x) + \sqrt{(1+(k-1)\mu)\sigma_k x} - 1,$$

and let $\hat{\psi}_k$ be the solution to $f_k(x) = 0$. Then $\hat{\psi}_k < 1$ for all positive k since $f_k(x)$ is strictly increasing in x, and $f_k(x) = 1$. Lemma 1 demonstrates the existence of a proper λ for Method 1 to identify the unobservable data attacks.

Lemma 1. *If it holds that*

$$(2-\hat{\psi}_1)\mu + (3-\hat{\psi}_1)\epsilon + \hat{\psi}_1 \leq 1, \quad (7)$$

then pick the largest positive integer k^* such that

$$\sigma_{k^*}^2\mu(k^* + (k^{*2}-k^*)\mu) + \sigma_{k^*}\epsilon(3-\hat{\psi}_k)\sqrt{k^* + (k^{*2}-k^*)\mu} + k^*\sigma_{k^*}\mu(1-\hat{\psi}_{k^*}) + \hat{\psi}_{k^*} \leq 1 \quad (8)$$

holds. Define

$$\lambda_{\min} = \frac{\epsilon(1-\psi^*)}{(1-k^*\sigma_{k^*}\mu)(1-\psi^*) - \sigma_{k^*}\epsilon\Gamma - \sigma_{k^*}^2\mu\Gamma^2},$$

$$\lambda_{\max} = \frac{1-\psi^*}{(2-\psi^*)\sigma_{k^*}\Gamma},$$

where $\psi^* := \hat{\psi}_{k^*}$, and $\Gamma := \sqrt{k^* + (k^{*2}-k^*)\mu}$. Then we have $\lambda_{\min} \leq \lambda_{\max}$.

Proof: If $k^* = 1$, then (8) is reduced to (7). Then condition (7) guarantees the existence of a positive k^*. Then (8) implies that $\lambda_{\min} \leq \lambda_{\max}$. ∎

Theorem 1. *As long as the column support of $\bar{C}$ has size at most k^*, for any $\lambda \in [\lambda_{\min}, \lambda_{\max}]$, where k^*, $\lambda_{\min}$ and $\lambda_{\max}$ are defined in Lemma 1, the output of Method 1 satisfies*

$$U^*U^{*\dagger} = \bar{U}\bar{U}^\dagger,$$

$$\mathcal{J}^* = \bar{\mathcal{J}} \text{ and } L^*_{\mathcal{J}^{*c}} = \bar{L}_{\bar{\mathcal{J}}^c}.$$

Theorem 1 guarantees that the "clean" PMU measurements could be identified. Furthermore, the subspace spanned by the actual phasors can be recovered, even though we do not recover the actual measurements that are under attack.

Method 1 is greatly inspired by [31]. In fact, after post-multiplying $W(W^TW)^{-1}$ to both sides of (1), we have

$$MW(W^TW)^{-1} = \bar{L}W(W^TW)^{-1} + \bar{C}, \quad (9)$$

where the right-hand side is the summation of a low-rank matrix plus a column-sparse matrix. Then, the results of [31] can be directly applied to (9). We do not follow this path due to two reasons. First, $MW(W^TW)^{-1}$ cannot be computed if some entries of M are missing, while Method 1 can be easily extended to scenarios with missing data by restricting the constraints in (6) to the observed measurements. Second, $(W^TW)^{-1}$ does not exist when W is a flat matrix, i.e., $p < n$, while Method 1 and Theorem 1 do not rely on the assumption that W is a tall matrix and can be applied to an arbitrary W. When W is an Identify matrix, Theorem 1 can be reduced to Theorem 1 of [31]. We skip this discussion here.

To the best of our knowledge, only one recent paper [28] considered the detection of unobservable attacks to SCADA data. The authors assume the measurements are i.i.d. random variables. Their method requires sufficient ambient data for statistical learning and does not have theoretical guarantee. Our method leverages the low-dimensional property of PMU measurements and is guaranteed to identify data attacks even though the system is under disturbance.

B. Proof sketch of Theorem 1

The proof of Theorem 1 follows the same line as the proof of Theorem 1 in [31]. With the additional projection matrix W, our proof is more involved than the one in [31].

We design the following Oracle Problem (10) by adding explicit constraints that the solution pair should have the correct column space of $\bar{L}$ and the correct column support of $\bar{C}$. The major step is to show that an optimal solution (L^*,C^*) to (6) is also an solution to Oracle problem (10). Note that Oracle problem is only designed for analysis, since $\bar{U}$ and $\bar{\mathcal{I}}$ are unknown to the operator.

$$\begin{aligned} \text{Oracle Problem} \quad & \min_{L,C} \|L\|_* + \lambda\|C\|_{1,2} \\ & \text{s.t.} \quad M = L + CW^T \\ & \quad\quad \mathcal{P}_{\bar{U}}(L) = L,\ \mathcal{P}_{\bar{\mathcal{I}}}(C) = C. \end{aligned} \quad (10)$$

Let (L', C') be an optimal solution to Oracle problem (10). We define $\mathcal{P}_{T'}(Q) := \mathcal{P}_{U'} + \mathcal{P}_{V'} - \mathcal{P}_{U'}\mathcal{P}_{V'}$, where the SVD of $L' = U'\Sigma'V'^\dagger$. Define

$$\mathfrak{G}(C') := \{H \in \mathbb{C}^{t\times k} \mid \forall i \in \mathcal{I}' : H_i = C'_i/\|C'_i\|; \quad \forall i \in \bar{\mathcal{I}} \cap (\mathcal{I}')^c : \|H_i\|_2 \leq 1\}$$

where $\mathcal{I}'$ is the column support of C'. From [31] we have

Lemma 2 (Lemma 4 and Lemma 5 of [31])**.**

$$U'U'^{\dagger} = \bar{U}\bar{U}^{\dagger}.$$

There exists an orthonormal matrix $\hat{V} \in \mathbb{C}^{t\times p}$ *such that*

$$U'V'^{\dagger} = \bar{U}\hat{V}^{\dagger}. \tag{11}$$

Also, we have

$$\mathcal{P}_{T'} := \mathcal{P}_{U'} + \mathcal{P}_{V'} - \mathcal{P}_{U'}\mathcal{P}_{V'} = \mathcal{P}_{\bar{U}} + \mathcal{P}_{\hat{V}} - \mathcal{P}_{\bar{U}}\mathcal{P}_{\hat{V}}.$$

The condition when a solution to Oracle problem (10) is also a solution to (6) is stated in the following lemma,

Lemma 3. *An optimal solution* (L', C') *to (10) is an optimal solution to (6) if there exists* $Q \in \mathbb{C}^{t\times p}$ *that satisfies*

$$\begin{aligned} &(a)\mathcal{P}_{T'}(Q) = U'V'^{\dagger}, \qquad (b)\|\mathcal{P}_{T'^{\perp}}(Q)\| \le 1, \\ &(c)(QW)_{\bar{\mathcal{I}}}/\lambda \in \mathfrak{G}(C'), \quad \text{and } (d)\|(QW)_{\bar{\mathcal{I}}^c}\|_{\infty,2} \le \lambda. \end{aligned} \tag{12}$$

If both (b) and (d) are strict, and $\mathcal{P}_{\bar{\mathcal{J}}} \cap \mathcal{P}_{V'} = \{0\}$*, then any optimal solution* (L^*, C^*) *to (6) satisfies* $\mathcal{P}_{\bar{U}}(L^*) = L^*$*,* $\mathcal{P}_{\bar{\mathcal{I}}}(C^*) = C^*$*.*

The major technical step is to construct Q, called the *dual certificate*, that satisfies (12). Our construction method is as follows. Pick $\hat{H} \in \mathfrak{G}(C')$ that satisfies

$$\hat{V}^{\dagger}W_{\bar{\mathcal{I}}} = \lambda\bar{U}^{\dagger}\hat{H}. \tag{13}$$

Define

$$\Phi := \lambda\hat{H}(W_{\bar{\mathcal{I}}}^{\dagger}W_{\bar{\mathcal{I}}})^{-1}W_{\bar{\mathcal{I}}}^{\dagger}, \tag{14}$$

$$\Delta_1 := \mathcal{P}_{\bar{U}}(\Phi),$$

$$\Delta_2 := \mathcal{P}_{\bar{U}^{\perp}}(I - \mathcal{P}_{W_{\bar{\mathcal{I}}}})\mathcal{P}_{\hat{V}}(I + \sum_{i=1}^{\infty}(\mathcal{P}_{\hat{V}}\mathcal{P}_{W_{\bar{\mathcal{I}}}}\mathcal{P}_{\hat{V}})^i)\mathcal{P}_{\hat{V}}(\Phi), \tag{15}$$

where

$$\mathcal{P}_{W_{\bar{\mathcal{I}}}}(X) := XW_{\bar{\mathcal{I}}}(W_{\bar{\mathcal{I}}}^{\dagger}W_{\bar{\mathcal{I}}})^{-1}W_{\bar{\mathcal{I}}}^{\dagger}.$$

$$Q := \bar{U}\hat{V}^{\dagger} + \Phi - \Delta_1 - \Delta_2. \tag{16}$$

We show that Q in 16 is well defined in Appendix-B. The following lemma shows that Q in (16) is the desired dual certificate.

Lemma 4. *If the column support of* $\bar{C}$ *has size no greater than* k^**, then for any* $\lambda \in [\lambda_{\min}, \lambda_{\max}]$*, where* k^**,* $\lambda_{\min}$ *and* $\lambda_{\max}$ *are defined in Lemma 1,* Q *defined in (16) satisfies (12).*

Theorem 1 follows when we combine Lemma 3 and Lemma 4. Due to the space limitation, we skip the proofs of these two key lemmas. Please refer to [30] for the proofs.

IV. Simulation

We explore the performance of data attack identification method on both synthetic data and actual PMU data. We use CVX [13] to solve (6).

A. Performance on synthetic data

We investigate the performance of Method 1 according to different rank r of $\bar{L}$ and the number of corrupted columns of $\bar{C}$, using randomly generated synthetic data. Fix $t = p = 50$. Given rank r, we generate matrices $A \in \mathbb{R}^{t\times r}$ and $B \in \mathbb{R}^{p\times r}$ with each entry independently drawn from Gaussian $\mathcal{N}(0,1)$ and set $\bar{L} := AB^T$. We generate matrix $W \in \mathbb{R}^{p\times n}$ with independent $\mathcal{N}(0,1)$ entries. To generate a column-sparse matrix $\bar{C} \in \mathbb{R}^{t\times n}$, we randomly select the column support and set the nonzero entries to be independent $\mathcal{N}(0,1)$. We simulate the observed measurement matrix M according to 5.

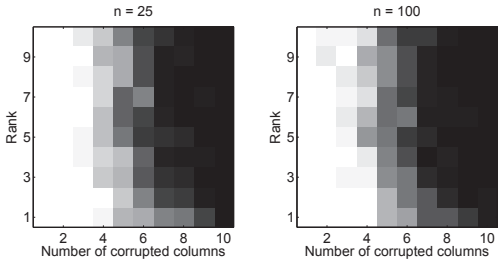

Fig. 6: Matrix decomposition performance for different n

We apply Method 1 to obtain the estimation (L^*, C^*). λ is set to be 0.9. We identify a column of C^* to be nonzero if its ℓ_2 norm exceeds some predefined threshold that is close to zero. Method 1 succeeds if $\|U^*U^{*\dagger} - \bar{U}\bar{U}^{\dagger}\|_2 \le \epsilon$ for some small positive ϵ, and the column supports of $\bar{C}$ and C^* are the same. We vary rank r and the number of corrupted columns and take 100 runs for each case. Fig. 6 shows the transition property of Method 1 in gray scale. White stands for 100% success while black denotes 100% failure. When n is 25, W is a tall matrix ($p > n$). When n is 100, W is a flat matrix ($p < n$). For both simulations, method succeeds even when rank r is eight, and $\bar{C}$ has two nonzero columns.

B. Performance on actual PMU data

We consider the PMU data shown in Section II-A. We consider the scenario that the intruder changes the measurements of I^{12} and I^{52} to inject unobservable attacks that will lead to an error in the estimation of the voltage of bus 2. Two 2-second PMU datasets are tested. One contains ambient data ($t = 17 \sim 19s$ in Fig. 2). The other one contain the data with abnormal event ($t = 2 \sim 4s$ in Fig. 2). λ is set to be 1.

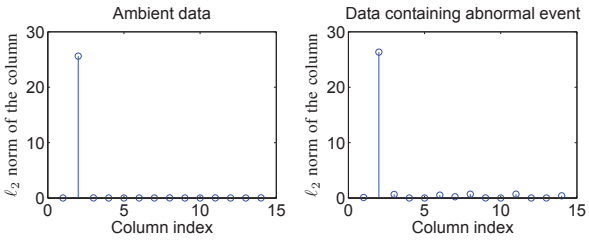

Fig. 7: ℓ_2 norm of each column of C^* for ambient and disturbed data

Because actual PMU data contain noise, we identify a column to be nonzero if its ℓ_2 norm exceeds some predefined threshold. Fig. 7 shows the ℓ_2 norm of each column of the resulting C^* matrix by Method 1. The column index Fig. 7 corresponds to the bus number in the power system. One can see our method successfully identifies that the estimation of the voltage of bus 2 is corrupted. Because the entries in W corresponding to I^{12} and I^{52} are nonzero for bus 2, then the PMU channels measuring I^{12} and I^{52} can also be correctly identified as corrupted through the relationship in (3).

V. Conclusion and Discussions

Existing cyber attack detection methods usually use the measurements at one time instant and only explore the spatial correlations. We further exploit the temporal correlations and develop a method that are guaranteed to correctly identify cyber data attacks that were considered unobservable. Although motivated by power system monitoring, our results can be applied to other scenarios like Internet traffic monitoring, dynamic MRI, face recognition, etc.

We have not considered data losses and measurement noise in this paper. We are currently developing the identification method of cyber data attacks and its theoretical guarantee when part of the measurements are lost during the communication to the central operator. We are also analyzing the performance when the measurement contain noise.

Acknowledgement

We thank New York Power Authority for providing PMU data for the Central NY Power System. This research is supported in part by the ERC Program of NSF and DoE under NSF Award EEC-1041877 and the CURENT Industry Partnership Program, and in part by NYSERDA Grants #36653 and #28815.

References

[1] A. Abur and A. G. Exposito, *Power system state estimation: theory and implementation*. CRC Press, 2004.

[2] R. Basri and D. W. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 2, pp. 218–233, 2003.

[3] R. B. Bobba, K. M. Rogers, Q. Wang, H. Khurana, K. Nahrstedt, and T. J. Overbye, "Detecting false data injection attacks on DC state estimation," in *Proc. the First Workshop on Secure Control Systems (SCS)*, 2010.

[4] 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, 2011.

[5] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky, "Rank-sparsity incoherence for matrix decomposition," *SIAM Journal on Optimization*, vol. 21, no. 2, pp. 572–596, 2011.

[6] J. Chen and A. Abur, "Placement of PMUs to enable bad data detection in state estimation," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1608–1615, 2006.

[7] Y. Chen, L. Xie, and P. Kumar, "Dimensionality reduction and early event detection using online synchrophasor data," in *Proc. IEEE Power and Energy Society General Meeting*, 2013, pp. 1–5.

[8] 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.

[9] G. Dán and H. Sandberg, "Stealth attacks and protection schemes for state estimators in power systems," in *Proc. IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2010, pp. 214–219.

[10] J. P. Finn, K. Nael, V. Deshpande, O. Ratib, and G. Laub, "Cardiac MR imaging: State of the technology1," *Radiology*, vol. 241, no. 2, pp. 338–354, 2006.

[11] H. Gao, J.-F. Cai, Z. Shen, and H. Zhao, "Robust principal component analysis-based four-dimensional computed tomography," *Physics in medicine and biology*, vol. 56, no. 11, p. 3181, 2011.

[12] P. Gao, M. Wang, S. Ghiocel, and J. H. Chow, "Modeless reconstruction of missing synchrophasor measurements," in *Proc. IEEE PES General Meeting*, 2014.

[13] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," http://cvxr.com/, Oct. 2010.

[14] E. Handschin, F. Schweppe, J. Kohlas, and A. Fiechter, "Bad data analysis for power system state estimation," *IEEE Trans. Power App. Syst.*, vol. 94, no. 2, pp. 329–337, 1975.

[15] C. Herridge, M. Levine, M. Emanuel, and M. Oinounou, "Sources: Staged cyber attack reveals vulnerability in power grid," http://www.foxnews.com/politics/2009/04/08/cyberspies-penetrate-power-grid-leave-software-disrupt/, 2009.

[16] T. Kim and H. Poor, "Strategic protection against data injection attacks on power grids," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 326–333, 2011.

[17] O. Kosut, L. Jia, R. Thomas, and L. Tong, "Malicious data attacks on smart grid state estimation: Attack strategies and countermeasures," in *Proc. IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2010, pp. 220–225.

[18] A. Lakhina, M. Crovella, and C. Diot, "Diagnosing network-wide traffic anomalies," in *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 4, 2004, pp. 219–230.

[19] Y. Li and D. Wang, "Separation of singing voice from music accompaniment for monaural recordings," *IEEE Trans. Audio, Speech, Language Process.*, vol. 15, no. 4, pp. 1475–1487, 2007.

[20] Z. Lin, A. Ganesh, J. Wright, L. Wu, M. Chen, and Y. Ma, "Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix," *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, vol. 61, 2009.

[21] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security (TISSEC)*, vol. 14, no. 1, p. 13, 2011.

[22] M. Mardani, G. Mateos, and G. Giannakis, "Recovery of low-rank plus compressed sparse matrices with application to unveiling traffic anomalies," *IEEE Trans. Inf. Theory*, vol. 59, no. 8, pp. 5186–5205, 2013.

[23] H. M. Merrill and F. C. Schweppe, "Bad data suppression in power system static state estimation," *IEEE Trans. Power App. Syst.*, no. 6, pp. 2718–2725, 1971.

[24] J. Meserve, "Sources: Staged cyber attack reveals vulnerability in power grid," *http://www.cnn.com/2007/US/09/26/power.at.risk/* , 2007.

[25] A. Monticelli and A. Garcia, "Reliable bad data processing for real-time state estimation," *IEEE Trans. Power App. Syst.*, no. 5, pp. 1126–1139, 1983.

[26] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Rev.*, vol. 52, no. 3, pp. 471–501, 2010.

[27] H. Sandberg, A. Teixeira, and K. H. Johansson, "On security indices for state estimators in power networks," in *Proc. the First Workshop on Secure Control Systems (SCS)*, 2010.

[28] H. Sedghi and E. Jonckheere, "Statistical structure learning of smart grid for detection of false data injection," in *Proc. IEEE Power and Energy Society General Meeting (PES)*, 2013, pp. 1–5.

[29] M. Thottan and C. Ji, "Anomaly detection in IP networks," *IEEE Trans. Signal Process.*, vol. 51, no. 8, pp. 2191–2204, 2003.

[30] M. Wang, P. Gao, S. Ghiocel, J. H. Chow, B. Fardanesh, G. Stefopoulos, and M. P. Razanousky, "Identification of "unobservable" cyber data attacks on power grids," 2014. [Online]. Available: http://ecse.rpi.edu/~wang/pub/SmartGridComm14.pdf

[31] H. Xu, C. Caramanis, and S. Sanghavi, "Robust pca via outlier pursuit," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3047–3064, May 2012.

[32] W. Xu, M. Wang, L. Lai, and A. Tang, "Sparse error correction from nonlinear measurements with applications in bad data detection for power networks," *IEEE Trans. Signal Process.*, vol. 61, no. 24, pp. 6175–6187, 2013.