This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2020.2982267, IEEE Transactions on Power Systems

A Synchrophasor Data-driven Method for Forced Oscillation Localization under Resonance Conditions

Tong Huang, *Student Member, IEEE,* Nikolaos M. Freris, *Senior Member, IEEE,* P. R. Kumar, *Fellow, IEEE,* and Le Xie, *Senior Member, IEEE*

***Abstract*—This paper proposes a data-driven algorithm for locating the source of forced oscillations and suggests a physical interpretation for the method. By leveraging the sparsity of forced oscillations along with the low-rank nature of synchrophasor data, the problem of source localization under resonance conditions is cast as computing the sparse and low-rank components using Robust Principal Component Analysis (RPCA), which can be efficiently solved by the exact Augmented Lagrange Multiplier method. Based on this problem formulation, an efficient and practically implementable algorithm is proposed to pinpoint the forced oscillation source during real-time operation. Furthermore, theoretical insights are provided for the efficacy of the proposed approach, by use of physical model-based analysis, specifically by highlighting the low-rank nature of the resonance component matrix. Without the availability of system topology information, the proposed method can achieve high localization accuracy in synthetic cases based on benchmark systems and real-world forced oscillations in the power grid of Texas.**

***Index Terms*—Forced oscillations (FOs), Phasor Measurement Unit (PMU), Resonant Systems, Robust Principal Component Analysis (RPCA), Unsupervised Learning, Big Data.**

I. Introduction

PHASOR measurement units (PMUs) enhance the transparency of bulk power systems by streaming the fast-sampled and synchronized measurements to system control centers. Such finely-sampled and time-stamped PMU measurements can reveal several aspects of the rich dynamical behavior of the grid which are invisible to conventional supervisory control and data acquisition (SCADA) systems. Among the system dynamical behaviors exposed by PMUs, *forced oscillations* (FOs) have attracted significant attention within the power community. FOs are driven by periodical exogenous disturbances that are typically injected by malfunctioning power apparatuses such as wind turbines, steam extractor valves of generators, or poorly-tuned control systems [1]–[3]. Cyclic loads, such as cement mills and steel plants, constitute another category of oscillation sources [1]. The impact of such injected periodic perturbation propagates through transmission lines and results in FOs throughout the grid; some real-world events of FOs since 1966 are reported in [1].

This work is supported in part by Ministry of Science and Technology of China Grant 2019YFB2102200, Anhui Department of Science and Technology Grant 201903a05020049, NSF Contract ECCS-1760554, NSF Science & Technology Center Grant CCF-0939370, CCF-1717207, Electric Reliability Council of Texas (ERCOT), the Power Systems Engineering Research Center (PSERC), and Department of Energy Grant DE-EE0009031. Preferred address for correspondence: Le Xie, Dept. of ECE, Texas A&M Univ., 3128 TAMU, College Station, TX 77843-3128.

The presence of FOs compromises the security and reliability of power systems. For example, FOs may trigger protection relays to trip transmission lines or generators, potentially causing uncontrollable cascading failures and unexpected load shedding [4]. Moreover, sustained FOs reduce device lifespans by introducing undesirable vibrations and additional wear and tear on power system components; consequently, failure rates and maintenance costs of compromised power apparatuses might increase [4]. Therefore, timely suppression of FOs is important to system operators.

One effective way of suppressing a forced oscillation is to locate the oscillation's source, a canonical problem that we call *forced oscillation localization*, and then to disconnect it from the power grid. A natural attempt to conduct forced oscillation localization could be tracking the largest oscillation over the power grid, under the assumption that measurements near the oscillatory source are expected to exhibit the most severe oscillations, based on engineering intuition. However, counter-intuitive cases may occur when the frequency of the periodic perturbation lies in the vicinity of one of the natural modes of the power system, whence a *resonance phenomenon* is triggered [5]. In such cases, PMU measurements exhibiting the most severe oscillations may be geographically far from where the periodic perturbation is injected, posing a significant challenge to system operators in pinpointing the forced oscillation source. It is worth noting that such counter-intuitive cases are more than a mere theoretical concern: one example occurred at the Western Electricity Coordinating Council (WECC) system on Nov. 29, 2005, when a 20-MW forced oscillation initiated by a generation plant at Alberta incurred a tenfold larger oscillation at the California-Oregon Inter-tie line that is 1100 miles away from Alberta [3]. Such a severe oscillation amplification significantly compromises the security and reliability of the power grid. Hence, it is imperative to develop a forced oscillation localization method that is effective even in the challenging but highly hazardous cases of resonance [6].

In order to pinpoint the source of FOs, several localization techniques have been developed. In [7], forced oscillation localization is achieved based on the following observation: the measurements near the source manifest distinct signatures in their magnitude or phase responses, in comparison to far away measurements. Such an observation is interpretable based on classic generator models, but whether it is valid or not in a power system with complex generator dynamics remains an open question [7]. In [2], the authors leverage the oscillation energy flows in power networks to locate the source of sustained oscillations. In this energy-based method, the energy flows can be computed using the preprocessed PMU data, and the power system components generating the oscillation energy are identified as the oscillation sources. In spite of the promising performance of the energy-based method [2], the rather stringent assumptions pertaining to knowledge of load characteristics and the grid topology may restrict its usefulness to specific scenarios [6], [8]. Reference [8] provides a comprehensive summary of FO localization methods. More recent research on FO localization is reported in [9] and [10]. In [9], the oscillation source is located by comparing the measured current spectrum of system components with one predicted by the effective admittance matrix. However, the construction of the effective admittance matrix requires accurate knowledge of system parameters that may be unavailable in practice. In [10], generator parameters are learned from measurements based on prior knowledge of generator model structures, and, subsequently, the admittance matrix is constructed and used for FO localization. Nevertheless, model structures of generators might not be known beforehand, owing to the unpredictable switching states of power system stabilizers [11]. Thus, it is highly desirable to design a FO localization method that does not heavily depend upon availability of the first-principle model and topology information of the power grid.

In this paper, we propose a *purely data-driven* yet *physically interpretable* approach to pinpoint the source of FOs in the challenging resonance case. By leveraging the sparsity of the FO sources and the low-rank nature of high-dimensional synchrophasor data, the problem of forced oscillation localization is formulated as computing the sparse and low-rank components of the measurement matrix using Robust Principal Component Analysis (RPCA) [12]. Based on this problem formulation, an algorithm for real-time operation is designed to pinpoint the source of FOs. The main merits of the proposed approach include the following: 1) It does not require any information on dynamical system model parameters or topology, thus providing an efficient and easily deployable practical implementation; 2) It can locate the source of FOs with high accuracy, even when resonance phenomena occur; and 3) Its efficacy can be interpreted by physical model-based analysis.

The rest of this paper is organized as follows: Section II elaborates on the forced oscillation localization problem and its main challenges; in Section III, the FO localization is formulated as a matrix decomposition problem and a FO localization algorithm is designed; Section IV provides theoretical justification of the efficacy of the algorithm; Section V validates the effectiveness of the proposed method in synthetic cases based on benchmark systems and real-world forced oscillations in the power grid of Texas; Section VI summarizes the paper and poses future research questions.

II. Localization of Forced Oscillations and Challenges

A. Mathematical Interpretation

The dynamic behavior of a power system in the vicinity of its operation condition can be represented by a continuous linear time-invariant (LTI) state-space model:

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

$$\mathbf{y}(t) = C\mathbf{x}(t) + D\mathbf{u}(t), \tag{1b}$$

where state vector $\mathbf{x} \in \mathbb{R}^n$, input vector $\mathbf{u} \in \mathbb{R}^r$, and output vector $\mathbf{y} \in \mathbb{R}^m$ collect the *deviations* of state variables, generator/load control setpoints, and measurements, from their respective steady-state values. Accordingly, matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times r}$, $C \in \mathbb{R}^{m\times n}$, and $D \in \mathbb{R}^{m\times r}$ are termed as the state matrix, the input matrix, the output matrix, and the feed-forward matrix, respectively. Typically, the input vector $\mathbf{u}$ is not streamed to control centers, so the feed-forward matrix D is assumed to be a zero matrix of appropriate dimension. Denote by $\mathcal{L} = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ the set of all eigenvalues of the state matrix A. The power system (1) is assumed to be stable, with all eigenvalues $\lambda_i \in \mathbb{C}$ being distinct, i.e., $\mathrm{Re}\{\lambda_i\} < 0$ for all $i \in \{1, 2, \dots, n\}$ and $\lambda_i \neq \lambda_j$ for all $i \neq j$. Note that the assumption on eigenvalue distinctness is only used for the purpose of simplifying the process of obtaining the time-domain solution of outputs in Section IV. Due to a large amount of symbols in this paper, the key symbols are summarized in the appendix for the convenience of readers.

We proceed to formally define the concepts of a forced oscillation source and source measurements. Suppose that the l-th input $u_l(t)$ in the input vector $\mathbf{u}(t)$ varies periodically due to malfunctioning components (generators/loads) in the grid. In such a case, $u_l(t)$ can be decomposed into J frequency components, viz.,

$$u_l(t) = \sum_{j=1}^{J} P_j \sin(\omega_j t + \theta_j), \tag{2}$$

where $\omega_j \neq 0$, $P_j \neq 0$ and θ_j are the frequency, amplitude and phase displacement of the j-th frequency component of the l-th input, respectively. Equation (2) is effectively equivalent to the Fourier series representation of a periodic signal [13]. As a consequence, the periodic input will result in sustained oscillations present in the measurement vector $\mathbf{y}$. The generator/load associated with input l is termed as the *forced oscillation source*, and the measurements at the bus directly connecting to the forced oscillation source are termed as *source measurements*.

In particular, suppose that the frequency ω_d of an injection component is close to the frequency of a poorly-damped mode, i.e., there exists $j^* \in \{1, 2, \dots, n\}$,

$$\omega_d \approx \mathrm{Im}\{\lambda_{j^*}\}, \quad \mathrm{Re}\{\lambda_{j^*}\} \approx 0. \tag{3}$$

In such a case, resonance phenomena can be observed [5]. Hence, (3) is adopted as the *resonance condition* in this paper. Studies on envelop shapes of FOs are reported in [14].

In a power system with PMUs, the measurement vector $\mathbf{y}(t)$ is sampled at a frequency of f_s (samples per second). Within a time interval from the FO starting time up to time instant t, the time evolution of the measurement vector $\mathbf{y}(t)$ can be discretized by sampling and represented by a matrix called a *measurement matrix* $Y_t = [y^t_{p,q}]$, which we formally define next. Without loss of generality, we assume that the FOs start at time 0. The following column concatenation defines the measurement matrix Y_t up to time t:

$$Y_t := \left[\mathbf{y}(0), \quad \mathbf{y}(1/f_{\mathrm{s}}), \quad \dots \quad \mathbf{y}(\lfloor t f_{\mathrm{s}} \rfloor / f_{\mathrm{s}})\right], \qquad (4)$$

where $\lfloor \cdot \rfloor$ denotes the floor operation. The i-th column of the measurement matrix Y_t in (4) suggests the "snapshot" of all synchrophasor measurements over system at the time $(i-1)/f_{\mathrm{s}}$. The k-th row of Y_t denotes the time evolution of the k-th measurement deviation in the output vector of the k-th PMU. Due to the fact that the output vector may contain multiple types of measurements (e.g., voltage magnitudes, frequencies, etc.), a normalization procedure is introduced as follows. Assume that there are K measurement types. Denote by $Y_{t,i} = [y^{t,i}_{p,q}] \in \mathbb{R}^{r_0 \times c_0}$ the measurement matrix of measurement type i, where $i = \{1, 2, \dots, K\}$. The normalized measurement matrix $Y_{\mathrm{nt}} = [y^{\mathrm{n},t}_{p,q}]$ is defined by

$$Y_{\mathrm{nt}} = \left[\frac{Y_{t,1}^\top}{\|Y_{t,1}\|_{\max}}, \quad \frac{Y_{t,2}^\top}{\|Y_{t,2}\|_{\max}}, \quad \dots \quad \frac{Y_{t,K}^\top}{\|Y_{t,K}\|_{\max}}\right]^\top, \qquad (5)$$

where $\|\cdot\|_{\max}$ returns the largest absolute element of a matrix.

The forced oscillation localization problem is equivalent to pinpointing the forced oscillation source using measurement matrix Y_t. Due to the complexity of power system dynamics, the precise power system model (1) may not be available to system operators, especially in real-time operation. Therefore, it is assumed that the only known information for forced oscillation localization is the measurement matrix Y_t. In brief, the first-principle model (1) as well as the perturbation model (2) is introduced mainly for the purpose of defining FO localization problem and theoretically justifying the data-driven method proposed in Section III, but is not needed for the proposed algorithm.

B. Main Challenges of Pinpointing the Sources of Forced Oscillation

The topology of the power system represented by (1) can be characterized by an undirected graph $G = (\mathcal{B}, \mathcal{T})$, where vertex set $\mathcal{B}$ comprises all buses in the power system, while edge set $\mathcal{T}$ collects all transmission lines. Suppose that the PMU measurements at bus $i_{\mathrm{s}} \in \mathcal{B}$ are the source measurements. Then bus j is said to be in the vicinity of the FO source if bus j is a member of the following vicinity set:

$$\mathcal{V}_0 = \{j \in \mathcal{B} | d_G(i_{\mathrm{s}}, j) \le N_0\}, \qquad (6)$$

where $d_G(i,j)$ denotes the i-j distance, viz., the number of transmission lines (edges) in a shortest path connecting buses (vertices) i and j; the threshold N_0 is a nonnegative integer. In particular, $\mathcal{V}_0 = \{i_{\mathrm{s}}\}$ for the source measurement at bus i_{s}, if N_0 is set to zero.

Intuitively, it is tempting to presume that the source measurement can be localized by finding the maximal absolute element in the normalized measurement matrix Y_{nt}, i.e., expecting that the most severe oscillation should be manifested in the vicinity of the source. However, a major challenge for pinpointing the FO sources arises from the following (perhaps counter-intuitive) fact: the most severe oscillation does not necessarily manifest near the FO source, in the presence of *resonance phenomena*. Following the same notation as in (4) and (6), we term a normalized measurement matrix Y_{nt} as *counter-intuitive case*, if

$$i^* \notin \mathcal{V}_0, \qquad (7)$$

where i^* can be obtained by finding the row index of the maximal element in the measurement matrix Y_t, i.e.,

$$[i^*, j^*] = \arg\max_{i,j} \left| y^{\mathrm{n},t}_{i,j} \right|. \qquad (8)$$

It is such counter-intuitive cases that make pinpointing the FO source challenging [5]. Figure 1 illustrates one such counter-intuitive case, where the source measurement (red) does not correspond to the most severe oscillation. Additional examples of counter-intuitive cases can be found in [6]. Although the counter-intuitive cases are much less likely to happen than the intuitive ones (in terms of frequency of occurrence), it is still imperative to design an algorithm to pinpoint the FO source even in the counter-intuitive cases due to the hazardous consequences of the FOs under resonance conditions.

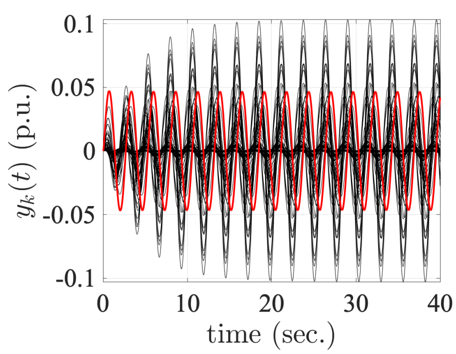


Fig. 1. One counter-intuitive case [6] from the IEEE 68-bus benchmark system [15]: the black curves correspond to the non-source measurements; the red curve corresponds to the source measurement.

III. Problem Formulation and Proposed Methodology

In this section, we formulate the FO localization problem as a matrix decomposition problem. Then, we present a FO localization algorithm for real-time operation.

A. Problem Formulation

Given a measurement matrix Y_t up to time t with one type of measurement (without loss in generality), the FO source localization is formulated as decomposing the measurement matrix Y_t into a low-rank matrix L_t and a sparse matrix S_t:

$$Y_t = L_t + S_t, \qquad (9a)$$

$$\operatorname{rank} L_t \le \gamma, \qquad (9b)$$

$$\|S_t\|_0 \le \beta, \qquad (9c)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2020.2982267, IEEE Transactions on Power Systems

where the pseudo-norm $\|\cdot\|_0$ returns the number of non-zero elements of a matrix; the non-negative integer γ is the upper bound of the rank of the low-rank matrix L_t, and the non-negative integer β is the upper bound on the number of non-zero entries in the sparse matrix S_t. Given non-negative integers γ and β, it is possible to numerically find $\{L_t, S_t\}$ via *alternating projections* [6]. The source measurement index p^* can be tracked by finding the largest absolute value in the sparse matrix S_t, viz.,

$$[p^*, q^*]^\top = \arg\max_{p,q} \left|s^t_{p,q}\right|. \tag{10}$$

The intuition behind the formulation (9) is as follows. As the power grid is an interconnected system, measurements at different buses have certain electrical couplings, resulting in correlations between the measurements. As a result, the measurements at different buses should exhibit a "general trend," [6] which can be captured by a low-rank matrix L_t. The measurements near the FO source are assumed to deviate most from its corresponding component in "general trend" (the low-rank matrix L_t). The deviation is supposed to be captured by the matrix S_t. As the number of the measurements near the FO source is limited, the matrix S_t is assumed to be sparse.

Due to the prior unavailability of the upper bounds γ and β [6], the matrix decomposition problem shown in (9) is reformulated as an instance of *Robust Principal Component Analysis (RPCA)* [12]:

$$\min_{S_t} \quad \|Y_t - S_t\|_\star + \xi \|S_t\|_1, \tag{11}$$

where $\|\cdot\|_\star$ and $\|\cdot\|_1$ denote the nuclear norm and l_1 norm, respectively; the tunable parameter ξ regulates the extent of sparsity in S_t. The formulation in (11) is a convex relaxation of (9). Under some assumptions, the sparse matrix S_t and the low-rank matrix L_t can be disentangled from the measurement matrix Y_t [12] by diverse algorithms [16]. The exact Lagrange Multiplier Method (ALM) is used for numerically solving the formulation (11). Recall that the measurement matrix Y_t has r_0 rows and c_0 columns. The tunable parameter ξ is suggested to be $1/\sqrt{k_0}$, where $k_0 = \max\{r_0, c_0\}$. Such selection of ξ is justified via the mathematical analysis in [12]. For a measurement matrix containing multiple measurement types, (11) can be modified by replacing Y_t with Y_{nt}.

B. FO Localization Algorithm for Real-time Operation

Next, we present a FO localization algorithm for real-time operation, using the formulation (11). In order to determine the starting time of forced oscillations, we can leverage the methods reported in [17], [18]. The method reported in [17] is used to detect FOs by comparing the periodogram of PMU measurements with a frequency-dependent threshold. In [18] the authors propose a method that uses geometric analysis on streaming synchrophasor data to estimate the starting and end times of FOs. Once periodic FOs are detected by the method reported in [17], the starting time of the FOs can be estimated by the time-localization algorithm proposed in [18]. A window of measurements with the starting time is collected into forming the measurement matrix. Then Algorithm 1 is triggered for pinpointing the FO source. In Algorithm 1, T_0 and ξ are user-defined parameters.

Algorithm 1 Real-time FO Localization

1: Update Y_{T_0} by (4);
2: Obtain Y_{nT_0} by (5);
3: Find S_t in (11) via the exact ALM for chosen ξ;
4: Obtain p^* by (10);
5: **return** p^* as the source measurement index.

Algorithm 1 can be leveraged to illustrate the intuition behind formulation (9) described in Section III-A. A measurement matrix Y_t can be formed based on the measurements visualized in Figure 1. Algorithm 1 can decompose Y_t into a low-rank matrix L_t and a sparse matrix S_t. Figure 2 visualizes Y_t, L_t, and S_t in a *normalized* fashion. For each matrix, we take the absolute values of their entries and normalize the absolute version of the entries by the maximal absolute entry in the corresponding matrix. The magnitudes of the *normalized* entries are represented by color: The bigger the magnitude of an entry, the yellower is its color, and conversely the smaller the magnitude of an entry, the bluer is its color. The "general trend" of the measurements is captured by the low-rank matrix L_t in Figure 2(b). The deviations from the "general trend" are captured by the sparse matrix S_t. In Figure 2(c), very few entries are colored with yellow, and these entries correspond measurements deviating most from the "general trend", while most entries are colored with dark blue, suggesting that most entries are close to zero. The entry colored with brightest color corresponds to Bus 65 which is the bus closest to the force oscillation source (Generator 13).

IV. Theoretical Interpretation of the RPCA-based Algorithm

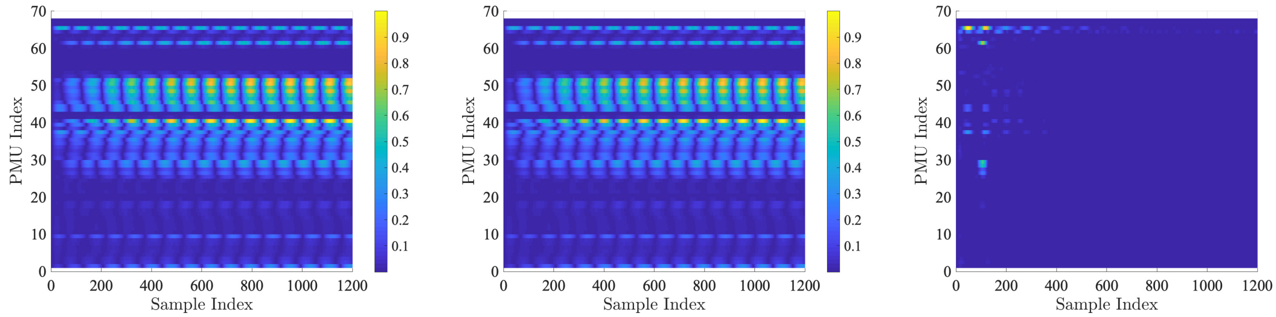
This section aims to develop a theoretical connection between the first-principle model in Section II and the data-driven approach presented in Section III. We start such an investigation by deriving the time-domain solution to PMU measurements in a power system under resonance conditions. Then, the resonance component matrix for the power grid is obtained from the derived solution to PMU measurements. Finally, the efficacy of the proposed method is interpreted by examining the rank of the resonance component matrix.

A. PMU Measurement Decomposition

For the power system with r inputs and m PMU measurements modeled using (1), the k-th measurement and the l-th input can be related by

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + \mathbf{b}_l u_l(t) \tag{12a}$$
$$y_k(t) = \mathbf{c}_k \mathbf{x}(t), \tag{12b}$$

where column vector $\mathbf{b}_l \in \mathbb{R}^n$ is the l-th column of matrix B in (1), and row vector $\mathbf{c}_k \in \mathbb{R}^n$ is the k-th row of matrix C. With the assumption on eigenvalue distinctness, let $\mathbf{x} = M\mathbf{z}$, where $\mathbf{z}$ denotes the transformed state vector and matrix M is



(a) (b) (c)

Fig. 2. Visualization of the measurement matrix Y_t (a), the low-rank matrix L_t (b), and the sparse matrix S_t (c)

.

chosen such that the similarity transformation of A is diagonal, then

$$\dot{\mathbf{z}}(t) = \mathbf{\Lambda}\mathbf{z}(t) + M^{-1}\mathbf{b}_l u_l(t) \tag{13a}$$

$$y_k(t) = \mathbf{c}_k M\mathbf{z}(t), \tag{13b}$$

where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = M^{-1}AM$ is a diagonal matrix stacking the eigenvalues of A. Denote by column vector $\mathbf{r}_i \in \mathbb{C}^n$ and row vector $\mathbf{l}_i \in \mathbb{C}^n$ the right and left eigenvectors associated with the eigenvalue λ_i, respectively. Accordingly, the transformation matrices M and M^{-1} can be written as $[\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n]$ and $[\mathbf{l}_1^\top, \mathbf{l}_2^\top, \dots, \mathbf{l}_n^\top]^\top$, respectively. The transfer function in the Laplace domain from l-th input to k-th output is

$$H(s) = \mathbf{c}_k M(sI - \Lambda)^{-1}M^{-1}\mathbf{b}_l = \sum_{i=1}^{n} \frac{\mathbf{c}_k\mathbf{r}_i\mathbf{l}_i\mathbf{b}_l}{s-\lambda_i}. \tag{14}$$

For simplicity, assume that the periodic injection u_l only contains one component with frequency ω_d and amplitude P_d, namely, $J = 1$, $\omega_1 = \omega_d$ and $P_1 = P_d$ in (2). Furthermore, we assume that before $t = 0^-$ the system is in steady state, viz., $\mathbf{x}(0^-) = \mathbf{0}$. Let sets $\mathcal{N}$ and $\mathcal{M}'$ consist of the indices of real eigenvalues, and the indices of complex eigenvalues with positive imaginary parts, respectively, viz.,

$$\mathcal{N} = \{i \in \mathbb{Z}^+ | \lambda_i \in \mathbb{R}\}; \quad \mathcal{M}' = \{i \in \mathbb{Z}^+ | \,\text{Im}(\lambda_i) > 0\}. \tag{15}$$

Then the Laplace transform for PMU measurement y_k is

$$\begin{aligned} Y_k(s) &= \left(\sum_{i=1}^{n} \frac{\mathbf{c}_k\mathbf{r}_i\mathbf{l}_i\mathbf{b}_l}{s-\lambda_i}\right)\frac{P_d\omega_d}{s^2+\omega_d^2} \\ &= \left[\sum_{i\in\mathcal{N}} \frac{\mathbf{c}_k\mathbf{r}_i\mathbf{l}_i\mathbf{b}_l}{s-\lambda_i} + \sum_{i\in\mathcal{M}'}\left(\frac{\mathbf{c}_k\mathbf{r}_i\mathbf{l}_i\mathbf{b}_l}{s-\lambda_i} + \frac{\mathbf{c}_k\bar{\mathbf{r}}_i\bar{\mathbf{l}}_i\mathbf{b}_l}{s-\bar{\lambda}_i}\right)\right]\frac{P_d\omega_d}{s^2+\omega_d^2} \end{aligned} \tag{16}$$

where $\bar{(\cdot)}$ denotes complex conjugation.

Next, we analyze the components resulting from the real eigenvalues and the components resulting from the complex eigenvalues, individually.

1) Components Resulting from Real Eigenvalues: In the Laplace domain, the component resulting from a real eigenvalue λ_i is

$$Y_{k,i}^{\text{D}}(s) = \frac{\mathbf{c}_k\mathbf{r}_i\mathbf{l}_i\mathbf{b}_l}{s-\lambda_i}\frac{P_d\omega_d}{s^2+\omega_d^2}. \tag{17}$$

The inverse Laplace transform of $Y_{k,i}^{\text{D}}(s)$ is

$$y_{k,i}^{\text{D}}(t) = \frac{\mathbf{c}_k\mathbf{r}_i\mathbf{l}_i\mathbf{b}_l P_d\omega_d}{\lambda_i^2+\omega_d^2}e^{\lambda_i t} + \frac{\mathbf{c}_k\mathbf{r}_i\mathbf{l}_i\mathbf{b}_l P_d}{\sqrt{\lambda_i^2+\omega_d^2}}\sin(\omega_d t + \phi_{i,l}) \tag{18}$$

where $\phi_{i,l} = \angle\left(\sqrt{\lambda_i^2+\omega_l^2} + j\lambda_i\right)$, and $\angle(\cdot)$ denotes the angle of a complex number.

2) Components Resulting from Complex Eigenvalues: In the Laplace domain, the component resulting from a complex eigenvalue $\lambda_i = -\sigma_i + j\omega_i$ is

$$Y_{k,i}^{\text{B}}(s) = \left(\frac{\mathbf{c}_k\mathbf{r}_i\mathbf{l}_i\mathbf{b}_l}{s-\lambda_i} + \frac{\mathbf{c}_k\bar{\mathbf{r}}_i\bar{\mathbf{l}}_i\mathbf{b}_l}{s-\bar{\lambda}_i}\right)\frac{P_d\omega_d}{s^2+\omega_d^2}. \tag{19}$$

The inverse Laplace transform of $Y_{k,i}^{\text{B}}(s)$ is

$$\begin{aligned} &y_{k,i}^{\text{B}}(t) = \\ &\frac{2P_d\omega_d|\mathbf{c}_k\mathbf{r}_i\mathbf{l}_i\mathbf{b}_l|}{\sqrt{(\sigma_i^2+\omega_d^2-\omega_i^2)^2+4\omega_i^2\sigma_i^2}}e^{-\sigma_i t}\cos(\omega_i t + \theta_{k,i} - \psi_i) + \\ &\frac{2P_d|\mathbf{c}_k\mathbf{r}_i\mathbf{l}_i\mathbf{b}_l|\sqrt{\omega_d^2\cos^2\theta_{k,i} + (\sigma_i\cos\theta_{k,i} - \omega_i\sin\theta_{k,i})^2}}{\sqrt{(\sigma_i^2-\omega_d^2+\omega_i^2)^2+4\omega_d^2\sigma_i^2}}\times \\ &\cos(\omega_d t + \phi_i - \alpha_i), \end{aligned} \tag{20}$$

where $\theta_{k,i} = \angle(\mathbf{c}_k\mathbf{r}_i\mathbf{l}_i\mathbf{b}_l)$; $\psi_i = \angle\left(\sigma_i^2+\omega_d^2-\omega_i^2 - j2\sigma_i\omega_i\right)$; $\phi_i = \angle(\sigma_i^2 - \omega_d^2 + \omega_i^2 - j2\omega_i\sigma_i)$, and $\alpha_i = \angle[\omega_d\cos\theta_{k,i} + j(\sigma_i\cos\theta_{k,i} - \omega_i\sin\theta_{k,i})]$.

3) Resonance Component: Under the resonance condition defined in (3), the injection frequency ω_d is in the vicinity of one natural modal frequency ω_{j^*}, and the real part of the natural mode is small. We define a new set $\mathcal{M} \subset \mathcal{M}'$ as $\mathcal{M} = \{i \in \mathbb{Z}^+ | \,\text{Im}(\lambda_i) > 0, |\omega_i - \omega_{j^*}| < \kappa_1, |\text{Re}(\lambda_i)| < \kappa_2\}$, where κ_1 and κ_2 are small and nonnegative real numbers. For $i \in \mathcal{M}$, the eigenvalue $\lambda_i = -\sigma_i + j\omega_i$ satisfies $\omega_i \approx \omega_d$ and $\sigma_i \approx 0$. Then $\psi_i \approx -\frac{\pi}{2}$, $\phi_i \approx -\frac{\pi}{2}$, and $\alpha_i \approx -\theta_{k,i}$. Therefore, equation (20) can be simplified as

$$y_{k,i}^{\text{B}}(t) \approx y_{k,i}^{\text{R}}(t) = \frac{P_d|\mathbf{c}_k\mathbf{r}_i\mathbf{l}_i\mathbf{b}_l|}{\sigma_i}(1-e^{-\sigma_i t})\sin(\omega_d t + \theta_{k,i}) \tag{21}$$

for $i \in \mathcal{M}$. In this paper, $y_{k,i}^{\text{R}}$ in (21) is termed the *resonance component* in the k-th measurement.

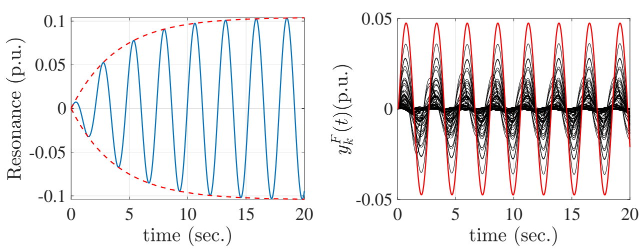
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2020.2982267, IEEE Transactions on Power Systems

In summary, a PMU measurement $y_k(t)$ in a power system (1) under resonance conditions can be decomposed into three classes of components, i.e.,

$$y_k(t) = \sum_{i\in\mathcal{N}} y_{k,i}^{\mathrm{D}}(t) + \sum_{i\notin\mathcal{M}\cup\mathcal{N}} y_{k,i}^{\mathrm{B}}(t) + \sum_{i\in\mathcal{M}} y_{k,i}^{\mathrm{R}}(t). \tag{22}$$

B. Observations on the Resonance Component and the Resonance-free Component

1) Severe Oscillations Arising from Resonance Component: Figure 3(a) visualizes the resonance component of a PMU measurement (at Bus 40[1]) in the IEEE 68-bus benchmark system. As can be observed from Figure 3(a), the upper envelop of the oscillation increases concavely at the initial stage before reaching a steady-stage value (about 0.1 in this case). The closed-form approximation for such a steady-state value is $P_d|\mathbf{c}_k\mathbf{r}_i\mathbf{l}_i\mathbf{b}_l|/\sigma_i$. For a small positive σ_{j^*} associated with eigenvalue λ_{j^*}, the steady-state amplitude of the resonance component may be the dominant one. If a PMU measurement far away from the source measurements is tightly coupled with the eigenvalue λ_{j^*}, it may manifest the most severe oscillation, thereby confusing system operators with regard to FO source localization. Therefore, the presence of resonance components may cause the counter-intuitive cases defined by (7), (8).



(a) (b)

Fig. 3. (a) Visualization of the resonance component of bus voltage magnitudes in the IEEE 68-bus benchmark system based on equation (21): the resonance components of the voltage magnitude measurement at Bus 40 (blue curve) and its envelopes (red-dash curves). (b) Resonance-free components of the source voltage magnitude measurement (red) and the non-source voltage magnitude measurement (black) in the IEEE 68-bus benchmark system.

2) Location Information on FO Source from the Resonance-free Component: As the resonance components of the set of all PMU measurements mislead system operators with respect to FO localization, we proceed by excluding the resonance component from (22), and checking whether if the remaining components exhibit any spatial information concerning the FO source. The superposition of the remaining components is termed as *resonance-free*. Specifically, for a power system with known physical model (1), the resonance-free component y_k^{F} in the k-th PMU measurement time series can be obtained by:

$$y_k^{\mathrm{F}}(t) = \sum_{i\in\mathcal{N}} y_{k,i}^{\mathrm{D}}(t) + \sum_{i\notin\mathcal{M}\cup\mathcal{N}} y_{k,i}^{\mathrm{B}}(t). \tag{23}$$

The visualization of the resonance-free component for all PMU measurements in the IEEE 68-bus system is shown in Figure 3(b) under a certain FO scenario[2]. Under the same FO scenario, Figure 1 visualizes all PMU measurements $y_k(t)$ in (22). In Figure 3(b), while the complete measurements $y_k(t)$ are counter-intuitive, the resonance-free components $y_k^{\mathrm{F}}(t)$ convey the location information on the FO source– the resonance-free component of the source measurement exhibits the largest oscillation. Such localized response of resonance-free components might be an extension of the no-gain property of an electric network rigorously justified in [19], [20]. Future work will examine what kinds of power systems possess localization property of resonance-free components in a theoretically rigorous fashion.

C. Low-rank Nature of Resonance Component Matrix

The physical interpretation of the efficacy of the RPCA-based algorithm is illustrated by examining the rank of the matrix containing all resonance components for all measurements, which we call the *resonance component matrix*, formally defined next. Similar to (4), the resonance component $y_k^{\mathrm{R}}(t)$ in the k-th measurement can be discretized into a row vector $\mathbf{y}_{k,t}^{\mathrm{R}}$:

$$\mathbf{y}_{k,t}^{\mathrm{R}} := \left[y_k^{\mathrm{R}}(0), \quad y_k^{\mathrm{R}}(1/f_{\mathrm{s}}), \quad \dots \quad y_k^{\mathrm{R}}(\lfloor tf_{\mathrm{s}} \rfloor / f_{\mathrm{s}}) \right]. \tag{24}$$

Then, the resonance component matrix Y_t^{R} can be defined as a row concatenation as follows:

$$Y_t^{\mathrm{R}} := \left[\left(\mathbf{y}_{1,t}^{\mathrm{R}}\right)^\top, \quad \left(\mathbf{y}_{2,t}^{\mathrm{R}}\right)^\top, \quad \dots \quad \left(\mathbf{y}_{m,t}^{\mathrm{R}}\right)^\top \right]^\top. \tag{25}$$

Theorem 1. *For the linear time-invariant dynamical system* (1)*, the rank of the resonance component matrix Y_t^R defined in* (25) *is at most* 2*.*

Proof. Based on (21), define $E_k := P_d|\mathbf{c}_k\mathbf{r}_i\mathbf{l}_i\mathbf{b}_l|/\sigma_i$. Then

$$\begin{aligned} y_{k,i}^{\mathrm{R}}(t) = & (1-e^{-\sigma_i t})\sin(\omega_d t) E_k \cos(\theta_{k,i}) + \\ & (1-e^{-\sigma_i t})\cos(\omega_d t) E_k \sin(\theta_{k,i}). \end{aligned}$$

We further define functions $f_1(t)$, $f_2(t)$ and variables $g_1(k)$, $g_2(k)$ as follows: $f_1(t) := (1-e^{-\sigma_i t})\sin(\omega_d t)$; $f_2(t) := (1-e^{-\sigma_i t})\cos(\omega_d t)$; $g_1(k) := E_k\cos(\theta_{k,i})$; and $g_2(k) := E_k\sin(\theta_{k,i})$. Then, $y_{k,i}^{\mathrm{R}}(t)$ can be represented by $y_{k,i}^{\mathrm{R}}(t) = f_1(t)g_1(k) + f_2(t)g_2(k)$.

The resonance component matrix Y_t^{R} up to time t can be factored as follows:

$$Y_t^{\mathrm{R}} = \begin{bmatrix} g_1(1) & g_2(1) \\ g_1(2) & g_2(2) \\ \vdots & \vdots \\ g_1(m) & g_2(m) \end{bmatrix} \begin{bmatrix} f_1(0) & f_1(\frac{1}{f_{\mathrm{s}}}) & \dots & f_1(\frac{\lfloor tf_{\mathrm{s}}\rfloor}{f_{\mathrm{s}}}) \\ f_2(0) & f_2(\frac{1}{f_{\mathrm{s}}}) & \dots & f_2(\frac{\lfloor tf_{\mathrm{s}}\rfloor}{f_{\mathrm{s}}}) \end{bmatrix}. \tag{26}$$

Denote by vectors $\mathbf{g}_1$ and $\mathbf{g}_2$ the first and second columns of the first matrix in the right hand side (RHS) of (26), respectively; and by vectors $\mathbf{f}_1$ and $\mathbf{f}_2$ the first and second rows

[1]The measurements at Bus 40 exhibit the largest oscillations but they are non-source measurements.

[2]A sinusoidal waveform with amplitude 0.05 per unit (p.u.) and frequency 0.38 Hz is injected into the IEEE 68-bus system via the voltage setpoint of generator 13. The information on the test system is elaborated in Section V.

of the second matrix in the RHS of (26). Then (26) turns to be

$$Y_t^{\mathrm{R}} = \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 \end{bmatrix} \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{bmatrix}. \tag{27}$$

Given (27), it is clear that the rank of the resonance component matrix Y_t^{R} is at most 2. □

Typically, for a resonance component matrix Y_t^{R} with m rows and $\lfloor tf_{\mathrm{s}} \rfloor$ columns, owing to $\min(m, \lfloor tf_{\mathrm{s}} \rfloor) \gg 2$, the resonance component matrix Y_t^{R} is a low-rank matrix, which is assumed to be integrated by the low-rank component L_t in equation (9). As discussed in Section IV-B2, the source measurement can be tracked by finding the maximal absolute entry of the resonance-free matrix $(Y_t - Y_t^{\mathrm{R}})$. According to (10), the PMU measurement containing the largest absolute entry in the sparse component S_t is considered as the source measurement. Then, it is reasonable to conjecture that the sparse component S_t in (9) captures the part of the resonance-free matrix that preserves the location information of FO source. Thereby, a theoretical connection between the proposed data-driven method in Algorithm 1 and the physical model of power systems described in equation (1) can be established. Although forced oscillation phenomena have been extensively studied in physics [21], the low-rank property, to the best of our knowledge, is first investigated in this paper.

Through the FO case shown in Figure 1, we next examine the entries corresponding to the largest amplitude channel (Bus 40) and the source measurement (Bus 65) in the measurement matrix Y_t, the low-rank matrix L_t, and the sparse matrix S_t. In Figure 4(a), the blue-dash curve and the red curve respectively present voltage magnitudes at the largest amplitude channel (Bus 40) and the source measurement (Bus 65). Figure 4(b) shows the components captured by the low-rank matrix L_t corresponding to measurements at Bus 40 (blue-dash) and Bus 65 (red). Figure 4(c) shows the components captured by the sparse matrix S_t corresponding to measurements at Bus 40 (blue-dash) and Bus 65 (red). As can be observed in Figure 4(a), the measurement at Bus 40 (blue-dash curve) comprises mainly the resonance component. As we have established in Theorem 1, the resonance component matrix is by nature low-rank. Therefore, the measurement at Bus 40 is better captured by the low-rank matrix than the measurement at Bus 65, as is shown in Figure 4(b). What is left in the sparse matrix pinpoints the forced oscillation source. Besides, in Figure 4, part of resonance-free component is also captured by the low-rank matrix, which cannot be explained by Theorem 1. Note that Theorem 1 offers one possible interpretation of the effectiveness of the proposed algorithm, but it is not claimed to be a fully rigorous interpretation of why the algorithm works, however as is verified by the above figure it indeed sheds a lot of light in its interpretation. As this paper focuses on the development of one possible data-driven localization algorithm, future work will investigate a broader category of possible algorithms and their theoretical underpinnings.

A natural question is if the robust-PCA procedure can pinpoint the source of other types of oscillations, such as natural oscillations. The difficulty to answering this question is that "source of natural oscillation" is not well defined. In a forced oscillation event, the FO source is defined as the power system component with external periodic perturbations, and one obvious solution to suppressing the oscillation is to disconnect the source from the grid. In a natural oscillation event, one may suppress it by tuning control apparatus of a set of generators or by decreasing the load level. In such a case, should the source be deemed the tuned generators or the decreased load? In brief, we believe it is challenging to consent on a definition of the "source" of natural oscillations. Due to the ambiguity in the definition of natural oscillation sources, this paper only focuses on the localization of forced oscillations.

V. Case Study

In this section, we validate the effectiveness of Algorithm 1 using data from IEEE 68-bus benchmark system and WECC 179-bus system. We first describe the key information on the test systems, the procedure for obtaining test data, the parameter settings of the proposed algorithm, and the algorithm performance over the obtained test data. Then the impact of different factors on the performance of the localization algorithm is investigated. Finally, we compare the proposed algorithm with the energy-based method reported in [2]. As will be seen, the proposed method can pinpoint the FO sources with high accuracy *without any information on system models and grid topology*, even when resonance exists.

A. Performance Evaluation of the Localization Algorithms in Benchmark Systems

1) IEEE 68-bus Power System Test Case: The system parameters of the IEEE 68-bus power system are reported in the Power System Toolbox (PST) [15] and its topology is shown in Figure 5. Let $\mathcal{V} = \{1, 2, \ldots, 16\}$ consist of the indices of all 16 generators in the 68-bus system. Based on the original parameters, the following modifications are made: 1) the power system stabilizers (PSS) at all generators, except the one at Generator 9, are removed, in order to create more poorly-damped oscillatory modes; 2) for the PSS at Generator 9, the product of PSS gain and washout time constant is changed to 250. Based on the modified system, the linearized model of the power system (1) can be obtained using the command "`svm_mgen`" in PST. There are 25 oscillatory modes whose frequencies range from 0.1 Hz to 2 Hz, which are shown in Figure 8(a). Denote by $\mathcal{W} = \{\omega_1, \omega_2, \ldots, \omega_{25}\}$ the set consisting all 25 modal frequencies of interest. The periodic perturbation u_l in (2) is introduced through the voltage setpoints of generators. The analytical expression of u_l is $0.05 \sin(\omega_d t)$, where $\omega_d \in \mathcal{W}$.

We create FOs in the 68-bus system according to set $\mathcal{V} \times \mathcal{W}$, where $\times$ is the Cartesian product. For element $(i, \omega_j) \in \mathcal{V} \times \mathcal{W}$, the periodic perturbation $u_l(t)$ with frequency ω_j is injected into the grid through the voltage setpoint of generator i at time $t = 0$. Then, the system response is obtained by conducting a 40-second simulation. The bus voltage magnitude deviations constitute the output/measurement vector $\mathbf{y}(t)$ in (1). Finally, the measurement matrix is constructed based on (4), where the sampling rate f_{s} is 60 Hz. By repeating the above procedure

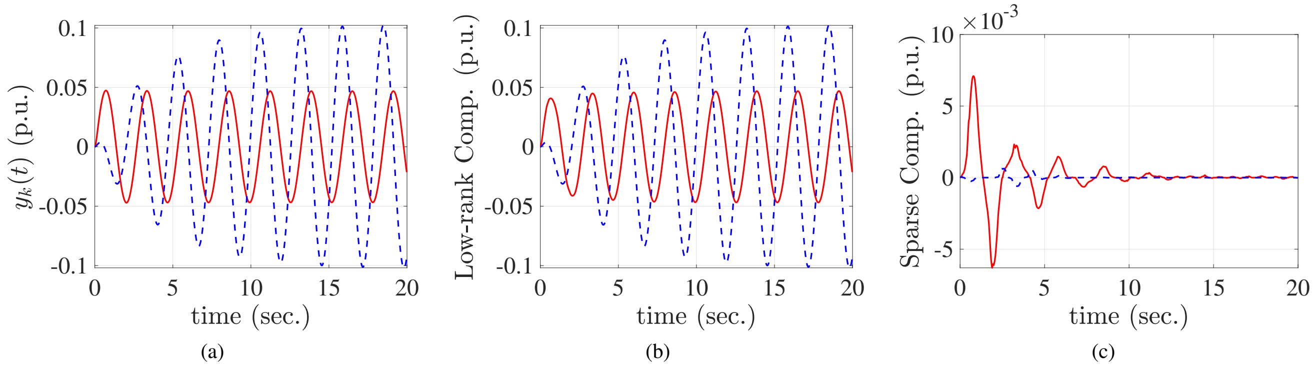


Fig. 4. Visualization of voltage magnitudes (a), components in low-rank matrix L_t (b) and components in sparse matrix S_t (c) at Bus 65 (red) and Bus 40 (blue dash): Bus 65 is the bus closest to the source, while the most severe oscillation appear at Bus 40.

for each element in set $\mathcal{V} \times \mathcal{W}$, we obtain 400 measurement matrices ($|\mathcal{V} \times \mathcal{W}|$). Among the 400 measurement matrices, 44 measurement matrices satisfy the resonance criteria (7), (8) with $N_0 = 0$ and they are marked as the counter-intuitive cases which are used for testing the performance of the proposed method. Some typical waveforms in the 44 test cases are shown in [6].

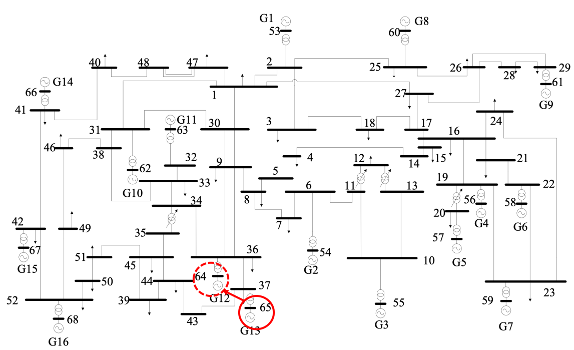


Fig. 5. The IEEE 68-bus power system [6]: the generator in the solid circle is the actual source generator; the generator in the dash circle is the identified source.

The tunable parameters T_0 and ξ in Algorithm 1 are set to 10 and 0.0408, respectively. Measurements of voltage magnitude, phase angle and frequency are used for constituting the measurement matrix. Then, we apply Algorithm 1 to the 44 counter-intuitive cases. Algorithm 1 pinpoints the source measurements in 43 counter-intuitive cases and, therefore, achieves 97.73% accuracy without any knowledge of system models and grid topology.

Next, we scrutinize the geographic proximity between the identified and actual source measurements in the single failed case. The algorithm outputs that the source measurement is located at Bus 64 (highlighted with a solid circle in Figure 5), when a periodic perturbation with frequency 1.3423 Hz is injected into the system through the generator directly connecting to Bus 65 (highlighted with a dash circle in Figure 5). As can be seen in Figure 5, the identified and actual source measurements are geographically close. Therefore, even in the failed cases, the proposed method can effectively narrow the search space.

2) WECC 179-bus System Test Case: This subsection leverages the open-source forced oscillation dataset [22] to validate the performance of the RPCA-based method. The offered dataset is generated via the WECC 179-bus power system [22] whose topology is shown in Figure 7(a). The procedure for synthesizing the data is reported in [22]. The available dataset includes 15 forced oscillation cases with single oscillation source, which are used to test the proposed method. The visualization for Case `F-3` is shown in Figure 6. In each forced oscillation case, the measurements of voltage magnitude, voltage angle and frequency at all generation buses are used to construct the measurement matrix Y_t in (4), from the 10-second oscillatory data, i.e., $T_0 = 10$. Then, the 15 measurement matrices are taken as the input for Algorithm 1, where the tunable parameter ξ is set to 0.0577.

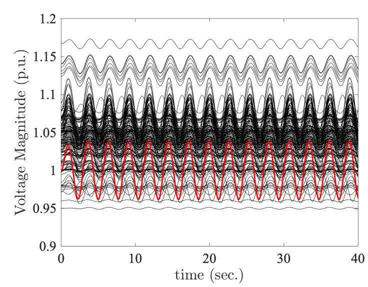
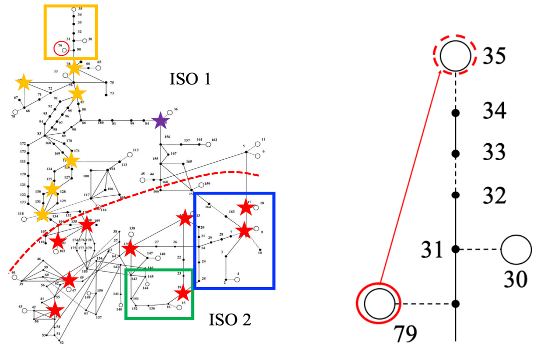


Fig. 6. Voltage magnitude visualization in Case `F-3`: the voltage magnitude of the bus connected with the forced oscillation source (red); the voltage magnitudes of the remaining buses (black).

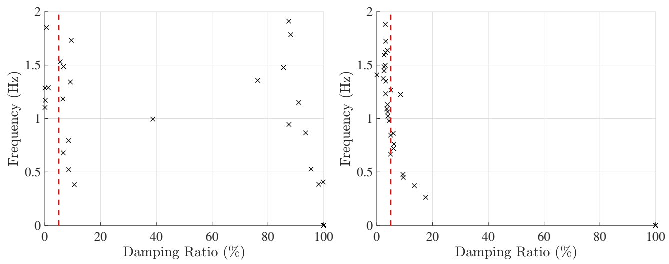
For the WECC 179-bus system, the proposed method achieved 93.33% accuracy. Next, we present how geographically close the identified FO sources are to the ground truth in the seemingly incorrect case. In Case `FM-6-2`, a periodic rectangular perturbation is injected into the grid through the governor of the generator at Bus 79 which is highlighted with a red solid circles in Figure 7(b). The source measurement identified by the proposed method is at Bus 35 which is highlighted by a red dash circle. As can be seen in Figure 7(b), the identified FO source is geographically close to the actual source. Again, even the seemingly wrong result can help system operators substantially narrow down the search space for FO sources.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2020.2982267, IEEE Transactions on Power Systems



(a) (b)

Fig. 7. WECC 179-bus power system [22]: (a) complete topology; (b) zoomed-in version of the area in the yellow box in the left figure.



(a) (b)

Fig. 8. Eigenvalues of the IEEE 68-bus system (a) and the WECC 179-bus system in Cases `F-1` and `FM-1` (b): the eigenvalues whose damping ratio less than 5% are in the left-hand side of the red-dash line.

3) ERCOT Forced Oscillation Event: We leverage the field measurements from a collaborative project with Electric Reliability Council of Texas (ERCOT), in order to test the localization algorithm in a realistic setting. Figure 9 shows the FOs observed by ERCOT. The FOs manifested themselves in seven PMU measurements on voltage magnitudes. For information privacy, the names of the PMU locations are replaced by indices $1, 2, \ldots, 7$, and the FO starting point is set to 0 seconds. In Figure 9, it can be observed that the PMU measurements contain high frequency components resulting from measurement noise and load fluctuation. We apply a band-pass filter from 0.1 Hz to 1 Hz to process the raw PMU measurements. Subsequently, we use a 10-second time window of the filtered data for forming the measurement matrix. Finally, the proposed algorithm indicates that PMU 4 is the one near the FO source. The localization result was reported to ERCOT, and ERCOT confirmed the correctness of the result. It is worth noting that no topology information was provided to our research team. Therefore, localization algorithms based on system topology, such as the Dissipating Energy Flow approach, are not applicable in this study.

B. Algorithm Robustness

The subsection focuses on testing the robustness of the proposed algorithm under different factors which include measurement types, noise, and partial coverage of PMUs. The impact of each factor on the algorithm performance will be demonstrated as follows.

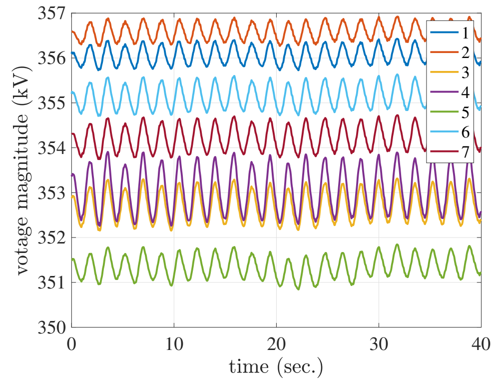


Fig. 9. Voltage Magnitudes during the ERCOT forced oscillation event.

TABLE I
IMPACT OF MEASUREMENT TYPES ON LOCALIZATION PERFORMANCE

Types	$\|V\|$	$\angle V$	$\|V\|, \angle V$	f
68-bus System	84.09%	50.00%	84.09%	52.27%
179-bus System	86.67%	33.33%	73.33%	20.00%
Types	$\|V\|, f$	$\angle V, f$	$\|V\|, \angle V, f$	N/A
68-bus System	93.18%	59.09%	**97.73%**	N/A
179-bus System	80.00%	46.67%	**93.33%**	N/A

1) Impact of Measurement Types on Algorithm Performance: Under all possible combinations of nodal measurements (voltage magnitude $|V|$, voltage angle $\angle V$ and frequency f), the localization accuracies of the proposed algorithm in the two benchmark systems are reported in Table I. As can be observed in Table I, the maximal accuracy is achieved when voltage magnitudes, voltage angles and frequencies are used to constitute the measurement matrix in (4).

2) Impact of Noise on Algorithm Performance: Table II records the localization accuracy under different levels of noise. In Table II, the signal-to-noise ratio (SNR) is defined as follows:

$$SNR = 10 \log(W_{\mathrm{s}}/W_{\mathrm{n}}) \quad \text{(dB)}$$

where W_{s} is the sum of squared measurement *deviations* over a period (10 seconds in this paper); and W_{n} is the sum of squared magnitudes of the corresponding noise over the same period. The noise superimposed upon each measurement has a Gaussian distribution with zero mean and variance σ_{n}. At each experiment for each measurement, the variance σ_{n} is chosen such that the corresponding SNR is achieved. From Table II, we conclude the proposed algorithm performs well under the cases with SNR less than 30 dB.

3) Impact of Partial Coverage of Synchrophasors on Algorithm Performance: In practice, not all buses are equipped with PMUs. Besides, available PMUs may be installed on buses near oscillation sources, instead of buses on which

TABLE II
IMPACT OF NOISE LEVEL ON LOCALIZATION PERFORMANCE

SNR	90dB	70dB	50dB	30dB	10dB
68-Bus	97.73%	97.73%	97.73%	97.73%	56.82%
179-Bus	93.33%	93.33%	93.33%	93.33%	73.33%

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2020.2982267, IEEE Transactions on Power Systems

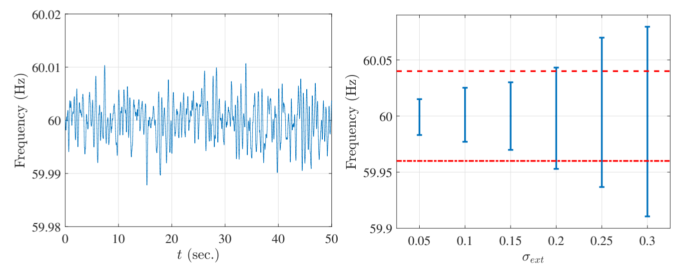
oscillation sources are directly connected. A test case is designed for testing the performance of the proposed algorithm in the scenario described above. In this test case, the locations of all available PMUs are marked with stars in Figure 7(a). The test result is listed in Table III. As illustrated in Table III, the proposed method can effectively identify the available PMUs that are close to oscillation sources, even though no PMU is installed on generation buses.

Independent System Operators (ISOs) may also need to know whether FO sources are within their control areas. However, ISOs might not be able to access PMUs near FO sources, limiting the usefulness of the proposed algorithm. For example, assume that there are two ISOs, i.e., ISO 1 and ISO 2, in Figure 7(a), where the red dash line is the boundary between the control areas of the two ISOs. It is possible that FO sources are at the ISO 1 control area, whereas ISO 2 only can access the PMUs at the buses marked with red stars. In order to apply the RPCA-based method, ISO 2 needs to access one PMU in the area controlled by ISO 1, say, the PMU marked with a purple star in Figure 7(a). In the `F-2` dataset, the FO source is located at Bus 79 which is marked with a red circle in Figure 7(a). With the data collected from PMUs marked with red and purple stars, the proposed algorithm outputs the bus marked with a purple star, indicating that the FO source is outside the control area of ISO 2.

4) Impact of External Excitation on Localization Performance: The external excitation is assumed to result mainly from load fluctuation. In order to introduce load fluctuation, load dynamics are included in the 68-bus benchmark system, and 33 real power setpoints along with 33 reactive power setpoints on load are considered as the augmented inputs. The above modification on the 68-bus system can be achieved by enabling load modulation in the Power System Toolbox (PST) [15]. Following the procedure described in Section V-A-1, 43 counter-intuitive cases are obtained. For the j-th case of the 43 counter-intuitive cases, we have a pair of numbers (i'_j, ω'_j), where ω'_j is the frequency of a periodic perturbation and i'_j is the source generator index. Let set $\mathcal{P}$ consist of such pairs, i.e., $\mathcal{P} = \{(i'_1, \omega'_1), (i'_2, \omega'_2), \ldots, (i'_j, \omega'_j), \ldots, (i'_{43}, \omega'_{43})\}$.

Note that the number of state variables in the 68-bus system with load dynamics is 268, whereas the number of state variables in the 68-bus system used in Section V-A is 202. Effectively, the 68-bus system in this subsection is a different system from the 68-bus system used in Section V-A, from the perspective of control theory, as the numbers of their state variables are distinct. Therefore, it is not surprising that the number of counter-intuitive cases in this subsection is different from that in Section V-A.

The 66 augmented setpoints fluctuate around their nominal values, which can be considered to be external excitations. Denote by $\Delta\mathbf{u}_{\text{Ld}}(t) \in \mathbb{R}^{66}$ the load setpoint deviations from their nominal values at time t. Assume that vector $\Delta\mathbf{u}_{\text{Ld}}$ has a Gaussian distribution with zero mean and covariance matrix $\sigma_{\text{ext}} I_{66}$, i.e., $\Delta\mathbf{u}_{\text{Ld}}(t) \sim \mathcal{N}(0, \sigma_{\text{ext}} I_{66})$, where σ_{ext} is a scalar, and I_{66} is a 66 by 66 identity matrix. Due to the excitation $\Delta\mathbf{u}_{\text{Ld}}$, the frequency fluctuates under normal operating condition as observed in Figure 10(a). Figure 10(b) shows how the system frequency range varies as scalar σ_{ext} changes. In Figure 10(b), each vertical line segment corresponds to the frequency range under a load fluctuation with parameter σ_{ext}: the upper terminal is the highest system frequency for each given σ_{ext}; and the lower terminal is the lowest system frequency with corresponding σ_{ext}. One observation from Figure 10(b) is that as scalar σ_{ext} increases, it is more likely that the system frequencies lie in a wider range. The normal range of frequency in power systems is from 59.96 Hz to 60.04 Hz [23], [24]. As shown in Figure 10(b), the range of system frequencies are out of the normal range under the excitation with $\sigma_{\text{ext}} = 0.2$. We use the random excitations $\Delta\mathbf{u}_{\text{Ld}}(t)$ with $\sigma_{\text{ext}} = 0.15$ to mimic real-world load fluctuation.



(a) (b)

Fig. 10. (a) Frequency at Bus 1 under normal operation condition with load fluctuation; (b) Ranges of system frequency (vertical blue-solid line segments) due to different levels of load fluctuation: the normal frequency range (59.96-60.04 Hz) is represented by two horizontal red-dash lines.

The random excitations $\Delta\mathbf{u}_{\text{Ld}}$ with $\sigma_{\text{ext}} = 0.15$ and set $\mathcal{P}$ are leveraged to obtain 43 test cases. The data acquisition procedure is described in what follows. For element $(i'_j, \omega'_j) \in \mathcal{P}$, the periodic perturbation $u_l(t)$ with frequency ω'_j is injected into the system via the voltage setpoint of generator i'_j at $t = 0$. At each experiment, the 68-bus system is excited by one realization of $\Delta\mathbf{u}_{\text{Ld}}$. Then, a 40-second simulation is conducted in order to obtain the system response. By repeating the above procedure for all elements, 43 test cases with load fluctuation are obtained. For these test cases, a 2-Hz low-pass filter is applied to process the measurements. The proposed algorithm achieves 90.70% localization accuracy.

5) Impact of Time-window Length on Localization Performance: In this section, we investigate the impact of the window width T_0 on the algorithm's performance. Fig. 11 summarizes the localization accuracy with different time-window widths T_0 in both the 68-bus and 179-bus systems. In Fig. 11, we observe a trade-off between the time required for decision making and the localization accuracy for the 68-bus system (the blue-dash line) with the given range of T_0: 100% accuracy can be achieved with $T_0 = 12$ (or 13) seconds; the price we pay for the high localization accuracy is a wider time window, i.e., more decision-making time.

In practice, the optimal window width T_0^* can be obtained by off-line studies on physical model-based simulations or historical FO events. Assume that we have N_1 options for the window width T_0, represented by $\mathcal{T}_0 := \{T_{01}, T_{02}, \ldots, T_{0N_1}\}$. For each window width option, say, T_{0i}, we run the localization algorithm on all available FO events and compute the localization accuracy η_i. The optimal window width T_0^* is the i^*-th element in the set $\mathcal{T}_0$, which maximizes η_i for

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2020.2982267, IEEE Transactions on Power Systems

TABLE III
IMPACT OF PARTIAL COVERAGE OF SYNCHROPHASOR ON ALGORITHM PERFORMANCE

Case Name	F-1	FM-1	F-2	F-3	FM-3	F-4-1	F-4-2	F-4-3	F-5-1	F-5-2	F-5-3	F-6-1	F-6-2	F-6-3	FM-6-2
Identified Source	8	8	78	69	69	69	78	78	78	78	78	78	78	78	78
Nearest PMU	8	8	78/69	78/69	78/69	78/69	78/69	78/69	78/69	78/69	78/69	78/69	78/69	78/69	78/69

$i = 1, 2, \ldots, N_1$. Such an optimal window width T_0^* is applied in the localization algorithm 1 during real-time operation.

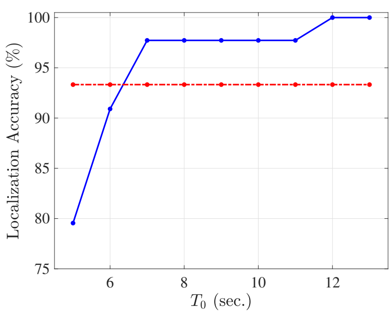


Fig. 11. Impact of T_0 on Localization Performance: the localization accuracy for the 68-bus system (blue-solid line) and the 179-bus system (red-dash line).

C. Comparison with Energy-based Localization Method

This subsection aims to compare the proposed localization approach with the Dissipating Energy Flow (DEF) approach [2]. We use the `FM-1` dataset (Bus 4 is the source measurement) [22] for the purpose of comparing DEF method with the proposed algorithm. PMUs are assumed to be installed at all generator buses except ones at Buses 4 and 15. Besides, Buses 7, 15 and 19 are also assumed to have PMUs. *Without any information on grid topology*, the RPCA-based method suggests the source measurement is at Bus 7 which is in the vicinity of the actual source. However, topology errors may cause DEF-based method to incur both false negative and false positive errors, as will be shown in the following two scenarios.

1) Scenarios 1: The zoomed-in version of the area within the blue box in Figure 7(a) is shown in Figure 12, where the left and right figures are the actual system topology and the topology reported to a control center, respectively. All available PMUs are marked with yellow stars in Figure 12. Based on these available PMUs, the relative magnitudes and directions of dissipating energy flows are computed according to the `FM-1` dataset and the method reported in [2]. With the true topology, the FO source cannot be determined, as the energy flow direction along Branch 8-3 cannot be inferred based on the available PMUs. However, with the topology error shown in Figure 12(b), i.e., it is mistakenly reported that Bus 29 (Bus 17) is connected to Bus 3 (Bus 9), it can be inferred that the energy flow with relative magnitude of 0.4874 is injected into the Bus 4, indicating that Bus 4 is *not* the source measurement. Such a conclusion contradicts the ground truth. Therefore, with such a topology error, the dissipating energy flow method leads to a false negative error.

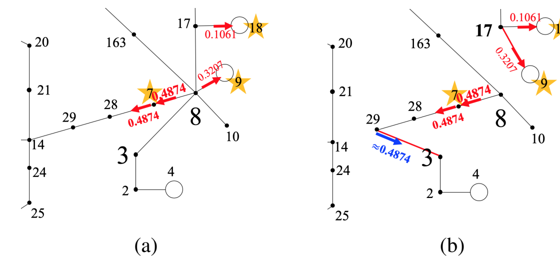


Fig. 12. Zoomed-in version of the area in the blue box at Figure 7 (a): actual topology (left); topology reported in a control center (right). Relative magnitudes and direction of energy flows are labeled with red numbers and arrows, respectively.

2) Scenario 2: Similar to Scenario 1, topology errors exist within the area highlighted by a green box in Figure 7(a), whose zoomed-in version is shown in Figure 13. As shown in Figure 13(a), it can be inferred that an energy flow with relative magnitude of 0.171 injects into Bus 15 with the information of actual topology and available PMUs, indicating Bus 15 is not a source. However, with the reported system topology, the generator at Bus 15 injects to the rest of grid an energy flow with magnitude of 0.0576, suggesting the source measurement is at Bus 15. Again, such a conclusion contradicts with the ground truth and, hence, incurs a false positive error.

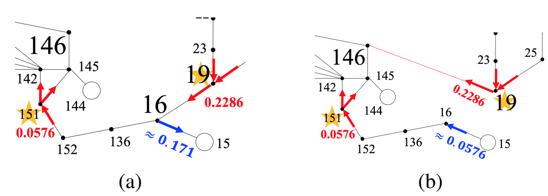


Fig. 13. Zoomed-in version of the area in the green box at Figure 7 (a): actual topology (left); topology reported in a control center (right).

VI. CONCLUSIONS

In this paper, a purely data-driven but physically interpretable method is proposed in order to locate forced oscillation sources in power systems. The localization problem is formulated as an instance of matrix decomposition, i.e., how to decompose the high-dimensional synchrophasor data into a low-rank matrix and a sparse matrix, which can be done using Robust Principal Component Analysis. Based on this problem formulation, a localization algorithm for real-time operation is presented. The proposed algorithm does not require any information on system models nor grid topology, thus providing an efficient and easily deployable solution for real-time operation. Without the availability of system topology, the proposed algorithm can achieve high localization accuracy in synthetic cases based on benchmark systems and real-world forced oscillation in the power grid of Texas. In addition, a possible theoretical interpretation of the efficacy of the algorithm is provided based on physical model-based analysis, highlighting the fact that the rank of the resonance component matrix is at most 2. Future work will test the proposed localization algorithm in conjunction with FO detection algorithms, and explore a broader set of algorithms and their theoretical performance analysis for large-scale realistic power systems.

References

[1] M. Ghorbaniparvar and N. Zhou, "A survey on forced oscillations in power system," *CoRR*, vol. abs/1612.04718, 2016. [Online]. Available: http://arxiv.org/abs/1612.04718

[2] E. L. S Maslennikov, B Wang, "Dissipating energy flow method for locating the source of sustained oscillations," *International Journal of Electrical Power & Energy Systems*, pp. 55–62, 2017.

[3] S. A. N. Sarmadi, V. Venkatasubramanian, and A. Salazar, "Analysis of november 29, 2005 western american oscillation event," *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 5210–5211, Nov 2016.

[4] S. Maslennikov, "Detection the source of forced oscillations," Tech. Rep. [Online]. Available: https://www.naspi.org/node/653

[5] S. A. N. Sarmadi and V. Venkatasubramanian, "Inter-area resonance in power systems from forced oscillations," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 378–386, 2016.

[6] T. Huang, N. M. Freris, P. R. Kumar, and L. Xie, "Localization of forced oscillations in the power grid under resonance conditions," in *2018 52nd Annual Conference on Information Sciences and Systems (CISS)*, March 2018, pp. 1–5.

[7] N. Zhou, M. Ghorbaniparvar, and S. Akhlaghi, "Locating sources of forced oscillations using transfer functions," in *2017 IEEE Power and Energy Conference at Illinois (PECI)*, Feb 2017, pp. 1–8.

[8] W. Bin and S. Kai, "Location methods of oscillation sources in power systems: a survey," *Journal of Modern Power Systems and Clean Energy*, vol. 5, no. 2, pp. 151–159, 2017.

[9] S. C. Chevalier, P. Vorobev, and K. Turitsyn, "Using effective generator impedance for forced oscillation source location," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6264–6277, Nov 2018.

[10] S. Chevalier, P. Vorobev, and K. Turitsyn, "A bayesian approach to forced oscillation source location given uncertain generator parameters," *IEEE Transactions on Power Systems*, pp. 1–1, 2018.

[11] "Var-501-wecc-3power system stabilizer," Tech. Rep. [Online]. Available: https://www.wecc.biz/Reliability/VAR-501-WECC-3.pdf

[12] 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.

[13] S. H. AV Oppenheim, Alan S. Willsky, *Signal and Systems*. Prentice Hall, 1997.

[14] H. Ye, Y. Liu, P. Zhang, and Z. Du, "Analysis and detection of forced oscillation in power system," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1149–1160, March 2017.

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

[16] Z. Lin, M. Chen, and Y. Ma, "The Augmented Lagrange Multiplier method for exact recovery of corrupted low-rank matrices," *arXiv preprint arXiv:1009.5055*, 2010. [Online]. Available: https://arxiv.org/abs/1009.5055

[17] J. Follum and J. W. Pierre, "Detection of periodic forced oscillations in power systems," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 2423–2433, May 2016.

[18] J. Follum and J. Pierre, "Time-localization of forced oscillations in power systems," in *2015 IEEE Power Energy Society General Meeting*, July 2015, pp. 1–5.

[19] M. Ilic-Spong, M. Spong, and R. Fischl, "The no-gain theorem and localized response for the decoupled p -hetapower network with active power losses included," *IEEE Transactions on Circuits and Systems*, vol. 32, no. 2, pp. 170–177, February 1985.

[20] M. Ilic-Spong, J. Thorp, and M. Spong, "Localized response performance of the decoupled q -v network," *IEEE Transactions on Circuits and Systems*, vol. 33, no. 3, pp. 316–322, March 1986.

[21] N. Wiener, *Cybernetics or Control and Communication in the Animal and the Machine*. MIT press, 1965, vol. 25.

[22] S. Maslennikov, B. Wang, Q. Zhang, a. Ma, a. Luo, a. Sun, and E. Litvinov, "A test cases library for methods locating the sources of sustained oscillations," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, July 2016, pp. 1–5.

[23] T. Huang, B. Satchidanandan, P. R. Kumar, and L. Xie, "An online detection framework for cyber attacks on automatic generation control," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6816–6827, Nov 2018.

[24] EPRI, *EPRI Power System Dynamic Tutorial*, Electric Power Research Institution, July 2009.

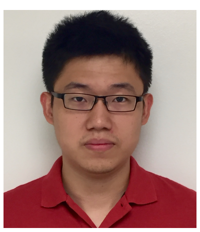
Appendix A
Notation

A	System matrix.
B	Input matrix.
$\mathcal{B}$	Set of buses.
$\mathbf{b}_l$	The l-th column of matrix B.
$\mathbf{c}_k$	The k-th row of matrix C.
D	Feed-forward matrix.
f_{s}	Sampling rate of synchrophasor.
i_{s}	Bus index of the source measurement.
J	Number of frequency components in u_l.
K	Number of measurement types.
λ_i	The i-th eigenvalue of A.
$\mathcal{L}$	A set collecting all eigenvalues of A.
L_t	Low-rank matrix.
$\mathbf{l}_i$	The i-th left eigenvector of A associated with eigenvalue λ_i.
m	Number of outputs.
M	Matrix chosen so that $M^{-1}AM$ is diagonal.
$\mathcal{M}$	Set of the indices of the eigenvalues of A with frequency near the injecting frequency ω_d.
n	Number of state variables.
$\mathcal{N}$	Set collecting the indexes of complex eigenvalues of A with positive imaginary part.
ξ	Regularization parameter in RPCA formulation.
ω_j	Frequency of j-th frequency component of u_l.
ω_d	Frequency of the perturbation (one frequency component).
P_d	Amplitude of the perturbation (one frequency component).
r	Number of inputs.
$\mathbf{r}_i$	Right eigenvector of A associated with eigenvalue λ_i.
S_t	Sparse matrix.
s	Variable of Laplace operator.
T_0	Length of time window.
$\mathcal{T}$	Set consisting of all transmission lines.
$\mathbf{u}$	Input vector.
$\mathcal{V}_0$	Vicinity set collecting the indexes of the buses near a forced oscillation source.
$\mathbf{x}$	State vector.
$\mathbf{y}$	Output/measurement vector.
Y_t	Measurement matrix.
$Y_{\mathrm{n}t}$	Normalized measurement matrix.
y_k^{F}	Resonance-free component of the k-th measurement.
y_k^{R}	Resonance component of the k-th measurement.

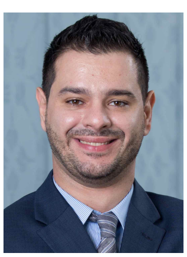
Y_t^{R}	Resonance component matrix.
$\mathbf{z}$	Transformed state vector.

Acknowledgment

The authors would like to thank Diran Obadina, Nemica Kadel, Patrick Gravois, Prabhu Gnanam and Sidharth Rajagopalan of Electric Reliability Council of Texas (ERCOT) for offering test cases and valuable suggestions. The authors also gratefully acknowledge all reviewers for their time and constructive suggestions on this paper during the process of peer review.

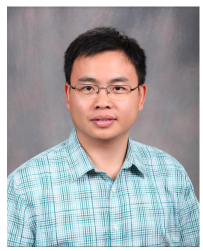


Tong Huang (S'16) received the B.E. degree in electrical engineering from North China Electric Power University, Baoding, China, in 2013 and the M.S. degree in electrical engineering from Texas A&M University, College Station, TX, USA, in 2017, where he is currently working toward the Ph.D. degree. He was a long-term visiting student in the Laboratory for Information and Decision Systems at MIT in 2018. His industry experience includes an internship at ISO-New England in 2018 and an internship at Mitsubishi Electric Research Laboratories in 2019. His research is focused on anomaly management in massively digitized power systems.



Nikolaos M. Freris is a Professor in the School of Computer Science and Technology and the Vice Dean of the International College at the University of Science and Technology of China (USTC). He received the Diploma in ECE from the National Technical University of Athens (NTUA), Greece, in 2005, and the M.S. degree in ECE, the M.S. degree in Mathematics, and the Ph.D. degree in ECE, all from the University of Illinois at Urbana-Champaign (UIUC) in 2007, 2008, and 2010, respectively. His research lies in Cyber-Physical Systems: machine learning, distributed optimization, and wireless networks, with applications in smart cities, intelligent transportation, and power systems. Dr. Freris has published papers in journals and conferences held by IEEE, ACM, and SIAM, and holds three patents. His research has been sponsored by the Ministry of Science and Technology of China, the Department of Science and Technology of Anhui province, China, and the U.S. National Science Foundation. He was recognized with the USTC Alumni Foundation Innovation Scholar award and the IBM High-Value Patent award. Previously, he was with the faculty of NYU, and before that, he held research positions at EPFL and IBM-Research. Dr. Freris is a senior member of IEEE, and a member of ACM, SIAM, and CCF.

P. R. Kumar B. Tech. (IIT Madras, 73), D.Sc.(Washington University, St. Louis, 77), was a faculty member at UMBC (1977-84) and Univ. of Illinois, Urbana-Champaign (1985-2011). He is currently at Texas A&M University. His current research is focused on stochastic systems, energy systems, wireless networks, security, automated transportation, and cyberphysical systems. He is a member of the US National Academy of Engineering and The World Academy of Sciences. He was awarded a Doctor Honoris Causa by ETH, Zurich. He has Award for Control Systems, the Donald P. Eckman received the IEEE Field Award of the AACC, Fred W. Ellersick Prize of the IEEE Communications Society, the Outstanding Contribution Award of ACM SIGMOBILE, the INFOCOM Achievement Award, and the SIGMOBILE Test-of-Time Paper Award. He is a Fellow of IEEE and ACM Fellow. He was Leader of the Guest Chair Professor Group on Wireless Communication and Networking at Tsinghua University, is a D. J. Gandhi Distinguished Visiting Professor at IIT Bombay, and an Honorary Professor at IIT Hyderabad. He was awarded the Distinguished Alumnus Award from IIT Madras, the Alumni Achievement Award from Washington Univ., and the Daniel Drucker Eminent Faculty Award from the College of Engineering at the Univ. of Illinois.



Le Xie (S'05-M'10-SM'16) received the B.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2004, the M.S. degree in engineering sciences from Harvard University, Cambridge, MA, USA, in 2005, and the Ph.D. degree from the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA, in 2009. He is currently a Professor with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA. His research interests include modeling and control of large-scale complex systems, smart grids application with renewable energy resources, and electricity markets.