Characterization of Qualified Grid Nodes for the Identification of Power Network Oscillations

Akash Kumar Mandal and Swades De
Department of Electrical Engineering and Bharti School of Telecommunication and Management
Indian Institute of Technology Delhi, New Delhi 110016, India

***Abstract*—Rapid incorporation of renewable energy sources and power electronic components in the conventional power network has resulted in increased sub-synchronous oscillations (SSOs) in the power network. As a consequence, accurate and exhaustive monitoring of SSOs is pertinent for reliable system operation. This has necessitated the requirement of a revised notion of power system stability and control. In this regard, this paper presents a benchmark analysis in disturbance study by appropriately identifying the most qualified buses that should be co-analyzed to capture the disturbances present in the power network. Our analysis shows that the proposed approach is computationally much faster and offers a significantly reduced number of buses required for disturbance identification over the state-of-the-art, without compromising on the system disturbance identification ability. The theoretical results are validated using the PMU data generated through RSCAD simulation, with 100% identification of critical oscillation frequencies.**

***Index Terms*—Integer optimization, phasor measurement unit, real-time structured computer-aided design, sub-synchronous oscillation.**

I. Introduction

With increasing renewable integration in the conventional power network and high voltage direct current transmission technology, the utilization of power electronics-interfaced components has increased manifolds [1]. This has resulted in an increased sub-synchronous and super-synchronous inter-harmonics injected into the power signals, causing sub-synchronous oscillations (SSOs). This can be attributed to the interaction between the inter-harmonics and torsional vibration frequencies of the generator shaft system [2]. These disturbance frequencies propagate to various parts of the system, causing system destabilization and degradation in power quality [3]. Therefore, real-time identification of such system oscillations is pertinent to the stability and control of modern power networks. Phasor measurement units (PMUs) serve this purpose with highly sampled and time-stamped values of important power system attributes.

A. Literature Review and Motivation

The research to date in the identification of SSOs can be broadly divided into two sets. The first set [4], [5] uses modal transformation techniques for the detection of SSO in the modern power networks. The work in [6] employed fast Fourier transform, [7] utilized wavelet theory, and [8] used Prony algorithm for generating detailed information about the oscillation modes. However, the accuracy of these methods depend on the data size and its spatial variability. It has been established in literature that, the signature from one PMU is not enough to capture the oscillation dynamics of the complete power network [9]. Alternatively, using multiple PMU data could be burdensome from data analysis and execution time standpoint. Therefore, the identification of the optimum number of PMUs, i.e., the PMUs monitoring the most qualified nodes is important for fast, accurate, and exhaustive detection of power system oscillations.

The second type of analysis [10], [11] involves the use of learning-based approaches in the identification of the power system oscillations. The study in [12] used wavelet-based feature extraction from the recorded real-time dataset followed by identification of the class of the disturbance using machine learning. An S-transform-based feature extraction strategy was proposed in [13], followed by an analysis using a combination of extracted features in [14]. The work in [15] used deep learning techniques to perform disturbance classification using an image converted form of the PMU dataset. In context to the SSO identification, the common learning strategies that have been utilized involve deep neural network, support vector machine, random forest, decision tree, artificial neural network, K-nearest neighbour, and naive Bayes [16].

Although machine learning has evolved as an effective tool in identification of such real-time events, these strategies often suffer from high computational complexity [17], over-fitting and local convergence issues [18], and the requirement of an prior data model [19]. Much of these issues arise with the requirement of a huge dataset in any learning-based methodology. Though this helps to reduce modeling error which results in an improved accuracy of classification/identification and faster convergence, it simultaneously leads to the aforementioned trade-offs. Therefore, identification of the most qualified nodes which can help in capturing the significant frequency components of the oscillations in the power network is crucial to an accurate and exhaustive detection of SSO events.

To this end, this paper proposes a novel optimization framework for the identification of most qualified buses that can help in capturing the critical SSO frequency components and their mode shapes, i.e., the spatial distribution of the oscillation across different components (e.g., buses, generators) of the power network. A PMU data-based optimization is utilized for the experimental validation of the obtained theoretical results.

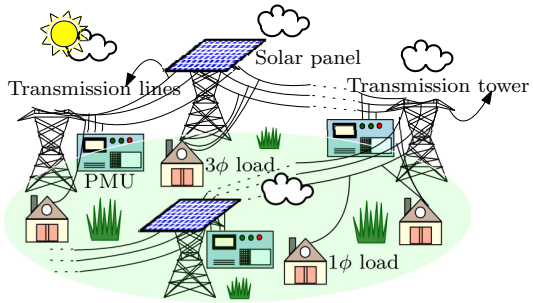


Fig. 1: System model for collaborative oscillation monitoring.

B. Contributions and Significance

The key contributions of this research are: 1) A theoretical system-based optimization problem is formulated for the identification of the important power system buses, that can help in capturing the critical frequency components of the SSO. 2) The solution to the optimization framework is explained for a standard IEEE 5-bus system. 3) A data-based characterization of most critical system parameter in the identification of SSO is done. 4) Finally, these parameters are utilized in the theoretical optimization, which is solved for IEEE 6, 9, 14, 30, and 57-bus networks, and the results are validated with the PMU data-based optimization problem to capture the disturbance signature for the complete power network.

Theoretical results corroborated by the results generated using real-time structured computer-aided deisgn (RSCAD) verify the appropriateness of the proposed optimization formulation and the results that are generated as a byproduct. This study helps in establishing a benchmark in defining the most qualified (important) buses, i.e., characterizing the minimum number of system nodes that must be co-analyzed to extract an accurate and exhaustive information about the SSOs. For example, in a 57-bus system, the number of disturbance identification location in the proposed approach reduces from 18 to 8, with a $\approx 95\%$ reduction in the execution time.

II. System Model

Let the power network as depicted in Fig. 1 has N nodes which are collected in the set $\mathscr{N}$, such that $\|\mathscr{N}\|_c = N$, where $\|\cdot\|_c$ denotes the cardinality operation. High renewable penetration and increased incorporation of power electronic components in the conventional power network injects oscillations of the form $u(t) = a_0 + \sum_{n=1}^{N} a_n e^{-\sigma_n t}\sin(\omega_n t + \delta_n)$, where a_n, σ, $\omega_n \geq 0$, and δ_n denote the amplitude, damping coefficient, frequency, and initial phase offset of the nth oscillatory component, respectively. The network is envisioned to supply power to various load types, requiring a quality power factor. For the sake of grid health monitoring, the PMUs are placed at strategic positions in the grid, that record data for various important power system features. This data is utilized for the identification of SSOs in the power network.

III. Theoretical Optimization for Detection of Power System Oscillation

Let the selection of ith bus be represented using a binary decision variable $x_i \in \{0,1\}$, $i \in \mathscr{N}$, such that $x_i = 1$ if the ith bus is chosen and 0 otherwise. Then, the selection of significant power system nodes/buses can be achieved through the following optimization problem:

$$\begin{aligned} \mathbf{P}1: \quad & \max_{x_i} \sum_i x_i r_i \\ \text{s.t. } \mathbf{C}11: \quad & \sum_i x_i \phi_i^{(m)} \geq 1, \quad \forall m \\ \mathbf{C}12: \quad & \sum_i x_i \theta_i^{(f)} \geq 1, \quad \forall f \\ \mathbf{C}13: \quad & \phi_i^{(m)}, \theta_i^{(f)}, x_i \in \{0,1\}, \quad \forall i, m, f \end{aligned} \tag{1}$$

where $r_i = \sum_{j=1}^{M_i} w_{ij}\lambda_{ij}$, M_i represents the total number of eigenvalues associated with bus i, and λ_{ij} denotes the jth eigenvalue associated with bus i. The weight w_{ij} represents the importance assigned to the jth eigenvalue of bus i. Furthermore, $\phi_i^{(m)}$ and $\theta_i^{(f)} \in \{0,1\}$ are binary indicators representing the presence of mth Eigen mode and fth frequency component in the state matrix of bus i.

Constraint $\mathbf{C}11$ ensures that every important mode of oscillation in the power network is associated to at least one of the chosen buses. Constraint $\mathbf{C}12$ ensures that the fth frequency band of oscillation is an Eigen frequency of at least one selected bus. Finally, constraint $\mathbf{C}13$ imposes a binary limit on all the relevant variables.

A. Definition of $\phi_i^{(m)}$ and $\theta_i^{(f)}$

The mathematical description of every dynamic system can be achieved using a state vector $\mathbf{x} \in \mathbb{R}^{V\times 1}$, input vector $\mathbf{u} \in \mathbb{R}^{V\times 1}$, and an output vector $\mathbf{y} \in \mathbb{R}^{W\times 1}$. In the case when the time derivatives of state variable $\dot{\mathbf{x}} \in \mathbb{R}^{V\times 1}$ and the output are not explicit functions of time t, the dynamic system is represented as $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\mathbf{u})$ and $\mathbf{y} = \mathbf{g}(\mathbf{x},\mathbf{u})$, where $\mathbf{f}(\cdot) \in \mathbb{R}^{V\times V}$ and $\mathbf{g}(\cdot) \in \mathbb{R}^{W\times V}$ are functionals relating the system state and input, to the change in state and output, respectively.

Introducing small perturbations in state vector and input vector and linearizing, we get

$$\begin{aligned} \Delta\dot{\mathbf{x}} =& \mathbf{F}_{\mathrm{S}}\Delta\mathbf{x} + \mathbf{F}_{\mathrm{I}}\Delta\mathbf{u} \\ \Delta\mathbf{y} =& \mathbf{G}_{\mathrm{S}}\Delta\mathbf{x} + \mathbf{G}_{\mathrm{I}}\Delta\mathbf{u} \end{aligned} \tag{2}$$

where $\mathbf{F}_{\mathrm{S}} \in \mathbb{R}^{V\times V}$, $\mathbf{F}_{\mathrm{I}} \in \mathbb{R}^{V\times V}$, $\mathbf{G}_{\mathrm{S}} \in \mathbb{R}^{W\times V}$, and $\mathbf{G}_{\mathrm{I}} \in \mathbb{R}^{W\times V}$ are the system matrix, input matrix, output matrix, and feed-forward matrices, respectively. In order to determine the Eigen frequencies, we solve the characteristic equation:

$$\|\mathbf{F}_{\mathrm{S}} - \lambda\mathbf{I}\| = 0 \tag{3}$$

where $\mathbf{I} \in \mathbb{R}^{V\times V}$ is an identity matrix. Let the Eigen values of the system be denoted by $\lambda_1, \cdots, \lambda_V$. Please note, if (3) is solved for the ith bus, the jth obtained Eigen value follows the nomenclature λ_{ij}. Next, we define an indicator function $\mathfrak{I}(\mathfrak{Z})$ such that the indicator evaluates to unity when the hypothesis $\mathfrak{Z}$ is true. Hereafter, we define two hypotheses for the modeling of $\phi_i^{(m)}$ and $\theta_i^{(f)}$, namely Hypotheses 1 and 2, respectively.

$$\begin{aligned} &\text{Hypothesis 1}: \mathfrak{Z}_{ij}^{(m)} \overset{(\text{def})}{\Longrightarrow} \lambda_m \in (\lambda_{ij} - \Delta\lambda, \lambda_{ij} + \Delta\lambda) \\ &\text{Hypothesis 2}: \mathfrak{Z}_{ij}^{(f)} \overset{(\text{def})}{\Longrightarrow} f \in (f_{ij} - \Delta f, f_{ij} + \Delta f) \end{aligned} \tag{4}$$

where $f_{ij} = \frac{\Im(\lambda_{ij})}{2\pi}$, $\Delta\lambda$ represents the mode variation, i.e., the proximity to a given mode shape as experienced by a given power system node, and $2(\Delta f)$ is the frequency bandwidth of oscillation observed as a virtue of the ijth Eigen mode. Therefore, based on the above-defined hypotheses, we have

$$\begin{aligned}\phi_i^{(m)} &= \mathfrak{I}\left(\mathfrak{Z}_{i1}^{(m)} \vee \mathfrak{Z}_{i2}^{(m)} \vee \cdots \vee \mathfrak{Z}_{iq}^{(m)}\right)\\ \theta_i^{(f)} &= \mathfrak{I}\left(\mathfrak{Z}_{i1}^{(f)} \vee \mathfrak{Z}_{i2}^{(f)} \vee \cdots \vee \mathfrak{Z}_{iq}^{(f)}\right)\end{aligned} \tag{5}$$

where $\vee$ represents the logical OR operation. Thus, $\phi_i^{(m)}$ and $\theta_i^{(f)}$ evaluate to unity if at least one of the involved hypotheses, i.e., $\mathfrak{Z}_{ij}^{(m)}$ and $\mathfrak{Z}_{ij}^{(f)}$, $\forall j = 1, \cdots, q$, respectively, holds true. The modeling of w_{ij} is done next.

B. Modeling of Weights w_{ij}

As per definition w_{ij} represents the importance assigned to the jth Eigen value of the ith bus. We bifurcate this weight into two disjoint product terms α_i and β_{ij}, such that $w_{ij} = \alpha_i\beta_{ij}$. In this notation, α_i models the ability of a node to contribute to the aggregate small-signal oscillations in the grid and β_{ij} captures the contribution of the jth Eigen mode of ith node in that aggregate oscillation. Therefore, β_{ij} can be written as

$$\beta_{ij} = \frac{\|\lambda_{ij}\|}{\sum_{j=1}^{M_i}\|\lambda_{ij}\|} \tag{6}$$

where $\|\cdot\|$ represents the norm operation. Let the ith node be associated with q features, such that $\boldsymbol{\varphi}_{i,\Delta t} = \{\varphi_{i,\Delta t}^{(1)}, \varphi_{i,\Delta t}^{(2)}, \cdots, \varphi_{i,\Delta t}^{(q)}\}$, where $\varphi_{i,\Delta t}$ is the feature set observed in the duration Δt. These features are related using the network equations as follows:

$$\begin{aligned}\varphi_{1,\Delta t}^{(1)} &= \psi_i^{(1)}\left(\varphi_{i,\Delta t}^{(1)}, \varphi_{i,\Delta t}^{(2)}, \cdots, \varphi_{i,\Delta t}^{(q)}\right)\\ \varphi_{i,\Delta t}^{(2)} &= \psi_i^{(2)}\left(\varphi_{i,\Delta t}^{(1)}, \varphi_{i,\Delta t}^{(2)}, \cdots, \varphi_{i,\Delta t}^{(q)}\right)\\ &\vdots\\ \varphi_{i,\Delta t}^{(q)} &= \psi_i^{(q)}\left(\varphi_{i,\Delta t}^{(1)}, \varphi_{i,\Delta t}^{(2)}, \cdots, \varphi_{i,\Delta t}^{(q)}\right)\end{aligned} \tag{7}$$

where the $\psi(\cdot)$ functions could be modeled through Kirchhoff's voltage/current laws, swing equations, power flow equations, etc. Applying Taylor's expansion in (7) for the modeling of small signal oscillations in the features of $\boldsymbol{\varphi}_{i,\Delta t}$ denoted as $\boldsymbol{\Delta\varphi}_{i,\Delta t}$, we have

$$\begin{aligned}\Delta\varphi_{i,\Delta t}^{(1)} &= \left[\boldsymbol{\Delta\psi}_i^{(1)}\left(\boldsymbol{\varphi}_{i,ss}\right)\right]^T \boldsymbol{\Delta\varphi}_{i,\Delta t}\\ \Delta\varphi_{i,\Delta t}^{(2)} &= \left[\boldsymbol{\Delta\psi}_i^{(2)}\left(\boldsymbol{\varphi}_{i,ss}\right)\right]^T \boldsymbol{\Delta\varphi}_{i,\Delta t}\\ &\vdots\\ \Delta\varphi_{i,\Delta t}^{(q)} &= \left[\boldsymbol{\Delta\psi}_i^{(q)}\left(\boldsymbol{\varphi}_{i,ss}\right)\right]^T \boldsymbol{\Delta\varphi}_{i,\Delta t}\end{aligned} \tag{8}$$

where the subscript ss denotes the set of steady state values, $\Delta(\cdot)$ represents the change in the parameter, and T represents the transpose operation. For dimensional compatibility, we must note that $\boldsymbol{\Delta\varphi}_{j,\Delta t}$, $\boldsymbol{\Delta\psi}_i^{(j)}\left(\boldsymbol{\varphi}_{i,ss}\right) \in \mathbb{R}^{q\times 1}$, and their product results in a scalar $\Delta\varphi_{i,\Delta t}^{(j)}$, $\forall j = 1, 2, \cdots, q$. (8) can be concisely represented as

$$\boldsymbol{\Delta\psi}_i\left(\varphi_{i,ss}\right)\boldsymbol{\Delta\varphi}_{i,\Delta t} = \boldsymbol{\Delta\varphi}_{i,\Delta t} \tag{9}$$

where

$$\boldsymbol{\Delta\psi}_i\left(\varphi_{i,ss}\right) = \left[\left[\boldsymbol{\Delta\psi}_i^{(1)}\left(\boldsymbol{\varphi}_{i,ss}\right)\right]^T, \cdots, \left[\boldsymbol{\Delta\psi}_i^{(q)}\left(\boldsymbol{\varphi}_{i,ss}\right)\right]^T\right]^T \in \mathbb{R}^{q\times q}$$

It is notable that (9) represents an Eigen relation, with an unit magnitude of the corresponding Eigen vector. Therefore, to define α_i, we find the contribution of the Eigen vector corresponding to magnitude 1 in the total q Eigen values obtained for the ith node, that is

$$\alpha_i = \frac{1}{\sum_{j=1}^{q}\|\lambda_{ij}^{'}\|} \tag{10}$$

where $\lambda_{ij}^{'}$, $\forall j = 1, 2, \cdots, q$ are the Eigen values of $\boldsymbol{\Delta\psi}_i\left(\boldsymbol{\varphi}_{i,ss}\right)$. The data-based optimization for selection of most qualified nodes is done next.

IV. Data-Based Optimization for Detection of Power System Oscillation

This section proposes a PMU data-dependent optimization for an overall oscillation monitoring in an N-node power network. This serves as a validation for the solution obtained for ($\mathbf{P}1$). Let $\mu_i \in \{0, 1\}$ and $\pi_i \in \{0, 1\}$, such that $i \in \mathscr{N}$, be binary decision variables indicating whether a PMU is installed at the ith grid node and whether or not its chosen for collaboration. If a PMU is installed at node i, $\mu_i = 1$ and 0 otherwise. Similarly, the PMU at node i is selected ($\pi_i = 1$) or not selected ($\pi_i = 0$) to be a part of the collaboration for oscillation detection. Since optimal placement of PMUs is not in the scope of this research, we assume that there are K optimally placed PMUs, with their locations stored in $\mathbf{L}$, such that $\|\mathbf{L}\|_c = K$. Then, we want to be able to detect the power system oscillations with least number of PMU involvement. Therefore, the optimal PMU selection problem to capture the oscillations in the power network is formulated as follows:

$$\begin{aligned}\mathbf{P}21: &\quad \min_{\pi_i}\sum_i \mu_i\pi_i\\ \text{s.t. } \mathbf{C}21: &\quad \chi_t \geq \chi_0\\ \mathbf{C}22: &\quad \sum_i \mu_i\pi_i a_i^{(m)} \geq 1, \quad \forall m\\ \mathbf{C}23: &\quad \sum_i \mu_i\pi_i b_i^{(f)} > 1, \quad \forall f\\ \mathbf{C}24: &\quad \sum_i \mu_i\pi_i \leq K\\ \mathbf{C}25: &\quad a_i^{(m)}, b_i^{(f)}, \pi_i \in \{0, 1\}, \quad \forall i, m, f\end{aligned} \tag{11}$$

where $a_i^{(m)}$ is a binary variable that indicates whether the PMU at node i captures measurements from oscillatory mode m. $a_i^{(m)} = 1$ iff PMU i captures measurements from oscillatory mode m, and 0 otherwise. $b_i^{(f)}$ is a binary variable that indicates whether PMU i captures measurements within frequency range f. $b_i^{(f)} = 1$ if the PMU installed at node i captures measurements within frequency range f, and 0 otherwise.

The χ_t-squared test criterion is defined as $\chi_t = \sqrt{\sum_{j=1}^{q}\left(\delta\varphi_{\Delta t}^{(j)} \cdot \sigma^{-1}_{\varphi_{\Delta t}^{(j)}}\right)^2}$, where $\delta\varphi_{\Delta t}^{(j)} = \frac{\sum_i \mu_i \pi_i \delta\varphi_{i,\Delta t}^{(j)}}{\sum_i \mu_i \pi_i}$, $\delta\varphi_{i,\Delta t}^{(j)} = \varphi_{i,\Delta t}^{(j)} - \varphi_{i,ss}^{(j)}$, with $\varphi_{i,\Delta t}^{(j)}$ representing the amplitude of the jth feature measured by the ith PMU in the time window Δt, subscript ss denotes the steady-state values of the respective parameters, and q denotes the total features monitored by the PMUs, viz., voltage magnitude, voltage phase, current magnitude, current phase, etc. Let these features for the ith node are stored in the feature set $\varphi_{i,\Delta t}$ and $\sigma_{\varphi_{\Delta t}^{(j)}}$ represents the standard deviation in the aggregated measurement of the jth feature.

Constraint $\mathbf{C}21$ ensures that the total represented deviation by the chosen PMUs captures a significant proportion of the potential oscillations in the power system. Constraint $\mathbf{C}22$ ensures that the selected PMUs collectively capture measurements from each oscillatory mode in the power system. Constraint $\mathbf{C}23$ focuses on capturing measurements from specific frequency ranges to enable accurate spectral analysis of oscillations. Constraint $\mathbf{C}24$ ensures that the total number of PMUs in collaboration are less than or equal to the maximum PMUs installed in the grid. Finally, the constraint $\mathbf{C}25$ limits $a_i^{(m)}$, $b_i^{(f)}$, and π_i to be a binary constants.

A. Mathematical Modeling of $a_i^{(m)}$

Let the ith PMU dataset be denoted as $\mathbf{R}_i \in \mathbb{R}^{p\times q}$, where p are the number of time instances over which the data is recorded. Next, we normalize each column of this dataset by subtracting the respective mean and dividing by the standard deviation of that particular data feature. Let $\tilde{\mathbf{R}}_i$ be the normalized dataset obtained from $\mathbf{R}_i$. We construct a collective data matrix $\mathbf{D}$ of all PMU datasets of size $\mathbb{R}^{p\times(q\times K)}$, such that $\mathbf{D} = [\tilde{\mathbf{R}}_1, \cdots, \tilde{\mathbf{R}}_K]$. We apply the singular value decomposition to obtain

$$\mathbf{D} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{12}$$

where $\mathbf{U} \in \mathbb{R}^{p\times p}$ is a upper triangular matrix, $\mathbf{\Sigma} \in \mathbb{R}^{p\times(q\times K)}$, and $\mathbf{V} \in \mathbb{R}^{(q\times K)\times(q\times K)}$ is a lower triangular matrix, and T represents the transpose operation. Next we define a correlation threshold value τ, such that $0 \leq \tau \leq 1$, to identify significant correlations between mode shapes and PMU datasets. For the mth oscillatory mode shape defined by the mth column vector $\mathbf{u}_m \in \mathfrak{C}(\mathbf{U})$ of the matrix $\mathbf{U}$, we calculate the correlation coefficient $\rho(u_m, R_{ij})$ between $\mathbf{u_k}$ and each column of the normalized PMU dataset $\mathbf{R}_i$, where $i \in 1, \cdots, K$, $j \in 1, \cdots, \|\mathfrak{C}(\mathbf{R}_i)\|_c$ is the column space of matrix $\mathbf{R}_i$, and $\mathfrak{C}(\mathbf{U})$ represents the column space of matrix $\mathbf{U}$. Mathematically

$$\rho_{ij}^{(m)} = \rho(u_m, R_{ij}) = \frac{\tilde{\mathbf{u}}_m^T \tilde{\mathbf{R}}_{ij}}{\sigma_{\mathbf{u}_m}\sigma_{\tilde{\mathbf{R}}_{ij}}} \tag{13}$$

where $\tilde{\mathbf{u}}_m = \mathbf{u}_m - \mu_{\mathbf{u}_m}\mathbf{f}_{p\times 1}$, with $\mathbf{f}_{p\times 1}$ being a vector of ones of size $p \times 1$. Therefore, for the indicator function $\mathfrak{I}(\cdot)$ defined before, we define the hypothesis, Hypothesis 3: $\mathfrak{Z} \overset{(\text{def})}{\Longrightarrow} \rho_{ij}^{(m)} \geq \tau$. Therefore, $a_i^{(m)}$ is defined as

$$a_i^{(m)} = \mathfrak{I}\left(\left\{\rho_{i1}^{(m)} \geq \tau\right\} \vee \left\{\rho_{i2}^{(m)} \geq \tau\right\} \vee \cdots \vee \left\{\rho_{iq}^{(m)} \geq \tau\right\}\right)$$

Thus $a_i^{(m)}$ evaluates to 1 if the mth oscillatory mode posses high correlation with at least one feature of the measurements by the ith PMU, i.e., with one column of the data matrix $\mathbf{R}_i$.

B. Mathematical Modeling of $b_i^{(f)}$

From the system deduced in (2) and the corresponding Eigen values achieved using (3), the frequency components of significance are given by the imaginary part of the Eigen values, i.e., the rth Eigen frequency component is written as $f_r = \frac{\Im(\lambda_r)}{2\pi}$, such that $\mathbf{f} = [f_1, \cdots, f_V]^T \in \mathbb{R}^{V\times 1}$ is a vector of Eigen frequencies, where $\Im(\cdot)$ denotes the imaginary part operation. Next, we take the row-wise Fourier transform of the ith PMU dataset. Mathematically

$$\mathfrak{R}_{ij} = \mathscr{F}\left(\tilde{\mathbf{R}}_{ij}\right), \quad \text{where } \tilde{\mathbf{R}}_{ij} \in \mathfrak{C}\left(\tilde{\mathbf{R}}_{ij}\right) \tag{14}$$

where $\mathscr{F}(\cdot)$ represents the Fourier transform operation. The power spectral density (PSD) $\mathfrak{P}_{ij}$ can be represented as

$$\mathfrak{P}_{ij} = \mathfrak{K}(\mathfrak{R}_{ij})\left(\|\mathfrak{K}(\mathfrak{R}_{ij})\|_c\right)^{-1} \tag{15}$$

where $\mathfrak{K}(\cdot)$ represents an element-wise norm-squared operation on the vector $\mathfrak{R}_{ij}$. We define the frequency component f from the $\mathfrak{P}_{ij}$ as $\mathfrak{P}_{ij}^{(f)}$ and a new hypothesis as, Hypothesis 4: $\mathfrak{Z} \overset{(\text{def})}{\Longrightarrow} \mathfrak{P}_{ij}^{(f)} \geq \mathfrak{P}_{\text{th}}$. Therefore, $b_i^{(f)}$ is defined as

$$b_i^{(f)} = \mathfrak{I}\left(\left\{\mathfrak{P}_{i1}^{(f)} \geq \mathfrak{P}_{\text{th}}\right\} \vee \left\{\mathfrak{P}_{i2}^{(f)} \geq \mathfrak{P}_{\text{th}}\right\} \vee \cdots \vee \left\{\mathfrak{P}_{iq}^{(f)} \geq \mathfrak{P}_{\text{th}}\right\}\right).$$

Therefore, $b_i^{(f)}$ evaluates to 1 if the frequency component f is dominant in the PSD of at least one data feature monitored by the ith PMU. The optimization problems $(\mathbf{P}1)$ and $(\mathbf{P}2)$ are solved using the *intlinprog* function of MATLAB.

V. Results and Discussions

This section presents the results for the optimization problems $(\mathbf{P}1)$ and $(\mathbf{P}2)$ using the simulated PMU data from RSCAD. The results are validated for IEEE 5, 6, 9, 14, 30, and 57-bus systems.

A. Example: IEEE 5 Bus (3-Machine) Test System

Fig. 2 depicts the structure and operating conditions of IEEE 5-bus (3-machine) system. Buses 1, 2, and 3 consists of fossil fuel-based, nuclear power-based, and smaller hydro power-based generating units. Machines 1 and 2 have static exciters and stabilizing signals fed-back depending on the rotor speed of the respective machines. Machine 3 has a type 1 exciter, with governor effects included in the simulation of all the three machines. System loads are represented as linear static elements located at buses 2, 3, and 4. Further, two dynamic equivalent models for induction motor loads are envisaged at buses 1 and 4. All per unit values mentioned in Fig. 2 are calculated as per the base values of 600 MVA and 24 kV. Based on the above system description, the detailed state

TABLE I: IEEE 5-bus (3-machine) test system's Eigen values at base condition.

Bus 1		Bus 2		Bus 3		Bus 4		Bus 5	
1	$-.542 \pm j1.854$	8	-2.823	15	-10.42	22	$-19.96 \pm j376.2$	29	-49.95
2	$-.628 \pm j5.943$	9	$-2.975 \pm j9.947$	16	-10.46	23	-20.58	30	-52.21
3	$-.691 \pm j.121$	10	-3.195	17	-13.21	24	-30.54	31	-152.6
4	$-.750$	11	$-3.218 \pm j2.004$	18	$-13.06 \pm j18.13$	25	-30.92	32	-167.8
5	-1	12	$-7.576 \pm j20.96$	19	-13.58	26	-33.02	33	$-239.5 \pm j1365$
6	-1.334	13	$-8.312 \pm j21.23$	20	$-14.41 \pm j376.5$	27	-34.01	34	-500.6
7	$-1.436 \pm j6.242$	14	$-8.555 \pm j27.10$	21	$-15.17 \pm j430.3$	28	$-34.39 \pm j560.9$	35	-502.1

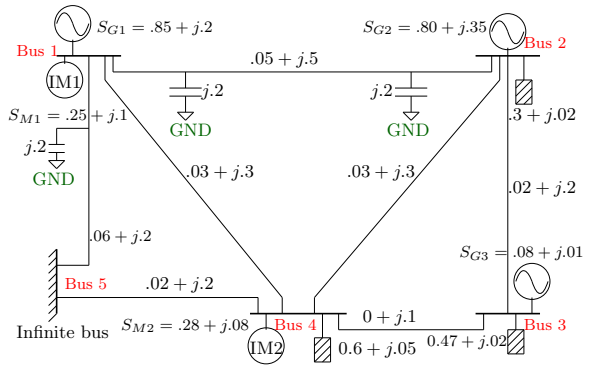
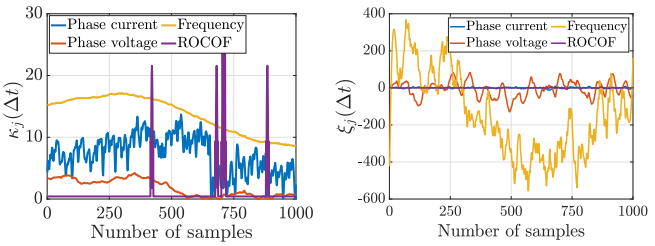


Fig. 2: Network diagram of IEEE 5-bus (3 machine) test system.



(a) (b)

Fig. 3: (a) Time variation of the jth grid feature and (b) variation of chi-squared metric with the change in the value of jth grid feature.

update equations for each bus is derived as detailed in [20] and the corresponding Eigen values are mentioned in Table I.

For the sake of covering all important Eigen frequency ranges, we choose the frequency ranges of importance as 0-1 Hz, 1-5 Hz, 50-80 Hz, and 190-220 Hz, with Δf for these ranges defined as half of the frequency range. For example, for the range 0-1 Hz, the central frequency is 0.5 Hz, with $\Delta f = 0.5$, i.e., the frequency bandwidth of oscillation. Therefore, using Table I, we get $\theta_1^{(f)} = 1$ for $f = 0.5$ Hz, while $\theta_i^{(f)} = 0, \forall i = 2, \cdots, 5$. Next, we define the mode range of importance starting from 0-5 in steps of 5 and $\Delta\lambda = 2.5$, i.e., 0-10, 10-20, and so on. The central mode value is naturally the mid-point of the mode range. Thus, $\phi_1^{(m)} = \phi_2^{(m)} = 1$ for $m = 5$, while $\phi_1^{(m)} = 0$, $\forall i = 3, 4, 5$. Similarly, all other ϕs and θs can be computed for the optimization problem ($\mathbf{P}1$).

On solving ($\mathbf{P}1$), we obtain the optimal solution as 2, 5. Further, on solving ($\mathbf{P}2$) using the PMU data generated using RSCAD, we obtain the same solution vector, i.e., 2, 5. This verifies the theoretical results obtained for ($\mathbf{P}1$) through the solution obtained for ($\mathbf{P}2$) using the simulated PMU dataset.

Remark 1. *It is notable that, since the optimization problem in ($\mathbf{P}1$) and ($\mathbf{P}2$) already obtains the most qualified nodes as per the frequency and mode ranges, the obtained solution helps in capturing the information about the disturbances/oscillations in the entire network in a least-time and tap manner, i.e., by tapping/analyzing least nodes/buses of the network.*

B. Characterization of Critical Grid Feature

From the joint analysis of the PMU dataset obtained through the RSCAD simulation, we plot the variation of $\kappa_j(\Delta t) = \frac{\delta\varphi_{\Delta t}^{(j)}}{\sigma_{\varphi_{\Delta t}^{(j)}}}$ and $\xi_j(\Delta t) = \lim_{\Delta t \to 0} \frac{d\chi_t}{d\alpha_j(\Delta t)}$ with the number of samples, denoting the time variation of the jth grid feature from the grid feature set and the variation of the chi-squared disturbance metric with the change in the value of jth grid feature. The α and β metrics define the time sensitivity of the jth parameter and the sensitivity of chi-squared metric with the jth parameter value. From Fig. 3(a), we note that frequency and ROCOF features demonstrate highest variability with time. Therefore, these features are most important in realizing the presence of disturbance in the power grid network. From Fig. 3(b), we observe that change in frequency and voltage-phase cause maximum impact to the chi-squared disturbance metric. Therefore, we conclude that frequency and voltage-phase are most critical features for the identification of the disturbance in the grid. It is notable that, this inference is consistent with the realization obtained from the swing equation.

C. Optimal Solution to Qualified Node Selection

Table II presents the solution to the 'qualified node selection' optimization problem. The results from the theoretical optimization in ($\mathbf{P}1$) are mentioned in the second column, which are verified by solving ($\mathbf{P}2$) using the simulated PMU data gathered from RSCAD. The solution to ($\mathbf{P}2$) is tabulated in the third column. First, we note that the results obtained from the theoretical optimization, i.e., ($\mathbf{P}1$), match the results obtained from the data-based optimization, i.e., ($\mathbf{P}2$). Next, it can be observed that, by the virtue of selecting the most qualified nodes, the time required in identifying the disturbances/oscillations in the power network reduces significantly.

The final column captures the amount of energy from the power network disturbances that is captured using the data from the selected qualified nodes. This is computed as the ratio of the total energy captured by joint analysis of the

TABLE II: Optimal Solution of the most qualified bus selection for standard IEEE test systems for capturing all Eigen frequencies

IEEE test system	Theoretical optimization	PMU data-based optimization	Strategic PMU placement [21]	Execution time (seconds)		% Reduction in required nodes
				Proposed	Conventional	
6-bus	4, 5	4, 5	4, 5	1.023	1.023	0
9-bus	4, 6	4, 6	4, 6, 8	1.009	7.11	33.33
14-bus	2, 7	2, 7	2, 6, 7, 9	1.013	8.74	50
30-bus	7, 11, 18, 26	7, 11, 18, 26	1, 7, 8, 10, 11, 12, 18, 23, 26, 30	1.793	23.44	60
57-bus	2, 6, 10, 22, 27, 32, 38, 57	2, 6, 10, 22, 27, 32, 38, 57	2, 6, 10, 12, 19, 22, 25, 27, 32 36, 38, 41, 45, 46, 49, 52, 55, 57	2.017	40.03	55.56

most qualified nodes obtained through optimizing ($\mathbf{P}1$) and ($\mathbf{P}2$) to the total energy of the disturbances present in the network. It was noted that the captured disturbance energy is close to 100%. Therefore, we conclude that, by choosing the most qualified nodes for identifying the disturbances in the power network, it is ensured that the disturbances at the critical frequencies, i.e., the Eigen frequencies are noted with certainty, resulting in a robust network protection and control. We further draw attention towards the reduction in the number of nodes required to be co-analyzed for the identification of disturbances. For example, for the IEEE 57-bus system, the number of required nodes is reduced by $\approx 55\%$, leading to a $\approx 95\%$ faster analysis time.

VI. Concluding Remarks

Identification of the oscillations/disturbances in the modern power networks is a crucial task for their reliable operation. This paper conducted a benchmark analysis in distinguishing the most qualified buses that can help in characterizing the disturbance profile of the entire power network. This study is important for reducing the time complexity in classical approaches and reducing model complexity in learning-based approaches. For this purpose, an optimization problem was formulated for the theoretical characterization of such nodes, with sole reliance on the small signal modeling of the power networks. A second optimization was executed to validate the results obtained from the first optimization using the PMU data obtained from the RSCAD. An intermediate result for the most critical feature of the grid was noted, signifying the grid feature with the most information about the disturbances in the network. The solution to the optimization problems demonstrated that by selecting the best qualified buses/nodes, the disturbance profile for the entire power network can be gathered in a reduced time frame while capturing a considerable fraction of the total energy in the disturbance wave. The results were validated for various standard IEEE test systems.

References

[1] Q. Hou, E. Du, N. Zhang, and C. Kang, "Impact of high renewable penetration on the power system operation mode: A data-driven approach," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 731–741, 2019.

[2] W. Du, B. Ren, H. Wang, and Y. Wang, "Comparison of methods to examine sub-synchronous oscillations caused by grid-connected wind turbine generators," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4931–4943, 2019.

[3] A. K. Mandal, A. Malkhandi, S. De, N. Senroy, and S. Mishra, "A multipath model for disturbance propagation in electrical power networks," *IEEE Trans. Circuits Syst. II: Express Briefs*, vol. 70, no. 4, pp. 1460–1464, 2022.

[4] X. Yang, J. Zhang, X. Xie, X. Xiao, B. Gao, and Y. Wang, "Interpolated dft-based identification of sub-synchronous oscillation parameters using synchrophasor data," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2662–2675, 2019.

[5] P. G. Estevez, P. Marchi, F. Messina, and C. Galarza, "Forced oscillation identification and filtering from multi-channel time-frequency representation," *IEEE Trans. Power Syst.*, vol. 38, no. 2, pp. 1257–1269, 2022.

[6] F. Zhang, L. Cheng, W. Gao, and R. Huang, "Synchrophasors-based identification for subsynchronous oscillations in power systems," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 2224–2233, 2018.

[7] E. Sezgin and Ö. Salor, "Analysis of power system harmonic subgroups of the electric arc furnace currents based on a hybrid time-frequency analysis method," *IEEE Trans. Indus. Appl.*, vol. 55, no. 4, pp. 4398–4406, 2019.

[8] T. Xia, Z. Yu, K. Sun, D. Shi, and Z. Wang, "Extended prony analysis on power system oscillation under a near-resonance condition," in *in Proc. IEEE Power Energy Society General Meeting (PESGM)*. IEEE, 2020, pp. 1–5.

[9] L. G. Meegahapola, S. Bu, D. P. Wadduwage, C. Y. Chung, and X. Yu, "Review on oscillatory stability in power grids with renewable energy sources: Monitoring, analysis, and control using synchrophasor technology," *IEEE Trans. Indus. Electron.*, vol. 68, no. 1, pp. 519–531, 2020.

[10] N. Severoglu and Ö. Salor, "Statistical models of eaf harmonics developed for harmonic estimation directly from waveform samples using deep learning framework," *IEEE Trans. Indus. Appl.*, vol. 57, no. 6, pp. 6730–6740, 2021.

[11] Y. Deng, L. Wang, H. Jia, X. Tong, and F. Li, "A sequence-to-sequence deep learning architecture based on bidirectional gru for type recognition and time location of combined power quality disturbance," *IEEE Trans. Indus. Informat.*, vol. 15, no. 8, pp. 4481–4493, 2019.

[12] H. Zhao, J. Liu, H. Chen, J. Chen, Y. Li, J. Xu, and W. Deng, "Intelligent diagnosis using continuous wavelet transform and gauss convolutional deep belief network," *IEEE Trans. Reliability*, 2022.

[13] P. K. Ray, S. R. Mohanty, N. Kishor, and J. P. Catalão, "Optimal feature and decision tree-based classification of power quality disturbances in distributed generation systems," *IEEE Trans. Sustain. Energy*, vol. 5, no. 1, pp. 200–208, 2013.

[14] A. Shahsavari, M. Farajollahi, E. M. Stewart, E. Cortez, and H. Mohsenian-Rad, "Situational awareness in distribution grid using micro-pmu data: A machine learning approach," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6167–6177, 2019.

[15] K. Chen, J. Hu, and J. He, "A framework for automatically extracting overvoltage features based on sparse autoencoder," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 594–604, 2016.

[16] M. Sahani and P. K. Dash, "Automatic power quality events recognition based on hilbert huang transform and weighted bidirectional extreme learning machine," *IEEE Trans. Indus. Informat.*, vol. 14, no. 9, pp. 3849–3858, 2018.

[17] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015.

[18] L. Rice, E. Wong, and Z. Kolter, "Overfitting in adversarially robust deep learning," in *Int. Conf. Machine Learning*. PMLR, 2020, pp. 8093–8104.

[19] Z. Wan, X. Xia, D. Lo, and G. C. Murphy, "How does machine learning change software development practices?" *IEEE Trans. Soft. Eng.*, vol. 47, no. 9, pp. 1857–1871, 2019.

[20] M. Z. H. El-Din, "An efficient approach for dynamic stability analysis of power systems-including load effects," Ph.D. dissertation, 1977.

[21] N. M. Manousakis and G. N. Korres, "Optimal PMU arrangement considering limited channel capacity and transformer tap settings," *IET Gener. Transm. Distrib.*, vol. 14, no. 24, pp. 5984–5991, 2020.