Topology Identification for Dynamic Equivalent Models of Large Power Systems using $\ell_1/\ell_2$ Optimization

Seyedbehzad Nabavi, Student Member, IEEE, Aranya Chakrabortty, Member, IEEE,

Abstract—In this paper we propose two algorithms to identify the equivalent topology of reduced-order models of power systems using measurements of phase angles and frequencies available from Phasor Measurement Units (PMUs). We first show that the topology identification problem can be posed as a parameter estimation problem. Thereafter, we extract the slow oscillatory component of the PMU measurements using subspace identification methods and use them for identifying the topological parameters via a $\ell_2$ minimization approach. Next, we consider the problem of identifying a sparse equivalent network topology using a $\ell_1/\ell_2$ minimization technique. We illustrate our results using IEEE 39-bus model and an Australian 59-bus model.

Index Terms—Reduced order systems, swing equation, phasor measurement, subspace identification, interarea oscillations

Notations:

- $A_{ij}$: $(i, j)^{th}$ element of matrix $A$
- $I_p$: identity matrix of size $p \times p$
- $0_{q \times s}$: $q \times s$ matrix of zero elements
- $|a|$: absolute value of scaler $a$
- $\|x\|_0$: $\ell_0$ norm of vector $x$
- $\|x\|_1$: $\ell_1$ norm of vector $x$
- $\|x\|_2$: $\ell_2$ norm (Euclidean norm) of vector $x$
- $\|A\|_2$: 2-norm of matrix $A$
- $\|A\|_2 = \max \{\|Ax\|_2 / \|x\|_2\}$

I. INTRODUCTION

Recent analyses of phase angle and frequency oscillations in the Western Electricity Coordinating Council (WECC) and the Eastern Interconnect (EI) have highlighted the importance of constructing real-time dynamic models of large power systems from Synchrophasor measurements [1]. For example, spectral analysis of phase angles in the California-Oregon Intertie [2], or frequency oscillations between New England and the Entergy grid in Florida [3] clearly indicate the potential use of such predictive models for critical applications such as oscillation monitoring, transient stability assessment, and voltage instability prediction. However, given the large size of any realistic power system, such as the WECC or EI, it is practically impossible to derive the pre-event or post-event dynamic model for the entire network in real-time. System operators are, rather, more interested in constructing reduced-order models of the power system that capture the dominant inter-area modes of oscillation, and, hence, can predict how the different parts of the system may oscillate with respect to each other in the face of a particular event. Such reduced-order models are often referred to as wide-area models.

Preliminary results for constructing wide-area models of two-area power systems using PMU measurements have recently been presented in [4]. The authors in this paper listed two main steps for deriving these models, namely 1. identification of the dynamic equivalent model for each area, and 2. identification of the topology of the equivalent reduced-order transmission network connecting these areas. The topology identification step was not addressed in [4] as the system under study was a simple two-area system connected by a single tie-line. For multi-area power systems, however, identifying the equivalent topology becomes absolutely imperative. The topology graph captures the effective impedance between the areas, and thereby retains the modal frequencies and the mode shapes for the inter-area oscillations. The idea is shown in Fig. 1a, where an IEEE 39-bus power system is divided into four areas using coherent grouping algorithms [5]. The generators marked as $G_1$ to $G_4$ in Fig. 1b represent the equivalent generators for each respective area, and the intermediate graph hypothetically indicates which area is connected to which other areas and with what equivalent weight.

Our objective is to identify this equivalent topology for any multi-area power system with well-defined inter-area modes using PMU measurements from inside the areas and from the boundaries. We first show that the topology identification problem can be posed as a parameter estimation problem. Thereafter, we extract the slow oscillatory component of the PMU measurements using subspace identification methods and use them for identifying the topology via $\ell_2$ minimization. Next, we consider the problem of identifying a sparse equivalent network topology using mixed $\ell_1/\ell_2$ optimization.

The uniqueness of our approach compared to traditional topology identification methods are as follows:

- First, the majority of network reduction methods used in the power system literature are model-based, such as the methods based on modal equivalencing [5], coherency [6], and decomposition algorithm assigning coupling factors to generators [7]. In contrast, our methods are completely measurement-based, and need only a few basic information...
about the underlying system model.

- Traditional methods such as SME [8] tend to capture the details of fast local oscillations that may not be necessary for wide-area monitoring but increases the computation time. Our methods, in contrast, are based on inter-area or slow oscillation only and, therefore, will be significantly faster.

- In computer science literature, topology identification of large complex networks is often formulated as a combinatorial optimization problem [9]. Our methods, however, are not based on combinatorial analysis, but follow from underlying system dynamics, thereby preserving all the system-level properties as reflected in the PMU measurements.

- Compared to recent works of [10]–[12] on topology identification of generic network dynamic systems using graph-theoretic methods, and of [13] where raw PMU data was used to estimate a Thevenin equivalent model of power systems in steady-state, our approach integrates model reduction with identification by considering separation of slow and fast dynamics. Details of this separation are shown in Section III.

Once constructed, the inter-area topology can be used for useful applications such as transient stability assessment and voltage stability assessment. For example, a common tool for assessing transient stability of a multi-area power system is the transient energy function consisting of kinetic and potential energies. As shown in [14], the potential energy function is dependent on the inter-area topology. Similarly, this topology can also serve as a critical parameter for tracking loadability limits for voltage, wide-area protection, and islanding schemes, several examples on which are provided in this paper in Section VI.

The remainder of the paper is organized as follows: Section II presents the power system models of interest. Section III describes model reduction and signal separation. Section IV formulates the topology identification problem as a parameter estimation problem, and proposes two methods to solve it using linear and nonlinear least squares methods. Section V presents a sparse topology realization algorithm. Section VI illustrates the results through simulations. Section VII concludes the paper.

II. PROBLEM FORMULATION

To motivate the problem of topology identification, we first recall the general ideas on how a multi-area power system model can be reduced to its dynamic equivalent via time-scale separation. For this, let us consider a power system network consisting of \( n \) synchronous generators and \( n_l \) loads connected by a given topology. Without loss of generality, we assume buses 1 through \( n \) to be the generator buses and buses \( n + 1 \) through \( n + n_l \) to be the load buses. Let \( P_m \) denote the vector of the mechanical power injection at generator buses, \( P_L \) be the vector of total active power consumed by the loads, and \( P_i \) be the total active power injected to the \( i^{th} \) bus of the network \( (i = 1, \ldots, n + n_l) \), where the superscript \( N \) indicates that this power is flowing in the network as opposed to the loads. This power is calculated as:

\[
P_i^N = \sum_{k=1}^{n+n_l} \left( V_i^2 r_{ik}/y_{ik}^2 + V_k V_i \sin(\theta_{ik} - \alpha_{ik})/y_{ik} \right),
\]

where, \( V_i \angle \theta_i \) is the voltage phasor at the \( i^{th} \) bus, \( \theta_i = \theta_i - \theta_k \), \( r_{ik} \) and \( x_{ik} \) are the resistance and reactance of the transmission line joining buses \( i \) and \( k \), \( y_{ik} = \sqrt{r_{ik}^2 + x_{ik}^2} \), and \( \alpha_{ik} = \tan^{-1}(r_{ik}/x_{ik}) \). Let \( P_{m}^N \) and \( P_{L}^N \) denote the vectors of \( P_i^N \) calculated for generators and loads, respectively. The electromechanical model of the power system can be described as a system of differential-algebraic equations (DAE) [15]:

\[
M\dot{\delta} = P_m - P_{G}^N - D\omega, \tag{2a}
\]

\[
P_L - P_{L}^N = 0, \tag{2b}
\]

where \( \delta \) is the vector of generator angles, \( \omega \) is the vector of the speed deviation of the generators from synchronous speed, and \( M = \text{diag}(M_i) \) and \( D = \text{diag}(D_i) \) are \( n \times n \) diagonal matrices of the generator inertias and damping factors, respectively. The DAE (2) can be converted to a system of pure differential equations by relating the algebraic variables \( V_i \) and \( \theta_i \) to the system state variables \( (\delta, \omega) \) from (2b), and then substituting them back in (2a) via Kron reduction. The resulting system is a fully connected network of \( n \) second-order oscillators with \( l \leq n(n-1)/2 \) tie-lines. Let the internal voltage phasor of the \( i^{th} \) machine be denoted as \( \dot{E}_i = E_i \angle \delta_i \). The electromechanical dynamics of the \( i^{th} \) generator in the Kron’s form, neglecting
line resistances, can be written as:
\[ \delta_i = \omega_i, \]  
\[ M_i \delta_i = P_{mi} - \sum_k \left( \frac{E_i E_k}{x_{ik}} \sin(\delta_{ik}) \right) - D_i \delta_i, \]  
for \( i = 1, \ldots, n \).

Linearizing (3) about the equilibrium \( (\delta_{i0}, 0) \) results in the small signal model:
\[ \begin{bmatrix} \Delta \delta \\ \Delta \omega \end{bmatrix} = \begin{bmatrix} 0 & I_n \\ -M^{-1}L & -M^{-1}D \end{bmatrix} \begin{bmatrix} \Delta \delta \\ \Delta \omega \end{bmatrix} + \begin{bmatrix} 0 \\ e_j \end{bmatrix} u_i, \]  
\[ \]  
where, \( \Delta \delta = [\Delta \delta_1 \cdots \Delta \delta_n]^T \), \( \Delta \omega = [\Delta \omega_1 \cdots \Delta \omega_n]^T \), \( I_n \) is the \( n \)-dimensional identity matrix, \( e_j \) is the \( j \)th unit vector with all elements zero but the \( j \)th element that is 1, considering that the input is modeled as a change in the mechanical power in the \( j \)th machine. However, since we are interested only in the oscillatory modes or eigenvalues of \( A \), this assumption is not necessary, and the input can be modeled in any other feasible way such as faults and excitation inputs. This fact will be illustrated in the simulations in Section VI. The matrix \( L \) in (4) is the \( n \times n \) Laplacian matrix of the form:
\[ \mathcal{L}_{ij} = (E_i E_j \cos(\delta_{i0} - \delta_{j0}))/x_{ij}, \ i \neq j, \]  
\[ \mathcal{L}_{ii} = -\sum_{k=1}^n (E_i E_k \cos(\delta_{i0} - \delta_{k0}))/x_{ik}. \]  
\[ \]  
It is obvious that \( \mathcal{L}_{ij} = \mathcal{L}_{ji} \). Let us denote the \( i \)th eigenvalue of the matrix \( M^{-1}L \) by \( \lambda_i \). The largest eigenvalue of this matrix is equal to 0, and all other eigenvalues are negative, i.e., \( \lambda_n \leq \cdots \leq \lambda_2 < \lambda_1 = 0 \). Assuming the magnitude of \( D_i/M_i \) are small, the eigenvalues of \( A \) can be approximated by \( \pm \Omega_i (j = \sqrt{-1}), \) where \( \Omega_i = \sqrt{|\lambda_i|} \).

Now, assuming that the entire network consists of \( r \) coherent areas that are sparsely-connected, and the \( i \)th area consists of \( m_i \) nodes with connections defined by \( M_i^{-1}L_i \), \( i = 1, \ldots, r \), one can then rewrite (4) as shown in (9) where, \( \Delta \delta^i_j \) means the angle of the \( j \)th machine in the \( i \)th area. Construction of matrices \( K_i \) follows from the definitions of \( A \) and \( M_i^{-1}L_i \) for \( i = 1, \ldots, r \). The off-diagonal block of the matrix \( M^{-1}L \) are shown by asterisks (*). The (\( i, j \))th off-diagonal block shows the connectivity between areas \( i \) and \( j \) in the Kron-reduced form. Due to the assumption of coherence following from the difference between the local and inter-cluster reactances and inertias, the oscillatory modes of the matrix \( A \) will be divided into sets of \((r-1)\) inter-area modes with eigenvalues \(-\sigma_i \pm j\Omega_i \) through \(-\sigma_{r-1} \pm j\Omega_{r-1} \). The remaining \((n-r)\) modes will be characterized by intra-area modes representing the local oscillations inside areas. In fact, as shown in [5], if the inertias \( M_i \forall i = 1, \ldots, n \) are of the same order of magnitude, then from (5a) it is clear that \( \mathcal{L}_{ij} \) will be a large positive number for nodes \( i \) and \( j \) that are connected by a short tie-line with small reactance \( x_{ij} \) while \( \mathcal{L}_{ij} \) will be a small positive number for nodes that are connected by a long transmission line with large reactance \( x_{ij} \), leading to a sharp separation of inter-area and intra-area frequencies. Our basic assumption for dynamic equivalencing is that the \((r-1)\) inter-area modes can be attributed to \( r \) equivalent machines, as shown in Fig. 1. Let \( E_k^E \), \( \delta_k^E(t) \), and \( \omega_k^E(t) \) denote the voltage, angle, and frequency of the \( k \)th equivalent machine, respectively. The equivalent small-signal model for the inter-area dynamics of (4) is shown in (10), where \( \Delta \delta^E \equiv [\Delta \delta^E_1 \cdots \Delta \delta^E_r]^T \), and \( \mathcal{M}^E \) and \( \mathcal{D}^E \) are \((r \times r)\) diagonal matrices of equivalent machine inertias and equivalent machine dampings, respectively. \( \mathcal{L}^E \) represents the connectivity of the \( r \) areas in the equivalent topology whose elements are as follows:
\[ \mathcal{L}^E_{ij} = \frac{(E_k^E E_j^E \cos(\delta_{i0}^E - \delta_{j0}^E))/x_{ij}^E}{x_{j0}^E}, \ i \neq j \]  
\[ \mathcal{L}^E_{ii} = -\sum_{k=1}^r (E_k^E E_k^E \cos(\delta_{i0}^E - \delta_{k0}^E))/x_{ik}^E, \]  
\[ \]  
where, \( x_{ij}^E \) is the equivalent reactance of the tie-line connecting areas \( i \) and \( j \) in the reduced-order model.

Our objective is to find the equivalent topology that connects these \( r \) equivalent machines using PMU data, which is equivalent to estimating the elements of the matrix \( \mathcal{L}^E \). Given the measurements of \( \Delta \delta \) and \( \Delta \omega \) from (9), our first task, therefore, is to apply modal decomposition techniques by which we can extract the slow components of these outputs, i.e., \( \Delta \delta^E \) and \( \Delta \omega^E \) respectively, as defined in (10), and, thereafter, use these slow components to estimate \( \mathcal{L}^E_{ij} \forall i, j = 1, \ldots, r \). We next review a modal decomposition technique using Hankel matrices to achieve the first task.

### III. Modal Extraction using Hankel Matrices

#### A. Modal Decomposition

As discussed in the previous section, the output measurements of (9) will contain the contribution of both inter-area and intra-area modes. The first step in topology identification of the dynamic equivalent model is to find the slow component of each measurement. Several methods for such modal extraction have recently been proposed both in the power system literature [16]–[20], and in the control systems literature [21] with applications to real-world power system models such as the WECC system [22]. All of these methods have their own advantages and disadvantages depending on applications. Among them, subspace identification methods such as Eigensystem Realization Algorithm (ERA) [20] have been shown to be very useful tools for identifying slow oscillation modes from PMU data [18]. The choice of ERA also follows from the fact that it is computationally fast and can be implemented in real-time. We next summarize the basic ERA algorithm as follows.

Let us consider a general continuous-time LTI system:
\[ \dot{x}(t) = Ax(t) + Bu(t), \ y(t) = Cx(t), \]  
\[ \]  
where \( A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times p}, \) and \( C \in \mathbb{R}^{r \times n} \) are unknown state space matrices and need to be identified from the output measurements \( y(t) \) and the input \( u(t) \). We assume the triplet \( (A, B, C) \) are controllable and observable. Also, \( u(t) \) is assumed to be persistently exciting. The discrete-time equivalent of (7) is written as
\[ x(k+1) = A_d x(k) + B_d u(k), \ y(k) = C x(k). \]  
\[ \]
The impulse response of (8) will be
\[ y(k) = CA_d^{k-1}B_d. \]  
(11)

Given measurement \( y(k) \) for \( k = 0, \ldots, m \), we next construct two \( l \times s \) Hankel matrices \( H_0 \) and \( H_1 \) as:
\[
H_0 \triangleq \begin{bmatrix} y_0^0 & y_1^0 & \cdots & y_m^0 \\ \end{bmatrix}, \quad H_1 \triangleq \begin{bmatrix} y_0^1 & y_1^1 & \cdots & y_s^1 \\ \end{bmatrix},
\]
(12a)
\[
\text{where } l \text{ and } s \text{ are positive integers satisfying } (n < l, n < s, s + l \leq m). \]
(12b)

where \( l \) and \( s \) are positive integers satisfying \( (n < l, n < s, s + l \leq m) \). \( y_j^j = \text{col} \{ y(i+j) \cdots y(i+j+l-1) \} \) for \( j = 0, 1 \) and \( i = 1, \ldots, s \). It can be easily shown that \( H_0 = \mathcal{O}C \) and \( H_1 = \mathcal{O}A_dC \), where \( \mathcal{O} \) and \( \mathcal{C} \) are observability and controllability matrices for (8), respectively. We next consider the truncated singular value decomposition of \( H_0 \) by retaining its largest \( n \) singular values as
\[
\hat{H}_0 = \hat{R} \hat{\Sigma} \hat{S}^T.
\]
(13)

Defining \( \mathcal{E}_1 \triangleq [I_p \ 0_{p \times (s-p)}]^T \) and \( \mathcal{E}_2 \triangleq [I_q \ 0_{q \times (l-q)}]^T \), the estimates for the triplet \( (A_d, B_d, C) \) up to a similarity transformation can be calculated as
\[
\hat{A}_d = \hat{\Sigma}^{-1/2} \hat{R}^T \hat{H}_1 \hat{S} \hat{\Sigma}^{-1/2},
\]
(14)
\[
\hat{B}_d = \hat{\Sigma}^{1/2} \hat{S}^T \mathcal{E}_1, \quad \hat{C} = \mathcal{E}_2^T \hat{R} \hat{\Sigma}^{1/2}.
\]
(15)

One can next convert \( (\hat{A}_d, \hat{B}_d, \hat{C}) \) to their continuous-time counterpart \( (\hat{A}, \hat{B}, \hat{C}) \) by zero-order hold.

For our application, since the eigenvalues of \( A \) satisfy the two-time-scale property of (9), the oscillatory components of the output response of (7) can be estimated as
\[
y(t) = \sum_{i=1}^{n-1} (\alpha_i + j\beta_i) e^{-\sigma_i t} e^{j\Omega_i t} + \sum_{k=r}^{s-1} (\alpha_k + j\beta_k) e^{-\sigma_k t} e^{j\Omega_k t},
\]
(16)
where \( -\sigma_i \pm j\Omega_i \) are the eigenvalues of \( \hat{A} \), and the residues \( \alpha_i \pm j\beta_i \) follow from the state-space structure of \( (\hat{A}, \hat{B}, \hat{C}) \) for \( i = 1, \ldots, n - 1 \). Therefore, given PMU measurements \( y(t) \) one can easily construct \( y^E(t) \) using (7)-(16). Note that if \( y(t) \) is corrupted by additive white Gaussian noise, one may use the stochastic variant of ERA [23] to preserve the accuracy of estimating \( y^E(t) \).

It is important to note that ERA alone cannot be used for solving the topology identification problem. This is because \( A_d \) will only be a similarity transform of \( \hat{A}_d \) and, therefore, may not capture the topology of the system through a Laplacian structure. Thus, we have to cast the problem using a parameter estimation framework as presented in Section IV.

**Remark 1:** The standard approach for doing modal decomposition, as shown in [17], [18], is to assume the input to be an impulse function. Accordingly, in (8) we assumed \( u(k) \) to be an impulse input. In fact, since we are interested in identifying only the inter-area modes, even if \( u(k) \) is not an impulse, ERA will still give accurate results provided that its frequency content do not lie in the inter-area frequency range.

**B. Choice of Measurements**

A plausible question for using the ERA algorithm is how to construct the signal \( y(t) \) from the available PMU measurements. This problem is addressed for three different cases based on the availability of the first-principle knowledge about the underlying system as well as the characteristics of the network. For all three cases we assume that the PMU measurements are collected from the generator buses and that they can be used to approximate the corresponding generator states due to the negligible internal reactances of the generators.

a) **Case 1:** Assume a well-defined time-scale separation between the intra-area and inter-area dynamics. In other words, generators inside an area are densely connected via short tie-lines, while the areas themselves are sparsely connected through longer tie-lines. Also, we assume that the machine inertias for any area are of the same order of magnitude. In such an ideal case, all machines in one area have almost similar slow oscillatory component of \( \delta(t) \) and \( \omega(t) \). Thus, measurements from any machine in an area can be used to construct \( y(t) \) followed by ERA to extract the inter-area component \( y^E(t) \) of the corresponding aggregated areas.

b) **Case 2:** Suppose PMUs are installed at \( N^2 \) number of generator buses in area \( j \) with \( M_j^E \) being the inertia of the \( i^{th} \) generator, \( i = 1, \ldots, N^2 \), which is assumed to be known. Let \( \delta_j^i(t) \) be the corresponding phase angle measurement from the \( i^{th} \) generator in area \( j \). Then, the representative measurement
for this area is obtained by defining an approximate center of inertia angle \( \delta_j(t) \):
\[
\hat{\delta}_j \triangleq \left( \sum_{i=1}^{N^L_j} M_i^j \delta_i \right) / \left( \sum_{i=1}^{N^L_j} M_i^j \right).
\]

If the measurements of angular velocity are also available, we can similarly define the approximate center of inertia velocity.

c) Case 3: In case the values of \( M \) for the nodes in any given area are not available, we will consider the slow dynamics of the boundary node as the equivalent slow dynamics of that area. The choice follows from the fact that boundary nodes are connected to long inter-area transmission lines and, therefore, the slow modes have higher participation factor at these nodes [24].

IV. Topology Identification of the Reduced-Order Network

It follows from (10) that the topology of the reduced-order model can be identified by estimating its parameters, which are embedded in the elements of \( L^E, M^E, \) and \( D^E \). To start, one can assume the equivalent topology to be a complete graph, meaning that every area is connected to every other area with equivalent tie-line weights (\( L^E_\Delta \)). The identified elements of \( L^E \) will then determine the required topology. A threshold \( L_T \) also can be set so that if \( L^E_\Delta_{ij} < L_T \), then \( L^E_\Delta_{ij} \) will be set to zero, i.e., areas \( i \) and \( j \) will not be connected. Next, we present a linear least squares and a nonlinear least squares approach to estimate \( L^E, M^E, \) and \( D^E \). We must have the following four apriori information before these optimizations can be executed: i) which generator belongs to which area, ii) the distribution of PMUs, iii) generator inertias corresponding to the PMUs (to satisfy Case 2 of Section III-B), iv) identity of the boundary buses of areas whose machine inertias are not known (to satisfy Case 3 of Section III-B).

A. Identification using Linear Least Squares (LLS)

The discrete-time representation of the continuous-time reduced-order power system of (10) is written as
\[
\dot{x}(k+1) = A_dx(k) + B_du(k),
\]
where \( B_d \in \mathbb{R}^{2r \times 1} \) and \( A_d \in \mathbb{R}^{2r \times 2r} \) is equal to \( e^{A^E T_s} \simeq I_{2r} + A^E T_s \) with \( A^E \) following from (10) as
\[
A^E = \begin{bmatrix}
0 & I_r \\
(M^E)^{-1}L^E & -(M^E)^{-1}D^E
\end{bmatrix},
\]
and \( T_s \) is the sampling period. The vector \( x(k) \) given by
\[
x(k) = [\Delta \delta^E_\Delta(k) \cdot \Delta \delta^E_r(k) \cdot \Delta \omega^E_\Delta(k) \cdot \Delta \omega^E_r(k)]^T,
\]
for \( k = 0, \ldots, m \), can be obtained by passing the output measurements of \( \Delta \delta(t) \) and \( \Delta \omega(t) \) from (9) through ERA, exactly in the same way as \( y^E(t) \) is obtained from \( y(t) \) in (16). One can rewrite (17) as
\[
x(k+1) = F(k)Y,
\]
where
\[
Y \triangleq [A_d u_{1}, A_d u_{2} \cdots A_d u_{r}, B_d 1 \cdots B_d r]^T
\]
and
\[
F(k) \triangleq \begin{bmatrix}
x^T(k) & \cdot & \cdot & u(k) \\
\cdot & \ddots & \cdot & \cdot \\
\cdot & \cdot & x^T(k) & u(k)
\end{bmatrix}.
\]

Let us concatenate \( F(k) \) and \( x(k) \) into matrices \( F \) and \( G \) as:
\[
F \triangleq \text{col}[F(0), \ldots, F(m-1)], \quad G \triangleq \text{col}[x(1), \ldots, x(m)].
\]
The LLS formulation for estimating \( Y \) can be posed as:
\[
\min_Y \|FY - G\|_2^2,
\]
Note that the measurements from all states of the reduced-order model (10) are needed to form matrices \( F \) and \( G \) in (21). The problem (22) is convex and can be solved using MATLAB cvx toolbox [25]. The drawback of this formulation, however, is that \( F \) is also dependent on the measurements that are usually perturbed by noise. In Appendix A we show how the perturbation in measurements will affect the estimation accuracy of (22). Equation (29) indicates that the estimation error for LLS grows in the order of \( \kappa^2(F) \), where \( \kappa(F) \triangleq \|F^\dagger\|_2 \|F\|_2 \) is the condition number of \( F \) and, therefore, can reduce the accuracy of the estimates if \( F \) is ill-conditioned [26]. Therefore, if the PMU data are noisy, a more accurate estimation method is nonlinear least squares, where \( \|\delta Y\|_2 \) grows linearly with \( \kappa(F) \). We next illustrate this approach.

B. Identification using Nonlinear Least Squares (NLS)

The NLS estimation can be posed as:
\[
\min_{\hat{F}} \frac{1}{2} \|\hat{F}Y - G\|_2^2,
\]
where \( G \) is formed as shown in (21), and \( \hat{F} \) is a function of \( Y \), that is equal to \( \hat{F} \triangleq \text{col}[\hat{F}(0), \ldots, \hat{F}(m-1)] \), where \( \hat{F}(k) \triangleq \begin{bmatrix} \text{diag}(\hat{x}^T(k)) & \text{diag}(u(k)) \end{bmatrix} \) and \( \hat{x}(k) \) is the reproduced system states of (17) at \( Y \) for \( k = 0, \ldots, m - 1 \). The NLS problem of (23) can be solved using MATLAB optimization toolbox. The perturbation analysis of (23) is shown in Appendix B. Equation (33) shows that the upper bound of the error in estimation of the step size \( \Delta \delta(k) \) at the trial iteration \( k \) is proportional to \( \kappa(\hat{F}(k)) \). The drawback of NLS, however, is that the problem formulation is not convex. Thus, there is no guarantee of convergence using a single run of a derivative-based method such as Gauss-Newton. Several runs of the optimization are needed, and the resulting parameter set with the minimum error is chosen as final estimates.

C. Uniqueness of the Topology Identification

Please see Appendix C for the uniqueness of \( A_d \) obtained from (22) and (23).

V. Sparse Identification of the Reduced-Order Network

This section considers identification of the sparsest possible structure of the equivalent network that matches the slow dynamics as close as possible. The sparsest structure denotes a model with minimum number of connectivity (edges) between the areas which may be of interest to operators for identifying the simplest possible equivalent model as well as for tracking the formation of electrical islands after any event. Islanding happens when a part of a network is isolated from the
rest of the network and swings independently with different frequencies [1]. This phenomenon can indeed be predicted using the sparse model of the slow modes that shows which edges can be removed without considerably affecting the slow dynamics of the entire network. For example, if any of the areas need to be islanded after a given contingency, the sparse model will enable the operator to achieve that by disconnecting much lesser number of lines without disturbing the inter-area oscillations between the other areas.

In general, the sparse identification of a set of parameters \( \mathbf{z} \) can be formulated as follows:

\[
\min_{\mathbf{z}} \| \mathbf{z} \|_0 \quad \text{subject to} \quad \| \mathbf{A} \mathbf{z} - \mathbf{b} \|_2 < \epsilon, \tag{24}
\]

where \( \mathbf{A}, \mathbf{b} \) and \( \mathbf{z} \) have proper dimensions, and \( \| \mathbf{z} \|_0 \) stands for the number of nonzero elements of \( \mathbf{z} \). The permissible deviation from the original dataset \( \mathbf{b} \) is determined by \( \epsilon \). This problem is NP-hard due to the term containing the \( \ell_0 \)-norm. A widely-used approach to solve this problem is to relax the \( \ell_0 \)-norm by the \( \ell_1 \)-norm. The constraint in (24) also can be replaced by a penalty term. Thus, the optimization problem of (24) can be rewritten as the following convex problem

\[
\min_{\mathbf{z}} \frac{1}{2} \| \mathbf{A} \mathbf{z} - \mathbf{b} \|_2^2 + \lambda \| \mathbf{z} \|_1. \tag{25}
\]

For our topology identification problem, we consider (25) in two different ways, namely, Centralized Sparse Identification and Decentralized Sparse Identification.

A. Centralized Sparse Identification

The identification of the sparsest structure of the equivalent network in this scenario is performed by the Independent System Operator (ISO) in a centralized way. Representative measurements from all areas are assumed to be available. The ISO will estimate all elements of the sparse \( \mathcal{L} \) at the same time using these measurements. The identification problem in this case can be formulated as:

\[
\min_{\mathcal{L}^E, \mathbf{B}_d} \frac{1}{2} \| \mathbf{X}^{m+1} - \mathbf{A}_d \mathbf{X}^m - \mathbf{B}_d \mathbf{u}^m \|_2^2 + \lambda \sum_{i,j} | \mathcal{L}^E_{ij} |, \tag{26}
\]

s.t. \( \sum_{i,j} \mathcal{L}^E_{ij} = 0 \), \( \mathcal{L}^E_{ij} = \mathcal{L}^E_{ji} \geq 0 \) \( i \neq j \), \( \mathcal{L}^E_{ii} < 0 \) \( \forall i \), where \( \mathbf{X}^m \triangleq [\mathbf{x}(0) \cdots \mathbf{x}(m-1)] \) and \( \mathbf{X}^{m+1} \triangleq [\mathbf{x}(1) \cdots \mathbf{x}(m)] \) are the concatenation of the measured states, and \( \mathbf{u}^m = [u(0) \cdots u(m-1)] \in \mathbb{R}^{1 \times m} \) is the concatenation of the input samples. \( \mathcal{L}^E \) is the Laplacian matrix of the identified graph as defined in (6a) and (6b). \( \mathbf{A}_d \) will be formed from \( \mathcal{L}^E \) in (17)–(18), assuming \( \mathcal{M}^E \) and \( \mathcal{D}^E \) are known a-priori from \( \ell_2 \) optimization of Section IV. The term \( \| \mathbf{z} \|_1 \) in (25) is replaced by \( \sum_{i,j} | \mathcal{L}^E_{ij} | \) due to the difference in the definition of norms for matrices and vectors. The listed constraints preserve the Laplacian structure of the identified \( \mathcal{L}^E \).

B. Decentralized Sparse Identification

In this scenario, the identification of the sparsest structure of the reduced-order network is performed separately by the Regional Transmission Operators (RTO) at each area. The main objective of the decentralized identification is to equip every RTO with the sense of its connection strength with only its neighbors. Thus, each operator identifies its corresponding row of the reduced-order Laplacian matrix (\( \mathcal{L}^E \)), and is required to have access to the aggregated measurement signals from all neighboring areas. Otherwise, the problem cannot be cast as a convex problem. The problem formulation for the \( i^{th} \) operator is as follows

\[
\min_{\mathcal{L}^E_{ii}, \mathbf{B}_d} \frac{1}{2} \| \mathbf{X}_{i}^{m+1} - \mathbf{A}_d \mathbf{X}^m - \mathbf{B}_d \mathbf{u}^m \|_2^2 + \lambda \sum_{j} | \mathcal{L}^E_{ij} |, \tag{26}
\]

s.t. \( \sum_{j} \mathcal{L}^E_{ij} = 0 \), \( \mathcal{L}^E_{ij} \geq 0 \) \( i \neq j \), where \( \mathcal{L}^E_{ij} \) corresponds to the \( i^{th} \) row of \( \mathcal{L}^E \). Similarly \( \mathbf{A}_d \), \( \mathbf{X}^{m+1} \), and \( \mathbf{B}_d \) correspond to the \( (i + r)^{th} \) rows of \( \mathbf{A}_d \), \( \mathbf{X}^{m+1} \), and \( \mathbf{B}_d \), respectively, where \( r \) is the number of the areas. The main difference between the resulting \( \mathcal{L}^E \) in centralized and decentralized methods is that, for the latter \( \mathcal{L}^E \) is not necessarily symmetric. This happens due to the separate identification of \( \mathcal{L}^E_{ij} \) and \( \mathcal{L}^E_{ji} \) at two different steps.

VI. SIMULATIONS RESULTS

A. IEEE 39-bus Model

1) \( \ell_2 \) Topology Identification: We consider the IEEE 39-bus model divided into four areas as shown in Fig. 1a. The model parameters are derived from [27]. The original system is modified by removing the line connecting buses 16 and 17. The generator partitioning is chosen based on the similarity of the angles and frequencies. Fig. 2 shows the generator frequencies after a three-phase fault occurs at the line connecting buses 4 and 5. The fault starts at \( t = 1 \) sec and clears at the near bus at \( t = 1.01 \) sec and at the remote bus at \( t = 1.03 \) sec. From Fig. 2, it can be seen that \( \omega_2 \) and \( \omega_3 \), \( \omega_4 \) through \( \omega_7 \), and \( \omega_8 \) through \( \omega_{10} \) are synchronized within 8-9 seconds. As a result, we group generator \( \mathcal{G}_1 \) as Area 1, \( \mathcal{G}_2 \) and \( \mathcal{G}_3 \) as Area 2, \( \mathcal{G}_4 \), \( \mathcal{G}_5 \), \( \mathcal{G}_6 \), and \( \mathcal{G}_7 \) as Area 3, and \( \mathcal{G}_8 \), \( \mathcal{G}_9 \), and \( \mathcal{G}_{10} \) as Area 4. All simulations are performed in Power System Toolbox (PST).

We consider four PMU measurements from buses 32, 33, 37, and 39 as shown in Fig. 1a. In order to apply the ERA algorithm for modal analysis of these PMU measurements, we need two prior information, namely, the total number of areas, which in this case is 4, and a reasonable range for the inter-area frequency, which is assumed to be 0.1–1 Hz. Eigenvalues of the original system are shown in Table I. The results of identification are shown in the first and second compartments of Table II. Fig. 3a and Fig. 3b show the identified topology using LLS and NLS, respectively. Figs. 5a–5c compare the
predicted outputs of the identified system using LLS and NLS with the measured outputs. It is clear that estimates of the NLS are more accurate than those of LLS. The second column of Table I also shows close matching of 3 inter-area modes with those of the original system, corresponding to four coherent areas. Another by-product of this estimation is the relative aggregate inertias of the areas. For example, if the inertia of \(G_1\) is set to 1 per unit, the second compartment of Table II indicates that the relative inertias of areas 2 through 4 are 0.2 per unit, 0.4 per unit and 0.2 per unit, respectively.

2) Identification using \(\theta(t)\): Since PMUs are often placed at terminal buses of generators, one may have the measurements of only the terminal bus phase angles \(\theta(t)\) instead of generator angles \(\delta(t)\). Fig. 4a compares \(\theta(t)\) at bus 33 and \(\delta(t)\) for \(G_4\) while Fig. 4b compares \(\theta(t)\) at bus 37 and \(\delta(t)\) for \(G_8\) for the 39-bus model. From these figures it can be seen that following an initial transient of 4 seconds, the differences between the trajectories of the angles are negligible. We, therefore, consider \(\theta(t)\) for \(4 \leq t \leq 10\) and repeat our estimation. The third compartment of Table II shows that the identified \((\mathcal{M}^E)^{-1}\mathcal{L}^E\) using this approximation is almost the same as that in the second compartment.

3) Sparse Identification: Fourth and fifth compartments of Table II show the simulation results of the sparse realization for two scenarios of centralized and decentralized estimations for the 39-bus system. Figs. 5d–5f show the respective dynamic responses of the identified model. Fig. 5c and Fig. 3d show the identified topology using these two methods. Arrows in Fig. 3d are shown due to the asymmetry of the zero elements in \((\mathcal{M}^E)^{-1}\mathcal{L}^E\). They indicate that the sparse identified topology depends on the corresponding operating node. For example, from the perspective of Area 1, it is connected to Area 2 but the reverse is not true. Figs. 3e–3h show the topology for each respective area.

4) Identification using Incomplete Set of Modes: Although modal estimation needs the incoming disturbance to be persistently exciting, in reality one may not always be able to verify that property due to which only a subset of slow modes may be identified by ERA. In that case, our method will simply result in a sparser topology for the reduced-order model. However, the eigenvalues of this sparse matrix will retain the identified modes with high residues, and some additional spurious modes with negligible residues indicating that these eigenvalues have almost no participation in the measurements. For example, if one assumes that only one pair of oscillatory modes was identified, the remaining entries of the matrix will reflect the contributions from the spurious modes. In order to confirm this, we compare the identified modes with those resulting from the original system. It turns out that the sparse identified modes result in a sparser topology for the reduced-order model. However, the eigenvalues of this sparse matrix will retain the identified modes with high residues, and some additional spurious modes with negligible residues indicating that these eigenvalues have almost no participation in the measurements.
Fig. 5. (a)-(c) Comparison of the slow components (blue) with the dynamics of the identified models using LLS (green) and NLS (red) for the 39-bus model, (d)-(f) Comparison of the slow components (blue) with the resulting model dynamics from centralized (green) and decentralized (red).

Fig. 6. (a) The 59-bus model power system, which is divided into four coherent areas shown by different colors, (b) Its graphical diagram

eigenvalues $(-0.1953 \pm 3.1281j)$ is identified by ERA in the 39-bus model, then the identified reduced-order topology is shown in the sixth compartment of Table I. It can be seen that the resulting topology is sparser than that in the second compartment of Table II. The oscillatory eigenvalues of the identified $A^E$ and their residues for $\Delta \omega^E_1$ are shown in Table III. The eigenvalues written in boldface capture the mode present in the measurements. The two spurious modes have almost zero participation.

Remark 2: It should be noted that the block diagonally dominant structure of $M^{-1}L$ in (9) is distinctly noticeable only for constant impedance loads. Our method, however, is applicable for any load as long as the swing dynamics are coherent. The 39-bus model consisting of constant power loads verifies this fact.

B. Australian 59-bus Model

To further illustrate the accuracy of our $\ell_2$ optimization method, we next consider the Australian SE 59-bus model that is divided into four areas as shown in Fig. 6a. The area partitioning is done based on the grouping algorithm described in [5]. The parameters of this model can be found in [28]. The choice behind this system follows from its well-defined radial inter-area structure which is obvious even by visual inspection. For example, the graphical connection diagram shown in Fig. 6b clearly indicates that Area 1 is in series with Area 3, while Area 4 is not connected to Area 3 at all. This fact is confirmed by the NLS estimates in Table V. We consider a three-phase fault at the line connecting buses 305 and 306. The fault starts at $t = 1$ sec, and clears at the near bus at $t = 1.01$ sec, and at the remote bus at $t = 1.03$ sec. PMU measurements from all generator buses were used for this estimation. The identification results are shown in Table V and Fig. 7. The identified topology using LLS and NLS are shown in Fig. 8.
Fig. 7. Comparison of the slow components of the measurements (blue) with the dynamics of the identified models using LLS (green) and NLS (red) for the 59-bus model

Fig. 8. Comparison of the identified topology of the reduced-order model (a) using LLS (b) using NLS for the 59-bus model

VII. CONCLUSIONS

In this paper, we developed $\ell_1$ and $\ell_2$ optimization techniques for topology identification of power system equivalent models. We showed that by extracting the slow oscillatory components of a measured PMU signal, we can identify the equivalent topology of a $r$-area network under assumptions on system connectivity and observability. The need for constructing the reduced-order network topology is not only to make decisions on specific lines and generators but also to create “monitoring metrics” that can be used for wide-area monitoring. For example, in [14] it was shown how PMU measurements can be used to construct inter-area energy functions for transient stability monitoring. The model in [14], however, was restricted to only two areas connected by a radial topology, and hence, there was no need to estimate the topology. The work done in this paper is a useful extension by which such energy functions can be estimated for any multi-area power systems connected by any arbitrary topology. One future direction of research will be to extend these optimization algorithms to cases when the operators have measurements only from their neighboring areas resulting in partial observability.

APPENDIX A

Perturbation Analysis of LLS Formulation (22)

Suppose the perturbation in $F$ and $G$ are given by $\delta F$ and $\delta G$, respectively. The upper-bound of deviation in estimation of $Y$, given by $\delta Y$ is calculated as follows:

$$Y + \delta Y = \arg\min\| (F + \delta F)Y - (G + \delta G) \|.$$  \hfill (27)

The normal equation for (27) implies:

$$(F + \delta F)^T ((F + \delta F)Y - (G + \delta G)) = 0,$$

$$\Rightarrow \delta Y \approx F^T (\delta G - \delta FY) + (F^T F)^{-1} \delta F^T r,$$ \hfill (28)

where $r \triangleq FY - G$, and $F^T \triangleq (F^T F)^{-1} F^T$. Taking the 2-norm of both sides of (28) results in:

$$\|\delta Y\|_2 \leq \|F^T\|_2 (\|\delta G\|_2 + \|\delta F\|_2 \|Y\|_2) + \|(F^T F)^{-1}\|_2 \|\delta F\|_2 \|r\|_2.$$ \hfill (29)

APPENDIX B

Perturbation Analysis of NLS Formulation (23)

Let us assume the Gauss-Newton method is chosen to solve (23) [26]. At a trial iteration $k$, we have:

$$Y^{(k+1)} = Y^{(k)} + \Delta^{(k)},$$ \hfill (30)

$$\Delta^{(k)} = -(J^{(k)})^T (J^{(k)})^{-1} (J^{(k)})^T (G - \hat{G}(Y^{(k)})),$$ \hfill (31)

where, $J^{(k)} \triangleq \partial R^{(k)} / \partial Y^{(k)} \triangleq \hat{F}^{(k)}$ and $R^{(k)} \triangleq \hat{F}^{(k)} Y^{(k)} - G$. Assuming the perturbation $\delta G$ in $G$, the upper bound of the perturbation of $\Delta^{(k)} (\delta^{(k)})$ will be calculated as

$$\Delta^{(k)} + \delta^{(k)} = -((\hat{F}^{(k)})) \hat{F}^{(k)} (G + \delta G - \hat{F}^{(k)} Y^{(k)}),$$ \hfill (32)

$$\Rightarrow \|\delta^{(k)}\|_2 \leq \|((\hat{F}^{(k)})^T \hat{F}^{(k)})^{-1} (\hat{F}^{(k)})^T \|_2 \|\delta G\|_2.$$ \hfill (33)
**APPENDIX C**

**UNIQUENESS OF TOPOLOGY IDENTIFICATION**

For the LLS problem (22), since every state of the reduced-order model is used as an output, we consider the initial values of the states to be equal to those of the measured outputs at \( t = 0 \). Also, the input \( u(k) \) is assumed to be an impulse function. With these assumptions, the solution of \( \mathcal{A}_f \) from (22) is guaranteed to be unique since the problem is convex.

The NLS problem of (23) is non-convex, but there is still a nice way to guarantee the uniqueness of the topology. This can be explained as follows. Rearranging the state vector \( \mathbf{x}(t) \) in (19) as \( \mathbf{x}'(t) \triangleq [\Delta \mathbf{E}^1(t) \omega^1(t) \ldots \Delta \mathbf{E}^p(t) \Delta \theta^p(t)] \) the state matrix for the reduced-order model (18) can be written as

\[
\mathbf{A}' = \begin{bmatrix}
0 & 1 & 0 & 0 & \ldots & 0 \\
\times & \times & \times & \ldots & & \times \\
0 & 0 & 1 & 0 & \ldots & 0 \\
\vdots & & & & \ddots & \vdots \\
0 & 0 & 0 & 0 & \ldots & 1 \\
\times & \times & \times & \ldots & & \times 
\end{bmatrix},
\]

where \( \times \) denotes the corresponding elements of \( (\mathcal{M}^E)^{-1} \mathcal{L}^E \) and \( -(\mathcal{M}^E)^{-1} \mathcal{D}^E \) blocks of \( \mathbf{A}^E \). Also, since every state of the reduced-order model (10) is measured, the output matrix

\[
\mathbf{C} = I_{2r}.
\]

Also, from (14) we have \( \text{rank}(\mathbf{S}) = r \) indicating that \( \text{rank}((\mathbf{S} \Sigma^{-1/2}) = r \). Since \( \Sigma^{-1/2} = \mathcal{S}^{-1} \), \( \text{rank}(\mathcal{S}) = r \) and the pair \( (\mathbf{A}_f, \mathbf{B}_f) \) is, therefore, controllable. Since \( (\mathbf{A}^E, \mathbf{B}^E) \) is a similarity transform of the continuous-time counterpart of \( (\mathbf{A}_f, \mathbf{B}_f) \), the pair \( (\mathbf{A}^E, \mathbf{B}^E) \) is also controllable. Given the structure of \( \mathbf{A}', \mathbf{C} = I_{2r} \) and controllability of the pair \( (\mathbf{A}^E, \mathbf{B}^E) \), from [29, Theorem 4A.2], the \( \times \) elements of \( \mathbf{A}' \) are guaranteed to be uniquely identifiable.

**REFERENCES**


Seyedehbazzar Nabhari received the B.S. degree from Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, in 2009 and the M.S. degree from North Carolina State University (NCSU), Raleigh, NC, in 2011, both in Electrical Engineering. He is pursuing his Ph.D. in Electrical Engineering in NCSU. His research interests are the applications of control theory in analysis of large-scale power systems.

Aranya Chakrabortty received his B.E. degree from Jadavpur University, India in 2004, and his M.S. and Ph.D. degrees from Rensselaer Polytechnic Institute, Troy, NY in 2005 and 2008, respectively, all in Electrical Engineering. He is currently an Assistant Professor in the Electrical and Computer Engineering department of North Carolina State University, Raleigh, NC.