Joint Topology Identification and State Estimation in Unobservable Distribution Grids

Hazhar Sufi Karimi, *Member, IEEE,* and Balasubramaniam Natarajan, *Senior Member, IEEE*

***Abstract*—Many distribution system operations (e.g., state estimation, control, fault detection/localization) rely on the assumption that the underlying topology is accurately defined. In general, topology identification is a challenging problem in distribution systems as these systems are unobservable with a very limited number of available measurements. In this paper, we tackle this problem by designing a compressive sensing framework that jointly estimates the systems states and network topology via an integrated mixed integer nonlinear program (MINLP) formulation. Two reformulations of the original MINLP problems are investigated. Firstly, in order to remove the nonlinearity in the MINLP formulation, a mixed integer linear programming (MILP) problem that employs auxiliary variables is derived. Furthermore, to achieve a faster solution, convex relaxation of the original formulation is derived. Finally, using a Markovian model for topology changes, prior information about system topology is used to improve topology identification particularly when a limited amount of measurements is available. Simulation results on IEEE 37-bus test feeder and IEEE 123-bus test feeder illustrate the efficiency and scalability of the proposed approaches from both state estimation and topology identification point of view (even with 30% of available data).**

***Index Terms*—Topology identification, state estimation, compressive sensing, distribution grids, unobservability**

I. Introduction

DISTRIBUTION system topology identification has become a popular research topic due to its critical role in state estimation, Volt-VAR control, fault detection, demand response, etc. [1]. Many of the proposed estimation, control and optimization solutions in distribution systems are not feasible if the underlying grid topology is not accurately identified [2]. Furthermore, the time-varying nature of distribution systems implies that the topology identification process must be frequently performed if relowiable results are expected. However, timely topology monitoring may not be completely possible since the required number of measurements for existing topology identification approaches are either unavailable (or unreliable) due to limited metering infrastructure and the sheer scale of distribution systems [3], [4]. Therefore, this paper tries to address the topology identification problem by integrating it with state estimation for highly unobservable distribution systems.

A. Related Work

The existing methods for topology identification can be classified into different categories. Traditional approaches employ all the available measurements for state estimation based on every possible topology configuration. Then, the topology that results in the minimum residual state estimation error is selected [5]. Correct configuration of distribution grid is identified in [6] based on a bank of critical network configurations. A notable drawback of these techniques is the high computational complexity that grows with the possible topologies. Instead of considering all possible topologies, another class of techniques focuses on the simultaneous estimation of switch statuses and system states [7] and [8]. Simultaneous estimates of state variables and switching device statuses are provided in [9] using both real and pseudo measurements as well as operational constraints for the switching devices. Recently, [10] develops a mathematical programming technique for joint estimation of operational topology and outages in unbalanced power distribution systems. Although, these methods could alleviate the burden of computational complexity, they require a large number of measurements. For topology error identification, [11] proposes a Bayesian hypothesis testing procedure in a generalized state estimation framework that eliminates the need for repeated state estimator runs for alternative hypothesis evaluation. For a low-voltage distribution network, [12] proposes a data-driven approach to identify the underlying network topology using time series of energy measurements. In [13], a mixed integer quadratic programming (MIQP) based topology identification model is proposed, that is suitable for radially operated distribution grids. To detect network topology, [14] presents a time-series signature verification approach which is based on the analysis of synchronized voltage phasor data. While most of the topology identification techniques rely on voltage and power information, [2] provides a new topology identification algorithm that is based on measurements from a few line current sensors, together with available pseudo-measurements. To continuously monitor the underlying topology of a smart distribution network, real-time topology identification has been targeted in some of the recent works [3], [15]. [16] proposes a data-driven technique for joint identification and parameter estimation. Using least absolute shrinkage and selection operator (LASSO) method, [17] proposes a joint estimation of admittance parameters and topology of a multi-phase distribution network. [18] develops algorithms to learn the structure of a radial distribution grid while [19] employs statistical learning methods to identify the topology of meshed distribution systems. Learning algorithms are devised for topology and parameter estimation using balanced voltage and injection measured only at the end-users [20], [21]. In [22], the authors provide algorithms to estimate the operational topology and injection statistics of missing nodes for radial distribution grids. Using non-synchronized voltage

Hazhar Sufi Karimi and Balasubramaniam Natarajan are with the Department of Electrical and Computer Engineering, Kansas State University, Manhattan, KS, 66502 USA e-mail: hazhars@ksu.edu , bala@ksu.edu.

data, [23] designs a statistical learning framework forth for verifying single-phase grid structures. Another category of the topology identification approaches employ graphical models to describe the probabilistic relationship among different voltage measurements [24]-[26]. It should be mentioned that most of these techniques rely on observability (i.e., having access to adequate number of measurements relative to the number of states) of the distribution systems. However, since the observability requirement may not always be fulfilled in distribution grids, most of the topology identification approaches can not be practically implemented.

Similar to topology identification techniques, most of the existing methods for distribution system state estimation (DSSE) are challenged by a lack of observability [27]. Although, an increase in the number of grid edge sensors and smart meters can cope with the unobservability issue, the large volume of the aggregated data imposes inevitable challenges to the underlying communication network and the data processing center [28]. Recently, compressive sensing-based DSSE techniques have been proposed to deal with the issues of unobservability and minimal data access [29]-[32]. The state estimation processes based on compressive sensing employ power flow equations that rely on the identified topology. Since the compressive sensing techniques are typically implemented for undersampled systems, it is likely that the required topology knowledge is impaired due to lack of observability. Consequently, the compressive sensing based state estimation fails to provide the correct states since the underlying topology is not accurately identified. On the other hand, the topology identification methods are not functional due to the lack of accurate state information. Therefore, our goal is to simultaneously estimate the states and the network topology with limited available measurements. Distribution system operators usually have access to information about physical distances among the buses, lines connecting buses and their parameters. Therefore, the topology relevant to the fixed lines is always known. However, in addition to the fixed lines that are always energized, there are lines with switches that establish a connection between the buses based on their statuses (open or closed). Therefore, topology identification can be mapped to the problem of switch status detection.

B. Contributions

The main contributions of this paper are summarized below:

- To simultaneously estimate the states and topology in an unobservable distribution grid, we first formulate an optimization problem with states of system and switch statuses as decision variables. Our framework projects the switch statuses on the underlying admittance matrix. Solution of this optimization problem return the estimated states and switch statuses. The proposed formulation contains nonlinear relationships and integer variables leading to a mixed integer nonlinear programming (MINLP) problem.
- In order to guarantee the solvability of the optimization formulations and reduce MINLP complexity, we modify the nonlinear constraints to yield a mixed integer linear programming (MILP) problem. To this end, we transform the existing nonlinearity by introducing auxiliary variables.
- To cope with the non-convexity of the original MINLP formulation and to reduce computational complexity, we relax the nonlinearity of MINLP by replacing integer variables with continuous decision variables. Then, we apply an alternative minimization approach that improves both state estimation and topology identification performance. Simulation results on IEEE 37-bus test feeder and IEEE 123-bus test feeder shows the efficiency of the proposed approach even with 30% of available data.
- Finally, a hybrid dynamic framework is introduced that incorporates prior information about system topology for current state/switch status estimation. More precisely, a Markov jump model is proposed for switch status that helps topology identification especially for high levels of compression. Besides, we employ the previous information about the support set that improves the fidelity of our method. Simulation results provided in section V demonstrate the performance of the proposed approaches on the IEEE 37-bus test feeder and IEEE 123-bus test feeder with practical load/generation data.

C. Paper Organization

The rest of this paper is organized as follows. Section II describes the system model of interest. In section III, we introduce the compressive sensing formulation for joint topology identification and state estimation. Then, we propose two methods to alleviate the complexities of the original MINLP problem. Section IV presents dynamical joint state estimation and topology identification in distributions systems. Section V provides numerical simulations that demonstrate the performance of the introduced methods.

II. System Model

In distribution grids, the conventional state estimation methods or topology identification techniques need to aggregate all the available sensor/smart meter measurements. This requirement implies that a large volume of measurements must be gathered and transmitted over a communication network. To this end, a communication network with large bandwidth and high reliability should be established. However, providing such a communication infrastructure is not feasible or affordable in most of the situations. Furthermore, the sensors may not be reliable. Therefore, our goal is to alleviate the burden on the communication network by reducing the required number of measurements for topology identification and state estimation while maintaining a reasonable fidelity in estimating both.

A. Background

In our earlier work , compressive sensing theory has been shown to be effective in recovering the distribution system states from a smaller number of random measurements [29], [33]. The performance of sparsity-based DSSE techniques rely on accurate power flow models which in turn require

precise knowledge about the system, including the system topology, parameters, power injection data, etc.. For a three-phase distribution system with $\mathcal{J}$ buses, let $\mathbf{z}$ denote the n-dimensional state vector of interest ($n = 9\mathcal{J}$), where $\mathbf{z} = \begin{bmatrix}\mathbf{p}^{'} & \mathbf{q}^{'} & \mathbf{v}^{'}\end{bmatrix}^{'}$. Here, $\mathbf{p} \in \mathbb{R}^{3\mathcal{J}}$, $\mathbf{q} \in \mathbb{R}^{3\mathcal{J}}$ and $\mathbf{v} \in \mathbb{C}^{3\mathcal{J}}$ are the real power, reactive power and complex voltage vectors, respectively. These vectors serve as the three phase states at every node in an unbalanced three phase distribution grid. Our goal is to recover $\mathbf{z}$ from an m-dimensional observation vector $\mathbf{y} = \mathbf{\Phi z} + \boldsymbol{\nu}$, where $\mathbf{\Phi}$ is an $m \times n$ measurement matrix with $m < n$; $\boldsymbol{\nu}$ represents the measurement noise vector. It should be noted that the linearity of measurement matrix $\mathbf{\Phi}$ does not indicate that the power and voltage values can be mixed as a measurement. In other words, all the elements in the measurement vector are either a function of power values or voltage values. To construct such a decoupled mechanism, the following should hold:

$$\mathbf{\Phi} = \begin{bmatrix}\mathbf{\Phi^p} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi^q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{\Phi^v}\end{bmatrix}$$

where, $\mathbf{\Phi^p}$, $\mathbf{\Phi^q}$ and $\mathbf{\Phi^v}$ are the measurement matrices correspond to real power, reactive power and voltage, respectively. As comprehensively discussed in [34] and [35], the state vector of interest $\mathbf{z}$ can be converted to a sparse vector using a known basis $\mathbf{\Psi}_{n\times n}$ (e.g., wavelet or Discrete Cosine Transform (DCT)). This implies $\mathbf{x} = \mathbf{\Psi z}$, is a sparse vector, (i.e, only S elements of $\mathbf{x}$ have non-zero values, where $S << n$). Using practical data, [36] shows that one can expect a certain level of correlation between power characteristics of the nodes in a distribution grid due to the geographical proximity. Such a correlation between the elements of the underlying signal suggests that the signal can have a concise representation when expressed in the proper basis $\mathbf{\Psi}$. It should be noted that choosing an appropriate transformation matrix $\mathbf{\Psi}$ plays an important role in the performance of compressive sensing problem. [35] and [37] demonstrate the effect of different transformation/sparsifying matrices on accuracy of power state reconstruction. Based on this transformation/sparsification, the measurement model is written as:

$$\mathbf{y} = \mathbf{Cx} + \boldsymbol{\nu}, \tag{1}$$

where, $\mathbf{C} = \mathbf{\Phi\Psi}; \mathbf{C} \in \mathbb{R}^{m\times n}$. To quantify the level of the measurement compression, we use the compression measurement ratio (CMR) that is defined as: $CMR = \frac{m}{n} \times 100$; i.e., a small CMR indicates a fewer number of available measurements. Also, let N denote the support set of the sparse state $\mathbf{x}$. That is, $N = \{i_1, i_2, ..., i_S\}$ where i_k denotes the non-zero coordinates of $\mathbf{x}$ in the $\mathbf{\Phi}$ space and $S = size(N)$. Here, N is time-varying and may not be completely known.

B. Compressive Sensing

If the transformed signal $\mathbf{x}$ is adequately sparse, compressive sensing theory allows us to estimate the sparse vector $\mathbf{x}$ using the available measurements $\mathbf{y}$ [33]. Obviously, the state of interest $\mathbf{z}$ can be straightforwardly recovered as $\hat{\mathbf{z}} = \mathbf{\Phi}\hat{\mathbf{x}}$, where $\hat{\mathbf{x}}$ is the estimate of $\mathbf{x}$. Direct compressive sensing based state estimation method is based on the following optimization problem [29]:

$$\begin{aligned}\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \parallel \mathbf{x} \parallel_1 + \lambda \parallel \mathbf{y} - \mathbf{\Phi\Psi x} \parallel_2 \\ \text{subject to :} \\ \mathbf{v} = f_{PF}(\mathbf{p}, \mathbf{q})\end{aligned} \tag{2}$$

where, f_{PF} represents the nonlinear power flow equations that captures the relationship between voltage states and power injected at the nodes. It should be mentioned that $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{v}$ are not new variables as the decision vector $\mathbf{x}$ is a linear transformation of $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{v}$. Nonlinearity of the AC power-flow equations poses significant challenges for the development of computationally efficient state estimation solutions. To avoid the associated computational complexity, we employ a linearized approximation of the power flow constraint that is applicable to multiphase distribution networks [38]. For a three-phase distribution system with $\mathcal{J}$ buses, let $\mathbf{Y}$ denote the admittance matrix: $\mathbf{Y} = \begin{bmatrix}\mathbf{Y}_{00} & \mathbf{Y}_{0L} \\ \mathbf{Y}_{L0} & \mathbf{Y}_{LL}\end{bmatrix}$ where, $\mathbf{Y}_{00} \in \mathbb{C}^{3\times3}$, $\mathbf{Y}_{L0} \in \mathbb{C}^{3(\mathcal{J}-1)\times3}$, $\mathbf{Y}_{0L} \in \mathbb{C}^{3\times3(\mathcal{J}-1)}$ and $\mathbf{Y}_{LL} \in \mathbb{C}^{3(\mathcal{J}-1)\times3(\mathcal{J}-1)}$ are submatrices of the admittance matrix. Therefore, the linearized voltage $\mathbf{v}$ can be expressed as follow:

$$\mathbf{v} \approx \mathbf{M}\begin{bmatrix}\mathbf{p} \\ \mathbf{q}\end{bmatrix} + \mathbf{w} \tag{3}$$

where,

$$\mathbf{M} = \begin{bmatrix}\mathbf{Y}_{LL}^{-1} diag(\tilde{\hat{\mathbf{v}}})^{-1} & -j\mathbf{Y}_{LL}^{-1} diag(\tilde{\hat{\mathbf{v}}})^{-1}\end{bmatrix} \tag{4}$$

$\mathbf{w} = -\mathbf{Y}_{LL}^{-1}\mathbf{Y}_{L0}\mathbf{v}_0$. Here, $\mathbf{v}_0$ is the slack bus voltage where the voltage level is assumed to be fixed or regulated (typically 1 per unit). This paper utilizes the flat voltage profile. $\hat{\mathbf{v}}$ denotes voltage associated with an operating point. Our prior works illustrate that state estimation based on compressive sensing and matrix/tensor completion is effective while employing the linearized power flow constraints [29], [30]. Other researchers have also confirmed the efficiency of using this linearization method for state estimation [32].

C. Topology

If the topology is correctly identified along with accurate data and parameters, one can employ different power flow methods for estimation/control purposes. However, topology identification is a challenging task as most of the existing distribution systems are equipped with switches and breakers that result in various topology configurations based on the status of the switches/breakers. Figure. 1 demonstrates IEEE 37-bus test feeder as an example. Here, the system is equipped with six switches. Node 1 serves as slack bus and the source of power. Let us consider node 27. In the normal situation, s_1 is closed and s_2 is open, therefore, node 27 is energized from node 3. Another possibility is when switch s_1 is open and s_2 is closed. This implies that node 27 is fed by node 8. Also, both s_1 and s_2 can be closed or open, simultaneously. Consequently, different voltage profiles will result based on the flow of power in the network. It should be noted that this example does not imply that a disconnected or meshed

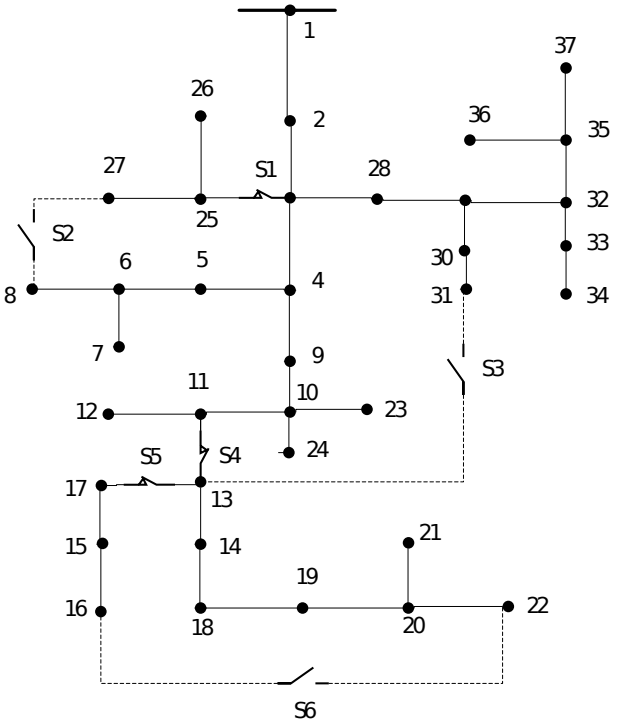

Fig. 1. IEEE 37-bus test feeder with six switches

structure is never allowed. The proposed approach is general and can accommodate any underlying topological structure. Impact of such a switching configuration is captured in the underlying admittance matrix. In other words, if we are able to identify the correct admittance matrix, we can determine the exact topology of the system. However, the number of possible topologies dramatically increases with the number of switches, i.e., if there are d switches in the network, 2^d different topology configurations are possible. Such large numbers of different possibilities impose a big load on the required computations; or they can lead to inaccurate and unreliable results. Therefore, estimation of switch statuses is a reasonable and scalable approach for topology identification [2], [8]. To this end, we write the admittance matrix as a function of the status of the switches as

$$\mathbf{Y}_{LL} = \bar{\mathbf{Y}} + s_1\mathbf{Y}_1 + s_2\mathbf{Y}_2 + ... + s_d\mathbf{Y}_d, \tag{5}$$

where $\bar{\mathbf{Y}}$ is the admittance matrix between the nodes with fixed permanent connection (i.e., the lines with no switch/breaker). $s_i \in \{0, 1\}$ denote the status of the i^{th} switch. If the switch s_1 is closed, then the corresponding line is activated. Mathematically, this implies that $s_1 = 1$. Therefore, the submatrix $\mathbf{Y}_1$ (the admittance matrix between the two nodes that are connected by s_1) is added to the general admittance matrix $\mathbf{Y}_{LL}$; otherwise, $\mathbf{Y}_{LL}$ does not contain submatrix $\mathbf{Y}_1$.

As it was mentioned earlier, this paper aims at estimating the state vector $\mathbf{z} = \begin{bmatrix}\mathbf{p}^T & \mathbf{q}^T & \mathbf{v}^T\end{bmatrix}^T$. Here, we create vector $\mathbf{S} = \begin{bmatrix}\mathbf{p}^T & \mathbf{q}^T\end{bmatrix}^T$ containing the real and reactive power values. The measurement mechanism is such that:

$$\begin{aligned}\mathbf{y}^{\mathbf{S}} &= \mathbf{\Phi}^{\mathbf{S}}\mathbf{S} + \boldsymbol{\nu}^{\mathbf{S}} \\ \mathbf{y}^{\mathbf{V}} &= \mathbf{\Phi}^{\mathbf{V}}\mathbf{v} + \boldsymbol{\nu}^{\mathbf{V}}\end{aligned} \tag{6}$$

where, $\mathbf{y}^{\mathbf{S}} \in \mathbb{R}^{m_{\mathbf{S}}}$ and $\mathbf{y}^{\mathbf{V}} \in \mathbb{C}^{m_{\mathbf{V}}}$ are the power measurements and voltage measurement, respectively. Here, $\boldsymbol{\nu}^{\mathbf{S}}$ and $\boldsymbol{\nu}^{\mathbf{V}}$ are the corresponding measurement noise vectors. $\mathbf{\Phi}^{\mathbf{S}} \in \mathbb{R}^{m_{\mathbf{S}} \times 6\mathcal{J}}$ and $\mathbf{\Phi}^{\mathbf{V}} \in \mathbb{R}^{m_{\mathbf{V}} \times 3\mathcal{J}}$ are measurement/projection matrices. These matrices are determined by the underlying measurement mechanisms (e.g., matrix elements distributed as i.i.d. Gaussian random variables with mean 0 and variance $1/m_{\mathbf{S}}$ or Bernoulli random variables). It should be noted that the measurement mechanism (i.e. $\mathbf{\Phi}^{\mathbf{S}}$ and $\mathbf{\Phi}^{\mathbf{V}}$) is not impacted by changes in topology of the grid.

III. Proposed Approach

The necessary condition for implementing compressive sensing approaches is that the state of interest exhibits sparsity or must be approximately sparse in a linear transformation basis. Let the state of interest $\mathbf{z} \in \mathbb{C}^{9\mathcal{J}}$ be a sparse vector in a linear transformation basis $\mathbf{\Psi} \in \mathbb{R}^{9\mathcal{J} \times 9\mathcal{J}}$ such that,

$$\mathbf{z} = \mathbf{\Psi}\mathbf{x} = \begin{bmatrix}\mathbf{\Psi}^{\mathbf{S}} & \mathbf{0}_{6\mathcal{J}\times 3\mathcal{J}} \\ \mathbf{0}_{3\mathcal{J}\times 6\mathcal{J}} & \mathbf{\Psi}^{\mathbf{V}}\end{bmatrix}\mathbf{x}$$

where, $\mathbf{x}$ has at most $S \ll 9\mathcal{J}$ significant coefficients i.e., $\mathbf{z}$ is S-sparse in the sparsifying basis $\mathbf{\Psi}$. Previous literature has validated the sparsity assumption for distribution systems [34], [39]. Along with the sparsity condition, the linearized relationship (3) between power and voltage serves as the link between the states and the network topology. For example, assume that none of the buses in the IEEE 37-bus test feeder have generation and all of them are energized through the slack bus 1. If switch s_1 is closed and s_2 is open, the voltage is larger than the case corresponding to s_1 being open and s_2 is closed. Therefore, (3) captures the relationship between the topology and states where the status of switches impact the states of the grid. Mathematically, (3) employs the matrix $\mathbf{M}$ and the bias term $\mathbf{w}$ to provide a linear relationship between states of the system (i.e., voltage and power). As indicated by equation (4), both $\mathbf{M}$ and $\mathbf{w}$ contain the admittance matrix $\mathbf{Y}$ which is determined by topology identification. However, both $\mathbf{M}$ and $\mathbf{w}$ include inverse of the admittance matrix which complicates the optimization problem since admittance matrix $\mathbf{Y}$ is influenced by the switch status and additional terms. To avoid this complexity, we multiply equation (4) by $\mathbf{Y}_{LL}$. Consequently, equation (7) represents the linearized power flow constraint in our optimization problem:

$$\mathbf{Y}_{LL}\mathbf{v} + \mathbf{Y}_{L0}\mathbf{v}_0 - \mathbf{M}_{yl}\mathbf{S} = 0 \tag{7}$$

where, $\mathbf{M}_{yl} = \begin{bmatrix}diag(\tilde{\mathbf{v}})^{-1} & -jdiag(\tilde{\mathbf{v}})^{-1}\end{bmatrix}$.

It should be noted that in many practical situations, opening/closing of a switch may depend on status of other switches. For example, let us consider the IEEE 37-bus test feeder in Figure 1. If switch s_1 (normally closed) is opened, nodes 25, 26 and 27 are not energized anymore. Obviously, to cope with this outage, switch s_2 must be closed, where the islanded nodes (25-27) can be connected to the grid and obtain their energy through node 8. To enhance resiliency in distribution systems, most of the utilities equip the distribution grids with line switches [40], [41]. Definitely, in such grids, the switch statuses are not completely independent. Therefore, we

introduce a set of constraints that captures the relationship between the switches as follow:

$$\mathcal{G}(s_1, s_2, ..., s_d) = \mathcal{C} \tag{8}$$

where, $\mathcal{C}$ is a vector with constant integer variables. For example, if the nodes (25-27) must be always energized and no mesh structure is allowed in the system, the constraint (8) is written as, $s_1 + s_2 = 1$. In other words, (8) enables us to enforce the desired structure of the network.

Using the information about the system, we setup the minimization (9) to jointly reconstruct the sparse state $\mathbf{x}$ and status of the switches $\mathbf{s} = [s_1, ..., s_d]^{'}$,

$$\begin{aligned} \underset{\mathbf{x}, s_i}{\operatorname{argmin}} \parallel \mathbf{x} \parallel_1 + \lambda^{\mathbf{S}} \parallel \mathbf{y}^{\mathbf{S}} - \mathbf{\Phi}^{\mathbf{S}} \mathbf{\Psi}^{\mathbf{S}} \mathbf{x}^{\mathbf{S}} \parallel_2 + \\ \lambda^{\mathbf{V}} \parallel \mathbf{y}^{\mathbf{V}} - \mathbf{\Phi}^{\mathbf{V}} \mathbf{\Psi}^{\mathbf{V}} \mathbf{x}^{\mathbf{V}} \parallel_2 \end{aligned} \tag{9}$$

subject to

$$\begin{aligned} (\bar{\mathbf{Y}} + s_1 \mathbf{Y}_1 + ... + s_d \mathbf{Y}_d,) \mathbf{\Psi}^{\mathbf{V}} \mathbf{x}^{\mathbf{V}} + \\ \mathbf{Y}_{L0} \mathbf{v}_0 - \mathbf{M}_{yl} \mathbf{\Psi}^{\mathbf{S}} \mathbf{x}^{\mathbf{S}} = 0 \\ \mathcal{G}(\mathbf{s}) = \mathcal{C} \end{aligned} \tag{10}$$

where, $\mathbf{x} = \begin{bmatrix} \mathbf{x}^{\mathbf{S}} \\ \mathbf{x}^{\mathbf{V}} \end{bmatrix}$ with $\mathbf{x}^{\mathbf{S}} \in \mathbb{R}^{6\mathcal{J}}$ and $\mathbf{x}^{\mathbf{V}} \in \mathbb{C}^{3\mathcal{J}}$.

Obviously, to guarantee a certain state estimation fidelity and accurate topology identification, the number of measurements can not be arbitrarily reduced (i.e., if we excessively reduce $m_{\mathbf{S}}$ and $m_{\mathbf{V}}$, both topology identification and state estimation may fail to perform accurately).

The introduced minimization problem (9) is nonlinear and nonconvex due to the multiplication of s_i and $\mathbf{v}$. Moreover, since the decision variables s_i are binary, the minimization problem (9) is a mixed integer problem. Therefore, our proposed compressive sensing based joint state and topology estimation approach results in a mixed integer nonlinear program (MINLP). In the remainder of this section, we aim to alleviate the complexity of the proposed optimization scheme by converting the problem into sub-optimal but less complex optimization problems.

A. MILP Formulation

Although, both MILP and MINLP are considered as NP-hard problems [42] and they share similar computational complexity behaviors, the availability of more advanced solvers motivates us to transform our MINLP formulation to MILP. To this end, the nonlinearity of (7) needs to be removed. Therefore, we consider an auxiliary variable $\mathbf{U}_i = s_i \mathbf{v}$. Now, if the corresponding switch is closed (i.e., $s_i = 1$) we have $\mathbf{U}_i = \mathbf{v}$, otherwise $\mathbf{U}_i = 0$. This relationship can be formulated using the linear constraints:

$$-(1 - s_i)F \leq \mathbf{U}_i - \mathbf{v} \leq (1 - s_i)F \tag{11}$$

$$-F s_i \leq \mathbf{U}_i \leq F s_i \tag{12}$$

Here, F should be large enough to impose the correct constraints. If $s_i = 0$, (11) is not binding and (12) enforces $\mathbf{U}_i$ to be 0. Similarly, if $s_i = 1$, (11) enforces $\mathbf{U}_i = \mathbf{v}$ while the bound (12) is loose to allow $\mathbf{U}_i$ to have flexibility. Algorithm 1 summarizes the procedure of MILP based topology identification and state estimation.

Algorithm 1 MILP Compressive Sensing based Joint Topology Identification and State Estimation

Step 1. Collect the compressed measurements into two vectors $\mathbf{y}^{\mathbf{S}}$ and $\mathbf{y}^{\mathbf{V}}$.

Step 2. Solve the following minimization problem

$$\begin{aligned} \hat{\mathbf{x}}, \hat{s}_i = \underset{\mathbf{x}, s_i}{\operatorname{argmin}} \parallel \mathbf{x} \parallel_1 + \lambda^{\mathbf{S}} \parallel \mathbf{y}^{\mathbf{S}} - \mathbf{\Phi}^{\mathbf{S}} \mathbf{\Psi}^{\mathbf{S}} \mathbf{x}^{\mathbf{S}} \parallel_2 + \\ \lambda^{\mathbf{V}} \parallel \mathbf{y}^{\mathbf{V}} - \mathbf{\Phi}^{\mathbf{V}} \mathbf{\Psi}^{\mathbf{V}} \mathbf{x}^{\mathbf{V}} \parallel_2 \end{aligned} \tag{13}$$

subject to

$$\begin{aligned} (\bar{\mathbf{Y}} + s_1 \mathbf{Y}_1 + ... + s_d \mathbf{Y}_d,) \mathbf{\Psi}^{\mathbf{V}} \mathbf{x}^{\mathbf{V}} + \\ \mathbf{Y}_{L0} \mathbf{v}_0 - \mathbf{M}_{yl} \mathbf{\Psi}^{\mathbf{S}} \mathbf{x}^{\mathbf{S}} = 0 \\ -(1 - s_i)F \leq \mathbf{U}_i - \mathbf{\Psi}^{\mathbf{V}} \mathbf{x}^{\mathbf{V}} \leq (1 - s_i)F \\ -F s_i \leq \mathbf{U}_i \leq F s_i \\ \mathcal{G}(s_1, s_2, ..., s_d) = \mathcal{C} \end{aligned} \tag{14}$$

Step 3. The original state vector is: $\hat{\mathbf{z}} = \mathbf{\Psi} \hat{\mathbf{x}}$, and the estimation of switch status is: $\hat{s}_i,\ i \in \{1, 2, ..., k\}$

B. Convex Relaxation

Both the MINLP and MILP formulation are nonconvex problems due to integer decision variables. To relax the nonconvexity, the decision variables must be chosen from a convex set. Therefore, we replace the integer variables with continuous variables. Assume that a continues variable represents the switch status (i.e., $(\bar{s}_i = s_i)$), where:

$$0 \leq \bar{s}_i \leq 1 \tag{15}$$

In the next step, we impose the power flow constraint in a second norm form, since the strict equality condition 7 can not guarantee any convexity. Furthermore, the equality may not always hold due to the linearization error.

$$||\mathbf{Y}_{LL} \mathbf{\Psi}^{\mathbf{V}} \mathbf{x}^{\mathbf{V}} + \mathbf{Y}_{L0} \mathbf{v}_0 - \mathbf{M}_{yl} \mathbf{\Psi}^{\mathbf{S}} \mathbf{x}^{\mathbf{S}}||_2 \leq \epsilon \tag{16}$$

where, ϵ can be tuned based on linearization error. Obviously, the constraint (16) is convex, since the norm function is a convex function defined on a convex set. However, the constraint (16) is still nonlinear. To tackle the nonlinearity, we employ an alternating minimization technique. In the alternating minimization method, we consider the switch statuses $\bar{s}_i$ as constant values and solve the minimization problem to find the voltage values. Then, the power and voltage values are held constant while we seek an optimal solution with switch statuses as decision variables.

Algorithm 2 summarizes the convex technique for joint estimation of states and switch statuses. In the first step, we arrange the power and voltage measurements into separate vectors. Then, we initialize the continuous switch statuses to be 0.5 in Step 2. This initialization policy implies that we have no idea or prior information about status of the switches.

Then, we start the iteration process. In Step 3, we find the states by solving the minimization problem while the $\hat{\bar{s}}_i^{j-1}$ are constant. Using the estimated states, we find the switch statuses in Step 4. Step 5 indicates the stopping criterion. It should be noted that the estimated switch status in Step 4 is a continuous value (i.e. a value between 0 and 1). However, a switch status should be an integer value. Therefore, Step 6 introduce the rounding of the soft switch status values to integer values. The estimated power/voltage is equal to the values calculated at the last iteration.

Algorithm 2 Convex Compressive Sensing based Joint Topology Identification and State Estimation

Step 1. Collect the compressed measurements into two vectors $\mathbf{y^S}$ and $\mathbf{y^V}$.
Step 2. Initialize the parameters $\hat{\bar{s}}_i^0 = 0.5 \quad \forall i \in \{1, 2, ..., k\}$; $\hat{\mathbf{v}}_i^0 = 1\angle 0$.
Set $j = 1$ and do:
Step 3. Fix the switch status by using $\hat{\bar{s}}_i^{j-1}$. Then estimate the voltage $\hat{\mathbf{V}}^j$ using:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \parallel \mathbf{x} \parallel_1 + \lambda^{\mathbf{S}} \parallel \mathbf{y^S} - \mathbf{\Phi^S \Psi^S x^S} \parallel_2 + \lambda^{\mathbf{V}} \parallel \mathbf{y^V} - \mathbf{\Phi^V \Psi^V x^V} \parallel_2 \qquad (17)$$

$$\text{subject to } (8), (16)$$

Step 4. In the following minimization, implement the voltage vector $\hat{\mathbf{V}}^j$ as fixed values and estimate switch status $\hat{\bar{s}}_i^j$ by solving:

$$\hat{\mathcal{B}} = \underset{\mathcal{B}}{\operatorname{argmin}} \parallel \mathcal{B} \parallel_1 + \lambda^{\mathbf{S}} \parallel \mathbf{y^S} - \mathbf{\Phi^S \Psi^S} \mathcal{B} \parallel_2$$

$$\text{subject to } (8), (15), (16)$$

Step 5. If $||\hat{\mathbf{v}}^j - \hat{\mathbf{v}}^{j-1}|| < \delta$ holds, $\hat{\mathbf{z}} = \mathbf{\Psi}\hat{\mathbf{x}}$ and proceed to Step 6.
Otherwise, $j \leftarrow j + 1$ and return to Step 3.
Step 6. Round the continuous switch status to 0 and 1:
$\hat{s}_i = \begin{cases} 1 & \hat{\bar{s}}_i \geq 0.5 \\ 0 & \hat{\bar{s}}_i < 0.5 \end{cases}$

It has been proved that the alternating minimization technique converges if the objective function and the constraints are convex and differentiable in their domain [43]. After convex relaxation, our framework completely fulfills this condition, which guarantees the convergence of Algorithm 2.

IV. Dynamic Hybrid Estimation

In this section, we focus on dynamic state estimation as well as real-time topology identification. Here, the possible topologies can be interpreted as discrete modes. That is, if the system operates in mode q, there is a unique topology relevant to that mode. Therefore, the terms "mode estimation" and "topology identification" can be used, interchangeably. In dealing with dynamical estimation of continuous states with discrete modes (hybrid systems), Interacting Multiple Model (IMM) is one of the popular techniques that combines all the estimates derived based on the possible underlying models [44]. However, implementing such a criterion for our problem may not be possible due to the large number of discrete modes. For example, let us consider the d switches system that leads to 2^d discrete modes. Obviously, real-time estimation of 2^d systems requires significant computational resources, particularly for larger values of d (number of switches). Furthermore, IMM algorithms estimate the probability of discrete modes. In this case, we have 2^d probability values that could be similar. Therefore, distinguishing the true discrete mode (topology) is not a straightforward task.

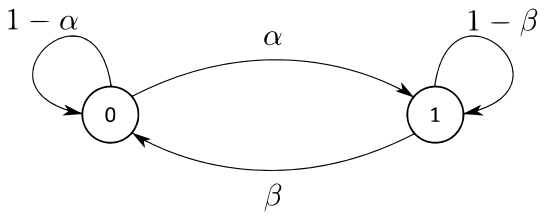

Fig. 2. Discrete time Markov chain model for switch status.

Similar to the small changes in the sparse state and its support [39], the switch statuses may be correlated across time. Here, we assume that we only know the probability of the switch status based on the previous statuses. To quantify this assumption, we consider the transition probability matrix denoted by (18). $s_{i,t}$ denotes the i^{th} switch status at time t, respectively. Here, $s_{i,t}$ can be considered as a discrete-time two-state Markov chain as demonstrated by Figure 2. This model is considered with prior work that captures the changes in a distribution system using a hierarchical Markov model [45]. The transition probability rates satisfy $0 \leq \alpha_i,\ \beta_i \leq 1$ and the transition matrix for this discrete time chain corresponds to:

$$\boldsymbol{\pi}_i = \begin{bmatrix} 1-\alpha_i & \alpha_i \\ \beta_i & 1-\beta_i \end{bmatrix}, \qquad (18)$$

where, $(\boldsymbol{\pi}_i)_{jk} = \mathbf{Pr}\big(s_{i,t} = k-1 | s_{i,t-1} = j-1\big)$. In practice, the transition probability rates (α and β) are identified using historical data from the distribution grid.

Remark. *The transition probability matrix* (18) *suggests that probabilities of switch status changes are completely independent. However, as previously stated by* (8)*, the switch statuses are not completely independent. To capture this dependency, let $\mathcal{F}$ indicate the status of other all the switches except the i^{th} switch (i.e., $\mathcal{F} = \big\{s_j | j \neq i\big\}$). Then, the transition probability matrix of the i^{th} switch based on status of other switches corresponds to,*

$$\boldsymbol{\pi}_{i|\mathcal{F}} = \begin{bmatrix} 1-\alpha_{i|\mathcal{F}} & \alpha_{i|\mathcal{F}} \\ \beta_{i|\mathcal{F}} & 1-\beta_{i|\mathcal{F}} \end{bmatrix}, \qquad (19)$$

where, $\alpha_{i|\mathcal{F}} \backslash \beta_{i|\mathcal{F}}$ are the probability rates of closing$\backslash$ opening of the i^{th} switch based on status of other switches.

Therefore, we can obtain a probability of every switch status at each time t. Then, we directly use this probability to start the alternative minimization. More precisely, the $\hat{s}_i^1$ is not selected to be 0.5 anymore. Instead, the initial estimation of i^{th} switch status corresponds to $\hat{s}_{i,t}^1 = \mathbf{Pr}(s_{i,t} = 1 | \hat{s}_{i,t-1})$. In other words, we use the level of our confidence about a switch status

to find more accurate guesstimate for switch statuses that lead to a better voltage estimation at the first iteration. Furthermore, we leverage the partial knowledge about the support of the sparse state to enhance the accuracy of the estimation process. With the assumption of slow support change which has also been confirmed with with partial data [39], we use $\hat{N}_{t-1}$ in the ℓ_1 minimization problem:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \parallel \mathbf{x}_{(\hat{\mathbf{N}}_{\mathbf{t-1}})^{\mathbf{c}}} \parallel_1 + \lambda^{\mathbf{S}} \parallel \mathbf{y}^{\mathbf{S}} - \mathbf{\Phi}^{\mathbf{S}}\mathbf{\Psi}^{\mathbf{S}}\mathbf{x}^{\mathbf{S}} \parallel_2 + \lambda^{\mathbf{V}} \parallel \mathbf{y}^{\mathbf{V}} - \mathbf{\Phi}^{\mathbf{V}}\mathbf{\Psi}^{\mathbf{V}}\mathbf{x}^{\mathbf{V}} \parallel_2 \quad (20)$$

Algorithm 3 summarizes our proposed approach for dynamic joint topology identification and state estimation. Based on the argument presented earlier in section III and [43], convergence of Algorithm 3 can also be guaranteed.

Algorithm 3 Dynamic Joint Topology Identification and State Estimation

Step 1. Collect the compressed measurements into two vectors $\mathbf{y}_{\mathbf{t}}^{\mathbf{S}}$ and $\mathbf{y}_{\mathbf{t}}^{\mathbf{V}}$.

Step 2. Initialize the parameters $\hat{s}_i^0 = \mathbf{Pr}\big(s_{i,t} = 1 | \hat{s}_{i,t-1}\big) \quad \forall i \in \{1, 2, ..., k\}$; $\hat{\mathbf{V}}_i^0 = 1\angle 0$. Set $j = 1$.

Step 3. Fix the switch status by using $\hat{\hat{s}}_{i,t}^{j-1}$. Then estimate the voltage $\hat{\mathbf{V}}_t^j$ using:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \parallel \mathbf{x} \parallel_1 + \lambda^{\mathbf{S}} \parallel \mathbf{y}^{\mathbf{S}} - \mathbf{\Phi}^{\mathbf{S}}\mathbf{\Psi}^{\mathbf{S}}\mathbf{x}^{\mathbf{S}} \parallel_2 + \lambda^{\mathbf{V}} \parallel \mathbf{y}^{\mathbf{V}} - \mathbf{\Phi}^{\mathbf{V}}\mathbf{\Psi}^{\mathbf{V}}\mathbf{x}^{\mathbf{V}} \parallel_2 \quad (21)$$

$$\text{subject to } (8), (16)$$

Step 4. In the following minimization, implement the voltage vector $\hat{\mathbf{V}}_t^j$ as fixed values and estimate switch status $\hat{\hat{s}}_{i,t}^j$ by solving:

$$\hat{\mathcal{B}} = \underset{\mathcal{B}}{\operatorname{argmin}} \parallel \mathcal{B} \parallel_1 + \lambda^{\mathbf{S}} \parallel \mathbf{y}_t^{\mathbf{S}} - \mathbf{\Phi}^{\mathbf{S}}\mathbf{\Psi}^{\mathbf{S}}\mathcal{B} \parallel_2$$

$$\text{subject to } (8), (15), (16)$$

Step 5. If $||\hat{\mathbf{v}}_t^j - \hat{\mathbf{v}}_t^{j-1}|| < \delta$ holds, $\hat{\mathbf{z}}_t = \mathbf{\Psi}\hat{\mathbf{x}}$ and proceed to Step 6.

Otherwise, $j \leftarrow j + 1$ and return to Step 3.

Step 6. Round the continuous switch status to 0 and 1:

$$\hat{s}_i = \begin{cases} 1 & \hat{\hat{s}}_i \geq 0.5 \\ 0 & \hat{\hat{s}}_i < 0.5 \end{cases}$$

V. Simulation Results

We first evaluate our proposed methods on the unbalanced three phase IEEE 37-bus test feeder with six switches as described in section II. For load data, we use the data collected from actual consumers that are publicly available at eGauge website [46]. Sparsifying matrix for both power and voltage vectors ($\mathbf{\Psi}^{\mathbf{S}}$ and $\mathbf{\Psi}^{\mathbf{V}}$) are DCT matrices. Here, we apply random measurement matrices $\mathbf{\Phi}^{\mathbf{S}}$ and $\mathbf{\Phi}^{\mathbf{V}}$ consisting of random Bernoulli entries which selects $(m/n)\%$ of the data. All the simulations are executed in MATLAB program equipped with MOSEK and cvx packages. The code was run on intel CORE i7 PC, with clock speed 3.4GHz and 32.0GB RAM.

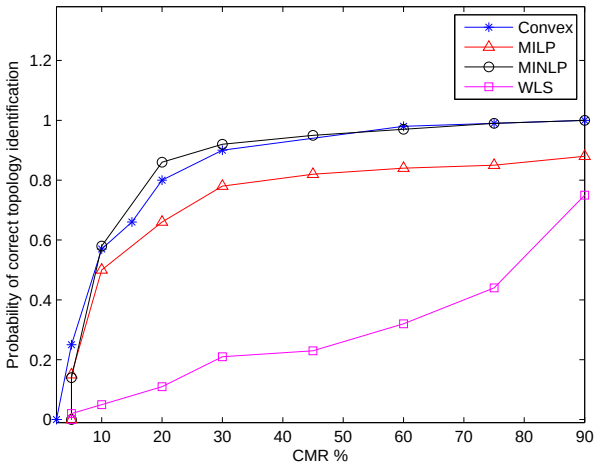

Fig. 3. MINLP, MILP and Convex topology identification for 6 switches system.

A. Performance of Topology Identification

The probability of correct topology identification averaged over 200 Monte-Carlo simulations is presented in Figure 3. Based on different CMR (percentage of measurements), Fig. 3 demonstrates and compares the probabilities of accurate topology identification performed by the original MINLP approach, MILP technique (Algorithm 1), the convex method (Algorithm 2) and the weighted least squares (WLS) method. Obviously, the MINLP has the best performance among the methods in terms of accurate topology identification. However, a major drawback of this technique is its high computational time. Furthermore, most of the existing solvers for MINLP fail to perform in presence of higher number of integer variables. As shown in Fig. 3, the convex method provides a very similar result to the MINLP while it considerably reduces the required time for optimization. Obviously, the convex method outperforms the MILP technique, i.e., Algorithm 2 has been more successful relative to the Algorithm 1. Additionally, the conventional WLS approach is not effective for topology identification as it does not exploit any underlying sparsity constraints. Therefore, the conventional methods fail to perform accurately in the presence of extreme low observability. Table I demonstrates the required times for implementing each method where the required time for MINLP approach is almost 40 times higher than the time required for executing the convex method.

TABLE I
Time requirement for 1 run

Algorithm 1 and 2	Time
MINLP	203.6s
MILP	29.2s
Convex	5.5s

To demonstrate scalability, the proposed techniques are implemented on the IEEE 123 bus test system and the results are shown in Fig. 4. The test system is equipped with 12 switches whose statuses determine the underlying topology.

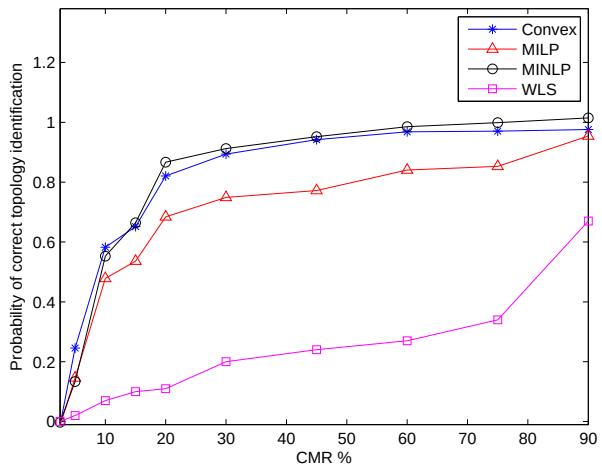

Fig. 4. MINLP, MILP and Convex topology identification for IEEE 123 bus test system.

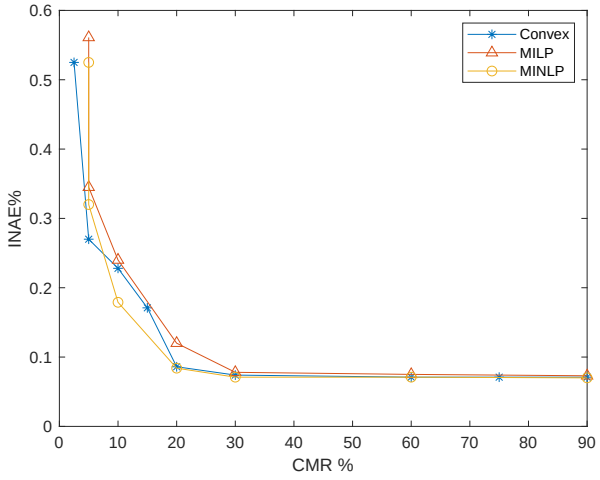

Fig. 5. Topology identification performance with different CMR

From Fig. 4, it is evident that the proposed algorithms successfully identify the underlying topologies while the WLS method fails to operate in an underdetermined situation. This example validates scalability and replicability of the proposed algorithms.

B. Performance of State Estimation

The performance of state estimation is compared using a percentage ratio, called *Integrated Normalized Absolute Error (INAE)* which is defined as [34],

$$INAE = \frac{\sum_{j=1}^{n} |\mathbf{x}_j - \hat{\mathbf{x}}_j|}{\sum_{j=1}^{n} |\mathbf{x}_j|} \times 100 \qquad (22)$$

where, $\mathbf{x}_j$ is the j^{th} element in the sparse vector $\mathbf{x}$ and $\hat{\mathbf{x}}_j$ is the corresponding estimate of $\mathbf{x}_j$ recovered by compressive sensing. Similar to topology identification, Fig. 5 demonstrates that the convex approach shows a better performance compared to the MILP method, in terms of state estimation accuracy. Also, there is not a notable improvement if we implement the original MINLP technique.

The dynamic topology identification technique (Algorithm 3) is evaluated based on the practical load data that is measured every minute. We assume that we have an accurate knowledge about the states and switches. Furthermore, we assume that the topology changes between the fifth and sixth minutes (i.e., $5 < t < 6$). Fig. 6 demonstrates the result of this setup. Interestingly, when we have partial information about the topology, the performance is improved. However, we witness a degradation in performance after the change in topology configuration of the IEEE 37 nodes test feeder. However, after three minutes the topology could be correctly identified. As a result, the dynamic technique outperforms the static techniques when as partial knowledge of switch statuses is better exploited in the former approach.

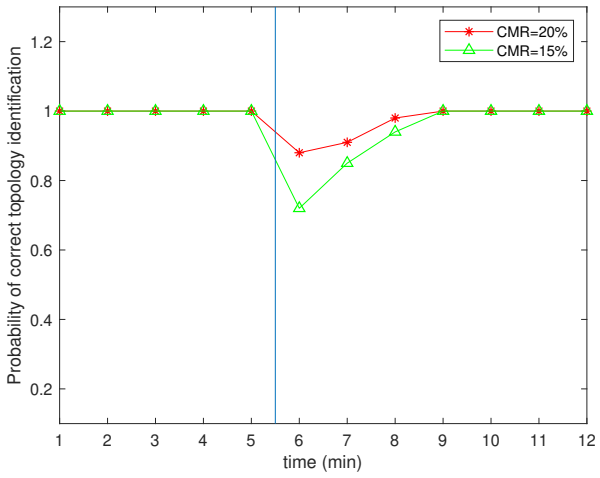

Fig. 6. Dynamic TI for two different CMR when the topology changes.

C. Effect of Measurement Noise

In addition to the level of compression and sparsity of the signal, the level of measurement noise is also another factor that considerably affects the performance of compressive sensing based approaches. Let $\delta_s(A)$ denote the restricted isometry constant (RIC) for a matrix $\mathbf{A}$, which is the smallest real number satisfying:

$$(1-\delta_s)\|\mathbf{b}\|_2^2 \leq \|\mathbf{Ab}\|_2^2 \leq (1+\delta_s)\|\mathbf{b}\|_2^2$$

for all s-sparse vectors $\mathbf{b}$. As shown in [47], for the original compressive sensing problem with the measurements model of $\mathbf{y} = \mathbf{Cx} + \boldsymbol{\nu}$ and $\|\boldsymbol{\nu}\|_2 \leq \xi$, if $\delta_s(C) < 0.207$, the reconstruction error is:

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 \leq G(2s)\xi \leq 7.5\xi$$

where $G(k) = \dfrac{4\sqrt{1+\delta_k}}{1-2\delta_k}$. This theoretical analysis implies that the error in reconstruction of a sparse signal heavily depends on measurement noise and level of compression. Fig.

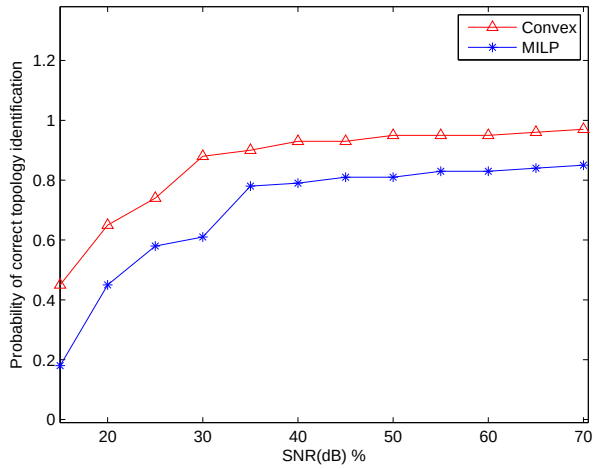

Fig. 7. MILP and Convex topology identification with different signal to noise ratios (SNR)

TABLE II
AVERAGED PROBABILITY OF ACCURATE TOPOLOGY IDENTIFICATION

Method	*IEEE 37*	*IEEE 123*	*559 node*
Algorithm 2	89.56	87.14	85.02
McCormick relaxation with NLP	88.72	82.91	73.47

7 shows the performance of Algorithms 1 and 2 in presence of different signal to noise(SNR). Here, the level of compression is set to be 0.5 (i.e. 50% of measurements are available). It is clear that our proposed techniques are still applicable even with high level of noises (low SNR).

D. Mccormick Relaxation with NLP

To compare the alternating minimization method (Algorithm 2) with a common technique, we implement the Mccormick relaxation that results in an NLP problem. To compare the performance of the methods in different scales, we evaluate our techniques using a larger distribution system introduced in [48]. This test feeder is a three-phase unbalanced distribution network consisting both the primary and secondary level distribution systems that totally includes 559 nodes. The results are summarized in Table II. While the performance of both techniques are similar for IEEE 37-node test feeder, it is obvious that the proposed algorithm outperforms the McCormick relaxation with the NLP method when the size of distribution systems becomes larger. Additionally, simulation results relevant to Algorithm 3 confirms that the iterative technique becomes more efficient when an accurate probability of transition between switches is available. However, since our technique is an iterative process, its average speed is lower than the speed of McCormick relaxation with NLP method.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed an MINLP framework that aims to jointly estimate the states and topology of the distribution system using a limited number of measurements. We removed the nonlinearity in the MINLP formulation by introducing auxiliary variables that led to a MILP. Furthermore, to provide a faster solution, we relaxed the nonlinearity of MINLP by replacing the integer variables with continuous decision variables. Then, we apply an alternative minimization approach that improves both state estimation and topology identification. Finally, we introduced a hybrid dynamic framework that incorporates the previous information about the switch statuses for current time estimation. More precisely, a Markov jump model was proposed for switch status that helps improve topology identification especially when the compression ratio is low. Prior support set information was shown to further improve the performance of our estimation method. Simulation results based on IEEE 37-bus test feeder and IEEE 123-bus test feeder illustrated the potency of the proposed techniques.

While this paper addresses some key problems related to joint topology identification and DSSE, there are several open questions that require further investigation. Specifically, this paper does not address status estimation of regulators and load tap changers (LTC). Furthermore, it should be mentioned that this paper focuses on primary level of the distribution systems. Investigating the status estimation of LTC as well as developing a hybrid model and data based approach to deal with lack of observability in the secondary system will be part of our future work. Additionally, to improve scalability, we plan to investigate distributed implementations of the proposed approach as part of our future work. Another critical factor in DSSE and topology identification is data aggregation mechanism (i.e., designing the measurement matrix Φ). To enhance observability and performance of state estimation in distribution grids, we are currently working on the problem of optimal sensor placement for data aggregation. As mentioned in section II, transformation/sparsifying matrix Ψ directly impacts the accuracy of state reconstruction. In general, determining the optimal Ψ is referred to as dictionary learning. While we have done some preliminary investigation (see [35]), a more detailed analysis will be a part of our future work.

ACKNOWLEDGMENT

This material is based upon work supported by the Department of Energy, Office of Energy Efficiency and Renewable Energy (EERE), Solar Energy Technologies Office, under Award Number DE-EE0008767.

REFERENCES

[1] A. Azizivahed, A. Arefi, S. Ghavidel, M. Shafie-khah, L. Li, J. Zhang, and J. P. Catalão, "Energy management strategy in dynamic distribution network reconfiguration considering renewable energy resources and storage," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 2, pp. 662–673, 2019.

[2] M. Farajollahi, A. Shahsavari, and H. Mohsenian-Rad, "Topology identification in distribution systems using line current sensors: An milp approach," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1159–1170, 2019.

[3] G. Cavraro, A. Bernstein, V. Kekatos, and Y. Zhang, "Real-time identifiability of power distribution network topologies with limited monitoring," *IEEE Control Systems Letters*, vol. 4, no. 2, pp. 325–330, 2019.

[4] S. Bhela, V. Kekatos, and S. Veeramachaneni, "Enhancing observability in distribution grids using smart meter data," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 5953–5961, 2017.

[5] F. F. Wu and W.-H. Liu, "Detection of topology errors by state estimation (power systems)," *IEEE Transactions on Power Systems*, vol. 4, no. 1, pp. 176–183, 1989.

[6] R. Singh, E. Manitsas, B. C. Pal, and G. Strbac, "A recursive bayesian approach for identification of network configuration changes in distribution system state estimation," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1329–1336, 2010.

[7] G. N. Korres and P. J. Katsikas, "Identification of circuit breaker statuses in wls state estimator," *IEEE Transactions on Power Systems*, vol. 17, no. 3, pp. 818–825, 2002.

[8] N. Da Silva, A. S. Costa, K. Clements, and E. Andreoli, "Simultaneous estimation of state variables and network topology for power system real-time modeling," *Electric Power Systems Research*, vol. 133, pp. 338–346, 2016.

[9] G. N. Korres and N. M. Manousakis, "A state estimation algorithm for monitoring topology changes in distribution systems," in *2012 IEEE Power and Energy Society General Meeting*. IEEE, 2012, pp. 1–8.

[10] A. Gandluru, S. Poudel, and A. Dubey, "Joint estimation of operational topology and outages for unbalanced power distribution systems," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 605–617, 2019.

[11] E. M. Lourenço, A. S. Costa, and K. A. Clements, "Bayesian-based hypothesis testing for topology error identification in generalized state estimation," *IEEE Transactions on power systems*, vol. 19, no. 2, pp. 1206–1215, 2004.

[12] S. J. Pappu, N. Bhatt, R. Pasumarthy, and A. Rajeswaran, "Identifying topology of low voltage distribution networks based on smart meter data," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5113–5122, 2017.

[13] Z. Tian, W. Wu, and B. Zhang, "A mixed integer quadratic programming model for topology identification in distribution network," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 823–824, 2015.

[14] G. Cavraro and R. Arghandeh, "Power distribution network topology detection with time-series signature verification method," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 3500–3509, 2017.

[15] N. Zhou, L. Luo, G. Sheng, and X. Jiang, "Power distribution network dynamic topology awareness and localization based on subspace perturbation model," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 1479–1488, 2019.

[16] J. Yu, Y. Weng, and R. Rajagopal, "Patopa: A data-driven parameter and topology joint estimation framework in distribution grids," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 4335–4347, 2017.

[17] O. Ardakanian, V. W. Wong, R. Dobbe, S. H. Low, A. von Meier, C. J. Tomlin, and Y. Yuan, "On identification of distribution grids," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 3, pp. 950–960, 2019.

[18] D. Deka, S. Backhaus, and M. Chertkov, "Structure learning in power distribution networks," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1061–1074, 2017.

[19] H. Doddi, D. Deka, and M. Salapaka, "Learning partially observed meshed distribution grids," in *2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*. IEEE, 2020, pp. 1–6.

[20] S. Park, D. Deka, and M. Chcrtkov, "Exact topology and parameter estimation in distribution grids with minimal observability," in *2018 Power Systems Computation Conference (PSCC)*. IEEE, 2018, pp. 1–6.

[21] S. Park, D. Deka, S. Backhaus, and M. Chertkov, "Learning with end-users in distribution grids: Topology and parameter estimation," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 3, pp. 1428–1440, 2020.

[22] D. Deka, M. Chertkov, and S. Backhaus, "Joint estimation of topology and injection statistics in distribution grids with missing nodes," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 3, pp. 1391–1403, 2020.

[23] G. Cavraro, V. Kekatos, and S. Veeramachaneni, "Voltage analytics for power distribution network topology verification," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 1058–1067, 2017.

[24] Y. Weng, Y. Liao, and R. Rajagopal, "Distributed energy resources topology identification via graphical modeling," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2682–2694, 2016.

[25] D. Deka, S. Talukdar, M. Chertkov, and M. V. Salapaka, "Graphical models in meshed distribution grids: Topology estimation, change detection & limitations," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4299–4310, 2020.

[26] T. Li, L. Werner, and S. H. Low, "Learning graphs from linear measurements: Fundamental trade-offs and applications," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 6, pp. 163–178, 2020.

[27] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1617–1626, 2012.

[28] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on information theory*, vol. 46, no. 2, pp. 388–404, 2000.

[29] S. Dahale, H. S. Karimi, K. Lai, and B. Natarajan, "Sparsity based approaches for distribution grid state estimation - a comparative study," *IEEE Access*, vol. 8, pp. 198 317–198 327, 2020.

[30] R. Madbhavi, H. S. Karimi, B. Natarajan, and B. Srinivasan, "Tensor completion based state estimation in distribution systems," in *2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 2020, pp. 1–5.

[31] A. Akrami, M. S. Asif, and H. Mohsenian-Rad, "Sparse distribution system state estimation: An approximate solution against low observability," in *2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 2020, pp. 1–5.

[32] P. L. Donti, Y. Liu, A. J. Schmitt, A. Bernstein, R. Yang, and Y. Zhang, "Matrix completion for low-observability voltage estimation," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2520–2530, 2019.

[33] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[34] S. S. Alam, B. Natarajan, and A. Pahwa, "Distribution grid state estimation from compressed measurements," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1631–1642, 2014.

[35] H. S. Karimi and B. Natarajan, "Compressive sensing based state estimation for three phase unbalanced distribution grid," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.

[36] S. K. Hazhar, "Design and analysis of dynamic compressive sensing in distribution grids," Ph.D. dissertation, 2020.

[37] A. Joshi, L. Das, B. Natarajan, and B. Srinivasan, "Effect of transformation in compressed sensing of smart grid data," in *2019 IEEE PES GTD Grand International Conference and Exposition Asia (GTD Asia)*. IEEE, 2019, pp. 177–182.

[38] A. Bernstein and E. Dall'Anese, "Linear power-flow models in multiphase distribution networks," in *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*. IEEE, 2017, pp. 1–6.

[39] H. S. Karimi and B. Natarajan, "Recursive dynamic compressive sensing in smart distribution systems," in *2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 2020, pp. 1–5.

[40] S. Ma, L. Su, Z. Wang, F. Qiu, and G. Guo, "Resilience enhancement of distribution grids against extreme weather events," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4842–4853, 2018.

[41] R. Billinton and S. Jonnavithula, "Optimal switching device placement in radial distribution systems," *IEEE transactions on power delivery*, vol. 11, no. 3, pp. 1646–1651, 1996.

[42] R. Kannan and C. L. Monma, "On the computational complexity of integer programming problems," in *Optimization and Operations Research*. Springer, 1978, pp. 161–172.

[43] A. Beck, "On the convergence of alternating minimization for convex programming with applications to iteratively reweighted least squares and decomposition schemes," *SIAM Journal on Optimization*, vol. 25, no. 1, pp. 185–209, 2015.

[44] H. A. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with markovian switching coefficients," *IEEE transactions on Automatic Control*, vol. 33, no. 8, pp. 780–783, 1988.

[45] R. E. Brown, S. Gupta, R. D. Christie, S. S. Venkata, and R. Fletcher, "Distribution system reliability assessment using hierarchical markov modeling," *IEEE Transactions on power Delivery*, vol. 11, no. 4, pp. 1929–1934, 1996.

[46] "egauge." [Online]. Available: http://egauge141.egaug.es/

[47] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes rendus mathematique*, vol. 346, no. 9-10, pp. 589–592, 2008.

[48] A. R. Malekpour and A. Pahwa, "Radial test feeder including primary and secondary distribution network," in *2015 North American Power Symposium (NAPS)*. IEEE, 2015, pp. 1–9.

Hazhar Sufi Karimi received his B.Sc. degree in Electrical Engineering from Sahand University of Technology in 2012, and the M.Sc. degree from Tarbiat Modares University, Iran, in 2014, and the Ph.D. degree in Electrical Engineering from Kansas State University, USA, in 2020. He is currently a postdoctoral fellow at the Harvard University, Boston, MA, USA. His research interests include statistical signal processing, estimation theory, machine learning, and power distribution systems.

Balasubramaniam Natarajan (SM08) received the B.E. degree (Hons.) in electrical and electronics engineering from Birla Institute of Technology and Science, Pilani, India, Ph.D. degree in electrical engineering from Colorado State University, Fort Collins, CO, USA, Ph.D. degree in Statistics from Kansas State University, Manhattan, KS, USA, in 1997, 2002, and 2018, respectively. He is currently a Clair N. Palmer and Sara M. Palmer Endowed Professor and the Director of the Wireless Communication and Information Processing Research Group. His research interests include statistical state processing, stochastic modeling, optimization, and control theories. He has worked on and published extensively on modeling, analysis and networked estimation and control of smart distribution grids and cyber physical systems in general. He has published over 200 refereed journal and conference articles and has served on the editorial board of multiple IEEE journals including IEEE Transactions on Wireless Communications.