Topology Identification in Distribution Systems using Line Current Sensors: An MILP Approach

Mohammad Farajollahi, *Student Member, IEEE*, Alireza Shahsavari, *Student Member, IEEE*, and Hamed Mohsenian-Rad, *Senior Member, IEEE*

***Abstract*— This study is motivated by the recent advancements in developing non-contact line sensor technologies that come at a low cost, but have limited measurement capabilities. While they are intended to measure current, they cannot measure voltage and power. This poses a challenge to certain distribution system applications, such as topology identification (TI), because they commonly use voltage and power measurements. To address this open problem, a new TI algorithm is proposed based on measurements from a few line current sensors, together with available pseudo-measurements for nodal power injections. A TI problem formulation is first developed in the form of a mixed integer nonlinear program (MINLP). Several reformulation steps are then adopted to tackle the nonlinearities to express the TI problem in the form of a mixed integer linear program (MILP). The proposed method is able to identify all possible topologies, including radial, loop, and island configurations, which extends the application of TI to identify switch malfunctions and to detect outages. In addition, recommendations are made with respect to the number and location of the line current sensors to ensure performance accuracy of the TI method. A novel multi-period TI algorithm is also proposed to use multiple measurement snapshots to improve the TI accuracy and robustness against errors in pseudo-measurements. The effectiveness of the proposed TI algorithms is examined on the IEEE 33-bus test case as well as a test case based on a real-world feeder in Riverside, CA.**

***Keywords*: Topology identification, line current sensors, distribution network, mixed integer linear program, single-period and multi-period optimization, radial topology, loop topology.**

Nomenclature

Sets

$\mathcal{K}$	Set of lines equipped with current sensor.
$\mathcal{L}$	Set of lines.
$\mathcal{N}$	Set of nodes.
$\mathcal{N}_i$	Set of lines which are connected to node i.
$\mathcal{T}$	Set of snapshots

Main Decision Variables

b_i	Binary variable for switching status of node i.
s_{ij}	Binary variable for switching status of line $\{i, j\}$.

Auxiliary Decision Variables

$\gamma[t]$	Probability of snapshot $[t]$.
$\Psi[t]$	TI objective function of snapshot $[t]$.
E_{ij},F_{ij}	Auxiliary variables used in Step 3 of Section II.B.
G_i, H_i	Auxiliary variables used in Section III.B.
I_{ij}	Current variable of line $\{i, j\}$.
I_i	Injection current variable of node i.
t	Time index for the snapshots.
U_{ij}	Auxiliary variable used in Step 1 of Section II.B.
V_i	Voltage variable of node i.
W_{ij}	Auxiliary variable used in Step 2 of Section II.B.

Operators and Parameters

$(\cdot)^*$	Conjugate transpose.
α_{ij}	Standard deviation of error in current measurement.
β_i	Standard deviation of error in pseudo-measurement.
η	Bound of measurements/pseudo-measurements error.
σ	Standard deviation of errors.
$\tilde{z}$	Contaminated measurement vector
I_{ij}^m	Current measurement at line $\{i, j\}$.
L	Number of lines in the network.
M	A large number in the big-M method.
N	Number of nodes in the network.
n	Number of nodes in an independent loop.
S_i	Load injection at node i.
y_{ij}	Admittance of line $\{i, j\}$.
z	True measurement vector.
Im$\{\cdot\}$	Operator for imaginary part.
Re$\{\cdot\}$	Operator for real part.

The authors are with the Department of Electrical and Computer Engineering, University of California, Riverside, CA, USA. This work was supported in part by DoE grant EE 0008001 and NSF grant 1253516. The corresponding author is H. Mohsenian-Rad, e-mail: hamed@ece.ucr.edu.

I. Introduction

Correct knowledge of the network topology is vital for distribution system operation, with applications to state estimation, fault location, Volt-VAR control, demand response, etc. If the network topology is known incorrectly, then these applications produce incorrect results. Therefore, the topology of the network must frequently be identified. However, in practice, topology identification (TI) is a challenging task for distribution systems due to the limited measurements as well as unavailable or unreliable information about the status of switches and circuit breakers across distribution feeders.

The most common TI approach in practice is to dispatch utility crew members to examine the status of switches. But this is costly and cannot be done frequently in order to continuously track changes in topology. An alternative approach is to use measurement-based TI algorithms that estimate the status of switches based on available sensor data.

Different measurement-based methods have been proposed in the literature. One class of TI algorithms uses various field measurements to do distribution system state estimation for every possible topology configuration, and then chooses the topology with the minimum residue errors in state estimation

[1], [2]. However, such exhaustive approach is often computationally complex. A remedy is to integrate the TI problem into the state estimation problem formulation to simultaneously estimate both the system states and the status of the switches [3], [4]. This can help improve computational efficiency. But the issue with any TI algorithm that involves state estimation is that one may not have sufficient field measurements to do full state estimation; even though the available measurements could be sufficient for the sole purpose of TI.

There are other TI methods in the literature that use rather limited measurements solely for the purpose of TI. A common approach in this line of work is to use *voltage measurements* [5]–[9]. For example, in [8], the authors reconstruct the grid topology by conducting a correlation analysis of voltage measurements. In [9], a time-series signature verification method is employed to identify the network topology using synchronized voltage phasor data. Another approach is to use smart meter data, e.g., in [10] where voltage measurements are used; in [11] where energy measurements are used; and in [12] where both voltage and energy measurements are used.

There are also TI methods that work based on measuring line power flow, such in [13]–[16]. In practice, measuring power at each measurement point requires installing instrumentation devices such as CTs and PTs, which are known to be expensive. It should be added that, for the method in [16], it uses the Distflow model which works only in radial topologies and cannot be used in loop configurations that can occur for example due to switching malfunction or cyber-attacks.

In this paper, a novel TI algorithm is proposed based on two types of measurements: 1) line current measurements at a few lines where line current sensors are installed; and 2) pseudo-measurements in form of measured or otherwise estimated nodal power injections at all buses. Specifically, this study is motivated by the recent advancements in developing non-contact line current sensors. These sensors have recently been adopted in the utility industry for installation on power distribution lines. They are cheaper in cost and easier to be installed, in comparison with the traditional utility sensors that measure voltage and power. Non-contact line sensors can also measure e-field, which can be used to estimate the relative phase angle for the current measurements. In other words, these sensors can estimate the phasor, just like micro-PMUs [17], but only for current and at a much lower cost. These line current sensors measure current greater than 1 *mA* with 1% error and phase angle with 1.5° error, with a second-by-second reporting rate [18]. There already exists a variety of commercial choices for non-contact line current sensors, e.g., see [19], [20], and a growing number of utilities, such as in Riverside, CA, are looking into installing such low-cost sensors on their medium and low voltage distribution lines.

While the aforementioned line current sensor technologies are capable of measuring electric field, they are not capable of measuring voltage. More precisely, they cannot measure voltage while keeping their cost down at a level that justifies their application in power distribution systems. As a result, they cannot measure the power that flows through the conductor either. Therefore, the TI methods that are built based on either measuring voltage, as mentioned in [8]–[11], or measuring line power flow, such as those in [14] and [15], are not applicable for the networks with non-contact line sensors.

Addressing the above open issues is the focus in this paper. The contributions in this paper can be summarized as follows:

1) A novel TI method is proposed for distribution networks that uses the measurements from line current sensors. To the best of our knowledge, this is the first paper proposing a TI method that is specifically concerned with utilizing the type of measurements that come from these emerging low-cost line current sensors.
2) A TI problem formulation is first developed in this context in the form of a mixed integer nonlinear program (MINLP). Then, proper reformulations is made to tackle the nonlinearities in order to express the problem in the form of a mixed integer linear program (MILP).
3) Some of the less expensive line current sensors do not directly measure the phase angle for current; instead, they measure the relative phase angle with respect to the electric field around the conductor. This results in less accurate readings of the phase angle. Importantly, our proposed TI algorithm can work accurately under such errors. There are also other sources of significant error in pseudo-measurements. Those errors too are tackled in this paper by developing a novel multi-period-based TI algorithm which uses multiple measurement snapshots, beyond the moment that a change in topology is detected, to improve the TI accuracy and robustness.
4) The proposed method can identify all possible topologies including radial, loop, and island configurations. Thus, the application of TI is extended to identifying switch malfunctions that are important concerns in practical operation of distribution systems, c.f. [21],[22].
5) Our case studies include both IEEE test cases and a real-world feeder in Riverside, CA. The performance of both the single-period and multi-period TI algorithms are evaluated for identifying radial, loop, and island configurations, as well as the impact of errors in measurements and pseudo-measurements are examined.
6) A theorem is also expressed with respect to the number and location of the line current sensors to ensure accurate TI performance.

II. Topology Identification Methodology

A. TI Problem Formulation in Non-linear Form

Suppose all switches are closed and all lines and nodes are in-service. According to the Kirchhoff Current Law (KCL), the current injection to each node is equal to the summation of the currents associated with the lines connected to that node:

$$I_i = \sum_{j \in \mathcal{N}_i} I_{ij}, \quad i \in \mathcal{N}. \tag{1}$$

The current that flows at a line between two nodes depends on the switching status of the line, the voltage of the two nodes, and the admittance of the line. According to the Kirchoff Voltage Law (KVL), we have:

$$I_{ij} = s_{ij}(V_i - V_j)y_{ij}, \tag{2}$$

If line $\{i, j\}$ is in-service, s_{ij} is one; otherwise it is zero, which ensures I_{ij} to be zero, indicating that the line is out-of-service.

According to the Power Law, the relationship between nodal voltage V_i and current injection I_i can be given as:

$$V_i I_i^* = b_i S_i. \tag{3}$$

If node i is connected to the grid, then b_i is one; otherwise, it is zero, indicating that there exists an island node. In such cases, the TI algorithm can determine an outage area; thus extending its use-cases beyond basic TI. The TI problem is expressed in form of the following optimization problem:

$$\begin{aligned} &\underset{\substack{s_{ij}, ij \in \mathcal{L} \\ b_i, i \in \mathcal{N}}}{\textbf{minimize}} \sum_{ij \in \mathcal{K}} |I_{ij} - I_{ij}^m|, \\ &\textbf{subject to } \text{Eqs (1)-(3).} \end{aligned} \tag{4}$$

In the above optimization problem, the main decision variables are the binary line switching status variables s_{ij} and the binary node switching status variables b_i. Other variables, i.e., the line current I_{ij}, the nodal current injection I_i, and the nodal voltage V_i, are auxiliary variables. All other notations, i.e, the complex load S_i and the line current measurement I_{ij}^m, are parameters. This problem tends to match the line current phasors that are calculated by (1)-(3) with those that are measured by the line sensors. The solutions to binary variables s_{ij} for all lines $\{i, j\}$ provide the status of line switches; thus they indicate the network topology. As for the solutions to binary variables b_i for all nodes i, they indicate whether a node is connected to the rest of the grid; thus they indicate outage areas. The auxiliary variables, such as V_i, can be interpreted as a rough estimation of the state of the system.

B. TI Problem Formulation in MILP Form

The optimization problem in (4) involves several nonlinear terms both in its objective function and in its constraints. In this section, it is explained how this problem can be converted into a more tractable MILP formulation.

Step 1: The non-linearity in (2) is due to the multiplication of binary variable s_{ij} with continuous phasor variables V_i and V_j. In order to overcome this nonlinearity, first, (2) is substituted with the following linear equation:

$$I_{ij} = y_{ij} U_{ij}, \tag{5}$$

where U_{ij} is an auxiliary variable which is equivalent to the multiplication of s_{ij} and $V_i - V_j$. If $s_{ij} = 1$, i.e, if line $\{i, j\}$ is in-service, then $U_{ij} = V_i - V_j$, otherwise $U_{ij} = 0$. The above relationship between U_{ij}, s_{ij}, and $V_i - V_j$ can be achieved by adding the following new linear constraints [23, Ch. 5]:

$$-M(1 - s_{ij}) \le \text{Re}\{U_{ij}\} - \text{Re}\{V_i - V_j\} \le M(1 - s_{ij}) \tag{6}$$

$$-M\ s_{ij} \le \text{Re}\{U_{ij}\} \le M\ s_{ij} \tag{7}$$

$$-M(1 - s_{ij}) \le \text{Im}\{U_{ij}\} - \text{Im}\{V_i - V_j\} \le M(1 - s_{ij}) \tag{8}$$

$$-M\ s_{ij} \le \text{Im}\{U_{ij}\} \le M\ s_{ij}. \tag{9}$$

Because U_{ij} and $V_i - V_j$ are in complex form, the equations in (6)-(7) and (8)-(9) are used to respectively achieve the intended real part and the intended imaginary part of U_{ij}. If $s_i = 0$, then (6) and (8) are not binding; from (7) and (9), we have $U_{ij} = 0$. From this, together with (5), we have $I_{ij} = 0$. Thus, the nonlinear equality constraint in (2) is enforced when $s_i = 0$. If $s_i = 1$, then (7) and (9) are not binding. From (6) and (8), we have $U_{ij} = V_i - V_j$. From this, together with (5), we have $I_{ij} = (V_i - V_j) y_{ij}$. Thus, the nonlinear equality constraint in (2) is enforced also when $s_i = 1$. Together, the linear constraints in (5)-(9) are equivalent to the nonlinear constraint in (2).

Step 2: Another nonlinearity term in (4) is associated with the multiplication of phasors V_i and I_i in (3), which is known as power flow equation nonlinearity. There are different options to relax such nonlinearity. For example, one may use Linearized DistFlow (LDF) approximation [24]. However, LDF works only in radial feeders and *cannot* be used for loop configuration. Another limitation with LDF is that *line current variables* must be removed in the process of deriving the linear approximation, which makes LDF inappropriate in the context of TI algorithm proposed in this paper. In addition, there are some other linear approximation methods which would rather increase computational complexity, such as the one in [25] which introduces multiplications of two or more binary variables and the one [26] which utilizes Newton's method in an iterative Benders decomposition framework. Therefore, in this paper, the method in [27] is used, which works based on Taylor expansion and engineering approximation. To this end, a linear approximation is developed on complex numbers as opposed to on real numbers as in the conventional load flow formulations. Suppose the voltage at each bus i is expressed in per unit and in relationship with the voltage at the reference bus at the substation. That is, suppose

$$V_i = 1 - \Delta V_i. \tag{10}$$

By applying the Taylor series around zero and neglecting the high order terms, we can write [27], [28, Ch. 5]:

$$\frac{1}{V_i} = \frac{1}{1 - \Delta V_i} \simeq 1 + \Delta V_i = 2 - V_i. \tag{11}$$

The accuracy of this approximation is validated in the domain of complex numbers in [29]; and in particular for power flow equations in [27]. The approximation error is calculated as

$$\Phi = \left| \frac{1}{V_i} - (2 - V_i) \right|. \tag{12}$$

The above error index is evaluated in each point inside the disc in Fig. 1(a), resulting in the area in Fig. 1(b). For example, at $\Delta V = 0.1$ p.u., the approximation error is only around 1%. In practice, ΔV is often 0.05 p.u. or less; therefore, the approximation error is only 0.3% or less; which is negligible.

Next, the approximation in (11) is substituted into (3) and constraint (3) is rewritten in a reordered form as

$$I_i^* = S_i (2 - V_i) b_i. \tag{13}$$

However, the above equation is still nonlinear due to the multiplication of binary variable b_i and phasor variable V_i.

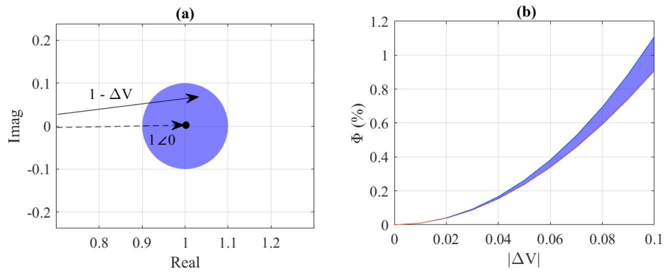

Fig. 1. Validating the accuracy of the linearization in (11). (a) Voltage phasor in complex domain; (b) Approximation error in percentage.

Hence, next, a new auxiliary variable W_i is introduced, and (13) is replaced with the following linear equation:

$$I_i^* = S_i(2b_i - W_i), \tag{14}$$

where W_i is equivalent to the multiplication of b_i and V_i. If b_i is one, i.e., if node i is connected to the grid, then $W_i = V_i$; otherwise $W_i = 0$, indicating the node is not connected to the grid. The mentioned relationship between W_i, b_i, and V_i can be achieved by enforcing the following new linear constraints:

$$-M(1-b_i) \leq \text{Re}\{W_i\} - \text{Re}\{V_i\} \leq M(1-b_i), \tag{15}$$

$$-M\, b_i \leq \text{Re}\{W_i\} \leq M\, b_i, \tag{16}$$

$$-M(1-b_i) \leq \text{Im}\{W_i\} - \text{Im}\{V_i\} \leq M(1-b_i), \tag{17}$$

$$-M\, b_i \leq \text{Im}\{W_i\} \leq M\, b_i. \tag{18}$$

For the same reason as mentioned for the constraints in (6)-(9), the equations in (15)-(16) and (17)-(18) can respectively achieve the intended real and imaginary parts of W_i. From (15)-(16), if $b_i = 1$, then $\text{Re}\{W_i\} = \text{Re}\{V_i\}$; and if $b_i = 0$, then $\text{Re}\{W_i\} = 0$. Similarly, the equations in (17)-(18) result in achieving the imaginary part of W_i. Together, the linear equality and inequality constraints in (14)-(18) are equivalent to the nonlinear equality constraint in (3).

It is worth mentioning that, from (14), the current injection associated with a disconnected node is assured to be zero. However, the equality in (14) does *not* require that the voltage of a disconnected node is also zero. Next, it is needed to also add the following new set of constraints to make sure that the disconnected node indeed has zero voltage:

$$-M\, b_i \leq \text{Re}\{V_i\} \leq M\, b_i, \tag{19}$$

$$-M\, b_i \leq \text{Im}\{V_i\} \leq M\, b_i. \tag{20}$$

Step 3: The objective function in (4) is an absolute value which is not linear. The absolute value in the objective function is substituted with a linear objective function as well as several linear inequality constraints using the technique introduced in [30, Ch.1]. In this regard, minimizing the objective function in (4) is equal to minimizing both the real part and the imaginary part of $(I_{ij} - I_{ij}^m)$, where $ij \in \mathcal{K}$. Let E_{ij} and F_{ij} denote two auxiliary variables, where $|\text{Re}\{I_{ij} - I_{ij}^m\}| \leq E_{ij}$ and $|\text{Im}\{I_{ij} - I_{ij}^m\}| \leq F_{ij}$. These auxiliary variables are integrated into the problem through the following constraints:

$$-E_{ij} \leq \text{Re}\{I_{ij}\} - \text{Re}\{I_{ij}^m\} \leq E_{ij}, \tag{21}$$

$$-F_{ij} \leq \text{Im}\{I_{ij}\} - \text{Im}\{I_{ij}^m\} \leq F_{ij}, \tag{22}$$

$$E_{ij} \geq 0 \text{ and } F_{ij} \geq 0. \tag{23}$$

Step 4: By combining Steps 1 to 3, the TI optimization problem is expressed in a linear form as follows:

$$\begin{aligned} &\underset{\substack{s_{ij}, ij \in \mathcal{L} \\ b_i, i \in \mathcal{N}}}{\textbf{minimize}} \sum_{ij \in \mathcal{K}} E_{ij} + F_{ij}, \\ &\textbf{subject to } \text{Eqs (1), (5)-(9), and (14)-(23).} \end{aligned} \tag{24}$$

The above problem is a MILP. It can be solved efficiently using various software packages, e.g., see [31].

III. Additional Notes and Extensions

A. Observability and Sensor Placement

Observability analysis and sensor placement is often studied for a particular application, such as in [32]–[35] where specific micro-PMUs allocation is considered for each application. In the context of this paper, observability analysis is concerned with the following. In order to determine the topology of a network, it is needed to know the status of switches. The lines whose switches are opened carry no current; thus, the status of switches is obtained by estimating the lines current; which ultimately leads to identifying the topology of the network. Therefore, observability analysis boils down to answering the following question: how many line sensors (and at what locations) are needed in order to estimate all lines current? The answer will automatically determine how many line sensors (and at what locations) are needed in order to solve TI problem in this paper.

Theorem 1. *If at least one line current sensor is placed in each independent loop in a distribution network, then the topology of the distribution network can be identifiable, i.e., problem (4) can be solved to provide the correct topology.*

Proof:

In a circuit with N nodes and L branches, L independent equations are needed in order to estimate the branch currents. From the N current injection equations that are available, $N-1$ equations are independent [36, Ch. 2]. Therefore, in order to reach a total of L independent equations, $L-(N-1)$ additional independent equations are needed, which is exactly the circuit nullity, and can be obtained by measuring $L-(N-1)$ line currents.

A *loop* is a closed path that starts from a node, passes through a set of nodes, and returns back to the initial starting node, without passing through any node more than once. A loop is said to be *independent* if it does not contain any loop in itself. In circuit theory, independent loops result in independent KVL equations, the number of which is $L-(N-1)$, which is known as *circuit nullity* [37]. Therefore, an immediate result of the nullity theorem is that all line currents can be estimated as long as the current of $L-(N-1)$ lines in independent loops are measured directly using line current sensors.

Next, consider an independent loop, as shown in Fig. 2. Suppose the nodal injection currents, i.e. $I_1, I_2, \cdots, I_n$, are

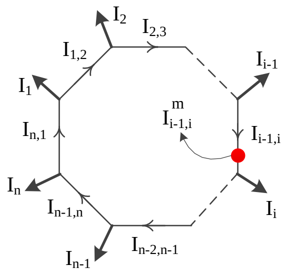

Fig. 2. An independent loop consisting of n buses. A current sensor is installed at line $\{i-1, i\}$ to measure $I^m_{i-1,i}$.

known. We want to calculate the current of each line segment in the loop, i.e. $I_{n,1}, I_{1,2} \cdots, I_{n-1,n}$. Due to KCL, we have:

$$\underbrace{\begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ & & & & \ddots & & & & \\ 0 & 0 & 0 & \cdots & 1 & -1 & \cdots & 0 & 0 \\ & & & & \ddots & & & & \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdot & 1 & -1 \\ -1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}}_{\text{Matrix}\,A} \begin{pmatrix} I_{n,1} \\ I_{1,2} \\ \vdots \\ I_{i-1,i} \\ \vdots \\ I_{n-2,n-1} \\ I_{n-1,n} \end{pmatrix} = \begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I_i \\ \vdots \\ I_{n-1} \\ I_n \end{pmatrix}, \tag{25}$$

where A is an $n \times n$ matrix. The rank of matrix A is $n-1$; thus, there exist infinite solutions for the choice of lines current. However, we can assure a unique solution for this system of linear equations if one row is replaced with an independent equation based on measurement of a line current sensor. Such additional equation is $I_{i-1,i} = I^m_{i-1,i}$, where $I^m_{i-1,i}$ denotes the measured line current on line segment between bus $i-1$ and bus i. This results in the following new system of linear:

$$\begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ & & & & \ddots & & & & \\ 0 & 0 & 0 & \cdots & 1 & 0 & \cdots & 0 & 0 \\ & & & & \ddots & & & & \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdot & 1 & -1 \\ -1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \begin{pmatrix} I_{n,1} \\ I_{1,2} \\ \vdots \\ I_{i-1,i} \\ \vdots \\ L_{n-2,n-1} \\ I_{n-1,n} \end{pmatrix} = \begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I^m_{i-1,i} \\ \vdots \\ I_{n-1} \\ I_n \end{pmatrix}. \tag{26}$$

Unlike in (25), the system of equations in (26) has a *unique* solution. It can be concluded that installing exactly one line current sensor anywhere in an independent loop provides an independent equation to be added to the KCL equations. Since there are $L-(N-1)$ independent loops in a network; installing $L-(N-1)$ line current sensors, one at each independent loop, results in solving the line current estimation problem.

Next, an illustrative example is used to demonstrate the observability results in the Theorem. Specifically, consider the network in Fig. 3 with $N = 4$ buses and $L = 5$ lines. In this network, loop *1231*, denoted by ℓ_1, and loop *1341*, denoted by ℓ_2, are two independent loops. However, loop *12341* is not an independent loop because it contains other loops, ℓ_1 or ℓ_2, inside itself. The set of equations for this network can be written as follows:

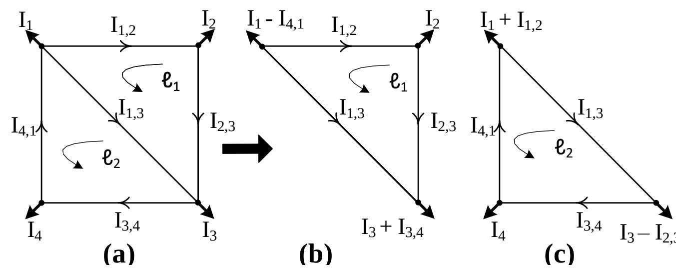

Fig. 3. A circuit with 4 buses and 2 loops. Nodal currents I_1, I_2, I_3, I_4 are known. Line currents $I_{1,2}, I_{1,3}, I_{2,3}, I_{3,4}, I_{4,1}$ are unknown.

TABLE I

SENSOR LOCATION SCENARIOS FOR THE NETWORK IN FIG. 3

Scen #	Lines Sensor	Loops	Added Rows to Matrix A	Rank
I	{1,2},{2,3}	ℓ_1,ℓ_1	$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$	4
II	{3,4},{4,1}	ℓ_2,ℓ_2	$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$	4
III	{1,2},{3,4}	ℓ_1,ℓ_2	$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$	5
IV	{1,2},{4,1}	ℓ_1,ℓ_2	$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$	5
V	{1,2},{1,3}	ℓ_1,ℓ_2	$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	5
VI	{2,3},{3,4}	ℓ_1,ℓ_2	$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$	5
VII	{2,3},{4,1}	ℓ_1,ℓ_2	$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$	5
VIII	{2,3},{1,3}	ℓ_1,ℓ_2	$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	5
IX	{3,4},{1,3}	ℓ_1,ℓ_2	$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	5
X	{4,1},{1,3}	ℓ_1,ℓ_2	$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	5

$$\underbrace{\begin{pmatrix} -1 & 0 & 0 & 1 & -1 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix}}_{\text{Matrix}\,A} \begin{pmatrix} I_{1,2} \\ I_{2,3} \\ I_{3,4} \\ I_{4,1} \\ I_{1,3} \end{pmatrix} = \begin{pmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{pmatrix} \tag{27}$$

Matrix A has rank 3, which means one row is dependent. In order to find the unique solution for all the 5 line currents, it is needed to add 2 more independent rows to matrix A. This is achieved by installing one line sensor at each of the two independent loops ℓ_1 and ℓ_2. Note that, out of the $\binom{5}{2}$=10 possible combinations for choosing the location to install the two sensors, 8 combinations result in measuring the line currents in both loops l_1 and l_2; as denoted by Scenarios III to X in Table I. These results are consistent with Theorem I. It should be noted that, if two independent loops share a line, such as line $\{1,3\}$ in Fig. 3 that is shared between loops ℓ_1 and ℓ_2, then placing a line sensor on such shared line will provide only one new independent equation. That is why two separate sensors are needed to be placed on two neighboring independent loops to obtain two independent equations in order to reach the number of independent equations that are needed to conduct topology identification.

Finally, there exist several graph theoretic methods to find the independent loops in a graph. Here, Kruskal's algorithm

[38] is used to identify the independent loops; which allowed us to place one line sensor at each independent loop.

B. Addressing Measurement Error

So far, it is assumed that the line measurements and nodal apparent power measurements, considered as pseudo-measurements, do not carry errors. However, this assumption may not hold in practice. In order to incorporate the impact of errors in measurements and pseudo-measurements, one needs to adjust the problem formulation in (24). To resolve this issue, we propose to move the injection current constraint in (1) to the objective function in form of a penalty term as

$$\underset{\substack{s_{ij}, ij\in\mathcal{L}\\ b_i, i\in\mathcal{N}}}{\textbf{minimize}} \sum_{i\in\mathcal{N}} \left| I_i - \sum_{j\in\mathcal{N}_i} I_{ij} \right|. \tag{28}$$

After that, by adopting a similar method as in Step 3 in Section II-B, the new nonlinearity can be tackled by defining auxiliary variables G_i and H_i, and adding the following new linear constraints:

$$-G_i \leq \mathrm{Re}\{I_i\} - \sum_{j\in\mathcal{N}_i} \mathrm{Re}\{I_{ij}\} \leq G_i, \tag{29}$$

$$-H_i \leq \mathrm{Im}\{I_i\} - \sum_{j\in\mathcal{N}_i} \mathrm{Im}\{I_{ij}\} \leq H_i, \tag{30}$$

$$G_i \geq 0 \text{ and } H_i \geq 0. \tag{31}$$

The TI problem in the presence of measurement and pseudo-measurement errors can be presented as a MILP as follows:

$$\begin{aligned} &\underset{\substack{s_{ij}, ij\in\mathcal{L}\\ b_i, i\in\mathcal{N}}}{\textbf{minimize}} \sum_{ij\in\mathcal{K}} (\alpha_{ij}^{Re} E_{ij} + \alpha_{ij}^{Im} F_{ij}) + \sum_{i\in\mathcal{N}} (\beta_i^{Re} G_i + \beta_i^{Im} H_i) \\ &\textbf{subject to} \text{ Eqs (5)-(9), (14)-(23), and (29)-(31),} \end{aligned} \tag{32}$$

where coefficients α_{ij} and β_i are calculated based on the standard deviation (SD) of measurements and pseudo-measurements [39]. Parameter α_{ij}^{Re} corresponds to the real part of I_{ij}^m. It is set to the inverse of the SD for $\mathrm{Re}\{I_{ij}^m\}$. Parameter α_{ij}^{Im} corresponds to the imaginary part of I_{ij}^m. It is set to the inverse of the SD for $\mathrm{Im}\{I_{ij}^m\}$. Coefficients β_i^{Re} and β_i^{Im} are set similarly.

The existence or uniqueness of the solution for the above optimization problem cannot be theoretically proved. But it *can* be confirmed that a feasible solution is always obtained even in the severe condition that measurements or pseudo-measurements are far from their true values due to major errors.

C. Multi-Period Optimization

The TI optimization problems in (24) and (32) are both defined for a single snapshot of available measurements. In fact, this is typical in the literature to run the TI algorithms based on one set of data, e.g., see [1], [14], [40]. However, in practice, the line current measurements are likely to continue to be available beyond the initial moment after the change in the network topology; thus, allowing the operator to run the TI algorithm in form of a multi-period optimization problem to better alleviate the impact of errors in measurements and pseudo-measurements. Note that, even if the line current measurements are accurate due to the use of higher precision line sensors, it is inevitable for any TI algorithm in practice to deal with the less accurate pseudo-measurements. Therefore, the use of multi-period optimization is likely to be advantageous, as it will be confirmed in a case study in Section IV-C.

Let T denote the number of available snapshots of measurements and pseudo-measurements. Let us use $[t]$ to indicate the data corresponding to each snapshot t, where $t = 1, \ldots, T$. Finally, let $\Psi[t]$ denote the objective value in problem (32) based on the data from snapshot t. The TI problem in a multi-period form can be presented as:

$$\begin{aligned} &\underset{\substack{s_{ij}, ij\in\mathcal{L}\\ b_i, i\in\mathcal{N}}}{\textbf{minimize}} \sum_{t=1}^{T} \gamma[t]\Psi[t], \\ &\textbf{subject to} \text{ Eqs. (5)-(9) at } [t], \quad t = 1, \ldots, T, \\ &\qquad\qquad\quad \text{Eqs. (14)-(23) at } [t], \quad t = 1, \ldots, T, \\ &\qquad\qquad\quad \text{Eqs. (29)-(31) at } [t], \quad t = 1, \ldots, T. \end{aligned} \tag{33}$$

Notation $\gamma[t]$ will be explained at the end of this section. An implicit assumption in problem (33) is that all the measurement snapshots are associated with the same topology. In other words, the topology does *not* change during measurement snapshots 1 to T. The multi-period TI algorithm is reset once a change in topology is detected. This is a reasonable assumption as long as there is a way also to *detect* a change in topology. There already exist several methods in the literature to detect the changes in topology, such as in [41]–[44].

Before ending this section, it is worth clarifying why problem (33) is referred to as a multi-period optimization problem. In principle, topology identification is an estimation problem to estimate the status of switches based on given measurements and pseudo-measurements as inputs. Each snapshot in (33) essentially provides redundancy for such estimation problem. Here, the status of the switches are fixed from one snapshot to another; however, there are random variations in the measurements and pseudo-measurements due to the randomness in the load at each bus. In this regard, each snapshot generates a new random scenario for such random variables; providing the redundancy in estimating the status of the switches. As a result, the problem in (33) can be seen as a multi-scenario-based optimization, where $\gamma[t]$ denotes the probability associated with snapshot t. If the randomness across different snapshot follows a uniform distribution, or simply it is unknown, then one can choose $\gamma[t] = 1/T,\ \forall t = 1, \ldots, T$. Otherwise, a known non-uniform probability distribution can be used.

IV. CASE STUDIES: PART I - IEEE TEST NETWORK

This section demonstrates the effectiveness of the proposed TI method by applying it to the IEEE 33-bus test system. The single line diagram of the feeder is shown in Fig. 4, and the relevant technical data can be found in [24]. The normally closed or normally open status of switches for all line segments are listed in Table II. The normally open switches are the five tie-lines that are shown using dashed lines. Since there are

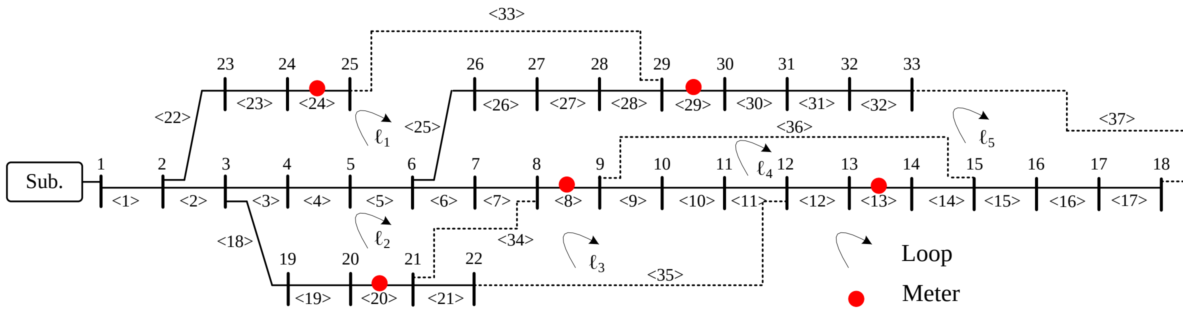

Fig. 4. The IEEE 33-bus test system that is used in our case studies. The tie-lines are shown by dotted lines. The feeder consists of five loops marked as ℓ_1 to ℓ_5. Unless stated otherwise, the line sensors are deployed on line segments <8>, <13>, <20>, <24>, and <29>.

TABLE II
LINE SEGMENTS WITH SWITCHES IN FIG. 4.

Normally Closed	Normally Open
4, 6, 7, 9, 10, 11, 12, 14, 15, 16, 17, 18, 26, 28, 30, 32	33, 34, 35, 36, 37

21 switches in this network, 2^{21} different topologies can be created, among which 65 topology configurations, comprising 50 radial, 10 loop, and 5 island, are selected. The criteria for choosing these topologies is mainly to make sure that all three types of topologies, i.e., radial, loop, and island, are covered while they involve all independent loops in the network.

Once the five tie-lines are in closed status, there exist five loops which are marked as $\ell_1, \ldots, \ell_5$. Recall from Section III-A that each loop has to be monitored with at least one line sensor to meet requirement for TI. An arbitrary choice of locations for the five line current sensors in this case study are as marked using red dots in Fig. 4. In addition to the current measurements, it is assumed that the load injections at all the buses are available through pseudo-measurements.

The network simulation and the implementation of the TI algorithm are done in MATLAB; and the optimization problems in (24), (32), and (33) are solved using `IntLinProg`. In simulations, the measurements are constructed as follows:

$$\tilde{z} = z + e, \tag{34}$$

where $\tilde{z}$ denotes the so-called contaminated measurement vector; z is the true value vector of the measurements; and e is the measurements error vector. It is assumed that the errors are normal distributed with zero mean value and standard deviation vector σ, i.e., $e = \mathcal{N}(0, \sigma^2)$. The standard deviation of the error can be computed as follows [16]:

$$\sigma_i = \frac{z_i \times \eta_i}{3}, \tag{35}$$

which guarantees that 99.7% of the e_i values fall within $\pm\eta_i$ percentage of the true value. Also, the accuracy of the TI algorithm is given in percentage as

$$\text{TI Accuracy} = 100 \times \frac{\text{Total Number of Correct TIs}}{\text{Total Number of TIs}} \tag{36}$$

A. Errors in Line Current Measurements

The line current phasors can be measured by line current sensors, either directly and precisely if the sensor is equipped with GPS, or indirectly and approximately if the sensor is not equipped with GPS; which in that case, it measures the relative phase angle by measuring e-field. Based on the different types of sensors that are available, the error in current magnitude can be 1% to 3%; and the error in current phase angle can be 1° to 5° [45]. Table III shows the result of TI algorithm with considering errors in line current measurements. The results demonstrate a satisfactory performance, with above 99% accuracy in almost all error levels. Thus, it can be concluded that the proposed TI algorithm is *robust* against errors in measurements; regardless of the exact line sensor technology.

TABLE III
TI ACCURACY VS. THE MEASUREMENT ERROR IN MAGNITUDE AND ANGLE

Magnitude		Angle	
Error	TI Accuracy	Error	TI Accuracy
1.0%	99.3%	1°	99.7%
1.5%	99.2%	2°	99.5%
2.0%	99.2%	3°	99.4%
2.5%	99.1%	4°	99.1%
3.0%	99.0%	5°	98.9%

B. Errors in Pseudo-Measurement

In practice, the utility's knowledge about pseudo-measurements is not precise. Pseudo-measurements are often obtained using short-term load forecasting by smart meter data or historical data. The robustness of the proposed TI algorithm is examined against any given level of measurements inaccuracy by using the Monte Carlo method to generate different scenarios, c.f. [46].

Here, there is not any hard requirement on how the pseudo-measurements are obtained. The range of uncertainty associated with pseudo-measurements may vary based the available information. These pseudo-measurements can be obtained through the aggregation of customer smart meter data as long as such metering data is available. Or they can be estimated solely based on the ratings of the load transformers when no metering data is available. A combination of metering data and transformer ratings may also be used. Depending on how the pseudo-measurements are obtained, they may carry a wide range of errors, as low as 10% [47]–[49], when smart meter data is available, or as high as 50%, when pseudo-measurements are synthesized/estimated purely based on load

TABLE IV
TI ACCURACY VS. ERROR IN PSEUDO-MEASUREMENTS

Pseudo-meas. error (%)	10	20	30	40	50
TI accuracy (%)	99.1	98.3	93.1	84.1	77.1
Run Time (s)	0.81	0.94	0.91	0.89	1.00

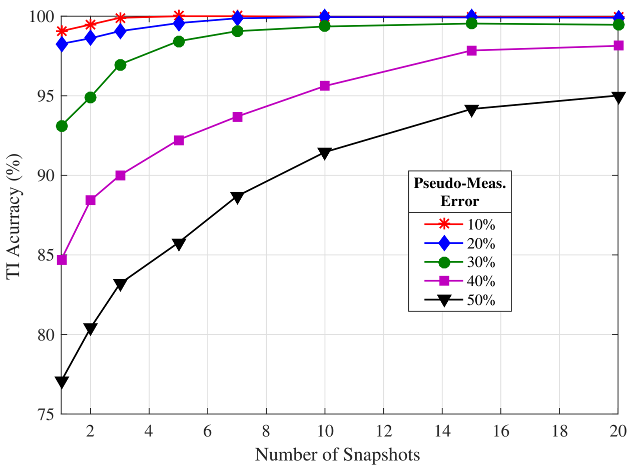

Fig. 5. The accuracy of the multi-period TI algorithm gradually improves as more measurement snapshots become available.

allocation according to the capacities of the low-voltage transformers. Both of the above ranges of errors are examined in the simulations in order to assess the performance of the proposed TI algorithm under a variety of possible scenarios for pseudo-measurements. The results of TI algorithm accuracy versus different levels of error in pseudo-measurements are presented in Table IV. If the error is limited to 30%, then the doing a correct TI is almost guaranteed. However, if the error increases to 50%, then the TI accuracy drops significantly; making the TI results unreliable. Of course, at such high errors, losing accuracy is inevitable.

C. Improved Accuracy with Multi-Period Optimization

Interestingly, the impact of errors is compensated in pseudo-measurements by making use of the additional snapshots of measurement data; together with the multi-period optimization method that was proposed in Section III-C.

In order to test the performance of the multi-period TI algorithm, a sequence of snapshots is generated using the pseudo-measurement generation method in [50], where we generate errors for pseudo-measurement based on a normal probability distribution, with zero mean and the following standard deviation [16]:

$$\sigma_i = \frac{S_i \times \eta_i}{3}. \tag{37}$$

The TI multi-period optimization in (33) is then solved and the TI results are updated every time that a new measurement snapshot becomes available. The results are shown in Fig. 5. The TI accuracy is improved as more measurement snapshots become available. This is true regardless of the level of errors in pseudo-measurements. For example, even with 50% error in pseudo-measurements, where the accuracy of the single-period TI algorithm is 77.1%, the multi-period version of our TI algorithm can enhance accuracy to 95.0% after 20 snapshots.

The computational complexity of solving the multi-period optimization problem in (33) grows as the number of snapshots increases. For example, the multi-period TI algorithm takes about 40 seconds and 125 seconds to run for 10 and 20 snapshots, respectively, which is considerably greater than the single-period TI run time provided in Table IV.

An illustrative example is used to compare the performance of single-period and multi-period TI algorithms on several snapshots. The results are shown in Table V. Note that, the true topology remains the same in all 15 snapshots. The single-period TI algorithm is applied *each snapshot separately*. In only four snapshots, i.e., snapshots 6, 9, 11, and 13, the correct topology is identified, as denoted by check marks. In 11 snapshots, the single-period TI algorithm results in incorrect topology identification, as denoted by cross marks. This is due to the change in loading and the large error in pseudo-measurements; which are intentional in this case study.

In contrast, when the multi-period TI algorithm is used, the correct topology is identified as soon as *sufficient* snapshots become available, which is six in this example. Note that, the multi-period TI algorithm is initially the same as the single-period TI method when it is applied to the first snapshot. Then, at the second snapshot, the multi-period TI algorithm is applied to the *combination* of both the first *and* the second measurement snapshots. At the third snapshot, the multi-period TI algorithm is applied to the *combination* of the first, the second, *and* the third measurement snapshots; and so on and so forth. As more measurement snapshots become available beyond the first six snapshots, the multi-period TI method continues to correctly identify the topology during snapshots 6 to 15.

D. MILP versus MINLP

Both MILP and MINLP are categorized as NP-hard problems [51]. Therefore, in principle, they are roughly similar in terms of their computational complexity. However, in practice, it is often very useful to convert an MINLP into an MILP. The reason is the considerable advancements in developing MILP solvers in the past few decades, such as CPLEX [31], also see [52]. In contrast, MINLP solvers are not matured enough, at least not yet, to be considered stable and reliable. In addition, MILP solvers are guaranteed to ultimately provide the exact optimal solution, by using relatively simple methods, such as branch and bound. Importantly, even when an MILP solver is terminated prior to obtaining the exact optimal solution, it can provide a provable bound on optimality gap [53]. When it comes to the proposed TI optimization problems, the original TI formulation in (4) is in fact a *non-convex* MINLP. Non-convex MINLPs are particularly difficult to solve; because even when the integer variables are handled using methods such as branch and bound, what is left to solve in each iteration is still a difficult non-linear non-convex optimization problem [54].

Three solvers were tested to solve the original MINLP problem in (4): NOMAD, SCIP, and BONMIN [55]. The first

TABLE V
COMPARING SINGLE-PERIOD AND MULTI-PERIOD TI ALGORITHMS FOR 15 SNAPSHOTS

Snapshots	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Single-Period TI	✗	✗	✗	✗	✗	✓	✗	✗	✓	✗	✓	✗	✓	✗	✗
Multi-Period TI	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

TABLE VI
RESULTS FOR MINLP TI FOR DIFFERENT INITIAL POINTS

Case	Result	OptGap	Iter	Time (s)	Open Lines
I	Correct	3%	27	100	33, 34, 35, 36, 14
II	Incorrect	87%	1131	1039	12, 28, 33, 34, 35
III	Incorrect	26%	43	386	33, 34, 35, 36, 16
IV	Incorrect	96%	732	438	33, 34, 35, 36, 7
V	Incorrect	76%	548	424	28, 34, 35, 37

two always failed to even find a feasible solution. As for the third one, by assuming 5% error in pseudo-measurements, BONMIN correctly identified 10 out of 65 toplogies. For those 55 topologies that were identified incorrectly, BONMIN neither could find a correct solution nor it could find any feasible solution at all. For instance, the network configuration with open lines <33>, <34>, <35>, <36>, and <14> is tested for several initial points. The results for six different initial points that converged to *some* feasible solutions are shown in Table VI. Only in one case, where the initial point is chosen very close to the optimal solution, the solution of the MINLP formulation was correct. Accordingly, the MILP reformulation is necessary.

E. Performance Comparison

Even though there is a rich literature for topology identification, there does not exist a prior study to address the use of line current sensors for topology identification. This is partly because the type of line current sensors that are of interest in this study are just starting to emerge only recently. With this general note in mind, in this section, the performance comparison is conducted with two references. The first reference is [40]. The method in this reference can in essence support utilizing line current sensors, but it is developed based on circuit connectivity and it does not involve the load flow equations. The results are shown in the second row of Table VII. It can be seen that the method in [40] is highly sensitive to errors in pseudo-measurements.

The second reference is [16]. The method in this reference is *not* directly comparable with our method in this study. Therefore, a somewhat new method that is rather *inspired* by the method in [16] is used. The new method is essentially the same as the proposed method in this study but it uses a different objective function. Specifically, it uses the objective function in [16], which results in a mixed integer quadratic program (MIQP), as opposed to an MILP. The results are shown in the third row of Table VII. It can be seen that the performance of this new method too degrades as the error in pseudo-measurements increases. This happens because minimizing squared errors as in MIQP, will pull the fit towards the outliers, i.e., the inaccurate pseudo-measurements, much more so than minimizing the absolute error as in this paper.

TABLE VII
COMPARING THE PERFORMANCE OF THE MILP-BASED TI ALGORITHMS PROPOSED IN [40] AND [16].

	Pseudo-Measurement Error (%)				
TI Accuracy (%)	10	20	30	40	50
MILP-based	99.9	98.8	94.4	86.5	78.4
[40]	79.8	75.1	74.8	67.5	61.8
MIQP-based [16]	99.2	96.3	87.9	77.2	68.0

TABLE VIII
TI ACCURACY VERSUS ERROR IN PSEUDO-MEASUREMENTS

Pseudo-meas. error (%)	10	20	30	40	50
TI accuracy (%)	99.6	95.4	90.1	83.6	71.2
Run Time (s)	10.1	12.0	15.4	16.8	16.9

Also, the computational time of MIQP is more that MILP. From the viewpoint of a cost-benefit analysis, the original MIQP method in [16] requires using 10 line power sensors that are often expensive and labor-intensive to install, while our method uses 5 line current sensors that are inexpensive and easy to install; yet it obtains almost the same performance. This demonstrates the advantages of our method in a simple cost-benefit analysis.

V. CASE STUDIES: PART II - REAL-LIFE NETWORK

In this section, the TI algorithm is applied to two long and interconnected real-life distribution feeders in Riverside, CA. The two feeders are isolated on two different transformers at the substation; however, they are inter-connected through tie-lines. The two feeders have about 400 nodes and 37 switches. The understudy network consists of 13 loops. It is assumed that 13 line current sensors are installed on these 13 loops, see Section III.A. A total of 20 different topologies are defined. Suppose the feeder-head voltages are measured for both feeders. Accurate models of these two feeders are simulated in CYME [56]. The results for TI accuracy are shown in Table VIII. It can be seen that the TI algorithm can successfully identify all the topologies for reasonably low pseudo-measurement errors. The results in this Table are comparable to those in Table IV for the IEEE test system in Section IV.

VI. ADDITIONAL DISCUSSIONS

A. Impact of Parameter M

In principle, M must be large enough such that for any combination of decision variables, the inequality constraints in (6)-(9) and (15)-(20) remain feasible. Therefore, since the voltage magnitudes are represented in per unit, M should be at least 1. The impact of parameter M on the TI accuracy is shown in Table IX. It can be seen that both $M = 1$ and

TABLE IX
THE IMPACT OF PARAMETER M

Choice of M	1	1.5	2	5	10
TI Accuracy	99.9%	99.9%	99.8%	99.6%	96.3%

TABLE X
TI ACCURACY VS. COMBINED ERROR IN
MEASUREMENTS AND PSEUDO-MEASUREMENTS(%)

		Error in Pseudo-Measurement				
		10%	**20%**	**30%**	**40%**	**50%**
Error in Measuring Current	**2%**	99.8	98.5	94.5	85.8	78.1
	4%	99.6	98.8	93.1	86.1	79.6
	6%	99.6	98.7	94.0	86.6	77.0
	8%	99.1	97.7	92.9	86.6	76.7
	10%	98.3	97.0	92.9	85.5	75.2

$M = 1.5$ are very good choices. As M increases, numerical issues start to show up; becoming severe when M is too large, e.g., when $M = 10$. To be on the safe side, so as to address any possible over-voltage scenario, we recommend to set $M = 1.5$.

B. Errors in Both Measurements and Pseudo-Measurements

In this section, the performance of the TI algorithm is examined for combinations of errors associated with both line current measurements and pseudo-measurements. In this regard, error in line current measurements is expressed in terms of Total Vector Error (TVE), which includes both magnitude and angle errors. The results are shown in Table X. As can be seen, the proposed TI algorithm is affected more by errors in pseudo-measurements than errors in line current measurements. Of course, here we considered much higher errors for pseudo-measurements because it is indeed the case in practice.

C. Further Discussion on Theorem 1

In the proof of Theorem 1, it was inherently assumed that measurements and pseudo-measurements are precise. Therefore, the estimated current associated with the lines that are switched off is precisely zero, which can be used to determine which lines are switched off; thereby identifying the topology. However, in a non-ideal situation, where the measurements and/or pseudo-measurements carry error, the numerically estimated current associated with a line that is switched off could be non-zero. This issue is addressed by using binary variables and by minimizing the absolute value of errors in the objective function of the TI optimization problem.

As a matter of fact, the proposed theorem determines a *requirement* for the placement of the line current sensors to support the TI application, rather than indicating the exact locations. That is, there are a large number of line current sensor placement options that would satisfy the requirements in the theorem, as long as there is at least one line sensor at each independent loop. Sensor placement methods however are often developed for specific applications with particular goals. For instance, in [57], authors developed a sensor placement method to improve the state estimation; also, several line sensors placement methods have been reported for outage detection and fault location, e.g., see [58], [59]. As a result, the proposed theorem can be used by itself or in conjunction with other sensor placement methods, as long as the requirement to place at least one sensor in each independent loop is satisfied.

VII. CONCLUSIONS

A distribution-level topology identification method is proposed that is concerned with utilizing the type of measurements that come from an emerging class of low-cost non-contact line current sensors. Three key challenges are addressed: 1) designing a TI algorithm that is compatible with the limitations of the aforementioned line current sensor technologies, such as their inability to measure voltage and power; 2) maintaining a low and tractable level of computational complexity for the TI algorithm; and 3) tackling the various sources of error in measurements and pseudo-measurements that are inevitable in the context of the study in this paper. The third challenge was particularly addressed by developing a novel multi-period TI algorithm which uses multiple measurement snapshots, beyond the moment that a change in topology is detected, in order to gradually improve the TI accuracy and robustness. Both the single-period and multi-period TI algorithms can identify all the possible topologies, including radial, loop, and island configurations. This extends the application of the TI algorithms to identify switch malfunctions and to detect outage. The performance of the proposed TI algorithms are studied in IEEE test cases and also a test case based on a feeder in Riverside, CA. Furthermore, observability analysis was made with respect to the number and location of the line current sensors that are needed to achieve accurate TI performance.

REFERENCES

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

[2] Y. Sharon *et al.*, "Topology identification in distribution network with limited measurements," in *Proc. of IEEE PES Innovative Smart Grid Technologies*, Washington, DC, 2012, pp. 1–6.

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

[4] V. Kekatos and G. B. Giannakis, "Joint power system state estimation and breaker status identification," in *Proc. of North American Power Symposium, 2012*, Champaign, IL, 2012, pp. 1–6.

[5] R. Arghandeh, M. Gahr, A. von Meier, G. Cavraro, M. Ruh, and G. Andersson, "Topology detection in microgrids with micro-synchrophasors," in *Proc. of IEEE PES General Meeting*, Denver, CO, 2015, pp. 1–5.

[6] D. Deka, S. Backhaus, and M. Chertkov, "Structure learning in power distribution networks," *IEEE Trans. on Control of Network Systems*, vol. 5, no. 3, pp. 1061–1074, Sept. 2018.

[7] G. Cavraro, V. Kekatos, and S. Veeramachaneni, "Voltage analytics for power distribution network topology verification," *IEEE Trans. on Smart Grid*, vol. 10, no. 1, pp. 1058–1067, Jan. 2019.

[8] S. Bolognani, N. Bof, D. Michelotti, R. Muraro, and L. Schenato, "Identification of power distribution network topology via voltage correlation analysis," in *Proc. of 52nd Annual Conference on Decision and Control*, Florence, Italy, 2013.

[9] G. Cavraro and R. Arghandeh, "Power distribution network topology detection with time-series signature verification method," *IEEE Trans. on Power Systems*, July 2018.

[10] S. Park, E. Lee, W. Yu, H. Lee, and J. Shin, "State estimation for supervisory monitoring of substations," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 406–410, Mar. 2013.

[11] S. J. Pappu *et al.*, "Identifying topology of low voltage distribution networks based on smart meter data," *IEEE Trans. on Smart Grid*, vol. 9, no. 5, pp. 5113–5122, Sept. 2018.

[12] Y. Liao, Y. Weng, M. Wu, and R. Rajagopal, "Distribution grid topology reconstruction: An information theoretic approach," in *Proc.of North American Power Symposium*, Charlotte, NC, 2015, pp. 1–6.

[13] Y. Gao, Z. Zhang, W. Wu, and H. Liang, "A method for the topology identification of distribution system," in *Proc. of IEEE PES General Meeting*, Vancouver, BC, 2013, pp. 1–5.

[14] E. Caro, A. J. Conejo, and A. Abur, "Breaker status identification," *IEEE Trans. on Power Systems*, vol. 25, no. 2, pp. 694–702, Dec. 2010.

[15] R. Sevlian and R. Rajagopal, "Distribution system topology detection using consumer load and line flow measurements," *CoRR*, vol. abs/1503.07224, 2015. [Online]. Available: http://arxiv.org/abs/1503.07224

[16] Z. Tian, W. Wu, and B. Zhang, "A mixed integer quadratic programming model for topology identification in distribution network," *IEEE Trans. Power Syst*, vol. 31, no. 1, pp. 823–824, Jan. 2016.

[17] H. Mohsenian-Rad, E. Stewart, and E. Cortez, "Distribution synchrophasors: pairing big data with analytics to create actionable information," *IEEE Power and Energy Magazine*, vol. 16, no. 3, pp. 26–34, May 2018.

[18] Lindsey Overhead Sensor. [Online]. Available: http://lindsey-usa.com/wp-content/uploads/2018/04/09B-002-GEN2-April-2018.pdf.

[19] Sentient MM3 Overhead Line Sensor. [Online]. Available: https://www.sentient-energy.com/products/mm3-intelligent-sensor-2.

[20] InHand Overhead Line Sensor. [Online]. Available: https://www.inhandnetworks.com/solutions/IWOS/overhead-line-sensor.html.

[21] M. Farajollahi, M. Fotuhi-Firuzabad, and A. Safdarian, "Optimal placement of sectionalizing switch considering switch malfunction probability," *IEEE Trans. on Smart Grid*, Jan. 2019.

[22] M. Ruh, A. Borer, and G. Andersson, "Condensed grid topology detection for a fully transparent distribution management system," in *Proc. of 17th Power Systems Computation Conference*, Stockholm, Sweden, 2011, p. 10.

[23] I. Griva *et al.*, "Linear and nonlinear programming: second edition," 2009.

[24] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Trans. on Power delivery*, vol. 4, no. 2, pp. 1401–1407, Apr. 1989.

[25] S. Bolognani, R. Carli, G. Cavraro, and S. Zampieri, "Distributed reactive power feedback control for voltage regulation and loss minimization," *IEEE Trans. on Automatic Control*, vol. 60, no. 4, pp. 966–981, Oct. 2015.

[26] E. Samani and F. Aminifar, "Tri-level robust investment planning of ders in distribution networks with ac constraints," *IEEE Trans. on Power Systems*, to be published.

[27] A. Garces, "A linear three-phase load flow for power distribution systems," *IEEE Trans. on Power Systems*, vol. 31, no. 1, pp. 827–828, Jan. 2016.

[28] V. A. Zorich, *Mathematical Analysis I*. Springer, 2015.

[29] R. B. Ash and W. Novinger, *Complex Variables*. Dover Books on Mathematics, 2015.

[30] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[31] I. Cplex, "11.0 user's manual," *ILOG SA, Gentilly, France*, p. 32, 2007.

[32] A. Shahsavari, M. Farajollahi, E. Stewart, E. Cortez, and H. Mohsenian-Rad, "A machine learning approach to event analysis in distribution feeders using distribution synchrophasors," *IEEE Trans. on Smart Grid*, to be published.

[33] M. Jamei, A. Scaglione, C. Roberts, E. Stewart, S. Peisert, C. McParland, and A. McEachern, "Anomaly detection using optimally-placed μpmu sensors in distribution grids," *IEEE Trans. on Power Systems*, July 2017.

[34] M. Farajollahi, A. Shahsavari, E. M. Stewart, and H. Mohsenian-Rad, "Locating the source of events in power distribution systems using micro-pmu data," *IEEE Trans. on Power Systems*, vol. 33, no. 6, pp. 6343–6354, Nov. 2018.

[35] A. Shahsavari, M. Farajollahi, and H. Mohsenian-Rad, "Individual load model parameter estimation in distribution systems using load switching events," *IEEE Trans. on Power Systems*, to be published.

[36] I. D. Mayergoyz and W. Lawson, *Basic electric circuit theory*. Gulf Professional Publishing, 1997.

[37] R. M. Foster, "Geometrical circuits of electrical networks," *Electrical Engineering*, vol. 51, no. 1, pp. 43–43, Jan. 1932.

[38] J. B. Kruskal, "On the shortest spanning subtree of a graph and the traveling salesman problem," in *Proc. of the American Mathematical society*, vol. 7, no. 1, 1956, pp. 48–50.

[39] H. Singh and F. L. Alvarado, "Weighted least absolute value state estimation using interior point methods," *IEEE Trans. on Power Systems*, vol. 9, no. 3, pp. 1478–1484, Aug. 1994.

[40] A. Abur, D. Shirmohammadi, and C. Cheng, "Estimation of switch statuses for radial power distribution systems," in *Proc. of International Symposium on Circuits and Systems*, Seattle, WA, 1995.

[41] X. Jiang, "Real-time power system topology change detection and identification," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2016.

[42] S. V. Wiel, R. Bent, E. Casleton, and E. Lawrence, "Identification of topology changes in power grids using phasor measurements," *Applied Stochastic Models in Business and Industry*, vol. 30, no. 6, pp. 740–752, Nov. 2014.

[43] 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 Trans. on Power Systems*, vol. 25, no. 3, pp. 1329–1336, Aug. 2010.

[44] G. N. Korres and N. M. Manousakis, "A state estimation algorithm for monitoring topology changes in distribution systems," in *Proc. of IEEE PES General Meeting*, San Diego, CA, 2012, pp. 1–8.

[45] Radial distribution Test Feeders, Distribution System Analysis Subcommittee Rep. [Online]. Available: http://ewh.ieee.org/soc/pes/dsacom/testfeeders.html.

[46] C. Z. Mooney, *Monte carlo simulation*. Sage Publications, 1997.

[47] F. Ni, P. Nguyen, J. Cobben, H. van den Brom, and D. Zhao, "Uncertainty analysis of aggregated smart meter data for state estimation," in *Proc. of Applied Measurements for Power Systems*, Aachen, Germany, 2016, pp. 1–6.

[48] J. Liu, J. Tang, F. Ponci, A. Monti, C. Muscas, and P. A. Pegoraro, "Trade-offs in pmu deployment for state estimation in active distribution grids," *IEEE Trans. on Smart Grid*, vol. 3, no. 2, pp. 915–924, June 2012.

[49] J. Wu, Y. He, and N. Jenkins, "A robust state estimator for medium voltage distribution networks," *IEEE Trans. on Power Systems*, vol. 28, no. 2, pp. 1008–1016, May 2013.

[50] C. S. Cheng and D. Shirmohammadi, "A three-phase power flow method for real-time distribution system analysis," *IEEE Trans. on Power Systems*, vol. 10, no. 2, pp. 671–679, May 1995.

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

[52] R. Bixby and E. Rothberg, "Progress in computational mixed integer programming—a look back from the other side of the tipping point," *Annals of Operations Research*, vol. 149, no. 1, pp. 37–41, 2007.

[53] J. T. Linderoth and M. W. Savelsbergh, "A computational study of search strategies for mixed integer programming," *INFORMS Journal on Computing*, vol. 11, no. 2, pp. 173–187, May 1999.

[54] A. Sadeghi-Mobarakeh and H. Mohsenian-Rad, "Optimal bidding in performance-based regulation markets: An mpec analysis with system dynamics," *IEEE Trans. on Power Systems*, vol. 32, no. 2, pp. 1282–1292, June 2017.

[55] P. Bonami, L. T. Biegler, A. R. Conn, G. Cornuéjols, I. E. Grossmann, C. D. Laird, J. Lee, A. Lodi, F. Margot, N. Sawaya, *et al.*, "An algorithmic framework for convex mixed integer nonlinear programs," *Discrete Optimization*, vol. 5, no. 2, pp. 186–204, 2008.

[56] CYME USER'S GUIDE REFERENCE MANUAL. [Online]. Available: http://www.cyme.com/software/cymdist/.

[57] N. Nusrat, M. Irving, and G. Taylor, "Novel meter placement algorithm for enhanced accuracy of distribution system state estimation," in *Proc. of IEEE PES General Meeting*, San Diego, CA, 2012, pp. 1–8.

[58] R. A. Sevlian, Y. Zhao, R. Rajagopal, A. Goldsmith, and H. V. Poor, "Outage detection using load and line flow measurements in power distribution systems," *IEEE Trans. on Power Systems*, vol. 33, no. 2, pp. 2053–2069, Mar. 2017.

[59] M. Farajollahi, M. Fotuhi-Firuzabad, and A. Safdarian, "Deployment of fault indicator in distribution networks: A mip-based approach," *IEEE Trans. on Smart Grid*, vol. 9, no. 3, pp. 2259–2267, May 2018.

Mohammad Farajollahi (S'15) received the B.Sc. degree in Electrical Engineering from University of Tehran, Tehran, Iran, in 2014, the M.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, in 2016, and is currently pursuing the Ph.D. degree at University of California, Riverside, CA, U.S. His research interests include power system planning, operation, relability, as well as optimization. He is specifically working on applications of micro-PMUs and data analysis in distribution system monitoring.

Alireza Shahsavari (S'17) received the M.Sc. degree in electrical engineering - power systems from the University of Tehran, Tehran, Iran, in 2013, and the Ph.D. degree in electrical and computer engineering at the University of California, Riverside, CA, USA, in 2019. His research interests include the data analysis in power systems and smart grid, distribution system planning and operation, and power system reliability.

Hamed Mohsenian-Rad (S'04–M'09–SM'14) received the Ph.D. degree in electrical and computer engineering from the University of British Columbia Vancouver, BC, Canada, in 2008. He is currently an Associate Professor of electrical engineering at the University of California, Riverside, CA, USA. His research interests include modeling, data analysis, and optimization of power systems and smart grids. He received the National Science Foundation CAREER Award, the Best Paper Award from the IEEE Power and Energy Society General Meeting, and the Best Paper Award from the IEEE International Conference on Smart Grid Communications. He serves as an Editor for the IEEE TRANSACTIONS ON SMART GRID and the IEEE POWER ENGINEERING LETTERS.