

# Reconstruction of Low-Voltage Networks with Limited Observability

Daniele MARULLI, Sébastien MATHIEU, Amina BENZERGA, Antonio SUTERA, Damien ERNST

*Department of Electrical Engineering and Computer Science*

*University of Liège*

Liège, Belgium

abenzerga, a.sutera, dernst@uliege.be

**Abstract**—This work addresses the problem of reconstructing topology and cable parameters of three-phase low-voltage networks when no a priori information about them is provided and not all the nodes in the grid are equipped with smart meters. This paper presents a methodology to obtain an estimation of the electrical model of each phase of the network, analysing voltage and current time-series measurements provided by the available meters. The proposed methodology involves an iterative algorithm developed to tackle the network reconstruction problem when unmetered nodes are located reasonably far from each other. The algorithm is tested on a 30-node network with different sets of metered nodes, providing relevant solutions in most scenarios having more than 80% of metered nodes.

**Index Terms**—network reconstruction, low-voltage network

## I. INTRODUCTION

With an ever-increasing penetration of distributed energy resources and electric vehicles, the efficient management of distribution networks becomes more complicated and distribution system operators (DSOs) might encounter serious difficulties in guaranteeing the safety of their network at Low-Voltage (LV) levels. In order to implement effective preventive or corrective measures, DSOs need to be able to assess the system's response to different realistic scenarios. These kinds of analyses are usually performed through power flow studies which require accurate information about the topology and physical characteristics of the network. However, most DSOs lack a reliable electrical model of their LV network. This lack of knowledge can hinder an efficient management and development of the system. Network reconstruction is a mathematical process that allows one to deduce this information. Effective network reconstruction methods for LV networks are essential for the development of smarter grids [1]. This work presents a methodology that allows one to infer the network topology (i.e. the connections between nodes) and to assess cable parameters of a LV network from time-series measurements (i.e. currents and voltages) provided by a limited number of smart-meters in the grid.

## II. LITERATURE REVIEW

In transmission systems, topology and cable parameters are usually known [1] and, thanks to a large number of

This research is supported by the public service of Wallonia within the framework of the Silver project.

978-1-6654-4875-8/21/\$31.00 ©IEEE

measurement devices, changes can be detected through well-established state estimation methods. Unfortunately, the same does not apply to LV distribution networks. With the increasing importance and complexity of distribution systems, the reconstruction problem of LV networks has gained more attention in scientific literature. Papers [1] and [2] focus on identifying the network topology when limited information is available. Algorithms that aim, as this work does, to simultaneously identify both network topology and cable parameters, are presented in [3]–[6]. Authors of [3] use the evaluation of voltage sensitivities with respect to active and reactive power injections to reconstruct the topology of small networks, assuming that only specific cable types and lengths are used for the cables. The identification problem in [4] takes the name of an inverse power flow problem, where the system admittance matrix is found by solving an unconstrained least square problem. The case with non-measured nodes in the grid, also referred to as hidden nodes, is also tackled in paper [4], both for meshed and radial topologies, with the assumption that these hidden nodes have zero net current injections. The inverse power flow problem is extended to poly-phase systems in [5], with the full-observability assumption. Finally, an algorithm to jointly estimate both admittances and topology, assuming that the measurements for all the non-zero power injecting nodes are available, is presented in [6].

## III. PROBLEM STATEMENT

Let us consider a three-phase four-wire LV radial distribution network with customer connected to the feeder through a single phase and the neutral wire. The three phases are denoted by  $a, b$  and  $c$ , and the neutral wire by  $d$ . Let  $\mathcal{P} = \{a, b, c\}$  and  $\mathcal{N}$  be the set of phase indexes and the set of nodes of the network, respectively. The system is observed over a finite time period discretized into  $T$  intervals of length  $\Delta t$ . The value of a variable at certain time  $t \in \{1, \dots, T\}$  is accessed with an additional subscript, e.g.  $x_{i,t}$  refers to the variable  $x_i$  at time  $t$ . The absence of such subscript denotes the entire time-series. Each phase of the network is considered separately as shown in Figure 1.

Phase  $p \in \mathcal{P}$  of the network is modelled as a rooted tree  $\mathcal{G}^p = (\mathcal{N}^p, \mathcal{E}^p)$  where  $\mathcal{N}^p$  is the set of nodes connected to phase  $p$  and  $\mathcal{E}^p$  the set of edges. Let  $e = (i, j) \in \mathcal{E}^p$  represent the edge connecting node  $i$  to node  $j$ . Let  $V_{n,t}^p$  and  $I_{n,t}^p$  be

the amplitudes of the voltage and the current injection of node  $n \in \mathcal{N}^p$  at phase  $p$  and time  $t \in \{1, \dots, T\}$ . Let  $\mathcal{M}^p \subseteq \mathcal{N}^p$  be the set of nodes in  $\mathcal{G}^p$  for which a meter is available. These meters are assumed to provide both voltage and current magnitude time-series measurements of the nodes to which they are connected. The set of hidden nodes is denoted by  $\mathcal{H}^p = \mathcal{N}^p \setminus \mathcal{M}^p$ .

The aim of a network identification problem is to retrieve, for each phase  $p$ , a model of the network describing the network topology,  $\mathcal{G}^p = (\mathcal{N}^p, \mathcal{E}^p)$ , impedance magnitudes  $Z_e, \forall e \in \mathcal{E}^p$  and the hidden node current injections  $I_h^p \forall h \in \mathcal{H}^p$ , assuming that:

- The topology of the network does not change during the observation period
- Customers are connected to the main feeder through single-phase two-wire connection
- The phase of the network to which metered customers are connected is provided (for instance using [7])
- At least one node at every phase of every feeder of the network is equipped with a meter
- Each phase of the MV/LV substation is metered.

#### IV. IDENTIFICATION ALGORITHM

The identification algorithm consists of three main parts, namely topology estimation, topology validation and hidden node detection, that are applied to each phase  $p \in \mathcal{P}$  independently. Let  $\hat{\mathcal{N}}^p \subseteq \mathcal{N}^p$  be the pool of nodes made available to the identification algorithm.  $\hat{\mathcal{N}}^p$  initially corresponds to the set of observed nodes  $\mathcal{M}^p$ .

##### A. Topology Estimation

The first step of the algorithm consists of the estimation of a rooted tree  $\hat{\mathcal{G}}^p = (\hat{\mathcal{N}}^p, \hat{\mathcal{E}}^p)$  with a topology as close as possible to  $\mathcal{G}^p = (\mathcal{N}^p, \mathcal{E}^p)$ . This operation is carried out using correlation analysis on voltage measurements to infer proximity between nodes. Correlation-based approaches have already been proved successful both in phase [7] and topology [1] identification methods. Load and production profiles at different nodes of the LV network can present similar patterns owing to comparable residential occupancy profiles and weather

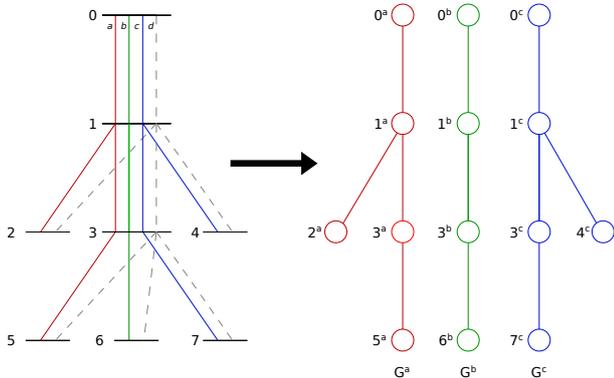


Figure 1: Graphical representation of  $\mathcal{G}^a$ ,  $\mathcal{G}^b$ ,  $\mathcal{G}^c$  for a 8-node network.

conditions. To be sure that this does not affect the results of the voltage correlation analysis, voltage time-series are pre-processed applying a high-pass filter, as suggested in [1]. Let  $w_{ij}$  be the Pearson Correlation Coefficient between the filtered voltage time-series of nodes  $i$  and  $j$ . Let  $\hat{\mathcal{G}}_w^p$  be a weighted complete graph built on  $\hat{\mathcal{N}}^p$ , where the edge  $(i, j)$  weight is equal to  $w_{ij}$ . The estimated topology  $\hat{\mathcal{G}}^p = (\hat{\mathcal{N}}^p, \hat{\mathcal{E}}^p)$  is obtained computing the *maximum spanning tree* on  $\hat{\mathcal{G}}_w^p$ .

##### B. Topology Validation

The second process checks each edge  $e = (i, j) \in \hat{\mathcal{E}}^p$  and suggests where missing nodes and wrong connections, if any, are. This process is performed evaluating the estimated impedance magnitude time-series  $\hat{Z}_e^p = \{\hat{Z}_{e,1}^p, \dots, \hat{Z}_{e,T}^p\}$  for each edge  $e = (i, j) \in \hat{\mathcal{E}}^p$  as,  $\forall t \in \{1, \dots, T\}$ :

$$\hat{Z}_{e,t}^p = \frac{V_{j,t}^p - V_{i,t}^p}{\hat{I}_{e,t}^p} \quad (1)$$

where the estimated current  $\hat{I}_{e,t}^p$  flowing in  $e$  is given by

$$\hat{I}_{e,t}^p = \sum_{n \in \hat{\mathcal{D}}_e^p} I_{n,t}^p \quad (2)$$

where  $\hat{\mathcal{D}}_e^p$  is the set of downstream nodes with respect to edge  $e$ . Since impedances are constants, the values  $\hat{Z}_e^p$  should be close to the same constant value. Edge  $e$  is considered valid if the relative standard deviation of time-series  $\hat{Z}_e^p$ ,  $RSD(\hat{Z}_e^p)$ , is less than an arbitrary threshold  $\lambda$ , e.g. determined by statistical tests for the largest accepted standard deviation.

##### C. Hidden Node Detection

The third process of the algorithm exploits the results provided from the previous steps to find the location of hidden nodes. This step is based on the assumption that unmetered nodes are sufficiently far apart from each other.

**Assumption 1.** *The distance on  $\mathcal{G}^p$ , defined as the minimum number of edges that connect two nodes, between any pair of hidden nodes  $i$  and  $j$  is greater than or equal to three.*

Consider a hidden node  $X \in \mathcal{H}^p$ . If  $X$  is a node with a non-zero net power injection, the current flowing through the path that connects it to the root computed by the topology validation step is not the correct one, since the contribution of the unobserved  $I_X^p$  is missing. This causes the rejection of the edges in that path, as shown in Figure 2. Note that since node 5 is hidden and it is not a terminal node, the algorithm mistakenly detects an edge connecting node 3 to node 6. This suggests that whenever there is a path of unvalidated edges in  $\hat{\mathcal{G}}^p$ , a missing node  $X$  is near the leaf node of such a path. The algorithm focuses on one of the hidden nodes at a time to estimate its location and, if possible, its voltage and current injection.

Let  $A \in \hat{\mathcal{N}}^p$  be the node, among all the nodes that are connected to a rejected edge, with the longest path to the root (node 6 in Figure 2). Let nodes  $B$  and  $C$  be the parent and grandparent node of  $A$ , respectively. Node  $X \in \mathcal{H}^p$  is

assumed to be adjacent to  $A$ . Three topological configurations can occur, as shown in Figure 3. The hidden node detection step examines these configurations to assess which suits best. Let  $\hat{I}_{A^+}^p = \hat{I}_{AB}^p$  be the contribution to the estimated  $\hat{I}_{BC}^p$  flowing through  $BC$  of  $A$  and its descendant nodes, and

$$\hat{I}_{B^+}^p = \hat{I}_{AB}^p - \hat{I}_{BC}^p \quad (3)$$

be the rest of the current. In order to detect the correct location of  $X$ , the algorithm solves three optimization problems:

**Configuration ‘‘Bridge’’:**

$$\delta^b = \min \sum_{t \in \mathcal{T}} |V_{B,t} - V_{C,t} - \hat{Z}_{CB}^p(\hat{I}_{A^+,t}^p + \hat{I}_{B^+,t}^p + \hat{I}_{X,t}^p)| \quad (4a)$$

subject to,  $\forall t \in \mathcal{T}$

$$V_{B,t} = V_{A,t} - \hat{Z}_{XA}^p \hat{I}_{A^+,t}^p - \hat{Z}_{BX}^p(\hat{I}_{A^+,t}^p + \hat{I}_{X,t}^p) \quad (4b)$$

with  $(\hat{I}_X^p, \hat{Z}_{XA}^p, \hat{Z}_{BX}^p, \hat{Z}_{CB}^p) \in \mathbb{R}^4$ .

**Configuration ‘‘Leaf’’:**

$$\delta^l = \min \sum_{t \in \mathcal{T}} |V_{B,t} - V_{C,t} - \hat{Z}_{CB}^p(\hat{I}_{A^+,t}^p + \hat{I}_{B^+,t}^p + \hat{I}_{X,t}^p)| \quad (5a)$$

subject to,  $\forall t \in \mathcal{T}$ ,

$$V_{B,t} = V_{A,t} - \hat{Z}_{BA}^p(\hat{I}_{A^+,t}^p + \hat{I}_{X,t}^p) \quad (5b)$$

with  $(\hat{I}_X^p, \hat{Z}_{BA}^p, \hat{Z}_{CB}^p) \in \mathbb{R}^3$ .

**Configuration ‘‘Common parent’’:**

$$\delta^{cp} = \min \sum_{t=0}^T |(V_{A,t}^p - \hat{Z}_{XA}^p \hat{I}_{A^+,t}^p) - (V_{B,t}^p - \hat{Z}_{XB}^p \hat{I}_{B^+,t}^p)| \quad (6)$$

with  $(\hat{Z}_{XA}^p, \hat{Z}_{XB}^p) \in \mathbb{R}^2$ .

Such problems are formulated by exploiting the currents and voltages relationship occurring in each configuration. Once problems (4 - 6) have been solved, the algorithm selects the configuration with the smallest  $\delta \in \{\delta^b, \delta^l, \delta^{cp}\}$ . Depending on the predicted location of  $X$ , the pool of nodes  $\hat{\mathcal{N}}^p$  is updated accordingly. If the algorithm picks the bridge configuration, an additional node  $X$  is added to  $\hat{\mathcal{N}}^p$ . The estimated current injection magnitude  $\hat{I}_X^p$  is extracted from solution of (4), along with the values of  $\hat{Z}_{XA}^p$  and  $\hat{Z}_{BX}^p$ . The estimation of the voltage time-series  $V_X^p$  is computed as:

$$\hat{V}_{X,t}^p = V_{A,t}^p - \hat{Z}_{XA}^p \hat{I}_{A^+,t}^p \quad \forall t \in \{1, \dots, T\} \quad (7)$$

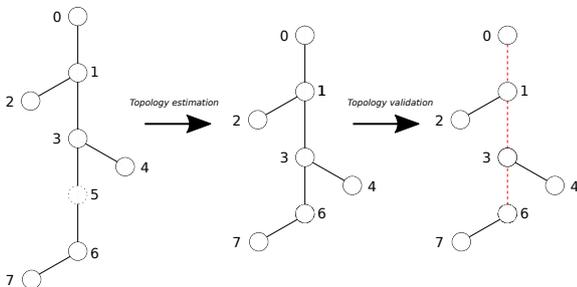


Figure 2: First two steps of the algorithm with node 5 as hidden node. Dashed edges are rejected in validation.

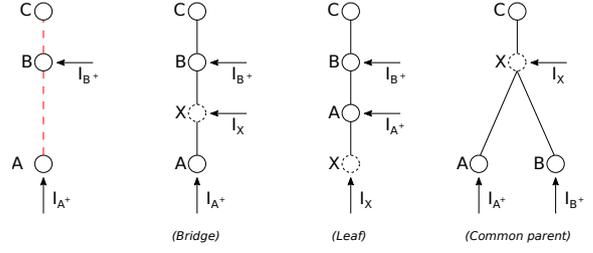


Figure 3: Possible locations for a hidden node  $X$  adjacent to node  $A$ .

If the leaf configuration is selected, the current injection  $\hat{I}_X^p$  is extracted from the solution of (5), while the voltage time-series of  $X$  cannot be estimated. To proceed with the identification process, node  $A \in \hat{\mathcal{N}}^p$  is substituted by an auxiliary node  $X'$  with the same voltage as  $A$  and its current injection is given by  $\hat{I}_{X'}^p = \hat{I}_A^p + \hat{I}_X^p$ . This allows one to fix the current flowing in the path to the root of  $A$  for the next validation step and to continue the identification process. If the common parent configuration is selected, this implies that two impedances  $\hat{Z}_{XA}^p$  and  $\hat{Z}_{XB}^p$  could be found, which leads to a common voltage  $\hat{V}_X^p$ . The voltage magnitude associated with  $\hat{V}_X^p$  is computed as the average voltage given by the two voltage drops:

$$\hat{V}_{X,t}^p = (V_{A,t}^p - \hat{Z}_{XA}^p \hat{I}_{A^+,t}^p + V_{B,t}^p - \hat{Z}_{XB}^p \hat{I}_{B^+,t}^p)/2 \quad (8)$$

A node  $X$  is added to  $\hat{\mathcal{N}}^p$  with voltage  $\hat{V}_X^p$  and a zero net current injection, since (6) voltages and currents in nodes  $A$  and  $B$  do not provide information to estimate any potentially missing current injection. If  $X$  is indeed a net zero-power injecting node, edge  $CX$  is accepted in the topology by the validation step. Otherwise, the correct  $\hat{I}_{CX}^p$  is obtained in the next validation iteration by solving problem (5), since node  $X$  is added in a leaf configuration.

Whenever Assumption 1 is not valid, the hidden node detection step might fail to find a node  $X$  that corresponds to a hidden node in the actual network, since none of the topological configurations of Figure 3 may be correct. Before updating the node pool with  $X$ , the algorithm checks if  $RSD(Z_{XC}^p) < \lambda$ . If not, the incorrect  $X$  is not added to the network since it would compromise the rest of the identification process. In this case, the algorithm proceeds considering the edge connecting  $A$  to  $B$  as a valid edge.

Once the pool of nodes  $\hat{\mathcal{N}}^p$  has been updated, topology validation and estimation steps are processed again. The algorithm is performed until all the edges in  $\hat{\mathcal{G}}^p$  are labelled as valid. Finally, the estimation of the edge impedance magnitude is given by the mean value of  $\hat{Z}_e^p$  in Equation (1).

## V. RESULTS

The developed algorithm is tested on a three-phase, four-wire, 30-node radial distribution network whose topology is known and shown in Figure 4a. Node S0 is the MV/LV substation, nodes from H1 to H24 are single-phase customers, and nodes from F1 to F6 are their connection point to the

main feeder. Household active power profiles are constructed from time-series measurements of residential Belgian smart-meters providing historical voltage and current amplitudes. Some households are equipped with PV units resulting in peaks and time patterns in power profiles which vary widely from a household to another. Two different cable types and various cable length have been selected for the modeling of the main feeder and the laterals. Voltages and currents in the grid are computed through a detailed unbalanced load flow algorithm. Voltage and current amplitude measurements are available for 15 days, at a 15-minutes resolution, resulting in 1440 time steps. The tolerated relative standard deviation for impedance magnitudes is arbitrarily set to  $\lambda = 0.1$ .

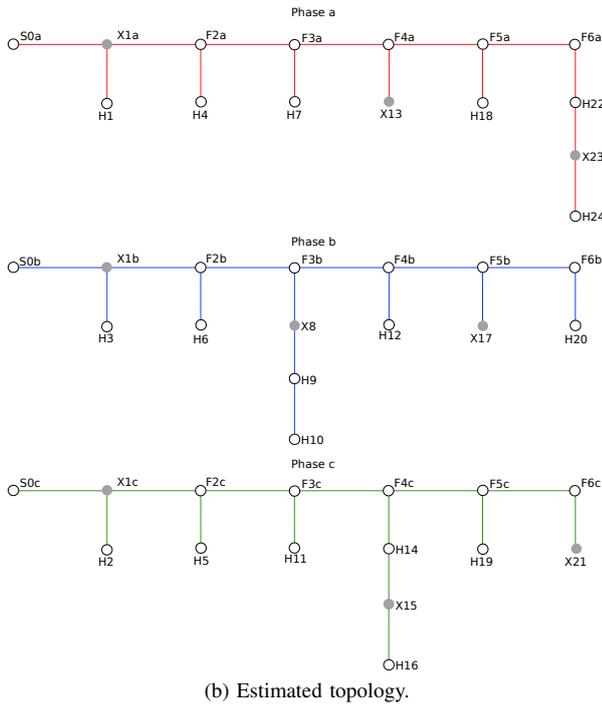
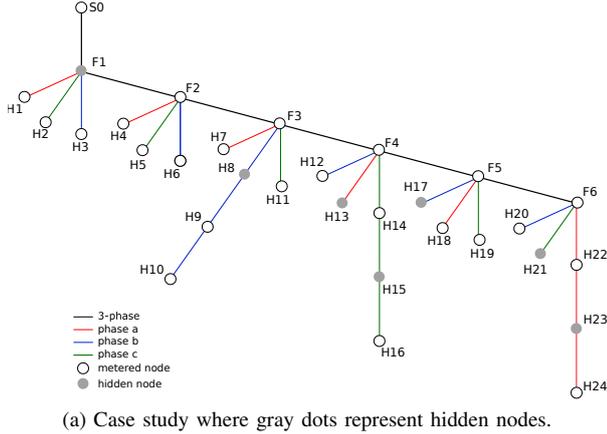


Figure 4: Case study and estimated topology.

Let us first focus on a case where Assumption 1 is satisfied. The case where nodes F1, H8, H13, H15, H17, H21 and H23 are hidden nodes has been selected. In this scenario the

algorithm was able to produce an exact representation of the network topology for each phase, as can be seen in Figure 4b. Missing nodes have been correctly detected and placed at their correct location.

Impedances are estimated for all the cables except for F6-H21, F5-H17 and F4-H13. The impedance magnitude of these cables cannot be determined because nodes H21, H17 and H13 were hidden leaf nodes, therefore the voltage drop of these edges cannot be computed. The estimated impedance magnitude for some cables is shown in Table I. This table shows that the impedance magnitude error is on average of 11%. Note that these impedance amplitudes could easily provide the resistive and reactive part of the impedance using common impedances ratios of low-voltage lines or cables.

Table I: Impedance modulus estimation of some cables of the case study.

Branch	Value	Estimated
<b>S0-F1</b>	0.0481	0.0440
<b>F1-F2</b>	0.0541	0.0503
<b>F1-H1</b>	0.0747	0.0661
<b>F2-H6</b>	0.0335	0.0291
<b>F3-H7</b>	0.0277	0.0241
<b>F4-H12</b>	0.0376	0.0331
<b>F4-F5</b>	0.0604	0.0557
<b>F5-H18</b>	0.0511	0.0431
<b>F6-H22</b>	0.0510	0.0451
<b>H23-H24</b>	0.0353	0.0311

To further assess the performances of the algorithm, a total of 830 cases have been examined, each of which is associated to a different set of hidden nodes. Among them, 30 cases represented the cases with one hidden node. The remaining 800 cases were equally divided in 8 groups with an increasing number of hidden nodes that ranges from 2 to 9. When a node is hidden, the identification process is provided with no information about their presence, voltages and current injections. If a hidden node is a feeder node, e.g. F1, none of the measurements for three phases is provided to the algorithm. The sets of hidden nodes have been randomly selected, implying that Assumption 1 is not valid for most of cases. In particular, none of the scenarios with more than 7 hidden nodes satisfies Assumption 1. Two metrics are considered to analyze the solutions provided by the algorithm in all 830 cases. The first metric is the percentage of edges that is correctly detected by the algorithm. This value assesses the quality of the estimated topology. Note that for this evaluation, a three-phase edge is considered as detected only if the nodes at its ends are directly connected by an edge in  $\hat{G}^a$ ,  $\hat{G}^b$  and  $\hat{G}^c$ . Figure 5 shows that the algorithm identifies the correct topology in most of the cases with less than 10% of hidden nodes. For cases with 10% to 20% hidden nodes, Assumption 1 is not valid in most cases. Nevertheless, the algorithm still performs well, accurately detecting more than 80% of network edges in most of the cases. Test cases with 7 and 8 hidden nodes

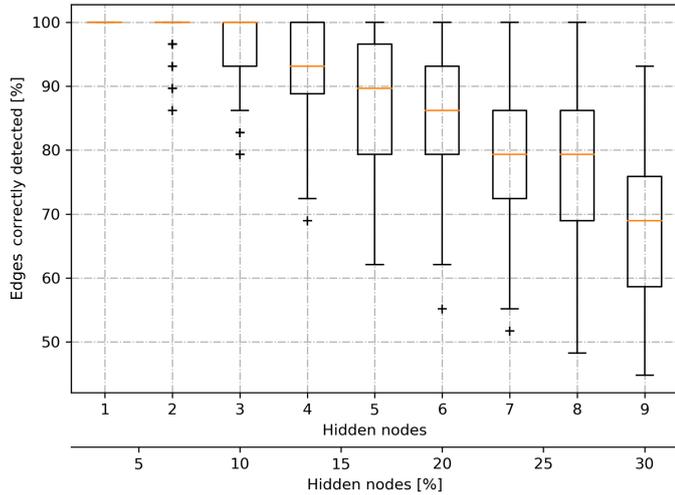


Figure 5: Percentage of edges correctly detected by the algorithm.

still produce the exact topology in some cases, but the average accuracy decreases in the others. The algorithm is not able to find the exact topology in any case with 9 hidden nodes, but still correctly identifies more than half of the edges in most cases.

The second metric evaluates the quality of the impedance magnitude estimation. This metric is given, for each metered node, by the error on the sum of the self-impedances in its path to the root. The mean absolute percentage error of the impedance magnitude is presented in Figure 6. In cases with less than 10% of hidden nodes, the cumulated impedance error is lower than 15%. For cases with more hidden nodes, the algorithm provides similar results when Assumption 1 is satisfied. When it is not, estimated topology and some current flows are miscalculated, resulting in larger impedance estimation errors.

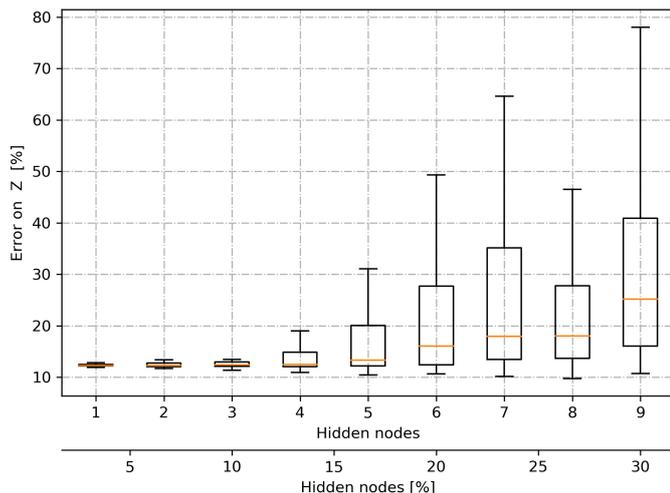


Figure 6: Error on the cumulative impedance of the path to the root of metered nodes.

## VI. CONCLUSION

In this work, the problem of reconstructing a model for LV networks, identifying topology and cable parameters, is addressed. The proposed algorithm tackles the network reconstruction problem by analysing available smart-meter data considering that some power injecting nodes in the network are not metered, assuming that such nodes are located reasonably far from each other. The algorithm is shown to provide relevant solutions in most of the scenarios with 20% or fewer hidden nodes on a 30-node test network.

This approach allows DSOs to have a better overview of the network, to better operate it and, for instance, be able to implement more adjusted preventive or corrective measures.

Future work could focus on merging the identified single-phase models into a three-phase one. Additional development would be needed to identify the topology of networks when a larger section of the grid cannot be observed, exploiting additional information available to DSOs, such as GIS data and conventional metering data.

## VII. ACKNOWLEDGMENTS

The authors thank RESA for providing data and constructive discussions that greatly contributed to this work. A. Sutera is supported via the Energy Transition Funds project EPOC 2030-2050 organised by the FPS economy, S.M.E.s, Self-employed and Energy.

## VIII. REFERENCES

- [1] S. Bolognani, N. Bof, D. Michelotti, R. Muraro, and L. Schenato, "Identification of power distribution network topology via voltage correlation analysis", in *52nd IEEE Conference on Decision and Control*, IEEE, 2013, pp. 1659–1664.
- [2] S. J. Pappu, N. Bhatt, R. Pasumathy, 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.
- [3] K. Soumalas, G. Messinis, and N. Hatziargyriou, "A data driven approach to distribution network topology identification", in *2017 IEEE Manchester PowerTech*, IEEE, 2017, pp. 1–6.
- [4] Y. Yuan, O. Ardakanian, S. Low, and C. Tomlin, "On the inverse power flow problem", *arXiv preprint arXiv:1610.06631*, 2016.
- [5] 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.
- [6] 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.
- [7] F. Olivier, A. Sutera, P. Geurts, R. Fonteneau, and D. Ernst, "Phase identification of smart meters by clustering voltage measurements", in *2018 Power Systems Computation Conference (PSCC)*, IEEE, 2018.