Identification of AC Distribution Networks with Recursive Least Squares and Optimal Design of Experiment

Emanuele Fabbiani, Pulkit Nahata, Giuseppe De Nicolao *Senior Member, IEEE*, and Giancarlo Ferrari-Trecate *Senior Member, IEEE*

***Abstract*— The increasing penetration of intermittent distributed energy resources in power networks calls for novel planning and control methodologies which hinge on detailed knowledge of the grid. However, reliable information concerning the system topology and parameters may be missing or outdated for temporally varying electric distribution networks. This paper proposes an online learning procedure to estimate the network admittance matrix capturing topological information and line parameters. We start off by providing a recursive identification algorithm exploiting phasor measurements of voltages and currents. With the goal of accelerating convergence, we subsequently complement our base algorithm with a design-of-experiment procedure which maximizes the information content of data at each step by computing optimal voltage excitations. Our approach improves on existing techniques, and its effectiveness is substantiated by numerical studies on a modified IEEE testbed.**

***Index Terms*— Power distribution, Power grids, Recursive estimation, Smart grids, System identification**

I. Introduction

Distribution networks serving as an interface between distribution substation and end-to-end customers are going through substantial transformations, attributable to an ever increasing deployment of demand-side technologies and distributed energy resources (DERs). While offering many advantages, DERs can compromise grid reliability due to added intermittency and creation of reverse power flows. In order to ensure safe and resilient operation of distribution systems, comprehensive monitoring and efficient control algorithms are necessary. Nevertheless, any meaningful grid optimization and monitoring task entails grid identification, that is gaining knowledge of grid topology and line parameters.

Research tackling the grid identification problem can broadly be classified into two main branches. On the one hand, works like [1]–[3] propose learning algorithms which draw on the statistical properties of nodal measurements to determine the operational structure and line impedances. This approach has the major advantage of accounting for buses with no available measurements (hidden nodes) [2] although restrictive assumptions are required, e.g. hidden nodes must not be adjacent to each other. Moreover, methods based on second-order statistics either make assumptions on the covariance of nodal injections [1] or assume its foreknowledge [2], [3], and apply only to radial feeders. The latter restriction is dropped in [3], but only for the purpose of topology estimation. In a realistic setting, these assumptions might not be satisfied; more so due to the rise of distributed generation and smart grids leading to meshed network structures.

On the other hand, in [4]–[7], network identification has been cast into the problem of learning the admittance matrix, where the position of non-zero elements provides topological information, while the values of these are related to the electrical parameters of the lines. Contrary to [1]–[3], this approach requires voltage, current, or power measurements at each bus of the grid. Nevertheless, it can be applied to both radial and meshed structures. In particular, Lasso and its variants have been widely adopted to enforce sparsity of the admittance matrix. In [5], a compressive sensing approach leads to a Lasso formulation to recover the connections of each bus. In [6], a probabilistic graphical model motivates the adoption of Lasso to identify the non-zero elements of the admittance matrix. However, no constraint on the symmetric structure of the admittance matrix is incorporated *a priori*, leading to an over-parameterized solution twice estimating each edge. As a partial remedy to this problem, estimates of the same edge are combined *a posteriori*. While both [5] and [6] focus on topology, neither considers the estimation of the electrical parameters of the lines. Finally, in [7], topology and line parameters are obtained at once owing to learning the admittance matrix using Adaptive Lasso. In addition, a procedure to cope with collinearity in measurements is also proposed.

Different from previously-stated works banking on passively recorded data, an active data collection paradigm is explored in [8]–[11]. Grid topology and parameter estimation are complemented with inverter probing in [9], [10]. Both works, besides assuming a resistive radial network and employing approximate linearized power-flow equations, lack a comprehensive framework for the optimal design of probing injections. A systematic procedure for maximizing the information content of data samples is explored in [11], wherein active power setpoints for generator nodes are provided by an online design-of-experiment (DoE) procedure [12]. Nonetheless, the proposed identification algorithm assumes the availability of line power flows, and neglects the structural constraints of the admittance matrix.

All the foregoing works adopt an offline approach, in the sense that they pivot on a batch of previously collected data to estimate grid topology and/or parameters. Distribution networks, unlike transmission networks, oftentimes undergo topological changes for maintenance, load balancing, and fault isolation. Furthermore, future distribution systems are envisaged as reconfigurable networks, wherein certain sections – just like microgrids – connect or disconnect to improve dispatch of DERs [13]–[15]. In the event of a topology change (often localised), a batch method shall discard valuable data, await new samples, and re-run the estimation afresh. On the contrary, an online, recursive identification methodology, encoding the relevant information carried by past data samples in its parameters, can provide new network topology and parameter estimates quickly and autonomously.

A. Paper Contributions and Organization

This article focuses on AC power distribution networks and introduces an online learning procedure, entirely based on nodal

This work was supported by the Swiss National Science Foundation under the NCCR Automation (grant agreement 51NF40_180545) and the COFLEX project (grant agreement 200021 169906).

Emanuele Fabbiani and Giuseppe De Nicolao are with the Identification and Control of Dynamic Systems Laboratory, Università degli Studi di Pavia, Italy. Email addresses: emanuele.fabbiani01@universitadipavia.it, giuseppe.denicolao@unipv.it.

Pulkit Nahata, and Giancarlo Ferrari-Trecate are with the Institute of Mechanical Engineering, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. Email addresses: {pulkit.nahata, giancarlo.ferraritrecate}@epfl.ch.

measurements, for estimating the admittance matrix, which embeds detailed information about grid topology and line parameters. The main novelties of this paper are on two fronts. First, different from previous works such as [4], [7], we propose an online recursive identification algorithm to estimate the admittance matrix, capable of automatically adapting the estimation in case of undetected changes in topology or faults. When the admittance matrix is symmetric or/and Laplacian, our algorithm does away with redundant parameters by means of a transformation matrix. Second, we tap into the principles of optimal DoE and discuss an approach to compute suitable generator voltages which, when complemented with the base recursive algorithm, accelerates the admittance matrix estimation. Methodological contributions are complemented by a simulation example demonstrating that our method can outperform existing schemes.

The remainder of Section I introduces relevant preliminaries and notation. Section II recaps network models and motivates the grid identification problem. Section III describes the recursive estimation algorithm whereas optimal DoE procedure is discussed in Section IV. Proposed algorithms are validated via numerical studies in Section V. Finally, conclusions are drawn in Section VI.

B. Preliminaries and Notation

Sets, vectors, matrices, and random variables: let $j = \sqrt{-1}$ represent the imaginary unit. For a finite set $\mathcal{V}$, $|\mathcal{V}|$ denotes its cardinality. An (m,n) matrix is one with m rows and n columns. Given $x \in \mathbb{C}^n$, $\overline{x}$ is its complex conjugate and $[x]$ the associated diagonal matrix of order n. Throughout, 1_n and 0_n are n-dimensional vectors of all ones and zeros, whereas $\mathbb{I}_n$ and $\mathbb{O}_{n\times m}$ represent (n,n) identity and (m,n) zero matrices, respectively. The unit vector $e_i,\ i = 1,...,n$ is the i^{th} column of $\mathbb{I}_n$. For a matrix A, $A^\top$ denotes its transpose, A^{H} its Hermitian (complex conjugate) transpose, and A_i its i^{th} column vector. The Kronecker product between matrices A and B is $A \otimes B$. A positive definite matrix A and a positive semidefinite matrix B verify $A \succ 0$ and $A \succeq 0$, respectively. We let $\mathcal{N}(x, A)$ designate a Gaussian random vector of dimension n, where x is the mean vector and A the covariance matrix.

Matrix vectorization operators: We indicate by $\mathrm{vec}(A) = [A_{1\cdot}^\top \cdots A_{n\cdot}^\top]^\top$ the mn-dimensional stacked column vector. Furthermore, if A is a square matrix, then the half-vectorization operator $\mathrm{vech}(A)$ provides the $n(n+1)/2$-dimensional vector obtained by eliminating all supradiagonal elements of A from $\mathrm{vec}(A)$. Furthermore, $\mathrm{ve}(A)$ is the $n(n-1)/2$-dimensional vector obtained by removing diagonal elements of A from $-\mathrm{vech}(A)$.

II. Network Modeling and Problem Formulation

In this section, we review relevant algebraic models for AC power networks, and detail the grid identification problem.

A. Distribution Network Modelling

An electric distribution network is modeled as an undirected, weighted, and connected graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{W})$, where the nodes in $\mathcal{V} = \{1, 2, \ldots, n\}$ represent buses, either generating units or loads, and edges represent power lines, each connecting two distinct buses and modeled after the standard lumped $\pi-$model [16]. To each edge $(i,k) \in \mathcal{E}$ we associate a complex weight equal to the line admittance $y_{ik} = g_{ik} + jb_{ik} \in \mathcal{W}$, where $g_{ik} > 0$ is the line conductance and $b_{ik} \in \mathbb{R}$ the line susceptance. The network is then completely represented by the admittance matrix $Y \in \mathbb{C}^{n\times n}$, with elements $Y_{ik} = -y_{ik}$ for $i \neq k$ and $Y_{ik} = \sum_{i=1, i\neq k}^{n} y_{ik} + y_{s,i}$, where $y_{s,i} \in \mathbb{C}$ is the shunt element at the i^{th} bus. If the network does not include phase-shifting transformers and power lines are not compensated by series capacitors, Y is symmetric. In addition, Y is Laplacian if the shunt elements $y_{s,i}$ are not present [17], [18]: this happens, for instance, in small- and medium-sized networks, with line lengths less than 60 km.

Throughout this work, we consider a phase-balanced power network operating in sinusoidal regime. To each bus $h \in \mathcal{V}$, we associate a voltage phasor $|v^{(h)}|e^{j\theta^{(h)}} \in \mathbb{C}$, where $|v^{(h)}|$ is the voltage magnitude and $\theta^{(h)} \in \mathbb{R}$ the voltage angle, a current injection phasor $|i^{(h)}|e^{j\phi^{(h)}} \in \mathbb{C}$, and a complex apparent power $s^{(h)} = p^{(h)} + jq^{(h)}$ with $p^{(h)}, q^{(h)} \in \mathbb{R}$. As standard in distribution networks, we assume the point of common coupling (PCC) to be the slack bus with fixed $|v^{(0)}| = 1$ and $\theta^{(0)} = 0$. The remaining buses are classified as generators $\mathcal{S}$ and loads $\mathcal{L}$, such that $\mathcal{V} = \mathcal{S} \cup \mathcal{L} \cup \{0\}$. For notational simplicity we set $|\mathcal{V}| = n$, $|\mathcal{S}| = g$, and $|\mathcal{L}| = l$, where $g, l \geq 1$. In active distribution networks, generators are DERs generally interfaced with inverters equipped with voltage and/or power control [19]. The current-voltage relation descending directly from Kirchhoff's and Ohm's laws is given by

$$i = Yv, \tag{1}$$

where $i \in \mathbb{C}^n$ is the vector of nodal current injections, and $v \in \mathbb{C}^n$ the vector of nodal voltages [20]. Similarly, one can deduce the relation between the vectors of nodal complex power injections s and nodal voltages v as

$$s = [v](\overline{Yv}). \tag{2}$$

B. Identification of AC distribution networks

The identification problem for AC distribution networks, defined in [4], [7], aims at reconstructing the admittance matrix from a sequence of voltage and current phasor measurements corresponding to different steady states of the system. Similar to [4], [7], our work makes the following assumption.

Assumption 1: The network is fully observable, i.e., voltage and current measurements are available at each node.

Let t be the number of samples collected up to a certain time instant, $v_\tau \in \mathbb{C}^n$ and $i_\tau \in \mathbb{C}^n$ the vectors of current injections and voltages for $\tau = 1, \ldots, t$. From (1), one can obtain

$$I_t = YV_t, \tag{3}$$

where $V_t = [v_1, v_2, \ldots, v_t] \in \mathbb{C}^{t\times n}$, and $I_t = [i_1, i_2, \ldots, i_t] \in \mathbb{C}^{t\times n}$. The admittance matrix Y, encoding both line parameters and topological information, is typically sparse as each bus is not connected to all the remaining nodes. Moreover, as explained in Section II-A, Y has other structural properties: for most distribution networks, which lack phase-shifting transformers and feature short lines, the following assumption is satisfied.

Assumption 2: The admittance matrix Y is symmetric and Laplacian, that is, $Y1_n = 0_n$.

Both the symmetric and Laplacian structures of admittance matrix greatly reduce the number of entries of Y to be estimated. This observation is further explored in the subsequent section.

III. Recursive Online Identification

In the ideal case of noiseless current and voltage measurements, the identification of Y reduces to solving the system of linear equations (3), once enough samples are collected. Unfortunately, μPMUs and other metering devices introduce an error commonly modeled as white noise [2], [7]. In the following, for sake of simplicity, it is assumed that the measurement error is distributed as $\mathcal{N}(0_n, \sigma^2\mathbb{I}_n)$, thus implying that the error at each bus has the same variance. As will be clear in the sequel, extensions to different covariance matrices are immediate.

Upon vectorizing either side of equation (3), one obtains

$$\mathrm{vec}(I_t) = \mathrm{vec}(YV_t) = \left(V_t^\top \otimes \mathbb{I}_n\right)\mathrm{vec}(Y). \tag{4}$$

Regression methods can be used to get a least squares estimate of $\mathrm{vec}(Y)$ – the vector representation of the admittance matrix. Before diving into the online estimation algorithm for the admittance matrix Y, we note that $\mathrm{vec}(Y)$ comprises n^2 entries of Y. Being symmetric, Y has at max $n(n+1)/2$ non-redundant entries, which further reduce to $n(n-1)/2$ under Assumption 2. Redundant entries in $\mathrm{vec}(Y)$ can be eliminated by means of $\mathrm{vech}(Y)$ – if Y is symmetric, or $\mathrm{ve}(Y)$ – when Assumption 2 holds. Relevant relations between the matrix vectorization operators are summarized in the following Lemma.

Lemma 1: Given n, there is a unique $(n^2, n(n+1)/2)$ matrix D, called duplication matrix, such that

$$\mathrm{vec}(Y) = D\,\mathrm{vech}(Y). \tag{5}$$

Furthermore, under Assumption 2, there exists a unique $(n(n+1)/2, n(n-1/2))$ matrix T such that

$$\mathrm{vech}(Y) = T\,\mathrm{ve}(Y). \tag{6}$$

Proof: Existence and uniqueness of D are shown in [21]. The proof of existence and uniqueness of T – as well as formulae to construct it – can be found in [22, Appendix A]. ■

Python and MATLAB codes for constructing D and T are available at [23]. Both T and D can be constructed given the number of nodes in the network n, therefore they must not be estimated from measurements.

Hereafter, we consider the case where Assumption 2 holds. Using Lemma 1, we recover the full vectorization of Y as

$$\mathrm{vec}(Y) = D\,\mathrm{vech}(Y) = DT\,\mathrm{ve}(Y). \tag{7}$$

By combining (4) and (7) we get

$$\mathrm{vec}(I_t) = \left(V_t^\top \otimes \mathbb{I}_n\right) DT\,\mathrm{ve}(Y). \tag{8}$$

Introducing the following matrices and vectors

$$A_t := \left(v_t^\top \otimes \mathbb{I}_n\right) DT, \tag{9a}$$

$$\underline{A_t} := \left(V_t^\top \otimes \mathbb{I}_n\right) DT, \tag{9b}$$

$$b_t := \mathrm{vec}(I_t), \text{ and} \tag{9c}$$

$$x := \mathrm{ve}(Y), \tag{9d}$$

the least squares estimation problem at time t writes as

$$\hat{x}_t = \arg\min_x \left\|b_t - \underline{A_t}x\right\|^2. \tag{10}$$

The formulation in (10) equally weights samples at any time instant, which can be detrimental for time-varying distribution networks and smart grids [24]. By introducing a forgetting factor $\lambda \in (0, 1]$ [25], we reformulate the estimation problem as

$$\hat{x}_t = \arg\min_x \sum_{i=1}^{t} \lambda^{t-i}\left\|i_i - A_i x\right\|^2. \tag{11}$$

Given an initial guess of the parameter vector $\hat{x}_0$ and the matrix $Z_0 := \sigma^{-2}\,\mathrm{Cov}[\hat{x}_0]$, estimates of $\hat{x}_t$ and $Z_t := \sigma^{-2}\,\mathrm{Cov}[\hat{x}_t]$ can be obtained by the recursive least squares (RLS) algorithm [25, p. 541]:

$$\hat{x}_t = \hat{x}_{t-1} + Z_t A_t^{\mathsf{H}}\left(i_t - A_t\hat{x}_{t-1}\right) \tag{12a}$$

$$Z_t = (\lambda Z_{t-1}^{-1} + A_t^{\mathsf{H}} A_t)^{-1} \tag{12b}$$

$$= \lambda^{-1}(Z_{t-1} - Z_{t-1}A_t^{\mathsf{H}}\left(\lambda\mathbb{I}_n + A_t Z_{t-1} A_t^{\mathsf{H}}\right)^{-1} A_t Z_{t-1}). \tag{12c}$$

From $\hat{x}_t$, one can derive the estimated admittance matrix $\hat{Y}_t = DT\hat{x}_t$. As shown in Section II, the complex elements of the admittance matrix capture both the conductance and the susceptance of the lines. In a real scenario, existing information or batch data can be used to improve the initial guess x_0 and Z_0.

The RLS algorithm with constant or bounded forgetting factor is known to have notable stability and convergence properties [26], [27]. For noisy measurements, RLS with constant forgetting factor is consistent under some excitation conditions only when the forgetting factor is 1 [26]. Otherwise, RLS has limited memory, preventing it from achieving consistency, which is generally traded off with the ability to follow changes in the parameters. In order to establish a basic degree of competency for the RLS estimator (12), we consider the case of a static network with noise-free measurements. In Section V we present numerical simulations to show how the identification method can tolerate noise and can adapt its estimation to changes in network topology.

Classical works establish that, when data are not affected by noise, the error on the parameters is bounded, and its projection onto the subspace for which persistent excitation holds – see [27] for a definition – converges to zero as the number of samples approaches infinity. Still, the arguments in [27] consider only real-valued, single-input-single-output settings. Here, we provide convergence results pertaining to our case, which involves complex inputs, outputs and parameters, and a multivariate output at each iteration.

Lemma 2: Consider the RLS algorithm (12). Assume that Y is constant in time – therefore, $x_t = x$, V_t is full-rank and measurement are not affected by noise. Define the error on the parameters $\tilde{x}_t := \hat{x}_t - x$. For any $\hat{x}_0$ and $Z_0 = Z_0^{\mathsf{H}} \succ 0$, (i) the norm of the error $\|\tilde{x}_t\|$ is bounded, and (ii) the projection of $\tilde{x}_t$ on the excitation subspace converges to zero as t approaches infinity.

Proof: See Appendix I. ■

Remark 1: Recursive least squares assumes that the matrix V_t is full-rank. If not, one can still apply the method to learn part of the admittance matrix, as shown in [7].

Remark 2: RLS algorithm can also be applied to three-phase unbalanced networks. As detailed in [7], the variables to be measured are line-to-ground voltages and current injections for each phase of the nodes, while the admittance matrix to be estimated shares the properties described in Section II.

IV. Optimal Design of Experiment

While several learning approaches only capitalize on uncontrolled inputs and outputs, identification algorithms appropriately probing controllable DERs can improve the estimation of the admittance matrix. In this work, each DER is assumed to be equipped with a voltage controller – necessary for networks with high photovoltaic integration [19]. Targeting these controllers, we henceforth propose a modified version of the recursive estimation algorithm (12) where, at each iteration, DER voltages are set according to a D-optimal design [12], the purpose of which is to maximize the determinant of the Fisher information matrix of the model parameters. With reference to the least squares problem (11), the Fisher information matrix [12] at time t is $F_t = (\mathrm{Cov}(x_t))^{-1}$. As the measurement noise is a

Gaussian vector $\mathcal{N}(\mathbf{0}_n, \sigma^2 \mathbb{I}_n)$, we have

$$F_t = \sigma^{-2} Z_t^{-1} = \sigma^{-2}(\lambda Z_{t-1}^{-1} + A_t^{\mathsf{H}} A_t). \tag{13}$$

We note that A_t depends on the nodal voltages V_t; see (9a). The D-optimal design is the result of the optimization problem

$$v_t^* = \arg\max_{v_t} \det(F_t). \tag{14}$$

We observe that σ does not influence the optimum and can thus be neglected. Moreover, upon applying the logarithm to the target function – a common practice for improving numerical properties [12, Chap. 10], we get

$$v_t^* = \arg\min_{v_t} -\log\det(\lambda Z_{t-1}^{-1} + A_t^{\mathsf{H}} A_t). \tag{15}$$

While formulating the DoE problem, we need to take into account voltage limits for all nodes, as well as the active and reactive power dispatched by DERs. Furthermore, the power requirements of loads, expressed by the power flow equations (2), must be satisfied. By adding these constraints, we get the optimization problem

$$(v_t^*, p_t^*) = \arg\min_{v_t, p_t} -\log\det(\lambda Z_{t-1}^{-1} A_t^{\mathsf{H}} A_t) \tag{16a}$$

$$\text{subject to:} \quad s_t = [v_t]\overline{(\hat{Y}_{t-1} v_t)} \tag{16b}$$

$$v_{\min}^{(i)} \le v_t^{(i)} \le v_{\max}^{(i)} \qquad \forall i \in \mathcal{V} \tag{16c}$$

$$\theta_{\min}^{(i)} \le \theta_t^{(i)} \le \theta_{\max}^{(i)} \qquad \forall i \in \mathcal{V} \tag{16d}$$

$$p_{\min}^{(k)} \le p_t^{(k)} \le p_{\max}^{(k)} \qquad \forall k \in \mathcal{S} \tag{16e}$$

$$q_{\min}^{(k)} \le q_t^{(k)} \le q_{\max}^{(k)} \qquad \forall k \in \mathcal{S}, \tag{16f}$$

where A_t depends on V_t as in (9a). It is worth noting that the computation of Z_{t-1}^{-1} in (16a) does not require the inversion of Z_{t-1}: from (12b), one has $Z_t^{-1} = \lambda Z_{t-1}^{-1} + A_t^{\mathsf{H}} A_t$, which allows for a recursive update of Z_t^{-1}. We also note that constraint (16b) depends on $\hat{Y}_{t-1}$, which is the most recent estimate of the unknown matrix Y. While this approximation makes it difficult to analyze the properties of the sequence $\{\hat{Y}_t\}_{t=0}^{\infty}$, numerical experiments described in Section V show that such an approach might be only slightly suboptimal with respect to using the real admittance matrix Y.

Remark 3: We note that the proposed DoE procedure helps achieve persistent excitation, which implies the information matrix of the parameters being full rank at each iteration [27]. Indeed, since DoE aims at maximizing the determinant of the information matrix, its objective is in contrast with a loss of rank. Therefore, setting voltages as per (16) helps satisfying the hypothesis of Lemma 2.
The DoE formulation (16) is flexible: one can append more constraints to the optimization problem to cope with technical limitations. For example, the voltage of some DERs may be fixed, or power limitations for certain lines can be introduced. The solution of problem (16) is the vector of all nodal voltages; however, voltage references are provided only to DERs as loads cannot generally be controlled.

Remark 4: The DoE problem (16) outputs both voltage magnitude and active power for each generating unit. In this work, we assume that the former is directly used as a control reference, however, the latter can be equivalently adopted in case of power-controlled DERs. When excited with the power reference signal, the generating units cause voltage variations in the network [9], [10]: the resulting current-voltage data can then be utilized in (12) for the admittance matrix estimation.
To summarize, given an initial guess of $\hat{x}_0$ and $\hat{Z}_0$, a value of λ, and active and reactive power demands for loads, the recursive estimation, enhanced with DoE, can be described by the following steps repeated at each time t.

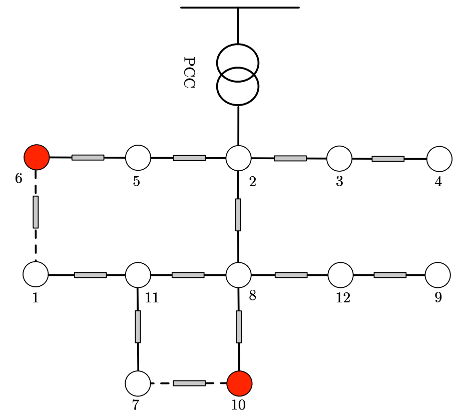


Fig. 1: Modified IEEE 13-bus feeder. Buses ● and ○ represent DERs and loads, respectively.

1) Solve the DoE problem (16) for the nodal voltages v_t^*, using the current estimation $\hat{x}_{t-1}$ and Z_{t-1}.
2) Provide the voltage set-point $|v_t^{*,(k)}|$ to DERs $k \in \mathcal{S}$.
3) Collect measurements of current and voltage phasors from each bus $i \in \mathcal{V}$.
4) Update the estimates of $\hat{x}_t$ and Z_t using the RLS algorithm (12).

V. Numerical Results

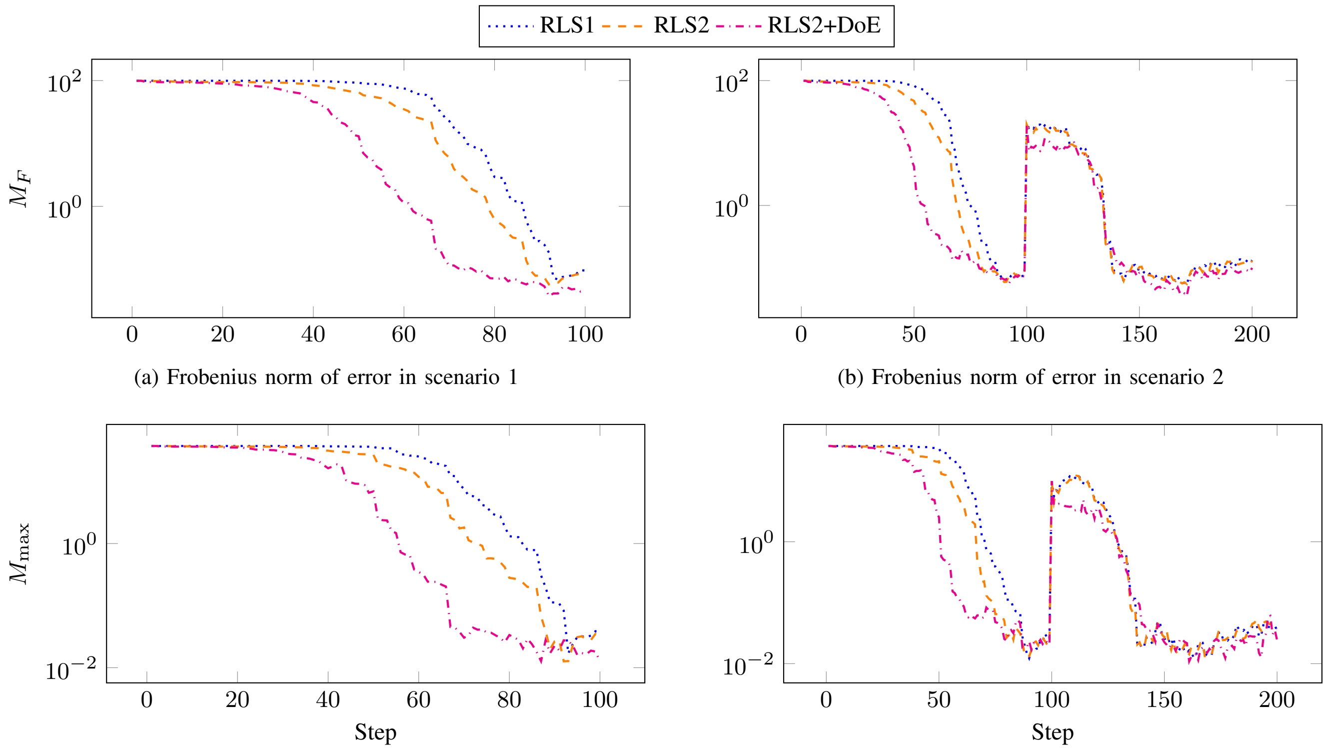
A. Setup

We validate the proposed methods on a modified version of the IEEE 13-bus feeder. In order to test the proposed algorithms on a meshed network, we add two lines, one connecting bus 6 with 1, and the other bus 7 with 10; see Fig. 1. As all the lines have negligible capacitance, the network verifies Assumption 2. Distributed generation is introduced through the addition of DERs to buses 6 and 10, whereas load profiles are extracted from the public Pecan Street dataset [28]. Since this dataset does not include reactive power, a random lagging power factor between 0.85 and 0.95 is considered. Following the procedure adopted in [7], we connect a random number of customers between 5 and 15 to each node. The granularity of the measurement is one minute. We use the AC Power flow solver MATPOWER to derive nodal current and voltage phasors.

We consider three different online estimation methods: (i) RLS1 solely imposing the symmetric structure of Y by adopting the parametrization $x = \mathrm{vech}(Y)$; (ii) RLS2 forcing a Laplacian structure of Y as in (9) and (12); and (iii) RLS2+DoE where the voltages of DER buses are set according to the DoE procedure presented in Section IV. The solution of the DoE problem (16) is computed using the BFGS interior-point non-convex solver implemented by the `fmincon` MATLAB function.

We consider two scenarios to asses the performance of our method in providing an accurate estimate of the admittance matrix. Scenario 1 looks at a network whose topology does not change over time, whereas scenario 2 considers time-varying networks. More specifically, scenario 2 simulates a fault between bus 7 and 10, leading to tripping of connecting line. In scenario 1, we also compare our online algorithm with two batch approaches: Ordinary Least Squares (OLS) and Adaptive Lasso [7]. However, batch algorithms are unable to reveal network changes and therefore cannot be applied in scenario 2. In this respect, scenario 2 illustrates the main value if the online method proposed in this paper over existing offline approaches.

In order to assess the identification performance, we use the error metrics $M_F := \|Y - \hat{Y}\|_F$, $M_{\max} := \|Y - \hat{Y}\|_{\max}$ and



(a) Frobenius norm of error in scenario 1

(b) Frobenius norm of error in scenario 2

(c) Max norm of error in scenario 1

(d) Max norm of error in scenario 2

Fig. 2: Error metrics in scenario 1 and 2

$M_R := \|Y - \hat{Y}\|_F / \|Y\|_F$, where subscripts F and max denote the Frobenius norm and the max norm, respectively. The metric M_F assesses the overall goodness of the estimation, $M_{\max}$ is intended to capture possible issues in the identification of single elements, while M_R provides a relative measure of the identification error.

In all the experiments, we introduce a Gaussian measurement error $\mathcal{N}(0_n, \sigma^2 \mathbb{I}_n)$, $\sigma = 10^{-5}$ on both the real and the imaginary part of the measurements, which approximately corresponds to an accuracy of 0.1% compatible with real metering devices [9]. The recursive estimation algorithms are initialized with $\hat{x}_0 = \delta 1_{n(n-1)/2}$, $\delta = 10^{-4}$ and $Z_0 = K\mathbb{I}_{n(n-1)/2}$, $K = 10^4$. The forgetting factor is set to $\lambda = 0.8$.

B. Results

For scenario 1, Table I shows the comparison with benchmarks after 100 iterations when the estimates provided by all online algorithms no longer improve. The error metrics can be noticed to be of the same order of magnitude for all methods; although RLS1 and RLS2 achieve poorer performance than OLS and Lasso. This is expected as both OLS and Lasso are batch estimators making simultaneous use of all the collected data, while online methods trade accuracy for the ability to adapt to changes. We also note RLS2+DoE outperforms all other methods, except for Lasso. All the methods are capable of estimating both the real and the imaginary part of the admittance matrix with comparable accuracy. For instance, the largest relative error of RLS+DoE on a line conductance is 0.42%, on a line susceptance is 0.33%.

In both scenarios 1 and 2, RLS2+DoE achieves faster convergence as well as better accuracy than other iterative methods; see Fig. 2. The downside is the stress on DER voltages, which are subjected to frequent changes (Fig. 3). Nevertheless, due to constraints in the formulation of the design problem (16), both voltage set-points and

	M_F $[\times 10^{-2}]$	$M_{\max}$ $[\times 10^{-2}]$	M_R
OLS (batch)	5.44	1.69	0.055%
Adaptive Lasso (batch)	2.58	0.87	0.026%
RLS1	9.55	3.84	0.095%
RLS2	7.97	3.26	0.080%
RLS2+DoE	4.74	1.27	0.047%

TABLE I: Error metrics for scenario 1 after 100 samples

realized voltages, stay within the prescribed voltage interval, which is $[0.95, 1.05]$ p.u, as highlighted in Fig. 3. In both scenarios, $M_{\max}$ follows the same trend as M_F until convergence to a low value, thus ruling out issues about the estimation of specific elements of Y.

In the context of scenario 2, the error in the estimation of $Y_{7,10}$ (Fig. 4) is worth a few comments. Only online algorithms RLS1, RLS2 and RLS2+DoE are here applicable, thus batch methods are not discussed. Note that $|Y_{7,10}| = 9.8$ up to $t = 100$, and subsequently drops to zero as a consequence of the simulated fault. All our recursive implementations are able to quickly adapt to a change in topology, thus proving the usefulness of online estimation. After mere two iterations ($t = 102$), the absolute value of the estimated line admittance is 2.21 for RLS1, 2.11 for RLS2, and 1.1 for RLS2+DoE. Moreover, after 7 iterations, the estimation is lower than 1 for all the online algorithms.

C. Sensitivity to voltage noise

In real applications, measurement noise affects both currents and voltages. A systematic discussion of this scenario is outside the scope of this paper, as it would require a radical change in the modelling approach, as discussed in Section VI. However, we assess the deterioration in performance experienced by the proposed algorithms when a zero-mean Gaussian noise with covariance matrix $\sigma_v^2 \mathbb{I}$ is applied to both the real and the imaginary part of voltage

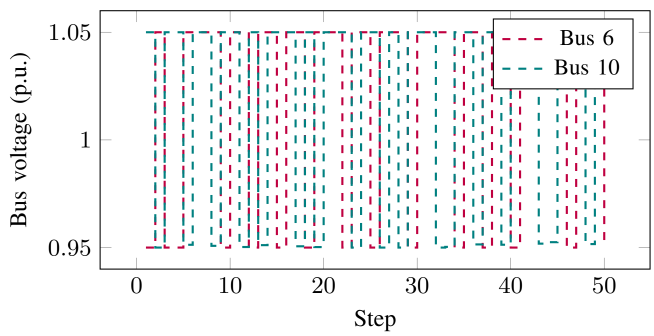


Fig. 3: DER voltages produced by DoE for the first 50 iterations in scenario 1.

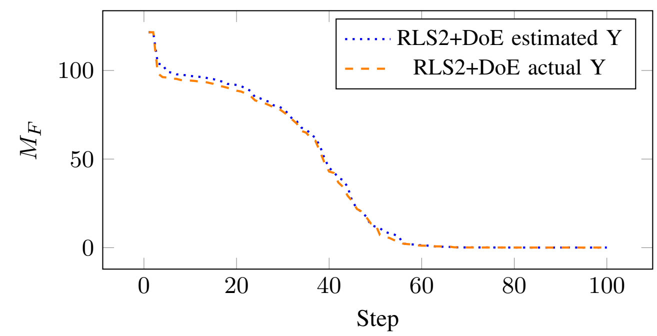


Fig. 6: Comparison between DoE with estimated and actual admittance matrix Y.

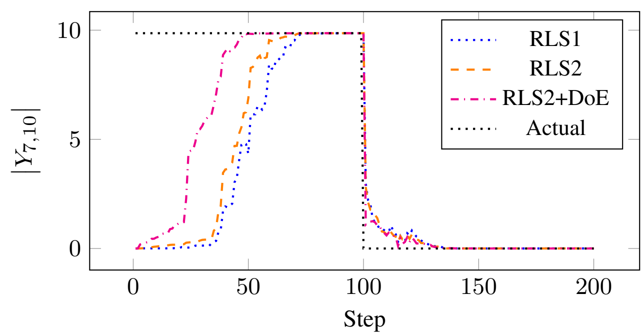


Fig. 4: Estimation of element $Y_{7,10}$ in scenario 2.

measurements. As displayed in Fig. 5, all methods suffer from input noise; however, RLS2+DoE is less affected than the others, and achieves an acceptable performance even when the noise on voltages is of the same order of magnitude as that on currents.

D. Effect of the DoE formulation

As noted in Section IV, the DoE formulation (16) has to rely on estimated admittance matrix $\hat{Y}_{t-1}$, instead of the unknown real admittance matrix Y. In order to show the effect of such an approximation on the identification algorithm, we run RLS2+DoE on scenario 1 by setting $\hat{Y}_{t-1} = Y$ in (16b). The results in Fig. 6 show that the procedure based on the real model of the network performs only marginally better.

VI. Conclusions

A frequent lack of detailed information such as grid topology and line parameters motivated the development of this work, which presents an online learning procedure for grid identification in AC networks. In contrast with batch methods for estimating the grid admittance matrix, our algorithm is online and recursive in nature, thus capable of adapting to both sudden changes in the network topology and slow drift in line parameters.

Notwithstanding the applicability of our methods to generic admittance matrices, we provide a transformation matrix that leverages the structural properties of symmetric Laplacian matrices. Furthermore, we propose a method based on optimal DoE for improving convergence of recursive identification algorithms.

Future developments will aim at coming up with novel identification techniques for networks where not all nodal electric variables can be measured. Effort will also be devoted to extending out identification framework to error-in-variable models, with a view to properly taking into account all sources of measurement error [29]. Further work will also explore the utility of grid identification schemes in the supervisory control of microgrids [14].

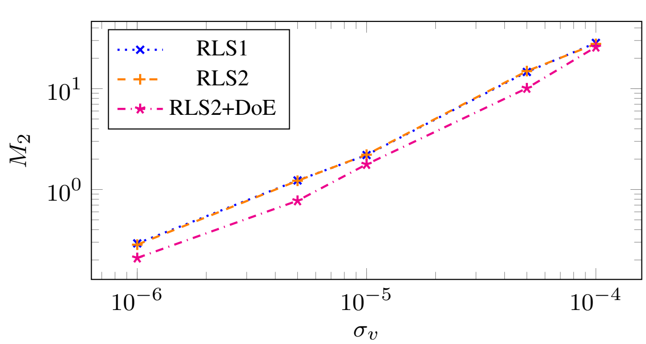


Fig. 5: Frobenius norm of estimation error for different levels of noise on voltage measurements in scenario 1.

Appendix I
Proof of Lemma 2

In order to show (i), we start by substituting $\tilde{x}_t = \hat{x}_t - x$ in (12) and considering that, in the noise-free case, $i_t = A_t x$. Then, we get the recursive formula

$$\tilde{x}_t = \hat{x}_{t-1} - Z_t A_t^{\mathsf{H}} A_t \tilde{x}_{t-1}. \tag{17}$$

For convenience, we define $\epsilon_t := A_t \tilde{x}_{t-1} = A_t \hat{x}_{t-1} - i_t$. Next, we introduce the Lyapunov-like, function $W_t := \tilde{x}_t^{\mathsf{H}} Z_t^{-1} \tilde{x}_t$. Note that W_t is a real-valued function, as Z_t is Hermitian, and so is Z_t^{-1}. By combining the definition of W_t with equations (17) and (12b), we derive

$$W_t = \lambda W_{t-1} - \epsilon_t^{\mathsf{H}} (\mathbb{I}_n - A_t Z_t A_t^{\mathsf{H}}) \epsilon_t. \tag{18}$$

It can be shown that $\mathbb{I}_n - A_t Z_t A_t^{\mathsf{H}} \succeq 0$; see e.g. Lemma 1 in [26]. Consider now the quantity $\epsilon_t^{\mathsf{H}} (\mathbb{I}_n - A_t Z_t A_t^{\mathsf{H}}) \epsilon_t$: it is real and non-negative because $\mathbb{I}_n - A_t Z_t A_t^{\mathsf{H}}$ is Hermitian and positive semidefinite. From (18), we obtain the inequality

$$W_t \leq \lambda W_{t-1}, \tag{19}$$

which one can recursively apply at each t to obtain

$$W_t \leq \lambda^t W_0. \tag{20}$$

Recalling the definition of W_t and (12b), we write:

$$\begin{aligned}\lambda^t W_0 \geq W_t &= \tilde{x}_t^{\mathsf{H}} Z_t^{-1} \tilde{x}_t \\ &= \tilde{x}_t^{\mathsf{H}} \left(\lambda Z_{t-1}^{-1} + A_t^{\mathsf{H}} A_t\right) \tilde{x}_t \\ &= \tilde{x}_t^{\mathsf{H}} \left(\lambda^t Z_0^{-1} + \sum_{i=1}^{t} \lambda^{t-i} A_i^{\mathsf{H}} A_i\right) \tilde{x}_t \\ &\geq \lambda^t \tilde{x}_t^{\mathsf{H}} \left(Z_0^{-1} + \sum_{i=1}^{t} A_i^{\mathsf{H}} A_i\right) \tilde{x}_t.\end{aligned} \tag{21}$$

Therefore, we conclude that

$$\tilde{x}_t^{\mathsf{H}} (Z_0^{-1} + \sum_{i=1}^{t} A_i^{\mathsf{H}} A_i) \tilde{x}_t \leq W_0. \tag{22}$$

As both $Z_0^{-1} \succ 0$ and $\sum_{i=1}^{t} A_i^{\mathsf{H}} A_i \succeq 0$, we have:

$$\tilde{x}_t^{\mathsf{H}} Z_0^{-1} \tilde{x}_t \leq W_0 \tag{23}$$

and

$$\tilde{x}_t^{\mathsf{H}} \left(\sum_{i=1}^{t} A_i^{\mathsf{H}} A_i\right) \tilde{x}_t \leq W_0. \tag{24}$$

Since $\tilde{x}_t^{\mathsf{H}} Z_0^{-1} \tilde{x}_t \geq \operatorname{mineig}(Z_0^{-1}) \|\tilde{x}_t\|^2$, equation (23) yields

$$\operatorname{mineig}(Z_0^{-1}) \|\tilde{x}_t\|^2 \leq W_0, \tag{25}$$

where $\operatorname{mineig}(X)$ is the minimal (real) eigenvalue of a Hermitian matrix X. Therefore, $\|\tilde{x}_t\|$ is bounded.

In order to show (ii), let G_t be the square root of $\sum_{i=1}^{t} A_i^{\mathsf{H}} A_i$, and $\tilde{x}_t^{(e)}$ and $\tilde{x}_t^{(u)}$ the projections of $\tilde{x}_t$ onto the subspaces where persistent excitation holds and does not hold, respectively. Then, (24) can be written as

$$\left\| G_t \tilde{x}_t^{(e)} + G_t \tilde{x}_t^{(u)} \right\| \leq W_0^{1/2}. \tag{26}$$

By the reverse triangular inequality, we have:

$$\underbrace{\left\| \frac{G_t \tilde{x}_t^{(e)}}{\left\| \tilde{x}_t^{(e)} \right\|} \left\| \tilde{x}_t^{(e)} \right\| \right\|}_{a} - \underbrace{\left\| \frac{G_t \tilde{x}_t^{(u)}}{\left\| \tilde{x}_t^{(u)} \right\|} \left\| \tilde{x}_t^{(u)} \right\| \right\|}_{b} \leq W_0^{1/2}. \tag{27}$$

We can now apply to (27) the same argument proposed in [27, Proof of Theorem 1]. Due to part (i) of the proof, the norms of $\tilde{x}_t^{(e)}$ and $\tilde{x}_t^{(u)}$ are bounded. Moreover, as $\tilde{x}_t^{(u)}$ is the projection of $\tilde{x}_t$ onto the subspace where persistent excitation does not hold, the term (b) of (27) is bounded. For the inequality (27) to hold, (a) must also be bounded. Yet, the lemma in the appendix of [27] shows that $G_t \tilde{x}_t^{(e)} / \left\| \tilde{x}_t^{(e)} \right\|$ is unbounded. Therefore, $\tilde{x}_t^{(e)}$ must converge to zero for (27) to be verified, proving part (ii). It is worth noting that the lemma in [27] involves sequences of real positive semidefinite matrices, but the proof holds without modification for complex Hermitian positive semidefinite matrices.

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

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

[4] Y. Yuan, S. Low, O. Ardakanian, and C. Tomlin, "Inverse power flow problem," *arXiv preprint arXiv:1610.06631*, 2016.

[5] M. Babakmehr, M. G. Simões, M. B. Wakin, and F. Harirchi, "Compressive sensing-based topology identification for smart grids," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 2, pp. 532–543, 2016.

[6] Y. Liao, Y. Weng, G. Liu, and R. Rajagopal, "Urban mv and lv distribution grid topology estimation via group lasso," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 12–27, 2018.

[7] 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.

[8] M. Angjelichinoski, Č. Stefanović, P. Popovski, A. Scaglione, and F. Blaabjerg, "Topology identification for multiple-bus dc microgrids via primary control perturbations," in *2017 IEEE Second International Conference on DC Microgrids (ICDCM)*. IEEE, 2017, pp. 202–206.

[9] G. Cavraro and V. Kekatos, "Graph algorithms for topology identification using power grid probing," *IEEE control systems letters*, vol. 2, no. 4, pp. 689–694, 2018.

[10] ——, "Inverter probing for power distribution network topology processing," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 3, pp. 980–992, 2019.

[11] X. Du, A. Engelmann, Y. Jiang, T. Faulwasser, and B. Houska, "Optimal experiment design for ac power systems admittance estimation," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 13 311–13 316, 2020.

[12] A. Atkinson, A. Donev, and R. Tobias, *Optimum experimental designs, with SAS*. Oxford University Press, 2007, vol. 34.

[13] D. Shelar, S. Amin, and I. Hiskens, "Resilience of electricity distribution networks-part ii: leveraging microgrids," *arXiv preprint arXiv:1812.01745*, 2018.

[14] P. Nahata, A. La Bella, R. Scattolini, and G. Ferrari-Trecate, "Hierarchical control in islanded dc microgrids with flexible structures," *IEEE Transactions on Control Systems Technology*, 2020.

[15] T. Dragičević, X. Lu, J. C. Vasquez, and J. M. Guerrero, "Dc microgrids—part i: A review of control strategies and stabilization techniques," *IEEE Transactions on power electronics*, vol. 31, no. 7, pp. 4876–4891, 2015.

[16] A. J. Wood, B. F. Wollenberg, and G. B. Sheblé, *Power generation, operation, and control*. John Wiley & Sons, 2013.

[17] P. Kundur, "Power system stability," *Power system stability and control*, pp. 7–1, 2007.

[18] M. Taleb, M. J. Ditto, and T. Bouthiba, "Performance of short transmission lines models," in *2006 IEEE GCC Conference (GCC)*. IEEE, 2006, pp. 1–7.

[19] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, and J. Lavaei, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2941–2962, 2017.

[20] F. Dörfler, J. W. Simpson-Porco, and F. Bullo, "Electrical networks and algebraic graph theory: Models, properties, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 977–1005, 2018.

[21] J. R. Magnus and H. Neudecker, "The elimination matrix: some lemmas and applications," *SIAM Journal on Algebraic Discrete Methods*, vol. 1, no. 4, pp. 422–449, 1980.

[22] E. Fabbiani, P. Nahata, G. De Nicolao, and G. Ferrari-Trecate, "Identification of ac networks via online learning," *arXiv preprint arXiv:2003.06210*, 2020.

[23] E. Fabbiani and P. Nahata, "Transformation matrix for non-redundant parametrization of laplacian matrices," Python: https://git.io/JfzHg, MATLAB: https://git.io/JfzH0, 2020, accessed September 27th, 2021.

[24] O. Ardakanian, Y. Yuan, R. Dobbe, A. von Meier, S. Low, and C. Tomlin, "Event detection and localization in distribution grids with phasor measurement units," in *2017 IEEE Power & Energy Society General Meeting*. IEEE, 2017, pp. 1–5.

[25] M. H. Hayes, *Statistical digital signal processing and modeling*. John Wiley & Sons, 2009.

[26] Y. Liu and F. Ding, "Convergence properties of the least squares estimation algorithm for multivariable systems," *Applied Mathematical Modelling*, vol. 37, no. 1-2, pp. 476–483, 2013.

[27] S. Bittanti and P. Bolzern, "Recursive least-squares identification algorithms with incomplete excitation: Convergence analysis," *IEEE Transactions on Automatic Control*, vol. 35, no. 12, 1990.

[28] "Pecan street inc," https://www.pecanstreet.org/dataport/, accessed May 2020.

[29] J.-S. Brouillon, E. Fabbiani, P. Nahata, F. Dörfler, and G. Ferrari-Trecate, "Bayesian error-in-variables models for the identification of power networks," *arXiv preprint arXiv:2107.04480*, 2021.