Bayesian Methods for the Identification of Distribution Networks

Jean-Sébastien Brouillon, Emanuele Fabbiani, Pulkit Nahata, Florian Dörfler, and Giancarlo Ferrari-Trecate

***Abstract*— The increasing integration of intermittent renewable generation, especially at the distribution level, necessitates advanced planning and optimisation methodologies contingent on the knowledge of the admittance matrix, capturing the topology and line parameters of an electric network. However, a reliable estimate of the admittance matrix may either be missing or quickly become obsolete for temporally varying grids. In this work, we propose a data-driven identification method utilising voltage and current measurements collected from micro-PMUs. More precisely, we first present a maximum likelihood approach and then move towards a Bayesian framework, leveraging the principles of maximum a posteriori estimation. In contrast with most existing contributions, our approach not only factors in measurement noise on both voltage and current data, but is also capable of exploiting available a priori information such as sparsity patterns and known line admittances. Simulations conducted on benchmark cases demonstrate that, compared to other algorithms, our method can achieve greater accuracy.**

I. Introduction

The key to realising green energy systems of tomorrow is the large-scale integration of renewable energy sources (RESs) at the distribution-grid level. While adding new capabilities, RES proliferation leads to additional risks such as reverse power flows and over-voltage—especially during periods of peak generation and low consumption [1]. Distribution grid operators are consequently required to put in place intelligent monitoring and control algorithms in order to maintain existing levels of grid safety and reliability [2], [3], [4], [5]. A perquisite for deploying such algorithms is grid identification, or the knowledge of admittance matrix embedding topology and line parameters of the distribution grid.

An exact estimate of the admittance matrix is hard to obtain for distribution grids, in particular as topological information and line parameter values either are unavailable for large chunks of the network, or become obsolete in the event of a topology change. To circumvent this issue, many recent contributions work out an up-to-date admittance matrix estimate by utilising data collected from micro-phasor measurement units (μPMUs). Although a more recent development than PMUs, which are commonly deployed on transmission systems, μPMUs have already been installed in distribution grids across America, Asia, and Europe; and their penetration is expected to steadily increase in the coming years [6].

In [7], [8], grid identification by means of inverter probing is explored. Both works, besides employing approximate linearized power-flow equations, are restricted to radial networks. Albeit requiring voltage and current (or power) measurements at each bus of the grid, identification approaches in [9], [10], [11], [12], [13] can be applied to both radial and meshed structures. Online design-of-experiment identification procedures are presented in [9], [10]; nonetheless, the proposed algorithms require control authority on the grid's state or additional measurements of line power flows, and neglect structural properties of the admittance matrix, such as symmetry and Laplacianity. In [12], [13], these properties are used to eliminate redundant admittance matrix parameters. More specifically, [12] details an identification method drawing on recursive least squares, while [13] proposes an adaptive Lasso algorithm promoting sparsity.

All the foregoing works suffer from two limitations. First, they either completely disregard measurement errors or assume errors solely on certain measurements. This can be limiting in practice, for μPMUs introduce an unavoidable error on all measured electric variables [14], [15]. Second, they do not capitalize on grid information which may already be available *a priori*, for instance sparsity patterns and known network sections and line parameters. To do away with the first limitation, [16], [17] introduce error-in-variable (EIV) models taking into consideration all sources of measurement errors. That notwithstanding, they leave aside all prior information, including structural properties of the admittance matrix, which can potentially improve grid identification.

In this work, we address the limitations of existing works by putting forth a novel Bayesian grid identification framework which incorporates EIV models and takes advantage of the principles of maximum likelihood estimation (MLE). Accounting for errors on both voltage and current data, our approach exploits not only the inherent structural properties of the admittance matrix, but also allows the use of information known *a priori*. We show that knowing whether lines are inductive can greatly increase the efficacy of a sparsity-promoting prior.

In order to substantiate the efficacy of our method, we conduct simulations with realistic voltage and current data, and practical μPMU noise levels. We also compare the performance with other grid identification methods. The presented analysis shows not only that EIV models are

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

J.S. Brouillon, P. Nahata, and G. Ferrari-Trecate are with the Institute of Mechanical Engineering, École Polytechnique Fédérale de Lausanne, Switzerland. Email addresses: {jean-sebastien.brouillon, pulkit.nahata, giancarlo.ferraritrecate}@epfl.ch.

E. Fabbiani is with the Identification and Control of Dynamic Systems Laboratory, University of Pavia, Italy. Email address: emanuele.fabbiani01@universitadipavia.it.

F. Dörfler is with the Automatic Control Laboratory, Swiss Federal Institute of Technology (ETH), Switzerland. Email address: dorfler@control.ee.ethz.ch.

needed to obtain reasonable grid estimates, but also that sparsity needs to be enforced if the topology of the network is unknown.

A. Preliminaries and Notation

Let $\boldsymbol{\jmath} = \sqrt{-1}$ denote the imaginary unit. For $x \in \mathbb{C}^n$, $\overline{x}$ is x's complex conjugate and $|x|$ its magnitude, both taken element-wise. $[x]$ is the diagonal matrix of order n associated with x. Throughout, $\mathbb{1}_n$ and $\mathbb{0}_n$ are n-dimensional vectors of all ones and zeros, whereas $\mathcal{I}_n$ and $\mathcal{O}_{m\times n}$ represent n-by-n identity and m-by-n zero matrices, respectively. For $i = 1, ..., n$, e_i is the i^{th} unit vector, A_i is the i^{th} row vector of a matrix A, and $\text{vec}(A) = [A_{1\cdot}^\top \cdots A_{n\cdot}^\top]^\top$ is A's column vectorization. Given a square matrix A, $\text{vech}(A)$ is the $n(n+1)/2$-dimensional vector obtained by eliminating all supradiagonal elements of A from $\text{vec}(A)$, and $\text{ve}(A)$ is the $n(n-1)/2$-dimensional vector obtained by removing diagonal elements from $-\text{vech}(A)$. The symbol $\propto$ represents a proportionality relation.

II. Grid model and data collection

A. Power grid model

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 [18]. To each edge $(i, k) \in \mathcal{E}$ we associate a complex weight equal to the line admittance $y_{ik} = g_{ik} + \boldsymbol{\jmath} b_{ik}$, 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.

Due to the structure of power networks, Y has peculiar properties. First, Y is typically sparse, as each bus is connected to few others. Second, without phase-shifting transformers and series capacitors, Y is symmetric. Third, for network where shunt elements are negligible, Y is Laplacian [19]. Standard distribution networks usually fulfil these three criterion and it is safe to assume the following [20].

Assumption 1. *The admittance matrix Y is symmetric and Laplacian (i.e. $L\mathbb{1}_n = \mathbb{0}_n$).*

We consider the network to be phase-balanced and operating in sinusoidal regime. To each bus $h \in \mathcal{V}$, we associate a phasor voltage $v_h e^{\boldsymbol{\jmath}\theta_h} \in \mathbb{C}$, where $v_h > 0$ is the voltage magnitude and $\theta_h \in \mathbb{R}$ the voltage angle, and a phasor current $i_h e^{\boldsymbol{\jmath}\phi_h} \in \mathbb{C}$, representing the injection at the node. We assume the point of common coupling (PCC) with the main grid to be the slack bus with fixed $v_0 = 1$ and $\theta_0 = 0$. 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 currents, and $v \in \mathbb{C}^n$ the vector of nodal voltages [21].

B. Data and measurement noise

We assume that each bus of the network is equipped with μPMU, while we do not require electrical variables to be measured on the lines.

Assumption 2. *The network is completely observable, that is current injections and voltages are measured at each node.*

μPMUs can provide up to 0.01% (proportional to their maximum ratings) and 0.01° accuracy, for sampling rates up to twice the grid's frequency [22]. By design, they measure current and voltage in polar coordinates, that is in terms of magnitude and phase [23]. With both theoretical and empirical arguments, studies have shown that the measurement noise is approximately Gaussian in polar coordinates, with zero mean and constant variance [15, Sec. 2.1].

In (1), the admittance matrix Y establishes a linear relationship between the real and the imaginary parts of i and v, but the equation becomes non-linear if magnitude and phase are considered. In order to preserve the linearity, one needs to transform the measurements and their associated noise from polar to Cartesian coordinates.

We consider a generic phasor measured by a μPMU. Without loss of generality, we will only discuss the case of a voltage phasor: the same arguments apply to the current. Let $\tilde{v}$ and $\tilde{\theta}$ denote the measured magnitude and phase, v and θ the actual unobservable variables. Then,

$$\tilde{v} = v + \epsilon, \tag{2a}$$

$$\tilde{\theta} = \theta + \delta, \tag{2b}$$

where $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon)$ and $\delta \sim \mathcal{N}(0, \sigma_\delta)$ are independent zero-mean Gaussian variables [15, Sec. 2.1].

Assumption 3. *The measurement noise is white, that is with no temporal correlation. The noise on the magnitude is independent from the one on the phase and from the noise on any other node of the network.*

Similarly to [17], to characterize the noise we compute the expected value and variance of $\Delta c + \boldsymbol{\jmath}\Delta d = \tilde{v}e^{-j\tilde{\theta}} - ve^{-j\theta}$:

$$\mathrm{E}[\Delta c] = v\cos\theta(e^{-\sigma_\delta^2/2} - 1), \tag{3a}$$

$$\mathrm{E}[\Delta d] = v\sin\theta(e^{-\sigma_\delta^2/2} - 1), \tag{3b}$$

$$\begin{aligned}\mathrm{Var}[\Delta c] = {} & v^2 e^{-\sigma_\delta^2}[\cos^2\theta(\cosh\sigma_\delta^2 - 1) + \sin^2\theta\sinh\sigma_\delta^2] \\ & + \sigma_\epsilon^2 e^{-\sigma_\delta^2}[\cos^2\theta\cosh\sigma_\delta^2 + \sin^2\theta\sinh\sigma_\delta^2],\end{aligned} \tag{4a}$$

$$\begin{aligned}\mathrm{Var}[\Delta d] = {} & v^2 e^{-\sigma_\delta^2}[\sin^2\theta(\cosh\sigma_\delta^2 - 1) + \cos^2\theta\sinh\sigma_\delta^2] \\ & + \sigma_\epsilon^2 e^{-\sigma_\delta^2}[\sin^2\theta\cosh\sigma_\delta^2 + \cos^2\theta\sinh\sigma_\delta^2],\end{aligned} \tag{4b}$$

$$\mathrm{Cov}[\Delta c, \Delta d] = \sin\theta\cos\theta e^{-2\sigma_\delta^2}[\sigma_\epsilon^2 + v^2(1 - e^{\sigma_\delta^2})]. \tag{4c}$$

The noise bias (3) and covariance matrix (4) are expressed in terms of the unobservable actual values v and θ but, for any realistic noise level, they can be replaced by the measurements $\tilde{v}$ and $\tilde{\theta}$ [17]. The noise bias will be disregarded, as it can be compensated and removed from the data.

Therefore, we will model the noise on a phasor measurement as

$$\begin{bmatrix}\Delta c\\ \Delta d\end{bmatrix} \sim \mathcal{N}\left(\mathbb{0}_2, \Sigma_v\right), \tag{5}$$

with the elements of Σ_v defined by (4). The covariance matrix Σ_v is not constant in time, but changes with the actual values of phase and magnitude: such property will be further discussed in Section IV, while presenting the estimation methods.

III. Problem Statement

Consider a power distribution network as described in Section II, fulfilling Assumptions 1 and 2. The identification problem amounts to reconstruct the admittance matrix from a sequence of voltage and current measurements corresponding to different steady states of the system [11], [13].

Let N be the number of samples, $v_t \in \mathbb{C}^n$ and $i_t \in \mathbb{C}^n$ the vectors of current injections and voltages for $t = 1, \ldots, N$. From (1), one can obtain

$$I = YV, \tag{6}$$

where $V = [v_1, v_2, \ldots, v_N] \in \mathbb{C}^{N\times n}$, and $I = [i_1, i_2, \ldots, i_N] \in \mathbb{C}^{N\times n}$.

Unfortunately, as described in Section II-B, the available current and voltage phasors are corrupted by measurement noise. Therefore, in place of the actual electrical variables V and I, only noisy samples $\tilde{V}$ and $\tilde{I}$ are available, where

$$\tilde{V} = V + \Delta V, \tag{7a}$$
$$\tilde{I} = I + \Delta I, \tag{7b}$$

with $\Delta V \in \mathbb{C}^{N\times n}$ and $\Delta I \in \mathbb{C}^{N\times n}$ denoting the complex measurement noise. The network identification problem then translates into the estimation of Y given $\tilde{V}$ and $\tilde{I}$.

IV. Maximum likelihood estimation

The high correlation between measurements may require a large number of samples to obtain an estimate with a satisfying variance. Therefore, we chose to use the maximum likelihood estimator for its efficiency and consistency properties [24, chapter 7, 10].

The variables V and I are deterministic and the noises ΔV and ΔI independent. The likelihood of $(\tilde{V}, \tilde{I})$, knowing the parameters and actual variables can be written as follows.

$$\begin{aligned} l(\tilde{V}, \tilde{I}|V, I, Y) &\propto p(\tilde{V}, \tilde{I}|V, I, Y), && \text{(8a)}\\ &\propto p(V + \Delta V|V, Y)p(I + \Delta I|I, Y), && \text{(8b)}\\ &\quad \text{s.t. } (\tilde{V} - \Delta V)Y = \tilde{I} - \Delta I. \end{aligned}$$

To work with real variables, we define

$$a = \begin{pmatrix}\Re\left(\mathrm{vec}(V)\right)\\ \Im\left(\mathrm{vec}(V)\right)\end{pmatrix}, \quad b = \begin{pmatrix}\Re\left(\mathrm{vec}(I)\right)\\ \Im\left(\mathrm{vec}(I)\right)\end{pmatrix}, \tag{9}$$
$$A = \begin{pmatrix}\Re(\mathcal{I}_n \otimes V) & -\Im(\mathcal{I}_n \otimes V)\\ \Im(\mathcal{I}_n \otimes V) & \Re(\mathcal{I}_n \otimes V)\end{pmatrix}.$$

The same transformations can be applied to $\tilde{V}$, $\tilde{I}$, ΔV, and ΔI, resulting in the vectors $\tilde{a}$, $\tilde{b}$, Δa, Δb, and the matrices $\tilde{A}$ and ΔA. Note that a and A contain the same elements but arranged differently. We will therefore use them interchangeably when describing an optimization problems over Δa or ΔA. The matrix A is introduced to represent the product VY with real and vectorized quantities as shown in (10) below.

Using the vectorized notations (9), we assume that $\Delta a \sim \mathcal{N}(0, \Sigma_a)$ and $\Delta b \sim \mathcal{N}(0, \Sigma_b)$, as per the approximate noise model discussed in Section II-B. The covariance matrices Σ_a and Σ_b are computed from (4) as explained in Appendix I. The likelihood (8) becomes

$$\mathcal{L}(\tilde{a}, \tilde{b}|a, b, Y) = -\Delta b^\top \Sigma_b^{-1} \Delta b - \Delta a^\top \Sigma_a^{-1} \Delta a, \tag{10a}$$
$$\text{s.t. } \tilde{b} - \Delta b = (\tilde{A} - \Delta A)\begin{pmatrix}\Re\left(\mathrm{vec}(Y)\right)\\ \Im\left(\mathrm{vec}(Y)\right)\end{pmatrix}. \tag{10b}$$

For a fixed, albeit unknown V and I, we use the shorthand notation $\mathcal{L}(Y, \Delta a, \Delta b)$. Minimizing $-\mathcal{L}$ for Δb, Δa and Y yields the MLE.

V. Including prior knowledge

A regularization term added to the MLE can be interpreted as the log-likelihood of a suitably defined Bayesian prior distribution [25], [26], [27]. This approach is not limited to sparsity, and can be applied to incorporate several other pieces of information.

A. Structural priors

Under Assumption 1, entries on and above the main diagonal of Y can be derived from the elements below the diagonal. Therefore, in order to avoid redundant variables, one can proceed as in [12] and use duplication and transformation matrices D and T to remove the redundant entries from the identification problem and solve for $y = \mathrm{ve}([\Re(Y), \Im(Y)])$ instead. In case some entries of Y are known to be zero, one can derive variants of D and T and also remove these zero entries from y by following a procedure similar to the one presented in [12, Appendix 2]. In both cases, $\left[\Re\left(\mathrm{vec}(Y)\right)^\top, \Im\left(\mathrm{vec}(Y)\right)^\top\right]^\top = DTy$ and the equation (10b) becomes

$$\tilde{b} - \Delta b = (\tilde{A} - \Delta A)DTy. \tag{11}$$

B. Bayesian estimation

Line admittances, even if measured, are known up to a tolerance. Some knowledge of Y's structure, such as its sparsity, may also not be certain or precisely defined, and therefore does not enable the use of the structural priors

as defined in Section V-A. This kind of uncertainty can be modeled via Bayesian prior distributions.

Following [28], we describe how to compute Maximum *A Posteriori* (MAP) estimates for the error-in-variables model (11). Using Bayes' rule, the posterior probability density is

$$p(V, I|\tilde{V}, \tilde{I}, Y) = p(\tilde{V}, \tilde{I}, Y|V, I)\frac{p(V, I)}{p(\tilde{V}, \tilde{I})}p(Y), \qquad (12)$$

where we assume that the line parameters y are independent of the grid state (V, I) or its measurement $(\tilde{V}, \tilde{I})$. The factor $\frac{p(V,I)}{p(\tilde{V},\tilde{I})}$ can be neglected as it is a quotient of non-informative priors [28], defined as uniform distributions over the finite set of feasible voltages and currents. The negative log-likelihood minimization problem is then written as

$$\min_{y,\Delta a,\Delta b} -\mathcal{L}(y, \Delta a, \Delta b) - \log\left(p(y)\right), \qquad (13)$$
$$\tilde{b} - \Delta b = \left(\tilde{A} - \Delta A\right) DTy,$$

with $\mathcal{L}$ defined by (10a). The optimizers of (13) provide maximum *a posteriori* (MAP) estimates y_{MAP}.

In order to obtain good robustness properties, we choose an element-wise Laplace distribution $p(y_i) \propto e^{-\lambda|y_i|}$. In vector form, this corresponds to $-\log(p(y)) = \lambda\|y\|_1$ the Lasso penalty [26]. Priors are centered on the believed value of y_i, which can be different from zero, e.g., in case of an existing line. More generally, one can also believe that a linear combination of y has a particular value [26], for example that two lines have the same admittance. The probability density $p(Ly - \mu)$ of a linear transformation $y \to Ly - \mu$ can describe such a belief. The penalty function is then

$$-\log(p(y)) = \lambda\|Ly - \mu\|_1. \qquad (14)$$

If L and μ are not chosen carefully, the prior can induce a large bias on the estimate estimate. This is often the case for a Lasso penalty with a too large λ. In the following section, we will explain how to limit this bias using a minimum amount of prior information or assumptions.

VI. Unbiased sparsity promotion

The most common sparsity promoting prior is the ℓ_1 penalty, corresponding to $L = \mathcal{I}$ and $\mu = \mathbb{0}$ [25]. However, this methods can bias the estimate $\hat{y}$ [29]. The adaptive Lasso method allows to remove this bias *asymptotically* (i.e. for $N \to \infty$) [29], [13]. With the representation (14), the adaptive Lasso penalty corresponds to $L = [|y_{\text{MLE}}|]^{-1}$.

The assumption that $N \to \infty$ is not practical, and with a lot of noise the bias of Lasso can remain high. To overcome this issue, the hyper-parameter λ can be interpreted as the Lagrange multiplier of the constraint $\|y\|_1 \leq \hat{\gamma}$, where the minimal $\hat{\gamma}$ can be estimated from the MLE as $\hat{\gamma} = \|\hat{y}_{\text{MLE}}\|_1$. The error $\hat{y}_{\text{MLE}} - y$ is a Gaussian random variable centered at zero [30]. Hence, for a zero element i, $|\hat{y}_{i,\text{MLE}} - y_i|$ follows a half-Gaussian distribution and is centered at $\frac{\sqrt{2}}{\sqrt{\pi}}\sigma_i > 0$ (assuming a standard deviation σ_i), which creates a bias on $\hat{\gamma}$. If one knows the sign $s_i \in \{-1, 1\}$ of all y_i, which is always positive for conductances and generally negative for susceptances (i.e. if lines are inductive), one can replace $|\hat{y}_{i,\text{MLE}}|$ by $s_i\hat{y}_{i,\text{MLE}}$ and keep the zero-centered Gaussian distribution. Therefore, an unbiased estimate $\hat{\gamma}$ is given by

$$\hat{\gamma} = \sum\nolimits_{i=1}^{n(n-1)} s_i\hat{y}_{i,\text{MLE}}. \qquad (15)$$

This estimate can be used to tune λ so as to obtain $\|\hat{y}_{\text{MAP}}\|_1 = \hat{\gamma}$. It can also be used to design a prior on the sum $\sum_{i=1}^{n(n-1)} s_i\hat{y}_i$ itself. In conclusion, the adaptive Lasso penalty corrected with $\hat{\gamma}$ is expressed as

$$-log(p(y|s)) = \lambda_s|s^\top y - \hat{\gamma}| + \lambda\|[|y_{\text{MLE}}|]^{-1}y\|_1. \qquad (16)$$

Note that the prior (16) only requires the knowledge of the sparsity of the network and which lines are not inductive.

VII. Numerical methods

The optimization problem (13) is NP-hard due to the multiplication of y and ΔA in the constraint. Similarly to the weighted TLS, no closed-form solution has been found yet [31]. To solve it, we will use a variant of the alternate block coordinate descent (BCD) algorithm, called broken adaptive ridge regression (BAR) [32], [33]. It is based on the following approximation, made at every iteration $k > 0$ with y_{k-1} the estimate at the previous iteration.

$$\|Ly - \mu\|_q^2 \approx (y - \mu)^\top L^\top[|Ly_{k-1} - \mu| + \alpha\mathcal{I}]^{q-2}(Ly - \mu),$$

where $0 \leq q \leq 2$ and $\alpha > 0$ is a small parameter for numerical stability. The update is written as

$$\Delta A_{k+1} = \arg\min_{\Delta A} -\mathcal{L}\left(y_k, \Delta A, \tilde{b} - \left(\tilde{A} - \Delta A\right)y_k\right), \qquad (17a)$$

$$y_{k+1} = \arg\min_{y} -\mathcal{L}\left(y, \Delta A_{k+1}, \tilde{b} - \left(\tilde{A} - \Delta A_{k+1}\right)y\right) + (Ly - \mu)^\top[|Ly_{k-1} - \mu| + \alpha\mathcal{I}]^{-1}(Ly - \mu). \qquad (17b)$$

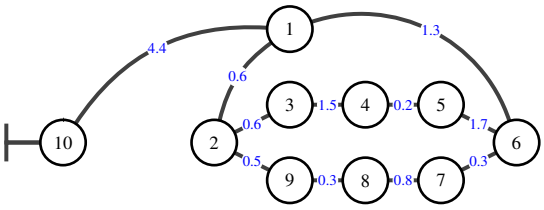
Empirical evidence suggests that, since both (17a) and (17b) have closed-form solutions, each iteration is computationally simpler than for the standard BCD.

VIII. Simulations

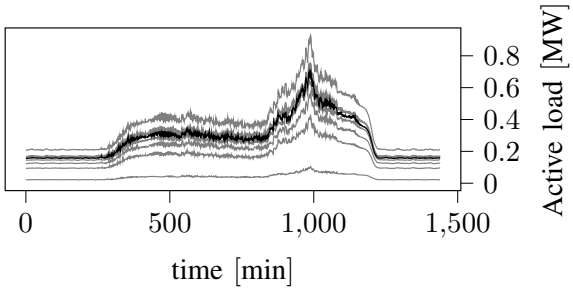
A. Setup

We test our methods on a medium voltage distribution grid [34, feeder 1] (Fig. 1a). Nodes 1 to 9 feed smaller grids where households are connected and node 10 is connected to an external sub-transmission grid. In order to provide a comparison, state-of-the-art estimators are also adopted to perform the same task. The load profiles are simulated following [35]: a sample of the resulting shapes is given in Fig. 1b. The load profiles of random individual households are summed until the nominal power of a node is reached. The voltage and current samples are then generated using the PandaPower framework [36].

To simulate μPMUs, current and voltage measurements are collected with a frequency of 50Hz. As load profiles have



(a) Graph of the feeder 1 of a medium voltage distribution network from [34] with line lengths in km. The connection to the external grid is made at node 10.



(b) Load profiles. Node 7's profile is displayed in black.

Fig. 1: Representation of the simulation settings.

a period of 1 minute, we interpolate the missing load data linearly. To increase the computational speed, we apply a low-pass filter on the data, implemented as a moving average. We adopt a window length of 17000 and under-sample the result by the same amount. This provides 500 samples of both current and voltages for each node as complex phasors.

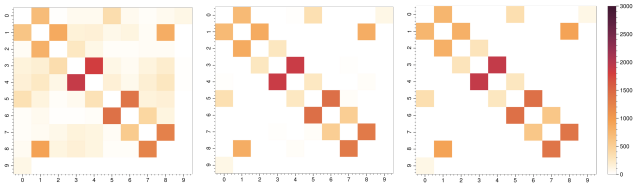
We corrupt both current and voltage measurements with a zero-mean Gaussian noise. In particular, a noise $\mathcal{N}(0, \sigma_v^2)$ is added to the magnitude and a noise $\mathcal{N}(0, \sigma_\theta^2)$ to the phase. Coherent with the specifications of state-of-the-art micro-synchrophasors (Section II-B), we pick $\sigma_v = 0.01\%$ and $\sigma_\theta = 0.01°$. Assuming that the µPMUs are adapted to their nodes, we choose a rating of four times the nominal power of the node.

B. Results

Thanks to the high accuracy of µPMUs, the TLS method manages to retrieve a fair estimate of Y (Fig. 2a). However, although the sparsity of the actual admittance matrix above 98%, the TLS estimate is almost dense. It can therefore not be used to identify the topology. With the noise levels used in our simulations, the MLE is almost equal to the TLS estimate is therefore not shown in Figs. 2 and 3 and Table I.

Noise	MAP	TLS	Lasso	OLS
5e-4	**26.64%**	48.77%	95.53%	96.15%
2e-4	**11.13%**	23.66%	84.71%	85.92%
1e-4	**5.81%**	12.89%	65.53%	67.30%
5e-5	**3.16%**	6.76%	38.75%	40.88%
2e-5	**1.62%**	2.79%	10.58%	12.54%
1e-5	**0.81%**	1.41%	2.80%	3.94%

TABLE I: Estimation error for various levels of noise, equal in phase and magnitude for current and voltages. The best performing estimator is shown in bold.



(a) TLS estimate. (b) MAP estimate. (c) Real admittance matrix.

Fig. 2: Heat maps of the magnitude of estimates from TLS and MAP with 0.01% noise. In all these, for the sake of clarity, elements on the diagonal are not shown.

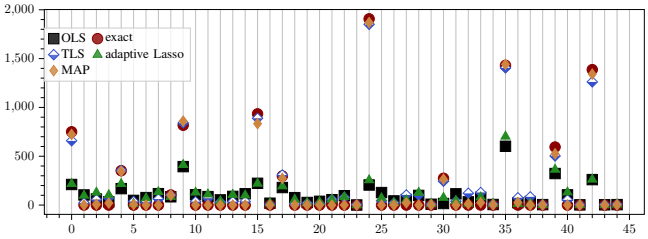


Fig. 3: Comparison of the magnitude of non-zero elements estimated with different methods with 0.01% noise.

Fig. 3 shows the estimated value of each unique entry of Y using the different methods. Table I shows the total relative error $\frac{\|\hat{Y}-Y\|_F}{\|Y\|_F}$ in terms of Frobenius norm for all methods and different noise levels. The quality of an estimate mainly depends on its use, but Figs. 2b and 2c show that the MAP estimate is much closer to the actual admittance matrix than the TLS output. However, Table I shows that it is close to its limits: with 0.05% noise, the MAP estimation yields 26.6% error. If one wants a better estimate without actually measuring lines, the only solution is to use more data for the identification, thus increasing the available information at the price of slowing the computation. This may therefore not be possible for big networks. Computational complexity is the first limit reached when identifying large networks.

IX. Conclusion

In this work, we proposed to exploit samples collected by micro-PMUs. Considering a realistic statistical model for the noise affecting both current and voltage measurements, we built maximum-likelihood and Bayesian estimators. We argued that the latter can outperform the former, due to their ability to exploit features of the grid, such as sparsity or inductiveness. Our argument is substantiated by numerical simulations on benchmark grids: even without any network-specific prior information, Bayesian methods outperformed state-of-the-art estimators with realistic noise levels.

Further research on Bayesian estimation for grid identification will focus on recursive methods to perform the estimation continuously and integrate both new samples and prior information.

APPENDIX I
COVARIANCE MATRIX

In order to solve the Maximum Likelihood problem (10) and all its subsequent refinements, one needs the covariance matrices Σ_a and Σ_b. The construction of the two is identical, thus we will focus on $\Sigma_a = \mathrm{Cov}[\Delta a] \in \mathbb{R}^{2nN \times 2nN}$ only.

From Section II-B, the Σ_a is sparse, having non-zero elements only on three diagonals: the main diagonal, hosting the variance, and the nNth super- and sub-diagonals, housing the covariance between the real and the imaginary part of the measurements. Such particular structure makes it possible to split Σ_a into four diagonal blocks and makes it easy to find Σ_a^{-1}.

It is also interesting to note that, up to a permutation of the elements in Δa, $\mathrm{Cov}[\Delta a]$ can be written as a block diagonal matrix where the 2-by-2 blocks are given by Σ_v in (5).

REFERENCES

[1] Y. Weng, Y. Liao, and R. Rajagopal, "Distributed energy resources topology identification via graphical modeling," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2682–2694, 2017.

[2] L. Schenato, G. Barchi, D. Macii, R. Arghandeh, K. Poolla, and A. Von Meier, "Bayesian linear state estimation using smart meters and pmus measurements in distribution grids," in *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2014, pp. 572–577.

[3] A. La Bella, S. Negri, R. Scattolini, and E. Tironi, "A two-layer control architecture for islanded ac microgrids with storage devices," in *2018 IEEE Conference on Control Technology and Applications (CCTA)*, 2018, pp. 1421–1426.

[4] A. Iovine, T. Rigaut, G. Damm, E. De Santis, and M. D. Di Benedetto, "Power management for a dc microgrid integrating renewables and storages," *Control Engineering Practice*, vol. 85, pp. 59–79, 2019.

[5] A. Parisio, E. Rikos, and L. Glielmo, "Stochastic model predictive control for economic/environmental operation management of microgrids: An experimental case study," *Journal of Process Control*, vol. 43, pp. 24–37, 2016.

[6] D. S. Kumar, J. Savier, and S. Biju, "Micro-synchrophasor based special protection scheme for distribution system automation in a smart city," *Protection and Control of Modern Power Systems*, vol. 5, no. 1, pp. 1–14, 2020.

[7] 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.

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

[9] X. Du, A. Engelmann, Y. Jiang, T. Faulwasser, and B. Houska, "Optimal experiment design for ac power systems admittance estimation," *arXiv preprint arXiv:1912.09017*, 2019.

[10] M. Angjelichinoski, C. 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)*, June 2017, pp. 202–206.

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

[12] E. Fabbiani, P. Nahata, G. D. Nicolao, and G. Ferrari-Trecate, "Identification of ac networks via online learning," *IEEE Transactions on Control Systems Technology*, to appear.

[13] O. Ardakanian, V. W. S. 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.

[14] B. Pinte, M. Quinlan, and K. Reinhard, "Low voltage micro-phasor measurement unit (μpmu)," in *2015 IEEE Power and Energy Conference at Illinois (PECI)*. IEEE, 2015, pp. 1–4.

[15] S. Sarri, "Methods and performance assessment of pmu-based real-time state estimation of active distribution networks," Ph.D. dissertation, EPFL, 2016.

[16] J. Yu, Y. Weng, and R. Rajagopal, "Patopa: A data-driven parameter and topology joint estimation framework in distribution grids," *IEEE Transactions on Power Systems*, vol. PP, 05 2017.

[17] A. Wehenkel, A. Mukhopadhyay, J.-Y. Le Boudec, and M. Paolone, "Parameter estimation of three-phase untransposed short transmission lines from synchrophasor measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6143–6154, 2020.

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

[19] P. Kundur, *Power System Stability and Control*. CRC Press New York, NY, USA, 2007.

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

[21] 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.

[22] A. Von Meier, E. Stewart, A. McEachern, M. Andersen, and L. Mehrmanesh, "Precision micro-synchrophasors for distribution systems: A summary of applications," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2926–2936, 2017.

[23] E. Dusabimana and S.-G. Yoon, "A survey on the micro-phasor measurement unit in distribution networks," *Electronics*, vol. 9, no. 2, p. 305, 2020.

[24] G. Casella and R. Berger, *Statistical Inference*. Duxbury Press, 01 2002.

[25] T. Park and G. Casella, "The bayesian lasso," *Journal of the American Statistical Association*, vol. 103, no. 482, pp. 681–686, 2008.

[26] M. Kyung, J. Gill, M. Ghosh, and G. Casella, "Penalized regression, standard errors, and bayesian lassos," *Bayesian Analysis*, vol. 5, pp. 369–412, 06 2010.

[27] P. Holland, "Weighted ridge regression: Combining ridge and robust regression methods," 10 1973.

[28] X. Fang, B. Li, H. Alkhatib, W. Zeng, and Y. Yao, "Bayesian inference for the errors-in-variables model," *Studia Geophysica et Geodaetica*, vol. 61, pp. 35–52, 02 2017.

[29] H. Zou, "The adaptive lasso and its oracle properties," *Journal of the American Statistical Association*, vol. 101, no. 476, pp. 1418–1429, 2006.

[30] J. Crassidis and Y. Cheng, "Error-covariance analysis of the total least squares problem," *Journal of Guidance, Control, and Dynamics*, vol. 37, 07 2014.

[31] "Overview of total least-squares methods," *Signal Processing*, vol. 87, no. 10, pp. 2283–2302, 2007, special Section: Total Least Squares and Errors-in-Variables Modeling.

[32] L. Dai, K. Chen, Z. Sun, Z. Liu, and G. Li, "Broken adaptive ridge regression and its asymptotic properties," *Journal of Multivariate Analysis*, vol. 168, 08 2018.

[33] F. Frommlet and G. Nuel, "An adaptive ridge procedure for l0 regularization," *PloS one*, vol. 11, 05 2015.

[34] K. Strunz, E. Abbasi, R. Fletcher, N. Hatziargyriou, R. Iravani, and G. Joos, *TF C6.04.02 : TB 575 – Benchmark Systems for Network Integration of Renewable and Distributed Energy Resources*, 04 2014.

[35] G. Hoogsteen, A. Molderink, J. L. Hurink, and G. J. M. Smit, "Generation of flexible domestic load profiles to evaluate demand side management approaches," in *2016 IEEE International Energy Conference (ENERGYCON)*, 2016, pp. 1–6.

[36] L. Thurner, A. Scheidler, F. Schafer, J. H. Menke, J. Dollichon, F. Meier, S. Meinecke, and M. Braun, "pandapower - an open source python tool for convenient modeling, analysis and optimization of electric power systems," *IEEE Transactions on Power Systems*, 2018.