

Identification of Power Line Outages

Shay Maymon, and Yonina C. Eldar, *Fellow, IEEE*

Abstract—This paper considers the problem of power line outages identification in its reformulation as a sparse recovery problem. Using only hourly basecase topology information and local (a.k.a internal) voltage phasor angle data available by phasor measurements unit, we seek the subset of line outages. We propose a least-squares formulation for solving the identification problem whose computational complexity is often reduced compared to existing approaches. A natural extension of the least-squares method leads to a generalization of the line outages identification problem in which the grid parameters are unknown. An iterative solution is developed to solve for the outaged lines in this general setting. Finally, we recognize that for internal line outages, partial information that is often assumed unknown can actually be computed from the data. We incorporate this knowledge into our algorithm, resulting in improved outages identification. Simulation results are provided to support our methods and observations.

Index Terms—Power line outages, phasor measurement units, sparsity, compressive sampling.

I. INTRODUCTION

A key aspect of situational awareness in the power grid is the knowledge of transmission line, transformer, and generator statuses. This is, in fact, the major component of the data shared via the north-American electric reliability corporation (NERC) system data exchange (SDX). Despite the clear need for information sharing across the network for situational awareness, there is limited real-time sharing of the legacy supervisory control and data acquisition (SCADA) measurements in the United States [1]. With the deployment of power measurements units (PMUs) [2] in the north American power-grid, significant efforts have been made to ensure that PMU data per local system is shared among all interested parties [3]. As opposed to SCADA measurements which provide power-related information and are available every 4 seconds, PMU also offers phase information and can provide 30-60 synchrophasors per second. Furthermore, using precise GPS timing, synchrophasors are time-stamped at the universal time coordinates, allowing data across large geographic areas to be consistently aggregated. Extensive research in applying PMU information to improve situational awareness has been conducted since their introduction, including applications in state estimation [4]–[6], dynamic security assessment [7]–[9], and visualization [10]–[12].

It is well known that major blackouts have occurred due in part to a lack of comprehensive situational awareness of the power grid [13]. Due to the high interconnection of modern

infrastructure systems, a change in conditions at any one location may have immediate impact over a wide area. Furthermore, the effect of a local disturbance can be magnified as it propagates through a network. Large-scale cascade failures can occur almost instantaneously and with consequences in remote regions. On the North American power grid, for example, where transmission lines link all electricity generation and distribution on the continent, wide-area outages in the late 1990s and summer of 2003 underscore the grid’s vulnerability to cascading effects.

Since the SDX updates can be provided only on an hourly basis [14], a method for timely identifying line outages throughout an electric interconnection is critical for wide-area monitoring in order to avoid failures from spreading quickly, leading to a grid-wide blackout. Most existing approaches for line outages identification use known system topology information together with real-time phase angle measurements that are typically obtained from PMUs, which can provide unique insights into the global operation of the grid. Current methods typically formulate the identification problem as a combinatorial complex problem, which can be computationally tractable only for single or, at most, double line outages [15]–[17]. However, in the face of cascading failures in recent blackouts, it is becoming increasingly crucial to cope with multiple line outages. A recent approach that can deal with multiple outages at affordable complexity adopts a Gauss-Markov graphical model of the power network. This method requires assumptions on the conditional independence among bus phasor angle measurements as well as the availability of real-time intersystem PMU data across the grid [18].

Zhu and Giannakis [19] have recognized the fact that the outaged lines represent a small fraction of the total number of lines and reformulated the problem of identifying line outages as a sparse vector estimation problem. Leveraging recent advances in compressive sampling [20]–[24], [28], they proposed a computationally efficient algorithm for identification of multiple line outages. Their approach uses only hourly basecase topology information and local real-time voltage phasor angle measurements obtained from PMUs.

This paper adopts the sparse linear model developed in [19] in which only a subset of voltage phasor angle data that corresponds to the subset \mathcal{N}_I of buses in the internal system is observed. The subset \mathcal{N}_E of buses in the external system is not observable. The sparse linear model relates the change in bus voltage phases to the sparse vector that captures information of line outages. Existing approaches for identifying line outages extract the model equations associated with the observed bus voltage phases while completely ignoring the model equations of the unobserved data [15]–[17]. We propose instead a least-squares approach in which both the unobserved phasor angle data and the sparse coefficient vector are jointly estimated.

Shay Maymon and Yonina C. Eldar are with the Department of Electrical Engineering, Technion, Haifa, Israel, e-mail: maymon.yonina@technion.ac.il. Manuscript received April 19, 2005; revised December 27, 2012.

This work is supported in part by an Aly Kaufman Fellowship, in part by the Israel Science Foundation under Grant no. 170/10, in part by the Ollendorf Foundation, in part by the SRC, and in part by the Intel Collaborative Research Institute for Computational Intelligence (ICRI-CI).

It turns out that estimating the unobserved phasor angle data using the least-squares approach and then substituting it back into the objective function yields a sparse vector estimation problem that is equivalent to the one we would have obtained if we recovered the sparse vector based on the extracted model equations associated with the observed data. Despite the equivalence between the sparse vector estimation problem obtained with the least-squares approach and the one resulted from just ignoring the unobserved data [19], least-squares is shown to be a more efficient approach for pre-processing the data. Namely, the computational complexity of eliminating the external phasor angle data by estimating it as a nuisance parameter in the least-squares approach is often much lower than that of extracting the model equations associated with the observed data. Simulations show that as the size of the internal network increases, the running time of our proposed algorithm is becoming significantly better than competing approaches.

The least-squares formulation also allows us to generalize the line-outages identification problem to the case in which the grid parameters are not accurately known. To this end, we develop an iterative solution based on the cyclic coordinate descent approach for estimating the grid parameters, the unobserved data, and the sparse vector. Simulations show that when the model of the grid parameters is taken into account in the estimation of the sparse vector, the percentage of correctly identified line outages is improved compared to the case in which the nominal values of the grid parameters are used instead. For the cases tested, we observed relative improvement in the range of 0.02% – 0.58%, where more noticeable improvement occurs as the perturbations in the grid parameters increase.

In solving the sparse vector estimation problem, both the support of the sparse vector as well as the values of its nonzero entries were assumed unknowns in [19]. We recognize that when the power outage is internal, the value of the corresponding entry in the sparse vector, often treated as unknown, can actually be computed from the data. When the outage is external we still treat the value of the corresponding entry in the sparse vector as unknown. To this end, we extend our sparse recovery algorithm to incorporate this information on the sparse vector for improved line outages identification. Two cases were simulated to test the impact of incorporating partial information of this form on the percentage of correctly identifying line outages. One in which all bus phasor angle measurements are available and another in which the identification is based only on a subset of phasor angle measurements. In both cases and for all levels of perturbation noise tested, better performance is achieved when incorporating the partial information in the estimation. For the first case we observed relative improvement of correct identification in the range of 1.77% – 3.13% and for the second case relative improvement of 0 – 20.06%.

The paper is organized as follows: Modeling of the power transmission network is given in Section II. Section III formulates the line outages identification problem. In Section IV the least-squares approach for estimating the unobserved phasor angle data is developed, the inverse approach is introduced, and the relation of these methods is discussed. Section V

generalizes the line outages identification problem to the case of unknown grid parameters. Section VI discusses sparse reconstruction when some of the power outages are internal. Finally, simulations are given in Section VII.

Notation: upper (lower) boldface letters will be used for matrices (column vectors); $()^T$ denotes transposition, \mathbf{I} is the identity matrix, $\|\cdot\|_p$ is the vector p -norm for $p \geq 1$, $\|\cdot\|_0$ is the l_0 seminorm, which is equal to the vector's number of nonzero entries, and $\|\mathbf{x}\|_{\mathbf{W}}^2 = \mathbf{x}^T \mathbf{W} \mathbf{x}$. The Moore-Penrose pseudo-inverse of \mathbf{A} is denoted by \mathbf{A}^\dagger , and an orthogonal projection onto $\mathcal{R}(\mathbf{A})$ is denoted by $P_{\mathbf{A}}^\perp$.

II. POWER TRANSMISSION NETWORK MODELING

Consider a power transmission network consisting of N buses (a.k.a nodes) denoted by the set $\mathcal{N} = \{1, 2, \dots, N\}$, and L transmission lines (a.k.a branches) that are represented by the set of edges $\mathcal{E} = \{(m, n)\} \subseteq \mathcal{N} \times \mathcal{N}$. A node can represent a generator or a load substation, whereas a line can stand for a transmission or distribution line. Power systems can be thought of as electric circuits of even continent-wide dimensions. Using Kirchoff's current law, the node currents $\mathbf{i} \in \mathcal{C}^{N \times 1}$ are shown to be related to the node voltages $\mathbf{v} \in \mathcal{C}^{N \times 1}$ through the following multivariate Ohm's law:

$$\mathbf{i} = \mathbf{Y} \mathbf{v}. \quad (1)$$

Here $\mathbf{Y} \in \mathcal{C}^{N \times N}$ is referred to as the bus-admittance matrix and is given by

$$[\mathbf{Y}]_{mn} = \begin{cases} \sum_{l \in \mathcal{N}_m} y_{ml} + y_{mm}, & m = n \\ -y_{mn}, & n \in \mathcal{N}_m \\ 0, & \text{o.w.}, \end{cases} \quad (2)$$

where \mathcal{N}_m is the set of neighboring buses linked to m . The line series admittance $y_{mn} = 1/z_{mn} = g_{mn} + jb_{mn}$, whose real and imaginary parts are called conductance and susceptance, is often used in place of the impedance z_{mn} . The bus-admittance matrix \mathbf{Y} is symmetric and more importantly sparse since its (m, n) -th off-diagonal entry is zero unless nodes m and n are directly connected. It is usually written in rectangular coordinates as $\mathbf{Y} = \mathbf{G}_Y + j\mathbf{B}_Y$.

Power flow models are useful for determining how injected power flows along all transmission lines. Let $S_m = P_m + jQ_m$ be the complex power injected at bus m whose real part P_m and its imaginary part Q_m are referred to as the active and reactive powers, respectively. Representing $S_m = \mathcal{V}_m \mathcal{I}_m^*$ where \mathcal{V}_m (\mathcal{I}_m) denotes the complex voltage (current) at bus m , the multivariate Ohm's law in (1) together with the representation of the complex nodal voltages in polar form, i.e., $\mathcal{V}_m = V_m e^{j\theta_m}$, yields

$$P_m = \sum_{n=1}^N V_m V_n ([\mathbf{G}_Y]_{mn} \cos(\theta_m - \theta_n) + [\mathbf{B}_Y]_{mn} \sin(\theta_m - \theta_n)) \quad (3a)$$

$$Q_m = \sum_{n=1}^N V_m V_n ([\mathbf{G}_Y]_{mn} \sin(\theta_m - \theta_n) - [\mathbf{B}_Y]_{mn} \cos(\theta_m - \theta_n)). \quad (3b)$$

Since P_m and Q_m depend on phase differences, power injections are invariant to phase shifts of bus voltages. This

explains why a selected bus called reference is conventionally assumed to have zero voltage phase without loss of generality. Assuming \mathbf{Y} is known, the standard power flow problem consists of fixing the pairs $\{P_m, V_m\}$ or the pairs $\{P_m, Q_m\}$ and solving the nonlinear equations in (3) for the remaining unknowns.

The linear DC power flow model [25], [26], whose importance is mostly due to its use in grid monitoring and optimization, provides a linear approximation of the power control flow model (3). This approximation is based on the following assumptions:

I. The power network is purely inductive; Resistances can be ignored and the conductance part \mathbf{G}_Y of \mathbf{Y} can be approximated by zero, i.e., $\mathbf{Y} = j\mathbf{B}_Y$.

II. The voltage phase differences across directly connected buses are small; $\theta_m - \theta_n \approx 0$ for every pair of neighboring buses (m, n) .

III. The magnitude of nodal voltages is approximated by one per unit.

Under these assumptions, the model in (3) reduces to

$$P_m = - \sum_{n \neq m} b_{mn} (\theta_m - \theta_n) \quad (4a)$$

$$Q_m = -b_{mm} - \sum_{n \neq m} b_{mn} (V_m - V_n), \quad (4b)$$

where b_{mn} is the susceptance of the (m, n) -th branch. Note that active powers depend on voltage phases, whereas reactive powers are solely expressible via voltage magnitudes.

Consider now the active subproblem in (4):

$$\begin{aligned} P_m &= - \sum_{n \neq m} b_{mn} (\theta_m - \theta_n) \\ &= \theta_m \left(- \sum_{n \neq m} b_{mn} \right) + \sum_{n \neq m} \theta_n b_{mn}, \end{aligned} \quad (5)$$

or in vector-matrix form

$$\mathbf{p} = \mathbf{B}\theta. \quad (6)$$

Here $\theta \in \mathcal{R}^N$ represents the voltage phasor angles of all buses in the network and $\mathbf{p} \in \mathcal{R}^N$ represents the corresponding injected power variables. The matrix \mathbf{B} is referred to as the weighted Laplacian matrix and is uniquely determined by the line susceptance parameters $\{b_{mn}\}$ and the topology-bearing information \mathcal{E} provided by the SDX, i.e.,

$$[\mathbf{B}]_{mn} = \begin{cases} - \sum_{k \in \mathcal{N}_m} b_{mk}, & m = n \\ b_{mn}, & n \in \mathcal{N}_m \\ 0, & \text{o.w.} \end{cases} \quad (7)$$

In vector-matrix form

$$\mathbf{B} = \mathbf{M}\mathbf{D}\mathbf{M}^T = - \sum_{l=1}^L b_l \mathbf{m}_l \mathbf{m}_l^T, \quad (8)$$

where the matrix \mathbf{M} , formed by columns $\{\mathbf{m}_l\}_{l=1}^L$, is referred to as the bus-line incidence matrix and is determined by the network topology. When l corresponds to the line connecting nodes m and n , the column vector \mathbf{m}_l has all its entries zero except the m th and n th, which take on the values 1 and

-1 , respectively. The diagonal matrix \mathbf{D} has its l th diagonal entry equal to $-b_l$, where here and after we abuse notation and simply denote by b_l the susceptance b_{mn} of the line l connecting nodes m and n . Similarly, the vector-matrix form of the reactive sub-problem in (4) is

$$\mathbf{q} = -\beta + \mathbf{B}\mathbf{v}, \quad (9)$$

where $\mathbf{q} \in \mathcal{R}^N$ represents the reactive power variables, and $\beta \in \mathcal{R}^N$ whose m -th entry is equal to b_{mm} . The next section mathematically formulates the problem of identifying line outages in the power network where it uses the linear DC power flow for solving it.

III. LINE OUTAGE IDENTIFICATION

Suppose that due to a failure in the grid, several outages occur and yield the post-event network $(\mathcal{N}, \mathcal{E}')$. Since the weighted Laplacian matrix \mathbf{B} is determined by the line susceptance parameters and the topology-bearing information, changes in the topology of the network due to line outages can be identified based on changes in \mathbf{B} . Clearly, if the k -th line is in outage, the post-event susceptance $b_k' = 0$ and the post-event weighted Laplacian matrix \mathbf{B}' becomes

$$\mathbf{B}' = - \sum_{l=1, l \neq k}^L b_l \mathbf{m}_l \mathbf{m}_l^T. \quad (10)$$

Thus, using pre- and post-event phasor angle data provided by PMUs together with the pre-event network-wide topology \mathbf{B} , line outages can be identified using the linear DC approximation of the active power (6).

We partition the network buses into two disjoint subsets: the subset \mathcal{N}_I of buses in the internal system, which are observable, and the subset \mathcal{N}_E of buses in the external system, which are unobservable. Similarly, we partition the pre- and post-event phasor angle data θ and θ' and the weighted Laplacian matrix \mathbf{B} as

$$\theta = \begin{bmatrix} \theta_I \\ \theta_E \end{bmatrix}, \theta' = \begin{bmatrix} \theta_I' \\ \theta_E' \end{bmatrix} \quad (11a)$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_I & \mathbf{B}_E \end{bmatrix}, \quad (11b)$$

where the subscripts I and E are associated with the subsets \mathcal{N}_I and \mathcal{N}_E , correspondingly. Let us also denote by \mathcal{E}_I the set of edges representing lines connecting internal nodes and by $\mathcal{E}_E = \mathcal{E} \setminus \mathcal{E}_I$ the set representing the remaining lines in the network. Given the pre- and post-event internal phasor angle data θ_I and θ_I' , as well as the pre-event network-wide topology \mathbf{B} , various approaches based on the DC model (4) are proposed in [15]–[17], [19] for unveiling the subset of line outages, denoted by $\tilde{\mathcal{E}} \subset \mathcal{E}$.

The power variables \mathbf{p}' of the quasi-stable post-event network can be expressed in terms of the pre-event power variables \mathbf{p} as

$$\mathbf{p}' = \mathbf{p} + \eta. \quad (12)$$

Here, the noise vector η accounts for the small perturbations due to e.g., variations in bus loads, which are usually modeled as a zero-mean vector, possibly Gaussian, with covariance

matrix $\sigma_\eta^2 \mathbf{I}$ [27]. Using the linear DC approximation for the active power (6) in (12) yields

$$\mathbf{B}'\theta' = \mathbf{B}\theta + \eta. \quad (13)$$

Then, introducing the difference $\tilde{\mathbf{B}} = \mathbf{B} - \mathbf{B}'$ denoting the weighted Laplacian matrix for the outage lines in $\tilde{\mathcal{E}}$ and representing \mathbf{B}' in terms of $\tilde{\mathbf{B}}$, the model in (13) becomes

$$\mathbf{B}(\theta' - \theta) = \tilde{\mathbf{B}}\theta' + \eta. \quad (14)$$

Denoting the change in bus voltage phases by $\tilde{\theta} = \theta' - \theta$ and using

$$\begin{aligned} \tilde{\mathbf{B}} &= \sum_{l \in \mathcal{E}} (-b_l) \mathbf{m}_l \mathbf{m}_l^T - \sum_{l \in \mathcal{E}'} (-b'_l) \mathbf{m}_l \mathbf{m}_l^T \\ &= - \sum_{l \in \tilde{\mathcal{E}}} b_l \mathbf{m}_l \mathbf{m}_l^T, \end{aligned} \quad (15)$$

which follows from (8), we obtain

$$\begin{aligned} \mathbf{B}\tilde{\theta} &= - \sum_{l \in \tilde{\mathcal{E}}} b_l (\mathbf{m}_l \mathbf{m}_l^T) \theta' + \eta \\ &= \sum_{l \in \tilde{\mathcal{E}}} \left(-b_l \mathbf{m}_l^T \theta' \right) \cdot \mathbf{m}_l + \eta \\ &= \sum_{l \in \tilde{\mathcal{E}}} s_l \cdot \mathbf{m}_l + \eta \end{aligned} \quad (16)$$

where the notation $s_l = -b_l \mathbf{m}_l^T \theta'$ is introduced for simplicity.

Identifying line outages using (16) amounts to unveiling the subset $\tilde{\mathcal{E}}$ given the internal phase difference vector $\tilde{\theta}_I = \theta'_I - \theta_I$ and the pre-event network topology \mathbf{B} . Existing approaches extract $\tilde{\theta}_I$, exhaustively check the least-squares error for each possible topology $\tilde{\mathcal{E}}$ and select the one that gives the minimum. Such an approach incurs combinatorial complexity since the number of possible topologies grows combinatorially with the number of line outages. These methods are thus limited to identifying single or at most double line outages [15]–[17].

Recognizing that the number of line outages is a small fraction of the total number of lines, line-outage identification can be formulated as sparse vector estimation [19], i.e.,

$$\mathbf{B}\tilde{\theta} = \mathbf{M}\mathbf{s} + \eta, \quad (17)$$

where $s[l] = s_l = -b_l \mathbf{m}_l^T \theta'$ for $l \in \tilde{\mathcal{E}}$ and zero otherwise. The sparse representation in (17) relates the vector $\tilde{\theta}$ to the sparse vector \mathbf{s} , whose support represents the subset of lines in outage, and as such, recovering it translates to identifying line outages. This sparse representation allows us to rely on the machinery of sparse recovery which can often solve such problems using polynomial time algorithms [24], [28].

Partitioning $\tilde{\theta}$ and \mathbf{B} as in (11), it then follows from (17) that

$$\mathbf{B}_I \tilde{\theta}_I + \mathbf{B}_E \tilde{\theta}_E = \mathbf{M}\mathbf{s} + \eta. \quad (18)$$

Note that in addition to the unknown sparse vector \mathbf{s} , the vector $\tilde{\theta}_E$ is also not available since external nodes are assumed unobservable. Although it is not of immediate interest it must be accounted for in the analysis. The next section addresses the problem of estimating the sparse vector \mathbf{s} based on the model in (18) where it introduces two approaches for treating

the unobserved phasor angle data $\tilde{\theta}_E$.

IV. SPARSE RECONSTRUCTION

Given the model in (18), where only a subset $\tilde{\theta}_I$ of voltage phasor angle data is available, we next discuss two methods for recovering the sparse vector \mathbf{s} : the least-squares approach and the method taken in the literature (e.g. [19]) which will be referred to here as the inverse approach.

A. Least-squares approach

In this approach, we propose to treat the unobservable phasor angle data $\tilde{\theta}_E$ as a nuisance parameter and jointly estimate it together with the desired sparse vector \mathbf{s} using a least-squares formulation. Specifically, solving the following optimization for both $\tilde{\theta}_E$ and the sparse vector \mathbf{s} :

$$\min_{\mathbf{s}, \tilde{\theta}_E} \|\mathbf{B}_I \tilde{\theta}_I + \mathbf{B}_E \tilde{\theta}_E - \mathbf{M}\mathbf{s}\|_2^2, \quad \text{s.t. } \|\mathbf{s}\|_0 \leq \kappa \quad (19)$$

where κ represents the sparsity level, i.e., the number of nonzero entries in \mathbf{s} . In the general case, optimizing (19) with respect to $\tilde{\theta}_E$ will result in more than one solution. Any two solutions will differ by a vector in the null-space of \mathbf{B}_E . When \mathbf{B}_E has full-column rank, a unique solution exists. Otherwise, a minimum norm solution of the least-squares problem is unique and is given by

$$\hat{\tilde{\theta}}_E = -\mathbf{B}_E^\dagger (\mathbf{B}_I \tilde{\theta}_I - \mathbf{M}\mathbf{s}), \quad (20)$$

where \mathbf{B}_E^\dagger is the Moore-Penrose pseudo-inverse of \mathbf{B}_E . Specifically, writing \mathbf{B}_E in its SVD form

$$\mathbf{B}_E = \mathbf{U} \begin{bmatrix} \boldsymbol{\Sigma} \\ \mathbf{0} \end{bmatrix} \mathbf{V}^T, \quad (21)$$

where \mathbf{U} and \mathbf{V} are unitary matrices and $\boldsymbol{\Sigma}$ is a diagonal matrix with nonnegative entries, the Moore-Penrose pseudo inverse of \mathbf{B}_E is given by

$$\mathbf{B}_E^\dagger = \mathbf{V} [\boldsymbol{\Sigma}^\dagger \quad \mathbf{0}] \mathbf{U}^T. \quad (22)$$

Substituting (20) into (19), the optimization reduces to

$$\min_{\mathbf{s}} \|\mathbf{P}_{\mathbf{B}_E^\perp}^\perp (\mathbf{B}_I \tilde{\theta}_I - \mathbf{M}\mathbf{s})\|_2^2, \quad \|\mathbf{s}\|_0 \leq \kappa \quad (23)$$

where

$$\mathbf{P}_{\mathbf{B}_E^\perp}^\perp = \mathbf{I} - \mathbf{B}_E \mathbf{B}_E^\dagger. \quad (24)$$

The operator $\mathbf{P}_{\mathbf{B}_E^\perp}^\perp$ is an orthogonal projection onto the left null space of \mathbf{B}_E , namely $\mathcal{N}(\mathbf{B}_E^T)$, which is the subspace orthogonal to the range space of \mathbf{B}_E . Denoting by \mathbf{Q}_I the orthonormal eigenvectors in \mathbf{U} that correspond to unit eigenvalues in the eigen decomposition of $\mathbf{P}_{\mathbf{B}_E^\perp}^\perp$, it follows that

$$\mathbf{P}_{\mathbf{B}_E^\perp}^\perp = \mathbf{Q}_I \mathbf{Q}_I^T. \quad (25)$$

Using the eigen-decomposition of $\mathbf{P}_{\mathbf{B}_E^\perp}^\perp$ (25) in (23) and noting that \mathbf{Q}_I consists of orthonormal columns, the optimization reduces to

$$\min_{\mathbf{s}} \|\mathbf{Q}_I^T (\mathbf{B}_I \tilde{\theta}_I - \mathbf{M}\mathbf{s})\|_2^2, \quad \|\mathbf{s}\|_0 \leq \kappa. \quad (26)$$

Denoting $\mathbf{y}_Q = \mathbf{Q}_I^T \mathbf{B}_I \tilde{\theta}_I$ and $\mathbf{A}_Q = \mathbf{Q}_I^T \mathbf{M}$, the optimization is further reduced to

$$\min_{\mathbf{s}} \|\mathbf{y}_Q - \mathbf{A}_Q \mathbf{s}\|_2^2, \quad \|\mathbf{s}\|_0 \leq \kappa. \quad (27)$$

Note that the formulation of (27) only requires an orthonormal basis \mathbf{Q}_I for the null space $\mathcal{N}(\mathbf{B}_E^T)$. Knowing \mathbf{y}_Q and \mathbf{A}_Q , efficient recovery of \mathbf{s} can then be obtained using approaches for reconstructing sparse coefficient vectors in a linear regression model [24], [28].

An alternative way to obtain the optimization in (27) is by applying the projection operator $\mathbf{P}_{\mathbf{B}_E^\perp}^\perp$ directly on (18). Noting that $\mathbf{B}_E \tilde{\theta}_E$ lies in $\mathcal{R}(\mathbf{B}_E)$, it is omitted when projected onto $\mathcal{N}(\mathbf{B}_E^T)$, and the model is reduced to

$$\mathbf{P}_{\mathbf{B}_E^\perp}^\perp \mathbf{B}_I \tilde{\theta}_I = \mathbf{P}_{\mathbf{B}_E^\perp}^\perp \mathbf{M} \mathbf{s} + \mathbf{P}_{\mathbf{B}_E^\perp}^\perp \boldsymbol{\eta}. \quad (28)$$

Using (25) and the fact that the matrix \mathbf{Q}_I is of full-column rank, the model in (28) is equivalent to

$$\mathbf{y}_Q = \mathbf{A}_Q \mathbf{s} + \mathbf{Q}_I^T \boldsymbol{\eta}, \quad (29)$$

where the random vector $\mathbf{Q}_I^T \boldsymbol{\eta}$ has zero mean and covariance $\sigma_\eta^2 \mathbf{I}$. Finally, solving for the sparse \mathbf{s} that minimizes the least squares error in (29) yields (27).

B. The inverse approach

Let us now introduce what we refer to as the inverse approach, which is taken in [15]–[17], [19] to cope with the fact that only a subset of voltage phasor angle data is observable. In this approach, it is assumed that the matrix \mathbf{B} is invertible, and the model in (17) is multiplied from the left with \mathbf{B}^{-1} to obtain

$$\begin{bmatrix} \tilde{\theta}_I \\ \tilde{\theta}_E \end{bmatrix} = \mathbf{B}^{-1} \mathbf{M} \mathbf{s} + \mathbf{B}^{-1} \boldsymbol{\eta}. \quad (30)$$

Then, the subset of equations corresponding to the observed bus voltage phases is extracted as

$$\tilde{\theta}_I = [\mathbf{B}^{-1}]_I \mathbf{M} \mathbf{s} + [\mathbf{B}^{-1}]_I \boldsymbol{\eta}. \quad (31)$$

Next, the compact singular value decomposition (SVD) of $[\mathbf{B}^{-1}]_I = \mathbf{U}_I \boldsymbol{\Sigma}_I \mathbf{V}_I^T$ is introduced in (31) to account for the colored perturbation introduced by the inverse, i.e.,

$$\tilde{\theta}_I = \mathbf{U}_I \boldsymbol{\Sigma}_I \mathbf{V}_I^T \mathbf{M} \mathbf{s} + \mathbf{U}_I \boldsymbol{\Sigma}_I \mathbf{V}_I^T \boldsymbol{\eta}. \quad (32)$$

Finally, multiplying (32) from the left by $\boldsymbol{\Sigma}_I^{-1} \mathbf{U}_I^T$ and defining $\mathbf{y}_V = \boldsymbol{\Sigma}_I^{-1} \mathbf{U}_I^T \tilde{\theta}_I$ and $\mathbf{A}_V = \mathbf{V}_I^T \mathbf{M}$, the following sparse linear regression model is obtained in [19]:

$$\mathbf{y}_V = \mathbf{A}_V \mathbf{s} + \mathbf{V}_I^T \boldsymbol{\eta}. \quad (33)$$

The noise term $\mathbf{V}_I^T \boldsymbol{\eta}$ is a zero-mean random vector with covariance $\sigma_\eta^2 \mathbf{I}$. The vector \mathbf{s} is then selected as the solution to

$$\min_{\mathbf{s}} \|\mathbf{y}_V - \mathbf{A}_V \mathbf{s}\|_2^2, \quad \|\mathbf{s}\|_0 \leq \kappa. \quad (34)$$

C. Discussion

Let us consider the case in which \mathbf{B} is invertible and compare the two approaches. This setting corresponds to the power network being fully connected [29].

Theorem: I. The optimization of (27) based on the least-squares approach is equivalent to the optimization of (34) based on the inverse approach, i.e., the same estimate is obtained with both approaches for the sparse vector \mathbf{s} , whose support represents the subset of lines in outage.

II. The sparse linear models introduced in (29) and in (33) are linearly related through the unitary transformation $\mathbf{T} = \mathbf{V}_I^T \mathbf{Q}_I$. Specifically, multiplying (29) by the unitary transformation \mathbf{T} yields (33):

$$\mathbf{y}_Q = \mathbf{A}_Q \mathbf{s} + \mathbf{Q}_I^T \boldsymbol{\eta} \xrightarrow{\mathbf{T}} \mathbf{y}_V = \mathbf{A}_V \mathbf{s} + \mathbf{V}_I^T \boldsymbol{\eta}. \quad (35)$$

The two models are identical iff the orthonormal bases satisfy $\mathbf{Q}_I = \mathbf{V}_I$.

Part I of the theorem is a direct consequence of part II since the least-squares objective is invariant under unitary transformations. To prove the second part of the theorem, we introduce the following lemma.

Lemma: Assuming \mathbf{B} is invertible, the square matrix $\mathbf{Q}_I^T \mathbf{B}_I$ is invertible and its inverse is given by

$$(\mathbf{Q}_I^T \mathbf{B}_I)^{-1} = [\mathbf{B}^{-1}]_I \mathbf{Q}_I. \quad (36)$$

Partition the matrices \mathbf{B} and \mathbf{B}^{-1} as in the following identity corresponding to internal and external measurements

$$\mathbf{B}^{-1} \mathbf{B} = \begin{bmatrix} [\mathbf{B}^{-1}]_I \\ [\mathbf{B}^{-1}]_E \end{bmatrix} \cdot [\mathbf{B}_I \quad \mathbf{B}_E] = \mathbf{I}. \quad (37)$$

It then follows that

$$[\mathbf{B}^{-1}]_I \mathbf{B}_I = \mathbf{I} \quad (38a)$$

$$[\mathbf{B}^{-1}]_I \mathbf{B}_E = \mathbf{0}. \quad (38b)$$

Specifically, $[\mathbf{B}^{-1}]_I$ is a left inverse of \mathbf{B}_I , and the rows of $[\mathbf{B}^{-1}]_I$ span the left null-space of \mathbf{B}_E (i.e., $\mathcal{N}(\mathbf{B}_E^T)$), which is orthogonal to the range space of \mathbf{B}_E (i.e., $\mathcal{R}(\mathbf{B}_E)$). We then conclude from (38a) that the matrix $\mathbf{P}_{\mathbf{B}_I} = \mathbf{B}_I [\mathbf{B}^{-1}]_I$ is an oblique projection onto the column space of \mathbf{B}_I whose null-space is \mathbf{B}_E . Similarly, its transpose $\mathbf{P}_{\mathbf{B}_I}^T = \mathbf{P}_{[\mathbf{B}^{-1}]_I^T} = [\mathbf{B}^{-1}]_I^T \mathbf{B}_I^T$ is an oblique projection onto the row space of $[\mathbf{B}^{-1}]_I$.

Multiplying $\mathbf{Q}_I^T \mathbf{B}_I$ by $[\mathbf{B}^{-1}]_I \mathbf{Q}_I$ from the right, we obtain

$$\begin{aligned} (\mathbf{Q}_I^T \mathbf{B}_I) ([\mathbf{B}^{-1}]_I \mathbf{Q}_I) &= \mathbf{Q}_I^T (\mathbf{B}_I [\mathbf{B}^{-1}]_I) \mathbf{Q}_I \\ &= \left(\mathbf{P}_{[\mathbf{B}^{-1}]_I^T} \mathbf{Q}_I \right)^T \mathbf{Q}_I \\ &= \mathbf{Q}_I^T \mathbf{Q}_I \end{aligned} \quad (39)$$

where the last equality follows from the fact that \mathbf{Q}_I is a basis for the left null space of \mathbf{B}_E , which is spanned by the rows of $[\mathbf{B}^{-1}]_I$. Recalling that \mathbf{Q}_I is an orthonormal basis, it follows from (39) that the matrix $[\mathbf{B}^{-1}]_I \mathbf{Q}_I$ is a right inverse of $\mathbf{Q}_I^T \mathbf{B}_I$.

To show that $[\mathbf{B}^{-1}]_I \mathbf{Q}_I$ is also a left inverse of $\mathbf{Q}_I^T \mathbf{B}_I$ and thus $(\mathbf{Q}_I^T \mathbf{B}_I)^{-1} = [\mathbf{B}^{-1}]_I \mathbf{Q}_I$, we use the fact that $\mathbf{P}_{\mathbf{B}_E^\perp}^\perp = \mathbf{Q}_I \mathbf{Q}_I^T$ is an orthogonal projection onto the left null space of

\mathbf{B}_E and the identity in (38a), i.e.,

$$\begin{aligned} ([\mathbf{B}^{-1}]_I \mathbf{Q}_I) (\mathbf{Q}_I^T \mathbf{B}_I) &= [\mathbf{B}^{-1}]_I (\mathbf{Q}_I \mathbf{Q}_I^T) \mathbf{B}_I \quad (40) \\ &= \left([\mathbf{B}^{-1}]_I \mathbf{P}_{\mathbf{B}_E^\perp}^\perp \right) \mathbf{B}_I \\ &= \left(\mathbf{P}_{\mathbf{B}_E^\perp}^\perp [\mathbf{B}^{-1}]_I^T \right)^T \mathbf{B}_I \\ &= [\mathbf{B}^{-1}]_I \mathbf{B}_I = \mathbf{I}, \end{aligned}$$

completing the proof of the lemma. \square

To prove the theorem, we show that applying the invertible linear transformation $\Sigma_I^{-1} \mathbf{U}_I^T (\mathbf{Q}_I^T \mathbf{B}_I)^{-1}$ on (29) yields the model in (33). Specifically,

$$\begin{aligned} \Sigma_I^{-1} \mathbf{U}_I^T \tilde{\theta}_I &= \\ \Sigma_I^{-1} \mathbf{U}_I^T (\mathbf{Q}_I^T \mathbf{B}_I)^{-1} \mathbf{Q}_I^T \mathbf{M} \mathbf{s} + \Sigma_I^{-1} \mathbf{U}_I^T (\mathbf{Q}_I^T \mathbf{B}_I)^{-1} \mathbf{Q}_I^T \eta &\stackrel{(1)}{=} \\ \Sigma_I^{-1} \mathbf{U}_I^T ([\mathbf{B}^{-1}]_I \mathbf{Q}_I \mathbf{Q}_I^T) \mathbf{M} \mathbf{s} + \Sigma_I^{-1} \mathbf{U}_I^T ([\mathbf{B}^{-1}]_I \mathbf{Q}_I \mathbf{Q}_I^T) \eta &\stackrel{(2)}{=} \\ \Sigma_I^{-1} \mathbf{U}_I^T [\mathbf{B}^{-1}]_I \mathbf{M} \mathbf{s} + \Sigma_I^{-1} \mathbf{U}_I^T [\mathbf{B}^{-1}]_I \eta &\stackrel{(3)}{=} \\ \Sigma_I^{-1} \mathbf{U}_I^T \mathbf{U}_I \Sigma_I \mathbf{V}_I^T \mathbf{M} \mathbf{s} + \Sigma_I^{-1} \mathbf{U}_I^T \mathbf{U}_I \Sigma_I \mathbf{V}_I^T \eta &\stackrel{(4)}{=} \\ \mathbf{V}_I^T \mathbf{M} \mathbf{s} + \mathbf{V}_I^T \eta, &\quad (41) \end{aligned}$$

where (36) is used in (1), (2) follows by noting that $\mathbf{P}_{\mathbf{B}_E^\perp}^\perp = \mathbf{Q}_I \mathbf{Q}_I^T$ is an orthogonal projection onto $\mathcal{N}(\mathbf{B}_E^T)$, the SVD decomposition of $[\mathbf{B}^{-1}]_I$ is used in (3), and (4) follows since \mathbf{U}_I is unitary.

Using (36) and the SVD decomposition of $[\mathbf{B}^{-1}]_I$, it follows that

$$\begin{aligned} \Sigma_I^{-1} \mathbf{U}_I^T (\mathbf{Q}_I^T \mathbf{B}_I)^{-1} &= \\ \Sigma_I^{-1} \mathbf{U}_I^T [\mathbf{B}^{-1}]_I \mathbf{Q}_I &= \\ \Sigma_I^{-1} \mathbf{U}_I^T \mathbf{U}_I \Sigma_I \mathbf{V}_I^T \mathbf{Q}_I &= \\ \mathbf{V}_I^T \mathbf{Q}_I &\triangleq \mathbf{T}. \quad (42) \end{aligned}$$

Noting that both $\mathbf{P}_{\mathbf{B}_E^\perp}^\perp = \mathbf{Q}_I \mathbf{Q}_I^T$ and $\mathbf{P}_{[\mathbf{B}^{-1}]_I^T}^\perp = \mathbf{V}_I \mathbf{V}_I^T$ are orthogonal projections onto the same subspace, $\mathcal{N}(\mathbf{B}_E^T)$, and that the orthogonal projection is unique, we obtain

$$\mathbf{Q}_I \mathbf{Q}_I^T = \mathbf{P}_{\mathbf{B}_E^\perp}^\perp = \mathbf{P}_{[\mathbf{B}^{-1}]_I^T}^\perp = \mathbf{V}_I \mathbf{V}_I^T, \quad (43)$$

from which it follows that \mathbf{T} is unitary.

Since the models are linearly related through the unitary transformation $\mathbf{T} = \mathbf{V}_I^T \mathbf{Q}_I$, i.e., applying \mathbf{T} on (29) results in (33), the models are identical when

$$\mathbf{V}_I^T \mathbf{Q}_I = \mathbf{I}. \quad (44)$$

Rearranging and multiplying both sides by \mathbf{V}_I from the left, we obtain

$$\begin{aligned} 0 &= \mathbf{V}_I (\mathbf{V}_I^T \mathbf{Q}_I - \mathbf{I}) \\ &= (\mathbf{V}_I \mathbf{V}_I^T) \mathbf{Q}_I - \mathbf{V}_I \\ &\stackrel{(1)}{=} (\mathbf{Q}_I \mathbf{Q}_I^T) \mathbf{Q}_I - \mathbf{V}_I \\ &= \mathbf{Q}_I (\mathbf{Q}_I^T \mathbf{Q}_I) - \mathbf{V}_I \\ &\stackrel{(2)}{=} \mathbf{Q}_I - \mathbf{V}_I \quad (45) \end{aligned}$$

where (1) follows from (43), and (2) follows since the columns of \mathbf{Q}_I are orthonormal. Since \mathbf{V}_I has full-column

rank, it follows from (45) that $\mathbf{V}_I^T \mathbf{Q}_I = \mathbf{I}$ iff $\mathbf{Q}_I = \mathbf{V}_I$, i.e., the two models are identical if and only if $\mathbf{Q}_I = \mathbf{V}_I$. \square

We proved that under both approaches we will obtain the same estimate for the sparse vector \mathbf{s} , whose support represents the subset of lines in outage. The computational complexity, though, is not the same for both methods. In [19], the matrix \mathbf{B} is first inverted, its N_I rows are then extracted and $[\mathbf{B}^{-1}]_I$ is finally decomposed into its compact SVD form. In contrast, the approach we propose requires only finding an orthonormal basis for the left null space of \mathbf{B}_E . The latter can be obtained using a QR decomposition. Specifically, decomposing \mathbf{B}_E into the product of an orthonormal matrix \mathbf{Q} with an upper triangular matrix \mathbf{R} , the first r columns of \mathbf{Q} where r is the rank of \mathbf{B}_E form an orthonormal set of basis vectors for $\mathcal{R}(\mathbf{B}_E)$ and the remaining columns of \mathbf{Q} form an orthonormal basis for $\mathcal{N}(\mathbf{B}_E^T)$. Furthermore, since \mathbf{B} is a sparse matrix, it follows that the matrix \mathbf{B}_E which is a sub matrix of \mathbf{B} is also sparse. This fact can be exploited in efficiently factorizing the matrix \mathbf{B}_E into its QR form. In the inverse approach, the sparsity of \mathbf{B} can be exploited when inverting it. The inverted matrix though is, in general, dense, and thus sparsity cannot be exploited in its SVD decomposition. A step-by-step summary of each of the approaches is given in Table I.

TABLE I
THE LEAST-SQUARES APPROACH VS. THE INVERSE APPROACH

Least-squares	Inverse
$\mathbf{B} = -\sum_{l=1}^L b_l \mathbf{m}_l \mathbf{m}_l^T$	$\mathbf{B} = -\sum_{l=1}^L b_l \mathbf{m}_l \mathbf{m}_l^T$
	$\mathbf{B}^{-1} = \text{inv}(\mathbf{B})$
$\mathbf{B} = [\mathbf{B}_I \mathbf{B}_E] - \text{sparse}$	$\mathbf{B}^{-1} = \begin{bmatrix} [\mathbf{B}^{-1}]_I \\ [\mathbf{B}^{-1}]_E \end{bmatrix} - \text{dense}$
\mathbf{Q}_I - Orthonormal basis for $\mathcal{N}(\mathbf{B}_E^T)$	$[\mathbf{B}^{-1}]_I = \mathbf{U}_I \Sigma_I \mathbf{V}_I^T$
$\mathbf{A}_Q = \mathbf{Q}_I^T \mathbf{M}$	$\mathbf{A}_V = \mathbf{V}_I^T \mathbf{M}$
$\mathbf{y}_Q = \mathbf{Q}_I^T \mathbf{B}_I \tilde{\theta}_I$	$\mathbf{y}_V = \Sigma_I^{-1} \mathbf{U}_I^T \tilde{\theta}_I$
$\min_{\mathbf{s}, \ \mathbf{s}\ _0 \leq \kappa} \ \mathbf{y}_Q - \mathbf{A}_Q \mathbf{s}\ _2^2$	$\min_{\mathbf{s}, \ \mathbf{s}\ _0 \leq \kappa} \ \mathbf{y}_V - \mathbf{A}_V \mathbf{s}\ _2^2$

V. UNCERTAINTY IN GRID PARAMETERS

Section III considers the problem of line outages identification where it assumes that the grid parameters, namely the susceptances $\{b_l\}$ of all branches are known. The least-squares formulation we proposed allows us to generalize the problem to the case in which the exact value of these parameters is not known, but rather prior information is available in the form:

$$\mathbf{b} = \mathbf{b}_0 + \xi. \quad (46)$$

Here, the vector \mathbf{b}_0 is deterministic whose values are known and may be based on estimates from past observations, and ξ is an error vector with zero mean and known positive-definite covariance matrix Λ_ξ . The prior information on \mathbf{b} can be incorporated into the least-squares optimization as a regularization term in the following form:

$$\begin{aligned} \min_{\mathbf{s}, \tilde{\theta}_E, \mathbf{b}} \frac{1}{\sigma_\eta^2} \|\mathbf{B}_I \tilde{\theta}_I + \mathbf{B}_E \tilde{\theta}_E - \mathbf{M} \mathbf{s}\|^2 + \|\mathbf{b} - \mathbf{b}_0\|_{\Lambda_\xi^{-1}}^2, \quad (47) \\ \text{s.t. } \|\mathbf{s}\|_0 \leq \kappa \end{aligned}$$

where $\|\mathbf{b} - \mathbf{b}_0\|_{\Lambda_\xi^{-1}}^2 = (\mathbf{b} - \mathbf{b}_0)^T \Lambda_\xi^{-1} (\mathbf{b} - \mathbf{b}_0)$ represents the weighted norm of $\mathbf{b} - \mathbf{b}_0$ with weighting matrix Λ_ξ^{-1} and

$$\mathbf{B} = [\mathbf{B}_I \ \mathbf{B}_E] = \sum_{l=1}^L -b_l \mathbf{m}_l \mathbf{m}_l^T. \quad (48)$$

The coupling between the unknowns \mathbf{b} and $\tilde{\theta}_E$ in (47) makes it difficult to find an exact solution for the optimization. However, noting that when fixing \mathbf{b} ($\tilde{\theta}_E$) the objective is a quadratic function of $\tilde{\theta}_E$ (\mathbf{b}), we propose an approximate iterative solution based on a cyclic coordinate descent approach. The latter is an optimization algorithm for finding a local minimum of a function. In its simplest case, one cyclically iterates through the directions, one at a time, minimizing the objective function with respect to each coordinate direction at a time. For the optimization of (47), we first assume that \mathbf{b} is fixed and jointly optimize over \mathbf{s} and $\tilde{\theta}_E$. Then, we assume that both \mathbf{s} and $\tilde{\theta}_E$ are fixed and optimize over \mathbf{b} . Since the regularization term is independent of \mathbf{s} and $\tilde{\theta}_E$, the optimization of (47) with respect to these parameters is equivalent to the following optimization:

$$\min_{\mathbf{s}, \tilde{\theta}_E} \frac{1}{\sigma_\eta^2} \|\mathbf{B}_I \tilde{\theta}_I + \mathbf{B}_E \tilde{\theta}_E - \mathbf{M}\mathbf{s}\|^2 \quad (49)$$

which yields the same equations as obtained in (20) and (26), i.e.,

$$\tilde{\theta}_E = \mathbf{B}_E^\dagger (\mathbf{M}\mathbf{s} - \mathbf{B}_I \tilde{\theta}_I), \quad (50a)$$

$$\min_{\mathbf{s}} \|\mathbf{Q}_I^T (\mathbf{B}_I \tilde{\theta}_I - \mathbf{M}\mathbf{s})\|_2^2, \quad \|\mathbf{s}\|_0 \leq \kappa. \quad (50b)$$

Introducing the matrix \mathbf{C} whose l -th column is given by $\mathbf{c}_l = \mathbf{m}_l \mathbf{m}_l^T \tilde{\theta}$ and noting that $\mathbf{B}\tilde{\theta} = -\mathbf{C}\mathbf{b}$, the objective function in (47), viewed as a function of \mathbf{b} , can be reduced to

$$J(\mathbf{b}) = \frac{1}{\sigma_\eta^2} \|\mathbf{C}\mathbf{b} + \mathbf{M}\mathbf{s}\|^2 + \|\mathbf{b} - \mathbf{b}_0\|_{\Lambda_\xi^{-1}}^2. \quad (51)$$

This equation is equivalent to

$$J(\mathbf{b}) = \|\mathbf{C}\mathbf{b}_0 + \mathbf{M}\mathbf{s}\|_{(\sigma_\eta^2 \mathbf{I} + \mathbf{C}\Lambda_\xi \mathbf{C}^T)^{-1}}^2 + \quad (52)$$

$$\|(\Lambda_\xi^{-1} + (\mathbf{C}^T \mathbf{C})/\sigma_\eta^2)^{1/2} \mathbf{b} - (\Lambda_\xi^{-1} + (\mathbf{C}^T \mathbf{C})/\sigma_\eta^2)^{-1/2} (\Lambda_\xi^{-1} \mathbf{b}_0 - \mathbf{C}^T \mathbf{M}\mathbf{s}/\sigma_\eta^2)\|^2.$$

It is straightforward to see that $J(\mathbf{b})$ in (52) is minimized for

$$\begin{aligned} \mathbf{b}_{\text{opt}} &= (\Lambda_\xi^{-1} + (\mathbf{C}^T \mathbf{C})/\sigma_\eta^2)^{-1} (\Lambda_\xi^{-1} \mathbf{b}_0 - \mathbf{C}^T \mathbf{M}\mathbf{s}/\sigma_\eta^2) \\ &= \mathbf{b}_0 - (\sigma_\eta^2 \Lambda_\xi^{-1} + (\mathbf{C}^T \mathbf{C}))^{-1} \mathbf{C}^T (\mathbf{C}\mathbf{b}_0 + \mathbf{M}\mathbf{s}), \end{aligned} \quad (53)$$

where

$$J(\mathbf{b}_{\text{opt}}) = \|\mathbf{C}\mathbf{b}_0 + \mathbf{M}\mathbf{s}\|_{(\sigma_\eta^2 \mathbf{I} + \mathbf{C}\Lambda_\xi \mathbf{C}^T)^{-1}}^2. \quad (54)$$

Note that the matrix $(\Lambda_\xi^{-1} + (\mathbf{C}^T \mathbf{C})/\sigma_\eta^2)$ is obtained by summing a positive-definite matrix with a positive-semidefinite matrix so it is invertible. The same applies for the matrix $(\sigma_\eta^2 \mathbf{I} + \mathbf{C}\Lambda_\xi \mathbf{C}^T)$.

We now summarize the procedure proposed for recovering the sparse vector \mathbf{s} given internal node measurements and partial information on the grid parameters.

Algorithm 1: Recovering \mathbf{s} with Unknown Grid Parameters

- 1 Initialization: $\mathbf{b} = \mathbf{b}_0$, $\tilde{\theta}_I = \theta'_I - \theta_I$, \mathbf{M} - the bus-line incidence matrix formed by columns $\{\mathbf{m}_l\}_{l=1}^L$
- 2 Compute $\mathbf{B} = -\sum_{l=1}^L b_l \mathbf{m}_l \mathbf{m}_l^T$ and extract \mathbf{B}_I and \mathbf{B}_E from it as in (11)
- 3 Recovering \mathbf{s} : Compute an orthonormal set of basis vectors \mathbf{Q}_I for the null-space $\mathcal{N}(\mathbf{B}_E^T)$, optimize $\min_{\mathbf{s}} \|\mathbf{Q}_I^T (\mathbf{B}_I \tilde{\theta}_I - \mathbf{M}\mathbf{s})\|_2^2, \quad \|\mathbf{s}\|_0 \leq \kappa$
- 4 Estimate $\tilde{\theta}_E$ according to $\tilde{\theta}_E = \mathbf{B}_E^\dagger (\mathbf{M}\mathbf{s} - \mathbf{B}_I \tilde{\theta}_I)$
- 5 Estimate \mathbf{b} : Compute the matrix \mathbf{C} whose l -th column is given by $\mathbf{c}_l = \mathbf{m}_l \mathbf{m}_l^T \tilde{\theta}$,
 $\mathbf{b}_{\text{opt}} = \mathbf{b}_0 - (\sigma_\eta^2 \Lambda_\xi^{-1} + (\mathbf{C}^T \mathbf{C}))^{-1} \mathbf{C}^T (\mathbf{C}\mathbf{b}_0 + \mathbf{M}\mathbf{s})$
- 6 Continue iterating steps (2)-(5) until convergence.

Note that at each iteration of the algorithm the objective function is decreased.

Simulations show that incorporating the model of the grid parameters into the optimization framework results in an improved percentage of correctly identified line outages. Specifically, we observed relative improvement of up to 1.67% with more significant improvement in the percentage of correct identification for larger perturbations in the grid parameters.

VI. EXPLOITING ADDITIONAL INFORMATION FOR SPARSE RECONSTRUCTION

In the approaches discussed in Section IV for reconstructing the sparse coefficient vector \mathbf{s} , both the support of \mathbf{s} as well as the values of its nonzero entries were assumed unknowns. Note, however, that nonzero entries of \mathbf{s} that correspond to lines connecting internal nodes can actually be computed from the measurement data.

Recall the definition of \mathbf{s} :

$$\mathbf{s}[l] = \begin{cases} s_l = -b_l \mathbf{m}_l^T \theta', & l \in \tilde{\mathcal{E}} \\ 0, & \text{o.w.} \end{cases} \quad (55)$$

or, alternatively,

$$\mathbf{s}[l] = \begin{cases} -b_{mn} (\theta'_m - \theta'_n), & l \in \tilde{\mathcal{E}} \\ 0, & \text{o.w.} \end{cases}, \quad (56)$$

where in (56) we used the fact that when l corresponds to the line connecting nodes m and n , the column vector \mathbf{m}_l has all its entries zero except the m th and n th, which take on the values 1 and -1 , respectively. Thus, when both nodes m and n are internal, the post-event bus voltage phases θ'_m and θ'_n are available, and the value of $\mathbf{s}[l]$ is either $-b_{mn} (\theta'_m - \theta'_n)$ or zero, depending whether the line l is in outage or not. When either node is external the value of $\mathbf{s}[l]$ is still treated as unknown in the case where the line l is outaged.

Recall that \mathcal{E}_I represents the set of edges connecting internal nodes and \mathcal{E}_E represents the set of remaining edges. Let us rewrite the sparse linear model from (29) in a simpler notation where \mathbf{y}_Q , \mathbf{A}_Q and $\mathbf{Q}_I \eta$ are replaced by \mathbf{y} , \mathbf{A} and η respectively, i.e.,

$$\mathbf{y} = \mathbf{A}\mathbf{s} + \eta. \quad (57)$$

Then each of the entries of \mathbf{s} in (57) whose index belongs to \mathcal{E}_I can take only one of two known values, as indicated in (56). As we next show, this observation can be exploited to improve the performance of recovering \mathbf{s} .

We next consider OMP [22], which is a popular approach for reconstructing sparse vectors in linear regression models. OMP relies on greedy approximation schemes and is mostly popular for its computational simplicity and guaranteed performance. It iteratively updates the support of the sparse subset by finding the column most correlated to the signal residual. Following an introduction of this approach, we discuss its adaptation to the case in which partial information on the sparse vector is given.

A. Overview of OMP

OMP first initializes both the subset \mathcal{L}^k of indices corresponding to nonzero entries in $\hat{\mathbf{s}}^k$ and the approximation error vector $\mathbf{r}^k = \mathbf{y} - \mathbf{A}\hat{\mathbf{s}}^k$:

$$\mathcal{L}^0 = \emptyset \quad (58a)$$

$$\mathbf{r}^0 = \mathbf{y}. \quad (58b)$$

Then, the column \mathbf{a}_l of \mathbf{A} whose correlation with the approximation error vector is largest is selected, and its index is added to the subset \mathcal{L}^k , i.e.,

$$l^k = \arg \max_l \frac{|\mathbf{a}_l^H \mathbf{r}^{k-1}|^2}{\|\mathbf{a}_l\|^2} \quad (59a)$$

$$\mathcal{L}^k = \mathcal{L}^{k-1} \cup l^k. \quad (59b)$$

Finally, a least-squares fitting of \mathbf{y} using all columns in \mathcal{L}^k is obtained to form the estimate

$$\hat{\mathbf{s}}^k = \arg \min_{\mathbf{s}[l]|l \in \mathcal{L}^k} \|\mathbf{y} - \mathbf{A}\mathbf{s}\|^2, \quad (60)$$

whose error vector $\mathbf{r}^k = \mathbf{y} - \mathbf{A}\hat{\mathbf{s}}^k$ is orthogonal to the subset \mathcal{L}^k of columns of \mathbf{A} . Note that in (60) each entry $\mathbf{s}[l]$ such that $l \notin \mathcal{L}^k$ is set to zero.

B. OMP extension for partial information on the sparse vector

We now suggest an adaptation of OMP in which partial information on the sparse vector \mathbf{s} , in the form discussed earlier, is utilized. We first examine a simple scenario in which the support of the sparse vector is one. The least-squares optimization of the model in (57) for this scenario is given by

$$\begin{aligned} \min_{\mathbf{s}} \|\mathbf{y} - \mathbf{A}\mathbf{s}\|_2^2 &= \min_i \begin{cases} \|\mathbf{y} - s_i \mathbf{a}_i\|_2^2 & i \in \mathcal{E}_I \\ \min_{\mathbf{s}[i]} \|\mathbf{y} - \mathbf{s}[i] \mathbf{a}_i\|_2^2 & i \in \mathcal{E}_E \end{cases} \\ &= \min_i \begin{cases} \|\mathbf{y}\|_2^2 - 2s_i \mathbf{a}_i^T \mathbf{y} + s_i^2 \|\mathbf{a}_i\|_2^2 & i \in \mathcal{E}_I \\ \min_{\mathbf{s}[i]} (\|\mathbf{y}\|_2^2 - 2\mathbf{s}[i] \mathbf{a}_i^T \mathbf{y} + \mathbf{s}[i]^2 \|\mathbf{a}_i\|_2^2) & i \in \mathcal{E}_E \end{cases} \\ &= \min_i \begin{cases} \|\mathbf{y}\|_2^2 - 2s_i \mathbf{a}_i^T \mathbf{y} + s_i^2 \|\mathbf{a}_i\|_2^2 & i \in \mathcal{E}_I \\ \|\mathbf{y}\|_2^2 - (\mathbf{a}_i^T \mathbf{y})^2 / \|\mathbf{a}_i\|_2^2 & i \in \mathcal{E}_E \end{cases} \end{aligned} \quad (61)$$

which results in

$$\hat{i} = \arg \max_i \left\{ [2s_i \mathbf{a}_i^T \mathbf{y} - s_i^2 \|\mathbf{a}_i\|_2^2]_{i \in \mathcal{E}_I}, [(\mathbf{a}_i^T \mathbf{y})^2 / \|\mathbf{a}_i\|_2^2]_{i \in \mathcal{E}_E} \right\}$$

and

$$\hat{\mathbf{s}}[\hat{i}] = \begin{cases} s_{\hat{i}} & \hat{i} \in \mathcal{E}_I \\ \frac{\mathbf{a}_{\hat{i}}^T \mathbf{y}}{\|\mathbf{a}_{\hat{i}}\|^2} & \hat{i} \in \mathcal{E}_E \end{cases}.$$

We next generalize the algorithm to the case in which the support of the sparse vector is greater than one. Since each entry of the sought sparse vector that is associated with index in \mathcal{E}_I can take either zero or another known value, the index selection step in (59a) is changed to reflect this. Similarly, in the least-squares fit of (60) the minimum should be taken over only the unknown entries of the sparse vector. Following is a summary of the algorithm.

Algorithm 2: Incorporating Partial Information on the Sparse Vector Using OMP

- 1 Initialization: $\mathcal{L}^0 = \emptyset$, $\mathbf{r}^0 = \mathbf{y}$,
 $\mathbf{B} = [\mathbf{B}_I \mathbf{B}_E] = -\sum_{l=1}^L b_l \mathbf{m}_l \mathbf{m}_l^T$,
 \mathbf{Q}_I is an orthonormal basis for $\mathcal{N}(\mathbf{B}_E^T)$,
 $\mathbf{y} = \mathbf{Q}_I^T \mathbf{B}_I \hat{\theta}_I$, $\mathbf{A} = [\{\mathbf{a}_l\}_l] = \mathbf{Q}_I^T \mathbf{M}$
- 2 For $k=1:\kappa$
- 3 $l^k = \arg \max_l \left\{ [2s_l \mathbf{a}_l^T \mathbf{r}^{k-1} - s_l^2 \|\mathbf{a}_l\|_2^2]_{l \in \mathcal{E}_I}, [(\mathbf{a}_l^T \mathbf{r}^{k-1})^2 / \|\mathbf{a}_l\|_2^2]_{l \in \mathcal{E}_E} \right\}$, where $s_l = -b_l \mathbf{m}_l^T \hat{\theta}'$
 $\mathcal{L}^k = \mathcal{L}^{k-1} \cup l^k$
- 4 $\hat{\mathbf{s}}^k = \arg \min_{\mathbf{s}[l]} \|\mathbf{y} - \sum_{l \in \mathcal{L}^k \cap \mathcal{E}_I} s_l \mathbf{a}_l - \sum_{l \in \mathcal{L}^k \cap \mathcal{E}_E} \mathbf{s}[l] \mathbf{a}_l\|^2$.
- 5 $\mathbf{r}^k = \mathbf{y} - \mathbf{A}\hat{\mathbf{s}}^k$

Simulation results show the benefit of exploiting the additional information on the sparse vector in the case of internal line outages for reconstructing it. As we see in Section VII, better performance is achieved when this information is utilized as compared to the standard OMP algorithm which treats both values and support of the nonzero entries of the sparse vector as unknowns. Moreover, the improvement of correctly identifying line outages with our proposed approach is much more noticeable in the case where only subset of the outages are internal as compared to the case where the phasor angle data of all nodes is observed.

VII. SIMULATIONS

In this section the IEEE 118-bus benchmark system [30] is used to test the proposed algorithms for outages identification. The software toolbox MATPOWER [31] is used throughout to generate the phasor angle measurements and the pertinent power flows. AC power flows are generated for both pre- and post-event systems.

We consider first the 118-bus system with the complete information on phasor angles. All possible candidate topologies with a single line outage are tested, excluding outaged lines that result in islanding the post-event system. For each line-outage topology tested, 100 realizations of the perturbation noise η are generated. The variance of η is set either equal to zero, or equal to 1%, 2%, or 5% of the average pre-event power injection. The percentage of correctly identified line outages is listed in Table II.

TABLE II
118-BUS SYSTEM WITH ALL BUS PHASOR ANGLE
MEASUREMENTS - ALGORITHM 2

	OMP	OMP with partial information
0%	0.9497	0.9665 (+1.77%)
1%	0.9063	0.9215 (+1.68%)
2%	0.8808	0.9056 (+2.82%)
5%	0.8416	0.8679 (+3.13%)

Table II shows the benefit of incorporating partial information on the sparse vector for reconstructing it. Note that since the complete information on phasor angles is available, all nodes are treated as internals and all entries of the sparse vector that correspond to outaged lines can be computed from the post-event phasor angles. For all levels of perturbation noise tested, Algorithm 2 is shown to achieve better performance in identifying line outages than the standard OMP procedure, in which both the support of the sparse vector as well as the values of its nonzero entries are assumed unknowns.

We next consider line-outage identification relying on a subset of phasor angles from the 118-bus system partitioned as in [17]. Specifically, the internal system contains buses with indexes in the set $\mathcal{N}_I = \{1 - 45, 113, 114, 115, 117\}$, and the external ones with those in $\mathcal{N}_E = \{46 - 112, 116, 118\}$. All other settings are identical to those specified in the previous case, where the complete information on phasor angles is available. The results, which are listed in Table III, show again the improved percentage of correctly identifying line outages when partial information on the sparse vector is used for reconstruction. The improvement is much more noticeable than in the first case, where all nodes are internal, and is increased as the perturbation noise level increases. In the first case, the estimate is fairly well since it is based on the complete information on phasor angles. Adding the partial information on the sparse vector is thus less significant as compared to the second case in which the estimate is based on only a subset of phasor angles.

TABLE III
118-BUS SYSTEM WITH PHASOR ANGLE MEASUREMENTS
FROM \mathcal{N}_I - ALGORITHM 2

	OMP	OMP with partial information
0%	0.4637	0.4637
1%	0.3886	0.4394 (+13.07%)
2%	0.3727	0.4315 (+15.78%)
5%	0.3469	0.4165 (+20.06%)

We next consider identification of line outages when the grid parameters are not accurately known and follow the model in (46). The 118-bus system with the complete information on phasor angles is considered. Again, all possible candidate topologies with a single line outage are tested, excluding outaged lines that result in islanding the post-event system. For each line-outage topology tested, 1000 realizations of the perturbation noise η and the error vector ξ are generated.

The error vector ξ is assumed to have uncorrelated entries and its covariance matrix is expressed by $\Lambda_\xi = \sigma_\xi^2 \mathbf{I}$. We compare the results of Algorithm 1 after a few iterations with those obtained after one iteration where the values of the grid parameters are set to their nominal values. The percentage of correctly identified line outages is listed in Table IV. As expected, when the model of the grid parameters is taken into account in the estimation of the sparse vector, the percentage of correctly identified line outages is improved compared to the base case in which the nominal values of the grid parameters are used instead. The simulations also show that for the same level of perturbation noise η , an increased improvement in the percentage of correct identification occurs for larger perturbations in the grid parameters.

TABLE IV
118-BUS SYSTEM WITH ALL BUS PHASOR ANGLE
MEASUREMENTS - ALGORITHM 1

	Itr 1	itr 3	itr 5
$\sigma_\eta = 1.2, \sigma_\eta/\sigma_\xi = 0.3$	0.5824	0.5895	0.5921 (1.67%)
$\sigma_\eta = 1.2, \sigma_\eta/\sigma_\xi = 0.4$	0.8385	0.8455	0.8478 (1.11%)
$\sigma_\eta = 1.2, \sigma_\eta/\sigma_\xi = 0.6$	0.9314	0.9316	0.9316 (0.02%)
$\sigma_\eta = 0.6, \sigma_\eta/\sigma_\xi = 0.1$	0.1736	0.1743	0.1746 (0.58%)
$\sigma_\eta = 0.6, \sigma_\eta/\sigma_\xi = 0.3$	0.9389	0.9400	0.9402 (0.14%)
$\sigma_\eta = 0.6, \sigma_\eta/\sigma_\xi = 0.4$	0.9476	0.9483	0.9483 (0.07%)

Finally, we compare the computational complexity of the least-squares and the inverse approaches for different partitions of the 118-bus system into internal and external subsets. Both algorithms are run using Matlab 2015a software, on a typical computer with Intel Core i5-4300U CPU @ 1.90GHz. The running times of both algorithms are listed on table V and illustrated in Figure VII. As we can see, when the number of observable nodes in the internal system is relatively small compared to the total number of nodes in the network, the running times of both algorithms is comparable, with a small favor towards the inverse approach. However, as the size of the internal network increases, the running time of our proposed algorithm is becoming significantly better than that of the inverse approach with a factor of up to 13.9 in the best case we tested.

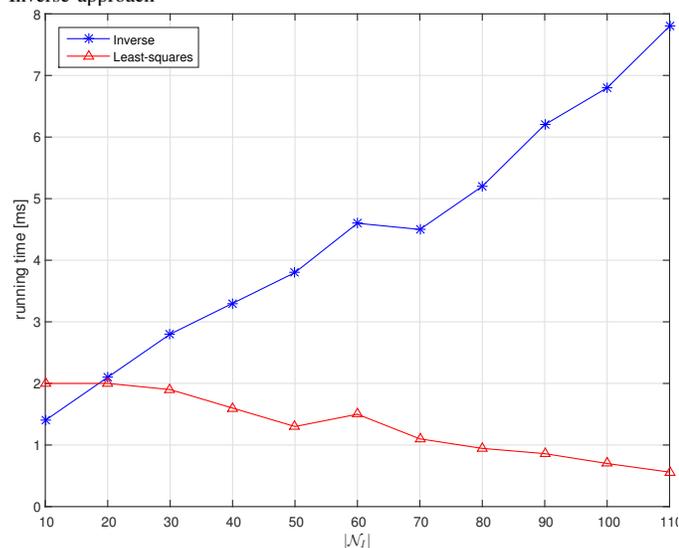
VIII. CONCLUSIONS

In this paper power line outages identification was reformulated as a least-squares optimization. As the size of the internal network increases, the running time of our proposed algorithm was shown to becoming significantly better than other competing approaches for outages identification. A natural extension of the least-squares method led to a generalization of the identification problem in which the grid parameters are unknown. Simulations show that when the model of the grid parameters is taken into account in the estimation of the sparse vector, the percentage of correctly identified line outages is improved, where more noticeable improvement occurs as the perturbations in the grid parameters

TABLE V
118-BUS SYSTEM - RUNNING TIMES OF BOTH APPROACHES

Size of Internal system	Inverse [ms]	Least-squares [ms]
10	1.4	2.0
20	2.1	2.0
30	2.8	1.9
40	3.3	1.6
50	3.8	1.3
60	4.6	1.5
70	4.5	1.1
80	5.2	0.945
90	6.2	0.86
100	6.8	0.7
110	7.8	0.56

Fig. 1. 118-BUS SYSTEM - Running times of Least-squares approach vs. Inverse approach



increase. We finally showed that extending our sparse recovery algorithm to incorporate partial information on the sparse vector, which is often assumed unknown, results in improved outages identification for both the case in which all bus phasor angle measurements are available and the case in which the identification is based only on a subset of phasor angle measurements.

REFERENCES

- [1] U.S. Department of Energy and Federal Energy Regulatory Commission, "Steps to establish a real-time transmission monitoring system for transmission owners and operators within the eastern and western interconnections," Tech. Rep., 2006.
- [2] A. G. Phadke, "Synchronized phasor measurements in power systems," *IEEE Comput. Appl. Power*, vol. 6, no. 2, pp. 10–15, Apr. 1993.
- [3] M. Donnelly, M. Ingram, and J. R. Carroll, "Eastern interconnection phasor project," in *Proc. 39th Hawaii Int'l Conf. System Sciences*, Jan 2006, pp. 336–342.
- [4] W. Jiang, V. Vittal, and G. T. Heydt, "A distributed state estimator utilizing synchronized phasor measurements," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 563–571, May 2007.
- [5] M. Zhou, V. A. Centeno, J. S. Thorp, and A. G. Phadke, "An alternative for including phasor measurements in state estimators," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1930–1937, November 2006.
- [6] J. S. Thorp, A. G. Phadke, and K. J. Karimi, "Real time voltage-phasor measurements for static state estimation," *IEEE Trans. Power Syst.*, vol. PAS-104, no. 11, pp. 3098–3106, November 1985.
- [7] A. P. S. Meliopoulos, G. J. Cokkinides, O. Wasynczuk, E. Coyle, C. Hoffmann, C. Nita-Rotaru, T. Downar, L. Tsoukalas, and R. Gao, "Pmu data characterization and application to stability monitoring," in *IEEE Power Engineering Society Power Syst. Conf. and Expo*, November 2006, pp. 151–158.
- [8] V. Vittal, V. S. Kolluri, K. Sun, S. Likhate and S. Mandal, "An online dynamic security assessment scheme using phasor measurements and an online dynamic security assessment scheme using phasor measurements and decision trees," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1935–1943, November 2007.
- [9] A. R. Khatib, R. F. Nuqui, M. R. Ingram, and A. G. Phadke, "Real-time estimation of security from voltage collapse using synchronized phasor real-time estimation of security from voltage collapse using synchronized phasor measurements," in *Proc. IEEE Power Eng. Soc. General Meeting*, June 2004, vol. 1, pp. 582–588.
- [10] P. Hirsch, G. Zhang and S. Lee, "Wide area frequency visualization using smart client technology," in *Proc. IEEE Power Eng. Soc. General Meeting*, June 2007, pp. 1–8.
- [11] R. E. Wilson, R. Klump and K. E. Martin, "Visualizing real-time security threats using hybrid scada / pmu measurement displays," in *Proc. 38th Hawaii Int'l Conf. Syst. Sciences*, Jan 2005.
- [12] S. Chun-Lien and J. Bo-Yuan, "Visualization of large-scale power system operations using phasor measurements," in *Proc. Int'l Conf. Power Syst. Tech.*, October 2006.
- [13] Dept. Energy, "U.s.-canada power system outage task force, final report on the august 14th blackout in the united states and canada," .
- [14] U.S.-Canada Power System Outage Task Force, "Final report on the implementation of the task force recommendation," .
- [15] J. E. Tate and T. J. Overbye, "Line outage detection using phasor angle measurements," *IEEE Transactions on Power Systems*, vol. 23, no. 4, pp. 1644–1652, November 2008.
- [16] J. E. Tate and T. J. Overbye, "Double line outage detection using phasor angle measurements," in *IEEE PES Gen. Meeting*, July 2009, pp. 1–5.
- [17] R. Emami and A. Abur, "Tracking changes in the external network model," September 2010.
- [18] M. He and J. Zhang, "A dependency graph approach for fault detection and localization towards secure smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 342–351, Jun 2011.
- [19] H. Zhu and G. B. Giannakis, "Sparse overcomplete representations for efficient identification of power line outages," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2215–2223, November 2012.
- [20] R. Tibshirani, "Regression shrinkage and selection via the lasso," *58*, vol. 58, pp. 267–288, 1996.
- [21] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comp.*, vol. 20, no. 33–61, Jan 1998.
- [22] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.
- [23] E. J. Candes and T. Tao, "Near optimal signal recovery from random projections; universal encoding strategies?," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec 2006.
- [24] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*, Cambridge University Press, 2012.
- [25] A. J. Wood and B. F. Wollenberg, *Power generation, Operation, and Control, 2nd ed.*, Wiley, 1996.
- [26] G. B. Giannakis, V. Kekatos, N. Gatsis, S. J. Kim, H. Zhu, and B. Wollenberg, "Monitoring and optimization for power grids: A signal processing perspectives," *Signal Processing Magazine*, vol. 30, no. 5, pp. 107–128, Sept. 2013 2013.
- [27] A. Schellenberg, W. Rosehart, and J. Aguado, "Cumulant-based probabilistic optimal power flow with gaussian and gamma distributions," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 773–781, May 2005.
- [28] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*, Cambridge University Press, April 2015.
- [29] *Introduction to Graph Theory*, Upper Saddle River, NJ: Prentice Hall, 2001.
- [30] Seattle [online].s Univ. Washington, "Power systems test case archive," .
- [31] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "Matpower: Steady-state operations, planning and analysis tools for power

systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb 2011.