IDENTIFICATION OF POWER LINE OUTAGES

Shay Maymon and Yonina C. Eldar

Department of Electrical Engineering
Technion, Haifa, Israel
{shay, yonina}@ee.technion.ac.il

ABSTRACT

This paper considers the problem of identifying power line outages throughout an electric interconnection based on changes in phasor angles observed at a limited number of buses. In existing approaches for solving the line outage identification problem the unobserved phasor angle data is ignored and identification is based on the observed phasor angles extracted from the data. We propose, instead, a least-squares approach for estimating the unobserved phasor angles, which is shown to yield a solution to the line outage identification problem that is equivalent to the solution obtained with existing approaches. This equivalence suggests an implementation of the solution to the line outage identification problem that is computationally more efficient than previous methods. A natural extension of the least-squares formulation leads to a generalization of the line outages identification problem in which the grid parameters are unknown.

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

1. INTRODUCTION

It is well known that major blackouts have occurred due in part to a lack of comprehensive situational awareness of the power grid. Timely identifying outages, or more generally abrupt changes in line parameters, is critical for wide-area monitoring in order to avoid unexpected events (such as a lightning strike or a tree fall) from spreading quickly, leading to a grid-wide blackout. Most existing approaches for line outages identification are challenged by the combinatorial complexity issues involved and are thus limited to identifying single or at most double line outages [1–3]. Zhu and Giannakis [1] 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, 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 phasor measurement units (PMUs).

This paper adopts the sparse linear model developed in [4] in which only a subset of voltage phasor angle data is observed. This 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 [1–3]. We propose instead to estimate the unobserved phasor angle data using a least-squares approach. It turns out that recovering the sparse vector with the unobserved data estimated using least-squares is equivalent to recovering the sparse vector based only on the extracted data. This equivalence suggests a more efficient implementation of the solution proposed in [4]. The least-squares formulation also allows us to easily 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 for estimating the grid parameters, the unobserved data, and the sparse vector.

The paper is organized as follows: Section 2 formulates the line outages identification problem. In Section 3 the least-squares approach for estimating the unobserved phasor angle data is developed, the deterministic approach is introduced, and the relation of these methods is discussed. Section 4 generalizes the line outages identification problem to the case of unknown grid parameters.

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 $||\mathbf{x}||_W^2 = \mathbf{x}^T W \mathbf{x}$. The Moonre-Penrose pseudo-inverse of $\mathbf{A}$ is denoted by $\mathbf{A}^\dagger$. An oblique projection onto the subspace $\mathcal{R}(A)$ is denoted by P_A, and an orthogonal projection onto $\mathcal{R}(\mathcal{A})$ is denoted by $P_A^\perp$.

2. LINE OUTAGE IDENTIFICATION

Consider a power transmission network $\mathcal{N}$ with N buses (a.k.a nodes) and L transmission lines (a.k.a branches) that

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).

are represented by the set $\mathcal{E}$. We partition the network buses into two subsets: the subset $\mathcal{N}_I$ of observable buses in the internal system, and the subset $\mathcal{N}_E$ of unobservable buses of the external system, where $\mathcal{N}_I \cap \mathcal{N}_E = \emptyset$. The vectors θ and $\mathbf{p}$ represent the voltage phasor angles of all buses in the network and the corresponding injected power variables. The linear DC power flow model [5, 6] provides a linear approximation of the actual AC system, in which the voltage phasor angle data θ is related to the injected power $\mathbf{p}$ through

$$\mathbf{p} = \mathbf{B}\theta. \tag{1}$$

The matrix $\mathbf{B}$ is referred to as the weighted Laplacian matrix and is uniquely determined by the line reactance parameters and the topology-bearing information. Specifically,

$$\mathbf{B} = \mathbf{M}\mathbf{D}\mathbf{M}^T = \sum_{l=1}^{L} \frac{1}{x_l}\mathbf{m}_l\mathbf{m}_l^T, \tag{2}$$

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. The diagonal matrix $\mathbf{D}$ has its lth diagonal entry equal to the inverse reactance $1/x_{mn}$ if l corresponds to the line (m, n). Note that the matrix $\mathbf{B}$ is rank deficient and has the vector $\mathbf{1}$ in its null space. If the power network is connected, the matrix $\mathbf{B}$ has rank N-1 [4]. With the reference bus convention, in which one generation bus is chosen as reference with its phasor angle set to zero, the $(N$-1$)\times(N$-1$)$ matrix B has full rank and can be formed by M as in (2) after removing the row corresponding to the reference bus in the incidence matrix. Given the pre-event network topology, and the pre- and post-event internal phasor angle vectors, namely θ_I and θ'_I, various approaches based on the DC model (1) are proposed in [1–4] for unveiling the subset $\tilde{\mathcal{E}} \subset \mathcal{E}$ of line outages.

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 [4], i.e.,

$$\mathbf{B}\tilde{\theta} \quad = \mathbf{M}\mathbf{s} + \eta, \tag{3}$$

where the sparse vector $\mathbf{s}$ captures information of line outages, and as such, recovering it translates to identifying line outages. The noise η is usually modeled as a zero-mean random vector with covariance $\sigma_\eta^2\mathbf{I}$. The sparse representation in (3) relates the vector $\tilde{\theta} = \theta^{'} - \theta$ representing the change in bus voltage phases due to an occurence of an event to the sparse vector $\mathbf{s}$, whose support represents the subset of lines in outage. This representation bypasses the combinatorial complexity incurred by methods based on exhaustive enumeration of all combinations to identifying single line outage or at most double line outages [1–3].

Partition $\tilde{\theta}$ and $\mathbf{B}$ as

$$\tilde{\theta} = \begin{bmatrix} \tilde{\theta}_I \\ \tilde{\theta}_E \end{bmatrix} \tag{4a}$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_I & \mathbf{B}_E \end{bmatrix} \tag{4b}$$

where the subscripts $_I$ and $_E$ are associated with the subsets $\mathcal{N}_I$ and $\mathcal{N}_E$, correspondingly. It then follows from (3) that

$$\mathbf{B}_I\tilde{\theta}_I + \mathbf{B}_E\tilde{\theta}_E = \mathbf{M}\mathbf{s} + \eta. \tag{5}$$

Since external nodes are assumed unobservable, the vector $\tilde{\theta}_E$ needs to be estimated as well for $\mathbf{s}$ to be recovered. Estimation of the model parameters is addressed in the next section.

3. SPARSE RECONSTRUCTION

Given the model in (5), 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 deterministic approach.

3.1. Least-squares approach

Using the least-squares criterion for estimating both the unobservable phasor angle data $\tilde{\theta}_E$ and the sparse vector $\mathbf{s}$, the following optimization is obtained:

$$\min_{\mathbf{s},\tilde{\theta}_E} ||\mathbf{B}_I\tilde{\theta}_I + \mathbf{B}_E\tilde{\theta}_E - \mathbf{M}\mathbf{s}||^2, \quad \text{s.t.} \quad ||\mathbf{s}||_0 \leq \kappa \tag{6}$$

where κ represents the sparsity level, i.e., the number of nonzero entries in $\mathbf{s}$. Optimizing with resect to $\tilde{\theta}_E$, we have

$$\tilde{\theta}_E = -\mathbf{B}_E^{\dagger}(\mathbf{B}_I\tilde{\theta}_I - \mathbf{M}\mathbf{s}), \tag{7}$$

where we assume that $\mathbf{B}_E$ has full column rank. Substituting (7) into (6), the optimization reduces to

$$\min_{\mathbf{s}} ||\mathbf{P}^{\perp}_{\mathbf{B}_E{}^{\perp}}(\mathbf{B}_I\tilde{\theta}_I - \mathbf{M}\mathbf{s})||^2, \quad ||\mathbf{s}||_0 \leq \kappa \tag{8}$$

where $\mathbf{P}^{\perp}_{\mathbf{B}_E{}^{\perp}} = (\mathbf{I} - \mathbf{B}_E(\mathbf{B}_E{}^T\mathbf{B}_E)^{-1}\mathbf{B}_E{}^T)$ is an orthogonal projection onto the left null space of $\mathbf{B}_E$.

Being an orthogonal projection, $\mathbf{P}^{\perp}_{B_E{}^{\perp}}$ is a symmetric matrix whose eigenvalues are either 0 or 1. In particular, vectors $\mathbf{v}_{\mathcal{R}(\mathbf{B}_E)} \in \mathcal{R}(\mathbf{B}_E)$ are projected to 0 and are associated with zero eigenvalues, whereas vectors $\mathbf{v}_{\mathcal{N}(B_E{}^T)} \in \mathcal{N}(\mathbf{B}_E{}^T)$ remain unaltered by this projection and are associated with unit eigenvalues. The eigen-decomposition of $P^{\perp}_{\mathbf{B}_E{}^{\perp}}$ can thus be represented by

$$\begin{aligned} \mathbf{P}^{\perp}_{\mathbf{B}_E{}^{\perp}} &= \begin{bmatrix} \mathbf{Q}_I & \mathbf{Q}_E \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_I^T \\ \mathbf{Q}_E^T \end{bmatrix} \\ &= \mathbf{Q}_I\mathbf{Q}_I^T \end{aligned} \tag{9}$$

where $\mathbf{Q}_I$ consists of orthonormal eigenvectors $\mathbf{v}_{\mathcal{N}(\mathbf{B}_E{}^T)}$ of $\mathbf{P}^{\perp}_{\mathbf{B}_E{}^{\perp}}$, which form a basis for $\mathcal{N}(\mathbf{B}_E{}^T)$. Using the eigen-decomposition of $\mathbf{P}^{\perp}_{\mathbf{B}_E{}^{\perp}}$ in (8) and the fact that the columns of $\mathbf{Q}_I$ are orthonormal, and 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 reduces to

$$\min_{\mathbf{s}} ||\mathbf{y}_Q - \mathbf{A}_Q\mathbf{s}||^2, \quad ||\mathbf{s}||_0 \leq \kappa. \tag{10}$$

Note that the formulation of (10) only requires an orthonormal basis $\mathbf{Q}_I$ to the null space $\mathcal{N}(\mathbf{B}_E{}^T)$. Efficient recovery of $\mathbf{s}$ can then be obtained using approaches for reconstructing sparse coefficient vectors in a linear regression model [7].

An alternative way to obtain the optimization in (8) is by applying the projection operator $\mathbf{P}^{\perp}_{\mathbf{B}_E{}^{\perp}}$ directly on (5). 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 then reduced to

$$\mathbf{P}^{\perp}_{\mathbf{B}_E{}^{\perp}}\mathbf{B}_I\tilde{\theta}_I = \mathbf{P}^{\perp}_{\mathbf{B}_E{}^{\perp}}\mathbf{M}\mathbf{s} + \mathbf{P}^{\perp}_{\mathbf{B}_E{}^{\perp}}\eta. \tag{11}$$

Using the eigen-decomposition of $P^{\perp}_{\mathbf{B}_E{}^{\perp}}$ and the fact that the matrix $\mathbf{Q}_I$ is of full column rank, the model in (11) is equivalent to

$$\mathbf{y}_Q = \mathbf{A}_Q\mathbf{s} + \mathbf{Q}_I^T\eta, \tag{12}$$

where the random vector $\mathbf{Q}_I^T\eta$ has zero mean and covariance $\sigma_\eta^2\mathbf{I}$. Finally, solving for the sparse $\mathbf{s}$ that minimizes the least squares error in (12) yields (10).

3.2. The deterministic approach

Let us now introduce what we refer to as the deterministic approach, which is taken in [1–4] to cope with the fact that only a subset of voltage phasor angle data is observable. In this approach, the model in (3) is first multiplied from the left with $\mathbf{B}^{-1}$, assuming the inverse exists, to obtain

$$\begin{bmatrix} \tilde{\theta}_I \\ \tilde{\theta}_E \end{bmatrix} = \mathbf{B}^{-1}\mathbf{M}\mathbf{s} + \mathbf{B}^{-1}\eta. \tag{13}$$

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

$$\tilde{\theta}_I \quad = \quad [\mathbf{B}^{-1}]_I\mathbf{M}\mathbf{s} + [\mathbf{B}^{-1}]_I\eta. \tag{14}$$

Next, the compact singular value decomposition (SVD) of $[\mathbf{B}^{-1}]_I = \mathbf{U}_I\mathbf{\Sigma}_I\mathbf{V}_I^T$ is introduced in (14) to account for the colored perturbation introduced by the inverse. Finally, defining $\mathbf{y_V} = \mathbf{\Sigma}_I^{-1}\mathbf{U}_I^T\tilde{\theta}_I$ and $\mathbf{A}_V = \mathbf{V}_I^T\mathbf{M}$, the following sparse linear regression model is obtained in [4]:

$$\mathbf{y}_V \quad = \quad \mathbf{A}_V\mathbf{s} + \mathbf{V}_I^T\eta, \tag{15}$$

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

$$\min_{\mathbf{s}} ||\mathbf{y}_V - \mathbf{A}_V\mathbf{s}||^2, \quad ||s||_0 \leq \kappa. \tag{16}$$

3.3. Discussion

We next discuss the connection between the least-squares approach and the deterministic approach.

Theorem: *The sparse linear models introduced in (12) and (15) are linearly related through a unitary transformation. The two models are identical iff the orthonormal bases* $\mathbf{Q}_I = \mathbf{V}_I$.

To prove the theorem, we introduce the following lemma.

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

$$(\mathbf{Q}_I^T\mathbf{B}_I)^{-1} = [\mathbf{B}^{-1}]_I\mathbf{Q}_I. \tag{17}$$

By partitioning the matrices $\mathbf{B}$ and $\mathbf{B}^{-1}$ in the following identity into two sub matrices, corresponding to internal and external measurements, i.e.,

$$\mathbf{B}^{-1}\mathbf{B} \quad = \quad \begin{bmatrix} [\mathbf{B}^{-1}]_I \\ [\mathbf{B}^{-1}]_E \end{bmatrix} \cdot \begin{bmatrix} \mathbf{B}_I & \mathbf{B}_E \end{bmatrix} = \mathbf{I} \tag{18}$$

it follows that

$$[\mathbf{B}^{-1}]_I\mathbf{B}_I = \mathbf{I} \tag{19a}$$

$$[\mathbf{B}^{-1}]_I\mathbf{B}_E = \mathbf{0}. \tag{19b}$$

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 (19a) that the matrix $P_{\mathbf{B}_I} = \mathbf{B}_I[\mathbf{B}^{-1}]_I$ is an oblique projection onto the column space of $\mathbf{B}_I$. 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

$$\begin{aligned}\left(\mathbf{Q}_I^T\mathbf{B}_I\right)\left([\mathbf{B}^{-1}]_I\mathbf{Q}_I\right) &= \mathbf{Q}_I^T\left(\mathbf{B}_I[\mathbf{B}^{-1}]_I\right)\mathbf{Q}_I \\ &= \left(\mathbf{P}_{[\mathbf{B}^{-1}]_I{}^T}\mathbf{Q}_I\right)^T\mathbf{Q}_I \\ &= \mathbf{Q}_I^T\mathbf{Q}_I \end{aligned} \tag{20}$$

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 (20) 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 $\mathbf{P}^{\perp}_{\mathbf{B}_E{}^{\perp}} = \mathbf{Q}_I\mathbf{Q}_I^T$ is an orthogonal projection onto the left null space of $\mathbf{B}_E$ and the identity in (19a), i.e.,

$$\begin{aligned}\left([\mathbf{B}^{-1}]_I\mathbf{Q}_I\right)\left(\mathbf{Q}_I^T\mathbf{B}_I\right) &= [\mathbf{B}^{-1}]_I\left(\mathbf{Q}_I\mathbf{Q}_I^T\right)\mathbf{B}_I \qquad (21)\\ &= \left([\mathbf{B}^{-1}]_I\mathbf{P}^{\perp}_{\mathbf{B}_E{}^{\perp}}\right)\mathbf{B}_I \\ &= \left(\mathbf{P}^{\perp}_{\mathbf{B}_E{}^{\perp}}[\mathbf{B}^{-1}]_I{}^T\right)^T\mathbf{B}_I \\ &= [\mathbf{B}^{-1}]_I\mathbf{B}_I = \mathbf{I}, \end{aligned}$$

completing the proof of the lemma. □

To prove the theorem, we show that applying the invertible linear transformation $\mathbf{T}_{Q\rightarrow V} = \mathbf{\Sigma}_I^{-1}\mathbf{U}_I^T(\mathbf{Q}_I^T\mathbf{B}_I)^{-1}$ on

(12) yields the model in (15). Specifically,

$$
\begin{aligned}
&\mathbf{\Sigma}_I^{-1}\mathbf{U}_I^T\tilde{\theta}_I = \\
&\mathbf{\Sigma}_I^{-1}\mathbf{U}_I^T(\mathbf{Q}_I^T\mathbf{B}_I)^{-1}\mathbf{Q}_I^T\mathbf{M}\mathbf{s} + \mathbf{\Sigma}_I^{-1}\mathbf{U}_I^T(\mathbf{Q}_I^T\mathbf{B}_I)^{-1}\mathbf{Q}_I^T\eta \overset{(1)}{=} \\
&\mathbf{\Sigma}_I^{-1}\mathbf{U}_I^T\left([\mathbf{B}^{-1}]_I\mathbf{Q}_I\mathbf{Q}_I^T\right)\mathbf{M}\mathbf{s} + \mathbf{\Sigma}_I^{-1}\mathbf{U}_I^T\left([\mathbf{B}^{-1}]_I\mathbf{Q}_I\mathbf{Q}_I^T\right)\eta \overset{(2)}{=} \\
&\mathbf{\Sigma}_I^{-1}\mathbf{U}_I^T[\mathbf{B}^{-1}]_I\mathbf{M}\mathbf{s} + \mathbf{\Sigma}_I^{-1}\mathbf{U}_I^T[\mathbf{B}^{-1}]_I\eta \overset{(3)}{=} \\
&\mathbf{\Sigma}_I^{-1}\mathbf{U}_I^T\mathbf{U}_I\mathbf{\Sigma}_I\mathbf{V}_I^T\mathbf{M}\mathbf{s} + \mathbf{\Sigma}_I^{-1}\mathbf{U}_I^T\mathbf{U}_I\mathbf{\Sigma}_I\mathbf{V}_I^T\eta \overset{(4)}{=} \\
&\mathbf{V}_I^T\mathbf{M}\mathbf{s} + \mathbf{V}_I^T\eta, \qquad (22)
\end{aligned}
$$

where (17) is used in (1), (2) follows by noting that $\mathbf{P}^{\perp}_{\mathbf{B}_E{}^{\perp}} = \mathbf{Q}_I\mathbf{Q}_I^T$ is a 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 (17) and the SVD decomposition of $[\mathbf{B}^{-1}]_I$, the linear transformation $\mathbf{T}_{Q\to V}$ reduces to

$$
\begin{aligned}
\mathbf{T}_{Q\to V} &= \mathbf{\Sigma}_I^{-1}\mathbf{U}_I^T(\mathbf{Q}_I^T\mathbf{B}_I)^{-1} \\
&= \mathbf{\Sigma}_I^{-1}\mathbf{U}_I^T[\mathbf{B}^{-1}]_I\mathbf{Q}_I \\
&= \mathbf{\Sigma}_I^{-1}\mathbf{U}_I^T\mathbf{U}_I\mathbf{\Sigma}_I\mathbf{V}_I^T\mathbf{Q}_I \\
&= \mathbf{V}_I^T\mathbf{Q}_I. \qquad (23)
\end{aligned}
$$

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

$$
\mathbf{Q}_I\mathbf{Q}_I^T = \mathbf{P}^{\perp}_{\mathbf{B}_E{}^{\perp}} = \mathbf{P}^{\perp}_{[\mathbf{B}^{-1}]_I{}^T} = \mathbf{V}_I\mathbf{V}_I^T, \qquad (24)
$$

from which it follows that $\mathbf{T}_{Q\to V}$ is unitary.

Since the models are linearly related through the unitary transformation $\mathbf{T}_{Q\to V} = \mathbf{V}_I^T\mathbf{Q}_I$, they are identical when

$$
\mathbf{V}_I^T\mathbf{Q}_I = \mathbf{I}. \qquad (25)
$$

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

$$
\begin{aligned}
0 &= \mathbf{V}_I\left(\mathbf{V}_I^T\mathbf{Q}_I - \mathbf{I}\right) \\
&= \left(\mathbf{V}_I\mathbf{V}_I^T\right)\mathbf{Q}_I - \mathbf{V}_I \\
&\overset{(1)}{=} \left(\mathbf{Q}_I\mathbf{Q}_I^T\right)\mathbf{Q}_I - \mathbf{V}_I \\
&= \mathbf{Q}_I\left(\mathbf{Q}_I^T\mathbf{Q}_I\right) - \mathbf{V}_I \\
&\overset{(2)}{=} \mathbf{Q}_I - \mathbf{V}_I \qquad (26)
\end{aligned}
$$

where (1) follows from (24), and (2) follows since the columns of $\mathbf{Q}_I$ are orthonormal. Since $\mathbf{V}_I$ has full-column rank, it follows from (26) that $\mathbf{V}_I^T\mathbf{Q}_I = I$ iff $\mathbf{Q}_I = \mathbf{V}_I$, i.e., the two models are identical if and only if $\mathbf{Q}_I = \mathbf{V}_I$. □

A direct consequence of the Theorem is that the optimzation based on the least-squares approach (10) is equivalent to the optimization formulated in [4] based on the deterministic approach (16). This equivalence shows that the deterministic approach is optimal in the least-squares sense. The computational complexity, though, is not the same for both methods. In [4], the matrix $\mathbf{B}$ is first inverted, its N_I rows are extracted and $[\mathbf{B}^{-1}]_I$ is finally decomposed into its SVD form, whereas the approach that we propose requires only to find an orthonormal basis for the left null space of $\mathbf{B}_E$, which can be obtained using a QR decomposition. Furthermore, the matrix $\mathbf{B}_E$ is sparse since it is a sub-matrix of the sparse matrix $\mathbf{B}$, whereas its inverse $\mathbf{B}^{-1}$ is, in general, dense.

4. UNCERTAINTY IN GRID PARAMETERS

Section 2 considers the problem of line outages identification where it assumes that the grid parameters, namely the susceptances $\{b_l\} = \{-1/x_l\}$ of all branches are known. We now 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:

$$
\mathbf{b} = \mathbf{b}_0 + \mathbf{v}. \qquad (27)
$$

The vector $\mathbf{b}_0$ is deterministic whose values are known and may be based on estimates from past observations and $\mathbf{v}$ is an error vector with zero mean and covariance matrix Λ. Note that by a proper choice of some elements in the covariance matrix Λ, we can also include cases in which no prior information is given for some of the parameters, or alternatively, when the exact value of some parameters is accurately known. The prior information on $\mathbf{b}$ can be incorporated into the optimization as a regularization term, i.e.,

$$
\min_{\mathbf{s},\tilde{\theta}_E,\mathbf{b}} \frac{1}{\sigma_\eta^2}||\mathbf{B}_I\tilde{\theta}_I + \mathbf{B}_E\tilde{\theta}_E - \mathbf{M}\mathbf{s}||^{\mathbf{2}} + ||\mathbf{b} - \mathbf{b_0}||^{\mathbf{2}}_{\mathbf{\Lambda^{-1}}}, \quad \text{s.t. } ||\mathbf{s}||_0 \le \kappa \qquad (28)
$$

where

$$
\mathbf{B} = [\mathbf{B}_I\ \mathbf{B}_E] = \sum_{l=1}^{L} -b_l\mathbf{m}_l\mathbf{m}_l^T. \qquad (29)
$$

Optimizing (28) with respect to $\mathbf{s}$ and $\tilde{\theta}_E$ is equivalent to optimizing the following objective function

$$
\begin{aligned}
J(\mathbf{s},\tilde{\theta}_E) &= \frac{1}{\sigma_\eta^2}||\mathbf{B}_I\tilde{\theta}_I + \mathbf{B}_E\tilde{\theta}_E - \mathbf{M}\mathbf{s}||^{\mathbf{2}} \qquad (30) \\
&= \frac{1}{\sigma_\eta^2}||\mathbf{P}^{\perp}_{B_E{}^{\perp}}(\mathbf{B}_I\tilde{\theta}_I - \mathbf{M}\mathbf{s})||^{\mathbf{2}} + \\
&\quad \frac{1}{\sigma_\eta^2}||\mathbf{B}_E\left(\tilde{\theta}_E - \mathbf{B}_E{}^{\dagger}(\mathbf{M}\mathbf{s} - \mathbf{B}_I\tilde{\theta}_I)\right)||^2,
\end{aligned}
$$

which yields the same equations as obtained in (7) and (8), i.e.,

$$
\tilde{\theta}_E = \mathbf{B}_E{}^{\dagger}(\mathbf{M}\mathbf{s} - \mathbf{B}_I\tilde{\theta}_I), \qquad (31a)
$$

$$\min_{\mathbf{s}} ||\mathbf{P}^{\perp}_{\mathbf{B}_{E^{\perp}}}(\mathbf{B}_I\tilde{\theta}_I - \mathbf{M}\mathbf{s})||^2, \quad ||\mathbf{s}||_0 \leq \kappa. \tag{31b}$$

Introducing the matrix $\mathbf{A}$ whose lth column is given by $\mathbf{a}_l = \mathbf{m}_l\mathbf{m}_l^T\tilde{\theta}$ and noting that $\mathbf{B}\tilde{\theta} = -\mathbf{A}\mathbf{b}$, the objective function in (28), viewed as a function of $\mathbf{b}$, can be reduced to

$$J(\mathbf{b}) = \frac{1}{\sigma_\eta^2}||\mathbf{A}\mathbf{b} + \mathbf{M}\mathbf{s}||^{\mathbf{2}} + ||\mathbf{b} - \mathbf{b_0}||^{\mathbf{2}}_{\mathbf{\Lambda^{-1}}}, \tag{32}$$

or alternatively to

$$\begin{aligned} J(\mathbf{b}) = {} & ||\mathbf{A}\mathbf{b}_0 + \mathbf{M}\mathbf{s}||^{\mathbf{2}}_{\left(\sigma_\eta^{\mathbf{2}}\mathbf{I}+\mathbf{A}\mathbf{\Lambda}\mathbf{A^T}\right)^{-\mathbf{1}}} + \\ & ||(\mathbf{\Lambda}^{-1} + (\mathbf{A}^T\mathbf{A})/\sigma_\eta^2)^{1/2}\mathbf{b} - \\ & (\mathbf{\Lambda}^{-1} + (\mathbf{A}^T\mathbf{A})/\sigma_\eta^2)^{-1/2}(\mathbf{\Lambda}^{-1}\mathbf{b}_0 - \mathbf{A}^T\mathbf{M}\mathbf{s}/\sigma_\eta^2)||^2. \end{aligned} \tag{33}$$

It is straightforward to see that $J(\mathbf{b})$ in (33) is minimized for

$$\begin{aligned} \mathbf{b}_{\text{opt}} &= (\Lambda^{-1} + (\mathbf{A}^T\mathbf{A})/\sigma_\eta^2)^{-1}(\mathbf{\Lambda}^{-1}\mathbf{b}_0 - \mathbf{A}^T\mathbf{M}\mathbf{s}/\sigma_\eta^2) \\ &= \mathbf{b}_0 - (\sigma_\eta^2\Lambda^{-1} + (\mathbf{A}^T\mathbf{A}))^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{b}_0 + \mathbf{M}\mathbf{s}), \end{aligned} \tag{34}$$

where

$$J(\mathbf{b}_{\text{opt}}) = ||\mathbf{A}\mathbf{b}_0 + \mathbf{M}\mathbf{s}||^{\mathbf{2}}_{\left(\sigma_\eta^{\mathbf{2}}\mathbf{I}+\mathbf{A}\mathbf{\Lambda}\mathbf{A^T}\right)^{-\mathbf{1}}}. \tag{35}$$

We now summarize the procedure proposed for recovering the sparse vector $\mathbf{s}$ given internal nodes measurements and when only partial information is given on the grid parameters. The algorithm starts with an initial estimate $\mathbf{b} = \mathbf{b}_0$. Given this initialization of $\mathbf{b}$, we seek for a sparse vector $\mathbf{s}$ which solves the optimization in (31b), and then estimate $\tilde{\theta}_E$ according to (31a). Having these estimates of $\mathbf{s}$ and $\tilde{\theta}_E$, we estimate $\mathbf{b}$ according to (34). We continue iterating between (31) and (34) until convergence. Note that at each iteration of the algorithm the objective function is decreased. Note also that in estimating the sparse vector $\mathbf{s}$, rather than building its support from the beginning at each iteration, we can exploit information from previous iterations.

REFERENCES

[1] 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.

[2] 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.

[3] R. Emami and A. Abur, "Tracking changes in the external network model," September 2010.

[4] 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.

[5] A. J. Wood and B. F. Wollenberg, *Power generation, Operation, and Control, 2nd ed.*, Wiley, 1996.

[6] 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.

[7] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*, Cambridge University Press, 2012.