Sparse Overcomplete Representations for Efficient Identification of Power Line Outages

Hao Zhu, *Student Member, IEEE*, and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—Fast and accurate unveiling of power-line outages is of paramount importance not only for preventing faults that may lead to blackouts, but also for routine monitoring and control tasks of the smart grid, including state estimation and optimal power flow. Existing approaches are either challenged by the combinatorial complexity issues involved and are thus limited to identifying single and double line-outages or they invoke less pragmatic assumptions such as conditionally independent phasor angle measurements available across the grid. Using only a subset of voltage phasor angle data, the present paper develops a near real-time algorithm for identifying multiple line outages at the affordable complexity of solving a sparse signal reconstruction problem via either greedy steps or coordinate descent iterations. Recognizing that the number of line outages is a small fraction of the total number of lines, the novel approach relies on reformulating the DC linear power flow model as a sparse overcomplete expansion and leveraging contemporary advances in compressive sampling and variable selection. This sparse representation can also be extended to incorporate available information on the internal system and more general line-parameter faults. Analysis and simulated tests on 118-, 300-, and 2383-bus systems confirm the effectiveness of identifying sparse power line outages.**

***Index Terms*—Cascading failures, compressive sampling, identification of line outages, Lasso, matching pursuit.**

I. Introduction

IT is well appreciated that major blackouts have occurred partly due to lack of comprehensive situational awareness of the power grid [22]. This speaks for the importance of timely monitoring the status of generators, transformers, and transmission lines. Identifying outages and, generally, changes of lines is particularly critical for a number of tasks, including state estimation, optimal power flow, real-time contingency analysis, and, thus, security assessment of power systems.

To appreciate the opportunities and challenges facing the line-change identification problem, it is prudent to think of the grid as a graph comprising topologically interconnected power systems. Phasor measurement units (PMUs) provide voltage and power data per local (a.k.a. internal) system in real time. Likewise, real-time data are telemetered internally per system to offer topology-bearing information on the connectivity status of local circuit breakers and switches. On the other hand, power flow conservation across interconnected systems allows for identifying changes even in external system lines—a critical task for comprehensive, wide-area monitoring. This would have been a nonissue if intersystem data were available. Unfortunately, the system data exchange (SDX) module of the North American Electric Reliability Corporation (NERC) can provide grid-wide interarea (i.e., basecase) topology information on an hourly basis [23], but the desiderata is near real-time monitoring of transmission lines. In a nutshell, the need arises for each internal system to identify, in a computationally efficient manner, line outages (and, generally, line changes) in its external counterpart relying only on basecase topology information and local PMU data.

Existing approaches typically formulate external line-outage identification as a combinatorially complex (integer programming) problem, which can be computationally tractable only for single or, at most, double line outages [5], [17], [18]. However, it is becoming increasingly crucial to cope with multiple line outages in the face of cascading failures in recent blackouts. This has really motivated several existing works [8], [9], which consider how to handle *multiple* line-outage distribution factors (LODFs) as compared with coping with a single LODF. In addition, an alternative recent approach to line-outage identification adopts a Gauss–Markov graphical model of the power network and can deal with multiple outages at affordable complexity, but assumes conditionally independent phasor angle measurements and requires intersystem PMU data to be available in real time *across* the grid [11].

This paper contributes a computationally efficient algorithm for near real-time identification of multiple external line outages (and generally changes) using only hourly basecase topology information and local voltage phasor angle data available by PMUs. It relies on the standard DC linear power flow model and leverages the fact that the outaged lines represent a small fraction of the total number of lines. The novel approach views the topology-bearing basecase information as the weighted Laplacian matrix of the grid-induced graph (Section II). This leads to an overcomplete representation of the outage-induced innovation, which in turn enables casting line-outage identification as a *sparse* vector estimation problem (Section III). Solving the latter draws from recent advances in compressive sampling and variable selection in linear regression problems [3], [4], [19], [20] and adopts the greedy orthogonal matching pursuit (OMP) and the convex least-absolute shrinkage and selection operator (Lasso) to identify line outages at affordable complexity (Section IV). Simulated tests corroborate the merits of the novel algorithm (Section V). Interesting extensions of

Manuscript received October 29, 2011; revised November 09, 2011; accepted March 20, 2012. Date of publication May 01, 2012; date of current version October 17, 2012. This work was supported in part by the National Science Foundation under Grant CCF-1016605 and a Doctoral Dissertation Fellowship from the University of Minnesota. Paper no. TPWRS-01023-2011.

The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: zhuh@umn.edu; georgios@umn.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2012.2192142

the sparse representation are also offered (Section VI), and the paper is wrapped up with a concluding summary (Section VII).

Notation: Upper (lower) boldface letters will be used for matrices (column vectors); $(\cdot)^T$ denotes transposition, $\mathbf{1}$ is the all-one vector, $\mathbf{0}$ is the all-zero vector, $\mathbf{I}$ is the identity matrix, $\|\cdot\|_p$ is the vector p-norm for $p \geq 1$, $|\cdot|$ is the magnitude of a variable or the cardinality of a set, $\text{sign}(\cdot)$ is the sign operator, and $\mathcal{S}_1 \backslash \mathcal{S}_2$ is the relative complement of the set $\mathcal{S}_2$ in the set $\mathcal{S}_1$.

II. Modeling and Problem Statement

Consider a power transmission network consisting of N buses denoted by the set of nodes $\mathcal{N} := \{1, \ldots, N\}$, and L transmission lines represented by the set of edges $\mathcal{E} := \{(m,n)\} \subseteq \mathcal{N} \times \mathcal{N}$. Collect all of the voltage phasor angles $\{\theta_n\}$, one per bus $n \in \mathcal{N}$, in the vector $\boldsymbol{\theta} \in \mathbb{R}^N$ and, correspondingly, the injected power variables $\{P_n\}$ in $\mathbf{p} \in \mathbb{R}^N$. The network buses comprise the union $\mathcal{N} = \mathcal{N}_I \bigcup \mathcal{N}_E$ with $\mathcal{N}_I \bigcap \mathcal{N}_E = \varnothing$, where $\mathcal{N}_I$ denotes the subset of observable buses in the internal system, and $\mathcal{N}_E$ stands for the unobservable buses of the external system. Accordingly, $\boldsymbol{\theta}$ comprises two subvectors $\boldsymbol{\theta}_I$ and $\boldsymbol{\theta}_E$, collecting phasor angles of voltages at $\mathcal{N}_I$ and $\mathcal{N}_E$ buses, respectively, and similarly for the $\mathbf{p}_I$ and $\mathbf{p}_E$ partitions of $\mathbf{p}$. Supposing that the network-wide injected-power vector $\mathbf{p}$ only changes gradually across time, the present work aims to unveil line changes, including (possibly multiple) line outages anywhere in the network, using data $\boldsymbol{\theta}_I$ acquired in real time from PMUs. Before the problem formulation, it is important to understand how the network topology dictates the relationship between $\mathbf{p}$ and $\boldsymbol{\theta}$.

A. Linear DC Power Flow Model and Its Graph Laplacian

Power flow models are useful for determining how injected power $\mathbf{p}$ flows along all transmission lines. Providing a linear approximation of the actual ac system, the DC power flow model turns out to facilitate a variety of power system monitoring tasks under normal operating conditions, including security-constrained contingency analysis, and system state estimation; see, e.g., [25, ch. 11–12]. In the linear DC model, the power injected to bus n balances all of the line flows originating from it, that is,

$$P_m = \sum_{n \in \mathcal{N}_m} P_{mn} = \sum_{n \in \mathcal{N}_m} \frac{1}{x_{mn}} (\theta_m - \theta_n) \tag{1}$$

where $x_{mn} = x_{nm}$ denotes the reactance along line (m,n), and $\mathcal{N}_m$ is the set of neighboring buses linked to bus m. Writing (1) in vector-matrix form yields

$$\mathbf{p} = \mathbf{B}\boldsymbol{\theta} \tag{2}$$

where the $N \times N$ matrix $\mathbf{B}$ has its (m,n)th entry given by

$$B_{mn} = \begin{cases} \frac{-1}{(x_{mn})}, & \text{if } (m,n) \in \mathcal{E} \\ \sum\limits_{\nu \in \mathcal{N}_m} \frac{1}{(x_\nu)}, & \text{if } m = n \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Matrix $\mathbf{B}$, relating the voltage-phasor angle $\boldsymbol{\theta}$ to the injected power $\mathbf{p}$ as in (2), is uniquely determined by the line reactance parameters $\{x_{mn}\}$ of the network $(\mathcal{N}, \mathcal{E})$, and topology-bearing information, namely $\mathcal{E}$, provided by the SDX. At this point, it is worth noting that each x_{mn} reactance is present in only four entries of $\mathbf{B}$, namely, (m,m), (m,n), (n,m), and (n,n). This observation about line's (m,n) presence in $\mathbf{B}$ intuitively describes the selective effect topology changes due to line outages can have on the angle vector $\boldsymbol{\theta}$ in (2). To formulate this topological effect concretely, $\mathbf{B}$ can be viewed as the weighted Laplacian matrix of the graph $(\mathcal{N}, \mathcal{E})$. To this end, consider the $N \times L$ bus-line incidence matrix $\mathbf{M}$; see, e.g., [24, p. 56], formed by columns $\{\mathbf{m}_\ell\}_{\ell=1}^L$ of length N. With subscript ℓ corresponding to the line (m,n), the column $\mathbf{m}_\ell$ has all its entries equal to 0 except the mth and nth, which take the value 1 and -1, respectively. Clearly, the rank-one matrix $\mathbf{m}_\ell \mathbf{m}_\ell^T$ has nonzero entries on its diagonal positions (m,m) and (n,n), which equal 1, and also on its (m,n) and (n,m) positions, which equal -1. Thus, when summing such rank-one matrices for distinct lines, those having common m or n are superimposed on the diagonal, while those off the diagonal still have values equal to -1. This argument and (3) establish readily the following representation of the network topology matrix:

$$\sum_{\ell=1}^{L} \frac{1}{x_\ell} \mathbf{m}_\ell \mathbf{m}_\ell^T = \mathbf{M}\mathbf{D}_x\mathbf{M}^T = \mathbf{B} \tag{4}$$

where the diagonal matrix $\mathbf{D}_x$ has its ℓth diagonal entry $1/x_\ell$ equal to the inverse reactance $1/x_{mn}$, if ℓ corresponds to the line (m,n). In addition, matrix $\mathbf{B}$ is symmetric and has rank $N-1$ if the power network is connected, since its null space is only spanned by $\mathbf{1}$; see, e.g., [24, p. 469]. Rank deficiency of $\mathbf{B}$ gives rise to multiple solutions for $\boldsymbol{\theta}$ in (2). To fix this ambiguity, one generation bus is typically chosen as reference with its phasor angle set to zero, case in which phasor angles of all other buses denote their differences relative to the reference bus; see e.g., [25, p. 76]. Clearly, with the reference bus convention, the $(N-1) \times (N-1)$ matrix $\mathbf{B}$ has full rank and can be formed by $\mathbf{M}$ as in (4) after removing the corresponding row of the incidence matrix.

B. Unveiling Power Line Outages

Given the linear DC power flow model (2) and the aforementioned reference bus convention, the pre-event phasor angles $\boldsymbol{\theta}$ can be uniquely characterized by the injected power vector $\mathbf{p}$ and the topology-dependent matrix $\mathbf{B}$. Suppose that, due to changes in the grid, e.g., cascading failures in an early stage, several outages occur on lines, collected in the subset $\tilde{\mathcal{E}} \subseteq \mathcal{E}$. Line outages on the transmission network yield the post-event graph $(\mathcal{N}, \mathcal{E}')$, with[1] $\mathcal{E}' := \mathcal{E} \backslash \tilde{\mathcal{E}}$.

As in [17], it is assumed that fast system dynamics are well damped and that the system settles down to a quasi-stable state following the line outages. The possibility of poor system damping can be accounted for by low-pass filtering of the measured sequential phasor angle data to smooth out any system oscillations, as detailed in [17]. Furthermore, as assumed in [18], the outaged lines in $\tilde{\mathcal{E}}$ will not result in islanding of the post-event system; that is, the underlying graph will not become disconnected. This precludes considerable changes between pre- and post-event bus power injections, such as

[1]As a mnemonic, post-event quantities are denoted with prime, and the differences relative to their pre-event counterparts are denoted with tilde.

those caused by generator outages or radial line outages. Under these considerations, the linear DC model for the quasi-stable post-event network is given by [cf. (2)]

$$\mathbf{p}' = \mathbf{p} + \boldsymbol{\eta} = \mathbf{B}'\boldsymbol{\theta}' \tag{5}$$

where $\mathbf{B}'$ is the post-event weighted Laplacian of $(\mathcal{N}, \mathcal{E}')$, and noise vector $\boldsymbol{\eta}$ accounts for the small perturbations between $\mathbf{p}'$ and $\mathbf{p}$ due to, e.g., variations in the bus loads, which can usually be modeled as zero-mean (possibly Gaussian) vector with covariance matrix $\sigma_\eta^2\mathbf{I}$; see e.g., [15].

Voltage phasor angles are only available at the subset of internal buses $\mathcal{N}_I$. Thus, unveiling outages amounts to identifying lines in $\tilde{\mathcal{E}}$, given the pre-event network-wide topology matrix $\mathbf{B}$, as well as pre- and post-event internal phasor angle vectors $\boldsymbol{\theta}_I$ and $\boldsymbol{\theta}'_I$. It will turn out that solving this problem incurs combinatorial complexity. Fortunately, using (2) and (5), the ensuing section leverages (3) of $\mathbf{B}$ as the weighted Laplacian of the underlying graph to derive an overcomplete representation for line outages, allowing for tractable solvers of the identification problem at hand. However, before this, it is important to point out the difference between the linear DC power flow model here and the probabilistic dependence graph model in [11]. The Gauss–Markov random field (GMRF) approach pursued in [11] requires the less pragmatic assumptions on the *conditional independence* among bus phasor angles, as well as the availability of $\boldsymbol{\theta}$ across the network. In contrast, these assumptions are not required for the DC model adopted here and in [5], [17], and [18].

Remark 1 (Nonrandom Errors Due to the DC Power Flow Model Approximation): Being an approximation of the real, ac physical system behavior, the linear DC flow model may introduce sizeable nonrandom errors on top of the random errors captured by $\boldsymbol{\eta}$, which are due to power injection variations. Bounds on the non-random errors have been derived in [12], and tested extensively in [16]. If sizeable, such errors will render the DC model less accurate, and thus challenge *all* approaches to line outage identification that rely on the DC model. Fortunately, effective remedies are available to decrease these errors, such as using "hot-start" modeling along with unequal branch sending and receiving flows [16]. In lieu of analytic means, the effect of DC approximation errors will be assessed in Section V via numerical tests.

III. Sparse Overcomplete Representation

Here, the line-outage identification task is formulated as that of recovering the coefficients of a sparse overcomplete representation of the pre- and post-event phase difference of the internal system's PMU data. To this end, the difference $\tilde{\mathbf{B}} := \mathbf{B} - \mathbf{B}'$ denoting the weighted Laplacian for the outage lines in $\tilde{\mathcal{E}}$, is written as [cf. (4)]

$$\tilde{\mathbf{B}} = \sum_{\ell\in\tilde{\mathcal{E}}} \frac{1}{x_\ell}\mathbf{m}_\ell\mathbf{m}_\ell^T. \tag{6}$$

Substituting the pre-event power flow model (2) into the post-event one in (5) leads to

$$\mathbf{B}\boldsymbol{\theta} + \boldsymbol{\eta} = \mathbf{B}'\boldsymbol{\theta}' = \mathbf{B}\boldsymbol{\theta}' - \tilde{\mathbf{B}}\boldsymbol{\theta}'. \tag{7}$$

Consider now the phasor angle change vector $\tilde{\boldsymbol{\theta}} := \boldsymbol{\theta}' - \boldsymbol{\theta}$, and partition it into the available $\tilde{\boldsymbol{\theta}}_I$ and the unavailable $\tilde{\boldsymbol{\theta}}_E$, each corresponding to buses in $\mathcal{N}_I$ and $\mathcal{N}_E$. With these notational conventions, substituting (6) into (7), yields

$$\begin{aligned}\mathbf{B}\tilde{\boldsymbol{\theta}} &= \tilde{\mathbf{B}}\boldsymbol{\theta}' + \boldsymbol{\eta}\\ &= \sum_{\ell\in\tilde{\mathcal{E}}} s_\ell\,\mathbf{m}_\ell + \boldsymbol{\eta}\end{aligned} \tag{8}$$

where $s_\ell := \mathbf{m}_\ell^T\boldsymbol{\theta}'/x_\ell$, $\forall \ell \in \tilde{\mathcal{E}}$. Solving (8) with respect to (wrt) $\tilde{\boldsymbol{\theta}}$, and extracting the rows corresponding to the buses in $\mathcal{N}_I$ leads to the following expression for the internal post- minus pre-event phase difference vector:

$$\tilde{\boldsymbol{\theta}}_I = [\mathbf{B}^{-1}]_I\left(\sum_{\ell\in\tilde{\mathcal{E}}} s_\ell\,\mathbf{m}_\ell\right) + [\mathbf{B}^{-1}]_I\boldsymbol{\eta} \tag{9}$$

where $[\mathbf{B}^{-1}]_I$ is constructed from the rows corresponding to buses in $\mathcal{N}_I$ of the inverse matrix $\mathbf{B}^{-1}$.

Suppose that the number of line outages, namely $L_o := |\tilde{\mathcal{E}}|$, is given. The number of possible topologies, each having L_o outages, is clearly the number of combinations $C := \binom{L}{L_o}$. Let $\{\tilde{\mathcal{E}}^{(c)}\}_{c=1}^C$ denote the sets of line outages corresponding to these candidate topologies. Existing approaches exhaustively check the ℓ_2-norm of the least-squares (LS) error in the internal phase difference vector for each candidate $\tilde{\mathcal{E}}^{(c)}$, and select the minimum, that is, given $\tilde{\boldsymbol{\theta}}_I$ and $\mathbf{B}$, the chosen topology index is

$$\hat{c} := \arg\min_{c=1,\ldots,C}\left\{\min_{\{s_\ell\}}\left\|\tilde{\boldsymbol{\theta}}_I - \left(\sum_{\ell\in\tilde{\mathcal{E}}^{(c)}}[\mathbf{B}^{-1}]_I\mathbf{m}_\ell s_\ell\right)\right\|_2^2\right\}. \tag{10}$$

Such an approach incurs combinatorial complexity and has limited existing methods based on exhaustive enumeration of all combinations to identifying single line outage $(L_o = 1)$ [17] or at most double line outages $(L_o = 2)$ [18]. Also, it will become clear soon that the covariance information of $\boldsymbol{\eta}$ is not exploited optimally in (10).

To bypass this combinatorial complexity, the fresh idea here is to consider an overcomplete representation capturing all possible line outages. However, notice that the inversion in (9) introduces colored perturbation. To account for this, consider the compact singular value decomposition (SVD) of the fat matrix $[\mathbf{B}^{-1}]_I = \mathbf{U}_I\boldsymbol{\Sigma}_I\mathbf{V}_I^T$, where the square diagonal matrix $\boldsymbol{\Sigma}_I$ comprises all of its $|\mathcal{N}_I|$ nonzero singular values. Upon defining $\mathbf{y} := \boldsymbol{\Sigma}_I^{-1}\mathbf{U}_I^T\tilde{\boldsymbol{\theta}}_I$, the data model in (9) can be reduced to a sparse linear regression one with unknown regression coefficients contained in the $L\times 1$ vector $\mathbf{s}$, whose ℓth entry equals s_ℓ, if $\ell \in \tilde{\mathcal{E}}$, and 0 otherwise. Thus, (8) can be written as

$$\begin{aligned}\mathbf{y} &= \mathbf{V}_I^T\left(\sum_{\ell\in\tilde{\mathcal{E}}} s_\ell\,\mathbf{m}_\ell\right) + \mathbf{V}_I^T\boldsymbol{\eta}\\ &= \mathbf{V}_I^T\mathbf{M}\mathbf{s} + \mathbf{V}_I^T\boldsymbol{\eta} \qquad (11)\\ &= \mathbf{A}\mathbf{s} + \mathbf{V}_I^T\boldsymbol{\eta} \qquad (12)\end{aligned}$$

where the transformed incidence (a.k.a. adjacency) matrix $\mathbf{A} := \mathbf{V}_I^T\mathbf{M}$, now viewed as a regression matrix, captures all of the pre-event transmission lines.

Remark 2 (Perturbation Whitening): After transforming $\tilde{\boldsymbol{\theta}}_I$ to $\mathbf{y}$, the effective perturbation $\mathbf{V}_I^T\boldsymbol{\eta}$ in (12) is still zero-mean with covariance $\sigma_\eta^2\mathbf{I}$. However, the enumeration-based approaches in [17], [18] or the mixed-integer programming one in [5] compute the ℓ_2-norm LS error for each candidate topology with line outages relying on the pre-whitened inverted linear system, and do not account for the nonwhite perturbation therein. Since optimal estimation (in the unbiased minimum variance sense) of linear regression models with nonwhite noise calls for *weighted* LS solvers, ignoring the noise color leads to suboptimal performance of the LS-based enumeration approaches that rely on (10), especially when the power perturbation is large. This will be illustrated further in the simulation results of Section V.

In essence, the overcomplete representation in (12) allows one to cast the problem of recovering the subset $\tilde{\mathcal{E}}$ as one of estimating the *sparse* vector $\mathbf{s}$. Different from (9), the line-outage set $\tilde{\mathcal{E}}$ is no longer present in (12), which eliminates the need for exhaustive search. The key *premise* is that the number of line outages is a small fraction of the total number of lines, i.e., $L_o := |\tilde{\mathcal{E}}| \ll L$. This holds even for multiple line outages such as those occurring during cascading failures, at least in the early stage when only a small number of lines start to fail. The same premise is adopted for contingency analysis, where usually single and double line outages are of primary concern; see, e.g., [25, ch. 11]. Under this sparsity constraint, the lines in outage are few, meaning that the vector $\mathbf{s}$ has only a few nonzero entries. In turn, this suggests recovering line outages by exploiting the sparsity in $\mathbf{s}$, which falls under the realm of sparse signal reconstruction approaches. Reconstruction of sparse signals has become very popular recently, especially after the emergence of compressive sampling (CS) theory; see, e.g., [3], [4], [20] and references therein, wherein several efficient algorithms have been proposed. A related sparse overcomplete representation has been used in [7] to address the fault estimation problem for electrical circuits, under a completely different system model based on Kirchhoff's laws.

Compared to previous works constrained to at most double line outages [5], [17], [18], the proposed formulation overcomes the inherent combinatorial complexity with a virtual "outage" at every possible transmission line. By leveraging the sparse signal representation framework, Section IV will focus on algorithms offering the potential to achieve high accuracy in recovering faults even with more than two line outages, while bypassing the exhaustive search which is computationally prohibitive.

IV. Sparse Line-Outage Identification

Approaches to reconstructing sparse coefficient vectors of linear regression models can be broadly grouped into two categories: those relying on greedy approximation schemes and those minimizing the ℓ_1-norm of the sparse vector.

The first category is rooted on the matching pursuit (MP) algorithm for signal approximation [13]. MP has undergone improvements which led to the orthogonal matching pursuit (OMP) algorithm [20], that draws its popularity from its computational simplicity and guaranteed performance. The second category relies on the ℓ_1-norm of the wanted vector offering a convex relaxation of the ℓ_0-norm based minimization problems, which are known to be NP-hard but can yield the sparsest solution to an underdetermined linear system of equations; see e.g., [3], [4]. Popular solvers of the resultant optimization problems are variants of the basis pursuit (BP) algorithm [4], or the Lasso [19]. Albeit approximating the optimal ℓ_0-norm, this relaxation has been shown to exhibit guaranteed performance of recovering sparse signal vectors; see e.g., [3] and references therein. As a convex quadratic program (QP), the global minimizer of the ℓ_1-norm regularized Lasso problem can be efficiently obtained under various sparsity levels using coordinate descent (CD) iterations [21].

A. Greedy Line-Outage Identification Via OMP

Given vector $\mathbf{y}$ obtained from $\tilde{\boldsymbol{\theta}}_I$ and $\mathbf{A}$, the greedy OMP algorithm will be outlined next for finding sparse solutions to (12) wrt a fixed number of nonzero entries (i.e., sparsity level) κ of the sought solution $\hat{\mathbf{s}}^\kappa$ to (12). The number κ can be directly related to the number of lines in outage, namely L_o, and its choice will be discussed in Section IV-C.

Greedy schemes generally aim to approximate the vector $\mathbf{y}$ by successively selecting columns $\{\mathbf{a}_\ell\}_{\ell=1}^L$ of $\mathbf{A}$ to superimpose with weights given by the wanted nonzero entries of $\mathbf{s}$. The selection criteria may be different but all share the common idea of choosing the next column as the one that correlates "best" with the current approximation error. Updating the sparsity level κ one by one, the approximation error vector per step is given by

$$\mathbf{r}^\kappa := \mathbf{y} - \mathbf{A}\hat{\mathbf{s}}^\kappa, \quad \kappa \geq 1, \quad \mathbf{r}^0 = \mathbf{y}. \tag{13}$$

Furthermore, with $\mathcal{L}^\kappa$ denoting the subset of indexes corresponding to nonzero entries in $\hat{\mathbf{s}}^\kappa$, once a new column ℓ^κ is chosen at step κ for inclusion in $\mathcal{L}^\kappa$, the ℓ^κth entry of $\hat{\mathbf{s}}^\kappa$ will be nonzero compared with $\hat{\mathbf{s}}^{\kappa-1}$. The orthogonality in OMP is manifested by updating the estimate $\hat{\mathbf{s}}^\kappa$ via LS fitting of $\mathbf{y}$ using all of the columns in $\mathcal{L}^\kappa$, such that $\mathbf{r}^\kappa$ in (13) becomes orthogonal to all the chosen columns in $\mathcal{L}^\kappa$. The OMP based line-outage solution to (12) with maximum sparsity level $\kappa_{\max}$ is tabulated as Algorithm 1.

Since Algorithm 1 chooses the column by comparing its correlation with the residual error vector, it is useful to normalize all columns first, so that comparison is in essence performed based on the correlation coefficient. Also, due to its greedy update, the sequence of solutions $\{\hat{\mathbf{s}}^\kappa\}_{\kappa=1}^{\kappa_{\max}}$ is nested in terms of their nonzero entries, but not necessarily in the exact entry values after the orthogonal LS fitting step. In addition to computational efficiency, the simulations in Section V will demonstrate that Algorithm 1 exhibits also quite competitive reconstruction performance. Analytically, the latter depends on the so-termed coherence of matrix $\mathbf{A}$ that is defined as

$$\mu := \max_{\ell \neq \ell'} \left|\mathbf{a}_\ell^T\mathbf{a}_{\ell'}\right|. \tag{14}$$

By adapting [20, Thm. B], the following proposition can be established for the present context of identifying line outages.

Proposition 1 (Exact Recovery of OMP): Consider the noise-free case, i.e., $\boldsymbol{\eta} = \mathbf{0}$ in (12), and suppose that the actual vector $\mathbf{s}_o$ has its number of nonzero entries L_o satisfying

$$L_o < \frac{(\mu^{-1}+1)}{2}. \tag{15}$$

Given the vector $\mathbf{y} = \mathbf{A}\mathbf{s}_o$ and matrix $\mathbf{A}$, the OMP in Algorithm 1 is capable of recovering the unknown $\mathbf{s}_o$ exactly, if the greedy steps stop once $\|\mathbf{r}^\kappa\|_2 = 0$.

Algorithm 1 (OMP): Input $\mathbf{y}$, $\mathbf{A}$, and $\kappa_{\max}$. Output $\{\hat{\mathbf{s}}^\kappa\}_{\kappa=1}^{\kappa_{\max}}$

Initialize $\mathbf{r}^0 = \mathbf{y}$, and the column subset $\mathcal{L}^0 = \emptyset$.

for $\kappa = 1, \ldots, \kappa_{\max}$ **do**

Choose $\ell^\kappa \in \arg\max_\ell |\mathbf{a}_\ell^T \mathbf{r}^{\kappa-1}|$ (arbitrarily breaking ties), and update $\mathcal{L}^\kappa := \mathcal{L}^{\kappa-1} \bigcup \{\ell^\kappa\}$.

Estimate $\hat{\mathbf{s}}^\kappa := \arg\min_{\mathbf{s}} \|\mathbf{y} - \mathbf{A}\mathbf{s}\|_2^2$, by setting its ℓ-th entry $s_\ell = 0$ for all $\ell \notin \mathcal{L}^\kappa$.

Update $\mathbf{r}^\kappa := \mathbf{y} - \mathbf{A}\hat{\mathbf{s}}^\kappa$.

end for

B. Lassoing Line Outages via CD

Lasso promotes sparse solutions by regularizing the LS error criterion with the ℓ_1-norm of $\mathbf{s}$. Specifically, the coefficient vector whose nonzero entries identify line outages is found as

$$\hat{\mathbf{s}}^\lambda := \arg\min_{\mathbf{s}} \|\mathbf{y} - \mathbf{A}\mathbf{s}\|_2^2 + \lambda\|\mathbf{s}\|_1 \tag{16}$$

where λ is a tuning parameter whose choice is detailed in Section IV-C. Thanks to the overcomplete representation (12) and the ℓ_1-norm relaxation, the line-outage identification is reduced to the convex QP in (16). Hence, its global minimizer can be efficiently obtained using general-purpose convex solvers, such as interior-point algorithms [2, ch. 11]. Instead of these off-the-shelf solvers, the CD iterative solver [21] will be adapted, because it has been shown to exploit the specific problem structure and can yield the so-termed regularization path of Lasso-based solutions as a function of λ [6]. This approach is advocated especially for sparse unknown vectors of high dimension, that is, for a large number of lines L.

To allow for a sequence of λ values, consider first (16) with a fixed λ. CD optimizes the regularized LS cost (16) by cyclically minimizing over the coordinates, namely, the scalar entries of $\mathbf{s}$. It yields successive estimates of each coordinate, while keeping the rest fixed. Suppose that the ℓth entry $s_\ell(i)$ is to be found. Precursor entries $\{s_1(i), \ldots, s_{\ell-1}(i)\}$ have been already obtained in the ith iteration along with postcursor entries $\{s_{\ell+1}(i-1), \ldots, s_L(i-1)\}$ are also available from the previous $(i-1)$st iteration. Thus, the effect of these given entries can be removed from $\mathbf{y}$ by forming

$$\mathbf{e}_\ell(i) := \mathbf{y} - \sum_{j=1}^{\ell-1} \mathbf{a}_j s_j(i) - \sum_{j=\ell+1}^{L} \mathbf{a}_j s_j(i-1) \tag{17}$$

where $\mathbf{a}_j$ denotes the jth column of matrix $\mathbf{A}$. Using (17), the vector optimization problem in (16) reduces to the following scalar one with $s_\ell(i)$ as unknown: $s_\ell(i) = \arg\min_{s_\ell}[\|\mathbf{e}_\ell(i) - \mathbf{a}_\ell s_\ell\|_2^2 + \lambda|s_\ell|]$. This is the popular *scalar* Lasso problem, which admits a closed-form solution expressed via a soft thresholding operator as (see, e.g., [6])

$$s_\ell(i) = \text{sign}\left(\mathbf{a}_\ell^T \mathbf{e}_\ell(i)\right)\left[\frac{|\mathbf{a}_\ell^T \mathbf{e}_\ell(i)|}{\|\mathbf{a}_\ell\|_2^2} - \frac{\lambda}{2\|\mathbf{a}_\ell\|_2^2}\right]_+ \quad \forall \ell \tag{18}$$

where $[\chi]_+ := \chi$, if $\chi > 0$, and zero otherwise.

Cycling through the closed forms in (17) and (18) explains why CD here is faster than, and thus preferable, over general-purpose convex solvers. The residual update in (17) is feasible iteratively, while the soft thresholding operation in (18) is also very fast. Following the basic convergence results in [21], the CD iteration is provably convergent to the global optimum $\hat{\mathbf{s}}^\lambda$ of (16), as asserted in the following proposition.

Proposition 2 (Convergence of CD): Given the parameter λ and arbitrary initialization, the iterates $\{\mathbf{s}(i)\}$ given by (18) converge monotonically to the global optimum $\hat{\mathbf{s}}^\lambda$ of the line outage identification problem in (16).

The identification path is further obtained by applying the CD method for a decreasing sequence of λ values. Larger λ's in (18) force more entries of $\mathbf{s}(i)$ to be nulled. Hence, if a sufficiently large parameter λ is picked, the corresponding $\hat{\mathbf{s}}^\lambda$ will eventually become $\mathbf{0}$. With a decreasing sequence of λ's, $\hat{\mathbf{s}}^\lambda$ for a large λ can be used as a *warm start* for solving (16) with the second largest λ. This way, the CD based line-outage solution path of (16), as tabulated in Algorithm 2, exploits both the efficient scalar solution in (18) as well as warm starts, to ensure reduced complexity and algorithmic stability.

Algorithm 2 (CD): Input $\mathbf{y}$, $\mathbf{A}$, and a decreasing sequence of λ values. Output $\{\hat{\mathbf{s}}^\lambda\}$ in **(16)** for each λ

Initialize with $i = -1$ and $\mathbf{s}(-1) = \mathbf{0}$.

for each λ value from the decreasing sequence **do**

repeat

Set $i := i + 1$.

for $\ell = 1, \ldots, L$ **do**

Compute the residual $\mathbf{e}_\ell(i)$ as in **(17)**.

Update the scalar $s_\ell(i)$ via **(18)**.

end for

until CD convergence is achieved.

Save $\hat{\mathbf{s}}^\lambda = \mathbf{s}(i)$ for the current λ value.

Initialize with $i = -1$ and $\mathbf{s}(-1) = \mathbf{s}(i)$ as the warm start for the next λ.

end for

Remark 3 (Computational Complexity): As mentioned in Section III, enumeration incurs combinatorial complexity while checking all C possible topologies $\{\tilde{\mathcal{E}}^{(c)}\}_{c=1}^{C}$ with a fixed L_o. Hence, its complexity is $\mathcal{O}(L^{L_o})$, which grows exponentially with L_o. In contrast, the recursive OMP Algorithm 1 has complexity $\mathcal{O}(\kappa_{\max} L)$, which is bilinear in L and $\kappa_{\max}$. For Algorithm 2, each CD iteration entails only L scalar operations,

while further acceleration is possible by exploiting sparsity and warm starts across CD iterations. Numerical running time comparisons will also be given in Section V to corroborate the linear complexity of the proposed algorithms.

C. Selection of Tuning Parameters

Algorithms 1 and 2 both yield a sequence of line-outage identification solutions. Selection of the parameter κ or λ, however, becomes a critical issue, as it either directly or indirectly relates to the underlying sparsity level, and thus the (generally unknown) L_o. With additional prior information, existing statistical tests can be adopted in order to choose the actual number of lines in outage.

1) Number of Line Outages L_o is Fixed or Upper Bounded by $\kappa_{\max}$: If $|\tilde{\mathcal{E}}| = L_o$ is given, the greedy OMP algorithm can automatically find the solution $\hat{\mathbf{s}}^{\kappa}$ with $\kappa = L_o$. As for the Lasso Algorithm 2, the parameter λ can also be determined if L_o is known. In fact, by directly inspecting the regularization path one can determine λ, so that the solution $\hat{\mathbf{s}}^{\lambda}$ has the number of nonzero entries equal to the given L_o. The indexes of nonzero entries in $\hat{\mathbf{s}}^{\lambda}$ also correspond to the identified lines in $\tilde{\mathcal{E}}$. Note that the number of line outages L_o is assumed known and fixed in all existing works [5], [11], [17], [18].

If the maximum number of line outages is prescribed, i.e., $L_o \leq \kappa_{\max}$, Algorithm 1 can output all of the greedy solutions $\{\hat{\mathbf{s}}^{\kappa}\}$ from $\kappa = 1$ to $\kappa_{\max}$. Likewise, Algorithm 2 can yield solution $\hat{\mathbf{s}}^{\lambda}$ along the identification path with the number of nonzero entries not exceeding $\kappa_{\max}$. In order to determine the actual cardinality, one can adopt a minimum description length (MDL) type test; see, e.g., [10]. Considering the OMP solutions $\{\hat{\mathbf{s}}^{\kappa}\}$ for example, the MDL test will specify a rating score per sparsity level κ, given by

$$\rho(\kappa) = \frac{1}{\sigma_{\eta}^2}\|\mathbf{y} - \mathbf{A}\hat{\mathbf{s}}^{\kappa}\|_2^2 + \kappa \log(N_I), \qquad \kappa = 1, \ldots, \kappa_{\max} \tag{19}$$

where $N_I := |\mathcal{N}_I|$. Based on (19), the chosen number of line outages will be $\hat{\kappa} := \arg\min_{\kappa} \rho(\kappa)$, thus yielding the identified lines in outage in the set $\mathcal{L}^{\hat{\kappa}}$ obtained from $\hat{\mathbf{s}}^{\hat{\kappa}}$.

When computing the MDL score for the Lasso solutions along the path, the LS fitting error in the first summand of (19) is obtained not by simply replacing $\hat{\mathbf{s}}^{\kappa}$ by $\hat{\mathbf{s}}^{\lambda}$. Instead, the zero entries of $\hat{\mathbf{s}}^{\lambda}$ are retained and the minimum LS cost corresponding to the nonzero entries is used in the ℓ_2-norm term in (19). This is because the ℓ_1-norm in the Lasso cost (16) introduces bias in the estimate $\hat{\mathbf{s}}^{\lambda}$, and this bias can be removed using the aforementioned LS fitting.

2) Variance of Injected Power Noise is Known: If only the variance σ_{η}^2 of entries in $\boldsymbol{\eta}$ is known, one can handle even an unknown number of outage lines as follows. Consider again $\hat{\mathbf{s}}^{\kappa}$ for each κ and find the equivalent LS reconstruction error to obtain the sample variance of $\boldsymbol{\eta}$ given by

$$\hat{\sigma}_{\eta}^2(\kappa) = \frac{\|\mathbf{y} - \mathbf{A}\hat{\mathbf{s}}^{\kappa}\|_2^2}{N_I}, \ \kappa = 1, \ldots, \kappa_{\max}. \tag{20}$$

This test selects the desired sparsity level κ as the one yielding the most accurate variance estimate as $\hat{\kappa} := \arg\min_{\kappa} |\hat{\sigma}_{\eta}^2(\kappa) - \sigma_{\eta}^2|$. The latter is used to identify the outage lines in $\mathcal{L}^{\hat{\kappa}}$ as described in Section IV-A. A similar method also applies to the Lasso approach after removing the bias in estimates $\hat{\mathbf{s}}^{\lambda}$, as explained earlier.

V. Numerical Tests

The proposed sparsity-exploiting fault diagnosis algorithms are tested in this section using first the IEEE 118-bus and 300-bus benchmark systems [14]. Both systems are initially tested with phasor angles available at all buses, in order to illustrate the merits of the proposed overcomplete representations in comparison with the optimal exhaustive search. Subsequently, line-outage identification relying on a subset of phasor angles is tested on the 118-bus system partitioned as in [5]. Specifically, the internal system contains buses with indexes in the set $\mathcal{N}_I := \{1 - 45, 113, 114, 115, 117\}$, and the external one with those in $\mathcal{N}_E := \{46 - 112, 116, 118\}$. The software toolbox MATPOWER [26] is used throughout to generate the phasor angle measurements as well as the pertinent power flows.

1) Test Case 1: The 118-bus system with the complete information on phasor angles available is considered first. AC power flows are generated for both pre- and post-event systems. All possible candidate topologies with a single line outage ($L_o = 1$), and 500 randomly chosen topologies with double line outages ($L_o = 2$) are tested. For each line-outage topology tested, ten realizations of the perturbation noise $\boldsymbol{\eta}$ in (5) are generated. Three methods are compared: Algorithms 1 and 2 and a statistically optimal exhaustive search (ES) using (8). Specifically, instead of solving for $\tilde{\boldsymbol{\theta}}_I$ in (9) and using the criterion (10) as proposed in [17] and [18], the optimal ES here utilizes the complete information of $\tilde{\boldsymbol{\theta}}$ and picks

$$\hat{c}_o := \arg \min_{c=1,\ldots,C} \left\{ \min_{\{s_\ell\}} \left\| \mathbf{B}\tilde{\boldsymbol{\theta}} - \left(\sum_{\ell \in \tilde{\mathcal{E}}^{(c)}} \mathbf{m}_\ell \, s_\ell \right) \right\|_2^2 \right\} \tag{21}$$

which is optimal because the noise $\boldsymbol{\eta}$ in (8) is white. (Contrast this with the enumeration in (10) used by [17] and [18], which is suboptimum because the color introduced by the matrix $[\mathbf{B}]_I^{-1}$ multiplying $\boldsymbol{\eta}$ is not accounted for.) The number L_o is assumed to be known to all three algorithms. For both single and double line-outage tests, the standard deviation of $\boldsymbol{\eta}$ is set either equal to zero, corresponding to a noise-free case, or, equal to 1%, 2%, or 5%, of the average pre-event power injection. The percentage of correctly identified line outages is listed[2] in Table I. Clearly, even under the ac flow model, Algorithms 1 and 2 exhibit performance comparable to the benchmark enumeration-based schemes of (21), while both incur complexity that grows linearly in L, as opposed to the combinatorial search, which, for $L_o = 2$, is $\mathcal{O}(L^2)$. This complexity comparison will be numerically demonstrated soon in Test Case 4.

For any fixed noise level, the performance of double line-outage tests does not degrade appreciably when compared to that of single line-outage tests. As the injection noise level increases, all three approaches experience performance degradation, which is expected. The OMP algorithm's performance resembles the optimal ES one thanks to its exact recovery asserted in Proposition 1. Interestingly, the ES method is not always the best one, while the Lasso solution of Algorithm 2 in certain

[2]In all tables, "ES" stands for the exhaustive search algorithm, and the number in boldface font indicates the best performance per row.

TABLE I
118-BUS SYSTEM WITH ALL BUS PHASOR ANGLE MEASUREMENTS

	ES	Alg. 1	Alg. 2
Single (0%)	**95.5**	**95.5**	**95.5**
Single (1%)	92.1	92.1	**93.0**
Single (2%)	88.0	88.0	**89.8**
Single (5%)	76.8	76.8	**76.9**
Double (0%)	**95.0**	**95.0**	94.9
Double (1%)	**91.8**	91.7	91.7
Double (2%)	87.7	87.7	**88.3**
Double (5%)	76.7	**76.8**	**76.8**

TABLE II
300-BUS SYSTEM WITH ALL BUS PHASOR ANGLE MEASUREMENTS

	ES	Alg. 1	Alg. 2
Single (0%)	**96.6**	**96.6**	**96.6**
Single (1%)	**82.4**	**82.4**	82.0
Single (2%)	39.6	39.6	**40.3**
Double (0%)	92.4	**92.6**	90.2
Double (1%)	79.1	**79.2**	74.2
Double (2%)	36.5	**36.7**	33.9

cases exhibits improved noise resilience. This can be attributed to two reasons. First, the *empirical* identification rates in Table I depend on the random line-outage topology and injection noise realizations and may not be identical with the exact rates. This explains why results for all three approaches are numerically comparable. Second, statistical optimality of the ES criterion (21) implicitly assumes that the variable s_ℓ is uniformly distributed. However, it corresponds to the post-outage power flow for any $\ell \in \mathcal{E}$ [cf. (8)] and differs from line to line. Hence, it is possible that the ℓ_1-norm in the Lasso cost (16) aids discerning the wrong nonzero entries induced by the noise.

2) Test Case 2: The IEEE 300-bus system is tested here, keeping all other parameters identical to those of Test Case 1. The percentage of correctly identified line outages is listed in Table II, which also shows that the proposed algorithms are quite competitive wrt the enumeration based benchmark. Similar observations are also made when comparing different noise levels and the results of single and double line-outage cases. This effect of injection noises to performance degradation becomes more pronounced, and should be attributed to the approximation induced by the linear DC model for larger systems (cf. Remark 1).

3) Test Case 3: Based on the aforementioned internal and external grid partition, the 118-bus system is tested here using only part of the phasor angles in $\boldsymbol{\theta}_I$. All other settings are as in Test Case 1, and the results are listed in Table III. The suboptimum exhaustive search in (10) without whitening the colored noise is also included for comparison, and performs much worse than other (near-) optimal algorithms, especially when the noise level is high (5%). On the contrary, the exhaustive search based on the whitened system (12), given by

$$\hat{c}_o := \arg \min_{c=1,\ldots,C} \left\{ \min_{\{s_\ell\}} \left\| \mathbf{y} - \left(\sum_{\ell \in \tilde{\mathcal{E}}^{(c)}} \mathbf{a}_\ell s_\ell \right) \right\|_2^2 \right\} \quad (22)$$

is optimal in the unbiased minimum variance sense, and for this reason it is used to benchmark the performance of others.

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

	ES without whitening	ES with whitening	Alg. 1	Alg. 2
Single (0%)	64.8	**67.2**	**67.2**	64.0
Single (1%)	34.9	**60.8**	**60.8**	57.7
Single (2%)	25.3	**56.9**	**56.9**	55.6
Single (5%)	12.5	**52.1**	**52.1**	51.2
Double (0%)	59.1	**63.9**	63.8	56.1
Double (1%)	45.2	**60.4**	**60.4**	52.3
Double (2%)	36.5	53.1	**53.2**	49.0
Double (5%)	27.3	**50.7**	**50.7**	45.6

Table III speaks for the importance of the whitening step (11). With partial phasor angle measurements, performance of the three (near-) optimal algorithms degrades relative to the ones in Test Case 1. This is expected because observability for the 118-bus system is compromised if only angle measurements from buses in $\mathcal{N}_I$ are available. It is also worth stressing that the proposed schemes, especially the OMP Algorithm 1, still perform comparably to the benchmark scheme based on (22).

To further illustrate the performance of the proposed algorithms, the lines in outage were considered drawn from the set $\tilde{\mathcal{E}} := \{(4,11),(42,49),(65,68)\}$, corresponding to lines indexed by $\ell = 10, 66, 103$. Notice that two of these three lines are linked to buses in the unobservable external system $\mathcal{N}_E$ at both sides. The noise level is set equal to 2% of the average pre-event power injection. Algorithm 1 is simulated with the maximum $\kappa_{\max} := 6$, and outputs the sequence of identified lines in $\mathcal{L}^6 = \{66, 10, 103, 42, 110, 34\}$. Clearly, the set $\tilde{\mathcal{E}}$ is identified exactly from the first three elements of $\mathcal{L}^6$. Subsequently, Algorithm 2 is applied to obtain the line outage identification path for $\hat{\mathbf{s}}^\lambda$ with an exponentially decreasing sequence of 20 values of λ.

When L_o is unknown, the tests based on the MDL and the sample variance deviation criterion are implemented with the number of outaged lines ranging from 1 to 5. The test scores are obtained by scaling relative to their respective maximum ones, and are plotted in Fig. 1(a). The plot demonstrates that the minimum scores are achieved at $|\hat{\mathcal{E}}| = 3$. Clearly, this is also the right κ for the OMP, which confirms that $\mathcal{L}^\kappa$ correctly identifies $\tilde{\mathcal{E}}$. For Algorithm 2, Fig. 1(b) depicts the absolute value of the $\hat{\mathbf{s}}^\lambda$ entries, corresponding to a λ value that yields three line outages. All entries in $\hat{\mathbf{s}}^\lambda$ are numerically equal to zero, except for three positions, coinciding exactly with lines in $\tilde{\mathcal{E}}$. This clearly demonstrates the effectiveness of the proposed algorithms.

4) Test Case 4: To gauge computational complexity of the proposed algorithms, their running times are provided for the 118-, 300-, and 2,383-bus systems. While the first two are the aforementioned IEEE benchmark systems from [14], the last one is from the Polish power system provided by MATPOWER "case2383wp" casefile. All algorithms are run using the MATLAB R2011a software, on a typical Windows XP computer with a 2.8-GHz CPU. The running times for both single and double line outages averaged over 500 Monte Carlo iterations are listed in Table IV. The OMP Algorithm 1 clearly outperforms the rest in running time, and scales well with the number of buses. For both Algorithms 1 and 2, running time for

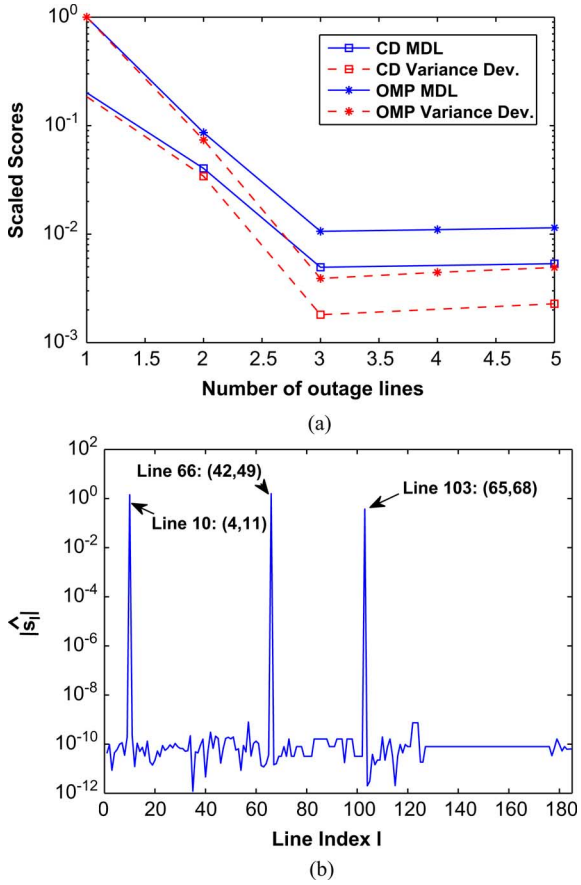


Fig. 1. (a) Scaled scores versus a variable L_o for MDL and variance tests. (b) Absolute value of entries of $\hat{s}_\ell^\lambda$ versus the index ℓ, for the λ value yielding three line outages.

TABLE IV
AVERAGE RUNNING TIMES IN SECONDS

	ES	Alg. 1	Alg. 2
Single, 118-bus	5.9e-2	1.3e-4	4.9e-2
Double, 118-bus	4.1	2.7e-4	9.7e-2
Single, 300-bus	0.15	4.0e-4	0.39
Double, 300-bus	22.5	1.1e-3	0.97
Single, 2383-bus	1.2	1.4e-2	2.6
Double, 2383-bus	$\sim$	2.9e-2	4.2

double line outages is about twice that for single line outages.[3] However, the combinatorially complex ES approach does not scale well with the system size and the number of line outages, resulting in hundreds of times more computational time for double line-outage tests with the 118- or the 300-bus system. As this increase grows to be in the order of thousands for the 2383-bus system, it is omitted here.

VI. GENERALIZING THE OVERCOMPLETE REPRESENTATION

This section generalizes the novel notion of sparse overcomplete representation for identifying line outages, along with Algorithms 1 and 2, to three directions of practical interest.

[3]Although Algorithm 2 may not appear as attractive, there are tailored programs in [6] to compute the entire identification path within 1 second for systems involving as many as 5000 unknowns.

A. Incorporating Information on Internal Noise

It is often the case that the internal system in $\mathcal{N}_I$ can acquire $\mathbf{p}_I$ in real time, using PMU data and state estimates obtained from legacy meters; see, e.g., [5]. As the noise vector $\boldsymbol{\eta}$ accounts for the perturbations present in the post-event power vector relative to the pre-event power vector, the internal system is able to recover exactly the part $\boldsymbol{\eta}_I$ of this noise vector, since it can measure both $\mathbf{p}_I$ and $\mathbf{p}'_I$ at the buses in $\mathcal{N}_I$. Upon forming $\boldsymbol{\eta}_I := \mathbf{p}'_I - \mathbf{p}_I$, this internal noise vector can be removed from both sides of (8). After this "internal noise" cancellation, $\boldsymbol{\eta}$ is no longer white, and likewise $\mathbf{V}_I^T\boldsymbol{\eta}$ in (12) is colored too. While the internal noise is absent ($\boldsymbol{\eta}_I = \mathbf{0}$), the external one ($\boldsymbol{\eta}_E$) is still zero-mean white with covariance $\sigma_\eta^2\mathbf{I}$.

To account for the absence of internal noise, let $[\mathbf{B}^{-1}]_{IE}$ denote the submatrix of $\mathbf{B}^{-1}$ formed by rows and columns from buses in $\mathcal{N}_I$ and $\mathcal{N}_E$, respectively. Using the latter, consider replacing the transformed noise $[\mathbf{B}^{-1}]_I\boldsymbol{\eta}$ in (9) with $[\mathbf{B}^{-1}]_{IE}\boldsymbol{\eta}_E$. It now becomes possible to employ a similar whitening procedure using the SVD $[\mathbf{B}^{-1}]_{IE} = \mathbf{U}_{IE}\boldsymbol{\Sigma}_{IE}\mathbf{V}_{IE}^T$ to obtain the data vector

$$\check{\mathbf{y}} := \boldsymbol{\Sigma}_{IE}^{-1}\mathbf{U}_{IE}^T\tilde{\boldsymbol{\theta}}_I = \boldsymbol{\Sigma}_{IE}^{-1}\mathbf{U}_{IE}^T[\mathbf{B}^{-1}]_I\mathbf{M}\mathbf{s} + \mathbf{V}_{IE}^T\boldsymbol{\eta}_E. \quad (23)$$

Apart from a different regression matrix, the modified data model (23) entails again a sparse overcomplete representation, and thus Algorithms 1 and 2 can be readily applied. It is expected that, by incorporating the extra information on the internal system, noise will markedly enhance accuracy without sacrificing computational efficiency.

B. From Line Outages to Line Changes

A second extension of interest is to allow for more general *line faults* including transmission line parameter changes. As suggested by the expression of $\tilde{\mathbf{B}}$ in (6), changes in the inverse reactance $\{1/x_\ell\}$ of lines in outage manifest themselves in changes of the difference matrix $\tilde{\mathbf{B}}$ in the pre- and post-event network Laplacians. Suppose that instead of having only line outages, a subset of lines experiences change in reactance parameters, giving rise to a modified set of lines $\tilde{\mathcal{E}}_{\text{mod}}$. Clearly, $\tilde{\mathcal{E}}_{\text{mod}}$ can also capture line outages which correspond to line reactance parameters going to infinity. Suppose that the ℓth line experiences a sudden change in its reactance from x_ℓ to x'_ℓ, for all $\ell \in \tilde{\mathcal{E}}$. This can be also captured by the matrix

$$\tilde{\mathbf{B}} := \mathbf{B} - \mathbf{B}' = \sum_{\ell \in \tilde{\mathcal{E}}_{\text{mod}}} \delta_\ell \mathbf{m}_\ell \mathbf{m}_\ell^T \quad (24)$$

where the difference in the inverse reactance $\delta_\ell := 1/x_\ell - 1/x'_\ell$.

Clearly, a line outage is a special type of a line change, since, if the ℓth line is in outage, it holds that $x'_\ell \to \infty$, and thus $\delta_\ell = 1/x_\ell$. However, it is worth noticing that the system model (12) still holds with $s_\ell := \delta_\ell \mathbf{m}_\ell^T\boldsymbol{\theta}'$ for $\ell \in \tilde{\mathcal{E}}_{\text{mod}}$. Thus, the general line change identification problem can again be solved under the sparse signal reconstruction framework [cf. (16)].

C. Decoupled Power Flow Model

The DC power flow model discussed so far actually coincides with the system obtained by linearizing the decoupled *nonlinear* ac power flow model obtained after the appropriate parameter simplifications; see, e.g., [1, Sec. 10.7]. To handle nonlinear

power flows, iterative approximations are typically employed. Among the most popular approximation methods, the fast decoupled power flow model is preferred as an efficient linearization method as it only needs four to seven iterations to converge in practice; see e.g., [1, Sec. 10.7]. Compared with the DC power flow involving phasor angles and real power injection, the fast decoupled model includes an additional system of equations, relating the phasor amplitudes to the reactive power injection. Interestingly, both transformation matrices in the fast decoupled model are also related to the Laplacian $\mathbf{B}$, and thus they can be used to infer changes in the grid topology.

VII. Conclusion and Current Research

Novel line-change identification algorithms were developed in this paper through overcomplete representations capturing the sparse innovation manifested by line changes on PMU measurements. The proposed approach allows for identification of multiple (even cascaded) power line outages using two schemes with complementary strengths. The first one entails successive (greedy) inclusion of the lines in outage and is computationally efficient. The second scheme regularizes the LS criterion with the ℓ_1-norm of the sparse vector comprising the overcomplete representation coefficients, thus enabling a coordinate descent based solver to compute the entire identification path. Practical statistical tests were also adapted to identify the actual number of line outages. The framework developed for linear DC power flow models was shown capable to accommodate available internal system noise information and cope with generic line changes. Numerical tests demonstrated the merits of the proposed schemes in unveiling multiple lines in outage, even when those reside at the external system for which phasor angle data are unavailable.

Future research directions include applying the sparse overcomplete representation to enhance the current power system contingency analysis from the "$(N-1)$ rule" to a general "$(N-L_o)$ rule," where L_o stands for a multiple line-outage contingency case.

References

[1] A. R. Bergen and V. Vittal, *Power System Analysis*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 2000.

[2] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ., 2004.

[3] E. J. Candès 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.

[4] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comp.*, vol. 20, pp. 33–61, Jan. 1998.

[5] R. Emami and A. Abur, "Tracking changes in the external network model," in *Proc. 42nd North Amer. Power Symp.*, Sep. 26–28, 2010, pp. 1–6.

[6] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *J. Stat. Software*, vol. 33, no. 1, pp. 1–22, Feb. 2010.

[7] D. Gorinevsky, S. Boyd, and S. Poll, "Estimation of faults in DC electrical power system," in *Proc. Amer. Control Conf.*, Jun. 10–12, 2009, pp. 4334–4339.

[8] T. Güler, G. Gross, and M. Liu, "Generalized line outage distribution factors," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 879–881, May 2007.

[9] J. Guo, Y. Fu, Z. Li, and M. Shahidehpour, "Direct calculation of line outage distribution factors," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1633–1634, Aug. 2009.

[10] M. H. Hansen and B. Yu, "Model selection and the principle of minimum description length," *J. Amer. Stat. Assoc.*, vol. 96, no. 454, pp. 746–774, Jun. 2001.

[11] 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.

[12] R. J. Kaye and F. F. Wu, "Analysis of linearized decoupled power flow approximations for steady-state security assessment," *IEEE Trans. Circuits Syst.*, vol. 31, no. 7, pp. 623–636, Jul. 1984.

[13] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[14] "Power systems test case archive," Univ. Washington, Seattle. [Online]. Available: http://www.ee.washington.edu/research/pstca/

[15] A. Schellenberg, W. Rosehart, and J. Aguado, "Cumulant-based probabilistic optimal power flow (P-OPF) with Gaussian and Gamma distributions," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 773–781, May 2005.

[16] B. Stott, J. Jardim, and O. Alsac, "DC power flow revisited," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1290–1300, Aug. 2009.

[17] J. E. Tate and T. J. Overbye, "Line outage detection using phasor angle measurements," *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1644–1652, Nov. 2008.

[18] J. E. Tate and T. J. Overbye, "Double line outage detection using phasor angle measurements," in *Proc. IEEE PES Gen. Meeting*, Jul. 26–30, 2009, pp. 1–5.

[19] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *J. Royal Stat. Soc. B*, vol. 58, pp. 267–288, 1996.

[20] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[21] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *J. Optim. Theory Applic.*, vol. 109, no. 3, pp. 475–494, Jun. 2001.

[22] "U.S.–Canada power system outage task force, final report on the August 14th blackout in the United States and Canada," Dept. Energy, Apr. 2004 [Online]. Available: https://reports.energy.gov/

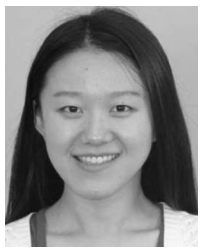
[23] "Final report on the implementation of the task force recommendation," U.S.–Canada Power System Outage Task Force, Sep. 2006. [Online]. Available: http://energy.gov/sites/prod/files/oeprod/DocumentsandMedia/Outage_Task_Force_-_DRAFT_Report_on_Implementation.pdf

[24] D. B. West, *Introduction to Graph Theory*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 2001.

[25] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*, 2nd ed. New York: Wiley, 1996.

[26] 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.

[27] H. Zhu and G. B. Giannakis, "Lassoing line outages in the smart power grid," in *Proc. 2nd Int. Conf. Smart Grid Commun.*, Brussels, Belgium, Oct. 17–20, 2011.



Hao Zhu (S'07) received the B.Sc. degree from Tsinghua University, Tsinghua, China, in 2006, and the M.Sc. degree from the University of Minnesota (UMN), Minneapolis, in 2009, both in electrical engineering. She is currently working toward the Ph.D. degree at the Department of Electrical and Computer Engineering, UMN.

She was a Summer Research Intern with NEC Labs America, Princeton, NJ, in 2011. Her research interests lie in compressive sampling, power system monitoring, the smart grid, and social network analysis.

Ms. Zhu received the two-year UMN Graduate School Fellowship in 2006, and the UMN Doctoral Dissertation Fellowship in 2011.

Georgios B. Giannakis (F'97) received the Diploma in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1981, and the M.Sc. degree in mathematics and M.Sc. and Ph.D. degrees in electrical engineering from the University of Southern California, Los Angeles, in 1983 and 1986, respectively.

Since 1999, he has been a Professor with the University of Minnesota, Minneapolis, where he now holds an ADC Chair in Wireless Telecommunications with the Electrical and Computer Engineering Department and serves as Director of the Digital Technology Center. His general interests span the areas of communications, networking and statistical signal processing—subjects on which he has authored or coauthored more than 325 journal papers, 525 conference papers, 20 book chapters, two edited books, and two research monographs. He is the inventor and co-inventor of 21 patents. His current research focuses on compressive sampling, cognitive radios, cross-layer designs, wireless sensors, social, gene-regulatory, and power grid networks.

Prof. Giannakis is a Fellow of EURASIP He is the (co-) recipient of eight paper awards from the IEEE Signal Processing (SP) and Communications Societies, including the G. Marconi Prize Paper Award in Wireless Communications. He also received Technical Achievement Awards from the SP Society (2000), from EURASIP (2005), a Young Faculty Teaching Award, and the G. W. Taylor Award for Distinguished Research from the University of Minnesota. He has served the IEEE in a number of posts, including that of a Distinguished Lecturer for the IEEE SP Society.