

Sparse Estimation of Faults by Compressed Sensing With Structural Constraints

Igal Rozenberg, *Student Member, IEEE*, Yuval Beck^{ID}, *Member, IEEE*, Yonina C. Eldar^{ID}, *Fellow, IEEE*, and Yoash Levron^{ID}, *Member, IEEE*

Abstract—This paper addresses the challenge of fault location in large power networks using a limited number of sensors. It was recently shown that power system faults may be modeled by sparse vectors, and hence, can be located efficiently using sparse recovery techniques. In this paper, we extend this approach and propose a sparse recovery algorithm that exploits both the sparsity constraints and additional structural constraints imposed by the faults physical models. To this end, faults are represented by sparse vectors that are subjected to nonconvex constraints. These constraints are shown to provide additional information that is exploited to reduce the number of measurements and to improve the location accuracy. The algorithm searches directly over these physical faults, and therefore, operates over a small solution space. Simulations on the IEEE 118 bus test-case network show that 4 to 20 sensors are sufficient to recover faults with adequate accuracy. With 20 PMU sensors, more than 99% of single-fault events and 84%–95% of two-fault events are located, depending on fault types and SNR.

Index Terms—Compressed sensing, fault location, sparse estimation, sparse representations, state estimation, wide area fault location.

I. INTRODUCTION

POWER system estimation problems often result in systems of equations that are underdetermined, due to a limited number of available sensors [1], [2]. To address the challenge of finding a meaningful estimator in such problems, several recent works exploit the sparse structure of power system events to estimate them using a limited number of sensors. Such estimation problems have been recently studied using sparse representations and compressed sensing techniques [3]–[5]. For example, [6] uses phasor measurements for representing AC power system models as linearly constrained least-squares problems. Work [7] solves the blind topology identification problem for power systems using power injection data only, at each bus. In addition, [8] suggests using complementary slackness to detect the structure from locational marginal prices, and [9] uses a graphical

Manuscript received June 11, 2017; revised October 30, 2017 and February 6, 2018; accepted March 28, 2018. Date of publication April 6, 2018; date of current version October 18, 2018. The work of Y. Levron was supported in part by the Grand Technion Energy Program. Paper no. TPWRS-00826-2017. (Corresponding author: Yoash Levron.)

I. Rozenberg Y. C. Eldar, and Y. Levron are with the Andrew and Erna Viterbi Faculty of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 3200003, Israel (e-mail: igalroz@tx.technion.ac.il; YOASHLEVRON@gmail.com; yonina@ee.technion.ac.il).

Y. Beck is with the Electrical Engineering Faculty, Holon Institute of Technology, Holon 58102, Israel (e-mail: beck@hit.ac.il).

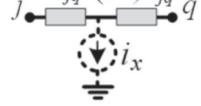
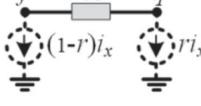
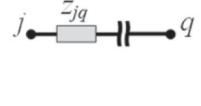
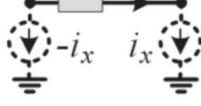
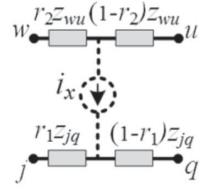
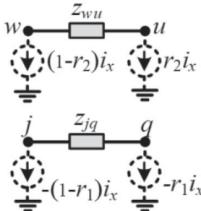
Digital Object Identifier 10.1109/TPWRS.2018.2823734

method to build the probabilistic relationships among different voltage measurements. The topology reconstruction problem is formulated as a regularized linear regression problem. A sparse recovery method for locating sources of harmonic distortion and polluting loads using ℓ_1 relaxation is proposed in [10]. Similar methods are utilized for data cleansing and detection of outliers in power system measurements [11], as well as for locating erroneous measurements [12].

Sparse methods may also reduce the complexity of non-linear state estimation problems by projecting the measurements into a linear space of lower dimension, thus exploiting the underlying sparse structure [13]. Evaluation of power system distribution factors is addressed in [14]. While this problem is not naturally sparse, it is shown that certain sets of buses share similar sensitivities, which allows a reduction of the solution space and hence utilization of sparse techniques. In [15], a sparsifying linear transform is utilized for full network state estimation; this approach relies on the small differences in voltages between adjacent power network buses in normal operation mode. Reference [16] suggests to estimate the network topology using a DC power-flow model. Samples of voltage phases over time on different buses allow recovery of non-zero entries in the sparse nodal admittance matrix by applying greedy sparse recovery techniques.

In addition, sparse representation methods are utilized for locating power system faults, such as hardware malfunctions, short circuits, or disconnected lines. Traditional approaches to fault location [17]–[27] typically assume, sometimes implicitly, that faults are rare and hence enable their location using few measurements. This concept is extended by modeling faults as sparse events, and locating them using sparse recovery algorithms. For instance, in [28] line outages are located based on a DC power model and a Laplacian decomposition of the line reactance matrix. The vector associated with disconnected lines is assumed sparse, and is recovered by compressed sensing techniques such as Lasso or Orthogonal-Matching Pursuit (OMP). It is suggested to improve the performance of the aforementioned technique by applying an additional transform on the sensing matrix, thus reducing its average coherence [29]. A similar approach for locating line outages is proposed in [30], where a technique based on cross-entropy stochastic optimization is utilized for sparse recovery. Sparse modeling of various network events such as faults, lightning strikes and power thefts is presented in [31]. A sparse modeling of network events and utilization of the Lasso method is discussed

TABLE I
MODELING FAULTS BY SPARSE VECTORS USING NETWORK TRANSFORMATIONS

Fault Type	Model	Equivalent Model	Elementary Sparse Vector
line-to-ground (LG)			$I^{sg,\nu}[j] = i_j = (1-r)i_x$ $I^{sg,\nu}[q] = i_q = ri_x$ i_x is complex r is real, $0 \leq r \leq 1$, otherwise $I^{sg,\nu}[n] = 0$
disconnected line (DL)			$I^{dl,\xi}[j] = i_j = -i_x$ $I^{dl,\xi}[q] = i_q = i_x$ i_x is complex otherwise $I^{dl,\xi}[n] = 0$
line-to-line short (LL)			$I^{ll,\gamma}[j] = i_j = -(1-r_1)i_x$ $I^{ll,\gamma}[q] = i_q = -r_1 i_x$ $I^{ll,\gamma}[w] = i_w = (1-r_2)i_x$ $I^{ll,\gamma}[u] = i_u = r_2 i_x$ i_x is complex, r_1, r_2 are real $0 \leq r_1, r_2 \leq 1$, otherwise $I^{ll,\gamma}[n] = 0$

in [32], [33]. A decentralized line outage detection that utilizes the Lasso is proposed [34]. In [35] sparse currents represent simultaneous events and are estimated using a Lars-Lasso ℓ_1 relaxation. In addition, the technique of [12] is implemented to detect erroneous measurements. This idea is further extended in [36] which also uses group Lasso to promote sparsity in pairs. This approach accounts for the fact that faults are represented by injected currents that appear in pairs, and may be overlapping.

This paper proposes a sparse recovery algorithm for locating power system faults, which highlights the information provided by the fault physical models. Three types of faults are analyzed: line-to-ground shorts, disconnected lines, and line-to-line shorts. Faults are represented by sparse vectors subjected to non-convex constraints. These constraints are shown to provide additional information, beyond the sparse structure, that is exploited by the proposed algorithm to reduce the number of measurements required for location. The algorithm employs principles of the Orthogonal-Matching-Pursuit (OMP) method for sparse recovery, and incorporates additional information provided by the non-linear constraints imposed on the solution vectors. As such, the algorithm utilizes the geometry of the power system, and only tests solutions that comply with the fault physical models. Simulations show that this additional information enables fault location using reduced number of sensors (4 to 20 Phasor Measurement Unit (PMU) sensors in the IEEE 118 bus test-case network). The algorithm is based on the positive sequence measurement only and employs per-phase data. A simple analysis and comparison between the various phases insures identification of faults of several types.

The paper is organized as follows: Sections II and III formulates fault recovery as a constrained sparse recovery problem. Section IV introduces a method for recovery of fault events, utilizing their special structure. Performance of the method is evaluated in Section V for various fault scenarios.

II. REPRESENTING FAULTS BY SPARSE VECTORS

Common faults in power systems are assumed to be of several types: line-to-ground short circuit, line-to-line short circuit, or disconnected lines. Each such fault is represented in this paper by equivalent current sources on the buses adjacent to the faulted line. The solution vector is then described by a sum of elementary sparse vectors, corresponding to these equivalent currents. This formulation leads to a sparse representation with additional non-convex equality constraints.

Assume a power network with N buses, M measurements, and a nodal impedance matrix Z_{bus} , such that $V = Z_{\text{bus}}I_{\text{bus}}$, where V is the vector of complex bus voltages, and I_{bus} is the vector of complex injected bus currents. The vector $Y_{M \times 1}$ defines the measurements, which may be the output of either voltage or current sensors. These measurements depend on the sensor locations and the network topology, and have linear dependency on the bus voltages. Consequently, the relation between the bus voltages and the measurements is defined by a matrix $B_{M \times N}$, such that $Y = BV$. These definitions lead to the following linear system of equations:

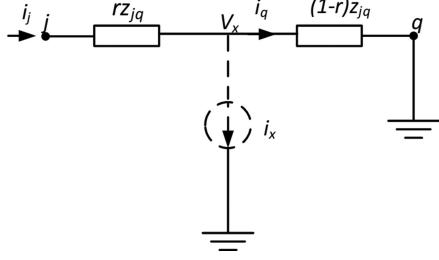
$$Y_{M \times 1} = A_{M \times N}(I_{\text{bus}})_{N \times 1} + \epsilon^y_{M \times 1} A = BZ_{\text{bus}}, \quad (1)$$

where A is the sensing matrix, and ϵ^y is an additive measurement noise. This noise can be attributed to measurement errors, small fluctuations in the steady-state behavior, and additional model uncertainties. The measurements in (1) are sampled before and after the fault (as proposed in [20], [21], [36]) resulting in the following vectors:

$$Y^a = AI^a + \epsilon^a, \quad Y^b = AI^b + \epsilon^b. \quad (2)$$

Subtraction of these equations leads to

$$Y^d = Y^b - Y^a = A(I^b - I^a) + (\epsilon^b - \epsilon^a) = AI^d + \epsilon, \quad (3)$$

Fig. 1. Equivalent circuit for calculating the injected current at bus j .

where Y^d is the difference in measured values, and $\epsilon = \epsilon^b - \epsilon^a$ is the resulting noise. The vector I^d in (3) has a specific structure that is determined by the fault types and locations (as shown in Table I).

Table I summarizes the various fault models. The models are obtained using network transformations and superposition, and consequently formulated as sparse vectors. For example, the first row of the table shows the derivation of the equivalent model for a line-to-ground (LG) fault. The circuit model describes a line between the j and q buses with an impedance of z_{jq} . The current source i_x represents the fault current at distance r from bus j , where both r and i_x are unknown. Next, the equivalent model shown in the third column is derived by superposition. The injected current at bus j is found by shortening V_q to ground, as shown in Fig. 1.

The voltage drop V_x on the current source i_x can be expressed as:

$$V_x = (i_j - i_x)(1 - r)z_{jq}, \quad (4)$$

and the current injected into bus j is now,

$$i_j = \frac{V_j - V_x}{r z_{jq}}. \quad (5)$$

Substituting (4) into (5) yields

$$i_j = \frac{V_j}{z_{jq}} + i_x(1 - r). \quad (6)$$

Equation (6) implies that the injected current at bus j is the sum of a current source with a value of $i_x(1 - r)$ and a current that flows through the line impedance z_{jq} . Repeating the same procedure for calculating the current injected to bus q yields

$$i_q = \frac{V_j}{z_{jq}} + i_x r. \quad (7)$$

The results of (6) and (7) provide the equivalent circuit as appears in Table I. The same procedure is applied to attain the other equivalent circuits.

Using the above definitions, the vector I^d is expressed by three sums, each corresponding to a possible type of fault:

$$I^d = \sum_{\nu=1}^{k_{sg}} I^{sg,\nu} + \sum_{\xi=1}^{k_{dl}} I^{dl,\xi} + \sum_{\gamma=1}^{k_{ll}} I^{ll,\gamma}, \quad (8)$$

where k_{sg} is the number of line-to-ground faults, k_{dl} is the number of disconnected lines, and k_{ll} is the number of

TABLE II
MEASUREMENTS CORRESPONDING TO A SINGLE FAULT IN EXPLICIT FORM

Fault Type	Γ	I_Γ	R_Γ	$Y^d \equiv h(I_\Gamma, R_\Gamma)$
LG	(j, q)	i_x	(r_1)	$((1 - r_1)A_j + r_1A_q)i_x$ $(A_j$ is the j^{th} column of A)
DL	(j, q)	i_x	-	$(-A_j + A_q)i_x$
LL	(j, q, w, u)	i_x	$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix}$	$\begin{aligned} & ((-1 - r_1)A_w - r_1A_u) + \\ & ((1 - r_2)A_j + r_2A_q)i_x \end{aligned}$

line-to-line shorts. The resulting vectors $I^{sg,\nu}$, $I^{dl,\xi}$, $I^{ll,\gamma}$ are sparse vectors that represent the equivalent injected currents.

III. PROBLEM FORMULATION

Classic methods for analyzing unbalanced faults are based on the well-known symmetrical components, and require data of the positive, negative, and zero sequences. The method presented in this paper requires only measurements data of the positive sequence since the main objective is to locate the fault and its type. It is evident from Table I that a per-phase analysis can indicate the type of fault by simply comparing the results from the various phases at the same measurement point. For example, a LG fault can be easily identified by comparing the current magnitudes of the various phases. In this scenario, injected currents with equal phase are expected on the two buses adjacent to the faulted line. In contrast, a DL fault results in inverse currents in the two adjacent buses. Therefore, although LG and DL faults on the same line may seem similar, they are identifiable by the aforementioned signs. In the same manner a simple analysis of the measurements in the case of LL fault can distinguish in a given R, S, T three phase system whether the fault occurred between RS, ST or TR lines. These simple rules are applicable for determining the various types of faults as well as detecting the faulted phase or phases.

With zero measurement noise ($\epsilon = 0$) the fault location problem is cast as a linear system of equations with additional constraints. For fixed values of k_{sg} , k_{dl} , k_{ll} , the set of candidate solution vectors is defined by:

$$\Psi_{k_{sg}, k_{dl}, k_{ll}} = \left\{ I^d \mid I^{sg,\nu}, I^{dl,\xi}, I^{ll,\gamma} \text{ are structured according to Table I} \right\}, \quad (9)$$

and the resulting fault location problem is to find a vector I_d such that

$$Y^d = AI^d \quad \text{and} \quad I^d \in \Psi_{k_{sg}, k_{dl}, k_{ll}}. \quad (10)$$

Due to the structure of the elementary vectors in Table I, the solution vector I^d is sparse, with a maximal number of $k = 2k_{sg} + 2k_{dl} + 4k_{ll}$ non-zero elements. The measurements for each specific fault type are given in Table II.

As an example, assume a single line-to-line short such that $(k_{sg}, k_{dl}, k_{ll}) = (0, 0, 1)$. The system of equations (6) then

becomes

$$\begin{aligned} Y^d &= A_{M \times N} I^d \quad \text{such that} \\ I^d &= (0, \dots, 0, i_j, 0, \dots, 0, i_w, 0, \dots, 0, i_u, 0, \dots, 0)^T, \\ i_j &= -(1 - r_1)i_x, \quad i_q = -r_1 i_x, \\ i_w &= (1 - r_2)i_x, \quad i_u = r_2 i_x, \end{aligned} \quad (11)$$

where i_x is complex, r_1, r_2 are real, $0 \leq r_1 \leq 1, 0 \leq r_2 \leq 1$. In this example neither the location of the non-zero elements in I^d nor their values are known. Moreover, the indices of the active entries in the vector should correspond to a legitimate fault, while their actual value must satisfy the structural constraints specified in Table I.

In the presence of noise ($\epsilon > 0$) the problem is cast as a Maximum-Likelihood (ML) estimation problem. It is assumed that ϵ is a Gaussian random vector with independent elements with zero mean and variance σ^2 . Following this assumption, the Maximum-Likelihood estimator is

$$I^{d, \text{opt}} = \arg \min_{I^d \in \Psi_{k_{\text{sg}}, k_{\text{dl}}, k_{11}}} \|Y^d - AI^d\|_2^2. \quad (12)$$

The challenge is to solve this minimization problem (12) with the constraint $I^d \in \Psi_{k_{\text{sg}}, k_{\text{dl}}, k_{11}}$ as in (10). To this end, the solution I^d consists of a sum of sparse vectors, each one admitting a specific structure as detailed in Table I.

IV. LINEAR STRUCTURED MATCHING PURSUIT ALGORITHM

This section presents a greedy fault location algorithm (StructMP), whose objective is to find the faults and their location using few measurements, by efficiently solving (12). The algorithm utilizes both the sparsity of the solution vector and the structural constraints in Table I.

Although the linear system of equations $Y^d = AI^d$ is under-determined ($M < N$), a unique sparse solution may still exist. Recent papers (e.g. [10], [28], [32]) solve such systems of equations by means of classic compressed sensing methods, such as greedy algorithms or ℓ_1 relaxation. These methods aim to approximate the sparse estimator for the problem $Y^d = AI^d + \epsilon$, which is given by:

$$I^{d, \text{opt}} = \arg \min_{I^d \in k-\text{sparse}} \|Y^d - AI^d\|^2. \quad (13)$$

However, with this approach, there is no guarantee that the solution of (13) complies with the fault structure, as defined in (9). To utilize this additional information, we propose an algorithm that incorporates both the sparsity constraints, and these additional structural constraints. Therefore, the proposed algorithm does not search for those columns in the sensing matrix A that best correlate with the residual, but scans directly over the faults defined in Table I. As a result, the algorithm searches over a smaller solution space that complies with the non-convex constraints.

A. Algorithm Overview and a Basic Example

This section presents a basic example, in which the algorithm is applied to find the location of two simultaneous faults. Assume

a network with two simultaneous line to ground shorts. The measurement associated with each short is given in Table II above, and the sensor measurement Y^d is given by their sum:

$$\begin{aligned} Y^d &= ((1 - r_1)A_{j1} + r_1 A_{q1})i_1 \\ &\quad + ((1 - r_2)A_{j2} + r_2 A_{q2})i_2 + \epsilon, \end{aligned} \quad (14)$$

where (j_1, q_1) and (j_2, q_2) are the faulted line indices and r_1, r_2 are the relative location of the shorts within each line respectively. The short currents are given by i_1, i_2 , respectively.

The algorithm chooses the shorted lines in an iterative manner, one fault in each step. In the first step, the algorithm is required to select two buses (j_1^*, q_1^*) corresponding to the faulted line. These are chosen to minimize the norm error:

$$(j_1^*, q_1^*) = \arg \min_{j, q} \min_{\substack{i_1 \\ 0 \leq r_1 \leq 1}} \|Y^d - ((1 - r_1)A_j + r_1 A_q)i_1\|. \quad (15)$$

This bi-level optimization problem solves for the estimated faulted line indices. In the inner minimization, the faulted line is fixed and an optimization is carried out with respect to the fault parameters, namely, the short current and relative location. The minimization is performed by alternating minimization as explained in Section IV-B. The outer minimization problem sweeps over legitimate index pairs (j, q) that correspond to actual transmission lines in the network. The candidate short which achieves the minimal error is selected.

In the next iteration, the algorithm selects a second pair of buses, corresponding to the other short. To do so, the algorithm searches for a second pair (j_2^*, q_2^*) that together with the previously selected pair, minimizes the norm error:

$$\begin{aligned} (j_2^*, q_2^*) &= \arg \min_{j, q} \min_{\substack{i_1, i_2 \\ 0 \leq r_1, r_2 \leq 1}} \|Y^d - ((1 - r_1)A_{j1} + r_1 A_{q1})i_1 \\ &\quad + ((1 - r_2)A_j + r_2 A_q)i_2\|. \end{aligned} \quad (16)$$

Here indices of the first estimated fault (j_1^*, q_1^*) are not modified and remain fixed. Yet, the parameters of all faults (short currents and relative location) are optimized by alternating minimization to achieve a minimal value. Consequently, the algorithm chooses a second pair associated with a shorted transmission line. The resulting short location estimate is: the two pairs $(j_1^*, q_1^*), (j_2^*, q_2^*)$ and their corresponding relative location r_1, r_2 .

B. General Description of the Fault Location Algorithm

The purpose of the algorithm is to locate k simultaneous faults of the types specified in Table I. The algorithm operates in steps, where in each step l one additional fault is selected.

To this end, two vectors I and R are defined. These vectors are constructed by concatenation of the fault parameters (Table II) up to the step l . In addition, the buses associated with the fault l are defined by the set Γ_l . The entries in the vectors I, R that are associated with Γ_l are designated by $I_{\Gamma_l}, R_{\Gamma_l}$. Each fault l is thus characterized by its corresponding bus set Γ_l and these two parameter sets $I_{\Gamma_l}, R_{\Gamma_l}$.

At each step l the algorithm selects a new fault, and the parameters of all previously selected faults, to minimize the

norm error:

$$\begin{aligned} \Gamma_l = \arg \min_{\Gamma_l} \min_{I, R} & \left\| Y^d - (h_l(I_{\Gamma_l}, R_{\Gamma_l}) \right. \\ & \left. + \sum_{p=1}^{l-1} h_p(I_{\Gamma_p}, R_{\Gamma_p})) \right\|, \\ \text{s.t. } & 0 \leq R_i \leq 1 \quad \forall i, \end{aligned} \quad (17)$$

where $h_p(I_{\Gamma_p}, R_{\Gamma_p})$ is the elementary sensor response associated with a fault, as defined in Table II. It is noted that entries of I associated with the inner optimization are complex, and the corresponding entries of R are real and takes value in the range $[0, 1]$. Also, the previously selected bus sets $\Gamma_p, p = 1, \dots, l-1$ are fixed in the l th optimization steps and are not reselected.

The fault selection rule (17) involves a bi-level optimization, comprising of an outer minimization with respect to the fault buses Γ_l , and an inner minimization with respect to the fault currents I and relative locations R . The outer minimization is performed by scanning over sets Γ_l that correspond to possible faults, as defined by the network topology. In most power networks, the number of such candidate sets is of order $O(N)$, since buses are typically connected only to a small group of neighboring buses.

The challenge in solving the inner minimization problem is that the objective function is non-convex with respect to I and R . This can be verified for instance in the simple example of (11). To address this problem, the algorithm approximates the solution by an alternate minimization scheme, operating over I and R . This approach utilizes the fact that the response functions $h_p(I_{\Gamma_p}, R_{\Gamma_p})$ are bi-linear. Due to this property, optimization with respect to either one of the vectors I or R , while fixing the other vector, results in a quadratic and convex least-squares problem. Thus, when R is fixed,

$$I = \arg \min_I \|Y^d - H_R I\|^2. \quad (18)$$

Similarly, when I is fixed,

$$\begin{aligned} r = \arg \min_R & \|Y' - H_I R\|^2 \\ \text{s.t. } & 0 \leq R_i \leq 1 \quad \forall i, \end{aligned} \quad (19)$$

where the matrices H_R , H_I and the vector Y' are given in Table III. Equation (19) is reformulated as

$$\begin{aligned} R = \arg \min_R & \left\| \begin{bmatrix} \Re\{Y'\} \\ \Im\{Y'\} \end{bmatrix} - \begin{bmatrix} \Re\{H_I\} \\ \Im\{H_I\} \end{bmatrix} R \right\|^2 \\ \text{s.t. } & 0 \leq R_i \leq 1 \quad \forall i. \end{aligned} \quad (20)$$

The optimization problems (18) and (20) are convex and can be efficiently solved. The algorithm alternates in each iteration between the two. Since in each such iteration the norm decreases, the algorithm is guaranteed to converge to a local minimum. This process is repeated for each set Γ_l , and the set for which the approximated objective achieves the minimal value is selected. The alternation procedure begins with (18).

TABLE III
FAULT SENSOR RESPONSES WITH FIXED FAULT CURRENTS OR
FIXED RELATIVE LOCATIONS

With fixed vector of relative locations R			
Fault Type	$H_{R,p}$ size	$H_{R,p}$	
LG	$M \times 1$	$H_{R,p} = (1 - r_p)A_{jp} + r_p A_{qp}$	
DL	$M \times 1$	$H_{R,p} = -A_{jp} + A_{qp}$	
LL	$M \times 1$	$H_{R,p} = -(1 - r_{p1})A_{wp} - r_{p1} A_{up}$ $+ (1 - r_{p2})A_{j\gamma} + r_{p2} A_{qp}$	
With fixed fault current vector I			
Fault Type	$H_{I,p}$ size	W_p size	H_I and W_p
LG	$M \times 1$	$M \times 1$	$H_{I,p} = (A_{qp} - A_{jp})i_p$ $W_p = A_{jp} i_p$
DL	-	$M \times 1$	$H_{I,p} = []$ $W_p = (A_{qp} - A_{jp})i_p$
LL	$M \times 2$	$M \times 1$	$H_{I,p} = [(A_{wp} - A_{up})i_p (A_{qp} - A_{jp})i_p]$ $W_p = (A_{jp} - A_{wp})i_p$

The initial fault location parameters, required to evaluate H_R (Table III), are determined from the previous stage. The relative location parameters of the new candidate fault are initialized to 1/2. The algorithm is outlined in the next section. In practice, the number of faults k_{sg} , k_{dl} , k_{ll} may be unknown. In this case, the fault location algorithm can be applied several times with different number of fault estimates.

C. The Fault Location Algorithm

The proposed algorithm requires the following parameters: $A_{M \times N}$, Y_N^d , (k_{lg}, k_{dl}, k_{ll}) , norm error minimal threshold ϵ_{th} (chosen in the order of the noise level σ_ϵ with a minimum of 10^{-4}), maximal iterations per candidate fault T_{max} , minimal norm reduction per iteration ϵ_{AltMin} (is selected to be 1/10 of ϵ_{th}).

Step 1. Initialization: the algorithm starts by initializing the following parameters: Iteration index $l = 0$, estimated faults count per type of fault, $(k_{sg}^0, k_{dl}^0, k_{ll}^0) = (0, 0, 0)$ and the initial norm error $\epsilon^0 = \|Y^d\|$.

Step 2. Main iteration (lth iteration): while the norm error ϵ^l is larger than a threshold ϵ_{th} , choose an additional fault Γ_l from the possible candidates, according to the conditions in Table IV.

Then solve optimization problem (17) to estimate Γ_l . For each candidate fault Γ_l :

TABLE IV
CONDITIONS FOR INCLUSION IN CANDIDATE FAULTS

Fault Type	Conditions for Inclusion in Candidate Faults
LG	$(k_{\text{sg}}^l < k_{\text{sg}})$ or *
DL	$(k_{\text{dl}}^l < k_{\text{dl}})$ or *
LL	$(k_{\text{ll}}^l < k_{\text{ll}})$ or *
*	$(k_{\text{sg}}^l \geq k_{\text{sg}}$ and $k_{\text{dl}}^l \geq k_{\text{dl}}$ and $k_{\text{ll}}^l \geq k_{\text{ll}})$

TABLE V
UPDATED k -VALUES

Fault Type	Updated k -values
LG	$(k_{\text{sg}}^l, k_{\text{dl}}^l, k_{\text{ll}}^l) \leftarrow (k_{\text{sg}}^{l-1} + 1, k_{\text{dl}}^{l-1}, k_{\text{ll}}^{l-1})$
DL	$(k_{\text{sg}}^l, k_{\text{dl}}^l, k_{\text{ll}}^l) \leftarrow (k_{\text{sg}}^{l-1}, k_{\text{dl}}^{l-1} + 1, k_{\text{ll}}^{l-1})$
LL	$(k_{\text{sg}}^l, k_{\text{dl}}^l, k_{\text{ll}}^l) \leftarrow (k_{\text{sg}}^{l-1}, k_{\text{dl}}^{l-1}, k_{\text{ll}}^{l-1} + 1)$

Step 2.1. Approximate candidate norm error: Define u as the alternate minimization iteration index, and $\epsilon^{l,(u)}(\Gamma_l)$ as the norm error of candidate Γ_l in iteration u . $\epsilon^{l,(u)}(\Gamma_l)$ is the minimal norm error of candidate Γ_l .

Step 2.1.1. Alternate Minimization iterations: While $(\|\epsilon^{l,(u-1)} - \epsilon^{l,(u)}\| \geq \epsilon_{\text{AltMin}} \& \epsilon^{l,(u)}(\Gamma_l) \geq \epsilon_{\text{th}} \& u < T_{\max})$:

- 1) Update fault current parameters I : compute H_R using Table III. Solve for fault currents I using (18).
- 2) Update fault relative location parameters R : compute H_I and Y' using Table III, Solve for the fault relative locations R using (20).
- 3) Compute norm error associated with Γ_l : $\epsilon^{l,(u)}(\Gamma_l) = \|Y' - H_I R\|$ using H_I , Y' and the solution R of previous stage ii.
- 4) Proceed to next alternate minimization iteration $u \leftarrow u + 1$.

Step 2.1.2. Set final candidate residual: set $\epsilon^l(\Gamma_l) = \epsilon^{l,(u_{\max})}(\Gamma_l)$ the minimal norm error associated with candidate Γ_l .

Step 2.2. Choose candidate with minimal residual: Select optimal fault candidate $\Gamma_l = \arg \min \{\epsilon^l(\Gamma_l)\}$, using the outer minimization of (17). Set $\epsilon^l = \min \{\epsilon^l(\Gamma_l)\}$. Save I , R values associated with the candidate of minimal norm error ϵ^l , for initialization in 2.1 for the next stage $l + 1$. Update values of $(k_{\text{sg}}^l, k_{\text{dl}}^l, k_{\text{ll}}^l)$ according to Table V and the selected fault type Γ_l .

Step 3. Final estimate of the fault location: If $(k_{\text{sg}}^l \leq k_{\text{sg}}$, and $k_{\text{dl}}^l \leq k_{\text{dl}}$ and $k_{\text{ll}}^l \leq k_{\text{ll}})$ then the estimated faults are given by the set $\{\Gamma_l, l = 1, \dots, l_{\max}\}$. Else:

- 1) Compute the response norm for each fault in $\{\Gamma_l\}$: $\|h_l(I_{\Gamma_l}, R_{\Gamma_l})\|$.
- 2) The estimated faults are the $(k_{\text{sg}}, k_{\text{dl}}, k_{\text{ll}})$ faults with the maximal response norm.

V. SIMULATION RESULTS

A. Simulation Settings

The fault location algorithms were tested on the IEEE 118 bus test-case network [37]. Faults of each one of the three categories of Table I (Line-to-ground fault, disconnected lines, line-to-line shorts) were generated according to the simulated scenarios. Sensors are assumed to be synchronized PMUs, and provide the bus voltages and the corresponding line currents. The policy for sensor placement is detailed in Section V-B.

The sensing matrix A which relates the injected currents to the sensor measurements is determined by the impedance matrix Z_{bus} and the sensor locations (as in (3)). Each column j in the sensing matrix corresponds to the sensor response due to a unit current source connected on bus j . To synthesize each fault response, the equivalent models of Table I are used; this enables to represent each of the faults by an equivalent current vector on network buses such as in example (11). The nominal measurement AI^d is normalized such that $\|AI^d\|^2 = 1$. This measurement is then contaminated with additive noise ϵ^y whose norm is determined by the SNR. The additive noise vector is thus re-normalized to achieve a target SNR (in dB scale):

$$10 \log_{10} \frac{\|AI^d\|^2}{\|\epsilon_{M \times 1}^y\|^2} = \text{SNR} \xrightarrow{\|AI^d\|^2 = 1} \|\epsilon_{M \times 1}^y\|^2 = 10^{-\text{SNR}/10}. \quad (21)$$

The corresponding SNR varies in different scenarios as indicated in Fig. 4. The SNR thus reflects the ratio of the total signal energy divided by the total noise energy.

A fault location estimate is considered successful if the following two conditions are met: *a*) the fault type is correctly identified, and *b*) the estimated location is either on the faulted line (or lines), or on a line adjacent to it. Such criterion is reasonable since oftentimes the fault location is close to a bus on the edge of the faulted line. Section V-C further examines the recovery performance with additional constraints on the maximal relative location error.

The results of the proposed StructMP algorithm were compared to the Dynamic-Search method, which is an exhaustive variant of StructMP. In Dynamic-Search the objective is to minimize (12) by testing any possible fault combination. The fault location parameters for each such combination are determined using the approximate alternate minimization scheme of (18) and (20). While Dynamic-Search yields a close to oracle fault location performance, this method requires exponential time in the number of faults. Hence, it is provided as a reference to the proposed technique.

The number of simulations per scenario is set between 450 and 550 for any given fault scheme, SNR, and fixed sensor deployment. An exception to the above is the Dynamic-Search: due to its long run-time, only 150 simulations per scenario were carried out.

B. Sensor Placement

As mentioned in Section III, the method presented in this paper is based on a per-phase positive sequence data. When

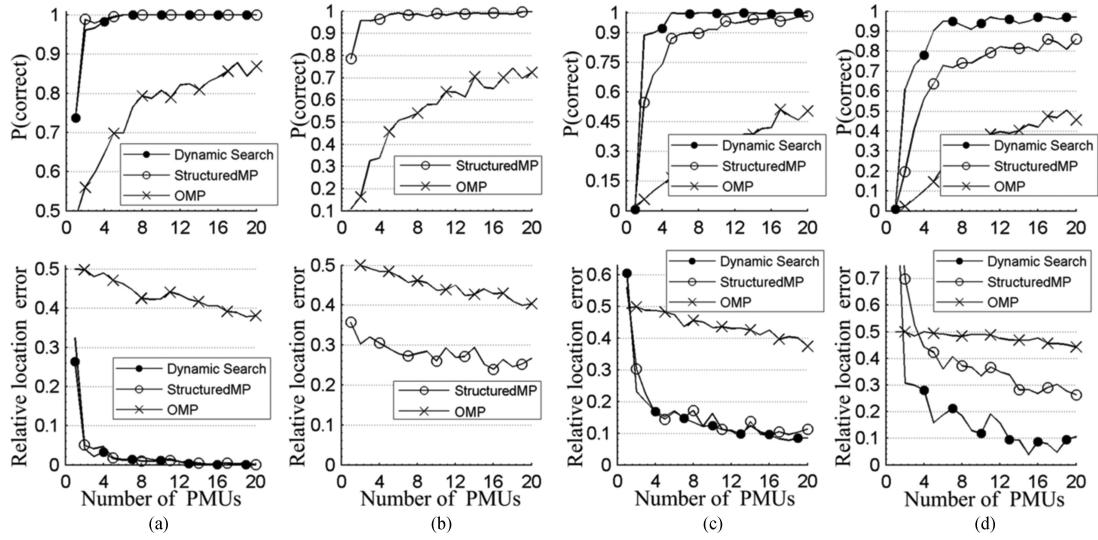


Fig. 2. Probability of fault recovery and mean maximal relative location error. SNR=Inf dB noise free scenario (from left to right: a single line-to-ground, a single line-to-line short, 1 line-to-ground fault and 1 disconnected line, 2 line to ground faults. (a) 1 short to ground [infinity symbol dB]. (b) 1 line-line short [infinity symbol dB]. (c) 1 short, 1 discon. line [infinity symbol dB]. (d) 2 short to ground [infinity symbol dB].

placing a sensor, it is assumed each PMU sensor used provides measurements for each one of the 3 phases.

We propose a greedy sensor placement method that stems from the coherence criterion. In addition to minimizing the correlation between columns of the sensing matrix, the proposed method also penalizes variability in the columns Euclidean norm. The objective of the sensor placement algorithm is to construct a sensing matrix A that enables accurate location of faults, while using a minimal number of sensors. Let A^{all} be a matrix with all possible physical measurements of the network variables; Each measurement corresponds to a single row in A^{all} . The sensing matrix A is formed by iteratively adding row subsets of A^{all} that correspond to the selected PMU sensor locations.

$$A^{\text{all}} = [(A^{\text{all}})^{S^{(1)}}; (A^{\text{all}})^{S^{(2)}}; \dots; (A^{\text{all}})^{S^{(m)}}]. \quad (22)$$

Sensor placement algorithms are often based on the sensing matrix coherence [10], [38]:

$$\mu(A) = \max_{\substack{j, k \\ j \neq k}} \frac{|\langle A_j, A_k \rangle|}{\|A_j\| \|A_k\|}. \quad (23)$$

However, experiments show that minimizing the coherence μ without additional constraints, makes some faults difficult to detect.

Specifically, with low norms of sensing matrix columns $\|A_j\|$, faults are more susceptible to noise, and may be screened by other correlated faults. The coherence criterion also performs poorly in greedy sensor placement methods, since each sensor measures responses that are restricted to a local and small part of the network. Consequently, with greedy sensor selection, the coherence measure becomes insensitive to different candidate locations, so that many different sensor locations show only negligible change in the coherence. As such, the coherence alone typically does not provide a good criterion to such an iterative sensor selection scheme.

To address these limitations, we used a sensor placement criterion that measures the absolute distance between different

response vectors, and in addition promotes uniformity among the response norms. The algorithm operates iteratively in a greedy manner. In each iteration l , one additional PMU is selected and its corresponding row subset $S(l)$ is augmented to A :

$$A = \begin{bmatrix} (A^{\text{all}})^{S_1} \\ \vdots \\ (A^{\text{all}})^{S_l} \end{bmatrix}, \quad (24)$$

where $S(l)$ is the subset of rows of the matrix A^{all} that are associated to the selected PMU at stage l . A PMU is selected in every iteration l such that following objective is maximized:

$$R(\lambda) = (1 - \lambda)\beta + \lambda P_{\text{ratio}}, \quad (25)$$

where λ is some constant in the range $\lambda \in (0, 1)$, and

$$\beta = \frac{\#\text{column pairs } \{j, k\} \text{ s.t. } (1 - \mu_{jk}^2) \|A_j\|^2 > d_{\min}^2}{\#\text{column pairs}}, \quad (26)$$

$$P_{\text{ratio}} = \frac{\#\text{column } j \text{ s.t. } \|A_j\|^2 / \max_k \|A_k\|^2 \geq r_{\min}^2}{\#\text{of columns}}. \quad (27)$$

Sensors should be placed such that β and P_{ratio} are as high as possible. Since these two objectives cannot be both optimized simultaneously, the weighted criterion $R(\lambda)$ of (19) is used as an objective. The first criterion β in (20) is analogous to the coherence [3],[4] but guarantees that different response columns have a non-negligible absolute distance from one another. $\mu_{jk} = \langle A_j, A_k \rangle / (\|A_j\| \|A_k\|)$ is the normalized correlation between the response vectors A_j and A_k of the sensing matrix. Eventually, β is defined as a relative count of the number of column pairs which comply with this condition. This relative count captures the local information provided by each candidate sensor.

The second criterion P_{ratio} measures the ratio between the highest and smallest response norms. It counts the relative number of columns in the sensing matrix A whose norm response is at least r_{\min} times the column norm with the highest response.

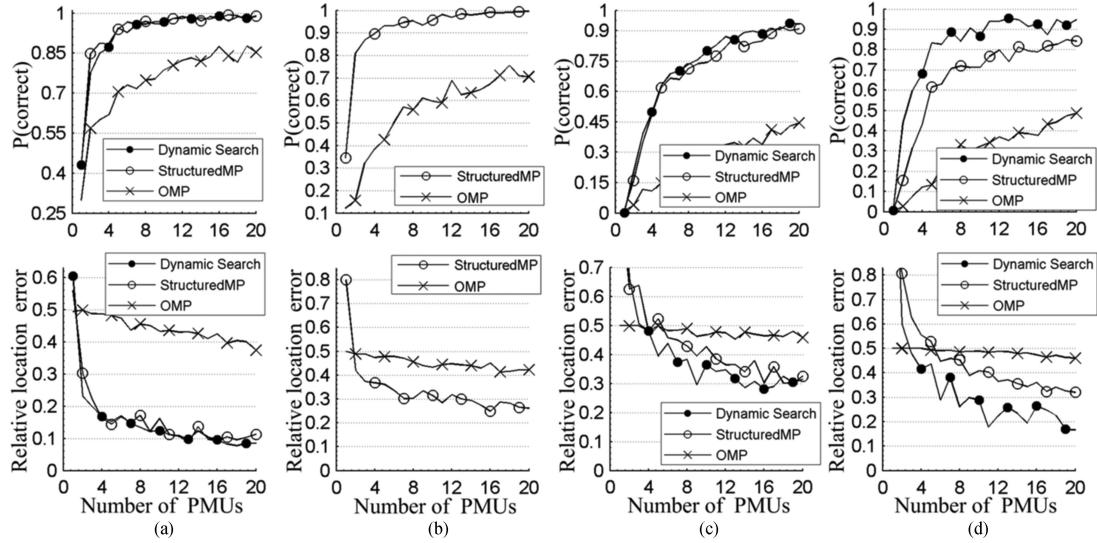


Fig. 3. Probability of fault recovery and mean maximal relative location error. SNR=50 dB (from left column to right column: a single line-to-ground short, a single line-line short, 1 Line-to-ground short and 1 disconnected line, 2 shorts to ground). (a) 1 short to ground [50 dB]. (b) 1 line-line short [50 dB]. (c) 1 short, 1 discon. line [50 dB]. (d) 2 shorts to ground [50 dB].

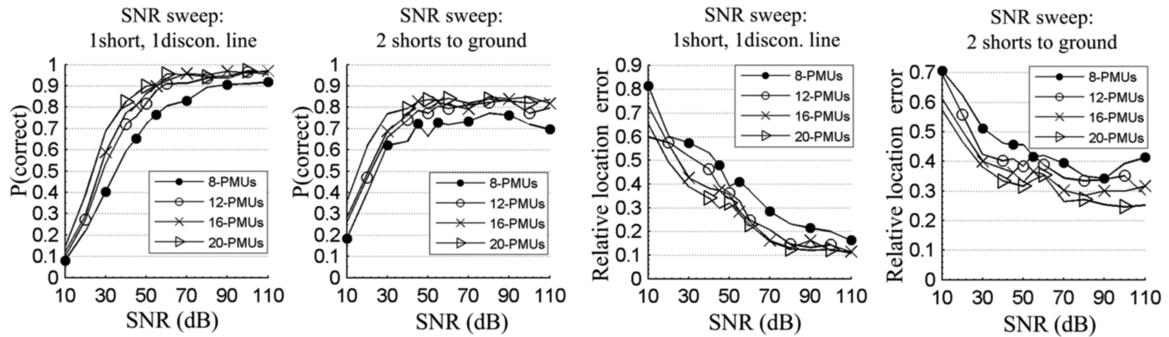


Fig. 4. SNR sweep for the StructMP: Probability of fault recovery and relative location error for varying SNRs.

The norm response to different faults are promoted to have the same order of magnitude, by selecting higher values of λ in (25). With this choice, the score of P_{ratio} becomes more significant and the resulting norm responses are closer to uniform.

It is observed that both criterions β and P_{ratio} are based on a relative measure of success rather than a global or worst case measure. This kind of objectives led to better properties of the resulting sensing matrix, and consequently to better recovery performance. This approach was adopted due to the physical nature of information provided by power network sensors. Since power network buses are sparsely connected, PMU sensors provide only local information on network variables. This property makes it almost impossible to optimize for a global objective (e.g. coherence), especially in a greedy stepwise selection of sensors and a relative improvement, measure gives better performance.

The values of $(d_{\min}, r_{\min}, \lambda)$ effect the performance of the StructMP algorithm as well as its robustness to noise, for a given number of PMU sensors. We choose the values $(d_{\min}, r_{\min}, \lambda) = (0.01, 0.39, 0.2)$ for the sensor positioning. The value $\lambda = 0.2$ from equation (25) is chosen to achieve a target r_{\min}^2 of 0.15 and specified d_{\min} of 0.01. The resulting sensor locations are used for the simulation below.

C. Simulated Scenarios and Results

Results for different fault schemes and a varying number of PMU sensors are presented in Figs. 2 and 3, both show a noise-free scenario with a 50 dB SNR. With 20 PMUs the proposed StructMP algorithm achieves recovery rate of up to 99% and higher for single fault scenarios even in the noisy case. With a simultaneous line to ground and disconnected line the recovery performance decreases from 98% in the noise-free scenario to 93% in the noisy case. For two line to ground faults the recovery rate is roughly the same in both scenarios, around 85%. In a single simulation run, the StructMP and Dynamic Search are successful only if the two conditions a and b of Section IV.A apply for all faults in the system. The OMP algorithm is considered successful if its estimated vector support contains at least one edge of a line associated with each fault. The mean relative location error is applied only for these successful estimations. The relative fault location error of the OMP simulations was set to 0.5 if one side of a faulted line is included in the support and 0 if both sides are included.

The StructMP robustness to noise was tested in the SNR sweep shown in Fig. 4. The plots provide information on the

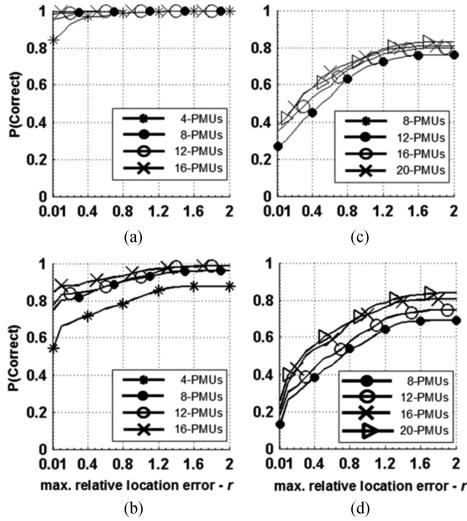


Fig. 5. StructMP recovery probability vs. allowed relative location error. (a) 1 short to ground [∞ symbol dB]. (b) 1 short to ground [50 dB]. (c) 2 short to ground [∞ symbol dB]. (d) 2 short to ground [50 dB].

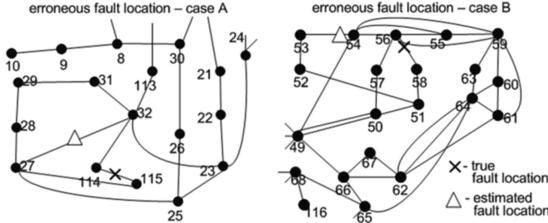


Fig. 6. Erroneous fault recovery of a single line-to-ground fault with 5 PMUs.

effect of the noise on the performance. At higher SNRs the performance improves and with 20 PMUs reaches around 85% to 95% depending on the fault schemes. For several SNR values a plateau in the fault recovery performance is observed.

The tolerance of the StructMP to relative location error is shown in Fig. 5, for two scenarios: a single line-to-ground and a simultaneous two line-to-ground short circuit fault. In these simulations, we adjust the maximal allowed relative location errors and measure the correct recovery rate with this additional constraint. It is evident that the StructMP recovery is less precise with multiple faults. The performance curve of Fig. 5 exhibits a knee pattern, so that the curves approximately align when the maximal relative location reaches 1.2. Below this approximate value, performance degrades substantially.

Analysis of unsuccessful fault location is outlined in Table IV for different scenarios. Five PMU sensors are positioned in the network. The sensor configuration is the same as used for the simulations reported in Figs. 2–4. The PMUs are located on buses 83, 38, 66, 11, 104. The table presents three fault scenarios (case A case C), in which the StructMP failed to correctly identify the fault locations. The inner table details the residuals obtained by the structMP inner minimization (13), and is used to select one fault among the candidate faults.

Case A (Fig. 5) reveals that some fault events induce the exact same sensor response and thus cannot be distinguished with 5 PMUs. For example, a short on the line connecting (115,114) is disguised as a short on a close, but not adjacent line (32,27). The non-observability of a single line-to-ground fault with 5 PMU

TABLE VI
ANALYSIS OF ERRONEOUS FAULT IDENTIFICATION OF THE
STRUCTURED MATCHING PURSUIT

Case A: A single line-to-ground short circuit without measurement noise			
Line-to-ground: line (115,114) $r = 0.487$, $i_x = -9.76 + i16.623$			
<i>Estimated:</i> line: (32,27) $r = 0.457$;			
Candidate residual norms after the 1st iteration:			
line estimate	(32,37)	(115,114)	(115,27)
resid. norm	$5.2 \cdot 10^{-6}$	$6.1 \cdot 10^{-6}$	$3.9 \cdot 10^{-4}$
line estimate	(114,32)		
resid. norm	$4.2 \cdot 10^{-4}$		
Case B: A single line-to-ground with measurement noise (SNR=50dB)			
Line-to-ground: line (58,56) $r = 0.789$, $i_x = -13.14 + i9.43$			
<i>Estimated:</i> line: (54,53) $r = 0.167$;			
Candidate residual norms after the 1st iteration without noise:			
line estimate	(58,56)	(54,49)	(57,56)
resid. norm	$2.2 \cdot 10^{-6}$	$9.5 \cdot 10^{-5}$	$4.6 \cdot 10^{-4}$
line estimate	(54,53)	(58,56)	(54,49)
resid. norm	$3.0 \cdot 10^{-3}$	$3.0 \cdot 10^{-3}$	$3.0 \cdot 10^{-3}$
line estimate	(57,56)		
resid. norm	$3.1 \cdot 10^{-3}$		
Case C: Simultaneous 1 line-to-ground and 1 disconnected line without measurement noise			
Line-to-ground: line (87,86) $r = 0.596$, $i_x = -3.23 + i13.31$			
Disconnected line: line (81,68), $i_x = -3.456 + i1.132$			
<i>Estimated:</i> Line-to-ground: line (85,83) $r = 0$			
Disconnected-line: line (81,68)			
Candidate residual norms following the 1st iteration (LG search):			
LG line cand.	(85,83)	(85,84)	(87,86)
resid. norm	$7 \cdot 10^{-3}$	$7 \cdot 10^{-3}$	$7.1 \cdot 10^{-3}$
LG line cand.	(86,85)		
resid. norm	$7.1 \cdot 10^{-3}$		
Candidate residual norms following the 2nd iteration (DL search):			
DL line cand.	(81,68)	(81,80)	(97,96)
resid. norm	$4.6 \cdot 10^{-5}$	$3.4 \cdot 10^{-3}$	$3.5 \cdot 10^{-3}$
DL line cand.	(96,80)		
resid. norm	$3.5 \cdot 10^{-3}$		

sensors can be verified using [39]. Case B (Fig. 5) introduces a scenario where the correct fault location might have been identified correctly in the absence of noise. In the presence of noise, however, the measurements are distorted, and as a result several shorts may explain the measurement (Fig. 6). An erroneous fault location of a simultaneous line to ground short circuit and a disconnected line in the absence of noise is addressed in Table VI, case C.

D. Computational Time

The StructMP algorithm has the advantage of reducing the number of sensors in the system as well as having good computation time. For comparison, the run time of StructMP algorithm is compared to the run time of the Dynamic-Search method, which is an exhaustive variant of StructMP. The two algorithms were tested under the same conditions on the standard IEEE 118 bus system.

The run times were computed by running 20 different fault combinations (at random) for StructMP and Dynamic Search. The results were as follows: for placement of 10 PMU's the run time of StructMP algorithm was 4.64s, while with Dynamic search it was 143.35 s. Repeating the simulation for 20 PMU a yield run time of 5.75 s in the case of StructMP and 146.15s in the case of Dynamic search method.

These results show a definite advantage of the StructMP algorithm over the Dynamic-Search algorithm. The run-time of the proposed StructMP algorithm is over 25 times faster than the Dynamic Search algorithm, which is a good reference base for comparison since as above-mentions scans the entire solution space.

VI. DISCUSSION AND CONCLUSION

This paper addresses the challenge of fault location using a limited number of sensors. We propose an algorithm that exploits both the sparsity of solution vectors and the structural constraints imposed by the faults physical models. These constraints are utilized for improved location of several simultaneous faults, using a greedy polynomial time algorithm (StructMP). The algorithm searches directly over physical faults, and as a result operates over a small solution space. The faults are characterized by a set of continuous parameters which enable to represent events on any point across the power network including on the transmission lines. An alternating minimization scheme extends the OMP method by minimizing the norm error with respect to these continuous parameters, and thus recovers faults with a small number of sensors. Simulations on the IEEE 118 bus test case network show that 4-20 PMU sensors are sufficient to recover faults with adequate accuracy and reasonable run times.

REFERENCES

- [1] A. Abur and A. G. Expósito, *Power System State Estimation: Theory and Implementation*. Boca Raton, FL, USA: CRC, 2004.
- [2] G. B. Giannakis, V. Kekatos, N. Gatsis, S.-J. Kim, H. Zhu, and B. F. Wollenberg, "Monitoring and optimization for power grids: A signal processing perspective," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 107–128, Sep. 2013.
- [3] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, U.K.: Cambridge Univ., 2012.
- [4] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York, NY, USA: Springer-Verlag, 2010.
- [5] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge, U.K.: Cambridge Univ., 2015.
- [6] V. Kekatos and G. B. Giannakis, "Joint power system state estimation and breaker status identification," in *Proc. North Amer. Power Symp.*, 2012, pp. 1–6.
- [7] X. Li, H. V. Poor, and A. Scaglione, "Blind topology identification for power systems," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, 2013, pp. 91–96.
- [8] V. Kekatos, G. B. Giannakis, and R. Baldick, "Online energy price matrix factorization for power grid topology tracking," *IEEE Trans. Smart Grid*, vol. 7, no. 3, pp. 1239–1248, May 2016.
- [9] Y. Liao, Y. Weng, and R. Rajagopal, "Urban distribution grid topology reconstruction via Lasso," in *Proc. Power Energy Soc. Gen. Meeting*, 2016, pp. 1–5.
- [10] H. Liao, "Power system harmonic state estimation and observability analysis via sparsity maximization," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 15–23, Feb. 2007.
- [11] G. Mateos and G. Giannakis, "Load curve data cleansing and imputation via sparsity and low rank," *IEEE Trans. Smart Grid*, vol. 4, no. 4, pp. 2347–2355, Dec. 2013.
- [12] N. Nguyen and T. Tran, "Robust Lasso with missing and grossly corrupted observations," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2036–2058, Apr. 2013.
- [13] S. Alam, B. Natarajan, and A. Pahwa, "Distribution grid state estimation from compressed measurements," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1631–1642, Jul. 2014.
- [14] Y. C. Chen, A. D. Dominguez-Garcia, and P. W. Sauer, "A sparse representation approach to online estimation of power system distribution factors," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1727–1738, Jul. 2015.
- [15] M. Majidi, M. Etezadi-Amoli, H. Livani, and M. Fadali, "Distribution systems state estimation using sparsified voltage profile," *Elect. Power Syst. Res.*, vol. 136, pp. 69–78, 2016.
- [16] M. Babakmehr, M. G. Simoes, M. B. Wakin, and F. Harirchi, "Compressive sensing-based topology identification for smart grids," *IEEE Trans. Ind. Inf.*, vol. 12, no. 2, pp. 532–543, Apr. 2016.
- [17] M. Korkali, H. Lev-Ari, and A. Abur, "Traveling-wave-based fault-location technique for transmission grids via wide-area synchronized voltage measurements," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 1003–1011, May 2012.
- [18] A. S. Dobakhshari and A. M. Ranjbar, "A novel method for fault location of transmission lines by wide-area voltage measurements considering measurement errors," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 874–884, Mar. 2015.
- [19] Y. Liao, "Fault location for single-circuit line based on bus-impedance matrix utilizing voltage measurements," *IEEE Trans. Power Del.*, vol. 23, no. 2, pp. 609–617, Apr. 2008.
- [20] Q. Jiang, X. Li, B. Wang, and H. Wang, "PMU-based fault location using voltage measurements in large transmission networks," *IEEE Trans. Power Del.*, vol. 27, no. 3, pp. 1644–1652, Jul. 2012.
- [21] Q. Jiang, B. Wang, and X. Li, "An efficient PMU-based fault-location technique for multiterminal transmission lines," *IEEE Trans. Power Del.*, vol. 29, no. 4, pp. 1675–1682, Aug. 2014.
- [22] 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.
- [23] S. Azizi and M. Sanaye-Pasand, "A straightforward method for wide-area fault location on transmission networks," *IEEE Trans. Power Del.*, vol. 30, no. 1, pp. 264–272, Feb. 2015.
- [24] S. Azizi, M. Sanaye-Pasand, and M. Paolone, "Locating faults on untransposed, meshed transmission networks using a limited number of synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4462–4472, Nov. 2016.
- [25] A. Salehi-Dobakhshari and A. M. Ranjbar, "Application of synchronised phasor measurements to wide-area fault diagnosis and location," *IET Gener. Transm. Distrib.*, vol. 8, no. 4, pp. 716–729, 2014.
- [26] C.-W. Liu, K.-P. Lien, C.-S. Chen, and J.-A. Jiang, "A universal fault location technique for n-terminal ($n \geq 3$) transmission lines," *IEEE Trans. Power Del.*, vol. 23, no. 3, pp. 1366–1373, Jul. 2008.
- [27] A. S. Dobakhshari, "Wide-area fault location of transmission lines by hybrid synchronized/unsynchronized voltage measurements," *IEEE Trans. Smart Grid*, Aug. 2016.
- [28] H. Zhu and G. B. Giannakis, "Sparse overcomplete representations for efficient identification of power line outages," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2215–2224, Nov. 2012.
- [29] M. Babakmehr, M. Godoy Simoes, A. Al-Durra, F. Harirchi, and Q. Han, "Application of compressive sensing for distributed and structured power line outage detection in smart grids," in *Proc. Amer. Control Conf.*, 2015, pp. 3682–3689.
- [30] J.-C. Chen, W.-T. Li, C.-K. Wen, J.-H. Teng, and P. Ting, "Efficient identification method for power line outages in the smart power grid," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1788–1800, Jul. 2014.
- [31] I. Rozenberg and Y. Levron, "Applications of compressed sensing and sparse representations for state estimation in power systems," in *Proc. IEEE Int. Conf. Microw., Commun., Antennas Electron. Syst.*, 2015.
- [32] M. Majidi, M. Etezadi-Amoli, and M. S. Fadali, "A sparse-data-driven approach for fault location in transmission networks," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 548–556, Mar. 2017.
- [33] M. Majidi, A. Arabali, and M. Etezadi-Amoli, "Fault location in distribution networks by compressive sensing," *IEEE Trans. Power Del.*, vol. 30, no. 4, pp. 1761–1769, Aug. 2015.
- [34] L. Zhao and W.-Z. Song, "Distributed power-line outage detection based on wide area measurement system," *Sensors*, vol. 14, no. 12, pp. 13 114–13 133, 2014.
- [35] G. Feng and A. Abur, "Robust fault location for two and three terminal lines using synchronized phasor measurements," in *IEEE Eindhoven PowerTech*, 2015, pp. 1–5.
- [36] G. Feng and A. Abur, "Fault location using wide-area measurements and sparse estimation," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2938–2945, Jul. 2016.
- [37] "Power systems test case archive," Dept. Elect. Eng., Univ. Washington, Seattle, WA, USA, 2018. [Online]. Available: <http://www.ee.washington.edu/research/pstca>
- [38] M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5695–5702, Dec. 2007.
- [39] I. Rozenberg and Y. Levron, "Observability challenges in sparse estimation of fault events," in *Proc. IEEE Int. Energy Conf.*, 2016, pp. 1–6.

Authors' photographs and biographies not available at the time of publication.