Line Outage Detection in Power Transmission Networks via Message Passing Algorithms

Jianshu Chen*, Yue Zhao†, Andrea Goldsmith‡ and H. Vincent Poor§
*Microsoft Research, Redmond, WA 98052
†Department of Electrical and Computer Engineering, Stony Brook University, Stony Brook, NY 11794
‡Department of Electrical Engineering, Stanford University, Stanford, CA 94305
§Department of Electrical Engineering, Princeton University, Princeton, NJ 08544
Email: cjs09@ucla.edu, yue.zhao.2@stonybrook.edu, andrea@ee.stanford.edu, poor@princeton.edu

***Abstract*—Detecting multiple simultaneous line outages in power transmission networks is known to be a challenging problem due to the number of hypotheses that grows exponentially with the network size. A low complexity message passing algorithm is proposed for multi-line outage identification, which exploits the underlying sparse structure of the network topology in power systems. First, a factor graph is established that characterizes jointly the power system and the sensor network monitoring it. For inferring line status, the mixed integer and continuous variables and the loopy structure of the factor graph make it difficult to use conventional message passing algorithms. Exploiting the power flow equations, efficient message representation and new techniques in message passing algorithms are developed. Simulation results demonstrate that the developed algorithm can effectively identify an arbitrary number of simultaneous line outages in real time.**

I. Introduction

The wide area monitoring system (WAMS) plays a crucial role in preventing power transmission networks from failing due to unexpected component outages. Major blackouts were often caused by the lack of real time knowledge of component failures that can quickly escalate into large-scale cascading failures [1]. While the power system is usually protected against the so called "$N-1$" scenarios (i.e., only one component fails), as failures accumulate, automatic protection is no longer guaranteed. In critical situations in which cascading failures start developing, real-time protective actions depend on correct and timely knowledge of the network status. In particular, since we may have already missed the first few component outages, the ability to identify in real time the network topology with an *arbitrary* number of outages becomes critical to prevent system collapse.

Previous works on transmission line outage detection using exhaustive search methods include [2], [3] and [4], which focus on detecting single and two-line outages. As the number of possible outage patterns grows exponentially with the number of components (e.g. lines) in the network, exhaustive search methods quickly run into computational complexity problems as the number of potential simultaneous outages increases. Beyond two-line outages, [5] has recently exploited the sparsity of outage patterns with overcomplete observations to detect sparse multi-line outages. A related work [6] relies on ergodicity of power injections and full phasor measurement unit (PMU) deployment to reconstruct a dependency graph of the power grid in order to detect multi-line outages.

In this paper, our objective is to identify the instantaneous topology of the power transmission network (i) in *real time*, (ii) regardless of how many line outages there are, (iii) with low complexity, and (iv) using undercomplete observations. To address the challenge of the exponentially large number of outage patterns (on the order of 2^L where L is the number of lines), we develop a *message passing* algorithm to provide fast identification of the instantaneous topology of the transmission network. We first establish a factor graph that captures the binary status of the lines, the continuous values of the physical quantities (power injections, power flows, etc.) in the system, and the noisy measurements from the sensors that monitor the network. We then develop a computationally efficient message representation for this mix of discrete and continuous variables. To address the loopy structure that is inherent in the topology of power transmission networks, we develop heuristics that significantly improve the performance of the message passing algorithms. We note that the proposed message passing algorithms for outage identification are quite different from those developed for power system state estimation in which all variables are continuous [7], [8].

We specifically study identifying multi-line outages based on measurements by PMUs of voltage phase angles at a subset of the buses. This is an undercomplete inference problem since the number of observations is less than the number of unknowns (i.e., the status of the lines). Nonetheless, the developed message passing algorithm is able to detect an arbitrary number of line outages with good performance. We evaluate the developed algorithm on the IEEE 14-bus system. Simulation results show that the algorithm can effectively identify the topology of the network in the presence of arbitrary line outages with undercomplete observations.

II. Problem Formulation

We consider the problem of identifying the instantaneous topology of a power transmission network of N buses and L transmission lines. We employ a DC power flow model [9].

This work was supported in part by the National Science Foundation under Grant CMMI-1435778, by the DTRA under Grant HDTRA1-08-1-0010, and by the NSF under CPS Synergy grant 1330081.

Let p, f and θ denote the $N \times 1$ vector that collects the power injection at all the buses, the $L \times 1$ vector that collects the power flows in all the lines, and the $N \times 1$ vector that collects the voltage phase angles at all the buses, respectively. We then have the following relations among these quantities:

$$p = Mf, \tag{1}$$
$$f = S_k \Gamma M^T \theta, \tag{2}$$

where M is the incidence matrix of the power grid topology without line outages, S_k is the diagonal status matrix of the lines (each diagonal entry indicates the status of a line; zero means disconnected and one means connected), and Γ is a diagonal matrix whose diagonal entries correspond to the inverses of the line reactance. We assume that the voltage phase angles on a subset of the buses are measured by PMUs, and we assume that the measurements are corrupted by additive Gaussian noise. We denote by $\mathcal{M}$ the set of all the buses equipped with PMUs. Accordingly, the measurement can be modeled as

$$y = U\theta + v, \tag{3}$$

where U is an $|\mathcal{M}| \times N$ selection matrix with one entry being 1 for each row (the (m, n)-th entry is one if the m-th PMU is deployed at the n-th bus) and all other entries being 0, and v is the measurement noise distributed according to $\mathcal{N}(0, \sigma_v^2 I)$.

Substituting (2) into (1), we have

$$\theta = (MS_k\Gamma M^T)^\dagger p, \tag{4}$$

where $(\cdot)^\dagger$ denotes the Moore-Penrose pseudoinverse. Substituting (4) into (3), we obtain the following hypothesis testing problem:

$$\mathcal{H}_k : \; y = U(MS_k\Gamma M^T)^\dagger p + v, \quad k = 1, \ldots, 2^L. \tag{5}$$

That is, the topology identification problem can be formulated as making the best conjecture over 2^L possible hypotheses based on the observation y. The 2^L possibilities are due to the fact that each line can be either connected or not. Specifically, the entries of the diagonal matrix S_k are binary variables, and our objective of identifying the network topology is equivalent to inferring the binary diagonal entries of S_k using y. This is a challenging problem for the following two reasons:

- The number of hypotheses is exponentially large. Therefore, the complexity is too high for any exhaustive search method.
- The number of observations $|\mathcal{M}|$ is smaller than the number of unknowns L, as we consider using voltage phase angle measurements on only a subset of the buses.

Nevertheless, since the number of possible hypotheses is *finite* due to the *binary* nature of the unknowns, it is still possible that a good conjecture over these finitely many possibilities can be made based on undercomplete observations.

III. Factor Graph for Transmission Networks

Our objective is to solve the hypothesis testing problem (5) with low complexity. However, it is unclear just from (5) how to develop a low complexity algorithm for this problem. Instead, we begin by exploiting the structures of (1)–(3) to develop a new graphical model that represents the relations among different quantities in the transmission network.

The three equations in (1)–(3) characterize the relations among the key quantities in power grids. Specifically, there are three types of unknowns: *voltage phase angles* θ, *power flows* f, and *line connectivities* S_k. The power injections p are assumed to be known a priori (based on load forecasts or state estimates). Furthermore, each equation corresponds to one type of constraint, and we call them *bus check* (1), *line check* (2) and *measurement check* (3) equations. Accordingly, we establish a factor graph [10] in which

- Entries in the vectors θ and f and diagonal entries in the matrix S_k correspond to variable nodes.
- Each row of equations in (1)–(3) corresponds to a check node.

For a bus check or line check node, i.e., a row in (1) or (2), the variable nodes connected to it are given by the variables that are selected according to the nonzero entries of M^T or $S_k\Gamma M$. For a measurement check node, i.e., a row in (3), the corresponding entry in y is the noisy measurement of the true phase angle variable in θ. Note that, y is fixed and known once the measurements are taken, and hence it does not correspond to any variable nodes. Since y is related to θ through a linear relation plus *unknown* measurement noises, as opposed to bus and line check nodes described by "hard constraints" (1)–(2), the relation enforced by the measurement check is a "soft relation". Such a soft relation is made precise with the Gaussian noise assumption as in (8) below. Based on these rules, an illustrative factor graph for a 3-bus example is given in Fig. 1. The relations among the variable nodes at the check nodes are given by:

$$\text{[Bus check]}: \quad p_i = \sum_l M_{il} f_l, \quad l \triangleq (i,j) \in \mathcal{E},\ j \in \partial i,\ i \in \mathcal{V}, \tag{6}$$
$$\text{[Line check]}: \quad f_l = s_l \cdot \gamma_l \cdot (\theta_i - \theta_j), \quad l \in \mathcal{E}, \tag{7}$$
$$\text{[Meas. check]}: \quad \psi_m(\theta_n) \cong \exp\left\{-\frac{1}{2\sigma_v^2}(y_m - \theta_n)^2\right\}, \quad m \in \mathcal{M}, \tag{8}$$

where $l = (i, j) \in \mathcal{E}$ denotes the index of the lines, $\mathcal{E}$ denotes the set of all lines, $\mathcal{V}$ denotes the set of all buses, the notation $\cong$ means "equal up to a normalization factor", and ∂i denotes the neighborhood of node i. Note that (8) characterizes the soft relation between the measurement y_m and the true phase angle θ_n, which is derived from the Gaussian linear model (3).

IV. Message Passing for Line Outage Detection

A. Sparsity of the Power Transmission Network Topology

From Fig. 1, we observe that the topology of the factor graph integrates the topology of the transmission network

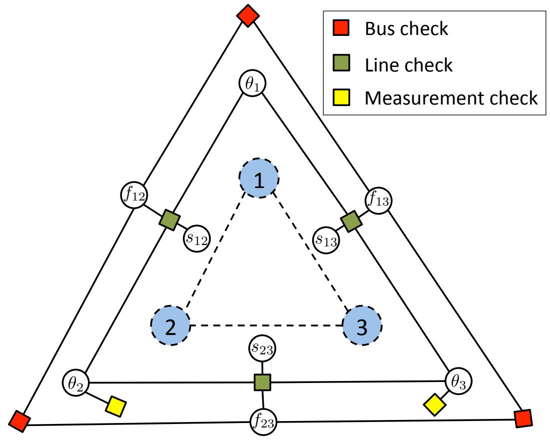


Fig. 1. Factor graph for a 3-bus example

and the placement of PMUs. Power transmission networks are sparsely connected networks, with an average degree per node typically below three. As a result, the associated factor graphs are also sparsely connected. Such sparsity enables us to develop a *low complexity* inference algorithm via message passing. We note that, unlike [5], we do not assume that the outage patterns are sparse, and we study identification of the network topology with an arbitrary number of line outages.

Next, we develop a message passing algorithm that performs inference of line status over the sparse factor graph, and address several challenges that arise in this inference problem.

B. Efficient Message Representation and Closed-form Iterations

In a factor graph, message passing (or belief propagation) algorithms infer the values of the variable nodes by computing their beliefs, i.e., the probabilities of these variables taking different values. This is achieved by having the variable and check nodes iteratively processing and passing messages over the edges. For a detailed tutorial on message passing over factor graphs, we refer the reader to [10].

The messages toward and from a variable node are probability distributions that characterize the beliefs of this variable from different information sources. Therefore, representing the messages is equivalent to representing a probability distribution of this variable. First, we observe from Fig. 1 that we have a mix of two types of variable nodes: discrete (binary) variables for the line status, and continuous variables for the phase angles and power flows. Therefore, we need to represent the probability distributions of both the discrete variables and the continuous variables. For any messages toward and from the binary variables, we can simply use the probability mass function (PMF) in the form of two-dimensional probability simplex vector to represent the messages, i.e., $[\Pr(s=0)\ \Pr(s=1)]$. On the other hand, the messages toward and from the continuous variables is a probability density function (PDF). A straightforward approach to representing the messages associated with the continuous variables is to discretize the PDF and represent it as a PMF as in

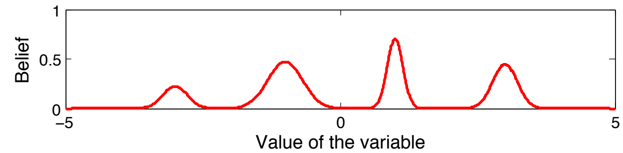


Fig. 2. Gaussian mixture model of the messages.

the discrete case. However, this turns out to be quite an inefficient method, because i) the range that each variable takes is fairly large, and ii) very small quantization error, and hence a large number of quantization levels are needed to achieve satisfactory performance in line status inference.

This motivates us to develop a different approach, in which we represent the messages for the continuous variables in a *parametric form*. This representation is based on the following observation from (6)–(8): the measurement noise is Gaussian, the relations from the bus checks are linear equations, and the relations from the line checks are linear equations multiplied by a binary variable indicating the corresponding line status. It follows that the messages over the edges in the factor graph *always obey a Gaussian mixture model* (GMM, cf. Fig. 2), which can be characterized by a few parameters: the probability, the mean and the variance of each component. This is a much more efficient way to represent the messages than discretization. Moreover, it allows us to develop closed-form expressions for iterations of message passing in the sum-product algorithm — a belief propagation method for computing the marginal probabilities of the variables in a factor graph. We summarize below the expressions for the messages that are iteratively passed in a sum-product algorithm. Their detailed derivations are omitted due to space limitations.

- Notation for messages: We denote a message from a check node b to a variable node i by $\hat{\nu}_{b\to i}$, and that from a variable node i to a check node a by $\nu_{i\to a}$. In what follows, we use a and b to index check nodes, and i, j, f and s to index variable nodes.
- Messages toward and from variable nodes: In each iteration, a variable node (i) collects and multiplies messages from all its neighboring check nodes ($b \in \partial i \backslash a$) but one ($a$), and sends this product as a message to the remaining check node (a). The product of GMM messages is also a GMM message. The mean (μ), variance (σ^2) and probability (q) of each GMM component in the output message, $\nu_{i\to a}$, can be computed from the means, variances and probabilities of one GMM component in each of the input messages, $\hat{\nu}_{b\to i}, b \in \partial i \backslash a$. Specifically,

$$\mu_{n_{i\to a}} = \sum_{b\in\partial i\backslash a} \frac{\mu_{n_{b\to i}}}{\sigma^2_{n_{b\to i}}} \Big/ \sum_{b\in\partial i\backslash a} \frac{1}{\sigma^2_{n_{b\to i}}}, \tag{9}$$

$$\sigma^2_{n_{i\to a}} = \left(\sum_{b\in\partial i\backslash a} \frac{1}{\sigma^2_{n_{b\to i}}} \right)^{-1}, \tag{10}$$

$$q_{n_{i\to a}} \cong \sqrt{\sigma^2_{n_{i\to a}}} \prod_{b\in\partial i\backslash a} \frac{q_{n_{b\to i}}}{\sqrt{2\pi\sigma^2_{n_{b\to i}}}}.$$

$$\exp\Big(-\sum_{b\in\partial i\setminus a}\frac{[\mu_{n_{i\to a}}-\mu_{n_{b\to i}}]^2}{2\cdot\sigma^2_{n_{b\to i}}}\Big), \tag{11}$$

where $n_{i\to a}$ and $n_{b\to i}$ are the indices of the GMM components in $\nu_{i\to a}$ and $\hat{\nu}_{b\to i}$, respectively.

- Messages from a measurement check node: Given the measurements at a check node a, and the noise variance σ_v^2, $\hat{\nu}_{a\to i}$ is a Gaussian distribution with

$$\mu_{a\to i} = y_a \tag{12}$$
$$\sigma^2_{a\to i} = \sigma_v^2. \tag{13}$$

- Messages from a bus check node: From (6), each GMM component $n_{a\to i}$ from a bus check node a to a power flow variable node i is computed as follows:

$$\mu_{n_{a\to i}} = \frac{1}{M_{ai}}\left(p_a - \sum_{j\in\partial a\setminus i} M_{aj}\mu_{n_{j\to a}}\right) \tag{14}$$
$$\sigma^2_{n_{a\to i}} = \frac{1}{M_{ai}^2}\sum_{j\in\partial a\setminus i} M_{aj}^2\sigma^2_{n_{j\to a}} \tag{15}$$
$$q_{n_{a\to i}} = \prod_{j\in\partial a\setminus i} q_{n_{j\to a}}, \tag{16}$$

where M is the network incidence matrix, and $n_{a\to i}$ and $n_{j\to a}$ are the indices of the GMM components in $\hat{\nu}_{a\to i}$ and $\nu_{j\to a}$, respectively.

- Messages from a line check node to a power flow node: From (2), depending on whether the probabilities that this line is connected and disconnected, i.e., $\nu_{s\to a}(1)$ and $\nu_{s\to a}(0)$, each GMM component from a line check node a to power flow variable node f is computed as follows:

$$\mu_{1,n_{a\to f}} = \gamma_a\cdot\big(\mu_{n_{i\to a}}-\mu_{n_{j\to a}}\big) \tag{17}$$
$$\sigma^2_{1,n_{a\to f}} = \gamma_a^2\cdot\big(\sigma^2_{n_{i\to a}}+\sigma^2_{n_{j\to a}}\big) \tag{18}$$
$$q_{1,n_{a\to f}} = \nu_{s\to a}(1)\cdot q_{n_{i\to a}}q_{n_{j\to a}} \tag{19}$$
$$\mu_{0,a\to f} = 0 \tag{20}$$
$$\sigma^2_{0,a\to f} = 0 \tag{21}$$
$$q_{0,a\to f} = \nu_{s\to a}(0), \tag{22}$$

where the meaning of the last three equations is that, if the line is disconnected, the power flow is deterministically zero (with a zero variance).

- Messages from a line check node to a phase angle node: From (2), each GMM component from a line check node a to a phase angle variable node i is computed as follows:

$$q_{0,a\to i} = \nu_{s\to a}(0) \tag{23}$$
$$\mu_{0,a\to i} = 0 \tag{24}$$
$$\sigma^2_{0,a\to i} = +\infty \tag{25}$$
$$q_{1,n_{a\to i}} = \nu_{s\to a}(1)\cdot q_{n_{j\to a}}\cdot q_{1,n_{p\to a}} \tag{26}$$
$$\mu_{1,n_{a\to i}} = \mu_{n_{j\to a}} + \frac{1}{\gamma_a}\cdot\mu_{1,n_{p\to a}} \tag{27}$$
$$\sigma^2_{1,n_{a\to i}} = \sigma^2_{n_{j\to a}} + \frac{1}{\gamma_a^2}\cdot\sigma^2_{1,n_{p\to a}}, \tag{28}$$

where the meaning of the first three equations is that, if the line is disconnected, then the phase angle at the other end of the line provides no information (infinite variance) to that at the other end.

- Messages from a line check node to a line status node: The estimated probability of line outage is returned as the soft-decision of this line status, i.e.,

$$\hat{\nu}_{a\to s}(0) = q_{0,p\to a},\ \hat{\nu}_{a\to s}(1) = q_{1,p\to a}. \tag{29}$$

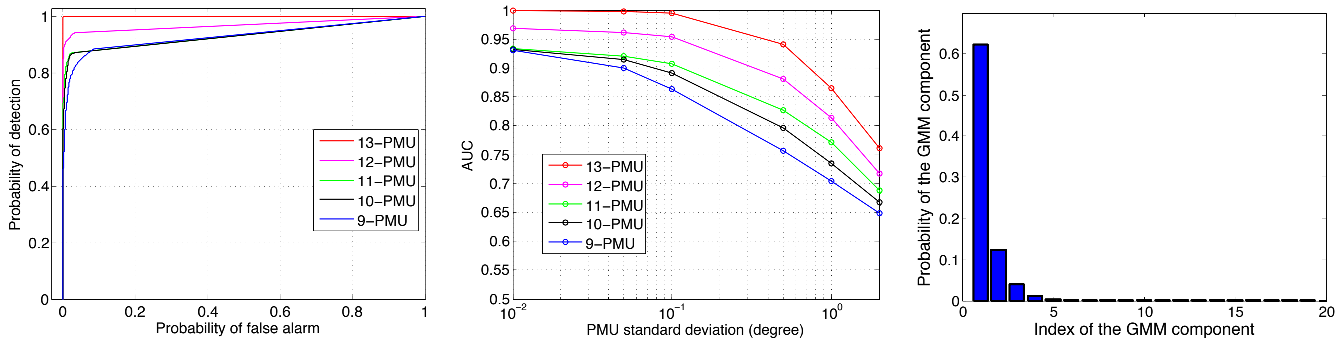
C. Loopy Message Passing

The effectiveness of the message passing algorithm is also challenged by the loopy nature of the factor graph as shown in Fig. 1. The issues caused by this loopy nature include the following. i) The parametric form for the variance update (10) is the harmonic mean of the variances from different edges. Thus, the more messages we aggregate, the smaller the variance is. This makes sense in a graph without loops in which messages are coming from *independent* information sources. In a loopy graph (especially with small loops as in Fig. 1), however, the *same* information will be passed along a small cycle, coming back and being aggregated again. With copies of identical information, intuitively we should not further reduce the variance of the GMM components because we do not have improvement in the accuracy the information. To address this issue, we incorporate a simple but effective solution in our implementation: we freeze the variance updates, and do not change them over time. ii) The number of GMM components grows exponentially with the number of iterations. This is because identical or perturbed information comes back via cycles. In reality, these identical pieces of information which come back have GMM components with very similar means. Accordingly, we just perform a quantization over the means, with which these similar GMM components get merged once they fall into the same quantization bin. We also observe that only a small fraction of the GMM components are dominant, and we hence discard the negligible components.

V. Simulation Results

In this section, we evaluate the performance of the proposed message passing algorithm on the IEEE 14-bus system. Typical power injections are employed [11], and the 7-th and the 8-th buses are equivalently merged into one bus because of the zero injections at these two adjacent buses. Therefore, we are dealing with an equivalent 13-bus system. We use the receiver operating curve (ROC), i.e., the probability of detection versus the probability of false alarm of the line outages, to evaluate the performance of the proposed algorithm, and we test the algorithm against up to 6-line outages. All the experimental results are averaged over 5000 Monte Carlo runs. Furthermore, only the top five GMM components are maintained during the message passing iterations.

We randomly generate from 1 to 6 line outages with a uniform distribution on the number of simultaneous line outages. We test against only those outages under which the network is still connected. In Fig. 3(a), we show the ROC curves for different numbers of PMUs with a measurement accuracy of 0.01 degree, which is the state-of-the-art PMU accuracy [12]. We observe that the message passing algorithm



(a) (b) (c)

Fig. 3. Outage detection performance of the message passing algorithm on the IEEE 14-bus system. (a) ROC curves for different numbers of PMUs, where the standard deviation of the PMU is 0.01 degree. (b) Noise performance for different numbers of PMUs. (c) Probabilities of Gaussian mixture components.

provides reasonably good performance even with 9 PMUs, which is significantly less than the number of unknowns (19 in the 14-bus case).

One question that arises is whether the algorithm is sensitive to different numbers of simultaneous line outages. To answer this question, we run extensive experiments comparing ROCs for identifying different numbers of line outages. Interestingly, we observe that the algorithm performs almost the same in all these settings. In other words, the developed message passing algorithm is equally powerful regardless of the number of simultaneous line outages to be identified.

Furthermore, we use AUC (area underneath the curve) of the ROC as the overall performance metric to evaluate the robustness of the message passing algorithm against measurement noise. A perfect detector would have an AUC of 1 and the random binary guess would give an AUC of 0.5, (and the higher the better.) Fig. 3(b) shows the AUC curves against different levels of PMU accuracy for different numbers of PMUs in the network, for which from 1 to 6 line outages are randomly generated. We can see that the proposed algorithm is quite robust against the level of PMU accuracy, and gives satisfactory performance when the accuracy is better than 0.1 degree.

Finally, we validate that using five GMM components to represent the messages is sufficient for the 14-bus system. To this end, we allow the algorithm to maintain the top 20 GMM components during the message passing iterations. Fig. 3(c) shows the average probability of different GMM components, from which we can see that there are only $3 \sim 5$ dominating GMM components during the message passing iterations. Accordingly, keeping just the top 5 components appears to be sufficient for the cases that we have tested.

VI. Conclusion

Based on the sparse topology of power transmission networks, we have proposed a low complexity message passing algorithm that can effectively identify in real time an arbitrary number of simultaneous line outages. We have first established a factor graph that characterizes jointly the power transmission network and the sensor network that monitors it. With representation of the messages in parametric forms, the message passing iteration can be computed very efficiently. We have further developed simple and effective heuristics to address the challenges brought by the loopy structure of the transmission network. Simulation results on the IEEE 14-bus system demonstrate that the proposed algorithm can identify multi-line outages even with undercomplete PMU measurements, and without any assumption on the sparsity of outage patterns. The results further show that the performance is robust to PMU accuracy.

References

[1] *US-Canada Power System Outage Task Force, Final Report on the August 14, 2003 Blackout in the United States and Canada*, 2004.

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

[3] J. E. Tate and T. J. Overbye, "Double line outage detection using phasor angle measurements," in *Proc. IEEE PES General Meeting*, July 2009.

[4] Y. Zhao, J. Chen, A. Goldsmith, and H. V. Poor, "Identification of outages in power systems with uncertain states and optimal sensor locations," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 6, pp. 1140–1153, Dec. 2014.

[5] H. Zhu and G. B. Giannakis, "Sparse overcomplete representations for efficient identification of power line outages," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2215–2224, Nov. 2012.

[6] M. He and J. Zhang, "A dependency graph approach for fault detection and localization towards secure smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 342–351, June 2011.

[7] Y. Hu, A. Kuh, T. Yang, and A. Kavcic, "A belief propagation based power distribution system state estimator," *IEEE Computational Intelligence Magazine*, vol. 6, no. 3, pp. 36–46, Aug. 2011.

[8] Y. Weng, R. Negi, and M. D. Ilic, "Graphical model for state estimation in electric power systems," in *Proc. IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2013, pp. 103–108.

[9] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*, John Wiley and Sons, New York, NY, USA, 2nd edition, 1996.

[10] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[11] 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 Transactions on Power Systems*, vol. 26, no. 1, pp. 12 – 19, Feb. 2011.

[12] A. von Meier, D. Culler, A. McEachern, and R. Arghandeh, "Micro-synchrophasors for distribution systems," *Proc. IEEE Innovative Smart Grid Technologies (ISGT)*, 2013.