Topology Identification in Distribution Network with Limited Measurements

Yoav Sharon
and Anuradha M. Annaswamy
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, USA
Email: {ysharon,aanna}@mit.edu

Alexis L. Motto
and Amit Chakraborty
Corporate Research
Siemens Corporation
Princeton, NJ 08540, USA
Email: {alexis.motto,amit.chakraborty}@siemens.com

***Abstract*—We consider here the problem of detecting changes in the status of switching devices, circuit breakers in particular, in distribution networks. The lack of measurements in distribution networks compared to transmission networks is the main challenge of this problem. Using expected values of power consumption, and their variance, we are able to quickly calculate the confidence level of identifying the correct topology, or the current status of switching devices, using any given configuration of real time measurements. This allows to compare between different configurations in order to select the optimal one. The main approach we propose relies on approximating the measurements as normal distributed random variables, and applying the maximum likelihood principle. We also discuss an alternative based on support vectors. Results are demonstrated using the IEEE 123 buses distribution test case.**

I. Introduction

The concept of electric power system state estimation was initially applied to transmission networks to determine the best estimate of node voltages, generator power outputs, load demands, and branch power and current flows at a given point in time based on real-time telemetered measurements [1]. This application has generally assumed imperfect but highly redundant measurements, as well as exact power system model topology and electrical parameters. Network topology estimation is an integral part of state estimation and a critical component of modern Energy Management Systems (EMS) or Distribution Management Systems (DMS). The conventional network topology processing (NTP) function monitors the statuses of switching and switchable devices, and determines the model input to the state estimator [2]. Circuit breaker statuses, isolator switch statuses, fuses statuses, and transformer tap positions are examples of real-time inputs used by the network topology processor. A conventional NTP determines the connectivity of the electrical network, taking as input a complete model of the network, comprising of nodes and switching devices. The NTP reduces the node-switching-device model to a "bus-branch" model, where the concept of bus defines a maximal subnetwork interconnecting nodes and closed switching devices only. The objective of the conventional NTP is to eliminate all switching devices from the network model, by instantiating their "open" or "closed" statuses. The NTP achieves this instantiation by processing the switching device user-defined, measured, scheduled or normal status, as available in this order of precedence. The conventional state estimation sub-program then solves and analyzes the resulting bus-branch model. Undetected switching device status errors during estimation show up as analog measurement errors in the solution, which are difficult to distinguish from actual analog measurement errors [3]. Hence, reliable and prompt detection of the switching device statuses is crucial for accurate state estimation. The output of the state estimator is a critical input to nearly all other network analysis, security, control and stability assessment applications.

In distribution grid management, a critical task of a system operator is to take quick actions to restore continuity of electric power supply following forced outages. For many distribution networks, however, the measurement redundancy is so low that the first and often only indications of an outage are telephone calls from customers reporting loss of supply. In the mostly radial topologies of distribution network, the opening of a normally-closed switching device generally results in some loss of electric power supply. Clearly, the analysis performed by aggregating and mapping multiple customer calls into a suspected common network device, such as a fuse, which is then suspected to be open, is an instance of network topology estimation. Many existing outage management system (OMS) are still based on the process of call aggregation, which can take from tens of minutes to hours (if happens at night for example) to identify the culprit device. Any automatic procedure that will dramatically reduce the detection time will lead to a much better quality of service and higher revenues to utility companies.

In typical distribution systems, at least in the current state of low penetration of distributed energy sources and communication devices, estimating the network topology is probably more important than estimating the analog variables [4].

That said, the tools for the detection of circuit breakers status in transmission networks do not apply to distribution networks. As the main goal in transmission networks is full state estimation, there is already a redundancy of measurements to state variables in the range of 1.7 to 2.2 (redundancy factor) [2]. In contrast distribution networks today may have only few measurements at the substation. On the other hand, while distribution networks usually have many more buses compared to transmission networks, relatively they have substantially fewer circuit breakers. Equipping every circuit breaker with a sensor would thus allow for immediate detection of the status while still using few measurements and certainly less than is needed for state estimation. We argue that with the decrease in the cost of sensors and communication, and the potential benefit to utility companies, there is room for adding more sensors in places other than the substation. Through this work we aim to provide a tool for optimizing the location of sensor placement, allowing for the detection of the status of all the circuit breakers while using fewer measurements than the number of such breakers.

While we don't provide an automatic tool for optimal placement of sensors, we provide a tool that given the placement of sensors, quickly reveals at what confidence one can detect the status of breakers. This then allows to compare easily several configurations of sensors placement, and select the one with the highest level of confidence.

The remaining sections are organized as follows. Section II presents the formulation of the proposed network topology identification problem. Section III outlines the steps we propose to take in order to address this problem. Section IV discusses an alternative approach that we investigated, using support vectors. Results using an IEEE PES 123-bus system are shown in Section V.

II. Problem Formulation

We are given a power network structure which includes buses, lines, the lines' impedance between the connected buses, the lines' admittance to the ground, the buses' admittance to the ground, the position of circuit breakers, and the nominal position of the breakers. We are also given a set of real time measurements which may include voltage, current and power flow readings. Finally, we are given the mean and covariance of the power injection (or extraction) for all buses except one, denoted as the slack bus. For distribution networks it is customary to choose the substation connecting the distribution network to the transmission network as the slack bus. Different sets of mean and covariance are expected for different times of the day (morning, afternoon, evening and night), as well as for different seasons. The objective is to find the breakers' actual position from the measurements, or identify the current network topology. Our goal is to measure the confidence of this identification, or in other words, calculate the probability of identifying one topology as the current one, when in fact it is not.

For any given topology $\mathcal{T}$, which includes the breakers' position, we use $L_{\mathcal{T}}$ to denote the set of buses which have direct or indirect connection to the slack bus. We use the notation $|L_{\mathcal{T}}|$ to denote the number of such buses. As we focus on distribution networks, we assume that buses with no connection to the slack bus (islanding) have zero voltage. We use $x_{\mathcal{T}}$ to denote the current state of the network. The state of the system includes the real and imaginary parts of the voltage at every bus that is in $L_{\mathcal{T}}$. Following [5], if two or more buses are connected to each other through a zero impedance line (or closed circuit breaker), then since they all have the same voltage, we have two state variables representing the real and imaginary parts of the voltage, and additional state variables representing the real and reactive power flow between the buses. Two zero-impedance lines connecting the same two buses are considered as one line. In any case the degrees of freedom in the system become $2\,|L_{\mathcal{T}}|$ ($x_{\mathcal{T}} \in \mathbb{R}^{2|L_{\mathcal{T}}|}$).

Remark: In this section and in the subsequent section, we consider every phase of a bus in a 3-phase system as a separate bus.

We use the classic convention and classify our measurements as real time measurements and as pseudo-measurements. Real time measurements are measurements we get from sensors readings. We use $y \in \mathbb{R}^m$ where m is the number of real time measurements to denote the real time measurements, and define the nonlinear measurement function f such that $y = f_{\mathcal{T}}(x_{\mathcal{T}})$. Pseudo-measurements are known constraints given by the topology of the network. For example, a bus to which no generator or load is connected, has zero real and reactive power injection (two pseudo-measurements). The voltage and the angle (which is arbitrarily set to 0) of the slack bus are another two pseudo-measurements which we use. We define the nonlinear constraint function g such that $g_{\mathcal{T}}(x_{\mathcal{T}}) = 0$ if and only if the state $x_{\mathcal{T}}$ satisfies the pseudo-measurements. Let r be the number of pseudo-measurements. We define n' as the degrees of freedom remaining after constraining the system to the pseudo-measurement. If all the pseudo-measurements are independent of each other (i.e. any single pseudo-measurement can be unsatisfied while all other pseudo-measurements are satisfied), then $n' = 2\,|L_{\mathcal{T}}| - r$.

If the number of real time measurements is larger than n', then in the absence of any measurements errors, each valuation of the measurements can only correspond to one topology. Another way to say this is that in this case

$$\left\{ y \in \mathbb{R}^m \,\middle|\, \exists x_{\mathcal{T}} \in \mathbb{R}^{2|L_{\mathcal{T}}|} : f_{\mathcal{T}}(x_{\mathcal{T}}) = y, g_{\mathcal{T}}(x_{\mathcal{T}}) = 0 \right\}$$

is a manifold of dimension strictly less than m and two such manifolds intersect each other over a set of measure zero. In such case theoretically one can identify the correct topology with a confidence level of 100%. In practice the computational issue of finding the right topology is still a major challenge (see [6] for a recent attempt at addressing this issue), and in addition measurement errors can bring the confidence level down. Nevertheless, here we are interested in the case where $m < n'$. While in transmission networks typical ratio of measurements to state variables is 1.7–2.2 [2], in distribution networks the number of measurements is indeed much lower than the number of buses.

Let $z_{\mathcal{T}} \in \mathbb{R}^{2|L_{\mathcal{T}}|}$ be a vector consisting of the real and imaginary parts of the slack bus voltage, and the $2\,(|L_{\mathcal{T}}| - 1)$ real and reactive power injections at all other buses in $L_{\mathcal{T}}$. Let $f_{z,\mathcal{T}}$ be the function such that $z_{\mathcal{T}} = f_{z,\mathcal{T}}(x_{\mathcal{T}})\ \forall x_{\mathcal{T}}$. While it is not hard to synthesize special cases where the following is not true, in practice $f_{z,\mathcal{T}}$ is almost always one-to-one. This implies there exists a function $h_{\mathcal{T}}$ such that $f_{z,\mathcal{T}}(h_{\mathcal{T}}(z_{\mathcal{T}})) = z_{\mathcal{T}}\ \forall z$.

Let $z \in \mathbb{R}^{2n}$ be the vector consisting of the real and imaginary parts of the slack bus voltage, and the $2(n-1)$ real and reactive power injections at all other buses. Let $I \in \mathbb{R}^{2n\times 2n}$ be the identity matrix, and let $I_{\mathcal{T}} \in \mathbb{R}^{2|L_{\mathcal{T}}|\times 2n}$ be a matrix derived from I by keeping only the rows whose indices are the same as the indices of the components of z corresponding to buses in $L_{\mathcal{T}}$. Thus if $\mathcal{T}$ is the active topology, then $z_{\mathcal{T}} = I_{\mathcal{T}}z$. By the problem description we are given the mean and covariance matrix of all real and reactive power injections except for the slack bus. The mean and the variance of the power injection for buses to which no load or generator are connected will both be zero naturally. Let $\mu_z \in \mathbb{R}^{2n}$ be a vector whose first two components are the real and imaginary parts of the slack bus, and its remaining components are the mean and reactive power injection for all other buses. Similarly, let $\Lambda_z \in \mathbb{R}^{2n\times 2n}$ be a matrix whose first two rows and first two columns are zeros, and its bottom right $2(n-1)\times 2(n-1)$ block equal the covariance matrix of the power injections. We can therefore, by assuming normal distribution, define the probability distribution function $\rho(\cdot;\mu_z,\Lambda_z) : \mathbb{R}^{2n} \to \mathbb{R}_{>0}$ using μ_z and Λ_z. We then say that z follows the normal distribution using the notation $z \sim \mathcal{N}(\mu_z,\Lambda_z)$. We note that μ_z and Λ_z do not depend on the topology. They represent the variation in demand or generation assuming the whole system is connected (no islanding). They do not represent the actual power delivered, which may be zero if the corresponding bus is disconnected from the slack bus.

Let $c : \mathbb{R}^m \to \{1,\ldots,p\}$ be the topology identification function, where p is the number of possible topologies: $\mathcal{T}_1,\ldots,\mathcal{T}_p$. We define the confidence level as:

$$1 - \max_i \operatorname{Prob}\left\{c\left(f_{\mathcal{T}_i}\left(h_{\mathcal{T}_i}\left(\mathcal{I}_{\mathcal{T}}z\right)\right)\right) \neq i \,|z \sim \mathcal{N}(\mu_z,\Lambda_z)\right\}. \tag{1}$$

III. Minimizing Classification Errors

Our first goal is for each topology $\mathcal{T}$, to approximate the random variable $y = f_{\mathcal{T}}(h_{\mathcal{T}}(I_{\mathcal{T}}z))$ as a normal distributed random variable. By linearizing, we can write

$$y \approx f_{\mathcal{T}}(h_{\mathcal{T}}(I_{\mathcal{T}}\mu_z)) + A_{\mathcal{T}}(z-\mu_z)$$

where

$$A_{\mathcal{T}} \doteq \frac{\partial f_{\mathcal{T}}}{\partial x_{\mathcal{T}}}_{|x_{\mathcal{T}}=h_{\mathcal{T}}(I_{\mathcal{T}}\mu_z)} \frac{\partial h_{\mathcal{T}}}{\partial z_{\mathcal{T}}}_{|z_{\mathcal{T}}=I_{\mathcal{T}}\mu_z} I_{\mathcal{T}},$$

and approximate y as $y \sim \mathcal{N}(\mu_{y,\mathcal{T}};\Lambda_{y,\mathcal{T}})$ where

$$\mu_{y;\mathcal{T}} = f_{\mathcal{T}}(h_{\mathcal{T}}(I_{\mathcal{T}}\mu_z))$$

and

$$\Lambda_{y;\mathcal{T}} = A_{\mathcal{T}}\Lambda_z A_{\mathcal{T}}^T.$$

While the functions $f_{\mathcal{T}}$ and $f_{z,\mathcal{T}}$ can be written explicitly as functions of $x_{\mathcal{T}}$, and thus $\partial f_{\mathcal{T}}/\partial x_{\mathcal{T}}$ can be easily calculated, this is not true for $h_{\mathcal{T}}$. Yet, since $\partial f_{z,\mathcal{T}}/\partial x_{\mathcal{T}}$ is full rank as shown in [7, Proposition 2] whenever $f_{z,\mathcal{T}}$ is one-to-one, $\partial h_{\mathcal{T}}/\partial z_{\mathcal{T}|z_{\mathcal{T}}} = \left(\partial f_{z,\mathcal{T}}/\partial x_{\mathcal{T}|x_{\mathcal{T}}=h_{\mathcal{T}}(z_{\mathcal{T}})}\right)^{-1}$.

Define

$$\begin{aligned} e_i &= \operatorname{Prob}\left\{c\left(f_{\mathcal{T}_i}\left(h_{\mathcal{T}_i}\left(\mathcal{I}_{\mathcal{T}}z\right)\right)\right) \neq i \,|z \sim \mathcal{N}(\mu_z,\Lambda_z)\right\} \\ \tilde{e}_i &= \operatorname{Prob}\left\{c(y) \neq i \,|y \sim \mathcal{N}(\mu_{y,\mathcal{T}_i},\Lambda_{y,\mathcal{T}_i})\right\} \end{aligned}$$

By our definition of confidence level (1), we would have wanted to minimize $\max_i e_i$. Due the complexities arising from the nonlinearities, instead we would focus on minimizing $\max_i \tilde{e}_i$. However, even for that we are not aware of a viable solution. Therefore we will minimize $\sum_i \tilde{e}_i$ for which the solution is the maximum likelihood (ML):

$$c(y) = \arg\max_i \rho(y;\mu_i,\Lambda_i) \tag{2}$$

where we used for short $\mu_i \doteq \mu_{y,\mathcal{T}_i}$ and $\Lambda_i \doteq \Lambda_{y,\mathcal{T}_i}$. To see why using (2) indeed minimizes $\sum_i \tilde{e}_i$, simply note that

$$\begin{aligned} \sum_i \tilde{e}_i &= \sum_i \int_{\{y\in\mathbb{R}^m|c(y)\neq i\}} \rho(y;\mu_i,\Lambda_i)\,\mathrm{d}\,y \\ &= \int_{\mathbb{R}^m} \sum_i \mathcal{I}_{c(y)\neq i}\rho(y;\mu_i,\Lambda_i)\,\mathrm{d}\,y \end{aligned} \tag{3}$$

where $\mathcal{I}_{a(y)}(y)$ is the indicator function which is equal to one if the conditional statement $a(y)$ is true and equal to zero if it is false.

Calculating (3) directly with c as defined in (2) by integrating the normal density distribution function ρ can be done numerically, but the computational complexity grows exponentially with the dimension of the measurement space, m, if accuracy is to be maintained. A good alternative is then to randomly generate enough samples of y for each topology i, and count for how many of these sample, $c(y) \neq i$. The computational complexity of this approach is still linear in the number of topologies, which in turn can be exponential in the number of switches. However, it now grows polynomially with the measurement space dimension.

A. Technicalities

In computing $\frac{\partial f}{\partial x}$, we found it is easier to use Cartesian (real and imaginary) coordinates rather the polar (magnitude and angle) coordinates. Let x_{1R} be the section of x corresponding to the 3 real parts of the voltages of bus 1 in a 3-phase system. Let x_{1I} be the imaginary counterpart. Let $Y = 1/Z$, $Y, Z \in \mathbb{C}^{3\times 3}$, be the complex admittance matrix of the line connecting bus 1 and bus 2, and $\operatorname{Re} Y$ and $\operatorname{Im} Y$ its real and imaginary part, respectively. The complex current flowing from bus 1 to bus 2 is given by:

$$I = Y\left(x_{2R} - x_{1R} + j\left(x_{2I} - x_{1I}\right)\right)$$

where here $j = \sqrt{-1}$. The power flow exiting bus 1 toward bus 2 is $P_1 = (x_{1R} + jx_{1I})\cdot\bar{I}$, $P_1 \in \mathbb{C}^3$, where $\cdot$ is elementwise multiplication and $\bar{I}$ is the conjugate of I. Thus,

$$\frac{\partial \operatorname{Re} P_1}{\partial x_{1R}} = \operatorname{diag}\left(\operatorname{Re} Y\left(x_{2R} - x_{1R}\right) - \operatorname{Im} Y\left(x_{2I} - x_{1I}\right)\right) - \operatorname{diag}\left(x_{2R}\right) \operatorname{Re} Y - \operatorname{diag}\left(x_{1I}\right) \operatorname{Im} Y,$$

$$\frac{\partial \operatorname{Re} P_1}{\partial x_{2R}} = \operatorname{diag}\left(x_{2R}\right) \operatorname{Re} Y + \operatorname{diag}\left(x_{1I}\right) \operatorname{Im} Y,$$

with similar expressions for $\partial \operatorname{Re} P_1 / \partial x_{1I}$, $\partial \operatorname{Im} P_1 / \partial x_{1R}$, $\partial \operatorname{Im} P_1 / \partial x_{1I}$, $\partial \operatorname{Re} P_1 / \partial x_{2I}$, $\partial \operatorname{Im} P_1 / \partial x_{2R}$, $\partial \operatorname{Im} P_1 / \partial x_{2I}$, where $\operatorname{diag}(x)$ is a matrix whose diagonal is the vector x and it is zero outside its diagonal. Note that by following this way we do not constrain ourselves to the standard linearization technique involving the decoupling of the 'voltage angle'-'real power' and 'voltage magnitude'-'reactive power' dependencies. Power injections are then just linear combinations of the lines power flows.

For power flow calculation, or solving for $x = h_{\mathcal{T}}(z)$ for which there is no explicit expression, we used the standard Newton-Raphson method, updating

$$\hat{x}_{k+1} = \hat{x}_k + \left(\frac{\partial f_z}{\partial x}_{|x=\hat{x}_k}\right)^{-1} \left(z - f_z\left(\hat{x}_k\right)\right) \tag{4}$$

through several iterations until convergence.

IV. An Alternative Approach

We report here an alternative approach to the maximum likelihood method, an approach we investigated during this research but found to be inferior to the ML method.

For each pair of topologies, i and j ($i \neq j$), we define the identification function $c_{ij} : \mathbb{R}^m \to \{i, j\}$ similarly to c except that it only distinguishes between topologies i and j. We then construct c as:

$$c(y) = i \text{ if and only if } \forall j \neq i : c_{ij}(y) = i \tag{5}$$

With this we can lower bound $1 - \max_i \tilde{e}_i$ which itself is an approximation to (1) $= 1 - \max_i e_i$ as

$$1 - \max_i \int_{\cup_{j \neq i}\{y \in \mathbb{R}^m | c_{ij}(y) = j\}} \rho\left(y; \mu_i, \Lambda_i\right) \mathrm{d}\, y \geq 1 - \max_i \sum_j \int_{\{y \in \mathbb{R}^m | c_{ij}(y) = j\}} \rho\left(y; \mu_i, \Lambda_i\right) \mathrm{d}\, y. \tag{6}$$

To simplify the calculation of this lower bound we construct c_{ij} using some $\alpha_{ij} \in \mathbb{R}^m$ and $\beta_{ij} \in \mathbb{R}$ as

$$c_{ij}(y) = \begin{cases} i & \alpha_{ij}^T y \leq \beta_{ij} \\ j & \alpha_{ij}^T y > \beta_{ij} \end{cases} \tag{7}$$

if $i < j$ and $c_{ij}(y) = c_{ji}(y)\ \forall y$ otherwise. By approximating y as a normal distributed random variable, $\alpha_{ij}^T y$ given topology $\mathcal{T}_i$ becomes a one dimensional normal distributed random variable with mean $\alpha_{ij}^T \mu_i$ and variance of $\alpha_{ij}^T \Lambda_i \alpha_{ij}$. In this case the bound in (6) becomes

$$1 - \max_i \sum_j e_i^{ij}, \qquad e_i^{ij} \doteq 1 - \Gamma\left(\beta_{ij}; \alpha_{ij}^T \mu_i, \alpha_{ij}^T \Lambda_i \alpha_{ij}\right) \tag{8}$$

where by convention, $\alpha_{ij} = -\alpha_{ji}$, $\beta_{ij} = -\beta_{ji}\ \forall i, j$, and $\Gamma\left(x; \mu, \sigma^2\right)$ is the one-dimensional normal *cumulative* distribution function with mean μ and variance σ^2. Having reduced to one-dimensional normal distribution, regardless of the dimension of y, allows us not only to quickly calculate (8). In contrast to (2) which can only minimizes the cost function $\sum_i \tilde{e}_i$, using (7) we have more freedom in choosing the cost function. In particular we can minimize $\max_i \tilde{e}_i$ over all the α's and β's using constrained nonlinear minimization where the gradients of the cost function and all the constraints have explicit analytic expressions. However, because the maximum likelihood function (2) is not in the family of functions defined by (7) and (5), it is possible, and indeed we found this to be the case, that using (2) will still result in a lower $\max_i \tilde{e}_i$ than have we used this alternative approach.

To demonstrate how to minimize $\max_i \tilde{e}_i$ over α's and β, consider we only have two topologies to distinguish between, i and j. Minimizing $\max\{e_i, e_j\}$ can be cast as an analytic nonlinear constrained optimization over $m + 2$ variables:

$$\begin{aligned} \text{minimize} \quad & e + \frac{1}{4}\left(\left\|\alpha_{ij}\right\|^2 - 1\right)^2 \\ \text{subject to} \quad & e > 1 - \Gamma\left(\beta_{ij}; \alpha_{ij}^T \mu_i, \alpha_{ij}^T \Lambda_i \alpha_{ij}\right) \\ & e > \Gamma\left(\beta_{ij}; \alpha_{ij}^T \mu_j, \alpha_{ij}^T \Lambda_j \alpha_{ij}\right). \end{aligned} \tag{9}$$

One can use for example MATLAB's *fmincon* function to solve for (9) using the sequential quadratic programming method (SQP) [8, Chapter 18]. For this function to run efficiently, we need to supply it with the derivatives of the cost function and the constraints with respect to all the variables. These are listed below:

$$\frac{\partial \frac{1}{4}\left(\|\alpha\|^2 - 1\right)^2}{\partial \alpha} = \left(\|\alpha\|^2 - 1\right) \alpha^T,$$

$$\frac{\partial \Gamma\left(\beta; \alpha^T \mu, \alpha^T \Lambda \alpha\right)}{\partial \beta} = \rho\left(\beta; \alpha^T \mu, \alpha^T \Lambda \alpha\right),$$

$$\frac{\partial \Gamma\left(\beta; \alpha^T \mu, \alpha^T \Lambda \alpha\right)}{\partial \alpha} = \rho\left(\beta; \alpha^T \mu, \alpha^T \Lambda \alpha\right) \left(-\mu^T - \frac{\beta - \alpha^T \mu}{\alpha^T \Lambda \alpha} \alpha^T \Lambda\right).$$

This approach follows closely that of support vector machine (SVM) [9]. A traditional SVM approach to our problem would proceed as follows. Generate s random power injection profiles based on μ_z and Λ_z. For each such profile, solve the nonlinear power flow calculation and find the measurement vector values. Then run the standard SVM between each two pair of topologies to best separate between the measurement points belonging to each topology.

The main advantage of the traditional SVM is that by working on the measurements derived from the nonlinear power flow calculation, there is no need to revert to linearization approximation. However, there are two disadvantages to the traditional SVM. The first is that for the results to be reliable, s must be very large. And while nonlinear power flow

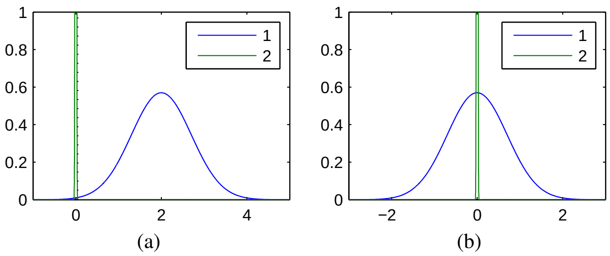

Fig. 1. Examples of a two bus system power flow measurements distribution

calculation can be evaluated quite fast, its computation time is not negligible, and when the time for a single calculation needs to be multiplied by s and by the number of topologies, it can come to a total computation time which is very substantial. In contrast in the approach described above only one power flow calculation per topology is needed. The second disadvantage is that ideally SVM would have minimized the number of misclassified points from each topology, which is a good proxy for the confidence level when s is large enough. However, SVM does not do that but rather minimize the distance from the support vector to the misclassified point which is the farthest away from the support vector. This will not necessarily lead to maximizing the confidence level, as there can still be many misclassified points.

The following example may explain why using the maximum likelihood will provide better results then the alternative approach. The support vector is good in separating distributions which are centered around distant means. Consider a system with two buses. One is the slack bus, and the other is connected to a consumer load. There is one power flow sensor on the circuit breaker connecting the two buses. In Figure 1a we show the probability density function of the sensor reading corresponding to the two possible topologies

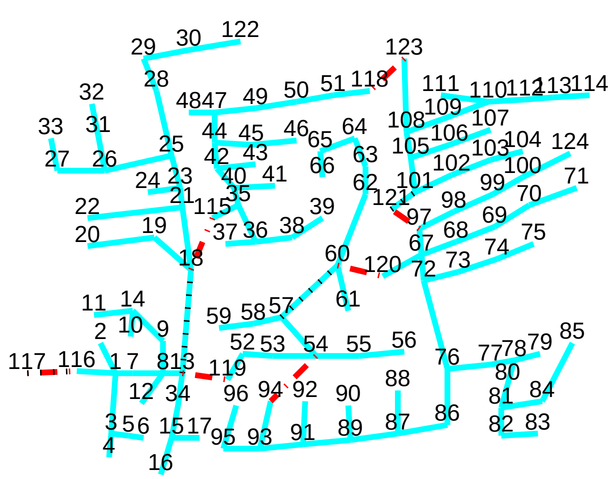

Fig. 2. IEEE PES 123-bus feeder test case. The numbers are the buses indices. Cyan continuous lines are regular transmission lines with impedance. Red dashed lines are where the circuit breakers are. Black dashed lines are where the power flow sensors are placed. Bus 117 at the bottom left corner is the slack bus.

TABLE I

ERROR RATE FOR EACH BREAKER: THE PROBABILITY THAT THE ALGORITHM WILL DECIDE THAT ANOTHER BREAKER CHANGED ITS POSITION, WHEN IN FACT IT WAS THE BREAKER LISTED THAT HAD CHANGED ITS POSITION. THE LETTER IN BRACKETS IN THE FIRST COLUMN REPRESENTS THE PHASE.

breaker	approximated error rate (%)	non-approximated error rate (%)
nominal	4.5	4
13-119 (a)	< 0.1	0
13-119 (b)	0.1	0
13-119 (c)	2.8	5
18-115 (a)	2.3	0
18-115 (b)	< 0.1	1
18-115 (c)	1.0	1
60-120 (a)	< 0.1	0
60-120 (b)	0.6	2
60-120 (c)	7.2	5
97-121 (a)	< 0.1	0
97-121 (b)	0.6	1
97-121 (c)	1.4	2
117-116 (a)	< 0.1	0
117-116 (b)	< 0.1	0
117-116 (c)	< 0.1	0
118-123 (a)	< 0.1	0
118-123 (b)	< 0.1	0
118-123 (c)	< 0.1	0
54-94 (a)	< 0.1	0

where in the first the circuit breaker is closed and in the second it is open. It is easy to see how we can get very good separation by placing the support vector just right of the peak corresponding to the second topology. Now consider that a distributed energy source (DER) such as photo-voltaic receptor is added to the second bus, and that on *average* the DER produces as much power as the consumer demands. The probability density functions corresponding to the two topologies are shown in 1b. Using the maximum likelihood approach would still wield high confidence even for the second case with DER, but with the support vector approach one cannot get above 50% confidence for both topologies using one support vector.

V. RESULTS

We tested this approach using the IEEE PES 123-bus feeder distribution test case [10], depicted in Figure 2, which is a 3-phase system. There are 7 pairs of buses connected with circuit breakers. Between buses 54 and 94 there is only one breaker on phase a. All other pairs of buses with circuit breakers have breakers on all 3 phases. The breakers connecting buses 119 and 123 in the top right corner, and the breaker connecting buses 54 and 94 in the middle bottom, are nominally open. All other breakers are nominally closed. The test case includes real and reactive load values at every bus. We took these values as mean power extraction. We divided these values by 2 and used that as the standard deviation of the power extraction.

We added 9 real power flow sensors where the dashed lines are shown in Figure 2: 3-phase measurements next to the substation (bus no. 117), 3-phase measurements between bus 121 and bus 101, 1-phase measurement (phase a) between bus

57 and bus 60, and 2-phase measurements (phases b and c) between bus 13 and 18. The decision on where to place the sensor was based on manual trial and error. We then used our approach to find the confidence level of identifying whether any of the breakers changed from its nominal position. We assumed that only one breaker may have changed its position. The results, using 1,000 samples drawn for each topology according to $y \sim \mathcal{N}(\mu_i, \Lambda_i)$, are listed in Table I. The results show a confidence level of 92.8%. We then compared our approximated confidence level to empirical evidence using Monte Carlo simulations without the linear approximation. We randomly generated power injection profiles according to $z \sim \mathcal{N}(\mu_z, \Lambda_z)$. For each breaker, we created the topology with this breaker changed from its nominal position, performed the nonlinear power flow calculation, derived the sensor readings, and checked whether our approach concludes that this was the breaker that changed position. The results are very close to our approximated values. Due to the much longer running time, arising from the power flow calculations, we only generated 100 samples per topology. And despite having 10 times less samples, the total running time was still about 10 times longer. This explains the coarseness of the results in the right column.

VI. Conclusions

We addressed here the issue of topology identification in distribution networks using as few measurements as possible. While not providing an automatic tool for optimal sensor placement, we proposed a tool for fast and reliable comparison between different sensor placements, thus allowing to find an optimal placement through trial and error. Results from an IEEE PES distribution feeder test case show the potential of our proposed tool. Extensions for this tool include the possibility to distinguish between topologies involving several circuit breakers changing their status simultaneously, and without having the running time increase exponentially. Another extension is an automatic tool for optimal sensor placement based on our proposed tool.

Acknowledgment

This work was supported by Siemens CKI fund.

References

[1] F. C. Schweppe and E. J. Handschin, "Static state estimation in electric power systems," *Proc. IEEE*, vol. 62, no. 7, pp. 972–982, 1974.

[2] A. Monticelli, "Electric power system state estimation," *Proc. IEEE*, vol. 88, no. 2, pp. 262–282, 2000.

[3] J. C. Pereira, J. T. Saraiva, V. Miranda, A. S. Costa, E. Lourengo, and K. Clements, "Comparison of approaches to identify topology errors in the scope of state estimation studies," in *2001 IEEE Power Tech Proc.*, vol. 3, Porto, Portugal, 2001.

[4] R. Hoffman, "Practical state estimation for electric distribution networks," in *Proc. Power Systems Conference and Exposition, 2006 (PSCE'06)*, 2006.

[5] O. Alsaç, N. Vempati, B. Stott, and A. Monticelli, "Generalized state estimation," *IEEE Trans. Power Syst.*, vol. 13, no. 3, pp. 1069–1075, 1998.

[6] E. Caro, A. J. Conejo, and A. Abur, "Breaker status identification," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 694–702, 2010.

[7] G. Tran, A. Kiani, A. Annaswamy, Y. Sharon, A. L. Motto, and A. Chakraborty, "Necessary and sufficient conditions for observability in power systems," in *Proc. 2012 IEEE Innovative Smart Grid Technologies Conference (ISGT'12)*, Washington DC, USA, submitted.

[8] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed., ser. Operation Research and Financial Engineering, T. V. Mikosch, S. I. Resnick, and S. M. Robinson, Eds. New York, NY: Springer, 2006.

[9] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[10] IEEE Power & Energy Society. Distribution test feeders. [Online]. Available: ewh.ieee.org/soc/pes/dsacom/testfeeders