Efficient Identification of Bus Split Events Using Synchrophasor Data

Yuqi Zhou, Student Member, IEEE, and Hao Zhu, Senior Member, IEEE

Abstract—Accurate grid topology information is of paramount importance for routine power system operations, while the growing availability of synchrophasor data offers the opportunity to identify topology changes in real time. Identification of bus split events, where the substation becomes electrically disconnected, is becoming increasingly important for maintaining the security of power systems. This paper aims to provide an efficient modeling and monitoring framework for bus split events. The linear sensitivity analysis is first performed to quickly evaluate the grid-wide impact of such events. Furthermore, the synchrophasor data enabled identification problem is formulated by matching the changes in bus phase angles (and possibly line flows). To address the resultant bilinear multiplication involving the binary connectivity variables, the McCormick relaxation technique is leveraged to attain an equivalent mixed-integer linear program reformulation that is efficiently solvable. Numerical studies on the IEEE 14-bus and 300-bus systems demonstrate the validity and efficiency of the proposed identification algorithm towards real-time implementation.

Index Terms—Bus split, sensitivity analysis, mixed-integer program, phasor measurement units.

I. INTRODUCTION

A CCURATE information of grid topology is crucial for performing various power system operation and maintenance tasks [1, Ch. 12]. Recently, the power grids have witnessed more frequent occurrence of unintentional circuit breaker (CB) actions, due to either misoperations of substation protection systems under renewable generation [2] or malicious attacks as in the 2015/16 Ukrainian blackouts [3]. Changes of CB status not only lead to the disconnection of power lines or generation/loads, but also can give rise to the bus split events when a substation’s bus bars become disconnected [4]. Recent reports [5]–[7] point out the increasing importance of developing efficient grid topology modeling and monitoring techniques that can include bus split events.

Traditionally, the grid topology is processed and updated by the energy management system (EMS) using the Supervisory Control and Data Acquisition (SCADA) inputs [8, Ch. 1]. A compact bus-branch model is typically obtained by examining the statuses of CBs and switching devices. Topology errors due to incorrect reports of CB statuses can be identified by the generalized state estimator using SCADA data; see e.g., [8, Ch. 8] and references therein. Recent deployment of phasor measurement units (PMUs) provides the synchrophasor data of high sampling rates and resolution, which can greatly facilitate the process of topology identification. Fast algorithms for identifying transmission line outages has been developed in e.g., [9]–[11], and also for distribution grids [12]–[14]. Unfortunately, the identification of CB status and bus split events has been rarely considered using synchrophasor data. A detailed node-breaker model is typically required to capture the CB statuses and substation connectivity, but this representation is known to increase the problem dimensionality and computational time [15], [16]. Detecting CB actions from the synchrophasor data has been considered in [17], where lack of identifiability has been observed therein due to the complicated node-breaker model. Thus, the problem of efficiently analyzing and monitoring bus split events using the latest synchrophasor data still remains open.

The goal of this paper is to develop an efficient modeling and monitoring framework for bus split events based on the concise bus-branch representation. For the modeling part, we approach the dc power flow model [18] based sensitivity analysis through simplifying the matrix computations. The sensitivity analysis results allow for constructing an equivalent bus-branch model for the post-split system, and are essential for executing fast contingency screening and security-constrained economic dispatch tasks; see e.g., [1, Ch. 11]. For the monitoring problem, it is possible identify the bus split event by searching for the post-split scenario that best matches the actual synchrophasor measurements. By enumerating all possible bus split scenarios, one can perform the sensitivity analysis for each scenario to compute the corresponding post-split phase angles. Nonetheless, the number of scenarios in this exhaustive search approach would grow exponentially with the number of substation connections, limiting it from real-time implementation. To address this complexity issue, we leverage the popular McCormick relaxation technique [19] to replace the nonconvex bilinear terms with equivalent convex linear inequalities. This reformulation leads to a tractable mixed-integer linear programming problem that is efficiently solvable for real-time identification. By systematically incorporating bus split events, our work directly addresses the lack of consideration for such events in the traditional grid analysis and monitoring framework.

The rest of the paper is organized as follows. Section II introduces the power flow model in matrix form and the bus split representation. Section III presents the linear sensitivity analysis for modeling the post-split line flows and bus phase angles. In Section IV, the synchrophasor data based identification problem is formulated, along with the development of the tractable reformulation via McCormick relaxation. Numerical studies using the IEEE 14-bus and 300-bus systems are presented in Section V to demonstrate the validity and efficiency of the proposed identification algorithm, and the paper is wrapped up in Section VI.

Notation: Upper (lower) boldface symbols stand for ma-
trices (vectors); \((\cdot)^{T}\) stands for transposition; \(I\) for identity matrix; \(|\cdot|\) denotes the cardinality of a set; \(||\cdot||\) denotes the vector norm; \(e_i\) denotes the standard basis vector with all entries being 0 except for the \(i\)-th entry equals to 1.

II. SYSTEM MODELING

We introduce the dc power flow model [18] for the modeling of bus split events. Consider a transmission system with \((N + 1)\) buses collected in the set \(\mathcal{N} := \{0, 1, \ldots, N\}\), and \(L\) transmission lines represented by the set \(\mathcal{L} := \{(i, j)\}\). For each bus \(i\), we use \(N_i\) to denote the set collecting its adjacent buses, and let \(\theta_i\) denote its voltage angle, as well as \(g_i\) and \(d_i\) as its connected generation and load, respectively. Hence, the power injection per bus \(i\) is \(p_i = g_i - d_i\). Without loss of generality (Wlog), we set bus 0 to be the reference angle bus with \(\theta_0 = 0\). All non-reference phase angles are concatenated into vector \(\theta \in \mathbb{R}^N\); similarly for \(p \in \mathbb{R}^N\). It is worthy mentioning that the dc power flow model is effective for the analysis and monitoring formulations considered in this paper, and our numerical results in Section V have corroborated this using the full ac power flow model.

For each line \((i, j)\), the power flow from bus \(i\) to bus \(j\) is denoted by \(f_{ij}\) and given by:

\[
f_{ij} = \frac{1}{x_{ij}}(\theta_i - \theta_j) = b_{ij}(\theta_i - \theta_j), \forall(i, j) \in \mathcal{L} \tag{1}
\]

where \(x_{ij}\) is the line reactance and its inverse equals to \(b_{ij} = 1/x_{ij}\). Concatenating (1) into matrix form gives rise to the flow vector \(f \in \mathbb{R}^L\), as:

\[
f = K\theta \tag{2}
\]

where matrix \(K \in \mathbb{R}^{L \times N}\) captures the network topology. The \(\ell\)-th row of \(K\) corresponding to line \((i, j)\) is given by \(b_{ij}(e_i - e_j)^T\), where \(e_i \in \mathbb{R}^N\) is the standard basis vector. Due to nodal flow conservation, one can sum up all the line flows to form the injected power \(p\), as given by:

\[
p = B\theta \tag{3}
\]

where the \textit{Bus matrix} \(B \in \mathbb{R}^{N \times N}\) is invertible with each entry:

\[
B_{ij} = \begin{cases} 
\sum_{k \in N_i} b_{ik}, & \text{if } i = j \\
-b_{ij}, & \text{if } (i, j) \in \mathcal{L} \\
0, & \text{otherwise}
\end{cases} \tag{4}
\]

Therefore, matrix \(B\) can be also given by:

\[
B = \sum_{(i, j) \in \mathcal{L}} b_{ij}(e_i - e_j)(e_i - e_j)^T. \tag{5}
\]

By solving for \(\theta\) in (3), one can write the line flow as \(f = (KB^{-1})p\), with the coefficients in \((KB^{-1})\) termed as the injection shift factors (ISFs) that can transform from the injection \(p\) to line flow \(f\).

We are interested in the grid topology changes due to \textit{bus split} within substations. Substations can be described as electrically connected nodes where multiple transmission lines terminate. For reliable and modular design, substations are equipped with switching equipment such as CBs and isolators

III. LINEAR SENSITIVITY ANALYSIS FOR BUS SPLIT

Consider the split of bus \(i\) leading to a new bus \(i'\), as illustrated by Fig. 3. After the split, let the subset \(\mathcal{J} \subseteq N_i\) collect the adjacent buses reconnected to the new bus \(i'\). Moreover, denote \(\tilde{\theta} = [\theta' \ \tilde{\theta}_{i'}] \in \mathbb{R}^{N+1}\) as the new angle vector after splitting; and similarly for \(\tilde{p}\) and \(\tilde{B}\). The vector/matrix dimension increases by one due to the new bus \(i'\). The goal is to update the new angle \(\tilde{\theta}\) and line flow \(\tilde{f}\). Note that bus splits may cause islanding in an interconnected system [24]. For simplicity, this paper assumes no island is formed from the split of bus \(i\), as needed for typical analysis of line outages.

Notice that the \((N+1) \times (N+1)\) matrix \(\tilde{B}\) follows a block structure:

\[
\tilde{B} := \begin{bmatrix} B' & \ell \\ \ell^T & d \end{bmatrix} \tag{6}
\]

where \(B'\) is the \(N \times N\) sub-matrix, while vector \(\ell\) and scalar \(d\) capture the rest. Based on Fig. 3, the lines \((i, j)\) with \(j \in \mathcal{J}\) to allow for flexible topology re-configurations during fault events. Due to CB actions, the original bus may become electrically disconnected, commonly termed as “bus split.” This type of topological changes is increasingly popular due to CB misoperations [17], [20] or malicious cyber intrusions [21]–[23]. Fig. 1 illustrates one such event, with solid squares representing closed breakers and hollow ones for open breakers. If the top right CBs open, bus \(i\) is split into two disconnected buses, \(i\) and \(i'\), in the same substation. Accordingly, the connectivity for transmission lines has changed, as well as for generation and load within this substation. Although the two buses are physically co-located in the same substation, they become electrically disconnected which affects the bus-branch model as shown in Fig. 2. Bus split events change the grid topology in a very different way from contingencies like line outages or generation/load disconnections. The ensuing section will first present the linearized sensitivity analysis, a powerful tool needed for various grid contingencies, for the bus split events.
denote those lines that are reconnected to the new bus $i'$ after the split. Similarly, we use buses $j \in \mathcal{J}'$ to match those lines $\{(i, j)\}$ that remain connected to bus $i$ after the split. Hence, the submatrix $B'$ is formed by eliminating these reconnected lines from $B$, as given by [cf. (5)]

$$B' = B - \sum_{j \in \mathcal{J}} b_{ij} \left( (e_i - e_j)(e_i - e_j)^T - e_j e_j^T \right).$$

The reconnected lines $\{(i', j)\}$ also affect the rest of $\tilde{B}$, as

$$\ell = -\sum_{j \in \mathcal{J}} b_{ij} e_j, \quad \text{and} \quad d = \sum_{j \in \mathcal{J}} b_{ij}.$$

With these definitions, one can find the inverse of $\tilde{B}$ using the popular matrix inverse lemma [25, p. 650]. To this end, define the following Schur complement of the entry $d$ to matrix $\tilde{B}$ as

$$B_d := \tilde{B} - (d^{-1}) \ell \ell^T.$$  

(9)

Using $B_d$, we can obtain the block structure of the inverse as

$$\tilde{B}^{-1} = \begin{bmatrix} B_d^{-1} & -d^{-1} B_d^{-1} \ell \\ -(d^{-1}) \ell^T B_d^{-1} & (d^{-2}) \ell^T \ell B_d^{-1} + d^{-1} \end{bmatrix}.$$  

(10)

Note that one can rewrite $B_d = B - d(\mathbf{u} \mathbf{u}^T)$ with the vector

$$\mathbf{u} := e_i + (d^{-1}) \ell.$$  

(11)

For example, if $|\mathcal{J}| = 1$, it becomes the case of single line reconnection as discussed in [26] where vector $\mathbf{u} = e_i - e_j$ with $\mathcal{J} = \{j\}$. Applying the Sherman-Morrison formula [27] leads to the inverse

$$B_d^{-1} = (I + \Delta_u)B^{-1} = B^{-1} + \frac{d B^{-1} \mathbf{u} \mathbf{u}^T}{1 - d \mathbf{u}^T B^{-1} \mathbf{u}} B^{-1},$$  

(12)

with $\Delta_u$ capturing the fractional term above. Clearly, the inverse $B_d^{-1}$ can be quickly formed with $B^{-1}$ available, and so are the other blocks in (10).

The next step is to form the post-split power injection vector $\mathbf{p}$. Fig. 3 indicates that generation and load originally connected to bus $i$ could become attached to the new bus $i'$ due to the split. Let $\tilde{p}_i$ denote the total injected power to bus $i'$ after the split, and thus the injected power to bus $i$ becomes $(\tilde{p}_i - \tilde{p}_i)$. For example, Fig. 3 shows the case of $\tilde{p}_i = \tilde{p}_i$ and $(\tilde{p}_i - \tilde{p}_i) = \tilde{d}_i$. In general, $\tilde{p}_i$ can also be the attached load, or, the combination of generation and load. Therefore, we can express the injected power as

$$\tilde{p} = \begin{bmatrix} p \\ 0 \end{bmatrix} + \begin{bmatrix} -\tilde{p}_i e_i \\ \tilde{p}_i \end{bmatrix} = \begin{bmatrix} p - \tilde{p}_i e_i \\ \tilde{p}_i \end{bmatrix}.$$  

(13)

Using (10) and (13), one can obtain the post-split $\mathbf{p}$ as

$$\begin{bmatrix} \mathbf{p}' \\ \tilde{p}_i \end{bmatrix} = \tilde{B}^{-1} \mathbf{p} = \begin{bmatrix} d^{-1} \ell^T \mathbf{p}' + \tilde{p}_i \end{bmatrix}$$  

(14)

where the angle vector for the original $N$ buses is

$$\mathbf{\theta} := d^{-1} (\mathbf{p} - \tilde{p}_i e_i - d^{-1} \tilde{p}_i \ell)$$  

(15)

Although the solution $\mathbf{\theta}$ can be readily computed once obtaining (12), it is possible to better interpret it by developing an equivalent model for the post-split system. As illustrated in Fig. 4, the new bus $i'$ can be eliminated by connecting every pair of buses in $\mathcal{J}$. For example, the new line connecting any buses $j_1, j_2 \in \mathcal{J}$ would have the equivalent line parameter $b_{j_1j_2} := (d^{-1}) b_{j_1j_2}$ by recalling $d$ from (8). This follows from the well-known Kron’s reduction approach [28]. Similarly, after eliminating bus $i'$, its injected power $\tilde{p}_i$ is reallocated to each of its neighboring buses, with $d^{-1} b_{ij} \tilde{p}_i$ to bus $j \in \mathcal{J}$.

Interestingly, the equivalent topology in Fig. 4 exactly matches the matrix $B_d$ as defined in (9). Substituting (7) into (10), one can decompose $B_d = B_d' + B_d''$, where the matrix

$$B_d' := \sum_{j \in \mathcal{J}} b_{ij}(e_i - e_j)(e_i - e_j)^T$$  

(16)

corresponds to the Bbus matrix with the outage of lines in $\{(i, j)\} \in \mathcal{J}$ from the original system, while the remaining part

$$B_d'' = \sum_{j \in \mathcal{J}} b_{ij} e_j e_j^T - (d^{-1}) \ell \ell^T.$$  

(17)

reflects the new equivalent lines among all buses in $\mathcal{J}$. Hence, the angle vector $\mathbf{\theta}'$ in (15) perfectly matches the equivalent system for the original $N$ buses with updated topology/Injection, as its total injection now includes the new ones at each $j \in \mathcal{J}$ due to eliminating bus $i'$. Using the updated angles in $\mathbf{\theta}'$, one can further recover the angle at bus $i'$ by solving the total power flow on the eliminated lines in $\{(i', j), \forall j \in \mathcal{J}\}$ as

$$\tilde{p}_i = \sum_{j \in \mathcal{J}} b_{ij} (\tilde{\theta}_i - e_j^T \mathbf{\theta}')$$  

(18)

leading to the same solution of $\tilde{\theta}_i$ as in (14). This equivalent analysis clearly explains how the post-split system can be related to the original one.

Upon solving for the post-split angle $\mathbf{\theta}$, one can formalize the sensitivity analysis in terms of the changes of bus angles and line flows, as given in the following proposition.
Proposition 1. For the split of bus $i$ with a re-connected injection of $\tilde{p}_i$, the phase angle difference at the original $N$ buses can be written as

$$\theta' - \theta = \Delta_u \theta - (I + \Delta_u)B^{-1}(\tilde{p}_i e_i + d^{-1}\tilde{p}_i \ell)$$

(19)

where $d$, $\ell$, and $\Delta_u$ are given in (8) and (12). Accordingly, the line flow difference due to the bus split is [cf. (2)]

$$\tilde{f} - f = K\Delta_u \theta - K(I + \Delta_u)B^{-1}(\tilde{p}_i e_i + d^{-1}\tilde{p}_i \ell).$$

(20)

Strictly speaking, (20) holds for the lines other than the reconnected ones $\{(i', j)\}_{j \in J}$, as the flows of the latter lines are related to $\theta'$. In addition, since matrix $\Delta_u$ is a linear transformation of $B^{-1}$ [cf. (12)], the matrix product $K\Delta_u$ in (20) can be quickly computed using the ISF matrix $KB^{-1}$.

Given the detailed post-split topology and injection, the sensitivity analysis in Proposition 1 enables explicit expression of the phase angle and line flow change. In addition to generalizing line outage sensitivity analysis [29], this result could benefit security-constrained economic dispatch by including bus split contingencies [26] and other related EMS tasks. The sensitivity analysis results could also be useful for the monitoring of bus split events, which is the subject of the ensuing section.

IV. TRACTABLE IDENTIFICATION VIA MCCORMICK RELAXATION

Increasing deployment of high-resolution sensors such as PMUs allows for real-time identification of anomalous events such as topology changes. It is thus possible to utilize synchrophasor data to efficiently identify the underlying CB statuses during bus split events. Based on the sensitivity analysis in Sec. III, an intuitive solution could be the exhaustive search by enumerating all number of possible scenarios. However, the complexity order of this approach is not ideal for real-time identification and motivates us to consider more efficient alternatives.

As in the identification of line outages [9], [10], synchrophasor data provides the difference between pre- and post-event phase angle measurements which can be used to locate bus splits. Consider again the split of bus $i$ as in Sec. III. This candidate bus can be selected based on the bus locations where significant angle changes have been observed, as detailed in Sec. V. Given bus $i$ and its subset $J$ that include all reconnected buses/generation/loads, the post-split phase angle $\theta$ can be computed in (14). This suggests a brute-force solution by enumerating all possible scenarios of sets $J$ and selecting the one best matching the measured changes. Towards this end, for the target bus $i$ let us define the binary variables $\{z^j\}_{j \in N_i}$ to indicate the post-event status for each line $(i, j)$, such that

$$z^j = \begin{cases} 1, & \text{if bus } j \text{ re-connects to bus } i', \\ 0, & \text{if bus } j \text{ stays connected to bus } i \end{cases}$$

(21)

Similarly, let the binary variables $\{z^g, z^d\}$ denote the post-event connection status for the generation and load located at bus $i$. For simplicity, only one generator and one load are considered here, which can be extended to the case of multiple number of them. Given all these binary variables in vector $z \in \{0, 1\}^{(|N_i|+2)}$, one can use the sensitivity analysis to compute the post-split angle $\tilde{\theta}(z)$. Denoting the measured post-split angle as vector $e$, the identification problem becomes to minimize the norm of mismatch error between the two, namely $\|\tilde{\theta}(z) - e\|$. To address the potential approximation error of dc power flow model, a more popular error objective is to compare the phase angle change instead of the post-split phase angle itself; see e.g., [10]. Hence, the bus split identification problem becomes

$$\min_{z \in \{0, 1\}^{(|N_i|+2)}} \|\tilde{\theta}(z) - \delta\|$$

(22)

where $\tilde{\theta} := [\theta; \theta_i]$ is pre-split phase angle solution by the dc model and $\delta$ is the observed phase angle difference by the PMUs. Common choice of norms such as $L_1$, $L_2$ or $L_\infty$ can be used for quantifying the mismatch error in the objective (22). In this paper, we select the $L_1$ norm error which is known to enjoy a nice reformulation to linear objective function. Note that if only partial angles are monitored or the line flow data is also available, one can modify the objective function of (22) by extracting corresponding entries or adding a linear transformation of the angle difference vector. Such generalization will be investigated numerically in Sec. V-B. Using (22), the naive exhaustive search approach works by enumerating all $2^{|N_i|+2}$ scenarios of $z$, and picking the one with the minimum error in (22). Such exponential complexity is not suitable for real-time implementation, and will be addressed here by developing a tractable solution.

The main challenge in directly optimizing for $z$ in (22) lies in the post-split power flow model, as

$$\tilde{\theta}(z) = \left[\tilde{B}(z)\right]^{-1}\tilde{p}(z).$$

(23)

The post-split Bbus matrix depends on the line status variables $\{z^j\}$, given by

$$\tilde{B}(z) = B + \sum_{j \in N_i} z^j D^j$$

(24)

where $B = [B 0; 0^T]$ corresponds to the original Bbus matrix for the augmented system including the new bus $i'$, while the change due to each reconnected line is captured by

$$D^j := b_{ij}[(e_{i'} - e_j)(e_{i'} - e_j)^T - (e_i - e_j)(e_i - e_j)^T].$$

(25)

Similarly, the post-split injection vector depends on the generation/load status variables, as

$$\tilde{p}(z) = p + z^g \delta^g - z^d \delta^d$$

(26)

where $p = [p; 0]$ also augments the dimension of $p$ to include the new bus $i'$, while the other two vectors are given by

$$\delta^g := g_i(e_{i'} - e_i) \quad \text{and} \quad \delta^d := d_i(e_{i'} - e_i).$$

(27)

Substituting (23) into the objective of (22) introduces the term $[B(z)]^{-1}$, which is the inverse of a matrix function of binary variables. Furthermore, its multiplication with $\tilde{p}(z)$ poses additional nonlinearity (bilinearity) to the problem. Both issues lead to the lack of tractability in solving (22).
To tackle these issues, we propose to adopt the McCormick relaxation technique [19], which is a powerful tool for dealing with bilinear terms of binary variables. Specifically, by introducing an additional matrix $\mathbf{X} \in \mathbb{R}^{(N+1) \times (N+1)}$ to represent $[\mathcal{B}(z)]^{-1}$, we can reformulate (22) as:

\[
\min_{z, \mathbf{X}} \left\| \mathbf{Xp}(z) - \hat{\theta} - \delta \right\| \\
\text{s.t.} \quad \hat{\mathcal{B}}(z) \mathbf{X} = \mathbf{I}.
\]  

(28a) and (28b)

This way, the matrix inversion is no longer needed, which is replaced by the bilinear products between the unknowns. Note that the bilinear constraint (28b) for enforcing the relation between $\mathbf{X}$ and $\mathcal{B}(z)$ now becomes [cf. 24]:

\[
\mathbf{BX} + \sum_{j \in N_i} \mathbf{D}_i^j \mathbf{Y}_j = \mathbf{I}
\]  

(29)

with the product $\mathbf{Y}_j := z^j \mathbf{X}$ defined for each line $(i, j)$. Given the binary $z_j \in \{0, 1\}$ and the $(m, n)$-th entry $X_{mn} \in [X_{mn}^{\text{min}}, X_{mn}^{\text{max}}]$, each entry of $\mathbf{Y}_j$ can be written as:

\[
Y_{mnj}^j = z^j X_{mn}, \quad \forall \ m, n.
\]  

(30)

Accordingly, the following four linear inequalities hold:

\[
\begin{align*}
Y_{mnj}^j &\geq z^j X_{mn}^{\text{min}}, \\
Y_{mnj}^j &\leq X_{mn} + z^j X_{mn}^{\text{max}} - X_{mn}^{\text{max}}, \\
Y_{mnj}^j &\geq z^j X_{mn}^{\text{max}}, \\
Y_{mnj}^j &\leq X_{mn} + z^j X_{mn}^{\text{min}} - X_{mn}^{\text{min}},
\end{align*}
\]  

(31a) - (31d)

Each inequality in (31) can be verified by substituting $z^j = 1$. Interestingly, the set of four linear inequalities in (31) jointly also guarantees the validity of (30) for any binary $z^j$. To demonstrate this, first consider the case of $z^j = 0$. The two inequalities (31a) and (31c) jointly lead to $Y_{mnj}^j = 0$, and thus (30) holds. Otherwise if $z^j = 1$, the other two inequalities (31b) and (31d) would enforce that $Y_{mnj}^j = X_{mn}$. Hence, for binary $z^j$ the set of inequalities in (31) is equivalent to the bilinear relation in (30). Reformulating (30) with the linear inequalities in (31) is known as the McCormick relaxation technique and has been usually used in power system topology designing problems [30], [31]. Thus, the original constraint (28b) can be replaced with (29) and (31a) - (31d).

For the objective function in (28a), we can introduce matrices $\mathbf{W}^p := z^p \mathbf{X}$ with $p \in \{g, d\}$ indicating either the generation or load, and similarly convert the matrix products into equivalent linear inequalities. We replace the resultant bilinear terms in (28a) for $p \in \{g, d\}$ using the follows:

\[
\begin{align*}
W_{mn}^p &\geq z^p X_{mn}^{\text{min}}, \\
W_{mn}^p &\leq X_{mn} + z^p X_{mn}^{\text{max}} - X_{mn}^{\text{max}}, \\
W_{mn}^p &\leq X_{mn} + z^p X_{mn}^{\text{min}} - X_{mn}^{\text{min}}, \\
W_{mn}^p &\geq z^p X_{mn}^{\text{max}},
\end{align*}
\]  

(32a) - (32d)

Proposition 2. Given the bounds $[X_{mn}^{\text{min}}, X_{mn}^{\text{max}}]$ for each entry $X_{mn}$, we can reformulate (28) as the following mixed-integer programming problem after substituting $\mathbf{p}(z)$ in (26):

\[
\min_{z, \mathbf{X}, \{Y_j\}, \{W^p\}} \left\| \mathbf{Xp}(z) - \hat{\theta} - \delta \right\| \\
\text{s.t.} \quad (29), (31) \text{ and } (32), \forall \ (m, n).
\]  

(33a) and (33b)

As mentioned earlier, the $L_1$ error norm in (33a) can lead to an equivalent linear objective cost; see e.g., [25, Ch. 4]. Meanwhile, all constraints in (33b) are linear, thanks to the McCormick relaxation technique. Hence, the original problem (22) is converted to a mixed-integer linear program (MILP), for which there exist several off-the-shelf efficient solvers such as CPLEX and Gurobi. This equivalent MILP reformulation (33) constitutes as a tractable solution for identifying the status of multiple connections within a specific substation. The setting of tight bounds $[X_{mn}^{\text{min}}, X_{mn}^{\text{max}}]$ for each system will be discussed soon.

It is possible to further reduce the number of decision variables in (33) by taking advantage of the sparse structure of matrix $\mathbf{D}_i^j$. To this end, let us define the product $\Delta^j := \mathbf{D}_i^j \mathbf{Y}_j = z^j \mathbf{D}_i^j \mathbf{X}$, which was used in (29). With the definition of $\mathbf{D}_i^j$ in (25), the product $\Delta^j$ would involve only three rows of $\mathbf{X}$, namely the $i$-th, $j$-th and $i'$-th rows. Accordingly, it suffices to form the corresponding submatrix $\mathbf{Y}_j^j \in \mathbb{R}^{3 \times (N+1)}$ as decision variables with $\forall n = 1, \ldots, N + 1$:

\[
\begin{align*}
Y_{1n_j}^j &= z^j X_{1n}, \\
Y_{2n_j}^j &= z^j X_{2n}, \\
Y_{3n_j}^j &= z^j X_{3n}.
\end{align*}
\]  

(34a) - (34c)

As a result, matrix $\Delta^j$ is a sparse matrix with all nonzero elements listed here:

\[
\begin{align*}
\Delta_{i,n}^j &= b_{ij} (Y_{2n_j}^j - Y_{1n_j}^j), \\
\Delta_{j,n}^j &= b_{ij} (Y_{1n_j}^j - Y_{3n_j}^j), \\
\Delta_{i',n}^j &= b_{ij} (Y_{3n_j}^j - Y_{2n_j}^j).
\end{align*}
\]  

(35a) - (35c)

The same technique can be applied to reduce the number of variables in the product of $\mathbf{W}^p$ and $\delta_p$ in (33a). For the example of generation, due to the definition of $\delta_g$ in (27), the operation only depends on the $i$-th and $i'$-th columns of $\mathbf{X}$. Hence, one only needs to introduce a submatrix $\mathbf{W}^g \in \mathbb{R}^{(N+1) \times 2}$ as decision variables and form the product similar to (35). Thanks to the sparsity of matrix/vector in (33), we can reduce the number of decision variables from $\mathcal{O}(N^2)$ to $\mathcal{O}(N)$, and same for the number of linear inequality constraints in (33b). This simplification step will allow for achieving the identification solution more efficiently by the MILP solvers.

A. Bounds on Continuous Variables $X_{mn}$

With the MILP reformulation given in (33), the problem remains to derive the upper/lower bounds of continuous variables in $\mathbf{X}$. In general, the tighter the bounds are, the faster the MILP can be solved; see e.g., [32]. Hence, we will aim to find reasonably good bounds for entries of $\mathbf{X}$.

To this end, first notice that each entry $X_{mn}$ is non-negative; i.e., the lower bound $X_{mn}^{\text{min}} = 0$ holds. This is because the Bbus matrix $\mathcal{B}(z)$ is an M-matrix [33] with its off-diagonal entries being non-positive and real eigenvalues being non-negative, and thus the entries of its inverse are non-negative. To obtain the upper bound $X_{mn}^{\text{max}}$, it is possible to directly maximize the
value of $X_{mn}$ under the same constraints in the optimization problem (33), as given by

$$
\begin{align}
\max_{z, X, \{Y\}, \{W_p\}, X_{\text{max}}} & \quad X_{\text{max}} \\
\text{s.t.} & \quad X_{mn} \leq X_{\text{max}}, \forall (m,n) (36a) \\
& \quad (29), (31) \text{ and } (32), \forall (m,n). (36c)
\end{align}
$$

Compared with (33), this new problem (36) differs only in the objective function and it is an MILP as well. Clearly, its optimum solution of $X_{\text{max}}$ defines an upper bound for every entry of $X$. Notice that to solve (36), we need to start with a rough estimate of the upper bound of $X_{mn}$ to tackle the inequality constraints in (31) and (32). This is possible by using a very large value of the upper bound estimate which is refined by (36). Since this upper bound holds for any choice of $z$, this refinement step can be performed off-line and does not affect the real-time identification time.

V. Numerical Results

In this section, we use IEEE 14-bus and 300-bus test cases to demonstrate the identification of bus split events. The 14-bus case allows to better illustrate the system-wide effects of bus split, while the 300-bus case is used to provide quantifiable identification error performance. The optimization problems have been implemented on a regular laptop with Intel® CPU @ 2.60 GHz and 12 GB of RAM in the MATLAB® R2018a simulator. The reformulated MILP-based identification problems are computed by Gurobi.

A. IEEE 14-Bus System Tests

The IEEE 14-bus system consists of 20 lines and 5 conventional generators. The proposed identification algorithm has been shown effective for all possible bus split scenarios (excluding the islanding ones) for this system. All types of bus injections, generation only, load only, or the combination of both, have been considered as well. Instead of listing all the results, we pick the split of bus $i = 13$ as an illustrative example. The first neighbor buses of bus $i = 13$ is given by $\mathcal{N}_i = \{6, 12, 14\}$. The line $(13, 14)$ and load $d_i$ are reconnected to the new bus $i' = 15$ after the bus split.

Upon solving the optimization problem (33) for every bus location, we plot the resultant minimum mismatch error in Fig. 5. As any but split at bus 8 leads to system islanding, it is excluded from the comparison. The mismatch error at bus 13 is the smallest among all possible buses, and thus it is identified as the correct location. Moreover, the optimal solution for this bus has $z = [0; 0; 1]$, which correctly indicates that line $(13, 14)$ is reconnected to the new bus while $\delta^d = 1$ identifies the reconnected load. Using the optimal solution for each bus location, we can find the corresponding system-wide phase angle difference based on dc model, namely $[\hat{B}(z)]^{-1}\hat{p}(z) - \hat{\theta}$. Fig. 6 plots the resultant angle difference solutions for selected bus locations as compared to the actual values $\delta$ from the ac power flow model as observed by PMUs. This comparison again confirms that the split at bus 13 has the closest match with the actual system responses.

It is worth mentioning that this test also points out how to efficiently select candidate bus locations for larger-sized systems. As shown by Fig. 6, the actual phase angles at bus 13 and its neighbor buses $\mathcal{N}_{13} = \{6, 12, 14\}$ change more evidently than other buses. As the inverse of $B_{bus}$ matrix is typically diagonally dominant, the effects of topology changes tend to reduce from the change location to buses further away; see e.g., [34]. This localized impact property enables to adopt a simple yet effective approach to prioritize candidate bus locations based on their angle differences, which will turn out to be very useful for large systems as detailed soon.

B. IEEE 300-Bus System Tests

We have performed more comprehensive validations using the IEEE 300-bus system that provide quantifiable identification performance. We first investigate the bus split at the substation of bus $i = 120$, and then provide more quantitative results for different bus locations. Finally, the impact of partial observation is also studied.

Test Case 1: We pick the split of bus $i = 120$ to investigate. It has 3 neighboring buses, and is connected with both generation and load. Compared to the 14-bus system, there are a larger number of buses in this system, making it difficult to enumerate all the buses as candidate buses. Using the empirical insights earlier on, we use a heuristic selection scheme that ranks the buses based on the observed value of phase angle difference, namely $|\delta_i|$ provided by the synchrophasor data. If not all buses are equipped with PMUs, as later in Test Case 3, the neighboring buses of the highly-ranked buses would also be included as candidate locations. For the split of bus $i = 120$, etc.
the top six buses in the ranking are buses 120, 153, 151, 152, 155, and 154, most of which are co-located in the same area to bus 120. These six buses are selected to run the problem (33). Accordingly, the bus with the smallest achievable error objective value is deemed as the location of bus split, along with its corresponding optimal solution.

This identification process has been run for every possible topology scenario (every choice of binary \( z \)) for the split of bus 120. The identification accuracy is evaluated based on the percentage of correctly recovered entries of vector \( z \), which is given in Fig. 7 for every topology scenario. For a majority of scenarios, the identification accuracy is perfect. This implies that the angle differences for each of these scenarios sufficiently differentiate from those of all other scenarios, and our proposed solution can effectively find the correct scenario. Nonetheless, this is not the case for scenario 3, where the connections for two out of the three lines have been erroneously identified. A closer look at this scenario reveals that the dc approximation error has led to that the mis-identified \( z \) solution has the smallest mismatch error with the nonlinear ac model. Overall, the proposed method can correctly find the optimal solution to (22), while the accuracy of latter may still depend on the approximation error of the dc model.

Test Case 2. We have further tested our proposed identification algorithm for the split events at 16 selected buses, giving rise to 116 topology scenarios in total. To quantify the identification accuracy for each bus, we average the percentage of correct recovery over all the scenarios under the split of that bus. The resultant accuracy results for the original 300-bus system are given in Fig. 8, along with those for a modified 300-bus system. Test Case 1 pointed out the impact of the dc power flow approximation, which is highly related to the line resistance-reactance ratio [18]. Some of the transmission lines in the original 300-bus system have a quite high ratio value, which is uncommon for high-voltage grids. Hence, we reduce the resistance values therein by half to attain a modified system. Overall, our proposed algorithm has achieved effective identification of bus split events in all the 16 selected buses. On average, the original system gives an accuracy of 97.6%, with a minimum accuracy of 83.3% at bus 45. By decreasing the resistance values, the modified system enjoys an increased average accuracy of 99.2%. Hence, our identification algorithm can attain accurate results for practical systems.

To evaluate the computational complexity of the proposed algorithm, the mean and median values for the run-time of all topology scenarios per bus for the original 300-bus system are given in Fig. 9. For this 300-bus system, it takes around 100 seconds to execute the proposed identification algorithm using a standard computer. This run-time scales nicely with the size of system. Improved parameter settings (e.g., the upper bound \( X_{\text{max}} \)) and high-performance computing resources can further facilitate the implementation of the proposed algorithm for real-time monitoring of bus split events.

Test Case 3. This test investigates the effects of partial observation using the modified system of reduced resistance. As mentioned in Sec. IV, the problem (22) can be easily modified to allow for partial angle or additional line flow measurements. Using the benchmark of full angle measurements, we compare three partial observation settings: 70% angle measurements, 85% angle measurements and a combination of 70% angle measurements and 50% line flow measurements. For the partial angle measurements, it suffices to use a selection matrix to extract the mismatch error vector in (22) for the metered locations. As for the line flow measurements, they can be incorporated into the error mismatch objective similar to angle measurements by recognizing the linear relation in (2). All the measurement locations have been randomly selected and the mismatch error vector is normalized to match the scaling difference between angle and flow changes. Specifically, the ratio between the average of the absolute angle changes and that of the absolute line flow change is used to scale these two types of measurements.

Fig. 10 plots the identification accuracy for the partial observation settings along with the benchmark. All the three settings have achieved satisfactory accuracy (on average 98.1%, 98.9% and 98.7% for each of the three settings), with a minimum accuracy of 89.6% at bus 8 under the 70% angle setting. Generally speaking, the accuracy consistently increases when there are more measurements of angles or line flows. Note that for the split events at bus 222, the inclusion of line flow data is very useful for achieving perfect identification results. At some other buses (bus 8 or 120), the additional angle information is more helpful for identifying certain split events. Therefore, the proposed algorithm can adapt to various measurement availability conditions and provide an accurate identification performance for large systems.
in the face of bus split events.

algorithm for enhancing power system situational awareness corroborated the performance of the proposed identification solvable for real-time identification. Numerical studies have equivalent the McCormick relaxation technique has been leveraged to To tackle the bilinear relations in the resultant problem, MILP reformulation that is efficiently convex optimization.

changes for given bus split event. In addition, synchrophasor connectivity changes. Based on the dc power flow analysis for the compact bus-branch model, the linear sensitivity analysis framework for power system bus split events due to substation identification in generalized state estimation,

VI. CONCLUSIONS

This paper developed an efficient modeling and monitoring framework for power system bus split events due to substation connectivity changes. Based on the dc power flow analysis for the compact bus-branch model, the linear sensitivity analysis was performed that can quickly compute the system-wide changes for given bus split event. In addition, synchronphasor data enabled monitoring of bus split events was cast as an optimization problem, with binary variables indicating the connectivity of lines, generation, loads within a substation. To tackle the bilinear relations in the resultant problem, the McCormick relaxation technique has been leveraged to attain an equivalent MILP reformulation that is efficiently solvable for real-time identification. Numerical studies have corroborated the performance of the proposed identification algorithm for enhancing power system situational awareness in the face of bus split events.

REFERENCES

[1] A. J. Wood, B. F. Wollenberg, and G. B. Sheblé, Power generation, operation, and control. John Wiley & Sons, 2013.
[2] S. Santosu, “Transmission line protection for enhanced security: Including circuit models in relays,” Tech. Rep., Dec. 2014.
[3] “Analysis of the cyber attack on the Ukrainian power grid,” Tech. Rep., Mar. 2016.
[4] A. Abur, H. Kim, and M. Celik, “Identifying the unknown circuit breaker statuses in power networks,” IEEE Trans. on Power Systems, vol. 10, no. 4, pp. 2029–2037, 1995.
[5] J. G. Kassakian, R. Schmalensee, G. Desrosiers et al., “The future of the electric grid,” Tech. Rep., 2011.
[6] “Node-breaker white paper,” Tech. Rep., Jan. 2014.
[7] “Proposal for development and use of node breaker topology representations for off-line study models,” Tech. Rep., Jul. 2019.
[8] A. Abur and A. G. Exposito, Power system state estimation: theory and implementation. CRC press, 2004.
[9] J. E. Tate and T. J. Overbye, “Line outage detection using phasor angle measurements,” IEEE Trans. on Power Systems, vol. 23, no. 4, pp. 1644–1652, 2008.
[10] H. Zhu and G. B. Giannakis, “Sparse overcomplete representations for efficient identification of power line outages,” IEEE Trans. Power Systems, vol. 27, no. 4, pp. 2215–2224, 2012.
[11] Y. C. Chen, T. Banerjee, A. D. Dominguez-Garcia, and V. V. Veeravalli, “Quickest line outage detection and identification,” IEEE Trans. on Power Systems, vol. 31, no. 1, pp. 749–758, 2015.
[12] D. Deka, S. Backhaus, and M. Chertkov, “Estimating distribution grid topologies: A graphical learning based approach,” in 2016 Power Systems Computation Conference (PSCC). IEEE, 2016, pp. 1–7.
[13] G. Cavraro and V. V. Vasilakos, “Graph algorithms for topology identification using power grid probing,” IEEE control systems letters, vol. 2, no. 4, pp. 689–694, 2018.
[14] O. Ardakanian, V. W. Wong, R. Dobbe, S. H. Low, A. von Meier, C. J. Tomlin, and Y. Yuan, “On identification of distribution grids,” IEEE Trans. on Control of Network Systems, vol. 6, no. 3, pp. 950–960, 2019.
[15] M. Heidarifar and H. Ghasemi, “A network topology optimization model based on substation and node-breaker modeling,” IEEE Trans. Power Systems, vol. 31, no. 1, pp. 247–255, 2015.
[16] B. Park, J. Holzer, and C. L. DeMarco, “A sparse tableau formulation for node-breaker representations in security-constrained optimal power flow,” IEEE Trans. Power Systems, vol. 34, no. 1, pp. 637–647, 2019.
[17] V. Kekatos and G. B. Giannakis, “Joint power system state estimation and breaker status identification,” in Proc. North American Power Symp., 2012.
[18] B. Stott, J. Jardim, and O. Alsac, “DC power flow revisited,” IEEE Trans. on Power Systems, vol. 24, no. 3, pp. 1290–1300, 2009.
[19] G. P. McCormick, “Computability of global solutions to factorable nonconvex programs: Part I-Convex underestimating problems,” Mathematical programming, vol. 10, no. 1, pp. 147–175, 1976.
[20] G. Korres, P. Katikas, and G. Chatzarakis, “Substation topology identification in generalized state estimation,” International Journal of Electrical Power & Energy Systems, vol. 28, no. 3, pp. 195–206, 2006.
[21] D. Deka, R. Baldick, and S. Vishwanath, “One breaker is enough: Hidden topology attacks on power grids,” in Proc. IEEE PES General Meeting, 2015.
[22] C.-W. Ten, K. Yamashita, Z. Yang, A. V. Vasilakos, and A. Ginter, “Impact assessment of hypothesized cyberattacks on interconnected bulk power systems,” IEEE Trans. Smart Grid, vol. 9, no. 5, 2017.
[23] Y. Zhou, J. Cisneros-Saldana, and L. Xie, “False analog data injection attack towards topology errors: Formulation and feasibility analysis,” in Proc. IEEE PES General Meeting, 2016.
[24] K. Sun, D.-Z. Zheng, and Q. Lu, “Splitting strategies for islanding operation of large-scale power systems using OBDD-based methods,” IEEE Trans. Power Systems, vol. 18, no. 2, pp. 912–923, 2003.
[25] S. Boyd and L. Vandenberghe, Convex optimization. Cambridge, U.K.: Cambridge University Press, 2004.
[26] Y. Zhou and H. Zhu, “Bus split sensitivity analysis for enhanced security in power system operations,” in Proc. North American Power Symp., 2019.
[27] J. Sherman and W. J. Morrison, “Adjustment of an inverse matrix corresponding to a change in one element of a given matrix,” The Annals of Mathematical Statistics, vol. 21, no. 1, pp. 124–127, 1950.
[28] F. Dorfler and F. Bullo, “Kron reduction of graphs with applications to electrical networks,” IEEE Trans. on Circuits and Systems I: Regular Papers, vol. 60, no. 1, pp. 150–163, 2012.
[29] T. Guler, G. Gross, and M. Liu, “Generalized line outage distribution factors,” IEEE Trans. on Power Systems, vol. 22, no. 2, pp. 879–881, 2007.
[30] S. Bhela, D. Deka, H. Nagarajan, and V. V. Vasilakos, “Designing power grid topologies for minimizing network disturbances: An exact MILP formulation,” in Proc. American Control Conference (ACC), 2019.
[31] B. Park and C. L. Demarco, “Optimal network topology for node-breaker representations with ac power flow constraints,” IEEE Access, vol. 8, pp. 64347–64355, 2020.
[32] J. Ostrowski, M. F. Anjos, and A. Vannelli, “Tight mixed integer linear programming formulations for the unit commitment problem,” IEEE Trans. on Power Systems, vol. 27, no. 1, pp. 39–46, 2011.
[33] R. J. Plemmons, “M-matrix characterizations. I- nonsingular m-matrices,” Linear Algebra and its Applications, vol. 18, no. 2, pp. 175–188, 1977.
[34] H. Zhu and T. J. Overbye, “Blocking device placement for mitigating the effects of geomagnetically induced currents,” IEEE Trans. on Power Systems, vol. 30, no. 4, pp. 2081–2089, 2014.