Power Grid State Estimation under General Cyber-Physical Attacks

Yudi Huang, Student Member, IEEE, Ting He, Senior Member, IEEE, Nilanjan Ray Chaudhuri, Senior Member, IEEE, and Thomas La Porta Fellow, IEEE

Abstract—Effective defense against cyber-physical attacks in power grid requires the capability of accurate damage assessment within the attacked area. While some solutions have been proposed to recover the phase angles and the link status (i.e., breaker status) within the attacked area, existing solutions made the limiting assumption that the grid stays connected after the attack. To fill this gap, we study the problem of recovering the phase angles and the link status under a general cyber-physical attack that may partition the grid into islands. To this end, we (i) show that the existing solutions and recovery conditions still hold if the post-attack power injections in the attacked area are known, and (ii) propose a linear programming-based algorithm that can perfectly recover the link status under certain conditions even if the post-attack power injections are unknown. Our numerical evaluations based on the Polish power grid demonstrate that the proposed algorithm is highly accurate in localizing failed links once the phase angles are known.

Index Terms—Power grid state estimation, cyber-physical attack, failure localization.

I. INTRODUCTION

Modern power grids are interdependent cyber-physical systems consisting of a power transmission system (power lines, substations, etc.) and an associated control system (Supervisory Control and Data Acquisition - SCADA and Wide-Area Monitoring Protection and Control - WAMPAC) that monitors and controls the status of the power grid. This interdependency raises a legitimate concern: what happens if an attacker attacks both the physical grid and its control system simultaneously? The resulting attack, known as a joint cyber-physical attack, can cause large-scale blackouts, as the cyber attack can blindfold the control system and thus make the physical attack on the power grid more damaging. For example, one such attack on Ukraine’s power grid left 225,000 people without power for days [2].

The potential severity of cyber-physical attacks has attracted efforts in counting these attacks [3], [4]. One of the challenges in dealing with such attacks is that as the cyber attack blocks measurements (e.g., phase angles, breaker status, and so on) from the attacked area, the control center is unable to accurately identify the damage caused by the physical attack (e.g., which lines are disconnected) and hence unable to make accurate mitigation decisions. To address this challenge, solutions have been proposed to estimate the state of the power grid inside the attacked area using power flow models. Specifically, [3] developed methods to estimate the grid state under cyber-physical attacks using the direct-current (DC) power flow model, and [4] developed similar methods using the alternating-current (AC) power flow model. Both works made the limiting assumption that either (i) the grid remains connected after the attack, or (ii) the control center is aware of the supply/demand in each island formed after the attack – both leading to known post-attack active power injection at each bus.

In practice, however, disconnecting lines within the attacked area may cause partitioning of the grid and change the active power injections, and such changes within the attacked area will not be directly observable to the control center due to the cyber attack. Our goal is thus to estimate the power grid state, especially the breaker status of lines, under cyber-physical attacks without the above assumption.

A. Related Work

Power grid state estimation, as a key functionality for supervisory control, has been extensively studied in the literature [5]. Secure state estimation under attack is of particular interest [6]. Specifically, the attackers can distort sensor data with noise [7] or inject false data [8] so that the control center cannot correctly estimate the phase angles [9] or the topology [10] of the power grid. Recently, joint cyber-physical attacks are gaining attention, as the physical effect of such attacks are harder to detect due to the cyber attack [3], [11], [12].

In particular, several approaches have been proposed for detecting failed links. In [13], [14], the problem is formulated as a mixed-integer program, which becomes computationally inefficient when multiple links fail. The problem is formulated as a sparse recovery problem over an overcomplete representation in [15], [16], where the combinatorial sparse recovery problem was relaxed to a linear programming (LP) problem. Based on this approach, the work in [3] further establishes graph-theoretic conditions for accurately recovering the failed links. All the algorithms in [3], [15], [16] aim to find the sparsest solution among the feasible solutions under the assumption that the power grid remains connected after failure.

All the state estimation solutions require the modeling of the relationship between the observable parameters and the unknown variables of interest. To this end, two types of models have been considered: DC power flow model and AC power flow model. The AC power flow model [17] is based on the AC power flow equations, which can represent the voltage
magnitude and phase angle at each bus in the system. The DC power flow model [13] is an approximation of the AC power flow model by neglecting the resistive losses and assuming a uniform voltage profile. In the literature of state estimation and particularly failure localization, most existing solutions are based on the DC power flow model [3], [13]–[16], with few exceptions [4], [19]. We adopt the DC power flow model in this work due to its simplicity and robustness, and leave extensions to the AC power flow model to future work.

B. Summary of Contributions

We aim at estimating the power grid state within an attacked area, focusing on the phase angles and the link status (i.e., breaker status of lines), with the following contributions:

1) We show that an existing rank-based condition for recovering the phase angles, previously established when the grid remains connected after the attack, still holds without this limiting assumption.

2) We show that existing graph-theoretical conditions for localizing the failed links, previously established under the above assumption of a connected grid, still hold without this assumption if the post-attack power injections are known.

3) When the post-attack power injections are unknown but the phase angles are known, we develop an LP-based algorithm that is guaranteed to correctly recover the status of failed/operational links under certain conditions.

4) Our evaluations on a large grid topology show that the proposed algorithm is highly accurate in localizing the failed links with few false alarms, while the rank-based condition for recovering the phase angles can be hard to satisfy, signaling the importance of protecting PMU measurements.

Roadmap. Section II formulates our overall problem, which is divided into three subproblems addressed in Sections III–V. Then Section VII evaluates our solutions on a real grid topology, and Section VII concludes the paper.

II. PROBLEM FORMULATION

A. Power Grid Model

We model the power grid as a connected undirected graph $G = (V, E)$, where $V$ is the set of nodes (buses) and $E$ the set of links (transmission lines). Each link $e = (s, t)$ is associated with a reactance $r_{st}$ ($r_{ts} = r_{st}$) and a status $\in \{\text{operational, failed}\}$ (assumed to be operational before attack). Each node $v$ is associated with a phase angle $\theta_v$ and an active power injection $p_v$. The phase angles $\theta := (\theta_v)_{v \in V}$ and the active powers $p := (p_v)_{v \in V}$ are related by

$$B\theta = p,$$

where $B := (b_{uv})_{u,v \in V} \in \mathbb{R}^{\mid V \mid \times \mid V \mid}$ is the admittance matrix, defined as:

$$b_{uv} = \begin{cases} 0 & \text{if } u \neq v, (u, v) \notin E, \\ -1/r_{uv} & \text{if } u \neq v, (u, v) \in E, \\ -\sum_{w \in V \setminus \{u\}} b_{uw} & \text{if } u = v. \end{cases}$$

Figure 1. A cyber-physical attack that blocks information from the attacked area $H$ while disconnecting certain lines within $H$.

| Notation | Description |
|----------|-------------|
| $G = (V, E)$ | power grid |
| $H, \bar{H}$ | attacked/unattacked area |
| $F$ | set of failed links |
| $B$ | admittance matrix |
| $\theta$ | vector of phase angles |
| $p$ | vector of active power injections |
| $\Gamma$ | $\text{diag}\{\frac{1}{r_e}\}_{e \in E}$ ($r_e$: reactance of link $e$) |
| $\Delta$ | vector of changes in active power injections |

Given an arbitrary orientation of the links, the topology of $G$ can also be represented by the incidence matrix $D \in \{-1, 0, 1\}^{\mid V \mid \times \mid E \mid}$, whose $(i, j)$-th entry is defined as

$$d_{ij} = \begin{cases} 1 & \text{if link } e_j \text{ comes out of node } v_i, \\ -1 & \text{if link } e_j \text{ goes into node } v_i, \\ 0 & \text{otherwise}. \end{cases}$$

We assume that each node is deployed with a phasor measurement unit (PMU) measuring the phase angle and remote terminal units (RTUs) measuring the active power injection at this node, and the status and the power flows of its incident links. These reports are sent to the control center via a SCADA or WAMPAC system. The PMU data is assumed to be communicated over a relatively secure WAMPAC network, and the RTU measurements over a more vulnerable SCADA network.

B. Attack Model

As illustrated in Fig. 1 an adversary attacks an area $H$ of the power grid by: (i) blocking reports from the nodes within $H$ (cyber attack on both SCADA and WAMPAC), and (ii) disconnecting a set $F$ ($\mid F \mid > 0$) of links within $H$ (physical attack). Formally, $H = (V_H, E_H)$ is a subgraph induced by a set of nodes $V_H$, where $E_H$ is the set of links for which both endpoints are in $V_H$.

C. State Estimation Problem

Notation. The main notations are summarized in Table I. Moreover, given a subgraph $X$ of $G$, $X$ and $E_X$ denote the subsets of nodes/links in $X$, and $x_X$ denotes the subvector of a vector $x$ containing elements corresponding to $X$. Similarly, given two subgraphs $X$ and $Y$ of $G$, $A_{X|Y}$ denotes the submatrix of a matrix $A$ containing rows corresponding to $X$ and columns corresponding to $Y$. We use $D_H \in$
\{−1,0,1\}^{|V_H| \times |E_H|} to denote the incidence matrix of the attacked area \(H\). For each quantity \(x\), we use \(x'\) to denote its value after the attack.

**Goal.** Our goal is to recover the post-attack phase angles \(\theta'_H\) and localize the failed links \(F\) within the attacked area, based on the state variables before the attack and the measurements from the unattacked area \(\bar{H}\) after the attack.

In contrast to the previous works, we consider cases where the attack may partition the grid into multiple islands, which can cause changes in active power injections to maintain the supply/demand balance in each island. Let \(\Delta = (\Delta_v)_{v \in V} := p - p'\) denote the change in active power injections, where \(\Delta_v > 0\) if \(v\) is a generator bus and \(\Delta_v \leq 0\) if \(v\) is a load bus.

III. RECOVERY OF PHASE ANGLES

Under the assumption that \(G\) remains connected after the attack and thus \(\Delta = 0\), [3] showed that the post-attack phase angles in the attacked area \(\theta'_H\) can be recovered if the submatrix \(B_{\bar{H}\mid H}\) of the admittance matrix has a full column rank. Below, we will show that the same condition actually holds without this limiting assumption.

Specifically, we have the following lemma that extends [3] Lemma 1 to the general case of arbitrary \(\Delta\). Here “supp” returns the indices of the non-zero entries in the input vector.

**Lemma III.1.** \(\text{supp}(B(\theta - \theta') - \Delta) \subseteq V_H\).

**Proof.** For a link \((s, t)\), define a column vector \(x_{st} \in \{-1, 0, 1\}^{|V|}\), which has 1 in \(s\)-th element, -1 in \(t\)-th element, and 0 elsewhere. The failure of links \(F\) changes the admittance matrix \(B'\)

\[
B' = B + \sum_{(s,t) \in F} b_{st} x_{st} x_{st}^T,
\]

where \(b_{st}\) is the \((s,t)\)-th element in \(B\). Before the attack, we have \(B\theta = p\). After the attack, we have \(B'\theta' = p' = p - \Delta\). Therefore, the following holds:

\[
B\theta - B'\theta' = \Delta \\
\Rightarrow B(\theta - \theta') - \Delta = \sum_{(s,t) \in F} b_{st} x_{st} x_{st}^T \theta' \\
\Rightarrow \text{supp}(B(\theta - \theta') - \Delta) \subseteq \bigcup_{(s,t) \in F} \{s, t\} \subseteq V_H,
\]

where \([\Theta]\) is obtained by plugging in \((4)\) into \((5)\).

Using Lemma III.1, we will prove that the recovery condition in [3] Theorem 1 remains sufficient even if its assumption of a connected post-attack grid (hence \(\Delta = 0\)) may not hold.

**Theorem III.1.** The phase angles \(\theta'_H\) within the attacked area can be recovered correctly if \(B_{\bar{H}\mid H}\) has a full column rank.

**Proof.** By Lemma III.1 we see that \(B_{\bar{H}\mid H}(\theta - \theta') - \Delta_{\bar{H}} = 0\). Writing this equation in more detail shows that

\[
B_{\bar{H}\mid H}(\theta_H - \theta'_H) + B_{\bar{H}\mid H}(\theta_H - \theta_H) - \Delta_{\bar{H}} = 0
\]

\[
\Rightarrow B_{\bar{H}\mid H}\theta'_H = B_{\bar{H}\mid H}\theta_H + B_{\bar{H}\mid H}(\theta_H - \theta'_H) - \Delta_{\bar{H}}.
\]

Since both \(B_{\bar{H}\mid H}\) and the righthand side of \([9]\) are known to the control center, we can uniquely recover \(\theta'_H\) if \(B_{\bar{H}\mid H}\) has a full column rank.

In the special case of \(\Delta = 0\), [3] gave an explicit graph-theoretical condition that was derived from the rank-based condition. As the rank-based condition remains valid in the general case by Theorem III.1, the graph-theoretical condition also holds in general.

**Corollary III.1 ([3]).** If there is a matching (i.e., a set of links without common endpoints) among the links with one endpoint in \(V_H\) and the other endpoint in \(V_{\bar{H}}\) that covers \(V_H\), then \(\theta'_H\) can be recovered almost surely.

IV. LOCALIZING FAILED LINKS WITH KNOWN ACTIVE POWERS

Now assume that the post-attack phase angles \(\theta'\) have been recovered. This can be achieved by solving \([9]\) when \(B_{\bar{H}\mid H}\) has a full column rank, or directly reported from PMUs – assuming that the cyber attack only affects SCADA but not the WAMPAC network carrying PMU measurements. We will show that as long as the change in active powers \(\Delta\) is known, the failed links can be uniquely localized under the same conditions as specified in [3].

First, we note that under practical assumptions, the conditions presented in Section III for recovering the phase angles greatly simplify the recovery of the active powers. To this end, we assume that the adjustment of active power injections at generator/load buses follows the proportional load shedding/generation reduction policy, where (i) either the load or the generation (but not both) will be reduced upon the formation of an island, and (ii) if nodes \(u\) and \(v\) are in the same island and of the same type (both load or generator), then \(p'_u/p_u = p'_v/p_v\). This policy models the common practice in adjusting load/generation in the case of islanding [22], [23].

We observe the following cases in which the active powers can be recovered via this policy.

**Lemma IV.1.** Let \(N(v; \bar{H})\) denote the set of all the nodes in \(\bar{H}\) that are connected to node \(v\) via links in \(E \setminus E_H\). Then under the proportional load shedding policy, \(\Delta_v\) for \(v \in V_H\) can be recovered unless \(N(v; \bar{H}) = \emptyset\) or every \(u \in N(v; \bar{H})\) is of a different type from \(v\) with \(\Delta_u = 0\).

**Proof.** As failures can only occur within \(E_H\), nodes in \(N(v; \bar{H})\) must be in the same island as \(v\) after the attack. Under the proportional load shedding policy, we know that (i) if \(\exists u \in N(v; \bar{H})\) of the same type as \(v\), then we can recover the post-attack active power at \(v\) by \(p'_v/p_u = p'_u/p_u\) and thus recover \(\Delta_v\); (ii) if \(\exists u \in N(v; \bar{H})\) of a different type from \(v\) (e.g., \(u\) is a generator bus but \(v\) is a load bus) and \(\Delta_u \neq 0\), then \(\Delta_v\) must be zero. This proves the claim.

**Remark I:** Under the condition of Theorem III.1 i.e., \(B_{\bar{H}\mid H}\) has a full column rank, each \(v \in V_H\) must be the neighbor of \(\bar{H} \setminus H\) link. This can occur in hybrid control systems, as the PMU measurements are reported via a modern WAMPAC system with stronger defenses and the other sensor measurements are reported via a legacy SCADA system that is more vulnerable to cyber attacks [20], [21].
at least one node in \( H \) (otherwise its corresponding column in \( B_{\tilde{H} | H} \) will be \( 0 \)), and thus \( N(v; \tilde{H}) \neq \emptyset \). Moreover, majority of the nodes in practice are load buses, and thus each node in \( H \) is likely to be a load bus neighboring to another load bus in \( \tilde{H} \). Thus, we can usually recover \( \Delta_H \) under the proportional load shedding policy if the condition for recovering \( \theta_H \) holds.

Remark 2: Besides the cases indicated in Lemma [IV.1], it is also easy to show that if \( H \) contains no generator bus or no load bus, and \( \sum_{v \in V_H} \Delta_v = 0 \), then we must have \( \Delta_H = 0 \).

Next, we will establish the conditions for localizing the failed links \( F \) with known \( \theta' \) and \( \Delta \). The basic observation is the following property of the set \( F \).

**Lemma IV.2.** There exists a vector \( x \in \mathbb{R}^{|E_H|} \) that satisfies \( \text{supp}(x) = F \), and

\[
D_H x = B_{H|G}(\theta - \theta') - \Delta_H. \tag{10}
\]

**Proof.** Note that by definition, \( x_{st} \) defined in the proof of Lemma [III.1] is the same as the column corresponding to link \( (s, t) \) in \( D \). Define a vector \( y \in \mathbb{R}^{|E|} \) by

\[
y_e = \begin{cases} b_{st}(\theta' - \theta) & \text{if } e = (s, t) \in F, \\
0 & \text{o.w.} \end{cases}
\]

Then it is easy to see that \( \sum_{(s, t) \in E} b_{st} x_{st}^* \theta' = D y \). By [6], we have \( B(\theta - \theta') - \Delta = Dy \). Considering only the equations corresponding to \( V_H \) in \( D_H \) yields

\[
B_{H|G}(\theta - \theta') - \Delta_H = D_H y_H, \tag{12}
\]

where we have used the fact that \( y_H = 0 \). Thus \( x = y_H \) satisfies the conditions in the lemma. \( \square \)

This lemma, which replaces [3] Lemma 2, implies that if one can find the conditions under which the solution to (10) is unique, then the links corresponding to non-zero elements of this solution must be the failed links. To this end, [3] gave a set of graph-theoretic conditions. As these conditions are only about the solution space of \( D_H x = y \), they remain valid in our setting as long as the righthand side is known. We summarize these conditions below for completeness.

**Theorem IV.1.** The failed links \( F \) within the attacked area can be localized correctly if:

1. \( H \) is acyclic (i.e., a tree or a set of trees), in which case (10) has a unique solution \( x \) for which \( \text{supp}(x) = F \), or
2. \( H \) is a planar graph satisfying (i) for any cycle \( C \) in \( H \), \(|C \cap F| < |C \setminus F| \), and (ii) \( F^* \) is \( H^* \)-separable, in which case the optimization \( \min \|x\|_1 \) s.t. (10) has a unique solution \( x \) for which \( \text{supp}(x) = F \).

**Proof.** Condition (1) is implied by [3] Lemma 3], which proved that \( D_H \) has a full column rank if and only if \( H \) is acyclic. This combined with Lemma [IV.2] shows that if \( H \) is acyclic, then (10) only has one solution, and hence the support of this solution must be \( F \).

Condition (2) is implied by the proof of [3] Theorem 2], which showed that if \( H \) satisfies this condition, then any solution \( x \) to (10) must satisfy \( \|x\|_1 \geq \|x^*\|_1 \), where \( x^* \) is a vector satisfying the conditions in Lemma [IV.2]. Moreover, it showed that \( \|x\|_1 = \|x^*\|_1 \) only if \( x = x^* \). Thus, \( x^* \), whose support equals \( F \), can be computed by minimizing \( \|x\|_1 \) s.t. (10). \( \square \)

Special cases satisfying the second condition in Theorem [IV.1] include that (i) \( H \) is a cycle in which majority of the links have not failed, and (ii) \( H \) is a planar bipartite graph in which each cycle contains fewer failed links than non-failed links [3].

**V. LOCALIZING FAILED LINKS WITH UNKNOWN ACTIVE POWERS**

Although providing strong theoretical guarantees, the solutions for localizing failed links given in Section [IV] are only applicable to small attacked areas with simple topologies (e.g., trees or cycles in which every node is connected to another node outside the attacked area). To deal with larger attacked areas for which \( \Delta_H \) cannot be recovered by Lemma [IV.1], we investigate alternative solutions by jointly estimating the set of failed links \( F \) and the changes in active power injections \( \Delta_H \). As in Section [IV], we assume that the post-attack phase angles \( \theta' \) are known, which can be either inferred or directly measured.

**A. Solution**

Our approach is to formulate the joint estimation problem as an optimization as follows.

**Constraints:** Let \( x \in \{0, 1\}^{|E|} \) be an indicator vector such that \( x_{st} = 1 \) if and only if \( e \in F \). Due to \( B = D^T D^T \) (see Table [I] for the definitions), we can write the post-attack admittance matrix as \( B' = B - D \text{diag}(x) D^T \), which implies

\[
\Delta_H = B_{H|G}(\theta - \theta') + D_H \Gamma_H \text{diag}\{D_H^T \theta'\} x_H, \tag{13}
\]

where \( D_H \in \{-1, 0, 1\}^{|V_H||E_H|} \) is the submatrix of the incidence matrix \( D \) only containing the columns corresponding to links in \( H \). For simplicity, we define

\[
\tilde{D} := D \text{diag}\{D^T \theta'\}. \tag{14}
\]

For link \( e_k = (i, j) \), \( \tilde{D}_{i,k} = \tilde{D}_{j,k} = \frac{\theta'_i - \theta'_j}{r_{ij}} \), where \( |\theta'_i - \theta'_j| \) indicates the post-attack power flow on link \( e_k \) if it is operational.

Besides [13], \( \Delta_H \) is also constrained as

\[
p_v \geq \Delta_v \geq 0, \quad \forall v \in \{u \mid u \in V_H, p_u > 0\}, \tag{15a}
\]
\[
p_v \leq \Delta_v \leq 0, \quad \forall v \in \{u \mid u \in V_H, p_u \leq 0\}, \tag{15b}
\]
\[
1^T \Delta = 0, \tag{15c}
\]

which ensures that a generator/load bus will remain of the same type after the attack, and the total power is balanced. It is worth noting that [15c] is ensured by (13), which implies that \( 1^T \Delta_H - 1^T B_{H|G}(\theta - \theta') = (1^T D_H) x_H = 0 \) since \( 1^T D_H = 0 \) by definition (14). This implies that any \( \Delta_H \) satisfying (13) will satisfy \( 1^T \Delta_H = 1^T B_{H|G}(\theta - \theta') = 1^T \Delta_H^* \) (the ground-truth load shedding values in \( H \)), and thus satisfy (15c). Hence, we will omit (15c) in the sequel.

**Objective:** The problem of failure localization aims at finding a set \( \tilde{F} \) that is as close as possible to the set \( F \) of failed links,
while satisfying all the constraints. The solution is generally not unique, e.g., if both endpoints of a link $l \in E_H$ are disconnected from $\hat{H}$ after the attack, then the status of $l$ will have no impact on any observable variable, and hence cannot be determined. To resolve this ambiguity, we set our objective as using the fewest failed links to satisfy all the constraints. This idea has been applied to failure localization in power grid in various forms [3], [16]. Mathematically, the problem is formulated as

\[
(P0) \quad \min_{x_H} \quad 1^T x_H \\
\text{s.t.} \quad \begin{cases} 
\Delta_H \geq 0, \\
\sum_{i=1}^n x_i = T, & \forall t \in T_H, \\
x_i \in \{0, 1\}, & \forall e \in E_H,
\end{cases}
\]  

(16a)

where the decision variables are $x_H$ and $\Delta_H$. Next, we characterize the complexity of (P0).

**Lemma V.1.** The optimization (P0) is NP-hard.

**Proof.** We will prove the claim by a reduction from the subset sum problem, which is known to be NP-hard [24]. Given any set of non-negative integers \( \{ f_i \geq 0 \}_{i=1}^n \) and a target value $T$, the subset sum problem determines whether there exists \( \{ x_i \in \{0, 1\} \}_{i=1}^n \) such that \( \sum_{i=1}^n x_i f_i = T \). For each subset sum instance, we construct the following star-shaped attacked area $H$: let $H = (V_H, E_H)$ such that $V_H$ is composed of $n + 1$ nodes, where node $u_0$ is the hub with $p_{u_0} = 0$ and $\theta_{u_0} = 0$, and node $u_i$ ($i \in [n]$) for $[n] := \{1, \ldots, n\}$ is incident to only one link $e_i = (u_0, u_i)$, with $p_{e_i} = -f_i$, $\theta_{e_i} = -f_i$, and $r_{e_i} = 1$. In addition, $u_0$ is connected to $v \in V_H$, with $\theta'_{e} = \sum_{i=1}^n f_i + T$, through link $e_0 = (u_0, v)$ with $r_{e_0} = 1$.

By substituting (13) and $p_{u_0} = 0$, (15b) for node $u_0$ becomes $D_{H,u_0} x_{H} = B_{u_0|G} \theta'$, where $D_{H,u_0}$ is the row of $D_{H}$ corresponding to node $u_0$. Since $(D_{H,u_0})_i = \theta_{e_0} - \theta_{e_i}$, it is easy to check that $D_{H,u_0} x_{H} = \sum_{i=1}^n f_i x_i$. Moreover, $B_{u_0|G} \theta' = \sum_{i=1}^n f_i + \theta_{e_0} - \theta_{e_i}$, which is equal to $T$. Since $u_i$ ($i \in [n]$) is connected to only one link $e_i = (u_0, u_i)$, we have that $D_{H,u_i} x_{H} = -f_i x_i$ and $B_{u_i|G} \theta' = -f_i$. Thus, (15b) for $u_i$ becomes $-f_i \leq -f_i x_i \leq 0$, which is satisfied whatever value $x_i$ takes. Therefore, a subset sum instance returns true if and only if the instance of (P0) constructed as above is feasible, which completes the proof. \[\square\]

By relaxing the integer constraint (16d), (P0) is relaxed into

\[
(P1) \quad \min_{x_H} \quad 1^T x_H \\
\text{s.t.} \quad \begin{cases} 
\Delta_H \geq 0, \\
\sum_{i=1}^n x_i = T, & \forall t \in T_H, \\
0 \leq x_i \leq 1, & \forall e \in E_H,
\end{cases}
\]  

(17a)

where $0 \leq x_H \leq 1$ denotes element-wise inequality. The problem (P1) is a linear program (LP) which can be solved in polynomial time. Based on (P1), we propose an algorithm for localizing the failed links, given in Algorithm [I] where the input parameter $\eta \in (0, 1)$ is a threshold for rounding the fractional solution of $x_H$ to an integral solution ($\eta = 0.5$ in our experiments). We will analyze how $\eta$ affects the trade-off between miss rate and false alarm rate of Algorithm [I] at the discussion after Theorem V.2 at Section V-B

**Algorithm 1: Failed Link Detection**

**Input:** $B, p, \Delta_H, \theta, \theta', D, \eta$  
**Output:** $\hat{F}$

1. Solve the problem (P1) to obtain $x_H$;
2. Return $\hat{F} = \{ e : x_e \geq \eta \}$.

**B. Analysis**

We now analyze when the proposed algorithm can correctly localize the failed links. In the sequel, $\Delta_H$ denotes the ground-truth load shedding values in $H$ and $x_H^*$ denotes the ground-truth failure indicators (i.e., $x_e^* = 1$ if $e \in F$ and $x_e^* = 0$ if $e \in E_H \setminus F$).

According to (15), we decompose $V_H$ into $V_{H,L}$ for nodes with $p_e \leq 0$ and $V_{H,G}$ for the rest. Define $E_1 \subseteq E_H$ as the set of links that operate normally after failure, and $F \subseteq E_H$ as the failed links. We make the following assumption:

**Assumption 1.** As in [3], we assume that for each link $(s, t) \in E_H$, $\theta_s^* \neq \theta_t^*$, as otherwise the link will carry no power flow and hence its status cannot be identified.

1) Main Results: First, we simplify (P1) into an equivalent but simpler optimization problem. To this end, we combine the decision variables $\Delta_H$ and $x_H$ of (P1) into a single vector $y_H = [\Delta_H^T, x_H^T] \in \mathbb{R}(|E_H|+|V_H|)$ (where $[A, B]$ denotes horizontal concatenation), and explicitly represent the solution to $y_H$ that satisfies (13). Notice that (13) can be written as $[I_{|V_H|}, -D_H] y_H = B_{H|G} (\theta - \theta')$ (where $I_{|V_H|}$ is the identity matrix). The ground-truth solution $y_H^* = [(\Delta_H^*)^T, (x_H^*)^T]^T$ certainly satisfies (13). Next, consider the null space of $[I_{|V_H|}, -D_H]$, whose dimension is $|E_H|$. It is easy to verify that $[d_e^T, u_e^T]^T (e \in E_H)$ are $|E_H|$ independent vectors spanning the null space of $[I_{|V_H|}, -D_H]$, where $d_e$ is the column vector of $D_{H}$ corresponding to link $e$, and $u_e$ is a unit vector in $\mathbb{R}^{|V_H|}$ with the $e$-th element being 1 and the other elements being 0. Therefore, any $y_H$ satisfying (13) can be expressed as

\[
y_H = \begin{bmatrix} \Delta_H^* \\ x_H^* \end{bmatrix} + \sum_{e \in E_H} c_e \begin{bmatrix} d_e \\ u_e \end{bmatrix}.
\]  

(18)

Based on the decomposition of $V_H$ into $V_{H,L}$ and $V_{H,G}$, $\hat{D}_H$ and $\Delta_H$ can be written as

\[
\hat{D}_H = \begin{bmatrix} V_{H,L} & \hat{D}_{H,L} \end{bmatrix}, \quad \Delta_H = \begin{bmatrix} \Delta_{H,L} \\ \Delta_{H,G} \end{bmatrix}.
\]  

(19a)

This assumption essentially means that we will ignore the existence of such links in failure localization.
Let $\tilde{D}_{H,L} := [\tilde{D}_{H,L,1}, \tilde{D}_{H,L,F}]$, $\tilde{D}_{H,G} := [\tilde{D}_{H,G,1}, \tilde{D}_{H,G,F}]$, and $c := (c_e)_{e \in E_H} \in \mathbb{R}^{|E_H|}$. Since $\Delta_{H,L}$ and $\Delta_{H,G}$ are constrained differently in (13a) and (13b), we introduce $\Delta_L = [I_{|V_H,L|}]$ and $\Delta_G = [0, I_{|V_H,G|}]$ such that $\Delta_{H,L} = \Delta_L \Delta_H$, $\Delta_{H,G} = \Delta_G \Delta_H$. According to (18), for Algorithm I to correctly localize the failed links, it suffices to have $x_e^* + c_e \geq \eta$ for all $e \in F$ and $x_e^* + c_e < \eta$ for all $e \notin F$. Equivalently, it suffices to ensure that the optimal solution $e^*$ to the following optimization problem satisfies $c_e^* \geq \eta - 1$ for all $e \in F$ and $c_e^* < \eta$ for all $e \notin F$:

$$\begin{align*}
\min & \quad 1^T c \\
\text{s.t.} & \quad \tilde{D}_{H,L} c \leq -\Delta_L^*, \quad (20a) \\
& \quad -\tilde{D}_{H,L} c \leq -(\Delta_L p_H - \Delta_{H,L}^*), \quad (20b) \\
& \quad -\tilde{D}_{H,G} c \leq \Delta_{H,G}, \quad (20c) \\
& \quad -\tilde{D}_{H,G} c \leq \Delta_G c - \Delta_{H,G}^*, \quad (20d) \\
& \quad e \leq x_e^*, \quad (20e) \\
& \quad e \leq 1 - x_e^*, \quad (20f) \\
& \quad 0 \leq c \leq \eta - 1. \quad (20g)
\end{align*}$$

This equivalent formulation of (P1) will help to simplify our analysis by eliminating the equality constraint in (15). For notational simplicity, we will omit the subscript $H$ in the sequel unless it causes confusion.

Next, we use (20) to analyze the accuracy of Algorithm I. Let $\hat{F}$ be the failed link set returned by Algorithm I. We first define $Q_m = F \setminus \hat{F}$ as the set of failed links that are not detected, and $Q_f = F \setminus \hat{F}$ as the set of operational links that are falsely detected as failed. Note that according to (20), a failed link $e \in F$ is missed if and only if $c_e^* > \eta - 1$. Similarly, an operational link $e \notin F$ is falsely detected as failed if and only if $c_e^* > \eta$. To express this in a vector form, we define $W_m \in \{0, 1\}^{Q_m \times |E_H|}$ as a binary matrix, where for each $i = 1, \ldots, Q_m$, $(W_m)_{i,j} = 1$ if the $i$-th missed link is link $e_j$ and thus we have $W_m e^* \leq (\eta - 1) 1$. Similarly, $W_f \in \{0, 1\}^{Q_f \times |E_H|}$ is defined such that $(W_f)_{i,j} = 1$ if the $i$-th false-alarmed link is link $e_j$, which leads to $-W_f c^* \leq -\eta 1$. For ease of presentation, we define

$$\begin{align*}
A_D^T & := [\tilde{D}_L, -\tilde{D}_L, -\tilde{D}_G, \tilde{D}_G]^T \in \mathbb{R}^{|E_H| \times 2|V_H|}, \quad (21a) \\
A_x^T & := [-I_{|E_H,1|}, I_{|E_H,1|}] \in \mathbb{R}^{|E_H| \times 2|E_H|}, \quad (21b) \\
W^T & := [W_m^T, -W_f^T] \in \mathbb{R}^{|E_H| \times |Q_m| + |Q_f|}, \quad (21c) \\
g_D^T & := [-A_D^T, 0, -p_L^T, (\eta e^*)^T, (\eta e^*)^T]^T, \quad (21d) \\
g_x^T & := [(\eta - 1)^T, -x_e^*]^T \in \mathbb{R}^{1 \times |Q_m| + |Q_f|}, \quad (21e) \\
g_w^T & := [(\eta - 1)^T, -\eta 1]^T \in \mathbb{R}^{1 \times |Q_m| + |Q_f|}. \quad (21f)
\end{align*}$$

Then the constraints for (20) can be succinctly written as $[A_D^T, A_x^T]^T c \leq [g_D^T, g_x^T]^T$, and the optimal solution must satisfy $W c \leq g_w$. Next, we will give a simple but important observation, which lays the groundwork for further analysis.

**Lemma V.2.** A link $e \in F$ cannot be missed ($e \notin F \setminus \hat{F}$) by Algorithm I if for any $Q_m$ containing $e$, there is always a solution $z \geq 0$ to

$$\begin{align*}
[A_D^T, A_x^T, W^T, 1] z &= 0, \quad (22a) \\
[g_D^T, g_x^T, g_w^T, 0] z &< 0. \quad (22b)
\end{align*}$$

Similarly, a link $e \in E_H \setminus F$ cannot be falsely detected as failed by Algorithm I if for any $Q_f$ with $e \in Q_f$, there always exists a solution $z \geq 0$ to (22).

**Proof.** First note that $c = 0$ is feasible to (20). If $\hat{F}$ is returned by Algorithm I with $e \in F \setminus \hat{F}$, there must exist an optimal solution $e^*$ to (20) such that $c_e^* \leq \eta - 1$. Thus, $e^*$ must be feasible to (20) with $1^T c^* \leq 0$, which can be formulated as

$$[A_D^T, A_x^T, W^T, 1]^T c^* \leq [g_D^T, g_x^T, g_w^T, 0]^T, \quad (23)$$

where $W$ and $g_w$ are defined for some $Q_m$ and $Q_f$ such that $e \in Q_m$. According to Gale’s theorem of alternative (25), there is a solution to (23) if and only if there is no solution $z \geq 0$ to (22). That is to say, $e \notin F \setminus \hat{F}$ if there is always a non-negative solution to (22) for any $Q_m$ containing $e$. The condition for $e \notin F \setminus \hat{F}$ can be proved similarly.

For ease of presentation, we will introduce a few notations as follows. Denote $D_u$ as the row in $D$ corresponding to node $u$, and $D_{u,e}$ as the entry in $D_u$ corresponding to link $e$. Recall that as defined in (14), if $e = (u, v)$, then

$$\tilde{D}_{u,e} = \frac{\theta_u - \theta_v}{r_{uv}}. \quad (24)$$

Define $z_D \in \mathbb{R}^{2|V_H|}$ as the subvector of $z$ corresponding to $A_D$. More specifically, since both $D$ and $-D$ are included in $A_D$, denote $z_D$ as the entry in $z_D$ corresponding to $D_u$ in $A_D$ and $z_D$, as the entry for $-D_u$ in $A_D$. Define $g_{D,u}$ and $g_{D,-u}$ as the entries in $g_D$ corresponding to $z_D$ and $z_D$, respectively. That is, if $p_u \leq 0$, then $g_{D,u} := -\Delta_u^*$ and $g_{D,-u} := -p_u^*$; if $p_u > 0$, then $g_{D,u} := p_u^*$ and $g_{D,-u} := \Delta_u^*$. For $A_D^T$, $W^T$ and $1$ in (22a), $z_x \in \mathbb{R}^{2|E_H|}$, $z_w \in \mathbb{R}^{|Q_m| + |Q_f|}$ and $z_s \in \mathbb{R}$ are their counterparts in $z$. Specifically, if link $e \in F$ is the $i$-th link in $Q_m$, we have $(W_m)_{i,j} = 1$ and $z_w, z_e$ is used to denote the entry in $z_w$ that corresponds to the $i$-th column of $W_m$; $z_w, z_e$ is defined similarly if $e \in Q_f$. For each link $e$, we denote $z_{x,-e}$ as the entry in $z_x$ corresponding to $x_e$ in $g_x$ and $z_{x,+e}$ as the entry corresponding to $(1 - x_e)$ in $g_x$.

Although Lemma V.2 can already be utilized as recovery conditions, it does not explicitly characterize what type of links are guaranteed to be correctly identified. To this end, we will show that a link will satisfy the conditions in Lemma V.2 (and can thus be correctly identified by Algorithm I) if its endpoints satisfy certain conditions. To make our conditions as general as possible, we introduce a generalization of node called hyper-node as follows (a single node is also a hyper-node):

**Definition V.1.** A set of nodes $U \subseteq V_H$ is a hyper-node if they induce a connected subgraph before attack.

We define a few properties of a hyper-node $U$. Define $E_U$ as the set of links with exactly one endpoint in $U$, i.e., $E_U := \{e \in E \mid u, v \in U \}$.
\[ \{ e | e = (s, t) \in E_H, s \in U, t \not\in U \} . \]

If \( E_U \cap F \neq \emptyset \), we define
\[
\hat{D}_{U,e} := \sum_{u \in U} \hat{D}_{u,e}, \quad (22a)
\]
\[
S_U := \{ e \in E_U \setminus F | \exists l \in E_U \cap F, \hat{D}_{U,l} \hat{D}_{U,l} > 0 \}, \quad (22b)
\]
\[
f_{U,0} := \max_{e \in S_U} \left| \hat{D}_{U,e} \right|, \quad (22c)
\]
\[
f_{U,1} := \min_{e \in \hat{E} \cap F} \left| \hat{D}_{U,e} \right|, \quad (22d)
\]
\[
f_{U,g} := \begin{cases} \sum_{u \in U} g_{D,u} & \text{if } \exists E \in U \cap F, \hat{D}_{U,l} < 0, \\ \sum_{u \in U} g_{D,-u} & \text{otherwise}. \end{cases} \quad (22e)
\]

**Example 1.** Consider an attacked area \( H \) as shown in Fig. 2, where blue circles denote nodes (buses) while the direction of each link indicates the direction of power flow. Suppose that \( F = \{ l_2, l_9 \} \) and all nodes are load buses. Nodes \( u_1, u_2 \) and \( u_3 \) form a hyper-node \( U \), where \( E_U = \{ l_2, l_4, l_6, l_7 \} \).

\[
S_U = \{ l_2 \}, \quad f_{U,0} = |\hat{D}_{U,l_2}|, \quad f_{U,1} = \min\{|\hat{D}_{U,l_4}|, |\hat{D}_{U,l_6}| \} \text{ and } f_{U,g} = -\sum_{e \in E_U} \Delta^e_{U} \hat{D}_{U,l_e} = \hat{D}_{U,l_4} + \hat{D}_{U,l_6} = 0 \text{ since } l_4 \neq E_U, \text{ while } \hat{D}_{U,l_2} = \hat{D}_{U,l_4} \neq 0 \text{ since } l_2 \in E_U.
\]

Based on these definitions and Lemma V.2, we are ready to present a condition under which a failed link \( l \in F \) will not be missed by Algorithm 1.

**Theorem V.1.** A failed link \( l \in F \) will be detected by Algorithm 1, i.e., \( l \in \hat{F} \), if there exists at least one hyper-node (say \( U \)) such that \( l \in E_U \), for which the following conditions hold:
1. \( \forall e, l \in E_U \cap F, \hat{D}_{U,e} \hat{D}_{U,l} > 0 \),
2. \( S_U = 0 \), and
3. \( f_{U,g} + (\eta - 1)|\hat{D}_{U,l}| < 0 \).

**Proof.** We will prove by showing that there is always a solution to (22a) for any possible \( Q_f \) and \( Q_m \) if \( l \in E_U \cap Q_m \). We prove this by directly constructing a solution \( z \) for (22a) as follows: \( \forall u \in U, \text{ if } \hat{D}_{U,l} < 0, \text{ set } z_{D,u} = 1; \text{ otherwise, set } z_{D,-u} = 1. \)

\[
\sum_{u \in U} \Delta^e_{U} \hat{D}_{U,l_e} = \hat{D}_{U,l_4} + \hat{D}_{U,l_6} = 0 \text{ since } l_4 \neq E_U, \text{ while } \hat{D}_{U,l_2} = \hat{D}_{U,l_4} \neq 0 \text{ since } l_2 \in E_U.
\]

The left-hand-side of (22a) corresponding to link \( l \) is expanded as \(-|\hat{D}_{U,l}| + \sum_{e \in E_U} \hat{D}_{u,e} \hat{D}_{u,l} = -|\hat{D}_{U,l}| + \sum_{e \in E_U} \hat{D}_{u,e} \hat{D}_{u,l} = -|\hat{D}_{U,l}| + \sum_{e \in E_U} \hat{D}_{u,e} \hat{D}_{u,l} = 0 \text{ due to condition 1). Second, since } S_U = 0, \text{ for all } e' \in E_U \setminus F, \text{ the corresponding row in (22a) is expanded into } |\hat{D}_{U,e_2}| - \sum_{e \in E_U} \hat{D}_{u,e} \hat{D}_{u,l} = 0. \text{ Other rows of (22a) hold trivially since they only involve the zero-entries in the constructed } z. \text{ Thus, (22a) holds under this assignment. As for (22b), its left-hand-side can be expanded as } f_{U,g} + (\eta - 1)|\hat{D}_{U,l}| < 0 \text{ due to condition 3). According to Lemma V.2, } l \in E_U \cap F \text{ will not be missed, which completes the proof.} \]

Based on similar arguments, the following condition can guarantee that an operational link \( l \in E_H \cap F \) will not be falsely detected by Algorithm 1, i.e., \( l \notin \hat{F} \), if there exists at least one hyper-node (say \( U \)) such that \( l \in E_U \), for which the following conditions hold:
1. \( \forall l, l' \in E_U \setminus F, \hat{D}_{U,l} \hat{D}_{U,l'} > 0 \),
2. \( S_U = 0 \) if \( E_U \cap F = \emptyset \), and
3. \( f_{U,g} - \eta|\hat{D}_{U,l}| < 0 \).

**Theorem V.2.** An operational link \( l \in E_H \setminus F \) will not be detected (as failed) by Algorithm 1, i.e., \( l \notin \hat{F} \), if there exists at least one hyper-node (say \( U \)) such that \( l \in E_U \), for which the following conditions hold:
1. \( \forall l, l' \in E_U \setminus F, \hat{D}_{U,l} \hat{D}_{U,l'} > 0 \),
2. \( S_U = 0 \) if \( E_U \cap F = \emptyset \), and
3. \( f_{U,g} - \eta|\hat{D}_{U,l}| < 0 \).

**Proof.** Similar to the proof of Theorem V.1, we will prove by showing that there is always a solution to (22a) for any possible \( Q_f \) and \( Q_m \), if \( l \in E_U \cap Q_f \). We construct the following \( z \):
\[
\forall u \in U, \text{ if } \hat{D}_{U,l} < 0, \text{ set } z_{D,u} = 1; \text{ otherwise, set } z_{D,-u} = 1. \text{ Set } z_{w,f,l} = |\hat{D}_{U,l}|, z_{e,-e'} = |\hat{D}_{U,e'}| \text{ for } e' \in E_U \setminus (F \cup \{l\}), \text{ and other entries of } z \text{ to 0. Then, it is easy to check that (22a) is satisfied. As for (22b), considering that } g_{D}^T z_D + g_{z}^T z_e - \eta g_{z,w,f,l} = f_{U,g} - \eta|\hat{D}_{U,l}| < 0, \]

where \( f_{U,g} = \sum_{u \in U} g_{D,u} \) if \( \hat{D}_{U,l} > 0 \) and \( f_{U,g} = \sum_{u \in U} g_{D,-u} \) if \( \hat{D}_{U,l} < 0 \), and the last inequality holds due to condition 3). Thus, according to Lemma V.2, \( l \notin \hat{F} \), which completes the proof.
proof of these theorems, which means that the optimality condition \( 1^T e^* \leq 0 \) formulated in (23) has not been exploited.

This results in the requirement of the condition "\( S_U = \emptyset \)" in these theorems. To better characterize the accuracy of Algorithm 1, we will establish a condition that exploits the optimality condition. To this end, we introduce a few further definitions as follows.

**Definition V.2.** A hyper-node \( U \) is a fail-cover hyper-node if \( E_U \cap F \neq \emptyset \) and \( \forall e, l \in E_U \cap F, D_{U,l} \neq D_{U,l} > 0 \).

Given a set of fail-cover hyper-nodes \( T \), we divide \( T \) into \( T_n = \{U_{n_i}\} \) and \( T_p = \{U_{p_i}\} \), such that each \( U_{n_i} \in T_n \) satisfies \( D_{U_{n_i},e} = 0 \) for all \( e \in F \), and each \( U_{p_i} \in T_p \) satisfies \( D_{U_{p_i},e} \geq 0 \) for all \( e \in F \). Then, we define:

\[
R_{U_l} := \max_{U \in T} \left\{ \frac{f_{U,l}}{f_{u,1}} \right\}, \quad \forall U_l \in T, \tag{28a}
\]

\[
\tilde{D}_{T,e} := \sum_{U_l \in T_n} R_{U_l} D_{U_l,e} + \sum_{U_l \in T_p} (-R_{U_l}) D_{U_l,e}, \tag{28b}
\]

\[
S_T := \{e' \notin F|3e \in F \text{ s.t. } \tilde{D}_{T,e'} > 0\}, \tag{28c}
\]

\[
f_{T,0} := \max_{e \in S_T} |\tilde{D}_{T,e'}|, \tag{28d}
\]

\[
f_{T,1} := \min_{e \in F} |\tilde{D}_{T,e'}|, \tag{28e}
\]

\[
f_{T,g} := \sum_{U_l \in T} R_{U_l} f_{U_l,g}. \tag{28f}
\]

Now we provide the conditions for Algorithm 1 to work correctly when condition 2) in Theorems V.1 and V.2 is relaxed.

**Theorem V.3.** Assume that there exists a set of fail-cover hyper-nodes \( T = \{U_i\} \) satisfying the following conditions:

1. \( \forall e \in F, f_{T,0} \geq f_{T,0} \), i.e., \( f_{T,1} \geq f_{T,0} \) and
2. \( F \subseteq \bigcup_{U_i \in T} E_{U_i} \)

Then, a failed link \( e \in F \) will be detected by Algorithm 1 \((e \in F)\) if

\[
f_{T,g} + (\eta - 1)(|\tilde{D}_{T,e'}| - f_{T,0}) < 0. \tag{29}
\]

In addition, an operational link \( e' \in E_H \setminus F \) will not be detected (as failed) by Algorithm 1 \((e' \notin F)\) if

\[
\begin{cases}
    f_{T,g} - \eta \left( f_{T,1} + |\tilde{D}_{T,e'}| \right) < 0 & \text{if } e' \notin S_T \\
    f_{T,g} - \eta \left( f_{T,1} - |\tilde{D}_{T,e'}| \right) < 0 & \text{if } e' \in S_T
\end{cases} \tag{30}
\]

**Proof.** We first prove the condition for a failed link \( e \in F \).

Based on Lemma V.2 we prove by constructing solution to (22) w.r.t any possible \( Q_f \) and \( Q_m \) if \( e \in F \) is the \( i^{th} \) link in \( Q_m \) and \( (W_m)_i = 1 \). We prove by directly constructing the following \( z \) for (22): (i) \( z_s = f_{T,0} \); (ii) \( \forall v \in V_H, z_{D,v} = \sum_{U_l \in T} \tilde{R}_{U_l}(v) R_{U_l} \); (iii) \( \forall l \in E_H \setminus F, z_{x_l} = |\tilde{D}_{T,e'}| + z_s \); (iv) \( \forall l \in E_H \setminus \{\} \), \( z_{x_l} \geq 0 \) due to the definition of \( f_{T,0} \). Furthermore, \( \forall l \in E_H \setminus \{\} \), \( z_{x_l} \geq 0 \) since \( |\tilde{D}_{T,e'}| \geq f_{T,0} \) by the assumption on \( T \), and \( z_{w,m,e} \geq 0 \) for a similar reason. We will show that (22) is satisfied under this assignment. First, note that according to the definition in (28b), \( D_{T,l} \leq 0 \) for all \( l \in F \), which implies that \( D_{T,l} \leq 0 \) for all \( l \in S_T \) and \( D_{T,l} \geq 0 \), \( \forall l \in (E_H \setminus F) \setminus S_T \). Thus, \( \forall l \in F \setminus \{\} \), the left-hand-side of (22a) can be expanded as \( \sum_{U_l \in T_n} R_{U_l} D_{U_l,l} + \sum_{U_l \in T_p} R_{U_l}(-D_{U_l,l}) + z_{x+l} + z_s = D_{T,l} + |\tilde{D}_{T,e'}| - z_s = 0 \). Similarly, the row of the left-hand-side of (22b) corresponding to \( e \) can be expanded as \( D_{T,e} + z_{w,m,e} + z_s = 0 \), while the rows corresponding to \( l \in E_H \setminus F \) can be expanded as \( D_{T,l} - z_{x-l} + z_s = 0 \). Moreover, the left-hand-side of (22b) can be expanded as \( f_{T,g} + (\eta - 1)(|\tilde{D}_{T,e'}| - f_{T,0}) \), which satisfies (22b) due to (29).

Note that this assignment of \( z \) is valid for any possible \( Q_m \) and \( Q_f \) with \( e \notin Q_m \). That is to say, there is always a non-negative solution to (22) if \( e \in Q_m \), which implies that \( e \) will not be missed by Algorithm 1 according to Lemma V.2.

Next, we prove the condition for an operational link \( e' \in E_H \setminus F \). Again, we prove by constructing a solution to (22) w.r.t any possible \( Q_f \) and \( Q_m \) with \( e' \in Q_f \). To this end, we construct the following assignment for \( z \): (i) \( z_s = f_{T,0} \); (ii) \( \forall v \in V_H, z_{D,v} = \sum_{U_l \in T} \tilde{R}_{U_l}(v) R_{U_l} \); (iii) \( \forall l \in E_H \setminus F \) and \( l \neq e' \), set \( z_{x-l} = D_{T,l} + z_s \); (iv) \( \forall l \in F, z_{x-l} = |\tilde{D}_{T,e'}| - z_s \); (v) \( z_{w,m,e} = |\tilde{D}_{T,e'}| + z_s \); (vi) the rest entries of \( z \) are set as 0. We will show that \( z \geq 0 \). For \( l \in F, z_{x+l} \geq 0 \) due to the definition of \( f_{T,0} \). Thus, \( \forall l \in E_H \setminus (F \cup \{e'\}) \), \( l \neq l_T \), \( z_{x-l} = 0 \) since \( D_{T,l} = 0 \) if \( l \in E_H \setminus F \).

For ease of understanding, we visualize the relationship among the results in this section in Fig. 5. Specifically,

- Algorithm 1 can return a \( \bar{F} \neq F \), whose corresponding \( e^* \) satisfies (23). This is because (P1) can have multiple optimal solutions while a typical linear programming solver can only return one of them, which can result in \( \bar{F} \neq F \). Thus, Lemma V.2 can only cover part of all the cases.

- The cases satisfying Lemma V.2 can be classified into four categories according to whether they satisfy Theorem V.1 or V.2.

- At a first glance, (29) in Theorem V.3 seems to include all the cases covered by Theorem V.1 due to the relaxed condition on \( S_U \). However, (29) depends on the value of \( f_{T,0} \), which is the weighted sum of all \( f_{U_l,g} \) for \( U_l \in T \). For a failed link, \( e \), it is possible to find a hyper-node \( U \) with \( f_{U_l,g} = 0 \) such that Theorem V.1 holds, while (29) is violated. This argument applies similarly to a link \( e' \in E_H \setminus F \). Thus, cases covered by Theorem V.1 or V.2 and those covered by Theorem V.3 partially overlap. A quantitative analysis will be shown in Fig. 7 through experiments.

**Example 2.** Consider Fig. 2 as an example, where all the nodes in \( H \) are load buses. If \( \Delta^2 = 0, \forall u \in V_2 := \{u_1, u_2, u_3, u_6, u_7\} \), then \( T = \{V_2\} \) satisfies that \( f_{T,g} = \)}
All Cases
Theo. v.1 and v.2

Figure 3. Relationship between cases covered by various theorems.

\[ \sum_{u \in V_2} g_{D,u} = -\sum_{u \in V_2} \Delta_u = 0 \text{ and } S_T = \emptyset, \text{ which leads to the satisfaction of (29) for all the failed links and (30) for all the operational links.} \]

Meanwhile, \( l_2 \) and \( l_6 \) are guaranteed to be correctly identified through Theorem V.2 by setting \( U = V_2 \), since \( S_{U} = \emptyset \) and \( f_{U,g} = 0 \). Furthermore, both \( l_3 \) and \( l_5 \) satisfy Theorem V.2 if \( U = \{u_3, u_6, u_7\} \). As can be seen, some failed links (\( l_2 \) and \( l_4 \)) can be covered by both Theorem V.2 and Theorem V.7. Also, some operational links (\( l_3 \) and \( l_5 \)) can be covered by both Theorem V.3 and Theorem V.2. However, \( l_1 \) can only be covered by Theorem V.2.

2) Special Cases: We now use these theorems to analyze the accuracy of Algorithm 1 in special cases of practical interest.

No Islanding: Equipped with Theorems V.1. and V.2, we can examine the accuracy of Algorithm 1 in the special case that the grid stays connected after failure, which has been studied in [3]. It is worthwhile to examine whether Algorithm 1 can correctly identify \( F \) in this case (where \( \Delta^* = 0 \)) because Algorithm 1 assumes no prior knowledge of \( \Delta^* \).

Corollary V.1. If the grid stays connected after failure, \( H \) is acyclic, and \( E \) contains either no load bus or no generator bus, then Algorithm 1 can correctly detect \( F \), i.e., \( \hat{F} = F \).

Proof. We only prove the case that \( H \) contains no generator bus since the other case can be proved similarly. We first prove that any failed link \( l \in F \) will not be missed (\( l \in \hat{F} \)). Under Assumption 1, link \( l \) must have one endpoint (say \( u \)) such that \( D_{u,l} < 0 \). Next, we will build a hyper-node \( U \) such that the induced subgraph is a tree rooted at node \( u \). Specifically, such hyper-node can be constructed by breadth-first search (BFS) starting from node \( u \). In the first iteration of BFS, we start with \( U = \{u\} \) and add a neighbor \( v_i \) of \( u \) into \( U \) if \( e = (u, v_i) \in F \) with \( D_{u,i} < 0 \). Then, we repeatedly add node \( v \) into \( U \) if \( \exists e = (s, v) \in E_U \cap F \) such that \( D_{U,i} < 0 \) or \( \exists e = (s, v) \in E_U \) such that \( D_{U,i} < 0 \). This procedure will terminate since \( H \) is acyclic, and the constructed \( U \) will satisfy condition 1) and condition 2) of Theorem V.1. Since all nodes \( u \in U \) are load buses, \( D_{u,l} < 0 \), and the grid stays connected after failure, we have \( f_{U,g} = -\sum_{u \in U} \Delta_u = 0 \), which satisfies condition 3) of Theorem V.1. Thus, we have \( F \subseteq \hat{F} \).

Next, we show that any operational link \( e \in E_H \setminus F \) will not be falsely detected by Algorithm 1 (\( e \notin \hat{F} \)). Under Assumption 1, link \( e \) must have one endpoint (say \( u \)) such that \( D_{u,e} > 0 \). The hyper-node \( U \) can be constructed as follows: start with \( U = \{u\} \), add node \( v \) into \( U \) if \( \exists e' = (s, v) \in E_U \cap F \) or \( \exists e' = (s, v) \in E_U \setminus F \) such that \( D_{U,e}, D_{U,e'} < 0 \). The resulting hyper-node must satisfy condition 1) and condition 2) of Theorem V.2. Again, we have \( f_{U,g} = -\sum_{u \in U} \Delta_u = 0 \), which leads to satisfaction of condition 3) in Theorem V.2. Therefore, we have \( \hat{F} \subseteq F \).

Remark: If the grid stays connected after failure, \( H \) contains either no load bus or no generator bus, and \( H \) contains cycles, then the status of a link \( l \in E_H \) is guaranteed to be correctly identified if \( l \) is not in any cycle. This is because in this case, we can construct a hyper-node satisfying the conditions in Theorem V.1 or V.2 as in the proof of Corollary V.1.

Islanding: Now, we study the case where the attacked area is decomposed into multiple islands. Formally, suppose that the failures partition \( H \) into \( K \) islands \( H_i = (V_i, E_i) \) (\( i = 1, \ldots, K \)), where \( V_H = \bigcup_{i=1}^{K} V_i \), \( V_i \cap V_j = \emptyset \) (\( i \neq j \)), and \( E_H = (\bigcup_{i=1}^{K} E_i) \cup E_c \), with \( E_c \) being the set of links between different islands. Then the following is implied by Theorem V.3.

Corollary V.2. Suppose that \( F = E_c \). Let \( E_{c,i} \subseteq E_c \) be the subset of failed links with one endpoint in \( H_i \), then, Algorithm 1 will correctly detect the failures (\( \hat{F} = F \)) if there exists a set \( \mathcal{L} \subseteq \{H_i\}_{i=1}^{K} \) with \( \bigcup_{H_i \in \mathcal{L}} E_{c,i} = E_c \), such that each \( H_i \in \mathcal{L} \) satisfies the following condition:

1) \( \forall e, e' \in E_{c,i}, \; D_{V_i,e} \cdot D_{V_j,e'} > 0 \),
2) \( \forall v \in V_i \) that is incident to a link in \( l \in E_{c,i} \), \( f_{v,l} < 0 \), and
3) \( \forall v \in V_i \) that is incident to a link in \( l \in E_{c,i} \), \( g_{D,v} > 0 \).

Proof. First, we construct a set of fail-cover hyper-nodes \( T \) as required in Theorem V.3. Let \( V_{L,c} \) contain all the nodes in \( \bigcup_{H_i \in \mathcal{L}} V_i \) that are incident to at least one link in \( E_c \). Formally, \( V_{L,c} := \{v_j\}_{j=1}^{K} \), where \( \forall v_j \in V_{L,c}, \; \exists l \in E_c \) such that \( D_{v_j,l} \neq 0 \). We construct \( T := \{U_j\}_{j=1}^{K} \), where \( U_j := \{v_j\} \), as the set of fail-cover hyper-nodes, which automatically satisfies condition 2) in Theorem V.3.

Next, we will show that the constructed \( T \) satisfies condition 1) in Theorem V.3 with strict inequality. To this end, consider any \( e = (u_1, u_2) \in S_{U} \). Suppose that \( u_1, u_2 \in V \). Recall from the proof of Theorem V.3 that \( D_{T,e} < 0 \), and hence at least one of \( U_1 := \{u_1\} \) and \( U_2 := \{u_2\} \) must be in \( T \) (as otherwise \( D_{T,e} = 0 \)). If only \( U_1 \in T \) (\( U_2 \notin T \)), then by (28a), \( D_{T,e} = R_{U_1} D_{U_1,e} \text{ if } D_{U_1,e} > 0, \forall l \in E_{c,i}, \) and \( D_{T,e} = -R_{U_1} D_{U_1,e} \text{ if } D_{U_1,e} > 0, \forall l \in E_{c,i} \). To satisfy \( D_{T,e} < 0 \), we must have \( D_{U_1,e} < 0 \) (\( \forall l \in E_{c,i} \)) and, thus \( e \in S_{U_1} \).

\[ |D_{T,e}| = R_{U_1} |D_{U_1,e}| \leq R_{U_1} f_{U_1,0} < R_{U_1} f_{U_1,1} \quad (31) \]

where \( |D_{U_1,e}| \leq f_{U_1,0} \) is because of the definition of \( f_{U_1,0} \) and that \( e \in S_{U_1} \) and \( f_{U_1,0} < f_{U_1,1} \) by condition 2) in this corollary. If \( U_1, U_2 \in T \), then \( e \) must be in one and only one of \( S_{U_1} \) and \( S_{U_2} \), as \( D_{U_1,e} > 0 \) for all \( l \in E_{U_1} \cap E_{c} \) and \( D_{U_2,e} > 0 \) for all \( l \in E_{U_2} \cap E_{c} \). Suppose that \( e \in S_{U_1} \). By (28b), \( D_{T,e} = R_{U_1} D_{U_1,e} + R_{U_2} D_{U_2,e} \text{ if } D_{U_2,e} < 0, \forall l \in E_{c,i}, \) and \( D_{T,e} = -R_{U_1} D_{U_1,e} - R_{U_2} D_{U_2,e} \text{ if } D_{U_2,e} > 0, \forall l \in E_{c,i} \).

\[ |D_{T,e}| = R_{U_1} |D_{U_1,e}| \leq R_{U_1} f_{U_1,0} < R_{U_1} f_{U_1,1} \quad (31) \]

See (25) for the definitions of notations.
\[ R_{U_1} \leq R_{U_2}, \] we will have \( \tilde{D}_{T,e} \geq 0 \) since \( \tilde{D}_{U_1,e} = -\tilde{D}_{U_2,e} \), which contradicts with \( e \in S_T \). Thus, \( R_{U_1} > R_{U_2} \), and hence

\[
|\tilde{D}_{T,e}| = (R_{U_1} - R_{U_2}) \cdot |\tilde{D}_{U_1,e}|
\leq R_{U_1} \tilde{D}_{U_1,e} < R_{U_1} f_{U_1,1},
\]

(32)

where the last inequality holds for the same reason as (31).

Then, it suffices to prove that \( \max_{U \in T} \{ R_U f_{U,1} \} \leq f_{T,1} \). To see this, note that \( \forall U \in T, R_U f_{U,1} = \max_{U \in T} f_{U,1} \).

Thus,

\[
f_{T,1} = \min_{\{ l \in E \}} \sum_{U \in T} I_{E_U}(l) R_U \tilde{D}_{U,l} \geq \min_{\{ l \in E \}} \sum_{U \in T} I_{E_U}(l) H_U f_{U,1}
\]

\[
\geq \min_{\{ l \in E \}} \sum_{U \in T} I_{E_U}(l) \max_{U \in T} f_{U,1} \geq \max_{U \in T} f_{U,1}.
\]

(33)

Combined with (31) and (32), this leads to \( |\tilde{D}_{T,e}| < f_{T,1} \) for any \( e \in S_T \). Therefore, we have \( f_{T,0} < f_{T,1} \).

Finally, due to assumption 3) of this corollary, we have \( f_{T,0} = 0, \forall U \in T \), which leads to \( f_{T,0} = 0 \). Thus, due to \( f_{T,1} > f_{T,0} \), (29) holds for all \( e \in E_c \) and (30) holds for all \( e^* \in E_H \setminus E_c \), which proves the corollary by Theorem V.3.

**Example 3.** We take the system shown in Fig. 2 as an example, where we assume all the nodes in \( H \) to be load buses. The failure of \( \{1_2,3\} \) partitions \( H \) into three islands, where \( V_1 = \{ u_4 \} \), \( V_2 = \{ u_1, u_3, u_6, u_7 \} \), and \( V_3 = \{ u_4, u_5 \} \). We provide several ways of constructing \( L \) and \( T \).

If we have \( \Delta_1^* = \Delta_2^* = 0 \) (satisfying condition 3 in Corollary V.2 for \( H_2 \)), and \( |\tilde{D}_{u_1,t_1}| < |\tilde{D}_{u_2,t_2}|, |\tilde{D}_{u_3,t_3}| < |\tilde{D}_{u_4,t_4}|, |\tilde{D}_{u_1,t_1}| < |\tilde{D}_{u_4,t_4}| \) (satisfying condition 2 in Corollary V.2 for \( H_2 \)), then we can construct \( L = \{ H_2 \} \) that satisfies all the conditions in Corollary V.2. Accordingly, \( T = \{ \{ u_1 \}, \{ u_2 \} \} \) satisfies all the conditions in Theorem V.3. Thus, \( F = E_c \).

Alternatively, if \( p_{u_4}^* = p_{u_8}^* = 0 \) (satisfying condition 3 in Corollary V.2 for \( H_1 \) and \( H_3 \)), we can construct \( L = \{ H_1, H_3 \} \) and \( T = \{ \{ u_4 \}, \{ u_8 \} \} \), where \( S_T = \emptyset \) and \( f_{T,g} = -p_{u_4}^* - p_{u_8}^* = 0 \). Then, it is easy to check that all the conditions of Corollary V.2 and Theorem V.3 are satisfied, which guarantees \( F = E_c \).

Remark: Corollary V.2 extends Theorem V.1 in [1], in the sense that Corollary V.2 only requires hypothetical power flows on failed links to be in the same direction, while Theorem V.1 of [1] requires these power flows to have specific directions. Specifically, if all the nodes in \( V_{L,e} \) are load buses and \( D_{u,e} > 0, \forall u \in V_{L,e}, e \in E_c \), then condition 3) of Corollary V.2 requires \( p_u^* = 0, \forall u \in V_{L,e} \), which is exactly the same as the conditions of Theorem V.1 in [1]. However, Corollary V.2 can additionally cover the cases that \( \tilde{D}_{u,e} < 0 \) and \( \Delta_u^* = 0, \forall u \in V_{L,e}, e \in E_c \).

**VI. PERFORMANCE EVALUATION**

We test our solutions on the Polish power grid (“Polish system - winter 1999-2000 peak”) [26] with 2383 nodes and 2886 links, where parallel links are combined into one link. We generate the attacked area \( H \) by randomly choosing one node as a starting point and performing a breadth first search to obtain \( H \) with a predetermined \( |V_H| \). We then randomly choose \( |F| \) links within \( H \) to fail. The phase angles of each island without any generator or load are set to 0, and the rest are computed according to (1). We vary \( |V_H| \) and \( |F| \) to explore different settings, and for each setting, we generate 70 different \( H \)’s and 300 different \( F \)’s per \( H \).

We evaluate two types of metrics: (1) how often the recovery conditions are satisfied, and (2) how accurate our algorithm is when its recovery conditions are not necessarily satisfied.

**A. Probability of Guaranteed Recovery**

First, we evaluate the fraction of randomly generated cases (combinations of \( H \) and \( F \)) satisfying the conditions in Theorem III.1 for recovering the phase angles, and Theorem IV.1(1) for localizing the failed links with known phase angles and active powers. We observe that: (i) the condition in Theorem III.1 is almost never satisfied under a nontrivial size of \( H \) (\( |V_H| = 20, \ldots, 40 \)), which emphasizes the importance of securing PMU measurements; (ii) the condition in Theorem IV.1(1) is only satisfied with a small probability as shown Fig. 4, which decreases with the size of \( H \) (note that Theorem IV.1(1) does not depend on \( F \)). In previous work [3], this issue was addressed by actively designing “zones” for reporting measurements such that each zone satisfies these stringent conditions, which only works for attacks limited to single zones. In contrast, our recovery conditions are much more general as shown below (Fig. 5 and Fig. 6), even though we do not assume to know the post-attack power injections in \( H \).

Then, we evaluate the fraction of links satisfying the recovery conditions in Lemma V.2 and Theorems V.1–V.3, together with the actual fraction of links whose status are correctly identified by Algorithm 1 under known phase angles and unknown active powers. The results are shown in Fig. 5 and Fig. 6, where “Experimental Results” is the actual performance of Algorithm 1 “Lemma V.2” indicates the failed/operational links satisfying Lemma V.2 and “Theorems V.1–V.3” indicates the failed links that satisfy either Theorem V.1 or (29) in

\[ |V_{H}| = 40. \]
Theorem V.3 or the operational links that satisfy either Theorem V.2 or (30) in Theorem V.3. More specifically, Figs. 5 (a) shows the fraction of correctly identified failed/operational links, averaged over all the randomly generated cases, where the bottom and the top edges of the error bar indicate the 25th and 75th percentiles. In addition, Figs. 5 (b) shows the fraction of cases with no miss/false alarm.

It is worth noting that checking whether Theorems V.1, V.2 hold for each link is hard since it requires testing all possible hyper-nodes, whose number is exponential in |VH|. To test whether Theorem V.1 holds for a given link, we heuristically construct a hyper-node through BFS starting from each of its endpoints. In each iteration of BFS, we add all the nodes that cause violation of condition 1 or condition 2 in Theorem V.1 into the hyper-node and test whether condition 3 holds. This procedure can be applied similarly to the testing of Theorem V.2 and the construction of T in Theorem V.3. Likewise, checking whether Lemma V.2 holds for each link is hard since it requires testing all possible Qm and Qf. To test whether a failed link e ∈ F satisfies Lemma V.2, we construct W according to Qm = {e}, Qf = ∅ and test if (22) has a non-negative solution. This procedure can be applied similarly to an operational link l ∈ EH \ F. Therefore, the fractions of links/cases covered by these theoretical conditions in Figs. 6 (a) are bounds (upper-bound for Lemma V.2 and lower-bound for Theorems V.1, V.3). Nevertheless, we see that (i) Theorems V.1, V.3 explain the success of Algorithm 1 quite well, especially when |F| is small, (ii) the theorems can explain most cases covered by Lemma V.2 and (iii) Algorithm 1 is actually highly accurate in identifying the operational links even though the theoretical conditions for doing so are appear stringent.

To better understand the relationship between Theorems V.1, V.3, we decompose all the failed links into 4 categories: (1) links satisfying both Theorem V.1 and (29) in Theorem V.3 (2) links satisfying only (29) in Theorem V.3 (3) links satisfying only Theorem V.1 and (4) links satisfying neither. Fig. 7 (a) shows the fraction of links in each category, averaged over all the simulated cases, where the bottom and the top edges of the error bar indicate the 25th and 75th percentiles. We observe that (i) many failed links satisfy both conditions; (ii) with growing |F|, the fraction of failed links satisfying only (29) in Theorem V.3 decreases, while the fraction of links satisfying only Theorem V.1 increases; (iii) the fraction of links covered by neither of the conditions remains small as |F| grows. Similarly, we look into the coverage of Theorems V.2 and (30) in Theorem V.3 for operational links, as shown in Fig. 7 (b). Similar phenomena are observed, except that the fraction of links covered by neither of these conditions increases with growing |F|, which is consistent with Fig. 6.

B. Accuracy of Failure Localization

Next, we focus on the accuracy of Algorithm 1 in comparison with benchmarks in localizing the failed links, assuming that post-attack phase angles are known. This assumption is justified by the better protection of PMU measurements from cyber attacks [21]. We consider two benchmarks: (i) the solution given in Theorem V.1 (extended from [3]), i.e., estimating F by \( \text{supp}(x) \) for the solution to \( \min_{x} ||x||_1 \text{ s.t. } (10) \), assuming the true \( \Delta_H \) to be known, and (ii) \( \min_{x} ||x||_1 \text{ s.t. } ||B_{H|G}(\theta - \hat{\theta}) - D_Hx||_2 \leq ||p_H||_2 \), which is extended from the solutions in [15], [16]. We note that the original solution in [3] (which assumes \( \Delta = 0 \)) is often infeasible for our problem, as shown in Fig. 8 and thus not used as a benchmark. Note that benchmark (i) is meant to be a “performance upper bound”, as it assumes more knowledge (i.e., ground-truth \( \Delta_H \)) than our proposed algorithm.

As shown in Fig. 9 benchmark (i) demonstrates the best performance with regard to both miss-detection rate and the probability of having no miss-detection, while Algorithm 1 performs much better than benchmark (ii). This confirms the importance of knowing or estimating load shedding values in failure localization. Regarding the false alarm as shown in Fig. 10 Algorithm 1 performs even better than benchmark (i). This is because the decision variable \( x \) in benchmark (i) combines the information of both the failed links and the phase angles \( \theta_H \), and thus does not fully exploit the knowledge of \( \theta_H \). Furthermore, from the specific number of false alarms/misses in Fig. 11 we see that Algorithm 1 correctly detects all the failed links with almost no false alarm majority of the time, while only missing a couple of failed links for the rest of the time.

Summary: The above results imply that while it is difficult to recover the phase angles from inference, if we can protect phase
Probability of no miss
Probability of no false alarm
Figure 9. Performance comparison on miss rate ($|V_H| = 40$).

Figure 10. Performance comparison on false alarm rate ($|V_H| = 40$).

Meanwhile, the existing condition for perfect phase angle measurements from the cyber attack, then our algorithm can be used to infer the link status in the attacked area with high accuracy, even under unknown changes in power injections.

VII. CONCLUSION

We investigated the problem of power grid state estimation under general cyber-physical attack that may disconnect the grid. First, we demonstrated that existing solutions and the corresponding recovery conditions for recovering phase angles and link status are still applicable with the knowledge of post-attack power injections. Second, for unknown post-attack power injections, we proposed an LP-based algorithm to identify link status within the attacked area with knowledge of recovered phase angles. We established sufficient conditions for the proposed algorithm to identify the link status correctly. Finally, our evaluations based on the Polish power grid demonstrated that the proposed algorithm is highly accurate in localizing the failed links, and the majority of cases are theoretically guaranteed by the presented conditions. Meanwhile, the existing condition for perfect phase angle recovery is hard to satisfy in practice, indicating the importance of safeguarding PMU measurements from cyber attack.

Figure 11. Number of false alarms/misses of Algorithm 1 ($|V_H| = 40$).

REFERENCES

[1] Y. Huang, T. He, N. R. Chaudhuri, and T. L. Porta, “Power grid state estimation under general cyber-physical attacks,” in IEEE SmartGridComm, 2020.

[2] P. Fairley, “Cybersecurity at U.S. utilities due for an upgrade: Tech to detect intrusions into industrial control systems will be mandatory,” IEEE Spectrum, vol. 53, no. 5, pp. 11–13, May 2016.

[3] S. Soltan, M. Yannakakis, and G. Zussman, “Power grid state estimation following a joint cyber and physical attack,” IEEE Transactions on Control of Network Systems, vol. 5, no. 1, p. 499–512, March 2018.

[4] S. Soltan and G. Zussman, “Power grid state estimation after a cyber-physical attack under the AC power flow model,” in IEEE PES-GM, 2017.

[5] Y.-F. Huang, S. Werner, J. Huang, N. Kashyap, and V. Gupta, “State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid,” IEEE Signal Processing Magazine, vol. 29, no. 5, pp. 33–43, 2012.

[6] Y. Liu, P. Ning, and M. K. Reiter, “False data injection attacks against state estimation in electric power grids,” ACM Transactions on Information and System Security (TISSEC), vol. 14, no. 1, pp. 1–33, 2011.

[7] Y. Shoukry, P. Nuzzo, A. Puggelli, A. L. Sangiovanni-Vincentelli, S. A. Seshia, and P. Tabuada, “Secure state estimation for cyber-physical systems under sensor attacks: A satisfiability modulo theory approach,” IEEE Transactions on Automatic Control, vol. 62, no. 10, pp. 4917–4932, 2017.

[8] G. Dan and H. Sandberg, “Stealth attacks and protection schemes for state estimators in power systems,” in 2010 first IEEE international conference on smart grid communications. IEEE, 2010, pp. 214–219.

[9] O. Vuković, K. C. Sou, G. Dan, and H. Sandberg, “Network-layer protection schemes against stealth attacks on state estimators in power systems,” in 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm). IEEE, 2011, pp. 184–189.

[10] J. Kim and L. Tong, “On topology attack of a smart grid: Undetectable attacks and countermeasures,” IEEE Journal on Selected Areas in Communications, vol. 31, no. 7, pp. 1294–1305, 2013.

[11] R. Deng, P. Zhuang, and H. Liang, “Cepa: Coordinated cyber-physical attacks and countermeasures in smart grid,” IEEE Transactions on Smart Grid, vol. 8, no. 5, pp. 2420–2430, 2017.

[12] S. Soltan, M. Yannakakis, and G. Zussman, “React to cyber attacks on power grids,” IEEE Transactions on Network Science and Engineering, vol. 6, no. 3, pp. 459–473, 2018.

[13] 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, 2008.

[14] ——, “Double line outage detection using phasor angle measurements,” in 2009 IEEE Power & Energy Society General Meeting. IEEE, 2009, pp. 1–5.

[15] 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, November 2012.

[16] J.-C. Chen, W.-T. Li, C.-K. Wen, J.-H. Teng, and P. Ting, “Efficient identification method for power line outages in the smart power grid,” IEEE Transactions on Power Systems, vol. 29, no. 4, pp. 1788–1800, 2014.

[17] W. Stevenson and J. Grainger, Power System Analysis. McGraw-Hill Education, 1994. [Online]. Available: https://books.google.com/books?id=NBlaoQAAMAAJ

[18] B. Scott, J. Jardim, and O. Alsac, “DC power flow revisited,” IEEE Transactions on Power Systems, vol. 24, no. 3, pp. 1290–1300, 2009.

[19] M. Garcia, T. Catanach, S. Vander Wiel, R. Bent, and E. Lawrence, “Line outage localization using phasor measurement data in transient state,” IEEE Transactions on Power Systems, vol. 31, no. 4, pp. 3019–3027, 2016.

[20] “WASA and the roadmap to WAMPAC at SDG&E,” PAC World magazine, September 2019.

[21] “Wide area monitoring, protection, and control systems (WAMPAC) standards for cyber security requirements,” National Electric Sector Cybersecurity Organization Resource (NESCOR), October 2012, https://smartgrid.cpi.org/doc/ESRFSD.pdf.

[22] B. Pal and B. Chaudhuri, Robust control in power systems. Springer Science & Business Media, 2006.

[23] M. Lu, W. ZainalAbidin, T. Masri, D. Lee, and S. Chen, “Under-frequency load shedding (ufsl) schemes-a survey,” International Journal of Applied Engineering Research, vol. 11, no. 1, pp. 456–472, 2016.
[24] S. Dasgupta, C. H. Papadimitriou, and U. V. Vazirani, *Algorithms*. McGraw-Hill Higher Education New York, 2008.
[25] O. L. Mangasarian, *Nonlinear programming*. SIAM, 1994.
[26] R. D. Zimmerman and C. E. Murillo-Sánchez, “Matpower 7.0 user’s manual,” *Power Systems Engineering Research Center*, vol. 9, 2019.