A MILP for Optimal Measurement Choice in Robust Power Grid State Estimation

Elizabeth Glista and Somayeh Sojoudi
University of California, Berkeley

***Abstract*— The reliability of the electric power grid is increasingly linked to the reliability of measured data which is used to understand the current state of the system. Determining the current state of the electric grid is the basis for decision-making related to the normal operation of the grid as well as operations in the case of an emergency scenario. When some of this data is corrupted in the case of a cyberattack, it is important that we can recover the true state of the system via state estimation (SE). Inspired by the work in [1] and [2], we propose a novel method using a notion in machine learning to optimize the choice of measurements in a given power network, formulating the problem as a mixed-integer linear program (MILP). Using this MILP, we study some test cases and show that it is impossible to certify that the network is fully robust in the case of bad data. However, we propose a method to optimally place the sensors in order to make the network more robust in the case of cyberattacks.**

I. Introduction

Power system state estimation (PSSE) is a critical problem for the reliability of the electric grid. PSSE uses data from sensors throughout a transmission or distribution network to monitor the state of the network [3]. The estimated state is in turn used to make decisions about real-time power dispatch, implement voltage control, and take action in the case of a contingency, such as a line or generator outage [4]. During the Northeast power blackout of 2003, which affected over 50 million people in the U.S. and Canada, the propagation of cascading failures could have been mitigated had the operators been able to recover the true state of the network [5]. Because sensor measurements may be subject to both random noise and intentional cyberattacks, it is important to consider a robust version of the SE problem [6]. Furthermore, as cyberattacks increase in frequency, robust PSSE will become more important in the design of algorithms for the future smart grid [7]–[9].

A special case of graph-structured quadratic sensing, PSSE is formulated as the minimization of a loss function representing the difference between the actual set of measurements and the measurements that would be observed for the estimated state. The state of a power network is defined by a complex voltage at each bus in the network. Due to the nonlinearity of alternating-current (AC) power flow, the classical PSSE problem is nonlinear, making the problem NP-hard. In practice, nonlinear SE is solved with local search algorithms such as Newton's method [10]. However, local methods may yield spurious local minima with no physical meaning since PSSE does not satisfy the restricted isometry property (RIP) from quadratic sensing that can be used to certify a lack of spurious local minima [11]. Because of this, there is growing interest in methods that can yield global solutions to the PSSE problem such as stochastic and convex methods [12]–[15]. The paper [1] proposes a two-step PSSE method which allows for the recovery of the true state of the system in the case without noise or bad data. Because this method involves solving a linear SE problem, it is convex and can be solved to global optimality efficiently with existing local search methods. Additionally, [1] introduces a sufficient condition to verify the robustness of PSSE that explicitly depends on the support of the bad data, and [2] extends this work to propose a method which certifies that a network is locally robust to bad data without any dependence on the bad data support.

A. Contributions

By leveraging the results of [2], this work proposes a novel MILP to optimize the placement of sensors in a network in order to satisfy a machine learning condition for PSSE robustness.

B. Notations

The symbol $\mathbb{R}$ denotes the set of real numbers, and $\mathbb{R}^N$ denotes the space of N-dimensional real vectors. The symbol $(\cdot)^T$ denotes the transpose of a vector or matrix. The symbol $|\cdot|$ is the absolute value operator if the argument is a scalar, vector, or matrix; otherwise, it is the cardinality of a measurable set. The imaginary unit is denoted by $\mathbf{j} = \sqrt{-1}$. The elementwise multiplication of two matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times n}$ is denoted as $A \odot B$. The symbol $\dagger$ denotes the left pseudoinverse of a matrix given as $A^\dagger \triangleq (A^TA)^{-1}A^T$. The notation $||A||_\infty$ corresponds to the matrix infinity norm, e.g. the maximum absolute column sum of matrix A. The expression $\mathbf{1}_n$ is a vector of ones of dimension n, and the expression $\mathbf{1}\{\zeta\}$ is the indicator function which is 1 if ζ is true and 0 otherwise. The notation $A[\mathcal{B},\mathcal{C}]$ or $A_{\mathcal{B},\mathcal{C}}$ represents a submatrix of matrix A formed by taking the rows and columns corresponding respectively to the sets $\mathcal{B}$ and $\mathcal{C}$. The notation $\mathcal{A}\setminus\mathcal{B}$ denotes the subtraction of set $\mathcal{B}$ from set $\mathcal{A}$, and $\mathcal{A}\cup\mathcal{B}$ denotes the union of sets $\mathcal{A}$ and $\mathcal{B}$. The notation $[n]$ denotes the index set $\{1,\dots,n\}$.

Elizabeth Glista and Somayeh Sojoudi are with the Department of Mechanical Engineering at the University of California, Berkeley. Emails: {glista, sojoudi}@berkeley.edu.
This work was supported by grants from NSF, AFOSR and ONR.

II. Background

A. Power System State Estimation (PSSE)

Let a power network be defined as the graph $\mathcal{G} = (\mathcal{N}, \mathcal{L})$, where $\mathcal{N}$ is the set of buses and $\mathcal{L}$ is the set of lines. The goal of PSSE is to recover the true state of the network, given as the complex voltage $v_i \triangleq |v_i|e^{\mathbf{j}\theta_i}$ at each bus $i \in \mathcal{N}$. We are given some set of measurements $\mathcal{M}$, which can include measurements of the real or reactive power flows p_{ij}, q_{ij} on line $(i, j) \in \mathcal{L}$, the real or reactive power injected p_i, q_i at bus $i \in \mathcal{N}$, or the voltage magnitude $|v_i|$ at bus $i \in \mathcal{N}$. We can also extend this method to include phase angle measurements θ_i for $i \in \mathcal{N}$ from phasor measurement units (PMUs). We will use the PSSE method from [1], which introduces a linear basis using the unknown state variables $x_i^{\text{mg}} \triangleq |v_i|^2$ for all $i \in \mathcal{N}$, $x_{ij}^{\text{re}} \triangleq |v_i||v_j|\cos(\theta_{ij})$ for all $(i, j) \in \mathcal{L}$, $x_{ij}^{\text{im}} \triangleq |v_i||v_j|\sin(\theta_{ij})$ for all $(i, j) \in \mathcal{L}$, where $\theta_{ij} \triangleq \theta_i - \theta_j$ for all $(i, j) \in \mathcal{L}$. We will take the set $\mathcal{X} = \left\{\{x_i^{\text{mg}}\}_{\forall \in \mathcal{N}}, \{x_{ij}^{\text{re}}\}_{\forall (i,j) \in \mathcal{L}}, \{x_{ij}^{\text{im}}\}_{\forall (i,j) \in \mathcal{L}}\right\}$ to be the set of all states for the network, which is fixed given the network topology.

Given this linear basis, the equations which relate the measurements to the state can be formulated as $\mathbf{m} = A\mathbf{x}$, where $A \in \mathbb{R}^{m \times n}$ is the sensing matrix that relates the unknown state $\mathbf{x} \in \mathbb{R}^n$ to the vector of measurements $\mathbf{m} \in \mathbb{R}^m$. We have that $n \triangleq |\mathcal{X}| = |\mathcal{N}| + 2|\mathcal{L}|$ and $m \triangleq |\mathcal{M}|$. Note that A is sparse due to the sparse nature of power networks (see [1] for the formulation of A). When $m > n$, the equation $\mathbf{m} = A\mathbf{x}$ represents an over-determined power flow problem. We will assume that we always have $m \geq n$. In a realistic scenario, the measurements $\mathbf{m}$ are corrupted with random noise and potentially other bad data, and therefore we cannot just solve this over-determined power flow to determine the true state. We can model the noisy and/or corrupted measurements $\mathbf{y} \in \mathbb{R}^m$ as:

$$\mathbf{y} = A\mathbf{x} + \mathbf{w} + \mathbf{b} \tag{1}$$

where $\mathbf{w} \in \mathbb{R}^m$ represents random noise and $\mathbf{b} \in \mathbb{R}^m$ represents the bad data vector. Typical assumptions on these vectors are that $\mathbf{w}$ follows a Gaussian distribution and that $\mathbf{b}$ is a sparse vector [16]. Note that the local recovery method in [2] is one of the most general methods as it does not make assumptions on the sparsity of $\mathbf{b}$.

The PSSE methods of [1] and [2] use a two-step process:

1) Solve SE problem defined by (1) to get an estimate $\hat{\mathbf{x}}$.
2) Recover an estimate of the complex voltages using the relations $|\hat{v}_i| = \sqrt{\hat{x}_i^{\text{mg}}}$, $\hat{\theta}_{ij} = \arctan\left(\hat{x}_{ij}^{\text{im}}/\hat{x}_{ij}^{\text{re}}\right)$, and $\hat{\theta} = \arg\min_{\theta \in \mathbb{R}^{|\mathcal{N}|}} \sum_{(i,j)\in\mathcal{L}} (\theta_i - \theta_j - \hat{\theta}_{ij})^2$.

If step 1 is able to recover the true state, then step 2 will recover the true complex voltage vector [1]. In the case of corrupted and/or noisy data, it will be impossible to recover the true state in step 1, but it is stated in [1] that the propagation of error is not too great in step 2. Thus, the focus of this paper for robust SE is on step 1, which we will call ℓ-PSSE (linearized PSSE) from this point forward.

In the case when both random Gaussian noise and sparse corruption are present, one version of ℓ-PSSE problem would be to solve the LASSO problem given in [1]:

$$\hat{x} = \min_{\mathbf{x}, \mathbf{b}} \frac{1}{2|\mathcal{M}|} ||\mathbf{y} - A\mathbf{x} - \mathbf{b}||_2^2 + \lambda ||\mathbf{b}||_1 \tag{2}$$

for some regularization parameter $\lambda > 0$ that promotes the sparsity of $\mathbf{b}$. As an alternative, the paper [2] proposes minimizing a Huber loss which is more robust to outliers.

B. Mutual Incoherence

Mutual coherence is a measure of the cross-correlation of the columns of a matrix $A \in \mathbb{R}^{m \times n}$, which is a powerful notion in the area of compressed sensing. The authors of [1] propose a new metric, which they call "mutual incoherence," a measure of the alignment of two particular submatrices of the sensing matrix A, one related to the clean data and one related to the corrupted data. As it is proposed in [1], this metric relies on the knowledge of the support of the bad data vector $\mathbf{b}$, denoted as $\mathcal{B} \subset \mathcal{M}$. The mutual incoherence metric $\rho(\mathcal{B})$ is then defined as $\rho(\mathcal{B}) = \left|\left|A_{\mathcal{B}^c}^{T\dagger} A_{\mathcal{B}}^T\right|\right|_\infty$, where $\mathcal{B}^c \triangleq \mathcal{M} \backslash \mathcal{B}$, $A_{\mathcal{B}}$ is the submatrix of A with rows corresponding to $\mathcal{B}$, and $A_{\mathcal{B}^c}$ is the submatrix of A with rows corresponding to $\mathcal{B}^c$. We need to make a few assumptions about the matrix A in order to use the mutual incoherence metric to certify the robustness of the ℓ-PSSE problem.

Assumption 1 (Preconditioning of sensing matrix)**.** *Each row of A is normalized so that $||\boldsymbol{a}_i||_2 = 1, \ \forall i \in [m]$, where $\boldsymbol{a}_i$ is the i^{th} row of A.*

Assumption 2 (Lower eigenvalue condition)**.**

$$\min\left\{\lambda_{\min}\left(A_{\mathcal{B}^c}^T A_{\mathcal{B}^c}\right), \lambda_{\min}\left(\begin{bmatrix} A \\ I_{\mathcal{B}} \end{bmatrix} \begin{bmatrix} A^T & I_{\mathcal{B}}^T \end{bmatrix}\right)\right\} > 0 \tag{3}$$

where $I_{\mathcal{B}}$ corresponds to a submatrix formed by the $\mathcal{B}$ rows of the identity matrix $I \in \mathbb{R}^{m \times n}$ and $\lambda_{\min}(\cdot)$ denotes the minimum eigenvalue of a matrix.

This second assumption implies that the true vector must be identifiable if the bad data support $\mathcal{B}$ were known. The authors of [1] show that under these assumptions on A, if $\rho(\mathcal{B}) < 1$, then problem (2) with a given choice of regularization parameter λ recovers an estimated state with a small error from the true state as well as a large degree of bad data detection with high probability. However, because this method relies on knowledge of the support of the bad data vector, its application is limited.

The paper [2] builds on [1] and proposes a way to avoid using the bad data support, by developing a method for certification which can be ensured locally for each line in the network $(i, j) \in \mathcal{L}$ without considering the actual attack set. This method partitions the graph into attack, boundary, and safe regions for a given line $(i, j) \in \mathcal{L}$ and then looks at the mutual incoherence metric defined on subsections of the partitioned boundary measurements, which are fixed for a given line $(i, j) \in \mathcal{L}$ and measurement set $\mathcal{M}$. During an actual attack, if measurements at a node i are attacked and if every line $(i, j) \in \mathcal{L}$ attached to node i satisfies the mutual incoherence condition, then the attack will not propagate through the network.

Algorithm 1 Sensing matrix partition for local attack $i \to j$

Inputs: $\mathcal{G}$, $\mathcal{M}$, $\mathcal{X}$, (i,j)
Compute sensing matrix A from $\mathcal{G}$
Set $\mathcal{X}_{\text{a}}^{ij} \leftarrow \{x_i^{\text{mg}}, x_{ij}^{\text{re}}, x_{ij}^{\text{im}}\}$
Set $\mathcal{M}_{\text{a}}^{ij} \leftarrow \{\text{all-zero rows of } A[:\ ,(\mathcal{X} \setminus \mathcal{X}_{\text{a}}^{ij})]\}$
Set $\mathcal{M}_{\text{db}}^{ij} \leftarrow \{\text{non-zero rows of} A[:\ ,\mathcal{X}_{\text{a}}^{ij}]\} \setminus \mathcal{M}_{\text{a}}^{ij}$
Set $\mathcal{X}_{\text{b}}^{ij} \leftarrow \{\text{non-zero columns of } A[\mathcal{M}_{\text{db}}^{ij}\ ,:]\} \setminus \mathcal{X}_{\text{a}}$
Set $\mathcal{M}_{\text{ib}}^{ij} \leftarrow \{\text{non-zero rows of } A[:\ ,\mathcal{X}_{\text{b}}^{ij}]\} \setminus \mathcal{M}_{\text{db}}^{ij}$
Set $\mathcal{M}_{\text{s}}^{ij} \leftarrow \mathcal{M} \setminus (\mathcal{M}_{\text{a}}^{ij} \cup \mathcal{M}_{\text{db}}^{ij} \cup \mathcal{M}_{\text{ib}}^{ij})$
Set $\mathcal{X}_{\text{s}}^{ij} \leftarrow \mathcal{X} \setminus (\mathcal{X}_{\text{a}}^{ij} \cup \mathcal{X}_{\text{b}}^{ij})$
Outputs: $\{\mathcal{X}_{\text{a}}^{ij}, \mathcal{X}_{\text{b}}^{ij}, \mathcal{X}_{\text{s}}^{ij}\}$, $\{\mathcal{M}_{\text{a}}^{ij}, \mathcal{M}_{\text{db}}^{ij}, \mathcal{M}_{\text{ib}}^{ij}, \mathcal{M}_{\text{s}}^{ij}\}$

In the next section, we present a modified version of the graph partitioning that was first introduced in [2]. While [2] partitions based on k^{th} connected neighbors in the network, this method partitions through variable coupling in the sensing matrix and thus takes into account the choice of measurements to determine the variable partition. Unlike that in [2], our method results in the minimum number of boundary variables and maximum number of safe variables and measurements. This version is effectively the same as that in [2], i.e. it does not change the mutual incoherence metric or results of [2], but it streamlines the partitioning process and results in a more intuitive partition for the application.

III. Graph partitioning for local certification

For a given line of attack $i \to j$, we aim to partition the set of state variables $\mathcal{X}$ into the sets of attacked variables $\mathcal{X}_{\text{a}}^{ij}$, boundary variables $\mathcal{X}_{\text{b}}^{ij}$, and safe variables $\mathcal{X}_{\text{s}}^{ij}$, where we use the superscript ij to indicate that the partition is specific to the chosen attack line $i \to j$. It is desirable to partition the measurement sets into the attacked measurements $\mathcal{M}_{\text{a}}^{ij}$ that depend only on $\mathcal{X}_{\text{a}}^{ij}$, the dependent boundary measurements that depend on both $\mathcal{X}_{\text{a}}^{ij}$ and $\mathcal{X}_{\text{b}}^{ij}$, the independent boundary measurements $\mathcal{M}_{\text{ib}}^{ij}$ that depend only on $\mathcal{X}_{\text{b}}^{ij}$, and the remaining safe measurements $\mathcal{M}_{\text{s}}^{ij}$ that can depend on both $\mathcal{X}_{\text{s}}^{ij}$ and $\mathcal{X}_{\text{b}}^{ij}$. We note that the "independent" and "dependent" boundary measurements are defined as dependent in relation to the attacked variables $\mathcal{X}_{\text{a}}^{ij}$. The algorithm to formulate the variable and measurement partitions is given in Algorithm 1. With this partition, we can rewrite the sensing matrix A as coupled through the boundary region.

If the matrix A satisfies some mutual incoherence condition for independent and dependent boundary measurement sets given by the partition in Algorithm 1, then line $i \to j$ is robust and bad data cannot propagate from i to j. In this case, if node i is part of the unknown attack set, then it will still be possible to recover a reasonable estimate of the state at node j with high probability. The required local mutual incoherence condition is given as:

$$\rho^{ij} \triangleq \left|\left|A^{T\dagger}_{\mathcal{M}_{\text{ib}}^{ij},\mathcal{X}_b^{ij}} A^{T}_{\mathcal{M}_{\text{db}}^{ij},\mathcal{X}_b^{ij}}\right|\right|_\infty < 1 \tag{4}$$

We can see that condition (4) depends on the measurement-variable partition. In this case, the mutual incoherence ρ^{ij} captures the alignment between measurements in the independent boundary set and the dependent boundary set. This condition ensures that attacked measurements do not propagate from the dependent boundary set to the independent boundary set.

Because condition (4) depends on the measurement set, it is apparent that we can optimize the choice of measurements $\mathcal{M}$ in order to decrease ρ^{ij} with the goal of finding measurements such that $\rho^{ij} < 1$. If we can find a measurement set $\mathcal{M}$ such that $\rho^{ij} < 1$ for all $i \to j$ and $j \to i$ for $(i,j) \in \mathcal{L}$, then we can say that the network is fully robust. If the network is fully robust, then we can find good estimates for local recovery of the safe and boundary region state variables via the method in [2]. In order to formalize the goal of placing sensors in a power network so that the network is robust, we will consider this mutual incoherence condition in an optimization framework, as presented in the next section.

IV. Problem Formulation

The goal is to find a minimum choice of measurements over the network such that the mutual incoherence condition is satisfied for all boundary measurement sets $\{\mathcal{M}_{\text{db}}^{ij}, \mathcal{M}_{\text{ib}}^{ij}\}$ in both $i \to j$ and $j \to i$ directions for every line $(i,j) \in \mathcal{L}$. Note that in the formulations below we will use the notation $(i,j) \in \mathcal{L}$ to denote lines in both $i \to j$ and $j \to i$ directions. Let $\phi \in \{0,1\}^m$ be a binary vector which indicates the choice of measurements such that $\phi_i = 1$ if measurement $i \in [m]$ is chosen and $\phi_i = 0$ otherwise. Note that m is equal to the total possible number of measurements for a given power network.

When we consider the mutual incoherence condition across a line $i \to j$, we can define a partition of all possible measurements $\tilde{\mathcal{M}}$, which is invariant to the choice of measurements ϕ and depends only on the graph topology. Given this partition, let $\tilde{\mathcal{M}}_{\text{db}}^{ij}$ be the set of total possible dependent boundary measurements and $\tilde{\mathcal{M}}_{\text{ib}}^{ij}$ be the set of total possible independent boundary measurements. Let $m_{\text{db}}^{ij} \triangleq |\tilde{\mathcal{M}}_{\text{db}}^{ij}|$, $m_{\text{ib}}^{ij} \triangleq |\tilde{\mathcal{M}}_{\text{ib}}^{ij}|$, and $n_{\text{b}}^{ij} \triangleq |\mathcal{X}_{\text{b}}^{ij}|$. We could formulate an optimization problem with condition (4) as a constraint. However, this problem may be infeasible if the constraints (4) cannot be satisfied for all lines $(i,j) \in \mathcal{L}$. Thus, it is more useful to consider the following mixed-integer nonlinear program (MINLP):

$$\min_{\substack{\beta\in\mathbb{R},\phi\in\{0,1\}^m \\ X^{ij},E^{ij},J^{ij},\forall(i,j)}} \quad \beta \tag{5a}$$

$$\text{s.t.} \quad \underline{M} \leq \sum\nolimits_{i=1}^{m} \phi_i \leq \overline{M} \tag{5b}$$

$$\forall (i,j) \in \mathcal{L}:$$

$$\left(R^{ij} \odot E^{ij}\right) X^{ij} = S^{ij} \odot J^{ij} \tag{5c}$$

$$E_k^{ij} = \phi[\tilde{\mathcal{M}}_{\text{ib}}^{ij}(k)]\, \mathbf{1}_{n_{\text{b}}^{ij}}, \forall k \in [m_{\text{ib}}^{ij}] \tag{5d}$$

$$J_k^{ij} = \phi[\tilde{\mathcal{M}}_{\text{db}}^{ij}(k)]\, \mathbf{1}_{n_{\text{b}}^{ij}}, \forall k \in [m_{\text{db}}^{ij}] \tag{5e}$$

$$||X^{ij}||_\infty \leq \beta \tag{5f}$$

where $R^{ij} \triangleq A^T_{\tilde{\mathcal{M}}_{\text{ib}}^{ij},\mathcal{X}_{\text{b}}^{ij}} \in \mathbb{R}^{n_{\text{b}}^{ij} \times m_{\text{ib}}^{ij}}$ and $S^{ij} \triangleq A^T_{\tilde{\mathcal{M}}_{\text{db}}^{ij},\mathcal{X}_{\text{b}}^{ij}} \in \mathbb{R}^{n_{\text{b}}^{ij} \times m_{\text{db}}^{ij}}$ are subsets of the transposed sensing matrix A^T.

978-1-6654-0823-3/22/$31.00 ©2022 IEEE

We have introduced the variable $X^{ij} \in \mathbb{R}^{m_{\text{ib}}^{ij} \times m_{\text{db}}^{ij}}$ in order to represent the mutual incoherence as $||X^{ij}||_\infty$ for each line $(i,j) \in \mathcal{L}$. The matrix variables $E^{ij} \in \mathbb{R}^{n_{\text{b}}^{ij} \times m_{\text{ib}}^{ij}}$ and $J^{ij} \in \mathbb{R}^{n_{\text{b}}^{ij} \times m_{\text{db}}^{ij}}$ are used to choose columns of the sensing matrix corresponding respectively to independent and dependent boundary measurements. E_k^{ij} and J_k^{ij} represent the k^{th} columns of E^{ij} and J^{ij}, respectively. The notation $\phi[\tilde{\mathcal{M}}_{\text{ib}}^{ij}(k)]$ represents the element of ϕ corresponding to the k^{th} entry of $\tilde{\mathcal{M}}_{\text{ib}}^{ij}$ (similarly for $\phi[\tilde{\mathcal{M}}_{\text{db}}^{ij}(k)]$). The given parameters $\underline{M}$ and $\overline{M}$ are respectively the minimum and maximum numbers of measurements, where we select $\underline{M}$ such that $\underline{M} \geq n$.

Theorem 1. *If the objective of (5) is strictly less than 1, then a measurement set can be found such that the network is robust in terms of the mutual incoherence condition (4).*

Proof: Using equations (5d) and (5e), we have that $R^{ij} \odot E^{ij}$ is equivalent to $A^T_{\mathcal{M}_{\text{ib}}^{ij}, \mathcal{X}_{\text{b}}^{ij}}$ and $S^{ij} \odot J^{ij}$ is equivalent to $A^T_{\mathcal{M}_{\text{db}}^{ij}, \mathcal{X}_{\text{b}}^{ij}}$, thus $X^{ij} = A^{T\dagger}_{\mathcal{M}_{\text{ib}}^{ij}, \mathcal{X}_{\text{b}}^{ij}} A^T_{\mathcal{M}_{\text{db}}^{ij}, \mathcal{X}_{\text{b}}^{ij}}$ by constraint (5c). We have that $||X^{ij}||_\infty$ corresponds to ρ^{ij} as defined in Equation (4), and if $\beta < 1$ then (5f) enforces that ρ^{ij} is under 1 for every line $(i,j) \in \mathcal{L}$. □

Note that Problem (5) is nonconvex due to both the discrete nature of the binary variables ϕ and the nonlinearity of the $E^{ij}X^{ij}$ term in constraint (5c). If we examine the constraint (5c) for some line $(i,j) \in \mathcal{L}$, we see that it is equivalent to: (for $\forall k \in [n_{\text{b}}^{ij}], \forall l \in [m_{\text{db}}^{ij}]$)

$$\sum_{r=1}^{m_{\text{ib}}^{ij}} R_{kr}^{ij} X_{rl}^{ij} \phi[\tilde{\mathcal{M}}_{\text{ib}}^{ij}(r)] = S_{kl}^{ij} \phi[\tilde{\mathcal{M}}_{\text{db}}^{ij}(l)] \tag{6}$$

We can relax the nonconvexity due to the nonlinearity by introducing new variables:

$$Z_{rl}^{ij} \triangleq X_{rl}^{ij} \phi[\tilde{\mathcal{M}}_{\text{ib}}^{ij}(r)] \in \mathbb{R}, \ \ \forall r \in [m_{\text{ib}}^{ij}], \forall l \in [m_{\text{db}}^{ij}] \tag{7}$$

Then we can reformulate (6) with linear relations:

$$\sum_{r=1}^{m_{\text{ib}}^{ij}} R_{kr}^{ij} Z_{rl}^{ij} = S_{kl}^{ij} \phi[\tilde{\mathcal{M}}_{\text{db}}^{ij}(l)], \ \ \forall k \in [n_{\text{b}}^{ij}], \forall l \in [m_{\text{db}}^{ij}] \tag{8}$$

With this reformulation, all the nonlinearity is in the constraints (7). If we relax (7), we have:

$$Z_{rl}^{ij} \leq X_{rl}^{ij} \phi[\tilde{\mathcal{M}}_{\text{ib}}^{ij}(r)], \quad \forall r \in [m_{\text{ib}}^{ij}], \ \forall l \in [m_{\text{db}}^{ij}] \tag{9}$$

We also note that $X_{rl}^{ij} = Z_{rl}^{ij}$. If $\phi[\tilde{\mathcal{M}}_{\text{ib}}^{ij}(r)] = 1$, this is obvious. If $\phi[\tilde{\mathcal{M}}_{\text{ib}}^{ij}(r)] = 0$, then the only constraint X_{rl}^{ij} appears in is (5f), and since we are minimizing the infinity norm of X^{ij}, we have that X_{rl}^{ij} will be equal to zero. Thus, we can substitute Z^{ij} into (5f) and (9) in place of X^{ij}. We can also reformulate constraint (9) using the big-M method by introducing some large constant $C > 0$ such that $Z_{rl}^{ij} \leq C$ for all $r \in [m_{\text{ib}}^{ij}]$, $l \in [m_{\text{db}}^{ij}]$, for all $(i,j) \in \mathcal{L}$ to yield the constraints:

$$Z_{rl}^{ij} \leq C\phi[\tilde{\mathcal{M}}_{\text{ib}}^{ij}(r)], \quad \forall r \in [m_{\text{ib}}^{ij}], \forall l \in [m_{\text{db}}^{ij}] \tag{10}$$

To reformulate the constraint (5f) in order to yield a MILP, we introduce a new variable Y_{rl}^{ij} corresponding to $|Z_{rl}^{ij}|$ for all $r \in [m_{\text{ib}}^{ij}]$ and $l \in [m_{\text{db}}^{ij}]$ which can be related to Z_{rl}^{ij} by the following constraints:

$$Y_{rl}^{ij} \geq \max\{-Z_{rl}^{ij}, Z_{rl}^{ij}\}, \quad \forall r \in [m_{\text{ib}}^{ij}], \forall l \in [m_{\text{db}}^{ij}] \tag{11}$$

We modify (10) to be upper bounds on Y^{ij}:

$$Y_{rl}^{ij} \leq C\phi[\tilde{\mathcal{M}}_{\text{ib}}^{ij}(r)], \quad \forall r \in [m_{\text{ib}}^{ij}], \forall l \in [m_{\text{db}}^{ij}] \tag{12}$$

This formulation allows (5f) to be recast in terms of Y^{ij}:

$$\sum_{l=1}^{m_{\text{db}}^{ij}} Y_{rl}^{ij} \leq \beta, \quad \forall r \in [m_{\text{ib}}^{ij}] \tag{13}$$

In order for the power flow solution to be fully defined in the case without noise, i.e. $\mathbf{m} = A\mathbf{x}$, we need A to be full rank, as the authors suggest in [1]. Instead of enforcing the rank constraint in this optimization problem, we can enforce a weaker constraint which says that every variable must appear in at least one of the measurement equations. We can model this by taking Φ_x to be the indicator variables corresponding to the set of measurements that depend on the variable $x \in \mathcal{X}$. The sets Φ_x are defined based on the structure of the graph and therefore can easily be incorporated into the constraints. To enforce that every variable appears at least once in the chosen measurement equations, we use the constraints:

$$\sum_{i=1}^{|\Phi_x|} \Phi_x[i] \geq 1 \qquad \forall x \in \mathcal{X} \tag{14}$$

where $\Phi_x[i]$ corresponds to the i^{th} element of Φ_x for $x \in \mathcal{X}$. Combining these constraints, we have the MILP of interest:

$$\begin{aligned} \min_{\substack{\beta \in \mathbb{R}, \phi \in \{0,1\}^m \\ Z^{ij}, Y^{ij}, \forall (i,j) \in \mathcal{L}}} \quad & \beta \\ \text{s.t.} \quad & (5b), (14) \\ & \forall (i,j) \in \mathcal{L}: \ \ (8), (11), (12), (13) \end{aligned} \tag{15}$$

In the case that it is impossible to recover a set of measurements that yields $\beta < 1$ for Problem (15), it will be more helpful to minimize the number of violations of the mutual incoherence condition, i.e. where $||Y^{ij}||_\infty \geq 1$. We can do this by solving the related MIP:

$$\begin{aligned} \min_{\substack{\phi \in \{0,1\}^m \\ Z^{ij}, Y^{ij}, \beta^{ij}, \forall (i,j) \in \mathcal{L}}} \quad & \sum_{(i,j) \in \mathcal{L}} \mathbf{1}\{\beta^{ij} \geq 1\} \\ \text{s.t.} \quad & (5b), (14) \\ & \forall (i,j) \in \mathcal{L}: \\ & (8), (11), (12) \\ & \sum_{l=1}^{m_{\text{db}}^{ij}} Y_{rl}^{ij} \leq \beta^{ij}, \quad \forall r \in [m_{\text{ib}}^{ij}] \end{aligned} \tag{16}$$

which is converted to a MILP by introducing binary variables α^{ij} corresponding to the indicators $\mathbf{1}\{\beta^{ij} \geq 1\}$ and using the big-M method to recast the constraints in linear form.

V. Simulations

The simulations are run on a standard laptop using the Pyomo modeling language in Python 3.8. The MILPs given by (15) and (16) are solved with the Gurobi solver, which uses a branch-and-bound method to determine the binary variables.

978-1-6654-0823-3/22/$31.00 ©2022 IEEE

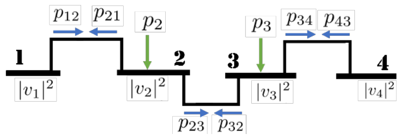

Fig. 1. Four bus network showing all possible voltage magnitude measurements, real power flow measurements, and real power injection measurements. Note that because real power injections at buses 1 and 4 are equivalent to p_{12} and p_{43} respectively, these measurements are not considered.

TABLE I
MUTUAL INCOHERENCE METRIC FOR FOUR BUS NETWORK

Line $i \to j$	Mutual incoherence ρ^{ij} from solving (15)	Mutual incoherence ρ^{ij} from solving (16)
$1 \to 2$	0.91	0.00
$2 \to 3$	1.02	0.90
$3 \to 4$	1.39	1.52
$2 \to 1$	1.39	1.52
$3 \to 2$	1.02	0.93
$4 \to 3$	0.91	0.19

A. Four-Bus Test Case

We first consider the four-bus test network shown in Figure 1. In [2], the authors considered this network and showed that different combinations of measurement choices yielded mutual incoherence metrics that were greater than 1 for certain lines in the network. By considering Problem (15), we formalize their guess-and-check process.

For the four bus network, we take the line parameters to be $G_{ij} = 5$, $B_{ij} = -20$, and $B_{ij}^{\text{sh}} = 0.5$ in per unit values. If we set $\underline{M} = 3|\mathcal{N}| - 2 = 10$, and $\overline{M} = m = 20$, then for the four bus network, we find that it is impossible to recover a set of measurements such that the mutual incoherence condition is satisfied in both directions for every line $(i, j) \in \mathcal{L}$, as shown in Table I. Instead, we can solve (16) to yield a choice of measurements that minimizes the number of violations of the mutual incoherence metric. By solving (16), we see that it is possible to create a measurement set such that 2 out of 3 of the lines are robust in both directions, as shown in the third column of Table I. We see that a mutual incoherence of 0 is obtained for line $1 \to 2$. This occurs because the chosen measurement set has no coupling between attack variables and the rest of the variables, resulting in $\mathcal{X}_{\text{b}}^{12} = \emptyset$.

B. IEEE Test Cases

We solve Problem (15) for some IEEE test cases [17], finding that there is no choice of measurements such that mutual incoherence is below 1 for every line on the network (see [18]). However, we can still solve problem (16) to yield the optimal choice of measurements for the mutual incoherence robustness condition. The results of (16) are given in Table II. Note that if the data for a part of the network is under attack, having more lines satisfy the mutual incoherence condition guarantees a reduction in the impact of the attack on the SE for nodes far away from the attacked region [2].

TABLE II
SOLUTION TO (16) FOR VARIOUS IEEE TEST CASES

Network	Fraction of chosen measurements	Fraction of lines with $\rho^{ij} < 1$	Solve time (s)
case5	30 / 39	6 / 12	1.89
case9	36 / 57	12 / 18	1.49
case14	92 / 120	18 / 40	120.5
case30	190 / 248	37 / 82	831.1

VI. CONCLUSIONS

This paper presented an original framework for optimizing the choice of measurements in a power system to protect against false data injection. By examining a local metric for robust PSSE, we were able to define a coupled optimization problem over all lines of the network. We showed that for some test cases, there is no choice of measurements such that every line can be certified as robust in both directions. However, this framework allows us to find subsets of measurements that are more optimal than others in terms of PSSE robustness.

REFERENCES

[1] M. Jin, I. Molybog, R. Mohammadi-Ghazi, and J. Lavaei, "Scalable and robust state estimation from abundant but untrusted data," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 1880–1894, 2020.
[2] M. Jin, J. Lavaei, S. Sojoudi, and R. Baldick, "Boundary defense against cyber threat for power system state estimation," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1752–1767, 2021.
[3] F. C. Schweppe and J. Wildes, "Power system static-state estimation, part i: Exact model," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, no. 1, pp. 120–125, 1970.
[4] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. CRC Press, 2004.
[5] U.S.-Canada Power System Outage Task Force, *Final Report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations*. U.S. Department of Energy, 2004.
[6] National Academies of Sciences, Engineering, and Medicine, *Enhancing the Resilience of the Nation's Electricity System*. National Academies Press, 2017.
[7] W. Wang and Z. Lu, "Cyber security in the smart grid: Survey and challenges," *Computer Networks*, vol. 57, p. 1344–1371, 04 2013.
[8] R. Deng, P. Zhuang, and H. Liang, "False data injection attacks against state estimation in power distribution systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2871–2881, 2019.
[9] Zong-Han Yu and Wen-Long Chin, "Blind false data injection attack using PCA approximation method in smart grid," *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1219 – 1226, May 2015.
[10] K. Madsen, H. Nielsen, and O. Tingleff, "Methods for non-linear least squares problems (2nd ed.)," p. 60, January 2004.
[11] I. Molybog, S. Sojoudi, and J. Lavaei, "Role of sparsity and structure in the optimization landscape of non-convex matrix sensing," *Mathematical Programming*, pp. 1–37, 2020.
[12] G. Wang, G. B. Giannakis, and J. Chen, "Robust and scalable power system state estimation via composite optimization," 2019.
[13] G. Wang, H. Zhu, G. B. Giannakis, and J. Sun, "Robust power system state estimation from rank-one measurements," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 4, pp. 1391–1403, 2019.
[14] K. Li, "State estimation for power distribution system and measurement impacts," *IEEE Transactions on Power Systems*, vol. 11, no. 2, pp. 911–916, 1996.
[15] R. Madani, J. Lavaei, and R. Baldick, "Convexification of power flow equations in the presence of noisy measurements," *IEEE Transactions on Automatic Control*, vol. 64, no. 8, pp. 3101–3116, 2019.
[16] J. Hao, R. J. Piechocki, D. Kaleshi, W. H. Chin, and Z. Fan, "Sparse malicious false data injection attacks and defense mechanisms in smart grids," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 5, pp. 1–12, 2015.
[17] "EGRET: electrical grid research and engineering tools." [Online]. Available: https://github.com/grid-parity-exchange/Egret
[18] E. Glista and S. Sojoudi, "A MILP for optimal measurement choice in robust power grid state estimation," 2021. [Online]. Available: https://people.eecs.berkeley.edu/~sojoudi/cyberattack_power_2021.pdf

978-1-6654-0823-3/22/$31.00 ©2022 IEEE