On the Exact Solution to a
Smart Grid Cyber-Security Analysis Problem

Kin Cheong Sou, Henrik Sandberg and Karl Henrik Johansson

Abstract—This paper considers a smart grid cyber-security problem analyzing the vulnerabilities of electric power networks to false data attacks. The analysis problem is related to a constrained cardinality minimization problem. The main result shows that an l_1 relaxation technique provides an exact optimal solution to this cardinality minimization problem. The proposed result is based on a polyhedral combinatorics argument. It is different from well-known results based on mutual coherence and restricted isometry property. The results are illustrated on benchmarks including the IEEE 118-bus, IEEE 300-bus and the Polish 2383-bus and 2736-bus systems.

Index Terms—Power network state estimation, security, operation research, optimization methods.

I. INTRODUCTION

Our society relies critically on the proper operation of the electric power distribution and transmission system, which is supervised and controlled through Supervisory Control And Data Acquisition (SCADA) systems. Using remote terminal units (RTUs), SCADA systems measure data such as transmission line power flows, bus power injections and part of the bus voltages, and send them to the state estimator to estimate the power network states. The estimated states are used for vital power network operations such as optimal power flow (OPF) dispatch and contingency analysis (CA) [1], [2]. See Fig. 1 for a block diagram of the above functionalities. Any malfunctioning of these operations can delay proper reactions in the control center, and lead to significant social and economical consequences such as the northeast US blackout of 2003 [3].

As the SCADA systems are increasingly dependent on the Internet, more access points are exposed to potential malicious attackers and hence more SCADA functionalities are subject to threat. For instance, the RTUs can be subjected to denial-of-service attacks (A1 in Fig. 1). The communicated data can be subjected to false data attacks (A2). The SCADA master itself can be attacked (A3). This paper focuses on the cyber security issue related to false data attacks (A2), where the communicated metered measurements are subjected to additive data attacks. A false data attack can potentially lead to erroneous state estimates by the state estimator, which can result in gross errors in OPF dispatch and CA. In turn, these can lead to disasters of significant social and economical
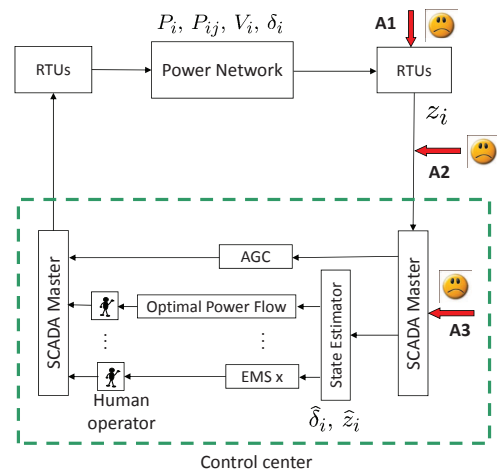
Fig. 1: Block diagram of power network control center and SCADA. RTUs connected to the substations transmit and receive data from the control center using the SCADA system. At the control center, a state estimate is computed and then used by Energy Management Systems (EMS) to send out commands to the power network. The human figures indicate where a human is needed in the control loop. This paper considers the false data attack scenario in A2.

consequences. False data attacks on communicated metered measurements have been considered in the literature (e.g., [4]–[11]). [4] was the first to point out that a coordinated intentional data attack can be staged without being detected by state estimation bad data detection (BDD) algorithm, which is a standard part of today's SCADA/EMS system. [4]–[6], [8]–[11] investigate the construction problem for such "unobservable" data attack, especially the sparse ones involving relatively few meters to compromise, under various assumptions of the network (e.g., DC power flow model [1], [2]). In particular, [4] poses the attack construction problem as a cardinality minimization problem to find the sparsest attack including a given set of target measurements. [5], [6], [9] set up similar optimization problems for the sparsest attack including a given measurement. [8], [11] seek the sparsest nonzero attack and [10] finds the sparsest attack including exactly two injection measurements. The solution information of the above optimization problems can help network operators identify the vulnerabilities in the network and strategically assign protection resources (e.g., encryption of meter measurements) to their best effect (e.g., [6], [7], [11]). On the other hand,

the unobservable data attack problem has its connection to another vital EMS functionality, namely observability analysis [1], [2]. In particular, solving the attack construction problem can also solve an observability analysis problem (this is to be explained in Section II-C). This connection was first reported in [8], and was utilized in [12] to compute the sparsest critical p-tuples for some integer p. This is a generalization of critical measurements and critical sets [1].

To perform the cyber-security analysis in a timely manner, it is important to solve the data attack construction problem efficiently. This effort has been discussed, for instance, in [4]–[6], [8]–[11]. The efficient solution to the attack construction problem in [5] is the focus of this paper. The matching pursuit method [13] employed in [4] and the basis pursuit method [14] (l_1 relaxation and its weighted variant) employed in [11] are common efficient (i.e., polynomial-time) approaches to suboptimally solve the attack construction problem. However, these methods do not guarantee exact optimal solutions, and in some cases they might not be sufficient (see for instance [9] for a naive application of basis pursuit and its consequences). While [8], [11] provide polynomial-time solution procedures for their respective attack construction problems, the problems therein are different from the one in this paper. Furthermore, the considered problem in this paper cannot be solved as a special case of [8], [11]. In particular, in [8] the attack vector contains at least one nonzero entry. However, this nonzero entry cannot be given a priori. [11] needs to restrict the number of nonzero injection measurements attacked, while there is no such requirement in the problem considered in this paper. In [5], [6] a simple heuristics is provided to find suboptimal solutions to the attack construction problem. This heuristics, however, might not be sufficiently accurate. [9], [12], [15] is most closely related to the current work. The distinctions will be elaborated in Section IV-B.

The main conclusion of this paper is that basis pursuit (i.e., l_1 relaxation) can indeed solve the data attack construction problem *exactly*, under the assumption on the network metering system that no injection measurements are metered. The limitations of this assumption will be discussed in Section IV-A. In fact, the main result identifies a class of cardinality minimization problems where basis pursuit can provide exact optimal solutions. This class of problems include as a special case the considered data attack construction problem, under the assumption above.

Outline: Section II describes the state estimation model and introduces the cyber-security analysis optimization problem considered in this paper. Section III describes the main result of this paper – the solution to the considered optimization problem. Section IV compares the proposed result to related works. Section V provides the proof of the main result. Section VI numerically demonstrates the advantages of the proposed result. The conclusion is made in Section VII.

II. State Estimation and Cyber-Security Analysis Optimization Problems

A. Power network model and state estimation

A power network with $n + 1$ buses and m_a transmission lines can be described as a graph with $n + 1$ nodes and m_a edges. The graph topology can be specified by the (directed) incidence matrix $B_0 \in \mathbb{R}^{(n+1) \times m_a}$, in which the direction of the edges can be assigned arbitrarily. The physical property of the network is described by a nonsingular diagonal matrix $D \in \mathbb{R}^{m_a \times m_a}$, whose nonzero entries are the reciprocals of the reactance of the transmission lines.

The states of the network include bus voltage phase angles and bus voltage magnitudes, the latter of which are typically assumed to be constant (equal to one in the per unit system). In addition, since one arbitrary bus is assigned as the reference with zero voltage phase angle, the network states considered can be captured in a vector $\theta \in [-\pi, \pi)^n$. The state estimator estimates the states θ based on the measurements obtained from the network. Under the DC power flow model [1], [2] the measurement vector, denoted as z, is related to θ by

$$z = H\theta + \Delta z, \quad \text{where} \quad H \triangleq \begin{bmatrix} PDB^T \\ QBDB^T \end{bmatrix}. \quad (1)$$

In (1), Δz can be either a vector of random error or intentional additive data attack (e.g., [4]). $B \in \mathbb{R}^{n \times m_a}$ is the truncated incidence matrix (i.e., B_0 with the row corresponding to the reference node removed), and P consists of a subset of rows of an identity matrix of appropriate dimension, indicating which line power flow measurements are actually taken. Together, $PDB^T\theta$ is a vector of the power flows on the transmission lines to be measured. Analogously, the matrix Q selects the bus power injection measurements that are taken. $QBDB^T\theta$ is a vector of power injections at the buses to be measured. Therefore, H is the measurement matrix, relating the measured power quantities to the network states. The number of rows of H is denoted m.

The measurements z and the network information H are jointly used to find an estimate of the network states denoted as $\hat{\theta}$. Assuming that the network is observable, it is well-established that the state estimate can be obtained using the weighted least squares approach [1, Chapter 5] [2, Chapter 8]:

$$\hat{\theta} = (H^T W H)^{-1} W H^T z, \quad (2)$$

where W is a positive definite diagonal weighting matrix, typically weighting more on the more accurate measurements. The state estimate $\hat{\theta}$ is subsequently fed to other vital SCADA functionalities such as OPF dispatch and CA. Therefore, the accuracy and reliability of $\hat{\theta}$ is of paramount concern.

To detect possible faults in the measurements z, the BDD test is commonly performed (see [1], [2]). In one typical strategy, if the norm of the residual

$$\text{residual} \triangleq z - H\hat{\theta} = (I - H(H^T W H)^{-1} W H^T)\Delta z \quad (3)$$

is too big, then the BDD alarm will be triggered.

B. Unobservable data attack and security index

The BDD test is in general sufficient to detect the presence of Δz if it contains a single random error [1], [2]. However, in face of a *coordinated* malicious data attack on multiple measurements the BDD test can fail. In particular, [4] considers unobservable attack of the form

$$\Delta z = H\Delta\theta \quad (4)$$

for an arbitrary $\Delta\theta \in \mathbb{R}^n$. Since Δz as defined in (4) would result in a zero residual in (3), it is unobservable from the BDD perspective. This was also experimentally verified in [16] in a realistic SCADA system testbed. To quantify the vulnerability of a network to unobservable attacks, [5] introduced the notion of security index for an arbitrarily specified measurement.

Definition 1: The **security index** is the optimal objective value of the following cardinality minimization problem:

$$
\begin{aligned}
\underset{\Delta\theta \in \mathbb{R}^n}{\text{minimize}} \quad & \|H\Delta\theta\|_0 \\
\text{subject to} \quad & H(k,:)\Delta\theta = 1,
\end{aligned}
\tag{5}
$$

where k is given, indicating that the security index is computed for measurement k. The symbol $\|\cdot\|_0$ denotes the cardinality of a vector and $H(k,:)$ denotes the kth row of H.

The security index is the minimum number of measurements an attacker needs to compromise in order to attack measurement k undetected. In particular, a small security index for a particular measurement k means that in order to attack k undetected it is necessary to alter only a small number of additional measurements. This can imply that measurement k is vulnerable target in an unobservable attack. As a result, the knowledge of the security indices allows the network operator to pinpoint the security vulnerabilities of the network, and to better protect the network with limited resource. To model the case where certain measurements are protected (hence cannot be attacked) [4], [6], [7], problem (5) becomes:

$$
\begin{aligned}
\underset{\Delta\theta \in \mathbb{R}^n}{\text{minimize}} \quad & \|H\Delta\theta\|_0 \\
\text{subject to} \quad & H(k,:)\Delta\theta = 1 \\
& H(\mathcal{I},:)\Delta\theta = 0,
\end{aligned}
\tag{6}
$$

where the protection index set $\mathcal{I} \subset \{1, 2, \ldots, m\}$ is given, $H(\mathcal{I},:)$ denotes a submatrix of H with rows indexed by $\mathcal{I}$. By convention, the constraint $H(\mathcal{I},:)\Delta\theta = 0$ is ignored when $\mathcal{I} = \emptyset$. Hence, (5) is a special case of (6).

C. Measurement set robustness analysis

Problem (5) is also motivated from another important state estimation analysis problem, namely observability analysis [1], [2]. The measurement set, described by H in (1), is observable if $\hat{\theta}$ can be uniquely determined by (2). An important question of observability analysis is as follows:

$$
\begin{aligned}
\underset{\mathcal{J}}{\text{minimize}} \quad & \|\mathcal{J}\|_0 \\
\text{subject to} \quad & \text{rank}(H(\bar{\mathcal{J}},:)) < n \\
& \text{rank}(H(\bar{\mathcal{J}} \cup \{k\},:)) = n \\
& k \in \mathcal{J}.
\end{aligned}
\tag{7}
$$

In above, k is a given index and $\bar{\mathcal{J}}$ denotes the complement of $\mathcal{J}$ (for index set $\mathcal{I}$ in the rest of the paper, $\bar{\mathcal{I}}$ denotes its complement). The meaning of (7) is as follows: $\mathcal{J}$ denotes a subset of measurements from the measurement system described by H. The condition that $\text{rank}(H(\bar{\mathcal{J}},:)) < n$ means that the measurement system becomes unobservable if the measurements associated with $\mathcal{J}$ are lost. That is, it becomes impossible to uniquely determine $\hat{\theta}$ from $H(\bar{\mathcal{J}},:)$. The problem in (7) seeks the minimum cardinality $\mathcal{J}$ which

must include a particular given measurement k. Therefore, if there exist a measurement k which leads to an instance of (7) with a very small objective value, then the measurement system is not robust against meter failure. Special cases of (7) have been extensively studied in the power system community. For instance, the solution label sets of cardinalities one and two are, respectively, referred to as critical measurements and critical sets containing measurement k [1]. Their calculations have been documented in, for example, [1], [17]–[20]. For the more general cases where the minimum cardinality is $p > 2$, the solution label set in (7) is a critical p-tuple which contains the specified measurement k [12], [21]. Solving (5) solves (7) as well. The justification is given by the following statement inspired by [8], and proved in Appendix:

Proposition 2.1: Let $H \in \mathbb{R}^{m \times n}$ and $k \in \{1, 2, \ldots, m\}$ be given for problems (5) and (7). Denote the two conditions:

I: $H(k,:) \neq 0$.
II: H has full column rank ($= n$).

The following three statements are true:

(a)   Problem (5) is feasible if and only if condition I is satisfied.
(b)   Problem (7) is feasible if and only if conditions I and II are satisfied.
(c)   If conditions I and II are satisfied, then (5) and (7) are equivalent (see Definition 2 in Section V-A).

Note that if condition I is not satisfied, then the corresponding measurement k should be removed from consideration. Also, since measurement redundancy is a common practice in power networks [1], [2], H can be assumed to have full column rank ($= n$). Therefore, conditions I and II in Proposition 2.1 can be justified in practice. Finally, note that Proposition 2.1 remains true for arbitrary matrix H (not necessarily defined by (1)).

III. PROBLEM STATEMENT AND MAIN RESULT

A. Problem statement

This paper proposes an efficient solution to the security index (i.e., attack construction) problem in (6). The proposed result focuses on a special case of (6), where H in (1) does not contain injection measurements:

$$
H = PDB^T.
\tag{8}
$$

The limitation of assumption (8) will be discussed in Section IV-A, after the main result is presented.

In Appendix, it is shown that (6) with assumption (8) is equivalent to

$$
\begin{aligned}
\underset{\Delta\theta \in \mathbb{R}^n}{\text{minimize}} \quad & \|P(\bar{\mathcal{I}},:)B^T\Delta\theta\|_0 \\
\text{subject to} \quad & P(k,:)B^T\Delta\theta = 1 \\
& P(\mathcal{I},:)B^T\Delta\theta = 0.
\end{aligned}
\tag{9}
$$

Instead of considering (9) directly, the proposed result pertains to a more general optimization problem associated with a *totally unimodular matrix* (i.e., the determinant of every square submatrix is either -1, 0, or 1 [22]). In particular, the

following problem is the main focus of this paper:

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & \left\| A(\bar{\mathcal{I}},:)x \right\|_0 \\ \text{subject to} \quad & A(k,:)x = 1 \\ & A(\mathcal{I},:)x = 0, \end{aligned} \qquad (10)$$

where $A \in \mathbb{R}^{m \times n}$ is a given totally unimodular matrix, and $k \in \{1, 2, \ldots, m\}$ and $\mathcal{I} \subset \{1, 2, \ldots, m\}$ are given. Since B in (9) is an incidence matrix, PB^T is a totally unimodular matrix. Therefore, (10) is a generalization of (9). However, neither (10) nor (6) includes each other as special cases since (6) restricts its consideration to a particular class of totally unimodular constraint matrices (i.e., graph incidence matrices).

B. l_1 relaxation

Problem (10) is a cardinality minimization problem. In general, no efficient algorithms have been found for solving cardinality minimization problems [23], so heuristic or relaxation based algorithms are often considered. The l_1 relaxation (i.e., basis pursuit [14]) is a relaxation technique which has received much attention. In l_1 relaxation, instead of (10), the following optimization problem is set up and solved:

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & \left\| A(\bar{\mathcal{I}},:)x \right\|_1 \\ \text{subject to} \quad & A(k,:)x = 1 \\ & A(\mathcal{I},:)x = 0, \end{aligned} \qquad (11)$$

where in the objective function in (11) the vector 1-norm replaces the cardinality in (10). Problem (11) can be rewritten as a linear programming (LP) problem in standard form [24, pp. 4-6, p.17]:

$$\begin{aligned} \underset{x_+, x_-, y_+, y_-}{\text{minimize}} \quad & \sum_{j=1}^{|\bar{\mathcal{I}}|} \big(y_+(j) + y_-(j) \big) \\ \text{subject to} \quad & A(\bar{\mathcal{I}},:)(x_+ - x_-) = y_+ - y_- \\ & A(k,:)(x_+ - x_-) = 1 \\ & A(\mathcal{I},:)(x_+ - x_-) = 0 \\ & x_+ \in \mathbb{R}_+^n, \; x_- \in \mathbb{R}_+^n, \; y_+ \in \mathbb{R}_+^{|\bar{\mathcal{I}}|}, \; y_- \in \mathbb{R}_+^{|\bar{\mathcal{I}}|}, \end{aligned} \qquad (12)$$

where $|\bar{\mathcal{I}}|$ denotes the cardinality of the index set $\bar{\mathcal{I}}$.

If (x_+, x_-, y_+, y_-) is a feasible solution to (12), then $x \triangleq x_+ - x_-$ is feasible to (10). Hence, an optimal solution to (12), if it exists, corresponds to a suboptimal solution to the original problem in (10). An important question is under what conditions this suboptimal solution is actually optimal to (10). An answer is provided by our main result, based on the special structure in (10) and the fact that matrix A is totally unimodular.

C. Statement of main result

Theorem 3.1: Let $(x_+^\star, x_-^\star, y_+^\star, y_-^\star)$ be an *optimal basic feasible solution* to (12), where A, k and $\mathcal{I}$ are defined in (10). Then $x^\star \triangleq (x_+^\star - x_-^\star)$ is an optimal solution to (10).

Remark 3.1: Theorem 3.1 provides a complete procedure for solving (10) via (12). If the standard form LP problem in (12) is feasible, then it contains at least one basic feasible solution (see the definition in Section V-A). Together with

the fact that the objective value is bounded from below (e.g., by zero), [24, Theorem 2.8, p. 66] implies that problem (12) contains at least one optimal basic feasible solution, which can be used to construct an optimal solution to (10) according to Theorem 3.1. Conversely, if the feasible set of (12) is empty, then the feasible set of (10) must also be empty because a feasible solution to (10) can be used to construct a feasible solution to (12).

Remark 3.2: To ensure that an optimal basic feasible solution to (12) is found if one exists, the simplex method (e.g., [24, Chapter 3]) can be used to solve (12).

The proof of Theorem 3.1 will be given in Section V. Before that, the related work are reviewed, and the assumption in (8) is discussed.

IV. RELATED WORK

A. Rationale of the no injection assumption in (8)

Consider the case of (6) where $\mathcal{I}$ corresponds only to line power flow measurements, then with the definition of H in (1) it can be verified that (6) is equivalent to the following:

$$\begin{aligned} \underset{\Delta\theta \in \mathbb{R}^n}{\text{minimize}} \quad & \left\| P(\bar{\mathcal{I}},:)B^T \Delta\theta \right\|_0 + \left\| QB^T DB\Delta\theta \right\|_0 \\ \text{subject to} \quad & P(k,:)B^T \Delta\theta = 1 \\ & P(\mathcal{I},:)B^T \Delta\theta = 0. \end{aligned} \qquad (13)$$

This indicates that the considered problem in (9) is a relaxation [24] of the general case in (13). [9] utilizes this observation and obtains satisfactory suboptimal solution to (6). Alternatively, [12] considers indirectly accounting for the term $\left\| QB^T DB\Delta\theta \right\|_0$ in the objective function of (13). [12] demonstrates that solving the following problem provides satisfactory suboptimal solution to (13)

$$\begin{aligned} \underset{\Delta\theta \in \mathbb{R}^n}{\text{minimize}} \quad & \left\| \tilde{P}(\tilde{\bar{I}},:)B^T \Delta\theta \right\|_0 \\ \text{subject to} \quad & \tilde{P}(\tilde{k},:)B^T \Delta\theta = 1 \\ & \tilde{P}(\tilde{\mathcal{I}},:)B^T \Delta\theta = 0, \end{aligned} \qquad (14)$$

with appropriately defined $\tilde{P}$, $\tilde{I}$ and $\tilde{k}$. Notice that (14) has the same form as our considered problem in (9). In conclusion, the "no injection assumption" in (8) which leads to (9) introduces limitation, but it need not be as restrictive as it might first appear. The proposed result in Theorem 3.1 still leads to an LP based approach to obtain suboptimal solutions to (13) (and hence (6)).

B. Relationship with minimum cut based results

The main strength of the current result lies in the fact that it solves problem (10) where the A matrix is totally unimodular. (10) includes (9) as a special case where the corresponding constraint matrix is a transposed graph incidence matrix. This distinguishes the current work with [9], [12], [15] which specialize in solving (9) using graph-based minimum cut algorithms (e.g., [25]). One example of A which is totally unimodular but not associated with a graph is the matrix with consecutive ones property (i.e., if either for each row or for each column, the 1's appear consecutively) [26]. For a possible application, consider a networked control system [27], [28]

with one controller and n sensor nodes. Each node contains a scalar state value, constant over a period of m time slots. The nodes need to transmit their state values through a shared channel to the controller. Each node can keep transmitting over an arbitrary period of consecutive time slots. At each time slot, the measurement transmitted to the controller is the sum of the state values of all transmitting nodes. Denote $z \in \mathbb{R}^m$ as the vector of measurements transmitted over all time slots, and $\theta \in \mathbb{R}^n$ as the vector of node state values. Then the measurements and the states are related by $z = A\theta$, where $A \in \mathbb{R}^{m \times n}$ is a $(0,1)$ matrix with consecutive ones in the each column. Solving the observability problem in (7) with $H = A$ can identify the vulnerable measurement slots, which should have higher priority in communication for such a networked control system.

C. Relationship with compressed sensing results

Problem (10) can be written in a form more common in the literature. Consider the case where the null space of A^T is not empty (otherwise $\mathrm{rank}(A) = m$ and (10) is trivial). With a change of decision variable $z = Ax$, (10) can be posed as:

$$
\begin{aligned}
\underset{z \in \mathbb{R}^m}{\text{minimize}} \quad & \|z(\bar{\mathcal{I}})\|_0 \\
\text{subject to} \quad & Lz = 0 \\
& z(\mathcal{I}) = 0 \\
& z(k) = 1,
\end{aligned} \tag{15}
$$

where L has full rank and $LA = 0$, and $z(\bar{\mathcal{I}})$ denotes a sub-vector of z containing the entries corresponding to the index set $\bar{\mathcal{I}}$. (15) can be written as the cardinality minimization problem considered, for instance, in [23], [29]–[31]:

$$
\begin{aligned}
\underset{\tilde{z}}{\text{minimize}} \quad & \|\tilde{z}\|_0 \\
\text{subject to} \quad & \Phi \tilde{z} = b,
\end{aligned} \tag{16}
$$

with appropriately defined matrix Φ and vector b. In this subsection, we restrict the discussion to the standard case. That is, (16) is feasible and Φ is a full rank matrix with more columns than rows. As (16) is well-studied, certain conditions regarding when its optimal solution can be obtained by l_1 relaxation are known. For example, [29], [32] report a sufficient condition based on mutual coherence, which is denoted as $\mu(\Phi)$ and defined as

$$
\mu(\Phi) = \max_{i \neq j} \frac{|\Phi(:,i)^T \Phi(:,j)|}{\|\Phi(:,i)\|_2 \|\Phi(:,j)\|_2}. \tag{17}
$$

The sufficient condition [31] states that if there exists a feasible solution $\tilde{z}$ in (16) which is sparse enough:

$$
\|\tilde{z}\|_0 < \frac{1}{2}\left(1 + \frac{1}{\mu(\Phi)}\right), \tag{18}
$$

then $\tilde{z}$ is the unique optimal solution to (16) and its l_1 relaxation (i.e., problem (16) with $\|\tilde{z}\|_1$ replacing $\|\tilde{z}\|_0$). Another well-known sufficient condition is based on the restricted isometry property (RIP) [23], [33]. For any integer s, the RIP constant δ_s of matrix Φ is the smallest number satisfying

$$
(1 - \delta_s)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta_s)\|x\|_2^2 \tag{19}
$$

for all vector x such that $\|x\|_0 \leq s$. The RIP-based sufficient condition [33] states that if for some s, Φ has a RIP constant $\delta_{2s} < \sqrt{2} - 1$, then any $\tilde{z}$ satisfying $\Phi\tilde{z} = b$ and $\|\tilde{z}\|_0 \leq s$ is necessarily the unique optimal solution to both (16) and its l_1 relaxation. It has been shown that certain type of randomly generated matrices satisfy the above conditions with overwhelming probabilities (e.g., [23] provides a RIP-related result). However, the above conditions might not apply to (10), which is the focus of this paper. For instance, consider A in (10) being a submatrix of the transpose of the incidence matrix of the 6-bus power network from [34]:

$$
A = \begin{bmatrix}
1 & -1 & 0 & 0 & 0 & 0 \\
0 & 1 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 \\
1 & 0 & 0 & -1 & 0 & 0 \\
0 & 1 & 0 & 0 & -1 & 0 \\
0 & 0 & 1 & 0 & 0 & -1
\end{bmatrix}.
$$

Let $k = 6$, and $\mathcal{I} = \emptyset$. Then the corresponding Φ in (16) and b are

$$
\Phi = \begin{bmatrix}
1 & 1 & -1 & -1 & -1 & 0 & 1 \\
1 & 0 & -1 & 0 & -1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix} \quad b = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{20}
$$

For this Φ, (17) implies that $\mu(\Phi) = 1$. Therefore, the sparsity bound in (18) becomes $\|\tilde{z}\|_0 < 1$. This is too restrictive to be practical. Similarly, for all $s \geq 1$, the RIP constants δ_{2s} are at least one because $\Phi(:,1) = -\Phi(:,3)$. Hence the RIP-based sufficient condition would not be applicable either. Nevertheless, the failure to apply these sufficient conditions here does not mean that it is impossible to show that l_1 relaxation can exactly solve (16). The mutual coherence and RIP-based conditions characterize when a *unique* optimal solution exists for both (16) and its l_1 relaxation, while in this paper uniqueness is not required. Indeed, for (16) with Φ and b defined in (20), both $\begin{bmatrix} -1 & 0 & 0 & -1 & 0 & 1 & 0 \end{bmatrix}^T$ and $\begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}^T$ are optimal (this can be verified by inspection). Using the CPLEX LP solver [35] in MATLAB to solve the l_1 relaxation leads to the first optimal solution. It is the main contribution of this paper to show that this is the case in general when (16) is defined by (10), even though the optimal solution might not be unique. The reason why the proposed result is applicable is that it is based on a polyhedral combinatorics argument, which is different from those of the mutual coherence and RIP based results.

V. PROOF OF THE MAIN RESULT

A. Definitions

The proof requires the following definitions:

Definition 2: Two optimization problems are **equivalent** if there is an one-to-one correspondence of their instances. The corresponding instances either are both infeasible, both unbounded or both have optimal solutions. In the last case, it is possible to construct an optimal solution to one problem from an optimal solution to the other problem and vice versa. In addition, the two problems have the same optimal objective value.

Definition 3: A **polyhedron** in $\mathbb{R}^p$ is a subset of $\mathbb{R}^p$ described by linear equality and inequality constraints. A standard form polyhedron (as associated with a standard form LP problem instance) is specified by $\{\theta \,|\, C\theta = d, \ \theta \geq 0\}$ for some given matrix C and vector d.

Definition 4: A **basic solution** [24, p. 50] of a polyhedron in $\mathbb{R}^p$ is a vector satisfying all equality constraints. In addition, out of all active constraints p of them are linearly independent. For a standard form polyhedron with a constraint matrix of full row rank, basic solutions can alternatively be defined by the following statement [24, p. 53]:

Lemma 5.1: Consider a polyhedron $\{\theta \,|\, C\theta = d, \ \theta \geq 0\}$, and assume that $C \in \mathbb{R}^{l \times p}$ and C has full row rank. A vector θ is a basic solution if and only if $C\theta = d$ and there exists an index set $\mathcal{J} \subset \{1, 2, \ldots, p\}$, with $|\mathcal{J}| = l$, such that $\det(C(:, \mathcal{J})) \neq 0$ and $\theta(i) = 0$ if $i \notin \mathcal{J}$.

Definition 5: A **basic feasible solution** [24, p. 50] of a polyhedron is a basic solution which is also feasible. By convention, the terminology "a basic feasible solution to a LP problem instance" should be understood as a basic feasible solution of the polyhedron which defines the feasible set of the instance.

B. Proof of Theorem 3.1

Two lemmas, key to the proof, are presented first. The first lemma states that problem (12), as set up by l_1 relaxation, has integer-valued optimal basic feasible solutions.

Lemma 5.2: Let $(x_+^\star, x_-^\star, y_+^\star, y_-^\star)$ be an optimal basic feasible solution to (12). Then it holds that $x^\star(i) \triangleq (x_+^\star(i) - x_-^\star(i)) \in \{-1, 0, 1\}$ for all $1 \leq i \leq n$. In addition, $y^\star(j) \triangleq |A(\bar{\mathcal{I}}(j), :)x^\star| \in \{0, 1\}$ for all $1 \leq j \leq |\bar{\mathcal{I}}|$, where $\bar{\mathcal{I}}(j)$ denotes the jth element of $\bar{\mathcal{I}}$.

Proof: Assume that the feasible set of (12) is nonempty, otherwise there is no basic feasible solution (cf. Definition 5). The following two claims are made:

(a) $A(k, :)$ cannot be a linear combination of the rows of $A(\mathcal{I}, :)$.

(b) There exists $\mathcal{I}' \subset \mathcal{I}$ such that either $\mathcal{I}' = \emptyset$ or the rows of $A(\mathcal{I}', :)$ are linearly independent. In addition, in both cases $A(\mathcal{I}', :)\theta = 0$ and $A(\mathcal{I}, :)\theta = 0$ define the same constraints.

Claims (a) and (b) together imply that problem (12) can be written as a standard form LP problem with a constraint matrix with full row rank (i.e., matrix C below):

$$\underset{\theta}{\text{minimize}} \quad f^T \theta$$
$$\text{subject to} \quad C\theta = d \tag{21}$$
$$\theta \geq 0,$$

with

$$C \triangleq \begin{bmatrix} A(\bar{\mathcal{I}}, :) & -A(\bar{\mathcal{I}}, :) & -I_{|\bar{\mathcal{I}}|} & I_{|\bar{\mathcal{I}}|} \\ A(\mathcal{I}', :) & -A(\mathcal{I}', :) & 0 & 0 \\ A(k, :) & -A(k, :) & 0 & 0 \end{bmatrix}$$

$$d \triangleq \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \theta \triangleq \begin{bmatrix} x_+ \\ x_- \\ y_+ \\ y_- \end{bmatrix} \quad f \triangleq \begin{bmatrix} 0_{n \times 1} \\ 0_{n \times 1} \\ \mathbf{1}_{|\bar{\mathcal{I}}| \times 1} \\ \mathbf{1}_{|\bar{\mathcal{I}}| \times 1} \end{bmatrix}, \tag{22}$$

where $I_{|\bar{\mathcal{I}}|}$ is an identity matrix of dimension $|\bar{\mathcal{I}}|$, and $\mathbf{1}$ is a vector of all ones.

To see the claims, first note that (a) is implied by the feasibility of (12). For (b), If $\mathcal{I} = \emptyset$ or $A(\mathcal{I}, :) = 0$, then set $\mathcal{I}' = \emptyset$. Otherwise, there exists $\mathcal{I}' \subset \mathcal{I}$ with the properties that $|\mathcal{I}'| = \text{rank}(A(\mathcal{I}, :))$, $A(\mathcal{I}', :)$ has linearly independent rows and $A(\mathcal{I}, :) = SA(\mathcal{I}', :)$ for some matrix S. On the other hand, $A(\mathcal{I}', :) = S'A(\mathcal{I}, :)$ for some matrix S', because $\mathcal{I}' \subset \mathcal{I}$. Hence, $A(\mathcal{I}, :)\theta = 0$ and $A(\mathcal{I}', :)\theta = 0$ define the same constraints. This shows (b).

The next step of the proof is to show that every basic solution of (21) has its entries being either -1, 0 or 1. Denote the matrix B_1 as the first $2n$ columns of C, and let $\tilde{B}_1$ be any square submatrix of B_1. If $\tilde{B}_1$ has two columns (or rows) which are the same or negative of each other, then $\det(\tilde{B}_1) = 0$. Otherwise, $\tilde{B}_1$ is a (possibly row and/or column permuted) square submatrix of A, and A is assumed to be totally unimodular. Hence, $\det(\tilde{B}_1) \in \{-1, 0, 1\}$ and B_1 is totally unimodular. Next consider the matrix B defined as

$$B \triangleq \begin{bmatrix} C & d \end{bmatrix} = \begin{bmatrix} A(\bar{\mathcal{I}}, :) & -A(\bar{\mathcal{I}}, :) & -I_{|\bar{\mathcal{I}}|} & I_{|\bar{\mathcal{I}}|} & 0 \\ A(\mathcal{I}', :) & -A(\mathcal{I}', :) & 0 & 0 & 0 \\ A(k, :) & -A(k, :) & 0 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} & -I_{|\bar{\mathcal{I}}|} & I_{|\bar{\mathcal{I}}|} & 0 \\ B_1 & 0 & 0 & 0 \\ & 0 & 0 & 1 \end{bmatrix}.$$

Denote the number of rows and the number of columns of B as m_B and n_B respectively. Let $\mathcal{J} \subset \{1, 2, \ldots, n_B\}$ be any set of column indices of B such that $|\mathcal{J}| = m_B$ (so that $B(:, \mathcal{J})$ is square). If $B(:, \mathcal{J})$ contains only columns of B_1, then $\det(B(:, \mathcal{J})) \in \{-1, 0, 1\}$ since B_1 is totally unimodular. Otherwise, by repeatedly applying Laplace expansion on the columns of $B(:, \mathcal{J})$ which are not columns of B_1, it can be shown that $\det(B(:, \mathcal{J}))$ is equal to the determinant of a square submatrix of B_1, which can only be -1, 0 or 1. Hence, by Cramer's rule the following holds: If v is the solution to the following system of linear equations

$$B(:, \mathcal{J})\, v = B(:, n_B), \ \mathcal{J} \subset \{1, \ldots, n_B - 1\}, \\ |\mathcal{J}| = m_B, \text{ and } \det(B(:, \mathcal{J})) \neq 0, \tag{23}$$

then

$$v(j) \in \{-1, 0, 1\}, \quad \forall\, j. \tag{24}$$

Lemma 5.1 and (23) together imply that the nonzero entries of all basic solutions to (21) are either -1, 0 or 1. Therefore, the basic feasible solutions, which are also basic solutions, to the polyhedron in (21) also satisfy this integrality property.

Finally, let $(x_+^\star, x_-^\star, y_+^\star, y_-^\star)$ be an optimal basic feasible solution. Then feasibility (i.e., nonnegativity) implies that

$$x_+^\star(j) \in \{0, 1\}, \quad x_-^\star(j) \in \{0, 1\}, \\ y_+^\star(j) \in \{0, 1\}, \quad y_-^\star(j) \in \{0, 1\}, \quad \forall\, j. \tag{25}$$

The minimization excludes the possibility that, at optimality, $y_+^\star(j) = y_-^\star(j) = 1$. Hence, it is possible to define $x^\star$ and $y^\star$ such that

$$x^\star(i) \triangleq (x_+^\star(i) - x_-^\star(i)) \in \{-1, 0, 1\} \quad \forall\, i \\ y^\star(j) \triangleq (y_+^\star(j) + y_-^\star(j)) = |A(\bar{\mathcal{I}}(j), :)\, x^\star| \in \{0, 1\} \quad \forall\, j. \tag{26}$$

The second lemma is concerned with a restricted version of (10) with an infinity norm bound as follows:

$$\begin{aligned}
\underset{x}{\text{minimize}} \quad & \left\| A(\bar{\mathcal{I}},:)\, x \right\|_0 \\
\text{subject to} \quad & A(k,:)\, x = 1 \\
& A(\mathcal{I},:)\, x = 0 \\
& \left\| Ax \right\|_\infty \leq 1.
\end{aligned} \tag{27}$$

Lemma 5.3: Optimization problems (10) and (27) are equivalent.

Proof: Suppose (10) is feasible, then it has an optimal solution denoted as $x^\star$. Let $\bar{\mathcal{I}}_{x^\star} \subset \bar{\mathcal{I}}$ be the row index set such that $A(j,:)x^\star \neq 0$ if and only if $j \in \bar{\mathcal{I}}_{x^\star}$. Then it is claimed that there exists a common optimal solution to both (10) and (27) with the same optimal objective value. The argument is as follows. The property of $x^\star$ implies the feasibility of $(12)'$, which is denoted as a variant of (12) with $\bar{\mathcal{I}}$ replaced by $\bar{\mathcal{I}}_{x^\star}$. By [24, Corollary 2.2, p. 65], problem $(12)'$, as a standard form LP problem, has at least one basic feasible solution. Furthermore, since the optimal objective value of $(12)'$ is bounded from below (e.g., by zero), [24, Theorem 2.8, p. 66] implies that $(12)'$ has an optimal basic feasible solution $(\tilde{x}_+^\star, \tilde{x}_-^\star, \tilde{y}_+^\star, \tilde{y}_-^\star)$ which is integer-valued as specified by Lemma 5.2. Denote $\tilde{x}^\star \triangleq (\tilde{x}_+^\star - \tilde{x}_-^\star)$. Then $\tilde{x}^\star$ is feasible to both (10) and (27) since $\bar{\mathcal{I}}_{x^\star} \subset \bar{\mathcal{I}}$, $|A(\bar{\mathcal{I}}_{x^\star},:)\tilde{x}^\star| \in \{0,1\}^{|\bar{\mathcal{I}}_{x^\star}|}$ and $k \in \bar{\mathcal{I}}_{x^\star}$. Also, $\left\| A(\bar{\mathcal{I}},:)\tilde{x}^\star \right\|_0 = \left\| A(\bar{\mathcal{I}}_{x^\star},:)\tilde{x}^\star \right\|_0 \leq \left\| A(\bar{\mathcal{I}}_{x^\star},:)x^\star \right\|_0 = \left\| A(\bar{\mathcal{I}},:)x^\star \right\|_0$, as the inequality is true because $\tilde{x}^\star$ is an optimal solution to $(12)'$. Hence $\tilde{x}^\star$ is optimal to both (10) and (27), with the same objective value.

Conversely, suppose (10) is infeasible, then (27) is also infeasible. This concludes that (10) and (27) are equivalent. ∎

Proof of Theorem 3.1: Let $(x_+^\star, x_-^\star, y_+^\star, y_-^\star)$ be an optimal basic feasible solution to (12). Then there exist $x^\star$ and $y^\star$ as defined in Lemma 5.2. In particular, $x^\star = (x_+^\star - x_-^\star)$. It can be verified that $(x^\star, y^\star)$ is an optimal solution to the following optimization problem:

$$\begin{aligned}
\underset{x,y}{\text{minimize}} \quad & \sum_{j=1}^{|\bar{\mathcal{I}}|} y(i) \\
\text{subject to} \quad & A(\bar{\mathcal{I}},:)x \leq y \\
& -A(\bar{\mathcal{I}},:)x \leq y \\
& A(k,:)x = 1 \\
& A(\mathcal{I},:)x = 0 \\
& 0 \leq y(j) \leq 1 \quad \forall\, j = 1,2,\ldots,|\bar{\mathcal{I}}|,
\end{aligned}$$

where the inequalities above hold entry-wise. Because of the property that $y^\star(j) \in \{0,1\}$ for all j, $(x^\star, y^\star)$ is also an optimal solution to

$$\begin{aligned}
\underset{x,y}{\text{minimize}} \quad & \sum_{j=1}^{|\bar{\mathcal{I}}|} y(i) \\
\text{subject to} \quad & A(\bar{\mathcal{I}},:)x \leq y \\
& -A(\bar{\mathcal{I}},:)x \leq y \\
& A(k,:)x = 1 \\
& A(\mathcal{I},:)x = 0 \\
& y(j) \in \{0,1\} \quad \forall\, j = 1,2,\ldots,|\bar{\mathcal{I}}|.
\end{aligned} \tag{28}$$

It can be verified that (28) is equivalent to (27). Then Lemma 5.3 states that (28) is also equivalent to (10). Consequently, $(x^\star, y^\star)$ being an optimal solution to (28) implies that (10) is feasible with optimal objective value being $\sum_{j=1}^{|\bar{\mathcal{I}}|} y^\star(j)$. A feasible solution to (10) is $x^\star$. Since $y^\star(j) = |A(\bar{\mathcal{I}}(j),:)x^\star| \in \{0,1\} \; \forall j$, it holds that $\left\| A(\bar{\mathcal{I}},:)x^\star \right\|_0 = \sum_{j=1}^{|\bar{\mathcal{I}}|} y^\star(j)$. Hence, $x^\star$ is an optimal solution to (10). ∎

VI. Numerical Demonstration

As a demonstration, instances of the restricted security index problem in (9) are solved with P being an identity matrix and $\mathcal{I}$ being empty. The incidence matrix B describes the topology of one of the following benchmark systems: IEEE 14-bus, IEEE 57-bus, IEEE 118-bus, IEEE 300-bus and Polish 2383-bus and Polish 2736-bus [36]. For each benchmark, (9) is solved for all possible values of k (e.g., 186 choices in the 118-bus case and 411 choices in the 300-bus case). Two solution approaches are tested. The first approach is the one proposed. It is denoted the l_1 approach, and includes the following steps:

1) Set up the LP problem in (12) with A being B^T.
2) Solve (12) using an LP solver (e.g., CPLEX LP). Let $(x_+^\star, x_-^\star, y_+^\star, y_-^\star)$ be its optimal solution.
3) Define $\Delta\theta^\star = x_+^\star - x_-^\star$. It is the optimal solution to (9), according to Theorem 3.1.

The second solution approach to (9) is standard, and it was applied also in [9], [12]. This second approach is referred to as the l_0 approach, as (9) is formulated into the following problem:

$$\begin{aligned}
\underset{\Delta\theta,\, y}{\text{minimize}} \quad & \sum_j y(j) \\
\text{subject to} \quad & B^T \Delta\theta \;\leq\; My \\
& -B^T \Delta\theta \;\leq\; My \\
& B(:,k)^T \Delta\theta \;=\; 1 \\
& y(j) \;\in\; \{0,1\} \quad \forall\, j,
\end{aligned} \tag{29}$$

where M is a constant required to be at least $\left\| B^T \right\|_\infty = \left\| B \right\|_1$ (i.e., the maximum column sum of the absolute values of the entries of B) [9]. Because of the binary decision variables in y, (29) is a mixed integer linear programming (MILP) problem [24]. It can be solved by a standard solver such as CPLEX. The correctness of the l_0 approach is a direct consequence that (29) is a reformulation of (9). As a result, both the l_1 and l_0 approaches are guaranteed to correctly solve (9). Fig. 2 shows the sorted security indices (i.e., optimal objective values of (9)) for the four larger benchmark systems. The security indices are computed using the l_1 approach. As a comparison, the security indices are also computed using the l_0 approach, and they are shown in Fig. 3. The two figures reaffirm the theory that the proposed l_1 approach computes the security indices exactly. Fig. 2 and Fig. 3 indicate that the measurement systems are relatively insecure, as there exist many measurements with very low security indices (i.e., equal to 1 or 2).

In terms of computation time performances, it is well-known that the l_0 approach is much more time-consuming than the l_1 approach since a MILP problem is much more difficult to
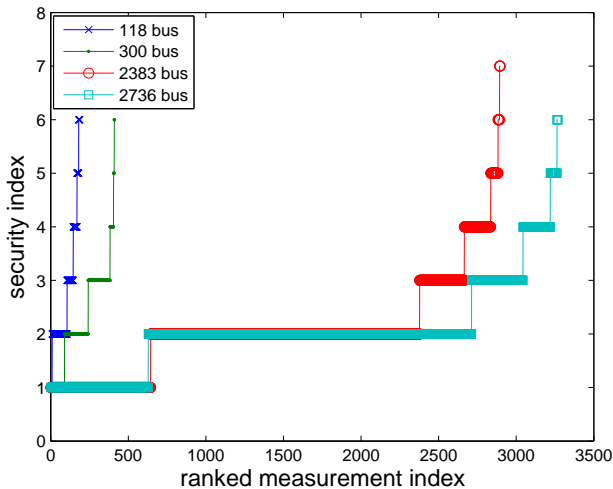
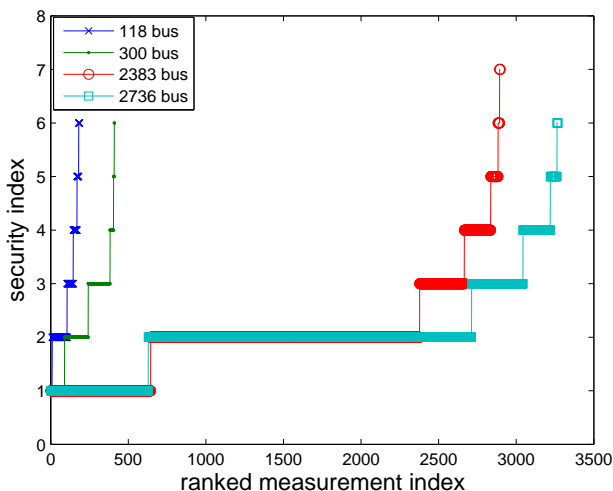Fig. 2: Security indices using the l_1 approach

Fig. 3: Security indices using the l_0 approach

Fig. 4: Solve-time for computing all security indices for different benchmark systems

solve than a LP problem of the same size [24]. Fig. 4 shows the computation time for computing all security indices for each benchmark system, using the l_1 and l_0 approaches. It verifies that the proposed l_1 approach is more effective. The computation time is about a magnitude shorter for the new l_1 approach as compared to the l_0 approach. In the above illustration, all computations are performed on a dual-core Windows machine with 2.4GHz CPU and 2GB of RAM.

VII. CONCLUSION

The cardinality minimization problem is important but in general difficult to solve. A problem example is shown in this paper as the smart grid security index problem in (6). The l_1 relaxation is promising but to establish the cases where it provides exact solutions is non-trivial. Well-known results based on mutual coherence and RIP provide sufficient conditions under which a unique optimal solution solves both the cardinality minimization problem and its l_1 relaxation. However, this pa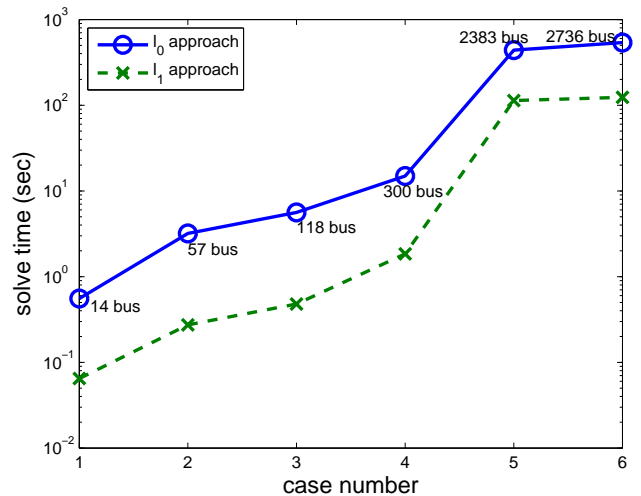per identifies a class of application motivated problems (as in (10)) which can be shown to be solvable by l_1 relaxation, even though results based on mutual coherence and RIP cannot make the assertion. In fact, the optimal solution to (10) might not be unique. The key property that leads to the conclusion of this paper is total unimodularity of the constraint matrix. The total unimodularity of matrix A in (10) leads to two important consequences. (10) is equivalent to its ∞-norm restricted version in (27). Furthermore, (27) can be solved exactly by solving the LP problem in (12), thus establishing the conclusion that l_1 relaxation exactly solves (10).

REFERENCES

[1] A. Abur and A. Expósito, *Power System State Estimation*. Marcel Dekker, Inc., 2004.

[2] A. Monticelli, *State Estimation in Electric Power Systems A Generalized Approach*. Kluwer Academic Publishers, 1999.

[3] G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa, P. Kundur, N. Martins, J. Paserba, P. Pourbeik, J. Sanchez-Gasca, R. Schulz, A. Stankovic, C. Taylor, and V. Vittal, "Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance," *Power Systems, IEEE Transactions on*, vol. 20, no. 4, pp. 1922 – 1928, nov. 2005.

[4] Y. Liu, M. Reiter, and P. Ning, "False data injection attacks against state estimation in electric power grids," in *16th ACM Conference on Computer and Communication Security*, New York, NY, USA, 2009, pp. 21–32.

[5] H. Sandberg, A. Teixeira and K.H. Johansson, "On security indices for state estimators in power networks," in *First Workshop on Secure Control Systems, CPSWEEK 2010*, 2010.

[6] G. Dán and H. Sandberg, "Stealth attacks and protection schemes for state estimators in power systems," in *IEEE SmartGridComm*, 2010.

[7] R. Bobba, K. Rogers, Q. Wang, H. Khurana, K. Nahrstedt, and T. Overbye, "Detecting false data injection attacks on dc state estimation," in *the First Workshop on Secure Control Systems, CPSWEEK 2010*, 2010.

[8] O. Kosut, L. Jia, R. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Transactions on Smart Grid*, vol. 2, pp. 645–658, 2011.

[9] K.C. Sou and H. Sandberg and K.H. Johansson, "Electric power network security analysis via minimum cut relaxation," in *IEEE Conference on Decision and Control*, December 2011.

[10] A. Giani, E. Bitar, M. McQueen, P. Khargonekar, and K. Poolla, "Smart grid data integrity attacks: Characterizations and countermeasures," in *IEEE SmartGridComm*, 2011.

[11] T.T. Kim and H.V. Poor, "Strategic protection against data injection attacks on power grids," *IEEE Transactions on Smart Grid*, vol. 2, pp. 326–333, June 2011.

[12] K.C. Sou, H. Sandberg and K.H. Johansson, "Computing Critical k-tuples in Power Networks," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1511–1520, 2012.

[13] S. Mallat and Z. Zhang, "Matching pursuit with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3397–3415, 1993.

[14] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, 1998.

[15] J. Hendrickx, K.H. Johansson, R. Jungers, H. Sandberg, and K.C. Sou, "An exact solution to the power networks security index problem and its generalized min cut formulation," in preparation, [Online]. Available: http://arxiv.org/abs/1204.6174.

[16] A. Teixeira, G. Dan, H. Sandberg, and K. Johansson, "Cyber security study of a scada energy management system: stealthy deception attacks on the state estimator," in *IFAC World Congress*, Milan, Italy, 2011.

[17] G. Korres and G. Contaxis, "Identification and updating of minimally dependent sets of measurements in state estimation," *IEEE Transactions on Power Systems*, vol. 6, no. 3, pp. 999 –1005, aug 1991.

[18] M. de Almeida, E. Asada, and A. Garcia, "Identifying critical sets in state estimation using gram matrix," in *PowerTech, 2009 IEEE Bucharest*, 28 2009-july 2 2009, pp. 1 –5.

[19] M. Ayres and P. H. Haley, "Bad data groups in power system state estimation," *IEEE Transactions on Power Systems*, vol. 1, no. 3, pp. 1 –7, aug. 1986.

[20] K. Clements, G. Krumpholz, and P. Davis, "Power system state estimation residual analysis: An algorithm using network topology," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-100, no. 4, pp. 1779 –1787, april 1981.

[21] J. London, J.B.A., L. Alberto, and N. Bretas, "Network observability: identification of the measurements redundancy level," in *Power System Technology, 2000. Proceedings. PowerCon 2000. International Conference on*, vol. 2, 2000, pp. 577 –582 vol.2.

[22] A. Schrijver, "A course in combinatorial optimization," CWI, Amsterdam, Netherlands, 2010, online document available from http://homepages.cwi.nl/~lex/.

[23] E. Candès and T. Tao, "Decoding by linear programming," *Information Theory, IEEE Transactions on*, vol. 51, no. 12, pp. 4203 – 4215, Dec. 2005.

[24] J. Tsitsiklis and D. Bertsimas, *Introduction to Linear Optimization*. Athena Scientific, 1997.

[25] M. Stoer and F. Wagner, "A simple min-cut algorithm," *J. ACM*, vol. 44, pp. 585–591, July 1997.

[26] A. Schrijver, *Theory of Linear and Integer Programming*. Wiley, 1998.

[27] J. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 138 –162, jan. 2007.

[28] A. Bemporad, M. Heemels, and M. Johansson, *Networked Control Systems*. Springer, 2010.

[29] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via l_1 minimization," *Proceedings of The National Academy of Sciences*, vol. 100, pp. 2197–2202, 2003.

[30] E. Candès, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted l_1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, pp. 877–905, 2008.

[31] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, 2009.

[32] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," *Information Theory, IEEE Transactions on*, vol. 49, no. 12, pp. 3320 – 3325, dec. 2003.

[33] E. J. Candés, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 910, pp. 589 – 592, 2008.

[34] A. Wood and B. Wollenberg, *Power generation, operation, and control*. J. Wiley & Sons, 1996.

[35] "Cplex," http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/.

[36] R. Zimmerman, C. Murillo-Sánchez, and R. Thomas, "MATPOWER Steady-State Operations, Planning and Analysis Tools for Power Systems Research and Education," *IEEE Transacations on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

APPENDIX

A. Proof of the equivalence between (6) and (9)

Note that the constraint $H(\mathcal{I},:)\Delta\theta = 0$ implies that $\|H\Delta\theta\|_0 = \|H(\bar{\mathcal{I}},:)\Delta\theta\|_0$. Since P consists of rows of an identity matrix and D is diagonal and nonsingular, for all $\mathcal{J} \subset \{1, 2, \ldots, m\}$, there exists a diagonal and nonsingular matrix $D_{\mathcal{J}}$ such that $P(\mathcal{J},:)D = D_{\mathcal{J}}P(\mathcal{J},:)$. In particular, let D_{kk} be a positive scalar such that $P(k,:)D = D_{kk}P(k,:) = P(k,:)D_{kk}$. The above implies that for all $\Delta\theta$,

$P(k,:)B^T\Delta\theta = 1$ if and only if
$P(k,:)D_{kk}B^T(D_{dd}{}^{-1}\Delta\theta) = P(k,:)DB^T(D_{dd}{}^{-1}\Delta\theta) = 1.$

In addition, for all $\Delta\theta$

$P(\mathcal{I},:)B^T\Delta\theta = 0$ if and only if
$D_{kk}{}^{-1}D_{\mathcal{I}}P(\mathcal{I},:)B^T\Delta\theta = P(\mathcal{I},:)DB^T(D_{dd}{}^{-1}\Delta\theta) = 0.$

Finally, for all $\Delta\theta$

$$\begin{aligned}\left\|P(\bar{\mathcal{I}},:)B^T\Delta\theta\right\|_0 &= \left\|D_{kk}{}^{-1}D_{\bar{\mathcal{I}}}P(\bar{\mathcal{I}},:)B^T\Delta\theta\right\|_0 \\ &= \left\|P(\bar{\mathcal{I}},:)DB^T(D_{dd}{}^{-1}\Delta\theta)\right\|_0.\end{aligned}$$

Applying the definition of H in (1) and a change of decision variable to $D_{kk}{}^{-1}\Delta\theta$ shows that (6) and (9) are equivalent.

B. Proof of Proposition 2.1

Part (a) is trivial.

For the necessary part of (b), condition I is necessary because if $H(k,:) = 0$ then $\text{rank}(H(\bar{\mathcal{J}},:)) = \text{rank}(H(\bar{\mathcal{J}} \cup \{k\},:))$ for all $\mathcal{J}$ (meaning that (7) is infeasible). Condition II is also necessary because if $\text{rank}(H) < n$. then there does not exist any $\mathcal{J}$ such that $\text{rank}(H(\bar{\mathcal{J}},:)) = n$.

For the sufficiency part of (b), assume that conditions I and II are satisfied. Then by part (a) problem (5) is feasible. Hence it has an optimal solution denoted as $\theta^\star$. Define $\mathcal{J}_{\theta^\star} \subset \{1, 2, \ldots, m\}$ such that $p \in \mathcal{J}_{\theta^\star}$ if and only if $H(p,:)\theta^\star \neq 0$. By definition of $\mathcal{J}_{\theta^\star}$, $\text{rank}(H(\overline{\mathcal{J}_{\theta^\star}},:)) < n$. Also, $k \in \mathcal{J}_{\theta^\star}$ because $H(k,:)\theta^\star = 1$. If $\text{rank}(H(\overline{\mathcal{J}_{\theta^\star}} \cup \{k\},:)) = n$, then $\mathcal{J}_{\theta^\star}$ is feasible to (7), thus showing that (7) is feasible. To show this, first consider the case when $\|H\theta^\star\|_0 = 1$. Then $\mathcal{J}_{\theta^\star} = \{k\}$ and $\text{rank}(H(\overline{\mathcal{J}_{\theta^\star}} \cup \{k\},:)) = \text{rank}(H) = n$ because of condition II (i.e., H has full column rank). Next consider the case when $\|H\theta^\star\|_0 > 1$ (i.e., $|\mathcal{J}_{\theta^\star} \setminus \{k\}| > 0$). If $\text{rank}(H(\overline{\mathcal{J}_{\theta^\star}} \cup \{k\},:)) < n$, then there exists $\tilde{\theta} \neq 0$ such that $H(\overline{\mathcal{J}_{\theta^\star}} \cup \{k\},:)\tilde{\theta} = 0$. In particular, $H(k,:)\tilde{\theta} = 0$. Also, condition II implies that $H(\mathcal{J}_{\theta^\star} \setminus \{k\},:)\tilde{\theta} \neq 0$ (since otherwise $H\tilde{\theta} = 0$). Let $q \in \mathcal{J}_{\theta^\star} \setminus \{k\}$ such that $H(q,:)\tilde{\theta} \neq 0$. Note also that by definition of $\mathcal{J}_{\theta^\star}$, $H(q,:)\theta^\star \neq 0$. Construct $\theta' \triangleq (H(q,:)\tilde{\theta})\theta^\star - (H(q,:)\theta^\star)\tilde{\theta}$. Then $H(k,:)\theta' = 1$, $H(p,:)\theta' = 0$ whenever $H(p,:)\theta^\star = 0$, but $H(q,:)\theta' = 0$ while $H(q,:)\theta^\star \neq 0$. This implies that θ' is feasible to (5) with a strictly less objective value than that of $\theta^\star$, contradicting the optimality of $\theta^\star$. Therefore, the claim that $\text{rank}(H(\overline{\mathcal{J}_{\theta^\star}} \cup \{k\},:)) = n$ is true. This implies that $\mathcal{J}_{\theta^\star}$ is feasible to (7), establishing the sufficiency of part (b).

For part (c), under conditions I and II both (5) and (7) are feasible. In addition, $\mathcal{J}_{\theta^\star}$ constructed in the proof of the sufficiency part of (b) satisfies $|\mathcal{J}_{\theta^\star}| = \|H\theta^\star\|_0$, for $\theta^\star$

being an optimal solution to (5). This means that the optimal objective function value of (7) is less than or equal to that of (5). For the converse, suppose that $\mathcal{J}^\star$ is optimal to (7), then the feasibility of $\mathcal{J}^\star$ implies that there exists $\theta_{\mathcal{J}^\star} \neq 0$ such that $H(\bar{\mathcal{J}}^\star, :)\theta_{\mathcal{J}^\star} = 0$. This also implies that $\|H\theta_{\mathcal{J}^\star}\|_0 \leq |\mathcal{J}^\star|$. If $H(k, :)\theta_{\mathcal{J}^\star} = 0$, then $H(\bar{\mathcal{J}}^\star \cup \{k\}, :)\theta_{\mathcal{J}^\star} = 0$. This implies that $\mathrm{rank}(H(\bar{\mathcal{J}}^\star \cup \{k\}, :)) < n$, contradicting the feasibility of $\mathcal{J}^\star$. Therefore, there exists a scalar α such that $H(k, :)(\alpha\theta_{\mathcal{J}^\star}) = 1$. Consequently, $\alpha\theta_{\mathcal{J}^\star}$ is feasible to (5) with an objective function value less than or equal to the optimal objective function value of (7).

Karl H. Johansson is Director of the ACCESS Linnaeus Centre and Professor at the School of Electrical Engineering, Royal Institute of Technology, Sweden. He is a Wallenberg Scholar and has held a Senior Researcher Position with the Swedish Research Council. He received MSc and PhD degrees in Electrical Engineering from Lund University in Sweden. He has held visiting positions at UC Berkeley (1998-2000) and California Institute of Technology (2006-07). His research interests are in networked control systems, hybrid and embedded control, and control applications in automation, communication, and transportation systems. He was a member of the IEEE Control Systems Society Board of Governors 2009 and the Chair of the IFAC Technical Committee on Networked Systems 2008-2011. He has been on the Editorial Boards of IEEE Transactions on Automatic Control and the IFAC journal Automatica, and is currently on the Editorial Boards of IET Control Theory and Applications and the International Journal of Robust and Nonlinear Control. He was the General Chair of the ACM/IEEE Cyber-Physical Systems Week (CPSWEEK) 2010 in Stockholm. He has served on the Executive Committees of several European and national research projects in the area of networked embedded systems. He was awarded an Individual Grant for the Advancement of Research Leaders from the Swedish Foundation for Strategic Research in 2005. He received the triennial Young Author Prize from IFAC in 1996 and the Peccei Award from the International Institute of System Analysis, Austria, in 1993. He received Young Researcher Awards from Scania in 1996 and from Ericsson in 1998 and 1999.

Kin Cheong Sou received a Ph.D. degree in Electrical Engineering and Computer Science at Massachusetts Institute of Technology in 2008. From 2008 to 2010 he was a postdoctoral researcher at Lund University, Lund, Sweden, and since September 2010 he has been a postdoctoral researcher at KTH Royal Institute of Technology, Stockholm, Sweden.

His research interests include power system cybersecurity analysis, environment aware building and community, convex/non-convex optimization and model reduction for dynamical systems.

Henrik Sandberg received the M.Sc. degree in Engineering Physics in 1999 and the Ph.D. degree in Automatic Control in 2004, both from the Lund University, Sweden. In 2005-2007, he was a postdoctoral scholar at the California Institute of Technology in Pasadena, USA. Since 2008, he is an Assistant Professor in the Automatic Control Laboratory at the KTH Royal Institute of Technology in Stockholm, Sweden. He has also held visiting appointments at the Australian National University and the University of Melbourne, Australia.

His research interests include secure networked control, power systems, model reduction, and fundamental limitations in control. Henrik Sandberg was the winner of the Best Student-Paper Award at the IEEE Conference on Decision and Control in 2004, and is an associate editor of Automatica.