Chapter 2

State estimation in power grid and false data injection attack

2.1 Introduction

An effective monitoring of the generation and transmission systems leads to efficient load frequency control, economic load dispatch etc. However, with the incorporation of large-scale renewable and electric vehicles, modern interconnected power grids have developed as more complex systems and their secure operation has thus become more difficult. Extensive deployment of SCADA at the control centres has developed data banks that ensure secure grid operation and also help to assess system operation if failure of equipment, or transmission line outages occur.

For the generation of appropriate control signals as well as for an efficient security assessment, an efficient and reliable estimation of the operating states is a necessity. For this, the number of acquired measurements can not be limited to only those required for the power flow module in the EMS. The inputs to the power flow module are the power injections at the load buses and the power and voltage values at the voltage-controlled buses. If any of the aforesaid measurements are unavailable, the power flow solutions from the EMS can not be obtained. Moreover, due to gross errors within the measurements due to noise, the power flow solutions from the EMS may become inaccurate. State estimation algorithms on the other hand incorporate power flows within transmission lines, and power injections at buses which on the other hand overcomes the limitation of the power flow module within EMS. Furthermore, gross errors are also filtered out by bad data detectors

in the control centre.

For a power grid, the term state is not unique and can be defined by those variables with which the operating conditions of the grid can be furnished. In practice, the operator at the control center undertakes the voltage magnitude along with its respective phase angles at all buses as the system states. As the acquired measurements at SCADA are generally contaminated with noise, hence an accurate estimation of the operating states is a necessity for determining the current operating scenario of the grid. An elaborate representation of power system state estimation can be found in more detail in [43, 44]. PMUs can be seen as advanced metering devices with enhanced security leading to a protection-based defence scheme. Such devices are generally used for measuring the voltage phasors of the respective buses in which they are installed followed by the current phasors of the respective lines. As PMUs are costly for large-scale deployment with difficulty in redeployment, an optimal allocation of such PMUs is a critical task. As real-time synchronous phasor measurements acquired from PMUs incorporate a common GPS time stamp for geographically dispersed nodes in the grid [97], the operator at the control centre receives measurements with their respective time stamps, hence it is difficult for the attackers to compromise them [118]. Hence it can be seen that measurements received from PMUs are secured and robust against FDIA. Although PMUs are reliable, due its high cost, traditional sensors are allocated at most of the lines and buses.

As most of the studies in this domain adopt placement of traditional sensors within the grid [202–206], this thesis has also adopted them within this study. The measurements (voltage and current) in their phasor form as acquired from the PMUs are sometimes incompatible with the conventional measurements received at SCADA due to varying sampling rates between SCADA and PMU-based measurements. Hence, recently a lot of works have shown effective strategies for real-time monitoring of the grid using PMU-based robust state estimation [207–209].

Primarily the robust state estimation techniques incorporating SCADA and PMU-based measurements can be broadly classified into the following categories [210, 211]:

- Single-stage power system state estimator that incorporates PMU measurements with the conventional measurements from the meters at SCADA [212–218]. The prime idea behind such a scheme is to map the traditional state vectors in polar form with the voltage and current phasors (in rectangular form as received from

the PMUs) using nonlinear transformation techniques. Such an augmented set of measurements can thus be fed to any static or dynamic state estimators, leading to a single-step state estimation. In comparison to the state estimators based on the conventional measurements acquired at SCADA, high-precision PMU-based measurements are used to enhance their performance. However, one of the key drawbacks of such an approach can be defined as the faster data rate of PMU measurements can not be essentially synchronised with the SCADA measurements and thus these measurements between two conventional SCADA measurements are generally dropped. Thus this can not lead to an effective tracking of system dynamics. To overcome this issue, two-stage schemes are deployed.

- A two-stage scheme is deployed where the states are initially estimated using the conventional measurements from SCADA and are improved using a second set of state estimators that invokes the PMU-based measurements. This indirectly can take advantage of the fast PMU measurements and also can reduce the computational burden of the state estimation algorithms, especially for large-scale grids [209]. Particularly, distributed state estimation using PMU measurements is adopted for large-scale networks [219]. Using tie lines within subsystems can effectively improve this aforesaid approach [220,221]. Local estimation of states using distributed state estimator at substation level using PMU measurements is also demonstrated in [207]. Multi-stage state estimators incorporating PMU measurements have also been demonstrated in [211].

It can be seen that using the graphical scheme for PMU placement, only 31.35% buses need to be integrated with PMUs for an effective attack detection without consideration of zero injection buses. Furthermore, when zero injection buses are considered, only 27.11% of the buses need to be incorporated with PMUs. A detailed overview of optimal locations of PMU placement for an efficient defence strategy against attacks along with the total number of allotted PMUs for the grid can be seen in [120, 222]. PMUs and PDCs are considered to be secured and robust phasor measurement units of the smart grid against FDIAs. It is evident that modern collusive FDIAs can compromise them. A decentralised homomorphic computation paradigm has demonstrated an efficient resilience against such advanced attacks [123]. Although PMUs are placed to defend the grid against

such attacks, still with access to a minimal number of attacked measurements, linear state estimators based on cartesian formulation in the presence of zero injection buses can be attacked successfully. To define and defend the minimal set of measurements, an exact and relaxed reformulation related to the cardinality minimization has also been proposed [223].

2.2 Power system modeling

The load-flow equations are the basis of the state estimation algorithms that relate the state variables with the power flow through the transmission lines along with the power injections at the respective buses. The following set of equations define the active and reactive power flow through the transmission line from bus i to j:

$$P_{[i,j]} = V_{[i]}^2 g_{[i,j]} - V_{[i]} V_{[j]} g_{[i,j]} cos(\theta_{[i,j]}) - V_{[i]} V_{[j]} b_{[i,j]} sin(\theta_{[i,j]}) \tag{2.1}$$

$$Q_{[i,j]} = -V_{[i]}^2 b_{[i,j]} + V_{[i]} V_{[j]} b_{[i,j]} cos(\theta_{[i,j]}) - V_{[i]} V_{[j]} g_{[i,j]} sin(\theta_{[i,j]}) \tag{2.2}$$

where, $V_{[.]}$ and $\theta_{[.]}$ respectively represent the voltage magnitude and its corresponding phase angles, which are the operating states that need to be estimated. Moreover, $g_{[i,j]}$ and $b_{[i,j]}$ respectively denote the conductance and the susceptance of the transmission lines between buses i and j. If bus i is considered as the injection bus, the active and reactive power injections can be defined as follows:

$$P_{[i]} = V_{[i]} \sum_{m'} (V_{[m']} [G_{[i,j]} cos(\theta_{[i,j]}) + B_{[i,j]} sin(\theta_{[i,j]})]) \tag{2.3}$$

$$Q_{[i]} = V_{[i]} \sum_{m'} (V_{[m']} [G_{[i,j]} sin(\theta_{[i,j]}) - B_{[i,j]} cos(\theta_{[i,j]})]) \tag{2.4}$$

where, m' denotes the total number of buses connected to bus i. G, B respectively represent the conductance and susceptance matrices, where the non-diagonal elements are represented as $G_{[i,j]} = -g_{[i,j]}$ and $B_{[i,j]} = -b_{[i,j]}$ respectively. Generally, the control center acquires measurements like power injections at the buses followed by power flows through the transmission lines, hence the operating states could be inter-related with the acquired measurements with the aforesaid set of equations (2.1-2.4) as follows:

$$z = h(x) + c \tag{2.5}$$

From (2.5), a nonlinear state estimation strategy is generally adopted for estimating x. After aggregating the raw measurement data from the PDCs/SCADA, the EMS modules

deploy bad data detection techniques to determine measurement quality and to cater for any bad data present in the measurement set due to meter failures, and noise in communication channels. A lot of nonlinear state estimation techniques have been already proposed like dynamic state estimators [224], Kalman filter [225]. To estimate the operating states from (2.5), a flat start approach is undertaken where all the bus angles are set to 0 with all the voltage magnitudes set to 1 p.u at the start as follows:

$$x[0] = [0\ 0\ 0 \ldots 1\ 1 \ldots 1]^T \tag{2.6}$$

An iterative approach has been adopted to solve the nonlinear state estimation algorithm using the honest Gauss-newton technique where the Jacobian matrix J ($J \in \mathcal{R}^{m \times n}$) is updated at each iteration [43]. Algorithm 1 in next page gives a brief overview of this strategy. It is seen that the algorithm converges with ϵ defined in the range of 10^{-7} for

Algorithm 1: Nonlinear State Estimation Algorithm

Input: $x[0] \in \mathcal{R}^n$: Initial set of state variables for flat start approach; $J[0] \in \mathcal{R}^{m \times n}$: Jacobian matrix formulated from $x[0]$ and z at start, $z \in \mathcal{R}^m$: Input measurement vector, $h(.) \in \mathcal{R}^m$: Nonlinear function connecting the measurements with the state variables, $W \in \mathcal{R}^{m \times m}$: Weight matrix of the meters

Output: $\hat{x}$: Optimally computed set of operating states $\in \mathcal{R}^n$

Initialization Define convergence criteria: ϵ ; `// parameter which defines the rate of convergence of the nonlinear state estimation algorithm`

while $x[n'+1] - x[n'] \leq \epsilon$ **do**

  Compute $\Delta x[n']$ as follows:

  $\Delta x[n'] = (J[n']^T W J[n'])^{-1} J[n']^T W (z - h(x[n']))$

  $x[n'+1] = x[n'] + \Delta x[n']$

  Update $J[n'+1]$ based on $x[n'+1]$ and $z : J[n'+1] = f(x[n'+1], z)$

end while

$\hat{x} = x[n'+1]$; `// Optimally computed set of states`

Return: Optimal set of operating states $\hat{x} \in \mathcal{R}^n$; `// termination of pseudo code`

the undertaken test bench (IEEE 14 bus) as shown in chapter 4. Conventional bad data detectors employ the statistical χ^2 based residue test which is defined as follows:

$$r = ||z - h(\hat{x})||_2 \leq \tau \tag{2.7}$$

τ is chosen on the basis of the degrees of freedom $(m - n)$ of the over-determined class of the system. Measurements showing the highest residues are successfully discarded and the cleaned set of measurements is subsequently fed to the state estimation algorithm within the EMS module. The drawback of such nonlinear honest Gauss-Newton based state estimation algorithms can be furnished as follows:

- The defined algorithm 1 is highly sensitive to initialization. Any random initialization may lead to the non-convergence of the algorithm. Flat start approaches and initializations from the load-flow solutions are the prevalent initialization schemes [43, 73].

- There is no definite proof of convergence of the algorithm under a varying range of initializations [43, 73].

- As it undertakes an iterative approach, exact fixed points of the solution of the estimation algorithms can not be guaranteed under all circumstances [43, 73].

2.3 Linear state estimation

To mitigate the aforesaid issues of the nonlinear state estimation algorithm along with a minimal computational burden, generally, the operators at the control center adopt a linearised model using the following assumptions [43, 226]:

- The magnitude of all bus voltages is assumed to be as 1 p.u.

- The difference between the phase angles of a connected set of buses is very small or close to zero.

- The transmission line resistances are considered negligible in comparison to the line reactances.

Under such aforementioned assumptions, the nonlinear power flow equations as shown in (2.1) and (2.2) can be linearised as follows:

$$\begin{aligned} P_{[i,j]} &= V_{[i]}^2 g_{[i,j]} - V_{[i]} V_{[j]} g_{[i,j]} cos(\theta_{[i,j]}) - V_{[i]} V_{[j]} b_{[i,j]} sin(\theta_{[i,j]}) \\ &= g_{[i,j]} - g_{[i,j]} cos(\theta_{[i,j]}) - b_{[i,j]} sin(\theta_{[i,j]}) \\ &= g_{[i,j]} - g_{[i,j]} - b_{[i,j]} \theta_{[i,j]} \\ &= -b_{[i,j]} \theta_{[i,j]} \end{aligned} \tag{2.8}$$

and

$$\begin{aligned} Q_{[i,j]} &= -V_{[i]}^2 b_{[i,j]} + V_{[i]} V_{[j]} b_{[i,j]} cos(\theta_{[i,j]}) - V_{[i]} V_{[j]} g_{[i,j]} sin(\theta_{[i,j]}) \\ &= -b_{[i,j]} + b_{[i,j]} cos(\theta_{[i,j]}) - g_{[i,j]} sin(\theta_{[i,j]}) \\ &= -b_{[i,j]} + b_{[i,j]} - g_{[i,j]} \theta_{[i,j]} \\ &\simeq 0 \end{aligned} \tag{2.9}$$

It can be seen from (2.9) that the reactive power flows within the transmission lines are negligible and can be assumed to be zero. Hence, for the linearised state estimation model, it can be assumed that only real power flows and injections are prevalent while the respective phase angles of the buses are existent which can be seen as per equation (2.10) and are considered as the operating states. The real power injections at the buses where there are no power generations are formulated by adding all the active power flows connected to that particular bus. The linear state estimation model can be shown as:

$$z = Hx + c \tag{2.10}$$

where, the acquired set of measurements is retrieved from the RTUs, PMUs, and the local PDCs and is generally comprised of the power flows through the transmission lines followed by the power injections at the buses. Generally, the power flow measurements are taken at one end or from sensors connected at both ends of the transmission line. Moreover, the set of operating states is lesser than the number of acquired measurements ($n << m$), hence equation (2.10) can be represented as an over-determined system with an adequate redundancy within the measurements. The estimated set of states (voltage phase angles) for the linear state estimation model can be defined with the help of the topology matrix $H \in \mathcal{R}^{m \times n}$ as shown in (2.10). To estimate the set of operating states, some popular approaches like maximum likelihood, minimum variance, weighted least

squares are also adopted. This thesis adopts the weighted least squares approach for the linear state estimation model to define the set of estimated states as follows:

$$\hat{x} = (H^T W H)^{-1} H^T W z \tag{2.11}$$

$$W = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & \dots \\ 0 & \frac{1}{\sigma_2^2} & \dots \\ \vdots & \vdots & \vdots \\ \dots & 0 & \frac{1}{\sigma_m^2} \end{bmatrix} \tag{2.12}$$

where, $\sigma_{(.)}$ represents the standard deviation of the respective meters along with their measurements. For the adopted linear state estimation model, the topology matrix $H \in \mathcal{R}^{m \times n}$ can be demonstrated as [43, 227]:

$$H = \begin{bmatrix} A^T D A \\ D A \\ -D A \end{bmatrix} \tag{2.13}$$

where, m represents the total number of available measurements whereas n denotes the total number of operating states of the grid. $A \in \mathcal{R}^{l \times n}$ represents the connectivity matrix of the current grid topology where l represents the total number of transmission lines. It can be inferred that the summation of each row $(i = 1, 2, \dots, l)$ for this connectivity matrix is zero. For any transmission line i, element corresponding to the i^{th} row with j^{th} column of A i.e. $[A_{ij}] = 1$ when bus i is directed from the bus j, $[A_{ij}] = -1$ when bus j is directed from bus i, whereas $[A_{ij}] = 0$ when bus i and j are not directly connected. $D \in \mathcal{R}^{l \times l}$ represents the diagonal matrix with admittances of the transmission lines. The term $(A^T D A)$ of (2.13) represents those equations that correspond to the power injections at the respective buses while (DA) and $(-DA)$ represent those terms related to the power flows through the transmission lines. Any unstructured FDIA along with the presence of bad data due to communication failure, meter malfunctions, etc. can be supposed if the following residual holds:

$$||z - H\hat{x}||_2 > \tau \tag{2.14}$$

2.4 Nonlinear state estimation for IEEE 14 bus system

For an effective estimation of the operating states of the undertaken IEEE 14 bus system, line active and reactive power flows along with active and reactive power injections at the respective buses are acquired as the available set of measurements.

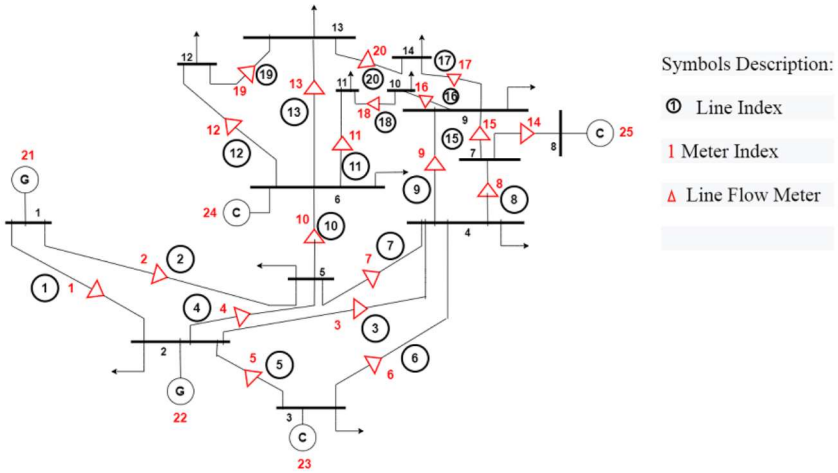

Figure 2.1: IEEE 14 bus system with labelled measurements

All the meters are labelled based on the present grid topology. Furthermore, the incorporated loads can be depicted as per [228]. The operating states comprise the bus voltage magnitudes along with their respective phase angles which is comparatively lesser than the number of acquired measurements. The weights of the corresponding meters incorporated for the nonlinear state estimation algorithm can be seen as per [228]. Fig. 2.1 represents the acquired set of measurements and the corresponding operating states. Some elaborate examples of nonlinear and linear state estimation techniques on the power grid can be seen in [43, 226].

2.5 Bad data detection

Bad data detection plays an important role in the case of state estimation algorithms residing within the EMS module in the control center. The primary objective of such kinds of detectors is to identify the set of faulty measurements owing to noise in communication channels, meter failure, etc, hence leading to higher measurement residuals. Primarily,

the bad data detection can be categorized into two sets of classes as follows:

- Distinct bad data: If the data received at the control center can be realized as potential bad data by observing its values like if the acquired power flow measurements at both ends of the line become positive or negative at the same time. It can be inferred that under such cases there is a distinct presence of bad data within the collected measurements.

- Indistinct bad data: Such kind of bad data can not be effectively distinguished at a glance. Hence, statistical residue test-based bad data detection schemes are undertaken to cater such cases. The presence of such data within measurements is possible due to inadvertent effects as furnished above or may be intentional due to a probable cyber-attack. It can be further differentiated into the following classes:

  - Isolated bad data: If a single data is demarcated as bad data, it can be easily detected as there is a prevalent inconsistency between the acquired measurements. Such kind of bad data may be present due to a single meter failure and can be easily recognized, hence leading to its successful discarding that it does not belong to a critical set of system observability.

  - Non-isolated multiple instances of bad data: Multiple instances of interacting or non-interacting bad data can be categorized into this class. If multiple non-interacting bad data are existing, they can be considered as multiple cases of isolated bad data and can be removed accordingly. In the case of interacting multiple bad data, the probability of detecting them is poor as they constitute a consistent system and hence may lead to grid unobservability.

Some common techniques deployed for detecting such bad data using statistical approaches can be defined as follows:

- The χ^2 distribution test: Such kind of statistical bad data detection tests compare the measurement residuals against a predefined threshold as defined from the χ^2 distribution. The predefined threshold on the basis of which such BDDs work is generally based on the degrees of freedom of the system $(m - n)$. For a predefined confidence interval $(90\% - 95\%)$ and degrees of freedom of the system, if the measurement residuals are lesser than the predefined threshold, it is assumed

that no bad data is present within the acquired set of measurements and hence can be effectively fed within the state estimation algorithms within the control center. However, such kind of a detection test furnishes whether the presence of bad data within measurements is prevalent or not and fails to identify the respective measurements which are faulty. To determine the presence as well as the respective faulty measurements, the normalized residue test as shown below is generally undertaken.

- Normalised residue test: The residue vector is formulated initially followed with the formulation of the sensitivity matrix as $S = I - H(H^T W H)^{-1} H^T W$. The covariance matrix ($\Omega = SW^{-1}$) is further defined followed with formulation of the normalised residue vector as $r_i^N = \frac{|r_i|}{\sqrt{\Omega_{ii}}}$, where r_i represents the i^{th} element of the measurement residue vector. If $\max(r_i^N)$ exceeds a predefined threshold, then the presence of bad data is evident and the measurement with the maximum normalized residue is discarded as potential bad data. Here $\max(.)$ defines the maximum operator, i.e. the maximum value of (.). Such kind of a predefined threshold depends on operator experience. Hence, it can be inferred that such an advanced bad data detection test leads to the identification of the presence of bad data along with the corresponding faulty measurements.

- Hypothesis test: The primary drawback of the aforesaid bad data detection technique is that the residuals may be highly correlated [43]. For multiple instances of interacting with bad data, the hypothesis test is favorable for the grid operator where the operator defines a definite hypothesis over the collected measurements based on the residue vector. Primarily, following are the two main approaches for hypothesis testing:

  - Hypothesis testing technique with a definite probability of false alarm rates.
  - Hypothesis testing technique with a definite probability of missing bad data.

2.6 FDIA on nonlinear state estimation algorithm

The sole idea behind such a class of attack is to circumvent the conventional bad data detector to accept a falsified set of measurements as a valid one and determine a forged set of estimates from them. FDIA has already shown its detrimental effects against the

linear state estimation model [42]. A representation of an effective implementation of FDIA for the nonlinear state estimation approach is shown below:

$$a = h(\hat{x}_a) - h(\hat{x}) \tag{2.15}$$

$$\hat{x}_a = \hat{x} + c' \tag{2.16}$$

$$z_a = z + a \tag{2.17}$$

where, $\hat{x}_a$ ($\hat{x}_a \in \mathcal{R}^n$) denotes the forged set of state estimates after the attack, derived from the falsified set of measurements z_a ($z_a \in \mathcal{R}^m$), while c' ($c' \in \mathcal{R}^n$) denotes the deviation of the state estimates after a successful attack. Residue test under such genre of attack vector formulation can be defined as:

$$\begin{aligned} r_a &= ||z_a - h(\hat{x}_a)||_2 \\ &= ||z + a - h(\hat{x}_a) + h(\hat{x}) - h(\hat{x})||_2 \\ &= ||z - h(\hat{x})||_2 = r \end{aligned} \tag{2.18}$$

It can be inferred that with such an attack vector modeling strategy, the attacker can successfully evade the BDDs and pose a critical scenario on the operating states of the grid.

2.7 FDIA on linear state estimation algorithm

Attacks against the linear state estimation model using the weighted least squares technique are primarily furnished in this section. The primary objective of such kind of an attack strategy is to bypass the bad data detection test as employed for such a state estimation algorithm to accept a corrupted set of measurements as a valid one and compute the falsified set of estimates from them. Such a falsified set of estimated states may lead to mal-operation of the grid as real-time operation, monitoring, and security of the grid along with optimal load dispatch depends on them [229]. Unstructured FDIA when formulated by the attacker has a high possibility of getting detected with the traditional bad data detection technique [230]. Hence, this work incorporates a structured FDIA which can inherently bypass the bad data detection test. To formulate a well-structured FDIA, the attacker must have access to information pertaining to grid topology, measurements, etc. A falsified set of measurements ($\hat{z} \in \mathcal{R}^m$) due to FDIA available at the control center

can be written as:

$$\hat{z} = z + a \tag{2.19}$$

These corrupted class of measurements lead to a falsified set of state estimates ($\hat{x}_1 \in \mathcal{R}^n$) as:

$$\hat{x}_1 = \hat{x} + c', \text{ where } c' \neq 0 \tag{2.20}$$

A well-structured FDIA is possible if the attacker can formulate the attack vector using the full column space of the topology matrix as:

$$a = Hc' \tag{2.21}$$

With such a well-structured attack vector, it can be seen that the attacker can inherently bypass the bad data detection technique based on the $\mathcal{L}_2$ norm of the measurement residuals as follows:

$$||\hat{z} - H\hat{x}_1||_2 = ||(z + a) - H(\hat{x} + c')||_2 = ||z - H\hat{x}||_2 \tag{2.22}$$

It can be well inferred from equation (2.22) that a falsified set of measurements can easily circumvent the bad data detection technique based on the measurement residuals. Although such formulation of attack vectors by the attacker seems feasible but accessing the topology information, location of meters, etc. remains a highly critical task as these pieces of information are kept highly confidential and secured. Such information are necessary to formulate the topology matrix H. With an assumption that with a limited set of resources of the attacker, such critical information of the grid could be accessible as shown in recent works [42, 130]. Hence, it can be concluded that only a small subset of measurements can be tampered, thus developing a stealthy sparse attack vector [49, 166, 231, 232].

2.8 Summary

This chapter has furnished some of the key state estimation techniques which are adopted by modern grid operators along with their vulnerabilities against structured FDIAs. It can be seen that attack vectors can effectively bypass the residue test. The following chapter

demonstrates effective stealthy attack vector formulation schemes against the linear state estimation algorithm using the low-rank structure of the topology matrix. It can be seen that with limited topology information along with some constraints, the attack vector can bypass the residue test.