# Chapter 3

# Reduced-rank based approaches for false data injection attack

## 3.1 Introduction

Energy management systems in the control center deal with various critical operations of the grid like contingency analysis, load dispatch, etc. Recently data-driven attack vector formulation schemes have demonstrated a profound impact on the modern grid. This chapter promotes a novel attack vector formulation policy for the linear state estimation algorithm by exploring the low-rank subspace of the topology matrix. With constraints over the state deviation vector, this work showcases that the developed attack vector can effectively bypass the residue test for the undertaken IEEE 14-bus test bench. An extensive analysis promotes the aforementioned propositions.

## 3.2 Preliminaries

Modern smart grids have already demonstrated their vulnerability to FDIAs targeting the power system state estimator. It is seen from [42, 166, 233–236] that most of the prevalent approaches of defining an attack vector incorporate the following assumptions of the grid:

- With a limited set of resources of the attacker, it is assumed that the attacker is capable of formulating the topology matrix.
- Furthermore, it is also assumed that measurements are accessible by the attacker

and it can be altered leading to a forged set of state estimates.

The aforesaid assumptions also hold for this chapter. This chapter furnishes two specific scenarios like when the attacker has an accurate knowledge of the topology matrix and when presence of outliers or large scale noises are prevalent within $H$. The following sections demonstrate an effective attack vector formulation strategy under the aforesaid cases.

## 3.3 Low-rank structure based attack vector formulation strategy

It is seen from [42, 166, 233–236] that most of the current approaches formulate a stealthy attack vector that is defined over the full column space of the topology matrix, hence leading to an optimal attack on the grid which is not localized. The current study also furnishes attack vectors which may be sparse due to limited resources of the attacker [166] which is developed on the span of the column space of the topology matrix ($H$). FDIAs based on the subspace defined by the covariance matrix of the acquired measurements have already demonstrated their effects on grid operation [88,91]. It is seen that such data-driven attack vectors still have a possibility of getting detected by the bad data detection algorithm. This chapter showcases that an attack vector defined on the low-rank subspace of the topology matrix can successfully bypass the BDD. A critical comparison between prevalent algorithms like CUR, SVD, and Go-Dec in computing the low-rank subspace under the presence and absence of outliers is showcased here. SVD although can give an optimal low-rank subspace, still it shows a higher computational burden than the bilateral random projection-based algorithms like Go-Dec while CUR develops a faster low-rank subspace with an optimal choice of sampling and rescaling the columns of $H$.

The physical significance of the low-rank subspace of the topology matrix furnishes the subspace information of the admittance matrix of the transmission lines. The low-rank subspace of $H$ defines a subspace of $D$ which spans a smaller set of independent information regarding the admittances of the transmission lines. The proposed strategy demonstrates the ability of the attacker to compromise the current operating scenario of the grid even if a small fraction of system information is leaked. This may lead to an

unobservable attack on the grid even if the rest of the system information is kept secured. With the low-rank subspace information of $D$ alongside the topological information of the grid ($A$), the attacker is capable of defining unobservable attacks on the grid. It is seen that most of the state-of-the-art approaches to defend the grid against FDIAs define a set of sensors whose data is protected via data authentication schemes while it is assumed that the attacker can access the system information [97, 121, 237, 238]. It is henceforth necessary for the utility to not only secure the sensor data but also to ensure that system data remain confidential. Hence, it can be seen that decomposing the topology matrix $H$ into a low-rank subspace leads to better subspace information which provides an efficient scheme for stealthy data-driven attack vector formulation. The attack vectors defined over the full column space poses a higher severity to modern grid operation. Although they are capable of bypassing the residue test, a large-scale sustained deviation from the set of uncorrupted estimated states may lead to a higher possibility of its detection. Attack vectors defined over the low-rank subspace although develops a minor attack vector formulation strategy, but are highly covert in nature and can not be effectively distinguished from the measurement noise.

If the original rank of the topology matrix $H$ be $r$ while $k << r$ represents its low-rank approximation, then if the attacker formulates an attack vector $a' \in \mathcal{R}^m$ based on the low-rank approximate $H_k \in \mathcal{R}^{m \times n}$ as shown:

$$a' = H_k c'' \tag{3.1}$$

With such an attack vector defined on the low-rank approximate, a compromised set of measurements $z'' \in \mathcal{R}^m$ followed by a forged set of estimated states $\hat{x}_2 \in \mathcal{R}^n$ are developed. $c'' \in \mathcal{R}^n$ denotes the very small deviation of the estimated states due to an attack defined on the low-rank subspace. Such an attack vector formulation strategy can effectively bypass the residue test employed for bad data detection as shown:

$$z'' = z + a', \quad \hat{x}_2 \simeq \hat{x} + c'' \tag{3.2}$$

$$\begin{aligned} ||z'' - H\hat{x}_2|| &= ||(z + a') - H(\hat{x} + c'')|| \\ &\leq ||z - H\hat{x}|| + ||(H_k - H)c''|| \end{aligned} \tag{3.3}$$

It can be clearly seen that the attack defined on the low-rank subspace of the topology matrix $H$ is bound to bypass the BDD if the vector $c''$ i.e the small non-zero vector

comprising of deviation of states due to an attack lies in the null space of the matrix $(H_k - H)$. Hence under this constraint, an attacker can define the attack vector in the low-rank subspace which is guaranteed to bypass the bad data detection algorithm.

In case of an accurate determination of the topology matrix by the attacker, this chapter furnishes that the CUR decomposition technique requires a minimal allocation of resources along with a reduced computational burden than SVD to define the low-rank subspace. Moreover, when there is a possibility of the presence of outliers within the acquired topology matrix by the attacker, matrix separation schemes like Go-Dec demonstrate an efficient performance. A detailed overview of the attack vector formulation scheme can be depicted in Fig. 3.1.

### 3.3.1 SVD based low-rank approximation

It is known from [239] that SVD can be used to obtain an optimal low-rank subspace of the full-rank topology matrix $H$ based on orthogonal vectors spanning its column and row spaces that leads to the best possible low-rank approximate $H_k$. Algorithm 2 presents the steps to obtain $H_k$ from $H$ using SVD [239]. $U_r \in \mathcal{R}^{m\times r}$, $\Sigma_r \in \mathcal{R}^{r\times r}$, $V_r \in \mathcal{R}^{n\times r}$ denote the left orthogonal matrix spanning the column space, the diagonal matrix with $r$ nonzero entries representing the set of singular values and the right orthogonal matrix spanning the row space of $H$, respectively. Though such an approximation yields an

**Algorithm 2:** Low-rank approximation using SVD

**Input:** Topology matrix $H \in \mathcal{R}^{m\times n}$

**Output:** Low-rank matrix approximate $H_k \in \mathcal{R}^{m\times n}$

1 **Compute**: SVD of $H$, $H = U_r\Sigma_r V_r^T$ ; `// Reduced order singular value decomposition is undertaken to reduce the computational burden`

2 **Compute**: $U_k \in \mathcal{R}^{m\times k}$, $\Sigma_k \in \mathcal{R}^{k\times k}$, $V_k \in \mathcal{R}^{n\times k}$ ; `// Top` $k$ `singular values along with their respective left and right orthogonal vectors`

3 **Compute**: $H_k = U_k\Sigma_k V_k^T$; `// Low-rank approximate of the topology matrix` $H$

4 **Return**: $H_k$

optimal low-rank subspace but offers an extensive computational burden while obtaining the FSVD. Even though using the RSVD as in Algorithm 2 promotes lesser computation,

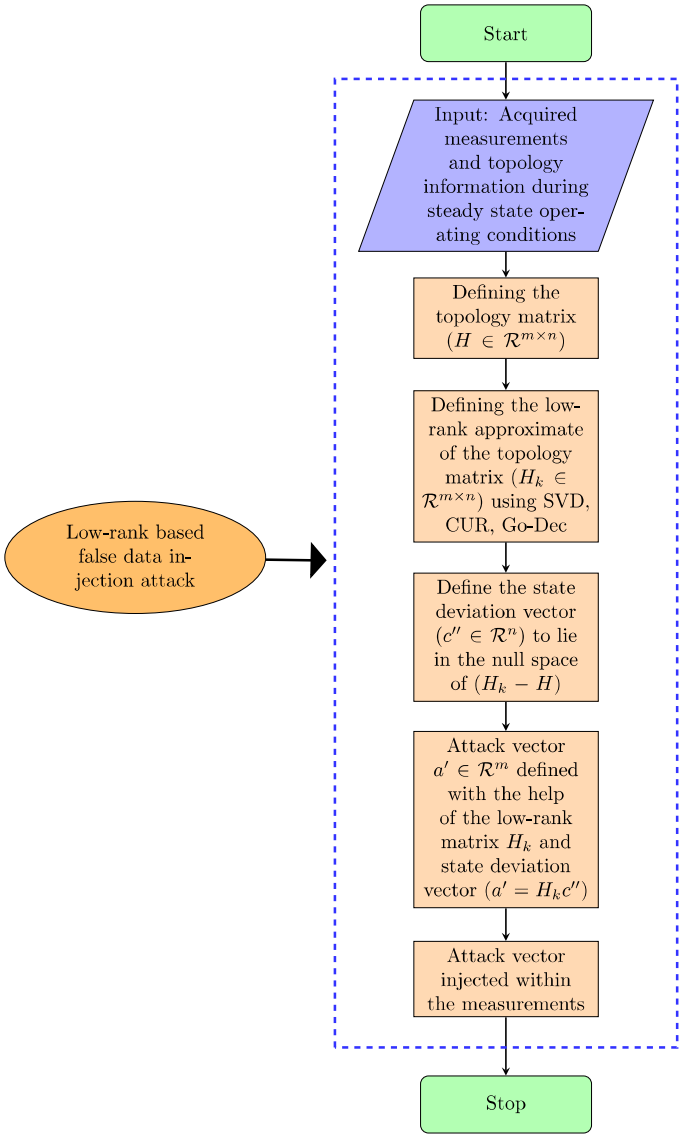


Figure 3.1: An overview of the attack vector formulation scheme

there still remains scope of improvement in obtaining a low-rank approximation using the CUR decomposition when an accurate knowledge of the topology matrix is present with the attacker.

### 3.3.2 Low-rank approximation using CUR decomposition

CUR decomposition demonstrates a faster low-rank matrix approximation with high efficacy [240]. The low-rank matrix $H_k$ is approximated as the product of three smaller matrices $C \in \mathcal{R}^{m \times c}$, $U \in \mathcal{R}^{c \times d}$ and $R \in \mathcal{R}^{d \times n}$, each of which can be formulated rapidly. $c$ and $d$ are positive constant integers. Clearly, it does not provide an orthogonal ba-

sis for the row and column spaces unlike SVD. Such an $H_k$ computed using the CUR decomposition has the following properties:

- $H_k = CUR$, where $C$ is formulated using $c$ randomly chosen columns of $H$. $R$ is formulated using $d$ randomly chosen rows of $H$, while $U$ is computed using $C$ and $R$.

- As $c$ and $d$ are constants, hence $U$ can be constructed with a computational burden which is linear in the order of $(m+n)$ [240].

- It is seen from [240] that for every $k > 0$ and $1 \leq k \leq r$, $c,\ d$ can be chosen such that $H_k$ satisfies the following with high probability, for a very small positive $\epsilon$.

$$\begin{aligned} ||H - H_k||_F \leq \ & \min_{D} \quad ||H - D||_F + \epsilon ||H||_F \\ & \text{s.t.} \quad rank(D) \leq k \end{aligned} \tag{3.4}$$

As the matrices $C,\ U,\ R$ can be formulated rapidly with a small constant number of passes over the topology matrix $H$, CUR decomposition demonstrates a faster low-rank approximation with an error bound in the order of $\epsilon||H||_F$ [240].

The proposed methodology (CUR) demonstrates a faster low-rank subspace that is defined with a minimal error with respect to the optimal low-rank subspace as defined by the singular value decomposition (SVD) technique [239, 241]. It must be noted that with modern grid restructuring along with topological changes due to the incorporation of renewables and advanced flexible ac transmission systems (FACTS) within the grid, the operator at the control center needs to store the topology matrix $H \in \mathcal{R}^{m \times n}$ where $\mathcal{O}(mn)$ elements need to be stored at the data center. Here $\mathcal{O}(.)$ represents the order of computation. Most of the studies in this domain assume that the attacker has access to the current grid topology and hence the topology matrix of the grid [42,54,121,166,234]. With the proposed scheme, the attacker is capable of rapidly formulating the low-rank structure ($H_k \in \mathcal{R}^{m \times n}$) of the topology matrix $H$ with the aid of three smaller matrices $C \in \mathcal{R}^{m \times c}$, $U \in \mathcal{R}^{c \times d}$ and $R \in \mathcal{R}^{d \times n}$ respectively by sampling and rescaling an appropriate number of rows and columns of the topology matrix $H$. Furthermore, with the limited resources of the attacker, the full order or reduced order singular value decomposition technique which defines an optimal low-rank subspace is a critical task as it requires $\mathcal{O}(mn)$ elements of the topology matrix to be stored in the random access memory (RAM). It is seen that

the demonstrated CUR decomposition algorithm can define a low-rank structure as a product of three smaller matrices without storing the $\mathcal{O}(mn)$ elements of the topology matrix in the RAM, provided the algorithm has access to the topology matrix in the data center. With two passes over the $\mathcal{O}(mn)$ elements of the topology matrix along with using additional $\mathcal{O}(m+n)$ RAM while keeping $c,\ d$ constants and independent of $m,\ n$ [240], the attacker can formulate the low-rank structure of the topology matrix $H$ with a reduced computational burden. The algorithm is capable of developing optimal probabilities [242] in the selection of the rows and columns of $H$ in the first pass while in the second pass over the topology matrix the appropriate rows and columns of $H$ are sampled and rescaled. It is seen that the low-rank approximation of the topology matrix using SVD can be formulated as per (3.5) [239, 241].

$$\begin{aligned} ||H - H_k||_F^2 \leq \min_{D} \quad & ||H - D||_F^2 \\ \text{s.t.} \quad & rank(D) \leq k \end{aligned} \tag{3.5}$$

The demonstrated scheme on the other hand develops a low-rank structure with a small error margin from the optimal low-rank subspace as shown in equation (3.4) [240].

Let $H_{k1}$ be defined using the singular values and right singular vectors of $C$ as $H_{k1} = [h^1\ h^2 \dots h^k] \in \mathcal{R}^{m \times k}$, where $h^t = Cy^t/\sigma_t(C)\ \forall\ t \in [1\ 2 \dots k]$, $\sigma(C)$ represents the singular values of $C$ with its right singular vectors $(y)$. With an optimal choice of probabilities along with the projection of the topology matrix over the low-rank subspace as spanned by the top $k$ right singular vectors of $C$ as $H_{k1}H_{k1}^T H$, the subspace is seen to be almost similar to that of the optimal low-rank subspace as defined by SVD [243]. It can be seen that such an approximation leads to an additional error which depends on $||HH^T - CC^T||_F$ as:

$$||H - H_{k1}H_{k1}^T H||_F^2 \ \leq\ ||H - H_k||_F^2 + 2\sqrt{k}||HH^T - CC^T||_F \tag{3.6}$$

*Proof.* As per the definition of the Frobenius norm of a matrix:

$$\begin{aligned} ||H - H_{k1}H_{k1}^T H||_F^2 \ &=\ Tr[(H - H_{k1}H_{k1}^T H)^T (H - H_{k1}H_{k1}^T H)] \\ &= Tr(H^T H - 2H^T H_{k1}H_{k1}^T H + H^T H_{k1}H_{k1}^T H_{k1}H_{k1}^T H) \\ &= Tr(H^T H) - Tr(H^T H_{k1}H_{k1}^T H) \\ &= ||H||_F^2 - ||H^T H_{k1}||_F^2 \end{aligned} \tag{3.7}$$

where, $Tr(.)$ represents the trace of a matrix and it is seen that as $H_{k1}^T H_{k1} = I_k$, hence $||H^T H_{k1}||_F^2$ can be related to the singular values of $C^T C$ using Cauchy–Schwarz inequality as follows:

$$\begin{aligned} |\ ||H^T H_{k1}||_F^2 - \sum_{t=1}^{k} \sigma_t^2(C)|\ &\leq \sqrt{k}[\sum_{t=1}^{k}(|H^T h^t|^2 - \sigma_t^2(C))^2]^{1/2} \\ &= \sqrt{k}[\sum_{t=1}^{k}(|H^T h^t|^2 - |C^T h^t|^2)^2]^{1/2} \\ &= \sqrt{k}[\sum_{t=1}^{k}(h^{t^T}(HH^T - CC^T)h^t)^2]^{1/2} \\ &\leq \sqrt{k}||HH^T - CC^T||_F \end{aligned} \tag{3.8}$$

The last inequality holds with respect to a basis incorporating $[h^t]_{t=1}^k$ for the matrices $HH^T$ and $CC^T$. By applying Cauchy–Schwarz and Hoffman–Wielandt inequality [244] the singular values of $C^T C$ and $H$ can be inter-related as:

$$\begin{aligned} |\ \sum_{t=1}^{k} \sigma_t^2(C) - \sum_{t=1}^{k} \sigma_t^2(H)| &\leq \sqrt{k}(\sum_{t=1}^{k}(\sigma_t^2(C) - \sigma_t^2(H))^2)^{1/2} \\ &= \sqrt{k}(\sum_{t=1}^{k}(\sigma_t(CC^T) - \sigma_t(HH^T))^2)^{1/2} \\ &\leq \sqrt{k}(\sum_{t=1}^{m}(\sigma_t(CC^T) - \sigma_t(HH^T))^2)^{1/2} \\ &\leq \sqrt{k}||CC^T - HH^T||_F \end{aligned} \tag{3.9}$$

Hence by combining the results of equations (3.8) and (3.9), $||H^T H_{k1}||_F^2$ and the singular values of the topology matrix ($\sum_{t=1}^{k} \sigma_t^2(H)$) can be inter-related as:

$$|\ ||H^T H_{k1}||_F^2 - \sum_{t=1}^{k} \sigma_t^2(H)|\ \leq 2\sqrt{k}||HH^T - CC^T||_F \tag{3.10}$$

Combining the results obtained from equations (3.10) and (3.7) proves the theorem as shown in (3.6). It is seen that the subspace developed by $C^T C$ can effectively capture nearly the top $k$ singular values of the topology matrix $H$ with minimal error bounds as shown in (3.6). Furthermore, it can be seen from [242] that with an optimal choice of the probability of sampling and rescaling the columns of the topology matrix of $H$ and with Jensen's inequality, $||HH^T - CC^T||_F$ can be approximated as follows:

$$E[||HH^T - CC^T||_F] \leq \frac{1}{\sqrt{c}}||H||_F^2 \tag{3.11}$$

where $E[.]$ represents the expectation of a random variable. Moreover if $\delta \in (0,1)$ and if $\eta = 1 + \sqrt{8log(\frac{1}{\delta})}$, then it can be inferred from [242] that with probability of at least (1-$\delta$), the following holds:

$$||HH^T - CC^T||_F \leq \frac{\eta}{\sqrt{c}}||H||_F^2 \tag{3.12}$$

This indirectly demonstrates that with minimal error and with optimal probabilities, the topology matrix can be approximated. The aforementioned theorem can also be proved for the spectral norm bound.

□

With the aforesaid equation (3.6), the three small matrices $C$, $U$ and $R$ can be used to approximate the topology matrix. To demonstrate bounds on the developed low-rank structure, it can be seen using triangle inequality that:

$$||H - CUR||_F \leq ||H - H_{k1}H_{k1}^T H||_F + ||H_{k1}H_{k1}^T H - CUR||_F \tag{3.13}$$

It can be seen from equations (3.7), (3.8), (3.9) and (3.12) that the first term of the triangle inequality is bounded by $||H||_F^2$. It is seen from [240] that the approximation of the topology matrix ($H_{k1}H_{k1}^T H$) satisfies the properties of CUR decomposition as shown above. With the use of sampling matrices $D_d \in \mathcal{R}^{d\times d}$ and $S_d \in \mathcal{R}^{d\times m}$, the followings can be inferred:

$$\hat{H_k^T} = H_{k1}^T (D_d S_d)^T \tag{3.14}$$

$$\hat{H} = D_d S_d H \tag{3.15}$$

where, $\hat{H_k^T}$ and $\hat{H}$ denote the column sampled and rescaled structures and row sampled and rescaled structures of $H_{k1}^T$ and $H$ respectively. It can be seen that [240]:

$$CUR = H_{k1}H_{k1}^T (D_d S_d)^T D_d S_d H \tag{3.16}$$

$$= H_{k1}\hat{H_k^T}\hat{H} \tag{3.17}$$

Furthermore as $H_{k1}^T H_{k1} = I_k$, the second term of the inequality of (3.13) can be represented as:

$$\begin{aligned} ||H_{k1}H_{k1}^T H - CUR||_F &= ||H_{k1}^T H - H_{k1}^T (D_d S_d)^T D_d S_d H||_F \\ &= ||H_{k1}^T H - \hat{H_k^T}\hat{H}||_F \end{aligned} \tag{3.18}$$

Furthermore with an optimal choice of probabilities with $c \geq \frac{1}{\epsilon^2}$ and using Jensen's inequality $||H_{k1}^T H - \hat{H}_k^T \hat{H}||_F$ can be approximated with high expectation and probability as follows [242, 243]:

$$E[||H_{k1}^T H - \hat{H}_k^T \hat{H}||_F] \leq \epsilon ||H_{k1}||_F ||H||_F \tag{3.19}$$

and if $c \geq \frac{\eta^2}{\epsilon^2}$, then with probability of at least $1 - \delta$ the following holds [242]:

$$||H_{k1}^T H - \hat{H}_k^T \hat{H}||_F \leq \epsilon ||H_{k1}||_F ||H||_F \tag{3.20}$$

This demonstrates the efficacy of the proposed scheme as the developed low-rank structure can be bounded with high probability. Moreover, with a small margin of error along with a faster computation than the SVD approach, the developed low-rank structure can be seen to provide the attacker a serious advantage in formulating the attack vector on the low-rank subspace of the topology matrix.

Algorithm 3 presents the steps involved in the CUR decomposition. $H_{(i)}$ and $H^{(j)}$, $[i = 1, \ldots, m]$, $[j = 1, \ldots, n]$, represent the $i^{th}$ row of $H$ as a row vector and $j^{th}$ column of $H$ as a column vector, respectively. First, the matrix $C$ is formulated with $c$ randomly chosen columns of $H$. These columns are selected with $c$ independent trials with the $\beta^{th}$ column of $H$ being selected with a probability of $q_\beta$. After selecting the $\beta^{th}$ column, it is rescaled by a factor of $1/\sqrt{cq_\beta}$ before it is included in $C$. Similarly, the matrix $R$ is formulated by randomly choosing $d$ rows of $H$ in $d$ independent trials. It is further rescaled by a factor of $1/\sqrt{dp_\beta}$. $p_\beta$ is the probability of choosing the $\beta^{th}$ row of $H$. In a similar fashion, the CUR decomposition algorithm formulates a matrix $P \in \mathcal{R}^{d \times c}$ from $C$, where $P_{i,j}$ represents its $i,\ j^{th}$ element. It can be represented as $P_{i,j} = H_{i_{t_1}, j_{t_2}} / \sqrt{cdp_{i_{t_1}} q_{j_{t_2}}}$. $i_{t_1}$ is the element chosen from $[1, \ldots, m]$ selected in $t_1^{th}$ row sampling trial, whereas $j_{t_2}$ represents the element chosen from $[1, \ldots, n]$ selected in $t_2^{th}$ column sampling trial.

Let us formulate the column sampling matrix $S_c \in \mathcal{R}^{n \times c}$ where $(S_c)_{i,j} = 1$ if the $i^{th}$ column of $H$ is chosen in the $j^{th}$ independent trial and zero otherwise. The re-scaled diagonal matrix is represented as $D_c \in \mathcal{R}^{c \times c}$ with its elements chosen as $(D_c)_{tt} = 1/\sqrt{cq_{j_t}}$, where $j_t \in [1, \ldots, n]$ in the $t^{th}$ sampling trial. Similarly, $S_d \in \mathcal{R}^{d \times m}$ and $D_d \in \mathcal{R}^{d \times d}$ are formulated. Hence, the followings can be inferred:

$$C = HS_cD_c, \text{ and } R = D_dS_dH \tag{3.21}$$

Thus, keeping the probability of selection and number of trials constant, $P \in \mathcal{R}^{d\times c}$ is formulated as shown:

$$P = D_d S_d C = D_d S_d H S_c D_c \tag{3.22}$$

Then a matrix $Q \in \mathcal{R}^{c\times c}$ can be constructed which integrates the top $k$ singular values of $C^T C$ ($\sigma_{t'}^2(C),\ t' = 1, \ldots, k$) with their corresponding singular vectors ($y^{t'},\ t' = 1, \ldots, k$) [240].

$$Q = \sum_{t'=1}^{k} \frac{1}{\sigma_{t'}^2(C)} y^{t'} y^{t'^T} \tag{3.23}$$

Finally, using $P$ and $Q$, $U \in \mathcal{R}^{c\times d}$ is computed as $U = QP^T$. For the CUR decomposition, an optimal choice of probabilities of sampling the columns $[q_j]_{j=1}^{n}$ and rows $[p_i]_{i=1}^{m}$ of the topology matrix can be defined as follows [242, 243]:

$$p_i = |H_{(i)}|^2 / ||H||_F^2 \quad \text{and} \quad q_j = |H^{(j)}|^2 / ||H||_F^2 \tag{3.24}$$

The CUR decomposition algorithm develops the set of optimal probabilities in one pass over the topology matrix $H$ with $\mathcal{O}(c+d)$ additional time and space [240]. In the following pass, the matrices $C,\ R,\ P,\ Q$ are formulated with an additional time and space of $\mathcal{O}(mc),\ \mathcal{O}(nd),\ \mathcal{O}(cd)$ and $\mathcal{O}(c^2 k)$, respectively. An additional space of $\mathcal{O}(mc)$ and additional time of $\mathcal{O}(mc^2)$ is required for computing $C^T C$ from $C$, while computing SVD of $C^T C$ takes an additional time of nearly $\mathcal{O}(c^3)$. $U$ is computed with an additional time of $\mathcal{O}(c^2 d)$. As $c,\ d,\ k$ are kept constant, the algorithm takes an overall additional time and space of $\mathcal{O}(m + n)$ [240]. An overview of the demonstrated approach can be seen in [239, 240]

It can be seen that the proposed low-rank attack vector formulation scheme using CUR decomposition technique can intrinsically bypass the residue test as employed by the BDD under varying attack strength and NASV. It must be noted that 500 Monte Carlo trials were used in this case. Such an approach performs well when large-scale noises and outliers are not prevalent within the topology matrix, thus furnishing an accurate determination of $H$ by the intruder. In case of the presence of large-scale noises and outliers within the topology matrix, matrix separation scheme (Go-Dec) as shown below is undertaken by the attacker for defining the robust low-rank subspace.

**Algorithm 3:** CUR Decomposition

**Input:** Topology matrix: $(H \in \mathcal{R}^{m \times n})$, Number of randomly selected rows: $d\ (1 \leq d \leq m),\ d \in \mathbb{N}$, Number of randomly selected columns: $c\ (1 \leq c \leq n),\ c \in \mathbb{N}$, Low-rank: $k\ (1 \leq k \leq min(d, c)),\ k \in \mathbb{N}$, Row sampling probability: $[p_i]_{i=1}^{m},\ \forall\ p_i \geq 0,\ \sum_{i=1}^{m} p_i = 1$, Column sampling probability: $[q_j]_{j=1}^{n},\ \forall\ q_j \geq 0,\ \sum_{j=1}^{n} q_j = 1$

**Output:** $C \in \mathcal{R}^{m \times c},\ U \in \mathcal{R}^{c \times d},\ R \in \mathcal{R}^{d \times n}$, Low-rank matrix approximate $H_k = CUR$

**for** *t = 1 to c* **do**
  Choose $j_t \in [1, \ldots, n]$ with probability $[j_t = \beta] = q_\beta,\ \beta = [1, \ldots, n]$
  Define $C^{(t)} = H^{(j_t)} / \sqrt{c q_{j_t}}$
**Compute:** $C^T C$ and its SVD; $C^T C = \sum_{t'=1}^{c} \sigma_{t'}^2(C) y^{t'} y^{t'^T}$
**If** for any $k$, $\sigma_k(C) = 0$, then $k = max[k' : \sigma_{k'}(C) \neq 0]$
**for** *t = 1 to d* **do**
  Choose $i_t \in [1, \ldots, m]$ with probability $[i_t = \beta] = p_\beta,\ \beta = [1, \ldots, m]$
  Define $R_{(t)} = H_{(i_t)} / \sqrt{r p_{i_t}}$
  Define $P_{(t)} = C_{(i_t)} / \sqrt{r p_{i_t}}$
**Define:** $Q = \sum_{t'=1}^{k} \frac{1}{\sigma_{t'}^2(C)} y^{t'} y^{t'^T}$ and $U = QP^T$
**Return**: $H_k = CUR$ and $C, U, R$

### 3.3.3 Go-Dec based low-rank approximation

To achieve a faster low-rank approximation, Go-Dec has already demonstrated its efficacy [245]. It incorporates bilateral random projections to compute the low-rank subspace instead of computing the orthogonal vectors corresponding to its row and column space. Such kind of projections makes this algorithm more robust to noise and outliers present in the topology matrix which can be decomposed as:

$$H = H_k + S + G \tag{3.25}$$

$$\text{Constraints: } rank\ (H_k) \leq k,\ card\ (S) \leq k' \tag{3.26}$$

where $S \in \mathcal{R}^{m \times n}$ represents the sparse matrix containing the outliers while $G \in \mathcal{R}^{m \times n}$ denotes the dense matrix comprising of noise in the topology matrix $H$. The aforesaid matrix separation scheme is formulated under the aforementioned constraints where $k'$

represents the cardinality [$card(.)$] of the sparse matrix $S$ i.e. the number of non-zero elements in $S$. The aforementioned matrix separation scheme is formulated as the following optimization problem:

$$\begin{aligned} \min_{H_k,S} \quad & ||H - H_k - S||_F^2 \\ \text{s.t.} \quad & rank\ (H_k) \leq k,\ card\ (S) \leq k' \end{aligned} \tag{3.27}$$

The above matrix separation problem is solved iteratively by solving the following two sub-problems until convergence [245]:

$$H_{k,[i]} = \underset{rank\ (H_k) \leq k}{\operatorname{argmin}} \quad ||H - H_k - S_{[i-1]}||_F^2 \tag{3.28}$$

$$S_{[i]} = \underset{card\ (S) \leq k'}{\operatorname{argmin}} \quad ||H - H_{k,[i]} - S||_F^2 \tag{3.29}$$

Although the aforementioned sub-problems have nonconvex constraints, still their global solution exists [245]. To improve the convergence speed of the algorithm for topology matrices having slow decay of singular values, power scheme modification is adopted.

$$\hat{H} = (HH^T)^q H,\ q > 0 \tag{3.30}$$

$$Y_1 = \hat{H} A_1,\ Y_2 = \hat{H}^T A_2 \tag{3.31}$$

where $\hat{H} \in \mathcal{R}^{m \times n}$ represents the topology matrix as developed after the power scheme modification. $Y_1 \in \mathcal{R}^{m \times k}$ denotes the random projection in the column space of $\hat{H}$ with the aid of a Gaussian random matrix $A_1 \in \mathcal{R}^{n \times k}$. $Y_2 \in \mathcal{R}^{n \times k}$ denotes the right random projection in the row space of $\hat{H}$ with the help of the matrix $A_2 \in \mathcal{R}^{m \times k}$ which is updated using the left random projection in the column space as:

$$A_2 = Y_1 \tag{3.32}$$

The low-rank matrix so formulated by the attacker using such bilateral random projections can be defined as:

$$\hat{H}_k = Y_1 (A_2^T Y_1)^{-1} Y_2^T \tag{3.33}$$

To obtain the low-rank approximate $H_k$ from $\hat{H}_k$, QR decomposition is applied as shown:

$$Y_1 = Q_1 R_1 \quad Y_2 = Q_2 R_2 \tag{3.34}$$

$$H_k = Q_1 [R_1 (A_2^T Y_1)^{-1} R_2^T]^{\frac{1}{2q+1}} Q_2^T \tag{3.35}$$

$\mathcal{P}_\Omega(.)$ represents the projection of the matrix over the set $\Omega$. $q$ can be so chosen to reduce the error in bilateral random projections as shown in algorithm 4. For an iterative update of (3.29), the first k largest entries of $|H - H_{k,[i]}|$ with nonzero elements are assigned to $S_{[i]}$ in the same position as follows:

$$S_{[i]} = \mathcal{P}_\Omega(H - H_{k,[i]}) \tag{3.36}$$

An overview of the demonstrated Go-Dec matrix separation algorithm can be defined as per algorithm 4. It is seen that subspace methods can define attack vectors that re-

**Algorithm 4:** Go-Dec Algorithm

**Input:** topology matrix : $(H \in \mathcal{R}^{m\times n})$, rank of the low-rank matrix : $k$, Cardinality (non zero entries) of sparse matrix : $k'$, Power scheme modification parameter : $q$, Convergence criteria : $\epsilon$

**Output:** Low-rank matrix approximation : $H_k \in \mathcal{R}^{m\times n}$, Sparse matrix : $S \in \mathcal{R}^{m\times n}$

**Initialize:** $H_{k,[0]} = H, S_{[0]} = 0, i = 0$

**while** $||H - H_{k,[i]} - S_{[i]}||_F^2 / ||H||_F^2 \leq \epsilon$ **do**

- $i = i + 1$
- $\hat{H}_k = [(H - S_{[i-1]})(H - S_{[i-1]})^T]^q (H - S_{[i-1]})$; `// Power scheme modification`
- $Y_1 = \hat{H}_k A_1$, $A_2 = Y_1$;
- $Y_2 = \hat{H}_k^T Y_1 = Q_2 R_2$; `// QR decomposition`
- $Y_1 = \hat{H}_k Y_2 = Q_1 R_1$ ; `// QR decomposition`
- **if** $rank\,(A_2^T Y_1) < k$ **then**
  - $k = \text{rank}\,(A_2^T Y_1)$, go to the first step
- **end if**
- $H_{k,[i]} = Q_1 [R_1 (A_2^T Y_1)^{-1} R_2^T]^{\frac{1}{2q+1}} Q_2^T$; `// QR decomposition is used to get` $H_k$ `from` $\hat{H}_k$
- $S_{[i]} = \mathcal{P}_\Omega(H - H_{k,[i]})$ ; `//` $\Omega$ `is the subset comprising of the first k largest entries of` $|H - H_{k,[i]}|$ `with nonzero elements`

**end while**

**Return**: $H_k$, $S$, $H - H_k$

main hidden in the system subspace and is capable of imposing critical scenarios on the grid [88,91]. Such an attack strategy furnishes a stealthy unobservable attack on the subspace defined by the covariance matrix of the acquired measurements, hence promoting a data-driven approach. It is seen that for such an attack vector formulation scheme, with increasing strength of the attack, the probability of detection using conventional BDD increases. Furthermore, there is a deterioration in the performance of such attacks when only a few meters with their corresponding measurements are accessible. Furthermore, with limited access to measurements, the estimated measurement subspace becomes erroneous. The error in the estimated eigenvectors of the measurement covariance matrix is seen to be inversely proportional to the eigenvalues. Hence, an erroneous estimation of the system subspace will lead to larger residuals leading to a higher probability of detection of attack using the conventional BDD.

For the linear state estimation model the covariance matrix of the acquired measurements ($C_1 \in \mathcal{R}^{m\times m}$) can be depicted as:

$$C_1 = (Z - \mu[z_t])(Z - \mu[z_t])^T \tag{3.37}$$

where, $Z = [z_1, z_2, \dots z_t] \in \mathcal{R}^{m\times t}$ denotes the acquired measurement matrix for a time period $t$ while the measurement vector at time $t$ is represented as $z_t$. $\mu(.)$ represents the mean operation. The theoretical expectation of the aforementioned covariance matrix can be defined as per (3.38) for the linear state estimation technique which can be defined as follows:

$$\mathbf{E}(C_1) = HC_{1,x}H^T + \sigma_m^2 I_m \tag{3.38}$$

where, $\sigma_m^2$ represents the variance of the measurement noise while $I_m \in \mathcal{R}^{m\times m}$ represents the identity matrix of order $m$. The data-driven approaches like [88, 91] try to estimate the column space of $H$ from the covariance matrix of the acquired measurements using (3.38). Such an approach can only be justified when the covariance matrix for the current set of acquired measurements is nearly equal to the theoretical expectation of $C_1$ [$\mathbf{E}(C_1)$] as shown in (3.38). Such an assumption is prevalent only when $t >> m$. When $t \simeq m$, the expected value of the error vector of the estimation model fails to converge to $\sigma_m^2 I_m$ as shown in (3.38). Hence, it can be seen that with a limited access to measurements, an accurate subspace estimation using RSVD [91] and FSVD [88] schemes can not be

achieved, hence leading to higher residues and a higher probability of detection using BDDs.

Furthermore, the low-rank subspace of the topology matrix defines a subspace or span of the topology matrix with a reduced basis with which the attacker can design stealthy attack vectors that can bypass the residue test. Moreover, with a lesser set of independent topological information of the grid, the attacker can formulate a low-rank subspace which provides better subspace information in comparison to the measurement covariance subspaces as shown in [88,91], hence leading to an unobservable stealthy attack.

## 3.4 Results

In comparison to the benchmark data-driven FDIA formulation schemes like [88,91] which undertakes the acquired measurement subspace, it can be seen that with limited availability of measurement data, the proposed approach demonstrates an effective attack vector formulation scheme as shown in Fig. 3.2.

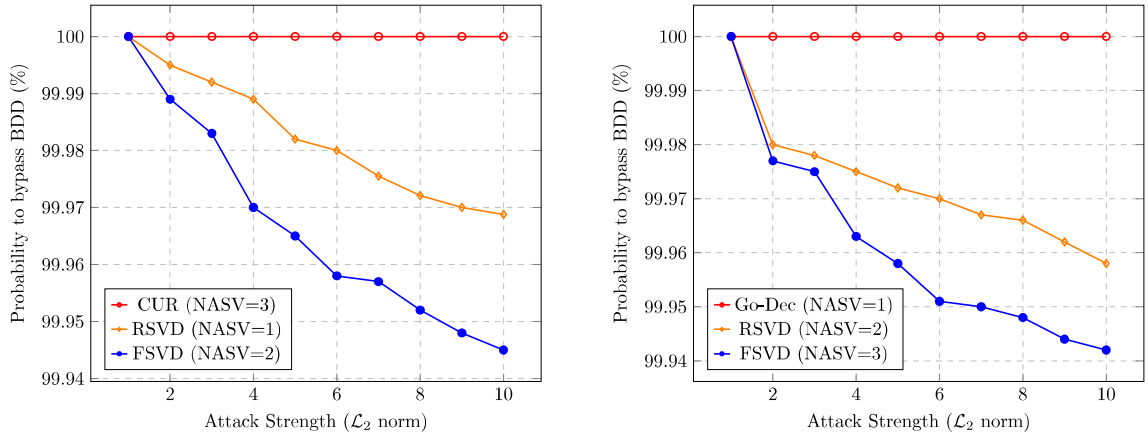


Figure 3.2: Probability to bypass bad data detector with varying attack strength (a) CUR (b) Go-Dec

It can be seen from Fig. 3.2 that attack vectors defined on the basis of the subspace of the measurement covariance matrix using FSVD [88] and RSVD models [91] can intrinsically bypass the residue test when the $\mathcal{L}_2$ norm of the attack vector is kept within the unit ball along with a minimal NASV. It can be inferred that with an increased attack strength and NASV, the probability of detection of attack using the conventional BDD increases.

The measurement samples are generally received at the control centre at a sampling rate of 100 Hz. Real-time measurements are acquired at nearly this sampling rate and state estimation techniques are invoked thereafter. As the mapping matrix does not take the measurements into account and only depends on the line connectivity and bus admittances, it can be seen that changing the frequency of estimation would not affect the defined low-rank subspace, and thus the attack vector. Hence, it can be inferred that changing the frequency of acquired measurements would not pose any significant changes in the attack vectors formulated. However, if parameter estimations are invoked from the solutions of the state estimates, then the mapping matrix would change. This will eventually change the full column and the low-rank subspace and hence the attack vector.

Furthermore, this thesis undertakes a steady-state operation of the grid with nearly constant loads. With a dynamic change in the system loads, the active and reactive power flows in the transmission lines would change significantly, but as the mapping matrix is defined using the admittances of each line and their corresponding connectivity, its entries would not change. This leads to a similar low-rank subspace. This on the other hand ensures that the attack vectors formulated using the full column space and the low-rank subspace remain similar. Hence, dynamic load changes can affect the state estimates significantly as they are based on the system measurements but not the attack vectors formulated. However, if parameter estimation methodologies are adopted that are based on state estimates, then dynamic conditions may lead to changes in the entries of the line admittances in the mapping matrix. This will eventually develop a different low-rank subspace. This ensures that the attack vectors developed on the basis of the low-rank subspace and the full column space under steady-state and dynamic conditions would differ. But as the formulated approach ensures that it can bypass the residue test, hence it can develop critical scenarios on the grid.

Hence, it can be seen that attack vector formulation over the low-rank subspace is a key research prospect. Thus, this chapter demonstrates an effective stealthy non-localized attack vector formulation on the low-rank subspace of the topology matrix using singular value decomposition, CUR decomposition, and robust matrix separation schemes. Moreover, it can be seen from [241] that SVD is vulnerable to large-scale noises and outliers. Under such circumstances, Go-Dec demonstrates a robust performance.

### 3.4.1 Computational burden

It can be seen from Fig. 3.3 that Go-Dec demonstrates a faster low-rank approximation than SVD using bilateral random projections. Furthermore, it is seen from Fig. 3.3 that Go-Dec and CUR showcase a satisfactory performance under the presence of noise as well, thus promoting a robust attack vector formulation scheme with the minimal computational burden. To effectively promote the efficacy of the proposed scheme in

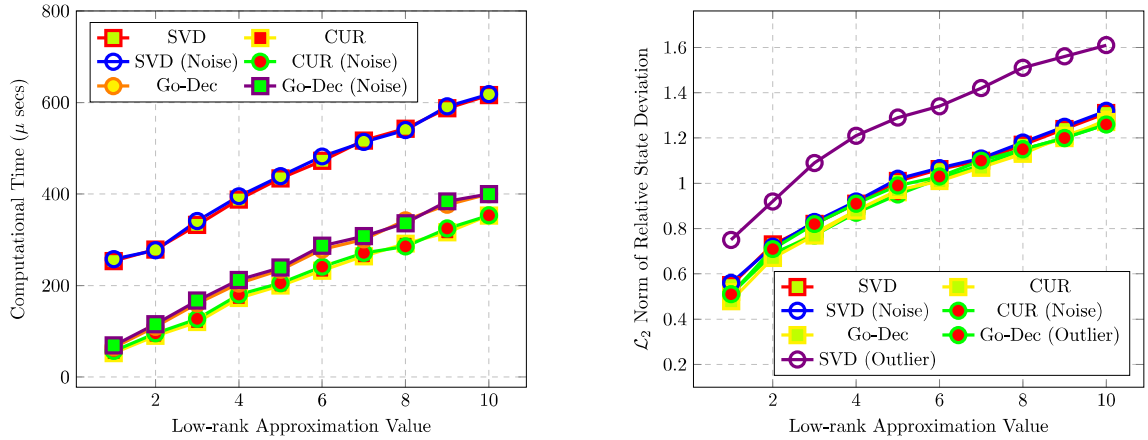


Figure 3.3: (a)Computational burden of the algorithms (b) Relative state deviation ($||c''||_2/||c'||_2$)

case of outliers and large-scale noises, this work has also incorporated them within the topology matrix. It can be seen from Fig. 3.3 that with the incorporation of outliers within the topology matrix, the $\mathcal{L}_2$ norm of the relative state deviation ($||c''||_2/||c'||_2$) increases for the low-rank structure as computed by SVD. As SVD performs poorly under outliers, hence to define the state deviation vector in the null space of $(H_k - H)$, a higher $\mathcal{L}_2$ norm is required. The low-rank structure as computed from Go-dec incorporates a power parameter ($q$) as 2, while $\epsilon$ is kept as 0.0001 while $c$ and $d$ are chosen as 10 and 20, respectively.

Moreover, the box plot as shown in Fig. 3.4 demonstrate the distribution of the measurement residuals after attack with a varying low-rank approximation and NASV for a net simulation time of 20 secs using CUR decomposition when accurate knowledge of the topology matrix is available while the distribution of the measurement residuals for Go-Dec for a net simulation time of 10 secs and with varying NASV is demonstrated in Fig. 3.5. It can be seen that the obtained residuals are nearly similar to that of the residuals

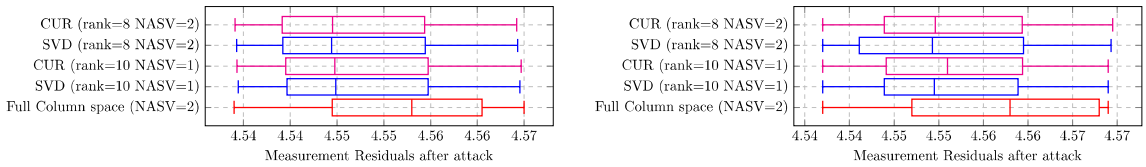


Figure 3.4: Measurement residuals after attack using CUR decomposition (a) under ideal conditions (b) with noise

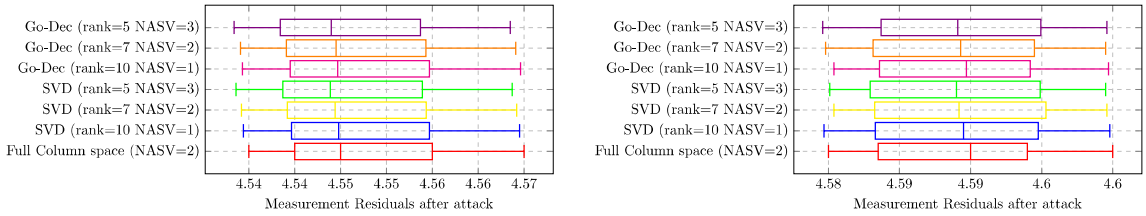


Figure 3.5: Measurement residuals after attack using Go-Dec (a) under ideal conditions (b) with noise

when the attack is defined using the full column space of the topology matrix under both ideal and noisy conditions. As shown in the prevalent attack vector formulation scheme which can effectively bypass the residue test, the measurement residuals are very much smaller than the bad data detection threshold $\tau$. It can be seen from Figs. 3.4 and 3.5 that the distribution of the measurement residuals is very much lesser than the detection threshold for attacks defined in the full column space and for the proposed low-rank attack vector formulation scheme. The measurement residuals lie in the range of 4.54-4.57 and 4.54-4.6 under ideal and noisy conditions respectively while $\tau$ is selected to be 31.32. This clearly demonstrates that the proposed scheme can bypass the residue test efficiently. Furthermore, to effectively demonstrate the performance of the algorithms under noise, Gaussian noise with mean zero and standard deviation one has been incorporated within the topology matrix. It is seen that both SVD and Go-Dec demonstrate satisfactory performance. SVD can handle small-scale noises but performs poorly when outliers are present in the topology matrix [241] while Go-Dec demonstrates a robust performance under all possible scenarios [245].

## 3.5 Summary

Hence, it can be seen that both under ideal and noisy conditions, the developed low-rank approximation algorithms (CUR, Go-Dec) can define an accurate low-rank subspace which inherently leads to stealthy attack vectors. Such low-rank attack vector formulation schemes can be implemented to test the prevalent FDIA detection algorithms. The following chapters develop some novel detection techniques that are capable of identifying the traditional attack vectors efficiently.