# Matrix-Completion-Based False Data Injection Attacks Against Machine Learning Detectors

Bo Liu, *Member, IEEE*, Hongyu Wu, *Senior Member, IEEE*, Qihui Yang, Hang Zhang, *Student Member, IEEE*, Yajing Liu, *Member, IEEE*, and Yingchen Zhang, *Senior Member, IEEE*

***Abstract*—False data injection (FDI) attacks can manipulate power system measurements, leading to system economic losses and security issues. Although machine-learning (ML) detectors can effectively detect FDI attacks, the current methods used to construct FDI attacks do not take into account the presence of ML detectors. To tackle this problem, we propose novel convex matrix-completion-based FDI (MC-FDI) attacks on DC and AC power flow models from an attacker's perspective, accounting for the temporal correlation between compromised and historical measurements. The proposed attacks minimize the nuclear norm of the compromised measurement matrix to make the compromised measurement consistent with the historical measurements, and also maximize the L1-norm of the incremental voltage angle to ensure a sufficient negative impact on the power system operation. Moving target defense (MTD) is proposed to detect the proposed MC-FDI attacks from the defender's standpoint. The idea is to actively change the line impedance to corrupt the spatial and temporal correlation of the compromised measurements in the MC-FDI attacks. Numerical results on the IEEE 14-bus and IEEE 118-bus systems show the stealthiness of the proposed attacks to both the Chi-square detector and ML detectors as well as the efficacy of MTD in detecting the MC-FDI attacks.**

***Index Terms*—False data injection, matrix completion, machine learning detector, moving target defense, state estimation.**

## I. Introduction

THE SMART grid integrates information and communication technology (ICT) enabled devices and Internet of Things (IoT) technologies to enable the transition to decarbonization and electrification. However, these devices also bring vulnerabilities to the cybersecurity of the smart grid. The U.S. Department of Energy received 362 power interruption reports related to cyber-physical attacks between 2011 and 2014 [1]. As significant threats to modern power systems, these attacks can undermine or even disrupt the control system of power grids, potentially resulting in tremendous economic loss and severe consequences.

The False Data Injection (FDI) attack is considered one of the most critical cyberattacks that can occur against smart grids due to its high-consequence nature. These attacks are intended to manipulate the state estimation (SE) results obtained by grid operators, thereby creating a significant risk to the grid's overall security and reliability. The elaborately constructed attack vector on the Supervisory Control and Data Acquisition (SCADA) measurements can bypass bad data detection by keeping consistent with physical laws like Kirchhoff's circuit laws. Since the power system state estimation is the basis of many power system operation applications in the energy management system (e.g., contingency analysis and economic dispatch), FDI attacks can result in serious consequences, such as economic loss, unstable system states, and even voltage collapse led to widespread blackouts [2]. The 2015 Ukraine cyberattack took over six months of infiltration and was successful in compromising the SCADA system and de-energizing a portion of the grid for a few hours [3]. In the Ukraine blackout, the attackers loaded malicious firmware into field gateway devices of the SCADA network to ensure that remote commands could not be issued to bring the substations back online, even when operator workstations were recovered [3]. Liang et al. argued that the circumstances of the Ukraine blackout highlight the plausibility of common assumptions in FDI attacks regarding the knowledge and capabilities required by an adversary [4].

Traditionally, state estimators obtain grid measurements from remote terminal units (RTUs), which measure the voltage magnitudes, power injections, and power flows. Most of the FDI attacks in the literature focus on RTU-based state estimation. Jin et al. analyzed the vulnerability of the power system AC-SE against FDI attacks, and demonstrated an attacker can plan a stealthy FDI attack in polynomial time with limited resources [5]. Different algorithms were proposed to construct stealthy FDI attacks by compromising the minimum number of sensors [6], [7]. Soltan et al. proposed joint cyber and physical attacks on the power grid based on a simplified AC power flow model [8]. Liang et al. introduced an FDI attack that can induce physical line overflows [9]. Yuan et al. proposed a special type of FDI, i.e., load redistribution attacks, which aim to cause load shedding in the power system [10]. In the modern power grid, Phasor Measurement Units (PMUs) are

Manuscript received 1 November 2022; revised 7 March 2023, 19 May 2023, and 9 August 2023; accepted 21 August 2023. Date of publication 28 August 2023; date of current version 21 February 2024. This work was supported in part by the U.S. National Science Foundation under Grant 1929147 and Grant 2146156, and in part by the United States Office of Naval Research under Grant N00014-23-1-2777. Paper no. TSG-01636-2022. *(Corresponding author: Bo Liu.)*

Bo Liu, Hongyu Wu, and Qihui Yang are with the Department of Electrical and Computer Engineering, Kansas State University, Manhattan, KS 66506 USA (e-mail: liubo1793@ksu.edu).

Hang Zhang is with the Department of Information Engineering, Henan University of Science and Technology, Luoyang, Henan, China.

Yajing Liu is with the Department of Mathematics, Colorado State University, Fort Collins, CO 80523 USA.

Yingchen Zhang is with the Product Solutions, Utilidata Inc., Providence, RI 02903 USA.

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/TSG.2023.3308339.

Digital Object Identifier 10.1109/TSG.2023.3308339

adopted to obtain a global view of the health of the grid [11]. PMUs can provide accurate and real-time synchronous phasor measurements with Global Positioning System (GPS) time sampled from geographically dispersed buses in the smart grid [12]. Pei et al. proposed a greedy placement algorithm of PMU devices to detect coordinated FDI attacks, in which the most vulnerable buses are protected [12]. However, PMUs are subjected to many vulnerabilities due to their dependency on external time sources for time synchronization and lack of strong authentication [13], [14]. Basumallik et al. investigated the vulnerability of PMU-based state estimation under the FDI attacks [15]. Alexopoulos et al. proposed a sparse FDI attack method against PMU-based state estimation using complementarity reformulation to compromise the minimum number of sensors [16].

However, there is no guide in selecting the malicious incremental voltage in the traditional FDI (TFDI) attacks, which can yield a distinct change of measurement values. These distinct measurement changes cause the FDI attacks to be easily detected by state-of-art anomaly detection methods, such as machine learning (ML) methods. ML methods have been proposed to detect FDI attacks in the smart grid by learning the spatial and temporal correlation of historical measurements. Supervised machine-learning-based binary classifiers were presented to check the distance between normal and compromised measurements [2]. Ozay et al. applied multiple supervised methods, including perceptron, k-nearest neighbor, support vector machine (SVM), and sparse logistic regression to detect FDI attacks [17]. Esmalifalak et al. first applied dimension reduction to the measurements and then utilized distributed SVM to classify the compromised measurements [18]. Sakhnini et al. tested three classification techniques using different heuristic feature selection techniques and concluded that the SVM and the k-nearest neighbor algorithms overperform artificial neural networks in detecting FDI attacks [19]. Semisupervised learning methods are also applied to detect FDI attacks, in which the information obtained from the unlabeled test samples is used for the learning models. The basic idea is to cluster the normal and compromised measurements into distinct regions in the feature spaces. Ozay et al. employed the semisupervised SVM algorithm to establish the relationship between supervised and semisupervised learning algorithms for detecting FDI attacks [17]. Esmalifalak et al. proposed the Gaussian abnormal detector to detect the deviation in measurements, and the outliers were identified as FDI attacks [18].

From the perspective of the attacker, the traditional FDI attacks have a problem because their construction methods only consider the spatial correlation of the historical measurement, not the temporal correlation, making them easily detectable by machine learning methods. Furthermore, there is a trade-off in selecting the malicious incremental voltage during constructing FDI attacks. An FDI attack with a large incremental voltage has a sufficient negative impact on the power system operation, but it is more likely to be detected by the ML detectors due to the distinct change of measurement values. Conversely, an FDI attack with a small incremental voltage could be stealthy to the ML detectors, but it has a trivial negative impact on the power system. In the literature, there is a research gap in developing a novel FDI attack that has a sufficient negative impact on the power system but remains stealthy to ML detectors.

Matrix completion (MC) is a promising technique to recover an intact matrix with low-rank properties from incomplete data [20]. MC applications include wireless communications, traffic sensing, integrated radar and communications, and power systems. Therefore it has received much attention in the past several years [21]. MC can learn the spatial or temporal correlation of incomplete data based on the low-rank property of matrices. If the matrix is formulated by the elements collected at a single time step, MC can recover the missing values by learning the spatial correlation of elements. Centralized and decentralized state estimation methods were proposed for the distribution system with low observability, which employed the conventional MC method augmented with noise-resilient power flow constraints [22], [23], [24]. If the matrix is formulated by the elements collected over multiple time steps, MC can estimate the missing elements by learning the temporal correlation of available elements. For example, MC was successfully applied to estimate missing phasor measurement unit (PMU) data [25] and detect bad data [26] using historical measurements.

This paper proposes novel FDI attack models which utilize matrix completion to determine the incremental voltage, such that the compromised measurements follow the temporal correlation of the historical measurements. Therefore, the proposed attack can remain stealthy to machine-learning attack detectors. In addition, the proposed attack models also maximize the incremental voltage to ensure a sufficient negative impact on the power system operation. The novelty of this paper is two-fold. From the attacker's perspective, this paper proposes novel FDI attack models which improve the stealthiness against ML detectors, while ensuring a sufficient negative impact on the power system operation. From the defender's perspective, we apply moving target defense (MTD) to detect the highly stealthy MC-FDI attacks. It is worth emphasizing that the goal of this paper is not to educate the attackers on how to construct highly crafted attacks, but to provide the grid defender a better understanding of how to design effective defense approaches against such new attacks.

- We propose a defense-attack framework consisting of a hybrid defense model and FDI attacks. The hybrid defense utilizes a model-based Chi-2 detector in conjunction with prevailing machine-learning detectors. To the best of our knowledge, this is the first piece of literature to focus on the development of FDI attacks designed to maintain stealthiness against a hybrid defense model.
- From the attacker's standpoint, we propose novel convex matrix-completion-based FDI (MC-FDI) attack algorithms in both DC and AC power flow models to optimize the malicious incremental voltage. The proposed attacks exhibit two distinctive features. Firstly, they utilize the MC technique to make the compromised measurements consistent with the temporal correlation of historical measurements. Secondly, the attacks use the TFDI model to enforce the spatial relationship of the compromised measurements. Furthermore, the proposed attacks can strike a

TABLE I
NOMENCLATURE

| Symbol | Definition |
|---|---|
| $\boldsymbol{\theta}$ | Voltage angle of buses excluding reference bus |
| $\mathbf{x}$ | System state vector |
| $\boldsymbol{a}$ | FDI attack vector |
| $\mathbf{z}$ | Measurement vector |
| $\mathbf{z}_a$ | Compromised measurement vector |
| $\mathbf{Z}_0$ | Historical measurement matrix |
| $\mathbf{Z}_a$ | Compromised measurement matrix |
| $\mathbf{H}$ | DC measurement matrix in state estimation |
| $b_{ij}$ | Susceptance of line $i$–$j$ (between bus $i$ and $j$) |
| $n$ | Total number of system buses |
| $m$ | Total number of measurements |
| $idx^t$ | Index array of time instants |
| $idx^{bus}$ | Index array of buses |
| $\mathrm{r}(\cdot)$ | Matrix rank operator |

balance between the attack's stealthiness against a hybrid defender and their malicious consequences on the power system operation, allowing the attackers to achieve their objectives while reducing the risks of being detected.

- From the perspective of the defender, we apply MTD, for the first time, in the physical layer of the power system to detect MC-FDI attacks by actively changing the system configuration. Our theoretical analysis proves that MTD can corrupt both the temporal and spatial correlation of compromised measurement in MC-FDI attacks. This represents a significant step towards detecting data-driven FDI attacks through the corruption of both temporal and spatial correlation of compromised measurements.

The rest of this paper is organized as follows. We provide preliminaries and related work in Section II. In Section III, we propose the DC- and AC- MC-FDI attack models. In Section IV, we propose MTD to detect the MC-FDI attacks. We conduct case studies in Section V. Conclusions are drawn in Section VI.

## II. PRELIMINARIES

In this section, we provide background knowledge of MC, SE, FDI attacks, and a machine-learning attack detector as preliminaries for the follow-up sections.

### A. Notation

Variables frequently used are summarized in Table I, where boldfaced lower-case and upper-case letters stand for vectors and matrices, respectively. From the attacker's perspective, subscript 0 denotes variables before attacks. For example, $\mathbf{z}_0$ and $\mathbf{z}_a$ stand for uncompromised and compromised measurement vector, respectively. From the defender's perspective, subscript 0 denotes variables before MTDs. For example, $\mathbf{H}_0$ represents the original measurement matrix before an MTD, and $\mathbf{H}_t$ stands for the one after the implementation of an MTD at time $t$. In addition, variables preceded by $\Delta$ represent changes in the variables. For example, $\Delta\mathbf{x}$ represents the malicious incremental voltage vector introduced by the attacker. This paper uses superscript $T$ to represent time instant and uses superscript $'$ to represent the transpose operator.

### B. Matrix Completion

Matrix completion technology aims to estimate the unknown elements in an incomplete matrix that has a low-rank property. Formally, let $\mathbf{M} \in \mathbb{R}^{n_1 \times n_2}$ be a real-valued data matrix to be recovered; let $\Psi \subseteq \{1, \ldots, n_1\} \times \{1, \ldots, n_2\}$ describe the index of the known elements in $\mathbf{M}$, and $\mathbf{M}_\Psi$ represents the observation matrix. The matrix completion problem can be formulated as a rank-minimization problem as follows:

$$\min_{\mathbf{D} \in \mathbb{R}^{n_1 \times n_2}} \ \mathrm{r}(\mathbf{D}) \quad s.t. \ \ \mathbf{D}_\Psi = \mathbf{M}_\Psi \tag{1}$$

where the decision variable $\mathbf{D}$ estimates the incomplete matrix $\mathbf{M}$. However, this problem is NP-hard due to the non-convexity of the rank function, and its solution algorithms are doubly exponential. Problem (1) can be modeled as a convex optimization problem by minimizing the nuclear norm using the convex relaxation technique [22].

$$\min_{\mathbf{D} \in \mathbb{R}^{n_1 \times n_2}} \ \|\mathbf{D}\|_* \quad s.t. \ \ \mathbf{D}_\Psi = \mathbf{M}_\Psi \tag{2}$$

where the nuclear norm $\|\mathbf{D}\|_*$ sums the singular values of $\mathbf{D}$. Problem (2) often has a unique minimizer $\mathbf{D}$ that equals $\mathbf{M}$, if there are sufficient randomly-sampled entries in $\mathbf{M}_\Psi$ [27].

### C. Power System State Estimation and FDI Attacks

In the DC-SE, nodal voltage angles, i.e., system states, $\mathbf{x} \in \mathbb{R}^{n-1}$are estimated by active nodal power injection (which can be positive or negative) and active branch power flow measurements $\mathbf{z} \in \mathbb{R}^m$. The measurement vector and states are related as $\mathbf{z} = \mathbf{H} \cdot \mathbf{x} + \mathbf{e}$, where $\mathbf{e} \in \mathbb{R}^m$ is the measurement noise assumed to be Gaussian distributed with zero mean, and a diagonal covariance matrix $\mathbf{W} = diag(\sigma_1^{-2}, \sigma_2^{-2}, \ldots, \sigma_m^{-2})$. DC-SE has a closed-form solution as follows: $\widehat{\mathbf{x}} = (\mathbf{H}'\mathbf{W}\mathbf{H})^{-1}\mathbf{H}'\mathbf{W}\mathbf{z}$. In the AC-SE, the system states $\mathbf{x} \in \mathbb{R}^{2n-1}$, i.e., nodal voltage angle and magnitude, are estimated by a set of measurements $\mathbf{z} \in \mathbb{R}^m$, including the power injection, power flow, and voltage magnitude measurements. The measurement vector and states are related as $\mathbf{z} = h(\mathbf{x}) + \mathbf{e}$, where $h(\cdot)$ is a vector of nonlinear function determined by the type of measurements. Since AC-SE doesn't have a closed-form solution, Gauss-Newton iterative algorithm is used to solve the following weighted least square problem: $min \ \ (\mathbf{z} - h(\mathbf{x}))'\mathbf{W}(\mathbf{z} - h(\mathbf{x}))$.

The Chi-2 detector is a widely used bad data detector (BDD) in DC-SE and AC-SE to detect bad data, as the estimation residual follows the Chi-2 distribution. When the system is free of bad data, the estimation residual is less than a preset threshold, i.e., $\gamma = ||\mathbf{z} - \mathbf{H} \cdot \widehat{\mathbf{x}}||_2 < \gamma_{th}$ in DC-SE and $\gamma = ||\mathbf{z} - h(\widehat{\mathbf{x}})||_2 < \gamma_{th}$ in AC-SE, where $\gamma_{th} = \chi^2_{(m-n),\alpha}$ is the threshold to ensure BDD has a false alarm rate at $1 - \alpha$.

An FDI attack aims to mislead the estimated states in the system operator's SE by injecting an attack vector $\mathbf{a}$ into the SCADA measurements, i.e., $\mathbf{z}_a = \mathbf{z} + \mathbf{a}$. In the DC-FDI attack [28], the estimation residual under attack in the defender's Chi-2 detector is zero in the noiseless condition,

i.e., $\gamma_a = ||(\mathbf{z}+\mathbf{a}) - \mathbf{H}\cdot(\mathbf{x}+\Delta\mathbf{x})||_2 = 0$, if the attacker calculates the FDI attack vector $\mathbf{a}$ as follows:

$$\mathbf{a} = \mathbf{H}\cdot\Delta\mathbf{x} \tag{3}$$

In the AC-FDI attacks, the estimation residual under attack in the defender's Chi-2 detector remains the same before and after FDI attacks, i.e., $\gamma_a = ||(\mathbf{z}+\mathbf{a}) - h(\mathbf{x}+\Delta\mathbf{x})||_2 = ||\mathbf{z} - h(\mathbf{x})||_2$, if the FDI attack vector $\mathbf{a}$ can be calculated by:

$$\mathbf{a} = h(\mathbf{x}+\Delta\mathbf{x}) - h(\mathbf{x}) \tag{4}$$

Since FDI attacks do not increase the estimation residual in the defender's Chi-2 detector compared with the situation free from attacks, FDI attacks remain stealthy to the defender's Chi-2 detector.

### D. Machine Learning Detectors

Machine learning methods have been used to detect FDI attacks based on the fact that normal data and compromised data tend to be separated in a certain projected space. In this paper, the SVM detector [18] is chosen to evaluate the stealthiness of the proposed FDI attacks for two reasons. First, it is the first machine learning method in the literature to demonstrate a decent capability to detect FDI attacks. Second, the SVM detector utilizes principal component analysis (PCA) as a preprocessing step, allowing for visualization during the attack detection. Traditional metrics such as precision, recall, and F1 score can be used to quantitatively measure the stealthiness of the proposed FDI attacks, while the visualization capabilities of the SVM detector enable qualitative demonstration of the proposed attack characteristics.

In the SVM detector, PCA is first applied to project the historical measurement data to a low-dimensional space. Dimension reduction solves the challenge brought by the high-dimensionality and redundancy of measurement data in practical power systems, and it is also beneficial for visualization. Then, the SVM method is proposed to detect stealthy FDI attacks [17], [18]. Generally, an SVM classifier constructs a hyperplane or a set of hyperplanes used for classification. Given a labeled training set $S = (x_l, y_l)$, $l = 1,\ldots,L$ of size $L$, with $y_l \in \{1, -1\}$ including both normal measurement data and compromised data, the SVM problem can be formulated by (5). The goal of SVM is to find the normal direction of a hyperplane $\boldsymbol{\omega}$ and parameter $b$ such that the prediction is correct for the most samples.

$$\begin{aligned} \min_{\boldsymbol{\omega},\xi,b}\ & \frac{1}{2}\boldsymbol{\omega}'\boldsymbol{\omega} + C\sum_{i=1}^{L}\xi_l \\ s.t.\ & y_i\left(\boldsymbol{\omega}'\phi(x_l) + b\right) \geq 1 - \xi_l \\ & \xi_l \geq 0, \quad l = 1,\ldots,L \end{aligned} \tag{5}$$

where $\phi(x_l)$ is a nonlinear transformation that maps $x_l$ in a higher dimensional space, and $\boldsymbol{\omega}'$ is the transpose of $\boldsymbol{\omega}$. The slack variable $\xi_l$ accounts for nonlinearly separable training sets, and $C$ is a tunable positive regularization parameter.

Artificial Neural Networks (ANN) consist of interconnected neurons organized in layers. The training process allows ANNs to learn complex patterns and relationships in the data, enabling them to make predictions or classifications [29]. Logistic Regression (LR) is a popular statistical modeling technique used for binary classification problems, which assigns probabilities to each outcome class [30]. Polynomial Logistic Regression extends the basic LR model by allowing for polynomial terms in the input variables, making it capable of capturing nonlinear relationships between the predictors and the response. Gaussian Naïve Bayes is a probabilistic machine learning algorithm based on Bayes' theorem that assumes a Gaussian (normal) distribution for the continuous-valued features. It calculates likelihoods for each class and combines them with prior probabilities to predict the output class [31]. These three ML methods are also used to evaluate the stealthiness of the proposed FDI attacks.

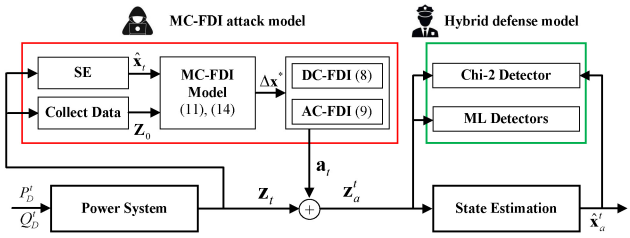



Fig. 1. The framework of MC-FDI attack against the hybrid defense model.

## III. Matrix Completion-Based FDI Attack

In this section, we will propose a defense-attack framework for MC-FDI attacks against a hybrid defense model and further define the capability and knowledge of the attacker in MC-FDI attacks. Finally, we will propose the mathematical model of MC-FDI attacks in the DC and AC power system models, respectively.

### A. Matrix Completion-Based FDI Attack Framework

We propose a defense-attack framework consisting of a hybrid defense model and an MC-FDI attacker, as shown in Fig. 1. In the framework, we define the hybrid defender as the system operator who utilizes both the model-based detector (Chi-2 detector) and ML detectors (such as SVM detector and ANN detector) to detect FDI attacks. We define the MC-FDI attacker as an attacker that utilizes the MC technology to remain attack stealthy to both the model-based and the ML detectors and ensure a sufficient negative impact on system operation. When the power system is under the load condition $P_D^t$ and $Q_D^t$ at time t, the SCADA measurement vector is $\mathbf{z}_t$. The MC-FDI attacker first collects SCADA measurements over time, and then conducts the SE to estimate $\widehat{\mathbf{x}}_t$. Based on the collected historical measurements $\mathbf{Z}_0$, the MC-FDI attacker utilizes MC-FDI models to calculate the optimal incremental voltage state $\Delta\mathbf{x}^*$ considering the characteristics of historical measurements. Finally, the traditional FDI models are utilized to calculate the malicious measurement $\mathbf{a}$ injected by the attacker.

The knowledge and capability of the MC-FDI attacker are summarized as follows: *i)* The attacker knows the grid topology and line parameters of the power systems. If the

**Algorithm 1** DC-MC-FDI Attack Algorithm

**Input:** Length of the historical measurement $T$, attacked buses $idx_a^{bus}$

**Output:** Compromised measurements $\mathbf{z}_a$

**Initialization:** A null historical measurement matrix $\mathbf{Z}_0 = \varnothing$
// Construct the historical measurement matrix $\mathbf{Z}_0$
**while** (the number of columns of $Z_0 < T$) **do**
  Eavesdrop SCADA measurements $\mathbf{z}_i$ at time $i$
  Add $\mathbf{z}_i$ to the last column of $\mathbf{Z}_0$, i.e.,$\mathbf{Z}_0 = [\mathbf{Z}_0,\ \mathbf{z}_i]$
  Keep eavesdropping SCADA measurements at time $i = i + 1$
**end while**
Run DC-MC-FDI model (11) to get $\Delta\mathbf{x}^*$
Calculate the compromised measurements according to (8)
**return** $\mathbf{z}_a$

**Algorithm 2** AC-MC-FDI Attack Algorithm

**Input:** Length of the historical measurement $T$, attacked buses $idx_a^{bus}$

**Output:** Compromised measurements $\mathbf{z}_a$ at time $t$

**Initialization:** A null historical measurement matrix $\mathbf{Z}_0 = \varnothing$
// Construct the historical measurement matrix $\mathbf{Z}_0$
**while** (number of columns of $Z_0 < T$) **do**
  Eavesdrop SCADA measurements $\mathbf{z}_i$ at time $i$
  Add $\mathbf{z}_i$ to the last column of $\mathbf{Z}_0$, i.e.,$\mathbf{Z}_0 = [\mathbf{Z}_0,\ \mathbf{z}_i]$
  Keep eavesdropping the SCADA measurements at time $i = i + 1$
**end while**
Apply state estimation to estimate the voltage at time $t$
Calculate the Jacobian matrix to linearize the AC-FDI attack model
Run AC-MC-FDI model (14) to get $\Delta\mathbf{x}^*$
Calculate the compromised measurements according to (9)
**return** $\mathbf{z}_a$

attacker only has limited information on grid topology and line parameters, the attack still can utilize Algorithms 1 and 2 to construct MC-FDI attacks using a modified **H** matrix and $h$(x) based on the principle in [32]; *ii)* The attacker has the write access to all measurements related to compromised states; *iii)* The attacker has the read access to all SCADA measurements, and can continuously eavesdrop on the SCADA measurements for a long time; and *iv)* The attacker knows the grid voltage in the AC power system model.

Regarding Assumption *i)*, the complete knowledge of the grid topology and line parameters is a common assumption in many prior works on FDI attacks [5], [7], [8], [9], [10], [16], [28], [33]. It is important to note that write access is a prerequisite for all FDI attacks, as stated in Assumption *ii)*. In FDI attacks, not all measurements need to be injected with malicious data, and only the measurements related to compromised states need to be injected. The attacker needs write access to all measurements related to compromised states. Assumption *iii)* of the read access is realistic, as evidenced by the Ukraine cyber-attack incident, where the attackers had long-term power system reconnaissance over six months or more without being noticed [3]. Since the SCADA layer communication network architecture has no cryptographically secure communication protocol, intercepting and forging communication messages is easy to achieve [4]. In addition, the existing blind FDI attacks [34], [35] assume that the attacker can collect historical measurements for a long period. Lakshminarayana et al. pointed out that the length of historical measurement should be more than 500 time instants in the IEEE 4-bus system to construct a stealthy blind FDI attack [36]. More historical measurements are needed to construct stealthy blind FDI attacks in a larger system. Assumption *iv)* of the grid voltage is a common assumption in most prior works on AC FDI attacks [5], [32], [33]. Hug and Giampapa first proposed this assumption in the AC-FDI attack model due to the nonlinear measurement-states relationship [33].

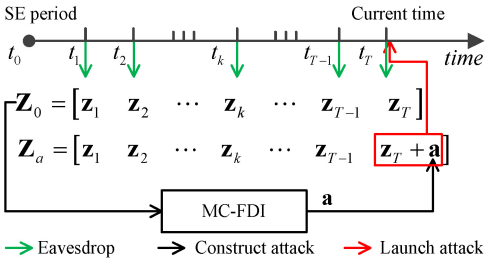



Fig. 2. Time sequence of measurement eavesdropping, attack construction, and attack injection.

We propose a DC-MC-FDI and an AC-MC-FDI attack algorithm, as shown in Algorithm 1 and Algorithm 2, respectively. The attacker first needs to decide the length of the historical measurement $T$ used in matrix completion and then decide on the attacked buses $idx_a^{bus}$ to manipulate their voltage. Since the matrix completion technology is used in the MC-FDI attacks, the attacker ought to construct a historical measurement matrix $\mathbf{Z}_0$ free of attacks in the first step of the MC-FDI attack. The attacker can continuously eavesdrop on the SCADA measurements for $T$ instants and then construct a historical measurement matrix $\mathbf{Z}_0$, as follows:

$$\mathbf{Z}_0 = \begin{bmatrix}\mathbf{z}_1 & \mathbf{z}_2 & \cdots & \mathbf{z}_{T-1} & \mathbf{z}_T\end{bmatrix} \tag{6}$$

where $\mathbf{z}_t \in \mathbb{R}^m$ is a vector of all SCADA measurements at time t, $\mathbf{Z}_0 \in \mathbb{R}^{m\times T}$ is composed of all SCADA measurements in the order of the time sequence, and $T$ represents the length of eavesdropping periods used in the historical measurement (the number of columns in $\mathbf{Z}_0$).

We present the time sequence of measurements eavesdropping, attack construction, and attack launch based on state estimation periods in Fig. 2. The attacker first keeps eavesdropping SCADA measurements until the attacker has $T$ historical measurement vectors. After collecting sufficient historical measurements, the attacker can launch an attack at any time. Here, we use $t_T$ to denote the current SE period when the attacker launches an attack. When the attacker decides to launch an attack on $t_T$, a malicious vector **a** calculated

by the MC-FDI model is injected into the SCADA measurement $\mathbf{z}_T$. Consequently, the compromised measurement matrix under attack can be defined as:

$$\begin{aligned}\mathbf{Z}_a\left(idx_a^t\right) &= \mathbf{Z}_0\left(idx_a^t\right) + \mathbf{a}\\ \mathbf{Z}_a\left(idx_0^t\right) &= \mathbf{Z}_0\left(idx_0^t\right)\end{aligned} \tag{7}$$

where $idx_a^t = \{T\}$, $idx_0^t = \{1,\ 2, \ldots,\ T-1\}$, and $\mathbf{a}$ is the malicious injection measurements determined by the malicious incremental voltage $\Delta\mathbf{x}$.

In the second step of MC-FDI attacks, MC-FDI models are proposed in Sections III-B and III-C to determine the optimal incremental voltage in the DC and AC power system models, respectively. Note that the voltage at the attack time $T$ needs to be estimated by the attacker through state estimation for constructing AC-MC-FDI attacks but is not necessary for DC-MC-FDI attacks. This voltage estimation requirement is the same as the construction of traditional AC-FDI attacks, according to (4).

In the last step of MC-FDI attacks, the attacker can calculate the malicious measurements at time $T$ using the optimal malicious incremental voltage based on the traditional FDI models. The compromised measurements in the DC-MC-FDI attack can be expressed as:

$$\mathbf{z}_a^T = \mathbf{z}_0^T + \mathbf{H}\cdot\Delta\mathbf{x}^* \tag{8}$$

Similarly, the compromised measurements in the AC-MC-FDI attack can be expressed as follows:

$$\mathbf{z}_a^T = \mathbf{z}_0^T + h\left(\mathbf{x}_T + \Delta\mathbf{x}^*\right) - h(\mathbf{x}_T) \tag{9}$$

### B. The DC-MC-FDI Attack Model

Since there is a trade-off between the attack stealthiness and the negative impact in the construction of FDI attacks against the machine learning detector, it is necessary to balance the trade-off in selecting the value of incremental voltage. Thus, we propose a novel MC-FDI model to calculate an optimal malicious incremental voltage in the FDI attacks, which considers the temporal correlation of the historical measurements and ensures sufficient negative impact on the power system operation. The MC-FDI model in the DC model is proposed in (10), which minimizes the nuclear norm of the compromised measurement matrix and maximizes the L1-norm of the malicious incremental voltage angle.

$$\min_{\Delta\mathbf{x}}\ \|\mathbf{Z}_a\|_* - \lambda\|\Delta\mathbf{x}\|_1 \tag{10}$$

$$\begin{aligned}
s.t.\ \ & \mathbf{Z}_a(i) = \mathbf{Z}_0(i) + \mathbf{a} && i \in idx_a^t && (10.1)\\
& \mathbf{Z}_a(i) = \mathbf{Z}_0(i) && i \in idx_0^t && (10.2)\\
& \mathbf{a} = \mathbf{H}\Delta\mathbf{x} && && (10.3)\\
& \Delta\mathbf{x}_{lb}(i) \le \Delta\mathbf{x}(i) \le \Delta\mathbf{x}_{ub}(i) && i \in idx_a^{bus} && (10.4)\\
& \Delta\mathbf{x}(i) = \mathbf{0} && i \in idx_0^{bus} && (10.5)
\end{aligned}$$

where $\mathbf{Z}_0(i)$ represents $i$-th column in $\mathbf{Z}_0$, and $\Delta\mathbf{x}(i)$ stands for the $i$-th element in the vector;$\lambda$ is the weight parameter; $idx_0^{bus}$ and $idx_a^{bus}$ is the index of buses free of attack and the index of attacked buses, respectively. Constraints (10.1) and (10.2) define the compromised measurement matrix under attack. Constraint (10.3) is the traditional DC-FDI attack model. Constraint (10.4) introduces lower bound $\Delta\mathbf{x}_{lb}$ and upper bounds $\Delta\mathbf{x}_{ub}$ to allow but limit the malicious incremental voltage angle in each of the attacked buses. Constraint (10.5) ensures that the buses in $idx_0^{bus}$ are free of attacks.

However, the proposed model is a non-convex optimization problem, as the objective function is a sum of a convex function $\|\mathbf{Z}_a\|_*$and a concave function $-\lambda\|\Delta\mathbf{x}\|_1$. By analyzing the problem structure, we transform the objective into the sum of a convex function $\|\mathbf{Z}_a\|_*$ and a plane $p'\Delta\mathbf{x}$. Specifically, we convexify the problem by introducing an attacker preference vector $p \in \mathbb{R}^{n\times 1}$. The elements in $p$ reflect the attacker's intention for each bus in the system. The attacker can set the value of$p(i)$ 1, $-1$, and 0 to decrease, increase and retain the voltage of Bus $i$, respectively. Multiple open-source packages can be used to solve the following convex problem (11), such as CVX [37] and CVXPY [38].

$$\begin{aligned}\min_{\Delta\mathbf{x}}\ & \|\mathbf{Z}_a\|_* + \lambda p'\Delta\mathbf{x}\\ s.t.\ & (10.1) - (10.5)\end{aligned} \tag{11}$$

Note that $\lambda$ is a positive weight parameter used to balance the trade-off between $\|\mathbf{Z}_a\|_*$ and $\|\Delta\mathbf{x}\|_1$. From the attacker's perspective, a larger $\lambda$ results in a larger incremental voltage $\|\Delta\mathbf{x}\|_1$ (a larger impact on power system operation), but a larger nuclear norm $\|\mathbf{Z}_a\|_*$ (indicating less following the temporal correlation of the historical measurements). The weight needs to be fine-tuned based on the specific system. Based on our numerical results, the range of $\lambda$ is usually between 1 and 20.

### C. The AC-MC-FDI Attack Model

We extend the DC-MC-FDI model to an AC-MC-FDI model, which integrates the traditional AC-FDI model (4) as a constraint. The differences between the DC-MC-FDI and AC-MC-FDI attacks are the essential differences between the DC and AC power system models. The decision variable$\Delta\mathbf{x}$ is composed of the incremental voltage magnitude and incremental voltage angle. The historical measurement matrix $\mathbf{Z}_0$ in the AC model contains more measurement types than those in the DC model. The most important difference is the nonlinear relationship $h(\cdot)$ between the measurements and states. However, the existence of nonlinear power flow constraints and power balance constraints makes the AC-MC-FDI model hard to solve. Therefore, we apply the first-order Taylor series expression on $h(\cdot)$ to linearize the relationship between the measurements and states at time $T$, as follows:

$$h(\mathbf{x}_T + \Delta\mathbf{x}) = h(\mathbf{x}_T) + \mathbf{H}(\mathbf{x}_T)\Delta\mathbf{x} \tag{12}$$

where $\mathbf{H}(\mathbf{x}_T) = \partial h(\mathbf{x}_T)/\partial\mathbf{x}$ is the Jacobian matrix of $h(\mathbf{x})$ at $\mathbf{x} = \mathbf{x}_T$. Note that the same Jacobian matrix is also needed in solving the AC weight least square (WLS) SE, which is a necessary step in the traditional AC-FDI attacks. Thus, the calculation of the Jacobian matrix in the AC-MC-FDI model is not an extra burden for the attacker. In this case, the injected measurements $\mathbf{a}$ can be expressed in a linearized formulation

using the Jacobian matrix:

$$\mathbf{a} = h(\mathbf{x}_T + \Delta\mathbf{x}) - h(\mathbf{x}_T) = \mathbf{H}(\mathbf{x}_T)\Delta\mathbf{x} \tag{13}$$

Therefore, we propose a convex AC-MC-FDI model in (14), which utilizes the linearized relationship between the measurements and states in the constraints. Note that we replace the voltage state $\mathbf{x}_T$ with $\widehat{\mathbf{x}}_T$ estimated by the attacker's SE, since $\mathbf{x}_T$ is unknown to the attacker.

$$\begin{aligned}
\min_{\Delta\mathbf{x}} \quad & \|\mathbf{Z}_a\|_* + \lambda p' \Delta\mathbf{x} \\
s.t. \quad & \mathbf{Z}_a(i) = \mathbf{Z}_0(i) + \mathbf{a} && i \in idx_a^t \\
& \mathbf{Z}_a(i) = \mathbf{Z}_0(i) && i \in idx_0^t \\
& \mathbf{a} = \mathbf{H}(\widehat{\mathbf{x}}_T)\Delta\mathbf{x} \\
& \Delta\mathbf{x}_{lb}(i) \le \Delta\mathbf{x}(i) \le \Delta\mathbf{x}_{ub}(i) && i \in idx_a^{bus} \\
& \Delta\mathbf{x}(i) = \mathbf{0} && i \in idx_0^{bus}
\end{aligned} \tag{14}$$

It is worth mentioning that this paper focuses on constructing MC-FDI attacks against state estimation with RTU measurements in transmission systems. FDI attacks against PMU-based state estimation can be effectively detected by a data-driven detector [26], which utilizes the temporal correlation of the historical synchronized measurements from the PMU devices. Therefore, from the perspective of the attacker, it is important to consider the temporal correlation of the synchronized measurements in the construction of FDI attacks against PMU-based state estimation. In the PMU-based state estimation, the synchronized measurements from the PMU devices $\mathbf{z} \in \mathbb{R}^m$ include the real and imaginary parts of voltage phasors and those of current phasors, and the state vector $\mathbf{x} \in \mathbb{R}^{2N}$ consists of the real and imaginary parts of $N$ bus voltage phasors. The PMU-based state estimation is a linear state estimator on Cartesian formulation and thus $\mathbf{z} = \mathbf{H}\mathbf{x}$ holds [16]. Since the linear relationship in the PMU-based state estimation is the same as that in the RTU-based state estimation in Section II-C, the DC-MC-FDI attack algorithm (Algorithm 1) can be directly applied to construct FDI attacks against PMU-based state estimation. However, the MC-FDI attacks against PMU-based state estimation is beyond the scope of this work and will be investigated in our future work.

## IV. Detection of MC-FDI Attacks

Since the proposed MC-FDI attacks are designed to be stealthy to the hybrid defense model, it is necessary to provide extra defense for the power systems. We propose to apply MTD in the physical layer of power systems to enhance the capability to detect MC-FDI attacks. In MTD, the system operator frequently and actively changes the transmission line reactance using the distributed flexible AC transmission system (D-FACTS) devices. The varying system configuration increases the barriers for the attackers to launch the MC-FDI attacks since the knowledge of the power system is one requirement for MC-FDI attacks.

### A. MTD Planning Model and Operation Model

MTD in power systems is composed of two essential steps, i.e., MTD planning and MTD operation. An MTD planning scheme determines a set of transmission lines to install D-FACTS devices, while an MTD operation delicately determines the D-FACTS setpoints under varying load conditions [39]. It has been proved that MTD planning largely determines the MTD detection effectiveness in the noiseless condition [40], and MTD operation slightly influences the MTD detection performance in noisy conditions [41].

The graph-based planning method can maximize the MTD detection effectiveness in the following two ways [41]. First, the graph-based planning method maximizes the rank of the composite matrix, the detection effectiveness metric of MTD, by eliminating the loops in the graphs composed of lines equipped with and without D-FACTS devices. Second, the graph-based planning method eliminates unprotected buses by covering all necessary buses with D-FACTS devices. Therefore, we adopt the graph-based planning method to determine the allocation of D-FACTS devices in MTD.

MTD operation determines the D-FACTS setpoints in real-time. The random MTD operation [42] is the simplest and most unpredictable MTD operation method, in which setpoints of each D-FACTS device are randomly selected based on uniform distribution within its operation range:

$$b_{ij} \sim U\Big((1-\eta)b_{ij}^0, (1+\eta)b_{ij}^0\Big) \tag{15}$$

where $\eta$ is the MTD magnitude that reflects the physical capability of D-FACTS devices; $b_{ij}^0$ is the original line susceptance; and $b_{ij}$ is the line susceptance modified by D-FACTS devices in MTD. As any other MTD operation method [43], [44] can be viewed as a subset of the random MTD, random MTD operation is generalized across all MTD operation methods. Without loss of generality, we use the random MTD operation under the graph-based planning to detect MC-FDI attacks. In the following subsections, we prove that MTD methods are able to corrupt both the temporal correlation and spatial correlation of compromised measurement in MC-FDI attacks. The conclusion drawn can be similarly extended to other MTD operation methods.

### B. Corruption of Temporal Correlation in MC-FDI Attacks

The MC-FDI model minimizes the nuclear norm of the compromised measurement matrix. The compromised measurement matrix in the DC noiseless condition can be reformulated in (16). As the measurement matrix $\mathbf{H}_0$ is a fixed matrix, the essential objective of the MC-FDI model is to minimize the nuclear norm of the compromised state matrix $\mathbf{X}_a$. Therefore, the MC-FDI model, in fact, optimizes the incremental state to follow the temporal correlations of the historical states.

$$\mathbf{Z}_a = \left[\mathbf{H}_0\mathbf{x}_1\ \mathbf{H}_0\mathbf{x}_2\ \cdots\ \mathbf{H}_0\mathbf{x}_{T-1}\ \mathbf{H}_0(\mathbf{x}_T + \Delta\mathbf{x})\right] = \mathbf{H}_0\mathbf{X}_a \tag{16}$$

where $\mathbf{X}_a = \left[\mathbf{x}_1\ \mathbf{x}_2\ \cdots\ \mathbf{x}_{T-1}\ (\mathbf{x}_T + \Delta\mathbf{x})\right]$ is the compromised state matrix.

MTD utilizes the random MTD operation method to change the susceptance of the lines identified by the graph-based planning method in each SE period. Accordingly, the measurement matrix in the state estimation at time t can be represented by $\mathbf{H}_t$. We define the uncertainties introduced by MTD for the attacker as the difference between the measurement matrix

before and after MTD, i.e., $\Delta\mathbf{H}_t = \mathbf{H}_t - \mathbf{H}_0$. In most MTD research in the cybersecurity of power systems, it is assumed that traditional attacker doesn't realize the existence of MTD deployed in the field [40], [41], [42], [44], [45]. It is possible that an attacker is aware of MTD defense techniques and knows D-FACTS devices are used in the power grid. Our recent work proposed three types of alert attackers who can detect the existence of MTD using bad data detection, unsupervised learning methods, or MTD operation models [46].

However, it is realistic to assume that the attacker doesn't have enough attack window to estimate the current system configuration under MTD, regardless of the traditional or alert attackers. This assumption holds based on two facts. First, the attacker may have the capability to track the changing system parameters resulting from MTD. However, the attacker must collect the historical measurements under this configuration for a long time period to estimate the current system configuration. For example, the length of historical measurements should be more than 500 time instants to accurately estimate the singular vector of the system configuration in the IEEE 4-bus system [36]. Much more historical measurements are required in a larger system. Second, as a proactive defense method, MTD's frequency can be determined by the defender. If the defender changes the system configuration more frequently, the attacker will have a shorter attack window to estimate the system configuration and launch attacks. Consequently, the attacker cannot collect enough historical measurements under the current system configuration in the short attack windows. The dynamic nature of MTD can make it more challenging for the attacker to identify the system parameters with insufficient historical measurements, as the MTD defense can make the attack surface more unpredictable and difficult to exploit.

If the attacker detects the existence of MTD but fails to accurately estimate the current system configuration under the MTD, a reasonable attacker will postpone launching FDI attacks. In other words, MTD prevents potential attackers from launching attacks, which is the advantage of proactive defense. To theoretically analyze the capability of MTD on detecting launched attacks, it is reasonable to assume that the attacker use the original system configuration without MTD ($\mathbf{H}_0$) to construct the MC-FDI attacks. This assumption is widely adopted in MTD works [40], [41], [42], [43], [44], [45], [46], [47] in the analysis of the detection effectiveness.

The compromised measurement matrix in the MC-FDI attack under the MTD can be expressed as (17).

$$\begin{aligned}\mathbf{Z}_a &= \left[\mathbf{H}_1\mathbf{x}_1\ \mathbf{H}_2\mathbf{x}_2 \cdots \mathbf{H}_T\mathbf{x}_T + \mathbf{H}_0\Delta\mathbf{x}\right] \\ &= \left[(\mathbf{H}_0+\Delta\mathbf{H}_1)\mathbf{x}_1\ (\mathbf{H}_0+\Delta\mathbf{H}_2)\mathbf{x}_2 \cdots (\mathbf{H}_0+\Delta\mathbf{H}_T)\mathbf{x}_T + \mathbf{H}_0\Delta\mathbf{x}\right] \\ &= \mathbf{H}_0\mathbf{X}_a + \left[\Delta\mathbf{H}_1\mathbf{x}_1\ \Delta\mathbf{H}_2\mathbf{x}_2\ \ldots\ \Delta\mathbf{H}_T\mathbf{x}_T\right] \\ &= \mathbf{H}_0\mathbf{X}_a + \Delta\mathbf{H}\odot\mathbf{X}_0 \end{aligned} \tag{17}$$

where $\mathbf{X}_0 = \left[\mathbf{x}_1\ \mathbf{x}_2 \cdots \mathbf{x}_{T-1}\ \mathbf{x}_T\right]$ is the state matrix, $\Delta\mathbf{H} = \left[\Delta\mathbf{H}_1\ \Delta\mathbf{H}_2\ \ldots\ \Delta\mathbf{H}_T\right]$ is the historical incremental $\mathbf{H}$ matrix introduced by MTD, and $\odot$ is the element-wise product for the submatrices in $\Delta\mathbf{H}$ and the columns in $\mathbf{X}_0$.

MTD corrupts the temporal correlation of the historical states in the MC-FDI model in two aspects. First, without considering any attacks, MTD corrupts the correlations of historical voltage $\mathbf{X}_0$. This is because the randomness introduced by MTD to the system configuration can cause irregular nodal voltage changes. Second, compared with (16), the MC-FDI model under the MTD no longer focuses on minimizing the nuclear norm of the compromised state matrix $\mathbf{X}_a$. The objective function of MC-FDI is also influenced by $\Delta\mathbf{H}\odot\mathbf{X}_0$ due to the varying system configurations in MTD. Thus, the compromised measurements calculated in MC-FDI attacks could not be consistent with the temporal correlation of historical measurements under the MTD.

### C. Corruption of Spatial Correlation in MC-FDI Attacks

The MC-FDI model takes the FDI model as constraints such that the compromised measurements satisfy the spatial correlation, i.e., the subject to the physical law of the power system. However, MTD can effectively break this spatial correlation of the compromised measurements. Assume that the attacker launches an MC-FDI attack at time $T$ using $\mathbf{H}_0$, and the actual measurement matrix is $\mathbf{H}_T$. Then, the compromised measurements of the MC-FDI attack at time $T$ is calculated as follows:

$$\mathbf{z}_a^T = \mathbf{H}_T\mathbf{x}_T + \mathbf{H}_0\Delta\mathbf{x}^* + \mathbf{e} \tag{18}$$

Note that the difference between (8) and (18) is $\mathbf{H}$ matrix. Under no MTD condition, the compromised measurements can be calculate by (8), where $\mathbf{z}_0^T = \mathbf{H}_0\mathbf{x}_T + \mathbf{e}$.

Under the noiseless condition, the estimation residual is zero, i.e., $\gamma_{MTD} = 0$, if and only if $\mathbf{H}_0\Delta\mathbf{x} \in col(\mathbf{H}_T)$ according to (3). As $\mathbf{H}_0 \neq \mathbf{H}_T$, the estimation residual is likely larger than zero, indicating the detection of MC-FDI attacks. Specifically, the estimation residual of the MC-FDI attack under the MTD can be expressed in (19). Essentially, MTD causes the MC-FDI attacker to use incorrect system configuration to calculate the attack vector, resulting in breaking the physical law of the power system, such as the imbalance of the nodal power injection in attacked buses. Therefore, MTD is able to corrupt the spatial correlation of the compromised measurement vector.

$$\begin{aligned}\gamma_{MTD} &= \left\| (\mathbf{H}_T\mathbf{x}_T + \mathbf{H}_0\Delta\mathbf{x}) - \mathbf{H}_T\left(\mathbf{H}_T^T\mathbf{H}_T\right)^{-1}\mathbf{H}_T^T(\mathbf{H}_T\mathbf{x}_T + \mathbf{H}_0\Delta\mathbf{x})\right\| \\ &= \left\| \left(I - \mathbf{H}_T\left(\mathbf{H}_T^T\mathbf{H}_T\right)^{-1}\mathbf{H}_T^T\right)\mathbf{H}_0\Delta\mathbf{x}\right\| \end{aligned} \tag{19}$$

## V. Numerical Results

### A. Test Systems and Simulation Setting

We perform numerical tests on the IEEE 14-bus system and the IEEE 118-bus system to evaluate the performance of MC-FDI attacks against the hybrid defense model and demonstrate the effectiveness of MTD in detecting MC-FDI attacks. The proposed MC-FDI attacks are modeled and solved by the CVX package in MATLAB [37]. The SE, Chi-2 detector, MTD, and TFDI attacks are all programmed in MATLAB. The algorithms are performed on a laptop with an Intel Core i5 processor CPU 2.70 GHz dual-core with 8 GB RAM. The measurements in the AC-SE include active and reactive power flow measurements, active and reactive power injection measurements, and voltage magnitude measurements. We

adopt a 2.5 redundant factor in the AC-SE. Specifically, we randomly select $2.5 \times (2N-1)$measurements among all possible measurements until the observability of the system is met.

We apply the hourly load profile of ERCOT [48] to the load buses in the IEEE 14-bus system, and the hourly load profile of WECC [49] to the load buses in the IEEE 118-bus system. Then, the power flow problem is solved by MATLAB MATPOWER in each power system. At each time instant, the SCADA measurements are collected from the solution of the power flow problem. These collected measurements serve as the normal (uncompromised) historical measurements in the training, validation, and testing dataset. We construct TFDI attacks with different attack magnitudes under multiple time instants, in which the voltage angles of randomly selected buses are compromised, and their incremental values are randomly selected according to their attack magnitude. These collected measurements under TFDI attacks serve as the compromised historical measurements in training, validation, and testing dataset. In addition, the proposed MC-FDI attacks are included in the testing dataset, since the objective of the experiments is to evaluate the stealthiness of the proposed MC-FDI attacks against ML detectors. In summary, the training dataset includes the normal measurements and measurements under TFDI attacks, and the testing dataset includes the measurements under the proposed attacks, measurements under the TFDI attacks for comparison, and the normal data to calculate the false positive (FP) rate.

The machine learning detectors, including SVM, logistic regression, artificial neural network, and Bayesian detector are trained and tested using the Sklearn package [50] in Python. 5-fold cross-validation is conducted in the training dataset to find the parameters of different ML detectors, such as the LR's penalization parameter, NB's probability threshold, and ANN's number of neurons in the hidden layer and the strength of the regularization. In the 5-fold cross-validation, the training dataset is randomly shuffled and then is equally split into five groups. For each unique group, we take the group as a validation set, take the remaining groups as a training data set, and then fit a ML detector on the training set and evaluate it on the validation set. Then, the average F1 score on the five validation sets is adopted as the metric to evaluate the performance of the ML detector under the given parameters.

### *B. Traditional FDI Attacks Against SVM Detector*

In this section, we first demonstrate the drawbacks of the TFDI attacks against the SVM detector and then show the importance of the nuclear norm of the historical measurement matrix in the construction of FDI attacks against the SVM detector. First, we evaluate the performance of TFDI attacks against the SVM detector under different attack magnitudes (AM), where the attack magnitude AM defines the range of incremental voltage, i.e., $\Delta \mathbf{x} \in [-\text{AM} \cdot \bar{\mathbf{x}}, \text{AM} \cdot \bar{\mathbf{x}}]$. Note that a larger attack magnitude reflects a larger selection range of the incremental voltage, but not necessarily ensures a larger incremental voltage angle due to the definition of the attack magnitude.

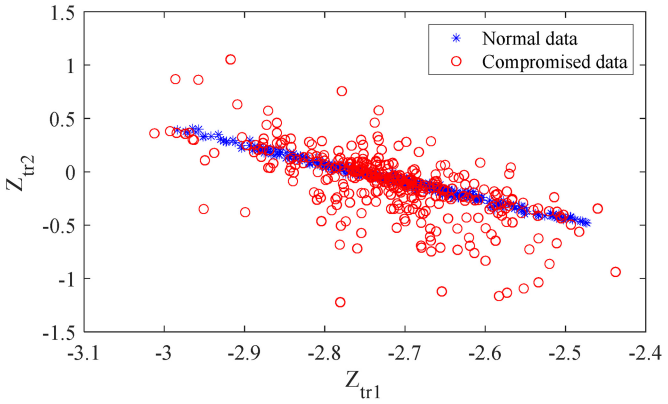



Fig. 3. Projection of normal and compromised data in the training set by PCA.

In the IEEE 14-bus system, the dimension of the data is 68. The training dataset includes 350 normal (uncompromised) measurement vectors and 350 compromised measurement vectors. The normal measurement vectors are collected from the SCADA system from the 1st time instant to 350th time instant. TFDI attacks randomly select three buses to compromise their voltage from the 100th time instant to the 300th time instant with 0.1-0.4 AMs. The compromised measurement vectors in the training set are sampled from the measurements under the TFDI attacks. The testing set includes 100 uncompromised measurement vectors collected from 351th time instant to 450th time instant, 200 compromised measurement vectors under MC-FDI attacks with $\lambda = \{3.1, 3.2\}$, and 500 compromised measurement vectors under TFDI attacks with AM=$\{0.01, 0.05, 0.1, 0.2, 0.3\}$. Specifically, we generate 100 attacks for each attack magnitude, and TFDI and MC-FDI attacks compromise the voltage of three buses. Note that five attack magnitudes of the TFDI attacks are selected based on the voltage situation of the IEEE 14-bus system, such that the incremental voltages in TFDI attacks are comparable to those in MC-FDI attacks. In the IEEE 14-bus system, the average value of L1-norm of the incremental voltage in MC-FDI attacks with two $\lambda$ weights are 0.057, and 0.140, respectively, and those in TFDI attacks with five AMs are 0.002, 0.008, 0.015, 0.029, and 0.049, respectively.

We apply the PCA dimension reduction to the training and testing dataset. With two principal components (convenient for the visualization), 99% of the signal variance will be retained. Figures 3, 4, and 5 show the projection of the training set, the projection of MC-FDI attacks in the testing set, and the projection of TFDI attacks with different attack magnitudes in the testing set, respectively. Note that the blue stars in Fig. 4 and 5 are the projection of the normal data in the training set, which serve as the reference.

Since the normal and compromised data are not linearly separable, the SVM detector with Gaussian kernel is applied to detect FDI attacks. The choice of kernel coefficient $\sigma$ and regularization parameter C can impact the efficiency of the SVM in detecting attacks. We train the SVM with different C and $\sigma$ values. Under each parameter pair, we conduct 5-fold cross-validation in the training dataset, and we adopt the average

TABLE II
AVERAGE F1 SCORE OF THE 5-FOLD CROSS-VALIDATION SET UNDER DIFFERENT C AND $\sigma$ VALUES

| C \ $\sigma$ | 0.3 | 1 | 3 | 10 | 30 | 100 | 300 |
|---|---|---|---|---|---|---|---|
| 0.3 | 0.50 | 0.50 | 0.52 | 0.57 | 0.59 | 0.66 | 0.71 |
| 1 | 0.50 | 0.53 | 0.55 | 0.64 | 0.65 | 0.74 | 0.77 |
| 3 | 0.51 | 0.54 | 0.61 | 0.67 | 0.72 | 0.76 | 0.79 |
| 10 | 0.54 | 0.56 | 0.65 | 0.71 | 0.75 | 0.79 | 0.80 |
| 30 | 0.55 | 0.63 | 0.67 | 0.73 | 0.78 | 0.80 | 0.80 |
| 100 | 0.57 | 0.66 | 0.71 | 0.77 | 0.80 | 0.82 | 0.78 |
| 300 | 0.61 | 0.68 | 0.74 | 0.79 | 0.81 | 0.81 | 0.76 |

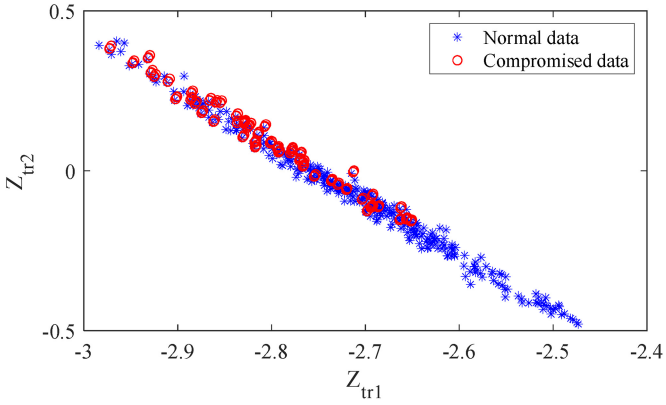



Fig. 4. Projection of MC-FDI attacks in the testing set by PCA.

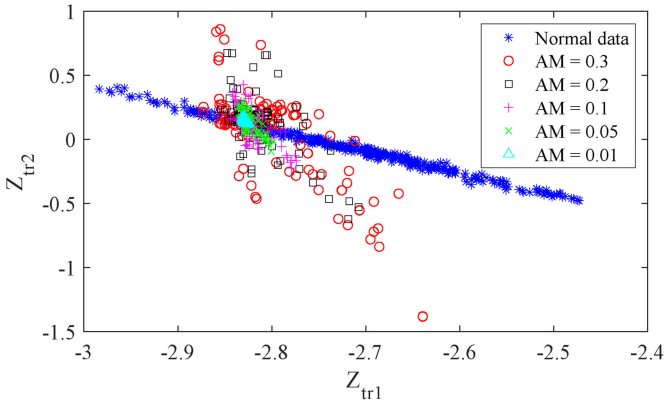
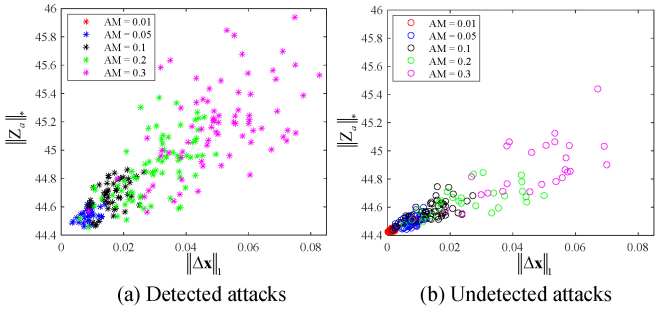



Fig. 5. Projection of TFDI attacks in the testing set by PCA.

of the F1 scores as the measure of accuracy. Specifically, the kernel coefficient $\sigma$ and regularization parameter C are optimized by searching in the set {0.01, 0.03, 0.1, 0.3, 1, 3, 10, 30, 100, 300}. Table II shows the F1 score of the cross-validation set under different C and $\sigma$ values. Thus, we set C=100 and $\sigma$ =100 in the SVM detector.

In Fig. 5, it is seen that most compromised data locate outside of the historical data area (blue star area), which can be treated as outliers (detected by the SVM detector). The outliers of the compromised data with a larger AM can be farther from the historical data area. This is consistent with the fact that a larger AM can drive the compromised data further deviated from the historical data. Accordingly, the compromised data with a smaller AM is more likely to remain inside the historical data area. However, we can observe that some compromised data with large AM is also located inside the historical data




(a) Detected attacks (b) Undetected attacks

Fig. 6. The nuclear norm and L1-norm of all TFDI attacks by SVM detector.

area. It is necessary to further investigate whether these FDI attacks inside the historical area have large AM but small incremental voltage.

We analyze the spatial and temporal correlation of each FDI attack to evaluate the performance of the SVM detector on detecting FDI attacks. Due to the definition of AM, AM cannot accurately and directly reflect the malicious modification on the measurements by attacks. Thus, we use the L1-norm of the incremental voltage angle ($\|\Delta \mathbf{x}\|_1$) as the metric for measuring the attack's strength on the spatial correlation. Then, we use the nuclear norm of the compromised historical measurement matrix as the metric for quantifying the attack's impact on the temporal correlation. The trained SVM detector is used to detect 500 TFDI attacks with different AMs. Then, we calculate the L1-norm and nuclear norm of 500 TFDI attacks. Finally, we project all detected TFDI attacks into $\mathbb{R}^2$ space in Fig. 6(a) and all undetected TFDI attacks in Fig. 6(b). We refer the $\mathbb{R}^2$ space as norm-norm space hereafter. Since TFDI attacks with small incremental voltage (less than 0.01 L1-norm value) slightly modify the measurements, these attacks have a very limited impact on the nuclear-norm values and thus are very likely to be stealthy to the SVM detector, as shown in Fig. 6(b). By comparing these two figures, we can see that for the TFDI attacks with large L1-norm values (more than 0.05 L1-norm value), most undetected TFDI attacks have comparatively lower nuclear norm values, and most detected TFDI attacks have comparatively higher nuclear norm values. Even though TFDI attacks with low nuclear norm values can also be detected, as shown in the lower right corner of Fig. 6(a), there are no undetected FDI attacks with high nuclear norm values, as shown in the upper right corner of Fig. 6(b).

We further equally divide the norm-norm space into 16 blocks and calculate the attack detection probability (ADP) of each block. The ADP of a given block is defined as the ratio of the number of detected attacks to the number of total attacks in the block. The heatmap of ADP in the norm-norm space is shown in Fig. 7. We can observe a low ADP in the low nuclear norm and low L1-norm value block (the lower left corner), a high ADP in the high nuclear norm and high L1-norm blocks (the upper right corner), and a low ADP in the high L1-norm but low nuclear norm blocks (the lower right corner). The lower right corner block and the upper left corner blocks are NaN, since no FDI attack falls into these blocks. Simulation results highlight the importance of temporal correlation (the nuclear norm of the historical

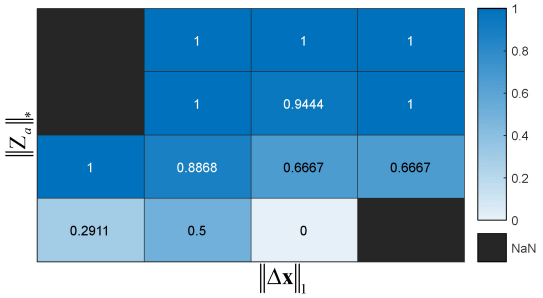
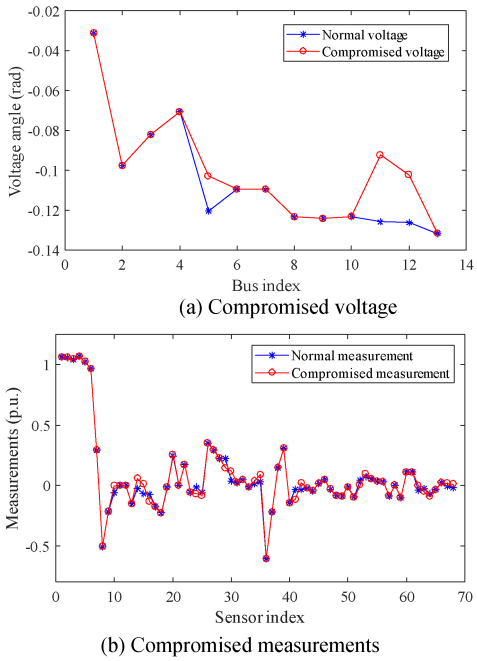



Fig. 7. ADP of the SVM detector against TFDI attacks.




(a) Compromised voltage

(b) Compromised measurements

Fig. 8. The MC-FDI attack on the 300th time instant.

measurement matrix), when the attacker aims to construct stealthy FDI attacks with large incremental voltage against the SVM detector.

Therefore, we can summarize the drawbacks of TFDI attacks: 1) there is no guide for selecting the incremental voltage; 2) TFDI attacks with small incremental voltage can be stealthy to the SVM detector, but they also have a low negative impact on the system; 3) attacks with large incremental voltage can be detected by the SVM detector without considering the temporal correlation of the historical measurements.

### C. Performance of MC-FDI Attacks

In this section, we evaluate the performance of MC-FDI attacks against the SVM detector. We assume the attacker utilizes 200 historical measurement vectors to construct the historical measurement matrix, intends to increase the voltage angle of Buses 6, 12, and 13, and adopts $\lambda = 3.1$. First, we show the compromised voltage angle and the compromised measurements in the 300th instant in Fig. 8. It is seen that the compromised voltage is very different from the normal voltage, and the attack manipulates the measurements related to the attacked buses. Even though the MC-FDI attack obviously manipulates the voltage, the MC-FDI attack doesn't yield a distinct change in measurement values. In Fig. 4, all MC-FDI attacks locate inside the normal data area. It indicates that the dimension reduction fails to separate the normal data and the compromised data. This is because the MC-FDI attacks consider the temporal correlation of the historical measurements by minimizing the nuclear norm of the historical measurement matrix. As the SVM detector can only detect the outlier, the proposed MC-FDI attacks ought to be stealthy to the SVM detector.

TABLE III
CPU TIME OF MC-FDI ATTACK WITH DIFFERENT NUMBER OF PERIODS IN THE HISTORICAL MEASUREMENT MATRIX

| Time (s) | T=50 | T=100 | T=200 | T=300 | T=400 |
|---|---|---|---|---|---|
| DC model | 4.3 | 16.8 | 97.7 | 116.9 | 299.8 |
| AC model | 8.1 | 27.3 | 110.9 | 262.6 | 488.8 |

The CPU time of the MC-FDI attacks using the different number of periods in the historical measurement matrix in the DC and AC model is summarized in Table III. The number of decision variables in the AC-MC-FDI attack is twice of those in the DC-MC-FDI attack. In addition, the size of the historical measurement in AC-MC-FDI attacks is larger than that in the DC-MC-FDI attack, since there are 34 measurements in the DC-SE and 68 measurements in the AC-SE. Thus, the CPU time of AC-MC-FDI attack is longer than that of the DC-MC-FDI attack. It is seen that the CPU time of solving the proposed MC-FDI models depends on the size of the historical measurement matrix. The CPU time greatly increases with the increasing number of historical measurement vectors in $\mathbf{Z}_a$. In order to reduce the CPU time and launch an FDI attack in time, it is suggested to reduce the number of the historical measurement vector in $\mathbf{Z}_a$. As shown in Fig. 4, 200 historical measurement vectors are sufficient to lead the MC-FDI attack stealthy to the SVM detector.

### D. Impact of Weights on the Performance of MC-FDI Attacks

In this section, we evaluate the impact of weights on the performance of MC-FDI attacks. Assume the attacker intends to increase the voltage angle of Buses 6, 12, and 13 in the 300th instant using 200 historical measurement vectors. We increase the weight $\lambda$ from 2.2 to 3.3 with an incremental of 0.05. First, we demonstrate the impact of weights on the L1-norm of incremental voltage and the nuclear norm of the compromised historical measurement matrix. We compare the MC-FDI attacks using different weights with the 500 TFDI attacks generated in the 300th instant in Section V-B. Specifically, we first combine the data points in Fig. 6(a) and Fig. 6(b) and then project the MC-FDI attacks into the norm-norm space, as shown in Fig. 9. With the increase of the weight, both the L1-norm of $\Delta\mathbf{x}$ and the nuclear norm of $\mathbf{Z}_a$ increase. Compared with TFDI attacks, the MC-FDI attacks always have the lowest nuclear norm value regardless of the L1-norm value. This comparison demonstrates the effectiveness of the proposed MC-FDI model in balancing the trade-off between maximizing the negative

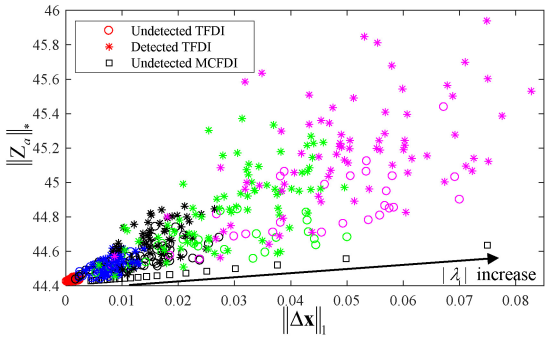



Fig. 9. Impact of weights on the performance of MC-FDI attacks.

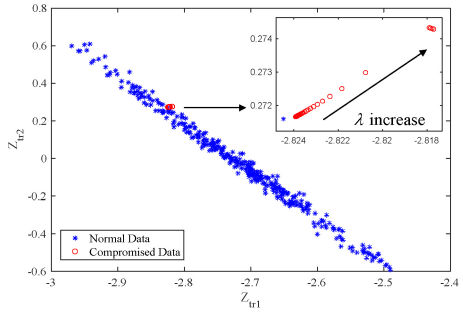



Fig. 10. Impact of weights on the stealthiness of MC-FDI attacks to the SVM.

impact and maintaining the temporal correlation of historical measurements.

Then, we demonstrate the impact of weights on the stealthiness of the MC-FDI attack on the SVM detector. We apply the PCA dimension reduction on 350 historical measurement vectors and the 22 MC-FDI attacks with increasing weights in the 300$^{\text{th}}$ instant, as shown in Fig. 10. We can see that all MC-FDI attacks succeed in locating inside the historical data area, indicating the stealthiness of MC-FDI attacks to the SVM detector. Moreover, as shown in the zoom-in figure, the compromised data with a larger $\lambda$ deviates farther from the historical data area. In summary, a larger weight results in a larger incremental voltage but also degrades the temporal correlation of the historical measurements, which increases the probability of being detected by the SVM detector.

### *E. Stealthiness of MC-FDI Attacks Against ML Detectors*

In this section, we further evaluate the stealthiness of MC-FDI attacks against other ML detectors, including artificial neural networks (ANN), polynomial logistic regression (LR), and Gaussian Naïve Bayes (NB). The performance of the four ML detectors on detecting TFDI and MC-FDI attacks are shown in Tables IV and V, respectively. Note that the column "Norm" in Tables IV and V represents the average L1-norm of the incremental voltage under the attacks with the given weight.

We apply artificial neural networks as the attack detector to evaluate the effectiveness of the proposed MC-FDI attacks. We construct an ANN with three hidden layers and conduct a grid search to optimize the number of neurons in each layer and the regularization parameters. Specifically, we search the

TABLE IV
PERFORMANCE OF MACHINE LEARNING DETECTORS ON DETECTING TFDI ATTACKS

| Detector | AM | Norm | Precision | Recall | F1 |
|---|---|---|---|---|---|
| SVM | 0.1 | 0.015 | 0.95 | 0.35 | 0.51 |
| | 0.3 | 0.049 | 0.97 | 0.74 | 0.84 |
| ANN | 0.1 | 0.015 | 0.94 | 0.52 | 0.67 |
| | 0.3 | 0.049 | 0.96 | 0.78 | 0.86 |
| LR | 0.1 | 0.015 | 1.00 | 0.26 | 0.42 |
| | 0.3 | 0.049 | 1.00 | **0.70** | 0.82 |
| NB | 0.1 | 0.015 | 0.75 | 0.18 | 0.29 |
| | 0.3 | 0.049 | 0.88 | 0.46 | 0.61 |

TABLE V
PERFORMANCE OF MACHINE LEARNING DETECTORS ON DETECTING MC-FDI ATTACKS

| Detector | $\lambda$ | Norm | Precision | Recall | F1 |
|---|---|---|---|---|---|
| SVM | 3.1 | 0.057 | 0.67 | 0.04 | 0.08 |
| | 3.2 | 0.140 | 0.78 | 0.07 | 0.13 |
| ANN | 3.1 | 0.057 | 0.67 | 0.12 | 0.20 |
| | 3.2 | 0.140 | 0.77 | 0.20 | 0.32 |
| LR | 3.1 | 0.057 | 1.00 | 0.02 | 0.04 |
| | 3.2 | 0.140 | 1.00 | **0.06** | 0.11 |
| NB | 3.1 | 0.057 | 0.81 | 0.26 | 0.39 |
| | 3.2 | 0.140 | 0.82 | 0.28 | 0.42 |

number of neurons in each hidden layer from 10 to 50 with an increment of 10, and the regularization parameter in the set {0.00001, 0.0001, 0.001, 0.01, 0.1, 1}. In the grid search, the average F1 score of the 5-fold cross-validation sets serves as the performance metric. In the IEEE 14-bus system, 50, 50, and 40 neurons are selected in three hidden layers with 0.01 regularization parameter. Due to the space limit, we only present the TFDI with 0.1 and 0.3 AM, and MC-FDI attacks with $\lambda = 3.1$ and $\lambda = 3.2$. It is seen that the ANN detector has a high precision value in detecting TFDI and MC-FDI attacks. ANN detector can detect **52%** and **78%** TFDI attacks with 0.1 AM and 0.3 AM, respectively, but it can only detect 12% and 20% MC-FDI attacks with $\lambda = 3.1$ and $\lambda = 3.2$, respectively.

Since the compromised and uncompromised data in the case study are not linearly separable, we adopt the polynomial logistic regression to deal with the nonlinear boundary. We perform a polynomial transformation on the original data with 4$^{\text{th}}$-order polynomials. The LR detector is solved using a Newton-CG solver. The penalization parameter C of the L2 penalty function is optimized by searching in the interval [0.01,100] using 5-fold cross-validation. The maximum number of iterations is chosen as 1000. The attack detection probability of the LR detector against TFDI attacks is **70%**, while that of the LR detector against MC-FDI attacks is **6%.**

In addition, we apply the Gaussian Naïve Bayes classifier to detect TFDI attacks and MC-FDI attacks. After the Gaussian Naïve Bayes classifier is trained, the probability threshold is optimized in the interval [0.01, 1] based on the average F1 score of the 5-fold cross-validation. It is seen that NB detector can detect 46% TFDI attacks with 0.88 precision, and detect 28% MC-FDI attacks. Even though the NB has the highest detection rate against MC-FDI attacks compared with other detectors, the low recall against TFDI attacks indicates

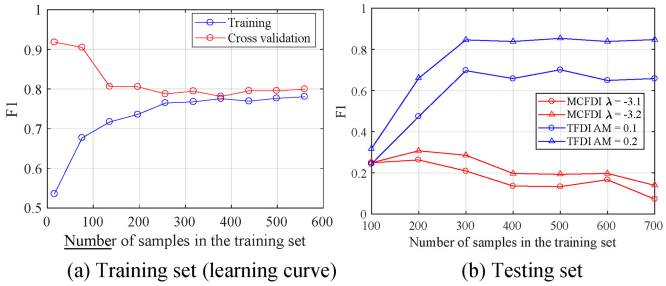



Fig. 11. The F1 score of the SVM detector using different numbers of training samples. Fig. 11(a) shows the performance of SVM in detecting TFDI attacks in the training process, and Fig. 11(b) shows the performance of the well-trained SVM in detecting TFDI attacks and MC-FDI attacks in the testing set.

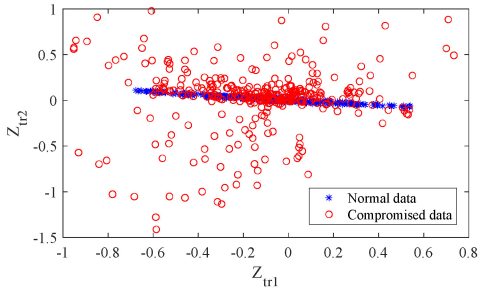



Fig. 12. Projection of the training set by MDS.

NB performs badly in determining whether a measurement vector is compromised or not.

The well-trained ML detectors have a low false positive (FP) rate, while the stealthiness of the proposed attacks against these detectors results in a low true positive (TP) rate. Consequently, the precision value of these detectors varies between [0, 1]. For example, the LR detector in Table V has a 1.0 precision value due to two detected attacks (TP=2) and zero misclassified normal data (FP = 0), while the LR detector in Table VII has a zero precision value due to zero detected attacks (TP=0). In Table V, the SVM detector detects four attacks (TP = 4) and incorrectly classifies two normal measurements as attacks (FP=2), resulting in a 0.67 precision value.

TFDI attacks with small AM are hidden to ML detectors, because these attacks merely inject tiny incremental voltage angle into the system. For TFDI attacks with 0.01 AM, its ADP against SVM, ANN, NB and LR are 0.01, 0.09, 0, and 0, respectively. Note that the L1-norm of incremental voltage in TFDI attacks with 0.01 AM is only 0.002. It is necessary to highlight that MC-FDI attacks with large L1-norm of the incremental voltage can also remain hidden to ML detectors. This is because the MC-FDI attacks consider the temporal correlation of the historical measurements, rather than injecting tiny incremental voltage angle.

We analyze the impact of the number of training data on the detection performance of SVM detector. The original training set includes 350 uncompromised historical measurement vectors collected from the 1st instant to 350th instant, and 350 compromised measurement vectors. We denote the training set by $\mathbf{Z}_{train} = \left[\mathbf{Z}^0_{train}\ \mathbf{Z}^a_{train}\right] \in \mathbb{R}^{m\times 700}$, where $\mathbf{Z}^0_{train}$ and $\mathbf{Z}^a_{train}$ are uncompromised and compromised historical measurements, respectively. When we decrease the size of the training samples to $k$, we keep the most recent historical measurements in the training set, i.e., $\mathbf{Z}^k_{train} = \left[\mathbf{Z}^0_{train}(idx)\ \mathbf{Z}^a_{train}(idx)\right]$ and $idx = 350-k/2+1{:}350$. Thus, the number of uncompromised measurement vectors is the same as that of compromised measurements in the training set.

Fig. 11(a) shows the learning curve of the SVM detector in the training process, in which the F1 score of the SVM in the training set and cross-validation set are calculated under different sizes of the training set. When the training set increases from 20 to 700 samples, we conduct 5-fold cross-validation in each given training set. Then, we calculate the average F1 score of the training set and cross-validation set, respectively. It is seen that the F1 scores of the training and cross-validation sets converge with a small gap, indicating no overfitting exists in the training process. Further increasing the sample size does not yield substantial enhancements in achieving higher accuracy. Figure 11(b) shows that, for the TFDI attacks, the detection effectiveness of the SVM detector can be improved when the number of samples in the training set increases from 100 to 300. After the number of samples in the training set is larger than 300, the performance becomes stable and cannot be improved by adding more training samples. The performance of the SVM detector in detecting TFDI attacks is consistent with the learning curve of the SVM detector proposed in [18]. The F1 score of the SVM detector in detecting MC-FDI attacks decreases with the increasing number of training samples. Specifically, the detector's precision increases but the recall decreases with the increasing number of the training samples. This is because, with more historical measurements, the project of MC-FDI attacks in $\mathbb{R}^2$ space is more likely to overlap with the historical measurements. Thus, MC-FDI attacks are more prone to be stealthy to the detector, i.e., a lower recall value.

In addition, we compare PCA with a global nonlinear technique for dimensionality reduction, multidimensional scaling (MDS) [51]. Then, we further evaluate the stealthiness of MC-FDI attacks using four ML detectors based on the projection data of MDS. PCA is a traditional linear technique for dimensionality reduction. PCA finds a linear basis of reduced dimensionality for the data, in which the amount of variance in the data is maximal. Multidimensional scaling (MDS) is one of the global nonlinear techniques for dimensionality reduction. Global nonlinear techniques attempt to preserve global properties of the data, and are capable of constructing nonlinear transformations between the high-dimensional data representation and its low-dimensional counterpart. Multidimensional scaling (MDS) maps the high-dimensional data representation to a low-dimensional representation while retaining the pairwise distances between the datapoints as much as possible. MDS is widely used for the visualization of data [52].

We evaluate the stealthiness of MC-FDI attacks against ML detectors under the global nonlinear dimensionality reduction technique, i.e., MDS. We apply the MDS dimension reduction to the training data set and the testing data set. Fig. 12 shows

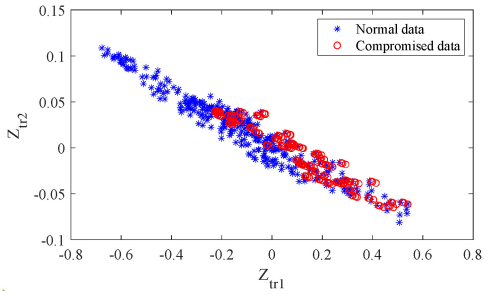



Fig. 13. Projection of MC-FDI attacks in the testing set by MDS.

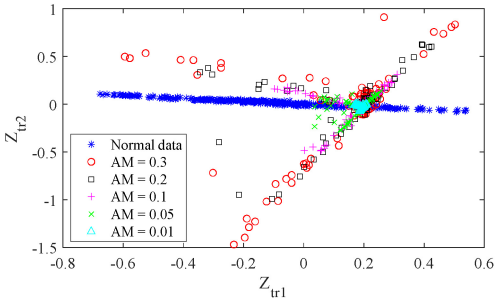



Fig. 14. Projection of TFDI attacks in the testing set by MDS.

TABLE VI
PERFORMANCE OF MACHINE LEARNING DETECTORS ON DETECTING TFDI ATTACKS USING MDS

| Detector | AM | Norm | Precision | Recall | F1 |
|---|---|---|---|---|---|
| SVM | 0.1 | 0.015 | 0.98 | 0.55 | 0.70 |
| | 0.3 | 0.049 | 0.99 | **0.79** | 0.88 |
| ANN | 0.1 | 0.015 | 1.00 | 0.53 | 0.69 |
| | 0.3 | 0.049 | 1.00 | 0.75 | 0.86 |
| LR | 0.1 | 0.015 | 1.00 | 0.34 | 0.51 |
| | 0.3 | 0.049 | 1.00 | **0.62** | 0.76 |
| NB | 0.1 | 0.015 | 1.00 | 0.21 | 0.35 |
| | 0.3 | 0.049 | 1.00 | 0.40 | 0.58 |

the projection of normal and compromised data in the training set, Fig. 13 shows the projection of MC-FDI attacks in the testing set, and Fig. 14 shows the projection of TFDI attacks with different attack magnitudes in the testing set. In Fig. 14, the projection of TFDI attacks by MDS deviated further from the normal data, compared with that by PCA in Fig. 5. It indicates more TFDI attacks can be detected by ML detectors under MDS. The projection of MC-FDI attacks under MDS locates inside the normal data area in Fig. 13, indicating the stealthiness of MC-FDI attacks.

Four ML detectors (SVM, ANN, LR, and NB) are trained and tested to detect TFDI and MC-FDI attacks using the low-dimensional data by MDS. The performance of the machine learning detectors on detecting TFDI and MC-FDI attacks are shown in Table VI and VII, respectively. In Table VI, it is seen that machine learning detectors have a better detection capability in detecting TFDI attacks with a larger AM. This is because TFDI attacks with a larger AM deviate further from the normal data compared with the TFDI attacks with a lower AM, as shown in Fig. 14. For TFDI attacks with 0.3 AM, SVM can detect **79%** attacks, and LR can detect **62%** attacks. In Table VII, it is seen that

TABLE VII
PERFORMANCE OF MACHINE LEARNING DETECTORS ON DETECTING MC-FDI ATTACKS USING MDS

| Detector | $\lambda$ | Norm | Precision | Recall | F1 |
|---|---|---|---|---|---|
| SVM | 3.1 | 0.057 | 0.86 | 0.06 | 0.11 |
| | 3.2 | 0.140 | 0.86 | 0.06 | 0.11 |
| ANN | 3.1 | 0.057 | 1.00 | 0.02 | 0.04 |
| | 3.2 | 0.140 | 1.00 | 0.02 | 0.04 |
| LR | 3.1 | 0.057 | 0.00 | 0.00 | 0.00 |
| | 3.2 | 0.140 | 0.00 | 0.00 | 0.00 |
| NB | 3.1 | 0.057 | 1.00 | 0.02 | 0.04 |
| | 3.2 | 0.140 | 1.00 | 0.06 | 0.11 |

the ADP of all four detectors is below 6%, indicating the stealthiness of MC-FDI attacks against machine learning detectors.

We further evaluate the effectiveness of MC-FDI attacks in the IEEE 118-bus system. We assume the attacker utilizes 200 historical measurement vectors to construct the historical measurement matrix, and continually launches MC-FDI attacks from the 300th instant to the 400th instant. Assume the MC-FDI attacker intends to compromise the voltage angle of Buses 60, 61, 62, 63, and 64.

In the IEEE 118-bus system, the dimension of the data is 292. The training dataset that includes 400 uncompromised measurement vectors and 400 compromised measurement vectors. These uncompromised measurement vectors are collected from the system under the WECC hourly load profile from the $1^{st}$ time instant to the 400th time instant. TFDI attacks randomly select five buses to compromise voltage from the 300th time instant to the 400th time instant with 0.6-1.0 AMs. The compromised measurements in the training set are sampled from the measurements under the TFDI attacks. The testing dataset includes 100 uncompromised measurement vectors collected from the $401^{st}$ time instant to the $500^{th}$ time instant, 400 compromised measurement vectors under MC-FDI attacks with $\lambda = \{6, 8, 10, 12\}$, and 500 compromised measurement vectors under TFDI attacks with AM $= \{0.6, 0.7, 0.8, 0.9, 1.0\}$. Specifically, we generate 100 attacks for each attack magnitude. Note that the attack magnitudes of the TFDI attacks are selected based on the voltage situation of the IEEE 118-bus system, such that the incremental voltages in TFDI attacks are comparable to those in MC-FDI attacks. Specifically, L1-norm of the incremental voltage in MC-FDI attacks with four $\lambda$ weights are 0.007, 0.019, 0.070, and 0.073, respectively, and those in TFDI attacks with five AMs are 0.066, 0.080, 0.088, 0.100, and 0.118, respectively. From the perspective of the attacker, the injected voltage in MC-FDI attacks and TFDI attacks are comparable. Figures 15, 16, and 17 show the projection of the dataset for training and cross-validation, the projection of MC-FDI attacks in the testing set, and the projection of TFDI attacks with different attack magnitudes in the testing set.

SVM, ANN, LR, and NB are trained and tested to evaluate the stealthiness of MC-FDI attacks in the IEEE 118-bus system. We set 30, 50, and 20 neurons in three hidden layers of the ANN and the regularization parameter is 0.01. Fig. 18 shows the learning curve of the ANN detector in the IEEE

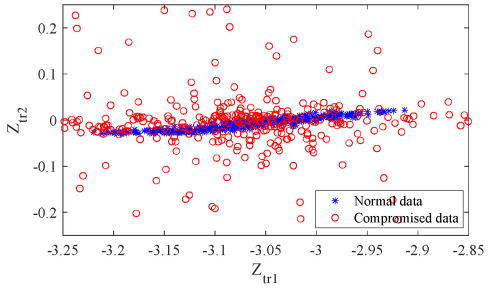



Fig. 15. Projection of the training set in the IEEE 118-bus system.

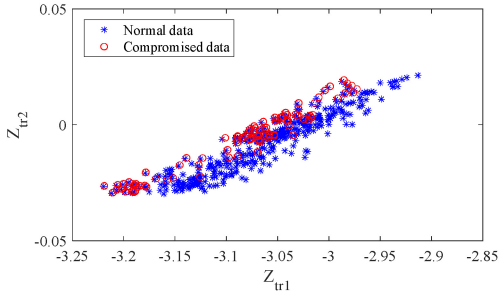



Fig. 16. Projection of MC-FDI attacks in the testing set in the IEEE 118-bus system.

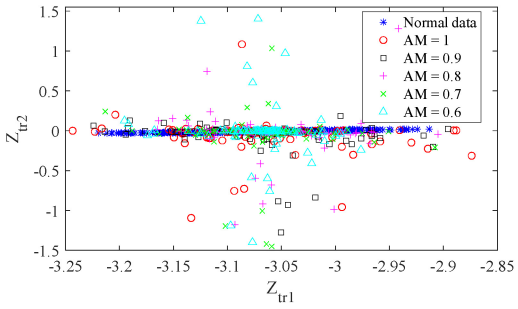



Fig. 17. Projection of TFDI attacks in the testing set in IEEE 118-bus system.

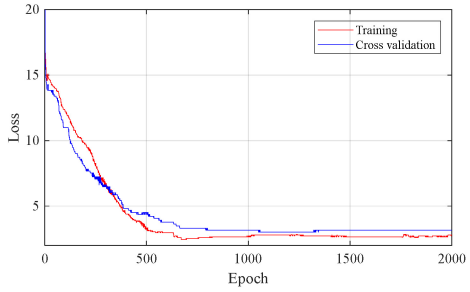



Fig. 18. The learning curve of the ANN detector in the IEEE 118-bus system.

118-bus system, in which the logistic loss of the training set and cross-validation set are calculated at each epoch. The learning curve indicates that ANN learns well, and the two curves converge after 1000 epochs.

Table VIII shows the performance of the machine learning detectors on detecting TFDI attacks. It is seen that the SVM has the best detection capability, which can detect **76%** TFDI attacks. The ADP of NB and LR are around 45%. This is because two detectors are not good at dealing with the overlap between the uncompromised and compromised data in the

TABLE VIII
PERFORMANCE OF MACHINE LEARNING DETECTORS ON DETECTING TFDI ATTACKS IN THE IEEE 118-BUS SYSTEM

| Detector | AM | Norm | Precision | Recall | F1 |
|---|---|---|---|---|---|
| SVM | 0.6 | 0.066 | 0.82 | 0.64 | 0.72 |
| | 0.9 | 0.100 | 0.84 | **0.76** | 0.80 |
| ANN | 0.6 | 0.066 | 0.80 | 0.62 | 0.70 |
| | 0.9 | 0.100 | 0.83 | 0.75 | 0.79 |
| LR | 0.6 | 0.066 | 0.69 | 0.34 | 0.46 |
| | 0.9 | 0.100 | 0.75 | 0.46 | 0.58 |
| NB | 0.6 | 0.066 | 1.00 | 0.27 | 0.43 |
| | 0.9 | 0.100 | 1.00 | 0.41 | 0.59 |

TABLE IX
PERFORMANCE OF MACHINE LEARNING DETECTORS ON DETECTING MC-FDI ATTACKS IN THE IEEE 118-BUS SYSTEM

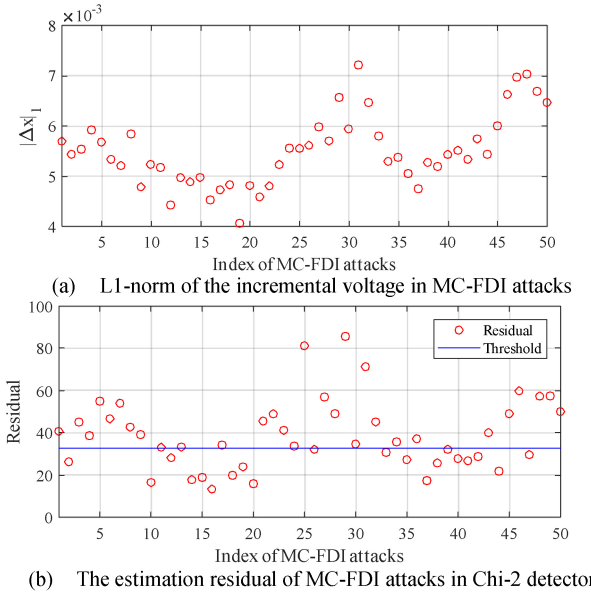
| Detector | $\lambda$ | Norm | Precision | Recall | F1 |
|---|---|---|---|---|---|
| SVM | 6 | 0.007 | 0.33 | 0.07 | 0.11 |
| | 12 | 0.073 | 0.36 | **0.08** | 0.13 |
| ANN | 6 | 0.007 | 0.38 | 0.09 | 0.14 |
| | 12 | 0.073 | 0.44 | 0.12 | 0.19 |
| LR | 6 | 0.007 | 0.32 | 0.07 | 0.11 |
| | 12 | 0.073 | 0.25 | 0.05 | 0.08 |
| NB | 6 | 0.007 | 0.00 | 0.00 | 0.00 |
| | 12 | 0.073 | 0.00 | 0.00 | 0.00 |

training set. Table IX shows the performance of the machine learning detectors on detecting MC-FDI attacks. The ADP of ANN is 12%, the ADP of SVM and LR are below **8%,** and the ADP of NB is zero. The simulation results verify the stealthiness of MC-FDI attacks against machine learning detectors.

### *F. Detection of MC-FDI Attacks Using MTD*

In this section, we evaluate the performance of MTD in detecting MC-FDI attacks. Similar to the previous sections, we assume the attacker continually launches MC-FDI attacks from the 300$^{\text{th}}$ instant to the 350$^{\text{th}}$ instant. The graph-based planning installs D-FACTS devices on nine lines, which are indexed by {1, 3, 4, 8, 10, 12, 13, 17, 18} in the IEEE 14-bus system. Note that we use the line index of the IEEE 14-bus system case in MATPOWER [53]. Generally, we set the MTD magnitude $\eta = 0.2$. In each instant from the 1$^{\text{st}}$ instant to the 350$^{\text{th}}$ instant, the random MTD operation method randomly selects the setpoints of D-FACTS in the range: $U(0.8b_{ij}^0, 1.2b_{ij}^0)$. Note that MC-FDI attacks under MTD are constructed using the original line impedance, and each measurement vector in the historical measurement matrix is based on a different system configuration.

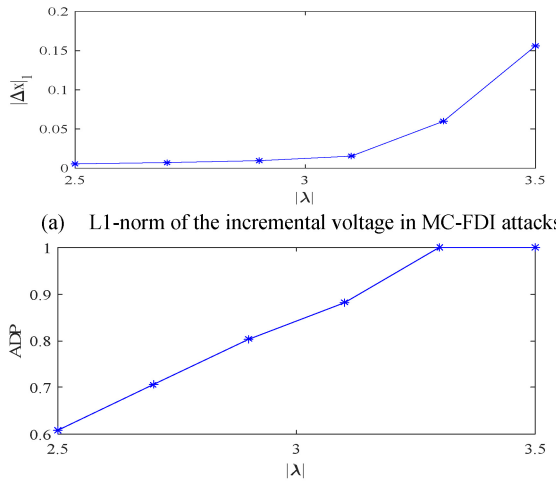
Assume the attacker adopts $\lambda = 2.5$ in the MC-FDI attack. For each MC-FDI attack, L1-norm of the incremental voltage and the estimation residual in the Chi-2 detector are shown in Fig. 19(a) and 19(b), respectively. The estimation residual is highly related to the incremental voltage. Specifically, a larger incremental voltage results in a larger estimation residual. With weight $\lambda = 2.5$, the incremental voltages in the MC-FDI attacks are relatively low (less than 0.01). The MTD method merely succeeds in detecting 61% of MC-FDI attacks. Even though the MTD detection effectiveness is mainly determined by the D-FACTS planning in the noiseless condition, it




(a) L1-norm of the incremental voltage in MC-FDI attacks

(b) The estimation residual of MC-FDI attacks in Chi-2 detector

Fig. 19. Performance of MTD in detecting MC-FDI attack with $\lambda = 2.5$.




(a) L1-norm of the incremental voltage in MC-FDI attacks

(b) ADP of MTD against MC-FDI attacks

Fig. 20. Performance of MTD in detecting MC-FDI attacks with different weights.

is also affected by FDI attack magnitude and MTD magnitude in noisy conditions. It is necessary to evaluate the effectiveness of MTD against MC-FDI attacks with different weights.

We further evaluate the performance of MTD in detecting MC-FDI attacks under different weights, as the weight essentially determines the attack magnitude of MC-FDI attacks. We increase the weight from 2.5 to 3.5 with an incremental of 0.2. For each weight, 50 MC-FDI attacks are constructed from the 300$^{\text{th}}$ instant to the 350$^{\text{th}}$ instant. In Fig. 20, it is seen that the L1-norm of the incremental voltage increases with an increase in the weight. Accordingly, the ADP of MTD increases with the weight due to the increasing L1-norm of the incremental voltage. It is necessary to highlight that a slight increase in the L1-norm of the incremental voltage from weight 2.5 to weight 3.1 results in a significant increase in the ADP. When the MC-FDI attacks with a small incremental voltage have a trivial negative impact on the system operation, the MTD method has a low ADP (around 60%). When the MC-FDI attacks start to have a large incremental voltage, the MTD method could reach a 100% detection probability. It reflects the effectiveness of MTD in detecting MC-FDI attacks and preventing the negative impact of MC-FDI attacks on the system operation.

## VI. Conclusion

In this paper, we propose convex MC-FDI attacks in the DC and AC power system model, respectively, which maximize the malicious incremental voltage and minimize the nuclear norm of the compromised historical measurement matrix. The proposed attack models first integrate the FDI attack model in the constraints to satisfy the spatial correlation of the malicious measurements, and then utilize the matrix completion to ensure malicious measurements consistent with the temporal correlation of the historical measurements. Therefore, the MC-FDI attacks are stealthy to both the model-based Chi-2 detector and the machine-learning detectors. Due to the high stealthiness of the MC-FDI attacks, we propose to apply MTD in the physical layer of power systems to detect MC-FDI attacks by actively changing the impedance of the lines with D-FACTS devices. We theoretically prove that MTD can corrupt both the temporal correlation and spatial correlation of the MC-FDI attacks. Simulation results show that the MC-FDI models are stealthy to both the Chi-2 detector and the machine learning detectors, and MTD is effective in detecting the MC-FDI attacks.

In the future, we will integrate matrix completion techniques into multiple blind FDI attacks, eliminating the attacker's need for grid topology and line parameters. These techniques will aid blind FDI attacks in determining the optimal malicious voltage increments. In addition, our future work will develop an alternating direction method of multipliers solver for the MC-FDI attacks to reduce the computational time.

## References

[1] A. S. Musleh, G. Chen, and Z. Y. Dong, "A survey on the detection algorithms for false data injection attacks in smart grids," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2218–2234, May 2020.

[2] H. Zhang, B. Liu, and H. Wu, "Smart grid cyber-physical attack and defense: A review," *IEEE Access*, vol. 9, pp. 29641–29659, 2021.

[3] Electricity Information Sharing and Analysis Center. "Analysis of the cyber attack on the Ukrainian power grid." Accessed: Aug. 11, 2019. [Online]. Available: https://ics.sans.org/media/E-ISAC

[4] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 Ukraine blackout: Implications for false data injection attacks," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3317–3318, Jul. 2017.

[5] M. Jin, J. Lavaei, and K. H. Johansson, "Power grid AC-based state estimation: Vulnerability analysis against cyber attacks," *IEEE Trans. Autom. Control*, vol. 64, no. 5, pp. 1784–1799, May 2019.

[6] J. Tian, B. Wang, and X. Li, "Data-driven and low-sparsity false data injection attacks in smart grid," *Security Commun. Netw.*, vol. 2018, pp. 1–11, Sep. 2018.

[7] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on smart grid state estimation: Attack strategies and countermeasures," in *Proc. 1st IEEE Int. Conf. Smart Grid Commun.*, Oct. 2010, pp. 220–225.

[8] S. Soltan, M. Yannakakis, and G. Zussman, "Power grid state estimation following a joint cyber and physical attack," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 1, pp. 499–512, Mar. 2018.

[9] J. Liang, L. Sankar, and O. Kosut, "Vulnerability analysis and consequences of false data injection attack on power system state estimation," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3864–3872, Sep. 2016.

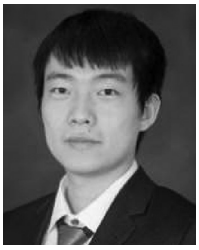
[10] Y. Yuan, Z. Li, and K. Ren, "Modeling load redistribution attacks in power systems," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 382–390, Jun. 2011.

[11] B. Sikdar and J. H. Chow, "Defending synchrophasor data networks against traffic analysis attacks," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 819–826, Dec. 2011.
[12] C. Pei, Y. Xiao, W. Liang, and X. Han, "PMU placement protection against coordinated false data injection attacks in smart grid," *IEEE Trans. Ind. Appl.*, vol. 56, no. 4, pp. 4381–4393, Jul. 2020.
[13] Z. Zhang, S. Gong, A. D. Dimitrovski, and H. Li, "Time synchronization attack in smart grid: Impact and analysis," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 87–98, Mar. 2013.
[14] X. Jiang, J. Zhang, B. J. Harding, J. J. Makela, and A. D. Domı'nguez-Garcı'a, "Spoofing GPS receiver clock offset of phasor measurement units," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3253–3262, Aug. 2013.
[15] S. Basumallik, S. Eftekharnejad, N. Davis, N. Nuthalapati, and B. K. Johnson, "Cyber security considerations on PMU-based state estimation," in *Proc. 5th Cybersecurity Symp.*, New York, NY, USA, 2018, pp. 1–4.
[16] A. Alexopoulos, G. N. Korres, and N. M. Manousakis, "Complementarity reformulations for false data injection attacks on PMU-only state estimation," in *Proc. Elect. Power Syst. Res.*, vol. 189, 2020, pp. 1–8.
[17] M. Ozay, I. Esnaola, F. T. Y. Vural, S. R. Kulkarni, and H. V. Poor, "Machine learning methods for attack detection in the smart grid," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 8, pp. 1773–1786, Aug. 2016.
[18] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, "Detecting stealthy false data injection using machine learning in smart grid," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1644–1652, Sep. 2017.
[19] J. Sakhnini, H. Karimipour, and A. Dehghantanha, "Smart grid cyber attacks detection using supervised learning and heuristic feature selection," in *Proc. IEEE 7th Int. Conf. Smart Energy Grid Eng. (SEGE)*, Aug. 2019, pp. 108–112.
[20] X. P. Li, L. Huang, H. C. So, and B. Zhao. "A survey on matrix completion: Perspective of signal processing." Accessed: Oct. 12, 2022. [Online]. Available: http://arxiv.org/abs/1901.10885
[21] Y. Zhang, A. Bernstein, A. Schmitt, and R. Yang, "State estimation in low-observable distribution systems using matrix completion," presented at the Hawaii Int. Conf. Syst. Sci., 2019.
[22] P. L. Donti, Y. Liu, A. J. Schmitt, A. Bernstein, R. Yang, and Y. Zhang, "Matrix completion for low-observability voltage estimation," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2520–2530, May 2020.
[23] A. Sagan, Y. Liu, and A. Bernstein, "Decentralized low-rank state estimation for power distribution systems," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3097–3106, Jul. 2021.
[24] B. Liu, H. Wu, Y. Zhang, R. Yang, and A. Bernstein, "Robust matrix completion state estimation in distribution systems," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2019, pp. 1–5.
[25] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stefopoulos, "Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1006–1013, Mar. 2016.
[26] K. Xie et al., "Recover corrupted data in sensor networks: A matrix completion solution," *IEEE Trans. Mobile Comput.*, vol. 16, no. 5, pp. 1434–1448, May 2017.
[27] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, p. 717, Apr. 2009.
[28] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans Inf. Syst. Security*, vol. 14, no. 1, pp. 1–33, Jun. 2011.
[29] I. J. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
[30] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York, NY, USA: Wiley, 2001.
[31] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, 2nd ed. New York, NY, USA: Springer, 2009.
[32] R. Deng and H. Liang, "False data injection attacks with limited susceptance information and new countermeasures in smart grid," *IEEE Trans. Ind. Informat.*, vol. 15, no. 3, pp. 1619–1628, Mar. 2019.
[33] G. Hug and J. A. Giampapa, "Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1362–1370, Sep. 2012.
[34] Z.-H. Yu and W.-L. Chin, "Blind false data injection attack using PCA approximation method in smart grid," *IEEE Trans. Smart Grid*, vol. 6, no. 3, pp. 1219–1226, May 2015.
[35] J. Kim, L. Tong, and R. J. Thomas, "Subspace methods for data attack on state estimation: A data driven approach," *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1102–1114, Mar. 2015.
[36] S. Lakshminarayana, F. Wen, and D. K. Y. Yau, "Trade-offs in data-driven false data injection attacks against the power grid," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 2022–2026.
[37] M. C. Grant and S. P. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*. London, U.K.: Springer, 2008, pp. 95–110.
[38] S. Diamond and S. Boyd, "CVXPY: A python-embedded modeling language for convex optimization," *J. Mach. Learn. Res.*, vol. 17, no. 83, pp. 1–5, 2016.
[39] B. Liu and H. Wu, "Optimal planning and operation of hidden moving target defense for maximal detection effectiveness," *IEEE Trans. Smart Grid*, vol. 12, no. 5, pp. 4447–4459, Sep. 2021.
[40] B. Liu and H. Wu, "Optimal D-FACTS placement in moving target defense against false data injection attacks," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4345–4357, Sep. 2020.
[41] B. Liu and H. Wu, "Systematic planning of moving target defence for maximising detection effectiveness against false data injection attacks," *IET Cyber Phys. Syst. Theory Appl.*, vol. 6, no. 3, pp. 151–163, 2021.
[42] M. A. Rahman, E. Al-Shaer, and R. B. Bobba, "Moving target defense for hardening the security of the power system state estimation," in *Proc. 1st ACM Workshop Moving Target Defense*, New York, NY, USA, 2014, pp. 59–68.
[43] J. Tian, R. Tan, X. Guan, and T. Liu, "Enhanced hidden moving target defense in smart grids," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 2208–2223, Mar. 2019.
[44] S. Lakshminarayana and D. K. Y. Yau, "Cost-benefit analysis of moving-target defense in power grids," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 1152–1163, Mar. 2021.
[45] H. Zhang, B. Liu, X. Liu, A. Pahwa, and H. Wu, "Voltage stability constrained moving target defense against net load redistribution attacks," *IEEE Trans. Smart Grid*, vol. 13, no. 5, pp. 3748–3759, Sep. 2022.
[46] B. Liu, H. Wu, Q. Yang, and H. Zhang, "Random-enabled hidden moving target defense against false data injection alert attackers," *Processes*, vol. 11, no. 2, p. 2, Feb. 2023.
[47] B. Li, G. Xiao, R. Lu, R. Deng, and H. Bao, "On feasibility and limitations of detecting false data injection attacks on power grid state estimation using D-FACTS devices," *IEEE Trans. Ind. Informat.*, vol. 16, no. 2, pp. 854–864, Feb. 2020.
[48] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," vol. 12, no. 85, pp. 2825–2830, 2011.
[49] "Electric reliability council of Texas." Accessed: Mar. 1, 2023. [Online]. Available: https://www.ercot.com/
[50] "WECC—Western Electricity Coordinating Council." Accessed: Mar. 1, 2023. [Online]. Available: https://www.wecc.org:443/Pages/home.aspx
[51] T. Cox and M. Cox, *Multidimensional Scaling*. London, U.K.: Chapman & Hall, 1994.
[52] M. S. Venkatarajan and W. Braun, "New quantitative descriptors of amino acids based on multidimensional scaling of a large number of physical–chemical properties," *Mol. Model. Annu.*, vol. 7, no. 12, pp. 445–453, Dec. 2001.
[53] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

**Bo Liu** (Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from the Harbin Institute of Technology, China, in 2013 and 2015, respectively, and the Ph.D. degree from the Mike Wiegers Department of Electrical and Computer Engineering, Kansas State University, Manhattan, KS, USA, in 2021, where he is a Research Assistant Professor. His current research interests include cyber–physical security of power systems, smart grid technologies, machine learning, and state estimation in smart grids. He serves as a Guest Editor of a special issue of *Processes* (MDPI).

**Hongyu Wu** (Senior Member, IEEE) received the B.S. degree in energy and power engineering and the Ph.D. degree in control science and engineering from Xi'an Jiaotong University, Xi'an, China. He is an Associate Professor and a Lucas-Rathbone Professor with the Mike Wiegers Department of Electrical and Computer Engineering, Kansas State University (K-State), Manhattan, KS, USA. Before joining K-State, he was a Research Engineer with the Power Systems Engineering Center, National Renewable Energy Laboratory, Golden, CO, USA. From 2011 to 2014, he was a Postdoctoral Researcher with the Robert W. Galvin Center for Electricity Innovation, Illinois Institute of Technology, Chicago, IL, USA. His research interests include cyber–physical security of smart grids, power system planning, operation and energy management, and power grid integration of renewable energy. He is a National Science Foundation (NSF) CAREER Awardee and an NSF EPSCoR Research Fellow. He serves on the IEEE-NERC Security Integration Project Committee and as an Associate Editor for IEEE TRANSACTIONS ON SMART GRID and IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS.

**Qihui Yang** received the B.S. and M.S. degrees in electrical engineering from the Harbin Institute of Technology, China, in 2012 and 2015, respectively, the M.S. degree in engineering policy and analysis from the Delft University of Technology, The Netherlands, in 2017, and the Ph.D. degree in electrical and computer engineering from Kansas State University, Manhattan, KS, USA, in 2021, where she is a Research Assistant Professor with the Department of Electrical and Computer Engineering. Her main research interests include modeling and analysis of complex systems.



**Hang Zhang** (Student Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical and computer engineering from Kansas State University, Manhattan, KS, USA, in 2017, 2018, and 2022, respectively. He is currently an Assistant Professor with the Department of Information Engineering, Henan University of Science and Technology, Henan, China. His research interests include cyber–physical security and resiliency of power systems, machine learning, and renewable energy.

**Yajing Liu** (Member, IEEE) received the Ph.D. degree in electrical engineering from Colorado State University, Fort Collins, CO, USA, in 2018. From 2018 to 2021, she was a Researcher with National Renewable Energy Laboratory, Golden, CO, USA. Since 2021, she has been a Research Scientist with Colorado State University. Her research interests include optimization algorithms, matrix/tensor completion with particular applications to power system state estimation and cyberattack, and geometries of learning.

**Yingchen Zhang** (Senior Member, IEEE) received the B.S. degree from Tianjin University, China, in 2003, and the Ph.D. degree from Virginia Polytechnic Institute and State University in 2010. He is the Vice President of Product Solutions with Utilidata Inc. He previously worked with NREL as a Research Group Manager pioneering artificial intelligence's applications in power systems. Before joining NREL, he was with California ISO developing and implementing State Estimator and Network Applications. He authored/coauthored over 150 peer-reviewed publications and holds two U.S. patents. His key areas of expertise lie in edge intelligence, advanced energy management system for future grids, and large-scale renewable integration. He serves as the Chair for IEEE PES Renewable Systems Integration Coordination Committee.