Analysis of Moving Target Defense Against False Data Injection Attacks on Power Grid

Zhenyong Zhang, Ruilong Deng, *Member, IEEE*, David K. Y. Yau, *Senior Member, IEEE*, Peng Cheng, *Member, IEEE*, and Jiming Chen, *Fellow, IEEE*

***Abstract*— Recent studies have considered thwarting false data injection (FDI) attacks against state estimation in power grids by proactively perturbing branch susceptances. This approach is known as moving target defense (MTD). However, despite of the deployment of MTD, it is still possible for the attacker to launch stealthy FDI attacks generated with former branch susceptances. In this paper, we prove that, an MTD has the capability to thwart all FDI attacks constructed with former branch susceptances only if (i) the number of branches l in the power system is not less than twice that of the system states n (i.e., $l \geq 2n$, where $n+1$ is the number of buses); (ii) the susceptances of more than n branches, which cover all buses, are perturbed. Moreover, we prove that the state variable of a bus that is only connected by a single branch (no matter it is perturbed or not) can always be modified by the attacker. Nevertheless, in order to reduce the attack opportunities of potential attackers, we first exploit the impact of the susceptance perturbation magnitude on the dimension of the *stealthy attack space*, in which the attack vector is constructed with former branch susceptances. Then, we propose that, by perturbing an appropriate set of branches, we can minimize the dimension of the *stealthy attack space* and maximize the number of covered buses. Besides, we consider the increasing operation cost caused by the activation of MTD. Finally, we conduct extensive simulations to illustrate our findings with IEEE standard test power systems.**

***Index Terms*— Power grids, cyber-physical system, false data injection attack, moving target defense, completeness, optimal protection.**

I. Introduction

MODERN power grid is becoming more scalable for new devices, more efficient for productions and smarter for operations. For the purpose to realize intelligent automation at all operation levels, numerous sensors and meters are distributed in this large-scale system for wide area monitoring, protection and control. However, these advanced information and communication technology (ICT) components make the power system prone to cyber attacks. It has been proved that the attacker can tamper with measurements in the field devices such as the smart meter [2] and the remote terminal unit (RTU) [3]. With increased connectivity of physical power grids to open systems such as the Internet (e.g., convergence between IT and operation technology or OT), it is imperative to enhance the security of power grids to keep intruders at bay [4]–[6].

Traditional supervisory control and data acquisition (SCADA) systems for power grids implement basic integrity and availability checks (e.g., bad data detection or BDD) for their data, to reject erroneous measurements due to failures or malicious attacks such as false data injection (FDI). However, research has shown that carefully designed FDI attacks can bypass the BDD and remain stealthy, when attackers utilize comprehensive knowledge of the system topology and branch susceptances of the power network to guide their actions [7]. Although stealthy, these FDI attacks [8]–[12] can be quite powerful. They may lead to large errors in the estimated system states and cause severe consequences such as prolonged interruption of power supply or equipment damage [13].

Addressing the imminent threats posed by stealthy FDI attacks, many recent research efforts have sought to characterize their properties and propose countermeasures against them [14]–[16]. For example, methods to secure meter measurements and critical state variables against tampering have been proposed [17], [18]. In practice, however, breach of the perimeter, including cryptographical safeguards, has been repeatedly demonstrated in the real world through persistent attempts by malicious attackers [2], [3], [19], [20]. Besides, since only partial of the measurements are trusted, this strategy may reduce the redundancy of the original monitoring system.

Observing that the construction of stealthy FDI attacks depends on the detailed knowledge of the power grid's configuration, an alternative defense approach is to change the system parameters by design for defeating the knowledgeable attacker. Existing work has typically implemented such *moving target defense* (MTD) [21] by proactively perturbing the impedance of certain branches of a power network using the distributed flexible AC transmission system (D-FACTS) devices, e.g., DSI and DSSC [22]. By modifying D-FACTS, the changes of branch parameters are unpredictable to attackers, thus, increasing uncertainties for them to execute stealthy FDI attacks on the power system, which complies with the definition of MTD.

Manuscript received January 7, 2019; revised April 27, 2019 and June 12, 2019; accepted July 2, 2019. Date of publication July 15, 2019; date of current version February 4, 2020. This work was supported in part by the National Key Research and Development Program of China under Grant 2016YFB0800204, in part by the National Natural Science Foundation of China under Grant 61833015, in part by the Singapore University of Technology and Design-Zhejiang University Innovation, Design and Entrepreneurship Alliance (SUTD-ZJU IDEA) under Award 201805, in part by the Nanyang Technological University (NTU) Internal Funding-Start-up Grant (SUG)-the College of Engineering (CoE) under Grant M4082287, and in part by the A*STAR- Nanyang Technological University-Singapore University of Technology and Design AI Partnership under Grant RGANS1906. This article was presented at the IEEE PES ISGT 2019 [1]. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Walid Saad. *(Corresponding author: Peng Cheng.)*

Z. Zhang, P. Cheng, and J. Chen are with the State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China, and also with the College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: zhangzhenyong@zju.edu.cn; saodiseng@gmail.com; cjm@zju.edu.cn).

R. Deng is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: rldeng@ntu.edu.sg).

D. K. Y. Yau is with the Information Systems Technology and Design, Singapore University of Technology and Design, Singapore 487372 (e-mail: david_yau@sutd.edu.sg).

Digital Object Identifier 10.1109/TIFS.2019.2928624

Prior works on MTD against stealthy FDI attacks have demonstrated success in defeating knowledgeable attackers. Morrow *et al.* [23] and Davis *et al.* [24] investigated the divergence of the system state due to bad/malicious data, by comparing expected results caused by branch impedance perturbations with actual observed responses of the system. Rahman *et al.* [21] presented the formal design of an MTD application, and show by simulations that arbitrary branch susceptance perturbations may not be effective in detecting FDI attacks. Tian *et al.* [25] proposed a notion of hidden MTD, which aims to make the defense stealthy to attackers. Liu *et al.* [26] extended the hidden MTD to the AC distribution system, and remarkably, considered the minimization of power losses and power flow differences before and after MTD. Lakshminarayana and Yau [27] presented analytical conditions for MTD to be truly effective, and presented an explicit cost-benefit analyis of the MTD, which can be viewed as a form of insurance. Liu *et al.* [28] considered the utility of MTD as an optimized goal, and solved a joint optimization problem with the generation cost loss. Although the authors in [25], [27], [28] have analyzed MTD's capability to thwart FDI attacks constructed with former branch parameters, they haven't given insight about the effectiveness of MTD and presented limitations of MTD related to both the power network structure and perturbed branches. Besides, the impact of the perturbation magnitude on the stealthy FDI attack space after MTD is not thoroughly investigated.

In fact, for understanding the impacts of the deployment of D-FACTS devices on the power system, pioneer works have devoted to studying the linear sensitivities of power system quantities such as voltage and power losses with respect to the branch impedance perturbation [29]–[31]. For example, Rogers and Overbye [29] analyzed the linear sensitivities with respect to branch impedance for solving the real power loss minimization and voltage control problems. They showed that, by perturbing the branch impedances of 5 branches within the range $+/-$ 20% of their original values in the IEEE 14-bus power system with D-FACTS devices, the power loss is 3.35 MW compared to 3.51 MW without any perturbation. Significantly, Morrow *et al.* [23] investigated the impacts of the use of D-FACTS on the power losses and voltages considering the defense effect of branch perturbations. They proved that if the perturbations are within 20% of the impedance in the IEEE 14-bus power system, then there are nearly 10,000 perturbation cases that can restrict the power losses within 1%. That is, the operating points after branch perturbations are limited to the small neighborhood around the optimum operation points. Other power systems were also tested. The results highlight the practicality of modifying D-FACTS to provide MTD in power grid.

In this paper, we focus on analyzing the completeness, deployment and the increasing operation cost of MTD in terms of thwarting FDI attacks constructed with old system information. Similar to prior studies [21], [25], [27], [28], we mainly consider FDI attacks of the form $\mathbf{a} = \mathbf{Hc}$ [7] with the DC power flow model. To begin with, we define a *stealthy attack space* as the intersection of the set of attack vectors generated with the former branch susceptances and the set of attack vectors generated with the current branch susceptances after MTD. Once an MTD is able to reduce the dimension of the *stealthy attack space* to 0, it is defined as a complete MTD. Based on these definitions, we analyze the conditions required for a complete MTD from the aspect of inherent power network structure and the branches that are perturbed. Besides, limitations for a power system to satisfy these conditions are also exploited. Further, for the case when the necessary conditions for a complete MTD are not met, we investigate methods to narrow down the attack opportunities of potential attackers by properly selecting a set of target-perturbation branches.[1] Moreover, we discuss the reduction of the additional operation cost caused by the activation of MTD. It should be clarified that some results have been presented in our conference paper [1], in which we have analyzed the topology limitation for achieving a complete MTD and the impacts of the perturbation magnitude on the reduction of the stealthy attack space. While in this journal version, more issues are discussed, including the special power network structure with which we can never achieve complete MTD, the impacts of the perturbation magnitude on the change (increase, invariant and decrease) of the dimension of the stealthy attack space, the optimal deployment of D-FACTS devices and the additional operation cost caused by MTD. We also present more simulations to illustrate and demonstrate our findings. In summary, the contributions are as follows:

- First, we prove that an MTD is complete only if (i) the number of branches l is larger than or equal to twice that of the system states n (i.e., $l \geq 2n$, where $n+1$ is the number of system buses), and (ii) the susceptances of more than n branches, which cover all buses, are perturbed. Besides, we prove that we can never realize a complete MTD if the power network contains a bus that is only connected by a single branch. The state variable of this bus can be injected arbitrary bias by the attacker.
- Moreover, we observe that, the change of the perturbation magnitude almost does not affect the dimension of the stealthy attack space. Based on this result, we propose an algorithm to compute a feasible set of target-perturbation branches that can minimize the dimension of the stealthy attack space and maximize the number of covered buses. Besides, we discuss the increase of the operation cost after MTD considering the security constraint, which is associated with the susceptance perturbation magnitude.
- Finally, we illustrate and demonstrate our results by conducting extensive simulations with the IEEE standard power systems.

The remainder of this paper is organized as follows. We introduce the system model and threat model in Section II-A. The problem statement is addressed in Section III. In Section IV, we analyze the MTD in terms of thwarting FDI attacks. We present simulation results of our findings in Section V. Section VI concludes the paper.

II. System Model and Threat Model

Throughout this paper, calligraphy font ($\mathcal{L}$) indicates a set or a subspace ($\mathcal{A}$), math boldface font ($\mathbf{H}$) indicates a matrix, bold lower case letter ($\mathbf{x}$) indicates an vector. $S(\cdot)$ denotes the (column) span of a matrix. $R(\cdot)$ is the rank of a matrix. $\cdot^T$ indicates the transpose of a matrix. All proofs in this paper are included in the Appendix. $|\cdot|$ denotes the cardinality of a set.

[1]The target-perturbation branch means that the susceptance of the branch will be perturbed.

A. System Model

To avoid intensive computation and obtain an optimal solution for large power systems, power system engineers usually utilize a linearized DC power flow model to approximate the AC power flow model [32], [33]. For its computational speed and simplicity, the DC model has been widely used for decades in both industry and academia [55]–[58]. Although it is less accurate, the DC model is more robust and often used in real-time operations such as the computation of marginal price [36]. In the DC model, the voltage magnitude is by default set as 1 p.u.. The state variables are reduced to the voltage phase angle. The measurements are reduced to active power flows. Moreover, since the phase angle difference is usually constrained to be small, the power flow equations can be reduced to

$$f_{ij} = -b_{ij}(\theta_i - \theta_j), \quad p_i = \sum_{j \in \mathcal{K}_i} f_{ij}, \tag{1}$$

where f_{ij} indicates the active power flow between bus i and bus j, b_{ij} is the equivalent susceptance on branch $\{i, j\}$, θ_i and θ_j are respectively the voltage phase angle of bus i and j, p_i is the active power injection of bus i, $\mathcal{K}_i$ is a set of neighboring buses connected to bus i.

We consider a classic power transmission network consisting of a set $\mathcal{N} = \{1, 2, \cdots, n+1\}$ of buses and a set $\mathcal{L} = \{k_1, k_2, \cdots, k_l\}$ of branches, where $n+1$ is the number of buses and l is the number of branches. With the DC model, by setting an arbitrary bus as the reference/slack bus, the remaining n phase angles $\theta_1, \theta_2, \cdots, \theta_n$ are taken to form the system states, typically denoted as $\mathbf{x} \in \mathbb{R}^n$. Each branch $k_t = \{i, j\} \in \mathcal{L}$ connects two buses i and j. Assuming that the power system is fully measured, i.e., each bus is monitored by one meter and each branch is monitored by two meters (both in the positive and negative direction). Then, the number of meter measurements is $m = 2l + n + 1$. The DC model can be derived as

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \boldsymbol{\eta}. \tag{2}$$

where $\mathbf{H}$ is termed as the measurement matrix, $\mathbf{z}$ denotes the measurements of active power injections and active power flows, $\boldsymbol{\eta}$ represents the independent measurement noises, which are typically assumed to be normally distributed [i.e., $\eta_i \sim \mathcal{N}(0, \sigma_i^2)$].

1) Construction of the Measurement Matrix: Let $\mathbf{A}$ denote the branch-bus incidence matrix. It is given by

$$a_{ti} = \begin{cases} 1, & \text{if branch } k_t \text{ starts from bus } i; \\ -1, & \text{if branch } k_t \text{ ends at bus } i; \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

where a_{ti} is the element in the position (t, i) of the matrix $\mathbf{A}$. Let $\mathbf{D}$ denote the diagonal branch susceptance matrix. Its diagonal element $\mathbf{D}_{tt}$ is $-b_{ij}$ with branch $k_t = \{i, j\}$. Thus, the invertible symmetric admittance matrix is $\mathbf{B} = \mathbf{A}^T\mathbf{D}\mathbf{A}$ and the branch-bus shift factor matrix is $\mathbf{S} = \mathbf{D}\mathbf{A}$. Considering the DC power flow equations, we have $\mathbf{f} = \mathbf{S}\boldsymbol{\theta}$ and $\mathbf{p} = \mathbf{B}\boldsymbol{\theta}$, where $\mathbf{f}$ denotes the active power flows, $\mathbf{p}$ denotes the power injections, $\boldsymbol{\theta}$ denotes the phase angles. Therefore, the measurement matrix $\mathbf{H} \in \mathbb{R}^{m \times n}$ can be derived as

$$\mathbf{H} = \begin{bmatrix} \mathbf{B} \\ \mathbf{S} \\ -\mathbf{S} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T\mathbf{D}\mathbf{A} \\ \mathbf{D}\mathbf{A} \\ -\mathbf{D}\mathbf{A} \end{bmatrix}, \tag{4}$$

We can see that the measurement matrix $\mathbf{H} \in \mathbb{R}^{m \times n}$ is a function of the system topology and branch susceptances.

In cases that the power systems are partially measured (i.e., some buses and branches are not monitored by meters), the corresponding measurement matrix is formed by selecting some rows from the measurement matrix of the fully measured case. For any measurement matrix, we assume that $m > n$ in this paper. Note that all results in this paper are satisfied whenever the system is fully measured or partially measured.

2) State Estimation (SE): State estimator, a fundamental tool for economically and dynamically routing power flows, is responsible to optimally estimate the state variables with noisy measurements collected by the underlying SCADA system [32]. The most common and concise mathematical description of SE is

$$\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}\mathbf{z} = \mathbf{K}\mathbf{z}, \tag{5}$$

where $\mathbf{W}$ is a diagonal matrix whose elements are reciprocals of the noise variances, $\mathbf{K}$ is referred to as the "pseudo-inverse" of $\mathbf{H}$ because $\mathbf{K}\mathbf{H} = \mathbf{I}$. $\hat{\mathbf{x}}$ is calculated by using certain statistical estimation criterions such as maximum likelihood, weighted least-squares and etc.

3) Bad Data Detection (BDD): Normally, in order to filter out abnormal/erronuous meter measurements, the bad data detection (BDD) checker is used in SE. Let $\mathbf{r} = \mathbf{z} - \mathbf{H}\hat{\mathbf{x}} = (\mathbf{I} - \mathbf{H}\mathbf{K})\mathbf{z}$ be the residues between the measurements $\mathbf{z}$ and their estimates $\hat{\mathbf{z}} = \mathbf{H}\hat{\mathbf{x}}$. BDD compares the Euclidean norm $\|\mathbf{r}\|_2$ with a predetermined threshold τ. If $\|\mathbf{r}\|_2 > \tau$, the abnormal alarm is triggered; otherwise, the measurements $\mathbf{z}$ are considered normal.

In the following, we assume that the measurements are noiseless (i.e., $\mathbf{e} = \mathbf{0}$) to simplify the discussions. Under this assumption, the DC model can be written as $\mathbf{z} = \mathbf{H}\mathbf{x}$. The corresponding state estimates are given by $\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{z} = \mathbf{x}$. Moreover, the threshold τ for the BDD is equal to 0. Nevertheless, the main results in the following derivations still keep valid with noisy measurements. In our simulations (see Section V), we will evaluate the impacts of measurement noises.

B. Threat Model

As a critical infrastructure, the security issue of power grid has attracted a lot of attention. For example, US National Electric Sector Cybersecurity Organization Resource (NESCOR) records safety incidents and negative impacts on the physical objects in power systems [34]. North American Electrical Reliability Council (NERC) reports lessons learned from failures and blackouts caused by cyber system faults [35]. These recordings show the security risks of power systems, which should be well addressed during daily maintenance and operation.

Adversarial Setting: In this paper, we consider the class of stealthy FDI attacks studied in [7]. Let $\mathbf{a} \in \mathbb{R}^m$ be the attack vector. The malicious measurements are given by $\mathbf{z}_a = \mathbf{z} + \mathbf{a}$. It has been proved that $\mathbf{z}_a$ cannot be detected by the BDD if $\mathbf{a} = \mathbf{H}\mathbf{c}$ [7], where $\mathbf{c} \in \mathbb{R}^n$ is an arbitrary vector. That is, $\mathbf{r}^a = \mathbf{z}^a - \mathbf{H}\hat{\mathbf{x}}^a = \mathbf{z} + \mathbf{a} - \mathbf{H}(\hat{\mathbf{x}} + \mathbf{c}) = \mathbf{z} - \mathbf{H}\hat{\mathbf{x}} = \mathbf{r}$, where $\mathbf{r}^a$ denotes the measurement residues after the FDI attack and $\hat{\mathbf{x}}^a$ denotes the modified states. We can see that the measurement residues are not changed after the FDI attack.

Thus, the bad data alarm is not triggered while the system states are subverted.

In other words, let $\mathcal{A} \in \mathbb{R}^m$ be a subspace that contains all FDI attacks with the form of $\mathbf{a} = \mathbf{Hc}$. Then, $\mathcal{A}$ is given by

$$\mathcal{A} = \{\mathbf{a} \mid \mathbf{a} = \mathbf{Hc},\ \mathbf{c} \in \mathbb{R}^n\}. \tag{6}$$

We can see that $\mathcal{A}$ is equal to the subspace $S(\mathbf{H})$, where $S(\cdot)$ denotes the (column) span of a matrix. Therefore, if $\mathbf{a} \in S(\mathbf{H})$ holds, then the malicious measurements $\mathbf{z}_a$ can bypass the BDD. In this paper, we assume the attacker has the following capabilities:

- The attacker is able to know a measurement matrix $\mathbf{H}$ based on his/her understanding, e.g., through topology-leaking attacks [37] or subspace attacks [38]. Since these topology attacks depend on historical measurements, the measurement matrix $\mathbf{H}$ cannot be learned by the attacker immediately. The learning/inferring process usually takes a sufficiently long time (hours or days) due to the exfiltration of an enough amount of historical measurement data [27], [38]–[41].
- The attacker is able to eavesdrop and tamper with the measurements by intruding into the communication network or IP-accessible field devices [42]. Practically, the attacker will have limited resources to compromise all the measurements [7]. However, we do not assume a priori what specific data can be compromised or not. Note also that although the attacker might be able to compromise the confidentiality of the raw data points, he/she will need non-trivial efforts and time to learn their higher system level relationships to guide his attack.
- The attacker is unable to take over (or have full access to) the control center or the SCADA systems. Once the attacker can exploit the control center or the SCADA systems, he/she is powerful enough to thwart most defensive mechanisms.

III. Problem Statement

Based on the system model and threat model, in this section, we first introduce the approach of MTD used in power grid. Then, we present the problems mainly investigated in this paper.

A. Moving Target Defense by Perturbing Branch Susceptances

1) D-FACTS: The distributed flexible AC transmission system (D-FACTS) devices are first introduced by Divan and Johal [22] for controlling the power flow. These devices can alter the impedances of power lines, and thus, control power flows to eliminate transmission constraints and bottlenecks. They are small and light enough to be suspended from the power line, floating both electrically and mechanically on it. Moreover, the equipped communication system enables them to receive control commands and transmit working states to remote control stations [43]. To date, a lot of researches have devoted to analyzing the performance of D-FACTS devices and investigating the use of them in different power system applications [29], [30], [44]–[46].

2) MTD: Recently, some pioneer works have exploited the adoption of D-FACTS devices for thwarting FDI attacks in power grid [21], [23]–[25], [27], [47]–[49]. Morrow *et al.* [23] [24] were the first researchers who proposed to perturb branch impedances for probing both the malicious and bad data in the power grid. The following works [21], [25], [27] named this branch perturbation strategy as moving target defense (MTD). Here we briefly introduce this defensive mechanism based on the DC model.

With D-FACTS devices, the defender is able to actively perturb a set of branch susceptances, an thus, increase system uncertainty for potential attackers. Supposing the susceptance of branch $k_t = \{i, j\}$ is perturbed, then we have

$$b_{ij} \rightarrow b'_{ij}, \tag{7}$$

where b'_{ij} is the susceptance of branch k_t after MTD. Actually, we cannot change b_{ij} as much as we can [23], [25], [29]. There are limits on the perturbed result, i.e.,

$$b_{ij}^{min} \leq b'_{ij} \leq b_{ij}^{max}. \tag{8}$$

As a result, the measurement matrix is changed. We assume that the control commands of MTD can be protected in the control center and transmitted through safeguarded communication channels. Thus, the attacker is unable to anticipate them. We use $\mathbf{H}'$ to denote the new measurement matrix after MTD. In fact, the attacker might try to learn the perturbed measurement matrix $\mathbf{H}'$. But the learning process usually takes a sufficiently long time (hours or days) due to the exfiltration of an enough amount of historical measurement data [27], [38]–[41]. In other words, the attacker cannot obtain the latest measurement matrix immediately. Thus, we can dynamically change the measurement matrix accordingly before it risks being exposed. The execution cycle of MTD has been discussed by authors in [25] and [27].

Since the current susceptance perturbations are not anticipated by the attacker, he/she only knows the former measurement matrix $\mathbf{H}$ (may not be the measurement matrix just before MTD). Thus, if the attacker still constructs the attack vector as $\mathbf{a} = \mathbf{Hc}$, it is possible that $\mathbf{a} \notin S(\mathbf{H}')$ after MTD. In other words, the malicious measurements $\mathbf{z}_a = \mathbf{z}' + \mathbf{Hc}$ may not bypass the BDD with $\mathbf{H}'$, where $\mathbf{z}'$ are the measurements after MTD. The defender's BDD after MTD is given by

$$r(\mathbf{z}') = \|\mathbf{z}' - \mathbf{H}'\mathbf{K}'\mathbf{z}'\|, \tag{9}$$

where $\mathbf{K}' = (\mathbf{H}'^T\mathbf{H}')^{-1}\mathbf{H}'^T$. Since $m > n$, we can prove that $\mathbf{H}'\mathbf{K}' \neq \mathbf{I}$. Therefore, $r(\mathbf{z}') = 0$ if and only if $\mathbf{z}' \in S(\mathbf{H}')$ with noiseless measurements [7]. Since $\mathbf{z}' \in S(\mathbf{H}')$, $r(\mathbf{z}_a) = 0$ if and only if $\mathbf{a} \in S(\mathbf{H}')$. This means that attack vector can bypass the BDD after MTD if and only if it belongs to the following *stealthy attack space*.

Definition 1: Let $\mathcal{A}_s$ denote the stealthy attack space. Then, $\mathcal{A}_s = S(\mathbf{H}) \cap S(\mathbf{H}')$, i.e.,

$$\mathcal{A}_s = \{\mathbf{a} \mid \mathbf{a} = \mathbf{Hc} \ \wedge\ \mathbf{a} = \mathbf{H}'\mathbf{c}',\ \mathbf{c}, \mathbf{c}' \in \mathbb{R}^n\}. \tag{10}$$

We can see that the stealthy attack space is the intersection of the column space of $\mathbf{H}$ and $\mathbf{H}'$. Therefore, the stealthy attack space is also a subspace and its dimension is given by

$$\begin{aligned} &dim(\mathcal{A}_s) \\ &\quad = dim(S(\mathbf{H}) \cap S(\mathbf{H}')) \\ &\quad = dim(S(\mathbf{H})) + dim(S(\mathbf{H}')) - dim\big(S(\mathbf{H}) \cup S(\mathbf{H}')\big) \\ &\quad = 2n - R([\mathbf{H}\ \mathbf{H}']), \end{aligned} \tag{11}$$

where $dim(\cdot)$ is the dimension of a subspace, $R(\cdot)$ is the rank of a matrix. We can see that the dimension of the stealthy attack space $\mathcal{A}_s$ is closely related to the rank of the combined

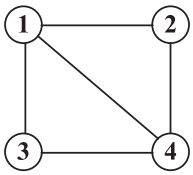

Fig. 1. A 4-bus power system with 5 branches.

matrix $[\mathbf{H}\ \mathbf{H}']$, i.e., $R([\mathbf{H}\ \mathbf{H}'])$. We define $R([\mathbf{H}\ \mathbf{H}'])$ as the security factor of MTD and denote it as γ.

B. Problem Statement

In this paper, we mainly analyze the effectiveness of MTD from two aspects: first, the capability of MTD to protect the system from any FDI attack constructed with the former measurement matrix; second, if the first capability cannot be achieved, the capability of MTD to narrow the attack opportunities of potential attackers by appropriately deploying the D-FACTS devices and setting the susceptance perturbation magnitudes.

As we know, as long as the stealthy attack space $\mathcal{A}_s$ is not $\{\mathbf{0}\}$, it is still possible for the attacker to successfully execute FDI attacks after MTD. The intuitive understanding is that, if $\exists\ \mathbf{c}, \mathbf{c}' \in \mathbb{R}^n$ satisfying $\mathbf{a} = \mathbf{Hc} = \mathbf{H}'\mathbf{c}' \neq \mathbf{0}$ [i.e., $\mathbf{a} \in S(\mathbf{H}) \cap S(\mathbf{H}')$], then the malicious measurements $\mathbf{z}_a = \mathbf{z}' + \mathbf{a}$ after MTD can circumvent the defender's BDD. We take the 4-bus power system (Fig. 1) as an example. Suppose that the branch susceptances known by the attacker are $b_{12} = -1$, $b_{13} = -2$, $b_{14} = -3$, $b_{23} = -4$ and $b_{34} = -5$. If we perturb the branches $\{1, 2\}$ and $\{2, 3\}$ as $\Delta b_{12} = 0.1$ and $\Delta b_{13} = 0.2$, the attacker still can construct an attack vector as $\mathbf{a} = [0, 0, 3c, 0, 5c]^T$ to bypass the BDD after MTD, where $c \in \mathbb{R}$ is the error injected into the state variable of bus 4. Here the dimension of the stealthy attack space is 1. In other words, the MTD may not be complete to thwart all FDI attacks with the form of $\mathbf{a} = \mathbf{Hc}$. We define the MTD's completeness as follows:

Definition 2: An MTD is said to be complete if and only if all attack vectors in $\mathcal{A}$ except the zero vector cannot bypass the BDD after MTD. This indicates that $\mathcal{A}_s = \{\mathbf{0}\}$.

Besides, once a complete MTD cannot be achieved, we need to exploit how to reduce the stealthy attack space as much as possible, thus, reducing the probability of stealthy FDI attacks after MTD. An intuitive idea is to perturb all branches in the power network. But it is an unrealistic assumption to deploy D-FACTS on all branches. On the other hand, we cannot randomly choose a set of target-perturbation branches and arbitrarily perturbed them. The impacts on the stealthy attack space might be different if we perturb different branches and set different perturbation values. For example, given the 4-bus power system shown in Fig. 1, Table I shows the change of the dimension of the attack space when we perturb different branches and set different perturbation values. From row 1 and row 2, we can see that the dimensions of the stealthy attack space are different when we perturb different branches. Form row 3 and row 4, we can see that the dimensions of the stealthy attack space are different when we perturb branch $\{3, 4\}$ to different values.

Moreover, due to the physical limitation of D-FACTS devices [30] and their induced operation costs [25], [27], the branch susceptance cannot be perturbed to any value. Therefore, a natural question emerges that, given a set of D-FACTS devices, how do they affect the stealthy attack space and the operation cost when we deploy them on different branches and set them with different values? This is a critical issue for us to carry out MTD.

TABLE I
THE DIMENSION OF THE STEALTHY ATTACK SPACE AFTER MTD

Branch Perturbations	$dim(\mathcal{A}_s)$
$\Delta b_{12} = 0.1, \Delta b_{13} = 0.1$	1
$\Delta b_{12} = 0.1, \Delta b_{23} = 0.1$	2
$\Delta b_{13} = 0.1, \Delta b_{14} = 0.3, \Delta b_{34} = 0.5$	2
$\Delta b_{13} = 0.1, \Delta b_{14} = 0.3, \Delta b_{34} = 0.2$	1

IV. Analysis of MTD to Thwart FDI Attacks

In this section, first, we analyze the detection of FDI attacks constructed with the former measurement matrix using MTD. Second, we analyze the completeness of MTD and discuss physical limitations of the power network to achieve this property. Third, we investigate the impact of the susceptance perturbation magnitude on the dimension of the stealthy attack space. Lastly, we provide guidance on effective MTD for minimizing the dimension of the stealthy attack space, maximizing the number of covered buses, and reducing the operation cost.

A. Detecting FDI Attacks With MTD

First of all, we discuss the detection of FDI attacks constructed with the former measurement matrix.

After MTD, from the measurement matrix $\mathbf{H}$, we select u columns which form a submatrix $\mathbf{H}^u$ such that $\mathbf{H}^u$ can be linearly represented by $\mathbf{H}'$. The rest v ($v = n - u$) columns in $\mathbf{H}$ form a submatrix $\mathbf{H}^v$, i.e., $\mathbf{H} = [\mathbf{H}^u\ \mathbf{H}^v]$. Assume that the attacker constructs an attack vector as $\mathbf{a} = \mathbf{Hc}$ with the purpose to bypass the BDD after MTD. Actually, $\mathbf{a} = \mathbf{H}^u\mathbf{c}_u + \mathbf{H}^v\mathbf{c}_v$, where $\mathbf{c}_u \in \mathbb{R}^u$ and $\mathbf{c}_v \in \mathbb{R}^v$. Then, we have the following conclusion.

Proposition 3: The BDD after MTD can detect the attack vector $\mathbf{a} = \mathbf{Hc}$ only if $\mathbf{c}_v \neq \mathbf{0}_{1\times v}$, where $\mathbf{0}_{1\times v}$ is a v-dimension zero vector.

Proposition 3 presents a necessary condition for the detection of FDI attacks constructed with the former measurement matrix. It also indicates the soundness of MTD. That is, if the BDD after MTD can detect the attack vector $\mathbf{a}$, then at least one of the state variables in $\mathbf{c}_v$ is modified. With this condition, if $\mathbf{H}^v$ cannot be linearly represented by $\mathbf{H}'$, then there might not exist $\mathbf{c}' \in \mathbb{R}^n$ such that $\mathbf{H}^v\mathbf{c}_v = \mathbf{H}'\mathbf{c}'$ for an arbitrarily selected $\mathbf{c}_v$. Thus, at most v state variables corresponding to $\mathbf{c}_v$ cannot be independently modified by the attacker. Since $v \geq n - dim(\mathcal{A}_s)$, if the dimension of the stealthy attack space $dim(\mathcal{A}_s)$ is smaller, the more state variables cannot be independently modified by the attacker after MTD. However, note that the attackers cannot anticipate/predict the new measurement matrix $\mathbf{H}'$, the "blind" attacker might try his/her luck to execute FDI attacks, and thus, increase the detection probability of these attacks with the BDD after MTD.

Specifically, to illustrate the effectiveness of MTD for thwarting FDI attacks, we give an illustration example (Fig. 2) about the change of the stealthy attack space with respect to the security factor γ. We can see that, the stealthy attack space is equal to the column space of $\mathbf{H}$ and $\mathbf{H}'$ $[S(\mathbf{H}) = S(\mathbf{H}')]$ when the security factor γ is n; the stealthy attack space is the intersection of the column space of $\mathbf{H}$ and $\mathbf{H}'$ $[S(\mathbf{H}) \neq S(\mathbf{H}')]$

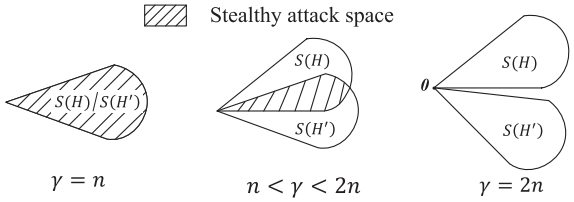

Fig. 2. An illustration example about the change of the stealthy attack space $\mathcal{A}_s$ with the security factor γ.

when the security factor γ is between n and $2n$; the stealthy attack space is $\{\mathbf{0}\}$ when the security factor γ is $2n$. Therefore, if the security factor γ is smaller, the "volume" of the stealthy attack space is smaller. A smaller stealthy attack space means a smaller probability of stealthy FDI attacks constructed with the former measurement matrix. In other words, after MTD, there are less successful opportunities for the attacker to execute stealthy FDI attacks with the old system information. Thus, we claim that MTD is more effective to thwart stealthy FDI attacks. For example, when $\gamma = 2n$, after MTD, no attack vector $\mathbf{a} = \mathbf{Hc}$ can bypass BDD, which means that the probability of stealthy FDI attacks constructed with the former measurement matrix is zero. We will evaluate the effectiveness of MTD with respect to the dimension of the stealthy attack space in Section V-A2.

B. Analysis of MTD's Completeness

Here we first give a sufficient and necessary condition for achieving a complete MTD mathematically. Then, we present physical constraints for achieving this property.

1) Mathematical Analysis: In fact, an MTD is not complete because the BDD after MTD misses detecting some FDI attacks, i.e., $\mathcal{A}_s \neq \{\mathbf{0}\}$. According to the Definition 2, to achieve a complete MTD, we need to make $\mathcal{A}_s = \{\mathbf{0}\}$. Since the subspace spanned by the zero vector has zero dimension, the dimension of $\mathcal{A}_s$ is zero if an MTD is complete. Thus, we have the following conclusion.

Proposition 4: An MTD is complete if and only if the security factor $\gamma = 2n$.

Proposition 4 presents a sufficient and necessary condition for achieving a complete MTD mathematically. Note that if $\gamma = R([\mathbf{H}\ \mathbf{H}']) = 2n$, the intersection of $S(\mathbf{H})$ and $S(\mathbf{H}')$ is $\{\mathbf{0}\}$. Thus, all FDI attacks constructed as $\mathbf{a} = \mathbf{Hc}$ ($\mathbf{a} \neq \mathbf{0}$) can be detected by the BDD after MTD.

Remark 5: Moreover, if an MTD is complete, then we can identify the attack vector constructed with the former measurement matrix. Suppose the attack vector is $\mathbf{a} = \mathbf{Hc}$. Then, the malicious measurements are $\mathbf{z}_a = \mathbf{H}'\mathbf{x}' + \mathbf{Hc} = [\mathbf{H}'\ \mathbf{H}][\mathbf{x}'\ \mathbf{c}]^T$, where $\mathbf{x}'$ denotes the state variables without FDI attacks. Since the combined matrix $[\mathbf{H}'\ \mathbf{H}]$ has full column rank, we can uniquely solve the variables $\mathbf{c}$ with the malicious measurements. Thus, we can identify the injected errors of the state variables.

2) Physical Limitations: Essentially, whether we can construct a complete MTD or not depends on the structure of the power network and the perturbed branches. Considering the fully measured case, the measurement matrix before and after MTD are

$$\mathbf{H} = \begin{bmatrix} \mathbf{A}^T\mathbf{DA} \\ \mathbf{DA} \\ -\mathbf{DA} \end{bmatrix} \quad \mathbf{H}' = \begin{bmatrix} \mathbf{A}^T\mathbf{D}'\mathbf{A} \\ \mathbf{D}'\mathbf{A} \\ -\mathbf{D}'\mathbf{A} \end{bmatrix}. \tag{12}$$

The combined matrix of $\mathbf{H}$ and $\mathbf{H}'$ can be written as $\mathbf{C} = [\mathbf{H}\ \mathbf{H}']$. Let $\widetilde{\mathbf{S}} = \left[\mathbf{S}\ \mathbf{S}'\right] = \left[\mathbf{DA}\ \mathbf{D}'\mathbf{A}\right]$, we have

$$\mathbf{C} = \begin{bmatrix} \mathbf{A}^T\widetilde{\mathbf{S}} \\ \widetilde{\mathbf{S}} \\ -\widetilde{\mathbf{S}} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T & \mathbf{0}_{n\times l} & \mathbf{0}_{n\times l} \\ \mathbf{0}_{l\times l} & \mathbf{I}_{l\times l} & \mathbf{0}_{l\times l} \\ \mathbf{0}_{l\times l} & \mathbf{0}_{l\times l} & -\mathbf{I}_{l\times l} \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{S}} \\ \widetilde{\mathbf{S}} \\ \widetilde{\mathbf{S}} \end{bmatrix}, \tag{13}$$

where $\mathbf{0}_{n\times l}$ is an n by l zero matrix, $\mathbf{I}_{l\times l}$ is an l by l identity matrix. Therefore, we can derive that $R(\mathbf{C}) \leq R(\widetilde{\mathbf{S}})$. Since $\widetilde{\mathbf{S}}$ is an l by $2n$ matrix, we have $R(\widetilde{\mathbf{S}}) \leq min\{l, 2n\}$. Therefore, only if $l \geq 2n$ can we make $R(\mathbf{C}) = 2n$. In other words, in order to make the security factor $\gamma = R(\mathbf{C}) = 2n$, the number of branches of the power transmission system must be larger than or equal to $2n$. We observe that this condition must be satisfied under the other measured cases (i.e., the system is not necessarily fully measured) as well.

Theorem 6: An MTD is complete only if the following two conditions are satisfied:

- $l \geq 2n$, where l is the number of branches and $n + 1$ is the number of buses in the power system;
- *The perturbed branches must cover all buses;*

Proof: Please see Appendix A.

Theorem 6 provides a necessary condition for realizing a complete MTD. That is, the completeness of MTD can be achieved only if the system topology and the perturbed branches meet certain requirements. Besides, we can derive another two points from the second condition. First, to cover all buses, at least n branches should be perturbed. Second, once a bus is not covered by the perturbed branches, the state variable of this bus can be modified stealthily if the attacker happens to attack it only. That is, let $\mathbf{a} = \mathbf{Hc}$ be an attack vector. Then, if $c_i = 0$ ($c_i \in \mathbf{c}$) for any $i \in \mathcal{M}_{mtd}$, then the malicious measurements $\mathbf{z}_a = \mathbf{z}' + \mathbf{a}$ can bypass the BDD after MTD, where $\mathcal{M}_{mtd}$ is a non-empty set of buses that are covered by the perturbed branches. Thus, $c_i \neq 0$ ($i \notin \mathcal{M}_{mtd}$) can be any values. This indicates that the attacker can arbitrarily modify the state variables of buses that are not covered by the perturbed branches. Therefore, for a complete MTD, we need to ensure that: (i) the power transmission system has more than $2n$ branches; (ii) more than n branches are perturbed; (iii) the perturbed branches cover all buses. For example, the 4-bus power system can support the realization of a complete MTD, because it has 6 branches ($2n = 2 * 3 = 6$), while the 4-bus power system with 5 branches shown in Fig. 4 cannot. If any one of the above three conditions is not satisfied, we cannot make an MTD complete.

Considering the first condition, if the power system has less than $2n$ branches, then the dimension of the stealthy attack space satisfies

$$dim(\mathcal{A}_s) = 2n - R([\mathbf{H}\ \mathbf{H}']) \geq 2n - R(\widetilde{\mathbf{S}}) \geq 2n - l. \tag{14}$$

That is, the smallest dimension of the stealthy attack space after MTD is larger than or equal to $2n - l$.

Remark 7: According to the aforementioned analysis, we find that is difficult to realize a complete MTD. For a complete MTD, the power transmission system must have at least 2n branches (i.e., $l \geq 2n$), and more than n branches, which cover all buses, should be perturbed. These conditions may not be held in practice. Particularly, we examine the number of branches in the IEEE test power systems provided in MATPOWER [50]; only three of all 41 cases have more than

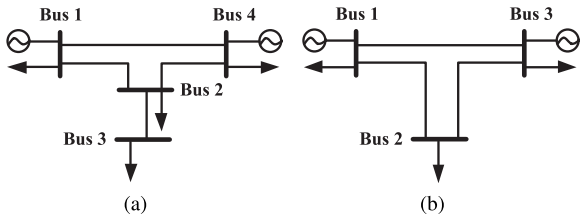

Fig. 3. (a) A 4-bus power system with 4 branches; (b) A 3-bus power system.

2n branches, namely case6ww (11 branches), case89pegase (210 branches) and case145 (453 branches). What's more, we discover that we can never make an MTD complete if the the power transmission system has a bus that is only connected by a single branch.

Theorem 8: It is impossible to make an MTD complete if the power transmission system contains a bus that is only connected by a single branch.

Proof: Please see Appendix B.

This special case of the power network structure limits the realization of a complete MTD. Besides, we find that the state variable of the bus that is only connected by a single branch can be arbitrarily modified by the attacker. Let t be the bus that is only connected by a single branch. Then, for any $c_t \in \mathbb{R}$, there exists $c_t' \in \mathbb{R}$ such that $\mathbf{h}_t c_t = \mathbf{h}_t' c_t'$, where $\mathbf{h}_t$ and $\mathbf{h}_t'$ are respective the tth column of the matrix $\mathbf{H}$ and $\mathbf{H}'$. Therefore, if the attack vector is $\mathbf{a} = \mathbf{h}_t c_t$, then $\mathbf{a} \in S(\mathbf{H}')$ always holds, which can definitely bypass the BDD after MTD. We present a simple example to illustrate that. In the 4-bus power system shown in Fig. 3(a), the state variable of bus 3 can be arbitrarily modified because it is only connected by branch $\{2, 3\}$. Therefore, as long as the system contains a bus that is only connected by a single branch, we can never realize a complete MTD.

3) Discussion: Nevertheless, it might not be necessary to achieve a complete MTD in some power systems. We take the 3-bus power system shown in Fig. 3(b) as an example. Even though we cannot realize a complete MTD because it only has 3 branches, which is less than $2 * 2 = 4$, we can protect all state variables from being arbitrarily modified by the attacker when we perturb the branches $\{1, 2\}$ and $\{2, 3\}$. Suppose the perturbations are Δb_{12} and Δb_{23}, and the attack vector is $\mathbf{a} = \mathbf{Hc}$, where $\mathbf{c} = [c_1\ c_2]^T$. If this attack can bypass the BDD after MTD, c_1 and c_2 must satisfy

$$\frac{c_2}{c_1} = \frac{b_{12}\Delta b_{23} - b_{23}\Delta b_{12}}{b_{12}\Delta b_{23} + \Delta b_{12}\Delta b_{23}}. \tag{15}$$

Therefore, the attacker must know the susceptance perturbations Δb_{12} and Δb_{23} to construct a coordinated attack vector $\mathbf{a}$. We can see that is almost impossible for the attacker to construct such an attack vector. Definitely, a complete MTD can thwart any attack vector $\mathbf{a} = \mathbf{Hc}$ when $\mathbf{H}$ is very different from the measurement matrix $\mathbf{H}'$ after MTD. But it might not be a necessary condition for some power systems.

C. Impact of the Susceptance Perturbation Magnitude on the Dimension of the Stealthy Attack Space

A practical and effective way to enhance the security of the power system is to reduce the stealthy attack space $\mathcal{A}_s$ and cover as many buses as possible with MTD, thus reducing the attack opportunities of potential attackers. We start by investigating the impact of the susceptance perturbation magnitude on the dimension of the stealthy attack space when one more branch susceptance is perturbed. Since the security factor γ [i.e., $R([\mathbf{H}\ \mathbf{H}'])$] determines $dim(\mathcal{A}_s)$, we focus on the change of γ in the following. As $\mathbf{H}' = \mathbf{H} + \Delta\mathbf{H}$, based on the sparse property of $\Delta\mathbf{H}$, we first present three cases about the change of γ when one more branch is perturbed. Second, we investigate the impact of the susceptance perturbation magnitude on the value of γ with these three cases.

Before go deeper into the analysis, we prove that $\Delta\mathbf{H}$ has a sparse structure.

Proposition 9: Suppose $\mathcal{L}_d$ is a set of perturbed branches. $\Delta\mathbf{H}$ is a sparse matrix with non-zero elements in the i_dth and j_dth columns, with $k_d = \{i_d, j_d\} \in \mathcal{L}_d$.

Proof: Please see Appendix C.

Based on the sparse structure of the matrix $\Delta\mathbf{H}$, we can draw a conclusion about the changing range of γ when one more branch is perturbed. Let $\mathbf{C} = [\mathbf{H}\ \mathbf{H}']$ and $\mathbf{C}' = [\mathbf{H}\ \mathbf{H}'']$ be the combined matrices before and after the perturbation of a new branch, respectively. We denote $\Delta\gamma = R(\mathbf{C}') - R(\mathbf{C})$ as the change of the security factor γ. Then, we obtain the following result.

Proposition 10: When perturbing one more branch, $\Delta\gamma$ changes -1, 0 or 1.

Proof: Please see Appendix D.

That is, there are three possible changes of the security factor γ when one more branch is perturbed: increasing 1, remaining the same or decreasing 1 (i.e., $\Delta\gamma = 1, 0$ or -1). Thus, a natural question emerges that, whether the value of $\Delta\gamma$ will change with the susceptance perturbation magnitude. Let $k_d = \{i_d, j_d\}$ be the new branch whose susceptance $b_{i_d j_d}$ is perturbed to be $\lambda b_{i_d j_d}$ with $\lambda > 0$ and $\lambda \neq 1$. We define λ as the perturbation ratio. Suppose $\mathcal{M}_{mtd}^q$ is an index set of q columns in $\mathbf{H}'$ that form a submatrix $\mathbf{H}^q$ such that $R([\mathbf{H}\ \mathbf{H}^q]) = R([\mathbf{H}\ \mathbf{H}']) = n+q$. The rest columns in $\mathbf{H}'$ form a submatrix $\mathbf{H}^p$, i.e., $\mathbf{H}' = [\mathbf{H}^q\ \mathbf{H}^p]$. We define $\mathcal{M}_{mtd}^q$ as a security set. Overall, there are three cases should be considered for the impacts on $\Delta\gamma$ when perturbing one more branch:

- ***Case 1:*** Neither bus i_d nor j_d are contained in the security set. That is, $i_d \notin \mathcal{M}_{mtd}^q$ and $j_d \notin \mathcal{M}_{mtd}^q$;
- ***Case 2:*** Either bus i_d or j_d is contained in the security set. That is, $i_d \in \mathcal{M}_{mtd}^q$ and $j_d \notin \mathcal{M}_{mtd}^q$, or $i_d \notin \mathcal{M}_{mtd}^q$ and $j_d \notin \mathcal{M}_{mtd}^q$;
- ***Case 3:*** Both bus i_d and j_d are contained in the security set. That is, $i_d \in \mathcal{M}_{mtd}^q$ and $j_d \in \mathcal{M}_{mtd}^q$.

For each case, we analyze the change of $\Delta\gamma$ by varying the perturbation ratio λ. Since there are limits on the susceptance perturbation [29], we assume $\lambda_{min} \leq \lambda \leq \lambda_{max}$. We find that the magnitude of the susceptance perturbation almost does not affect the change of the security factor. Especially, if the security factor increases 1 (i.e., $\Delta\gamma = 1$), this result remains the same regardless of the perturbation magnitude. We present the details of the analysis in the following.

*Proposition 11: Under **Case 1**, if $i_d \notin \mathcal{M}_{mtd}^q$ and $j_d \notin \mathcal{M}_{mtd}^q$, then $\Delta\gamma$ remains the same regardless of the change of λ.*

Proof: Please see Appendix E.

Proposition 11 implies that, under ***Case 1***, the perturbation magnitude does not affect the change of $dim(\mathcal{A}_s)$. Therefore, if we find that the dimension of the stealthy attack space

$dim(\mathcal{A}_s)$ decreses 1 when we perturb a branch to a certain value, then this result will not change if we perturb the branch to the other values. That is, we can determine the value of $dim(\mathcal{A}_s)$ by only testing one perturbation ratio under ***Case 1***. Considering the other two cases, we obtain the following result.

*Proposition 12: Under **Case 2** and **Case 3**, only if there exists a value* λ^* ($\lambda_{min} \leq \lambda^* \leq \lambda_{max}$, $\lambda^* \neq 1$) *and* $\lambda = \lambda^*$, *we obtain* $\Delta\gamma = -1$.

Proof: Please see Appendix F.

Proposition 12 implies that, under ***Case 2*** and ***Case 3***, only if there exists an unique perturbation ratio and the target-perturbation branch is perturbed to that value, the dimension of the stealthy attack space $dim(\mathcal{A}_s)$ increases 1. Based on this result, if $\Delta\gamma = 0$ when we perturb a target-perturbation branch to a certain value, then this result almost remains the same regardless of the perturbation magnitude. Further, we obtain a result about the increase of the security factor.

Theorem 13: Under all cases, $\Delta\gamma = 1$ *for any* $\lambda > 0$ ($\lambda \neq 1$) *if and only if there exists* $\lambda = \lambda^*$ ($\lambda^* \neq 1$) *such that* $\Delta\gamma = 1$.

Proof: The proof is similar to that for Proposition 12.

Theorem 13 implies that if the security factor γ increases 1 when we perturb one more branch, then this result will not be changed with the variation of the susceptance perturbation magnitude. In other words, once we find that the dimension of the stealthy attack space $dim(\mathcal{A}_s)$ decreases 1 in a trial, then this result will not change in the other trials. Therefore, we can determine the security factor γ using only one tested perturbation ratio when $\Delta\gamma = 1$.

In summary, the susceptance perturbation magnitude almost does not affect the change of the dimension of the stealthy attack space $dim(\mathcal{A}_s)$ when one more branch is perturbed. Thus, we can determine the dimension of the stealthy attack space after we have perturbed the branches to certain ratios, without worrying about the impact of the perturbation magnitude, especially the case when $\Delta\gamma = 1$. Note that this is very useful for the selection of the target-perturbation branches. We present details about this issue in the next subsection.

D. Guidance on the Construction of an Effective MTD

In this section, we first consider the optimal selection of the set of target-perturbation branches. Then, we discuss the increasing generation cost caused by MTD.

1) Target-Perturbation Branch Selection: Here we propose an algorithm for maximizing the security factor γ (i.e., minimizing the dimension of the stealthy attack space $\mathcal{A}_s$) and covering the largest number of buses, with a given number of D-FACTS devices. We briefly introduce the algorithm in the following.

For the sake of power transfer quantity and quality, the branch parameters of some branches cannot be perturbed. Therefore, the set of candidate branches input to Algorithm 1 may not be the set of all branches in the power network. Thus, we only need to seek for an optimal deployment strategy in the set of branches that can be perturbed. With the loop (form line 2 to line 12) in Algorithm 1, we traverse all perturbable branches until we obtain the maximum γ. This process computes a set of target-perturbation branches that can minimize the stealthy attack space. In the loop, we determine whether the branch should be selected or not from line 8 to line 11. Once $\Delta\gamma = 1$, the branch is selected as a target-perturbation branch; otherwise, it is not. When the iterations in Algorithm 1 have finished, we determine whether we should go into Algorithm 2 with the condition given in line 13. That is, if γ is maximized and all the D-FACTS devices have been used, then the algorithm is closed and the output $\mathcal{L}_d^1$ is the set of target-perturbation branches. If γ is maximized but there still exist unused D-FACTS devices, the algorithm enters Algorithm 2 for maximizing the covered buses.

Algorithm 1 Minimizing the Dimension of the Stealthy Attack Space

Data: The number of D-FACTS devices n_d; The set of branches that can be perturbed $\mathcal{L}_p$
1 Initialization: randomly rearrange $\mathcal{L}_p$ as $\mathcal{L}_p = \{k'_1, k'_2, \cdots, k'_p\}$, $r = 0$, $r' = 0$;
2 **for** *each* $k'_t \in \mathcal{L}_p$ **do**
3 Perturb k'_t to a random ratio λ;
4 **if** $r = n_d$ **then**
5 break ;
6 **end**
7 Compute $\Delta\gamma = R([\mathbf{H}\ \mathbf{H}'']) - R([\mathbf{H}\ \mathbf{H}'])$;
8 **if** $\Delta\gamma = 1$ **then**
9 put k'_t in the branch set $\mathcal{L}_d^1$;
10 $r = r + 1$;
11 **end**
12 **end**
13 **if** $r < n_d$ **then**
14 Enter Algorithm 2;
15 $\mathcal{L}_d = \mathcal{L}_d^1 \cup \mathcal{L}_d^2$;
16 **else**
17 $\mathcal{L}_d = \mathcal{L}_d^1$;
18 **end**
Result: The set of target-perturbation branches $\mathcal{L}_d$

Algorithm 2 starts with the rest candidate branches except those have been selected for maximizing γ. It incrementally searches for branches that cover new buses. The target-perturbation branch is selected according to the following process. First of all, we select the branches that cover two new buses (line 7 to line 10). Then, if there still exist unused D-FACTS devices, we select the branches that cover one new bus (line 16 and line 17). At last, if there still exist unused D-FACTS devices, we select the branches from the rest candidate branches (line 19 to line 21).

The computation time requirement for the rank operation of $\Delta\gamma = R([\mathbf{H}\ \mathbf{H}'']) - R([\mathbf{H}\ \mathbf{H}'])$ is $\mathcal{O}(mn^2)$, where m is the number of measurements and n is the number of buses in the power network. The runtime for Algorithm 2 is $\mathcal{O}(l)$, where l is the number of branches in the power network. Considering the worst case that the rank operation would be executed l times in the loop, the time complexity for Algorithm 1 is $\mathcal{O}(lmn^2)$. Note that Algorithm 2 is not executed if the condition given in line 13 of Algorithm 1 is not satisfied. In fact, the above algorithm only outputs an alternative set of target-perturbation branches. There may be a lot of candidates that are also satisfied (see Section V-B1). Therefore, we can dynamically change the perturbed branches and maintain the effectiveness of MTD. On the other hand, we can realize MTD on the basis of already deployed D-FACTS devices for saving the additional infrastructure cost.

Algorithm 2 Covering the Largest Number of Buses

```
Compute: L'_p = L_p \ L^1_d; the set of buses I covered by branches in L^1_d;
for each k''_t = {i''_t, j''_t} ∈ L'_p do
    if r + r' = n_d then
        break ;
    end
    if i''_t or j''_t ∉ I then
        if i''_t and j''_t ∉ I then
            put k''_t in the branch set L^2_d ;
            r' = r' + 1 ;
        else
            put k''_t in the branch set L^3_d ;
        end
    end
end
if r + r' < n_d then
    if n_d − r − r' ≤ | L^3_d | then
        Select n_d − r − r' branches from L^3_d and put them into the branch set L^2_d;
    else
        Put all branches in L^3_d into L^2_d;
        Select n_d − r − r' − | L^3_d | branches from L'_p \ L^2_d and put them into L^2_d;
    end
end
Return L^2_d ;
```

2) Reducing the Operation Cost: After determining the set of branches that should be perturbed, a following problem that we should consider is to reduce the operation cost by appropriately setting their perturbation magnitudes. Optimal power flow (OPF) seeks to optimize the operation of an electric power system subject to the physical constraints imposed by electrical laws and engineering limits [51]. It outputs the minimum generation cost for the given loads by adjusting the power flows. This generation cost can represent the operation cost caused by the system changes of MTD. Here, OPF is stated as follows:

$$\begin{aligned} \min_{p_i^g, i\ \in\mathcal{G}} \ & \sum_{i\in\mathcal{G}} C_i(p_i^g) \\ s.t.\ & p_i^g - p_i^l = \mathbf{B}\boldsymbol{\theta}, \quad i \in \mathcal{N} \\ & p_i^{min} \le p_i^g \le p_i^{max}, \quad i \in \mathcal{G} \\ & f_{ij}^{min} \le f_{ij} \le f_{ij}^{max}, \quad \{i,j\} \in \mathcal{L} \\ & \lambda_{ij}^{min} \le \lambda_{ij} \le \lambda_{ij}^{max}, \{i,j\} \ \in \mathcal{L} \end{aligned} \tag{16}$$

where $C_i(p_i^g)$ is the cost function, $\mathcal{G}$ is the set of generators, p_i^g is the real output power, p_i^l is the load, f_{ij} is the branch active power flow, λ_{ij} is the susceptance perturbation ratio. In the above optimization problem, the decision variable is the output generation level of each generator, and the cost function is a quadratic function. The first constraint is about the nodal power balance constraint, i.e., the power injections must be equal to the power consumptions. The second and third constraints are about the limits about the output generation and branch active power flow, respectively. The fourth constraint is newly added here for the limits of the susceptance perturbation. We can see that the matrix $\mathbf{B}$ contained in the first constraint contains the branch perturbation parameters. Therefore, this optimization problem is correlated with the branch perturbation magnitude. Since the objective function is convex and the constraints are differentiable, the OPF problem is a typically convex optimization problem [32], [52], which can be solved by the non-linear programming solver *fmincon* in MATLAB.

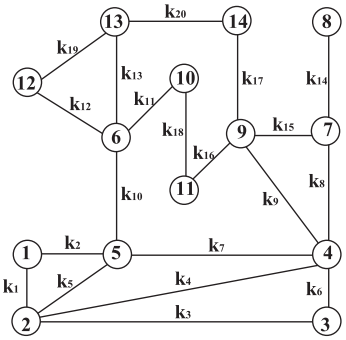

Fig. 4. The IEEE 14-bus power system.

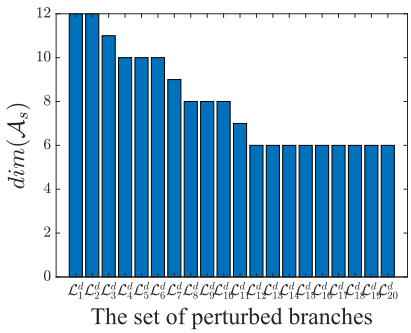

Fig. 5. The dimension change of the stealthy attack space $\mathcal{A}_s$ with respect to the perturbed branches.

V. Simulation Results

In this section, we evaluate our findings about MTD using an illustrative IEEE 14-bus power system (Fig. 4) and the IEEE 30-bus, 57-bus, 118-bus and 145-bus power systems. All simulations are based on the fully measured power system and carried out in MATLAB. We assume that all meters are subject to the same noise distribution, namely the Normal distribution $\mathcal{N}(0, \sigma^2)$, if the measurement noises are considered. The threshold τ (see Section II-A) is set as $\sigma\sqrt{\chi^2_{m-n,\alpha}}$ [32], where $m-n$ is the freedom degree of the Chi-square distribution and α is the false alarm rate (which is 0.05). In practice, each branch susceptance must be within given limits, namely $b_{ij}^{min} \le b_{ij} \le b_{ij}^{max}$ [25]. The authors in [23] have proved that if the perturbations are within 20% of the impedance, then there are sufficiently large number of perturbation cases that can restrict power losses within 1%. Thus, it is feasible to perturb the branch susceptances within 20% maximum change. In this paper, for the perturbation ratio λ, it is constrained to be within [0.8, 1.2].

A. Effectiveness of MTD

1) The Dimension of the Stealthy Attack Space vs. Branch Perturbations: Taking the IEEE 14-bus power system (Fig. 4) as an example, we analyze the dimension change of the stealthy attack space ($\mathcal{A}_s$) when we increase the number of perturbed branches. We successively perturb the set of branches as $\mathcal{L}_1^d = \{k_1\}$, $\mathcal{L}_2^d = \{k_1, k_2\}$, $\mathcal{L}_3^d = \{k_1, k_2, k_3\}$, $\cdots$. Fig. 5

TABLE II
THE DIMENSION OF $\mathcal{A}_s$ AFTER PERTURBING ALL BRANCHES

IEEE test system	30-bus	57-bus	118-bus	145-bus
# branches	41	80	186	453
$dim(\mathcal{A}_s)$	18	31	40	10

shows the simulation results. We can see that the dimension of the stealthy attack space decreases when more branches are perturbed. But $dim(\mathcal{A}_s)$ cannot reach 0, because the 14-bus power system only has 20 branches, which is less than $2n = 26$. Even though we perturb all branches, the dimension of the stealthy attack space is 6, which is the smallest stealthy attack space we can achieve in this 14-bus power system. Consistent with the equation (14), the smallest dimension of the stealthy attack space is equal to $2n - l = 2 * 13 - 20 = 6$.

Moreover, we perturb all branches in the IEEE 30-bus, 57-bus, 118-bus and 145-bus power systems. Table II shows the smallest dimension of the stealthy attack space we obtain. We also give the number of branches for each power system in this table. We can see that all the smallest dimensions of the stealthy attack space are not zero, which indicates that none of the power systems support constructing complete MTD. Even though the IEEE 145-bus power system has 453 branches, which is larger than $2n = 288$, the dimension of the stealthy attack space is 10 after perturbing all its branches. In our opinion, one reason for this result is that the IEEE 145-bus power system contains 7 buses that are only connected by a single branch. The simulation result indicates that is difficult to meet the conditions required for achieving a complete MTD in practice.

2) MTD's Effectiveness for Thwarting FDI Attacks: Next, we exploit MTD's effectiveness for thwarting FDI attacks with respect to the dimension of the stealthy attack space. The attack vector is constructed with the form of $\mathbf{a} = \mathbf{Hc}$, where $\mathbf{H}$ is a measurement matrix before MTD. We sample the value of the element c_i in $\mathbf{c}$ from a uniform distribution $\mathcal{U}(-dm, dm)$, where dm is the maximum magnitude of the injected bias into the state variable. Here dm is 0.1. The modified state variables are uniformly selected from the bus set, i.e., the non-zero elements in $\mathbf{c}$ are uniformly selected. We assume that all branches in the power system can be perturbed, and the measurements are noiseless. The perturbation ratio is randomly chosen within [0.8, 1.2]. For each setting, we repeat random attacks for 1000 times based on Monte Carlo simulations, and estimate the detection probability of the FDI attack constructed with $\mathbf{H}$ as

$$\Pr = \frac{\#\text{ of detected trials}}{1000}, \tag{17}$$

where *# of detected trials* means the number of FDI attack vectors detected by the BDD after MTD. For consistency, we use the metric $\frac{dim(\mathcal{A}_s)}{n}$ to measure the change of the dimension of the stealthy attack space. Note that for a fixed dimension of the stealthy attack space, there are several branch-perturbation schemes for realizing MTD.

We use the IEEE 14-bus, 30-bus and 57-bus power systems with default settings and data in MATPOWER. Fig. 6 shows the simulation results. We can see that if the dimension of the stealthy attack space is smaller, the detection probability of FDI attacks is larger. For example, in the 14-bus power system,

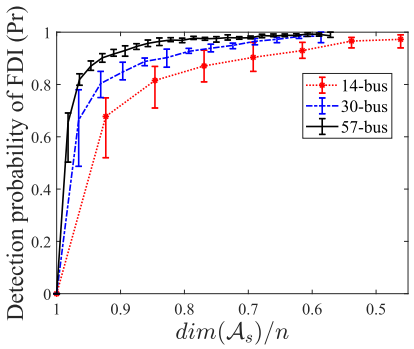

Fig. 6. The detection probability of FDI attacks with respect to the dimension of the stealthy attack space.

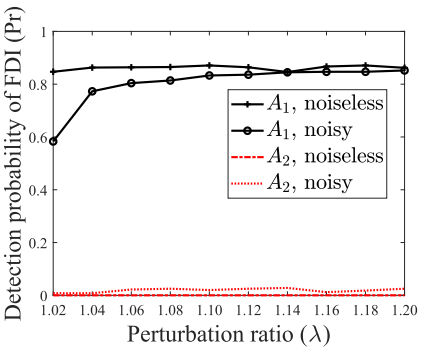

Fig. 7. The detection probability of FDI attacks with $c_i \neq 0$ for $i \in \mathcal{M}_{mtd}$ (A_1) and $c_i = 0$ for all $i \in \mathcal{M}_{mtd}$ (A_2).

the detection probability of FDI is more than 90% when $\frac{dim(\mathcal{A}_s)}{n}$ is reduced to 0.6. Moreover, we find that if the system size is larger, the curve is more smooth and the detection probability of FDI attacks is larger given the same $\frac{dim(\mathcal{A}_s)}{n}$. The simulation results highlight that the smaller the dimension of the stealthy attack space is, the better performance the MTD achieves in terms of thwarting FDI attacks.

Moreover, we investigate the impact of FDI attacks constructed with the former measurement matrix on the buses that are not covered by the perturbed branches. We select k_1, k_3, k_4 and k_7 as the target-perturbation branches. They are perturbed to the same ratio during simulations. The perturbation ratios are set within [1.02, 1.2] spaced by 0.02. The attack vector $\mathbf{a} = \mathbf{Hc}$ is constructed by considering two cases: $c_i \neq 0$ for $i \in \mathcal{M}_{mtd}$ and $c_i = 0$ for all $i \in \mathcal{M}_{mtd}$, where $\mathcal{M}_{mtd}$ is a set of buses that are covered by the perturbed branches. We denote these two cases as "A_1" and "A_2," respectively. The measurement noise is fixed as $\sigma^2 = 0.01$. For each setting, we repeat the simulation for 1000 times. For each time, we uniformly selected the non-zero elements in $\mathbf{c}$. The simulation results are given in Fig. 7. We can see that, for case A_1, the detection probability of the FDI attack is more than 84% when the measurements are noiseless. We believe that this large detection probabilities is because $\mathbf{a} \notin S(\mathbf{H}')$ holds for most cases. When the measurements are noisy in case A_1, the detection probability of FDI attacks is disturbed by the noise. But it is still more than 60% and increases with the perturbation ratio. For case A_2, we can see that, whether the measurements are noisy or not, the detection probability of FDI attacks is almost negligible. This indicates that the buses that are not covered by the perturbed branches are vulnerable to FDI attacks. Therefore, to improve the effectiveness of MTD, the perturbed branches should cover as many buses as possible.

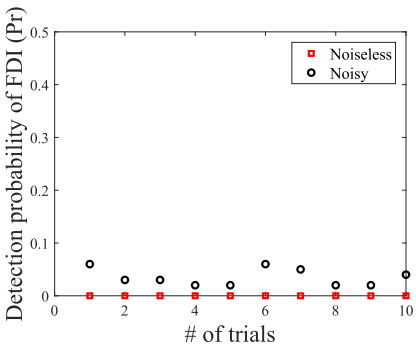

Fig. 8. The detection probability of FDI attacks vs. different sets of perturbed branches with only bus 8 is attacked.

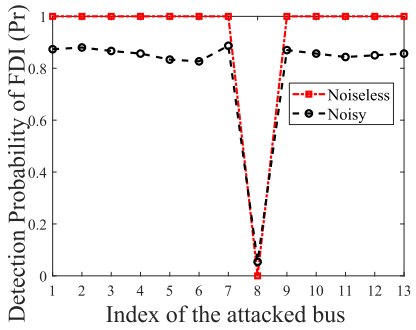

Fig. 9. The attacks on bus 8 can always bypass the BDD after MTD in the IEEE 14-bus power system.

3) Impact of the Bus That Is Only Connected by a Single Branch: Next, we consider the attack on the bus that is only connected by a single branch. Here we validate our finding using the IEEE 14-bus power system. We find that bus 8 is only connected by branch $\{7, 8\}$. First, we only inject errors into the state variable of bus 8. We test this attack for 10 trials. For each trial, we select 1000 sets of perturbed branches except for branch $\{7, 8\}$ randomly. The noise variance is fixed as 0.01. Fig. 8 shows the simulation results. We can see that the attacks on bus 8 can always bypass the BDD after MTD when the measurements are noiseless. And for the noisy case, the detection probability of this attack is almost negligible (around 0.05). By contrast, we conduct another 13 trials but focus on attacking different buses. For the ith trial, we only modify the state variable of bus i. In these trials, we perturb all branch susceptances. Fig. 9 shows the simulation results. We can see that, no matter the measurements are noiseless or noisy, we have a negligible detection probability of the FDI attack when the state variable of bus 8 is modified. But it can be detected when we attack the other buses. The above results highlight the weakness of the bus that is only connected by a single branch and prove that we can never achieve a complete MTD when the power network contains such a bus.

Next, we consider the induced generation cost when attacking the bus that is only connected by a single branch. Considering the IEEE 30-bus power system, there are 3 buses that are only connected by a single branch. And one of them is a load bus (bus 26). In the following, we exploit the increasing generation cost when the attacker compromises this load bus. The objective function is linear with the value of generation, i.e., $C_i(p_i^g) = \mu_i p_i^g$ [see the OPF problem (16)]. The parameters about the generators are shown in Table III. And all active power flows are limited to 500 MW. In this case, we perturb all branches with the perturbation ratios sampled within [0.8, 1.2]. The default setting and load data provided in MATPOWER are used. First, we use *fmincon* in MATLAB to compute an optimal generation dispatch result. Then, the measurements associated with bus 26 are corrupted. It results in deceiving the amount of the load on bus 26 transmitted to the control center for solving the OPF problem. The simulation results are shown in Table IV. L_0 is the original load at bus 26. We can see the generation cost increases with the corrupted load. Actually, the load attack may also lead the system to a non-optimal generation dispatch, and the worst, may cause load shedding [53].

TABLE III
PARAMETERS OF THE GENERATORS

Generation bus	1	2	13	22	23	27
P_{max}^g	100	100	100	100	100	100
μ_i(\$/MWh)	20	30	20	20	30	20

TABLE IV
INCREASING IN GENERATION COST WHEN THE LOAD BUS THAT IS ONLY CONNECTED BY A SINGLE BRANCH IS ATTACKED

Communicated load	L_0	$1.2L_0$	$1.4L_0$	$1.6L_0$	$1.8L_0$	$2.0L_0$
Generation cost ($10^3\$/h$)	3.784	3.798	3.812	3.826	3.840	3.854
Increasing rate	0	0.37%	0.74%	1.11%	1.48%	1.85%

B. Guidance on Constructing an Effective MTD

1) Target-Perturbation Branch Selection:

a) Evaluation of our algorithm: First, with the IEEE 14-bus power system, we evaluate the selection of the set of target-perturbation branches with Algorithm 1 and Algorithm 2. A total number of 10 D-FACTS devices are given for constructing/realizing MTD. We assume that all branches in this power system can be perturbed. With the aim to minimize the dimension of the stealthy attack space and maximize the number of covered buses, we adopt Algorithm 1 and 2 in Section IV-D1. From Algorithm 1, we obtain a set of target-perturbation branches as $\{k_1, k_3, k_4, k_7, k_8, k_{11}, k_{12}\}$. From Algorithm 2, we obtain a set of target-perturbation branches as $\{k_{16}, k_{17}, k_{20}\}$. We get the final set of the target-perturbation branches by combining these two sets. With this MTD, the dimension of the stealthy attack space is 6 and the number of covered buses is 13. The target-perturbation branches deployed with D-FACTS devices are shown in Fig. 10(a). For comparison, if we optionally perturb the set of branches $\{k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8, k_9, k_{10}\}$ (Fig. 10(b)), the dimension of stealthy attack space is 8 and the number of the covered buses is 8. This indicates that the output from our algorithm is better than the optional selections.

In fact, if we start Algorithm 1 from different branches, we can obtain different sets of target-perturbation branches. We take the IEEE 14-bus power system as an example. Suppose that there are 9 D-FACTS devices for constructing MTD, and all branches in this 14-bus power system can be perturbed. We start Algorithm 1 from branch k_2 and k_{16}, respectively. Correspondingly, we obtain the sets of target-perturbation branches as $\mathcal{L}_d^{\text{from } k_2} = \{k_2, k_3, k_7, k_9, k_{10}, k_{11}, k_{15}, k_{16}, k_{19}\}$ and $\mathcal{L}_d^{\text{from } k_{16}} = \{k_2, k_4, k_6, k_8, k_{10}, k_{11}, k_{16}, k_{19}, k_{20}\}$, respectively. We can see that the sets of target-perturbation branches are different when we start Algorithm 1 from branch k_2 and k_{16}. But we find that the values of γ and the number of covered buses are the same, i.e., $\gamma = 7$ and the number of

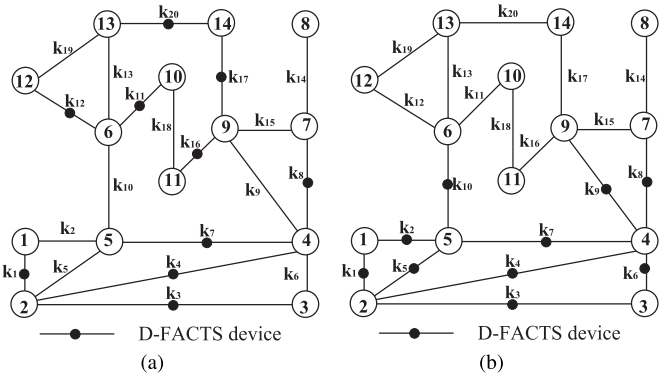

Fig. 10. Comparison of the optimal deployment of D-FACTS devices using our algorithm and the optionally selected target-perturbation branches. (a) Output from the Algorithms given in Section IV-D1; (b) Optionally selected target-perturbation branches. In the first case, $\dim(\mathcal{A}_s) = 6$ and the number of covered buses is 13; while in the second case, $\dim(\mathcal{A}_s) = 8$ and the number of covered buses is 8.

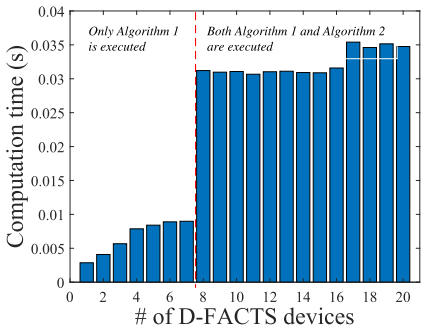

Fig. 11. The variation of the computation time of the proposed algorithm with different numbers of D-FACTS devices (14-bus power system).

covered buses is 13. This indicates that, even though the sets of target-perturbation branches output from our algorithm are different with different start points, they result in the same dimension of stealthy attack space and the same number of covered buses. Thus, we can dynamically change the perturbed branches and maintain the effectiveness of MTD. On the other hand, we can select the branches that are already deployed with D-FACTS devices for realizing MTD, which can help reducing the infrastructure cost.

b) Computation time of the algorithm: Moreover, we evaluate the computation time of the algorithm for the selection of target-perturbation branches. Note that Algorithm 2 is not executed if the condition given in line 13 of Algorithm 1 is not satisfied. We run the algorithm in a core i7 laptop, which has a 2.4GHz CPU and 8.0G memory. We assume that all branches can be perturbed in the adopted power systems. First, we use the IEEE 14-bus power system as an example. We vary the number of perturbed branches from 1 to 20. The computation time of our algorithm is shown in Fig. 11. We can see that the computation time is less than 10ms when the number of D-FACTS devices is less than 8, while it is more than 30ms when the number of D-FACTS devices is larger than 8. It seems that the computation time increases 3 times when the number of D-FACTS devices increases from 8 to 9. This is because only Algorithm 1 is executed when the number of D-FACTS devices is less than 8, i.e., the condition in line 13 of Algorithm 1 is not satisfied. While both Algorithm 1 and Algorithm 2 are executed when the number of D-FACTS devices is larger than 8. This indicates that sometimes it takes longer for executing Algorithm 2.

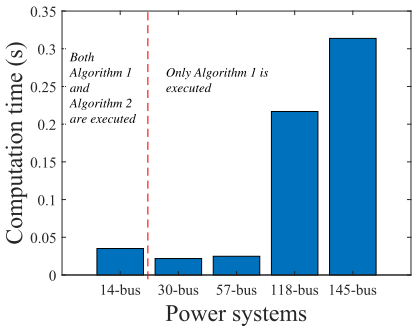

Fig. 12. The variation of the computation time of the proposed algorithm with different power systems (given 10 D-FACTS devices).

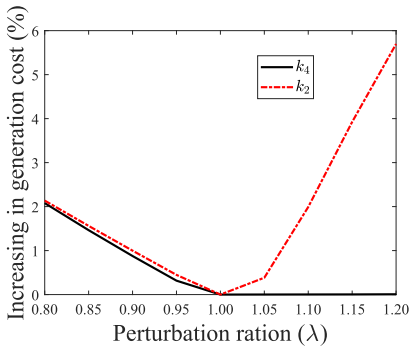

Fig. 13. The increasing in generation cost by perturbing branch k_2 and k_4 to different ratios.

TABLE V
PARAMETERS OF THE GENERATORS

Generation bus	1	2	3	6	8
P^g_{max}	300	50	30	50	20
μ_i(\$/MWh)	20	30	40	50	35

Second, we change the size of the power system while fix the number of D-FACTS devices as 10. Here we adopt the IEEE 14-bus, 30-bus, 57-bus, 118-bus and 145-bus power systems. The computation time of our algorithm is given in Fig. 12. We can see that the computation time increases from around 30ms (14-bus) to more than 300ms (145-bus), which indicates that the computation time of our algorithm increases with the system size. Moreover, we find that the computation time with the 14-bus power system is a little larger than that with the 30-bus power system. In our opinion, the reason is that, given 10 D-FACTS devices, both Algorithm 1 and Algorithm 2 are executed with the 14-bus power system, while only the Algorithm 1 is executed with the 30-bus, 57-bus, 118-bus and 145-bus power systems. Overall, we can efficiently compute the result (within 350ms in all given power systems).

2) Reducing the Operation Cost: Furthermore, we evaluate the impact of the perturbed branch and the susceptance perturbation magnitude on the increasing of the operation cost. For the OPF problem, we use the objective function as $C_i(p_i^g) = \mu_i p_i^g$, which is a linear generation cost model. With the IEEE 14-bus power system, the generators are installed at bus 1, 2, 3, 6 and 8, and their parameters are shown in Table V. The active power flow limits of branch 1 is 160 MW and the other active power flows are limited to 60 MW. We assume that the optimal results are obtained at the beginning. Then, we analyze the increase of the generation cost when we perturb branch k_2 and k_4 to different ratios, respectively. The simulation

results are plotted in Fig. 13. We can see that the increasing of the generation cost under these two cases are different. When the branch susceptance is decreased, the increasing of generation cost by perturbing k_4 is slightly lower than that of by perturbing k_2. But the generation cost increases as the perturbation ratio decreases in both cases. When the branch susceptance is increased, the generation cost of perturbing k_4 almost remains invariant, while it increases with the perturbation ratio by perturbing k_2. The simulation result indicates that, by appropriately selecting the perturbed branch and the perturbation magnitude, we can reduce the increasing operation cost of MTD.

VI. Conclusion

In this paper, with the DC power flow model, we analyzed the completeness, deployment and the increasing operation cost of MTD in terms of thwarting stealthy FDI attacks constructed with old system information. To begin with, we proved that an MTD is complete to defeat all FDI attacks constructed with former branch parameters only if the number of branches l is larger than or equal to twice that of the system states n (i.e., $l \geq 2n$, where $n+1$ is the number of system buses), and the susceptances of more than n branches, which cover all buses, are perturbed. Besides, we prove that we can never realize a complete MTD if the power transmission system has a bus that is only connected by a single branch. Further, we prove that the susceptance perturbation magnitude almost does not affect the dimension of the stealthy attack space after MTD. Based on this result, we presented guidance on effective MTD for minimizing the dimension of the stealthy stealthy attack space, maximizing the number of covered buese and reducing the operation cost. Finally, we illustrated and demonstrated our findings with the IEEE standard test power systems.

Appendix A

Proof of Theorem 6: Considering the first condition, we resort to a contradiction. That is, suppose that the MTD is complete but $l < 2n$. Let $\widehat{\mathbf{H}}$ and $\widehat{\mathbf{H}}'$ denote the measurement matrices before and after MTD with a fully measured power system, respectively. Thus, with any partially measured power system, the measurement matrix $\mathbf{H}$ is formed by selecting some rows from $\widehat{\mathbf{H}}$. Therefore, the combined matrix $[\mathbf{H}\ \mathbf{H}']$ can be formed by selecting some rows from the combined matrix $[\widehat{\mathbf{H}}\ \widehat{\mathbf{H}}']$. It follows that $R([\mathbf{H}\ \mathbf{H}']) \leq R([\widehat{\mathbf{H}}\ \widehat{\mathbf{H}}'])$. We have proved that $R([\widehat{\mathbf{H}}\ \widehat{\mathbf{H}}']) = 2n$ only if the number of branches l is greater than or equal to $2n$. Therefore, we can derive that $R([\mathbf{H}\ \mathbf{H}']) \leq R([\widehat{\mathbf{H}}\ \widehat{\mathbf{H}}']) < 2n$ if $l < 2n$. This indicates that the MTD is not complete, which contradicts to the original assumption. Therefore, we must guarantee $l \geq 2n$ for achieving a complete MTD.

Next, we consider the second condition. Let $\mathcal{M}_{mtd}$ be the set of buses covered by the perturbed branches and n_s is the size of $\mathcal{M}_{mtd}$. $\mathbf{H}^q$ is a submatrix formed by selecting q columns from $\mathbf{H}'$ such that the combined matrix $[\mathbf{H}\ \mathbf{H}^q]$ has full column rank. Let $\mathcal{M}^q_{mtd}$ be an index set of these q columns in $\mathbf{H}'$. Since $\mathcal{M}^q_{mtd} \subseteq \mathcal{M}_{mtd}$, we have $q \leq n_s$. If the MTD is complete, then we have $q = n$. Therefore, $n_s \geq n$, that is, the perturbed branches must cover all buses.

Appendix B

Proof of Theorem 8: Suppose t is the bus that is only connected by a single branch. And the susceptance of the connected branch $k = \{t', t\}$ is $b_{t't}$. Let $\mathbf{e}_i \in \{0,1\}^n$ be a vector with a unit in the ith position and zero elsewhere, and let $\mathbf{u}_{ij} = \mathbf{e}_i - \mathbf{e}_j$ ($k = \{i, j\} \in \mathcal{L}$) be a vector with "1" in the ith position and "-1" in the jth position and zero elsewhere. Considering the fully measured case, let the tth column of the branch-bus shift factor matrix $\mathbf{S}$ and the symmetric admittance matrix $\mathbf{B}$ be $\mathbf{s}_t$ and $\mathbf{b}_t$, respectively. Then, we have $\mathbf{s}_t = -b_{t't}\mathbf{e}_t$ and $\mathbf{b}_t = b_{t't}\mathbf{u}_{t't}$. Suppose $\mathbf{s}'_t = -b'_{t't}\mathbf{e}_t$ and $\mathbf{b}'_t = b'_{t't}\mathbf{u}_{t't}$ after MTD. Then, we can derive that $\mathbf{s}_t = \frac{b_{t't}}{b'_{t't}}\mathbf{s}'_t$ and $\mathbf{b}_t = \frac{b_{t't}}{b'_{t't}}\mathbf{b}'_t$. Let $\mathbf{h}_t$ and $\mathbf{h}'_t$ be the tth column of $\mathbf{H}$ and $\mathbf{H}'$ before and after MTD, respectively. Then, we have $\mathbf{h}_t = \frac{b_{t't}}{b'_{t't}}\mathbf{h}'_t$. Since any $\mathbf{H}$ can be formed by selecting some rows from the measurement matrix under the fully measured case, we can derive $\mathbf{h}_t = \frac{b_{t't}}{b'_{t't}}\mathbf{h}'_t$ under the partially measured case as well. It follows that $S(\mathbf{H}') = S(\mathbf{H})$ if we only perturb the branch k. Therefore, we can never obtain $\gamma = R([\mathbf{H}\ \mathbf{H}']) = 2n$.

Appendix C

Proof of Proposition 9: Let $\mathbf{e}_i \in \{0,1\}^n$ be a vector with a unit in the ith position and zeros elsewhere, and let $\mathbf{u}_{ij} = \mathbf{e}_i - \mathbf{e}_j$ ($k = \{i, j\} \in \mathcal{L}$) be a vector with "1" in the ith position and "-1" in the jth position and zeros elsewhere. Then, we can rewrite $\mathbf{A}$ and $\mathbf{D}$ as

$$\mathbf{A} = \sum_{\substack{k\in\mathcal{L}\\ k=\{i,j\}}} \mathbf{e}_k\mathbf{u}_{ij}^T, \quad \mathbf{D} = \sum_{\substack{k\in\mathcal{L}\\ k=\{i,j\}}} -b_{ij}\mathbf{e}_k\mathbf{e}_k^T. \tag{18}$$

Supposing the diagonal branch susceptance matrix after MTD is $\mathbf{D}'$ and $\Delta\mathbf{D} = \mathbf{D}' - \mathbf{D}$, then we can derive $\Delta\mathbf{D}$ as

$$\begin{aligned}\Delta\mathbf{D} &= \sum_{\substack{k_d\in\mathcal{L}_D\\ k_d=\{i_d,j_d\}}} -(b'_{i_dj_d} - b_{i_dj_d})\mathbf{e}_{k_d}\mathbf{e}_{k_d}^T \\ &= \sum_{\substack{k_d\in\mathcal{L}_D\\ k_d=\{i_d,j_d\}}} -\Delta b_{i_dj_d}\mathbf{e}_{k_d}\mathbf{e}_{k_d}^T. \end{aligned}\tag{19}$$

We can see that there are non-zero elements in positions (k_d, k_d) with $k_d \in \mathcal{L}_D$ and zeros elsewhere in $\Delta\mathbf{D}$. Since $\mathbf{e}_i^T\mathbf{e}_j = 0$ if $i \neq j$ and $\mathbf{e}_i^T\mathbf{e}_j = 1$ if $i = j$, we can derive that the change of the branch-bus shift factor matrix is

$$\begin{aligned}\Delta\mathbf{S} = \Delta\mathbf{D}\mathbf{A} &= \sum_{\substack{k_d\in\mathcal{L}_D\\ k_d=\{i_d,j_d\}}} -(\Delta b_{i_dj_d}\mathbf{e}_{k_d}\mathbf{e}_{k_d}^T)\left(\mathbf{e}_{k_d}\mathbf{u}_{i_dj_d}^T\right) \\ &= \sum_{\substack{k_d\in\mathcal{L}_D\\ k_d=\{i_d,j_d\}}} -\Delta b_{i_dj_d}\mathbf{e}_{k_d}\mathbf{u}_{i_dj_d}^T \end{aligned}\tag{20}$$

We can observe that $\Delta\mathbf{S}$ is a sparse matrix with the k_dth row is non-zero and other rows are all zero. That is, if $i \in \mathcal{L}_D$, then the ith row of $\Delta\mathbf{S}$ is $\Delta\mathbf{S}_i = [\cdots - \Delta b_{i_dj_d} \cdots \Delta b_{i_dj_d} \cdots]$ with $-\Delta b_{i_dj_d}$ in the ith column and $\Delta b_{i_dj_d}$ in the jth column; if $i \notin \mathcal{L}_D$, $\Delta\mathbf{S}_i = [0 \cdots 0 \cdots 0]$. Similarly, since $\mathbf{B} = \mathbf{A}^T\mathbf{S}$,

we have $\Delta\mathbf{B} = \mathbf{A}^T \Delta\mathbf{S}$

$$\Delta\mathbf{B} = \mathbf{A}^T \Delta\mathbf{S} = \sum_{\substack{k_d \in \mathcal{L}_D \\ k_d = \{i_d, j_d\}}} -\Delta b_{i_d j_d} (\mathbf{u}_{i_d j_d} \mathbf{e}_{k_d}^T)(\mathbf{e}_{k_d} \mathbf{u}_{i_d j_d}^T) = \sum_{\substack{k_d \in \mathcal{L}_D \\ k_d = \{i_d, j_d\}}} -\Delta b_{i_d j_d} \mathbf{u}_{i_d j_d} \mathbf{u}_{i_d j_d}^T. \quad (21)$$

$\Delta\mathbf{B}$ is a sparse matrix with the i_dth and j_dth columns are non-zero and the other columns are all zero.

On the basis of the measurement matrix $\mathbf{H}$ (before MTD) and $\mathbf{H}'$ (after MTD). Since $\mathbf{H}' = [\mathbf{B} + \Delta\mathbf{B}; \mathbf{S} + \Delta\mathbf{S}; -\mathbf{S} - \Delta\mathbf{S}]$, we can derive that the difference of the measurement matrix is

$$\Delta\mathbf{H} = \mathbf{H}' - \mathbf{H} = \begin{bmatrix} \Delta\mathbf{B} \\ \Delta\mathbf{S} \\ -\Delta\mathbf{S} \end{bmatrix}. \quad (22)$$

Obviously, we can see that $\Delta\mathbf{H}$ is a sparse matrix with non-zero elements in the i_dth and j_dth column with $k_d = \{i_d, j_d\} \in \mathcal{L}_D$.

For example, if we only perturb a single branch, that is, $\mathcal{L}_d = \{k_d\}$ with $k_d = \{i_d, j_d\}$, then $\Delta\mathbf{H}$ under a fully measured case is stated as

$$\Delta\mathbf{H} = \begin{bmatrix} \mathbf{0} & \vdots & \mathbf{0} & \vdots & \mathbf{0} & \\ \cdots & -\Delta b_{i_d j_d} & \cdots & \Delta b_{i_d j_d} & \cdots & \} i_d \\ \mathbf{0} & \vdots & \mathbf{0} & \vdots & \mathbf{0} & \\ \cdots & \Delta b_{i_d j_d} & \cdots & -\Delta b_{i_d j_d} & \cdots & \} j_d \\ \mathbf{0} & \vdots & \mathbf{0} & \vdots & \mathbf{0} & \\ \cdots & -\Delta b_{i_d j_d} & \cdots & \Delta b_{i_d j_d} & \cdots & \} n + k_d \\ \mathbf{0} & \vdots & \mathbf{0} & \vdots & \mathbf{0} & \\ \cdots & \Delta b_{i_d j_d} & \cdots & -\Delta b_{i_d j_d} & \cdots & \} n + l + k_d \\ \mathbf{0} & \vdots & \mathbf{0} & \vdots & \mathbf{0} & \\ & \underbrace{\quad}_{i_d} & & \underbrace{\quad}_{j_d} & & \end{bmatrix}, \quad (23)$$

where $\Delta b_{i_d j_d}$ is the susceptance perturbation of branch k_d, $\mathbf{0}$ is a zero matrix/vector. We can see that all non-zero elements in $\Delta\mathbf{H}$ are in the i_dth column and the j_dth column. Since any measurement matrix is formed by selecting some rows from the measurement matrix $\mathbf{H}$ under the fully measured case, we can derive that $\Delta\mathbf{H}$ is also a sparse matrix with non-zero elements in the i_dth and j_dth column with $k_d = \{i_d, j_d\} \in \mathcal{L}_D$.

APPENDIX D

Proof of Proposition 10: We have $\mathbf{C}' = [\mathbf{H}\ \mathbf{H}'] + [\mathbf{0}_{m\times n}\ \Delta\mathbf{H}] = \mathbf{C} + [\mathbf{0}_{m\times n}\ \Delta\mathbf{H}]$, where $\mathbf{0}_{m\times n}$ is an m by n zero matrix, $\Delta\mathbf{H}$ is a matrix corresponding to the susceptance perturbation of the new branch. According to the equation (23), $R(\Delta\mathbf{H}) = 1$ when one more branch is perturbed. Since $R(\mathbf{C}') \leq R(\mathbf{C}) + R(\Delta\mathbf{H})$, we have $R(\mathbf{C}') \leq R(\mathbf{C}) + 1$. Conversely, we have $\mathbf{C} = [\mathbf{H}\ \mathbf{H}'] + [\mathbf{0}_{m\times n}\ -\Delta\mathbf{H}] = \mathbf{C}' + [\mathbf{0}_{m\times n}\ -\Delta\mathbf{H}]$. Thus, we can derive that $R(\mathbf{C}) \leq R(\mathbf{C}') + R(-\Delta\mathbf{H}) \Rightarrow R(\mathbf{C}') \geq R(\mathbf{C}) - 1$. Therefore, we have $R(\mathbf{C}) - 1 \leq R(\mathbf{C}') \leq R(\mathbf{C}) + 1$.

APPENDIX E

Proof of Proposition 11: Here we only analyze the case when $i_d \notin \mathcal{M}_{mtd}^q$, because we can draw a same conclusion when $j_d \notin \mathcal{M}_{mtd}^q$. We denote $\det(\cdot)$ as the determinant of a matrix. $\mathbf{a}_i$ is the ith column of a matrix $\mathbf{A}$. Since $i_d \notin \mathcal{M}_{mtd}^q$ and $j_d \notin \mathcal{M}_{mtd}^q$, the i_dth column $\mathbf{h}'_{i_d}$ and the j_dth column $\mathbf{h}'_{j_d}$ of $\mathbf{H}'$ can be linearly represented by the columns in $[\mathbf{H}\ \mathbf{H}^q]$. It follows that $R([\mathbf{H}\ \mathbf{H}^q\ \mathbf{h}'_{i_d}]) = R([\mathbf{H}\ \mathbf{H}^q]) = R([\mathbf{H}\ \mathbf{H}'])$. After the susceptance perturbation of branch k_d, $\mathbf{h}'_{i_d}$ and $\mathbf{h}'_{j_d}$ respectively become $\mathbf{h}''_{i_d}$ and $\mathbf{h}''_{j_d}$ according to the proof of Proposition 9. We rearrange the matrix $\mathbf{C}'$ as $\mathbf{C}' = [\mathbf{H}\ \mathbf{H}^q\ \mathbf{h}''_{i_d}\ \mathbf{h}''_{j_d}\ \mathbf{H}''_r]$, where $\mathbf{H}''_r$ is a submatrix containing the columns in $\mathbf{H}^p$ ($\mathbf{H}' = [\mathbf{H}^q\ \mathbf{H}^p]$) that are not changed after the perturbation of branch k_d. Let $\mathbf{C}'' = [\mathbf{H}\ \mathbf{H}^q\ \mathbf{h}'_{i_d}\ \mathbf{h}''_{i_d}\ \mathbf{h}''_{j_d}\ \mathbf{H}''_r]$ be a matrix by filling the column $\mathbf{h}'_{i_d}$ into $\mathbf{C}'$. Then, we can derive that $R(\mathbf{C}'') = R(\mathbf{C}')$.

We calculate $R(\mathbf{C}'')$ through elementary column operations. Since only the element relating to branch k_d is changed in column $\mathbf{h}''_{i_d}$, by subtracting $\mathbf{h}''_{i_d}$ from $\mathbf{h}'_{i_d}$, we have $\Delta\mathbf{h}''_{i_d} = \mathbf{h}''_{i_d} - \mathbf{h}'_{i_d} = [0\ \cdots\ 0\ \pm(\lambda-1)b_{i_d j_d}\ 0\ \cdots\ 0\ \pm(\lambda-1)b_{i_d j_d}\ 0\ \cdots\ 0]^T$, where $(\lambda - 1)b_{i_d j_d}$ and $-(\lambda - 1)b_{i_d j_d}$ are the only non-zero elements in the vector $\Delta\mathbf{h}''_{i_d}$. The number of non-zero elements in this column depends on the measured case of the power system. But it is less than 4 according to the example given in the equation (23). For the elementary column operation, we transfer the column $\mathbf{h}'_{i_d}$ to $\Delta\mathbf{h}''_{i_d}$ in $\mathbf{C}''$. Since the other elements of the column $\Delta\mathbf{h}''_{i_d}$ are zero, any other column operations on this column are only related to $(\lambda - 1)b_{i_d j_d}$. Therefore, it is equivalent to operate on $(\lambda - 1)b_{i_d j_d}$ for elementary column operations for any λ if $\lambda \neq 1$. For example, we perturb the susceptance of branch k_d to $\lambda_1\ b_{i_d j_d}$ and $\lambda_2\ b_{i_d j_d}$, respectively. Assume that only the branches are monitored by meters and each branch is monitored by one meter. Then, we can derive that

$$\Delta\mathbf{h}''_{i_d} = [0\ \cdots\ 0\ \underbrace{(1-\lambda)b_{i_d j_d}}_{k_d th\ element}\ 0\ \cdots\ 0]^T, \quad (24)$$

If the required column operation is to multiply $\Delta\mathbf{h}''_{i_d}$ with a constant η, then the result is $\mathbf{h} = \eta\Delta\mathbf{h}''_{i_d}$. We find $\mathbf{h} = \eta[0\ \cdots\ 0\ (1-\lambda_1)b_{i_d j_d}\ 0\ \cdots\ 0]^T = \eta \times \frac{\lambda_1}{\lambda_2}[0\ \cdots\ 0\ (1-\lambda_2)b_{i_d j_d}\ 0\ \cdots\ 0]^T$. That is, we can always obtain the same column operation result $\mathbf{h}$ regardless of the value of λ.

Therefore, the impact of $\Delta\mathbf{h}''_{i_d}$ on $R(\mathbf{C}'')$ is not related to the value of λ. Thus, the impact of the column $\mathbf{h}''_{i_d}$ on $R(\mathbf{C}'')$ is not related to the value of λ. Since $R(\mathbf{C}'') = R(\mathbf{C}')$, the impact of the column $\mathbf{h}''_{i_d}$ on $R(\mathbf{C}')$ is not related to the value of λ. As for the column $\mathbf{h}''_{j_d}$, we can draw the same conclusion by using elementary column operations. Therefore, $\Delta\gamma = R(\mathbf{C}') - R(\mathbf{C})$ does not change with λ ($\lambda > 0$ and $\lambda \neq 1$).

APPENDIX F

Proof of Proposition 12: We prove this proposition in ***Case 2*** and ***Case 3***, respectively. We denote $\det(\cdot)$ as the determinant of a matrix. $\mathbf{a}_i$ is the ith column of a matrix $\mathbf{A}$.

Case 2: $i_d \in \mathcal{M}_{mtd}^q$ ***and*** $j_d \notin \mathcal{M}_{mtd}^q$

Since $\Delta\gamma = -1$, we have $R([\mathbf{H}\ \mathbf{H}'']) = R([\mathbf{H}\ \mathbf{H}']) - 1$. It follows that we must have $R([\mathbf{H}\ \mathbf{H}_r^q\ \mathbf{h}''_{i_d}]) = R([\mathbf{H}\ \mathbf{H}^q]) - 1$ when $\lambda = \lambda^*$, where $\mathbf{H}_r^q$ is a submatrix formed by deleting the column $\mathbf{h}'_{i_d}$ from $\mathbf{H}^q$, $\mathbf{h}'_{i_d}$ is the i_dth column of $\mathbf{H}'$, $\mathbf{h}''_{i_d}$ is the i_dth column of $\mathbf{H}''$ after the perturbation of branch k_d. If we select any $n+q$ rows that contain the perturbed rows (i.e., there are elements $\lambda b_{i_d j_d}$ and/or $-\lambda b_{i_d j_d}$ in this row) and any $n+q$ columns from the combined matrix $[\mathbf{H}\ \mathbf{H}_r^q\ \mathbf{h}''_{i_d}]$, we can form an $n+q$ by $n+q$ square matrix $\mathbf{Q}$. Using the Leibniz formula, $\det(\mathbf{Q})$ is a linear function of $\lambda b_{i_d j_d}$, that is, $\det(\mathbf{Q}) = ab_{i_d j_d}\lambda + b$, where a and b are values calculated by other elements. Specifically, a and b are different if we choose different rows and columns to form $\mathbf{Q}$. Because the combined matrix $[\mathbf{H}\ \mathbf{H}_r^q\ \mathbf{h}''_{i_d}]$ has full column rank when $\lambda = 1$, there exists $\mathbf{Q}$ such that $\det(\mathbf{Q}) = ab_{i_d j_d} + b \neq 0$. Therefore, a and b cannot be 0 at the same time. If $a = 0$, then $\det(\mathbf{Q}) = ab_{i_d j_d}\lambda + b \neq 0$ for any λ. Therefore, there does not exist λ such that $\det(\mathbf{Q}) = 0$. If $a \neq 0$, then $\det(\mathbf{Q}) = ab_{i_d j_d}\lambda + b = 0$ has an unique solution $\frac{-b}{ab_{i_d j_d}}$. As $\Delta\gamma = -1$ when $\lambda = \lambda^*$, we have $R([\mathbf{H}\ \mathbf{H}_r^q\ \mathbf{h}''_{i_d}]) = n + q - 1$. It follows that, for any $\mathbf{Q}$, we have $\det(\mathbf{Q}) = ab_{i_d j_d}\lambda^* + b = 0$. This indicates that λ^* must be equal to $\frac{-b}{ab_{i_d j_d}}$. Therefore, $\Delta\gamma = -1$ only if there exists λ^* such that $\lambda_{min} \leq \lambda^* \leq \lambda_{max}$ and $\lambda^* \neq 1$ and $\lambda = \lambda^*$.

Case 3: $i_d \in \mathcal{M}_{mtd}^q$ ***and*** $j_d \in \mathcal{M}_{mtd}^q$

After branch k_d is perturbed, we can rearrange $\mathbf{C}' = [\mathbf{H}\ \mathbf{H}'']$ as $\mathbf{C}' = [\mathbf{H}\ \mathbf{H}_r^q\ \mathbf{h}''_{i_d}\ \mathbf{h}''_{j_d}\ \mathbf{H}^p]$, where $\mathbf{H}_r^q$ is a matrix containing the rest columns of $\mathbf{H}^q$ that are not changed, $\mathbf{h}''_{i_d}$ and $\mathbf{h}''_{j_d}$ are the i_dth column and j_dth column in $\mathbf{H}''$ that are perturbed. If we select any $n+q$ rows that contain the perturbed rows (i.e., there are elements $\lambda b_{i_d j_d}$ and/or $-\lambda b_{i_d j_d}$ in this row) and any $n+q$ columns from the combined matrix $[\mathbf{H}\ \mathbf{H}_r^q\ \mathbf{h}''_{i_d}\ \mathbf{h}''_{j_d}]$, we can from an $n+q$ by $n+q$ square matrix $\mathbf{Q}$. Using the Leibniz formula, the determinant of $\mathbf{Q}$ is $\det(\mathbf{Q}) = ab_{i_d j_d}\lambda + b$, where a and b are coefficients after the calculation. As there exists $\mathbf{Q}$ such that we have $\det(\mathbf{Q}) = ab_{i_d j_d} + b \neq 0$ when $\lambda = 1$, a and b cannot be 0 at the same time. Then, if $a = 0$, there exists $\mathbf{Q}$ such that $\det(\mathbf{Q}) \neq 0$ for any λ. Therefore, $\Delta\gamma \neq -1$ for any λ. If $a \neq 0$, $\det(\mathbf{Q}) = ab_{i_d j_d}\lambda + b = 0$ has an unique solution $\frac{-b}{ab_{i_d j_d}}$. Thus, if $R(\mathbf{Q}) = n + q - 1$ (i.e., $\Delta\gamma = -1$) when $\lambda = \lambda^*$, then λ^* must be equal to $\frac{-b}{ab_{i_d j_d}}$. Therefore, $\Delta\gamma = -1$ only if there exists λ^* such that $\lambda_{min} \leq \lambda^* \leq \lambda_{max}$ and $\lambda^* \neq 1$ and $\lambda = \lambda^*$.

REFERENCES

[1] Z. Zhang, R. Deng, D. Yau, P. Cheng, and J. Chen, "On effectiveness of detecting FDI attacks on power grid using moving target defense," in *Proc. IEEE PES Innov. Smart Grid Technol. Conf. (ISGT)*, Feb. 2019, pp. 1–5.

[2] A. G. Illera and J. V. Vidal. (2014). *Lights Off! The Darkness of the Smart Meters*. [Online]. Available: http://youtube.be/Z_y_vjYtAWM

[3] C. Konstantinou and M. Maniatakos, "A case study on implementing false data injection attacks against nonlinear state estimation," in *Proc. 2nd ACM Workshop Cyber-Phys. Syst. Secur. Privacy (CPS-SPC)*, Oct. 2016, pp. 81–91.

[4] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber-physical system security for the electric power grid," *Proc. IEEE*, vol. 100, no. 1, pp. 210–224, Jan. 2012.

[5] M. N. Kurt, Y. Yilmaz, and X. Wang, "Distributed quickest detection of cyber-attacks in smart grid," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 8, pp. 2015–2030, Aug. 2018.

[6] M. N. Kurt, Y. Yılmaz, and X. Wang, "Real-time detection of hybrid and stealthy cyber-attacks in smart grid," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 2, pp. 498–513, Feb. 2019.

[7] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grid," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, May 2011, Art. no. 13.

[8] R. Deng, P. Zhuang, and H. Liang, "CCPA: Coordinated cyber-physical attacks and countermeasures in smart grid," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2420–2430, May 2017.

[9] R. Tan *et al.*, "Optimal false data injection attack against automatic generation control in power grids," in *Proc. ACM/IEEE Int. Conf. Cyber-Phys. Syst. (ICCPS)*, Apr. 2016, pp. 1–10.

[10] M. Liu, C. Zhao, R. Deng, P. Cheng, W. Wang, and J. Chen, "Nonzero-dynamics stealthy attack and its impacts analysis in DC microgrids," in *Proc. IEEE Amer. Control Conf. (ACC)*, Jul. 2019, pp. 1–5.

[11] Y. Yuan, Z. Li, and K. Ren, "Quantitative analysis of load redistribution attacks in power systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 9, pp. 1731–1738, Sep. 2012.

[12] R. Zhang and P. Venkitasubramaniam, "Stealthy control signal attacks in linear quadratic Gaussian control systems: Detectability reward tradeoff," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 7, pp. 1555–1570, Jul. 2017.

[13] M. Giannini, "Improving cyber-security of power system state estimators," M.S. thesis, School Elect. Eng., KTH, Stockholm, Sweden, Feb. 2014.

[14] R. Tan *et al.*, "Modeling and mitigating impact of false data injection attacks on automatic generation control," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 7, pp. 1609–1623, Jul. 2017.

[15] 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.

[16] R. Deng, G. Xiao, R. Lu, H. Liang, and A. V. Vasilakos, "False data injection on state estimation in power systems—Attacks, impacts, and defense: A survey," *IEEE Trans. Ind. Informat.*, vol. 13, no. 2, pp. 411–423, Apr. 2017.

[17] Q. Yang, D. An, R. Min, W. Yu, X. Yang, and W. Zhao, "On optimal PMU placement-based defense against data integrity attacks in smart grid," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 7, pp. 1735–1750, Jul. 2017.

[18] T. T. Kim and H. V. Poor, "Strategic protection against data injection attacks on power grids," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 326–333, Jun. 2011.

[19] J. Yao, P. Venkitasubramaniam, S. Kishore, L. V. Snyder, and R. S. Blum, "Network topology risk assessment of stealthy cyber attacks on advanced metering infrastructure networks," in *Proc. 51st Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2017, pp. 1–6.

[20] S. Pal, B. Sikdar, and J. H. Chow, "Classification and detection of PMU data manipulation attacks using transmission line parameters," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5057–5066, Sep. 2018.

[21] 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 (MTD)*, Nov. 2014, pp. 59–68.

[22] D. Divan and H. Johal, "Distributed FACTS—A new concept for realizing grid power flow control," *IEEE Trans. Power Electron.*, vol. 22, no. 6, pp. 2253–2260, Nov. 2007.

[23] K. L. Morrow, E. Heine, K. M. Rogers, R. B. Bobba, and T. J. Overbye, "Topology perturbation for detecting malicious data injection," in *Proc. Int. Conf. Syst. Sci.*, Jan. 2012, pp. 2104–2113.

[24] K. R. Davis, K. L. Morrow, R. Bobba, and E. Heine, "Power flow cyber attacks and perturbation-based defense," in *Proc. IEEE 3rd Int. Conf. Smart Grid Commun. (SmartGridComm)*, Nov. 2012, pp. 342–347.

[25] J. Tian, R. Tan, X. Guan, and T. Liu, "Enhanced hidden moving target defense in smart grids," *IEEE Trans. Smart Grids*, vol. 10, no. 2, pp. 2208–2223, Dec. 2018.

[26] B. Liu, H. Wu, A. Pahwa, F. Ding, E. Ibrahim, and T. Liu, "Hidden moving target defense against false data injection in distribution network reconfiguration," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2018, pp. 1–5.

[27] S. Lakshminarayana and D. K. Yau, "Cost-benefit analysis of moving-target defense in power grids" 2018, *arXiv:1804.01472*. [Online]. Available: https://arxiv.org/abs/1804.01472

[28] C. Liu, J. Wu, C. Long, and D. Kundur, "Reactance perturbation for detecting and identifying FDI attacks in power system state estimation," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 4, pp. 763–776, Aug. 2018.

[29] K. M. Rogers and T. J. Overbye, "Some applications of distributed flexible AC transmission system (D-FACTS) devices in power systems," in *Proc. IEEE 40th North Amer. Power Symp. (NAPS)*, Nov. 2009, pp. 1–8.

[30] A. Brissette, "Performance analysis and design of distributed static series compensators for transmission line susceptance control," Ph.D. dissertation, Dept. Elect., Comput. Energy Eng., Univ. Colorado, Boulder, Boulder, CO, USA, 2014.
[31] S.-C. Lee, J.-Y. Kim, J.-H. Lee, J.-W. Lim, and S.-I. Moon, "Hybrid linearization of a power system with FACTS devices for a small signal stability study," in *Proc. IEEE Power Eng. Soc. Summer Meeting*, Jul. 2009, pp. 1566–1572.
[32] A. Wood, B. Wollenberg, and G. B. Sheblé, *Power Generation, Operation, and Control*, 3rd ed. Hoboken, NJ, USA: Wiley, 2013.
[33] A. Abur and A. G. Expósito, *Power System State Estimation: Theory and Implementation*. New York, NY, USA: Marcel Dekker, 2004.
[34] NESCOR. *Electric Sector Failure Scenarios and Impact Analyses Version 3.0*. Accessed: 2019. [Online]. Available: http://smartgrid.epri.com/doc/NESCOR%20Failure%20Scenarios%20v3%2012-11-15.pdf
[35] NERC. *Learned Lessons*. Accessed: 2019. [Online]. Available: https://www.nerc.com/pa/rrm/ea/Pages/Lessons-Learned.aspx
[36] F. Li, Y. Wei, and S. Adhikari, "Improving an unjustified common practice in EX post LMP calculation: An expanded version," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Jul. 2010, pp. 1–4.
[37] I. Markwood, Y. Liu, K. Kwiat, and C. Kamhoua, "Electric grid power flow model camouflage against topology leaking attacks," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Oct. 2017, pp. 1–9.
[38] 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.
[39] 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, Jan. 2015.
[40] Y. Chen, S. Huang, F. Liu, Z. Wang, and X. Sun, "Evaluation of reinforcement learning-based false data injection attack to automatic voltage control," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 2158–2169, Mar. 2019.
[41] 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. 1–5.
[42] M. Stephen, P. Dmitry, M. Sergei, D. Adam, and M. Patrick, "Multi-vendor penetration testing in the advanced metering infrastructure," in *Proc. ACM 26th Annu. Comput. Secur. Appl. Conf. Trans. (ACSAC)*, Dec. 2010, pp. 107–116.
[43] H. Johal and D. Divan, "Design Considerations for Series-Connected Distributed FACTS Converters," *IEEE Trans. Ind. Appl.*, vol. 43, no. 6, pp. 1609–1618, Dec. 2018.
[44] J. Urquizo, P. Singh, N. Kondrath, R. Hidalgo-León, and G. Soriano, "Using D-FACTS in microgrids for power quality improvement: A review," in *Proc. IEEE Educ. Tech. Chapters Meeting (ETCM)*, Oct. 2017, pp. 1–6.
[45] D. Mehta, A. Ravindran, B. Joshi, and S. Kamalasadan, "Graph theory based online optimal power flow control of power grid with distributed flexible AC transmission systems (D-FACTS) devices," in *Proc. IEEE North Amer. Power Symp.*, Oct. 2015, pp. 1–6.
[46] A. Hamidi, S. Golshannavaz, and D. Nazarpour, "D-FACTS cooperation in renewable integrated microgrids: A linear multiobjective approach," *IEEE Trans. Sustain. Energy*, vol. 10, no. 1, pp. 355–363, Jan. 2017.
[47] K. Kuntz, M. Smith, K. Wedeward, and M. Collins, "Detecting, locating, & quantifying false data injections utilizing grid topology through optimized D-FACTS device placement," in *Proc. North Amer. Power Symp. (NAPS)*, Sep. 2014, pp. 1–6.
[48] H. Salehghaffari and F. Khorrami, "Resilient power grid state estimation under false data injection attacks" in *Proc. IEEE PES Innov. Smart Grid Technol. Conf.*, Feb. 2018, pp. 1–5.
[49] 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.*, to be published.
[50] R. D. Zimmerman, C. E. Murillo-Sánchez, 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.
[51] S. Frank and S. Rebennack, "An introduction to optimal power flow: Theory, formulation, and examples," *IIE Trans.*, vol. 48, no. 12, pp. 1172–1197, Aug. 2016.
[52] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[53] 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.
[54] W.-L. Chin, C.-H. Lee, and T. Jiang, "Blind false data attacks against ac state estimation based on geometric approach in smart grid communications," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6298–6306, Nov. 2018.
[55] M. Awad, K. E. Casey, A. S. Geevarghese, J. C. Miller, A. F. Rahimi, A. Y. Sheffrin, M. Zhang, E. Toolson, G. Drayton, B. F. Hobbs, and F. A. Wolak, "Economic assessment of transmission upgrades: Application of the California ISO approach," in *Restructured Electric Power Systems: Analysis of Electricity Markets With Equilibrium Models*. Hoboken, NJ, USA: Wiley, 2010, pp. 241–270.
[56] GE Energy. *MAPS Software-For Informed Economic Decisions*. [Online]. Available: https://www.gepower.com/prod_serv/products/utility_software/en/downloads/10320.pdf
[57] B. Stott, J. Jardim, and O. Alsac, "DC power flow revisited," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1290–1300, Aug. 2009.
[58] PLEXOS Software. (2019). *Plexos Overview & Tutorial*. [Online]. Available: https://www.plexos.info

Zhenyong Zhang received the bachelor's degree in control science and engineering from Central South University, Changsha, China, in 2015. He is currently pursuing the Ph.D. degree with the School of Control Science and Engineering, Zhejiang University, Hangzhou, China. He is currently a Visiting Ph.D. Student with the Singapore University of Technology and Design. His research interests include mobile computing and cyber-physical system security.

Ruilong Deng (S'11–M'14) received the B.Sc. and Ph.D. degrees in control science and engineering from Zhejiang University, Hangzhou, Zhejiang, China, in 2009 and 2014, respectively. He was a Research Fellow with Nanyang Technological University, Singapore, from 2014 to 2015, and an AITF Post-Doctoral Fellow with the University of Alberta, Edmonton, AB, Canada, from 2015 to 2018. He is currently an Assistant Professor with the School of Computer Science and Engineering, Nanyang Technological University. His research interests include smart grid, cyber security, and wireless networking.

David K. Y. Yau (M'10–SM'13) received the B.Sc. degree in computer science from the Chinese University of Hong Kong, and the M.S. and Ph.D. degrees in computer science from The University of Texas at Austin. He was Associate Professor of computer science with Purdue University, West Lafayette. He has been a Professor with the Singapore University of Technology and Design since 2013. His research interests include cyber-physical system and network security/privacy, wireless sensor networks, and smart grid IT.

Peng Cheng (M'10) received the B.Sc. degree in automation and the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2004 and 2009, respectively. From 2012 to 2013, he was a Research Fellow with Information System Technology and Design Pillar, Singapore University of Technology and Design. He is currently a Professor with the College of Control Science and Engineering, Zhejiang University. His research interests include networked sensing and control, cyber-physical systems, and control system security.

Jiming Chen (M'08–SM'11–F'18) received the B.Sc. and Ph.D. degrees in control science and engineering from Zhejiang University, Hangzhou, China, in 2000 and 2005, respectively. He was a Visiting Researcher with the University of Waterloo, from 2008 to 2010. He is currently a Changjiang Scholars Chair Professor (MOE) with the College of Control Science and Engineering, the Deputy Director of the State Key Laboratory of Industrial Control Technology, and a member of the Academic Committee with Zhejiang University. His research interests include the Internet of Things, sensor networks, networked control, and control system security. He is currently an IEEE VTS Distinguished Lecturer.