Resilient Distributed Control against False Data Injection Attacks for Demand Response

Shaohua Yang, *Student Member, IEEE,* Keng-Weng Lao, *Senior Member, IEEE,* Yulin Chen, *Member, IEEE,* and Hongxun Hui, *Member, IEEE*

***Abstract*—To maintain the power system balance, flexible load resources have been widely employed to provide operating reserves, which is named demand response (DR). The heating, ventilation, and air conditioning (HVAC) loads cover a high percentage of the total power consumption and have a significant regulation potential that deserves further investigation. To dispatch such dispersed HVACs, distributed control techniques are developed in the DR field due to their flexibility and scalability. However, the distributed DR system is vulnerable to potential cyber-attacks, i.e., false data injection (FDI) attacks, which may lead to DR failure. This paper proposes a resilient distributed controller to protect HVACs against FDI attacks. Firstly, an HVAC-based DR system is built by using distributed control to provide operating reserves for the power grid. Then, the adverse effect of FDI attacks is quantified with mathematical derivation. It is found that a small FDI attack can lead to severe power output deviations, which is a lethal problem for DR. On this basis, an FDI attack-resilient distributed controller is designed for HVACs so as to meet the grid's operating reserve requirements even under attacks. Moreover, the convergence of the proposed controller is proved based on the Laplace transform and final value theorem. Finally, case studies validate the effectiveness of the proposed resilient controller.**

***Index Terms*—Demand response, HVAC loads, distributed control, false data injection attack, resilient control.**

I. Introduction

ACCORDING to the statistics in 2021, global renewable energy generation capacity increased by more than 260 GW last year, accounting for roughly 80% of all new electricity capacity [1], which creates unprecedented power fluctuations on the grid due to power mismatch [2], [3]. Therefore, maintaining the real-time power balance between the demand-side and supply-side is important to the power grids [4]. The operating reserve is the generating capability that is "standing by" ready for service in the event that something happens on the power system [5], which plays a pivotal role in maintaining the system balance due to the ability to deal with power fluctuations [6]. Facing the massive increase in renewable energies, more flexible load resources on the demand-side are utilized to provide operating reserves [7], which is named demand response (DR) and has become a worldwide hot research topic [8].

This work was partly Funded by The Science and Technology Development Fund, Macau SAR (File/Project no. SKL-IOTSC-2021-2023, FDCT/0022/2020/A1, 0003/2020/AKP). (Corresponding author: *Kengweng Lao*, *Yulin Chen*.)

S. Yang, K. Lao and H. Hui are with the State Key Laboratory of Internet of Things for Smart City and Department of Electrical and Computer Engineering, University of Macau, Macao, 999078 China (email: yc17436@um.edu.mo, johnnylao@um.edu.mo, hongxunhui@um.edu.mo).

Y. Chen is with the Hainan Institute of Zhejiang University, Sanya, 572025 China (email: cheny12017@zju.edu.cn).

Among different flexible load resources, the heating, ventilation, and air conditioning (HVAC) loads can be one of the most potential candidates. We focus on HVACs to provide DR services for three reasons. i) The HVAC accounts for more than 40% of the overall power consumption in modern cities [9], which implies HVACs have a significant regulation potential. ii) Due to the air's thermal inertia, the indoor temperature can be kept within the comfortable range when the HVAC is involved in DR [10]. iii) The HVACs of nearby neighbors are close to each other in apartments. The fact that wireless networks can cover communication among HVACs makes distributed control for numerous HVACs a reality [11].

In fact, many works have been developed to validate the effectiveness of HVAC involvement in DR for system balancing. For example, in [12], the HVACs are regulated to provide operating reserves for the power grid considering the daily demand profile. In [13], HVACs as important flexible load resources are modeled and regulated to provide operating reserves for multi-area power grids. In addition, some DR-related demonstration projects have also been implemented. For examples, in Bornholm, Denmark, the EcoGrid EU project helps the grid with high penetration of renewables maintain system balance [14]. In the United States, a smart HVAC program involving thousands of customers is originated by the company *Avangrid*, which shows that DR for power systems can be performed by controlling the power of HVACs [15]. In Jiangsu province, China, the DR project named the friendly interactive grid has been implemented to decrease pear-valley loads and maintain the supply and demand balance by controlling HVACs [16]. These existing practical applications show no practical technological obstacles in equipping HVACs with control modules to provide the DR to power systems. As a result, the HVAC has become an important demand-side resource to participate in DR.

In general, the control methods of HVAC-based DR can be classified into two categories, i.e., centralized control and distributed control. The key difference between them is the communication mode [17]. In centralized control methods, a control center communicates with numerous HVACs directly [18]. However, this method requires high investment in communication infrastructure between the control center and large-scale HVACs [19]. In contrast, in distributed control methods, most HVACs only require point-to-point communication with their neighbors (but not the control center), so heavy communication requirements can be avoided [20]. In fact, applying distributed control technologies in DR has become a new research direction. For example, by utilizing a distributed control, [21] shows that the communication and computational burden on aggregators can be relieved significantly. An inge-

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2023.3287205

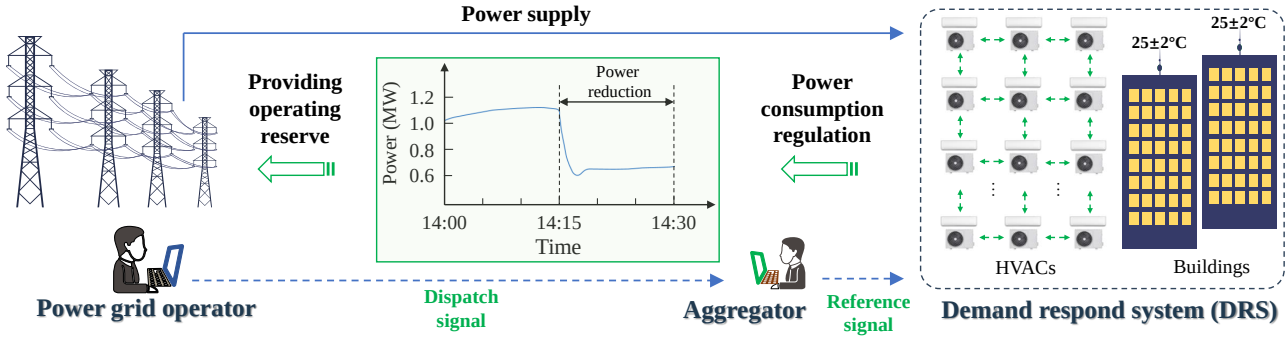

Fig. 1. The overall dispatch framework of the HVAC-based DRS.

nious distributed consensus control for HVACs is presented in [22] to meet the grid's requirements while protecting the customer's data privacy. A two-layer distributed controller for HVACs is designed in [23] to decrease the power variation caused by solar generators and loads in a building-microgrid community. All studies above have made remarkable progress in the distributed control of HVACs. However, potential cyber-attacks are not considered in the literature above.

Most of the existing works assume that the distributed control of HVACs can be implemented under a safe cyber environment [24]. However, cyber-attacks are ever-present in the control system, which has led to many security problems and has grown to be a critical issue for industrial control systems [25]. For example, in the U.S., the control system of Davis–Besse nuclear power station was penetrated by a cyber-attack in 2003 [26]. In 2010, the Natanz enrichment plant in Iran was compromised by the Stuxnet worm [27]. In 2015, the supervisory control and data acquisition (SCADA) network of the Ukrainian electric power system was compromised by false data injection (FDI) attacks [28], leading to several hours of power outages that affected about 225,000 consumers [29]. Control systems can be disrupted by various types of cyber-attacks, such as denial of service (DoS) attacks, replay attacks, FDI attacks, etc [30], [31]. Among them, the FDI attack is considered the most typical and threatening risk for the distributed control system of HVACs since the control performance can be seriously undermined by malicious false data [32].

In fact, many works have been developed to verify the destructiveness of FDI attacks. Liu et al. [33] present an FDI attack from the attacker's perspective and indicate that malicious attacks can undermine the state estimation of power grids. In addition, FDI attacks can also destroy the distributed control systems for secondary frequency restoration in micro-grids [34]. Considering HVAC-based DR with distributed control, state estimation is hardly performed, since there is not enough global information to establish a relationship between the measurement data and each HVAC's power state. As a result, the technique in state estimation (e.g., bad data detection) can not be employed to detect the FDI attack. Therefore, we address the adverse effects of FDI attacks from the perspective of designing an attack-resilient controller. To this end, there are still two significant research gaps to be filled in the HVAC-based demand response system (DRS) under FDI attacks. (i) The principle of how FDI attacks affect the power output of DRS is unclear, so the seriousness of adverse effects caused by FDI attacks cannot be quantified. (ii) So far, there is a lack of research on the defense of DRS against FDI attacks.

To address these problems, we quantify the impact of FDI attacks on HVACs mathematically and design a novel resilient distributed controller for the HVAC-based DRS against FDI attacks. The major contributions of this paper are threefold:

1) An HVAC-based DRS with distributed control is built to provide operating reserves for the power grid. In this system, dispersed HVACs are aggregated under the condition that both power and comfort states are shared fairly among all HVACs.
2) The relationship between the arbitrary FDI attack and the HVAC's power output is derived mathematically for the first time to quantify the impact of the FDI attack on DRS. It is found that even if only one HVAC is attacked, all the HVACs' power states will deviate, and the overall power output of the DRS cannot meet the grid's requirement.
3) A resilient distributed controller is proposed for HVACs against the FDI attacks (the attack vector can be arbitrary vector) so as to ensure that the grid's operating reserves can be provided by the DRS. Moreover, the convergence of the proposed controller is proved strictly, and the steady-state value solved shows that adverse effects caused by FDI attacks can be completely eliminated.

The remainder of this article is organized as follows. An HVAC-based DRS with distributed control is developed in Section II. In Section III, the FDI attack model is given first, then the impact of FDI attacks on DRS is analyzed and quantified. In Section IV, a resilient distributed controller is proposed against FDI attacks. Case studies are presented in Section V. Finally, Section VI concludes this paper.

II. HVAC-based DRS with Distributed Control

In this section, the model of HVAC is given first. Then, a distributed control is developed for the HVAC-based DRS, so that large-scale HVACs can be aggregated to provide operating reserves to power systems. The overall dispatch framework of HVAC-based DRS can be illustrated in Fig. 1. In general, the regulation capacity of one HVAC is not enough to engage in DR directly. Therefore, to provide significant operating reserves, numbers of HVACs have to be aggregated. To maintain the system power balance, the grid operator sends dispatch signals to the aggregator (DR provider) for operating reserves.

After receiving dispatch signals, the aggregator adjusts the overall power output of the DRS by the distributed control method. In the distributed manner, just a few HVACs need to receive the aggregator's reference signals, and most HVACs only require communication peer-to-peer with their neighbors. Then, all the HVACs can cooperatively complete the DR task with a sparse communication topology. As a result, the DSR's power consumption is regulated, and the grid's requirements for operating reserves are met.

A. Thermodynamic Model of HVACs

Indoor comfort is a major concern for customers when participating in DR. Therefore, the thermodynamic model of the room needs to be developed. The room's heat variation can be calculated by heat gains H_{gain} and heat losses H_{loss}. The thermodynamics of the room corresponding to HVAC i can be expressed by an ordinary differential equation as below [9]:

$$c_A \rho_A V \frac{dT_i(t)}{dt} = H_{\text{gain}} - H_{\text{loss}} \ , \tag{1}$$

where c_A is the air heat capacity; ρ_A is the air density, V is the room's volume; $T_i(t)$ is the indoor temperature of room i at time t. Take the cooling mode as an example. The heat gain can be calculated by heat transfer from air leakages and the building envelop, which can be expressed as follows:

$$H_{\text{gain}} = U_h A_s (T_o - T_i(t)) + c_A \rho_A V n (T_o - T_i(t)) \ , \tag{2}$$

where U_h is the heat transfer coefficient, A_s is the envelope's surface area, T_o is the ambient temperature, n denotes the air exchange times. The heat loss can be expressed as follows:

$$H_{\text{loss}} = \eta \cdot \alpha_i(t) P_{\text{N}} \ , \tag{3}$$

where η is the coefficient of performance of HVAC, which implies the relationship between the input power and heat supply (cooling or heating); α_i is the power state of HVAC i, which stands for the percentage of rated power and can be expressed as below:

$$\alpha_i(t) = \frac{P_i(t)}{P_{\text{N}}}, \tag{4}$$

where $P_i(t)$ is the power of HVAC i at time t; P_{N} is the rated power of HVAC. As well, the denominator also could be the available regulation capacity quoted by the customer, taking into account customers' personal willingness.

Remark 1: The indoor temperature T_i can be controlled by adjusting the power state of HVAC i (α_i), as shown in (1)-(3).

To ensure customers' thermal comfort, the indoor temperature should be maintained within a comfortable range $[T_{\min}, T_{\max}]$, where $T_{\min} = T_s - \Delta T$ and $T_{\max} = T_s + \Delta T$ are the lower and upper indoor temperature bounds, respectively, where T_s is the set temperature and ΔT is the customers' tolerable temperature change. On this basis, the index of thermal comfort can be formulated as follows:

$$\beta_i(t) = \frac{T_i(t) - T_s}{\Delta T} \ , \tag{5}$$

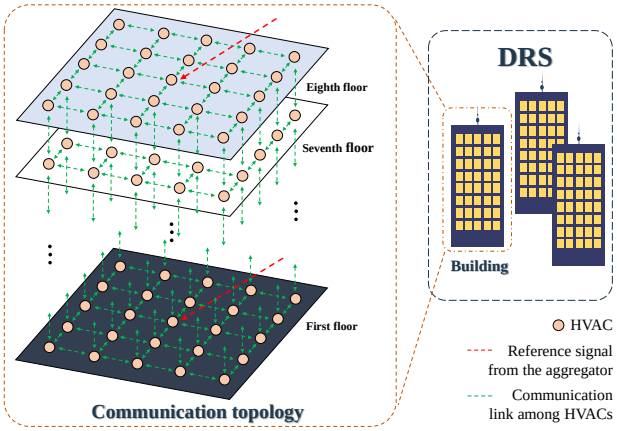

Fig. 2. The communication graph of HVACs in the DRS.

where β_i is the comfort state of room i. The value range of this index is $\beta_i \in [-1, 1]$. The lower bound -1 and upper bound 1 represent cold and heat tolerance limits, respectively.

Substituting (5) to (1)-(3), the variation of comfort state can be formulated as follows:

$$\begin{aligned} \frac{d\beta_i(t)}{dt} = &- \frac{\eta \cdot P_{\text{N}}}{\Delta T c_A \rho_A V} \alpha_i(t) - \frac{(U_h A_s + c_A \rho_A V n)}{c_A \rho_A V} \beta_i(t) \\ &+ \frac{(U_h A_s + c_A \rho_A V n)(T_o - T_s)}{\Delta T c_A \rho_A V} \ . \end{aligned} \tag{6}$$

According to (6), the relationship between the indoor comfort state and HVAC's power state is established.

For brevity, in the following, $G_{th} = U_h A_s + c_A \rho_A V n$ and $C_{th} = c_A \rho_A V$ are denoted as the thermal conductance and thermal capacitance coefficient, respectively. On this basis, the state-space equation of the HVAC i can be given as follows:

$$\begin{aligned} \underbrace{\begin{bmatrix} \dot{\alpha}_i(t) \\ \dot{\beta}_i(t) \end{bmatrix}}_{\dot{\boldsymbol{x}}_i} = &\underbrace{\begin{bmatrix} 0 & 0 \\ -\frac{\eta \cdot P_{\text{N}}}{\Delta T C_{th}} & -\frac{G_{th}}{C_{th}} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \alpha_i(t) \\ \beta_i(t) \end{bmatrix}}_{\boldsymbol{x}_i} \\ &+ \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{B} u_i(t) + \underbrace{\begin{bmatrix} 0 \\ \frac{G_{th}(T_o - T_s)}{\Delta T C_{th}} \end{bmatrix}}_{C} \ , \end{aligned} \tag{7}$$

where $\boldsymbol{x}_i = [\alpha_i, \beta_i]^T$ is the state variable vector, including the power state and comfort state defined above; u_i is the control input needed to design; A is the state transition matrix; B is the input matrix; C is the supplemental matrix.

B. Distributed Control for HVAC-based DRS

For the DRS in Fig. 1, the communication topology of HVACs can be shown in Fig. 2. For example, there are some buildings in the DRS, each with eight floors and 25 rooms on each floor. Each room is assumed to have one HVAC, which is equipped with a communication and control module. As shown in Fig. 2, only a few HVACs need to receive the aggregator's reference signal (i.e., the red communication links). Moreover, for the green communication topology, most HVACs just require peer-to-peer communication with neighbors, including the upper, lower, left, and right neighbors. By the cooperative

control of HVACs with this communication topology, the DRS can achieve the global objective of the grid's dispatch assignment.

Based on the feedback linearization, the distributed control for HVACs' power regulation can be formulated as $\dot{\alpha}_i = u_i$. The control input of HVAC i (i.e., u_i), using the local and neighboring information as well as the reference signal from the aggregator, can be expressed as follows:

$$u_i = \dot{\alpha}_i = -k_\alpha \sum_{j \in \mathcal{N}_i} a_{ij}(\alpha_i - \alpha_j) + b_i(\alpha_i - \alpha_{ref}) \ , \quad (8)$$

where $k_\alpha > 0$ is a coupling gain; $\mathcal{N}_i$ is the set of all the neighbors of HVAC i; a_{ij} is the entry of the adjacency matrix; α_{ref} is the reference signal; b_i is the pinning gain, where $b_i = 1$ implies the HVAC i can receive the reference signal from the aggregator and $b_i = 0$ otherwise. The motivation for selecting α as the control is to regulate (reduce or increase) the power of HVACs to provide a DR to the power system. The consensus control of α can enable HVAC to participate in DR in a fair way, i.e., providing operating reserve services in equal proportions. The corresponding matrix form of (8) can be represented as follows:

$$\boldsymbol{u} = \dot{\boldsymbol{\alpha}} = -k_\alpha(\mathcal{L} + \mathcal{B})\boldsymbol{\alpha} + k_\alpha \alpha_{ref} \mathcal{B} \mathbf{1}_N \ , \quad (9)$$

where $\boldsymbol{u} = [u_1, u_2, \ldots, u_N]^T$ is the designed control input vector; $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_N]^T$ is the power state vector; $\mathcal{L}$ is the Laplacian matrix; $\mathcal{B} = \text{diag}\{\boldsymbol{b}\} \subseteq \mathbb{R}^{N \times N}$ is the pinning matrix with $\boldsymbol{b} = [b_1, b_2, \ldots, b_N]^T$; $\mathbf{1}_N$ is the N-dimensional vector filled with 1 entries.

Under the distributed control manner in (9), the dynamics of the DRS with the HVAC aggregation ($\forall i \in \mathcal{I}^1$) can be represented as follows:

$$\underbrace{\begin{bmatrix} \dot{\boldsymbol{\alpha}} \\ \dot{\boldsymbol{\beta}} \end{bmatrix}}_{\dot{\boldsymbol{x}}} = \underbrace{\begin{bmatrix} -k_\alpha(\mathcal{L} + \mathcal{B}) & 0 \\ -\frac{\eta \cdot P_N}{\Delta T C_{th}} \boldsymbol{I}_N & -\frac{G_{th}}{C_{th}} \boldsymbol{I}_N \end{bmatrix}}_{T} \underbrace{\begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{bmatrix}}_{\boldsymbol{x}} + \underbrace{\begin{bmatrix} k_\alpha \alpha_{ref} \mathcal{B} & 0 \\ 0 & \frac{G_{th}(T_o - T_s)}{\Delta T C_{th}} \boldsymbol{I}_N \end{bmatrix}}_{W} \underbrace{\begin{bmatrix} \mathbf{1}_{2N} \end{bmatrix}}_{\boldsymbol{c}} \ , \quad (10)$$

where T and W are the state transition matrix and supplemental matrix of this DRS, respectively; $\boldsymbol{I}_N \subseteq \mathbb{R}^{N \times N}$ is the identity matrix; $\boldsymbol{x} \subseteq \mathbb{R}^{2N}$ is the state variable vector of the DRS including the power and comfort states; $\boldsymbol{c}$ is the $2N$-dimensional vector filled with 1 entries. The state-space equation of overall HVAC-based DRS with distributed control is established as shown in (10).

Through separation principle, the power state $\boldsymbol{\alpha}$ and the comfort state $\boldsymbol{\beta}$ can be analyzed separately, since the matrix T in (10) is a triangular matrix [35]. For the power state $\boldsymbol{\alpha}$, the dynamics of the DRS containing N HVACs can be shown in (8). According to this dynamics, since $k_\alpha > 0$ and the matrix $\mathcal{L} + \mathcal{B}$ is a positive definite matrix, the steady-state value of $\boldsymbol{\alpha}$ can converge to $\alpha_{ref}\mathbf{1}_N$ [36]. This implies the power state of each HVAC can reach a steady-state consensus according to the aggregator's reference signal. On this basis, from the comfort state part in (10), it can be known that for all HVACs, the steady-state consensus of $\boldsymbol{\beta}$ can also be reached. Therefore, by adjusting the reference signal α_{ref}, the DRS can ensure the customer's comfort, as well as provide operating reserves to the power system.

[1] $\mathcal{I} = \{1 \leq i \leq \text{N} \mid i \in \mathbb{Z}\}$ is the set of all HVACs in the DRS with the distributed control.

III. Quantification of FDI Attack Impact on the HVAC-based DRS

In this section, an FDI attack model is defined first. In addition, how it affects the HVAC-based DRS is also quantified.

A. FDI Attack Modeling for HVAC's Controller

According to the white paper 'Cybersecurity for Industrial Automation and Control Environments', cyber-attack has become a serious threat to industrial control systems [25]. In fact, a malicious FDI attack on the DRS's control system can lead to the failure of DR. The controller of each HVAC, as a component of computing and information exchange, is indeed vulnerable to FDI attacks. When the HVAC's controllers are compromised by FDI attacks, the corrupted control input vector ($\boldsymbol{u}_\xi$) of the DRS contains injected malicious data [33] and can be expressed as follows:

$$\boldsymbol{u}_\xi = \boldsymbol{u} + \boldsymbol{\Xi}, \quad (11)$$

where $\boldsymbol{u} = [u_1, u_2, \ldots, u_N]^T$ is the vector of each HVAC's original control input; $\boldsymbol{\Xi} = [\nu_1\xi_1, \nu_2\xi_2, \ldots, \nu_N\xi_N]^T$ is an FDI attack vector, i.e., malicious injected data added to the original control inputs; ν_i is a binary number where $\nu_i = 1$ indicates the presence of attack, otherwise, $\nu_i = 0$; ξ_i is the injected data for the HVAC i.

Under the FDI attack, for the specific controller of HVAC i, the original control input u_i is replaced with a phony control input $u_i + \nu_i\xi_i$.

Attack on the u_i of one HVAC i can indirectly affect the calculation of the control signals of all HVACs. This is because the control signal is calculated from distributed control by using the neighboring information. It is worth noting that this neighboring information would be affected by the attack. In addition, the adverse impact caused by an attack can propagate through the information exchange, which can make neighboring information of each HVAC affected by the attack. Therefore, the control signal calculated using distributed control is affected by the cyber-attack.

B. Impact of FDI Attacks on the DRS

The attacker can choose an arbitrary non-zero vector as an attack vector $\boldsymbol{\Xi}$. From (11), we know that more non-zero entries in the attack vector $\boldsymbol{\Xi}$ imply more controllers of HVAC are attacked, which is more difficult for attackers. However, by theoretical analysis, we find that even if only one controller of HVAC is attacked, the DR task will also be failed. Theorem 1 will provide the result.

Theorem 1: Suppose the arbitrary attack vector is $\boldsymbol{\Xi}$. The steady-state value of power state vector converges to $\boldsymbol{\alpha} =$

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2023.3287205

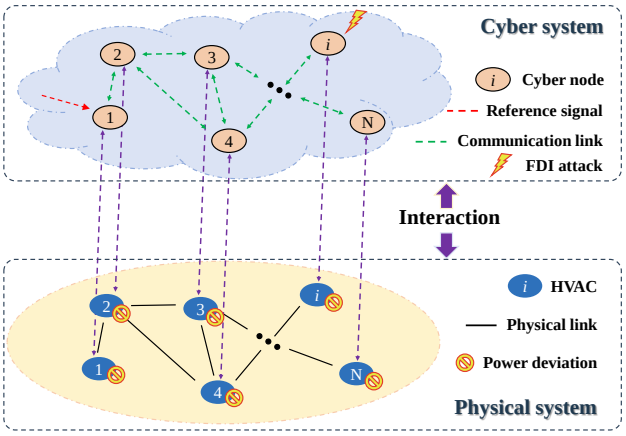

Fig. 3. Only one FDI attack can deviate the power of all HVACs.

$\alpha_{ref}\mathbf{1}_N + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon} = [k_\alpha(\mathcal{L}+\mathcal{B})]^{-1}\boldsymbol{\Xi}$ is the error vector. This implies all HVACs' power states will deviate from the aggregator's reference signal, even if only one controller is attacked.

The description of Theorem 1 can be illustrated as Fig. 3. Only one FDI attack can make all HVACs have power deviations, which means the DR task fails.

Proof 1: Please see Appendix B-A and Appendix B-B. ■

To further analyze the adverse impact on each HVAC's power state caused by the FDI attack, we define the injection error matrix $\boldsymbol{\Theta} = [k_\alpha(\mathcal{L}+\mathcal{B})]^{-1}$.

Theorem 1 implies that the error vector $\boldsymbol{\epsilon}$ is the combination of columns of the matrix $\boldsymbol{\Theta}$, which can be represented by (12):

$$\begin{aligned}\boldsymbol{\epsilon} &= \boldsymbol{\Theta}\boldsymbol{\Xi} \\ &= \sum\nolimits_{i\in\mathcal{I}} \nu_i\xi_i\boldsymbol{\theta}_i \\ &= \sum\nolimits_{i\in\mathcal{I}_{attack}} \nu_i\xi_i\boldsymbol{\theta}_i + \sum\nolimits_{i\notin\mathcal{I}_{attack}} \nu_i\xi_i\boldsymbol{\theta}_i \\ &= \sum\nolimits_{i\in\mathcal{I}_{attack}} \xi_i\boldsymbol{\theta}_i, \end{aligned} \tag{12}$$

where the vector $\boldsymbol{\theta}_i$ represents the ith column of the matrix $\boldsymbol{\Theta}$; the symbol $\mathcal{I}_{attack}$ denotes the set of target HVACs chosen by the attacker. For the specific HVAC k, the power state's error ϵ_k caused by the FDI attack can be expressed as follows:

$$\epsilon_k = \sum\nolimits_{i\in\mathcal{I}_{attack}} \xi_i\theta_{ik},\ \forall k\in\mathcal{I}, \tag{13}$$

where θ_{ik} is the kth entry of the column vector $\boldsymbol{\theta}_i$. According to (13), even if the set $\mathcal{I}_{attack}$ has only one entry (i.e., only one attacked controller), all the HVACs' power states will have errors, which are determined by the entry θ_{ik}.

The fact that all the IACs' power states deviate implies that the DRS's overall power output is no longer under the control of the aggregator. Under a attack, the DRS's overall power deviation ΔP can be expressed as follows:

$$\Delta P = \sum_{k\in\mathcal{I}} \epsilon_k. \tag{14}$$

According to (14), the overall power of DRS has a significant deviation, which implies the grid's operating reserve requirement cannot be met, and the DR is failed.

In addition, each room's comfort state β is also affected by the FDI attack. Combining Theorem 1 with dynamic characteristics in (10), the comfort state's error δ_k for the specific room k can be derived as follows:

$$\delta_k = \sum\nolimits_{i\in\mathcal{I}_{attack}} \xi_i \frac{-\eta\cdot P_{\mathrm{N}}\theta_{ik}}{G_{th}\Delta T},\ \forall k\in\mathcal{I}. \tag{15}$$

The equation (15) illustrates that there are deviations in the comfort states, which may make each indoor temperature out of its comfortable range.

Remark 2: A small FDI attack can make customers uncomfortable, and the operating reserve providing failed, which is a lethal problem for the DRS and must be addressed.

IV. Attack-resilient Controller Design

A. Resilient Distributed Control for HVAC-based DRS

To go against FDI attacks, a resilient distributed controller is proposed to protect the HVAC-based DRS to provide operating reserves for the power grid. The proposed resilient distributed controller for the HVAC i is designed as follows:

$$\begin{aligned} u_i &= \dot{\alpha}_i \\ &= -k_\alpha \int [\sum_{j\in\mathcal{N}_i} a_{ij}(\alpha_i-\alpha_j) + b_i(\alpha_i-\alpha_{ref})]\mathrm{d}t - \alpha_i, \end{aligned} \tag{16}$$

Oriented to a large number of HVACs, this controller in (16) needs to be expressed in matrix form, which can be expressed as follows:

$$\boldsymbol{u} = \dot{\boldsymbol{\alpha}} = -k_\alpha \int [(\mathcal{L}+\mathcal{B})(\boldsymbol{\alpha} - \alpha_{ref}\mathbf{1}_N)]\mathrm{d}t - \boldsymbol{\alpha}. \tag{17}$$

Theorem 2: For the arbitrary FDI attack vector $\boldsymbol{\Xi}$, the steady-state convergence result of the proposed controller is $\boldsymbol{\alpha} = \alpha_{ref}\mathbf{1}_N$, i.e., $\forall i\in\mathcal{I}, \alpha_i = \alpha_{ref}$. This implies by using the proposed resilient distributed controller, the steady-state values of the power states are completely unaffected by attacks, and hence all HVACs can track the aggregator's reference signal even under the FDI attack.

Proof 2: Please see Appendix B-C and Appendix B-D. ■

Theorem 2 implies that with the proposed resilient distributed controller, the power state of each HVAC can reach a steady-state consensus. On this basis, according to the comfort state part in (10), it can be known that for all the HVACs, both the power state sharing and comfort state sharing can be satisfied, shown as follows:

$$\lim_{t\to\infty} \|\alpha_i(t) - \alpha_{ref}\| = 0,\ \forall i\in\mathcal{I}\ , \tag{18}$$

$$\lim_{t\to\infty} \|\beta_i(t) - \beta_k(t)\| = 0,\ \forall i,k\in\mathcal{I}\ . \tag{19}$$

Combining (18) with the comfort state's dynamics in (6), the steady-state value of comfort state β_∞ is derived as follows:

$$\beta_\infty = \lim_{t\to\infty}\beta_i(t) = \frac{T_o - T_s}{\Delta T} - \frac{\eta\cdot P_{\mathrm{N}}}{G_{th}\Delta T}\alpha_{ref},\ \forall i\in\mathcal{I}. \tag{20}$$

From (18)-(20), it can be found that steady-state values of both the power state α and comfort state β are in fact controlled by the aggregator's reference signal α_{ref}. In other words, the DRS can provide operating reserves to power systems and guarantee the comfort of customers by the reference signal. Therefore, with the help of our proposed

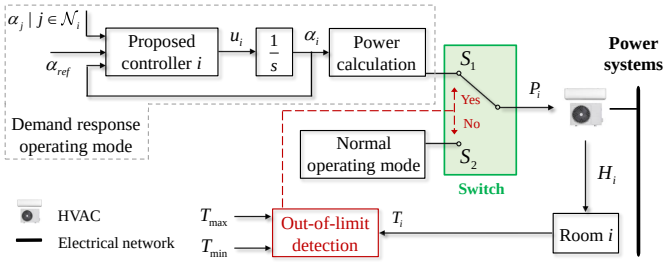

Fig. 4. The withdrawal mechanism of HVACs for ensuring the customer's comfort requirements.

resilient distributed controller, the DR task can be completed successfully even under FDI attacks.

The proposed control can be implemented in practice. This is because the compressor is the main power-consuming component of the HVAC, and there is a corresponding relationship between the power and speed of the compressor [13]. As a result, the target power could be reached by controlling the speed of the compressor. In fact, some large HVAC companies (e.g., GREE, 20.6% global market share of residential air conditioners) have been promoting to the market intelligent air conditioning products that can control the compressor's speed and the HVAC's power [37].

B. Withdrawal Mechanism to Ensure Comfort Requirement

There are two operating modes for HVAC, i.e., (i) normal operating mode, and (ii) demand response operating mode. In the normal operating mode, the HVAC's power is mainly determined by the deviation between the temperature setpoint and the current indoor temperature, which can be shown as below [13]:

$$\Delta P_s(s) = C(s) \cdot \Delta T_{dev}(s), \tag{21}$$

$$C(s) = \sigma + \mu/s, \tag{22}$$

$$\Delta T_{dev}(s) = T(s) - T_s(s), \tag{23}$$

where $\Delta P_s(s)$ is the power adjustment to ensure the set temperature; $C(s)$ is the PI-based temperature controller of the HVAC; σ and μ are the constant coefficients of the controller, respectively; and ΔT_{dev} is the deviation between the indoor temperature T and the set temperature T_s. In this normal operating mode, the indoor temperature can be controlled to the set temperature T_s.

In the demand response operating mode, the proposed controller described in (16) is activated to meet the regulation requirements of the power system. In general, for a given available regulation capacity from the evaluation, the proposed controller can regulate distributed HVACs to satisfy the power system's DR requirement even under cyber-attacks. In some emergency situations, for example, when an individual HVAC's regulation capability is evaluated incorrectly, or when an individual HVAC suffers unexpected physical damage, the temperature may be affected negatively and be out of the tolerable temperature range. To address this issue, a withdrawal mechanism is added in control, i.e., when the comfort limit is touched for a particular customer, the corresponding HVAC need to withdraw from the demand response operating mode to the normal operating mode. In this way, the indoor temperature can be recovered back to the comfort range.

TABLE I
TYPICAL PARAMETERS FOR HVACS AND CORRESPONDING ROOMS

Symbols	Parameters	Values	Units
H	Height of room	3	m
S	Living area	100	m^2
c_A	Heat capacity of air	1.005	kJ/(kg°C)
ρ_A	Density of air	1.205	kg/m^3
U_h	Heat transfer coefficient	7.69	W/(m^2 °C)
n	Air exchange times	0.5	1/h
η	Coefficient of performance	3	–
P_N	Rated power	8	kW
P_{ini}	Initial power	$\mathcal{U}(2,8)$[1]	kW
T_{ini}	Initial indoor temperature	$\mathcal{U}(23.4,26.6)$	°C
T_s	Set temperature	25	°C
ΔT	Tolerable temperature change	±2	°C
T_o	Ambient temperature	31	°C

[1] $\mathcal{U}$ denotes uniform distributions.

The withdrawal mechanism of HVACs can be illustrated in Figure 4. In Figure 4, H_i is the cooling capacity of the HVAC i; S_1 and S_2 are switches. If the indoor temperature is within the comfort range, switch S_1 remains closed, and switch S_2 remains open. In this scenario, the HVAC provides the operating reserve for power systems. In contrast, if the indoor temperature is detected to touch the customer's comfort limits (i.e., $T_{\min}$ or $T_{\max}$), switch S_1 will be opened and switch S_2 will be closed. This means that the consensus control output is disabled for this HVAC while the normal operating state is enabled. In other words, the corresponding HVAC withdraws from providing service to power systems and returns to the original operating state to ensure the comfort requirement. It is worth noting that the withdrawal of HVAC is different from the attack on HVAC. The key difference is that when an HVAC is compromised by cyber-attacks, the control system is unaware of it, thus leading to failures in the control and the providing operating reserve. However, when an HVAC withdraws actively, this event can be known, and an appropriate adjustment can be done. To be specific, the adjustment is as follows: the withdrawn HVAC stops uploading its own power state α_i as exchange information. Instead, the withdrawn HVAC functions as an intermediary for its neighbors, and transfers the neighboring HVACs' information. As a result, the normal information exchange in distributed control is not affected. Moreover, for the aggregation of HVACs, despite the withdrawal of individual HVACs, the rest of the HVACs will still provide significant operating reserves for power systems. Therefore, this withdrawal mechanism can ensure the comfort requirement of an individual customer (who is in an emergency) while still maintaining the service quality of the DRS to the power system.

V. CASE STUDY AND VERIFICATION

A. Test System

Both the quantitative analysis of FDI attack impact and the proposed resilient distributed controller are validated in an HVAC-based DRS. It is assumed that the aggregator controls 200 HVACs in this DRS to meet the grid's operating reserve

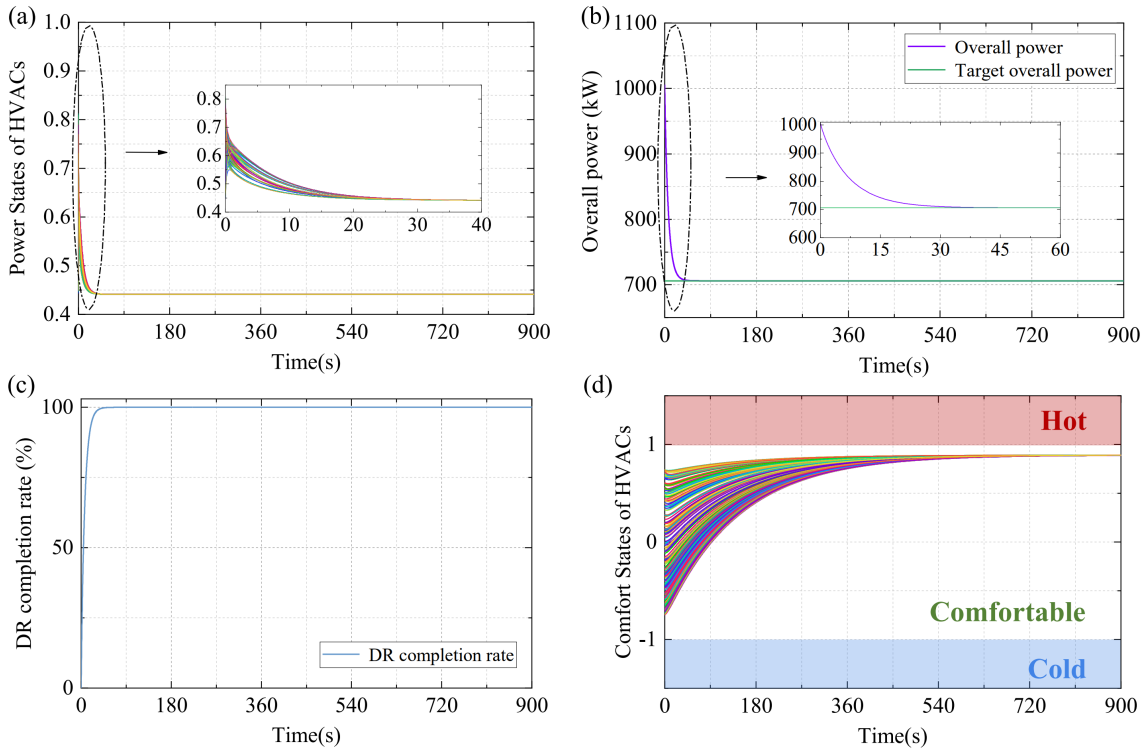

Fig. 5. Benchmark case 1 without any attack: the performance of HVAC-based DRS with the original distributed controller: (a) Power states of HVACs; (b) Overall power of the DRS; (c) DR completion rate; (d) Comfort states of HVACs.

requirements, and its associated communication topology is shown in Fig. 2. The communication frequency of HVACs in this test DRS is set to 0.01 kHz, which is well suited to general wireless communication [38]. The parameters of HVACs and corresponding rooms are detailed in Table I. The parameters of ambient temperature are realistic test data in Macao on July 1st, 2022 [39]. More parameters are The CPU model for case studies is Intel Core i7-10700 at 2.90 GHz, and the simulation environment is Matlab R2022a.

The test is conducted according to the following process: At 14:15, the aggregator receives a dispatch signal from the grid operator with information on the regulation capacity (300 kW) and duration time (15 minutes). The aggregator, as the manager of the DRS, has the responsibility to provide operating reserves to the grid. Therefore, the aggregator needs to adjust the DRS's overall power output as soon as possible until the dispatch signal is met. Attacks are also launched at 14:15, which can lead to power deviation and thus failure of the DR task. The total test time is 15 minutes.

Two types of cyber-attacks are launched in case studies, i.e., single-action cyber-attack and series-of-actions cyber-attacks. The single-action cyber-attack means that at the beginning of DR, the hacker launches a cyber-attack by injecting false data, and this attack lasts until the end of the DR (the attack lasts 15 minutes). The series-of-actions cyber-attack means that at the beginning of DR, the hacker maliciously launches a cyber-attack by injecting false data, and this round of attack lasts 3 minutes (from 0 to 3 minutes). At 3 minutes, the first round of attack is stopped, and the second round of attack is launched, which also lasts 3 minutes (from 3 to 6 minutes). In this way, within 15 minutes of the DR, there are 5 times of cyber-attack actions in total. In addition, the value of the injected data under each attack action can be different. Thus dynamic false data can be injected by this series-of-actions cyber-attack.

There are two benchmark cases and three scenarios with cyber-attacks. Symbols B-1, B-2, S-1, S-1, and S-3 represent benchmark case 1, benchmark case 2, attack scenario 1, attack scenario 2, and attack scenario 3, respectively, which can be listed as follows:

[B-1] Benchmark case without any cyber-attack and without the proposed resilient controller.

[B-2] Benchmark case without any cyber-attack but with the proposed resilient controller.

[S-1] Attack scenario with a single-action cyber-attack without the proposed resilient controller.

[S-2] Attack scenario with series-of-actions cyber-attacks without the proposed resilient controller.

[S-3] Attack scenario with series-of-actions cyber-attacks with the proposed resilient controller.

B. Benchmark Cases without Any Attack

The benchmark cases without any attack (B-1 and B-2) are shown as Figure 5 and 6.

Benchmark case 1 is shown as Figure 5, which is the performance of HVAC-based DRS without any attack, and based on the original distributed controller described in (8). As shown in Figure 5 (a) and (d), the consensus control of different HVAC's power states and comfort states can be achieved. In addition, as shown in Figure 5 (b), the DRS's overall power can be reduced and converged to the target value at about 38s (less than 1 minute), which implies the DRS can regulate the overall power according to the aggregator's signal.

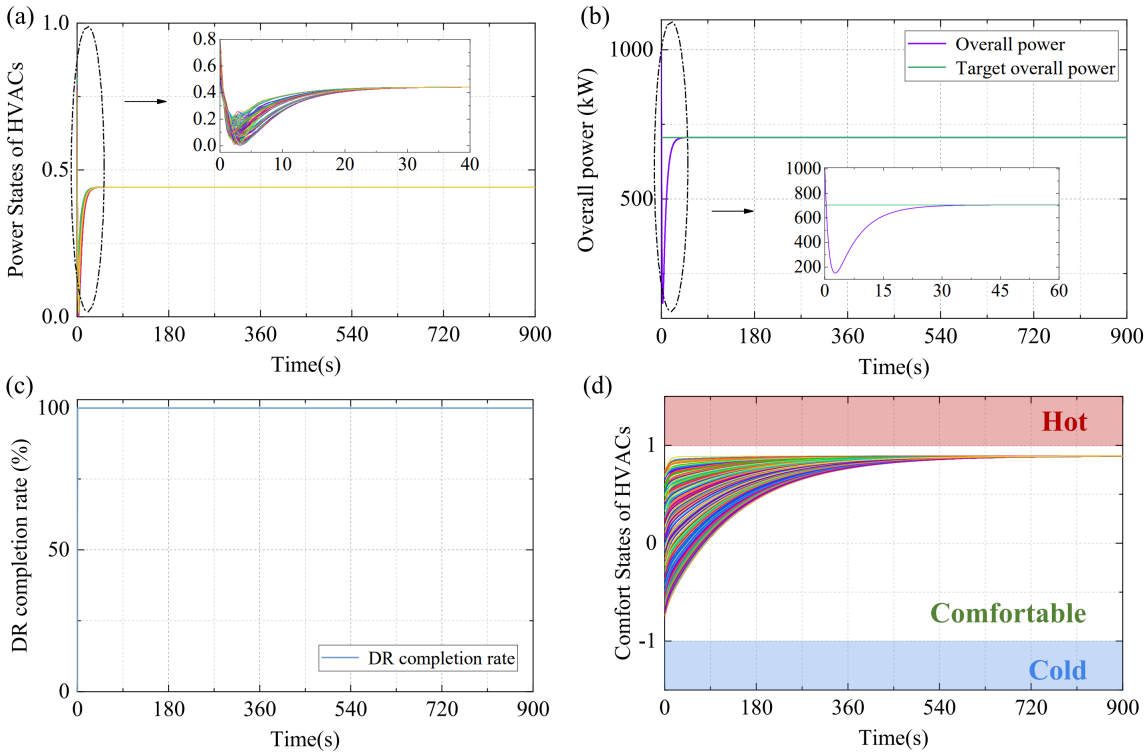

Fig. 6. Benchmark case 2 without any attack: the performance of HVAC-based DRS with the proposed resilient distributed controller: (a) Power states of HVACs; (b) Overall power of the DRS; (c) DR completion rate; (d) Comfort states of HVACs.

Moreover, as shown in Figure 5 (c), the DR completion rate can reach 100%, which implies the DR can be achieved.

Benchmark case 2 is shown as Figure 6, which is also the performance of HVAC-based DRS without any attack, but based on the proposed resilient distributed controller described in (16). From Figure 6 (a) and (d), the consensus control of different HVAC's power states and comfort states can also be achieved with the proposed controller described in equation (22). In addition, as shown in Figure 6 (b), the DRS's overall power can be reduced and converged to the target value (i.e., from 1006 kW to 706 kW) at about 37s, which implies the response speed is also less than 1 minute. Moreover, from Figure 6 (c), the DR completion rate can also reach 100%, which implies the DR can also be achieved with the proposed controller.

It is worth noting that both response speeds in two benchmark cases are sufficient to meet the requirement of providing reserve capacity (i.e., the response time should be less than 5 minutes [40]). Therefore, without any attack, the proposed controller can also accomplish the power system's DR task.

C. Scenario 1: Result Analysis of the DRS with a Single-action Cyber-attack

According to Theorem 1, we know that even if only one HVAC is attacked, the DRS's overall power can have a severe deviation. It has been proved theoretically in Section III-B. To further validate the effectiveness of Theorem 1, it is assumed that only one HVAC (Take No. 13 HVAC as an example) has been compromised by an FDI attack in this case. The injected false data for No.13 HVAC's controller is set to 0.8, and the attack vector can be described as follows:

$$\Xi = [\underbrace{0}_{No.1}, 0, \ldots, 0, \underbrace{0.8}_{No.13}, 0, \ldots, 0, \underbrace{0}_{No.200}]^T. \quad (24)$$

Under this small attack, the performance of the HVAC-based DRS is shown in Fig. 7 from four different perspectives. Fig. 7 (a) shows that after receiving the dispatch signal from the grid operator, the power states of all the HVACs are adjusted in anticipation of a consistent convergence to the aggregator's reference value. Moreover, from the steady-state results in Fig. 7 (a), it is found that even if only one HVAC is attacked, the power states of all the HVACs can have different degrees of deviation and can no longer track the reference signal. For this reason, the overall power of DRS cannot converge to the target value, as shown in Fig. 7 (b).

The operating reserve is based on the overall power consumption, which can be calculated by the difference between the overall power consumption before and after the DR. In this case, the overall power needs to be reduced from 1006 kW to 706 kW to provide operating reserves to the grid. However, under a small attack, the overall power of DRS can only be reduced to 757 kW and cannot meet the grid's DR requirement.

Furthermore, the DR completion rate for tests, which implies the completion rate of the DR task, is defined as follows:

$$CR(t) = \frac{P_{\text{orig}} - P(t)}{P_{\text{reg}}}\%, \quad (25)$$

where $CR(t)$ is the DR completion rate at time t; P_{orig} is the original overall power of the DRS before participating the

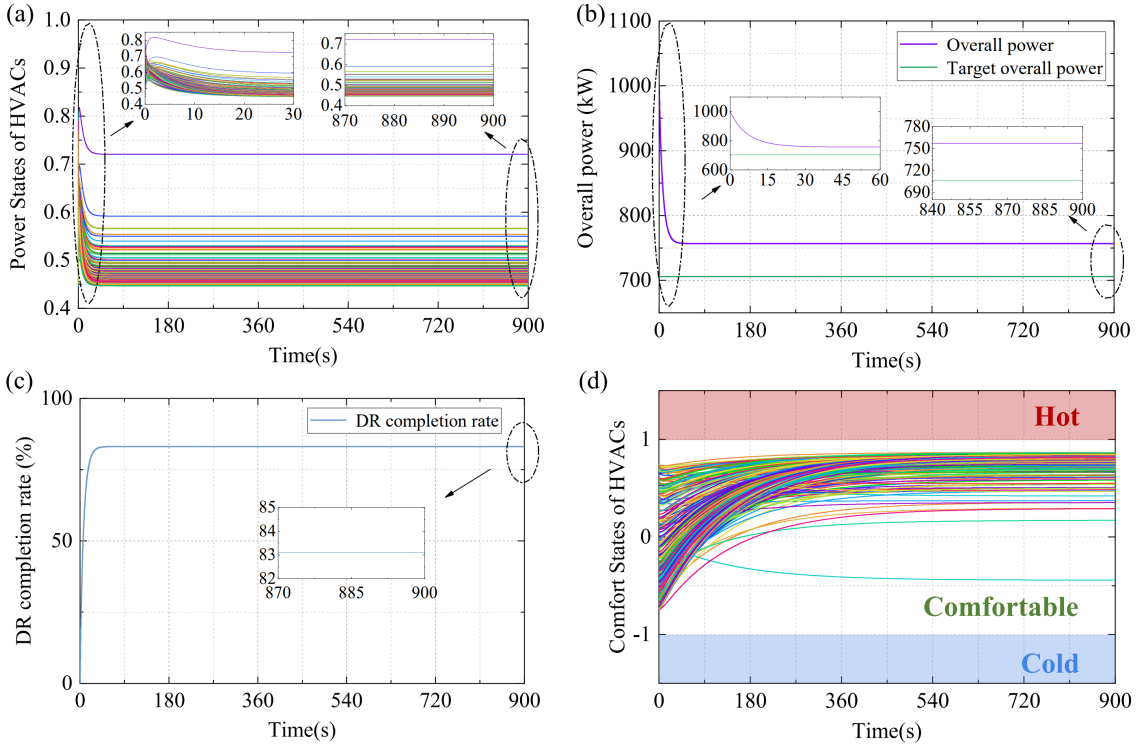

Fig. 7. The performance of HVAC-based DRS with a single-action cyber-attack: (a) Power states of HVACs; (b) Overall power of the DRS; (c) DR completion rate; (d) Comfort states of HVACs.

DR; $P(t)$ is the overall power of the DRS at time t; P_{reg} is the power regulation requirement of grid. Even if only one HVAC has been attacked, the regulation completion rate of the DRS can achieve only 83.10% of the target requirement, as shown in Fig. 7 (c). This indicates that a small FDI attack has a significant adverse effect on the DRS and leads to the failure of the DR task. Fig. 7 (d) demonstrates that the comfort state of each room is within the comfortable range because the attack is very mild. However, from the fairness perspective, the comfort states of different rooms cannot converge to a consensus. This indicates that the FDI attack can also have a different degree of impact on each room's comfort state. It is worth noting that deviation values in this test satisfy calculation results by using (13), (14), and (15). Therefore, the quantitative analyses of the impact of FDI attacks in Section III are verified to be correct.

D. Scenario 2: Result Analysis of the DRS with Series-of-actions Dynamic Cyber-attacks

Without loss of generality, this case further explores the impact of different attacks on the DRS. It is assumed that three HVACs (No. 12, 13, and 14 HVACs) are compromised and attack vectors are time-varying in this case. In particular, the injected data for the three HVACs' controllers at each time period are different, as shown in Table II. Take time periods I and IV as examples. Attack vectors at I and IV periods can be expressed as follows:

$$\Xi_{\text{I}} = [\underbrace{0}_{No.1}, \ldots, 0, \underbrace{0.8, 0.8, 0.8}_{No.12,\ 13,\ \&\ 14}, 0, \ldots, \underbrace{0}_{No.200}]^T, \quad (26)$$

$$\Xi_{\text{IV}} = [\underbrace{0}_{No.1}, \ldots, 0, \underbrace{0.2, 0.2, 0.2}_{No.12,\ 13,\ \&\ 14}, 0, \ldots, \underbrace{0}_{No.200}]^T. \quad (27)$$

TABLE II
TIME-VARYING ATTACK VECTORS AT DIFFERENT TIME PERIODS

Time Period	I	II	III	IV	V
No.12 HVAC	0.8	0.8	0	0.2	0.8
No.13 HVAC	0.8	0	0.8	0.2	0.8
No.14 HVAC	0.8	0	0	0.2	0.8
Time (min)	1-3	4-6	7-9	10-12	13-15

Under these time-varying attacks in Table II, the performance of the DRS is shown in Fig. 8. Fig. 8 (a) shows that each HVAC's power deviations under time-varying attack vectors. We find that no matter which attack vector is launched, the power states of all HVACs are always deviated and cannot converge to the reference signal congruously. For the specific analysis with Table II and Fig. 8 (a), comparing time periods I and II, when the injected data is the same, the more HVACs are attacked, the larger power state deviations (i.e., attack result) will be. Comparing time periods II and III, when both the injected data and the number of attacked HVACs are the same, choosing to attack the different HVAC, the power state deviations are still different. In particular, it can be found that the DR completion rate during time periods III is lower than that during time periods II from Fig. 8 (c). This is because No.13 HVAC is a key one that can receive the aggregator's reference signal. Comparing time periods I and IV, when attacking the same HVAC, the degree of power deviation depends on the injected data. A larger injected data results in a more severe power deviation, and vice versa. Comparing time periods I and V, the steady-state values are the same when both the attacked HVACs and the injected data are the same, which verifies the effectiveness of the test results. In fact, the

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2023.3287205

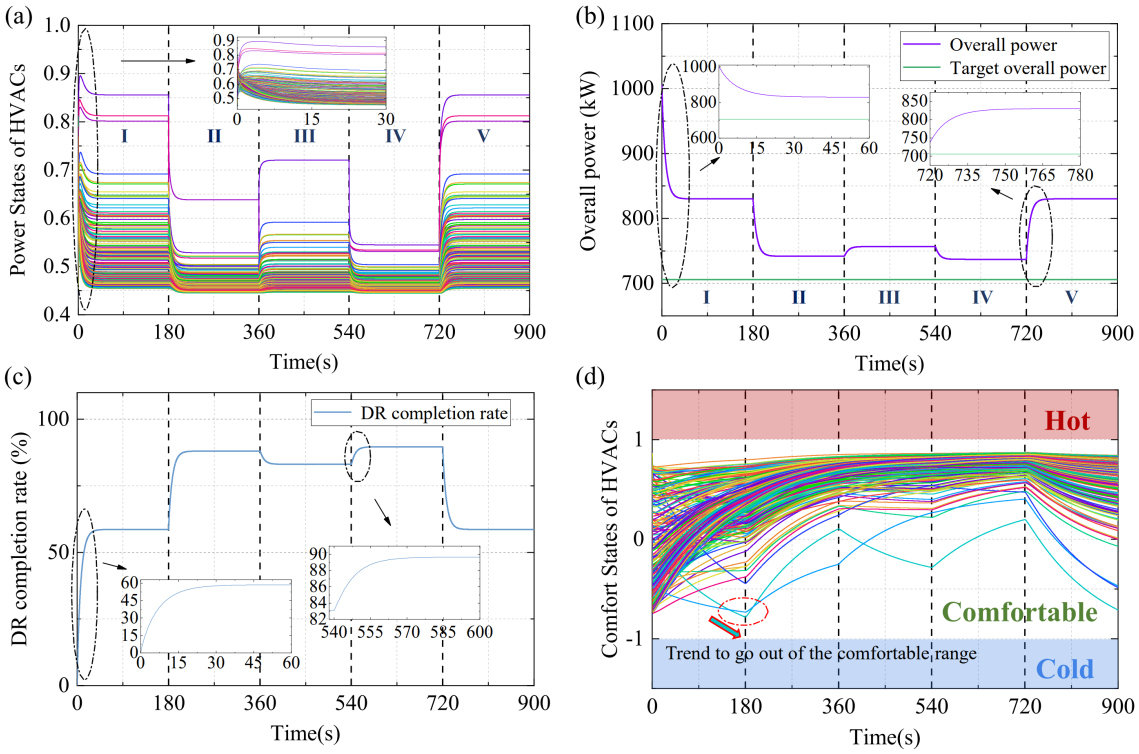

Fig. 8. The performance of HVAC-based DRS with series-of-actions dynamic cyber-attacks: (a) Power states of HVACs; (b) Overall power of the DRS; (c) DR completion rate; (d) Comfort states of HVACs.

results for all the attack vectors mentioned above satisfy the theoretical derivations in Section III.

In addition, as shown in Fig. 8 (b), under any attack vector, the overall power is deviated and cannot converge to the target value (706 kW), which makes it difficult to meet the grid's operating reserve requirement. The deviation of the overall power also differs for varying attack vectors. In particular, deviation becomes larger as the value of the injected data or the number of attacked HVACs increases.

Moreover, when all three HVACs are attacked, the DRS's DR completion rate can only achieve 58.63% of the target. This means that the FDI attack can have a significant impact on the DRS. The DR completion rate of DRS also becomes different when the attack vector is changed, as shown in Fig. 8 (c). This indicates that the attacker can control the impact of the attack by adjusting the attack vector.

The comfort states are shown in Fig. 8 (d). During the time period I, the indoor temperature trend can be out of the comfortable range. However, as shown in time period II, by adjusting the attack vector and reducing the HVACs' power deviation, the indoor temperature is avoided from being too cold, which increases the stealthiness of the attack in a sense (i.e., customers cannot perceive the problem directly). This indicates that attacks can impact the regulation results within the customer's comfortable indoor temperature.

It is worth noting that the test results under time-varying multiple attacks are also consistent with the theoretical analysis in Section III. Therefore, the effectiveness of quantitative analyses is validated.

E. Scenario 3: Result Analysis of the DRS with Series-of-actions Dynamic Cyber-attacks and with the Proposed Resilient Distributed Controller

To go against FDI attacks, a resilient distributed controller is proposed to protect the DRS to provide operating reserves for the power grid. The convergence of this resilient distributed controller is proved mathematically in Section IV. To further test the effectiveness, this case explores the performance of DRS under the different FDI attacks in Table II by using the resilient distributed controller. The results are shown in Fig. 9.

It is shown in Fig. 9 (a) that under different FDI attacks, the steady-state values of each HVAC's power state can converge to the aggregator's signal congruously. This indicates that the adverse effect caused by FDI attacks is eliminated, which conforms with the theoretical analysis in Section IV.

Furthermore, with the help of the resilient distributed controller, the DRS's overall power can be reduced and converged to the target value (i.e., from 1006 kW to 706 kW). When new FDI attacks are launched, the overall power can immediately converge to the target again after a minor adjustment, as shown in Fig. 9 (b). This shows that with the proposed controller, the DRS can still regulate the overall power according to the aggregator's signal even under FDI attacks.

Moreover, as shown in Fig. 9 (c), the DR completion rate can reach 100% quickly by using the proposed controller. When new attacks are launched, the DR completion rate can also immediately return to 100% completion again after a minor perturbation. This shows that with the proposed controller, the DRS is still able to fulfill the grid's operating reserve requirements even under FDI attacks.

According to Fig. 9 (d), by using the proposed controller, the comfort state of each room can be within the comfortable

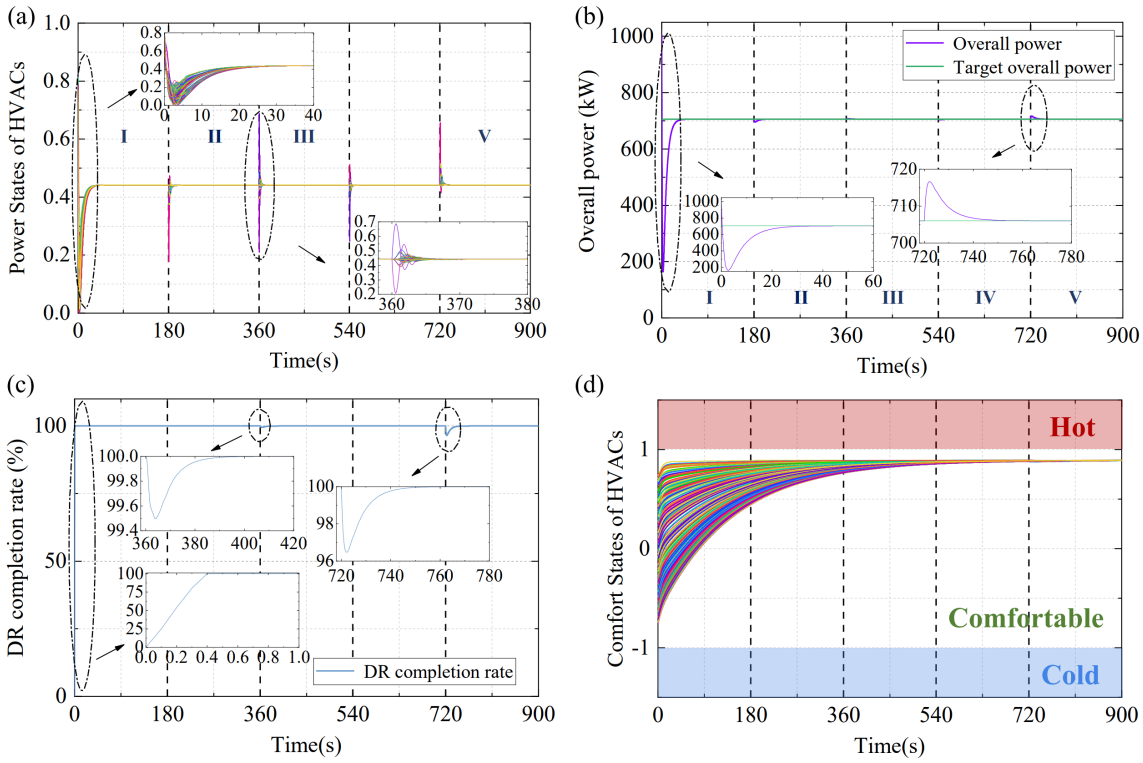

Fig. 9. The performance of HVAC-based DRS with series-of-actions dynamic cyber-attacks and with the proposed resilient distributed controller: (a) Power states of HVACs; (b) Overall power of the DRS; (c) DR completion rate; (d) Comfort states of HVACs.

TABLE III
RESULT COMPARISON OF HVAC-BASED DRS IN SCENARIOS WITH CYBER-ATTACKS

Scenarios	Con_α	Con_β	Rec	$\Delta P_{\max}$(kW)	$CR_{\min}$
S-1	No	No	No	51	83.10%
S-2	No	No	No	124	58.63%
S-3	Yes	Yes	Yes	0	100.00%

range and can converge congruously even under different attacks. This indicates that the FDI attack's adverse effect on the indoor temperature can be avoided, and thus both customer comfort and inter-customer fairness are ensured.

Compared with the severe deviations in four perspectives shown in Fig. 8, the DRS with the proposed resilient distributed controller still works well even under FDI attacks. The aggregator's control signals can be tracked and the grid's operating reserve requirements can be met by the HVAC-based DRS. Moreover, the fair sharing of both power and comfort states can be ensured. The test results show that adverse effects of FDI attacks can be avoided by using the proposed controller.

F. Comparative Analyses for Scenarios with Cyber-attacks

To improve readability, a table is illustrated to show the results in all three scenarios with cyber-attacks, especially including whether the consensus control can be achieved, whether DR can recover from the cyber-attacks, maximum overall power deviation, and minimum DR completion rate, illustrated in Table III.

In Table III, Con_α denotes whether HVACs' power states can be controlled to the aggregator's reference signal; Con_β denotes whether the consensus control of HVACs' comfort states can be achieved; Rec indicates whether the DRS can recover from the cyber-attacks; $\Delta P_{\max}$ denotes the DRS's maximum power deviation from the target overall power at steady state; $CR_{\min}$ denotes the minimum DR completion rate at steady state due to cyber-attacks; 'Yes' means the corresponding condition variable can be achieved, and 'No' means the condition variable cannot be achieved.

As shown in Table III (S-1), the consensus control of HVACs' power states and comfort states cannot be achieved under a single-action cyber-attack. In addition, the DRS cannot recover from the cyber-attack. The maximum power deviation in this scenario is about 51 kW, and the minimum DR completion rate is about 83.10%. This means the HVACs' power states cannot be controlled to the aggregator's reference signal, and the DRS cannot meet the operating reserve requirement of power systems due to the cyber-attack.

From Table III (S-2), under series-of-actions cyber-attacks, the consensus control of HVACs' power states and comfort states remains unachievable. The DRS still cannot recover from cyber-attacks. Moreover, the maximum power deviation reaches about 124 kW, and the minimum DR completion rate is only about 58.63% in this scenario. This means the performance of DRS degraded severely caused of the series-of-actions attacks.

As shown in Table III (S-3), with the proposed resilient distributed controller, the consensus control of HVACs' power states and comfort states can be achieved even under series-of-actions attacks. In addition, the DRS can recover from cyber-attacks by using the proposed resilient distributed controller. As a result, there is no power deviation in this scenario, and

the DR completion rate can achieve 100%. This means adverse effects caused by cyber-attacks can be eliminated effectively, and the operating reserve requirement of power systems can still be met by using the proposed controller.

VI. Conclusion

DR is necessary for the power balance of grids due to the ability to deal with power fluctuations. In this paper, we consider an HVAC-based DRS to provide operating reserves to the grid. Moreover, we quantify the adverse impact caused by an arbitrary FDI attack on the DRS theoretically. Furthermore, we propose a novel resilient distributed controller for the DRS against FDI attacks. Test results indicate that the power output of DRS can be disturbed by FDI attacks and the degree of impact can be controlled by attackers, which are consistent with the theoretical analyses. In addition, the case studies show that by using the proposed resilient distributed controller, the completion rate of DR task can be recovered from 58.63% to 100% even under FDI attacks, which verifies the proposed controller's effectiveness.

Appendix A: Basic Graph Theory

The graph theory is fundamental to distributed control. In this HVAC-based DRS problem, a communication topology of HVACs $\mathcal{I} = \{1 \leq i \leq \mathrm{N} \mid i \in \mathbb{Z}\}$ can be described by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_1, v_2, \ldots, v_N\}$ is the set of vertices and $\mathcal{E}$ is the set of edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. Each vertex v_i is associated with an HVAC i. Each edge $(v_i, v_j) \in \mathcal{E}$ represents a communication link for information exchange from vertex v_i to vertex v_j (vertex v_i and vertex v_j are neighbors). The set of all the neighbors of vertex v_i is defined as $\mathcal{N}_i = \{v_j \mid (v_i, v_j) \in \mathcal{E}\}$. The adjacency matrix $\mathcal{A} = [a_{ij}]$ of a graph $\mathcal{G}$ can be defined with N dimensions as follows:

$$a_{ij} = \begin{cases} 1, \ \forall (v_i, v_j) \in \mathcal{E}, \\ 0, \ \text{otherwise}. \end{cases} \tag{28}$$

The Laplacian matrix $\mathcal{L}$ of a graph $\mathcal{G}$ is defined by:

$$\mathcal{L} = \mathcal{D} - \mathcal{A} \ , \tag{29}$$

where $\mathcal{D}$ is the in-degree matrix defined as $\mathcal{D} = \mathrm{diag}\{\boldsymbol{d}\} \subseteq \mathbb{R}^{N \times N}$ with $\boldsymbol{d} = [d_1, d_2, \ldots, d_N]^T$ and $d_i = \sum_{j \in \mathcal{N}_i} a_{ij}$.

Appendix B: Proof of Theorems

A. Proof of Theorem 1 from the Time Domain Perspective

Since the Laplacian matrix $\mathcal{L}$ has a zero eigenvalue (i.e., $\lambda_0 = 0$), and the corresponding eigenvector is $\mathbf{1}_N$, the following equation can be derived as follows:

$$\mathcal{L}\mathbf{1}_N = \lambda_0 \times \mathcal{L} = 0. \tag{30}$$

Combining (9), (11) and the property in (30), the dynamic behaviors of the power state vector under the attack can be reformulated as follows:

$$\begin{aligned} \dot{\boldsymbol{\alpha}} &= \boldsymbol{u}_\xi = \boldsymbol{u} + \boldsymbol{\Xi} \\ &= -k_\alpha(\mathcal{L} + \mathcal{B})\boldsymbol{\alpha} + k_\alpha \alpha_{ref} \mathcal{B}\mathbf{1}_N + \boldsymbol{\Xi} \\ &= -k_\alpha[(\mathcal{L} + \mathcal{B})\boldsymbol{\alpha} - \alpha_{ref}\mathcal{L}\mathbf{1}_N - \alpha_{ref}\mathcal{B}\mathbf{1}_N] + \boldsymbol{\Xi} \\ &= -k_\alpha[(\mathcal{L} + \mathcal{B})\boldsymbol{\alpha} - \alpha_{ref}(\mathcal{L} + \mathcal{B})\mathbf{1}_N] + \boldsymbol{\Xi} \\ &= -k_\alpha(\mathcal{L} + \mathcal{B})(\boldsymbol{\alpha} - \alpha_{ref}\mathbf{1}_N) + \boldsymbol{\Xi}. \end{aligned} \tag{31}$$

When the power state $\boldsymbol{\alpha}$ reaches a steady-state, the derivative of $\boldsymbol{\alpha}$ (i.e., $\dot{\boldsymbol{\alpha}}$) is equal to zero. Therefore, for calculating steady-state values, this dynamic can directly be equated to zero to obtain the result as follows:

$$\begin{aligned} \boldsymbol{\alpha} &= \alpha_{ref}\mathbf{1}_N + [k_\alpha(\mathcal{L} + \mathcal{B})]^{-1}\boldsymbol{\Xi} \\ &= \alpha_{ref}\mathbf{1}_N + \boldsymbol{\epsilon}, \end{aligned} \tag{32}$$

where $\boldsymbol{\epsilon} = [k_\alpha (\mathcal{L} + \mathcal{B})]^{-1}\boldsymbol{\Xi}$ is the error vector. In this case, HVAC's power states have steady-state errors, and the steady-state value of power state converges to $\alpha_{ref}\mathbf{1}_N + \boldsymbol{\epsilon}$.

The proof of Theorem 1 is complete. ■

B. Proof of Theorem 1 from the Frequency Domain Perspective

By performing the Laplace transform on (31), the dynamics of the power state vector can be transferred from the time domain to the complex frequency domain as follows:

$$\begin{aligned} &s\boldsymbol{\alpha}(s) - \boldsymbol{\alpha}(0) \\ &= -k_\alpha(\mathcal{L} + \mathcal{B})\boldsymbol{\alpha}(s) + \frac{k_\alpha}{s}\alpha_{ref}(\mathcal{L} + \mathcal{B})\mathbf{1}_N + \frac{\boldsymbol{\Xi}}{s}. \end{aligned} \tag{33}$$

where s is the Laplace operator; the $\boldsymbol{\alpha}(s)$ is the form of $\boldsymbol{\alpha}(t)$ in the complex frequency domain; the $\boldsymbol{\alpha}(0)$ is the initial value vector of $\boldsymbol{\alpha}(t)$. Since $[\mathcal{L}+\mathcal{B}]$ is a diagonally dominant matrix, $[s\boldsymbol{I}_N + k_\alpha(\mathcal{L} + \mathcal{B})]$ is also a diagonally dominant matrix, and then it is non-singular. Therefore, the power state vector in the complex frequency domain can be reformulated as follows:

$$\begin{aligned} \boldsymbol{\alpha}(s) = &[s\boldsymbol{I}_N + k_\alpha(\mathcal{L} + \mathcal{B})]^{-1} \\ &\times \left[\boldsymbol{\alpha}(0) + \frac{k_\alpha \alpha_{ref}}{s}(\mathcal{L} + \mathcal{B})\mathbf{1}_N + \frac{\boldsymbol{\Xi}}{s}\right]. \end{aligned} \tag{34}$$

By applying the final value theorem (FVT) to (34), the steady-state value of HVAC's power state is driven as follows:

$$\begin{aligned} &\lim_{t \to \infty} \boldsymbol{\alpha}(t) = \lim_{s \to 0} s\boldsymbol{\alpha}(s) \\ &= \lim_{s \to 0} [s\boldsymbol{I}_N + k_\alpha(\mathcal{L} + \mathcal{B})]^{-1} [k_\alpha \alpha_{ref}(\mathcal{L} + \mathcal{B})\mathbf{1}_N \\ &+ s\boldsymbol{\alpha}(0) + \boldsymbol{\Xi}] \\ &= \alpha_{ref}\mathbf{1}_N + [k_\alpha (\mathcal{L} + \mathcal{B})]^{-1}\boldsymbol{\Xi} \\ &= \alpha_{ref}\mathbf{1}_N + \boldsymbol{\epsilon}. \end{aligned} \tag{35}$$

Therefore, the error vector is $\boldsymbol{\epsilon} = [k_\alpha (\mathcal{L} + \mathcal{B})]^{-1}\boldsymbol{\Xi}$ and the steady-state value of power state converges to $\boldsymbol{\alpha} = \alpha_{ref}\mathbf{1}_N + \boldsymbol{\epsilon}$ under the attack.

The proof of Theorem 1 is complete. ■

C. Proof of Theorem 2 from the Time Domain Perspective

With the proposed controller in (17), the dynamic behaviors of the power state under an arbitrary attack vector $\boldsymbol{\Xi}$ can be described as follows:

$$\begin{aligned} \dot{\boldsymbol{\alpha}} &= \boldsymbol{u}_\xi = \boldsymbol{u} + \boldsymbol{\Xi} \\ &= -k_\alpha \int [(\mathcal{L} + \mathcal{B})(\boldsymbol{\alpha} - \alpha_{ref}\mathbf{1}_N)]\mathrm{d}t - \boldsymbol{\alpha} + \boldsymbol{\Xi}. \end{aligned} \tag{36}$$

The difference between the power state and the reference signal about power state is defined as the tracking error $\boldsymbol{\delta}$, which can be shown as below:

$$\boldsymbol{\delta} = \boldsymbol{\alpha} - \alpha_{ref}\mathbf{1}_N. \tag{37}$$

Then, combined with the definition of the tracking error $\boldsymbol{\delta}$, the dynamics in (36) can be reformulated as follows:

$$\dot{\boldsymbol{\delta}} = -k_\alpha \int [(\mathcal{L}+\mathcal{B})\boldsymbol{\delta}]\mathrm{d}t - \boldsymbol{\delta} - \alpha_{ref}\mathbf{1}_N + \boldsymbol{\Xi}. \tag{38}$$

When the power state $\boldsymbol{\alpha}$ reaches a steady state, the tracking error $\boldsymbol{\delta}$ can also reach a steady state, and the derivative of $\boldsymbol{\delta}$ (i.e., $\dot{\boldsymbol{\delta}}$) is equal to zero. Therefore, for calculating steady-state values, this dynamic described in (38) can directly be equated to zero to obtain the result, as follows:

$$k_\alpha \int [(\mathcal{L}+\mathcal{B})\boldsymbol{\delta}]\mathrm{d}t + \boldsymbol{\delta} + \alpha_{ref}\mathbf{1}_N - \boldsymbol{\Xi} = 0. \tag{39}$$

The derivative of dynamics in (39) yields the following equation:

$$\dot{\boldsymbol{\delta}} + k_\alpha(\mathcal{L}+\mathcal{B})\boldsymbol{\delta} = 0. \tag{40}$$

Since k_α is a positive constant and the matrix $\mathcal{L}+\mathcal{B}$ is a positive definite matrix, the tracking error $\boldsymbol{\delta}$ in (40) can converge to zero ($\boldsymbol{\delta} = \boldsymbol{\alpha} - \alpha_{ref}\mathbf{1}_N = 0$) [36], which implies the steady-state value of power state converges to the reference signal, which can be shown as below:

$$\begin{aligned} \lim_{t\to\infty} \boldsymbol{\delta}(t) = 0 \Rightarrow \quad & \lim_{t\to\infty} [\boldsymbol{\alpha}(t) - \alpha_{ref}\mathbf{1}_N] = 0 \\ \Rightarrow \quad & \lim_{t\to\infty} \boldsymbol{\alpha}(t) = \alpha_{ref}\mathbf{1}_N. \end{aligned} \tag{41}$$

According to (41), it is shown that the power state's steady-state value of all the HVAC can converge to the reference value, i.e., α_{ref}, from the power system operator even under the attack.

The proof of Theorem 2 is complete. ■

D. Proof of Theorem 2 from the Frequency Domain Perspective

By the Laplace transform of (36), the dynamics of the power state in the complex frequency domain can be shown as follows:

$$\begin{aligned} s\boldsymbol{\alpha}(s) - \boldsymbol{\alpha}(0) = & -k_\alpha \left[\frac{1}{s}(\mathcal{L}+\mathcal{B})\boldsymbol{\alpha}(s) + \frac{1}{s}(\mathcal{L}+\mathcal{B})\boldsymbol{\alpha}^{-1}(0)\right] \\ & + \frac{k_\alpha}{s^2}\alpha_{ref}(\mathcal{L}+\mathcal{B})\mathbf{1}_N - \boldsymbol{\alpha}(s) + \frac{\boldsymbol{\Xi}}{s}. \end{aligned} \tag{42}$$

where $\boldsymbol{\alpha}^{-1}(0) = \left[\int \boldsymbol{\alpha}(t)\mathrm{d}t\right]_{t=0}$. Since $[\mathcal{L}+\mathcal{B}]$ is non-singular, the $[(s^2+s)\boldsymbol{I}_N + k_\alpha(\mathcal{L}+\mathcal{B})]$ is also non-singular. On this basis, to obtain the $\boldsymbol{\alpha}(s)$, the power state vector in (42) can be reformulated by shifting the terms as follows:

$$\begin{aligned} \boldsymbol{\alpha}(s) = & \left[s^2 + s + k_\alpha(\mathcal{L}+\mathcal{B})\right]^{-1}[-k_\alpha(\mathcal{L}+\mathcal{B})\boldsymbol{\alpha}^{-1}(0) \\ & + \frac{k_\alpha}{s}\alpha_{ref}(\mathcal{L}+\mathcal{B})\mathbf{1}_N - s\boldsymbol{\alpha}(s) + \boldsymbol{\Xi}]. \end{aligned} \tag{43}$$

According to the FVT, the steady-state value of HVAC's power state is driven as follows:

$$\begin{aligned} \lim_{t\to\infty} \boldsymbol{\alpha}(t) &= \lim_{s\to 0} s\boldsymbol{\alpha}(s) \\ &= \lim_{s\to 0} \left[s^2 + s + k_\alpha(\mathcal{L}+\mathcal{B})\right]^{-1}[-sk_\alpha(\mathcal{L}+\mathcal{B})\boldsymbol{\alpha}^{-1}(0) \\ &+ k_\alpha\alpha_{ref}(\mathcal{L}+\mathcal{B})\mathbf{1}_N - s^2\boldsymbol{\alpha}(s) + s\boldsymbol{\Xi}] \\ &= [k_\alpha(\mathcal{L}+\mathcal{B})]^{-1}[k_\alpha\alpha_{ref}(\mathcal{L}+\mathcal{B})\mathbf{1}_N] \\ &= \alpha_{ref}\mathbf{1}_N. \end{aligned} \tag{44}$$

The result from (44) shows that there is no error caused by the FDI attack. Moreover, for all the HVACs $i \in \mathcal{I}$, the power state's steady-state value converges to the reference value from the aggregator, i.e., α_{ref}.

The proof of Theorem 2 is complete. ■

APPENDIX C: DISCUSSIONS

A. Discussion on Transients

The cyber-attacks are the direct reason for transients. The processes of power deviations caused by cyber-attacks and rapid recovery due to the attack-resilient controller are presented as transients. In addition, transients are indeed also related to the control gain k_α. When the selected control gain is large, the power states of HVACs are less affected by cyber-attacks, resulting in smaller transients, and vice versa. To demonstrate transients under different control gains, additional cases have also been supplemented, which can be shown in Figure 10.

Figure 10 shows the power states of HVACs under different control gains (i.e., control gain k_α = 2, 1, 0.5, and 0.1, respectively). In Case 1, the control gain is equal to 2, and the deviations caused by cyber-attacks can be recovered quickly. Therefore, the power states of HVACs are less affected by cyber-attacks, and for example, the maximum deviation at 360s is about 0.188. As a result, the transients are smaller. In Case 2, the control gain is equal to 1. In this case, the maximum deviation at 360s is about 0.246. In Case 3, the control gain is equal to 0.5, and the deviations caused by cyber-attacks are recovered slowly. Therefore, the power states of HVACs are more affected by cyber-attacks, and for example, the maximum deviation at 360s is about 0.315. As a result, the transients become larger. As shown in Figure 10 (d), the control gain is equal to 0.1, which is the minimum value for these cases. In this case 4, the transients are large enough to cause the control saturation effects. Through the saturation function module, the range of HVAC's power state is constrained to the interval from 0 to 1 (i.e., $\alpha \in [0,1]$). In reality, the power of HVAC is generally below the rated power and above zero, i.e., below 1 and above 0. Therefore, this saturation function module makes practical physical sense and can be consistent with reality.

B. Discussion on Withdrawal Mechanism

A case is given to show that when the customer enters the hot range, the withdrawal mechanism can satisfy the customer's comfort requirements, as below.

As shown in Figure 11 (Case 1), with an appropriate evaluated regulation capacity, the maximum fluctuation of the corresponding comfort state index can be maintained at less than 1, which means the customer's comfort requirement can be satisfied. In an extreme case, the evaluated regulation capacity may become inaccurate. With the regulation command with inaccurate evaluation, a certain room's comfort state

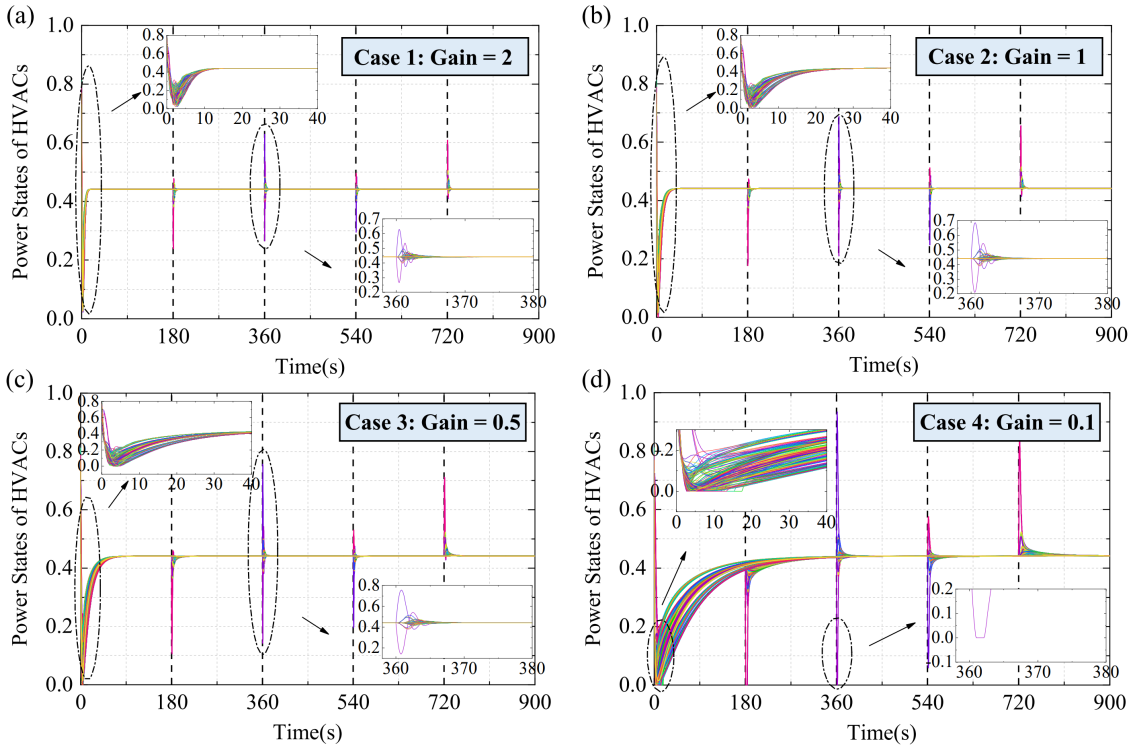

Fig. 10. Power states of HVACs with different control gains: (a) gain = 2; (b) gain = 1; (c) gain = 0.5; (d) gain = 0.1.

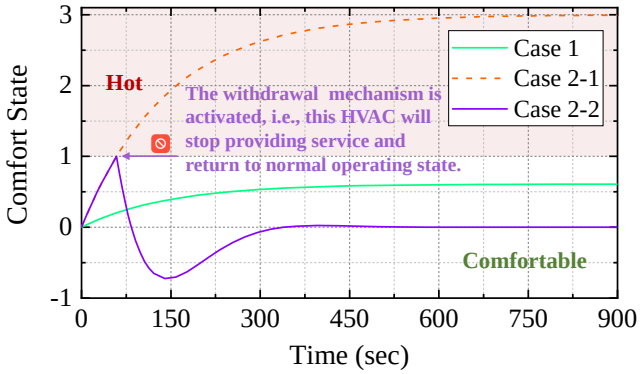

Fig. 11. Fluctuations of the room's comfort state: (i) Case 1: with an appropriate regulation capacity evaluation; (ii) Case 2-1: with an inaccurate regulation capacity evaluation but without the withdrawal mechanism; (iii) Case 2-2: with an inaccurate regulation capacity evaluation and with the activated withdrawal mechanism.

may be out of the hot tolerance limit during the process of providing service to power systems (Case 2-1). To ensure the customer's comfort, the withdrawal mechanism is activated, i.e., the corresponding HVAC will stop providing service and return to the normal operating state, as shown in Figure 11 (Case 2-2). However, there would be few HVACs whose comfort states reach the upper bound during the process of providing service, since the regulation capacity can be evaluated accurately in general. For an aggregation of HVACs, the rest of the HVACs will still provide significant operating reserves for power systems, which means the service quality can still be maintained.

C. Discussion on HVAC Receiving the Reference Signal

To make the distributed control work, the communication topology should be constructed as a spanning tree[2], and at least one HVAC should receive the signal from the aggregator. Therefore, the number of HVAC (i.e., n_o) that should communicate with and receive signals from the aggregator should satisfy $1 \leq n_o \leq N$, where N is the total number of HVACs considered.

The performance of DRS can be affected by the change in HVACs that receive the reference signal. In general, the distributed control converges faster and the DRS responds faster when there are more HVACs receiving the reference signal, and vice versa. In addition, when the number of HVACs receiving the reference signal is unchanged, the distributed control converges faster and the DRS responds faster when the HVACs receiving the reference signal are located at the middle position of the communication topology, and vice versa.

Additional case studies are given to verify how the performance is affected by the change in the HVACs that receive the reference signal. There are 3 cases, namely (i) case 1: benchmark case; (ii) case 2: number change, i.e., reduction in the number of HVACs receiving the reference signal; (iii) case 3: position change, i.e., moving the position of HVACs selected to receive the reference signal from the middle position to the end position. In these case studies, the performance of HVACs' power states, DRS's overall power, and DRS's DR completion rate are simulated under different HVACs receiving the reference signal, which are illustrated in Figure 12, Figure 13 (a), and Figure 13 (b), respectively.

Figure 12 shows power states of HVACs under different HVACs receiving the reference signal (3 cases mentioned above). In case 1 (benchmark case), the power states of HVACs can converge to the reference signal at about 38s. However, in case 2 (number change) and case 3 (position change), the convergence time is at about 171s and 116s, respectively. The

[2]A spanning tree is a tree that connects all of the vertices of the graph.

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2023.3287205

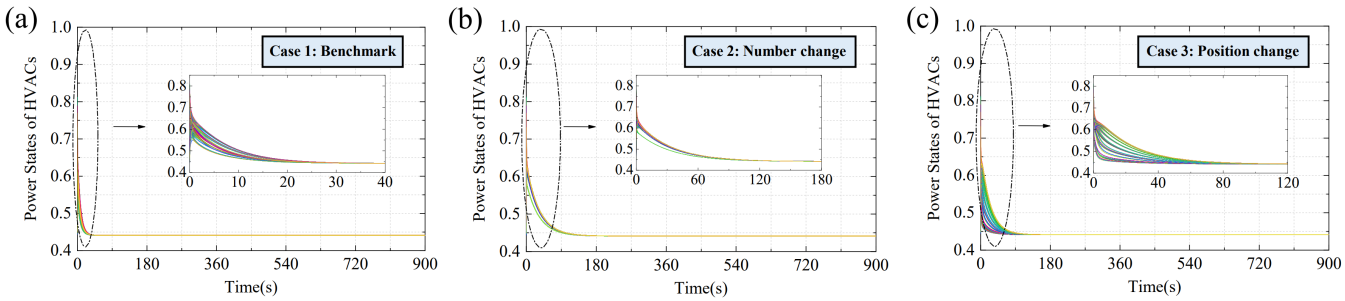

Fig. 12. The performance of HVACs' power states in different cases: (a) Case 1: Benchmark; (b) Case 2: Number change; (3) Case 3: Position change.

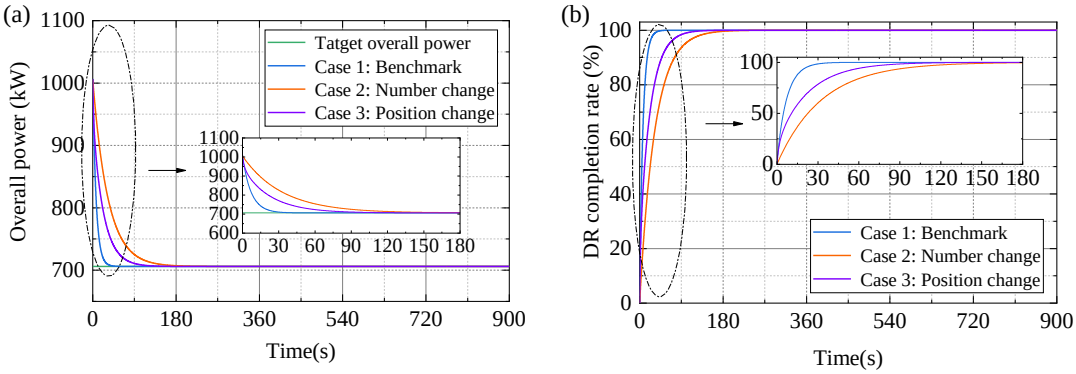

Fig. 13. The performance of DRS with different HVAC receiving the reference signal: (a) Overall power of the DRS; (b) DR completion rate of the DRS.

results imply that the convergence speed of distributed control decreases obviously, when the number of HVACs receiving the reference signal is reduced, or when the position of HVACs selected to receive the reference signal is moved from the middle position to the end position.

Figure 13 shows the performance of DRS in different cases, especially including the overall power of DRS and the DR completion rate of DRS. From Figure 13 (a), the overall power of DRS can be regulated to the target overall power required by the power system operator. This response time of DRS in case 1 (benchmark case) is about 38s, while in case 2 (number change) and case 3 (position change), the response time is about 171s and 116s, respectively. The results imply that the response time of DRS becomes longer obviously, when the number of HVACs receiving the reference signal is reduced, or when the position of HVACs selected to receive the reference signal is moved from the middle position to the end position. How the topology and number of HVACs receiving the reference signal affect the performance deserves further investigation. The response time in different cases is also illustrated in Figure 13 (b). In addition, the DR completion rate can achieve 100% in different cases, which implies the response time of DRS can be affected by the change in HVACs receiving the reference signal, and yet DRS can still meet the power system's requirements for operating reserves.

REFERENCES

[1] International Recommended Energy Agency, "Renewable capacity statistics 2021," tech. rep., 2021. [Online]. Available: https://www.irena.org/publications/2021/March/Renewable-Capacity-Statistics-2021.

[2] P. Wang, Z. Zhang, T. Ma, Q. Huang, and W.-J. Lee, "Parameter calibration of wind farm with error tracing technique and correlated parameter identification," *IEEE Trans. Power Syst.*, Early Access, 2022.

[3] S. Yang, K.-W. Lao, H. Hui, Y. Chen, and N. Dai, "Real-time harmonic contribution evaluation considering multiple dynamic customers," *CSEE J. Power Energy Syst.*, Early Access, 2023.

[4] S. Wang, H. Hui, Y. Ding, C. Ye, and M. Zheng, "Operational reliability evaluation of urban multi-energy systems with equivalent energy storage," *IEEE Trans. Ind. Appl.*, vol. 59, no. 2, pp. 2186–2201, 2023.

[5] Operations Planning Division of PJM, "PJM Manual 10: Pre-scheduling Operations," tech. rep., 2022. [Online]. Available: https://www.pjm.com/-/media/documents/manuals/m10.ashx.

[6] X. Liu, Y. Li, X. Lin, J. Guo, Y. Shi, and Y. Shen, "Dynamic bidding strategy for a demand response aggregator in the frequency regulation market," *Appl. Energy*, vol. 314, p. 118998, 2022.

[7] S. Wang, J. Zhai, H. Hui, Y. Ding, and Y. Song, "Operational reliability of integrated energy systems considering gas flow dynamics and demand-side flexibilities," *IEEE Trans. Ind. Informat.*, Early Access, 2023.

[8] P. Siano, "Demand response and smart grids—a survey," *Renew. Sust. Energ. Rev.*, vol. 30, pp. 461–478, Feb. 2014.

[9] H. Hui, P. Siano, Y. Ding, P. Yu, Y. Song, H. Zhang, and N.-Y. Dai, "A transactive energy framework for inverter-based HVAC loads in a real-time local electricity market considering distributed energy resources," *IEEE Trans. Ind. Informat.*, Early Access, 2022.

[10] P. Siano and D. Sarno, "Assessing the benefits of residential demand response in a real time distribution energy market," *Appl. Energy*, vol. 161, no. 1, pp. 533–551, Jan. 2016.

[11] N. Saxena, A. Roy, and H. Kim, "Efficient 5G small cell planning with embms for optimal demand response in smart grids," *IEEE Trans. on Ind. Informat.*, vol. 13, no. 3, pp. 1471–1481, 2017.

[12] Q. Shi, F. Li, G. Liu, D. Shi, Z. Yi, and Z. Wang, "Thermostatic load control for system frequency regulation considering daily demand profile and progressive recovery," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6259–6270, Feb. 2019.

[13] H. Hui, Y. Ding, Z. Lin, P. Siano, and Y. Song, "Capacity allocation and optimal control of inverter air conditioners considering area control error in multi-area power systems," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 332–345, Jun. 2019.

[14] Y. Ding, S. Pineda, P. Nyeng, J. Østergaard, E. M. Larsen, and Q. Wu, "Real-time market concept architecture for EcoGrid EU—a prototype

for european smart grids," *IEEE Trans. Smart Grid*, vol. 4, pp. 2006–2016, May. 2013.

[15] Avangrid Corporation, "UI demand response programs," tech. rep., 2016. [Online]. Available: https://portal.ct.gov/-/media/DEEP/energy/CES/UnitedIlluminatingDemandResourcesPresentation102716pdf.pdf.

[16] H. Hui, Y. Ding, K. Luan, T. Chen, Y. Song, and S. Rahman, "Coupon-based demand response for consumers facing flat-rate retail pricing," *CSEE J. Power Energy Syst.*, Early Access, 2022.

[17] F. Chen, M. Chen, Q. Li, K. Meng, J. M. Guerrero, and D. Abbott, "Multiagent-based reactive power sharing and control model for islanded microgrids," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 1232–1244, Jul. 2016.

[18] J. Su, H. Zhang, H. Liu, L. Yu, and Z. Tan, "Membership-function-based secondary frequency regulation for distributed energy resources in islanded microgrids with communication delay compensation," *IEEE Trans. on Sustain. Energy*, Early Access, 2023.

[19] Y. Chen, D. Qi, H. Hui, S. Yang, Y. Gu, Y. Yan, Y. Zheng, and J. Zhang, "Self-triggered coordination of distributed renewable generators for frequency restoration in islanded microgrids: A low communication and computation strategy," *Adv. Appl. Energy*, vol. 10, p. 100128, 2023.

[20] H. Hui, Y. Chen, S. Yang, H. Zhang, and T. Jiang, "Coordination control of distributed generators and load resources for frequency restoration in isolated urban microgrids," *Appl. Energy*, vol. 327, p. 120116, 2022.

[21] Y. Ding, D. Xie, H. Hui, Y. Xu, and P. Siano, "Game-theoretic demand side management of thermostatically controlled loads for smoothing tie-line power of microgrids," *IEEE Trans. Power Syst.*, vol. 36, pp. 4089–4101, Sep. 2021.

[22] J. Hong, H. Hui, H. Zhang, N. Dai, and Y. Song, "Event-triggered consensus control of large-scale inverter air conditioners for demand response," *IEEE Trans. Power Syst.*, vol. 37, no. 6, pp. 4954–4957, 2022.

[23] Y. Wang, Y. Tang, Y. Xu, and Y. Xu, "A distributed control scheme of thermostatically controlled loads for the building-microgrid community," *IEEE Trans. Sustain. Energy*, vol. 11, no. 1, pp. 350–360, Jan. 2020.

[24] X. Zhang, M. Pipattanasomporn, T. Chen, and S. Rahman, "An IoT-based thermal model learning framework for smart buildings," *IEEE Internet Things J.*, vol. 7, pp. 518–527, Jan. 2020.

[25] "Cybersecurity for Industrial Automation & Control Environments," tech. rep., 2013. [Online]. Available: https://www.se.com/ww/en/download/document/998-2095-04-13-13AR0_EN/.

[26] W. Ahn, M. Chung, B.-G. Min, and J. Seo, "Development of cyber-attack scenarios for nuclear power plants using scenario graphs," *Int. J. Distrib. Sens. Netw.*, vol. 11, no. 9, p. 836258, Sep. 2015.

[27] S. Weinberger, "Clash over Iran's capability: effects of sanctions and computer worm on uranium production are disputed," *Nature*, vol. 470, no. 7335, pp. 443–445, 2011.

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

[29] "Cyber-attack Against Ukrainian Critical Infrastructure," tech. rep., NCCIC/ICS-CERT, Jun. 2016. [Online]. Available: https://ics-cert.us-cert.gov/alerts/IR-ALERT-H16-056-01.

[30] S. Mousavian, M. Erol-Kantarci, L. Wu, and T. Ortmeyer, "A risk-based optimization model for electric vehicle infrastructure response to cyber attacks," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6160–6169, Nov. 2018.

[31] S. Acharya, Y. Dvorkin, and R. Karri, "Causative cyberattacks on online learning-based automated demand response systems," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3548–3559, Jul. 2021.

[32] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1630–1638, Jul. 2017.

[33] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inf. Syst. Security*, vol. 14, no. 1, pp. 1–33, May. 2011.

[34] Y. Chen, D. Qi, H. Dong, C. Li, Z. Li, and J. Zhang, "A FDI attack-resilient distributed secondary control strategy for islanded microgrids," *IEEE Trans. Smart Grid*, vol. 12, pp. 1929–1938, May. 2021.

[35] G. F. Franklin, J. D. Powell, A. Emami-Naeini, and J. D. Powell, *Feedback control of dynamic systems*, vol. 4. Prentice hall Upper Saddle River, 2002.

[36] Y. Chen, K.-W. Lao, D. Qi, H. Hui, S. Yang, Y. Yan, and Y. Zheng, "Distributed self-triggered control for frequency restoration and active power sharing in islanded microgrids," *IEEE Trans. Ind. Informat.*, Early Access, 2023.

[37] Gree Official Website, "Facts about GREE," tech. rep., 2023. [Online]. Available: https://global.gree.com/.

[38] "IEEE standard for information technology - telecommunications and information exchange between systems - local and metropolitan networks - specific requirements - part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications: Higher speed physical layer (phy) extension in the 2.4 GHz band - corrigendum 1," *IEEE Std 802.11b-1999/Cor 1-2001*, pp. 1–24, 2001.

[39] "Results of Meteorological Observations," tech. rep., Macao Meteorological and Geophysical Bureau, 2022. [Online]. Available: https://www.smg.gov.mo/zh/subpage/348/report/download-pdf.

[40] German Transmission System Operators, "Prequalification process for balancing service providers in germany," tech. rep., 2020. [Online]. Available: https://www.regelleistung.net/ext/download/PQ_Bedingungen_FCR_aFRR_mFRR_en.