Robust Restoration from Cyber-Physical Attacks in Active Distribution Grids with Grid-Edge IBRs

Xue Gao, *Student Member, IEEE*, and Wei Sun, *Senior Member, IEEE*

***Abstract*—The inverter-based resources (IBRs) have enabled the integration of renewable energy at the grid edge with enhanced control capabilities to support the reliable operation of power grids. Different control frameworks, such as hierarchical or distributed architecture, have been proposed with the expansion of cyber networks for real-time monitoring and control. This evolution of critical infrastructure into cyber-physical systems also brings more vulnerabilities for the broadened attack surfaces, and significantly increases the possibility of physical system failures or outages caused by cyberattacks. Among tremendous efforts in the defense-in-depth approach, it remains challenging to provide prompt detection and accurate location of attack entry points or paths. Therefore, the prevailing restoration framework may struggle to fully consider the cyber-physical interdependence, successfully isolate the compromised cyber and physical components, and safely recover the systems without the potential risks leading to secondary outages. This paper is motivated to develop a cyber-physical restoration framework for distribution grids to recover from cyber attacks by harnessing grid-edge IBRs. The framework is first built on the operational guidelines of IBRs considering the compromised cyber layer. Then, an ambiguity set is established to represent the uncertainty of attack scenarios and their possibility levels. Next, a distributionally robust optimization model is developed to provide the optimal load restoration strategy across all scenarios. The effectiveness of the proposed model is demonstrated through various use cases on the modified IEEE 13-node and 123-node test systems. Simulation results demonstrate the effectiveness and advancement of developed post-attack restoration strategies.**

***Index Terms*—Cybersecurity, Distributed energy resources, Distributionally robust optimization, Inverter-based resources, Power system resilience, Post-attack restoration**

I. Introduction

THE inverter-based resources (IBRs) have emerged as the new types of distributed energy resources (DERs) at the edge of distribution grids. These grid-edge IBRs are equipped with the capability of enhanced power quality, rapid response, and expanded functionalities, serving as the main interfaces for renewable integration. Given their inherent low inertia and the uncertain nature of renewable energy, smart sensors and controllers are necessary to facilitate the real-time monitoring and control of IBRs. While this evolving infrastructure is essential for ensuring the secure and reliable operation of distribution grids with IBRs, this expansion significantly broadens the attack surfaces of existing cyber networks. Moreover, non-utility owners of IBR might encounter challenges in establishing sufficient security policies and consequently elevating the risk of cyber threats. Furthermore, the design of the cyber network allows accessibility to third parties such as IBR owners, manufacturers, and aggregators, posing safety concerns for unsecured network connections. Therefore, the increased exposure of the cyber networks of IBRs to potential attacks emerges as a significant security concern.

Xue Gao and Wei Sun are with the Department of Electrical and Computer Engineering, University of Central Florida, Orlando, FL. Email: xue.gao@ucf.edu, and sun@ucf.edu.

In literature, [1], [2] indicated that the common cyber attacks, such as denial of services (DoS) or false data injection (FDI), targeting IBR cyber networks, might lead to severe consequences at the physical layer. These include branch overflow, posing risks of equipment damage, stability issues such as voltage/frequency violation, and even power outages. Strengthening distribution system resilience against cyber attacks involves two main strategies. One way is to identify and strengthen the vulnerable components accordingly, thereby enhancing system resistance to cyber threats. The other way is to develop post-attack restoration methods, improving the system's ability to recover from cyber attacks rapidly. This paper is motivated to focus on the latter one.

The majority of existing restoration algorithms mainly tackle the disruptions within the physical layer, and usually assume the full functionality of the cyber layer, treating restoration as an optimization problem [3]–[6]. However, the post-attack restoration necessitates accounting for the compromised cyber layer. Given the compromised components infiltrated by the malicious entity, if not identified and isolated successfully, attackers can launch subsequent attacks, continuously impacting the physical layer and potentially leading to a secondary outage. Therefore, the prevailing restoration algorithms from component failures, system malfunctions, or natural disasters are inadequate for handling post-attack restoration problems.

Furthermore, due to the intricate cyber-physical interdependence, it is challenging or usually impossible to pinpoint and identify the compromised components shortly after observing physical system failures, given very limited or implicit information of cyber-physical system status. Recent efforts have made some progress in detecting and locating potential attacks. For instance, [7] proposed a graph neural network to identify the presence and location of FDI attacks. [8] introduced a clustering method to determine the compromised phasor measurement units. [9] presented a deep-learning method to identify active attack locations. However, these approaches mainly utilize the data-based learning method, which requires massive data and powerful computational resources. Moreover, the training data are generally confined to specific devices or attack types. Consequently, when faced with incomplete attack information, these methods might struggle to identify the entry points of attacks within the cyber layer. This limitation could impede the ability of these algorithms to effectively identify and isolate the compromised components.

This paper is motivated to address these challenges, by in-

troducing a distributionally robust optimization (DRO) method to solve the post-attack load restoration problem. Instead of the deterministic identification of compromised components based on assumptions, the proposed method develops an ambiguity set, reflecting the uncertainty of the probability of compromised components, leveraging the limited data from the cyber layer and the observations from the physical layer. Subsequently, based on this ambiguity set, the DRO algorithm can provide the optimal restoration strategy, ensuring the minimum expectation of lost load in the worst cases.

The main contributions of this paper are threefold, i) Established the operational guidelines to ensure the secure operation of IBRs while considering the compromised cyber layer; ii) Built up an ambiguity set to model the uncertainty level regarding the probability of cyber components being compromised; and iii) Developed a tractable DRO model for providing the optimal restoration strategy from cyber attacks.

The remainder of the paper is structured as follows: Section II provided an overview of the post-attack restoration framework. Section III presented a deterministic model under given attack scenarios. Section IV introduced the DRO model, focusing on more practical scenarios where the attack entry points are unknown. SectionV introduced the DRO model reformulation strategy and the corresponding solution algorithm. Section VI presents simulation results and discussions on the performances of the proposed DRO model. Finally, Section VII concludes this paper.

II. Post-Attack Restoration Framework with IBRs

The grid-forming IBRs are equipped with the ability to regulate their terminal voltages and system frequency, which enables them to independently establish a stable grid and function effectively in both grid-connected and islanded scenarios, significantly enhancing the adaptability of IBR systems [10]–[13]. Ensuring the secure operation of IBRs heavily relies on a highly efficient communication network that enables real-time monitoring and control. However, the extensive presence of cyber components, such as sensors and routers, has notably broadened the attack surface of the supporting communication network. Moreover, non-utility owners, often lacking expertise in configuring secure policies, inadvertently introduce more vulnerabilities into these networks. Furthermore, due to the operational requirements of IBRs, which necessitate monitoring and support from manufacturers, aggregators, and maintainers, inadequate access policies may heighten the risk of insecure network connection. Hence, the escalating cyber threats within current IBR communication networks pose a significant risk to the secure operation of IBRs.

The real-world attack cases highlighted by [14], [15] emphasize the susceptibility of IBRs to cyber threats. These cases reveal that current cyber layer vulnerabilities provide attackers with potential intrusion points, potentially enabling them to terminate essential communications, manipulate IBR control parameters, and issue false commands to IBRs. Given the low inertia of IBRs, such interruptions have the potential to induce severe system stability issues, trigger IBR trips and cause cascade failures, and consequently lead to load shedding and blackouts. Therefore, in anticipation of potential blackouts due to cyber attacks, it is crucial to develop secure and reliable post-attack restoration strategies.

Post-attack restoration using IBRs requires coordination among multiple stakeholders [16]. First, utilities must secure approval and access from IBR owners or aggregators to involve non-utility-owned IBRs in the restoration process. Then, collaboration with regulatory agencies is necessary to validate IBR capabilities and establish operational guidelines, ensuring secure restorations. Utilities also need to work with vendors to accurately model IBRs and develop appropriate control algorithms for the restoration. Finally, coordination with regional ISO/RTO is essential to integrate upstream resources and facilitate comprehensive T&D system restoration. Besides, utilities need to deploy additional devices to support the restoration process, such as integrating IBR status into existing situational awareness (SCADA) systems for monitoring and fault alerts, deploying smart breakers to manage sub-grid processes to isolate faults, and coordinating IBR operations.

A typical restoration process after system outages can be described as follows: First, the protection breakers isolate the faulted devices, whose statuses are sent to the control center through the fault awareness (SCADA) system. Then, the utility activates the restoration decision support system and sends the fault/outage information to the restoration optimization module to determine the restoration strategy. Finally, the restoration commands generated by this module are sent to device actuators for execution.

Existing restoration optimization models are built upon the assumption of a fully functional cyber layer during the restoration process. Nevertheless, these models prove inadequate for post-attack restoration scenarios. The rapid recovery of cyber networks poses substantial challenges, wherein compromised components within the cyber layer can impede the execution of the restoration process, consequently giving rise to potential system instability and subsequent outages. In light of these considerations, there is a pressing need to enhance existing restoration frameworks to address the intricacies associated with post-cyber attack recovery. To overcome this limitation, a deterministic post-attack restoration model that considers compromised cyber networks is first developed, as illustrated in Fig. 1. This model is designed to restore as much load as possible by utilizing available IBRs and, at the same time isolating the identified compromised components. Besides the traditional grid-related constraints, this model incorporates operational constraints for IBRs, considering compromised components in the cyber layer. A detailed explanation of this deterministic model is provided in Section III.

However, the timely identification of compromised components presents significant challenges with current detection algorithms, posing a feasibility issue for the deterministic model. To address this limitation, a DRO model is designed to handle the uncertainty inherent in attack scenarios, i.e., compromised components. This model integrates the deterministic model with an ambiguity set to address the attack scenario uncertainty. This ambiguity set is a Wasserstein ball-based distribution set, presenting the possibility level of each attack scenario, and ensuring a high confidence level

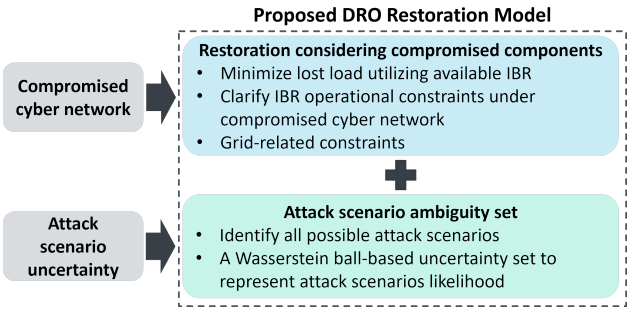

Fig. 1. Post-attack restoration framework

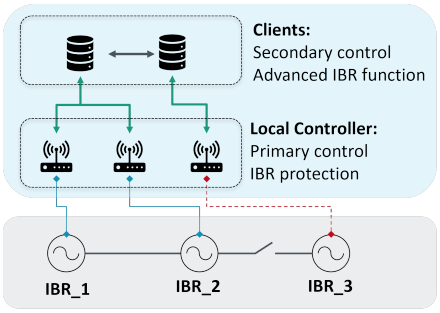

Fig. 2. IBR cyber networks

that the actual probability distribution within this set. As a result, the proposed DRO model can generate restoration plans that guarantee maximized load pickup across all scenarios. Further details on the development of the DRO model will be presented in Section IV.

III. Deterministic Post-Attack Restoration Model

In this section, the IBR operational guidelines considering compromised cyber networks will be established first, followed by the formulation of the restoration model.

A. IBR Operational Guidelines under Cyber Attacks

1) IBR control framework: Among various control algorithms designed for grid-forming IBRs, the load-sharing algorithm stands out for its cost-effectiveness and high reliability. This algorithm utilizes a distributed two-tier control framework. The secondary control gathers local and neighboring measurements, computes the deviation in their power ratios, and transmits it to the primary control, at the frequency of seconds. Operating with a faster cycle time within milliseconds, the primary control adjusts the terminal output to compensate for the power ratio deviations. As the system stabilizes in the steady state, the power ratio among each IBR unit becomes uniform, effectively achieving load sharing.

2) Cyber attacks on IBR communication system: The IBR control algorithm necessitates a two-layer communication network, as shown in Fig. 2. The first layer includes IBR Clients, communicating with local controllers and neighboring clients. These IBR clients handle the secondary control and other advanced functions, such as emergency response. The second layer comprises local controllers, exclusively communicating with their associated clients. These controllers oversee the execution of primary control and IBR protection, such as over- and under-voltage protection.

Table I presents the common vulnerabilities that attackers can exploit in the IBR cyber layer, according to the National Vulnerability Database (NVD). In this network, the protocol used for communication among clients is DNP3, while the Modbus is used between the local controller and the client. In the client nodes, the DNP3 master and the gateway are deployed for packet transportation and protocol conversion. Their vulnerabilities allow attackers to launch replay attacks, jamming attacks, and FDI attacks, which can lead to IBR misoperation, causing large-scale IBR trips, grid stability issues, and even power outages. The local controller is generally provided by the inverter manufacturer, and its vulnerabilities are inherent in the users' applications. Once local controllers are compromised, attackers can execute unauthorized commands, modify or delete data packets, and cause similar consequences to those on client nodes in the physical layer, although typically limited to local areas. Regarding the communication channel, both DNP3 and Modbus pose vulnerabilities that attackers can exploit to launch man-in-the-middle (MITM) attacks, leading to packet delays, dropping, or malicious modification of payloads. The attack consequences in the physical layer include grid loss and stability issues [2].

TABLE I
Vulnerabilities in IBR Cyber Layer

Component	Vulnerability
Client	Unrestricted upload of file Improper access control Improper certificate validation Missing authentication for critical function
Local Controller	Improper restriction of excessive authentication attempts Download of code without integrity check
Channel (DNP3)	Remotely expose network through TCP crafted packets
Channel (Modbus)	Issue harmful command to devices Reply to illegal function

Due to the low cost and high impact, the increasing frequency of cyber attacks on the IBR communication network could significantly disrupt IBR operations. In this work, we focus on attacks targeting cyber nodes, specifically client nodes and local controllers, due to their significant impact compared to channel attacks. We will delve into the attacks that disrupt connectivity between cyber nodes, such as FDI and DoS attacks. These attacks pose severe threats to IBR operations under load-sharing algorithms. To ensure secure operation, additional guidelines must be established, considering the disconnectivity among IBRs in the cyber layer.

3) IBR operational guideline considering compromised cyber network: Take Fig. 2 as an example, and assume the communication channel between IBR_3 and its local controller is compromised. To ensure the secure operation of IBRs, one option is to trip IBR_3 and terminate the compromised channel, thus the remaining IBRs can operate securely. However, it is worth noting that IBR is capable of operating in autonomous mode, which it utilizes the default configuration to execute

voltage/frequency control, serves local load within its capacity, and guarantees system state variables in normal range. Yet, without communication with other IBRs, physical connection between multiple autonomous IBRs may cause unregulated power flow, leading to state variables exceeding the normal range, and potentially causing IBR trips. Hence, physical isolation between autonomous IBRs becomes necessary when they are operating in this mode. Therefore, an alternative option is to isolate IBR_3 in both physical and cyber layers and let it operate in autonomous mode, as shown in Fig. 2. Nevertheless, this strategy restricts the power flow within two sub-grids, potentially leading to a reduced load restoration.

The choice between these two options relies on comprehensive analysis. In this example, communication between IBR_2 and IBR_3 is disrupted, and the first solution involves turning off IBR_3 to prevent potential disturbances. In the second solution, although IBR_3 is turned on, it is physically disconnected from IBR_2. In summary, to ensure the secure operation of IBRs under compromised cyber networks, it is essential to maintain consistency in their cyber and physical connectivity, regardless of their operation strategy.

B. Deterministic Model Formulation

Consider a distribution grid with $\mathcal{N}$ buses and $\mathcal{K}$ lines, alongside a set of $\mathcal{G}$ IBRs and $\mathcal{L}$ loads. The corresponding cyber network consists of $\mathcal{M}$ nodes and $\mathcal{V}$ links. Let binary variable $\boldsymbol{\xi}$ indicate the cyber node status, which is predefined in the deterministic model. $\xi_i = 0$ refers to the i^{th} node is compromised; otherwise, it is secure. Incorporating with the IBR operational constraint, binary vectors $\boldsymbol{x}$ and $\boldsymbol{z}$ are introduced to enable modification of IBR physical connections. Here, $\boldsymbol{x}^{Sw}$ denotes the states of line switches, while $\boldsymbol{z}^{IBR}$ refers to ON/OFF status of IBRs. $\mathcal{S}$ indicates the swtichable lines. The binary vector $\boldsymbol{y}$ indicates the load status. The objective function of this model can be formulated as:

$$\min_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}} \sum_{l\in\mathcal{L}} (1-y_l)\cdot P_l^{Load} \tag{1}$$

1) Grid-related constraints: Let p, q refer to the real and reactive power, u indicate bus voltage square, and binary variable e refer to the energization status of buses. The grid-related constraints can be formulated as follows:

$$\underline{P_g}^{IBR} z_g^{IBR} \le p_g^{IBR} \le \overline{P_g}^{IBR} z_g^{IBR} \quad \forall g \in \mathcal{G} \tag{2}$$

$$\underline{Q_g}^{IBR} z_g^{IBR} \le q_g^{IBR} \le \overline{Q_g}^{IBR} z_g^{IBR} \quad \forall g \in \mathcal{G} \tag{3}$$

$$|p_{ij}^{Line}| \le P_{ij}^{Line} x_{ij}^{Sw} \quad \forall ij \in \mathcal{K} \tag{4}$$

$$|q_{ij}^{Line}| \le Q_{ij}^{Line} x_{ij}^{Sw} \quad \forall ij \in \mathcal{K} \tag{5}$$

$$q_c^{Cap} \le Q_c^{Cap} \quad \forall c \in \mathcal{C} \tag{6}$$

$$u_i x_{ij}^{Sw} = [u_j - 2\widetilde{r}_{ij} p_{ij}^{line} + \widetilde{x}_{ij} q_{ij}^{line})] x_{ij}^{Sw} \quad \forall ij \in \mathcal{K} \tag{7}$$

$$p_{ij}^{Line} = y_i P_i^{Load} + \sum_{k\in C_i} p_{ki}^{Line} - p_i^{IBR} \quad \forall ij \in \mathcal{K} \tag{8}$$

$$q_{ij}^{Line} = y_i Q_i^{Load} + \sum_{k\in C_i} q_{ki}^{Line} - q_i^{IBR} - q_i^{Cap} \quad \forall ij \in \mathcal{K} \tag{9}$$

$$\underline{U}_i e_i \le u_i \le \overline{U}_i e_i \quad \forall i \in \mathcal{N} \tag{10}$$

$$x_{ij}^{Sw} = 1 \quad \forall ij \in (\mathcal{K} - \mathcal{S}) \tag{11}$$

Constraints (2)-(6) represent the capacity limits of IBRs, line branches, and capacitor banks. (7) refers to voltage drop equation, where $\widetilde{r}, \widetilde{x}$ can be referred to [17]. (8) and (9) indicate the power flow equation, and C_i indicates the children nodes of node i. (10) specifies the permissive voltage range. The determination of e_i and load-sharing constraints will be discussed next. (11) indicates the switchable lines.

2) IBR operational constraints considering compromised cyber nodes: The IBR connectivity is established among every pair of IBRs. Let $\mathcal{G}^{pair}$ denote all possible pairs, the binary variable Con^p indicates the IBR physical connectivity and Con^c denotes the IBR cyber connectivity. Based on the previous discussion, the IBR cyber and physical connectivity should be consistent to ensure secure operation. This constraint can be formulated as follows:

$$Con_{ij}^p = Con_{ij}^c \quad \forall ij \in \mathcal{G}^{pair} \tag{12}$$

The IBR physical connectivity is determined as follows: assuming ρ^{ij} refers to the set of all possible physical paths from IBR_i to IBR_j, which can be derived by the depth-first search algorithm [18]. Let binary variable $S_m^{p,ij}$ indicate the status of the m^{th} path in ρ^{ij}. Therefore, $S_m^{p,ij}$ is active only when all on-path switches are closed, expressed as follows:

$$S_m^{p,ij} = \prod_{kq\in\rho_m^{ij}} x_{kq}^{Sw} \tag{13}$$

where kq denotes the line from bus k to bus q.

The physical connectivity of IBRs is determined by the path status, as well as the ON/OFF status of IBRs. It's active when both IBRs are online, and at least one path between them is active, which can be expressed as follows:

$$Con_{ij}^p = (1 - \prod_m (1 - S_m^{p,ij})) \cdot z_i^{IBR} \cdot z_j^{IBR} \tag{14}$$

Similarly, $e_i = 1$ if the i^{th} bus is connected to at least one IBR in the ON state, as determined in the following equation:

$$e_i = 1 - \prod_{k\in\mathcal{G}} [1 - (1 - \prod_m (1 - S_m^{p,ki})) \cdot z_k^{IBR}] \tag{15}$$

The cyber connectivity can be derived similarly. Assuming the cyber node is directly connected to IBR_i and IBR_j at node k and node q respectively, the set of paths in between is σ^{kq}, then the p^{th} path status $S_p^{c,kq}$ is positive when all on-path cyber nodes (including starting and end nodes) is in normal status, which can be formulated as:

$$S_p^{c,kq} = \prod_{v\in\sigma_p^{kq}} \xi_v \tag{16}$$

Therefore, the cyber connection between IBR_i and IBR_j is active if there exists at least one active path in σ^{kq}:

$$Con_{ij}^c = (1 - \prod_p (1 - S_p^{c,kq})) \quad \forall ij \in \mathcal{G}^{pair} \tag{17}$$

In addition, the load-sharing algorithm ensures that the output power ratio among interconnected IBRs remains identical in the steady state, formulated as follows:

$$\frac{p_i^{IBR}}{\overline{P_i}^{IBR}} Con_{ij}^p = \frac{p_j^{IBR}}{\overline{P_j}^{IBR}} Con_{ij}^p \quad \forall ij \in \mathcal{G}^{pair} \tag{18}$$

$$\frac{q_i^{IBR}}{\overline{Q_i}^{IBR}} Con_{ij}^p = \frac{q_j^{IBR}}{\overline{Q_j}^{IBR}} Con_{ij}^p \quad \forall ij \in \mathcal{G}^{pair} \tag{19}$$

In summary, the deterministic model is formulated as:

$$\min_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}} \sum_{l\in\mathcal{L}}(1-y_l)\cdot P_l^{Load} \tag{20}$$

$$\text{s.t. Grid-related constraints: (2)-(11)} \tag{20a}$$

$$\text{Connectivity consistency:(12)} \tag{20b}$$

$$\text{IBR physical connectivity: (13)-(15)} \tag{20c}$$

$$\text{IBR cyber connectivity: (16)-(17)} \tag{20d}$$

$$\text{Load sharing: (18)-(19)} \tag{20e}$$

There are multiple nonlinear constraints in this problem, as indicated by equations (13)-(17). These constraints contain high order binary variable terms and can be reformulated into linear constraints. For instance, equation (13) can be expressed as a set of equivalent linear constraints:

$$S_m^{p,ij} \le x_{kq}^{Sw} \quad \forall kq \in \rho_m^{ij} \tag{21}$$

$$S_m^{p,ij} \ge \sum_{kq\in\rho_m^{ij}} (x_{kq}^{Sw}-1)+1 \tag{22}$$

Similarly, constraints (14)-(17) can be transformed into linear forms. Consequently, the deterministic model is converted into a standard MILP problem.

IV. Proposed DRO Model

In reality, it is significantly challenging to fully identify the compromised nodes in real time, which notably reduces the practicality of the proposed deterministic model (20) in Section III. Therefore, we have to embrace the uncertainty of compromised nodes in developing the post-attack restoration model. Approaches such as stochastic optimization (SO), robust optimization (RO), and the recently proposed distributionally robust optimization (DRO) can be employed to tackle this challenge. However, the SO method necessitates complete knowledge of attack probability distribution, which may be difficult to obtain or inaccurate due to limited historical attack data. While the RO method does not require exhaustive probability distribution information, it may yield overly conservative solutions. In contrast, the DRO model demonstrates superior performance compared to SO and RO. This model utilizes an ambiguity set to address the uncertainty of compromised nodes, encompassing empirical attack probability information and ensuring the real attack probability distribution falls within this set with a considerable confidence level. This feature enhances the model's reliability compared to SO. Compared to RO, this approach extensively integrates probability information, thereby yielding a more efficient yet robust solution [19]. The DRO model enables the derivation of the sub-grid plans to maximize load pickup expectations across all possible attack scenarios in worst-case scenarios.

The development of the proposed DRO model is illustrated in Fig. 3. This model is established through two steps. First, the ambiguity set is defined. The input data include physical layer observations (e.g., outage information), ambiguity set parameters (e.g., confidence level), historical attack data, and the cyber-physical interdependence model. This process is detailed in Subsection A. Then, the previous deterministic model is revised and combined with the ambiguity set to formulate the DRO model. The revised deterministic model requires the cyber and physical configuration and the sub-grid breaker profile as inputs. By replacing the regular variable ξ as the random variable and combining it with the ambiguity set, the DRO model is finally established. This process is presented in Subsection B. In Subsection C, a detailed comparison of RO, SO, and DRO models is presented to demonstrate the overall best performance of the DRO model.

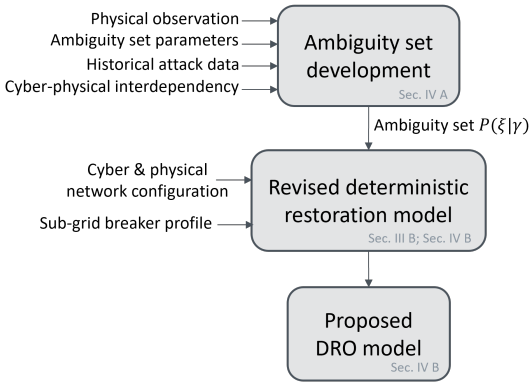

Fig. 3. Flowchart of DRO model development

A. Ambiguity Set

Given the condition of observed physical outages denoted as $\boldsymbol{\gamma}$, the probability of scenario $\boldsymbol{\xi}$ can be represented as $P(\boldsymbol{\xi}|\boldsymbol{\gamma})$. While one cannot directly obtain the information of this probability, it can be determined through the Bayesian inference as follows:

$$P(\boldsymbol{\xi}|\boldsymbol{\gamma}) = \frac{P(\boldsymbol{\gamma}|\boldsymbol{\xi})P(\boldsymbol{\xi})}{\sum_{\Xi} P(\boldsymbol{\gamma}|\boldsymbol{\xi}^i)P(\boldsymbol{\xi}^i)} \tag{23}$$

where Ξ represents the supporting space of $\boldsymbol{\xi}$. $P(\boldsymbol{\xi}|\boldsymbol{\gamma})$ can be determined by the following steps.

1) Determination of $P(\boldsymbol{\gamma}|\boldsymbol{\xi})$: $P(\boldsymbol{\gamma}|\boldsymbol{\xi})$ indicates the probability of γ given attack scenario $\boldsymbol{\xi}$. This probability represents the problem faced by attackers in determining their targets, i.e. which IBR to trip, and it can be addressed using the Quantal Response model [20]. This model suggests that, owing to incomplete information or subjective preferences, attackers may not necessarily choose the node with the highest reward U. Instead, they might target all IBRs. To capture this scenario, a non-negative parameter μ is introduced to assess their rationality level. A higher μ signifies a greater likelihood of selecting targets with larger U. Specifically, the probability of each node being chosen is determined by the following:

$$P(\boldsymbol{\gamma}|\boldsymbol{\xi}) = \frac{\exp(\mu_{\boldsymbol{\xi}} U_{\boldsymbol{\gamma}})}{\sum_{\Gamma^{\boldsymbol{\xi}}} \exp(\mu_{\boldsymbol{\xi}} U_{\boldsymbol{\gamma}^j})} \tag{24}$$

In this paper, U is defined as the capacity of tripped IBRs, expressed as follows:

$$U_{\boldsymbol{\gamma}} = (\mathbf{1}-\boldsymbol{\gamma})^T \boldsymbol{Cap}^{IBR} \tag{25}$$

Let Γ denote all possible scenarios of IBR tripping. Then $\Gamma^{\boldsymbol{\xi}}$ is a sub-set of Γ, indicating all IBR tripping scenarios under $\boldsymbol{\xi}$, which is determined by analyzing the cyber-physical interdependence model.

Based on our previous work [2], the cyber-physical interdependence can be modeled as a graph-based mapping function $M(t)$, to demonstrate the data exchange patterns between the control and physical system layers. $M(t)$ is derived from the cyber node functionality matrix denoted as $F(t)$, the data packets transition path matrix $P(t)$, and the data packets starting node incidence matrix S. The mapping function of i^{th} cyber node $\boldsymbol{M}^i$ is determined by the following equation:

$$\boldsymbol{M}^i = (diag(\boldsymbol{P}^i \odot \boldsymbol{F}))^T \cdot \boldsymbol{S}^i \tag{26}$$

Accordingly, the cyber sensitivity matrix, indicating which IBR can be tripped by each node, is defined as:

$$\boldsymbol{Sen} = [\frac{d\boldsymbol{M}^1}{d\boldsymbol{\lambda}^1} ... \frac{d\boldsymbol{M}^{\mathcal{M}}}{d\boldsymbol{\lambda}^{\mathcal{M}}}] \tag{27}$$

where λ refers to the input data of each node. In this model, it refers to the IBR control command. Replacing all non-zero elements in $\boldsymbol{Sen}$ as 1, and denotes the new matrix as $\boldsymbol{Sen}'$. Thus, all possible tripped IBRs under $\boldsymbol{\xi}$, denoting as $\hat{\gamma}^{\boldsymbol{\xi}}$ is determined by following equation:

$$\hat{\gamma}^{\boldsymbol{\xi}} = \boldsymbol{Sen}'(1 - \boldsymbol{\xi}) \tag{28}$$

Thus, $\Gamma^{\boldsymbol{\xi}}$ can be determined accordingly.

2) Determination of $P(\boldsymbol{\xi})$: In response to the escalating cyber threats to power systems, the deployment of honeypot systems [21] as intrusion detection techniques in power systems supporting communication networks has recently become more popular. These systems meticulously monitor and log intrusion attempts, regardless of whether they would lead to physical impacts or not. These records can be used to identify attack types and determine whether they will disrupt cyber connectivity. However, given the limited implementation, the data collected by the honeypot system might not be sufficient to establish a convincing attack probability distribution. Nevertheless, it can contribute to the development of a Wasserstein ball-based probability distribution set, which presents the desirable out-of-sample performance [22]. This set ensures that the true $P(\boldsymbol{\xi})$ falls within the distribution set with a considerable confidence level, which is suitable for approximating $P(\boldsymbol{\xi})$.

Assuming the honeypot system has collected N historical data. In this framework, it's crucial to ensure that N is not too small. A small N could result in a relatively large ambiguity set, leading to a "worst-case" scenario that significantly diverges from the actual probability and tends to be more conservative. Moreover, given the rapid pace of technological evolution, outdated data may fail to accurately reflect current attacker behaviors. Therefore, we recommend utilizing data no older than two years. Based on [23], within this framework, N is selected within the range of at least 1 month but no more than 2 years. This dataset is updated weekly. The more data available, the more accurate estimation of the worst-case scenario. These data can be expressed as: $\hat{\boldsymbol{\xi}}^1$, $\hat{\boldsymbol{\xi}}^2$,...$\hat{\boldsymbol{\xi}}^N$, then the empirical distribution of $\boldsymbol{\xi}$ is formulated as:

$$P_0(\boldsymbol{\xi}) = \frac{1}{N} \sum_{i=1}^{N} \delta_{\hat{\boldsymbol{\xi}}^i}(\boldsymbol{\xi}) \tag{29}$$

where $\delta_{\hat{\boldsymbol{\xi}}^i}$ is defined by following equation:

$$\delta_{\hat{\boldsymbol{\xi}}^i} = \begin{cases} 1, & \hat{\boldsymbol{\xi}}^i = \boldsymbol{\xi} \\ 0, & \text{otherwise} \end{cases} \tag{30}$$

The supporting space Ξ can be approximated as the sampling space. The confidence level β is determined by the subjective preference of the users. For a given confidence level β, there is always a Wasserstein ball centered with $P_0(\boldsymbol{\xi})$ satisfying that:

$$P(D_W(P(\boldsymbol{\xi}), P_0(\boldsymbol{\xi})) \leq \nu) \geq \beta \tag{31}$$

where ν indicates the Wasserstein ball radius and can be derived by the following equation [24]:

$$\nu = D\sqrt{\frac{-2\log(1-\beta)}{N}} \tag{32}$$

where D denotes the diameter of $\boldsymbol{\xi}$ supporting space. Given historical data and a predefined confidence level, ν is fixed. $D_W(P(\boldsymbol{\xi}), P_0(\boldsymbol{\xi}))$ in (31) refers to the Wasserstein distance:

$$D_W(P(\boldsymbol{\xi}), P_0(\boldsymbol{\xi})) = \min_{\Xi^2} \sum_{\boldsymbol{\xi}}^{\Xi} \sum_{\boldsymbol{\xi}'}^{\Xi} ||\boldsymbol{\xi} - \boldsymbol{\xi}'|| p(\boldsymbol{\xi}, \boldsymbol{\xi}') \tag{33}$$

where $p(\boldsymbol{\xi}, \boldsymbol{\xi}')$ is the combined probability of $(\boldsymbol{\xi}, \boldsymbol{\xi}')$, with the marginal probability of $P(\boldsymbol{\xi})$ and $P_0(\boldsymbol{\xi})$, respectively. Then, (33) can be reformulated as follows [25]:

$$\sum_{\boldsymbol{\xi}}^{\Xi} \sum_{\boldsymbol{\xi}'}^{\Xi} ||\boldsymbol{\xi} - \boldsymbol{\xi}'|| p(\boldsymbol{\xi}, \boldsymbol{\xi}') \leq \nu \tag{34}$$

$$\sum_{\boldsymbol{\xi}'}^{\Xi} p(\boldsymbol{\xi}, \boldsymbol{\xi}') = P(\boldsymbol{\xi}) \tag{34a}$$

$$\sum_{\boldsymbol{\xi}}^{\Xi} p(\boldsymbol{\xi}, \boldsymbol{\xi}') = P_0(\boldsymbol{\xi}') \tag{34b}$$

Thus, the ambiguity set of $P(\boldsymbol{\xi}|\boldsymbol{\gamma})$, denoted as $\mathcal{P}$, can be summarized as:

$$\mathcal{P} = \{P(\boldsymbol{\xi}|\boldsymbol{\gamma}) | (23)(24)(34)\} \tag{35}$$

In summary, the development of the ambiguity set is presented in Fig. 4. First, the model utilizes our previously developed cyber-physical interdependence model to determine attack reward U given any γ. Then the Quantal Response model is used to determine the marginal likelihood $P(\boldsymbol{\gamma}|\boldsymbol{\xi})$. Next, an uncertainty set is constructed to represent the prior probability $P(\boldsymbol{\xi})$, which is a Wasserstein ball-based set centered with empirical distribution $P_0(\boldsymbol{\xi})$. Finally, according to Bayesian inference, the ambiguity set representing the probability level of all possible scenarios $P(\boldsymbol{\xi}|\boldsymbol{\gamma})$ can be derived.

B. DRO Formulation

The DRO model aims to generate a sub-grid plan to minimize the expected lost load across all possible scenarios in the worst-case situations. Consequently, $\boldsymbol{\xi}$ becomes a variable in this case, with the ambiguity set of (35). The decision variable $\boldsymbol{x}$ pertains to the sub-grid plan. Specifically, $\boldsymbol{x}^{Iso}$ denotes whether the cyber node requires isolation, particularly if its likelihood of being compromised is relatively high. $\boldsymbol{x}^{Link}$ refers to the decision to terminate the communication channel

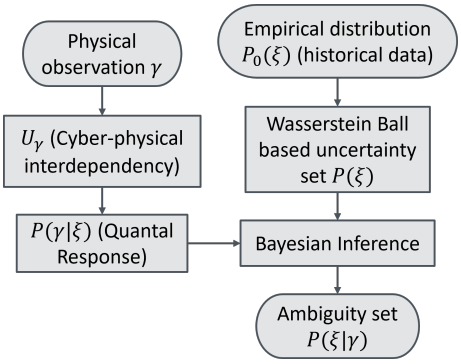

Fig. 4. Development of ambiguity set

within the cyber network to facilitate its sub-grid. Meanwhile, $\boldsymbol{x}^{Sw}$ still denotes the line relay status, serving the purpose of the physical distribution system sub-grid.

In the DRO model, the operational constraints remain identical to the deterministic model, except for cyber connectivity, which is determined as follows. Assuming the set of paths between node k and node q is still σ^{kq}, then the p^{th} path status is active only when there are no isolated nodes, compromised nodes, or terminated links on the path, as follows:

$$S_p^{c,kq} \leq \prod_{v\in\sigma_p^{kq}} \xi_v \tag{36}$$

$$S_p^{c,kq} \leq \prod_{v\in\sigma_p^{kq}} x_v^{Iso} \tag{37}$$

$$S_p^{c,kq} \leq \prod_{N_\sigma^{kq}-1}^{i} x^{Link}_{\sigma^{kq}_{p,i},\sigma^{kq}_{p,i+1}} \tag{38}$$

$$S_p^{c,kq} \geq \sum_{v\in\sigma_p^{kq}} (\xi_v + x_v^{Iso} - 2) + \sum_{N_\sigma^{kq}-1}^{i} (x^{Link}_{\sigma^{kq}_{p,i},\sigma^{kq}_{p,i+1}} - 1) + 1 \tag{39}$$

These constraints can be linearized in a similar way to the method presented in the deterministic model. Therefore, the proposed DRO model can be formulated as follows:

$$\min_{\boldsymbol{x}} \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[\min \sum_{l\in\mathcal{L}} (1 - y_l^{\boldsymbol{\xi}}) \cdot P_l^{Load}] \tag{40}$$

$$\text{s.t.} \quad \text{Operational Constraints:} \quad (2)\text{-}(11),(12)\text{-}(15),(36)\text{-}(39),(17) \quad \forall \boldsymbol{\xi} \in \Xi \tag{40a}$$

$$\text{Uncertainty Constraints: } (23)(24)(34) \tag{40b}$$

C. Comparison with SO and RO

For the SO algorithm, the ambiguity set is reduced to a single distribution, which is determined as follows: the Wasserstein ball-based set $D_W(P(\boldsymbol{\xi}), P_0(\boldsymbol{\xi}))$ is replaced solely by the empirical distribution $P_0(\boldsymbol{\xi})$. The attackers rational level is also determined by historical data, resulting in a fixed value rather than a range. Consequently, the probability distribution of all attack scenarios given physical outage $\boldsymbol{\gamma}$ can still be derived based on (23) and (24), denoted as $\mathbb{P}_0$. Thus, the objective function in the SO algorithm becomes:

$$\min_{\boldsymbol{x}} \mathbb{E}_{\mathbb{P}_0}[\min \sum_{l\in\mathcal{L}} (1 - y_l^{\boldsymbol{\xi}}) \cdot P_l^{Load}] \tag{41}$$

Thus, this algorithm will prioritize the scenarios with higher probabilities in $\mathbb{P}_0$. However, due to the limited historical data, this probability distribution may not accurately reflect the real likelihood of each attack scenario. For example, if $\mathbb{P}_0$ indicates a higher probability of attacks on the local controller, the solution derived from SO will most likely only isolate the local controller and shut down the corresponding IBR. However, if the actual probability distribution suggests that the client nodes or other critical nodes are of high probability to be targeted in the cyber layer, the SO algorithm's focus on the local controller alone will leave the actual compromised nodes unidentified. These unidentified nodes could launch continuous attacks and cause secondary outages. Therefore, the solution derived from the SO algorithm is unreliable.

For the RO algorithm, only the worst-case attack scenario is considered. It is worth noting that the "worst-case" here differs from that in the DRO model. Instead, it refers to a particular attack scenario, whereas the "worst-case" in the DRO model pertains to a specific attack probability distribution. This distribution encompasses all possible attack scenarios, with the expectation of load pick-up across all scenarios being the worst. The objective function of RO is modeled as follows:

$$\min_{\boldsymbol{x}} \sup_{\boldsymbol{\xi}\in\boldsymbol{\Xi}} \min \sum_{l\in\mathcal{L}} (1 - y_l^{\boldsymbol{\xi}}) \cdot P_l^{Load} \tag{42}$$

The restoration strategy derived from the RO algorithm only considers the scenario that the compromised node is the most critical one and will cause the most severe disconnection among IBRs. Thus, to ensure consistent cyber and physical connectivity among IBRs, the main grid is divided into multiple sub-grids. However, this approach blocks power flow and results in the most lost load. Compared to SO and DRO, this algorithm is the most "robust," avoiding secondary outages to the greatest extent. However, if the probability of this scenario is relatively low, the algorithm is overly conservative and may lead to significantly higher economic losses due to the increased lost load.

Based on the discussion above, the DRO model can adjust the restoration strategy according to the probability level of each attack scenario, resulting in better performance compared to both SO and RO models.

V. Reformulation of DRO Model

Let $Q_{\boldsymbol{\xi}}(\boldsymbol{x})$ indicates the inner problem of (40), i.e.:

$$\begin{aligned} Q_{\boldsymbol{\xi}}(\boldsymbol{x}) = \min_{\boldsymbol{y^\xi}, \boldsymbol{z^\xi}} \{ &\sum_{l\in\mathcal{L}} (1 - y_l^{\boldsymbol{\xi}}) \cdot P_l^{Load} : \\ &(2) - (11)(12) - (15)(36) - (39)(17)\} \end{aligned} \tag{43}$$

This problem is similar to the previously presented deterministic model and can be easily solved by the off-the-shelf solver for each given $\boldsymbol{\xi}$ and $\boldsymbol{x}$. Thus, the proposed DRO problem can be reformulated as:

$$\begin{aligned} \min_{\boldsymbol{x}} \sup_{\mathbb{P}\in\mathcal{P}} &\sum_{\boldsymbol{\Xi}} P(\boldsymbol{\xi}|\boldsymbol{\gamma}) Q_{\boldsymbol{\xi}}(\boldsymbol{x}) \\ \text{s.t.} \quad &(23)(24)(34) \end{aligned} \tag{44}$$

In the following sections, (44) is reformulated step by step and transformed into a tractable problem.

A. Reformulation of Quantal Response Constraint

The first step is to transform the Quantal Response constraint (24) into a linear constraint. Given that $\mu_{\boldsymbol{\xi}} U_{\boldsymbol{\gamma}}$ in (24) is non-negative, $\exp(\mu_{\boldsymbol{\xi}} U_{\boldsymbol{\gamma}})$ is positive. Dividing both numerator and denominator of equation (24) by $\exp(\mu_{\boldsymbol{\xi}} U_{\boldsymbol{\gamma}})$, it can be rewritten as follows:

$$P(\boldsymbol{\gamma}|\boldsymbol{\xi}) = \frac{1}{\sum_{\Gamma^{\boldsymbol{\xi}}} \exp(\mu_{\boldsymbol{\xi}}(U_{\boldsymbol{\gamma}^j} - U_{\boldsymbol{\gamma}}))} \tag{45}$$

Let $f(\mu_{\boldsymbol{\xi}}) = \sum_{\Gamma^{\boldsymbol{\xi}}} \exp(\mu_{\boldsymbol{\xi}}(U_{\boldsymbol{\gamma}^j} - U_{\boldsymbol{\gamma}}))$, then the second order derivative of $f(\mu_{\boldsymbol{\xi}})$ with respect to $\mu_{\boldsymbol{\xi}}$ is expressed as follows:

$$\frac{d^2 f}{d\mu_{\boldsymbol{\xi}}^2} = \sum_{\Gamma^{\boldsymbol{\xi}}} (U_{\boldsymbol{\gamma}^j} - U_{\boldsymbol{\gamma}})^2 \exp(\mu_{\boldsymbol{\xi}}(U_{\boldsymbol{\gamma}^j} - U_{\boldsymbol{\gamma}})) \tag{46}$$

It is easy to prove that $\frac{d^2 f}{d\mu_{\boldsymbol{\xi}}^2}$ is positive definite, therefore $f(\mu_{\boldsymbol{\xi}})$ is convex. Giving $\mu_{\boldsymbol{\xi}}$ within the non-negative range $[\underline{\mu}_{\boldsymbol{\xi}}, \overline{\mu}_{\boldsymbol{\xi}}]$, which can be estimated by reviewing the log information of the compromised nodes, $f(\mu_{\boldsymbol{\xi}})$ is bounded and positive. Consequently, $P(\boldsymbol{\gamma}|\boldsymbol{\xi})$ is also bounded and positive. Since $P(\boldsymbol{\gamma}|\boldsymbol{\xi})$ depends solely on $\mu_{\boldsymbol{\xi}}$, and $\mu_{\boldsymbol{\xi}}$ is independent of all other variables, it is possible to treat $P(\boldsymbol{\gamma}|\boldsymbol{\xi})$ as an independent variable, with its bounds determined by $\mu_{\boldsymbol{\xi}}$. Let $b_{\boldsymbol{\xi}} = P(\boldsymbol{\gamma}|\boldsymbol{\xi})$, based on the previous discussion, (24) is equivalent to the following linear constraints, with lower and upper bounds of the variable determined offline using Algorithm 1.

$$\begin{gathered} \underline{b}_{\boldsymbol{\xi}} \leq b_{\boldsymbol{\xi}} \leq \overline{b}_{\boldsymbol{\xi}} \\ \underline{b}_{\boldsymbol{\xi}} = \frac{1}{\max f(u_{\boldsymbol{\xi}})}, \overline{b}_{\boldsymbol{\xi}} = \frac{1}{\min f(u_{\boldsymbol{\xi}})} \end{gathered} \tag{47}$$

Algorithm 1 is based on the classical gradient descent method to solve convex optimization problems. The basic idea is to search for the minimum value along the direction in which the gradient decreases the fastest. In Algorithm 1, α refers to the user-defined step size. A smaller α tends to improve the convergence of algorithm, but increases the execution time. δ indicates a minimal value, forcing the gradient to approach 0, thereby bringing the searched minimum value as close as possible to the actual minimum value.

Algorithm 1 Gradient Descent Search

if $f'(\underline{\mu}_{\boldsymbol{\xi}}) \geq 0$ **then**
  $\underline{b}_{\boldsymbol{\xi}} \leftarrow \frac{1}{f(\overline{\mu}_{\boldsymbol{\xi}})}, \overline{b}_{\boldsymbol{\xi}} \leftarrow \frac{1}{f(\underline{\mu}_{\boldsymbol{\xi}})}$
else if $f'(\overline{\mu}_{\boldsymbol{\xi}}) \leq 0$ **then**
  $\underline{b}_{\boldsymbol{\xi}} \leftarrow \frac{1}{f(\underline{\mu}_{\boldsymbol{\xi}})}, \overline{b}_{\boldsymbol{\xi}} \leftarrow \frac{1}{f(\overline{\mu}_{\boldsymbol{\xi}})}$
else
  $\mu_{\boldsymbol{\xi}} \leftarrow \underline{\mu}_{\boldsymbol{\xi}}, g \leftarrow |f'(\mu_{\boldsymbol{\xi}})|$
  while $g \geq \delta$ **do**
    $\mu_{\boldsymbol{\xi}} \leftarrow \mu_{\boldsymbol{\xi}} - \alpha * g, g \leftarrow |f'(\mu_{\boldsymbol{\xi}})|$
  end while
  $\underline{b}_{\boldsymbol{\xi}}, \overline{b}_{\boldsymbol{\xi}} \leftarrow \min, \max\{\frac{1}{f(\mu_{\boldsymbol{\xi}})}, \frac{1}{f(\overline{\mu}_{\boldsymbol{\xi}})}, \frac{1}{f(\underline{\mu}_{\boldsymbol{\xi}})}\}$
end if

B. Transformation of Proposed DRO Problem Formulation

We first convert (23) into a linear constraint and then transform the DRO formulation into a tractable convex form.

1) Transformation of (23): Based on (47), (23) can be reformulated as:

$$P(\boldsymbol{\xi}|\boldsymbol{\gamma}) = \frac{b_{\boldsymbol{\xi}} P(\boldsymbol{\xi})}{\sum_{\Xi} b_{\boldsymbol{\xi}^i} P(\boldsymbol{\xi}^i)} \tag{48}$$

Introduce new variable $k_{\boldsymbol{\xi}} = b_{\boldsymbol{\xi}} P(\boldsymbol{\xi})$. Given that $P(\boldsymbol{\xi})$ is non-negative, the following constraint can be derived based on (47):

$$b_{\boldsymbol{\xi}} P(\boldsymbol{\xi}) \leq b_{\boldsymbol{\xi}} P(\boldsymbol{\xi}) = k_{\boldsymbol{\xi}} \leq \overline{b}_{\boldsymbol{\xi}} P(\boldsymbol{\xi}) \tag{49}$$

In addition, (48) can be rewritten as:

$$P(\boldsymbol{\xi}|\boldsymbol{\gamma}) = \frac{k_{\boldsymbol{\xi}}}{\sum_{\boldsymbol{\Xi}} k_{\boldsymbol{\xi}^i}} \tag{50}$$

Therefore, variable $b_{\boldsymbol{\xi}}$ can be replaced by $k_{\boldsymbol{\xi}}$ in this model. Insert (50) into the objective function of (44) and replace (49) with (47), the proposed DRO model can be reformulated as:

$$\begin{aligned} &\min_{\boldsymbol{x}} \sup_{\mathbb{P} \in \mathcal{P}} \frac{\sum_{\boldsymbol{\Xi}} k_{\boldsymbol{\xi}} Q_{\boldsymbol{\xi}}(\boldsymbol{x})}{\sum_{\boldsymbol{\Xi}} k_{\boldsymbol{\xi}}} \\ &\text{s.t.} \quad (49)(34) \end{aligned} \tag{51}$$

2) Convexification of DRO Formulation: (51) is an optimization problem with linear fraction objective function and linear constraints, constituting a quasi-convex problem, and it can be transformed into a convex form. Let $s_{\boldsymbol{\xi},\boldsymbol{\xi}'} = \frac{p(\boldsymbol{\xi},\boldsymbol{\xi}')}{\sum_{\boldsymbol{\Xi}} k_{\boldsymbol{\xi}}}$, $u_{\boldsymbol{\xi}} = \frac{P(\boldsymbol{\xi})}{\sum_{\boldsymbol{\Xi}} k_{\boldsymbol{\xi}}}$, $m_{\boldsymbol{\xi}} = \frac{k_{\boldsymbol{\xi}}}{\sum_{\boldsymbol{\Xi}} k_{\boldsymbol{\xi}}}$, $v = \frac{1}{\sum_{\boldsymbol{\Xi}} k_{\boldsymbol{\xi}}}$, then the inner sup problem of (51) can be rewritten as follows:

$$\sup_{\boldsymbol{s},\boldsymbol{u},\boldsymbol{m},v} \quad \boldsymbol{m}^T \boldsymbol{Q}(\boldsymbol{x}) \tag{52}$$

$$\text{s.t.} \quad \boldsymbol{a}^T \boldsymbol{s} - \nu v \leq 0 \tag{52a}$$

$$b_{\boldsymbol{\xi}} u_{\boldsymbol{\xi}} \leq m_{\boldsymbol{\xi}} \leq \overline{b}_{\boldsymbol{\xi}} u_{\boldsymbol{\xi}} \quad \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{52b}$$

$$\sum_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} s_{\boldsymbol{\xi},\boldsymbol{\xi}'} = u_{\boldsymbol{\xi}'} \quad \forall \boldsymbol{\xi}' \in \boldsymbol{\Xi} \tag{52c}$$

$$\sum_{\boldsymbol{\xi}' \in \boldsymbol{\Xi}} s_{\boldsymbol{\xi},\boldsymbol{\xi}'} = P_0(\boldsymbol{\xi}) v \quad \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{52d}$$

$$\mathbf{1}^T \boldsymbol{u} = v \tag{52e}$$

$$\mathbf{1}^T \boldsymbol{m} = 1 \tag{52f}$$

where $a^{\boldsymbol{\xi},\boldsymbol{\xi}'} = ||\boldsymbol{\xi} - \boldsymbol{\xi}'||$. Thus, the proposed DRO model becomes a tracable optimization problem.

C. C&CG-based Solution Algorithm

The proposed DRO problem is a tri-level optimization problem. Given that the dual function of the inner sup problem is a quadratic form, the cutting-plane method will be employed to solve this problem. Specifically, the C&CG algorithm [26] will be utilized, due to its efficient convergence characteristics. This algorithm decomposes the original problem into two sub-problems. The master problem is formulated as follows:

$$\boldsymbol{MP}: \quad \min_{\boldsymbol{x}} \eta \tag{53}$$

$$\text{s.t.} \quad \eta \geq \sum_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} m_{\boldsymbol{\xi}}^{t,*} [(\mathbf{1} - \boldsymbol{y}^{\boldsymbol{\xi},t})^T \boldsymbol{P}^{Load}] \quad \forall t \leq k \tag{53a}$$

$$(2) - (11)(12) - (15)(36) - (39)(17) \quad \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{53b}$$

The sub-problem is a bi-level problem, which can be referred to (52), where the second level problem $\boldsymbol{Q}(x)$ is defined by (43).

The implementation of the C&CG algorithm for solving the reformulated DRO problem is outlined in Algorithm 2. The overall iteration process can be described as follows: Upon initializing $\boldsymbol{m}^{k,*}$, the master problem transforms into a single-level MILP problem. The lower bound LB is updated as its objective value, and the optimal sub-grid plan $\boldsymbol{x}^*$ is obtained. Subsequently, (43) is solved for any given $\boldsymbol{x}^*$ to derive $\boldsymbol{Q}(x)$ under each attack scenario $\boldsymbol{\xi}$. The sub-problem (52) then becomes a linear problem. The upper bound UB is updated with the objective value of (52), and $\boldsymbol{m}^{k+1,*}$ is determined for the next iteration. This iteration loop continues until LB and UB are sufficiently close.

Algorithm 2 C&CG Algorithm

1: $UB \leftarrow +\infty, LB \leftarrow -\infty, k \leftarrow 0$
2: initiate $\boldsymbol{m}^{k,*} = \boldsymbol{m}^0$
3: **while** UB-LB$\leq \epsilon$ **do**
4: Solve **MP**, update $LB, \boldsymbol{x}^*$
5: Solve (43) $\forall \boldsymbol{\xi} \in \boldsymbol{\Xi}$, update $\boldsymbol{Q}^*(\boldsymbol{x}^*)$
6: Solve **SP**, update $UB, \boldsymbol{m}^{k+1,*}$
7: $k \leftarrow k+1$
8: **end while**

The comprehensive process of DRO reformulation and solution algorithm can be summarized as follows: first, reformulate (24) into an equivalent linear box constraint (47). Then, the inner problem can be transformed into a linear fraction form. This can be further converted into a convex problem through variable substitution. As a result, this DRO problem is transformed into a tractable tri-level optimization problem that can be solved using the C&CG algorithm.

VI. Simulation Results

A. IEEE 13-node Test Cases

1) Test system: The standard test system is modified to include IBRs at Nodes 680, 633, 692, and 675, with different capacities of 700kW, 1000kW, 600kW, and 1000kW, respectively. The circuit breakers are located at Lines 670671, 671692, and 692675. The corresponding cyber network is shown in Fig. 5. The cyber nodes can be divided into two tiers: IBR clients and local controllers. In this test system, Nodes 1, 2, and 3 are designed as clients, while the remaining nodes serve as local controllers.

Let us assume IBR at Node 675 is reported to be tripped. The historical data identifies 7 distinct attack scenarios, each targeting one unique cyber node. The confidence level β is set to 95%, data size $N = 1,000$, and the range of regional index $[\underline{\mu}, \overline{\mu}] = [5, 20]$. Two cases have been developed and tested, each with different empirical distributions. Case 1 depicts a scenario where local controllers have a higher probability of being attacked. This may be due to insufficient security configuration in local controllers. Case 2 portrays a scenario where IBR clients are more susceptible to attacks, likely due to attackers prioritizing client nodes that could potentially cause more significant impacts on the physical layer.

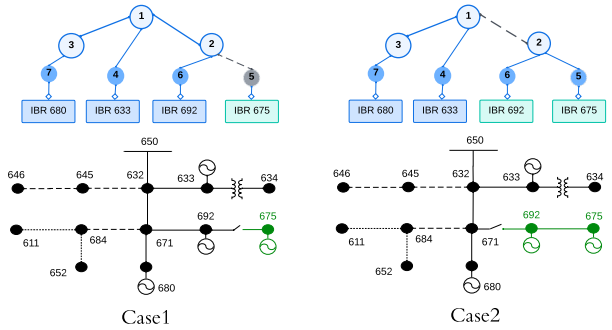

Fig. 5. Sub-grid strategy in IEEE 13-node test cases (top: cyber network, bottom: physical network)

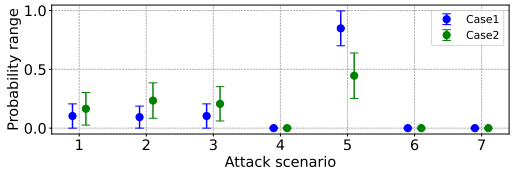

Fig. 6. Probability range of attack scenario in IEEE 13-node test cases

2) Attack probability: Fig. 6 indicates the probability range of each attack scenario. The scenario index corresponds to the compromised node. It can be observed that Scenarios 4, 6, and 7 exhibit a probability of 0, as Nodes 4, 6, and 7 function as local controllers without the authorization to manage other IBRs. Consequently, they are unable to trigger IBR 675 to trip. Conversely, Nodes 1, 2, and 3 operate as client nodes, possessing the necessary authorization to initiate trip commands for IBR 675. Meanwhile, Node 5 acts as the local controller of IBR 675, empowered to issue a trip command when specific protection mechanisms are activated.

In Case 1, the most probable scenario is the compromise of local controller at Node 5, with a probability range between 0.7 and 1. In Case 2, there is a higher probability of the client node being attacked, ranging approximately from 0.1 to 0.3. Besides, the probability range for Node 5 experiences a noticeable decrease compared to Case 1. It is worth noting that this range is derived from the entire ambiguity set. For a specific distribution, the sum of the probabilities of each attack scenario equals 1.

3) Restoration result: In Case 1, due to the dominant probability of an attack on Node 5, the DRO model proposes a strategy to isolate Node 5 (marked gray in Fig. 5) directly within the cyber layer. Accordingly, Node 675 is disconnected from the main grid in the physical layer. The distribution grid is therefore divided into two subgrids. During restoration, IBR 675 operates in autonomous mode, eliminating the impact of the compromised local controller. In Case 2, the compromised nodes within the client nodes hold a higher likelihood, which potentially impacts more IBRs. To respond to this risk, the proposed DRO method provides a strategy to divide the distribution grid into two subgrids, one with IBRs 692 and 675 and the other with IBRs 633 and 680, restricting the potential impact of compromised client nodes within smaller regions, thereby reducing the overall lost load.

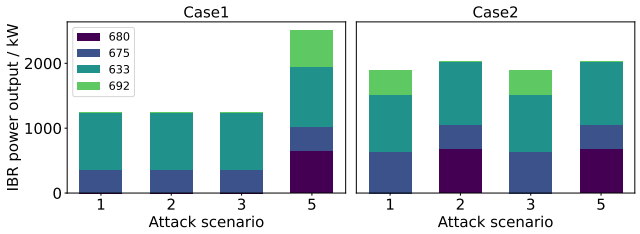

Fig. 7. IBR output power in IEEE 13-node test cases

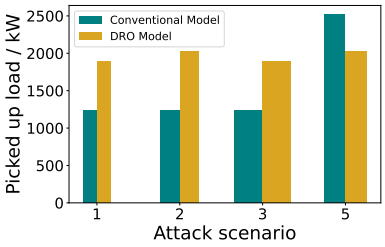

Fig. 8. Comparisons of DRO model and conventional restoration method

TABLE II
LOAD PICK-UP IN IEEE 13-NODE TEST CASES

Scenario	1	2	3	5	Expectation
Case1 [kW]	1243	1243	1243	2514	
Prob. (worst case)	0.057	0.038	0.204	**0.700**	2131.46
Case2 [kW]	1898	2030	1898	2030	
Prob. (worst case)	0.167	0.227	0.347	0.259	1962.15

Fig. 7 illustrates the IBR output across all scenarios. In Case 1, the compromise of Nodes 1, 2, or 3, as per the sub-grid plan, impacts the sub-grid composed of IBRs 680, 633, and 692. To ensure optimal operations amid a compromised communication network, the most secure strategy is to select an IBR with the largest capacity for operation while keeping the others offline. Simultaneously, isolating IBR 675 from the remaining IBRs enables its autonomous operation in every scenario. In Scenario 5, the accurate isolation of the compromised node allows the sub-grid comprising IBRs 680, 633, and 692 to securely operate in a load-sharing mode, resulting in increased load pickup. In Case 2, in Scenarios 1 and 3, IBRs 692 and 675 establish a secure connection, allowing them to operate efficiently in a load-sharing mode. Conversely, only one unit between IBRs 680 and 633 can remain online. In Scenarios 2 and 5, the situation reverses. In summary, this sub-grid plan accommodates the operation of 3 IBRs and ensures balanced load pick-up in each scenario.

For load restoration, the medium-level loads are typically recoverable across all cases and scenarios. However, the restoration of larger loads depends on the sub-grid capability.

Essentially, the proposed DRO model provides adaptive sub-grid plans customized to specific probability profiles, ensuring optimal load recovery across all potential scenarios.

4) Comparison to conventional restoration: To demonstrate the performance of the proposed model, this section compares the DRO method with the conventional restoration approach. Following the reported IBR 675 incident, the conventional restoration algorithm ignores cyber-physical interdependence and directly isolates IBR 675, allowing it to operate autonomously, which is the same restoration strategy offered by the proposed DRO in Case 1. Fig. 8 illustrates the load restoration derived by the proposed DRO and the conventional model in the worst case.

In Fig. 8, the x-axis represents attack scenarios. The bar heights depict the load pick-up in each scenario, while the bar widths denote the probability of each scenario. In Scenarios 1, 2, and 3, the conventional method displays less load pick-up. This is because, within the sub-grid of IBRs 680, 633, and 692, only one IBR can securely operate in autonomous mode, thereby limiting the total available IBR capacity under this plan. Conversely, the DRO model offers an alternative sub-grid plan, as previously discussed, enabling more IBRs to operate securely and consequently pick up more load in these scenarios. In Scenario 5, the conventional restoration outperforms by accurately identifying the compromised node. The DRO's load pick-up is slightly less due to the offline status of IBR 692, prioritizing the secure operation of IBR 675. In summary, the conventional method picks up approximately 800 kW less load in Scenarios 1, 2, and 3 compared to the proposed DRO method. These three scenarios collectively account for around 80% probability. Additionally, it picks up 500 kW more load in Scenario 5 compared to the DRO method. However, this scenario only accounts for around 20% probability. Thus, considering the load pickup expectations across all scenarios, the DRO method demonstrates superior overall performance.

Moreover, when there are more local controllers, the probability of attacks on the client nodes is relatively lower. The DRO algorithm will prioritize scenarios where attacks occur on the local controllers, making the restoration strategy tend to resemble that of the conventional algorithm.

B. IEEE 123-node Test Cases

1) Test system: The standard test system is modified to include 7 IBRs and 11 switchable lines (L24, L13, L55, L61, L117, L68, L76, L105, L45, L96, L10). The IBR location, capacity, and the cyber network topology can be referred to Fig. 9. This network is delineated into two clusters, represented by different colors. The cyber system includes two tiers: clients and local controllers, the same as the previous test system. However, the client nodes encompass two types: regular nodes (3, 4, 5, 6, 7) and cluster leads (1, 2). Regular nodes have control capabilities confined to their respective in-cluster IBRs. In contrast, cluster leads possess the ability to manage in-cluster IBRs and coordinate with IBRs outside their cluster. Within this system, historical data reveals the existence of 14 distinct scenarios, each involving an attack on a specific node.

Considering the reported trip of IBR 77, three cases are designed: Case 1 assumes a scenario of random attack behavior. The local controllers, being generally more vulnerable, present a higher probability of being targeted in this case. Case 2 assumes attackers with a more regional focus. Given that regular clients are relatively easier to compromise and the impact surface is substantial, they attract the largest

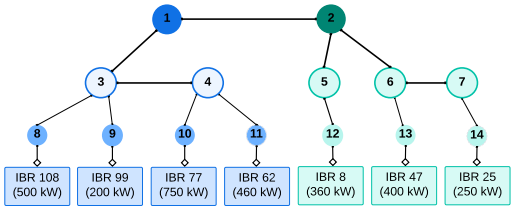

Fig. 9. The cyber network of IEEE 123-node test system

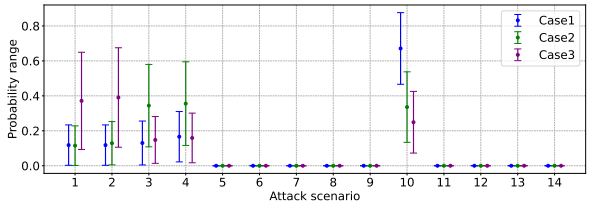

Fig. 10. Probability range of attack scenario in 123-node test system

number of attackers. In Case 3, the attackers adopt a more radical approach, leaning toward targeting cluster leads as they possess the ability to impact a greater number of IBRs.

2) Probability of attack scenario: Fig. 10 illustrates the probability ranges for attack scenarios in each case. As the regular client nodes possess control over in-cluster IBRs, Nodes 5, 6, and 7 cannot trip IBR 77, resulting in a zero probability of attack for these nodes. Furthermore, aside from local controller 10, other local controllers (8, 9, 11, 12, 13, 14) cannot also trip IBR 77. In Case 1, the local controller at Node 10 exhibits the highest probability of attack, ranging between 0.47 and 0.88. Each of Nodes 1, 2, 3, and 4 presents a relatively even probability, approximately ranging from 0 to 0.2. In Case 2, regular clients at Nodes 3 and 4 exhibit considerable probabilities, along with the local controller at Node 10, ranging from approximately 0.1 to 0.6. The probability of attack for cluster leads 1 and 2 remains relatively low, peaking at less than 0.3. In Case 3, the cluster leads display the highest probability of being targeted, ranging from 0.1 to 0.7. Conversely, the probability of local controller attack is notably lower in this case, approximately ranging from 0.1 to 0.4. Nodes 3 and 4 depict the lowest probability of being attacked within this case.

3) Restoration result: Table III outlines the load restoration under each scenario across three cases, with the worst-case probability labeled accordingly. In Case 1, Scenario 10 dominates in the worst case. Hence, the grid deployed the strategy to divide the main grid into two sub-grids to maximize load pick-up. In Case 2, Scenarios 2, 3, 4, and 10 exhibit relatively equal probabilities, with each around 0.25. In response, the DRO model strategically divides the grid into three regions, ensuring optimal load restoration across these scenarios. It is noticed that compared to Case 1, this plan significantly enhances load pick-up in Scenario 2 particularly. For Case 3, Scenarios 1 and 2 take precedence. Here, the DRO model divides the main grid into three sub-grids, optimizing load restoration plans for Scenarios 1 and 2. This sub-grid strategy differs from Case 2, focusing on enhancing load restoration

TABLE III
LOAD PICK-UP IN IEEE 123-NODE TEST CASES (WORST CASE) / KW

Scenario	Case1	Case2	Case3
(Opened lines)	L55	L24, L55	L55, L45
1	3470 (0.027)	3430 (0.026)	3365 (**0.343**)
2	3195 (0.234)	3315 (**0.253**)	3365 (**0.374**)
3	3470 (0.046)	3430 (**0.222**)	3365 (0.069)
4	3470 (0.070)	3430 (**0.233**)	3365 (0.107)
10	3470 (**0.623**)	3430 (**0.267**)	3365 (0.107)
Expectation	3405.65	3400.91	3365

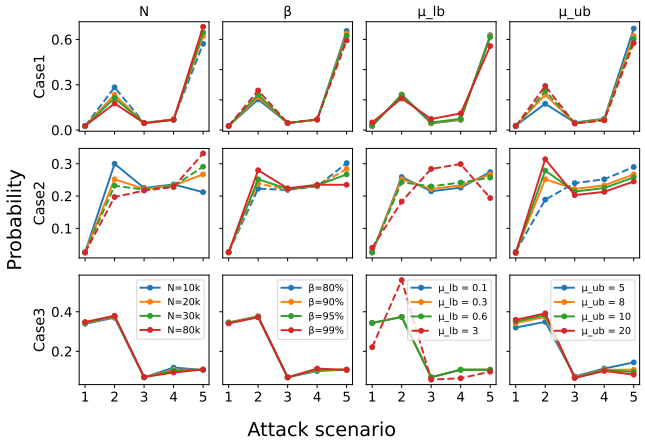

Fig. 11. Parameter sensitivity analysis in IEEE 123-node test cases

specifically in Scenarios 1 and 2.

Overall, when client nodes carry a higher probability of being targeted and potentially impacting more IBRs, the DRO model tends to segment the grid into more sub-grids, mitigating their impact within smaller areas. Nevertheless, this segmentation restricts the transfer of power within the grid, leading to lower load pick-up expectations in comparison to cases with fewer sub-grids.

The computational time for this test system varies from 20 to 50 seconds among different cases. In the previous 13-node test system, the computational time is around 2 seconds across all cases. This data confirms the efficacy of the proposed algorithm for rapid system restoration.

4) Sensitivity analysis: Fig. 11 illustrates the sensitivity of the ambiguity set parameters (confidence level β, data size N, and the range of regional index $[\underline{\mu}, \overline{\mu}]$) in all cases. The default parameters are as follows: $N = 20k, \beta = 95\%, \mu \in [0.3, 8]$. In this analysis, only one parameter changes while others remain at their default values. The x-axis in this figure represents the attack scenario, while the y-axis indicates the probability in the worst cases. Each row signifies a specific case, whereas each column denotes the influence of a particular parameter across all cases. The dashed line indicates deviations in the sub-grid plan compared to the benchmark case.

It's noteworthy that an increase in N results in a contraction of the ambiguity set, leading to an "improved" worst case. This can be translated to a reduced probability of client nodes being attacked, potentially impacting fewer IBRs. It is observed that Cases 1 and 2 are more sensitive to this parameter. Case 3 exhibits robustness to changes in this parameter.

On the other hand, an increase in β leads to an extension of the ambiguity set, resulting in a "worse" worst case. Cases 1 and 2 continue to display high sensitivity to β. Contrarily, Case 3 shows negligible sensitivity to this parameter.

When $\underline{\mu}$ varies, notable sensitivity to this parameter is observed in Cases 2 and 3, particularly in Case 3. This is attributed to the impact of higher rational levels when client nodes are compromised, which compels attackers to target larger IBRs. However, this effect doesn't impact local controllers, as their accessibility is confined to the local IBR. Consequently, Case 1 exhibits limited sensitivity to this parameter. On the contrary, when $\overline{\mu}$ fluctuates, Cases 1 and 2 show considerable sensitivity. The increase in $\overline{\mu}$ shifts the probability from local controllers to client nodes, resulting in "worse" worst-case scenarios. In Case 3, where the most probable compromised nodes are assumed to be cluster leads, this scenario inherently represents the "worst" worst-case situation. Consequently, as $\overline{\mu}$ increases, there's no space for the worst-case scenario to deteriorate further. This results in minimal variations in the probability distribution in Case 3.

In summary, each case exhibits unique sensitivity to these parameters. Thus, the variations in these parameters may necessitate corresponding adjustments to the sub-grid strategy.

VII. Conclusion

This paper introduces a distributionally robust optimization model designed to tackle the restoration problem utilizing inverter-based resources following cyber attacks. The model first delineates the operational guidelines for IBR amidst the compromised cyber network. Then, an ambiguity set is constructed using Bayesian inference to encompass potential attack scenarios and their likelihoods. Next, the DRO model is formulated with advanced transformation techniques for tractable solution algorithms with guaranteed convergence. Simulation results across various test cases and scenarios substantiate the model's efficacy in managing post-attack restoration processes.

This work addresses the post-attack restoration, clarifying essential data to be collected, resources to be coordinated, and the algorithm to be implemented. It provides a comprehensive guideline for post-attack recovery, which can be integrated into the system restoration platform to enable rapid restoration after attacks. This framework can be combined with existing system restoration strategies to tackle increasing cyber-physical coordinated attacks. Additionally, it can be enhanced with advanced detection algorithms in the future. By identifying the attack type and estimating the cyber restoration time, this approach can offer a more effective restoration plan, thus improving the overall cyber-physical resilience of the system.

References

[1] J. Qi, A. Hahn, X. Lu, J. Wang, and C.-C. Liu, "Cybersecurity for distributed energy resources and smart inverters," *IET Cyber-Physical Systems: Theory & Applications*, vol. 1, no. 1, pp. 28–39, 2016.

[2] X. Gao, M. Ali, and W. Sun, "A risk assessment framework for cyber-physical security in distribution grids with grid-edge ders," *Energies*, vol. 17, no. 7, 2024.

[3] W. Liu, J. Zhan, C. Y. Chung, and L. Sun, "Availability assessment based case-sensitive power system restoration strategy," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 1432–1445, 2020.

[4] P. Li, J. Ji, H. Ji, G. Song, C. Wang, and J. Wu, "Self-healing oriented supply restoration method based on the coordination of multiple sops in active distribution networks," *Energy*, vol. 195, p. 116968, 2020.

[5] X. Gao and Z. Chen, "Optimal restoration strategy to enhance the resilience of transmission system under windstorms," in *2020 IEEE Texas Power and Energy Conference (TPEC)*, 2020, pp. 1–6.

[6] W. Liu and F. Ding, "Collaborative distribution system restoration planning and real-time dispatch considering behind-the-meter ders," *IEEE Transactions on Power Systems*, vol. 36, no. 4, pp. 3629–3644, 2021.

[7] O. Boyaci, M. R. Narimani, K. R. Davis, M. Ismail, T. J. Overbye, and E. Serpedin, "Joint detection and localization of stealth false data injection attacks in smart grids using graph neural networks," *IEEE Transactions on Smart Grid*, vol. 13, no. 1, pp. 807–819, 2022.

[8] Z. S. Khalafi, M. Dehghani, A. Khalili, A. Sami, N. Vafamand, and T. Dragičević, "Intrusion detection, measurement correction, and attack localization of pmu networks," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 5, pp. 4697–4706, 2022.

[9] A. Presekal, A. Ştefanov, V. S. Rajkumar, and P. Palensky, "Attack graph model for cyber-physical power systems using hybrid deep learning," *IEEE Transactions on Smart Grid*, vol. 14, no. 5, pp. 4007–4020, 2023.

[10] Y. Lin, J. H. Eto, B. B. Johnson, J. D. Flicker, R. H. Lasseter, H. N. Villegas Pico, G.-S. Seo, B. J. Pierre, and A. Ellis, "Research roadmap on grid-forming inverters," 11 2020.

[11] R. H. Lasseter, Z. Chen, and D. Pattabiraman, "Grid-forming inverters: A critical asset for the power grid," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 8, no. 2, pp. 925–935, 2020.

[12] G. Song, B. Cao, and L. Chang, "Review of grid-forming inverters in support of power system operation," *Chinese Journal of Electrical Engineering*, vol. 8, no. 1, pp. 1–15, 2022.

[13] K. Ahmed, M. Seyedmahmoudian, S. Mekhilef, N. M. Mubarak, and A. Stojcevski, "A review on primary and secondary controls of inverter-interfaced microgrid," *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 5, pp. 969–985, 2021.

[14] I. Zografopoulos, N. D. Hatziargyriou, and C. Konstantinou, "Distributed energy resources cybersecurity outlook: Vulnerabilities, attacks, impacts, and mitigations," *IEEE Systems Journal*, vol. 17, no. 4, pp. 6695–6709, 2023.

[15] T. O. Olowu, S. Dharmasena, A. Hernandez, and A. Sarwat, *Impact Analysis of Cyber Attacks on Smart Grid: A Review and Case Study*. Singapore: Springer Singapore, 2021, pp. 31–51.

[16] C. Chen, J. Wang, and D. Ton, "Modernizing distribution system restoration to achieve grid resiliency against extreme weather events: An integrated solution," *Proceedings of the IEEE*, vol. 105, no. 7, pp. 1267–1288, 2017.

[17] R. Roofegari Nejad and W. Sun, "Distributed load restoration in unbalanced active distribution systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5759–5769, 2019.

[18] A. A. Hagberg, D. A. Schult, and P. J. Swart, "Exploring network structure, dynamics, and function using networkx," in *Proceedings of the 7th Python in Science Conference*, G. Varoquaux, T. Vaught, and J. Millman, Eds., Pasadena, CA USA, 2008, pp. 11 – 15.

[19] H. Rahimian and S. Mehrotra, "Frameworks and Results in Distributionally Robust Optimization," *Open Journal of Mathematical Optimization*, vol. 3, 2022.

[20] R. D. McKelvey and T. R. Palfrey, "Quantal response equilibria for normal form games," *Games and Economic Behavior*, vol. 10, no. 1, pp. 6–38, 1995.

[21] I. Kuwatly, M. Sraj, Z. Al Masri, and H. Artail, "A dynamic honeypot design for intrusion detection," in *The IEEE/ACS International Conference onPervasive Services, 2004. ICPS 2004. Proceedings.*, 2004, pp. 95–104.

[22] P. Mohajerin Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations," *Mathematical Programming*, vol. 171, 2018.

[23] L. Spitzner, *Honeypots: Tracking Hackers*. USA: Addison-Wesley Longman Publishing Co., Inc., 2002.

[24] C. Zhao and Y. Guan, "Data-driven risk-averse two-stage stochastic program with ζ-structure probability metrics," 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:19022668

[25] P. Mohajerin Esfahani, "Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations," *Mathematical Programming*, vol. 171, pp. 115–166, 2018.

[26] B. Zeng and L. Zhao, "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Operations Research Letters*, vol. 41, no. 5, pp. 457–461, 2013.