A Planning Model for Optimal Sizing of Integrated Power and Gas Systems Capturing Frequency Security

Yi Wang, *Member*, *IEEE*, and Goran Strbac, *Member*, *IEEE*

***Abstract*—Large renewable penetration has been witnessed in power systems, resulting in reduced level of system inertia and increasing requirements for frequency response services. There have been plenty of studies developing frequency-constrained operation models for power system security. However, most existing literature only focuses on operational level rather than planning level. To fill this gap, this paper proposes a novel planning model for the optimal sizing problem of integrated power and gas systems, capturing both under and over frequency security requirements. A detailed unit commitment setup considering different ramping rates is incorporated into the planning model to accurately represent the scheduling behavior of each individual generator and accurate inertia calculation. The power importing and exporting behaviors of interconnectors are considered, which can influence the largest loss of generation and demand, accounting for under and over frequency security, respectively. Additionally, a deep learning-based clustering method featured by concurrent and integrated learning is introduced in the planning model to effectively generate representative days. Case studies have been conducted on a coupled 6-bus power and 7-node gas system as well as a 14-bus power and 14-node gas system to verify the effectiveness of the proposed planning model in accurate clustering performance and realistic investment decision making.**

***Index Terms*—Deep learning, frequency security, integrated power and gas systems, optimal sizing, unit commitment.**

NOMENCLATURE

A. Abbreviation

RES	Renewable energy resource.
FR	Frequency response.
UC	Unit commitment.
P2G	Power-to-gas.
WT	Wind turbine.
ES	Energy storage.
PV	Photovoltaic.
IC	Interconnector.
RoCoF	Rate-of-Change-of-Frequency.
q-s-s	Quasi-steady-state.
DEC	Deep embedded clustering.
KL	Kullback-Leibler.
DAE	Deep autoencoder.

B. Indices and Sets

$t \in T$	Index and set of time steps.
$g \in \mathcal{G}$	Index and set of generators.
$w \in \mathcal{W}$	Index and set of generator types.
$k \in \mathcal{ES}$	Index and set of ESs.
$g \in \mathcal{RES}$	Index and set of RESs.
$g \in \mathcal{P}2\mathcal{G}$	Index and set of P2Gs.
$i \in \mathcal{IC}$	Index and set of ICs.
$d \in \mathcal{ED}$	Index and set of demands.
$b \in \mathcal{EB}$	Index and set of electric buses.
$bp \in \mathcal{EL}$	Index and set of transmission lines.
$b \in \mathcal{GB}$	Index and set of gas nodes.
$bp \in \mathcal{GL}$	Index and set of gas pipelines.

C. Parameters

Δt	Time resolution.
$c^{\mathrm{gen}}_{g,w}$	Generation cost of generator g with type w.
$c^{\mathrm{no}}_{g,w}$	No load cost of generator g with type w.
$c^{\mathrm{st}}_{g,w}$	Start-up cost of generator g with type w.
$c^{\mathrm{ic}}_{i,t}$	Electricity cost of IC i at time t.
$c^{\mathrm{co_2}}$	Carbon price or tax.
c^{gw}_{g}	Gas cost of gas well g.
C^{cap}_{w}	Capital cost of generator type w.
AF^{gen}_{w}	Annuity factor of generator type w.
$C^{\mathrm{fix,gen}}_{w}$	Fixed operating and maintenance cost of generator type w.
CI^{w}_{g}	Carbon intensity of generator g with type w.
$P^{\mathrm{ed}}_{d,t}$	Electric load d at time t.
$\overline{P}^{w}_{g}, \underline{P}^{w}_{g}$	Active power output limits of generator g with type w.
$\overline{S}^{w}_{g}$	Apparent power capacity of generator g with type w.
δ^{w}_{g}	Rated power factor of generator g with type w.
$RU^{w}_{g,k}$	Ramp-up rate k of generator g with type w.
$RD^{w}_{g,k}$	Ramp-down rate k of generator g with type w.
$\overline{P}^{\mathrm{es}}_{k}$	Active power capacity of ES k.
$\overline{S}^{\mathrm{es}}_{k}$	Apparent power capacity of ES k.
$\overline{E}^{\mathrm{es}}_{k}$	Energy capacity of ES k.
η^{es}_{k}	Charging/discharging coefficient of ES k.

Manuscript received January 11, 2024; revised April 15, 2024; accepted May 27, 2024. Date of online publication January 10, 2025; date of current version February 28, 2025. This work was supported by the National Grid ESO Project 'Tool for Co-Optimisation of Energy and Frequency-Containment Services (COEF)'.

Y. Wang (corresponding author, email: yi.wang18@imperial.ac.uk) and G. Strbac are with the Department of Electrical and Electronic Engineering, Imperial College London, London, UK.

DOI: 10.17775/CSEEJPES.2024.00240

$\overline{P}_g^{\mathrm{res}}$ Active power capacity of RES g.
$\overline{S}_g^{\mathrm{res}}$ Apparent power capacity of RES g.
δ_g^{res} Rated power factor of RES g.
$\overline{P}_g^{\mathrm{p2g}}$ Active power capacity of P2G g.
η_g^{p2g} Energy coefficient of P2G g.
f_0 Nominal grid frequency.
$\overline{\mathrm{RoCoF}}$ Maximum Rate of Change of Frequency.
$\underline{\Delta f}$ Frequency nadir.
$\overline{R}_g^w$ FR capacity of generator g with type w.
$\overline{R}_k^{\mathrm{es}}$ FR capacity of ES k.
$\overline{P}_i^{\mathrm{ic}}, \underline{P}_i^{\mathrm{ic}}$ Maximum/minimum capacity of IC i.
$\overline{V}, \underline{V}$ Maximum/minimum permissible voltage.
G_{bp} Conductance of transmission line (b,p).
B_{bp} Susceptance of transmission line (b,p).
$\overline{S}_{bp}$ Capacity limit of transmission line (b,p).
$\overline{\rho}_b, \underline{\rho}_b$ Maximum/minimum gas pressure of node b.
$\overline{G}_{bp}, \underline{G}_{bp}$ Capacity limits of gas pipeline (b,p).

D. Variables

$P_{g,t}^w$ Active power output of generator g with type w at time t.
$Q_{g,t}^w$ Reactive power output of generator g with type w at time t.
$y_{g,t}^w$ Commitment state of generator g with type w at time t.
$y_{g,t}^{w,sg}$ 'Start generating' state of generator g with type w at time t.
$y_{g,t}^{w,st}$ Startup variable for generator g with type w at time t.
$y_{g,t}^{w,sd}$ Shutdown variable for generator g with type w at time t.
$G_{g,t}^w$ Gas consumption of generator g with type w at time t.
$P_{i,t}^{\mathrm{ic}}$ Active power output of IC i at time t.
$P_{k,t}^c$ Charging power of ES k at time t.
$P_{k,t}^d$ Discharging power of ES k at time t.
$Q_{k,t}^{\mathrm{es}}$ Reactive power of ES k at time t.
$E_{k,t}^{\mathrm{es}}$ Energy content of ES k at time t.
$u_{k,t}^{\mathrm{es}}$ Binary indicating the status of ES k at time t.
$P_{g,t}^{\mathrm{res}}$ Active power output of RES g at time t.
$Q_{g,t}^{\mathrm{res}}$ Reactive power output of RES g at time t.
$P_{g,t}^{\mathrm{p2g}}$ Power consumption of P2G g at time t.
$G_{g,t}^{\mathrm{p2g}}$ Gas output of P2G g at time t.
H_t System inertia at time t.
P_t^{GL} Largest power infeed loss at time t.
$R_{g,t}^w$ FR of generator g with type w at time t.
$R_{k,t}^{\mathrm{es}}$ FR of ES k at time t.
$V_{b,t}$ Voltage of bus b at time t.
$P_{b,t}^{\mathrm{ex}}$ Active power exchange between bus b and other buses at time t.
$Q_{b,t}^{\mathrm{ex}}$ Reactive power exchange between bus b and other buses at time t.
$P_{bp,t}$ Active power flow of line (b,p) at time t.
$Q_{bp,t}$ Reactive power flow of line (b,p) at time t.
$G_{g,t}^{\mathrm{gw}}$ Gas output of gas well g at time t.
$G_{bp,t}$ Gas flow of pipeline (b,p) at time t.
$\rho_{b,t}$ Gas pressure of node b at time t.

I. Introduction

In response to the low-carbon requirement, power systems are transitioning significantly from synchronous fossil fuel resources to converter-interfaced renewable energy resources (RESs) such as wind and solar [1]. However, the high penetration of RESs can cause the reduced inertia level of future power systems, which increases the risk of frequency instability issues [2]. To ensure frequency security, system operators have to procure various frequency response (FR) services from energy-intensive industries, leading to much research on frequency-constrained power system operation problems. However, the frequency security issues of power systems from the planning perspective have not been well investigated. Furthermore, to accommodate more renewable generation, gas-fired power plants have been widely deployed in power systems, highlighting the interdependence between power and gas systems [3]. In this context, the state variations of natural gas systems will directly impact the operation states of gas-fired power plants and then the system inertia level. To account for this, this paper aims to propose a novel planning model for the optimal sizing of integrated power and gas systems capturing frequency security.

Previous work has intensively investigated the planning problems of power systems from the steady-state perspective. For example, in [4], a carbon-oriented planning model of integrated power and gas systems is proposed to incorporate various low-carbon technologies and high penetration of wind generation; nevertheless, this paper does not consider uncertainties associated with load profiles. In [5], a planning model is proposed for integrated power and heat systems, taking into account investment and operation timescales at both national and local levels. However, similar to [4], this paper assumes perfect information about RES and load uncertainties. To capture the influence of uncertainties on planning decisions, a risk-averse planning model is proposed in [6] for large-scale distribution system expansion, accounting for both reliability and resilience. In [7], a robust and stochastic optimization method is proposed for the planning of integrated power and hydrogen systems, considering hydrogen production and storage technologies. In [8], a stochastic expansion planning model based on carbon emission flow is proposed for the construction of fast-charging stations in integrated power and gas systems. However, the above work only focuses on the steady-state operation of power systems and ignores the influence of transient system dynamics on investment decision-making, which can lead to impractical and insecure optimization results [2]. Given the highly scarce inertia of future power systems, post-fault transient system dynamics should be considered in the planning stage to ensure system frequency security.

To capture the influence of frequency security, there have been numerous studies focused on frequency-constrained power system operation problems. For example, in [9], the analytical formulation of system frequency nadir and the frequency security margin are derived based on the frequency

regulation characteristics of conventional generators and RESs. However, the uncertainties associated with RESs and demand profiles are not considered. In [10], a frequency-constrained stochastic unit commitment (UC) model is developed to co-optimize energy and various frequency ancillary services as well as capture uncertainties associated with RESs via a quantile-based scenario generation method. However, the capabilities of RESs on FR service provisions are not investigated in detail. In [11], a frequency-constrained UC problem based on conservative sparse neural networks is developed to facilitate system reliable operation with high RES penetration. However, the benefits of optimized largest loss are not investigated in this paper. In [12], the participation of grid-forming and grid-following inverter-based technologies in the ancillary service market are considered for frequency control; nevertheless, the flexibility from advanced zero-carbon technologies such as hydroelectricity is not considered. In [13], a transient frequency-constrained two-stage stochastic scheduling model is proposed to investigate the day-ahead operation of thermal-hydro-wind-demand response systems. However, the above papers only focus on power sectors, while the operation flexibility from gas sectors is not considered. To fill this gap, a co-optimization model is proposed in [14] to capture both frequency constraints and gas system operation constraints; nevertheless, this paper only focuses on the short-term operation stage and ignores long-term investment decisions.

With respect to the planning stage, several papers have examined frequency security requirements in investment decisions. In [15], simulation-based studies are conducted for the optimal sizing of batteries in a microgrid capturing frequency security metrics; nevertheless, no mathematical formulations are derived for frequency dynamics. In [16], an investment planning model based on stochastic-robust optimization is developed to design a resilient microgrid, which captures dynamic frequency response during the islanding period. However, this paper focuses on the distribution level rather than the transmission level, while the UC of generators is not considered. In [17], a frequency stability-constrained battery sizing model considering the UC setup of generators is proposed to enhance microgrid controllability, where both the grid-connected and stand-alone operating modes are modeled. However, this paper only considers the optimal sizing of batteries and still focuses on the distribution level, while not accounting for investment decisions at the transmission level. In [18], the mathematical formulation of frequency constraints and transmission network expansion planning is introduced for the economical investment of transmission lines under frequency stability. However, this paper does not consider detailed planning strategies for different resources. In [19], a stochastic planning model capturing frequency dynamics is proposed for the optimal sizing of batteries in power systems; nevertheless, the investment decisions of conventional generators and RESs are not considered. In [20], an expansion planning model capturing frequency security constraints is developed for the optimal sizing of RESs such as wind and solar. However, the investment decisions on conventional generators and batteries are ignored. In [21], a frequency-constrained stochastic planning method is proposed for a power system with high penetration of RESs, capturing the frequency response support from wind farms. However, this paper employs a simplified UC model based on aggregated generation, which may suffer a flexibility inadequacy problem and make it difficult to accurately estimate the number of online generators, leading to inaccurate inertia calculations. Additionally, it is notable that each individual generator might even have different ramp-up/down rates at different generation levels in real-world applications [22], which are ignored in the above frequency-constrained problems and should be considered in the planning stage for accurate decision-making.

Furthermore, the above papers on frequency-constrained planning only consider the power system, while ignoring the influence of gas system operation on power system investment decisions. Nowadays, the number of gas-fired generators has significantly increased, which further increases the dependence of power systems on gas systems. For realistic planning decisions, the operation of integrated power and gas systems should be considered. In the existing literature, only one reference [14] considers both integrated power-gas system operation and frequency security requirements; nevertheless, as aforementioned, this paper focuses on the operational level rather than the planning level, while power-to-gas (P2G) technologies are not considered. Finally, all the existing papers on frequency-constrained planning only focus on dealing with under-frequency security issues, while ignoring the over-frequency security requirements. Note that the number of large-scale interconnections (ICs) has gradually increased in modern power systems [23], leading to requirements for both under and over frequency security. Specifically, the power-importing behaviors of ICs can cause the largest loss of generation and under-frequency issues, while the power-exporting behaviors of ICs can cause the largest loss of demand and over-frequency issues.

From the perspective of data preparation, it is natural to collect data spanning an entire year or multiple years to formulate a system planning problem for realistic decision making [24]. However, directly using all the data profiles for the planning problem can be extremely time-consuming, due to the long scheduling horizon. As such, it is a common practice in existing literature (e.g., [6]) to generate several representative days for the problem formulation instead of using all the data profiles [25]. In recent years, various distance-based clustering approaches have been used to generate representative days, including K-means, hierarchical clustering, etc. [26]. However, these conventional distance-based methods depend on a similarity matrix to classify different profiles, which may not be capable of extracting adequate representative features from the yearly dataset and eventually lead to poor clustering performance. In this context, integrating deep learning techniques into traditional distance-based methods can be insightful for the clustering of the large-scale yearly dataset, due to their powerful capabilities of feature extraction. Specifically, deep learning-based clustering methods can enhance the clustering performance and obtain a more accurate cluster label for each data profile through a better representative similarity matrix [26].

To address the above research gaps, this paper proposes a novel planning model based on deep learning techniques for the optimal sizing of integrated power and gas systems while capturing frequency security. The detailed contributions can be summarized as follows.

1) A novel planning model capturing both under and over frequency security constraints is proposed for integrated power and gas systems. A detailed UC model considering different ramp-up/down rates is developed to capture the scheduling behaviors of each individual generator, while both gas-fired power plants and P2G technologies are included in the model.

2) The impacts of the power importing and exporting behaviors of large-scale ICs are considered in the proposed planning model, which are linked with the largest loss of generation and demand, accounting for both under and over frequency security requirements.

3) A deep embedded clustering (DEC) method based on deep learning is introduced in the planning model to effectively extract features of the yearly dataset and generate representative days, increasing the use of important attributes encompassed in the dataset and ensuring a dense distribution of data points within each cluster.

Extensive case studies have been developed to verify the performance of the proposed DEC method in generating representative days as well as the performance of the developed planning model in accurate and secure optimal sizing decision-making across various low-carbon technologies. Additionally, frequency response support from wind turbines has been considered in the case studies, which can lead to investment cost savings. The rest of this paper is organized as follows. Section II provides an overview of the proposed planning model, while Section III details the mathematical formulations of the proposed planning model. Then, the developed linearization and data preparation methods, which are based on deep learning techniques, for solving the frequency-constrained planning problem are presented in Section IV. Section V presents the case studies involving different network structures. Finally, conclusions and future extensions of this paper are drawn in Section VI.

II. Outline of the Proposed Planning Model

A. Problem Descriptions

This paper focuses on the optimal sizing problem of an integrated power and gas system with the consideration of post-fault frequency security. As depicted in Fig. 1, both gas-fired power plants and P2G technologies are considered for the bidirectional interactions between power and gas systems. Specifically, gas-fired plants supply electricity as a source to the power system and consume natural gas as a demand in the gas system, while P2G plants can be realized as an indispensable medium transforming the surplus wind power into synthetic natural gas for the gas system [27]. Except for the coupling components, the sizing and positioning of wind turbines (WTs), photovoltaics (PVs), and energy storage systems (ESs) are also considered at the planning stage. To ensure secure system operations, post-fault transient system dynamics including both under and over frequency security requirements are captured in planning decisions, where the power importing and exporting behaviors of ICs can influence the largest loss of generation and demand, respectively.

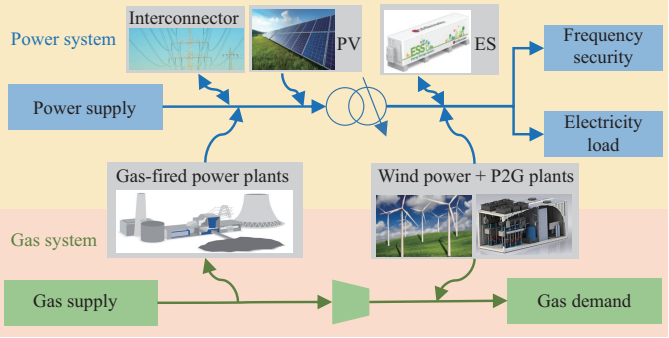

Fig. 1. Illustration of the proposed planning model for the integrated power and gas systems capturing frequency security.

B. Research Challenges and Solutions

To achieve the accurate and efficient frequency-constrained planning decisons of integrated power and gas systems, three main challenges need to be appropriately addressed:

1) How to respect the reality of the low-level operation model and ensure accurate optimization results? To reach this target, a detailed UC model allowing complete parameters and different ramp-up/down rates is required to capture the scheduling behaviors of each individual generator for accurate decision making and inertia calculation.

2) How to make cost-effective investment decisions, while ensuring post-fault frequency security? To reach this target, detailed frequency security constraints are required in the planning model, accounting for both under and over frequency issues, e.g., frequency Nadir and Rate-of-Change-of-Frequency (RoCoF), and quasi-steady-state (q-s-s) constraints.

3) How to ensure the computing efficiency of the proposed planning model while maintaining the decision reality? To reach this target, a deep learning-based clustering method can be adopted to improve the feature extraction capability of traditional distance-based clustering methods, generating more realistic representative days.

III. General Mathematical Formulations of the Proposed Planning Model

The planning model should take into account various operational constraints, with the objective of minimizing the whole system cost of the integrated power and gas system while meeting frequency security requirements. Overall, there are three sets of constraints that should be considered, including the UC constraints of conventional generators, frequency-related constraints, and the energy flow constraints of the integrated power and gas system. Detailed mathematical formulations can be found below.

A. Resource Model

1) Detailed Generator Model

The generation and ramp-up/down limits of generators are modeled in constraints (1)–(7). The formulation has two

distinguishing features compared with the simplified UC used in existing literature:

- Existing literature on power system planning uses aggregated generation or simplified UC for synchronous generators, which cannot accurately model the schedules and inertia provision from each generator [21]. Compared to this, the proposed planning model is capable of capturing the scheduling behaviors of each individual generator, leading to accurate inertia calculation.
- Commonly-used UC assumes the same ramp-up/down rate for each generator, which may not be realistic [22]. In reality, each generator may have up to three different ramping rates, which brings much larger flexibility, as depicted in Fig. 2. Compared to existing literature, the proposed planning approach accurately models the different ramp-up/down rates for each generator.

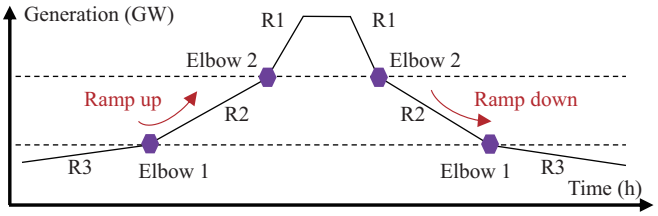

Fig. 2. Illustration of one generator with three different ramping rates.

In detail, the scheduling behaviors of each individual generator can be modeled as:

$$y^w_{g,t}\underline{P}^w_g \le P^w_{g,t} \le y^w_{g,t}\overline{P}^w_g,\ \forall w\in\mathcal{W},\ \forall g\in\mathcal{G},\ \forall t\in T \quad (1)$$

$$P^w_{g,t}-P^w_{g,t-1}\le RU^w_{g,k}y^w_{g,t}, if\ P^w_{g,t-1}\in[\underline{P}^{ru,w}_{g,k},\overline{P}^{ru,w}_{g,k})$$
$$\forall w\in\mathcal{W},\ \forall g\in\mathcal{G},\ \forall k\in\mathcal{RU}_g,\ \forall t\in T \quad (2)$$

$$P^w_{g,t-1}-P^w_{g,t}\le RD^w_{g,k}y^w_{g,t}, if\ P^w_{g,t-1}\in[\underline{P}^{rd,w}_{g,k},\overline{P}^{rd,w}_{g,k})$$
$$\forall w\in\mathcal{W},\ \forall g\in\mathcal{G},\ \forall k\in\mathcal{RD}_g,\ \forall t\in T \quad (3)$$

$$y^w_{g,t}=y^w_{g,t-1}+y^{w,sg}_{g,t-1}-y^{w,sd}_{g,t},\ \forall w\in\mathcal{W},\ \forall g\in\mathcal{G},\ \forall t\in T \quad (4)$$

$$y^{w,sg}_{g,t}=y^{w,st}_{g,t-T^{\rm st}_g},\ \forall w\in\mathcal{W},\ \forall g\in\mathcal{G},\ \forall t\in T \quad (5)$$

$$y^{w,st}_{g,t}\le(1-y^w_{g,t-1})-\sum_{\tau=t-T^d_g}^{t}y^{w,sd}_{g,\tau}$$
$$\forall w\in\mathcal{W},\ \forall g\in\mathcal{G},\ \forall t\in T, \quad (6)$$

$$y^{w,sd}_{g,t}\le y^w_{g,t-1}-\sum_{\tau=t-T^u_g}^{t}y^{w,sg}_{g,\tau},\ \forall w\in\mathcal{W},\ \forall g\in\mathcal{G},\ \forall t\in T \quad (7)$$

$$(P^w_{g,t})^2+(Q^w_{g,t})^2\le(\overline{S}^w_g)^2,\ \forall w\in\mathcal{W},\ \forall g\in\mathcal{G},\ \forall t\in T \quad (8)$$

$$|Q^w_{g,t}|\le P^w_{g,t}\tan(\cos^{-1}\delta^w_g),\ \forall w\in\mathcal{W},\ \forall g\in\mathcal{G},\ \forall t\in T \quad (9)$$

$$P^w_{g,t}=G^w_{g,t}/\eta^w_g,\ \forall w\in\mathcal{W},\ \forall g\in\mathcal{G},\ \forall t\in T \quad (10)$$

$$G^{\rm cap}_w\ge\sum_{g\in\mathcal{G}}y^w_{g,t}\overline{P}^w_g,\ \forall w\in\mathcal{W},\ \forall t\in T \quad (11)$$

where (1) corresponds to the output limits of generator g with type w. Constraints (2)–(3) refer to the ramp-up and ramp-down limits of generator g with type w at a specific generation range, respectively. In other words, the ramp-up/down rates can change for each time step t, depending on the current generation level [22]. For example, when the generation level $P^w_{g,t-1}$ at time step $t-1$ is within the range of $[\underline{P}^{ru,w}_{g,k},\overline{P}^{ru,w}_{g,k})$, the k^{th} ramping rate $RU^w_{g,k}$ will be used to limit the ramp-up behaviors of generator g with type w at time step t. Accordingly, when the generation level moves to another range, a different ramping rate will be selected for constraints (2)–(3). Note that conditional constraints (2)–(3) include non-linear features and can be linearized by introducing auxiliary binary variables [28], where the detailed procedure can be found in Section IV. Constraint (4) describes that the commitment state of generator g at time step t is 'on' if it was generating at time step $t-1$ or has started generating, unless it has been shut down at time step t. Constraint (5) shows that generator g starts generating after it started up for '$T^{\rm st}_g$' hours. Constraints (6)–(7) correspond to the conditions of start-up and shut-down of generator g with type w, respectively. Additionally, constraints (8)–(9) limit the reactive power of generator g with type w, where δ^w_g is its rated power factor [29]. The energy conversion between gas consumption and power generation of generator g is expressed in (10), where η^w_g is the conversion coefficient. Furthermore, constraint (11) calculates the required generation capacity of type w, which will be utilized for the objective function construction (50) of the planning model.

2) Battery Model

ESs in power system are modeled as:

$$0\le P^c_{k,t}\le u^{\rm es}_{k,t}\overline{P}^{\rm es}_k,\ \forall k\in\mathcal{ES},\ \forall t\in T \quad (12)$$

$$(u^{\rm es}_{k,t}-1)\overline{P}^{\rm es}_k\le P^d_{k,t}\le 0,\ \forall k\in\mathcal{ES},\ \forall t\in T \quad (13)$$

$$(P^d_{k,t}+P^c_{k,t})^2+(Q^{\rm es}_{k,t})^2\le\overline{S}^{\rm es}_k,\ \forall k\in\mathcal{ES},\ \forall t\in T \quad (14)$$

$$\underline{E}^{\rm es}_k\le E^{\rm es}_{k,t}\le\overline{E}^{\rm es}_k,\ \forall k\in\mathcal{ES},\ \forall t\in T \quad (15)$$

$$E^{\rm es}_{k,t+1}=E^{\rm es}_{k,t}+(P^c_{k,t}\eta^{\rm es}_k+P^d_{k,t}/\eta^{\rm es}_k)\Delta t$$
$$\forall k\in\mathcal{ES},\ \forall t\in T \quad (16)$$

$$ES^{\rm cap}_k\ge\overline{P}^{\rm es}_k,\ \forall k\in\mathcal{ES} \quad (17)$$

where constraints (12)–(16) limit the charging/discharging behaviors and energy content of ES k. $u^{\rm es}_{k,t}$ indicates the charging/discharging status of ES k at time step t ($u^{\rm es}_{k,t}=1$ if charging; 0, otherwise). Constraint (17) calculates the required power capacity of ES k.

3) RES and P2G Model

The operating model of RESs (e.g., WTs and PVs) and P2G plants can be expressed as:

$$0\le P^{\rm res}_{g,t}+P^{\rm p2g}_{g,t}\le\alpha_{g,t}\overline{P}^{\rm res}_g,\ \forall g\in\mathcal{RES},\ \forall t\in T \quad (18)$$

$$(P^{\rm res}_{g,t})^2+(Q^{\rm res}_{g,t})^2\le(\overline{S}^{\rm res}_g)^2,\ \forall g\in\mathcal{RES},\ \forall t\in T \quad (19)$$

$$|Q^{\rm res}_{g,t}|\le P^{\rm res}_{g,t}\tan(\cos^{-1}\delta^{\rm res}_g),\ \forall g\in\mathcal{RES},\ \forall t\in T \quad (20)$$

$$P^{\rm p2g}_{g,t}\le\overline{P}^{\rm p2g}_g,\ \forall g\in\mathcal{P2G},\ \forall t\in T \quad (21)$$

$$P^{\rm p2g}_{g,t}=G^{\rm p2g}_{g,t}/\eta^{\rm p2g}_g,\ \forall g\in\mathcal{P2G},\ \forall t\in T \quad (22)$$

$$RES^{\rm cap}_g\ge\overline{P}^{\rm res}_g,\ \forall g\in\mathcal{RES} \quad (23)$$

$$P2G^{\rm cap}_g\ge\overline{P}^{\rm p2g}_g,\ \forall g\in\mathcal{P2G} \quad (24)$$

where constraint (18) limits the active power output of P2G g and RES g at time step t, which is also affected by weather

conditions (e.g., wind speed) contained in $\alpha_{g,t}.P_{g,t}^{\mathrm{res}}$ and $P_{g,t}^{\mathrm{p2g}}$ correspond to the RES generation and P2G consumption at time step t, where the P2G plant can provide synthetic natural gas for the gas system. Constraints (19)–(20) limit the apparent power of RES g, taking into account both the nameplate capacity $\overline{S}_g^{\mathrm{res}}$ and rated power factor δ_g^{res} [29]. Constraint (21) limits the power consumption of P2G plant g, while the energy conversion is expressed in (22) with coefficient η_g^{p2g}. Finally, constraints (23)–(24) calculate the required generation capacity of RES g and P2G g, respectively.

B. Frequency Security Constraints

To maintain post-fault system security, frequency-related constraints should be included in the planning model. Specifically, the proposed frequency-related constraints outlined below have two distinguishing features compared with existing literature:

- Existing literature on frequency-constrained UC only focuses on under-frequency security, while the over-frequency security is ignored [10]. Compared to this, the proposed planning model is capable of capturing both under and over frequency security requirements, including two sets of frequency Nadir, RoCoF, and q-s-s constraints for low and high frequency issues, respectively.
- Existing literature only models the influence of thermal generators on the largest generation loss. In contrast, the proposed planning approach models both the largest loss of generation and demand, which can be influenced by the power importing and exporting behaviors of ICs, respectively.

In detail, the frequency-related constraints are expressed as:

$$H_t = \sum_{g\in\mathcal{G}}\sum_{g\in\mathcal{W}} H_g^w \overline{P}_g^w y_{g,t}^w - z_t^{\mathrm{GL}} H^{\mathrm{GL}} P_t^{\mathrm{GL}},\ \forall t\in T \tag{25}$$

$$z_t^{\mathrm{GL}} = 0,\ \text{if } P_t^{\mathrm{GL}} = P_{i,t}^{\mathrm{ic}},\ \forall i\in\mathcal{IC},\ \forall t\in T \tag{26}$$

$$(P_t^{\mathrm{GL}} f_0)/(2H_t) \le \overline{\mathrm{RoCoF}},\ \forall t\in T \tag{27}$$

$$\sum_{g\in\mathcal{G}}\sum_{w\in\mathcal{W}} R_{g,t}^w + \sum_{k\in\mathcal{ES}} R_{k,t}^{\mathrm{es}} \ge P_t^{\mathrm{GL}},\ \forall t\in T \tag{28}$$

$$\left(\frac{H_t}{f_0} - \frac{\sum_{k\in\mathcal{ES}} R_{k,t}^{\mathrm{es}} T^{\mathrm{es}}}{4\Delta\overline{f}}\right)\frac{\sum_{g\in\mathcal{G}}\sum_{w\in\mathcal{W}} R_{g,t}^w}{T^{\mathrm{gen}}} \ge \frac{(P_t^{\mathrm{GL}} - \sum_{k\in\mathcal{ES}} R_{k,t}^{\mathrm{es}})^2 T^{\mathrm{gen}}}{4\Delta\overline{f}},\ \forall t\in T \tag{29}$$

$$0 \le R_{g,t}^w \le y_{g,t}^w \overline{P}_g^w - P_{g,t}^w,\ \forall w\in\mathcal{W},\ \forall g\in\mathcal{G},\ \forall t\in T \tag{30}$$

$$0 \le R_{g,t}^w \le \overline{R}_g^w,\ \forall w\in\mathcal{W},\ \forall g\in\mathcal{G},\ \forall t\in T \tag{31}$$

$$0 \le R_{k,t}^{\mathrm{es}} \le \overline{R}_k^{\mathrm{es}},\ \forall k\in\mathcal{ES},\ \forall t\in T \tag{32}$$

$$0 \le R_{k,t}^{\mathrm{es}} \le (1-u_{k,t}^{\mathrm{es}})\overline{P}_k^{\mathrm{es}} + P_{k,t}^d + P_{k,t}^c,\ \forall k\in\mathcal{ES},\ \forall t\in T \tag{33}$$

where the system inertia H_t is calculated in (25) by aggregating the inertia of generators, except for the inertia from the largest power infeed loss P_t^{GL} at time step t. Note that there is no inertia loss ($z_t^{\mathrm{GL}} = 0$) if the largest power infeed is from the IC i, which can be captured by the conditional constraint (26). The RoCoF limit is guaranteed by (27), while the q-s-s security is assured in (28), where f_0 and $\overline{\mathrm{RoCoF}}$ are the nominal grid frequency and the maximum admissible RoCoF, respectively [10]. Note that constraint (28) is used to ensure the sum of FR is greater than the largest loss so that frequency drop will be finally arrested [30]. The frequency nadir limit is guaranteed by (29), where $\Delta\overline{f}$ is the maximum admissible frequency deviation. $T^{\mathrm{es}}, T^{\mathrm{gen}}$ correspond to the FR delivery speed from ESs and generators, respectively. The FR quantity of generator g is limited by the headroom in (30) and FR capacity in (31). The FR quantity from ES k is limited by (32) and (33), where $\overline{R}_k^{\mathrm{es}}$ corresponds to the FR capacity of ES k. Regarding over-frequency security, constraint set (25)–(33) can be duplicated to account for the largest loss of demand P_t^{LL}, while the FR services from generators and ESs for over-frequency security will be limited by the footroom rather than headroom. For notational simplicity, frequency constraints accounting for over-frequency issues are omitted.

Finally, by considering the influence of ICs on frequency security, the largest loss of generation and demand can be respectively modeled as:

$$P_t^{\mathrm{GL}} \ge \max[P_{i,t}^{\mathrm{ic}}, P_{g,t}^w],\ \forall i\in\mathcal{IC},\ \forall g\in\mathcal{G},\ \forall t\in T \tag{34}$$

$$P_t^{LL} \ge \max[-P_{i,t}^{\mathrm{ic}}, P_{d,t}^{\mathrm{ed}}],\ \forall i\in\mathcal{IC},\ \forall d\in\mathcal{ED},\ \forall t\in T \tag{35}$$

$$\underline{P}_i^{\mathrm{ic}} \le P_{i,t}^{\mathrm{ic}} \le \overline{P}_i^{\mathrm{ic}},\ \forall i\in\mathcal{IC},\ \forall t\in T \tag{36}$$

where constraints (34)–(35) shows that the largest loss of generation or demand should not be lower than the power importing or exporting of the IC i at time step t, respectively. $\mathcal{ED}$ refers to the set of loads in the power system, while the power output of IC i is limited by (36).

C. Integrated Power and Gas Operation

1) Power System Operation

The power system operation is modeled by a linearized AC optimal power flow [25] capturing voltage and power losses, which can be expressed as:

$$\sum_{i\in B_{\mathrm{ic}}} P_{i,t}^{\mathrm{ic}} + \sum_{g\in B_{\mathrm{eg}}}\sum_{w\in\mathcal{W}} P_{g,t}^w + \sum_{g\in B_{\mathrm{res}}} P_{g,t}^{\mathrm{res}} = P_{b,t}^{\mathrm{ex}} + \sum_{d\in B_{\mathrm{ed}}} P_{d,t}^{\mathrm{ed}} + \sum_{k\in B_{\mathrm{es}}}(P_{k,t}^c - P_{k,t}^d),\ \forall b\in\mathcal{EB},\ \forall t\in T \tag{37}$$

$$\sum_{g\in B_{\mathrm{eg}}}\sum_{w\in\mathcal{W}} Q_{g,t}^w + \sum_{g\in B_{\mathrm{res}}} Q_{g,t}^{\mathrm{res}} = Q_{b,t}^{\mathrm{ex}} + \sum_{d\in B_{\mathrm{ed}}} Q_{d,t}^{\mathrm{ed}} + \sum_{k\in B_{\mathrm{es}}} Q_{k,t}^{\mathrm{es}},\ \forall b\in\mathcal{EB},\ \forall t\in T \tag{38}$$

$$P_{b,t}^{\mathrm{ex}} = \sum_{bp\in\mathcal{EL}} P_{bp,t} + \left(\sum_{p\in\mathcal{EB}} G_{bp}\right)V_{b,t}^2,\ \forall b\in\mathcal{EB},\ \forall t\in T \tag{39}$$

$$Q_{b,t}^{\mathrm{ex}} = \sum_{bp\in\mathcal{EL}} Q_{bp,t} - \left(\sum_{p\in\mathcal{EB}} B_{bp}\right)V_{b,t}^2,\ \forall b\in\mathcal{EB},\ \forall t\in T \tag{40}$$

$$P_{bp,t} = G_{bp}(V_{b,t}^2 - V_{p,t}^2)/2 - B_{bp}\delta_{bp,t} + P_{bp,t}^{lo},\ \forall bp\in\mathcal{EL},\ \forall t\in T \tag{41}$$

$$Q_{bp,t} = -B_{bp}(V_{b,t}^2 - V_{p,t}^2)/2 - G_{bp}\delta_{bp,t} + Q_{bp,t}^{lo},\ \forall bp\in\mathcal{EL},\ \forall t\in T \tag{42}$$

$$\underline{V}^2 \le V_{b,t}^2 \le \overline{V}^2,\ \forall b \in \mathcal{EB},\ \forall t \in T \tag{43}$$

$$P_{bp,t}^2 + Q_{bp,t}^2 \le \overline{S}_{bp}^2,\ \forall bp \in \mathcal{EL},\ \forall t \in T \tag{44}$$

where constraints (37)–(38) indicate the active and reactive power balances at bus b. Sets $B_{\rm ic}$, $B_{\rm eg}$, $B_{\rm ed}$, $B_{\rm res}$, and $B_{\rm es}$ correspond to the IC i, generator g, load d, RES g, and ES k connected with bus b, respectively. Nodal power exchanges $P_{b,t}^{\rm ex}, Q_{b,t}^{\rm ex}$ at bus b can be linearized by (39)–(40). The active and reactive power flows $P_{bp,t}, Q_{bp,t}$ of line $b-p$ are expressed in (41)–(42), where $P_{bp,t}^{lo}, Q_{bp,t}^{lo}$ correspond to active and reactive power losses that are linearized by loss factors [31]. The nodal voltage level is limited by (43), while the thermal capacity of line $b-p$ is limited by (44).

2) Gas System Operation

The gas system operation is modeled by steady gas flow [32]. Since this paper mainly focuses on developing a planning model to maintain both under and over frequency security of the power sector, the dynamic behaviors of gas flow are not considered. In addition, using the steady-state gas flow model is suitable for high-pressure transmission pipelines [33], meeting the requirements of the proposed planning model on solution quality. Furthermore, the steady-state gas flow model could significantly reduce the computation burden and then contribute to model simplicity and computing efficiency [33]. Specifically, the gas system operation can be realized as:

$$\sum_{g\in B_{\rm gw}} G_{g,t}^{\rm gw} + \sum_{g\in B_{\rm p2g}} G_{g,t}^{\rm p2g} = \sum_{d\in B_{\rm gd}} G_{d,t}^{gd} + \sum_{g\in B_{\rm eg}} G_{g,t}^{w} + \sum_{pb\in\mathcal{GL}} G_{\rm pb,t} - \sum_{bp\in\mathcal{GL}} G_{bp,t},\ \forall b \in \mathcal{GB},\ \forall t \in T \tag{45}$$

$$G_{bp,t}^2 = sign(G_{bp,t})\Pi_{bp}^2(\rho_{b,t} - \rho_{p,t})\quad \forall bp \in \mathcal{GL},\ \forall t \in T \tag{46}$$

$$\underline{\rho}_b \le \rho_{b,t} \le \overline{\rho}_b,\ \forall b \in \mathcal{GB},\ \forall t \in T \tag{47}$$

$$\underline{G}_{bp} \le G_{bp,t} \le \overline{G}_{bp},\ \forall bp \in \mathcal{GL},\ \forall t \in T \tag{48}$$

where the gas flow balance of node b at time step t is expressed in (45). $B_{\rm gw}$, $B_{\rm gd}$, $B_{\rm eg}$, and $B_{\rm p2g}$ refer to the sets of gas wells (GWs), loads, gas-fired power plants, and P2G plants located at node b, respectively. The relation between gas pressure and gas flow at pipeline $b-p$ is represented by (46), given the Weymouth constant Π_{bp}. Nodal gas pressure $\rho_{b,t}$ and gas flow $G_{bp,t}$ at time step t are limited by (47)–(48), respectively.

3) Carbon Target

In response to the low-carbon transition of future power systems, the overall carbon emissions in the power system should not exceed the regulated limit of carbon emissions, which can be expressed as the product of the annual energy consumption and the carbon target ($\overline{CO_2}$) [5]. The carbon constraint can be expressed as:

$$\sum_{w\in\mathcal{W}}\sum_{g\in\mathcal{G}}\sum_{t\in T} P_{g,t}^w CI_g^w \le \overline{CO_2} \sum_{d\in\mathcal{ED}}\sum_{t\in T} P_{d,t}^{\rm ed} \tag{49}$$

where CI_g^w corresponds to the carbon intensity factor of generator g with type w.

D. Objective Function

The objective function includes both annuitized infrastructure investment costs and annual system operation costs, which can be expressed as:

$$\begin{aligned}\min \sum_{w\in\mathcal{W}} & G_w^{\rm cap}(C_w^{\rm gen} AF_w^{\rm gen} + C_w^{\rm fix,gen}) \\ & + \sum_{g\in\mathcal{RES}} RES_g^{\rm cap}(C_g^{\rm res} AF_g^{\rm res} + C_g^{\rm fix,res}) \\ & + \sum_{g\in\mathcal{P}2\mathcal{G}} P2G_g^{\rm cap}(C_g^{\rm p2g} AF_g^{\rm p2g} + C_g^{fix,p2g}) \\ & + \sum_{k\in\mathcal{ES}} ES_k^{\rm cap}(C_k^{\rm es} AF_k^{\rm es} + C_k^{\rm fix,es}) + F^{op}\end{aligned} \tag{50}$$

where the investment costs include capital costs of generators, RESs (WTs and PVs), P2G plants, and ESs. Annuity factors (e.g., $AF_w^{\rm gen}$) are considered for different assets according to their discount rates and life spans, while $C_w^{\rm fix,gen}$ corresponds to the fixed operating and maintenance cost of generator type w [5]. The annual operation costs F^{op} of the integrated power and gas system are defined as the operating cost of generators and gas wells, which can be written as:

$$\begin{aligned}F^{op} = & \sum_{w\in\mathcal{W}}\sum_{g\in\mathcal{G}}\sum_{t\in T}(c_{g,w}^{\rm gen} P_{g,t}^w + \lambda^{\rm co_2} CI_g^w P_{g,t}^w) \\ & + \sum_{w\in\mathcal{W}}\sum_{g\in\mathcal{G}}\sum_{t\in T}(c_{g,w}^{\rm no} y_{g,t}^w + c_{g,w}^{\rm st} y_{g,t}^{w,st}) \\ & + \sum_{i\in\mathcal{IC}}\sum_{t\in T} c_{i,t}^{\rm ic} P_{i,t}^{\rm ic} + \sum_{g\in\mathcal{GW}}\sum_{t\in T} c_g^{\rm gw} G_{g,t}^{\rm gw}\end{aligned} \tag{51}$$

where $c_{g,w}^{\rm gen}$, $c_{g,w}^{\rm st}$, and $c_{g,w}^{\rm no}$ correspond to the marginal generation cost, start-up cost, and no load cost of generator g with type w, respectively. $\lambda^{\rm co_2}$ indicates the carbon price/tax. $c_{i,t}^{\rm ic}$ indicates the electricity price of IC i at time step t, while $c_g^{\rm gw}$ is the gas cost of gas well g.

IV. Linearization and Data Preparation

A. Linearization of Conditional Constraints

Equations (2), (3), and (26) belong to conditional constraints that are only enforced if their associated conditions are met. To make the optimization solvable, auxiliary binary variables can be introduced for the linearization of these constraints [28].

Take the ramp-up constraint (2) as an example, if the generator g with type w has three different ramp-up rates (e.g., R_1, R_2, and R_3) and two elbows (e.g., E_1 and E_2), the conditional constraint (2) can be rewritten as:

$$\begin{cases} P_{g,t}^w - P_{g,t-1}^w \le R_1 \cdot y_{g,t}^w, & \text{if } P_{g,t-1}^w \in [0, E_1) \\ P_{g,t}^w - P_{g,t-1}^w \le R_2 \cdot y_{g,t}^w, & \text{if } P_{g,t-1}^w \in [E_1, E_2) \\ P_{g,t}^w - P_{g,t-1}^w \le R_3 \cdot y_{g,t}^w, & \text{if } P_{g,t-1}^w \in [E_2, \overline{P}_{g,t}^w] \end{cases} \tag{52}$$

where $E_1 < E_2 < \overline{P}_{g,t}^w$. To linearize the conditional constraint (52), binary variables $z_{g,t-1}^{1,w}$, $z_{g,t-1}^{2,w}$, $z_{g,t-1}^{3,w}$, and $z_{g,t-1}^{4,w}$ are introduced to reformulate (52) as:

$$P_{g,t-1}^w + M \cdot z_{g,t-1}^{1,w} \ge E_1$$

$$P_{g,t}^w - P_{g,t-1}^w - M \cdot (1 - z_{g,t-1}^{1,w}) \le R_1 \cdot y_{g,t}^w$$

$$
\begin{aligned}
&P^{w}_{g,t-1} - M \cdot z^{2,w}_{g,t-1} \le E_1 - m \\
&P^{w}_{g,t} - P^{w}_{g,t-1} - M \cdot (1 - z^{2,w}_{g,t-1}) \\
&\quad - M \cdot z^{3,w}_{g,t-1} \le R_2 \cdot y^{w}_{g,t} \\
&P^{w}_{g,t-1} - M \cdot z^{3,w}_{g,t-1} <= E_2 - m \\
&P^{w}_{g,t} - P^{w}_{g,t-1} - M \cdot (1 - z^{3,w}_{g,t-1}) \le R_3 \cdot y^{w}_{g,t} \\
&P^{w}_{g,t-1} + M \cdot z^{4,w}_{g,t-1} \ge E_2 \\
&z^{2,w}_{2,g,t-1} + z^{3,w}_{g,t-1} - M \cdot z^{4,w}_{g,t-1} \le 1
\end{aligned} \tag{53}
$$

where M refers to a very big number (e.g., 10^5), while m is a very small number (e.g., 10^{-5}) to avoid numerical issues. Conditional constraints (3) and (26) can be linearized in a similar matter.

It is worth noting that different ramping rates are typically ignored in existing literature due to their highly nonlinear nature. To better reflect real-world scenarios, the above conditional constraint and its linearization formulation are implemented in the proposed planning model. In this case, the ramping rate used at time step t only depends on the generation output level at time step $t-1$. According to constraint (52), when the generation output at step $t-1$ is within the range of $[0, E_1)$, ramping rate R_1 should be used. Even though it still cannot fully simulate the real-time scheduling behaviors of generators (e.g., choosing different ramping rates every minute), it has been a much better solution than only using one ramping rate in the existing literature. Furthermore, the solution reality can be further improved by adopting a shorter time resolution (e.g., every 30 minutes). Finally, the big-M method used in constraint (53) may influence the solution quality, which is related to the value choice of M and m. Specifically, when the M is sufficiently big (e.g., 10^5) and m is small enough (e.g., 10^{-5}) with respect to decision variables in the model, the solution quality can be guaranteed [30].

B. Deep Learning-based Data Clustering

To improve the feature extraction capability of conventional distance-based clustering methods, this paper introduces an integrated DEC method for the generation of representative days, which accomplishes the simultaneous and integrated learning of feature extraction and clustering loss reduction [26]. The DEC method is capable of fully extracting the informative features encompassed in the yearly dataset, thereby enabling the dense distribution of data profiles in each cluster and addresses the clustering loss issue in conventional distance-based clustering methods. In this context, the DEC method is capable of generating more realistic representative days compared with distance-based clustering methods, ensuring the final planning decisions (e.g., optimal sizing of various resources) are obtained based on realistic daily profiles and reflect real-world investment scenarios.

To generate K representative days, we first define the daily profile pool of load and RES as:

$$
d_n \in \mathcal{D}^{\text{pool}}, \ \forall n \in \mathcal{N} \tag{54}
$$

where the yearly dataset includes N daily profiles. Then, K clusters can be generated from the profile pool, which is expressed as:

$$
D_k \in \mathcal{D}^{\text{cluster}}, \ \forall k \in \mathcal{K} \tag{55}
$$

where $\{D_k\} \in \mathbb{R}^{N_k \times T}$. Cluster k includes N_k daily profiles. As a result, the centroids of $\mathcal{D}^{\text{cluster}}$ are calculated as:

$$
\overline{D}_k \in \mathcal{D}^{\text{centriod}}, \ \forall k \in \mathcal{K} \tag{56}
$$

where $\overline{D}_k$ refers to the centroid of cluster D_k. The centroids of each cluster will be realized as the representative days in the proposed planning approach.

As illustrated in Fig. 3, the developed DEC framework includes two interacted steps: 1) network parameter initialization via a deep autoencoder network; 2) network parameter update via the minimization of the clustering loss function calculated by Kullback-Leibler (KL) divergence [26]. These two steps are detailed as below:

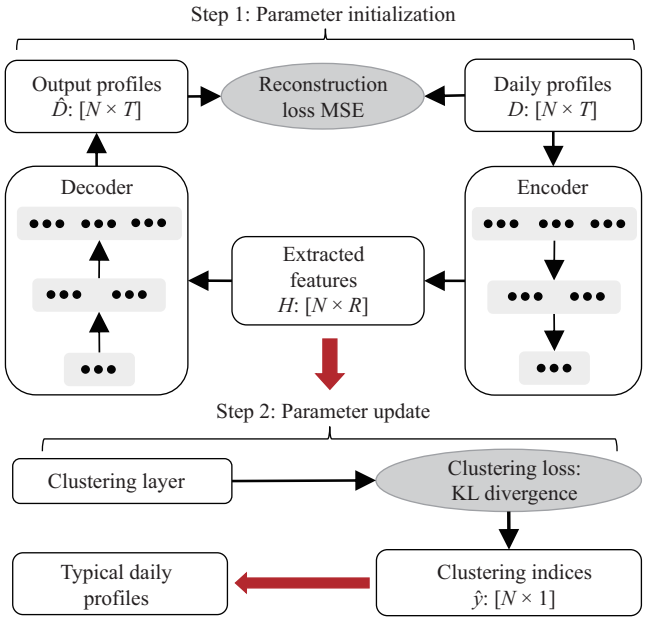

Fig. 3. Illustration of the DEC method, where both the encoder and decoder include several fully connected layers.

Step 1: Parameter Initialization: The objective of step 1 is to initialize the DEC network and retain the well-trained encoder layers for step 2. As one of the powerful unsupervised learning methods, a deep autoencoder network (DAE) can be introduced for the initialization of the DEC method towards effective feature extraction of the input data (daily profiles) [26]. The objective function of the DAE is to minimize the discrepancy between original data inputs and rebuilt data outputs [34], as depicted in Fig. 3.

Specifically, the encoder and decoder functions of the DAE network are expressed as:

$$
h = f_w(d) = \sigma(wd + b) \tag{57}
$$

$$
\hat{d} = g_{w^T}(h) = \sigma(w^T h + b^T) \tag{58}
$$

where σ corresponds to the sigmoid activation function [26]. The weights and biases of the DAE network are denoted by w and b, respectively. The set of network layers in the DAE is defined as $l \in \mathcal{L}$, where the number of neurons in each layer is denoted by $z_l \in \mathcal{Z}_L$.

During the initialization process, each layer in the DAE network can be realized as a de-noising autoencoder integrated with random corruption, which is stated as:

$$\tilde{h_l} \leftarrow \mathrm{Drop}(h_l) \tag{59}$$

$$h_{l+1} = s(W_l \tilde{h_l} + b_l),\ \forall l \in \mathcal{L} \tag{60}$$

where $\mathrm{Drop}(\cdot)$ denotes the Dropout mapping function that can randomly reset several input dimensions to zero [35]. The de-noising DAE aims to minimize the reconstruction error J between the decoding output $\hat{d}$ and the original input d, which is written as:

$$J = \|\hat{d} - d\|_2^2 \tag{61}$$

where $\|\cdot\|_2$ corresponds to the norm operation based on Euclidean distance [25].

To speed up the training process, the deployed clustering framework uses greedy layerwise pretraining to effectively unfold and initialize the encoder and decoder networks [36]. Specifically, $|\mathcal{L}|$ blocks of de-noising autoencoder are trained via this method, where the output of block l serves as the input of block $l+1$. Additionally, in order to further optimize the parameters, the global fine-tuning technique is deployed via backpropagation with the goal of minimizing reconstruction loss [37]. Finally, the decoder is discarded, while the encoder is retained as feature extractor [26], as presented in Fig. 3.

Step 2: Parameter Update: The objective of step 2 is to update the parameters of the deep encoder layers and then update the cluster centroids by backpropagation until reaching convergence. After extracting the features $h_n \in \mathcal{H}_N$ in step 1, conventional distance-based methods (e.g., K-means) can be used to output cluster centroids $\mathcal{D}^{\text{centriod}} = \{\mu_1, \mu_2, \cdots, \mu_K\}$.

The Student's t-distribution is used to calculate the similarity metric $q_{n,k}$ between feature h_n and centroid μ_k, which can be expressed as:

$$q_{n,k} = \frac{(1 + \|h_n - \mu_k\|^2)^{-1}}{\sum_k (1 + \|h_n - \mu_k\|^2)^{-1}},\ \forall n \in \mathcal{N},\ \forall k \in \mathcal{K} \tag{62}$$

where the probability of sample n in cluster k is denoted by $q_{n,k}$. In this context, the encoder layers can be updated by KL divergence minimization, which is written as:

$$L = KL(P\|Q) = \sum_n \sum_k p_{n,k} \log \frac{p_{n,k}}{q_{n,k}} \tag{63}$$

$$p_{n,k} = \frac{q_{n,k}^2 / \sum_n q_{n,k}}{\sum_n (q_{n,k}^2 / \sum_n q_{n,k})},\ \forall n \in \mathcal{N},\ \forall k \in \mathcal{K} \tag{64}$$

where L is the KL divergence between the auxiliary distribution p_n and the probability q_n. After obtaining the optimal index k^* with the highest probability $q_{n,k}$, the cluster label y_n is assigned for sample d_n, which is expressed as:

$$y_n = \arg\max_k (q_{n,k}),\ \forall k \in \mathcal{K} \tag{65}$$

To update deep encoder layers, the stochastic gradient descent method is used for the calculation of gradients of L in (63), which is expressed as:

$$\frac{\partial L}{\partial h_n} = 2 \sum_k (2 + \|h_n - \mu_k\|^2)^{-1} (p_{n,k} - q_{n,k})(h_n - \mu_k) \tag{66}$$

$$\frac{\partial L}{\partial \mu_k} = -2 \sum_k (2 + \|h_n - \mu_k\|^2)^{-1} (p_{n,k} - q_{n,k})(h_n - \mu_k) \tag{67}$$

where $\frac{\partial L}{\partial h_n}$ and $\frac{\partial L}{\partial \mu_k}$ correspond to the gradients of L in (63) with respect to h_n and cluster centroid μ_k, respectively. This process will be repeated iteratively until meeting the convergence criterion ϵ. The detailed training process of the DEC method is presented in Algorithm 1.

Algorithm 1: Deep Embedded Clustering

Input: Yearly dataset: $\mathcal{D}^{\text{pool}}$; Profile number: N; Cluster number: K; Maximum iterations: $\overline{\text{iter}}$; Convergence criterion: ϵ; Update interval of the target distribution: I_U.

Output: Cluster labels: $\hat{y} \in \mathcal{Y} = \{\hat{y}_1, \hat{y}_2, \cdots, \hat{y_N}\}$.

Step 1: Train a deep de-noising autoencoder network with L layers to initialize the DEC model. Retain the well-trained encoder layers $f_w(\cdot)$.
Step 2: Use K-means to initialize centroids $\mathcal{D}^{\text{centriod}}$ based on extracted features $\mathcal{H} = f_w(\mathcal{D}^{\text{sample}})$.
$\mathcal{Y} = \mathrm{Kmeans}(\mathcal{H}, K)$.
for iter $= 1 : \overline{\text{iter}}$ **do**
  if *Modulo*$(\text{iter}, I_U) == 0$ **then**
    Extract feature $\mathcal{H}$ from $\mathcal{D}^{\text{pool}}$.
    for $n = 1 : N$ **do**
      $h_n = f_W(d_n)$
    end
    Update P via (62) and (64).
    $\hat{y}^0 = \hat{y}$
    Update cluster labels.
    for $n = 1 : N$ **do**
      $\hat{y} = \arg\max_k (q_{n,k})$
    end
    if $\frac{\#(\hat{y}^0 \neq \hat{y})}{N} \leq \epsilon$ **then**
      Stop
    end
  end
  Select sample data $\mathcal{D}^{\text{sample}}$ from $\mathcal{D}^{\text{pool}}$.
  Update the DAE parameters w and b, and the cluster centroids $\mathcal{D}^{\text{centriod}}$ via (66) and (67).
end

V. Case Studies

A. Experimental Setup

1) Network and Data Descriptions

Case studies are conducted on a coupled 6-bus power and 7-node gas system, where the network structure is illustrated in Fig. 4. The yearly dataset used to generate representative days are collected from [38]. Cost and operation data for different types of generation technologies including CCGT, OCGT, NG CCS, WTs, and PVs are collected from [5], [10], while the cost and operation data of ESs and P2G plants can be found in [25], [27], respectively. The carbon target is set as 100 g/kWh [5].

2) Frequency Security Setup

Regarding post-fault under-frequency limits, the frequency nadir should never be lower than 49.2 Hz to avoid activating under-frequency load shedding, i.e., $\Delta\overline{f} = 0.8$ Hz [10]. Since the RoCoF limit is being relaxed to 1 Hz/s in the GB power

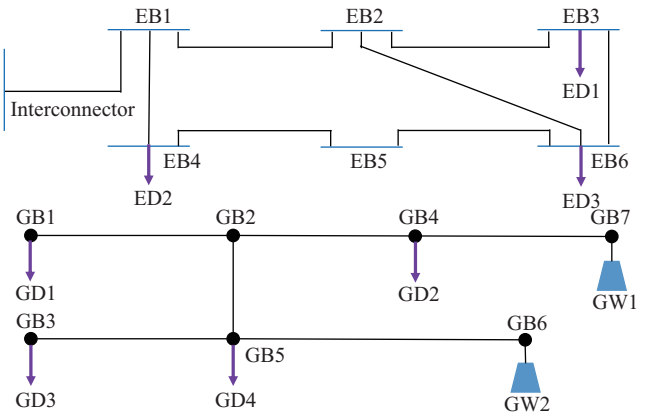

Fig. 4. The studied integrated system including a 6-bus power network and a 7-node gas network.

system [39], the RoCoF should be below 1 Hz/s to prevent the tripping of RoCoF-sensitive protection relays, i.e., $\overline{\text{RoCoF}} = 1$ Hz/s. The delivery speed of frequency response from ESs and generators is set as 1 s and 10 s, respectively [10]. Similarly, the frequency Nadir and RoCoF for over-frequency security are also set as 0.8 Hz and 1 Hz/s, respectively.

3) DEC Network Setup

Both the deep encoder and decoder networks in the DAE network contain five fully connected layers [26]. The learning rate of the Adam optimizer is 0.001. The batch size is set as 1000, while the convergence threshold is set as 0.1%. The maximum iteration and update interval are set as $2e^5$ and 2000, respectively.

B. Clustering Performance

To verify the clustering performance of the introduced DEC model on generating representative days, four different clustering validity indicators are used to quantify the quality of the DEC clustering method and three traditional distance-based methods, i.e., K-means, Gaussian Mixture Model (GMM), and hierarchical clustering. The four indicators are clustering dispersion indicator (CDI), Davies-Bouldin index (DBI), modified Dunn index (MDI), and mean index adequacy (MIA), where their detailed mathematical formulations can be found on [40]. To capture the trade-off between computing efficiency and result reality, the number of representative days (clusters) is set to 6. The clustering results from the DEC method are depicted in Figs. 5 and 6, while the comparison among four different methods is depicted in Fig. 7.

It can be observed in Figs. 5 and 6 that the DEC model can classify the yearly data profiles into 6 different clusters with different shapes. Specifically, the obtained 6 representative daily profiles have different RES and load levels, which demonstrates the effectiveness of the DEC clustering method. Regarding the 4 indicators provided in Fig. 7, the DEC method outperforms the other three distance-based clustering methods in terms of different metrics. In the case of the CDI metric, the DEC model (green) obtains 21.5%, 25.5%, and 13.9% lower values than the K-means method (yellow), the GMM model (red), and the hierarchical method (blue), respectively. The reason that the DEC method can achieve good performance is the simultaneous and integrated learning of feature extraction and cluster-oriented loss minimization. As compared to GMM and K-means approaches, the hierarchical clustering method also obtains high performance in terms of the MDI and DBI metrics. However, the hierarchical clustering method does not perform well with respect to the other two metrics (MIA and CDI), because of the difficulties in handling large-scale yearly datasets. Furthermore, it can be observed from Fig. 7 that the GMM method has the worst performance in terms of all 4 metrics, underscoring the issue with conventional multivariate

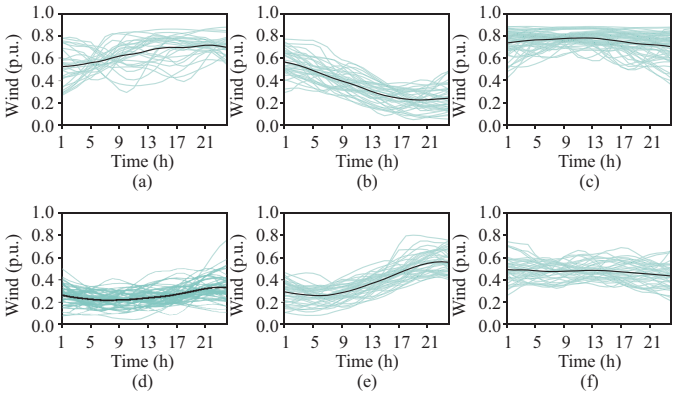

Fig. 5. Illustration of the wind profiles in the representation days. (a) Representative day 1. (b) Representative day 2. (c) Representative day 3. (d) Representative day 4. (e) Representative day 5. (f) Representative day 6.

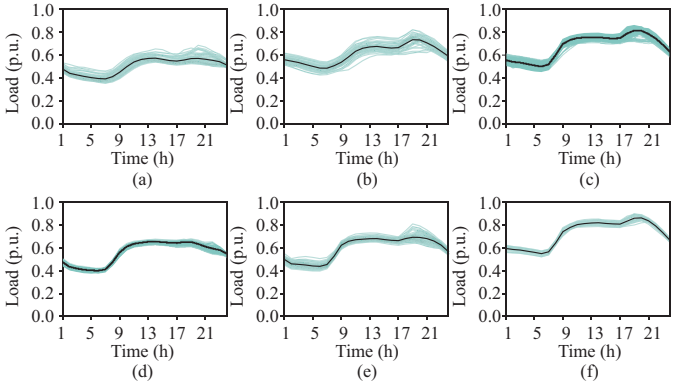

Fig. 6. Illustration of the load profiles in the representation days. (a) Representative day 1. (b) Representative day 2. (c) Representative day 3. (d) Representative day 4. (e) Representative day 5. (f) Representative day 6.

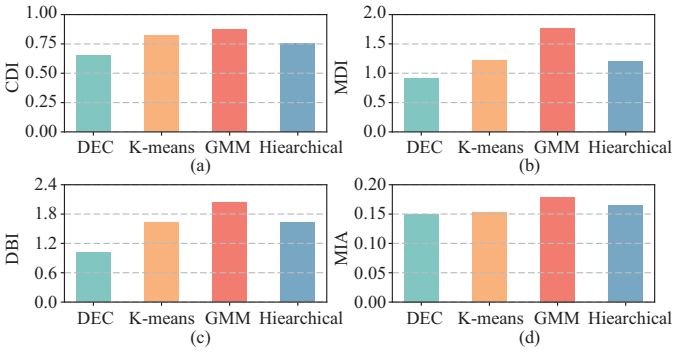

Fig. 7. Comparisons between DEC and three distance-based clustering methods employing CDI, MDI, DBI, and MIA as evaluation metrics. (a) CDI. (b) MDI. (c) DBI. (d) MIA.

model-based solutions.

Finally, to further verify the clustering performance of the proposed DEC method, a comparison between DEC and three density-based clustering methods including Density-Based Spatial Clustering of Applications with Noise (DBSCAN), Ordering Points to Identify the Clustering Structure (OPTICS), and Hierarchical Density-based Clustering (HDBSCAN) is conducted in this section [41]. Two indicators suitable for both DEC and density-based clustering methods are adopted for evaluation, i.e., DBI and Silhouette Score (SC) [41], where the comparison results are illustrated in Fig. 8. It can be observed that the proposed DEC method can achieve the best performance with regard to these two indicators (lowest value on DBI and highest value on SC), further verifying the reliable clustering performance of the DEC method.

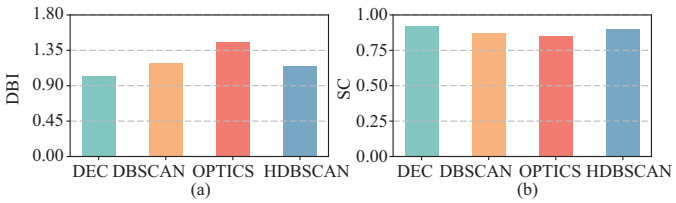

Fig. 8. Comparisons between DEC and three density-based clustering methods employing DBI and SC as evaluation metrics. (a) DBI. (b) SC.

C. Analysis of the Planning Results

To demonstrate the effectiveness and benefits of the proposed planning model in accurate and reliable decision making, four different cases are developed for detailed comparison:

Case 1: The optimal sizing of the integrated power and gas system is conducted, considering frequency security requirements and P2G technologies.

Case 2: The optimal sizing of the integrated power and gas system is conducted with P2G technologies, while frequency security requirements are not considered.

Case 3: The optimal sizing of the integrated power and gas system is conducted with frequency security requirements, while P2G technologies are not considered.

Case 4: The optimal sizing of the integrated power and gas system is conducted with frequency security requirements and P2G technologies, while frequency support from WTs is considered.

The comparisons among the above four cases with respect to the whole system costs, installed capacities of gas-fired plants, P2G plants, ESs, and RESs are illustrated in Fig. 9. The operating behaviors of the interconnector and its influence on frequency security are depicted in Fig. 10, while the detailed scheduling behaviors of one gas-fired generator during one representative day are presented in Fig. 11.

1) Investment Results

It can be observed from Fig. 9(a) that considering frequency security requirements (Case 1) in the proposed planning model can increase the whole system cost, compared with Case 2 that does not consider frequency security. This is because more conventional generators (e.g., gas-fired power plants) are required in Case 1 to provide sufficient inertia and frequency response services, as depicted in Fig. 9(b). Accordingly, when frequency security is considered in Case 1, the capability of the integrated system to accommodate RESs is limited, causing less RES installment compared with Case 2, as depicted in Fig. 9(e) and (f). Furthermore, the higher penetration of RESs under Case 2 raises larger uncertainties, requiring more battery units to be installed for demand-supply balance, as depicted in Fig. 9(d).

Regarding the influence of P2G plants on planning results, Fig. 9(a) and (c) show that considering P2G technologies in Case 1 leads to significant savings in investment costs, compared with Case 3. This is because P2G plants can transform additional wind power into synthetic natural gas, reducing wind curtailment and the operation cost of the gas

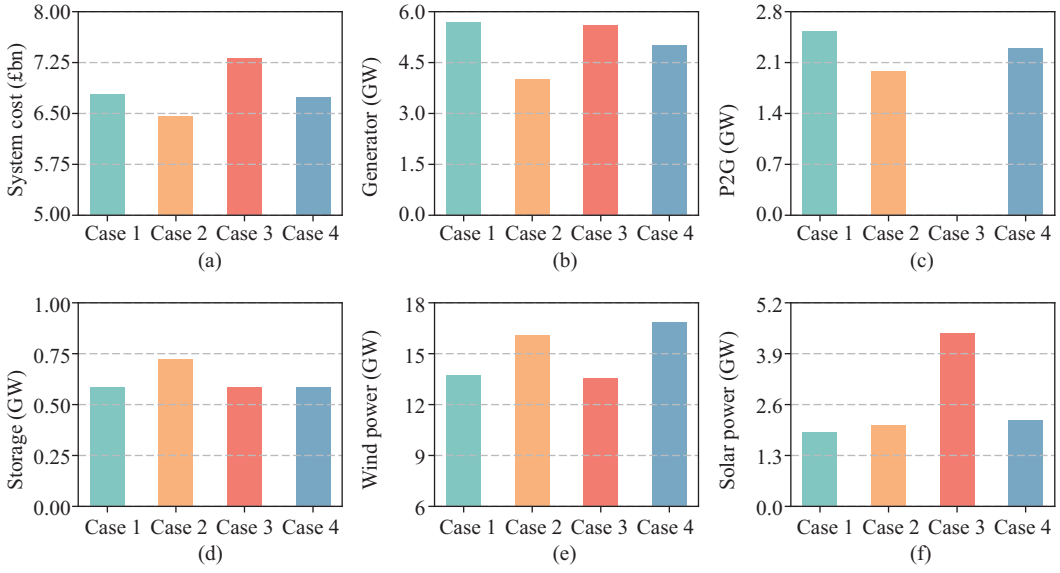

Fig. 9. Comparisons among 4 different cases with respect to the whole system cost as well as the installed capacities of conventional generation, P2G plants, ESs, and renewable generation. (a) Whole system cost. (b) Capacity of gas plants. (c) Capacity of P2G plants. (d) Capacity of ESs. (e) Capacity of WTs. (f) Capacity of PVs.

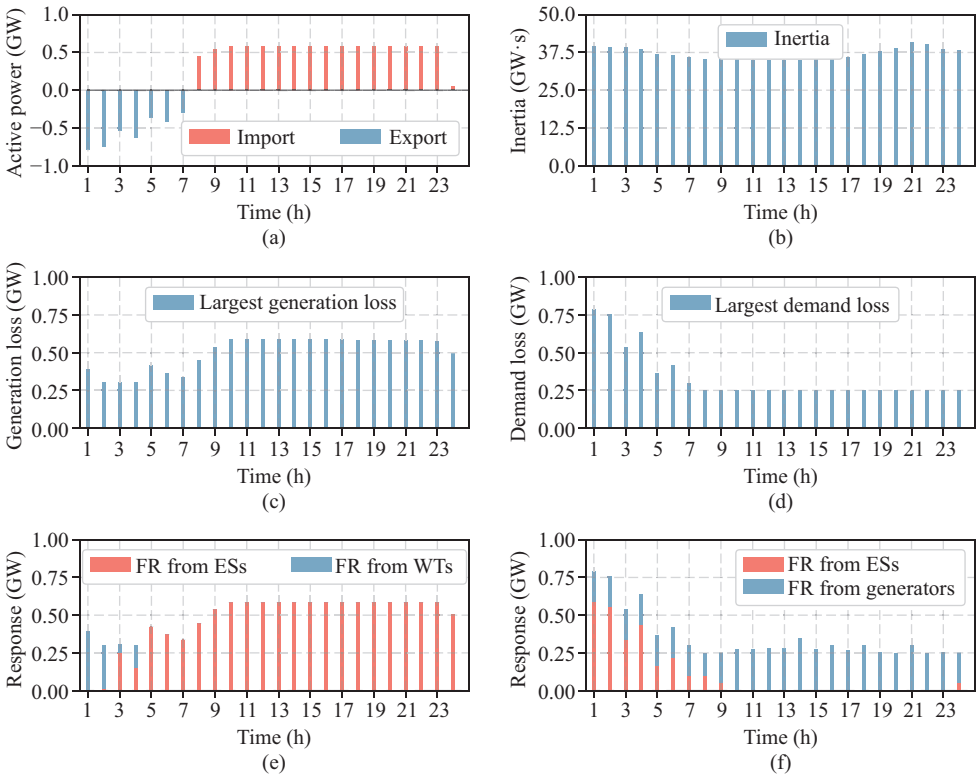

Fig. 10. Illustration of the scheduling behaviors of the interconnector, overall system inertia, the largest loss of generation and demand as well as the under and over FR services from different resources. (a) Output of the interconnector. (b) Overall system inertia. (c) Largest loss of generation. (d) Largest loss of demand. (e) Under-frequency response. (f) Over-frequency response.

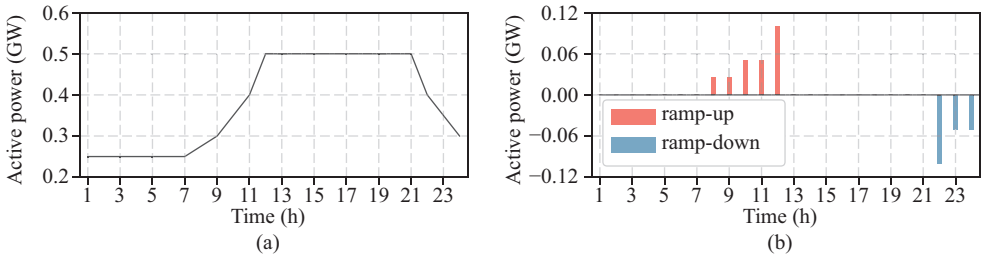

Fig. 11. Illustration of the scheduling behaviors of one gas-fired generator with three different ramping rates. (a) Gas-fired generator 1. (b) Ramping rates at each hour.

system. Furthermore, it is observed from Fig. 9(a) and (b) that considering frequency response from WTs in Case 4 can slightly decrease the whole system cost and then reduce the installed capacity of conventional gas-fired plants, while the installed capacity of RESs is significantly increased compared with Case 1, as depicted in Fig. 9(e).

2) Operation Results

The scheduling behaviors of the interconnector on one selected representative day are illustrated in Fig. 10(a), while the overall system inertia is shown in Fig. 10(b). It can be found that the interconnector is exporting power in the first few hours of the day and then importing power after 8:00 am. The power exporting and importing behaviors influence the largest loss of demand and generation, as depicted in Fig. 10(c) and (d). Specifically, over 0.5 GW largest loss could occur on both the generation and demand sides, which leads to much more procurement of FR services from different resources, as shown in Fig. 10(e) and (f).

On the other hand, the detailed generation behaviors of one gas-fired generator are illustrated in Fig. 11. It can be clearly found that the generator experiences three different ramp-up rates and two elbows during the selected representative day. The advantages of the proposed planning model in capturing detailed scheduling behaviors of each individual generator have been verified.

D. Cost Benefits of Adding More Technologies

To further verify the performance of the proposed planning model on capturing different types of technologies, two low-carbon resources including hydroelectricity and biomass are considered in this section. Specifically, the cost and operation data of hydroelectricity are collected and estimated from [13] and [42], while the cost and operation data of biomass are collected from [5]. To investigate the benefits of adding these two new technologies, a comparison case study is carried

out: 1) Case 1: Generation technologies including CCGT, OCGT, NG CCS, WTs, PVs, and P2Gs are considered, while hydroelectricity and biomass are not considered; 2) Case 2: Hydroelectricity and biomass are considered in the optimal sizing problem of integrated power and gas systems.

It can be observed from Fig. 12(a) that adding these low-carbon technologies can significantly reduce the whole system costs, due to the zero-carbon nature and stability of hydroelectricity that can displace certain amounts of conventional generators. As depicted in Fig. 12(b), biomass is not installed in Case 2, due to its high investment costs. This finding demonstrates that large-scale deployment of biomass-based resources is economically unfeasible compared to the generation types selected by the model, from the whole-system point of view, which is also aligned with the conclusion in [5]. However, it is worth noting that long-term planning decisions for hydroelectricity are complicated and should take into account various factors, such as environmental impact, reservoir management, geographical constraints, etc. This is out of the scope of this paper and could be further investigated in future research. Additionally, the carbon intensity of biomass varies depending on different technologies, which could provide renewable and sustainable energy supply and deserve further exploration.

E. Scalability Analysis in a Larger System

This section is used as a further illustration of the proposed planning approach on scalability through a larger integrated power and gas system, i.e., the reduced 14-bus GB power system [43] and the 14-node gas system [27]. With respect to the four cases provided in Section V-C, the whole system cost and installed capacity of RESs in the large system are illustrated in Fig. 13. Additionally, to investigate the influence of carbon targets and RoCoF limits, relevant sensitivity analysis is conducted, where the results can be found in Fig. 14.

It can be observed from Fig. 13 that considering frequency security (Case 1) can lead to larger investment costs than the case without frequency security (Case 2), while considering P2G technologies can significantly reduce the whole system costs, compared with the case without P2G technologies (Case 3). Additionally, considering frequency response from WTs (Case 4) can further reduce the whole system costs and increase the accommodation of RESs, compared with the case without wind frequency response (Case 1). Finally, Fig. 14 illustrates that changing the carbon target from 100 g/kWh to 25 g/kWh can slightly increase the whole system costs, due to the higher requirement for generation technologies with low carbon intensity (e.g., Gas CCS and RESs). On the other hand,

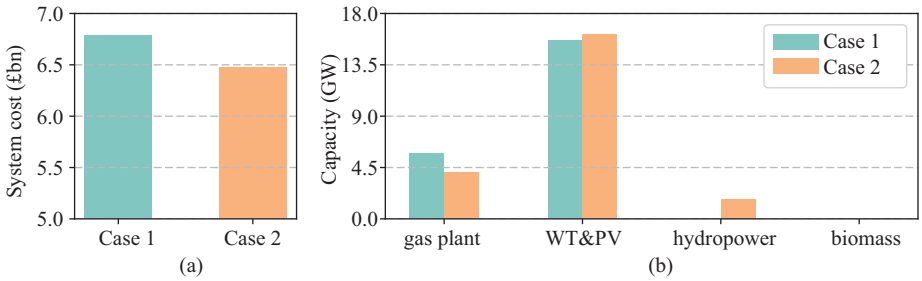

Fig. 12. Comparisons among 2 different cases with respect to the whole system cost and the installed capacities of conventional generation and renewable generation. (a) Whole system cost. (b) Capacities of installed technologies.

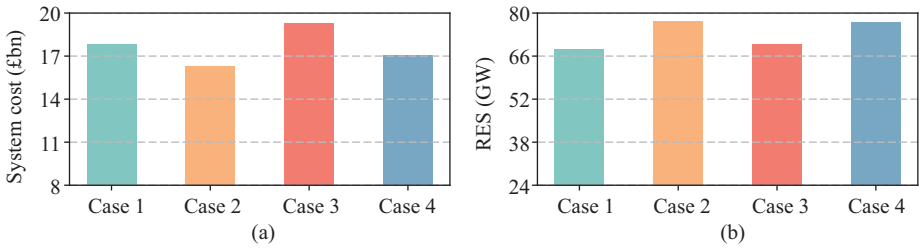

Fig. 13. Comparisons among 4 different cases with respect to the whole system cost and the installed capacity of renewable generation in the larger integrated power and gas system. (a) Whole system cost. (b) Capacity of RESs.

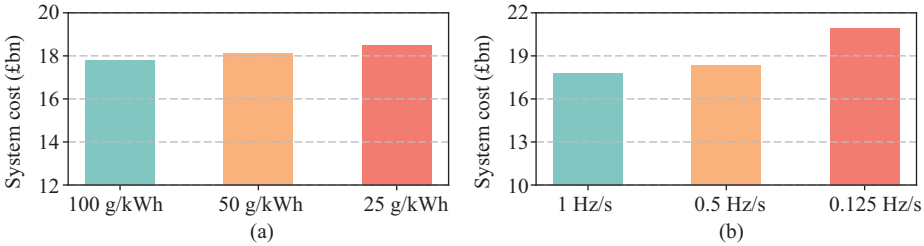

Fig. 14. Sensitivity analysis on the influence of different carbon targets and RoCoF limits in the larger integrated power and gas system. (a) Influence of carbon target. (b) Influence of RoCoF limit.

tightening the RoCoF limit from 1 Hz/s to 0.125 Hz/s can significantly increase the whole system costs, due to the much higher requirement for system inertia.

VI. Conclusion

A novel planning approach is proposed in this paper to solve the optimal sizing problem of integrated power and gas systems while capturing frequency security. The proposed approach is based on a detailed UC model with complete parameters and different ramping rates, which is capable of simulating the detailed scheduling behaviors of each individual generator and leading to accurate inertia calculation. Additionally, the power importing and exporting behaviors of interconnectors are modeled to account for both the under and over frequency security requirements. Furthermore, the DEC clustering method based on deep learning techniques is introduced to generate representative days for the problem formulation of the proposed planning model, fully exploiting the informative attributes encompassed in the dataset. Experiments are conducted on two integrated systems including a 6-bus power and 7-node gas system as well as a 14-bus power and 14-node gas system to evaluate the superiority of the proposed novel planning approach in clustering performance and investment decision-making accuracy. Finally, the scheduling behaviors of the integrated system are analyzed and compared in four different cases to verify the advantages of the proposed method in improving decision-making accuracy.

To summarize, the proposed planning model is capable of capturing detailed scheduling behaviors of each generator and then calculating accurate inertia provision, while both under and over frequency security requirements are incorporated. The research limitations of this paper and potential future works are discussed as below. First, the proposed planning model and developed case studies mainly consider a limited number of technologies (e.g., gas-fired generators, RESs, P2G plants, etc.). However, in response to the low-carbon transition, future power systems may be rapidly shifted to involve various types of technologies, e.g., electric vehicles, hydrogen and CCS, flexible nuclear power plants, etc. In this context, future work will extend the proposed planning model to involve different types of zero-carbon technologies to reflect real-world scenarios and contribute to the low-carbon future. Second, the proposed planning model covers the integration of power and gas systems, while the influence of heating systems on investment costs is not considered. Since flexibility from heating sectors such as heat pumps and thermal storages could significantly enhance the energy efficiency of the whole system and contribute to investment cost savings, future work will develop planning approaches for the whole system operations across power, gas, and heating sectors. Third, the proposed planning model mainly focuses on the post-fault frequency security of power sectors via frequency-related constraints; nevertheless, the dynamic behaviors of gas systems are not considered, which may lead to conservative solutions. Future work will focus on developing planning approaches that consider the dynamic behaviors of gas flow via dynamic partial differential equations towards more reliable investment decisions.

References

[1] G. Strbac, D. Pudjianto, M. Aunedi, D. Papadaskalopoulos, P. Djapic, Y. J. Ye, R. Moreira, H. Karimi, and Y. Fan, "Cost-effective decarbonization in a decentralized market: the benefits of using flexible technologies and resources," *IEEE Power and Energy Magazine*, vol. 17, no. 2, pp. 25–36, Mar./Apr. 2019.

[2] F. Teng, V. Trovato, and G. Strbac, "Stochastic scheduling with inertia-dependent fast frequency response requirements," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1557–1566, Mar. 2016.

[3] Z. Y. Zeng, T. Ding, Y. T. Xu, Y. H. Yang, and Z. Y. Dong, "Reliability evaluation for integrated power-gas systems with power-to-gas and gas storages," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 571–583, Jan. 2020.

[4] Y. Xiang, Y. T. Guo, G. Wu, J. Y. Liu, W. Sun, Y. T. Lei, and P. L. Zeng, "Low-carbon economic planning of integrated electricity-gas energy systems," *Energy*, vol. 249, pp. 123755, Jun. 2022.

[5] X. Zhang, G. Strbac, N. Shah, F. Teng, and D. Pudjianto, "Whole-system assessment of the benefits of integrated electricity and heat system," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 1132–1145, Jan. 2019.

[6] A. Moreira, M. Heleno, A. Valenzuela, J. H. Eto, J. Ortega, and C. Botero, "A scalable approach to large scale risk-averse distribution grid expansion planning," *IEEE Transactions on Power Systems*, vol. 39, no. 1, pp. 2115–2128, Jan. 2024.

[7] G. S. Pan, W. Gu, Y. P. Lu, H. F. Qiu, S. Lu, and S. Yao, "Optimal planning for electricity-hydrogen integrated energy system considering power to hydrogen and heat and seasonal storage," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 4, pp. 2662–2676, Oct. 2020.

[8] T. Wu, X. Wei, X. Zhang, G. B. Wang, J. Qiu, and S. W. Xia, "Carbon-oriented expansion planning of integrated electricity-natural gas systems with EV fast-charging stations," *IEEE Transactions on Transportation Electrification*, vol. 8, no. 2, pp. 2797–2809, Jun. 2022.

[9] Z. Y. Zhang, E. S. Du, F. Teng, N. Zhang, and C. Q. Kang, "Modeling frequency dynamics in unit commitment with a high share of renewable energy," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4383–4395, Nov. 2020.

[10] L. Badesa, F. Teng, and G. Strbac, "Simultaneous scheduling of multiple frequency services in stochastic unit commitment," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3858–3868, Sep. 2019.

[11] L. W. Sang, Y. L. Xu, Z. K. Yi, L. Yang, H. Long, and H. B. Sun, "Conservative sparse neural network embedded frequency-constrained unit commitment with distributed energy resources," *IEEE Transactions on Sustainable Energy*, vol. 14, no. 4, pp. 2351–2363, Oct. 2023.

[12] L. Badesa, C. Matamala, Y. J. Zhou, and G. Strbac, "Assigning shadow prices to synthetic inertia and frequency response reserves from renewable energy sources," *IEEE Transactions on Sustainable Energy*, vol. 14, no. 1, pp. 12–26, Jan. 2023.

[13] Y. Yin, T. Q. Liu, L. Wu, C. He, and Y. K. Liu, "Frequency-constrained multi-source power system scheduling against *N-1* contingency and renewable uncertainty," *Energy*, vol. 216, pp. 119296, Feb. 2021.

[14] L. Yang, Y. L. Xu, J. G. Zhou, and H. B. Sun, "Distributionally robust frequency constrained scheduling for an integrated electricity-gas system," *IEEE Transactions on Smart Grid*, vol. 13, no. 4, pp. 2730–2743, Jul. 2022.

[15] A. Jawad, Nahid-Al-Masood, and S. Munim, "Optimal sizing of BESS for attaining frequency stability under high PV penetration," in *2021 International Conference on Technology and Policy in Energy and Electric Power (ICT-PEP)*, 2021, pp. 348–353.

[16] A. M. Nakiganda, S. Dehghan, U. Markovic, G. Hug, and P. Aristidou, "A stochastic-robust approach for resilient microgrid investment planning under static and transient islanding security constraints," *IEEE Transactions on Smart Grid*, vol. 13, no. 3, pp. 1774–1788, May 2022.

[17] M. Javadi, Y. Z. Gong, and C. Y. Chung, "Frequency stability constrained BESS sizing model for microgrids," *IEEE Transactions on Power Systems*, vol. 39, no. 2, pp. 2866–2878, Mar. 2024.

[18] X. T. Ni, Z. J. Liu, D. Liu, H. Y. Yu, K. Li, and S. C. Cui, "A frequency-constrained Co-planning model of renewable energy and transmission network towards energy efficient power systems," in *2023 IEEE 6th International Electrical and Energy Conference (CIEEC)*, 2023, pp. 380–385.

[19] T. Q. Zhao, N. Raghunathan, A. Yogarathnam, M. Yue, and P. B. Luh, "A scalable planning framework of energy storage systems under frequency dynamics constraints," *International Journal of Electrical Power & Energy Systems*, vol. 145, pp. 108693, Feb. 2023.

[20] H. Y. Zhang, K. Liao, J. W. Yang, S. W. Zheng, and Z. Y. He, "Frequency-constrained expansion planning for wind and photovoltaic power in wind-photovoltaic-hydro-thermal multi-power system," *Applied Energy*, vol. 356, pp. 122401, Feb. 2024.
[21] H. Li, Y. Qiao, Z. X. Lu, B. S. Zhang, and F. Teng, "Frequency-constrained stochastic planning towards a high renewable target considering frequency response support from wind power," *IEEE Transactions on Power Systems*, vol. 36, no. 5, pp. 4632–4644, Sep. 2021.
[22] National Grid. (2023). Grid code documents. [Online]. Available: https://www.nationalgrideso.com/industry-information/codes/grid-code-gc/grid-code-documents
[23] Q. W. Pang, A. De Paola, V. Trovato, and G. Strbac, "Value of interconnectors operating in simultaneous energy-frequency response markets," *IEEE Transactions on Power Systems*, vol. 37, no. 5, pp. 3381–3393, Sep. 2022.
[24] W. W. Li, T. Qian, Y. Zhang, Y. Q. Shen, C. H. Wu, and W. H. Tang, "Distributionally robust chance-constrained planning for regional integrated electricity–heat systems with data centers considering wind power uncertainty," *Applied Energy*, vol. 336, pp. 120787, Apr. 2023.
[25] Y. Wang, A. O. Rousis, and G. Strbac, "A three-level planning model for optimal sizing of networked microgrids considering a trade-off between resilience and cost," *IEEE Transactions on Power Systems*, vol. 36, no. 6, pp. 5657–5669, Nov. 2021.
[26] M. Y. Sun, Y. Wang, F. Teng, Y. J. Ye, G. Strbac, and C. Q. Kang, "Clustering-based residential baseline estimation: a probabilistic perspective," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6014–6028, Nov. 2019.
[27] H. S. Zhou, J. H. Zheng, Z. G. Li, Q. H. Wu, and X. X. Zhou, "Multi-stage contingency-constrained co-planning for electricity-gas systems interconnected with gas-fired units and power-to-gas plants using iterative benders decomposition," *Energy*, vol. 180, pp. 689–701, Aug. 2019.
[28] R. Sioshansi and A. J. Conejo, *Optimization in Engineering: Models and Algorithms*, Cham: Springer International Publishing, 2017.
[29] L. Q. Bai, J. H. Wang, C. S. Wang, C. Chen, and F. X. Li, "Distribution locational marginal pricing (DLMP) for congestion management and voltage support," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 4061–4073, Jul. 2018.
[30] L. Badesa, F. Teng, and G. Strbac, "Pricing inertia and frequency response with diverse dynamics in a mixed-integer second-order cone programming formulation," *Applied Energy*, vol. 260, pp. 114334, Feb. 2020.
[31] Z. F. Yang, H. W. Zhong, A. Bose, T. X. Zheng, Q. Xia, and C. Q. Kang, "A linearized OPF model with reactive power and voltage magnitude: a pathway to improve the MW-only DC OPF," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1734–1745, Mar. 2018.
[32] V. S. Saravi, M. Kalantar, and A. Anvari-Moghaddam, "Resilience-constrained expansion planning of integrated power–gas–heat distribution networks," *Applied Energy*, vol. 323, pp. 119315, Oct. 2022.
[33] E. Raheli, Q. W. Wu, M. L. Zhang, and C. Y. Wen, "Optimal coordinated operation of integrated natural gas and electric power systems: a review of modeling and solution methods," *Renewable and Sustainable Energy Reviews*, vol. 145, pp. 111134, Jul. 2021.
[34] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.
[35] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, Jan. 2014.
[36] G. C. Ruan, Z. K. Yu, S. T. Pu, S. T. Zhou, H. W. Zhong, L. Xie, Q. Xia, and C. Q. Kang, "Open-access data and toolbox for tracking COVID-19 impact on power systems," *IEEE Transactions on Power Systems*, vol. 38, no. 2, pp. 1619–1631, Mar. 2023.
[37] G. C. Ruan, J. X. Wang, H. W. Zhong, Q. Xia, and C. Q. Kang, "Improving sample efficiency of deep learning models in electricity market," *IEEE Transactions on Power Systems*, vol. 38, no. 5, pp. 4761–4773, Sep. 2023.
[38] National Grid. (2022). Data finder and explorer, demand data, historic demand data. [Online]. Available: https://www.nationalgrideso.com/document/162271/download
[39] NESO.(2022). Security and quality of supply standards. [Online]. Available: https://www.nationalgrideso.com/industry-information/codes/security-and-quality-supply-standard-sqss
[40] M. Y. Sun, I. Konstantelos, and G. Strbac, "C-vine copula mixture model for clustering of residential electrical load pattern data," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2382–2393, May 2017.
[41] M. A. Ahmed, H. Baharin, and P. N. Nohuddin, "Analysis of K-means, DBSCAN and OPTICS cluster algorithms on AL-quran verses," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 8, pp. 248–254, Jan. 2020.
[42] G. A. Oladosu, J. Werble, W. Tingen, A. Witt, M. Mobley, and P. O'Connor, "Costs of mitigating the environmental impacts of hydropower projects in the united states," *Renewable and Sustainable Energy Reviews*, vol. 135, pp. 110121, Jan. 2021.
[43] D. W. Qiu, Y. Wang, J. K. Wang, C. W. Jiang, and G. Strbac, "Personalized retail pricing design for smart metering consumers in electricity market," *Applied Energy*, vol. 348, pp. 121545, Oct. 2023.

Yi Wang received the B.Eng. degree and the M.Eng. degree from Tianjin University in 2015 and 2018, respectively, and the Ph.D. degree from Imperial College London in 2022. He is currently employed as a Research Associate in the Department of Electrical and Electronic Engineering at Imperial College London. His research interests include mathematical programming and learning approaches applied to the planning and operation of networked microgrids, the resilience enhancement of future power systems, and multi-energy system integration.

Goran Strbac is a Professor of Energy Systems at Imperial College London, London, U.K. He led the development of novel advanced analysis approaches and methodologies that have been extensively used to inform industry, governments, and regulatory bodies about the role and value of emerging new technologies and systems in supporting cost effective evolution to smart low carbon future. He is currently the Director of the joint Imperial-Tsinghua Research Centre on Intelligent Power and Energy Systems, Leading Author in IPCC WG 3, Member of the European Technology and Innovation Platform for Smart Networks for the Energy Transition, and Member of the Joint EU Programme in Energy Systems Integration of the European Energy Research Alliance.