Chance-constrained regulation capacity offering for HVAC systems under non-Gaussian uncertainties with mixture-model-based convexification

Ge Chen, *Graduate Student Member, IEEE,* Hongcai Zhang, *Member, IEEE,* Hongxun Hui, *Member, IEEE,* and Yonghua Song, *Fellow, IEEE*

***Abstract*—Heating, ventilation, and air-conditioning (HVAC) systems are ideal demand-side flexible resources to provide regulation services. However, finding the best hourly regulation capacity offers for HVAC systems in a power market ahead of time is challenging because they are affected by non-Gaussian uncertainties from regulation signals. Moreover, since HVAC systems need to frequently regulate their power according to regulation signals, numerous thermodynamic constraints are introduced, leading to a huge computational burden. This paper proposes a tractable chance-constrained model to address these challenges. It first develops a temporal compression approach, in which the extreme indoor temperatures in the operating hour are estimated and restricted in the comfortable range so that the numerous thermodynamic constraints can be compressed into only a few ones. Then, a novel convexification method is proposed to handle the non-Gaussian uncertainties. This method leverages the Gaussian mixture model to reformulate the chance constraints with non-Gaussian uncertainties on the left-hand side into deterministic non-convex forms. We further prove that these non-convex forms can be approximated by mixed-integer second-order cone constraints that can be efficiently solved by off-the-shelf solvers. The optimality gap because of this approximation is marginal under mild conditions. Numerical experiments are conducted to validate the superiority of the proposed method.**

***Index Terms*—HVAC systems, demand-side flexibility, regulation capacity, chance-constrained programming, Gaussian mixture model, convexification.**

NOMENCLATURE

Indices

i	Index of buildings.
j	Index of Gaussian components.
k	Index set of elements in vector $\boldsymbol{\alpha}(\boldsymbol{x})$.
l	Index set of short time slots in a hour (time step is two seconds).
n	Index sets of regulation signals.
t	Index sets of hours.
τ	Index set of the long duration in a hour.

Parameters

A	Coefficient matrix.
C_i	Heat capacity of building i (MWh/°C).
COP_i	Coefficient of performance of the HVAC system in building i (MW/MW).
g_i	Heat transfer coefficient between indoor and outdoor environments in building i (MW/°C).
h_i	Heat load from indoor sources in building i (MW).
$p_i^{\max}$, $p_i^{\min}$	Upper and lower bounds of HVAC's power in building i (MW).
$R_{t,i}^{\text{da}}$	Day-head regulation capacity offer of building i in hour t (MW).
r_t^{m}, r_t^{rc}	unit revenues for regulation capacity and millage in hour t (\$/MW).
γ_n, γ_n^R	Intercepts of lines introduced by piecewise linearization.
ϵ	Risk parameter.
η_t	Unit price for electricity purchasing (\$/MWh).
$\theta_i^{\max}$, $\theta_i^{\min}$	Upper and lower bounds of the thermal comfort region in building i (°C).
θ_t^{out}	Outdoor temperature at hour t (°C).
λ_n, λ_n^R	slopes of lines introduced by piecewise linearization.
$\boldsymbol{\mu}_j$, π_j, $\boldsymbol{\Sigma}_j$	Expectation, weighted coefficient, and covariance matrix of the j-th Gaussian component.
σ_k	Stand derivation of the k-th element of $\boldsymbol{\xi}_j$.

Uncertainties

m_t	Regulation millage in hour t.
$s_{t,l}$	The l-th regulation signal in hour t.
$\overline{u}_\tau$, $\underline{u}_\tau$	Auxiliary uncertain parameters.
$\boldsymbol{\xi}_j$	The j-th Gaussian component introduced by Gaussian mixture model.
$\boldsymbol{\omega}$	Generic form of uncertainties.

Variables

EC_t	Total cost in hour t (\$).
$p_{t,l,i}^{\text{HV}}$	Actual power of the HVAC system in building i at the l-th time slot in hour t (MW).
$p_{t,i}^{\text{ha}}$	Hour-ahead power schedule of the HVAC system in building i (MW).
$q_{t,l,i}$	Cooling supply to building i at the l-th time slot in hour t (MW).
$R_{t,i}^{\text{ha}}$	Hour-ahead regulation capacity offer of building i in hour t (MW).
$\boldsymbol{x}$	Generic decision variable.
y_j	Auxiliary variable for reformulating chance constraints.
$\boldsymbol{z}$	Auxiliary binary variable introduced by piecewise linearization.
$\theta_{t,l,i}^{\text{in}}$	Indoor temperature of building i in the l-th time slot in hour t (°C).
ρ_j	Auxiliary variable for expressing vector $\boldsymbol{\alpha}(\boldsymbol{x})$ in an exponential manner.

This paper is funded in part by the Science and Technology Development Fund, Macau SAR (File no. SKL-IOTSC(UM)-2021-2023, and File no. 0003/2020/AKP), and in part by the National Natural Science Foundation of China under Grant 52007200. (Corresponding author: *Hongcai Zhang*.)

G. Chen, H. Zhang, H. Hui, and Y. Song are with the State Key Laboratory of Internet of Things for Smart City and Department of Electrical and Computer Engineering, University of Macau, Macao, 999078 China; H. Zhang is also with the Smart City Research Center, Zhuhai UM Science & Technology Research Institute, Zhuhai, 519031 China (email: hczhang@um.edu.mo).

I. INTRODUCTION

The growing penetration of renewable energies in power systems reduces fossil fuel consumption and carbon emissions.

However, the intermittent and stochastic characteristics of renewable energies may lead to supply and demand imbalance, which severely threatens the stability of power systems [1]. To support the stable and economic operation of power systems, more demand-side flexible resources are needed for providing regulation services by controlling their power schedules [2].

Heating, ventilation, and air conditioning (HVAC) loads are one of the most promising demand-side flexible resources because of the building's inherent ability to store heating/cooling power [3], [4]. To utilize the flexible HVAC systems for regulation services, the corresponding regulation capacity offers need to be reported to the power market in advance [5]. The power market can collect all the regulation capacity offers to design regulation signals. Then, HVAC systems can follow regulation signals and adjust their power scheduling to earn regulation revenue [6], [7]. Since this revenue is in proportion to the regulation capacity offers, increasing attention has been paid to quantifying the potential regulation capacity for HVAC systems. For example, reference [8] proposed a geometric approach to characterize the aggregated regulation capacity of HVAC systems. Reference [9] proposed a robust-based method to quantify the HVAC's regulation capacity in distribution networks. Reference [10] leveraged deep learning techniques to develop a model-free method to determine the best regulation capacity for HVAC systems. Because one HVAC system's thermal inertia is limited, its regulation capacity can get affected by uncertain and biased regulation operations. However, most of the aforementioned papers do not consider the impacts of regulation signals, which may overly estimate HVAC's regulation capacity and violate the corresponding building's indoor thermal comforts.

Nevertheless, taking regulation signals into consideration is challenging because they are highly stochastic and unpredictable. To address this issue, some papers treat signals as uncertainties and leverage robust optimization to design perfectly safe scheduling strategies for flexible resources [11], [12]. However, robust optimization methods do not allow any constraint violation for all realizations of uncertainties [13], so their solutions are usually overly conservative. An alternative choice is chance-constrained programming (CCP). CCP only requires that the probability of meeting critical constraints is above a certain level, while low-probability constraint violations are allowed [14], [15]. Unfortunately, applying CCP still faces the following two challenges:

1) According to our statistics for the whole-year regD signals from the PJM regulation market in 2020 [5], the uncertain regulation signals do not follow Gaussian distribution. However, the most widely used reformulation for CCP is based on the Gaussian assumption of uncertainties [15]–[17]. Thus, applying this Gaussian-assumption-based method may lead to infeasible solutions. Some other scholars proposed distributionally robust chance-constrained methods (DRCC) to handle non-Gaussian uncertainties. Based on some statistical information of the uncertainties, e.g., moments or Wasserstein distances, DRCC constructs an ambiguity set to cover possible distributions and requires the probabilistic constraints to be robust to this ambiguity set [18], [19]. Since ambiguity sets can cover non-Gaussian distributions, DRCC can handle non-Gaussian uncertainties and has been used for scheduling flexible sources under uncertain regulation signals [20], [21]. However, DRCC may result in overly conservative solutions and dramatically reduce the regulation revenue because the ambiguity set may cover some distributions that are much different from the true one.
2) The regulation signal updates frequently (e.g. RegD signal updates every two seconds in the PJM market [5]). Since the HVAC power needs to be regulated to follow these signals, numerous constraints need to be involved in order to ensure thermal comforts corresponding to each regulation signal. This will make the CCP computationally expensive, especially for the methods that involve many additional variables and constraints for each chance constraint (e.g. scenario approach [22], [23], sample average approximation [24], conditional value-at-risk approximation [25]).

To overcome the first challenge, several papers combined the Gaussian mixture model (GMM) with CCP. GMM is a universal approximator of probability densities, and any non-Gaussian distribution can be approximately fitted with a finite number of Gaussian components [26]. In [27], [28], GMM was used to fit the non-Gaussian renewable energy uncertainties. Then, the chance constraints were directly reformulated into tractable forms based on the quantile of uncertainties. In [29], an online-offline double-track approach was developed to accelerate the fitting of GMM for the uncertainties of gas demands. However, these GMM-based methods are only suitable for the chance constraints with right-hand side (RHS) uncertainties.[1] Considering that the regulation signal uncertainties are on the left-hand side (LHS), these methods are still inapplicable.

Unlike the first challenge, there are only very few papers that have tried to tackle the second challenge. In fact, most published papers, including [11], [12], [21], only required the satisfaction of constraints with low temporal resolutions, while the intermediate variable variations between two neighboring time slots were ignored. Thus, although this manner can reduce the constraint number, it cannot always guarantee feasibility. For example, reference [21] only restricted that constraints should be satisfied every five minutes, e.g., at $t \in \{5\text{min}, 10\text{min}, \cdots, 55\text{min}, 60\text{min}\}$. However, it cannot guarantee that there is no violation within every five minutes, e.g., at t=7min.

To overcome the aforementioned two challenges, we propose a tractable chance-constrained model to optimize the regulation capacity offering for HVAC systems. The specific contributions are threefold:

1) We propose a chance-constrained model to determine the hour-ahead regulation capacity offers for HVAC systems in the PJM market. This model considers the impacts of non-Gaussian uncertainties from regulation signals. Moreover, the thermodynamic constraints are

[1]Consider a linear constraint $\boldsymbol{a}^\top \boldsymbol{x} \leq b$. If the uncertainty is the vector $\boldsymbol{a}$, then we call it "left-hand side (LHS) uncertainty"; if the uncertainty is in the constant b, then it is called "right-hand side (RHS) uncertainty".

built according to the updated frequency of regulation signals (i.e. every two seconds) so that indoor thermal comforts can be properly maintained.

2) To address the intractability from non-Gaussian uncertainties on the LHS, we propose a mixture-model-based convexification method. It first leverages GMM to reformulate each chance constraint with non-Gaussian LHS uncertainties into a deterministic non-convex form. Then, this non-convex reformulation is equivalently re-expressed as an exponential form. Based on piecewise linearization, we further prove that this exponential form can be approximated by a SOCP constraint with only a few binary variables, which can be efficiently solved by off-the-shelf solvers. Moreover, the optimality loss introduced by this approximation is also marginal under mild conditions. To the best of our knowledge, this is the first time that GMM-based methods can be extended to chance constraints with LHS uncertainties from regulation signals.
3) To reduce the computational burden brought by the huge number of thermodynamic constraints, we propose a temporal compression method. In this method, we first estimate the maximum and minimum indoor temperatures over a long time duration based on the monotonicity of the thermodynamic model. Then, by restricting the estimated extreme temperatures in the comfortable range, all thermodynamic constraints in this long time duration can be replaced by only a few ones, which significantly enhances computational efficiency.

Besides, numerical experiments are conducted to validate the benefits of the proposed approach. The results show that it can achieve much less conservative results than the state-of-art DRCC methods with desirable optimality.

The remaining parts are organized as follows. Section II describes the problem formulation. Section III presents the details of the proposed mixture-model-based convexification method. Section IV shows simulation results and Section V concludes this paper.

II. Problem Formulation

We consider an aggregator strategically operating a couple of HVAC systems to provide regulation capacities in the PJM market. The aggregator offers the regulation capacity to the market, and the market sends regulation signals to the aggregator to guide the power regulation of HVAC systems, as shown in Fig. 1(a). As required, the aggregator needs to offer the regulation capacity to the market at least one hour ahead, as shown in Fig. 1(b). For example, the capacity offer for 3:00pm-4:00pm should be reported before 2:00pm. In order to maximize its regulation revenue, the aggregator need to properly design the power schedule of HVAC systems, i.e., p_t^{ha}, and accurately estimate their corresponding regulation capacities, i.e., R_t^{ha}. Because buildings have limited thermal inertia, their regulation capacities are significantly affected by the uncertain regulation signals, which should be explicitly considered. This paper aims to determine the optimal hour-ahead regulation offers for HVAC systems under the impact

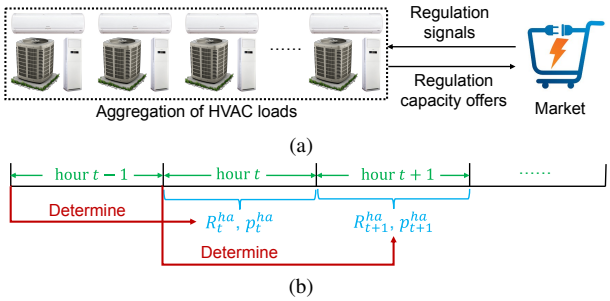


Fig. 1. (a) System configuration of scheduling HVAC systems to provide regulation services and (b) schema of hour-ahead regulation capacity offers in the PJM market.

of uncertainties from regulation signals. Meanwhile, thermal comforts should be also maintained to guarantee the HVAC systems' quality of services.

Since the regulation signal in PJM updates every two seconds, we divide the operating hour into 1800 time slots with $l \in \mathcal{L}$ as their indexes and $\Delta L = 2s$ as the time step size. By using $i \in \mathcal{I}$ to index HVAC systems, the thermodynamic model of one building can be expressed as:

$$C_i \frac{d\theta_i^{\text{in}}}{dt} = g_i(\theta^{\text{out}} - \theta_i^{\text{in}}) + h_i - q_i, \forall i \in \mathcal{I}, \tag{1}$$

where C_i and g_i are the building heat capacity and heat transfer coefficient between indoor and outdoor environments, respectively; θ_i^{in} and θ^{out} are the indoor and outdoor temperatures, respectively; h_i and q_i denote the heat load contributed by indoor sources (e.g. electronic devices) and cooling supply from HVAC systems, respectively.

Since the outdoor temperature θ^{out} and indoor heat load h_i vary slowly, we assume that they keep unchanged in the operating hour. When the HVAC system participates in regulation services, their power need to respond to the regulation signals. Thus, the actual HVAC power at the l-th time slot in hour t, i.e., $p_{t,l,i}^{\text{HV}}$, and corresponding cooling supply, i.e., $q_{t,l,i}$, are expressed as:

$$p_{t,l,i}^{\text{HV}} = p_{t,i}^{\text{ha}} - R_{t,i}^{\text{ha}} s_{t,l}, \quad \forall l \in \mathcal{L}, \forall i \in \mathcal{I}, \tag{2}$$

$$q_{t,l,i} = \text{COP}_i \cdot p_{t,l,i}^{\text{HV}}, \quad \forall l \in \mathcal{L}, \forall i \in \mathcal{I}, \tag{3}$$

where $s_{t,l}$ is the l-th signal in hour t; COP_i is the coefficient of performance of the HVAC system in building i. Eqs. (2)-(3) indicate that the cooling supply keeps constant in each Δl, so (1) can be directly integrated from time l to $l + \Delta l$:

$$\begin{aligned}\theta_{t,l,i}^{\text{in}} =& a_i^{\text{in}} \theta_{t,l-1,i}^{\text{in}} + a_i^{\text{out}} \theta_t^{\text{out}} + a_i^{\text{h}} h_t \\ &+ a_i^{\text{q}} (p_{i,t}^{\text{ha}} - R_{t,i} s_{t,l-1}), \quad \forall l \in \mathcal{L}/\{0\}, \forall i \in \mathcal{I},\end{aligned} \tag{4}$$

where parameters a_i^{in}, a_i^{out}, a_i^{h}, and a_i^{q} are calculated by:

$$\begin{cases} a_i^{\text{in}} = e^{-\frac{g_i}{C_i}\Delta l}, \quad a_i^{\text{out}} = 1 - e^{-\frac{g_i}{C_i}\Delta l}, \\ a_i^{\text{h}} = \frac{1}{g_i}(1 - e^{-\frac{g_i}{C_i}\Delta l}), \ a_i^{\text{q}} = -\frac{\text{COP}_i}{g_i}(1 - e^{-\frac{g_i}{C_i}\Delta l}). \end{cases} \tag{5}$$

Eq. (4) indicates that buildings have thermal inertia: The indoor temperature $\theta_{t,l,i}^{\text{in}}$ is affected by the heat loads and cooling supplies in previous time slots. In other words, buildings have inherent ability to store heating/cooling energy. Thus, we can properly control HVAC systems' power schedules following regulation signals to provide regulation services while satisfy-

ing thermal comfort requirement [30]. Since $s_{t,l}$ is uncertain, both the actual HVAC power $p^{\text{HV}}_{t,l,i}$ and indoor temperature $\theta^{\text{in}}_{t,l,i}$ are also uncertain according to (2)-(4). Considering that small thermal discomforts can be temporarily tolerated, we employ CCP to describe the thermal comfort requirement. CCP controls the feasible region of the optimization problem to guarantee that the probability of meeting critical constraints is above a certain level [14], [15], while low-probability constraint violation is allowable. By applying CCP, desirable energy efficiency can be achieved, while customer experiences can be also guaranteed with a high probability. The chance-constrained thermal comfort requirement is expressed as:

$$\begin{cases} \mathbb{P}\left(\theta^{\text{in}}_{t,l,i} \le \theta^{\max}_{i}\right) \ge 1-\epsilon, \\ \mathbb{P}\left(\theta^{\text{in}}_{t,l,i} \ge \theta^{\min}_{i}\right) \ge 1-\epsilon, \end{cases} \quad \forall l \in \mathcal{L}, \quad \forall i \in \mathcal{I}, \tag{6}$$

where $\theta^{\max}$ and $\theta^{\min}$ are the upper and lower bounds of the thermal comfort range, respectively; ϵ is the risk parameter. The device limit requires that the HVAC power should always stay in the allowable range, so a robust manner is used to describe this limitation:

$$\begin{cases} \max_{s_{t,l}} p^{\text{HV}}_{t,l,i} \le p^{\max}_{i}, \\ \min_{s_{t,l}} p^{\text{HV}}_{t,l,i} \ge p^{\min}_{i}, \end{cases} \quad \forall l \in \mathcal{L}, \quad \forall i \in \mathcal{I}, \tag{7}$$

where $p^{\max}_i$ and $p^{\min}_i$ are the upper and lower bounds of the HVAC power, respectively. Since the regulation signal is restricted in [-1,1], Eq. (7) can be reformulated as follows according to (2):

$$p^{\text{ha}}_{t,i} + R^{\text{ha}}_{t,i} \le p^{\max}_{i}, \quad p^{\text{ha}}_{t,i} - R^{\text{ha}}_{t,i} \ge p^{\min}_{i}, \quad \forall i \in \mathcal{I}. \tag{8}$$

In the PJM market, the hour-ahead regulation capacity offers are only allowed to be reduced from the corresponding day-ahead offers, i.e., $R^{\text{da}}_{t,i}$, as follows:

$$0 \le R^{\text{ha}}_{t,i} \le R^{\text{da}}_{t,i}, \quad \forall i \in \mathcal{I}, \tag{9}$$

where $R^{\text{da}}_{t,i}$ is determined one day ahead, so it is a known parameter when we optimize the hour-ahead offer $R^{\text{ha}}_{t,i}$.

Our objective is to minimize the total cost EC_t, which equals to the energy cost minus the regulation revenue:

$$EC_t = \sum_i \sum_{l \in \mathcal{L}} \eta_t p^{\text{HV}}_{t,i,l} \Delta l - \sum_i (r^{\text{rc}}_t + r^m_t m_t) R^{\text{ha}}_{t,i}, \tag{10}$$

where the first and second terms on the RHS of (10) represent the energy cost of HVAC systems and revenue from the HVAC power regulation. Symbol $\mathbb{E}$ denotes the expectation operator; η_t is the price for electricity purchasing at hour t; $\Delta t = 1h$ is the operating time duration; r^{rc}_t and r^m_t are the unit revenues for regulation capacity and millage, respectively; m_t is the regulation millage, which is defined as:

$$m_t = \sum_l |s_{t,l+1} - s_t|. \tag{11}$$

Considering s_t is uncertain regulation signals, the regulation millage m_t is also uncertain in this optimization problem.

Finally, the optimization problem is formulated as:

$$\min_{R^{\text{ha}}_{t,i}, p^{\text{ha}}_{t,i}, \forall l, \forall i} \mathbb{E}(EC_t), \text{ s.t.: Eqs. (2)-(6) and (8)-(10).} \quad \textbf{(P1)}$$

Solving **P1** is quite challenging. On the one hand, due to the high update frequency of regulation signals, Eq. (6) introduces numerous thermodynamic constraints, leading to computational intractability. On the other hand, in (6), the signal uncertainties do not follow the Gaussian distribution. Therefore, the Gaussian-assumption-based models used in [15]–[17] can not be directly applied. Moreover, they are LHS uncertainties due to the multiplication term $R_{t,i}s_{t,l-1}$, where $R_{t,i}$ is a decision variable and $s_{t,l-1}$ is an uncertain parameter. As a result, the GMM-based methods proposed in [27]–[29] are also inapplicable.[2]

III. Solution Methodology

To overcome the aforementioned challenges, we first propose a temporal compression approach to reduce the thermodynamic constraint number. Then, a mixture-model-based convexification method is developed to reformulate the chance constraint with non-Gaussian uncertainties, i.e., Eq. (6), into second-order cone programming (SOCP) forms. For simplicity, we omit the subscripts t and i in this section.

A. Temporal compression

The key idea of the proposed temporal compression is to estimate the maximum and minimum indoor temperatures in the operating hour. Then, by restricting these extreme indoor temperatures in the comfort range, the thermodynamic constraint number in (6) can be significantly reduced. For example, the first line in (6), i.e., $\mathbb{P}\left(\theta^{\text{in}}_l \le \theta^{\max}\right) \ge 1-\epsilon, \forall l \in \mathcal{L}$, contains 1800 constraints. However, once the maximum indoor temperature in the operating hour, i.e., $\overline{\theta} = \max_{l \in \mathcal{L}} \theta_l$ is estimated, the 1800 constraints can be replaced by only one single constraint $\mathbb{P}\left(\overline{\theta} \le \theta^{\max}\right) \ge 1-\epsilon$. Thus, the key problem is how to estimate these maximum and minimum indoor temperatures.

According to (4), the l-th indoor temperature in hour t, i.e., θ^{in}_l, is expressed as:

$$\begin{aligned} \theta^{\text{in}}_l =& a^{\text{in}}\theta^{\text{in}}_{l-1} + a^{\text{out}}\theta^{\text{out}} + a^{\text{h}}h + a^{\text{q}}(p^{\text{ha}} - R^{\text{ha}}s_{l-1}) \\ =& a^{\text{in}}\left(a^{\text{in}}\theta^{\text{in}}_{l-2} + a^{\text{out}}\theta^{\text{out}} + a^{\text{h}}h + a^{\text{q}}(p^{\text{ha}} - R^{\text{ha}}s_{l-2})\right) \\ & + a^{\text{out}}\theta^{\text{out}} + a^{\text{h}}h + a^{\text{q}}(p^{\text{ha}} - R^{\text{ha}}s_{l-1}) \\ =& \cdots\cdots \\ =& R^{\text{ha}}[\boldsymbol{A}\boldsymbol{s}]_l + f_l, \quad \forall l \in \mathcal{L}/\{0\}, \end{aligned} \tag{12}$$

[2]Consider a chance constraint $\mathbb{P}(\boldsymbol{\alpha}^\intercal \boldsymbol{x} \le b) \ge 1-\epsilon$. If it only contains RHS uncertainties, i.e., b is a non-Gaussian uncertainty while a is a deterministic parameter, then it can be equivalently reformulated as $\alpha^\intercal \boldsymbol{x} \le F_b^{-1}(\epsilon)$, where $F_b^{-1}(\cdot)$ is the inverse of the cumulative density function of b. In this case, the GMM-based methods proposed in [27]–[29] can employ GMM to estimating the value of $F_b^{-1}(\cdot)$ in advance. If the previous chance constraint contains LHS uncertainty, i.e., a is uncertain, it is equivalent to $F^{-1}_{\boldsymbol{\alpha}^\intercal \boldsymbol{x}}(1-\epsilon) \le b$. However, it is hard for GMM to estimate the value of $F^{-1}_{\boldsymbol{\alpha}^\intercal \boldsymbol{x}}(1-\epsilon)$ before solving the optimization problem because it is relevant to the decision variable $\boldsymbol{x}$. As a result, the GMM-based methods are inapplicable to the case with LHS uncertainties.

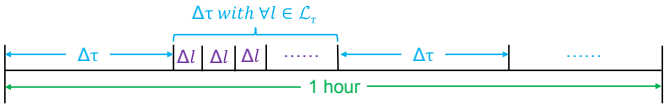


Fig. 2. Schematic diagram for uniformly splitting the operating hour into multiple shorter duration $\Delta\tau$, where $\Delta l = 2s$ denotes the update interval of regulation signals; set $\mathcal{L}_\tau$ contains the indexes of l in the τ-th time duration.

where $\boldsymbol{A}$ is a coefficient matrix and defined as follows:

$$\boldsymbol{A} = -a^{\mathrm{q}} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_1^{\mathrm{in}} & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ (a_1^{\mathrm{in}})^{|\mathcal{L}|-1} & (a_1^{\mathrm{in}})^{|\mathcal{L}|-2} & \cdots & 1 \end{bmatrix}. \tag{13}$$

Symbol $|\mathcal{L}|$ represents the length of set $\mathcal{L}$. Vector $\boldsymbol{s}$ represents $[s_0, s_1, \cdots, s_{|\mathcal{L}|-1}]$. Symbol $[\boldsymbol{As}]_l$ denotes the l-th element of the product $\boldsymbol{As}$. Function f_l is defined as:

$$\begin{aligned} f_l =& (a^{\mathrm{in}})^l \theta_0^{\mathrm{in}} + a^{\mathrm{out}} \frac{1-(a^{\mathrm{in}})^l}{1-a^{\mathrm{in}}} \theta^{\mathrm{out}} \\ &+ a^{\mathrm{h}} \frac{1-(a^{\mathrm{in}})^l}{1-a^{\mathrm{in}}} h + a^{\mathrm{q}} \frac{1-(a^{\mathrm{in}})^l}{1-a^{\mathrm{in}}} p^{\mathrm{ha}}. \end{aligned} \tag{14}$$

Based on (12), we must have:

$$\begin{cases} \max_{l\in\mathcal{L}} \theta_l^{\mathrm{in}} \le \max_{l\in\mathcal{L}} f_l + R^{\mathrm{ha}} \max_{l\in\mathcal{L}} [\boldsymbol{As}]_l, \\ \min_{l\in\mathcal{L}} \theta_l^{\mathrm{in}} \ge \min_{l\in\mathcal{L}} f_l + R^{\mathrm{ha}} \min_{l\in\mathcal{L}} [\boldsymbol{As}]_l. \end{cases} \tag{15}$$

Thus, the maximum and minimum indoor temperatures can be approximated by the RHS terms of (15). However, since $\max_{l\in\mathcal{L}} f_l$ and $\max_{l\in\mathcal{L}} [\boldsymbol{As}]_l$ may appear at different moments, Eq. (15) may lead to overly conservative solutions. To mitigate this conservativeness, we uniformly split the operating hour $\Delta t = 1h$ into multiple shorter time duration $\Delta\tau$, as shown in Fig. 2. Then, the extreme indoor temperatures in each $\Delta\tau$ can be approximated by the RHS terms of (16):

$$\begin{cases} \max_{l\in\mathcal{L}_\tau} \theta_l^{\mathrm{in}} \le \max_{l\in\mathcal{L}_\tau} f_l + R^{\mathrm{ha}} \max_{l\in\mathcal{L}_\tau} [\boldsymbol{As}]_l, \\ \min_{l\in\mathcal{L}_\tau} \theta_l^{\mathrm{in}} \ge \min_{l\in\mathcal{L}_\tau} f_l + R^{\mathrm{ha}} \min_{l\in\mathcal{L}_\tau} [\boldsymbol{As}]_l, \end{cases} \quad \forall \tau \in \mathcal{T}, \tag{16}$$

where $\tau \in \mathcal{T}$ is the index of the shorter duration; $\mathcal{L}_\tau$ denotes the index set of l in the τ-th duration, which is obtained by uniformly splitting $\mathcal{L}$ into $|\mathcal{T}|$ parts. By adding a maximum/minimum operator over $\tau \in \mathcal{T}$ on both sides of (16), we have:

$$\begin{cases} \max_{l\in\mathcal{L}} \theta_l^{\mathrm{in}} \le \max_{\tau\in\mathcal{T}} \{\max_{l\in\mathcal{L}_\tau} f_l + R^{\mathrm{ha}} \max_{l\in\mathcal{L}_\tau} [\boldsymbol{As}]_l\}, \\ \min_{l\in\mathcal{L}} \theta_l^{\mathrm{in}} \ge \min_{\tau\in\mathcal{T}} \{\min_{l\in\mathcal{L}_\tau} f_l + R^{\mathrm{ha}} \min_{l\in\mathcal{L}_\tau} [\boldsymbol{As}]_l\}. \end{cases} \tag{17}$$

Then, the extreme indoor temperatures in the operating hour can be estimated by the RHS terms in (17).

Proposition 1. *The approximation in (17) is less conservative than that in (15).*

Proof: See Appendix A.

Based on (14), function f_l is monotone with respect to l because θ_0^{in}, θ^{out}, h, and p^{ha} keep unchanged in the operating hour. Thus, the maximum and minimum values of f_l in each $\Delta\tau$ must appear at the boundaries, i.e., at $l = \tau\frac{|\mathcal{L}|}{|\mathcal{T}|}$ or $l = (\tau+1)\frac{|\mathcal{L}|}{|\mathcal{T}|}$. As for the terms $\max_{l\in\mathcal{L}_\tau}[\boldsymbol{As}]_l$ and $\min_{l\in\mathcal{L}_\tau}[\boldsymbol{As}]_l$, their values are uncertain but independent of decision variables. Therefore, we can directly treat them as two new uncertain parameters, i.e., $\overline{u}_\tau$ and $\underline{u}_\tau$, as follows:

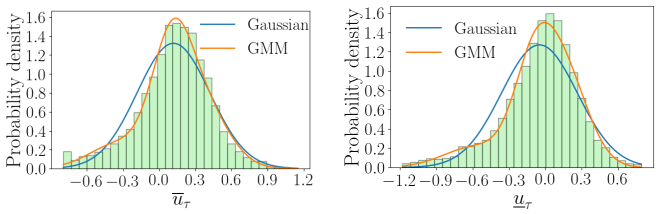
$$\overline{u}_\tau = \max_{l\in\mathcal{L}_\tau} [\boldsymbol{As}]_l, \quad \underline{u}_\tau = \min_{l\in\mathcal{L}_\tau} [\boldsymbol{As}]_l, \quad \forall \tau \in \mathcal{T}. \tag{18}$$

Finally, Eq. (6) can be replaced by:

$$\begin{cases} \mathbb{P}\left(f_{l=\tau\frac{|\mathcal{L}|}{|\mathcal{T}|}} + \overline{u}_\tau R^{\mathrm{ha}} \le \theta^{\max} \right) \ge 1-\epsilon, \\ \mathbb{P}\left(f_{l=(\tau+1)\frac{|\mathcal{L}|}{|\mathcal{T}|}} + \overline{u}_\tau R^{\mathrm{ha}} \le \theta^{\max} \right) \ge 1-\epsilon, \\ \mathbb{P}\left(f_{l=\tau\frac{|\mathcal{L}|}{|\mathcal{T}|}} + \underline{u}_\tau R^{\mathrm{ha}} \ge \theta^{\min} \right) \ge 1-\epsilon, \\ \mathbb{P}\left(f_{l=(\tau+1)\frac{|\mathcal{L}|}{|\mathcal{T}|}} + \underline{u}_\tau R^{\mathrm{ha}} \ge \theta^{\min} \right) \ge 1-\epsilon, \end{cases} \quad \forall \tau \in \mathcal{T}. \tag{19}$$

Based on (17), any feasible solution of (19) must be also feasible for the chance constraint (6). Thus, Eq. (19) is a safe approximation of (6). Moreover, the thermodynamic constraint number is reduced from $2\cdot|\mathcal{L}|$ to $4\cdot|\mathcal{T}|$ (note that $|\mathcal{T}| \ll |\mathcal{L}|$), which significantly reduces the computational burden.

Based the whole-year regD signals from PJM in 2020 [5], the uncertainties $\overline{u}_\tau$ and $\underline{u}_\tau$ do not follow Gaussian distribution, as shown in Fig. 3. Moreover, these uncertainties are on the LHS in (19) because they are multiplied with a decision variable R^{ha}. Thus, Eq. (19) is still hard to deal with.



(a) (b)

Fig. 3. The probability distribution of (a) $\overline{u}_\tau$ and (b) $\underline{u}_\tau$ with $\tau = 2$ when $|\mathcal{T}| = 10$. The blue and orange lines represent the fitting results of the Gaussian-assumption-based model and GMM, respectively. Obviously, the uncertainties do not follow Gaussian distribution but can be well fitted by GMM.

B. Mixture-model-based convexification approach

To address the intractability caused by the non-Gaussian LHS uncertainties in (19), a mixture-model-based convexification approach is proposed. We first introduce GMM to fit the original non-Gaussian uncertainties with multiple Gaussian distributions. Then, we reformulate (19) into deterministic non-convex forms. Finally, we propose tractable SOCP approximations for these non-convex constraints.

1) Introduction of GMM: GMM can approximate the distribution of non-Gaussian variable $\boldsymbol{\omega}$ with multiple Gaussian distributions [26]:

$$p^{\mathrm{NG}}(\boldsymbol{\omega}) = \sum_{j\in\mathcal{J}} \pi_j p\left(\boldsymbol{\xi}_j | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j\right), \tag{20}$$

where $p^{\mathrm{NG}}(\boldsymbol{\omega})$ is the probability density function (PDF) of $\boldsymbol{\omega}$; $j \in \mathcal{J}$ is the index of the Gaussian component; π_j is the weight of component j, and $\sum_{j\in\mathcal{J}} \pi_j = 1$; $p(\boldsymbol{\xi}_j|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ represents the PDF of a Gaussian uncertainty $\boldsymbol{\xi}_j$; $\boldsymbol{\mu}_j$ and $\boldsymbol{\Sigma}_j$ are the expectation and covariance of $\boldsymbol{\xi}_j$, respectively. Based on the historical samples of $\boldsymbol{\omega}$, the three parameters π_j, $\boldsymbol{\mu}_j$, and Σ_j can be estimated based on the Expectation Maximization algorithm [15]–[17]. Fig. 3 provides an example to demonstrate the excellent fitting power of GMM.

2) Deterministic reformulations of chance constraints: The generic form of the chance constraints in (19) can be expressed as follows:[3]

$$\mathbb{P}\left(\boldsymbol{\alpha}(\boldsymbol{x})^{\intercal}\boldsymbol{\omega} \le \beta(\boldsymbol{x})\right) \ge 1-\epsilon. \tag{21}$$

The detail expressions of $\boldsymbol{\alpha}(\boldsymbol{x})$ and $\beta(\boldsymbol{x})$ for each constraint is given in Appendix B.

To reformulate (21), we introduce the following **Lemma**.

Lemma 1. *If the PDF of the uncertainty $\boldsymbol{\omega}$ is approximated by GMM, i.e., Eq. (20), then we have [31]*

$$\mathbb{P}\left(\boldsymbol{\alpha}(\boldsymbol{x})^{\intercal}\boldsymbol{\omega} \le \beta(\boldsymbol{x})\right) = \sum_{j\in\mathcal{J}} \pi_j \mathbb{P}\left(\boldsymbol{\alpha}(\boldsymbol{x})^{\intercal}\boldsymbol{\xi}_j \le \beta(\boldsymbol{x})\right). \tag{22}$$

By introducing an auxiliary variable y_j for each Gaussian component, Eq. (21) can be converted into:

$$\mathbb{P}\left(\boldsymbol{\alpha}(\boldsymbol{x})^{\intercal}\boldsymbol{\xi}_j \le \beta(\boldsymbol{x})\right) \ge y_j, \quad \forall j \in \mathcal{J}, \tag{23}$$

$$\sum_{j\in\mathcal{J}} \pi_j y_j \ge 1-\epsilon, \tag{24}$$

where $y_j \in [0,1]$ because it is used to restrict a probability. The uncertainty $\boldsymbol{\xi}_j$ is the Gaussian component of GMM and follows Gaussian distribution. Thus, based on the Gaussian-assumption-based reformulations used in [15]–[17], Eq. (23) can be further reformulated into the following deterministic non-convex form when $y_j \ge 0.5$:

$$\Phi^{-1}(y_j)\sqrt{\boldsymbol{\alpha}(\boldsymbol{x})^{\intercal}\boldsymbol{\Sigma}_j\boldsymbol{\alpha}(\boldsymbol{x})} + \boldsymbol{\alpha}(\boldsymbol{x})^{\intercal}\boldsymbol{\mu}_j \le \beta(\boldsymbol{x}), \forall j \in \mathcal{J}, \tag{25}$$

$$y_j \ge 0.5, \forall j \in \mathcal{J}, \tag{26}$$

where $\Phi^{-1}(\cdot)$ is the inverse of the cumulative distribution function of the stand normal distribution. Note that an extra constraint (26) is introduced, which may bring additional conservativeness.

Remark 1. *The reformulation (25) is intractable and cannot be effectively handled by off-the-shelf solvers. On the one hand, its first term $\Phi^{-1}(y_j)\sqrt{\boldsymbol{\alpha}(\boldsymbol{x})^{\intercal}\boldsymbol{\Sigma}_j\boldsymbol{\alpha}(\boldsymbol{x})}$ is non-convex because both y_j and $\boldsymbol{\alpha}(\boldsymbol{x})$ are variables. On the other hand, it is hard to give an analytical formulation for function $\Phi^{-1}(\cdot)$.*

3) Convexification for deterministic reformulations: We propose a convexification method to approximately reformulate the deterministic reformulation (25) into a tractable SOCP form. Firstly, observing that all elements of $\boldsymbol{\alpha}(\boldsymbol{x})$ in our problem are always nonegative (See Appendix B), we can re-express each element of $\boldsymbol{\alpha}(\boldsymbol{x})$ in an exponential manner:

$$\alpha_k = e^{\rho_k}, \quad \forall k \in \mathcal{K}, \tag{27}$$

where α_k is the k-th element of $\boldsymbol{\alpha}(\boldsymbol{x})$; ρ_k is an auxiliary variable; $\mathcal{K}$ is the corresponding index set. Since $y_j \ge 0.5$, function $\Phi^{-1}(y_j)$ is nonegative, so it can be also expressed as an exponential form, i.e., $\Phi^{-1}(y_j) = e^{\ln\Phi^{-1}(y_j)}$. According to the statistics for the whole-year regulation signals in 2020 [5], the off-diagonal elements in the covariance matrix $\boldsymbol{\Sigma}_j$ is close to zero and much smaller than the diagonal ones. Thus, we can use zero to replace the off-diagonal elements of $\boldsymbol{\Sigma}_j$. Then, by substituting the above exponential expressions, the first term in (25) can be converted into an L2-norm form:

$$\begin{aligned}\Phi^{-1}(y_j)\sqrt{\boldsymbol{\alpha}(\boldsymbol{x})^{\intercal}\boldsymbol{\Sigma}_j\boldsymbol{\alpha}(\boldsymbol{x})} &= \sqrt{\sum_k (\sigma_k e^{\rho_k+\ln\Phi^{-1}(y_j)})^2} \\ &= \|\sigma_k e^{\rho_k+\ln\Phi^{-1}(y_j)}, \quad \forall k\in\mathcal{K}\|_2, \quad \forall j\in\mathcal{J}, \end{aligned} \tag{28}$$

where σ_k is the standard deviation of the k-th element of $\boldsymbol{\omega}$, and its value is always non-negative. Since both the exponential and L2-norm functions are convex and element-wise monotonically increasing, the L2-norm term in (28) is convex if its power, $\rho_k + \ln\Phi^{-1}(y_j)$, is convex according to convex condition for composite functions [32]. Unluckily, this power is non-convex in its domain. Nevertheless, we can find a piecewise-linearization-based safe approximation to convexify this non-convex power based on the following **Proposition**.

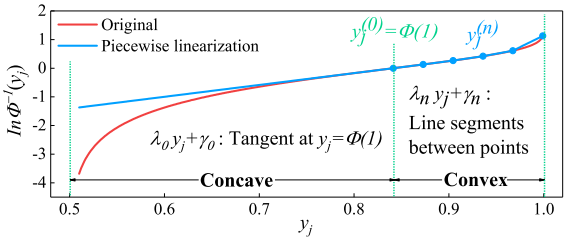


Fig. 4. Graphs of function $\ln\Phi^{-1}(y_j)$ (red line) and the proposed inner approximation based on piecewise linear function $\max_{n\in\mathcal{N}}\{\lambda_n y_j+\gamma_n\}$ (blue line). The blue dots, i.e., $(y_j^{(n)}, \ln\Phi^{-1}(y_j^{(n)}))$, are points on function $\ln\Phi^{-1}(y_j)$. The first line, i.e., $\lambda_0 y_j+\gamma_0$ is the tangent of $\ln\Phi^{-1}(y_j)$ at $y_j=\Phi(1)$. The rest lines are segments between two adjacent blue dots. Note the piecewise linear function $\max_{n\in\mathcal{N}}\{\lambda_n y_j+\gamma_n\}$ is convex.

Proposition 2. *In the domain of function $\ln\Phi^{-1}(y_j)$, i.e., $y_j > 0.5$, we have*

$$\ln\Phi^{-1}(y_j) \le \max_{n\in\mathcal{N}}\{\lambda_n y_j+\gamma_n\}, \quad \forall j\in\mathcal{J}, \tag{29}$$

where $\mathcal{N}=\{0,1,\cdots,N\}$ is the index set of lines. Symbol $\lambda_n y_j+\gamma_n, \forall n\in\mathcal{N}$ represents different lines constructed by piecewise linearization (shown in Fig. 4), i.e., $\lambda_0 y_j+\gamma_0$ is the tangent at $y_j=\Phi(1)$, while the rest lines, i.e., $\lambda_n y_j+\gamma_n, \forall n\in\mathcal{N}/\{0\}$, are line segments by connecting two points on function $\ln\Phi^{-1}(y_j)$ in sequence. Note the piecewise linear function $\max_{n\in\mathcal{N}}\{\lambda_n y_j+\gamma_n\}$ is convex because its epigraph is a convex polyhedron.

Proof: See Appendix C.

Observing that both the exponential and L2-norm functions are monotonically increasing, the following inequality holds

[3]Note here we only consider the uncertainties from regulation signals but do not involve the uncertainties from heat loads h_t. Nevertheless, even if h_t is uncertain, we can still express the chance constraint (19) into the generic form (21). As a result, the proposed convexification method can easily handle the uncertainties from h_t.

according to **Proposition** 2:

$$\|\sigma_k e^{\rho_k+\ln\Phi^{-1}(y_j)}, \forall k\in\mathcal{K}\|_2 \leq \|\sigma_k e^{\rho_k+\max_{n\in\mathcal{N}}\{\lambda_n y_j+\gamma_n\}}, \forall k\in\mathcal{K}\|_2, \forall j\in\mathcal{J}. \quad (30)$$

Based on (28) and (30), the deterministic non-convex reformulation (25) can be convexified as the following SOCP form:

$$\|\sigma_k e^{\rho_k+\max_{n\in\mathcal{N}}\{\lambda_n y_j+\gamma_n\}}, \forall k\in\mathcal{K}\|_2+\boldsymbol{\alpha}(\boldsymbol{x})^\intercal\boldsymbol{\mu}_j\leq\beta(\boldsymbol{x}), \quad \forall j\in\mathcal{J}, \quad (31)$$

which is also equivalent to

$$\|\sigma_k e^{\rho_k+\lambda_n y_j+\gamma_n}, \forall k\in\mathcal{K}\|_2+\boldsymbol{\alpha}(\boldsymbol{x})^\intercal\boldsymbol{\mu}_j\leq\beta(\boldsymbol{x}), \quad \forall n\in\mathcal{N}, \quad \forall j\in\mathcal{J}. \quad (32)$$

Based on the convex condition for composite functions [32], the first L2-norm term in (32) is convex because both the exponential and the L2-norm functions are convex and element-wise monotonically increasing. Thus, it is tractable for off-the-shelf solvers. According to (30), the proposed reformulation (32) is a conservative approximation of the non-convex deterministic constraint (25). In other words, the feasible set of (32) is a subset of the feasible set of (25). Thus, the proposed reformulation can guarantee the feasibility of solutions.

Remark 2. *During the convexification for the original chance constraint (21), we introduce two conservative approximations. The first one is the additional constraint (26), while the second one is the piecewise linearization used in (30)-(32). These conservative approximations may affect the optimality of the proposed method. Since Eq. (24) requires that the weighted average of y_j should be no smaller than $1-\epsilon$, the lower bound of y_j can be calculated by:*

$$y_j \geq \frac{1}{\pi_j}\left(1-\epsilon-\sum_{i\in\mathcal{J},i\neq j}\pi_i y_i\right) \geq \frac{1}{\pi_j}\left(1-\epsilon-\sum_{i\in\mathcal{J},i\neq j}\pi_i\right)=1-\frac{\epsilon}{\pi_j}, \quad (33)$$

where the first "$\geq$" is based on Eq. (24); the second "$\geq$" holds because the domain of y_j is $[0,1]$; the "$=$" holds due to $\sum_{j\in\mathcal{J}}\pi_j$=1. In practice, the risk parameter ϵ is usually small. Moreover, based on (21)-(22), a small π_j indicates that the i-th Gaussian component has very limited impact on the original chance constraint. This implies that the j-th Gaussian component may be redundant. Thus, we can decrease the hyper-parameter $|\mathcal{J}|$ and retrain the GMM until the weight of every Gaussian component is significant. With a small ϵ and big π_i, the value of y_j is large according to (33). In this case, constraint (26) can be naturally satisfied without involving any additional conservativeness. Meanwhile, according to Fig. 4, the approximation error introduced by the piecewise linearization (30)-(32) is also insignificant. As a result, with a small ϵ, the optimality loss introduced by the proposed convexification method is commonly marginal.

In the previous convexification, we introduce another non-convex constraint, i.e., Eq. (27). Nevertheless, the vector $\boldsymbol{\alpha}(\boldsymbol{x})$ only contains one single variable R_t^{ha} (See Appendix B), leading to only one non-convex constraint, i.e., $R_t^{\text{ha}}=e^{\rho_2}$. We can also use piecewise linearization to reformulate this non-convex constraint into:

$$R_t^{\text{ha}}=\max_{n\in\mathcal{N}^R}\lambda_n^R\rho_2+\gamma_n^R, \quad (34)$$

where $\mathcal{N}^R$ is the index set of lines for piecewise linearization. The n-th line, i.e., $\lambda_n^R\rho_2+\gamma_n^R$, is constructed by connecting the n-th and $(n+1)$-th points on function e^{ρ_2}. Unlike (31), the maximum operator appears on the RHS of "$=$" in (34). To realize this strict equivalence, we need to employ the Big-M method with auxiliary binary variable z_n to eliminate the maximum operator so that (34) can be reformulated into a solvable mixed integer form:

$$\begin{cases} R_t^{\text{ha}}\geq\lambda_n^R\rho_2+\gamma_n^R, & \forall n\in\mathcal{N}^R,\\ R_t^{\text{ha}}\leq\lambda_n^R\rho_2+\gamma_n^R+M(1-z_n), & \forall n\in\mathcal{N}^R,\\ \mathbf{1}^\intercal\boldsymbol{z}=1, \quad \boldsymbol{z}\in\{0,1\}^{|\mathcal{N}^R|}, \end{cases} \quad (35)$$

where M is a big number and $|\mathcal{N}^R|$ is the length of set $\mathcal{N}^R$.

Remark 3. *The auxiliary binary variable number, i.e., $|\mathcal{N}^R|$, can be very small because only one single variable R_t^{ha} needs to be linearized based on (35), which guarantees computational tractability.*

Finally, since (10) is linear, the expectation of the total cost, $\mathbb{E}(EC_t)$, can be calculated by:

$$\mathbb{E}(EC_t)=\sum_{i\in\mathcal{I}}\eta_t(p_{t,i}^{\text{ha}}-R_{t,i}^{\text{ha}}s_t^{\text{avg}})\Delta t - \sum_{i\in\mathcal{I}}(r_t^{\text{rc}}+r_t^m m_t^{\text{avg}})R_{t,i}^{\text{ha}}, \quad (36)$$

where $s_t^{\text{avg}}=\mathbb{E}(s_{t,l})=\frac{\sum_{l\in\mathcal{L}}s_{t,l}}{|\mathcal{L}|}$ and $m_t^{\text{avg}}=\mathbb{E}(m_t)$. Then, **P1** can be reformulated into a mixed-integer SOCP problem:

$$\begin{aligned} &\min \quad \text{Eq. (36)}, \\ &\text{s.t.: Eqs. (8)-(9), } \{(24), (32)\}_{(19)}, \text{ and (35)}, \end{aligned} \quad (\textbf{P2})$$

where $\{(24), (32)\}_{(19)}$ represents that each chance constraint in (19) is reformulated into (24) and (32) in **P2**.

C. Whole procedure

The whole procedure of the proposed method is summarized in Fig. 5. Specifically, the intractability of the original thermodynamic constraint (6) is from the heavy computational burden caused by the large-scale constraints and non-convexity brought by non-Gaussian regulation signals. To make them tractable, we first propose a temporal compression approach to reduce the number of constraints, in which the extreme indoor temperatures for a long duration are estimated and restricted in the comfortable range. Then, the original thermodynamic constraints can be replaced by only a few ones, i.e., Eq. (19). In the next step, we develop a mixture-model-based convexification method to handle the non-convexity from the non-Gaussian uncertainties. The GMM is first employed to fit the non-Gaussian uncertainties with multiple Gaussian uncertainties. Then, the chance constraint (19) is reformulated

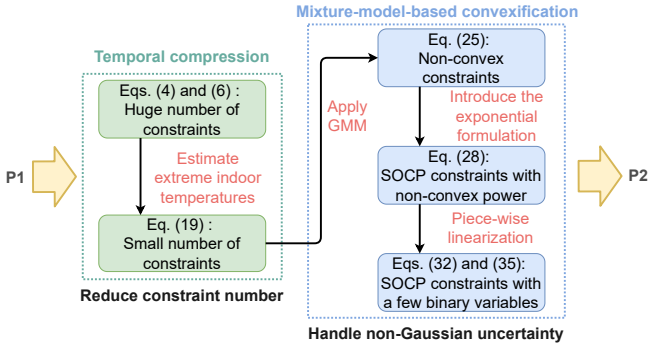


Fig. 5. The whole procedure of the proposed method.

TABLE I
PARAMETERS IN CASE STUDY

Parameters	Value	Parameters	Value
C_i	1.75 MWh/°C	$\theta_i^{\min}$	22°C
g_i	0.2MW/°C	$\theta_i^{\max}$	28°C
COP_i	5	$P_i^{\max}$	2MW
$\lvert\mathcal{N}\rvert^*$	10	$P_i^{\min}$	0
$\lvert\mathcal{N}^R\rvert^*$	50	$\lvert\mathcal{T}\rvert^*$	10
$\lvert\mathcal{J}\rvert^*$	3		

* Symbol $|\cdot|$ denotes the length of the set.

into a deterministic non-convex form (25). We further exponentially express these non-convex terms and equivalently convert them into L2-norm forms with non-convex powers. Finally, a piecewise linearized approximation is designed to handle these non-convex power, resulting in a mixed-integer SOCP reformulation, i.e., Eqs. (32) and (35). Since the proposed reformulation contains only a few binary variables, it is tractable and can be efficiently solved by off-the-shelf solvers.

IV. Case Study

A. System Configuration

We validate the proposed method based on one large-capacity HVAC system. The daily heat loads, outdoor temperature, and unit prices for electricity purchasing and regulation revenue are demonstrated in Fig. 6. The thermal parameters of building and other parameters are listed in Table I. Here the outdoor temperature is the air temperature of one summer day in Macau [33]. The unit prices for electricity purchasing/selling and regulation revenue are from the PJM market [34]. The thermal parameters of building and heat loads are generated based on a real building at the University of Macau. We collect the whole-year regD signals from PJM in 2020 [5] as historical data. Based on these signal data, we construct the samples of the uncertain parameters used in **P2**, including m_t in (11), $\overline{u}_\tau$ and $\underline{u}_\tau$ in (18). The whole dataset has been uploaded in [35].

All simulations are implemented based on an Intel(R) Core(TM) 8700 3.20GHz CPU with 16GB memory. The corresponding optimization problem is built by CVXPY and solved by MOSEK.

B. Benchmarks

To demonstrate the superiority of the proposed approach, three benchmarks are introduced:

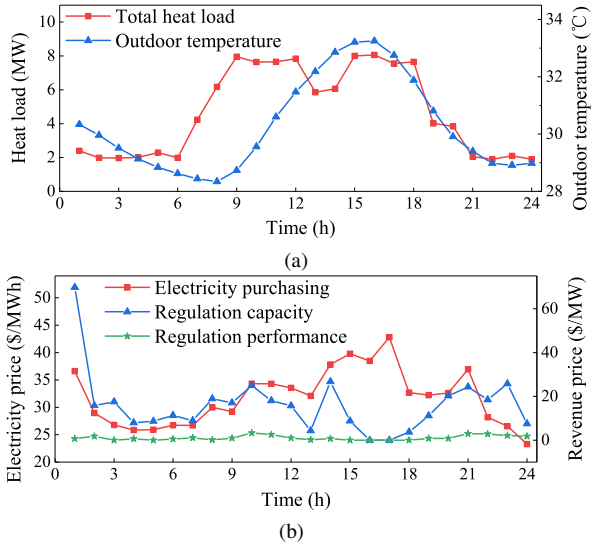


Fig. 6. (a) Heat load h_t and outdoor temperature θ^{out} and (b) electricity purchase price e_t, regulation capacity revenue price r_t^{rc}, and regulation performance revenue price r_t^{m}.

1) **B1**: The Gaussian-assumption-based CCP used in [15]–[17];
2) **B2**: The moment-based DRCC method used in [20];
3) **B3**: The Wasserstein-distance-based DRCC method used in [21]. In this approach, the proposed temporal compression method is also applied; otherwise, the computational burden will be too heavy and the out-of-memory issue will occur. A total of 1000 samples are used for constructing the ambiguity set:[4]

C. Model comparison

1) Optimality, feasibility, and computational efficiency: Fig. 7 shows the results of the whole-day total costs (i.e. $\sum_{t=1}^{24} EC_t$), solving times, and maximum probability violation under different risk parameters. The Gaussian-assumption-based method **B1** can achieve good optimality performance. However, since the original uncertainties do not follow Gaussian distributions, it may derive infeasible solutions, i.e., the maximum violation probability is larger than the given risk parameter. The two DRCC methods, i.e. **B2** and **B3**, can always ensure the feasibility of solutions. However, their total costs are much higher than that of the proposed method because they need to ensure the feasibility of all the possible distributions in their ambiguity sets, including some distributions that are quite different from the actual one. The total cost of the proposed method is the lowest among all methods when the risk parameter is small. With the increase of the risk parameter, the optimality loss introduced by the proposed method may be not negligible according to **Remark** 2. Nevertheless, its optimality performance is still much better than those of the two DRCC methods, **B2** and **B3**. Moreover, unlike **B1**, its maximum violation probability is also always

[4]The optimality of **B3** can be improved by increasing the sample number for constructing the ambiguity set [18]. However, the computational burden is also proportional to this sample number. According to our test, if this sample number is larger than 1000 (e.g. 1500), the out-of-memory issue occurs.

This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2022.3182000

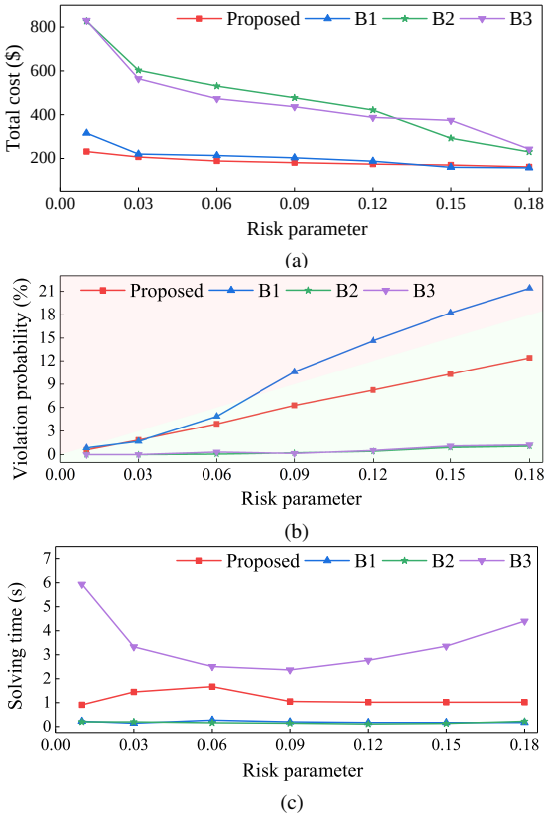


Fig. 7. Results of (a) whole-day total cost, i.e., $\sum_{t=1}^{24} EC_t$, (b) maximum violation probability, and (c) solving time obtained by different models. In (b), the green and red regions represent the safe (i.e. the maximum violation probability is smaller than the risk parameter) and unsafe regions (i.e. the maximum violation probability is larger than the risk parameter).

lower than the given risk parameter. These results confirm the superior optimality and feasibility performances of the proposed method.

In **B1** and **B2**, each chance constraint is reformulated into only one single SOCP constraint but introduces no additional constraint, while some additional constraints, such as (35), are necessary for the proposed method. Thus, the solving times of **B1** and **B2** are lower than that of the proposed model. Nevertheless, the solving time of the proposed method is only around 1s, which is also acceptable in practice. In **B3**, numerous additional constraints have to be introduced [18], so its computational performance is the worst among all methods.

2) Hour-ahead regulation capacity offers: Fig. 8 summarizes the hour-ahead regulation capacity offers obtained by different methods under $\epsilon = 0.15$. Note the results of **B1** are not listed here because **B1** can not ensure the feasibility of solutions. In all time, the hour-ahead regulation capacity offer of the proposed method is much larger compared to the rest models. As aforementioned, **B2** and **B3** have to satisfy constraints for all distributions in their ambiguity sets, so they are more conservative. As a result, large margins need to be reserved for the uncertainties in both **B2** and **B3**, which shrinks the potential regulation capacity. This result validates the better optimality of the proposed method.

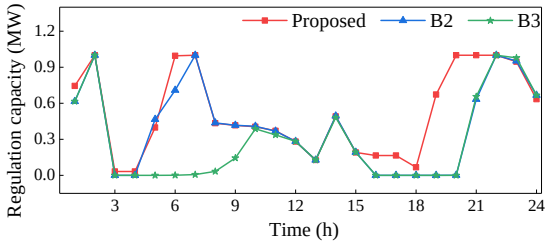


Fig. 8. Results of hour-ahead regulation capacity offers under $\epsilon = 0.15$.

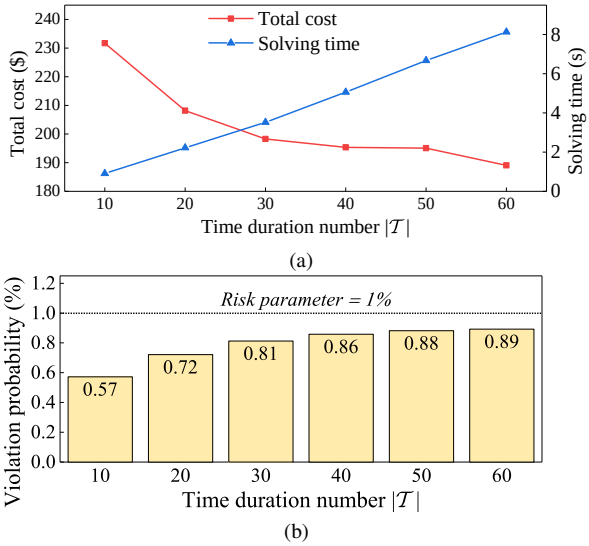


Fig. 9. Results of (a) daily total cost and solving time, and (b) maximum violation probability under different time duration numbers, i.e., $|\mathcal{T}|$, for splitting the operating hour with $\epsilon = 0.01$.

D. Sensitivity analysis

1) Time duration number $|\mathcal{T}|$ for splitting the operating hour: We implement a case study with different $|\mathcal{T}|$ to investigate its effects, and the corresponding results are shown in Fig. 9. The risk parameter is 0.01, while the line number for the piecewise linearization $|\mathcal{N}|^R$ is set as 50. With the growth of $|\mathcal{T}|$, the total cost decreases. According to **Proposition** 1, if one time duration $\Delta\tau$ is split into smaller ones, the additional conservativeness introduced by the inner approximation used in (16) can be further reduced. That is to say, increasing $|\mathcal{T}|$ can improve the optimality of the proposed method. Conversely, the solving time becomes larger because more constraints are introduced according to (32). Because increasing $|\mathcal{T}|$ decreases the conservativeness, the maximum violation probability also grows with the increase of $|\mathcal{T}|$. Nevertheless, its value always keeps in the safe region, which demonstrates the great feasibility of the proposed method.

2) Line number $|\mathcal{N}|^R$ in (34): In (34), we employ piecewise linearization to approximate the regulation capacity offer R_t^{ha}. Fig. 10 demonstrates the maximum and average approximation errors, whole-day total cost, solving time, and maximum violation probability under different line numbers $|\mathcal{N}|^R$. With the increase of $|\mathcal{N}|^R$, the piecewise linearization used in (34) becomes more accurate, so both the maximum and average approximation errors decrease, as shown in Fig. 10(a). Once $|\mathcal{N}|^R$ reaches 100, the impacts of the approximation error on the optimal solution becomes insignificant. Thus, even if we further increase $|\mathcal{N}|^R$ from 100 to 500, the obtained total cost

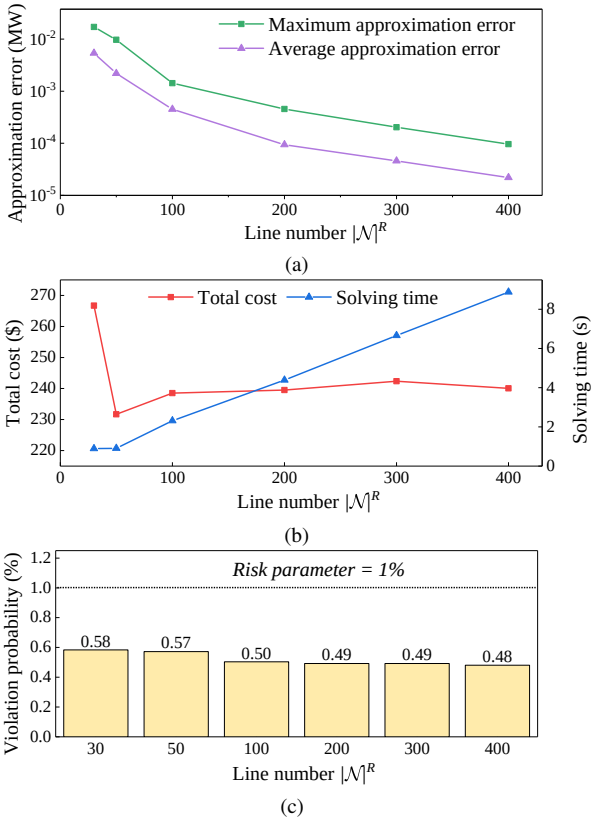


Fig. 10. Results of (a) approximation errors caused by the piecewise linearization in (34), (b) whole-day total cost and solving time, and (c) maximum violation probability under different line numbers, i.e., $|\mathcal{N}|^R$ with $\epsilon = 0.01$. Note the error is defined as $|e^{\rho_2} - R_t^{\text{ha}}|$.

and maximum violation probability keep almost unchanged, as illustrated in Figs. 10(b) and (c). Increasing $|\mathcal{N}|^R$ introduces more binary variables according to (35), so the solving time grows rapidly. Nevertheless, in all these cases, the total cost is much less compared to **B1**, **B2**, and **B3**. Moreover, the maximum violation probability is also always lower than the risk parameter. These results further confirm the benefits of the proposed method.

E. Effects of temporal compression

We implement a new case study to demonstrate the effectiveness of the proposed temporal compression method on computational efficiency. In this case study, four scenarios are compared. The first three employ the proposed temporal compression method, while the last one does not. As aforementioned, the time duration Δt affects the optimality and computational efficiency. Thus, here we also change the length of Δt in the first three scenarios to evaluate its effects. The corresponding results of solving times, total costs, and maximum violation probabilities are illustrated in Fig. 11. Obviously, employing temporal compression can significantly reduce the computational burden. For example, when the risk parameter equals 0.12, the solving times of the first three scenarios are around 1s, 4s, and 7s. However, it reaches 317s in the scenario without temporal compression. As introduced in Section III-A, the temporal compression method introduces additional conservativeness. Therefore, the total costs of the

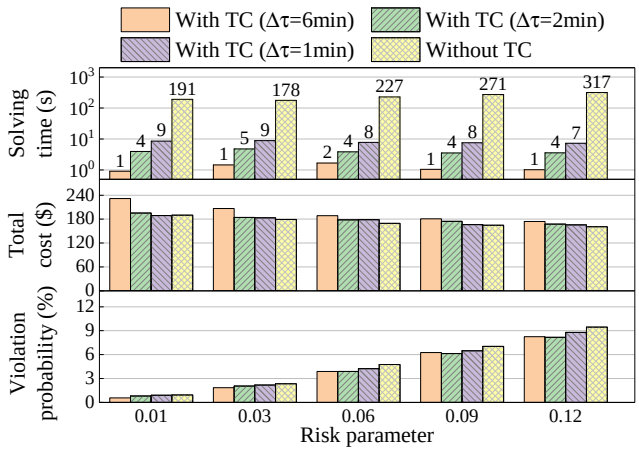


Fig. 11. Results of solving times, total costs, and maximum violation probabilities in the cases with and without the proposed temporal compression. Here "TC" represents "Temporal Compression." Symbol $\Delta\tau$ is the shorter time duration introduced in Fig. 2.

scenarios with temporal compression are larger than that of the one without temporal compression. Nevertheless, reducing the length of the time duration $\Delta\tau$ can mitigate this conservativeness. For instance, the total costs of the third scenario with $\Delta\tau = 1$min are only slightly higher than those without temporal compression. The average diffidence of the total costs between the third and last scenarios is only around 2%. For every scenario, its maximum violation probability is always lower than the corresponding risk parameter. These results confirm that the proposed temporal compression method can significantly improve the computational efficiency with desirable optimality and feasibility.

V. Conclusions

This paper proposes a tractable chance-constrained model to optimize the hour-ahead regulation capacity offers for HVAC systems. It first proposes a temporal compression method to compress the numerous thermodynamic constraints introduced by frequently regulated HVAC power into only a few constraints. Then, a novel mixture-model-based convexification approach is developed to overcome the intractability caused by the non-Gaussian uncertainties from regulation signals. By applying this approach, the chance constraints with these non-Gaussian uncertainties on the LHS can be approximated by tractable SOCP forms. The optimality loss introduced by this approximation is also marginal under mild conditions. Numerical experiments confirm that the proposed method can achieve better feasibility performance compared to the widely used Gaussian-assumption-based methods, while its solution is also less conservative than the DRCC methods.

Appendix A

Proof of Proposition 1: We first define two new vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, as follows:

$$\boldsymbol{v}_1 = \left[\max_{l\in\mathcal{L}_\tau} f_l, \forall\tau\in\mathcal{T}\right]^\intercal, \boldsymbol{v}_2 = \left[R\min_{l\in\mathcal{L}_\tau}[\boldsymbol{A}\boldsymbol{s}]_l, \forall\tau\in\mathcal{T}\right]^\intercal. \tag{37}$$

Then, based on the Minkowski's inequality, we must have:

$$\underbrace{\|\boldsymbol{v}_1+\boldsymbol{v}_2\|_\infty}_{\text{RHS term of (17)}} \le \underbrace{\|\boldsymbol{v}_1\|_\infty+\|\boldsymbol{v}_2\|_\infty}_{\text{RHS term of (15)}}. \tag{38}$$

By substituting (37) into (38), we prove the maximum indoor temperature estimated by (17) is no more than that of (15). Based on the same way, we can also prove that the minimum indoor temperature estimated by (17) is no less than that of (15). This completes the proof.

APPENDIX B

For the first constraint in (19), the detail expressions of $\boldsymbol{\alpha}(\boldsymbol{x})$ and $\beta(\boldsymbol{x})$ are as follows:

$$\text{1st}\begin{cases}\boldsymbol{\alpha}(\boldsymbol{x})=[a_\tau^{\text{in}},R^{\text{ha}}]^\intercal, \quad \boldsymbol{\omega}=[\theta_0^{\text{in}},\overline{u}_\tau]^\intercal,\\ \beta(\boldsymbol{x})=\theta^{\max}-a_\tau^{\text{out}}\theta^{\text{out}}-a_\tau^{\text{h}}\theta^{\text{h}}-a_\tau^{\text{q}}p^{\text{ha}},\end{cases} \tag{39}$$

where 1st represents the first constraint; parameters $a_\tau^{\text{in}}=(a^{\text{in}})^{\tau\frac{|\mathcal{L}|}{|\mathcal{T}|}}$, $a_\tau^{\text{out}}=a^{\text{out}}\frac{1-a_\tau^{\text{in}}}{1-a^{\text{in}}}$, $a_\tau^{\text{h}}=a^{\text{h}}\frac{1-a_\tau^{\text{in}}}{1-a^{\text{in}}}$, and $a_\tau^{\text{q}}=a^{\text{q}}\frac{1-a_\tau^{\text{in}}}{1-a^{\text{in}}}$. Similarly, for the rest chance constraints, we have

$$\text{2nd}\begin{cases}\boldsymbol{\alpha}(\boldsymbol{x})=[a_{\tau+1}^{\text{in}},R^{\text{ha}}]^\intercal, \quad \boldsymbol{\omega}=[\theta_0^{\text{in}},\overline{u}_\tau]^\intercal,\\ \beta(\boldsymbol{x})=\theta^{\max}-a_{\tau+1}^{\text{out}}\theta^{\text{out}}-a_{\tau+1}^{\text{h}}\theta^{\text{h}}-a_{\tau+1}^{\text{q}}p^{\text{ha}},\end{cases} \tag{40}$$

$$\text{3rd}\begin{cases}\boldsymbol{\alpha}(\boldsymbol{x})=[a_\tau^{\text{in}},R^{\text{ha}}]^\intercal, \quad \boldsymbol{\omega}=[-\theta_0^{\text{in}},-\underline{u}_\tau]^\intercal,\\ \beta(\boldsymbol{x})=a_\tau^{\text{out}}\theta^{\text{out}}+a_\tau^{\text{h}}\theta^{\text{h}}+a_\tau^{\text{q}}p^{\text{ha}}-\theta^{\min},\end{cases} \tag{41}$$

$$\text{4th}\begin{cases}\boldsymbol{\alpha}(\boldsymbol{x})=[a_{\tau+1}^{\text{in}},R^{\text{ha}}]^\intercal, \quad \boldsymbol{\omega}=[-\theta_0^{\text{in}},-\underline{u}_\tau]^\intercal,\\ \beta(\boldsymbol{x})=a_{\tau+1}^{\text{out}}\theta^{\text{out}}+a_{\tau+1}^{\text{h}}\theta^{\text{h}}+a_{\tau+1}^{\text{q}}p^{\text{ha}}-\theta^{\min}.\end{cases} \tag{42}$$

Note that only one single variable, i.e., R^{ha}, is contained in the four $\boldsymbol{\alpha}(\boldsymbol{x})$ in (39)-(42).

APPENDIX C

Proof of Proposition 2: The convexity of function $\ln\Phi^{-1}(y_j)$ can be analyzed by its second-order derivative:

$$\frac{d^2\ln\Phi^{-1}(y_j)}{(dy_j)^2}=\frac{-\phi(v_j)-v_j\phi'(v_j)}{v_j^2(\phi(v_j))^3}, \quad \forall j\in\mathcal{J}, \tag{43}$$

where $v_j=\Phi^{-1}(y_j)$; $\phi(\cdot)$ is the PDF of the standard normal distribution and $\phi'(\cdot)$ is its first-order derivative. Note the PDF $\phi(\cdot)$ is always nonnegative. Considering that a convex function has a nonnegative second-order derivative, we can get the convex condition for function $\ln\Phi^{-1}(y_j)$, as follows:

$$\begin{aligned}&\frac{-\phi(v_j)-v_j\phi'(v_j)}{v_j^2(\phi(v_j))^3}\ge 0\Leftrightarrow\phi(v_j)+v_j\phi'(v_j)\le 0\\ &\Leftrightarrow\frac{e^{-v_j^2/2}}{\sqrt{2\pi}}-v_j^2\frac{e^{-v_j^2/2}}{\sqrt{2\pi}}\le 0\Leftrightarrow v_j^2\ge 1.\end{aligned} \tag{44}$$

According to the definition of chance constraints, we have $y_j\ge 0.5$. Thus, variable v_j is non-negative. As a result, the above inequality can be further converted into:

$$v_j^2\ge 1\Leftrightarrow v_j\ge 1\Leftrightarrow y_j\ge\Phi(1). \tag{45}$$

Similarly, its concave condition can be obtained by:

$$\frac{-\phi(v_j)-v_j\phi'(v_j)}{v_j^2(\phi(v_j))^3}\le 0\Leftrightarrow v_j^2\le 1\Leftrightarrow y_j\le\Phi(1). \tag{46}$$

We uniformly select $N+1$ points on function $\ln\Phi^{-1}(y_j)$ (recorded by $y_j^{(n)},\forall n\in\mathcal{N}$) and let $y_j^{(0)}=\Phi(1)<y_j^{(1)}<\cdots<y_j^{(N)}$. With points $y_j^{(1)},\cdots,y_j^{(N)}$, N-1 line segments can be constructed by connecting these points in sequence (denoted by $\lambda_n y_j+\gamma_n,\forall n\in\mathcal{N}/\{0\}$). Since function $\ln\Phi^{-1}(y_j)$ is convex when $y_j\ge\Phi(1)$, according to the definition of convex functions, we must have:

$$\ln\Phi^{-1}(y_j)\le\max_{n\in\mathcal{N}/\{0\}}\{\lambda_n y_j+\gamma_n\}, \quad \text{for } y_j\ge\Phi(1). \tag{47}$$

For the region $y_j\le\Phi(1)$, function $\ln\Phi^{-1}(y_j)$ is concave. According to the concave function's first-order condition, function $\ln\Phi^{-1}(y_j)$ should always be equal to or below its tangent. Thus, by letting $\lambda_0 y_j+\gamma_0$ as the tangent of function $\ln\Phi^{-1}(y_j)$ at $y_j=\Phi(1)$, we have:

$$\ln\Phi^{-1}(y_j)\le\lambda_0 y_j+\gamma_0, \quad \text{for } y_j\le\Phi(1). \tag{48}$$

By combining (47) and (48), we prove **Proposition** 2.

REFERENCES

[1] X. Liang, "Emerging power quality challenges due to integration of renewable energy sources," *IEEE Trans. Ind. Appl.*, vol. 53, no. 2, pp. 855–866, 2017.

[2] S. Impram, S. V. Nese, and B. Oral, "Challenges of renewable energy penetration on power system flexibility: A survey," *Energy Strategy Rev.*, vol. 31, p. 100539, 2020.

[3] P. Kohlhepp, H. Harb, H. Wolisz, S. Waczowicz, D. Müller, and V. Hagenmeyer, "Large-scale grid integration of residential thermal energy storages as demand-side flexibility resource: A review of international field studies," *Renew. Sust. Energ. Rev.*, vol. 101, pp. 527–547, 2019.

[4] J. E. Contreras-Ocaña, M. A. Ortega-Vazquez, D. Kirschen, and B. Zhang, "Tractable and robust modeling of building flexibility using coarse data," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5456–5468, 2018.

[5] P. Dispatch, "PJM manual 12: Balance operations," 2019.

[6] L. Fabietti, T. T. Gorecki, F. A. Qureshi, A. Bitlislioğlu, I. Lymperopoulos, and C. N. Jones, "Experimental implementation of frequency regulation services using commercial buildings," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1657–1666, 2018.

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

[8] L. Zhao, W. Zhang, H. Hao, and K. Kalsi, "A geometric approach to aggregate flexibility modeling of thermostatically controlled loads," *IEEE Trans. Power Syst.*, vol. 32, no. 6, pp. 4721–4731, 2017.

[9] X. Chen, E. Dall'Anese, C. Zhao, and N. Li, "Aggregate power flexibility in unbalanced distribution systems," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 258–269, 2020.

[10] G. Chen, H. Zhang, H. Hui, N. Dai, and Y. Song, "Scheduling thermostatically controlled loads to provide regulation capacity based on a learning-based optimal power flow model," *IEEE Trans. Sustain. Energy*, vol. 12, no. 4, pp. 2459–2470, 2021.

[11] E. Yao, V. W. S. Wong, and R. Schober, "Robust frequency regulation capacity scheduling algorithm for electric vehicles," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 984–997, 2017.

[12] Y. Wang, C. Wan, Z. Zhou, K. Zhang, and A. Botterud, "Improving deployment availability of energy storage with data-driven agc signal models," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4207–4217, 2018.

[13] D. Bertsimas, V. Gupta, and N. Kallus, "Data-driven robust optimization," *Math. Programming*, vol. 167, no. 2, pp. 235–292, 2018.

[14] X. Geng and L. Xie, "Data-driven decision making in power systems with probabilistic guarantees: Theory and applications of chance-constrained optimization," *Annu Rev Control*, vol. 47, pp. 341–363, 2019.

[15] G. Chen, B. Yan, H. Zhang, D. Zhang, and Y. Song, "Time-efficient strategic power dispatch for district cooling systems considering the spatial-temporal evolution of cooling load uncertainties," *CSEE J. Power Energy Syst.*, pp. 1–11, 2021.

[16] L. Roald and G. Andersson, "Chance-constrained ac optimal power flow: Reformulations and efficient algorithms," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 2906–2918, 2018.

[17] A. Peña-Ordieres, D. K. Molzahn, L. A. Roald, and A. Wächter, "Dc optimal power flow with joint chance constraints," *IEEE Trans. Power Syst.*, vol. 36, no. 1, pp. 147–158, 2021.

[18] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations," *Math. Programming*, vol. 171, no. 1, pp. 115–166, 2018.

[19] G. Chen, H. Zhang, H. Hui, and Y. Song, "Fast wasserstein-distance-based distributionally robust chance-constrained power dispatch for multi-zone hvac systems," *IEEE Trans. Smart Grid*, vol. 12, no. 5, pp. 4016–4028, 2021.

[20] H. Zhang, Z. Hu, E. Munsing, S. J. Moura, and Y. Song, "Data-driven chance-constrained regulation capacity offering for distributed energy resources," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2713–2725, 2019.

[21] L. Liu, Z. Hu, X. Duan, and N. Pathak, "Data-driven distributionally robust optimization for real-time economic dispatch considering secondary frequency regulation cost," *IEEE Trans. Power Syst.*, vol. 36, no. 5, pp. 4172–4184, 2021.

[22] A. Venzke, L. Halilbasic, U. Markovic, G. Hug, and S. Chatzivasileiadis, "Convex relaxations of chance constrained ac optimal power flow," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 2829–2841, 2018.

[23] A. Venzke and S. Chatzivasileiadis, "Convex relaxations of probabilistic ac optimal power flow for interconnected ac and hvdc grids," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 2706–2718, 2019.

[24] X. Cao, J. Wang, and B. Zeng, "Networked microgrids planning through chance constrained stochastic conic programming," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6619–6628, 2019.

[25] E. Dall'Anese, K. Baker, and T. Summers, "Chance-constrained ac optimal power flow for distribution systems with renewables," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3427–3438, 2017.

[26] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*, vol. 1. MIT press Cambridge, 2016.

[27] W. Sun, M. Zamani, M. R. Hesamzadeh, and H.-T. Zhang, "Data-driven probabilistic optimal power flow with nonparametric bayesian modeling and inference," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1077–1090, 2020.

[28] Y. Yang, W. Wu, B. Wang, and M. Li, "Analytical reformulation for stochastic unit commitment considering wind power uncertainty with gaussian mixture model," *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 2769–2782, 2020.

[29] J. Wang, C. Wang, Y. Liang, T. Bi, M. Shafie-khah, and J. P. S. Catalao, "Data-driven chance-constrained optimal gas-power flow calculation: A bayesian nonparametric approach," *IEEE Trans. Power Syst.*, pp. 1–1, 2021.

[30] H. Hui, P. Siano, Y. Ding, P. Yu, Y. Song, H. Zhang, and N.-Y. Dai, "A transactive energy framework for inverter-based hvac loads in a real-time local electricity market considering distributed energy resources," *IEEE Trans. Industr. Inform.*, pp. 1–1, 2022.

[31] Z. Hu, W. Sun, and S. Zhu, "Chance constrained programs with mixture distributions," 2018.

[32] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

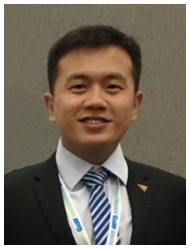
[33] "Actual Weather 24-Hour Time Series." [Online]. https://www.smg.gov.mo/zh/subpage/73/time_series/DP.

[34] "Real-Time Hourly LMPs." [Online]. https://dataminer2.pjm.com/feed/rt_hrl_lmps/definition.

[35] "Samples of uncertainties." [Online]. https://github.com/lelouchsola/NonGaussianCCPConvexification.

Ge Chen (S'20) received the B.S. degree from Huazhong University of Science and Technology, Wuhan, China, in 2015 and the M.S. degree from Xi'an Jiaotong University in 2018, both in thermodynamic engineering. He is currently working toward the Ph.D. degree at University of Macau, Macau, China. His research interests include Internet of Things for smart energy, optimal operation and data-driven optimization under uncertainty.

Hongcai Zhang (S'14–M'18) received the B.S. and Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2013 and 2018, respectively. He is currently an Assistant Professor with the State Key Laboratory of Internet of Things for Smart City and Department of Electrical and Computer Engineering, University of Macau, Macao, China. In 2018-2019, he was a postdoctoral scholar with the Energy, Controls, and Applications Lab at University of California, Berkeley, where he also worked as a visiting student researcher in 2016. His current research interests include Internet of Things for smart energy, optimal operation and optimization of power and transportation systems, and grid integration of distributed energy resources.



Hongxun Hui (S'17–M'20) received both the Ph.D. and B. Eng degrees in electrical engineering from Zhejiang University in 2020 and 2015, respectively. He is currently a Post-doctoral Fellow with the State Key Laboratory of Internet of Things for Smart City, University of Macau. His research interests include modelling and optimal control of demand side resources in smart grid, the electricity market considering demand response, and the uncertainty analysis brought by flexible loads and renewable energies.

Yonghua Song (F'08) received the B.E. and Ph.D. degrees from the Chengdu University of Science and Technology, Chengdu, China, and the China Electric Power Research Institute, Beijing, China, in 1984 and 1989, respectively, all in electrical engineering. He was awarded DSc by Brunel University in 2002, Honorary DEng by University of Bath in 2014 and Honorary DSc by University of Edinburgh in 2019. He is currently a Chair professor of University of Macau and Director of the State Key Laboratory in the Internet of Things for Smart City. He is also the Vice President of Chinese Electrotechnical Society, an international member of Academia Europaea, a fellow of the Royal Academy of Engineering (UK) and an IEEE fellow. He has long been engaged in renewable electrical power systems and smart energy research. He won a second prize of the State Scientific and Technological Progress Award and the prize for Scientific and Technological Progress from the Ho Leung Ho Lee Foundation.