A Novel Surrogate Polytope Method for Day-ahead Virtual Power Plant Scheduling with Joint Probabilistic Constraints

Yihong Zhou *, Chaimaa Essayeh † and Thomas Morstyn ‡
* School of Engineering
University of Edinburgh, Edinburgh, EH9 3FB, U.K.
yihong.zhou@ed.ac.uk
† Departement of Engineering
Nottingham Trent University, Nottingham, NG1 4FQ, U.K.
chaimaa.essayeh@ntu.ac.uk
‡ Department of Engineering Science
University of Oxford, Oxford, OX1 3PJ, U.K.
thomas.morstyn@eng.ox.ac.uk

***Abstract*—Virtual Power Plant (VPP) is an emerging concept that can effectively manage a large number of distributed energy resources (DERs). However, the inherent uncertainty of load and renewable generation poses a challenge to reliable VPP power scheduling. This manuscript proposes a novel method for day-ahead VPP scheduling with a joint probabilistic guarantee on its power availability without violating the DER constraints and network constraints. A surrogate polytope is first used to find the inner approximation of the VPP power, implicitly including the low-level DER power, DER constraints, and network constraints. Then, a multivariate Gaussian distribution is used to fit the random parameters of the surrogate polytope, after which the iterative supporting hyperplane algorithm is used to solve the VPP scheduling problem. Extensive case studies based on real-world renewable generation scenarios demonstrate the superior performance of the proposed method in out-of-sample cost and reliability, with a manageable computing complexity.**

***Index Terms*—Distributed Energy Resources (DERs), Joint Probabilistic Constraints, Network Constraints, Virtual Power Plant (VPP).**

I. Introduction

The proliferation of distributed energy resources (DERs), including photovoltaic units (PV), wind turbines (WT), and energy storage (ES), is transforming the traditionally passive distribution network into an active one [1]. The Virtual Power Plant (VPP) is a concept that aggregates and utilizes a collection of DERs for tasks such as energy cost minimization, frequency regulation, and reserve provision [1]. Pilot projects for VPP deployment are implementing in several countries [2]. Based on the objective of aggregation, existing literature divides VPPs into two categories: commercial VPPs (CVPPs) and technical VPPs (TVPPs). CVPPs primarily focus on financial activities, and the managed DERs are not restricted to a single distribution grid [3]. TVPPs additionally consider the distribution network constraints and the managed DERs are typically within a single distribution grid [3]. TVPPs can also be referred to Active Distribution Networks (ADNs) [4]. This paper focuses on the scheduling of a TVPP, and all the "VPP" terms in the later content refer to "TVPP" in existing literature for simplicity.

In a standard market engagement process [5], a VPP determines its schedule of the aggregated power for the next day; in real-time operation, this VPP needs to fulfil the scheduled aggregated power by adjusting its managed DERs. Nevertheless, the inherent uncertainty in DERs like PV and WT introduces challenges to reliably scheduling the aggregated power. Ensuring reliable power availability when scheduling a VPP becomes crucial, especially for safety-critical applications and when DERs constitute a larger portion of the overall system in the future. To tackle the uncertainty, robust optimization (RO) was widely applied due to its solution robustness and mathematical tractability [6]. Ref. [7] applied adaptive robust optimization (ARO), a variant of RO considering the multi-stage decision making process, to derive the real-reactive feasible region of the aggregated power of a VPP that facilitates the hour-ahead or day-ahead scheduling. Ref. [8] also applied ARO for this objective and additionally modelled the time-coupling DER constraints. However, as RO fights against the worst-case scenario, a drawback is its resulting over-conservativeness [9].

Another alternative is the chance constraint (CC) approach, which explicitly models the probability of constraint satisfaction and reduces conservativeness [10]. The CC-based day-ahead (DA) power scheduling was studied in [11], using a Gaussian mixture model (GMM) to model the wind forecast error. Ref. [12] also applied GMM and CC to derive the feasible region of the VPP aggregated power, which can be used for day-ahead scheduling purposes as [7]. However, the fitted

Submitted to the 23rd Power Systems Computation Conference (PSCC 2024).

distribution for random variables can be inaccurate, thereby limiting the out-of-sample performance [13]. Distributionally robust chance constraints (DRCC) consider the potential error in the estimated distribution. This property leads to higher robustness than CC and less conservativeness than RO, making DRCC attractive in power system scheduling [14].

The aforementioned CC and DRCC studies [12], [14], [15] used individual CC (ICC), controlling the violation probability for individual constraints. In power system applications, a joint probability guarantee is desired to ensure that all system constraints can be met simultaneously [16], [17], calling for a joint chance constraint (JCC) modelling. By assuming that the random variables are normally distributed, [17] used an iterative supporting hyperplane method to exactly solve the JCC. The exact reformulation of distributionally robust joint chance constraint (DRJCC) was studied in [5]. However, the tractability of the exact solution methods in [17] and [5] depends on the size of the JCC, i.e., the number of constraints included in the JCC. When numerous uncertain DERs and multiple network constraints require a joint probability satisfaction guarantee, tractability becomes a concern. Bonferroni approximation is a commonly used tractable approach for tackling the JCC [16]. It achieves a safe approximation to the JCC by using an ICC model with a low violation probability. Ref. [18] applied Bonferroni approximation to the intractable JCC in the VPP scheduling. However, this approximation introduces over-conservativeness once again.

To address the issue of the scalability challenge in exact JCC solutions and the over-conservativeness in the Bonferroni JCC approximation, this paper introduces a novel surrogate polytope method for solving the VPP scheduling problems with JCC. The proposed method leverages the insight that the result of interest for a VPP in a DA scheduling phase is its aggregated power. The availability of the VPP power depends on its low-level power elements (DERs) while adhering to internal constraints. A surrogate polytope is employed as an inner approximation of the VPP's power availability. This allows for avoiding the explicit modelling of the extensive number of DER power variables, DER constraints, and network constraints. The JCC is applied to the surrogate polytope, and the problem is addressed using the supporting hyperplane method [19]. Since our approach is built upon the surrogate polytope for VPP power alone, the JCC becomes tractable, and the overall computational complexity remains manageable. Note that [20] used conditional value-at-risk approximations to convert the energy dispatch problem with DRJCC into a tractable convex problem. However, their method needs an iterative approach for hyperparameter tuning, which is not required here. In summary, this paper makes the following contributions:

1) We propose a novel surrogate polytope method for VPP DA scheduling with a joint probabilistic guarantee on the VPP power availability. The method is less conservative than the Bonferroni approximation and is still reliable.
2) The proposed method has a manageable computing complexity. For varying DER numbers up to 120, the worst-case scheduling time without parallelization is around 12 minutes.
3) The proposed method is verified by extensive case studies on real-world non-Gaussian DER generation scenarios and comparisons with benchmarks.

The manuscript is organised as follows: Section II outlines the basic formulation of VPP scheduling. The proposed method is detailed in Section III. Section IV presents the case studies, and Section V concludes the manuscript.

II. VPP Scheduling Problem

We consider a distribution network that takes the form of a VPP, wherein a single distribution substation functions as the point of interface with the upper-level system. For mathematical notation, a lowercase letter a denotes a scalar, a lowercase bold letter $\boldsymbol{a}$ signifies a vector, and an uppercase bold letter $\boldsymbol{A}$ denotes a matrix.

A. Model of Distributed Energy Resources

For a scheduling horizon $[T] := \{1, \cdots, T\}$, we consider the following DERs:

1) Renewable Generator: The renewable generator (RG) including PV and WT can be modelled as:

$$\forall t \in [T], \forall i_{\text{RG}} \in [N^{\text{RG}}],$$
$$0 \le p^{\text{RG},\,\psi}_{t,i_{\text{RG}}} \le \overline{p}^{\text{RG},\psi}_{t,i_{\text{RG}}} \tag{1}$$
$$(p^{\text{RG},\psi}_{t,i_{\text{RG}}}, q^{\text{RG},\psi}_{t,i_{\text{RG}}}) \in \mathcal{PQ}^{\text{RG},\psi}_{t,i_{\text{RG}}} \tag{2}$$

where $\overline{p}^{\text{RG},\psi}_{t,i_{\text{RG}}}$ denotes the maximum active power capability of the RG connected to the $i_{\text{RG}}{}^{\text{th}}$ bus at time-step t, which depends on the weather and the generator capacity limit. Here, $\psi \in \Psi$ represents the phase index, where Ψ comprises a, b, c for Wye connections or ab, bc, ca for Delta connections. The set $[N^{\text{RG}}]$ consists of buses with RG connections. Eq. (1) means that the RG generation $p^{\text{RG},\psi}_{t,i_{\text{RG}}}$ can be curtailed to zero. As modelled in (2), an RG can also provide reactive power supports within their active-reactive power capability $\mathcal{PQ}^{\text{RG},\psi}_{t,i_{\text{RG}}}$. Fig. 1(a) and (b) illustrate the PQ-charts for PV and WT based on [21], which consist of the capacity limit (green curves), active power limits (red curves), and reactive power limits (blue curves).

2) Energy Storage: The power capability of the energy storage (ES) can be modelled as

$$\forall t \in [T], \forall i_{\text{ES}} \in [N^{\text{ES}}]$$
$$p^{\text{ES},\psi}_{t,i_{\text{ES}}} = \widehat{p}^{\text{ES},\psi}_{t,i_{\text{ES}}} + \check{p}^{\text{ES},\psi}_{t,i_{\text{ES}}} \tag{3}$$
$$\underline{p}^{\text{ES},\psi}_{i_{\text{ES}}} \le \check{p}^{\text{ES},\psi}_{t,i_{\text{ES}}} \le 0 \tag{4}$$
$$0 \le \widehat{p}^{\text{ES},\psi}_{t,i_{\text{ES}}} \le \overline{p}^{\text{ES},\psi}_{i_{\text{ES}}} \tag{5}$$
$$e_{t,i_{\text{ES}}} = e_{t-1,i_{\text{ES}}} + \sum_{\psi} (\widehat{p}^{\text{ES},\psi}_{t,i_{\text{ES}}} \widehat{\eta} + \check{p}^{\text{ES},\psi}_{t,i_{\text{ES}}} / \check{\eta}) \cdot \Delta t \tag{6}$$
$$\underline{e}_{i_{\text{ES}}} \le e_{t,i_{\text{ES}}} \le \overline{e}_{i_{\text{ES}}} \tag{7}$$
$$(p^{\text{ES},\psi}_{t,i_{\text{ES}}}, q^{\text{ES},\psi}_{t,i_{\text{ES}}}) \in \mathcal{PQ}^{\text{ES},\psi}_{t,i_{\text{ES}}} \tag{8}$$

Similarly, $[N^{\text{ES}}]$ is the set of buses with ES connections. The charging and discharging power variables $\widehat{p}^{\text{ES},\psi}_{t,i_{\text{ES}}}$, $\check{p}^{\text{ES},\psi}_{t,i_{\text{ES}}}$

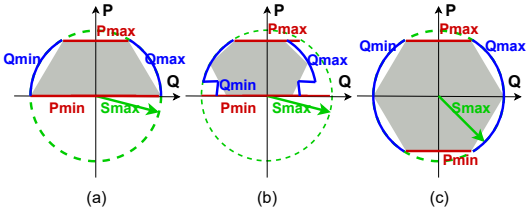

Figure 1. The PQ capability chart for (a) PV, (b) WT, and (c) ES.

are modelled separately to account for different charging (discharging) efficiencies $\widehat{\eta}$ ($\check{\eta}$). Eq. (4)-(5) model the charging and discharging power limits. Eq. (6) describes the relationship between the power of the ES and its energy level $e_{t,i_{\text{ES}}}$, which needs to be within the energy limits in (7). Here, Δt is the length of one time-step. ES can also provide reactive power as (8), and its PQ chart is visualised in Fig. 1(c).

Following [21], the PQ-charts of RG and ES are linearized as the grey areas in Fig. 1 to avoid non-linearity.

3) Controllable Load: The VPP manages a set of flexible loads that can be curtailed:

$$\begin{aligned} &\forall t \in [T], \forall i_{\text{CL}} \in [N^{\text{CL}}], \\ &\sigma_{\min} \tilde{p}^{\text{CL},\psi}_{t,i_{\text{CL}}} \le p^{\text{CL},\psi}_{t,i_{\text{CL}}} \le \tilde{p}^{\text{CL},\psi}_{t,i_{\text{CL}}} \qquad (9) \\ &q^{\text{CL},\psi}_{t,i_{\text{CL}}} = \alpha \cdot p^{\text{CL},\psi}_{t,i_{\text{CL}}} \qquad (10) \end{aligned}$$

where (9) means the CL can be curtailed to a $\sigma_{\min}$ proportion of its base load $\tilde{p}^{\text{CL},\psi}_{t,i_{\text{CL}}}$. We assume that CL has a fixed power factor α as (10). Lastly, non-controllable loads (NL) exist within the VPP, characterized by $\sigma_{\min} = 1$.

B. Network Model

The VPP has internal network constraints. A linearised unbalanced distribution network model proposed in [22] is adopted, with a global performance better than the well-known first-order Taylor's approximation. Let $\boldsymbol{x}_t$ collect the active and reactive power of all the DERs plus NLs at time-step t as

$$\begin{aligned} \boldsymbol{x}_t \coloneqq &[p^{d,\psi}_{t,i_d}, q^{d,\psi}_{t,i_d}, \tilde{p}^{\text{NL},\psi}_{t,i_{\text{NL}}}, \alpha\tilde{p}^{\text{NL},\psi}_{t,i_{\text{NL}}}] \qquad (11) \\ &i_d \in [N^d], d \in \{\text{RG}, \text{ES}, \text{CL}\}, i_{\text{NL}} \in [N^{\text{NL}}], \psi \in \Psi \end{aligned}$$

Let the vector $\boldsymbol{v}_t$ gather the magnitudes of three-phase nodal voltages for all buses, and let $\boldsymbol{i}_t$ encompass all the line currents at time-step t. The network constraints can be expressed as:

$$\begin{aligned} \forall t \in [T],\ \boldsymbol{i}_t &= \boldsymbol{J}_t \boldsymbol{x}_t + \boldsymbol{j}^0_t \qquad (12) \\ \boldsymbol{v}_t &= \boldsymbol{K}_t \boldsymbol{x}_t + \boldsymbol{k}^0_t \qquad (13) \\ p^{\text{VPP}}_t &= \boldsymbol{g}^\top_t \boldsymbol{x}_t + g^0_t \qquad (14) \\ \boldsymbol{i}_{\min} \le \boldsymbol{i}_t &\le \boldsymbol{i}_{\max} \qquad (15) \\ \boldsymbol{v}_{\min} \le \boldsymbol{v}_t &\le \boldsymbol{v}_{\max} \qquad (16) \end{aligned}$$

where $\boldsymbol{J}_t$, $\boldsymbol{j}^0_t$, $\boldsymbol{K}_t$, $\boldsymbol{k}^0_t$, $\boldsymbol{g}_t$, g^0_t are coefficients that characterize the approximated linear relationship encompassing the DER power $\boldsymbol{x}_t$, VPP power p^{VPP}_t at time-step t, nodal voltage $\boldsymbol{v}_t$, and line current $\boldsymbol{i}_t$. These coefficients are derived from a given operational point and a no-load point, as detailed in [22]. Eq. (15) and (16) describe the nodal voltage limits and the line current limits (thermal limits).

C. VPP Scheduling

During the DA phase, a VPP needs to determine its (aggregated) power $\boldsymbol{p}^{\text{VPP}} \coloneqq [p^{\text{VPP}}_{t=0}, \cdots, p^{\text{VPP}}_{t=T}]$ through the participation in the DA market. Its objective is to maximize its profits in the market considering DERs' running cost. When the next day comes, the VPP dispatches its managed DERs so that their aggregated power meets the DA schedule [5], [23], while minimizing the DER operation cost. Note that we do not focus on generators whose on/off status needs to be determined at the DA stage as well. The VPP scheduling problem in the DA stage with no uncertainty can be formulated as:

$$\begin{aligned} \min \quad & -\mathcal{J}(\boldsymbol{p}^{\text{VPP}}) + \mathcal{C}(\boldsymbol{x}) \qquad (17a) \\ \text{s.t.} \quad & (1) - (16) \qquad (17b) \end{aligned}$$

The constraints are the DER operational constraints and the network constraints (1)-(16). The objective consists of

1) The revenue of the VPP in engaging in the market $\mathcal{J}(\boldsymbol{p}^{\text{VPP}})$. For the DA energy market participation as a price taker, we have

$$\mathcal{J}(\boldsymbol{p}^{\text{VPP}}) = -\Delta t \sum_t \pi^{\text{DA}}_t p^{\text{VPP}}_t \qquad (18)$$

where π^{DA} is the DA market price.

2) The DER operational cost $\mathcal{C}(\boldsymbol{x})$. Here we only consider the cost for CL and ES, and assume that RG has zero marginal costs. We follow [5] and [16] and model the DER operational cost as:

$$\begin{aligned} \mathcal{C}(\boldsymbol{x}) = &\omega^{\text{ES}} \Delta t \sum_{t,\psi,i^{\text{CL}}} (\tilde{p}^{\text{ES},\psi}_{t,i_{\text{ES}}} - \check{p}^{\text{ES},\psi}_{t,i_{\text{ES}}}) \\ &+ \omega^{\text{CL}} \Delta t \sum_{t,\psi,i^{\text{CL}}} (\tilde{p}^{\text{CL},\psi}_{t,i_{\text{CL}}} - p^{\text{CL},\psi}_{t,i_{\text{CL}}}) \end{aligned} \qquad (19)$$

where ω^{ES} is the coefficient for ES degradation penalty cost, and ω^{CL} accounts for the load shedding cost.

III. Methodology

In the DA scheduling stage, the power capability of RG, CL, and NL is uncertain. Therefore, a VPP operator wants to ensure that the scheduled VPP power $\boldsymbol{p}^{\text{VPP}}$ is available in real-time operation without violating DER operational constraints and network constraints with a high probability. This objective can be expressed in the following symbolic way:

$$\begin{aligned} \min \quad & -\mathcal{J}(\boldsymbol{p}^{\text{VPP}}) + \mathcal{C}(\boldsymbol{x}) \qquad (20a) \\ \text{s.t.} \quad & \mathbb{P}(\boldsymbol{p}^{\text{VPP}} \text{ is implementable}) \ge 1 - \epsilon \qquad (20b) \end{aligned}$$

Note that the uncertain parameters can also affect the objective function, motivating the minimization of expectation or worst-case cost considering multi-stage recourse actions [23]. However, the key focus of this study is the power availability of the scheduled VPP power, so the multi-stage setting is not considered here.

A. Standard Route

The standard route to ensure the power availability, like [5], [14], [16], would formulate all the low-level constraints containing random variables into a large JCC. Here, the random variables include the maximum active power capability of RG $\overline{p}^{\text{RG}}_{t,i_{\text{RG}}}$, the base load of CL $\tilde{p}^{\text{CL},\psi}_{t,i_{\text{CL}}}$ and NL $\tilde{p}^{\text{NL},\psi}_{t,i_{\text{NL}}}$, and thus a part of $\boldsymbol{x}$. The VPP scheduling problem becomes:

$$\min \quad -\mathcal{J}(\boldsymbol{p}^{\text{VPP}}) + \mathcal{C}(\boldsymbol{x}) \tag{21a}$$

$$\text{s.t.} \quad \mathbb{P}((1),\ (9),\ (15),\ (16)) \geq 1-\epsilon, \tag{21b}$$

$$(2),\ (3)-(8),\ (10) \tag{21c}$$

where $1-\epsilon$ is the joint constraint satisfaction probability. Eqs. (12) and (13) can be merged into their thermal limit (15) and voltage limit (16) by changing variables and are thus omitted. The relationship between VPP power $\boldsymbol{p}^{\text{VPP}}$ and DER power $\boldsymbol{x}$ in (14) can be integrated into the objective function (18) as well. Due to the large number of constraints in the JCC (21b), the Bonferroni-based method becomes overly conservative and exact reformulation becomes intractable.

B. Surrogate Polytope

For a VPP, the decision to be made at the DA stage is the VPP power $\boldsymbol{p}^{\text{VPP}}$, while the specific DER power can be dispatched in the real-world operation process with the gradual realisation of uncertain DERs' outputs. This motivates us to isolate the VPP power $\boldsymbol{p}^{\text{VPP}}$ from the low-level DER power $\boldsymbol{x}$. Notice that the VPP's power availability constraints under a deterministic case can be inner-approximated by specific surrogate polytope constraints with a pre-selected shape $\boldsymbol{A}$ [12]. In other words, define the set of feasible VPP power $\boldsymbol{p}^{\text{VPP}}$ governed by the DER constraints and network constraints as $\Omega = \{\boldsymbol{p}^{\text{VPP}} | (1)-(16) \text{ (i.e., } \boldsymbol{p}^{\text{VPP}} \text{ is implementable)}\}$, we have

$$\exists \Omega^{\text{P}} \subseteq \Omega,\ \Omega^{\text{P}} = \{\boldsymbol{p}^{\text{VPP}} | \boldsymbol{A}\boldsymbol{p}^{\text{VPP}} \leq \boldsymbol{b}\} \tag{22}$$

where the vector $\boldsymbol{b}$ needs to be inferred. As we can observe, Ω^{P} in (22) is defined for the low-dimension vector $\boldsymbol{p}^{\text{VPP}}$ only without considering the DER power $\boldsymbol{x}$, but the relationship $\Omega^{\text{P}} \subseteq \Omega$ ensures that all the VPP power $\boldsymbol{p}^{\text{VPP}}$ in Ω^{P} will also be within the set Ω, which means the numerous DER constraints and network constraints (1)-(16) are implicitly satisfied.

A common and intuitive polytope is the virtual battery (VB), whose constraint form is like a battery with energy and power constraints. The VB polytope can be written as:

$$\Omega^{\text{VB}} = \{\boldsymbol{p}^{\text{VPP}} | \boldsymbol{A}^{\text{VB}}\boldsymbol{p}^{\text{VPP}} \leq \boldsymbol{b}^{\text{VB}}\} \tag{23}$$

$$\boldsymbol{A}^{\text{VB}} = [\boldsymbol{I}, -\boldsymbol{I}, \boldsymbol{\Lambda}, -\boldsymbol{\Lambda}]^{\top} \tag{24}$$

$$\begin{aligned}\boldsymbol{b}^{\text{VB}} = [&\overline{p}^{\text{VB}}_{t=1}, \cdots, \overline{p}^{\text{VB}}_{t=T}, -\underline{p}^{\text{VB},}_{t=1}, \cdots, -\underline{p}^{\text{VB},}_{t=T}, \\ &\overline{e}^{\text{VB}}_{t=1}, \cdots, \overline{e}^{\text{VB}}_{t=T}, -\underline{e}^{\text{VB}}_{t=1}, \cdots, -\underline{e}^{\text{VB}}_{t=T}]\end{aligned} \tag{25}$$

where $\boldsymbol{I} \in \mathbb{R}^{T\times T}$ is the identity matrix and $\boldsymbol{\Lambda} \in \mathbb{R}^{T\times T}$ is a lower-triangle matrix with ones. $\boldsymbol{b}^{\text{VB}}$ encompasses the VB parameters including the maximum power $\overline{p}^{\text{VB}}_t$, minimum power $\underline{p}^{\text{VB}}_t$, maximum energy $\overline{e}^{\text{VB}}_t$, and minimum energy $\underline{e}^{\text{VB}}_t$ at each time-step $t \in [T]$.

Eq. (22) or its special case (23) is only imposed on the VPP power $\boldsymbol{p}^{\text{VPP}}$ for the scheduling horizon $[T]$ and implicitly includes DER operational constraints and network constraints. For a pre-selected shape $\boldsymbol{A}$, the number of linear constraints in (22) only depends on $\text{Dim}(\boldsymbol{p}^{\text{VPP}})$=$T$.

C. Proposed Solution to Joint Chance Constraint

Due to the uncertain DERs, the VPP power availability set Ω is also uncertain and so is the surrogate polytope Ω^{P}. As the shape $\boldsymbol{A}$ is pre-selected, the random part in (26) is $\boldsymbol{b}$; the surrogate polytope with a joint probabilistic guarantee can be expressed using the following JCC:

$$\mathbb{P}(\boldsymbol{A}\boldsymbol{p}^{\text{VPP}} \leq \boldsymbol{b}) \geq 1-\epsilon \Leftrightarrow F_{-\boldsymbol{b}}(-\boldsymbol{A}\boldsymbol{p}^{\text{VPP}}) \geq 1-\epsilon \tag{26}$$

where $F(\cdot)$ is the cumulative distribution function (CDF) for the random variable $-\boldsymbol{b}$. Due to the inner approximation of the surrogate polytope, for any solution $\hat{\boldsymbol{p}}^{\text{VPP}}$ feasible to (26), we have the following implication:

$$F_{-\boldsymbol{b}}(-\boldsymbol{A}\hat{\boldsymbol{p}}^{\text{VPP}}) \geq 1-\epsilon \Leftrightarrow \mathbb{P}(\boldsymbol{A}\hat{\boldsymbol{p}}^{\text{VPP}} \leq \boldsymbol{b}) \geq 1-\epsilon \tag{27}$$

$$\Leftrightarrow \mathbb{P}(\hat{\boldsymbol{p}}^{\text{VPP}} \in \Omega^{\text{P}}) \geq 1-\epsilon \tag{28}$$

$$\Rightarrow \mathbb{P}(\hat{\boldsymbol{p}}^{\text{VPP}} \in \Omega) \geq 1-\epsilon \tag{29}$$

which means any solution satisfying our surrogate-polytope JCC (26) also satisfies the original JCC (20b).

We fit $\boldsymbol{b}$ by a multivariate Gaussian distribution; then based on Theorem 10.2.1 in [19], the CDF $F_{-\boldsymbol{b}}$ is a log-concave function (and also quasiconcave) and subsequently the surrogate JCC (26) is convex based on the definition of quasiconcave function in Chapter 3.4.1 of [24]. The optimization problem with this JCC can then be solved by an iterative supporting hyperplane algorithm [19]. The core idea is to replace the surrogate polytope JCC in an abstract form (26) with a set of iteratively tightened linear constraints denoted by $\mathcal{K}^j$, with j denoting the index of the iteration. The iteration process ends when the scheduling problem based on $\mathcal{K}^j$ leads to a solution $\boldsymbol{p}^{\text{VPP}*}$ satisfying $F_{-\boldsymbol{b}}(-\boldsymbol{A}\boldsymbol{p}^{\text{VPP}*}) \geq 1-\epsilon$.

At the start of the algorithm, $\mathcal{K}^0$ needs to be initialized to be large enough to cover the convex region defined by the JCC (26). One option is defining a set of ICCs of the polytope constraints $\boldsymbol{A}\boldsymbol{p}^{\text{VPP}} \leq \boldsymbol{b}$ with the same probability level as the JCC:

$$\mathcal{K}^0 = \{\mathbb{P}(\boldsymbol{a}_i^{\top}\boldsymbol{p}^{\text{VPP}} \leq b_i) \geq 1-\epsilon,\ \forall i \in [\text{Dim}(\boldsymbol{b})]\} \tag{30}$$

$$= \{\boldsymbol{a}_i^{\top}\boldsymbol{p}^{\text{VPP}} \leq F^{-1}_{b_i}(\epsilon),\ \forall i \in [\text{Dim}(\boldsymbol{b})]\} \tag{31}$$

where $\boldsymbol{a}_i$ (b_i) is the i^{th} row of $\boldsymbol{A}$ ($\boldsymbol{b}$ respectively). $F^{-1}_{b_i}(\epsilon)$ is the inverse CDF (percentile function) of a one-dimension random variable b_i that can be trivially evaluated. In addition, the initialization involves finding a feasible solution $\tilde{\boldsymbol{p}}^{\text{VPP}}$ to the JCC (26), which can be obtained by solving the following feasibility problem constrained by the conservative Bonferroni approximation to the surrogate JCC (26):

$$\min_{\boldsymbol{p}^{\text{VPP}}} \quad 0 \tag{32a}$$

$$\text{s.t.} \quad \boldsymbol{a}_i^{\top}\boldsymbol{p}^{\text{VPP}} \leq F^{-1}_{b_i}(\epsilon/\text{Dim}(\boldsymbol{b})),\ \forall i \in [\text{Dim}(\boldsymbol{b})] \tag{32b}$$

An iterative process starts after initialization. Over the j^{th} iteration with $j \geq 0$, the following optimization problem is solved:

$$\min_{\boldsymbol{p}^{\text{VPP}}} \quad -\mathcal{J}(\boldsymbol{p}^{\text{VPP}}) + \mathcal{C}'(\boldsymbol{p}^{\text{VPP}}) \tag{33a}$$

$$\text{s.t.} \quad \mathcal{K}^j \tag{33b}$$

where $\mathcal{C}'$ is the transformed DER cost function and will be introduced in Section III-E. Denote the optimal solution to (33) as $\boldsymbol{p}^{\text{VPP}*,j}$, we then evaluate constraint (26) by calculating the CDF $F_{-\boldsymbol{b}}(-\boldsymbol{A}\boldsymbol{p}^{\text{VPP}*,j})$. If constraint (26) is satisfied, then the problem is solved; otherwise, the algorithm finds a new boundary solution $\boldsymbol{p}^{\text{VPP,B},j}$ between the optimal solution $\boldsymbol{p}^{\text{VPP}*,j}$ over the j^{th} iteration and the initialized feasible solution $\tilde{\boldsymbol{p}}^{\text{VPP}}$:

$$\boldsymbol{p}^{\text{VPP,B},j} = \lambda \boldsymbol{p}^{\text{VPP}*,j} + (1-\lambda)\tilde{\boldsymbol{p}}^{\text{VPP}} \tag{34}$$

such that $F_{-\boldsymbol{b}}(-\boldsymbol{A}p^{\text{VPP,B},j}) = 1-\epsilon$. The value of $\lambda \in [0,1]$ can be obtained by bisection [19]. Note that the boundary point $\boldsymbol{p}^{\text{VPP,B},j}$ is a feasible solution to the JCC (26) but not necessarily an optimal solution. Based on this boundary point, a new linear constraint representing the hyperplane tangent to the JCC boundary at $\boldsymbol{p}^{\text{VPP,B},j}$ is constructed and added to $\mathcal{K}^j$. The new constraint is:

$$\nabla_{p^{\text{VPP}}} F_{-\boldsymbol{b}}(-\boldsymbol{A}\boldsymbol{p}^{\text{VPP,B},j})^\top (\boldsymbol{p}^{\text{VPP}} - \boldsymbol{p}^{\text{VPP,B},j}) \geq 0 \tag{35}$$

where the gradient of the multivariate Gaussian CDF can be calculated based on the Lemma 2.7.5 in [25].

The above procedures have a convergence guarantee [19], which means there will finally be one iteration such that the solution $\boldsymbol{p}^{\text{VPP}*,j}$ satisfies the surrogate JCC constraint (26). Suppose this is the M^{th} iteration, and the final VPP scheduling problem with our surrogate polytope JCC (26) can be expressed as the following problem with a set of linear constraints:

$$\min_{\boldsymbol{p}^{\text{VPP}}} \quad -\mathcal{J}(\boldsymbol{p}^{\text{VPP}}) + \mathcal{C}'(\boldsymbol{p}^{\text{VPP}}) \tag{36a}$$

$$\text{s.t.} \quad \mathcal{K}^M \tag{36b}$$

with

$$\mathcal{K}^M = \left\{ \begin{array}{l} \boldsymbol{a}_i^\top \boldsymbol{p}^{\text{VPP}} \leq F_{b_i}^{-1}(\epsilon),\ \forall i \in [\text{Dim}(\boldsymbol{b})], \\ \nabla_{p^{\text{Agg}}} F_{-\boldsymbol{b}}(-\boldsymbol{A}\boldsymbol{p}^{\text{VPP,B}j})^\top (\boldsymbol{p}^{\text{VPP}} - \boldsymbol{p}^{\text{VPP,B}j}) \geq 0, \\ \qquad\qquad\qquad\qquad \forall j \in [M] \end{array} \right\} \tag{37}$$

The optimal solution $\boldsymbol{p}^{\text{VPP}*}$ ($\boldsymbol{p}^{\text{VPP}*,j}$ with $j=M$) to (36) is the DA VPP scheduled power with a $1-\epsilon$ joint probabilistic guarantee. The whole process of the supporting hyperplane algorithm is summarized in Algorithm 1.

Our method differs from existing work with exact JCC solutions like [5] and [17] in that we reformulate the surrogate polytope JCC for the VPP's aggregated power, rather than the JCC for the low-level DERs and network constraints. For a pre-selected polytope shape $\boldsymbol{A}$, the number of random variables in (26), i.e., len($\boldsymbol{b}$), only depends on Dim($\boldsymbol{p}^{\text{VPP}}$)=$T$, leading to a manageable complexity to exactly solve the JCC as will be demonstrated in Section IV-H.

Algorithm 1 Supporting hyperplane algorithm

```
Initialize K^j ← K^0 as (31);
Initialize the feasible solution p̃^VPP by solving (32);
Initialize j ← 0;
while True do
    Get the j^th optimal solution p^{VPP*,j} by solving (33);
    if F_{-b}(−A p^{VPP*,j}) ≥ 1 − ε then
        Stop; output p^{VPP*} ← p^{VPP*,j};
    end if
    Find the boundary solution p^{VPP,B,j} (34) by bisection;
    Append the new linear constraint (supporting hyperplane) (35) to K^j;
    j ← j + 1;
end while
```

Remark 1: Based on Theorem 10.2.1 in [19], the convexity of our proposed surrogate JCC (26) holds for all probability distributions with log-concave density, including skewed distributions like Weibull, multivariate Gamma, and Dirichlet (the multivariate generalisation of the Beta) distributions with shape parameters greater than or equal to 1, and the heavily tailed hyperbolic secant distribution. These distributions are applicable for modelling a wide range of uncertain variables. However, only a subset of log-concave functions, like multivariate Gaussian, Gamma, and Dirichlet, have efficient gradient calculation when applying our supporting hyperplane algorithm [25]. Other solution algorithms, such as [26], which avoids the gradient evaluation of the CDF, may be investigated. Gaussian mixtures under certain conditions may also be applied, as they can preserve log-concavity and their CDF is the weighted summation of Gaussian CDFs; thus the gradient can be obtained. Note that although we fit $\boldsymbol{b}$ with a multivariate Gaussian distribution, the low-level random variables, i.e., DER power capability, do not necessarily need to follow a Gaussian distribution. Our case studies will verify this by using non-Gaussian real-world wind and solar scenarios.

Remark 2: The convexity of (26) and the gradient calculation in the supporting hyperplane algorithm require the distribution of $\boldsymbol{b}$ to be non-degenerate. Non-degeneracy can be guaranteed by adding σI to the covariance matrix, where σ is a small positive number and I is the identity matrix. Since all covariance matrices are positive semi-definite, σ can be set arbitrarily small to minimize the disturbance on the distribution while maintaining the positive definite property of the covariance, so as the non-degeneracy of the distribution.

Remark 3: This paper applies the surrogate-polytope JCC in (26) to substitute the original JCC in (20b). Another popular alternative in power system applications is DRJCC, known for its improved robustness and better tractability. However, since the use of the inner approximation polytope in (22) already makes our surrogate-polytope JCC an inner approximation to the original JCC, further enhanced robustness by (potentially more tractable) DRJCC may introduce unnecessary conservativeness. Future work can explore this trade-off.

D. Fitting the Gaussian

We collect a set of possible scenarios $\boldsymbol{b}^k$ of $\boldsymbol{b}$ to fit the multivariate Gaussian distribution of $\boldsymbol{b}$. To collect the scenarios, we first decompose $\boldsymbol{b}$ into a deterministic part $\boldsymbol{b}^{\text{fore}}$ plus a random part $\boldsymbol{b}^{\text{error}}$. This decomposition captures the dependency of $\boldsymbol{b}$ by a specific scheduling day by $\boldsymbol{b}^{\text{fore}}$, and is commonly used in stochastic power system scheduling formulations [16].

The deterministic part $\boldsymbol{b}^{\text{fore}}$ needs to be evaluated in the scheduling stage based on the forecasted DERs' power capability. The scenarios for the random part $\boldsymbol{b}^{\text{error}}$ can be obtained in the following offline manner: for each historical scheduling day k, once that historical day k is passed (uncertainty revealed), 1) calculate the historical deterministic part $\boldsymbol{b}^{k,\text{fore,h}}$ for day k using the forecasted DER power capability; the superscript ^{h} means the value is for historical days; 2) calculate the vector $\boldsymbol{b}^{k,\text{h}}$ for the day k using real values; 3) store their difference as the scenario $\boldsymbol{b}^{k,\text{error}} = \boldsymbol{b}^{k,\text{h}} - \boldsymbol{b}^{k,\text{fore,h}}$ for our random part $\boldsymbol{b}^{\text{error}}$. The bound shrinking algorithm proposed in [12] can infer the values of $\boldsymbol{b}$ given certain values of DERs' power capability, and is used to calculate the aforementioned $\boldsymbol{b}^{\text{fore}}$, $\boldsymbol{b}^{k,\text{fore,h}}$, and $\boldsymbol{b}^{k,\text{h}}$.

After adding each of the collected scenarios $\boldsymbol{b}^{k,\text{error}}$ for the random part $\boldsymbol{b}^{\text{error}}$ to the derived deterministic part $\boldsymbol{b}^{\text{fore}}$ in the scheduling phase, we obtain a set of scenarios for $\boldsymbol{b}^k$, which are used to fit a multivariate Gaussian distribution for $\boldsymbol{b}$.

Note that an accurate probabilistic fitting of the random part $\boldsymbol{b}^{\text{error}}$ needs to model the dependency on $\boldsymbol{b}^{\text{fore}}$, namely a conditional probability function, which may need more advanced forecasting techniques and more historical data. The applied method here may be less accurate, but has the advantages of requiring less data and being easier to implement.

E. Cost Function

The use of surrogate polytope for the VPP power $\boldsymbol{p}^{\text{VPP}}$ makes DER power hidden, so the DER cost function $\mathcal{C}(\boldsymbol{x})$ in the objective can not be straightforwardly calculated. However, we can find the aggregated DER cost function $\mathcal{C}'(\boldsymbol{p}^{\text{VPP}})$ by establishing a cost mapping from $\boldsymbol{x}$ to $\boldsymbol{p}^{\text{VPP}}$, via the piece-wise linear fitting in [21]. The intuition is to find the piecewise-linear function that fits the minimum achievable DER cost for each VPP power $\boldsymbol{p}^{\text{VPP}}$. For each $t \in [T]$, we first find the minimum and maximum active output power of the VPP, denoted as $\underline{p}_t^{\text{VPP}}$ and $\overline{p}_t^{\text{VPP}}$. Then, P points in total $\{p_{t,0}^{\text{VPP}}, \cdots, p_{t,p}^{\text{VPP}}, \cdots, p_{t,P}^{\text{VPP}}\}$ are sampled equally in interval $[\underline{p}_t^{\text{VPP}}, \overline{p}_t^{\text{VPP}}]$. For each of the sampled $p_{t,p}^{\text{VPP}}$, we solve the following problem to find the corresponding minimum total DER operational cost $C'_{t,p}$:

$$C'_{t,p} = \min_{\boldsymbol{x}, p_t^{\text{VPP}}} \quad \mathcal{C}(\boldsymbol{x}) \tag{38a}$$

$$\text{s.t.} \quad (1) - (16), \tag{38b}$$

$$p_t^{\text{VPP}} = p_{t,p}^{\text{VPP}} \tag{38c}$$

The pair $(p_{t,p}^{\text{VPP}}, C'_{t,p})$ establishes the one-to-one mapping from the sampled VPP power $p_{t,p}^{\text{VPP}}$ to the achievable minimum DER cost $C'_{t,p}$. Repeat solving (38) for all the P sample points leads to P pairs of $(p_{t,p}^{\text{VPP}}, C'_{t,p})$, based on which a convex piece-wise linear function can be fitted to approximate the aggregated DER cost function $\mathcal{C}'(\boldsymbol{p}^{\text{VPP}})$ with $P-1$ partitions. The above process is repeated for each time step $t \in [T]$. The fitted convex piece-wise linear function can be expressed as:

$$\mathcal{C}'(\boldsymbol{p}^{\text{VPP}}) = \min_{\boldsymbol{p}^{\text{VPP}}} \quad \sum_{t\in[T]} c_t \tag{39a}$$

$$\text{s.t.} \quad a_{t,p} p_t^{\text{VPP}} + b_{t,p} \le c_t, \quad p \in [P-1] \tag{39b}$$

where $a_{t,p}$ and $b_{t,p}$ are parameters of the fitted peice-wise linear function derived via the method in [21]. This expression (39) can be easily incorperated into other optimization problems. For example, in our VPP scheduling (36), we only need to replace $\mathcal{C}'(\boldsymbol{p}^{\text{VPP}})$ in the objective with the linear term $\sum_{t\in[T]} c_t$, define continuous variables c_t, and add the linear constraints (39b) into the constraints in (36), after which our VPP scheduling (36) is still a tractable.

Figure 2 summarized the flowchart of our proposed method, covering the content from Section III-C to III-E.

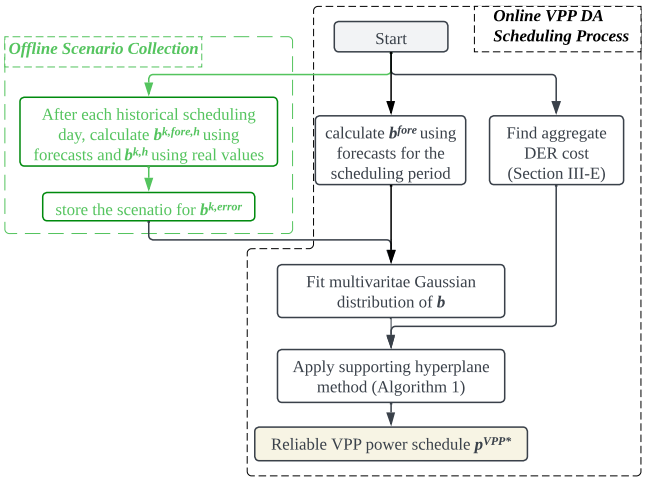

Figure 2. Flowchart of the proposed VPP DA scheduling with joint probability guarantee on VPP power availability.

F. DER Power Dispatch

After determining the scheduled VPP power $\boldsymbol{p}^{\text{VPP}*}$ in the DA phase, we need to retrieve the DER power $\boldsymbol{x}$ in real-time operation. An optimization problem with a similar form to (17) can be set up, with the only difference being that the decision variable $\boldsymbol{p}^{\text{VPP}}$ in (17) is now replaced with the optimal solution $\boldsymbol{p}^{\text{VPP}*}$ to (36). Note that, in the real-world VPP dispatch process, the uncertainty still exists (although gradually revealed). Here, we do not address the disaggregation process under uncertainty, which can be explored in future work.

IV. Case Studies

A. Settings of the Network and the Proposed Method

The case studies are carried out on the IEEE-123 distribution feeder [27] modelled in OPEN [28]. Codes are written in Python, and we use the `approxcdf` package [29] and the

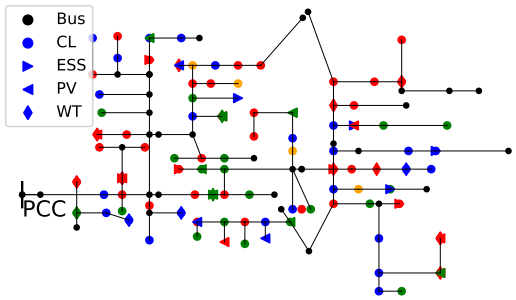

Figure 3. DER locations on IEEE-123 distribution network: red for phase a or ab, green for phase b or bc, and blue for the remaining phases.

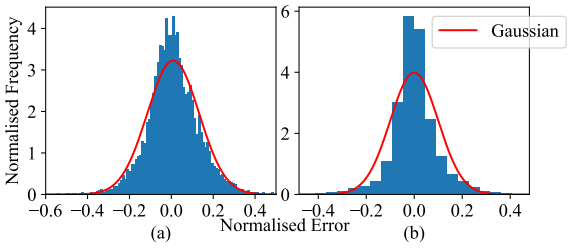

Figure 4. Histogram of the real-world (a) wind and (b) solar forecasting error.

`stats.mvn.mvnun` function in `SciPy` package to evaluate the CDF of multivariate Gaussian distributions. There are 20 PVs, 20 WTs, 10 ESs, 30 CLs, and 65 NLs. The DER location and phase connections are plotted in Fig. 3. We assume that CL can be curtailed by 40%, which is the upper bound of the forecasted demand curtailment potential in the UK by 2050 [30]. The voltage limits are set to be 1 ± 0.05 p.u. Line current limits are not considered as the voltage limits can be the main issue for certain distribution networks [31]. The peak of the base demand in the whole network is around 3500 kW. To fully validate the proposed framework, we consider three case studies: 1) 'LowRG': PV and WT capacities are both around 800 kW in total; 2) 'HighWT': the total WT capacity is set to 3500 kW; and 3) 'HighPV': the total PV capacity is set to 3500 kW. In the latter two cases, the distribution network is self-sufficient with local RG. The total ES capacity remains 3600 kWh across all case studies.

For the proposed surrogate-polytope method, we set the shape of the polytope to be the VB in (23), whose parameter $\boldsymbol{b}$ is inferred by the bound-shrinking method in [12].

B. Benchmarks

Two benchmarks are set up for our case studies. They are both based on the standard route as (21). The first one (B1) uses safe Bonferroni approximation to approximate the JCC in (21b), i.e., replacing the JCC with ICCs of reliability $1 - \epsilon/(\text{number of ICCs})$. The second one (B2) is a standard ICC with risk levels identical to those in the JCC, which can be less reliable. Exact JCC reformulations like [5] and [17] are not applied due to the excessive complexity under the large number of DERs and the large network size.

C. DER Data

The data for RG and network load (CL and NL) consists of two parts. The first is their prediction, which is the known information in the scheduling phase. The second is the forecasting error scenarios, which cannot be perfectly informed in advance. The prediction curves for PV and WT come from real-world measurement in [32], and the prediction for network load comes from a US commercial building dataset [33]. The RG forecasting error scenarios come from real-world statistics in [34], [35], whose histogram is plotted in Fig. 4. We can see that these RG error scenarios are not normally distributed by comparing to the corresponding Gaussian fits. On the other hand, the network load error scenarios are synthesized from a Gaussian distribution with a standard deviation equal to 5% of the load predictions. 200 samples are used as the training set for our proposed method and the benchmarks.

D. Evaluation Setting

We consider a DA cost minimization problem for a VPP, where $\mathcal{J}(\boldsymbol{p}^{\text{VPP}})$ is set to (18). We consider an hourly resolution and hence $T = 24$. The DER cost coefficients ω^{ES} and ω^{CL} are both set to 8 p/kWh as [5]. The random variables in the objective function are replaced with their forecast mean values. We consider two evaluations using 200 out-of-sample error scenarios for each of the 24 time-steps:

1) Out-of-sample reliability of the optimal solution $\boldsymbol{p}^{\text{VPP}*}$ to (36), which is defined as the proportion of out-of-sample scenarios where the disaggregation process outlined in Section III-F is feasible. We employ a nonlinear power flow simulation using the Z-Bus method [36] to assess violations of network constraints. To account for the linear network model's inaccuracies, we introduce a penalty term for the small distance between the voltage magnitude and its limits in the disaggregation objective in Section III-F. The penalty coefficient is set to 0.1 for both our proposed method and the benchmarks. The solution is reliable as long as the out-of-sample reliability is greater or equal to the desired level $1 - \epsilon$.
2) Out-of-sample cost of the optimal solution $\boldsymbol{p}^{\text{VPP}*}$ to (36), calculated by averaging the minimized cost of the disaggregation problem in Section III-F for all the disaggregatable scenarios. The lower the less conservative.

E. Out-of-sample Reliability Results

Fig. 5 shows the out-of-sample solution reliability for (a) 'LowRG', (b) 'HighPV', and (c) 'HighWT', where the black diagonal lines are the desired out-of-sample reliability levels. We see that both our proposed method and B1 can meet the reliability requirements for all the case studies, but B2 cannot. This result suggests that ICC (B2) can lead to unreliable solutions, highlighting the need for JCC.

F. Out-of-sample Cost Results

Fig. 6 presents the out-of-sample cost for the three RG case studies. While B2 achieves the lowest cost for all cases, its unreliability, as demonstrated in Section IV-E, compromises its suitability. Our proposed method, being the second lowest

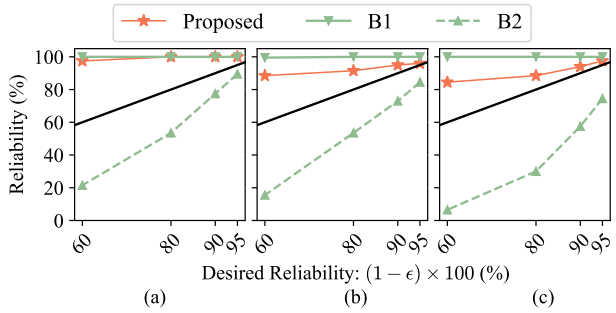

Figure 5. Out-of-sample reliability evaluation for (a) 'LowRG', (b) 'HighPV', and (c) 'HighWT'. The black lines indicate the desired reliability levels.

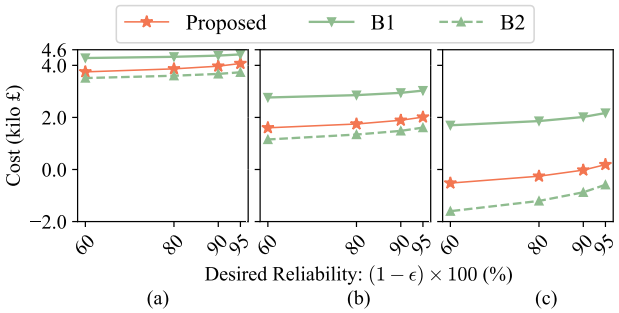

Figure 6. Out-of-sample DA cost evaluation for (a) 'LowRG', (b) 'HighPV', and (c) 'HighWT'.

in terms of cost, closely approaches the lowest costs achieved by B2. When considering the reliability evaluation detailed in Section IV-E, we conclude that our proposed method not only offers reliable solutions but also exhibits less conservativeness.

G. Scalability

To verify the applicability of our proposed method in wide-range problem scales, we evaluate the out-of-sample cost and reliability under: the total DER number (excluding NL) being 1) 50, 2) 80, and 3) 120. Figs. 7 (a), (b), (c), and (d) display the results for the 'HighWT' case study. We can see the conclusions from previous sections still hold: our proposed method can constantly meet the desired reliability requirements and is less conservative than B1. B2 has a lower cost but cannot meet the reliability requirements for all cases.

H. Overall Computing Time

Based on the flowchart in Fig. 2, the online calculations involve four steps: 1) computing $\boldsymbol{b}^{\text{fore}}$ using the bound shrinking algorithm in [12], which involves several mixed-integer linear programmes (MILP) solvable in 100-200 seconds in total for DER numbers from 50 to 120 with an AMD Ryzen9 5900X CPU. Throughout the case studies we have relaxed the stopping bounds, capping the MILP solution time at 40 seconds; 2) employing the piece-wise linear approximation method from [21] to determine the aggregated DER cost function through $T \times P$ parallelizable linear programs (39) taking 1.5-3.5 seconds each, leading to a total of 200-500 seconds without parallelization; 3) fitting a multivariate Gaussian distribution for $\boldsymbol{b}$ with dimensions Dim($\boldsymbol{b}$)=96, trivial due to the small scale of the VB polytope ($T = 24$); and 4)

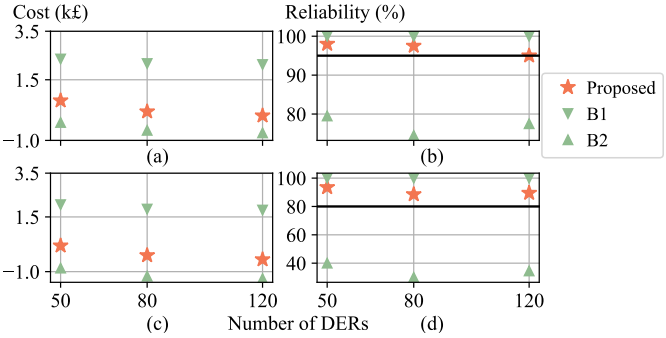

Figure 7. First column: out-of-sample cost under different numbers of DERs with desired reliability levels (a) $1-\epsilon = 0.95$ and (c) $1-\epsilon = 0.8$; Second column: out-of-sample reliability with desired reliability levels (b) $1-\epsilon = 0.95$ and (d) $1-\epsilon = 0.8$, where the black horizontal lines indicate the desired reliability levels.

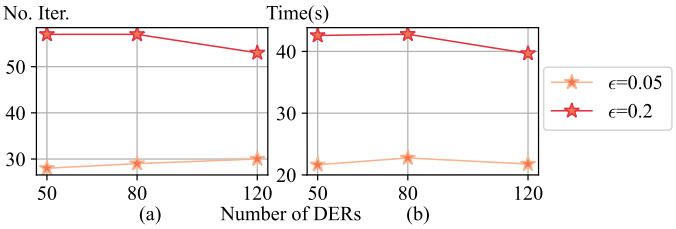

Figure 8. (a) The number of iterations (No. Iter.) in applying the proposed supporting hyperplane algorithm to solve the proposed surrogate JCC under different numbers of DERs; (b) The computing time in applying the proposed supporting hyperplane algorithm to solve the proposed surrogate JCC.

applying the supporting hyperplane method in Algorithm 1. Fig. 8 illustrates the number of iterations and the computing time in applying the supporting hyperplane method to solve the JCC. We can observe that both the number of iterations and the computing time are independent of the number of DERs. A lower level of reliability (a greater ϵ) requires more iterations due to the larger feasible region defined by the JCC. The computing time is around 20 seconds for a 95% reliability ($\epsilon = 0.05$), and around 40 seconds for an 80% reliability.

Implementing all the calculations sequentially results in a worst-case total computation time of approximately 12 minutes for 120 DERs in the IEEE-123 network with $T = 24$, manageable for a DA scheduling problem. Note that parallelization could be explored to further speed up the calculation.

V. Conclusion

This paper proposes a novel method for ensuring joint probability-guaranteed VPP power availability in DA scheduling. It starts from an inner surrogate polytope approximation for the VPP's power availability, which focuses on the VPP power only while implicitly modelling the DER constraints and network constraints. A multivariate Gaussian distribution is used to fit the distribution of the parameters in the surrogate polytope. Finally, an iterative algorithm is used to solve the VPP scheduling problem with our surrogate polytope JCC. Through case studies with different RG settings and varying problem sizes, we demonstrate that the proposed method has an overall better performance in out-of-sample reliability and

out-of-sample cost than an ICC benchmark and a Bonferroni benchmark. The computing complexity is acceptable, with a worst-case total computation time of approximately 12 minutes for 120 DERs in the IEEE-123 network with $T = 24$.

Future work includes 1) incorporating the multi-stage power dispatch process with a recourse cost modelled in the objective; 2) investigating a real-time strategy that dispatches the VPP power to its internal DERs, subject to gradually revealed uncertainty, given the VPP power that is available with (1-ϵ) joint probability scheduled by our proposed method; 3) comparison and synergy with other data-driven methods. Recent work [37] has demonstrated the promising performance of data-driven reinforcement learning methods in VPP scheduling but did not tackle the JCC as ours. There are potential synergies of our approach with data-driven RL methods to combine their improved performance in VPP scheduling with our method's JCCs.

VI. Acknowledgments

The work was supported by the Engineering Studentship from the University of Edinburgh and was also supported by the UK Engineering and Physical Sciences Research Council (EPSRC) (Project references EP/S031901/1, EP/T028564/1 and EP/W027321/1).

References

[1] T. Morstyn, N. Farrell, S. J. Darby, and M. D. McCulloch, "Using peer-to-peer energy-trading platforms to incentivize prosumers to form federated power plants," *Nature Energy*, 2018.

[2] Tesla. (2023) Tesla virtual power plant with pg&e. [Online]. Available: https://www.tesla.com/support/energy/tesla-virtual-power-plant-pge

[3] D. Pudjianto, C. Ramsay, and G. Strbac, "Virtual power plant and system integration of distributed energy resources," *IET Renewable power generation*, vol. 1, no. 1, pp. 10–16, 2007.

[4] M. Braun and P. Strauss, "A review on aggregation approaches of controllable distributed energy units in electrical power systems," *International Journal of Distributed Energy Resources*, vol. 4, no. 4, pp. 297–319, 2008.

[5] H. Liu, J. Qiu, and J. Zhao, "A data-driven scheduling model of virtual power plant using wasserstein distributionally robust optimization," *International Journal of Electrical Power & Energy Systems*, 2022.

[6] M. Shabanzadeh, M.-K. Sheikh-El-Eslami, and M.-R. Haghifam, "The design of a risk-hedging tool for virtual power plants via robust optimization approach," *Applied Energy*, vol. 155, pp. 766–777, 2015.

[7] Z. Tan, H. Zhong, Q. Xia, C. Kang, X. S. Wang, and H. Tang, "Estimating the robust p-q capability of a technical virtual power plant under uncertainties," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4285–4296, 2020.

[8] B. Cui, A. Zamzam, and A. Bernstein, "Network-cognizant time-coupled aggregate flexibility of distribution systems under uncertainties," in *2021 American Control Conference (ACC)*. IEEE, 2021, pp. 4178–4183.

[9] W. Wiesemann, D. Kuhn, and M. Sim, "Distributionally robust convex optimization," *Operations Research*, vol. 62, no. 6, pp. 1358–1376, 2014.

[10] A. Shapiro and A. Philpott, "A tutorial on stochastic programming," *www2. isye. gatech. edu/ashapiro/publications. html*, vol. 17, 2007.

[11] 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 on Power Systems*, 2020.

[12] S. Wang and W. Wu, "Aggregate flexibility of virtual power plants with temporal coupling constraints," *IEEE Trans. on Smart Grid*, 2021.

[13] S. Babaei, C. Zhao, and L. Fan, "A data-driven model of virtual power plants in day-ahead unit commitment," *IEEE Trans on Power Systems*, vol. 34, no. 6, pp. 5125–5135, 2019.

[14] A. Bagchi and Y. Xu, "Distributionally robust chance-constrained bidding strategy for distribution system aggregator in day-ahead markets," in *2018 SmartGridComm*. IEEE, 2018, pp. 1–6.

[15] H. Wu, M. Shahidehpour, Z. Li, and W. Tian, "Chance-constrained day-ahead scheduling in stochastic power system operation," *IEEE Trans on Power Systems*, vol. 29, no. 4, pp. 1583–1591, 2014.

[16] Y. Ding, T. Morstyn, and M. D. McCulloch, "Distributionally robust joint chance-constrained optimization for microgrids considering contingencies and renewable uncertainty," *IEEE Trans. on Smart Grid*, 2022.

[17] W. van Ackooij *et al.*, "An exact solution method for the hydrothermal unit commitment under wind power uncertainty with joint probability constraints," *IEEE Trans on Power Systems*, 2018.

[18] F. H. Aghdam, M. S. Javadi, and J. P. Catalão, "Optimal stochastic operation of technical virtual power plants in reconfigurable distribution networks considering contingencies," *International Journal of Electrical Power & Energy Systems*, vol. 147, p. 108799, 2023.

[19] A. Prékopa, *Stochastic programming*. Springer Science & Business Media, 2013, vol. 324.

[20] C. Ordoudis, V. A. Nguyen, D. Kuhn, and P. Pinson, "Energy and reserve dispatch with distributionally robust joint chance constraints," *Operations Research Letters*, vol. 49, no. 3, pp. 291–299, 2021.

[21] S. Wang, W. Wu, Q. Chen, J. Yu, and P. Wang, "Stochastic flexibility evaluation for virtual power plant by aggregating distributed energy resources," *CSEE Journal of Power and Energy Systems*, 2022.

[22] A. Bernstein and E. Dall'Anese, "Linear power-flow models in multiphase distribution networks," in *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*. IEEE, 2017, pp. 1–6.

[23] M. Rahimiyan and L. Baringo, "Strategic bidding for a virtual power plant in the day-ahead and real-time markets: A price-taker robust optimization approach," *IEEE Trans on Power Systems*, 2016.

[24] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[25] W. Van Ackooij, "Chance constrained programming: with applications in energy management," Ph.D. dissertation, Ecole Centrale Paris, 2013.

[26] W. Van Ackooij and W. de Oliveira, "Convexity and optimization with copulæ structured probabilistic constraints," *Optimization*, vol. 65, no. 7, pp. 1349–1376, 2016.

[27] K. P. Schneider *et al.*, "Analytic considerations and design basis for the ieee distribution test feeders," *IEEE Trans on power systems*, 2017.

[28] T. Morstyn *et al.*, "Open: An open-source platform for developing smart local energy system applications," *Applied Energy*, vol. 275, 2020.

[29] D. Cortes, E. Fuentes, and N. Belrose, "Approximate MVN CDF," accessed: 2024-02-08. [Online]. Available: https://approxcdf.readthedocs.io/en/latest/

[30] National Grid ESO, "Uk future energy scenarios." [Online]. Available: https://www.nationalgrideso.com/future-energy/future-energy-scenarios

[31] C. Heinrich *et al.*, "Pv-integration strategies for low voltage networks," in *2016 IEEE International Energy Conference*. IEEE, 2016, pp. 1–6.

[32] S. Pfenninger and I. Staffell, "Renewables.ninja." [Online]. Available: https://www.renewables.ninja/

[33] NREL, "End-use load profiles for the u.s. building stock," 10 2021. [Online]. Available: https://data.openei.org/submissions/4520

[34] B.-M. Hodge *et al.*, "Wind power forecasting error distributions: An international comparison," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2012.

[35] Y. Zhang, M. Beaudin, R. Taheri, H. Zareipour, and D. Wood, "Day-ahead power output forecasting for small-scale solar photovoltaic electricity generators," *IEEE Trans on Smart Grid*, 2015.

[36] M. Bazrafshan and N. Gatsis, "Comprehensive modeling of three-phase distribution systems via the bus admittance matrix," *IEEE Trans on Power Systems*, vol. 33, no. 2, pp. 2015–2029, 2018.

[37] B. Feng, Z. Liu, G. Huang, and C. Guo, "Robust federated deep reinforcement learning for optimal control in multiple virtual power plants with electric vehicles," *Applied Energy*, vol. 349, p. 121615, 2023.