

# Tractable Data Enriched Distributionally Robust Chance-Constrained CVR

Qianzhi Zhang, *Graduate Student Member, IEEE*, Fankun Bu, *Graduate Student Member, IEEE*,  
Yi Guo, *Member, IEEE*, and Zhaoyu Wang, *Senior Member, IEEE*,

**Abstract**—This paper proposes a tractable distributionally robust chance-constrained conservation voltage reduction (DRCC-CVR) method with enriched data-based ambiguity set in unbalanced three-phase distribution systems. The increasing penetration of distributed renewable energy not only brings clean power but also challenges the voltage regulation and energy-saving performance of CVR by introducing high uncertainties to distribution systems. In most cases, the conventional robust optimization methods for CVR only provide too conservative solutions. To better consider the impacts of load and PV generation uncertainties on CVR implementation in distribution systems and provide less conservative solutions, this paper develops a data-based DRCC-CVR model with tractable reformulation and data enrichment method. Even though the uncertainties of load and photovoltaic (PV) can be captured by data, the availability of smart meter (SMs) and micro-phasor measurement units (PMUs) is restricted by cost budget. The limited data access may hinder the performance of the proposed DRCC-CVR. Thus, we further present a data enrichment method to statistically recover the high-resolution load and PV generation data from low-resolution data with Gaussian Process Regression (GPR) and Markov Chain (MC) models, which can be used to construct a data-based moment ambiguity set of uncertainty distributions for the proposed DRCC-CVR. Finally, the nonlinear power flow and voltage dependent load models and DRCC with moment-based ambiguity set are reformulated to be computationally tractable and tested on a real distribution feeder in Midwest U. S. to validate the effectiveness and robustness of the proposed method.

**Index Terms**—Conservation voltage reduction, data enhancement, distributionally robust chance-constrained optimization, distribution systems, tractable reformulation.

## I. INTRODUCTION

CONSERVATION voltage reduction (CVR) can reduce the voltage for peak load shaving and long-term energy-saving [1]. To achieve system-wide optimal performance, voltage/var optimization-based CVR (VVO-CVR) is previously studied [1]–[4], which can be cast into an optimal power flow program. While the previous works have contributed valuable insights to VVO-CVR, there are problems remaining open, summarized as follows:

(1) *The impact of load and renewable uncertainties on VVO-CVR*: In [2], a linear least-squares centralized optimization

model is developed for coordinating combinations of voltage regulating devices and PVs to implement CVR in distribution systems. In [3] and [4], several alternating direction method of multipliers (ADMM)-based distributed optimization algorithms are developed for CVR implementation, which can decompose a large-scale VVO-CVR problem into several small-scale problems with improved scalability. However, the above centralized or distributed VVO-CVR works [2]–[4] are developed based on deterministic optimization methods, which assume the perfect forecasts of load and renewable power. Neglecting those prediction errors could result in potential violations of operational constraints, such as bus voltage constraints in VVO-CVR. To consider the impacts of load and renewable uncertainties on voltage regulation for CVR implementation, stochastic programming (SP) and robust optimization (RO) are applied in some existing works [5]–[8]. In [5] and [6], the scenario-based SP methods aim at optimizing reactive power dispatch for voltage regulation with expected performance and an accurate probability distribution model. While the SP methods need an accurate probability distribution model and a large sampling number of scenarios, it requires heavy computational efforts. In [7] and [8], the RO methods are developed to handle the uncertain load and renewable power production. However, the RO methods can only give a feasible solution for the worst-case scenario, which is too conservative and hinders the performance of CVR.

Recently, the distributionally robust optimization (DRO) has been considered a more effective way to handle uncertainty in power systems, such as economic and power dispatch [9] and [10]. The DRO can construct an ambiguity set of probability distributions based on historical data, including all possible uncertainty distributions. Thus, the DRO can ensure constraints are satisfied for any distributions in the ambiguity set built upon distribution moments or structure information. Also, the distributionally robust chance-constrained (DRCC) models [11] and [12] are developed in power systems, which integrate chance constraints to enforce certain events within a probability threshold. Even though DRO and DRCC models have some advantages over conventional deterministic and stochastic approaches, there are still challenges to use the DRCC method for CVR implementation, including how to formulate DRCC in a tractable way for CVR implementation and how to construct an ambiguity set with limited access to historical and real-time load and PV generation data.

(2) *Intractable DRCC and nonlinear voltage-dependent load models*: The VVO-CVR problem is nonlinear and intractable, which makes a distributionally robust stochastic reformula-

This work was supported in part by the U.S. Department of Energy Wind Energy Technologies Office under Grant DE-EE0008956, and in part by the National Science Foundation under ECCS 1929975. This work was partially supported by an ETH Zürich Postdoctoral Fellowship. (*Corresponding author: Zhaoyu Wang*).

Qianzhi Zhang, Fankun Bu, and Zhaoyu Wang are with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA (e-mail: qianzhi@iastate.edu; fbu@iastate.edu; wzy@iastate.edu).

Yi Guo is with Power Systems Laboratory at ETH Zürich, Zürich, 8092, Switzerland (e-mail: guo@eeh.ee.ethz.ch).

tion of the VVO-CVR problem even more challenging. In [13], an approximation method is proposed for conventional DRO problems, while the approximation for DRCC problems is not considered. Also, the loads are considered as voltage-independent models in above reference of DRO and DRCC [9]–[12]. In practice, the nature of CVR is lowering network voltages to reduce the voltage, and the literature on the CVR problem validates the necessity of voltage-dependent loads, such as ZIP load and exponential load models [3]. In existing research works [14] and [15], some approximation methods are applied to linearize ZIP load model in deterministic optimization. However, the linearization and convexification for ZIP load model in DRO or DRCC with uncertainties of load are not considered in previous works [9]–[15].

(3) *Availability of data for constructing ambiguity set of load and renewable uncertainties*: Even though the ambiguity set of uncertainty in DRCC can characterize a group of possibility distributions, defining a high-quality ambiguity set is non-trivial, as one needs to decide the trade-off between the conservativeness of decisions and the operational efficiency, while considering the mathematical tractability. To construct an ambiguity set of uncertainties in DRCC, the conventional way is using statistical inference and data analysis methods with the historical data and system feedback measurements. For example, the existing DRO works [16] and [17] use data-based method to construct ambiguity set. However, the aforementioned data-based ambiguity set relies heavily on either sufficiently high-resolution data or complex machine learning methods.

To capture the uncertainties of load and PV generation through data-based methods in modern distribution networks, micro-phasor measurement units (PMUs) and smart meters (SMs) are implemented to record load and PV generation data, where micro-PMUs<sup>1</sup> can record high-resolution data (1-second resolution or higher), and SMs<sup>2</sup> can record low-resolution data (typically 1-hour resolution). However, due to the cost issue, micro-PMUs are only installed at limited locations in real distribution networks, while SMs are widely installed in real distribution networks. Therefore, the access to available data for constructing an ambiguity set is limited by the number of micro-PMUs and SMs. If we only use limited data to construct an ambiguity set, it may hinder the performance of the DRCC program. The data from SMs and micro-PMUs are necessary but need modification to support our approaches. Therefore, we need to enrich the load and PV generation data, then construct the data-based ambiguity set for uncertainties of load and PV generation.

To address the above challenges, this paper proposes a tractable data-based DRCC-CVR model under the uncertainties of voltage-dependent load and PV generation. Inspired by our previous work [18], we apply the data enrichment method to both load and PV generation data from SMs and

micro-PMU, then construct the ambiguity set of uncertainties with the enriched data for the proposed DRCC-CVR. This data enrichment method enables a strong connection between the moment-based representation of the limited data set and the moment-based ambiguity set of the DRCC model, which avoids over- or under-conservativeness of the decisions. The main contributions of this paper are three-fold:

- *DRCC-CVR model with uncertainties of load and solar PV generation*: To consider the impacts of load and renewable uncertainties on voltage regulation and energy-saving performance of CVR in the unbalanced three-phase distribution systems, we present a deterministic VVO-CVR model and extend it by introducing chance constraints for possible voltage violations due to the load and the renewable uncertainties. To obtain proper conservative solutions robust to the high-resolution dataset and ensure a better performance for energy-saving and voltage regulation of CVR, we further integrate the chance constraints with DRO techniques to propose a DRCC-CVR model.
- *Linearization and tractable reformulation of DRCC-CVR model*: To make our proposed DRCC-CVR model tractable, we present the linearized version of the power flow model and voltage-dependent ZIP load model. Then, we reformulate the chance constraints with a moment-based ambiguity set of load and PV generation uncertainties in a tractable way.
- *Data-enriching moment-based ambiguity set with SM and micro-PMU data*: To guarantee the performance of the data-based DRCC-CVR model, we leverage the data enrichment method to recover high-resolution load and PV generation data from SMs and micro-PMUs. The moment information of load and PV generation uncertainties are extracted from the enriched data to construct the moment-based ambiguity sets for the proposed DRCC-CVR model.

## II. THE PROBLEM FORMULATION AND THE PROPOSED METHOD

### A. Solving a VVO-CVR Problem in the Unbalanced Three-phase Distribution Networks

In this paper, we consider an unbalanced three-phase radial distribution network that consists of  $N$  buses denoted by a set  $\mathcal{N}$  and  $N - 1$  branches denoted by a set  $\mathcal{E}$ . Let  $bp(i)$  denote the bus that immediately precedes bus  $i$  along the radial network headed by the feeder head bus. The three-phase indices  $\phi_a, \phi_b, \phi_c$  are simplified as  $\phi$ . The time instance is represented by  $t$ . Distributed assets are located at different buses including voltage dependent ZIP loads and solar PV distributed generators. We assume that the customers are either equipped with SMs or micro-PMUs, which monitor the active and reactive load power and active PV generation power with proper time resolution. For each bus  $i \in \mathcal{N}$ , let  $p_{i,\phi,t}^{\text{ZIP}}, q_{i,\phi,t}^{\text{ZIP}} \in \mathbb{R}^{3 \times 1}$  denote the vector of three-phase active and reactive voltage-dependent ZIP loads at time  $t$ . For each bus  $i \in \mathcal{G}$ , let  $p_{i,\phi,t}^g, q_{i,\phi,t}^g \in \mathbb{R}^{3 \times 1}$  denote the vector of three-phase active and reactive power outputs of the  $i$ -th PV inverter at time  $t$ ;  $V_{i,\phi,t} \in \mathbb{R}^{3 \times 1}$  repre-

<sup>1</sup>Micro-PMUs are synchrophasor devices that high-speed record real-time stamped data measurement of power and energy consumption. Micro-PMUs have a high sampling rate, e.g., one sample per second or higher [18]

<sup>2</sup>SMs are electronic devices that record power and energy consumption and can communicate remotely with utility. SMs have a relatively low sampling rate compared to micro-PMUs, e.g., one sample per hour [18].

sents the vector of three-phase voltage magnitude at time  $t$ ,  $v_{i,\phi,t} := V_{i,\phi,t} \odot V_{i,\phi,t} \in \mathbb{R}^{3 \times 1}$  represents the vector of three-phase squared voltage magnitude at time  $t$ . For each branch  $(i, j) \in \mathcal{E}$ , let  $z_{ij} = r_{ij} + \mathbf{i}x_{ij} \in \mathbb{C}^{3 \times 3}$  denotes matrix of the three-phase impedance of line  $ij$ , where  $r_{ij}$  and  $x_{ij}$  are the matrices of the three-phase resistance and reactance, respectively. Let  $S_{ij,\phi,t} = P_{ij,\phi,t} + \mathbf{i}Q_{ij,\phi,t} \in \mathbb{C}^{3 \times 1}$  denotes the vector of three-phase apparent power, where  $P_{ij,\phi,t}$  and  $Q_{ij,\phi,t}$  are the vector of three-phase active and reactive power flow through line  $ij$  from bus  $i$  to bus  $j$  at time  $t$ . Let the line active and reactive power flows, nodal active and reactive power injections, and squared voltage magnitudes be denoted by the following column vectors:  $P := \{P_{bp(i),i,\phi,t}, \forall i, t, \phi\}$ ,  $Q = \{Q_{bp(i),i,\phi,t}, \forall i, t, \phi\}$ ,  $p = \{p_{i,\phi,t}, \forall i, t, \phi\}$ ,  $q = \{q_{i,\phi,t}, \forall i, t, \phi\}$ , and  $v = \{v_{i,\phi,t}, \forall i, t, \phi\}$ .  $\odot$  and  $\oslash$  denote the element-wise multiplication and division.

The classic VVO-CVR program can be formulated as a deterministic problem (1), which aims to reduce the total power consumption of the entire distribution network while maintaining a feasible voltage profile within the predefined bounds across the distribution network as follows:

$$\min_{P, Q, p, q, v} \sum_{t \in [t, t+T]} \sum_{j: 0 \rightarrow j} \sum_{\phi \in \{a, b, c\}} \text{Re}\{S_{0j,\phi,t}\}, \quad (1a)$$

s.t.

$$P_{ij,\phi,t} = \sum_{k: j \rightarrow k} P_{jk,\phi,t} - p_{j,\phi,t}^g + p_{j,\phi,t}^{\text{ZIP}} + \varepsilon_{ij,\phi,t}^p, \quad (1b)$$

$$Q_{ij,\phi,t} = \sum_{k: j \rightarrow k} Q_{jk,\phi,t} - q_{j,\phi,t}^g + q_{j,\phi,t}^{\text{ZIP}} + \varepsilon_{ij,\phi,t}^q, \quad (1c)$$

$$v_{j,\phi,t} = v_{i,\phi,t} - 2(\bar{r}_{ij} \odot P_{ij,\phi,t} + \bar{x}_{ij} \odot Q_{ij,\phi,t}) + \varepsilon_{ij,\phi,t}^v, \quad (1d)$$

$$p_{i,\phi,t}^{\text{ZIP}} = p_{i,\phi,t}^L \odot (k_{i,1}^p \cdot v_{i,\phi,t} + k_{i,2}^p \cdot \sqrt{v_{i,\phi,t}} + k_{i,3}^p), \quad (1e)$$

$$q_{i,\phi,t}^{\text{ZIP}} = q_{i,\phi,t}^L \odot (k_{i,1}^q \cdot v_{i,\phi,t} + k_{i,2}^q \cdot \sqrt{v_{i,\phi,t}} + k_{i,3}^q), \quad (1f)$$

$$-Q_{i,\phi,t}^{\text{cap}} \leq q_{i,\phi,t}^g \leq Q_{i,\phi,t}^{\text{cap}}, \forall i \in \mathcal{G}, \quad (1g)$$

$$Q_{i,\phi,t}^{\text{cap}} = \sqrt{(S_{i,\phi,t}^{\text{cap}})^2 - (p_{i,\phi,t}^g)^2}, \forall i \in \mathcal{G}, \quad (1h)$$

$$v_{i,\phi,t}^{\min} \leq v_{i,\phi,t} \leq v_{i,\phi,t}^{\max}, \forall i \in \mathcal{N}. \quad (1i)$$

In objective function (1a), the three-phase active power supplied from the substation of the feeders  $\text{Re}\{S_{0j,\phi,t}\}$  is minimized over a moving finite horizon  $[t, t+T]$  for energy-saving with CVR implementation. Constraints (1b)-(1d) are defined by the unbalanced three-phase version of DistFlow model [3]. Constraints (1b) and (1c) guarantee the nodal active and reactive power balance. Constraint (1d) calculates the voltage difference between bus  $i$  and bus  $j$ . The detailed formulations of nonlinear terms  $\varepsilon_{ij,\phi,t}^p$ ,  $\varepsilon_{ij,\phi,t}^q$  and  $\varepsilon_{ij,\phi,t}^v$  can be found in [4]. If the network is not too severely unbalanced, the voltage magnitudes between the phases are similar and the relative phase unbalance  $\alpha$  is small [3], then we can use unbalanced three-phase resistance matrix  $\bar{r}_{ij}$  and reactance matrix  $\bar{x}_{ij}$  in constraint (1d). More details about  $\bar{r}_{ij}$  and  $\bar{x}_{ij}$  for unbalanced three-phase distribution systems can be referred to [3]. The implementation of CVR requires the modeling of voltage-dependent ZIP active and reactive loads  $p_{i,\phi,t}^{\text{ZIP}}$  and  $q_{i,\phi,t}^{\text{ZIP}}$  as

shown in (1e) and (1f).  $p_{i,\phi,t}^L, q_{i,\phi,t}^L \in \mathbb{R}^{3 \times 1}$  are the vectors of three-phase active and reactive load time-series multipliers on bus  $i$  at time  $t$ , respectively.  $k_{i,1}^p, k_{i,2}^p, k_{i,3}^p$  and  $k_{i,1}^q, k_{i,2}^q, k_{i,3}^q$  are constant-impedance (Z), constant-current (I) and constant-power (P) coefficients for active and reactive ZIP loads on bus  $i$ . Constraint (1g) limits the reactive power output  $q_{i,\phi,t}^g$  of the PV inverters by the available reactive power capacity  $Q_{i,\phi,t}^{\text{cap}}$ . Constraint (1h) calculates  $Q_{i,\phi,t}^{\text{cap}}$  with the total capacity of the PV inverter  $S_{i,\phi,t}^{\text{cap}}$  and the active power output of PV inverter  $p_{i,\phi,t}^g$ . Based on IEEE 1547-2018 Standard [19], the PV inverters can provide reactive power injection or absorption  $q_{i,\phi,t}^g$  to achieve fast voltage regulation. In this work, we focus on proposing a CVR model by optimally controlling the injection or absorption of reactive power in the PV inverters against the uncertainties of loads and renewable powers. While the dispatches of the on-load tap changers (OLTCs) and the capacitor banks (CBs) are slow and limited by a certain number of switching operation, which cannot response the uncertainties of loads and renewable powers. To consider the impacts of those conventional voltage regulation devices on CVR performance, a hierarchical control method [4] and [20] can be easily implemented to coordinate PV inverters, OLTC and CBs from different control stages. Note that the coordination between PV inverters, OLTC and CBs is out of the scope of this paper. In constraint (1i), the squared bus voltage magnitude  $v_{i,\phi,t}$  is limited by  $v_{\min}$  and  $v_{\max}$ , which are typically  $[0.95^2, 1.05^2]$  p.u., respectively.

The deterministic VVO-CVR problem (1) has an underlying assumption that the load and PV generation predictions are perfect, which means  $p_{i,\phi,t}^L, q_{i,\phi,t}^L$  in constraints (1e) and (1f), and  $p_{i,\phi,t}^g$  in constraint (1h) are predefined constant parameters. The more realistic setting is to take the load and PV generation prediction errors into account. To do this, we can replace the deterministic parameters  $p_{i,\phi,t}^L, q_{i,\phi,t}^L, p_{i,\phi,t}^g$  by uncertainty variables. Particularly, we introduce the regularized uncertainty variables  $p_{i,\phi,t}^{L,\xi} \in [0, 1]$ ,  $q_{i,\phi,t}^{L,\xi} \in [0, 1]$  to replace deterministic load multipliers  $p_{i,\phi,t}^L, q_{i,\phi,t}^L$ , as shown in (2a) and (2b). We reserve the super-script  $\xi$  to define the random variables, which also apply to the rest of the definition blow. Then, we introduce an auxiliary variables  $\alpha_{i,\phi,t}^q \in [-1, 1]$ , which represents the ratio between reactive power output  $q_{i,\phi,t}^g$  and reactive power capacity. Here, we use  $Q_{i,\phi,t}^{\text{cap},\xi}$  to represent the square root term  $\sqrt{(S_{i,\phi,t}^{\text{cap}})^2 - (p_{i,\phi,t}^{g,\xi})^2}$ , so that the constraint (1h) can be reformulated as constraint (2c).

$$p_{i,\phi,t}^{\text{ZIP}} = p_{i,\phi,t}^{L,\xi} \odot (k_{i,1}^p \cdot v_{i,\phi,t} + k_{i,2}^p \cdot \sqrt{v_{i,\phi,t}} + k_{i,3}^p), \quad (2a)$$

$$q_{i,\phi,t}^{\text{ZIP}} = q_{i,\phi,t}^{L,\xi} \odot (k_{i,1}^q \cdot v_{i,\phi,t} + k_{i,2}^q \cdot \sqrt{v_{i,\phi,t}} + k_{i,3}^q), \quad (2b)$$

$$q_{i,\phi,t}^g = \alpha_{i,\phi,t}^q Q_{i,\phi,t}^{\text{cap},\xi}, \forall i \in \mathcal{G}. \quad (2c)$$

We can define a uncertainty variable vector  $\xi_{i,\phi,t} = [(p_{i,\phi,t}^{L,\xi})^\top, (q_{i,\phi,t}^{L,\xi})^\top, (p_{i,\phi,t}^{g,\xi})^\top, (Q_{i,\phi,t}^{\text{cap},\xi})^\top]^\top$  to include all the uncertainty variables. To be simplified, we avoid the indices of  $i, \phi, t$  in vector  $\xi$ . To consider the impacts of uncertainty  $\xi$  on voltage regulation performance, we can extend the deterministic maximum/minimum voltage constraint (1i) to two chance constraints (3a) and (3b) as follows:

$$\mathbb{P}\{v - v_{\max} \leq 0\} \geq 1 - \epsilon, \quad (3a)$$

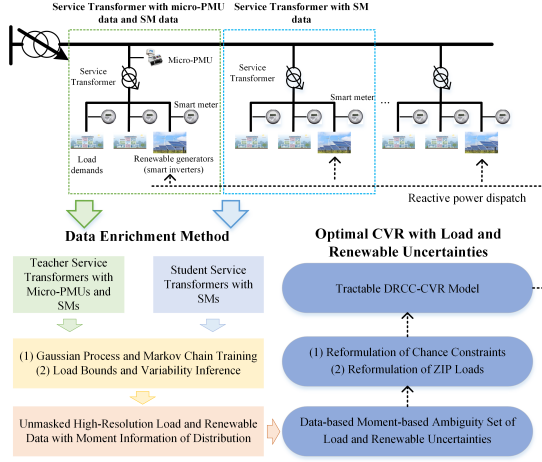


Fig. 1. Overall framework of the tractable DRCC-CVR with data enrichment method and enriched data-based moment-based ambiguity set.

$$\mathbb{P}\{-v + v_{\min} \leq 0\} \geq 1 - \epsilon. \quad (3b)$$

where  $\epsilon$  is a pre-defined risk level of failing to satisfy bus voltage constraint against uncertainties in  $\xi$ . To further make the solution robust to a group of probability distributions with controllable conservativeness, we introduce DRO and an ambiguity set of uncertainty to chance-constrained CVR and compactly formulate a DRCC-CVR problem (4a)-(4c) as follows:

$$\min_x \max_{\xi \sim \mathbb{P} \in \mathcal{P}} E_{\mathbb{P}}\{f(x, \xi)\}, \quad (4a)$$

$$\text{s.t. } g_1(x) \leq 0, \quad (4b)$$

$$\mathbb{P}\{g_2(x, \xi) \leq 0\} \geq 1 - \epsilon, \quad (4c)$$

where  $x$  represents the decision variable vector (i.e. reactive power dispatch of PV inverter) and  $\xi \sim \mathbb{P} \in \mathcal{P}$  in objective (4a) means that the uncertainty variable vector  $\xi$  following the distribution  $\mathbb{P}$  within an ambiguity set of distributions  $\mathcal{P}$ . Constraints (1b)-(1h) can be represented by the compact constraint (4b) and the chance constraints (3a) and (3b) can be represented by the compact constraint (4c).

### B. Our Proposed Method

The challenges of solving this DRCC-CVR problem (4a)-(4c) can be summarized as follows: (i) the nonlinear power flow model and voltage-dependent ZIP load model, and the chance constraints with the random variables make the DRCC-CVR problem (4a)-(4c) intractable to be solved; and (ii) even though we can reformulate the DRCC problem in a tractable way, the limited access to the high-resolution load and PV data will lead to an ill-posed DRCC-CVR problem and hinder the performance of CVR implementation. To address those challenges, we propose a solution to address these challenges, as shown in Fig. 1, which includes the following two-fold:

1) *Tractable DRCC-CVR model*: To reformulate a tractable DRCC-CVR model, we leverage the linearized Distflow model [3] and linearized voltage-dependent ZIP load model with Binominal Approximation method [15]. Then, we reformulate the chance constraints of voltage by a tractable DRCC model with a moment-based ambiguity set. Compared to other types

of DRO, the DRCC with a moment-based ambiguity set has higher computational efficiency for tractable reformulation. We will show the linearized version of the power flow model and voltage-dependent ZIP load model, as well as the tractable reformulation of the DRCC-CVR problem in Section III.

2) *Data enrichment method and moment-based Ambiguity set*: As shown in Fig. 1, there are a large number of service transformers that can collect low-resolution load and PV generation data by SMs, while only a few service transformers are installed with micro-PMUs with access to high-resolution load and PV generation data. To capture the uncertainty of load and PV generation, we use SMs and micro-PMUs to collect load and PV generation data. Then we enrich the load and PV generation data and extract the corresponding moment information of probability distributions from the enriched load and PV data. Finally, we can construct the ambiguity set with the first two moment information i.e., mean and variance, and implement the ambiguity set in our proposed DRCC-CVR model. The purpose of introducing the data enrichment method in a data-based ambiguity set is to avoid potential over- or under-conservativeness. We will show the data-enrichment method and the construction of a moment-based ambiguity set for DRCC-CVR in Section IV.

## III. TRACTABLE REFORMULATION OF DRCC-CVR MODEL

This section presents the linearized version of the power flow model and voltage-dependent ZIP load model, and reformulates a tractable version of the DRCC-CVR problem.

### A. Linearized Reformulation of Power Flow and Voltage-Dependent ZIP Loads

In power flow constraints (1b)-(1d), the nonlinear terms  $\varepsilon_{ij,\phi,t}^p$ ,  $\varepsilon_{ij,\phi,t}^q$  and  $\varepsilon_{i,\phi,t}^v$  make the optimization problems non-convex and NP hard. In practice, those nonlinear terms are much smaller than the linear terms in power flow constraints (1b)-(1d). Therefore, the constraints (1b)-(1d) can be reformulated as constraints (5a)-(5c) with linearized Distflow model by neglecting those nonlinear terms.

$$P_{ij,\phi,t} = \sum_{k:j \rightarrow k} P_{jk,\phi,t} - p_{j,\phi,t}^g + p_{j,\phi,t}^{\text{ZIP}}, \quad (5a)$$

$$Q_{ij,\phi,t} = \sum_{k:j \rightarrow k} Q_{jk,\phi,t} - q_{j,\phi,t}^g + q_{j,\phi,t}^{\text{ZIP}}, \quad (5b)$$

$$v_{j,\phi,t} = v_{i,\phi,t} - 2(\bar{r}_{ij} \odot P_{ij,\phi,t} + \bar{x}_{ij} \odot Q_{ij,\phi,t}). \quad (5c)$$

This linear form of DistFlow has been verified in many previous studies, such as [3]. The nonlinear term  $\sqrt{v_{i,\phi,t}}$  of ZIP loads also introduces non-convexity to the problem. Because the voltage magnitudes of all buses in a distribution network stay close to 1 p.u. under normal operating conditions [3] and [4], the active and reactive ZIP loads can be linearized by Binominal Approximation Method [15]. Therefore, the squared deviation of voltage  $(\Delta V)^2$  is very small, so it can be neglected. Then we have the following approximations (6) of squared voltage magnitude, as follows:

$$v = V \odot V = (1 + \Delta V) \odot (1 + \Delta V) \approx 1 + 2\Delta V,$$

$$\begin{aligned} v &= 1 + \Delta v \approx 1 + 2\Delta V, \\ \Delta v &\approx 2\Delta V, \end{aligned} \quad (6)$$

where  $v$  and  $\Delta v$  are the vectors of squared voltage magnitude and the derivation from the nominal value, respectively;  $V$  and  $\Delta V$  are the vectors of voltage magnitude and the derivation from the nominal value, respectively. By introducing equation (6) and  $\sqrt{v} = V = (1 + \Delta V)$  to equations (2a) and (2b), we have the linear approximation of voltage-dependent active and reactive ZIP loads as follows:

$$p_{i,\phi,t}^{\text{ZIP}} \approx p_{i,\phi,t}^{L,\xi} \odot \left( (k_{i,1}^p + \frac{k_{i,2}^p}{2})v_{i,\phi,t} + (k_{i,3}^p + \frac{k_{i,2}^p}{2}) \right), \quad (7a)$$

$$q_{i,\phi,t}^{\text{ZIP}} \approx q_{i,\phi,t}^{L,\xi} \odot \left( (k_{i,1}^q + \frac{k_{i,2}^q}{2})v_{i,\phi,t} + (k_{i,3}^q + \frac{k_{i,2}^q}{2}) \right). \quad (7b)$$

### B. Tractable Reformulation of DRCC-CVR with Load and Renewable Uncertainties

To achieve the tractable reformulation of DRCC-CVR, the power flow constraints (1b)-(1d) can be compactly formulated as follows:

$$-AP = p, \quad (8a)$$

$$-AQ = q, \quad (8b)$$

$$-A_0 v_0 - A^\top v = -2D_r P - 2D_x Q, \quad (8c)$$

where  $A_0$  and  $A$  are the incidence matrices of unbalanced radial distribution network,  $A_0$  represents the connection structure between substation (the feeder head bus) and each of the line segments in  $\mathcal{E}$ ,  $A$  represents the connection structure between the remaining buses and each of the line segment in  $\mathcal{E}$ .  $v_0$  is vector of square nominal voltage magnitudes.  $D_r = \text{blkdiag}[R_{bp(1)1}, \dots, R_{bp(N)N}]$  and  $D_x = \text{blkdiag}[X_{bp(1)1}, \dots, X_{bp(N)N}]$  are block diagonal matrices of line segment resistance and reactance, respectively. In equations (8a) and (8b), the nodal active and reactive power injections can be calculated based on ZIP loads and PV generations. Based on the compact power flow formulations (8a)-(8c), we have the compact formulation (9) to represent the relationship between bus voltage  $v$  and bus power injections  $p$  and  $q$ , as follows:

$$v = Rp + Xq + \tilde{v}, \quad (9)$$

with

$$\begin{aligned} R &= 2[A^\top]^{-1}D_r A^{-1}, \\ X &= 2[A^\top]^{-1}D_x A^{-1}, \\ \tilde{v} &= -[A^\top]^{-1}A_0 v_0. \end{aligned}$$

By introducing PV inverter reactive power output equation (2c), linearized ZIP load equations (7a) and (7b) as bus power injections  $p$  and  $q$  into the compact formulation (9), we have an equation as shown in equation (10), which represents the relationship between vector of squared voltage  $v$  and all the uncertainty variables in vector  $\xi$ . If we select an appropriate value as a power base and use per unit to represent  $p^{L,\xi}$  and  $q^{L,\xi}$ , because the load level of distribution systems is usually not heavy at kW level, then we have  $p^{L,\xi}$  and  $q^{L,\xi}$  are much smaller than identity matrix  $I$ , the network resistance

and reactance matrices  $R, X$  are also small in per unit, thus the equation (11) is valid.

$$I - p^{L,\xi} \left( Rk_1^p + R\frac{k_2^p}{2} \right) - q^{L,\xi} \left( Xk_1^q + X\frac{k_2^q}{2} \right) > 0. \quad (11)$$

After we introduce the equation (10) into the deterministic constraint (1g) on bus voltages, and because equation (11) is valid, we can obtain equation (12), which represents the impacts of load and PV generation uncertainties on bus voltage constraints (1i) with  $v_{\min}$  and  $v_{\max}$ . For simplicity, we also avoid the indices of  $i, \phi, t$  in equations (10) and (12). The compact formulation (4c) can be reformulated in a linear form  $a(x)^\top \xi + b(x) \leq 0$ . We can obtain the formulations of  $a(x)$  and  $b(x)$  for chance constraints (3a) and (3b) by introducing the equation (12) into chance constraints (3a) and (3b). Therefore, the  $a(x)$  and  $b(x)$  of chance constraint (3a) can be formulated as (13a) and (13b), respectively.

$$a(x) = \begin{bmatrix} \text{diag}(v_{\max}) \left( Rk_1^p + R\frac{k_2^p}{2} \right) + \left( Rk_3^p + R\frac{k_2^p}{2} \right) \\ \text{diag}(v_{\max}) \left( Xk_1^q + X\frac{k_2^q}{2} \right) + \left( Xk_3^q + X\frac{k_2^q}{2} \right) \\ -R \\ -Xa_j \end{bmatrix}, \quad (13a)$$

$$b(x) = \tilde{v} - v_{\max}. \quad (13b)$$

Similarly, the  $a(x)$  and  $b(x)$  for another chance constraint (3b) can be formulated as (14a) and (14b), respectively.

$$a(x) = \begin{bmatrix} -\text{diag}(v_{\min}) \left( Rk_1^p + R\frac{k_2^p}{2} \right) - \left( Rk_3^p + R\frac{k_2^p}{2} \right) \\ -\text{diag}(v_{\min}) \left( Xk_1^q + X\frac{k_2^q}{2} \right) - \left( Xk_3^q + X\frac{k_2^q}{2} \right) \\ R \\ Xa_j \end{bmatrix}, \quad (14a)$$

$$b(x) = -\tilde{v} + v_{\min}. \quad (14b)$$

A moment-based ambiguity set of load and PV generation uncertainties can be constructed in (15), as follows:

$$\mathcal{D}_\xi = \{ \xi \sim \mathbb{P} \in \mathcal{P} : \mathbb{E}_{\mathbb{P}_\xi}[\xi] = \mu, \mathbb{E}_{\mathbb{P}_\xi}[\xi\xi^\top] = \Sigma \} \quad (15)$$

where  $\mu$  and  $\Sigma$  represent the mean and covariance of the uncertain variables of load and PV generation. Finally, based on equations (13a)-(14b) for  $a(x)$  and  $b(x)$ , we can obtain a second-order conic reformulation (16) for the DRCC (3a) and (3b) with moment information, mean  $\mu$  and covariance  $\Sigma$  of uncertainty variable vector  $\xi$ , as follows: [21] and [22]:

$$a(x)^\top \mu + b(x) + \sqrt{\frac{1-\epsilon}{\epsilon}} \|\Sigma^{\frac{1}{2}} a(x)\|_2 \leq 0. \quad (16)$$

Even though the mean  $\mu$  and covariance  $\Sigma$  of load and PV generation in (16) can be extracted from recorded data of SMs and micro-PMUs, the reality is that we only have limited access to high-resolution load and PV generation data. The limited data will lead to the potential ill-posed of DRCC (16), and further hinder the performance of DRCC-CVR. Therefore, we introduce a data enrichment method for high-resolution data recovery and ambiguity set construction in Section IV.

$$v = R \left( p^{L,\xi} \left( (k_1^p + \frac{k_3^p}{2})v + (k_3^p + \frac{k_2^p}{2}) \right) - p^{g,\xi} \right) + X \left( q^{L,\xi} \left( (k_1^q + \frac{k_3^q}{2})v + (k_3^q + \frac{k_2^q}{2}) \right) - \alpha^q Q^{\text{cap},\xi} \right) + \tilde{v}, \quad (10)$$

$$v_{\min} \leq \frac{p^{L,\xi} \left( Rk_3^p + R\frac{k_2^p}{2} \right) + q^{L,\xi} \left( Xk_3^q + X\frac{k_2^q}{2} \right) + Rp^{g,\xi} + X\alpha^q Q^{\text{cap},\xi} + \tilde{v}}{I - p^{L,\xi} \left( Rk_1^p + R\frac{k_2^p}{2} \right) - q^{L,\xi} \left( Xk_1^q + X\frac{k_2^q}{2} \right)} \leq v_{\max}, \quad (12)$$

#### IV. DATA ENRICHMENT METHOD AND MOMENT-BASED AMBIGUITY SET

This section presents the data enrichment method to recover high-resolution data of load and PV generation for those service transformers with only SMs. Then, the ambiguity set with moment information of probability distributions is constructed based on the enriched load and PV generation data.

##### A. Data Enrichment Method with Micro-PMU and SM Data

As shown in Fig. 1, the majority of service transformers are installed with SMs to record low-resolution load and PV generation data, and only a few service transformers are also installed with micro-PMU to record high-resolution data. To enrich the data of load and PV generation, we consider the service transformers with micro-PMU data as a teacher repository and the service transformers with only SM data as a student repository. The teacher service transformers train two models capturing the statistical relationship between high-resolution data and low-resolution data. The trained models are utilized to perform data enrichment for those service transformers with only SMs. There are four main steps:

**Step. I** *Train the load/PV generation data maximum and minimum boundary inference models:* The first step is to train probabilistic models using known high-resolution load/PV generation data from teacher service transformers with micro-PMUs. The Gaussian Process Regression (GPR) technique [18] is used to capture the relationship between the maximum/minimum bounds and the average values for load/PV generation. More specifically, two GPR models  $GPR_{s,1}^*$  and  $GPR_{s,2}^*$  are trained for the  $s$ -th teacher service transformer:

$$GPR_{s,1}^* : P_a(t) \rightarrow \bar{P}(t), \quad (17a)$$

$$GPR_{s,2}^* : P_a(t) \rightarrow \underline{P}(t), \quad (17b)$$

where  $P_a(t)$  denotes the average load/PV generation over the  $t$ -th hour,  $\bar{P}(t)$  and  $\underline{P}(t)$  denote the upper and lower bounds of instantaneous load/PV generation within the  $t$ -th hour, respectively.

**Step. II** *Train the load/PV generation variability inference models:* The second step is to model the probabilistic transition of instantaneous load/PV generation within their maximum and minimum bounds using the second-order Markov Chain (MC) model. Specifically, one MC model  $MC_s^*$  is trained for each service transformer:

$$MC_s^* : \{P_t(m-2), P_t(m-1)\} \rightarrow P_r(P_t(m)), \quad (18)$$

where  $P_t(m-2)$ ,  $P_t(m-1)$ , and  $P_t(m)$  denote the  $(m-2)$ -th,  $(m-1)$ -th, and  $m$ -th high-resolution load/PV generation samples within the  $t$ -th hour. Formulation (18) outputs the probability of  $P_t(m)$  based on  $P_t(m-2)$  and  $P_t(m-1)$ . Note that Step. I and II build the load/PV generation data boundary inference models and the load/PV generation data variability inference models with a small number of high-resolution micro-PMU data of teacher service transformers. Steps. III and IV will extend the trained load/PV generation probabilistic models to service transformers with only SMs, so it recovers the high-resolution load/PV generation data masked by the low-resolution load/PV generation data.

**Step. III** *Determine the learning weights of student service transformers with respect to teacher service transformers:* The third step evaluates the low-resolution data similarity between the teacher with micro-PMUs and student service transformers with SMs by determining the learning weights as shown in (19a) and in (19b). The weights  $W_s$  and  $W'_s$  can represent the confidence of a student service transformer to learn from multiple teacher service transformers for enriching the low-resolution load/PV generation data to high-resolution load/PV generation data.

$$W_s = \frac{W'_s}{\sum_{s=1}^{N_t} W'_s}, \quad (19a)$$

$$W'_s = \frac{1}{N_c N_c^s} \sum_{i=1}^{N_c} \sum_{j=1}^{N_c^s} \|P_i - P_j^s\|, s = \{1, \dots, N_t\}, \quad (19b)$$

where  $N_c$  and  $N_c^s$  denote the number of customers served by a student service transformer and the  $s$ -th teacher service transformer, respectively. We can obtain  $N_c$  daily load/PV generation patterns for that service transformer,  $\{P_1, \dots, P_{N_c}\}$ . Similarly, the load/PV generation patterns for the  $s$ -th teacher service transformer are denoted by  $\{P_1^s, \dots, P_{N_c^s}^s\}$ ,  $s = 1, \dots, N_t$ . The weights are used to linearly combining the estimated bounds in (17) and the probabilistic transition matrices in (18).

**Step. IV** *Extend the trained load/PV generation data probabilistic model:* The fourth step extends the trained probabilistic models of teacher service transformers in Steps. I and II to student service transformers that only have SMs for enriching low-resolution load/PV generation data. Specifically, the  $m$ -th high-resolution load/PV generation sample is randomly generated based on the following Bernoulli distribution:

$$P_t(m) \sim Be(Pr(P_t)), \quad (20)$$

where  $Be$  denotes the Bernoulli distribution [18]. Note that for each hourly load/PV generation sample, (20) can give us  $N'$  high-resolution load/PV generation samples, i.e.,  $P_t(m), m = 1, \dots, N'$ . Then, the enriched high-resolution data samples are employed to optimize the mean and standard deviation of Gaussian distribution using maximum likelihood estimation:

$$(\mu^*, \Sigma^*) = \underset{\mu, \Sigma}{\operatorname{argmin}} \prod_{m=1}^{N'} f(\mu, \Sigma; P_t(m)), \quad (21)$$

where,  $f(\cdot)$  denotes the probability density function of Gaussian distribution. Therefore, the mean and covariance of load and PV generation uncertainties  $\mu$  and  $\Sigma$  can be extracted from enriched data in (21) for the construction of the moment-based ambiguity set (15). i.e., substitute  $\mu^*$  and  $\Sigma^*$  in (21) into (16).

### B. Enriched Data-Based Ambiguity Set and DRCC-CVR Method

The enriched high-resolution data from (20) can recover instantaneous uncertainties of load and PV generation, which can be further extracted for the two moment information and construct moment-based ambiguity set in (15). By considering the data-based moment-based ambiguity set and the reformulation of tractable DRO and chance constraints, the DRCC-CVR problem can be compactly reformulated as follows:

$$\text{Objective (1),} \quad (22a)$$

$$\text{s.t. PV generations } \{(2c)\}, \quad (22b)$$

$$\text{OPF constraints } \{(5)\}, \quad (22c)$$

$$\text{Linearized ZIP loads } \{(7)\}, \quad (22d)$$

$$\text{Tractable reformulation of DRCC } \{(13) - (14), (16)\}. \quad (22e)$$

Note that problem (22a)-(22e) is a tractable linear programming problem, where the mean values of load and PV are used in compact constraints (22b)-(22d) and the mean and covariance values of load and PV are used in the compact constraint (22e) of DRCC reformulation. To summarize the above steps, the detailed procedure of the proposed DRCC-CVR with an enriched-based ambiguity set of the uncertainties of load and PV generation is shown in Algorithm 1.

## V. CASE STUDIES

This section presents the simulation results, including the enriched data of load and PV generation for ambiguity set, comparison results of the benchmark methods and the proposed DRCC-CVR method, and the impacts of hyper-parameter on the proposed DRCC-CVR method.

### A. Simulation Setup

A real-world distribution feeder [23] in Fig. 2 is used to test our proposed DRCC-CVR method, which is located in Midwest U.S. and shared by our utility partner. The reason for choosing this real-world distribution feeder as our test system is that the service transformers of customers in this feeder are either equipped with SMs or micro-PMUs, which

### Algorithm 1 DRCC-CVR Model with Enriched-based Ambiguity Set of Uncertainty of Load and PV Generation

- 1: **Input:** High-resolution data from micro-PMUs and low-resolution data from SMs
- 2: **Initialization:** Choose hyper-parameters in DRCC-CVR
- 3: **For:**  $i = 1, 2, \dots, N$ .
- 4: Train load/PV generation upper and lower boundary inference models from teacher service transformers with micro-PMUs by (17a) and (17b).
- 5: Train load/PV generation variability inference model from teacher service transformers with micro-PMUs by (18).
- 6: Determine the learning weights of student service transformers with respect to teacher service transformers in (19a) and (19b).
- 7: Extend the trained load/PV generation data to student service transformers with SMs in (20).
- 8: **End for.**
- 9: Extract the first two moment information of load and PV generation uncertainties from enriched data in (21)
- 10: Construct an ambiguity set with the first two moment information in (15).
- 11: Solve the DRCC-CVR problem (22) with objective (1) and constraints (2c),(5),(7), (13)-(14), and (16).
- 12: **Output:** Reactive power dispatches of PV inverters

can record low-resolution (1-hour) data and high-resolution (1-second) data to construct an ambiguity set of load and PV generation uncertainties for our DRCC-CVR method. More information on this real-world distribution feeder and the data from SMs and micro-PMUs can be found in [23]. In Fig. 2, the yellow bot-ted boxes represent the buses' service transformers installed with micro-PMUs, and the rest buses' service transformers are installed with SMs; the blue dots represent the buses installed with single-phase or three-phase PV generators; the solid, dashed and dotted lines represent three-phase overhead lines, three-phase underground cables and single-phase overhead lines, respectively. In this real-world distribution feeder, the total capacity of PVs can serve around 20% to 30% load. Adopted from our industrial partner [4], we use the following coefficients  $[k_1^p, k_2^p, k_3^p] = [0.96, -1.17, 1.21]$  and  $[k_1^q, k_2^q, k_3^q] = [6.28, -10.16, 4.88]$  for active and reactive ZIP loads in our test cases. The base voltage and base power values are 13.8 kV and 100 kVA. The prescribed risk level parameter  $\epsilon$  is set to 0.05 for quantifying the 5% violation probability of chance constraints in our proposed DRCC-CVR model. We demonstrate the advantages and effectiveness of the enriched data-based ambiguity set and the proposed DRCC-CVR method through numerical comparisons of several benchmark methods. The following simulations are built-in MATLAB R2019b, which integrate YALMIP Toolbox with IBM ILOG CPLEX 12.9 solver for optimization. All case studies are simulated on a PC with Intel Core i7-4790 3.6 GHz CPU and 16 GB RAM.

### B. Original and Enriched Data of Load and PV Generation

As shown in Fig. 2, there are 8 service transformers installed with micro-PMUs and the rest 34 service transform-



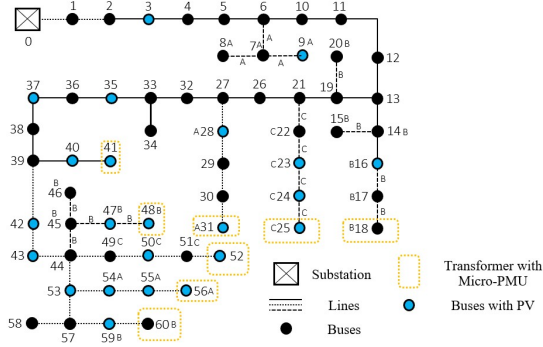


Fig. 2. A real distribution feeder in Midwest U.S. [23]

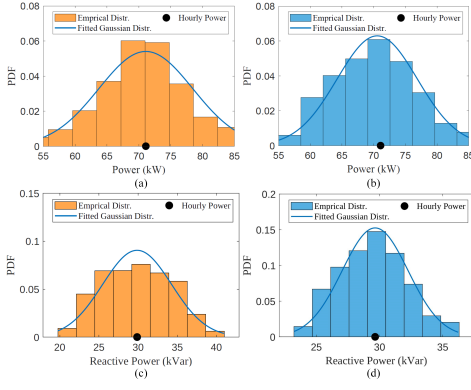


Fig. 3. Empirical distribution and fitted Gaussian distribution: (a) active load (original data); (b) active load (enriched data); (c) reactive load (original data); (d) reactive load (enriched data).

ers installed with SMs, which can record high-resolution (1-second) and low-resolution (1-hour) load and PV generation data, respectively. In this sub-section, we use two ways to obtain the mean and covariance of the uncertainty variables of load and PV generation for ambiguity sets: (i) we use a statistical method to obtain mean and variance information of the original load and PV generation data (few high-resolution data from 8 micro-PMUs and a lot of low-resolution data from 34 SMs). (ii) We use the data enrichment method to enrich the original load and PV generation data, then we obtain mean and variance information of the enriched data. To verify the effectiveness of the data enrichment method, we show the empirical distributions and their according fitted Gaussian distributions of active load, reactive load, and active power output of PV generation in Fig. 3 and Fig. 4, respectively. Note that the distributions in the right column and left column of Fig. 3 and Fig. 4 are obtained from original data and enriched data, respectively. In Fig. 3 and Fig. 4, each dot denotes the hourly average power corresponding to the data samples, which means the hourly power is the average of observations. It can be observed that the distributions obtained from enriched data have a highly similar fitting to the distributions obtained from original data. Therefore, the first two moment information extracted from enriched data can also accurately construct the ambiguity set for the DRCC-CVR method.

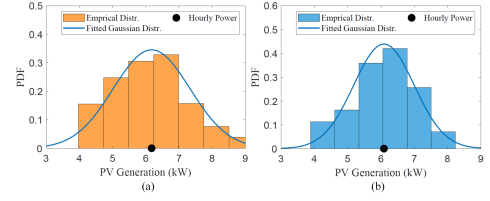


Fig. 4. Empirical distribution and fitted Gaussian distribution: (a) PV generation (original data); (b) PV generation (enriched data).

### C. Voltage Reduction and Power/energy-saving Through CVR Implementation

To serve as a reference to investigate the performance of VVO-CVR, a base case is firstly built by setting the unity-power factor control model for all PV generators, which means there is no reactive power support from PV generators. In this sub-section, we use three ways to implement VVO-CVR: (i) The deterministic VVO-CVR (Deter-CVR) is solved by deterministic optimization, where the load and PV prediction errors and uncertainties are neglected. (ii) The robust VVO-CVR (RO-CVR) is solved by the robust optimization [7] and [8], where the uncertainties of load and PV generation are considered with 10% variance from the predictions. (iii) The VVO-CVR is solved by our proposed DRCC-CVR with an enriched data-based ambiguity set of load and PV generation uncertainties. The performance through CVR implementation can be evaluated from three aspects: voltage profile, active power supply from the substation, and total energy consumption. To show the time-series simulation, the VVO-CVR is performed in a daily operation of the real-world distribution feeder with different control strategies. In Fig. 5, the voltage profiles for a selected bus (bus 23 on phase c) are shown, which are generated from the base case (without control) and the proposed DRCC-CVR. In Fig. 5, the blue bar and red bar represent the voltage profiles of the base case and DRCC-CVR, respectively. It can be observed that all the nodal voltages can maintain within the predefined range  $[0.95, 1.05]$  p.u., while the voltage profiles of DRCC-CVR are overall lower than the voltage profiles of the base case. Because the DRCC-CVR can optimally dispatch the reactive power from PV inverters to achieve maximum voltage reduction while still satisfying the voltage constraints. During the midnight period from 00:00 to 6:00, the voltage reduction of DRCC-CVR is obviously higher than the voltage reduction of the base case, as shown in the circled part in Fig. 5. It is because the active power outputs of PV generators from 00:00 to 6:00 are nearly zero, and according to the calculations of reactive power output and capacity in constraints (1g) and equation (1h), the DRCC-CVR has more reactive power supports to maximize voltage reductions.

The active power supplies from the substation of the base case, Deter-CVR, RO-CVR, and DRCC-CVR are shown as different curves in Fig. 6, which represent the overall active power consumption of the base case, and those VVO-CVR benchmarks. It can be observed that the proposed DRCC-CVR can effectively reduce the power supply from the substation compared to the base case and other methods. In comparison the power saving of RO-CVR is less than the pro-



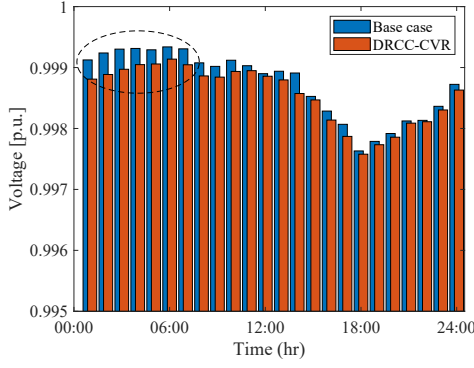


Fig. 5. Voltage profiles on selected bus 23 on phase c with different control strategies.

posed DDRC-CVR and only slightly better than Deter-CVR. On the one hand, the proposed DDRC-CVR has an ambiguity set of uncertainties to balance the trade-off between the conservativeness of decisions and operational efficiency. On the other hand, RO-CVR is too conservative and hinders the performance of CVR implementation. The numerical comparisons of total energy consumption over one day and the energy reduction percentage are presented in Table I among base case, Deter-CVR, RO-CVR, and DDRC-CVR. The total energy consumption of base case, Deter-CVR, RO-CVR, and DRCC-CVR are 958.045 kWh, 944.048 kWh, 934.178 kWh, and 898.616 kWh, respectively. Therefore, compared to the original energy consumption in the base case, the Deter-CVR, RO-CVR, and DRCC-CVR can achieve 1.461%, 2.491%, and 6.203% of energy savings, respectively. The computational times of Deter-CVR, RO-CVR, and DRCC-CVR for this test system are 12.968 seconds, 18.312 seconds, and 21.911 seconds, respectively. According to the above results in Fig. 5, Fig. 6 and Table I, we can summarize the differences between Deter-CVR, RO-CVR, and DRCC-CVR: (i) Among all the methods, DRCC-CVR can achieve the lowest total energy consumption and highest energy-saving. While the differences between the total energy consumption and energy-saving results of Deter-CVR and RO-CVR are small, because Deter-CVR does not consider the errors of load and PV generation predictions, and RO-CVR provides a conservative solution. Both Deter-CVR and RO-CVR cannot fully explore the benefit of CVR implementation, while the proposed DDRC-CVR can better explore the potential of CVR implementation. (ii) Because of the different ways of handling uncertainties of load and PV generation, the computation time of DRCC-CVR is slightly slower than Deter-CVR and RO-CVR. However, their differences in computational time are very small, which can be neglected for a day-ahead operational application. Thus the computational efficiency of the DRCC-CVR can be acceptable.

#### D. Impact of Hyper-Parameters in Performance of DRCC-CVR

It is obvious that more micro-PMUs installed in service transformers can record more high-resolution data of load and PV generation. Also, more high-resolution data can be benefi-

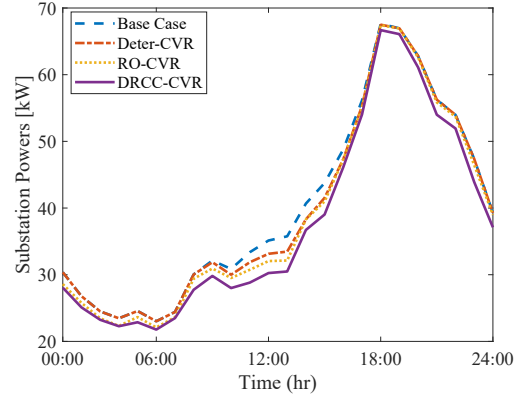


Fig. 6. Substation feed-in active power with different control strategies.

TABLE I  
ENERGY CONSUMPTION AND ENERGY-SAVING RESULTS WITH DIFFERENT CONTROL STRATEGIES

|           | Energy (kWh) | Reduction (%) | Computation (sec) |
|-----------|--------------|---------------|-------------------|
| Base Case | 958.045      | -             | -                 |
| Deter-CVR | 944.048      | 1.461%        | 12.968            |
| RO-CVR    | 934.178      | 2.491%        | 18.312            |
| DRCC-CVR  | 898.616      | 6.203%        | 21.911            |

cial for recording the uncertainties of load and PV generation, which is helpful for the data enrichment method and construction of ambiguity set in DDRC-CVR. To show the impact of micro-PMUs on the proposed DRCC-CVR, the total energy consumption and energy-saving of DRCC-CVR with the different numbers of micro-PMUs are presented in Table II. There are three different tests: the first test only constructs the ambiguity set based on low-resolution data from SMs without any high-resolution data from micro-PMU; the second test implements the data enrichment method with high-resolution data from 4 micro-PMUs and low-resolution data from SMs, then the ambiguity set is constructed based on the enriched data; the third test has the high-resolution data from 8 micro-PMUs as the same setting in Section V-B and Section V-C. It can be observed that a higher number of micro-PMUs can help to construct a high-quality ambiguity set for DRCC-CVR and achieve better performance of CVR implementation through DRCC-CVR.

The pre-defined risk level  $\epsilon$  is selected as 0.05 in the above simulation tests. While the different values of  $\epsilon$  will influence the confidence level on chance constraints in the proposed

TABLE II  
ENERGY CONSUMPTION AND ENERGY-SAVING RESULTS WITH DIFFERENT NUMBER OF MICRO-PMUS.

| Number of micro-PMUs | Location of micro-PMUs                                           | Energy (kWh) | Saving (%) |
|----------------------|------------------------------------------------------------------|--------------|------------|
| 0 (only SM data)     | -                                                                | 933.889      | 2.521%     |
| 4                    | $B_{18}, B_{25}, B_{41}, B_{60}$                                 | 927.817      | 3.155%     |
| 8                    | $B_{18}, B_{25}, B_{31}, B_{41}, B_{48}, B_{52}, B_{56}, B_{60}$ | 898.618      | 6.203%     |

TABLE III  
ENERGY CONSUMPTION AND ENERGY-SAVING RESULTS WITH  
DIFFERENT VIOLATION RATES

|                   | Energy (kWh) | Reduction (%) |
|-------------------|--------------|---------------|
| $\epsilon = 0.02$ | 904.980      | 4.803%        |
| $\epsilon = 0.05$ | 898.616      | 6.203%        |
| $\epsilon = 0.1$  | 894.754      | 6.606%        |

DRCC-CVR and further influence the benefits of CVR implementation. Therefore, we also test the DRCC-CVR with three different violation rates (i.e. 0.02, 0.05, 0.1), as shown in Table III. It can be observed that a larger value of  $\epsilon$  leads to lower energy consumption and a higher energy reduction, which will benefit the CVR implementation. However, this kind of benefit is achieved by increasing the violation rate and sacrificing the reliability of operational constraints. For example, the energy reduction (6.606%) of  $\epsilon = 0.1$  is only slightly higher than the energy reduction (6.203%) with  $\epsilon = 0.05$ , but the risk of constraint violation also become higher when  $\epsilon$  increases from 0.05 to 0.1. The results indicate that there is a trade-off between the maximization of CVR benefit and the reliability of operational constraints in the proposed DRCC-CVR.

## VI. CONCLUSION

To better consider the impacts of load and PV generation uncertainties on voltage regulating performance while implementing CVR in unbalanced three-phase distribution systems, a DRCC-CVR model with an enriched data-based and moment-based ambiguity set is proposed to optimally dispatch the reactive power of PV inverters. The original and intractable DRCC-CVR is approximated by linearized ZIP load models and reformulated in a tractable way with the first two moment information of load and PV generation uncertainties. We further implement a data enrichment method with low-resolution data from SMs and high-resolution data from micro-PMUs to recover instantaneous uncertainties of load and PV generation. An ambiguity set is constructed based on enriched data for DRCC-CVR. Simulation results on a real Midwest U.S. distribution feeder have validated the effectiveness and robustness of the proposed DRCC-CVR. According to the case studies, we have shown that: (i) The data enrichment can construct an accurate ambiguity set of load and PV generation uncertainties. (ii) Compared to other benchmark methods, the proposed DRCC-CVR shows a better trade-off between the conservativeness of decisions and operational efficiency. Thus, the proposed DRCC-CVR can achieve better CVR performance on voltage reduction and power/energy-saving. (iii) With a reasonable number of micro-PMUs, a high-quality ambiguity set can be constructed for DRCC-CVR and better performance of CVR can be achieved. (iv) To consider the benefits of CVR implementation and the reliability of operational constraints, a proper risk and confidence level needs to be tuned in DRCC-CVR.

## REFERENCES

- [1] Z. Wang and J. Wang, "Review on implementation and assessment of conservation voltage reduction," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1306–1315, 2014.
- [2] T. V. Dao, S. Chaitusaney, and H. T. N. Nguyen, "Linear least-squares method for conservation voltage reduction in distribution systems with photovoltaic inverters," *IEEE Trans. Smart Grid*, vol. 8, no. 3, pp. 1252–1263, 2017.
- [3] Q. Zhang, K. Dehghanpour, and Z. Wang, "Distributed CVR in unbalanced distribution systems with PV penetration," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5308–5319, Sept. 2019.
- [4] Q. Zhang, Y. Guo, Z. Wang, and F. Bu, "Distributed optimal conservation voltage reduction in integrated primary-secondary distribution systems," *IEEE Trans. Smart Grid*, vol. 12, no. 5, pp. 3889–3900, 2021.
- [5] J. A. Momoh, S. S. Reddy, and Y. Baxi, "Stochastic voltage/var control with load variation," in *2014 IEEE PES General Meeting (PESGM)*, 2014, pp. 1–5.
- [6] A. Samimi and A. Kazemi, "Scenario-based stochastic programming for Volt/Var control in distribution systems with renewable energy sources," *IETE Technical Review*, vol. 33, no. 6, pp. 638–650, 2016.
- [7] R. A. Jabr, "Robust volt/var control with photovoltaics," *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 2401–2408, 2019.
- [8] M. Azarnia and M. Rahimiyan, "Robust volt-var control of a smart distribution system under uncertain voltage-dependent load and renewable production," *Int. J. Electr. Power Energy Syst.*, vol. 134, p. 107383, 2022.
- [9] Y. Yang and W. Wu, "A distributionally robust optimization model for real-time power dispatch in distribution networks," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3743–3752, 2019.
- [10] L. Dong, J. Li, T. Pu, and N. Chen, "Distributionally robust optimization model of active distribution network considering uncertainties of source and load," *J. Mod. Power Syst. Clean Energy*, vol. 7, no. 6, pp. 1585–1595, 2019.
- [11] B. Li, R. Jiang, and J. L. Mathieu, "Distributionally robust chance-constrained optimal power flow assuming unimodal distributions with misspecified modes," *IEEE Trans. Control. Netw. Syst.*, vol. 6, no. 3, pp. 1223–1234, 2019.
- [12] L. Yang, Y. Xu, W. Gu, and H. Sun, "Distributionally robust chance-constrained optimal power-gas flow under bidirectional interactions considering uncertain wind power," *IEEE Trans. Smart Grid*, vol. 12, no. 2, pp. 1722–1735, 2021.
- [13] L. Yang, Y. Xu, H. Sun, and W. Wu, "Tractable convex approximations for distributionally robust joint chance constrained optimal power flow under uncertainties," *IEEE Trans. Power Syst.*, pp. 1–1, 2021.
- [14] Z. Shen, Z. Wei, G. Sun, and S. Chen, "Representing ZIP loads in convex relaxations of optimal power flow problems," *Int. J. Electr. Power Energy Syst.*, vol. 110, pp. 372–385, 2019.
- [15] F. U. Nazir, B. C. Pal, and R. A. Jabr, "Approximate load models for conic OPF solvers," *IEEE Trans. Power Syst.*, vol. 36, no. 1, pp. 549–552, 2021.
- [16] Y. Guo, K. Baker, E. Dall'Anese, Z. Hu, and T. H. Summers, "Data-based distributionally robust stochastic optimal power flow—part I: Methodologies," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1483–1492, 2019.
- [17] —, "Data-based distributionally robust stochastic optimal power flow—part II: Case studies," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1493–1503, 2019.
- [18] F. Bu, K. Dehghanpour, and Z. Wang, "Enriching load data using micro-PMUs and smart meters," *IEEE Trans. Smart Grid*, vol. 12, no. 6, pp. 5084–5094, 2021.
- [19] "IEEE standard for interconnection and interoperability of distributed energy resources with associated electric power systems interfaces," *IEEE Std 1547-2018 (Revision of IEEE Std 1547-2003)*, pp. 1–138, 2018.
- [20] R. Cheng, Z. Wang, Y. Guo, and Q. Zhang, "Online voltage control for unbalanced distribution networks using projected newton method," *IEEE Trans. Power Syst.*, pp. 1–1, 2022.
- [21] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Oper. Res.*, vol. 58, no. 3, pp. 595–612, 2010.
- [22] G. C. Calafiore and L. El Ghaoui, "On distributionally robust chance-constrained linear programs," *J Optim Theory Appl.*, vol. 130, no. 1, pp. 1–22, 2006.
- [23] F. Bu, Y. Yuan, Z. Wang, K. Dehghanpour, and A. Kimber, "A time-series distribution test system based on real utility data," in *Proc. North American Power Symposium (NAPS)*, 2019, pp. 1–6.