A Data-Driven Stochastic Reactive Power Optimization Considering Uncertainties in Active Distribution Networks and Decomposition Method

Ding, Tao; Yang, Qingrun; Yang, Yongheng; Li, Cheng; Bie, Zhaohong; Blaabjerg, Frede

Published in:
IEEE Transactions on Smart Grid

DOI (link to publication from Publisher):
10.1109/TSG.2017.2677481

Publication date:
2018

Document Version
Accepted author manuscript, peer reviewed version

A Data-Driven Stochastic Reactive Power Optimization Considering Uncertainties in Active Distribution Networks and Decomposition Method

Tao Ding, *Member, IEEE*, Qingrun Yang, *Student Member, IEEE*, Yongheng Yang, *Member, IEEE*,
Cheng Li, *Student Member, IEEE*, Zhaohong Bie, *Senior Member, IEEE*, Frede Blaabjerg, *Fellow, IEEE*

***Abstract*— To address the uncertain output of distributed generators (DGs) for reactive power optimization in active distribution networks, the stochastic programming model is widely used. The model is employed to find an optimal control strategy with minimum expected network loss while satisfying all the physical constraints. Therein, the probability distribution of uncertainties in the stochastic model is always pre-defined by the historical data. However, the empirical distribution can be biased due to a limited amount of historical data and thus result in a suboptimal control decision. Therefore, in this paper, a data-driven modeling approach is introduced to assume that the probability distribution from the historical data is uncertain within a confidence set. Furthermore, a data-driven stochastic programming model is formulated as a two-stage problem, where the first-stage variables find the optimal control for discrete reactive power compensation equipment under the worst probability distribution of the second stage recourse. The second-stage variables are adjusted to uncertain probability distribution. In particular, this two-stage problem has a special structure so that the second-stage problem can be directly decomposed into several small-scale sub-problems, which can be handled in parallel without the information of dual problems. Numerical study on two distribution systems has been performed. Comparisons with the two-stage stochastic and robust approaches demonstrate the effectiveness of the proposal.**

***Index Terms*—Stochastic optimization; reactive power optimization; column-and-constraint generation algorithm; active distribution network; distributed generation**

NOMENCLATURE

Indices and Sets

i, j, k	Index for buses
t	Index for time period
B	Set of buses
E	Set of branches
Θ	Set of branches with transformers
Ω	Set of buses for reactive power compensators
Ω_D	Set of buses for shunt capacitors/reactors
$\pi(j)$	Set of all parents of bus j
$\delta(j)$	Set of all children of bus j
ψ	Confidence set of the probability distribution
Y^s	Feasible region of continuous variables under s-th scenario

Parameters

M	A large number
T	Time horizons
N_w	Cardinality of Θ
N_c	Cardinality of Ω_D
N_r	Cardinality of Ω_D
N_s	Number of scenarios
K	Number of observations for uncertain parameters
n_{ij}	Number of tap ratios at transformer branch (i, j)
r_{ij}, x_{ij}	Resistance/reactance of branch (i, j)
$b_{s,j}$	Shunt susceptance from j to ground
$C^{\max}_j/C^{\min}_j$	Upper/lower bound of shunt capacitors/reactors capacity at bus j
$W^{\max}_{ij}/W^{\min}_{ij}$	Upper/lower bound of transformer ratio limit at branch (i, j)
$\eta_{c,j}$	Specified operational times for shunt capacitors/reactor at bus j
$\eta_{w,ij}$	Specified operational times for transformer (i, j)
s_j	Step size of shunt capacitors/reactors at bus j
$w_{ij,k}$	Tap ratio on k-th level of the transformer (i, j)
$U_j^{\max}/U_j^{\min}$	Upper/lower bound of voltage magnitude at bus j
$I_l^{\max}$	Current capacity limit of branch (i, j)
$Q_{c,j}^{\max}/Q_{c,j}^{\min}$	Upper/lower bound of reactive power compensation for continuous reactive power compensators at bus j
$\boldsymbol{u}$	Uncertain parameters
$\boldsymbol{u}^s$	Uncertain parameters under s-th scenario
τ_j	Number of auxiliary binary variables $\lambda_{j,0},...,\lambda_{j,\tau_j}$
$\boldsymbol{p}_0$	Probability from the historical data
θ	A parameter that can control the size of the confidence set
θ_1	θ using 1-norm to control the size of the confidence set
θ_∞	θ using inf-norm to control the size of the confidence set
α	Confidence level

Variables

H_{ij}, G_{ij}	Active/reactive power flow from bus i to j

This work was supported in part by National Key Research and Development Program of China (2016YFB0901904), in part by National Natural Science Foundation of China (Grant 51607137), in part by China Postdoctoral Science Foundation (2015M580847), in part by Natural Science Basis Research Plan in Shaanxi Province of China (2016JQ5015), the project of State Key Laboratory of Electrical Insulation and in part by Power Equipment in Xi'an Jiaotong University (EIPE16301).
T. Ding (e-mail: tding15@mail.xjtu.edu.cn), Q. Yang, C. Li and Z. Bie are with the State Key Laboratory of Electrical Insulation and Power Equipment, Department of Electrical Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, China;
Y. Yang and F. Blaabjerg are with the Department of Energy Technology, Aalborg University, Aalborg DK-9220, Denmark (e-mail: yoy@et.aau.dk; fbl@et.aau.dk).

U_j	Voltage magnitude of bus j
P_j, Q_j	Injected active/reactive power of bus j
l_{ij}	Squared branch current at branch (i, j)
w_{ij}	Tap ratio of the transformer branch (i, j)
C_j	Value of shunt capacitors/reactors at bus j
o_{ij}^{k}	Optimal 0-1 decision on k-th level of the transformer (i, j)
ρ_j	Optimal step of shunt capacitors/reactors at bus j
$Q_{c,j}$	Value of reactive power compensation for continuous reactive power compensators at bus j
v_j	Squared voltage magnitude of bus j
$\boldsymbol{z}$	Discrete decision variables
$\boldsymbol{y}$	Continuous decision variables
$\boldsymbol{y}^s$	Continuous decision variables under s-th scenario
$\lambda_{j,0},...,\lambda_{j,\tau_j}$	Auxiliary binary variables to express the integer variable ρ_j by binary code

I. INTRODUCTION

DISTRIBUTED networks, characterized by their mostly radial topology, are featured with heavily fluctuating loads, which may lead to large power losses and voltage drop near the end of feeders, adversely affecting industrial manufactures and daily lives. To improve the power quality, reactive power optimization, serving for tertiary voltage control (TVC), aims to minimize the total transmission losses and improve the voltage profile by controlling reactive power compensators and transformer tap ratios over several periods, while satisfying specific physical and operating constraints.

Generally, the controlled equipment can be classified as continuous and discrete controllable devices. The discrete controllable devices are controlled via switching on/off and they should not be adjusted quite frequently due to their service lifetime and existing manufacture techniques. Thus, the total number of operating times of discrete controllable devices is limited, which leads to the development of the dynamic reactive power optimization (DRPO) model [1-3]. This model is actually a large-scale mixed-integer nonlinear programming and several techniques including intelligent searches and standard branch-and-bound/cut methods were proposed to solve this complex model [4]-[6]. With a proposal of a two-stage multi-period mixed-integer convex model, [7] analyzed the tradeoff between risk mitigation and investment cost minimization. In [8], a voltage security constrained multi-period optimal reactive power flow model was proposed based on the generalized Benders decomposition method with an optimal condition decomposition approach to solve it. However, the size of data arises as a result of large-scale mixed-integer nonlinear programming problems with multi-periods, increasing the computational burden and time. Recently, the conic relaxation technique was studied in distribution networks, which gives a sound solution while significantly improving the computational performance [9]-[11]. For instance, in [12], the second-order cones to relax the non-convex power flow equations were proposed in order to obtain a mixed integer second order coned programming model, after which a sensitivity-based relaxation and decomposition method was introduced to further improve the computation. After determining the total size of the distributed energy storage (DES, e.g., batteries) and optimal locations for the DES, [13] applied the second order cone programming relaxation to obtain the globally optimal solution and avoid the problem of NP-hardness. Furthermore, [14] dealt with a joint problem of reactive power optimization and network reconfiguration to minimize power losses and improve the voltage profile, the original non-convex model of which was converted into a mixed integer second order cone programming model using the second-order cone relaxation, the big-M method and the piecewise linearization techniques.

Nevertheless, an increasing number of distributed generators (DGs) including wind power and photovoltaic (PV) is coming into distributed networks nowadays. The distributed networks integrated with DGs, termed as active distributed networks, are facing critical technical challenges to traditional operation due to the stochastic nature of DGs, which may result in uncertain output, and thus severer voltage violations.

To cope with the uncertain output of DGs in the optimization operation in active distribution networks, stochastic programming [15-17], chance-constrained based stochastic programming [18-20] and robust optimization [21-23] have been extensively explored. For example, a multi-scenario framework for optimal power flow under the worst wind scenario and transmission N-1 contingency to properly address the uncertain wind power generation was proposed in [24]. A stochastic multi-objective framework for distribution feeder reconfiguration was employed in [25], firstly converting it into specific deterministic scenarios among random scenarios of wind/load forecast variations and then implementing multi-objective formulation for each deterministic scenario in the first stage. In [26], a chance-constrained programming for optimal power flow under uncertainty considering nonlinear model with multiple uncertain inputs was studied, where a back-mapping approach and linear approximation of nonlinear model equations were performed. Furthermore, [27] converted the chance-constrained stochastic programming formulation into a linear deterministic problem and a decomposition-based method to solve the day-ahead scheduling problem. Although linearized models enable to improve computational efficiency, the accuracy of linearization should be ensured.

Generally, stochastic programming methods cannot cover all the possible realization of uncertainties. In order to address this problem, robust optimization was proposed to immunize against the solution within a given uncertainty set. As presented in [28], a two-stage robust reactive power optimization to coordinate the discrete and continuous reactive power compensators was set up, while hedging against any possible realization within uncertain wind power output. A mixed-integer two-stage robust optimization formulation and a decomposition algorithm in a master-slave structure to achieve minimum network losses were discussed in [29], considering the worst conditions over uncertainty sets. Although the robust optimization can protect the system against a pre-defined uncertainty set, it always gives a more conservative solution than the stochastic approach.

In practice, historical data of DG outputs may be available at ISOs/RTOs. Therefore, it is possible to derive a more efficient solution that is robust while less conservative, which incorporates the superiority of both stochastic and robust approaches. According to the historical data, a confidence set is constructed for the probability distribution of the uncertainties to find an optimal solution under the worst probability distribution [30]-[35]. Therefore, a data-driven two-stage stochastic dynamic reactive power optimization model is developed in this work to coordinate the discrete and continuous controllable devices, while addressing the uncertain DG output. The contributions of the paper are summarized as follows:

1) It is the first time to set up a data-driven stochastic programming model in the distribution networks, where the second order cone programming relaxation is utilized to relax the nonconvex feasible region caused by the branch flow equations. Furthermore, the dynamic reactive power optimization can be termed as a large-scale mixed-integer second order cone programming model.

2) It is found that the proposed model has a special structure in the second-stage bi-level model, where the feasible region of the uncertainty set is disjoint with the operating region. As a result, a new column-and-constraint generation algorithm is proposed to decompose the bi-level problem into several small-scale sub-problems to be handled in parallel, which does not require the duality information as the traditional method.

The rest of the paper is organized as follows: Section II presents a general dynamic reactive power optimization based on second order cone programming relaxation for active distribution networks. In Section III, a data-driven stochastic reactive power optimization model is proposed with the consideration of uncertain DG output. Furthermore, a new duality-free based column-and-constraint generation algorithm is presented to solve the proposed reactive power optimization model in Section IV. In Section V, numerical results obtained on a 33-bus system demonstrate the effectiveness of the proposal, which is also compared with the two traditional approaches. Finally, conclusions are drawn in Section VI.

II. Reactive Power Optimization Model in Active Distribution Networks

A. Formulation of Reactive Power Optimization Model

Distribution networks, different from transmission networks, have the property that the topology is radial, so it is very common to utilize the branch flow formulation for describing the power flow in distribution networks [12], [28], [36].

$$\begin{cases} P_j = \sum_{k\in\delta(j)} H_{jk} - \sum_{i\in\pi(j)} \left(H_{ij} - r_{ij} l_{ij}\right), \ \forall j \in B \\ Q_j = \sum_{k\in\delta(j)} G_{jk} - \sum_{i\in\pi(j)} \left(G_{ij} - x_{ij} l_{ij}\right) + b_{s,j} U_j^2 \ \ \forall j \in B \\ U_j^2 = U_i^2 - 2\left(r_{ij} H_{ij} + x_{ij} G_{ij}\right) + \left(r_{ij}^2 + x_{ij}^2\right) l_{ij}, \ \forall (i,j) \in E \setminus \Theta \\ \dfrac{U_j^2}{w_{ij}^2} = U_i^2 - 2\left(r_{ij} H_{ij} + x_{ij} G_{ij}\right) + \left(r_{ij}^2 + x_{ij}^2\right) l_{ij}, \ \forall (i,j) \in \Theta \\ H_{ij}^2 + G_{ij}^2 = l_{ij} U_i^2, \ \forall (i,j) \in E \end{cases} \quad (1)$$

where $(i,j)\in E\setminus\Theta$ denotes $(i,j)\in E$, but $(i,j)\notin\Theta$. The first and second equations describe the active and reactive power balance at each bus; the third and fourth equations describe the voltage drop at each line and transformer; the last equation describes the relationship among voltage, current and power.

The reactive power optimization problem essentially aims to minimize total power losses by controlling the reactive power compensators and transformer tap ratios over a given number of time horizons while satisfying various physical constraints. Here, the reactive power compensators can be classified as continuous adjustment equipment such as DG output, and discrete adjustment equipment including capacitor banks. It is common that the electric devices including transformer tap ratios and switched capacitor banks cannot be adjusted very frequently due to the limitation of their service lifetime and existing manufacture techniques. Therefore, the maximum allowable operational times should be considered in the model and the reactive power optimization model can be exactly written as follows

$$\min_{Q_c(t),\rho(t),o(t)} \sum_{t=1}^{T} \sum_{(i,j)\in E} \left(r_{ij} l_{ij}(t)\right) \quad (2)$$

$$s.t \quad P_{DG,j}(t) - P_{L,j}(t) = \sum_{k\in\delta(j)} H_{jk}(t) - \sum_{i\in\pi(j)} \left(H_{ij}(t) - r_{ij} l_{ij}(t)\right), \quad \forall j \in B,\ t=1,\ldots,T \quad (3)$$

$$\frac{1}{2}U_j^2(t) C_j(t) + Q_{c,j}(t) - Q_{L,j}(t) = \sum_{k\in\delta(j)} G_{jk}(t) - \sum_{i\in\pi(j)} \left(G_{ij}(t) - x_{ij} l_{ij}(t)\right) + b_{s,j} U_j^2(t), \ \forall j \in B,\ t=1,\ldots,T \quad (4)$$

$$U_j^2(t) = U_i^2(t) - 2\left(r_{ij} H_{ij}(t) + x_{ij} G_{ij}(t)\right) + \left(r_{ij}^2 + x_{ij}^2\right) l_{ij}(t), \quad \forall (i,j) \in E/\Theta,\ t=1,\ldots,T \quad (5)$$

$$\sum_{k=0}^{n_{ij}} \frac{o_{ij,k}(t)}{\left(w_{ij,k}\right)^2} U_j^2(t) = U_i^2(t) - 2\left(r_{ij} H_{ij}(t) + x_{ij} G_{ij}(t)\right) + \left(r_{ij}^2(t) + x_{ij}^2(t)\right) l_{ij}(t), \ \forall (i,j) \in \Theta,\ t=1,\ldots,T \quad (6)$$

$$H_{ij}^2(t) + G_{ij}^2(t) = l_{ij}(t) U_i^2(t), \ \forall (i,j) \in E \quad (7)$$

$$\sum_{k=0}^{n_{ij}} o_{ij,k}(t) = 1, \quad \forall (i,j) \in \Theta \quad (8)$$

$$U_j^{\min} \le U_j(t) \le U_j^{\max}, \quad \forall j \in B \quad (9)$$

$$0 \le l_{ij}(t) \le \left(I_{ij}^{\max}\right)^2, \quad \forall (i,j) \in E \quad (10)$$

$$Q_{c,j}^{\min} \le Q_{c,j}(t) \le Q_{c,j}^{\max}, \quad \forall j \in \Omega \setminus \Omega_D \quad (11)$$

$$C_j(t) = C_j^{\min} + s_j \rho_j(t), \quad \forall j \in \Omega_D \quad (12)$$

$$C_j^{\min} \le C_j(t) \le C_j^{\max}, \quad \forall j \in \Omega_D \quad (13)$$

$$\sum_{t=2}^{T} \left|\rho_j(t+1) - \rho_j(t)\right| \le \eta_{c,j}, \quad \forall j \in \Omega_D \quad (14)$$

$$\sum_{t=2}^{T} \left|\sum_{k=0}^{n_{ij}} 2^k \left(o_{ij}^k(t+1) - o_{ij}^k(t)\right)\right| \le \eta_{w,ij}, \quad \forall (i,j) \in \Theta \quad (15)$$

$$o_{ij,k}(t) \in \{0,1\}, \quad \forall (i,j) \in \Theta, \quad (16)$$

$$\rho_j(t) \in Z^+, \quad \forall j \in \Omega_D \quad (17)$$

where (2) aims to minimize total network loss over T time periods; (3)-(4) denote the power balance at each bus; (5)-(7) show the Ohm's law for each branch, including (6) for transformer branch; (8) shows a choice constraint by which only one trap ratio level is chosen; (9)-(10) are constraints for voltage magnitude and branch current; (11) is the constraint for the continuous reactive power compensators; (12)-(13) are the constraints for discrete reactive power compensators; (14)-(15) are restrictions that the total allowable operational times by discrete adjustment equipment should be limited.

However, the model (2)-(17) is a mixed integer nonlinear nonconvex programming which is very difficult to solve. However, the non-convexity comes only from the nonlinear power flow constraints. To address this issue, the semi-definite programming (SDP) and second order cone programming (SOCP) were proposed to convexify the feasible region enclosed by the power flow constraints [9]-[10]. It was shown in [9]-[10] that SOCP and SDP relaxation methods are equivalent for the radial network, but the computational time from the former one is much less than the latter one. This is because both SOCP and SDP are solved by the standard primal-dual interior point method, but SOCP has much better worse-case complexity than SDP [37]. Theoretically, the complexity of SOCP is $O(n^3)$ whereas $O(n^4)$ of SDP. Here, n is the number of variables. Thus, for a large power system with numerous variables, SOCP would perform much faster than SDP and thus is selected in this work.

B. SOCP Relaxation for Reactive Power Optimization Model

At first, let $U_j^2(t) = v_j(t)$ for $\forall j \in B$ and then constraints (4)-(7), (9) will become

$$\frac{1}{2} v_j(t) C_j(t) + Q_{c,j}(t) - Q_{L,j}(t) = \sum_{k\in\delta(j)} G_{jk}(t) - \sum_{i\in\pi(j)} \left(G_{ij}(t) - x_{ij} l_{ij}(t)\right) + b_{s,j} v_j(t), \quad \forall j \in B,\ t = 1,...,T \tag{18}$$

$$v_j(t) = v_i(t) - 2\left(r_{ij} H_{ij}(t) + x_{ij} G_{ij}(t)\right) + \left(r_{ij}^2 + x_{ij}^2\right) l_{ij}(t), \quad \forall (i,j) \in E/\Theta,\ t = 1,...,T \tag{19}$$

$$\sum_{k=0}^{n_{ij}} \frac{o_{ij,k}(t)}{\left(w_{ij,k}\right)^2} v_j(t) = u_i(t) - 2\left(r_{ij} H_{ij}(t) + x_{ij} G_{ij}(t)\right) + \left(r_{ij}^2(t) + x_{ij}^2(t)\right) l_{ij}(t), \quad \forall (i,j) \in \Theta,\ t = 1,...,T \tag{20}$$

$$H_{ij}^2(t) + G_{ij}^2(t) = l_{ij}(t) v_i(t), \quad \forall (i,j) \in E \tag{21}$$

$$\left(U_j^{\min}\right)^2 \le v_j(t) \le \left(U_j^{\max}\right)^2, \quad \forall j \in B \tag{22}$$

The constraint in (21) is a nonlinear equality, resulting in the nonconvex problem. To address this issue, the second order cone relaxation is performed by relaxing the quadratic equality into inequality, yielding

$$\left\| \begin{matrix} 2H_{ij}(t) \\ 2G_{ij}(t) \\ l_{ij}(t) - u_i(t) \end{matrix} \right\|_2 \le l_{ij}(t) + v_i(t), \quad \forall (i,j) \in E \tag{23}$$

After this relaxation, the original reactive power optimization model will lead to be a mixed integer second order cone programming model, but not a standard mixed integer second order cone programming model since there are still many bilinear terms in the above model, and we can simplify them by reformulations in the appendix, leading to (A10)-(A12).

Subsequently, the reactive power optimization model in can be mathematically formulated as a general problem as

$$\min_{\boldsymbol{z},\boldsymbol{y}\in\Upsilon} \ \boldsymbol{a}^T \boldsymbol{y} \tag{24}$$

$$s.t. \quad \boldsymbol{Az} \ge \boldsymbol{b}, \quad \boldsymbol{z} \in \{0,1\} \tag{25}$$

$$\Upsilon = \left\{ \begin{matrix} \boldsymbol{y} \,|\, \boldsymbol{Cy} \le \boldsymbol{f},\ \left\|\boldsymbol{Q}_i \boldsymbol{y} + \boldsymbol{q}_i\right\|_2 \le \boldsymbol{c}_i^T \boldsymbol{y} + d_i, i = 1,...,n \\ \boldsymbol{Dy} = \boldsymbol{g} - \boldsymbol{Gz}, \quad \boldsymbol{Ey} = \boldsymbol{u} \end{matrix} \right\} \tag{26}$$

where $\boldsymbol{C}$, $\boldsymbol{Q}_i$, $\boldsymbol{q}_i$, $\boldsymbol{c}_i$, $\boldsymbol{D}$, $\boldsymbol{g}$, $\boldsymbol{G}$, $\boldsymbol{E}$ and d_i are matrix/vector form with respect to the original model.

III. Data-Driven Stochastic Reactive Power Optimization Considering Uncertainties

In the last section, the reactive power optimization model is only conducted under a given load demand curve over multiple time periods. However, to address the uncertain generation output of the distributed generators (i.e., $\boldsymbol{u}$ in (26)), the stochastic programming is employed to coordinate the discrete and continuous reactive power compensators. Specifically, the discrete decision variables (i.e., $\boldsymbol{z}$ in (24)) should be determined before the uncertainty is revealed since such equipment should not be adjusted quite frequently, whereas the continuous decision variables (i.e., $\boldsymbol{y}$ in (24)) can be flexible with the revealed uncertainty. This framework gives a two-stage framework and for the N_s scenarios of uncertainties from discretizing the given probability distribution, such that $\boldsymbol{u}^1,\dots,\boldsymbol{u}^{N_s}$ and the corresponding probability is ($p^1,\dots,p^{N_s}$). The objective function minimizes the total expected network loss. Then, the general data-driven stochastic reactive power optimization model is formulated as

$$\min_{\boldsymbol{z},\boldsymbol{y}^s\in\Upsilon^s} \ \sum_{s=1}^{N_s} p_s \boldsymbol{a}^T \boldsymbol{y}^s \tag{27}$$

$$s.t. \quad \boldsymbol{Az} \ge \boldsymbol{b}, \quad \boldsymbol{z} \in \{0,1\} \tag{28}$$

$$\Upsilon^s = \left\{ \begin{matrix} \boldsymbol{y}^s \,|\, \boldsymbol{Cy}^s \le \boldsymbol{f},\ \left\|\boldsymbol{Q}_i \boldsymbol{y}^s + \boldsymbol{q}_i\right\|_2 \le \boldsymbol{c}_i^T \boldsymbol{y}^s + d_i, i = 1,...,n \\ \boldsymbol{Dy}^s = \boldsymbol{g} - \boldsymbol{Gz}, \quad \boldsymbol{Ey}^s = \boldsymbol{u}^s \end{matrix} \right\} \tag{29}$$

Due to the limited information from the historical data, the probability distribution of uncertainties cannot be exactly determined by the data. As a result, we allow the probability distribution of uncertainties to be arbitrary within a pre-defined confidence set constructed from the historical data. Thus, the proposed data-driven stochastic reactive power optimization model aims to find the optimal solution under the worst-case probability distribution, such that

$$\min_{\boldsymbol{z}} \max_{\boldsymbol{p}\in\psi} \min_{\boldsymbol{y}\in\Upsilon} \ \sum_{s=1}^{N_s} p_s \boldsymbol{a}^T \boldsymbol{y}^s \tag{30}$$

$$s.t. \quad \boldsymbol{Az} \ge \boldsymbol{b}, \quad \boldsymbol{z} \in \{0,1\} \tag{31}$$

$$\Upsilon^s = \left\{ \begin{matrix} \boldsymbol{y}^s \,|\, \boldsymbol{Cy}^s \le \boldsymbol{f},\ \left\|\boldsymbol{Q}_i \boldsymbol{y}^s + \boldsymbol{q}_i\right\|_2 \le \boldsymbol{c}_i^T \boldsymbol{y}^s + d_i, i = 1,...,n \\ \boldsymbol{Dy}^s = \boldsymbol{g} - \boldsymbol{Gz}, \quad \boldsymbol{Ey}^s = \boldsymbol{u}^s \end{matrix} \right\} \tag{32}$$

In [35], two popular confidence sets based on norm-1 and norm-inf were presented for ψ, which can be expressed as

$$\psi_1 = \left\{ \boldsymbol{p} \in R_+^{N_s} \,\middle|\, \left\| \boldsymbol{p} - \boldsymbol{p}_0 \right\|_1 \le \theta \right\} = \left\{ \boldsymbol{p} \in R_+^{N_s} \,\middle|\, \sum_{k=1}^{N_s} \left(p_k - p_{0,k} \right) \le \theta \right\} \quad (33)$$

$$\psi_\infty = \left\{ \boldsymbol{p} \in R_+^{N_s} \,\middle|\, \left\| \boldsymbol{p} - \boldsymbol{p}_0 \right\|_\infty \le \theta \right\} = \left\{ \boldsymbol{p} \in R_+^{N_s} \,\middle|\, \max_{1 \le k \le N_s} \left| p_k - p_{0,k} \right| \le \theta \right\} \quad (34)$$

Supposing N_s scenarios from K observations, we have the following relationship between the number of historical data and θ:

$$\Pr\left\{ \left\| \boldsymbol{p} - \boldsymbol{p}_0 \right\|_1 \le \theta \right\} \ge 1 - 2N_s e^{-2K\theta/N_s} \quad (35)$$

$$\Pr\left\{ \left\| \boldsymbol{p} - \boldsymbol{p}_0 \right\|_\infty \le \theta \right\} \ge 1 - 2N_s e^{-2K\theta} \quad (36)$$

It can be found that the right-hand side of (35)-(36) is actually the confidence level of the confidence set. Then, the relationship between confidence level (i.e., the right-hand side of (35)-(36)) α and the value of θ is given by

$$\theta_1 = \frac{N_s}{2K} \ln \frac{2N_s}{1-\alpha} \quad (37)$$

$$\theta_\infty = \frac{1}{2K} \ln \frac{2N_s}{1-\alpha} \quad (38)$$

Furthermore, (37) and (38) show that with the increase of the number of historical data, i.e., M, the estimated probability distribution will be closer to its true distribution. That means, θ will become smaller until to zero. Moreover, for the same α, θ_∞ is smaller than θ_1.

IV. Column-and-Constraint Generation Algorithm

The proposed data-driven stochastic reactive power optimization model can be cast as a two-stage optimization problem which generally can be solved by the Benders decomposition method or standard column-and-constraint generation method (C&CG). These methods are implemented in a master-subproblem framework: sub-problem (SP) aims to find the critical scenario of the uncertain set for a given first-stage decision variable that provides an upper bound; then new variables and constraints are added to the master problem (MP) to obtain a lower bound. The MP and SP are solved iteratively and the method stops until the gap between the upper and lower bounds is smaller than a pre-set convergence tolerance.

A. C&CG-Sub-problem

For a given specific first-stage variables in the k-th iteration as $\boldsymbol{z}^{k*}$, we can set up a second-stage bi-level "max-min" model from (30)-(32) to find the worst-case scenario, yielding

$$\max_{\boldsymbol{p} \in \psi} \min_{\boldsymbol{y}^s \in \mathrm{Y}^s} \sum_{s=1}^{N_s} p_s \boldsymbol{a}^T \boldsymbol{y}^s \quad (39)$$

$$\text{s.t.} \quad \mathrm{Y}^s = \left\{ \boldsymbol{y}^s \,\middle|\, \begin{array}{l} \boldsymbol{C}\boldsymbol{y}^s \le \boldsymbol{f},\ \left\| \boldsymbol{Q}_i \boldsymbol{y}^s + \boldsymbol{q}_i \right\|_2 \le \boldsymbol{c}_i^T \boldsymbol{y}^s + d_i, i = 1,\dots,n \\ \boldsymbol{D}\boldsymbol{y}^s = \boldsymbol{g} - \boldsymbol{G}\boldsymbol{z}^{k*}, \quad \boldsymbol{E}\boldsymbol{y}^s = \boldsymbol{u}^s \end{array} \right\}$$

$$s = 1,\dots,N_s \quad (40)$$

It can be observed that the model (39)-(40) has some special properties: (i) the sub-feasible regions ($\mathrm{Y}^1,\dots,\mathrm{Y}^s,\dots,\mathrm{Y}^{Ns}$) are separable; (ii) the decision variables $\boldsymbol{p}$ are all nonnegative; (iii) the feasible region of ψ and Y^s are absolutely disjoint.

For the first and second properties that the sub-feasible regions ($\mathrm{Y}^1,\dots,\mathrm{Y}^s,\dots,\mathrm{Y}^{Ns}$) are separable and the decision variables $\boldsymbol{p}$ are all nonnegative, we can exchange the summation operator "Σ" and "min" operator, so the second-stage "max-min" problem can be reformulated as

$$\max_{\boldsymbol{p} \in \psi} \sum_{s=1}^{N_s} p_s \min_{\boldsymbol{y}^s \in \mathrm{Y}^s} \boldsymbol{a}^T \boldsymbol{y}^s \quad (41)$$

$$\text{s.t.} \quad \mathrm{Y}^s = \left\{ \boldsymbol{y}^s \,\middle|\, \begin{array}{l} \boldsymbol{C}\boldsymbol{y}^s \le \boldsymbol{f},\ \left\| \boldsymbol{Q}_i \boldsymbol{y}^s + \boldsymbol{q}_i \right\|_2 \le \boldsymbol{c}_i^T \boldsymbol{y}^s + d_i, i = 1,\dots,n \\ \boldsymbol{D}\boldsymbol{y}^s = \boldsymbol{g} - \boldsymbol{G}\boldsymbol{z}^{k*}, \quad \boldsymbol{E}\boldsymbol{y}^s = \boldsymbol{u}^s \end{array} \right\}$$

$$s = 1,\dots,N_s \quad (42)$$

For convenience, let $h_s = \min_{\boldsymbol{y}^s \in \mathrm{Y}^s} \boldsymbol{a}^T \boldsymbol{y}^s$ and the above model becomes

$$\max_{\boldsymbol{p} \in \psi} \sum_{s=1}^{N_s} p_s h_s \quad (43)$$

$$\text{s.t.} \quad h_s = \arg\min_{\boldsymbol{y}^s \in \mathrm{Y}^s} \boldsymbol{a}^T \boldsymbol{y}^s \quad (44)$$

$$\text{s.t.} \quad \mathrm{Y}^s = \left\{ \boldsymbol{y}^s \,\middle|\, \begin{array}{l} \boldsymbol{C}\boldsymbol{y}^s \le \boldsymbol{f},\ \left\| \boldsymbol{Q}_i \boldsymbol{y}^s + \boldsymbol{q}_i \right\|_2 \le \boldsymbol{c}_i^T \boldsymbol{y}^s + d_i, i = 1,\dots,n \\ \boldsymbol{D}\boldsymbol{y}^s = \boldsymbol{g} - \boldsymbol{G}\boldsymbol{z}^{k*}, \quad \boldsymbol{E}\boldsymbol{y}^s = \boldsymbol{u}^s \end{array} \right\}$$

$$s = 1,\dots,N_s \quad (45)$$

According to the property (iii), the feasible region Y^s for variables $\boldsymbol{y}^s$ and the feasible region ψ for variables p^s are absolutely disjoint. That means, the feasible region of upper-level model ψ doesn't affect the lower-level model and for any given value p^s, the optimal solution $\boldsymbol{y}^s$ is unique. As a result, the bi-level model can be solved by sequentially solving upper-level and lower-level models, respectively. Moreover, the first property tells that the sub-feasible regions ($\mathrm{Y}^1,\dots,\mathrm{Y}^s,\dots,\mathrm{Y}^{Ns}$) are separable, so lower-level model of the bi-level model can be further decomposed into N_s independent optimization models. This gives the fact that the bi-level model can be decoupled by the following structure:

For each $\boldsymbol{u}^s$, it generates a second order cone programming model, such that

$$h_s^* = \arg\min_{\boldsymbol{y}^s \in \mathrm{Y}^s} \boldsymbol{a}^T \boldsymbol{y}^s \quad (46)$$

$$\text{s.t.} \quad \mathrm{Y}^s = \left\{ \boldsymbol{y}^s \,\middle|\, \begin{array}{l} \boldsymbol{C}\boldsymbol{y}^s \le \boldsymbol{f},\ \left\| \boldsymbol{Q}_i \boldsymbol{y}^s + \boldsymbol{q}_i \right\|_2 \le \boldsymbol{c}_i^T \boldsymbol{y}^s + d_i, i = 1,\dots,n \\ \boldsymbol{D}\boldsymbol{y}^s = \boldsymbol{g} - \boldsymbol{G}\boldsymbol{z}^{k*}, \quad \boldsymbol{E}\boldsymbol{y}^s = \boldsymbol{u}^s \end{array} \right\} \quad (47)$$

It can be observed that the above second order cone programming models are N_s small models, comparing to the original model (43)-(45), since (46)-(47) only contains variables $\boldsymbol{y}^s$ for each model whereas (43)-(45) contains ($\boldsymbol{y}^1,\dots,\boldsymbol{y}^s,\dots,\boldsymbol{y}^{Ns}$) simultaneously in one model. Moreover, the N_s small models can be handled in parallel.

After obtaining the optimal solution ($h_1^*,\dots,h_{N_s}^*$) for the above N_s small models, we have

$$\max_{\boldsymbol{p} \in \psi} \sum_{s=1}^{N_s} h_s^* p_s \quad (48)$$

Thus, we can see that the original bi-level model can be solved by N_s small second order cone programming models that can be handled in parallel and one small linear programming.

When the SP is solved, an optimal value $Q\left(\boldsymbol{z}^{k*}\right)$ and the worst-case probability $\boldsymbol{p}^{k*}$ are obtained, which in fact gives an upper bound for the original model. Then, a set of extra variables $\boldsymbol{y}^{s,k+1}$ and associated constraints are generated and added into master problem by fixing the optimal probability $\boldsymbol{p}^{k*}$ from the above model in (48).

If the SP is feasible, we can create variables $\boldsymbol{y}^{s,k+1}$ and assign the following constraints to C&CG-master problem, which is called "optimality cuts".

$$\eta \geq \sum_{s=1}^{N_s} p_s^{k*} \boldsymbol{a}^T \boldsymbol{y}^{s,k+1} \tag{49}$$

$$\boldsymbol{C}\boldsymbol{y}^{s,k+1} \leq \boldsymbol{f},\ \left\|\boldsymbol{Q}_i \boldsymbol{y}^{s,k+1} + \boldsymbol{q}_i\right\|_2 \leq \boldsymbol{c}_i^T \boldsymbol{y}^{s,k+1} + d_i, i = 1,...,n \tag{50}$$

$$\boldsymbol{D}\boldsymbol{y}^{s,k+1} = \boldsymbol{g} - \boldsymbol{G}\boldsymbol{z},\quad \boldsymbol{E}\boldsymbol{y}^{s,k+1} = \boldsymbol{u}^s,\ s = 1,..., N_s \tag{51}$$

where η is a dummy continuous variable.

If the SP is infeasible, it is possible to create variables $\boldsymbol{x}^{k+1}$ and assign the following constraints to C&CG-master problem, which is called "feasibility cuts".

$$\boldsymbol{C}\boldsymbol{y}^{s,k+1} \leq \boldsymbol{f},\ \left\|\boldsymbol{Q}_i \boldsymbol{y}^{s,k+1} + \boldsymbol{q}_i\right\|_2 \leq \boldsymbol{c}_i^T \boldsymbol{y}^{s,k+1} + d_i, i = 1,...,n \tag{52}$$

$$\boldsymbol{D}\boldsymbol{y}^{s,k+1} = \boldsymbol{g} - \boldsymbol{G}\boldsymbol{z},\quad \boldsymbol{E}\boldsymbol{y}^{s,k+1} = \boldsymbol{u}^s,\ s = 1,..., N_s \tag{53}$$

B. C&CG-Master Problem

The MP aims to relax the original optimization model and provide a lower bound. After K iterations have been preceded, the master problem can be described as follows:

$$\min_{\boldsymbol{z}} \quad \eta \tag{54}$$

$$s.t. \qquad \boldsymbol{A}\boldsymbol{z} \geq \boldsymbol{b},\quad \boldsymbol{z} \in \{0,1\} \tag{55}$$

$$\eta \geq \sum_{s=1}^{N_s} p_s^{k*} \boldsymbol{a}^T \boldsymbol{y}^{s,k+1},\ k = 1,2,...K \tag{56}$$

$$\boldsymbol{C}\boldsymbol{y}^{s,k+1} \leq \boldsymbol{f},\ \left\|\boldsymbol{Q}_i \boldsymbol{y}^{s,k+1} + \boldsymbol{q}_i\right\|_2 \leq \boldsymbol{c}_i^T \boldsymbol{y}^{s,k+1} + d_i, i = 1,...,n,\quad k = 1,2,...K,\ s = 1,..., N_s \tag{57}$$

$$\boldsymbol{D}\boldsymbol{y}^{s,k+1} = \boldsymbol{g} - \boldsymbol{G}\boldsymbol{z},\quad \boldsymbol{E}\boldsymbol{y}^{s,k+1} = \boldsymbol{u}^s,\ k = 1,2,...K,\ s = 1,..., N_s \tag{58}$$

The above MP is a standard mixed integer second order cone programming model that can be easily handled by the off-the-shell commercial solvers, such as MOSEK, CPLEX, GUROBI, *etc*.. Solving the MP gives the optimal discrete variables $\boldsymbol{z}^*$ and optimal continuous variables ($\boldsymbol{y}^{s,1*}$,…,$\boldsymbol{y}^{s,k*}$) that are generated in SP for the uncertainty set. The SP and the MP are solved iteratively until the given convergence criteria is satisfied and thus the global optimal solution is obtained. As presented in [24], the column-and-constraint generation algorithm can be converged in finite iterations where all possible realizations of are needed to be enumerated [24]. However, it should be noted that the proposed column-and-constraint generation algorithm is a little different from that in [24], where the proposed model has a special structure, so that there is no need as [24] to dualize the inner "max-min" bi-level model into a single level model to solve it. With respect to the special structure, we propose a novel C&CG decomposition method with the information of duality. For the given convergence error ε, the implementation steps of the proposed algorithm are given in Table I.

TABLE I PROCEDURE OF C&CG ALGORITHM

	C&CG algorithm
1.	Set LB= $-\infty$, UB= $+\infty$, k=0;
2.	Solve the master problem (54)-(58). Derive an optimal solution $\boldsymbol{z}^*$ and (η^* , $\boldsymbol{y}^{s,1*}$,…,$\boldsymbol{y}^{s,k+1*}$) for s=1,…,N_s. Then, update the lower bound $LB = \boldsymbol{c}^T \boldsymbol{y}^* + \eta^*$;
3.	Fix $\boldsymbol{z}^*$ and solve the subproblems (46)-(47) in parallel, yielding an optimal value $Q(\boldsymbol{z}^*)$ and worst-case probability $\boldsymbol{p}^*$.
4	Update UB=min{UB , $\boldsymbol{c}^T\boldsymbol{z}^* + Q\left(\boldsymbol{z}^*\right)$ }. If UB-LB≤ε, terminate; else, go to step 5.
5	Generate variables ($\boldsymbol{y}^{1,k+1}$,…, $\boldsymbol{y}^{s,k+1}$,…, $\boldsymbol{y}^{N_s,k+1}$). Add the new variables and constraints to master problem according to (49)-(53). Update k=k+1, and go to step 2.

V. NUMERICAL ANALYSIS

A. Test System and Data Collection

In this section, a 33-bus distribution network that is plotted in Fig. 1 is analyzed to verify the proposed method. We consider the step of tap ratio (TR) of the transformer in the substation is 0.01 and the range is [0.94, 1.06]. Two switchable capacitors/reactors (SCRs) are connected to buses {#3, #9} whose capacity are both [-0.01, +0.01] MVar, where the steps are 0.002 and 0.005 MVar. The maximum operating times over 24 hours for SCRs are 8 and 6, respectively. Besides, five DGs are installed at buses {#19, #25, #28, #31, #33} with the capacity being 0.1 MW, 0.2 MW, 0.3 MW, 0.3 MW and 0.3 MW respectively. The forecasted load demand and DG generation factors over 24 hours are depicted in Fig. 2., where it is assumed that the uncertain DG output follows a multivariate normal distribution with the variance equivalent to 1/5 of the mean value (a.k.a., forecasted value). We randomly generate 1000 samples by Monte Carlo simulation to simulate the set of the historical data. Taking the θ_1 determined by (37) for example, the relationship among θ_1, N_s and α is shown in Fig. 3. This reveals that for the given number of samples, with the increase of the number of scenarios N_s and confidence level α, θ_1 becomes larger and the uncertainty set will become larger as well. It is obvious that the size of uncertainty set will affect the optimal solution, so in the following study, we will choose different N_s and α to show the impact of uncertainty set on the reactive power optimization model.

The computational tasks were performed on a 2.0 GHz personal computer with 4 GB RAM, and the proposed method was programmed in MATLAB where the mixed integer second order cone programming were solved using CPLEX 12.5.

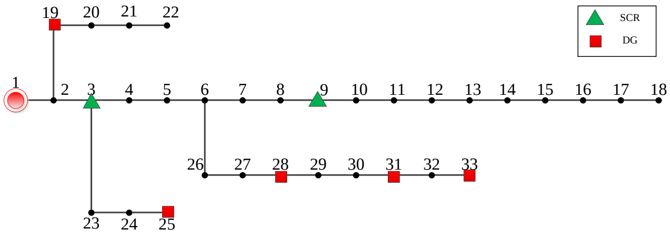

Fig. 1. Topology of 33-bus system

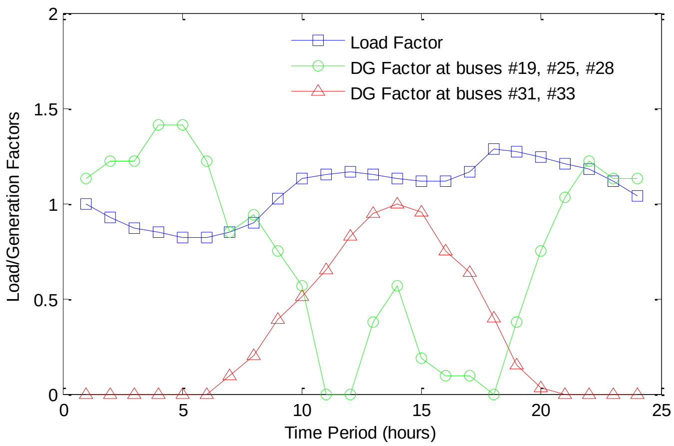

Fig. 2. Load/Generation factors over 24 hours.

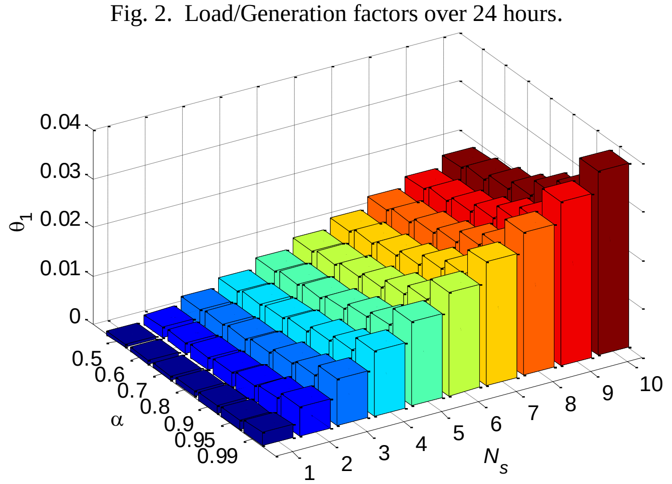

Fig. 3. The relationship among θ_1, N_s and α.

B. Results and Comparison on A 33-Bus Test System

The proposed method is compared with the traditional two popular methods, two-stage stochastic and robust optimization approaches, denoted by 'S' and 'R' respectively. To compare with the traditional stochastic programming model with the deterministic multivariate normal distribution, we solve the traditional model and fix the first-stage decision variables. Then, we randomly choose 10000 different probabilities from the uncertain set and solve the second-stage problem for each given probability, where it is found that the solution with the maximum network loss is served as the worst-case scenario for the stochastic approaches, denoted as 'Swst'.

Furthermore, the comparison of the three methods is presented in Table II. The results show that the two-stage robust optimization method yields the highest network loss (2.8545 MW) and the two-stage stochastic optimization method arrives at the lowest network loss (2.0117 MW). The two-stage robust optimization method optimizes the optimal solution under the worst-case for all the possible realizations, which leads to the largest optimal solution. The two-stage stochastic optimization method neglects the uncertainty of probability of each scenario, which leads to the smallest optimal solution. Moreover, The two methods always yield the same solution for different α. The proposed method under both θ_1 and θ_∞ gives a mild optimal solution and α can be termed as a budget that can control the size of uncertainty set and further affects the optimal solution.

Moreover, the network loss from the worst case of stochastic approach (i.e, Swst) considering uncertain probability distribution is about 20%~30% larger than the traditional two-stage stochastic programming. Increasing confidence level α leads to a larger uncertainty set, so that the worst-case solution will become larger. Comparing the proposed method with the traditional stochastic approach, it can be observed that the network loss from the proposed method under both uncertain sets $\psi_{\inf}$ and ψ_1 is larger than the traditional stochastic programming, while it is smaller than that from the worst case of stochastic approach. In particular, a smaller confidence level α leads to a larger gap between Swst and the proposed method.

Besides, the network loss by the proposed method under different uncertainty sets gives different values, but for the same confidence level α, the optimal solution is very close and the optimal solution under ψ_1 is slightly smaller than that under $\psi_{\inf}$.

Finally, the discrete control actions by the three methods are studied and compared. Take the first SCR for illustration and Fig. 4 depicts that four and eight operating times of SCR1 are obtained by robust and stochastic optimization methods, whereas the proposed method is operated between 4 and 8 times. Here, we only choose α=0.5 and α=0.99 for comparison due to the limited space. It observes that with the increase of α, the optimal control action over 24 hours is closer to that of robust optimization method. This is because the increase of α will enlarge the uncertainty set, which is closer to the uncertainty set of robust optimization approach.

Table II. Comparison of network loss by three methods under different α

α	Network Loss (MW)				
	$\psi_{\inf}$	ψ_1	R	S	Swst
0.5	2.0813	2.0575	2.8545	2.0117	2.3888
0.6	2.1280	2.1076	2.8545	2.0117	2.4305
0.7	2.1720	2.1522	2.8545	2.0117	2.4732
0.8	2.2409	2.2210	2.8545	2.0117	2.5170
0.9	2.3072	2.2941	2.8545	2.0117	2.5618
0.95	2.3702	2.3614	2.8545	2.0117	2.6076
0.99	2.5066	2.5088	2.8545	2.0117	2.6309

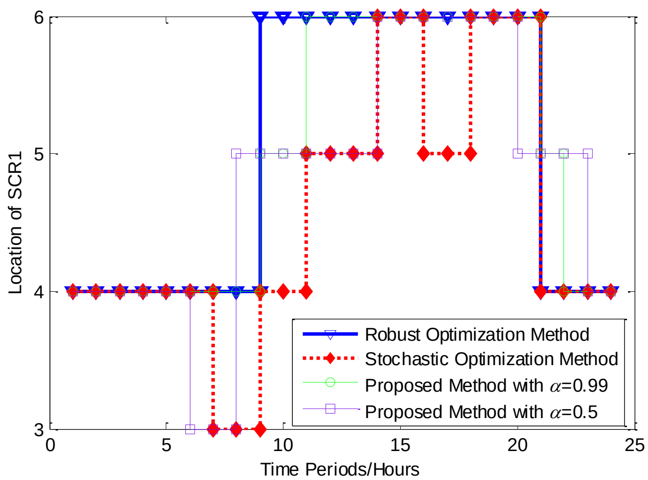

Fig. 4. Comparison of SCR1 on the three methods

C. Comparison of Computational Performance Between the Proposed Method and Traditional Approaches

The comparison of computational performance among the three approaches is shown in Table III, presenting the iterations (Iter.) and computational time (Time) of each method.

For the two-stage stochastic programing model, it needs to solve a large-scale mixed integer second order cone programming, which is actually a single-level model that can be directly handled by the off-the-shelf solvers. However, the computational time increases significantly with the increase of the number of scenarios.

For the two-stage robust optimization model, the number of iterations is only 3, where a new worst-case scenario is identified at each iteration. It is very time-consuming because solving the inner bi-level "max-min" problem needs to take dual and furthermore to solve a large-scale mixed integer second order cone programming model.

In contrast, the proposed data-driven stochastic programming model has a very special structure, in which the feasible region of second-stage problem is disjoint with the uncertainty set, so a new column-and-constraint generation algorithm is proposed to decompose the SP into N_s small-scale SPs that can be solved in parallel. Meanwhile, the SPs are several second order cone programming models, different from the robust optimization model where the SP is a large-scale mixed integer second order cone programming. The computational time can be further reduced significantly. It is observed from Table III that the proposed method is much faster than the two-stage robust optimization method.

Moreover, when N_s is small, two-stage stochastic programming model is a little faster than the proposed method, but with increasing N_s, the two-stage stochastic programming model becomes significantly slower due to the large number of variables and constraints from the scenarios, whereas the computational time of the proposed model increases only slightly thanks to the decomposition method. Therefore, the proposed method performs faster than the two-stage stochastic programming model especially for the case with a large number of scenarios.

Another test system is from a 123-bus test system with 10 DG and five switchable capacitors/reactors (SCRs) connected to bus 12, 35, 54, 76, and 108, which is shown in Fig. 5 . The detailed information can be available from [28]. The comparison of computational performance among the three approaches is shown in Table IV, where it can be observed that the robust optimization needs six iterations for convergence by use of column-and-constraint generation algorithm and the total time is about 12473s. The computational time of the stochastic optimization will increase significantly with increasing the number of scenarios. This is because the stochastic optimization model contains N_s sets of decision variables and constraints. Large N_s will significantly increase the number of total decision variables and constraints and thus need more computational time. As for the proposed method, the computational speed is more than 20 times faster than the robust and stochastic optimization models when N_s is large. Since the increase of N_s will enlarge the uncertainty set. Therefore, it needs more iterations for convergence and the total computational time will increase as well.

Finally, it should be mentioned that the maximum gap of conic relaxation for any test system is smaller than 10^{-4} MW, suggesting that the second order cone programming relaxation is always exact to the original nonconvex model.

Table III. Comparison of computational efficiency by three methods on 33-bus test system

N_s	Proposed		Robust		Stochastic	
	Iter.	Time (s)	Iter.	Time (s)	Iter.	Time (s)
5	3	11.3	3	2897.4	1	8.5
10	3	16.6	3		1	23.4
15	3	23.5	3		1	50.3
20	4	29.9	3		1	156.8
25	4	37.5	3		1	363.7

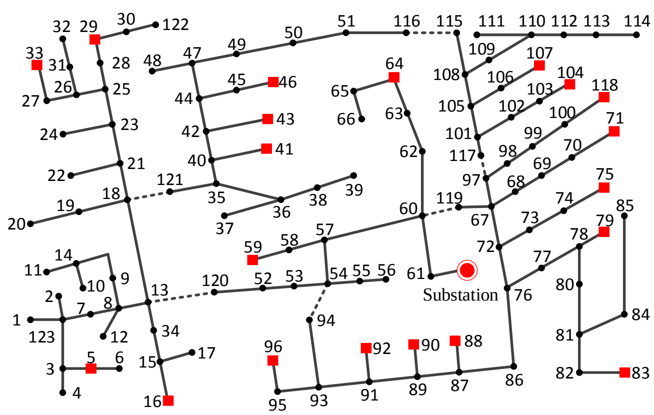

Fig. 5. Topology of 123-bus system

Table IV. Comparison of computational efficiency by three methods on 123-bus test system

N_s	Proposed		Robust		Stochastic	
	Iter.	Time (s)	Iter.	Time (s)	Iter.	Time (s)
5	6	76.5	6	12473	1	125.3
10	7	108.2	6		1	323.4
15	7	113.9	6		1	987.6
20	8	208.8	6		1	2092.1
25	8	235.4	6		1	5754.3

VI. Conclusions

This work proposes a data-driven stochastic reactive power optimization model to address uncertain distributed generators integrated into active distribution networks. According to the historical data, the proposed method constructs a confidence set for the probability distribution of the uncertainties and aims to find an optimal solution under the worst probability distribution. Furthermore, conic relaxation is employed to utilize to relax the feasible region enclosed by power flow equations. It is noted that the proposed model has a special structure, so that a new column-and-constraint generation algorithm is proposed to decompose the second-stage bi-level inner problem into several small-scale subproblems that can be handled in parallel. The comparison with the traditional two-stage stochastic and robust approaches on two test systems shows that the proposed model can achieve better optimal solution and computational performance than traditional methods.

Appendix

As discussed in Section II, the original reactive power optimization model is not a standard mixed integer second order cone programming model since there are still many bilinear terms in the above model. Now, we can simplify them so as to construct a standard mixed integer second order cone programming model.

(i) Reformulations for constraints in (12)-(14) and (18)

The discrete reactive power compensators in (12)-(14) and (18) are nonnegative integers, rather than 0-1 binary variables. For the standard mixed integer programming model, it is expected to formulate the model with binary variables. Therefore, we should reformulate each integer variable $\rho_j(t)$ into a combination of 0-1 binary variables. Since any integer number has a unique binary code, the binary code of $\rho_j(t)$ can be expressed by the combination of binary variables $\lambda_{j,0}(t),\lambda_{j,1}(t),...,\lambda_{j,\tau_j}(t)$ as

$$\rho_j(t)=2^0\lambda_{j,0}(t)+2^1\lambda_{j,1}(t)+...+2^{\tau_j}\lambda_{j,\tau_j}(t) \qquad \text{(A1)}$$

According to the bound constraints in (12) and (13) that $\begin{cases} C_j^{\min}\le C_j(t)\le C_j^{\max} \\ C_j(t)=C_j^{\min}+s_j\rho_j(t)\end{cases}$, we can derive

$$s_j\left(2^0\lambda_{j,0}(t)+2^1\lambda_{j,1}(t)+...+2^{\tau_j}\lambda_{j,\tau_j}(t)\right)\le C_j^{\max}-C_j^{\min} \qquad \text{(A2)}$$

Since $\lambda_{j,0}(t),\lambda_{j,1}(t),...,\lambda_{j,\tau_j}(t)\in\{0,1\}$, the maximum value should be 1. Therefore, the maximum value of τ_j should be

$$\log_2\left(\frac{C_j^{\max}-C_j^{\min}}{s_j}+1\right)-1\le\tau_j\le\log_2\left(\frac{C_j^{\max}-C_j^{\min}}{s_j}+1\right) \qquad \text{(A3)}$$

According to (A1) and (12), $v_j(t)C_j(t)$ becomes

$$\begin{aligned} &v_j(t)C_j(t)\\ &=v_j(t)\left(C_j^{\min}+s_j\rho_j(t)\right)\\ &=C_j^{\min}v_j(t)+s_j\left(2^0\lambda_{j,0}(t)v_j(t)+2^1\lambda_{j,1}v_j(t)+...+2^{\tau_j}\lambda_{j,\tau_j}(t)v_j(t)\right)\end{aligned} \qquad \text{(A4)}$$

For convenience, let $\sigma_{j,k}(t)=\lambda_{j,k}(t)v_j(t)$, (18) derives

$$\begin{aligned} &\frac{1}{2}\left(C_j^{\min}+s_j\left(2^0\sigma_{j,0}(t)+2^1\sigma_{j,1}(t)+...+2^{\tau_j}\sigma_{j,\tau_j}(t)\right)\right)+Q_{c,j}(t)-Q_{L,j}(t)\\ &=\sum_{k\in\delta(j)}G_{jk}(t)-\sum_{i\in\pi(j)}\left(G_{ij}(t)-x_{ij}l_{ij}(t)\right)+b_{s,j}v_j(t)\end{aligned},$$

$$\forall j\in\Omega_D,\ t=1,...,T \qquad \text{(A5)}$$

Furthermore, $\sigma_{j,k}(t)=\lambda_{j,k}(t)v_j(t)$ can be linearized by means of the big-M approach, such that

$$\begin{cases}-M\left(1-\lambda_{j,k}(t)\right)\le\sigma_{j,k}(t)-v_j(t)\le M\left(1-\lambda_{j,k}(t)\right)\\ -M\lambda_{j,k}(t)\le\sigma_{j,k}(t)\le M\lambda_{j,k}(t)\end{cases},$$

$$\forall j\in\Omega_D,\ k=1,..,\tau_j \qquad \text{(A6)}$$

For (A6), since $\lambda_{j,k}(t)\in\{0,1\}$, we can find that

$\lambda_{j,k}(t)=0$ gives $\begin{cases}-M\le\sigma_{j,k}(t)-v_j(t)\le M\\ 0\le\sigma_{j,k}(t)\le 0\end{cases}\Rightarrow\sigma_{j,k}(t)=0$;

$\lambda_{j,k}(t)=1$ gives $\begin{cases}0\le\sigma_{j,k}(t)-v_j(t)\le 0\\ -M\le\sigma_{j,k}(t)\le M\end{cases}\Rightarrow\sigma_{j,k}(t)=v_j(t)$.

Therefore, (A6) is equivalent to $\sigma_{j,k}(t)=\lambda_{j,k}(t)v_j(t)$ and (12)-(14) and (18) can be expressed as (A5) with additional constraints (A6).

Meanwhile, taking (A1) into (14) leads to

$$\sum_{t=2}^{T}\left|\sum_{k=0}^{\tau_j}2^k\left(\lambda_{j,k}(t+1)-\lambda_{j,k}(t)\right)\right|\le\eta_{c,j},\quad\forall j\in\Omega_D \qquad \text{(A7)}$$

(ii) *Reformulations for constraints in (20)*

Similar to the method for linearizing bilinear terms $v_j(t)C_j(t)$, the bilinear terms $o_{ij,k}(t)v_j(t)$ can be also linearized using the big-M approach. Let $h_{ij,k}(t)=o_{ij,k}(t)v_j(t)$ and the constraints (20) containing bilinear terms $o_{ij,k}(t)v_j(t)$ will become

$$\sum_{k=0}^{n_{ij}}\frac{h_{ij,k}(t)}{\left(w_{ij,k}\right)^2}=v_i(t)-2\left(r_{ij}H_{ij}(t)+x_{ij}G_{ij}(t)\right)+\left(r_{ij}^2(t)+x_{ij}^2(t)\right)l_{ij}(t),$$

$$\forall(i,j)\in\Theta,\ t=1,...,T \qquad \text{(A8)}$$

$$\begin{cases}-M\left(1-o_{ij,k}(t)\right)\le h_{ij,k}(t)-v_j(t)\le M\left(1-o_{ij,k}(t)\right)\\ -Mo_{ij,k}(t)\le h_{ij,k}\le Mo_{ij,k}(t)\end{cases},$$

$$\forall(i,j)\in\Theta,\ k=1,..,n_{ij} \qquad \text{(A9)}$$

According to the above reformulations and relaxation in section II, the reactive power optimization model can be cast as a standard 0-1 mixed integer second order cone programming as follows:

$$\min_{Q_c(t),\lambda(t),o(t)} \quad \sum_{t=1}^{T} \sum_{(i,j)\in E} \left(r_{ij} l_{ij}(t) \right) \tag{A10}$$

$s.t$ (3), (8), (10)-(11), (15)-(16), (19), (22)-(23), (A5)-(A9) (A11)

$$\boldsymbol{\lambda}(t) \in \{0,1\},\ \boldsymbol{o}(t) \in \{0,1\},\ \boldsymbol{Q}_c(t) \in Continous \tag{A12}$$

References

[1] S. Corsi, P. Marannino, N. Losignore, G. Moreschini *et al*., "Coordination between the reactive power scheduling function and the hierarchical voltage control of the EHV ENEL system," *IEEE Trans. Power Syst.*, vol. 10, no. 2, pp. 686-694, 1995.

[2] T. Ding, Q. Guo, H. Sun, et al., "A quadratic robust optimization model for automatic voltage control on wind farm side," *IEEE Power and Energy Society General Meeting*, Vancouver, Canada, 2013, pp. 1-5.

[3] T. Ding, Q. Guo, H. Sun, et al., "A robust two-level coordinated static voltage security region for centrally integrated wind farms," *IEEE Trans. Smart Grid*, vol. PP, no. 99, pp. 1-1, 2015.

[4] J. Zhao, L. Jud n Z. Dai, et al., "Voltage stability constrained dynamic optimal reactive power flow bsed on branch-bound and primal-dual interior point method," *Int. J. Electr. Power Energy Syst.*, vol. 7, no.3, pp. 601-607, 2015.

[5] B. Zhao, C. X. Guo, Y. J. Cao, "A multiagent-based particle swarm optimization approach for optimal reactive power dispatch," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 1070-1078, 2005.

[6] T. Ding, R. Bo, F. Li, et al., "A bi-level branch and bound method for economic dispatch with disjoint prohibited zones considering network losses", *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 2841-2855, 2015.

[7] J. Lopez, J. Contreras and J. R. S Mantovani, "Reactive power planning under conditional-value-at-risk assessment using chance-constrained optimization," *IET Gener. Transm. Distrib.*, vol. 9, no. 3, pp. 231-240, 2014.

[8] A. Rabiee and M. Parniani, "Voltage security constrained multi-period optimal reactive power flow using benders and optimality condition decompositions", *IEEE Trans. Power Syst.*, vol. 28, no. 2, 2013.

[9] M. Farivar, and S. H. Low, "Branch Flow Model: Relaxations and Convexification-Part I," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2554-2564, 2013.

[10] M. Farivar, and S. H. Low, "Branch Flow Model: Relaxations and Convexification-Part II," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2565-2572, 2013.

[11] B. Mohamadreza, H. M. Reza and G. Mehrdad, "Second-order cone programming for optimal power flow in VSC-type-AC-DC grids", *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4282-4291, 2013.

[12] T. Ding, S. Liu, Z. Wu and Z. Bie, "Sensitivity-based relaxation and decomposition method to dynamic reactive power optimization considering DGs in active distribution networks", *IET Gener. Transm. Distrib.*,vol. PP, no. 99, pp. 1-1, 2016.

[13] Q. Li, R. Ayyanar and V. Vittal, "Convex optimization for DES planning and operation in radial distribution systems with high penetration of photovoltaic resources," *IEEE Trans. Sustain. Energy*, vol. 7, no.3, 2016.

[14] Z. Tian, W. Wu, B. Zhang and A. Bose, "Mixed-integer second-order cone programming model for VAR optimization and network reconfiguration in active distribution," *IET Gener. Transm. Distrib.*, vol. 10, no. 8, pp. 1938-1946, 2016.

[15] J. Birge and F. Louveaux, "Introduction to stochastic programming", New York, NY,USA: *Springer-Verlag*,1997

[16] A. Papavasiliou, S. Oren and R. O'Neill, "Reserve requirements for wind power integration: a scenario-based stochastic programming framework", *IEEE Trans. Power Syst.*, vol. 255, no. 4, pp.2197-2206, 2011.

[17] C. Sahin, M. Shahidehpour and I. Erkmen, "Allocation of hourly reserve versus demand response for security-constrained scheduling of stochastic wind energy", *IEEE Trans. Sustain. Energy*, vol. 4, no.1, pp.219-228, 2013.

[18] A. Charnes, W. W. Cooper, "Chance-constrained programming", *Manage. Sci.*, vol. 6, no. 1, pp. 73-79, 1959.

[19] L. Roald, S. Misra, T. Krause and G. Andersson, "Corrective control to handle forecast uncertainty: a chance constrained optimal power flow", *IEEE Trans. Power Syst.*, vol. PP, no. 99, pp. 1-1, 2016.

[20] H. Hejazi and H. M. Rad, "Energy storage planning in active distribution grids: a chance-constrained optimization with non-parametric probability functions", *IEEE Trans. Smart Grid*, vol. PP, no. 99, pp. 1-1, 2016.

[21] J. Lopez, D. Pozo, J. Contreras and J. Mantovani, "A multi-objective minimax regret robust VAr planning model", *IEEE Trans. Power Syst.*, vol. PP, no. 99, pp. 1-1, 2016.

[22] R. A. Jabr, "Adjustable robust OPF with renewable energy sources", *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4742-4751, 2013.

[23] L. Zhihuan, Y. Lin and X. Duan, "Non-dominated sorting genetic algorithm-II for robust multi-objective optimal reactive power dispatch", *IET Gener. Transm. Distrib.*, vol. 4, no. 9, pp. 1000-1008, 2010.

[24] X. Fang, F. Li, Y. Wei, R. Azim and Y. Xu, "Reactive power planning under high penetration of wind energy using Benders decomposition", *IET Gener. Transm. Distrib.*, vol. 9, no. 14, pp. 1835-1844, 2015.

[25] T. Niknam, A. Kavousifard and J. Aghaei, "Scenario-based multi-objective distribution feeder reconfiguration considering wind power using adaptive modified particle swarm optimization", *IET Renewable Power Gener.*, vol. 6, no. 4,pp. 236-247, 2012.

[26] H. Zhang and P. Li. "Chance constrained programming for optimal power flow under uncertainty", *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2417-2424, 2011.

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

[28] T. Ding, S. Liu, W. Yuan Z. Bie and B. Zeng, "A two-stage robust reactive power optimization considering uncertain wind power integration in active distribution networks," *IEEE Trans. Sustain. Energy*, vol. 7, no.1, pp. 301-311,2016.

[29] H. Haghighat and B. Zeng, "Distribution system reconfiguration under uncertain load and renewable generation", *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp.2666-2675, 2016.

[30] Z. Wang, Q. Bian, H. Xin and D. Gan, "A distributionally robust co-ordinated reserve scheduling model considering CVaR-based wind power reserve requirements," *IEEE Trans. Sustain. Energy*, vol. 7, no. 2, pp. 625-636, 2016.

[31] P. Xiong, P. Jirutitijaroen and C. Singh, "A distributionally robust optimization model for unit commitment considering uncertain wind power generation," *IEEE Trans. Power Syst.*, vol. PP, no. 99, pp.1-1, 2016.

[32] Y. Zhang, S. Shen and J. L. Mathieu, "Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads," *IEEE Trans. Power Syst.*, vol. PP, no. 99, pp. 1-1, 2016.

[33] W. Wei, F. Liu and S. Mei, "Distributionally robust co-optimization of energy and reserve dispatch," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 289-300, 2016.

[34] Q. Bian, H. Xin, Z. Wang, D. Gan and K. P. Wong, "Distributionally robust solution to the reserve scheduling problem with partial information of wind power," *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2822-2823, 2015.

[35] C. Zhao and Y. Guan, "Data-driven stochastic unit commitment for integrating wind generation", *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2587-2596, 2016.

[36] N. Li, L. Chen, S. H. Low, "Exact convex relaxation of OPF for radial networks using branch flow model," *IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, pp. 7-12, 2012.

[37] M. S. Lobo, L. Vandenberghe, S. Boyd and et al, "Applications of second-order cone programming," *Linear algebra and its applications*, vol. 284, no. 1, pp. 193-228, 1998.

Tao Ding (S'13–M'15) received the B.S.E.E. and M.S.E.E. degrees from Southeast University, Nanjing, China, in 2009 and 2012, respectively, and the Ph.D. degree from Tsinghua University, Beijing, China, in 2015. During 2013~2014, he was a visiting scholar with the Department of Electrical Engineering and Computer Science, The University of Tennessee, Knoxville (UTK), TN, USA. He received the excellent master and doctoral dissertation from Southeast University and Tsinghua University, respectively, and outstanding graduate award of Beijing City. Dr. Ding has published more than 60 technical papers and authored by "Springer Theses" recognizing outstanding Ph.D. research around the world and across the physical sciences – *Power System Operation with Large Scale Stochastic Wind Power Integration*.

He is currently an associate professor in the State Key Laboratory of Electrical Insulation and Power Equipment, the School of Electrical Engineering, Xi'an Jiaotong University. His current research interests include electricity markets, power system economics and optimization methods, and power system planning and reliability evaluation.

QingrunYang received the B.S. degree from the School of Electrical Enginee ring, Xi'an Jiaotong University, Xi'an, China, in 2017. He is currently pursuin g the M.S. degree at Xi'an Jiaotong University. His major research interests in clude power system optimization and demand response.

Yongheng Yang (S'12 - M'15) received the B.Eng. degree in 2009 from Northwestern Polytechnical University, China and the Ph.D. degree in 2014 from Aalborg University, Denmark.

He was a postgraduate with Southeast University, China, from 2009 to 2011. In 2013, he was a Visiting Scholar with Texas A&M University, USA. Since 2014, he has been with the Department of Energy Technology, Aalborg University, where currently he is an Assistant Professor. His research interests are focused on grid integration of renewable energy systems, power converter design, analysis and control, harmonics identification and mitigation, and reliability in power electronics. Dr. Yang has published more than 80 technical papers and co-authored a book – *Periodic Control of Power Electronic Converters* (London, UK: IET, 2017).

Dr. Yang is a Member of the IEEE Power Electronics Society (PELS) Students and Young Professionals Committee, where he serves as the Global Strategy Chair and responsible for the IEEE PELS Students and Young Professionals Activities. He served as a Guest Associate Editor of **IEEE JOURNAL OF EMERGING AND SELECTED TOPICS IN POWER ELECTRONICS**, and has also been invited as a Guest Editor of *Applied Sciences*. He is an active reviewer for relevant top-tier journals.

Cheng Li received the B.S. degree from the School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China, in 2016. He is currently pursuing the M.S. degree at Xi'an Jiaotong University. His major research interests include power system optimization and renewable energy integration.

Zhaohong Bie (M'98–SM'12) received the B.S. and M.S. degrees from the Electric Power Department of Shandong University, Jinan, China, in 1992 and 1994, respectively, and the Ph.D. degree from Xi'an Jiaotong University, Xi'an, China, in 1998. Currently, she is a Professor in the State Key Laboratory of Electrical Insulation and Power Equipment and the School of Electrical Engineering, Xi'an Jiaotong University. Her main interests are power system planning and reliability evaluation, as well as the integration of the renewable energy.

Frede Blaabjerg (S'86 - M'88 - SM'97 - F'03) was with ABB-Scandia, Randers, Denmark, from 1987 to 1988. From 1988 to 1992, he was a Ph.D. Student with Aalborg University, Aalborg, Denmark. He became an Assistant Professor in 1992, an Associate Professor in 1996, and a Full Professor of power electronics and drives in 1998 at Aalborg University. His current research interests include power electronics and its applications such as in wind turbines, PV systems, reliability, harmonics and adjustable speed drives.

Prof. Blaabjerg has received 17 IEEE Prize Paper Awards, the IEEE Power Electronics Society (PELS) Distinguished Service Award in 2009, the EPE-PEMC Council Award in 2010, the IEEE William E. Newell Power Electronics Award in 2014 and the Villum Kann Rasmussen Research Award in 2014. He was an Editor-in-Chief of the **IEEE TRANSACTIONS ON POWER ELECTRONICS** from 2006 to 2012. Prof. Blaabjerg was nominated in 2014 and 2015 by Thomson Reuters to be between the most 250 cited researchers in Engineering in the world.