Data-driven Distributionally Robust Energy-Reserve-Storage Dispatch

Chao Duan, Lin Jiang, Wanliang Fang, Jun Liu,, Shiming Liu

***Abstract*—This paper proposes distributionally robust energy-reserve-storage co-dispatch model and method to facilitate the integration of variable and uncertain renewable energy. The uncertainties of renewable generation forecasting errors are characterized through an ambiguity set which is a set of probability distributions consistent with observed historical data. The proposed model minimizes the expected operation costs corresponding to the worst-case distribution in the ambiguity set. Distributionally robust chance constraints are employed to guarantee reserve and transmission adequacy. The more historical data is available, the smaller the ambiguity set is and the less conservative the solution is. The formulation is finally cast into a mixed integer linear programming whose scale remains unchanged as the number of historical data increases. Inactive constraint identification and convex relaxation techniques are introduced to reduce the computational burden. Numerical results and Monte Carlo simulations on IEEE 118-bus systems demonstrate the effectiveness and efficiency of the proposed method.**

***Index Terms*—economic dispatch, energy storage, distributionally robust optimization, chance constraints, reserve scheduling**

NOMENCLATURE

$\mathcal{B}, \mathcal{L}, \mathcal{T}$	Set of all buses, lines and time periods.
c_{gi}	Generation price of the i_{th} generator.
c^c_{si}/c^d_{si}	Charge/discharge price of the i_{th} storage.
d^+_{gi}/d^-_{gi}	Upward/downward reserve availability price of the i_{th} generator.
d^+_{si}/d^-_{si}	Upward/downward reserve availability price of the i_{th} storage.
f^+_{gi}/f^-_{gi}	Upward/downward reserve utilization price of the i_{th} generator.
f^+_{si}/f^-_{si}	Upward/downward reserve utilization price of the i_{th} storage.
$\underline{P}_{gi}/\overline{P}_{gi}$	Lower/upper limits of output power of the i_{th} generator.
$\overline{R}^+_{gi}/\overline{R}^-_{gi}$	Ramp up/down rate limits of the i_{th} generator.
$\overline{P}^c_{si}/\overline{P}^d_{si}$	Upper limits for the charge/discharge power of the i_{th} storage.
C_l	Capacity of transmission line l.
π^l_{gi}/π^l_{si}	Load shift factor from generator/storage/load i to line l.
π^l_{di}	Load shift factor from load i to line l.
η^c_{si}/η^d_{si}	Charge/discharge efficiency of the i_{th} storage.
$\tilde{\epsilon}_i(t)$	Random forecasting error of composite load at the i_{th} bus in time t.
$P_{di}(t)$	Forecasted composite load at the i_{th} bus in time t.
$\tilde{P}_{di}(t)$	Random composite load at the i_{th} bus in time t.
$P_{gi}(t)$	Output power setting point of the i_{th} generator in time t.
$\tilde{P}_{gi}(t)$	Random actual output power of the i_{th} generator in time t.
$\tilde{S}_l(t)$	Random line power flow on line l in time t.
$\tilde{P}_{si}(t)$	Random actual output power of the i_{th} storage in time t.
$E_{si}(0)$	Initial stored energy of the i_{th} storage.
$\alpha_{gi}(t)$	Participation factor of the i_{th} generator in time t.
$\alpha_{si}(t)$	Participation factor of the i_{th} storage in time t.
$P^c_{si}(t)/P^d_{si}(t)$	Charge/discharge power setting point of the i_{th} storage in time t.
$R^+_{gi}(t)/R^-_{gi}(t)$	Upward/downward reserve of the i_{th} generator in time t.
$R^+_{si}(t)/R^-_{si}(t)$	Upward/downward reserve of the i_{th} storage in time t.
$\mathbb{P}$	A probability measure/distribution.
$\mathbb{E}_{\mathbb{P}}$	Expectation respect to probability measure $\mathbb{P}$.
$\mathcal{P}_0(\mathcal{S})$	Set of all probability measures with support $\mathcal{S}$.
$I_{\mathcal{S}}(\cdot)$	Indicator function of set $\mathcal{S}$, i.e. $I_{\mathcal{S}}(x) = 1$ when $x \in \mathcal{S}$ and 0 otherwise.
$(x)^+$	$\max\{x, 0\}$.

C. Duan, W. Fang and J. Liu are with the Department of Electrical Engineering, Xi'an Jiaotong University, Xi'an 710049, China. C. Duan is also with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, U.K. (e-mail:duanchao@stu.xjtu.edu.cn; eewlfang@mail.xjtu.edu.cn; eeliujun@mail.xjtu.edu.cn)

L. Jiang is with Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, U.K. (e-mail: ljiang@liverpool.ac.uk)

S. Liu is with the Department of Electrical Engineering, Shandong University, Jinan 250002, China. (e-mail: lsm@sdu.edu.cn)

I. INTRODUCTION

Large-scale integration of renewable energy has brought a high level of variability and uncertainty into power system operation, which poses a great challenge to system schedule and dispatch. Energy storage systems (ESS) are recognized as underpinning technologies to meet such challenge due to their ability to provide time-varying energy management and alleviate the intermittence of renewable generation [1]. To minimize operation costs while guarantee system reliability under variability and uncertainty, the operation of ESSs must be integrated into the conventional economic dispatch and reserve scheduling problems.

The investigation of the optimal operation of power system with ESSs requires the multi-period optimal power flow (OPF) models because the operation of ESS is strongly coupled over time by charge/discharge dynamics. Both DC and AC power flow models are applied to this problem. Jabr *et al* developed a robust multi-period DC OPF with ESS to address the uncertainties of renewable generation [2]. DC flow based multi-period OPF is also employed in [3] to optimize storage allocation and portfolio. Since DC power flow neglects voltage magnitude and reactive power, its results might be unreasonable for practical operation. Thus, full AC formulations of multi-period OPF are also introduced to address the optimal operation of distribution networks in [4] and [5]. But due to the non-convexity of the problem formulation, only local optimality is guaranteed. By extending the seminar work of Lavaei and Low [6], semi-definite program (SDP) relaxations of AC flow based multi-period OPF are discussed in [7] where global optimal solutions are achievable in some cases.

Conventionally, spinning reserve is not explicitly handled in economic dispatch but treated separately in the reserve scheduling problem [8], [9]. Due to the ever-increasing level of uncertainty and the commercialization of spinning reserve as an auxiliary service, co-optimization of energy and reserve becomes a trend in recent literature [10]–[12]. When the load and renewable generation deviate from the predicted values, spinning reserve needs to be utilized to maintain real-time power balance. The process of reserve procurement is usually handled in two different approaches. The first is the affinely adjustable approach [2], [13] where a generator uses its reserve according to the associated participation factor. The second is the fully adjustable approach [11], [14] in which the reserve procurement is treated as a sub-level optimization problem after the realization of uncertainties. The advantages of affinely adjustable approach are its compatibility with existing automatic generation systems (AGC) and the numerical tractability of the optimization model. Nevertheless, the affine policy is more restrictive thus brings some conservatism compared with full recourse strategy.

Besides above-mentioned aspects related to problem formulation, a more prominent issue is how to deal with uncertainties. Stochastic programming (SP) [15], [16], robust optimization (RO) [2], [10], [11], [13], [14], [17], [18] and distributionally robust optimization (DRO) [9], [12], [19], [20] have been employed to tackle uncertainties in power system operation. SP assumes operational uncertainties follow a pre-specified probability distribution and characterizes the uncertainties by scenarios sampled from that distribution. In contrast to SP, RO does not require any probabilistic information of the uncertainties. Instead, randomness is represented by a deterministic uncertainty set, and RO seeks strategies that are immune against all realizations of the uncertainty set. In practice, the probability distribution of uncertainties truly exists but must be estimated from historical data and is therefore itself uncertain. To bridge the gap between the specificity of SP and conservatism of RO, DRO assumes that the true distribution lies in an ambiguity set and immunizes the operation strategies against all distributions in the ambiguity set. The ambiguity set employed in [9], [12], [19], [20] is the set of all probability distributions sharing given mean and covariance. Paper [19] further assumes the unimodality of the distribution to reduce conservatism. However, the DRO approach discussed in [9], [12], [19] has the following drawbacks. First, the ambiguity set characterized only by the first two moments is in fact very large thus the method is still very conservative. Second, moments also need to be estimated from historical data therefore are uncertain as well. Third, the problems are finally cast into a sequence of semi-definite programmings which are very computationally intensive.

In this paper, we propose novel formulation and method for co-optimization of energy, reserve, and storage under the spirit of DRO [21]. The contributions are threefold:

1) **Problem formulation:** We extend the robust multi-period OPF formulation [2] to co-dispatch of energy, reserve, and storage. Distributionally robust chance constraints (DRCC) are employed to provide explicit reliability guarantee for reserve and transmission adequacy. The objective is to minimize the expected operation costs w.r.t. the worst-case distribution in the constructed ambiguity set, which provides robustness for economical system operation.
2) **Data-driven and data-exploiting features:** The proposed method is data-driven in the sense that the ambiguity set for DRO is constructed from historical data without any prior knowledge about the distribution. The method can automatically extract and exploit the probabilistic information contained in the data set. The more historical data is available, the less conservative the solution is.
3) **Efficient solution approach:** The problem is finally formulated as a mixed integer linear programming (MILP) for which off-the-shelf solvers are available. The scale of the MILP remains unchanged as the number of available data increases. Computational issues, including the elimination of inactive line capacity constraints and convex relaxation of binary variables, are considered in the solution approach to significantly improve the numerical tractability.

II. Distributionally Robust Optimization

A. Basic Concepts

In power system, some operational strategy x is needed to minimize the cost function $f(x,\tilde{\xi})$ while satisfy the technical and security constraints $g(x,\tilde{\xi}) \leq 0$ where both the cost function and constraints are affected by some uncertainty represented by random variable $\tilde{\xi}$. In practice, the probability distribution of $\tilde{\xi}$ is unknown and only some historical data is available. Theoretically, the precise description of the probability distribution cannot be obtained from finite sample data. Therefore, the probability distribution itself is uncertain. However, the historical data does provide us some reliable information about the distribution, and based on these information we can construct an ambiguity set $\mathcal{P}$, i.e. a set of probability distributions consistent with the observed historical data. Hence the *distributionally robust optimization* seeks decisions that are immune against all distributions and perform best in view of the worst-case distribution from the ambiguity set, i.e.

$$\begin{aligned} &\min_{x} \max_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[f(x,\tilde{\xi})] \\ &\text{s.t. } \mathbb{P}[g(x,\tilde{\xi}) \leq 0] \geq 1-\beta, \forall \mathbb{P}\in\mathcal{P}. \end{aligned} \tag{1}$$

The performance and numerical tractability of the above problem largely rely on the structure of the ambiguity set. A desirable ambiguity set should possess the following properties: 1) $\mathcal{P}$ contains the underlying true probability distribution; 2) to reduce conservatism, $\mathcal{P}$ can be made as small as possible by incorporating more observed data; 3) the structure of $\mathcal{P}$

allows the reformulation of distributionally robust optimization problem (1) into tractable deterministic problem.

B. Construction of Ambiguity Set

In this subsection, we provide one approach to construct the ambiguity set $\mathcal{P}$ based on confidence bands for cumulative distribution function (CDF) from non-parametric statistics. Consider a one-dimensional random variable $\tilde{\xi}$ whose probability distribution is unknown whereas the ascendingly ordered sample set $S = \{\hat{\xi}^{(1)}, \hat{\xi}^{(2)}, \cdots, \hat{\xi}^{(n)}\}$ is available. Let $F(x) = \mathbb{P}^*\{\xi \leq x\}$ be the CDF of true distribution $\mathbb{P}^*$. The $1-\alpha$ confidence bands for $F(x)$ is a pair of sample-dependent functions $\underline{P}(x)$ and $\overline{P}(x)$ for which $\underline{P}(x) \leq F(x) \leq \overline{P}(x),\ \forall x \in \mathbb{R}$ with probability $1-\alpha$ over the choice of sample set S. One approach to obtain the confidence bands for $F(x)$ is the Dirichlet method [22] summarized blow:

***Lemma** 1:* Let $S = \{\hat{\xi}^{(1)}, \hat{\xi}^{(2)}, \cdots, \hat{\xi}^{(n)}\}$ be the ascendingly ordered sample set of random variable ξ generated independently according to true distribution $\mathbb{P}^*$ with continuous CDF $F(x)$. $B^{\alpha}_{k,n}$ denote the α-quantile of the $\beta(k, n+1-k)$ distribution. For given n and α, define $\underline{p}_k = B^{\tilde{\alpha}/2}_{k,n}$ and $\overline{p}_k = B^{1-\tilde{\alpha}/2}_{k,n}$ where

$$\tilde{\alpha} = \exp\left(-c_1(\alpha) - c_2(\alpha)\sqrt{\ln[\ln(n)]} - c_3(\alpha)[\ln(n)]^{c_4(\alpha)}\right) \tag{2}$$

with $c_1(\alpha) = -2.75 - 1.04\ln(\alpha)$, $c_2(\alpha) = 4.76 - 1.20\alpha$, $c_3(\alpha) = 1.15 - 2.39\alpha$, and $c_4(\alpha) = -3.96 + 1.72\alpha^{0.171}$. Add $\hat{\xi}^{(0)} = -\infty$ and $\hat{\xi}^{(n+1)} = \infty$ to the ascending sequence of the sample set S, and define $\underline{p}_0 = 0$ and $\overline{p}_{n+1} = 1$. Then

$$\underline{P}(x) = \max\{\underline{p}_k : \hat{\xi}^{(k)} \leq x\} \tag{3}$$

$$\overline{P}(x) = \min\{\overline{p}_k : \hat{\xi}^{(k)} \leq x\} \tag{4}$$

are the $1-\alpha$ confidence bands for $F(x)$.

Note that $\underline{P}(x)$ and $\overline{P}(x)$ have the following properties: 1) they are stair-step functions that take values $\underline{p}_k$ and $\overline{p}_k$ at $\hat{\xi}^{(k)}$, respectively; 2) the empirical CDF $\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{\hat{\xi}^{(i)} \leq x\}}$ is lower and upper bounded by $\underline{P}(x)$ and $\overline{P}(x)$, i.e. $\underline{P}(x) \leq \hat{F}(x) \leq \overline{P}(x)$; 3) as the size of the sample set $n \to \infty$, $\sup|\underline{P}(x) - \overline{P}(x)| \to 0$. In other words, $\underline{P}(x)$ and $\overline{P}(x)$ represent the reliable information that can be extracted from finite samples and the information becomes more and more accurate as the size of the sample set grows.

Based on the confidence bands for CDF, the ambiguity set $\mathcal{P}$ employed in this paper takes the form

$$\mathcal{P} = \left\{\mathbb{P} \in \mathcal{P}_0([\underline{\xi}, \overline{\xi}]) \middle| \mathbb{P}[\tilde{\xi} \leq \hat{\xi}^{(k)}] \in [\underline{p}_k, \overline{p}_k],\ k = 1, \cdots, n\right\} \tag{5}$$

where $\mathcal{P}_0([\underline{\xi}, \overline{\xi}])$ denotes the set of all probability measures whose supports are the interval $[\underline{\xi}, \overline{\xi}]$. The proposed **CDF-based ambiguity set** $\mathcal{P}$ is designed to encode the information from confidence bands for CDF and does not assume any prior knowledge about the distribution type. Due to the convergence property of the confidence bands, the ambiguity set $\mathcal{P}$ is made smaller and smaller by incorporating more and more historical data. Moreover, the structure defined in (5) allows very efficient reformulation of distrituionally robust optimization problems, which will be analyzed in section IV.

III. Problem Formulation

The problem formulation of distributionally robust energy-reserve-storage dispatch proposed in this paper is directly extended from Jabr's robust multi-period OPF with storage and renewables [2]. The extension is made from two aspects. First, spinning reserve is explicitly handled in the formulation, and the availability and utilization costs of the spinning reserve are thus reflected in the objective function. Second, instead of robust optimization, the problem is formulated as a distributionally robust optimization problem as (1).

The uncertainties of system operation mainly originate from the the load and renewable forecasting errors uniformly represented by the forecasting errors of composite load:

$$\tilde{P}_{di}(t) = P_{di}(t) + \tilde{\epsilon}_i(t). \tag{6}$$

To maintain real-time power balance under uncertainties, the proposed formulation inherits the affinely adjustable approach from [2] where both the conventional generators and the energy storages participate in the frequency regulation according to the associated participation factors:

$$\tilde{P}_{gi}(t) = P_{gi}(t) + \alpha_{gi}(t)\sum_{k\in\mathcal{B}}\tilde{\epsilon}_k(t), \tag{7}$$

$$\tilde{P}_{si}(t) = P^d_{si}(t) - P^c_{si}(t) + \alpha_{si}(t)\sum_{k\in\mathcal{B}}\tilde{\epsilon}_k(t). \tag{8}$$

Although the affinely adjustable approach [2], [13], [17], [18] is only conservative approximation to the fully adjustable approach [12], [14], affine policy is directly compatible to the AGC and numerically more tractable. To reduce the dimension of the random variables, we further define

$$\tilde{\phi}_t = \sum_{k\in\mathcal{B}}\tilde{\epsilon}_k(t) \tag{9}$$

$$\tilde{\theta}_{lt} = \sum_{k\in\mathcal{B}}\pi^l_{dk}\tilde{\epsilon}_k(t). \tag{10}$$

Under above definitions, we have $\tilde{P}_{gi}(t) = P_{gi}(t) + \alpha_{gi}(t)\tilde{\phi}_t$ and $\tilde{P}_{si}(t) = P^d_{si}(t) - P^c_{si}(t) + \alpha_{si}(t)\tilde{\phi}_t$. Furthermore, by using equation (6)$\sim$(10), the power flow on line l at time t takes the form as in equation (11).

Let $\mathbb{P}^l_t$ denote the probability distribution for 2-dimensional random variable $(\tilde{\phi}_t, \tilde{\theta}_{lt})$ with marginal distributions $\mathbb{P}^{\phi}_t$ and $\mathbb{P}^{\theta^l}_t$. Using the historical data of $\tilde{\phi}_t$ and $\tilde{\theta}_{lt}$, we can construct the ambiguity sets as defined in (5) for $\mathbb{P}^{\phi}_t$ and $\mathbb{P}^{\theta_l}_t$, denoted as $\mathcal{P}^{\phi}_t$ and $\mathcal{P}^{\theta_l}_t$, respectively. Then the ambiguity set of joint distribution $\mathbb{P}^l_t$ is defined as $\mathcal{P}^l_t = \left\{\mathbb{P}^l_t \in \mathcal{P}_0(\mathcal{R}^2) \middle| \mathbb{P}^{\phi}_t \in \mathcal{P}^{\phi}_t, \mathbb{P}^{\theta_l}_t \in \mathcal{P}^{\theta_l}_t\right\}$.

Under the affinely adjustable framework described above, the real-time power balance is guaranteed by

$$\sum_{i\in\mathcal{G}}P_{gi}(t) + \sum_{i\in\mathcal{S}}(P^d_{si}(t) - P^c_{si}(t)) = \sum_{i\in\mathcal{B}}P_{di}(t) \tag{12}$$

$$\sum_{i\in\mathcal{G}}\alpha_{gi}(t) + \sum_{i\in\mathcal{S}}\alpha_{si}(t) = 1 \tag{13}$$

$$\alpha_{gi}(t) \geq 0,\ \alpha_{si}(t) \geq 0, \tag{14}$$

$$\tilde{S}_l(t) = \sum_{i\in\mathcal{G}} \pi_{gi}^l \tilde{P}_{gi}(t) + \sum_{i\in\mathcal{S}} \pi_{si}^l \tilde{P}_{si}(t) - \sum_{i\in\mathcal{B}} \pi_{di}^l \tilde{P}_{di}(t) \tag{11a}$$

$$= \sum_{i\in\mathcal{G}} \pi_{gi}^l (P_{gi}(t) + \alpha_{gi}(t)\tilde{\phi}_t) + \sum_{i\in\mathcal{S}} \pi_{si}^l (P_{si}^d(t) - P_{si}^c(t) + \alpha_{si}(t)\tilde{\phi}_t) - \sum_{i\in\mathcal{B}} \pi_{di}^l P_{di}(t) - \tilde{\theta}_{lt} \tag{11b}$$

where the setting values of generator power, storage charge/discharge power and generator/storage participator factors need to be dynamically adjusted to minimize costs while ensure system security. The followed constraints (15)∼(17)

$$\underline{P}_{gi} + R_{gi}^-(t) \le P_{gi}(t) \le \overline{P}_{gi} - R_{gi}^+(t) \tag{15}$$

$$R_{gi}^-(t) \ge 0,\ R_{gi}^+(t) \ge 0 \tag{16}$$

$$\begin{aligned} -\overline{R}_{gi}^- + R_{gi}^-(t+1) + R_{gi}^+(t) &\le P_{gi}(t+1) - P_{gi}(t) \\ &\le \overline{R}_{gi}^+ - R_{gi}^+(t+1) - R_{gi}^-(t). \end{aligned} \tag{17}$$

together ensure the procurability of upward and downward spinning reserve from conventional generators considering generator capacity limits (15) and ramp rate limits (17). By introducing binary decision variable $\omega_i(t)$, the following constraints

$$0 \le P_{si}^c(t) \le \overline{P}_{si}^c \omega_i(t) \tag{18a}$$

$$0 \le P_{si}^d(t) \le \overline{P}_{si}^d (1 - \omega_i(t)) \tag{18b}$$

$$\omega_i(t) \in \{0,1\},\ i \in \mathcal{S},\ t \in \mathcal{T} \tag{18c}$$

not only set the storage charge/discharge power limits but also avoid simultaneous charging and discharging. The followed constraints (19)∼(22)

$$-\overline{P}_{si}^c + R_{si}^-(t) \le P_{si}^d(t) - P_{si}^c(t) \le \overline{P}_{si}^d - R_{si}^+(t) \tag{19}$$

$$R_{si}^-(t) \ge 0,\ R_{si}^+(t) \ge 0 \tag{20}$$

$$E_{si}(0) + \sum_{\tau=1}^{t} \left(\eta_{si}^c P_{si}^c(\tau) - \frac{1}{\eta_{si}^d} P_{si}^d(\tau) + \eta_{si}^c R_{si}^-(\tau) \right) \Delta t \le \overline{E}_{si} \tag{21}$$

$$E_{si}(0) + \sum_{\tau=1}^{t} \left(\eta_{si}^c P_{si}^c(\tau) - \frac{1}{\eta_{si}^d} P_{si}^d(\tau) - \frac{1}{\eta_{si}^d} R_{si}^+(\tau) \right) \Delta t \ge \underline{E}_{si} \tag{22}$$

together guarantee the procurability of spinning reserve from energy storages by considering the charge/discharge power limits (19) and the stored energy upper/lower limits (21)(22). The following cycling constraint

$$E_{si}(0) + \sum_{\tau=1}^{T} \left(\eta_{si}^c P_{si}^c(\tau) - \frac{1}{\eta_{si}^d} P_{si}^d(\tau) \right) \Delta t = E_{si}(0) \tag{23}$$

sets the final stored energy to be the initial values.

In addition, the adequacy of downward/upward spinning reserve and line capacity is ensured by distributionally robust chance constraint (DRCC)

$$\mathbb{P}_t^\phi \begin{bmatrix} -R_{gi}^-(t) \le \alpha_{gi}(t)\tilde{\phi}_t,\ \forall i \in \mathcal{G} \\ -R_{si}^-(t) \le \alpha_{si}(t)\tilde{\phi}_t,\ \forall i \in \mathcal{S} \end{bmatrix} \ge 1 - \beta_1, \forall \mathbb{P}_t^\phi \in \mathcal{P}_t^\phi \tag{24}$$

$$\mathbb{P}_t^\phi \begin{bmatrix} \alpha_{gi}(t)\tilde{\phi}_t \le R_{gi}^+(t),\ \forall i \in \mathcal{G} \\ \alpha_{si}(t)\tilde{\phi}_t \le R_{si}^+(t),\ \forall i \in \mathcal{S} \end{bmatrix} \ge 1 - \beta_2, \forall \mathbb{P}_t^\phi \in \mathcal{P}_t^\phi \tag{25}$$

and

$$\mathbb{P}_t^l \begin{bmatrix} -C_l \le \sum_{i\in\mathcal{G}} \pi_{gi}^l (P_{gi}(t) + \alpha_{gi}(t)\tilde{\phi}_t) \\ + \sum_{i\in\mathcal{S}} \pi_{si}^l (P_{si}^d(t) - P_{si}^c(t) + \alpha_{si}(t)\tilde{\phi}_t) \\ - \sum_{i\in\mathcal{B}} \pi_{di}^l P_{di}(t) - \tilde{\theta}_{lt} \le C_l \end{bmatrix} \begin{matrix} \ge 1 - \gamma, \\ \forall \mathbb{P}_t^l \in \mathcal{P}_t^l \end{matrix} \tag{26}$$

where the parameter β_1, β_2 and γ are pre-specified allowable probability for renewable curtailment, load shedding and transmission line overload, respectively. In order for (24)∼(24) to be feasible, it requires $\gamma \ge \beta_1 + \beta_2$. We will shed more light on the DRCC (24)∼(26) by looking at their deterministic counterparts in section-IV-A.

The objective function consists of the costs of conventional generation (27a), storage charge/discharge (27b), upward/downward reserve availability of generators (27c), upward/downward reserve availability of storages (27d) and the expected costs of reserve utilization (27e), formally stated as

$$F(\boldsymbol{P}_g, \boldsymbol{P}_s^d, \boldsymbol{P}_s^c, \boldsymbol{R}_g^+, \boldsymbol{R}_g^-, \boldsymbol{R}_s^+, \boldsymbol{R}_s^-, \boldsymbol{\alpha}_g, \boldsymbol{\alpha}_s)$$

$$= \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{G}} c_{gi} P_{gi}(t) \tag{27a}$$

$$+ \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{S}} (c_{si}^d P_{si}^d(t) + c_{si}^c P_{si}^c(t)) \tag{27b}$$

$$+ \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{G}} (d_{gi}^+ R_{gi}^+(t) + d_{gi}^- R_{gi}^-(t)) \tag{27c}$$

$$+ \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{S}} (d_{si}^+ R_{si}^+(t) + d_{si}^- R_{si}^-(t)) \tag{27d}$$

$$+ \sum_{t\in\mathcal{T}} \max_{\mathbb{P}_t^\phi \in \mathcal{P}_t^\phi} \mathbb{E}_{\mathbb{P}_t^\phi} [Q_t(\boldsymbol{\alpha}_g(t), \boldsymbol{\alpha}_s(t), \tilde{\phi}(t))]. \tag{27e}$$

In light of (1), the expectation in (27e) is evaluated w.r.t. the worst-case distribution in the ambiguity set, which guarantees that the obtained strategy can perform well in the absence of precise knowledge about the underlying true probability distribution. The explicit formula for reserve utilization costs $Q_t(\boldsymbol{\alpha}_g(t), \boldsymbol{\alpha}_s(t), \tilde{\phi}_t)$ is given by

$$\begin{aligned} & Q_t(\boldsymbol{\alpha}_g(t), \boldsymbol{\alpha}_s(t), \tilde{\phi}_t) = \\ & \sum_{i\in\mathcal{G}} \Big(f_{gi}^{+}(\alpha_{gi}(t)\text{min}\{\tilde{\phi}_t, \overline{\phi}_t'\})^{+} + f_{gi}^{-}(-\alpha_{gi}(t)\text{max}\{\tilde{\phi}_t, \underline{\phi}_t'\})^{+} \Big) \\ & + \sum_{i\in\mathcal{S}} \Big(f_{si}^{+}(\alpha_{si}(t)\text{min}\{\tilde{\phi}_t, \overline{\phi}_t'\})^{+} + f_{si}^{-}(-\alpha_{si}(t)\text{max}\{\tilde{\phi}_t, \underline{\phi}_t'\})^{+} \Big) \end{aligned} \tag{28}$$

where $[\underline{\phi}', \overline{\phi}']$, that will be analyzed in section-IV-A, is the dispatchable range of total forecasting error determined by the DRCC (24)∼(26).

To sum up, the proposed distributionally robust energy-reserve-storage dispatch (ERSD) problem takes the form:

$$\begin{aligned} \min \ & F(\boldsymbol{P}_g, \boldsymbol{P}_s^d, \boldsymbol{P}_s^c, \boldsymbol{R}_g^{+}, \boldsymbol{R}_g^{-}, \boldsymbol{R}_s^{+}, \boldsymbol{R}_s^{-}, \boldsymbol{\alpha}_g, \boldsymbol{\alpha}_s) \\ \text{s.t.} \ & (12) \sim (26). \end{aligned} \tag{29}$$

IV. Solution Approach

The major obstacles to solve problem (29) are the DRCCs (24)∼(26) and the worst-case expectation in the objective function (27). In this section, we show the DRCCs can be replaced by some deterministic linear constraints and the worst-case expectation can be evaluated by a linear programming. Therefore, the problem (29) can be cast into a MILP.

A. Deterministic Reformulation of Distributionally Robust Chance Constraints

Let $\underline{\Phi}_t(x)$ and $\overline{\Phi}_t(x)$ be the confidence bands for the CDF of random variable $\tilde{\phi}_t$, and further define $\underline{\phi}_t' = \overline{\Phi}_t^{-1}(\beta_1)$ and $\overline{\phi}_t' = \underline{\Phi}_t^{-1}(1-\beta_2)$. Then DRCC (24) is satisfied if

$$\begin{cases} -R_{gi}^{-}(t) \le \alpha_{gi}(t)\underline{\phi}_t', \ \forall i \in \mathcal{G} \\ -R_{si}^{-}(t) \le \alpha_{si}(t)\underline{\phi}_t', \ \forall i \in \mathcal{S}, \end{cases} \tag{30}$$

which can be easily seen from $\mathbb{P}_t^{\phi}[\tilde{\phi}_t < \underline{\phi}_t'] \le \overline{\Phi}_t(\underline{\phi}_t') = \beta_1$, $\forall \mathbb{P}_t^{\phi} \in \mathcal{P}_t^{\phi}$. Similarly, DRCC (25) is satisfied if

$$\begin{cases} \alpha_{gi}(t)\overline{\phi}_t' \le R_{gi}^{+}(t), \ \forall i \in \mathcal{G} \\ \alpha_{si}(t)\overline{\phi}_t' \le R_{si}^{+}(t), \ \forall i \in \mathcal{S}, \end{cases} \tag{31}$$

since $\mathbb{P}_t^{\phi}[\tilde{\phi}_t > \overline{\phi}_t'] \le 1 - \underline{\Phi}_t(\overline{\phi}_t') = \beta_2$. In addition, let $\underline{\Theta}_t^l(x)$ and $\overline{\Theta}_t^l(x)$ be the confidence bands for the CDF of random variable $\tilde{\theta}_{lt}$, and $\underline{\theta}_{lt}' = (\overline{\Theta}_t^l)^{-1}(\frac{\gamma-\beta_1-\beta_2}{2})$ and $\overline{\theta}_{lt}' = (\underline{\Theta}_t^l)^{-1}(1-\frac{\gamma-\beta_1-\beta_2}{2})$. Then we have $\forall \mathbb{P}_t^l \in \mathcal{P}_t^l$,

$$\begin{aligned} & \mathbb{P}_t^l \left[\tilde{\phi}_t \notin [\underline{\phi}_t', \overline{\phi}_t'] \text{ or } \tilde{\theta}_{lt} \notin [\underline{\theta}_{lt}', \overline{\theta}_{lt}'] \right] \\ \le & \mathbb{P}_t[\tilde{\phi}_t < \underline{\phi}_t'] + \mathbb{P}_t[\tilde{\phi}_t > \overline{\phi}_t'] + \mathbb{P}_t[\tilde{\theta}_{lt} < \underline{\theta}_{lt}'] + \mathbb{P}_t[\tilde{\theta}_{lt} > \overline{\theta}_{lt}'] \\ = & \mathbb{P}_t^{\phi}[\tilde{\phi}_t < \underline{\phi}_t'] + \mathbb{P}_t^{\phi}[\tilde{\phi}_t > \overline{\phi}_t'] + \mathbb{P}_t^{\theta_l}[\tilde{\theta}_{lt} < \underline{\theta}_{tl}'] + \mathbb{P}_t^{\theta_l}[\tilde{\theta}_{lt} > \overline{\theta}_{lt}'] \\ \le & \overline{\Phi}_t(\underline{\phi}_t') + 1 - \underline{\Phi}_t(\overline{\phi}_t') + \overline{\Theta}_t^l(\underline{\theta}_{lt}') + 1 - \underline{\Theta}_t^l(\overline{\theta}_{lt}') = \gamma. \end{aligned} \tag{32}$$

Therefore, DRCC (26) is satisfied if the line overload does not happen when $(\tilde{\phi}_t, \tilde{\theta}_{lt})$ takes values at the vertices of polyhedron $\{(\tilde{\phi}_t, \tilde{\theta}_{lt}) | \tilde{\phi}_t \in [\underline{\phi}_t', \overline{\phi}_t'] \text{ and } \tilde{\theta}_{lt} \in [\underline{\theta}_{lt}', \overline{\theta}_{lt}']\}$, which can be further written as deterministic constraint (33) by noticing $\underline{\theta}_{lt}' < \overline{\theta}_{lt}'$.

As revealed in deterministic constraints (30), (31) and (33), the system can safely respond to random variable $\tilde{\phi}_t$ in the range of $[\underline{\phi}', \overline{\phi}']$ which we call the dispatchable range of total forecasting error. To ensure the system security, the system operator resorts to load shedding when $\tilde{\phi}_t$ exceeds $\overline{\phi}_t'$ and renewable curtailment when $\tilde{\phi}_t$ goes below $\underline{\phi}_t'$. Therefore, the reserve utilization costs (28) are the saturating linear function of total forecasting error $\tilde{\phi}_t$.

B. Evaluation of Worst-case Expectation

In the objective function (27), we need to evaluate the worst-case expectation of the reserve utilization costs which is a piece-wise linear function of the decision variables. The structure of the ambiguity set (5) allows the reformulation of the worst-case expectation as a LP, formally stated as

***Lemma** 2:* Let $\hat{\phi}_t^{(1)}, \hat{\phi}_t^{(2)}, \cdots, \hat{\phi}_t^{(n)}$ be the ascendingly ordered samples of random variable $\tilde{\phi}_t$. Without loss of generality, assume $\hat{\phi}_t^{(k)} \le 0, \forall k \le m$ and $\hat{\phi}_t^{(k)} > 0, \forall k > m$. The ambiguity set $\mathcal{P}_t^{\phi}$ is constructed as in (5), i.e. $\mathcal{P}_t^{\phi} = \Big\{ \mathbb{P} \in \mathcal{P}_0([\underline{\phi}_t, \overline{\phi}_t]) \Big| \mathbb{P}[\tilde{\phi}_t \le \hat{\phi}_t^{(k)}] \in [\underline{p}_t^k, \overline{p}_t^k], k = 1, \cdots, n \Big\}$. For notational convenience, let $\hat{\phi}_t^{(0)} = \underline{\phi}_t$, $\hat{\phi}_t^{(n+1)} = \overline{\phi}_t$ and $\underline{p}_t^{n+1} = \overline{p}_t^{n+1} = 1$. The worst-case expectation in (27) is equal to the optimum of the following LP [23]:

$$\begin{aligned} & \max_{\mathbb{P}_t^{\phi} \in \mathcal{P}_t^{\phi}} \mathbb{E}_{\mathbb{P}_t^{\phi}}[Q_t(\boldsymbol{\alpha}_g(t), \boldsymbol{\alpha}_s(t), \tilde{\phi}(t))] \\ = & \mathop{\arg\min}_{\substack{\underline{\lambda}_t^k, \overline{\lambda}_t^k \\ k=m-1,\cdots,m+2}} \sum_{k=m-1}^{m+2} (\overline{\lambda}_t^k \overline{p}_t^k - \underline{\lambda}_t^k \underline{p}_t^k) \\ & + \sum_{k=1}^{m-2} (f_t^{dn}(\hat{\phi}_t^{(k-1)}) - f_t^{dn}(\hat{\phi}_t^{(k)}))\overline{p}_t^k \\ & + \sum_{k=m+3}^{n} (f_t^{up}(\hat{\phi}_t^{(k)}) - f_t^{up}(\hat{\phi}_t^{(k+1)}))\underline{p}_t^k + f_t^{up}(\hat{\phi}_t^{(n+1)})\overline{p}_t^{n+1} \\ & \text{s.t.} \begin{cases} \underline{\lambda}_t^k \ge 0, \ \overline{\lambda}_t^k \ge 0, k = m-1, \cdots, m+2 \\ \sum_{i=k}^{m+2} (\overline{\lambda}_t^i - \underline{\lambda}_t^i) + f_t^{up}(\hat{\phi}_t^{(m+3)}) \ge f_t^{dn}(\hat{\phi}_t^{(k-1)}), \\ \qquad\qquad k = m-1, m, m+1 \\ \sum_{i=k}^{m+2} (\overline{\lambda}_t^i - \underline{\lambda}_t^i) + f_t^{up}(\hat{\phi}_t^{(m+3)}) \ge f_t^{up}(\hat{\phi}_t^{(k)}), \\ \qquad\qquad k = m+1, m+2 \end{cases} \end{aligned} \tag{34}$$

where

$$f_t^{up}(\tilde{\phi}_t) = g_t^{up}\tilde{\phi}_t - g_t^{up}(\tilde{\phi}_t - \overline{\phi}_t')^{+} \tag{35}$$

$$f_t^{dn}(\tilde{\phi}_t) = -g_t^{dn}\tilde{\phi}_t - g_t^{dn}(-\tilde{\phi}_t + \underline{\phi}_t')^{+} \tag{36}$$

$$g_t^{up} = \sum_{i\in\mathcal{G}} f_{gi}^{+}\alpha_{gi}(t) + \sum_{i\in\mathcal{S}} f_{si}^{+}\alpha_{si}(t) \tag{37}$$

$$g_t^{dn} = \sum_{i\in\mathcal{G}} f_{gi}^{-}\alpha_{gi}(t) + \sum_{i\in\mathcal{S}} f_{si}^{-}\alpha_{si}(t). \tag{38}$$

$$\sum_{i\in\mathcal{G}}\pi_{gi}^l(P_{gi}(t)+\alpha_{gi}(t)\underline{\phi}_t')+\sum_{i\in\mathcal{S}}\pi_{si}^l(P_{si}^d(t)-P_{si}^c(t)+\alpha_{si}(t)\underline{\phi}_t')-\sum_{i\in\mathcal{B}}\pi_{di}^lP_{di}(t)-\underline{\theta}_{lt}'\le C_l,\ \forall l\in\mathcal{L} \tag{33a}$$

$$\sum_{i\in\mathcal{G}}\pi_{gi}^l(P_{gi}(t)+\alpha_{gi}(t)\overline{\phi}_t')+\sum_{i\in\mathcal{S}}\pi_{si}^l(P_{si}^d(t)-P_{si}^c(t)+\alpha_{si}(t)\overline{\phi}_t')-\sum_{i\in\mathcal{B}}\pi_{di}^lP_{di}(t)-\underline{\theta}_{lt}'\le C_l,\ \forall l\in\mathcal{L} \tag{33b}$$

$$-C_l\le\sum_{i\in\mathcal{G}}\pi_{gi}^l(P_{gi}(t)+\alpha_{gi}(t)\underline{\phi}_t')+\sum_{i\in\mathcal{S}}\pi_{si}^l(P_{si}^d(t)-P_{si}^c(t)+\alpha_{si}(t)\underline{\phi}_t')-\sum_{i\in\mathcal{B}}\pi_{di}^lP_{di}(t)-\overline{\theta}_{lt}',\ \forall l\in\mathcal{L} \tag{33c}$$

$$-C_l\le\sum_{i\in\mathcal{G}}\pi_{gi}^l(P_{gi}(t)+\alpha_{gi}(t)\overline{\phi}_t')+\sum_{i\in\mathcal{S}}\pi_{si}^l(P_{si}^d(t)-P_{si}^c(t)+\alpha_{si}(t)\overline{\phi}_t')-\sum_{i\in\mathcal{B}}\pi_{di}^lP_{di}(t)-\overline{\theta}_{lt}',\ \forall l\in\mathcal{L} \tag{33d}$$

C. Deterministic MILP Formulation

In summary, evaluating the worst-case expectation in (27) with the LP (34) and replacing the DRCC (24)$\sim$(26) with the deterministic reformulation (30)$\sim$(33), we can cast the proposed formulation into a **MILP model** for which off-the-shelf solvers are available.

D. Eliminating Inactive Line Capacity Constraints

The number of line capacity constraints (33) are $4\times|\mathcal{L}|\times|\mathcal{T}|$ which could be prohibitively large for real-world power systems with small dispatch intervals. Fortunately, in practice, most of the line capacity constraints are inactive thus redundant for the optimization model. If the inactive constraints can be identified and eliminated before solving the problem, the computational burden can be significantly reduced. Here we extend the fast identification method in [24] to the problem formulation in this paper.

Consider the following problems:

$$\begin{aligned}&\Lambda_{t,max}^l(\tilde{\phi}_t,\tilde{\theta}_{lt})\left(\Lambda_{t,min}^l(\tilde{\phi}_t,\tilde{\theta}_{lt})\right)\\&=\underset{\boldsymbol{P}_g(t),\boldsymbol{P}_s(t),\boldsymbol{\alpha}_g(t),\boldsymbol{\alpha}_s(t)}{\arg\max(\min)}\sum_{i\in\mathcal{G}}\pi_{gi}^l(P_{gi}(t)+\alpha_{gi}(t)\tilde{\phi}_t)\\&\qquad+\sum_{i\in\mathcal{S}}\pi_{si}^l(P_{si}(t)+\alpha_{si}(t)\tilde{\phi}_t)\\&\qquad-\sum_{i\in\mathcal{B}}\pi_{di}^lP_{di}(t)-\tilde{\theta}_{lt}\end{aligned} \tag{39}$$

subject to

$$\begin{aligned}&\sum_{i\in\mathcal{G}}(P_{gi}(t)+\alpha_{gi}(t)\tilde{\phi}_t)+\sum_{i\in\mathcal{S}}(P_{si}(t)+\alpha_{si}(t)\tilde{\phi}_t)\\&\qquad=\sum_{i\in\mathcal{B}}P_{di}(t)+\tilde{\phi}_t\end{aligned} \tag{40a}$$

$$\underline{P}_{gi}\le P_{gi}(t)+\alpha_{gi}(t)\underline{\phi}_t' \tag{40b}$$

$$P_{gi}(t)+\alpha_{gi}(t)\overline{\phi}_t'\le\overline{P}_{gi} \tag{40c}$$

$$-\overline{P}_{si}^c\le P_{si}(t)+\alpha_{si}(t)\underline{\phi}_t' \tag{40d}$$

$$P_{si}(t)+\alpha_{si}(t)\overline{\phi}_t'\le\overline{P}_{si}^d. \tag{40e}$$

where (40a) is obtained by multiplying (13) by $\tilde{\phi}_t$ and adding to (12); (40b)$\sim$(40c) are deduced from (15), (30) and (31). Therefore, the feasible sets of the above optimization problems are relaxations of the feasible set of the original MILP model. Minimization (maximization) w.r.t. the feasible set defined by (40) yields a lower (upper) bound of the minimum (maximum) w.r.t the feasible set of the original MILPmodel. The objective function (39) is just the possible line flow at each line in each time period. Similar to the analysis in [24], [25], we have the following lemma.

***Lemma** 3:* For any $l\in\mathcal{L}$ and $t\in\mathcal{T}$, we have

- If $\Lambda_{t,max}^l(\underline{\phi}_t,\underline{\theta}_t^l)\le C_l$, constraint (33a) is inactive;
- If $\Lambda_{t,max}^l(\overline{\phi}_t,\underline{\theta}_t^l)\le C_l$, constraint (33b) is inactive;
- If $\Lambda_{t,min}^l(\underline{\phi}_t,\overline{\theta}_t^l)\ge -C_l$, constraint (33c) is inactive;
- If $\Lambda_{t,min}^l(\overline{\phi}_t,\overline{\theta}_t^l)\ge -C_l$, constraint (33d) is inactive.

To simplify the above LP (39)(40), we further define $\mathcal{H}=\mathcal{G}\cup\mathcal{S}$ and for any $i\in\mathcal{H}$:

$$P_i(t)=\begin{cases}P_{gi}(t)+\alpha_{gi}(t)\tilde{\phi}_t-\underline{P}_{gi},\ i\in\mathcal{G}\\P_{si}(t)+\alpha_{si}(t)\tilde{\phi}_t+\overline{P}_{si}^c,\ i\in\mathcal{S}\end{cases} \tag{41a}$$

$$\pi_i^l=\begin{cases}\pi_{gi}^l,\ i\in\mathcal{G}\\\pi_{si}^l,\ i\in\mathcal{S}\end{cases} \tag{41b}$$

$$\overline{P}_i=\begin{cases}\overline{P}_{gi}-\underline{P}_{gi},\ i\in\mathcal{G}\\\overline{P}_{si}^d+\overline{P}_{si}^c,\ i\in\mathcal{S}\end{cases} \tag{41c}$$

$$D_t=\sum_{i\in\mathcal{B}}P_{di}(t)-\sum_{i\in\mathcal{G}}\underline{P}_{gi}+\sum_{i\in\mathcal{S}}\overline{P}_{si}^c \tag{41d}$$

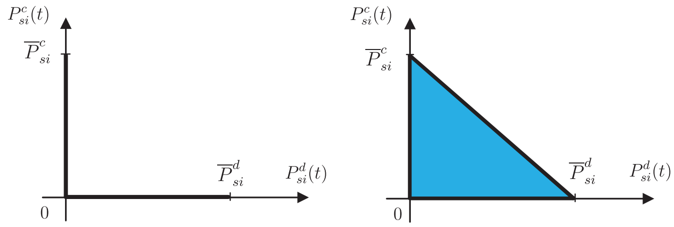
$$K_t^l=-\sum_{i\in\mathcal{B}}\pi_{di}^lP_{di}(t)+\sum_{i\in\mathcal{G}}\pi_{gi}^l\underline{P}_{gi}-\sum_{i\in\mathcal{S}}\pi_{si}^l\overline{P}_{si}^c \tag{41e}$$

Then we have

$$\begin{aligned}&\Lambda_{t,max}^l(\tilde{\phi}_t,\tilde{\theta}_{lt})\left(\Lambda_{t,min}^l(\tilde{\phi}_t,\tilde{\theta}_{lt})\right)\\&=\arg\underset{\boldsymbol{P}(t)}{\max(\min)}\sum_{i\in\mathcal{H}}\pi_i^lP_i(t)+K_t^l-\tilde{\theta}_{lt}\\&\text{s.t.}\begin{cases}\sum_{i\in\mathcal{H}}P_i(t)=D_t+\tilde{\phi}_t\\0\le P_i(t)\le\overline{P}_i\end{cases}\end{aligned} \tag{42}$$

LP (42) has a analytical solution according to the analysis in [24]. Let $i_1,i_2,\cdots,i_{|\mathcal{H}|}$ be a permutation of $1,2,\cdots,|\mathcal{H}|$ such that $\{\pi_{i_1}^l,\pi_{i_2}^l,\cdots,\pi_{i_{|\mathcal{H}|}}^l\}$ are in descending (ascending) order, and there exists an integer $1\le m\le|\mathcal{H}|$ such that $\sum_{k=1}^{m-1}\overline{P}_{i_k}\le D_t+\tilde{\phi}_t\le\sum_{k=1}^{m}\overline{P}_{i_k}$. Then

$$\begin{aligned}&\Lambda_{t,max}^l(\tilde{\phi}_t,\tilde{\theta}_{lt})\left(\Lambda_{t,min}^l(\tilde{\phi}_t,\tilde{\theta}_{lt})\right)\\&=\sum_{k=1}^{m-1}(\pi_{i_k}^l-\pi_{i_m}^l)\overline{P}_{i_k}+\pi_{i_m}^l(D_t+\tilde{\phi}_t)+K_t^l-\tilde{\theta}_{lt}.\end{aligned} \tag{43}$$



(a) Feasible set of constraint (18) (b) Convex hull defined by (44)

Fig. 1: Illustration of Convex Relaxation of (18) to (44).

Based on the analytical expression (43), lemma 3 gives a computationally cheap way to identify most of the inactive line capacity constraints in (33).

E. Convex Relaxation

The non-convexity of the proposed optimization model comes only from the binary variables introduced in (18) to avoid simultaneous charging and discharging of energy storages. The analysis in [26] and [27] show that simultaneous charging and discharging could only happen when the local marginal price goes below a negative threshold value, which is very unusual in practical operation. Considering such feature, we employ an iterative scheme similar to the successive constraint enforcement in [2], [17]. In our implementation, we first relax all the constraint (18) to its convex hull:

$$P^c_{si}(t) \geq 0,\ \ P^d_{si}(t) \geq 0 \tag{44a}$$

$$P^c_{si}(t)/\overline{P}^c_{si} + P^d_{si}(t)/\overline{P}^d_{si} \leq 1 \tag{44b}$$

which is illustrated in Fig. 1. In this way, we obtain a LP relaxation to the original MILP. If the solution of relaxed model happens to lie in the feasible set of the original MILP model, we can conclude that this solution is also the global optimal solution of the original MILP problem. Therefore, after solving the LP model, we check whether $P^c_{si}(t)P^d_{si}(t) = 0$ for all $i \in \mathcal{S}, t \in \mathcal{T}$. If so, the solution to the orginal MILP model is found. Otherwise, the constraint (44) has to be changed back to constraint (18) for those i and t where $P^c_{si}(t)P^d_{si}(t) > 0$. Hence we obtain a relaxed MILP model which is still much simpler than the orignal MILP model. Again we solve the relaxed model and then check the exactness of the relaxation. This process is repeated until $P^c_{si}(t)P^d_{si}(t) = 0$ for all $i \in \mathcal{S}, t \in \mathcal{T}$.

V. Numerical Results

The proposed formulation and method were programmed in MATLAB with Gurobi as the MILP and LP solver running on a Win 8 PC with a 3.0GHz CPU and 24 GB RAM. The simulation was carried out on IEEE-118 bus system modified according to [28]. Five wind farms were installed at bus 16, 37, 48, 75 and 83 with each capacity of 100MW. Storage was assumed to be installed at each non-generator bus. The energy capacity for each storage was 32 MW, and the charge & discharge power capacity were both set to be 8 MW/h. The charge & discharge efficiency were both 0.9. The ramp rates over each hour for conventional generator were set to be 50% of the rated capacity. We considered a time horizon of 24 h with each time step 30 min. The half-hourly forecasting load and wind generation profile were obtained from [2] and NREL WIND Toolkit. All uncertainties were assumed to originate from the forecasting errors of wind generation. Different types of probability distributions were employed to generate "realistic" data of wind power forecasting errors whose mean and variance were set according to the typical day-ahead forecasting errors in U.S. reported in [29]. The confidence level α for the confidence bands of CDF is set to 0.05. In addition, the reserve availability and utilization prices of generators and storages were randomly selected from the ranges shown in Table I.

TABLE I: Reserve Availability and Utilization Prices

R^+_{gi}/R^-_{gi}	f^+_{gi}/f^-_{gi}	R^+_{si}/R^-_{si}	f^+_{si}/f^-_{si}	c^d_{si}/c^c_{si}
0.1∼0.3 c_{gi}	0.8∼1.1 c_{gi}	0.1∼0.3 c^d_{si}/c^c_{si}	0.8∼1.1 c^d_{si}/c^c_{si}	15/10 $/WMh

The proposed method is data-driven and distribution-free, so it can deal with wind power forecasting errors following any probability distributions. Thus, we tested the method with wind power forecasting errors generated from normal, laplace, beta and hyperbolic distributions, and Monte Carlo simulations with 10^6 samples were employed to assess the performance of the proposed method. Fig. 2 illustrates the evolution of optimization objective function (solid blue line) and the operation costs from Monte Carlo simulation (dotted red line) as the number of available historical data increases. Firstly, we observe that the operation costs from Monte Carlo simulation are always upper bounded by the objective function of the optimization model. This exhibits the distributional robustness of the proposed method: the objective function represents the costs w.r.t the worst-case distribution in the ambiguity set whereas the underlying true distribution could be different from the worst-case one. Secondly, as more historical data is available, both the objective functions and the simulated costs decrease, which shows the value of data: the more historical data is employed, the less conservative the solution is. Finally, the gaps between the objective functions and the simulated costs are narrowed by incorporating more historical data which reveals that the ambiguity set shrinks to the underlying true probability distribution as the number of historical data increases. Table II further lists the percentage gaps between objective functions and simulated costs.

Since the proposed problem formulation is a direct extension of the robust multi-period OPF [2], we compared the proposed CDF-based DRO approach with the RO approach in [2]. The RO approach can be implemented in our problem formulation by: 1) replacing the DRCC (24)∼(26) with linear robust constraints using the support of the random variable as the uncertainty set; 2) eliminating the reserve utilization term in the objective function (27). The operation costs of the strategies obtained by both methods are compared in Fig. 3. In Fig. 3a, the wind power forecasting follows beta distribution with different variances. The proposed CDF-based DRO approach achieves lower operation costs than RO approach under each level of forecasting errors. The difference of the operation costs tends to increase as the variances of the forecasting error increase. In Fig. 3b, the similar comparison is carried out under different types of distributions of the forecasting errors. It is also shown that the proposed CDF-based DRO approach obtains more economical operation strategies than the RO approach. In short, the proposed CDF-based DRO approach captures the detailed probabilistic information from historical data while the RO approach ignores such information. When more data is available, the proposed method can produce a more economic strategy whereas the RO does not have a mechanism to take advantage of more data.

Another major competitor of the CDF-based DRO is the moment-based DRO appears in [9], [12], [19], [20], etc. The ambiguity set for moment-based DRO is the set of all probability distributions with given mean and covariance (usually sample mean and sample covariance). To initiate a meaningful comparison, we implement the moment-based DRO in our problem formulation as follows: 1) the chance constraints (24)∼(26) can be reformulated as SOCP constraints by leveraging the Chebyshev inequality $\mathbb{P}\{|\tilde{\xi}-\mu| \geq \sqrt{1/\rho}\sigma\} \leq \rho$ [20]; 2) the evaluation of worst-case expectation in (27e) can be also cast into a SOCP using duality theory of moment problem. Fig. 4 compares the operational costs by two methods when different number of historical data is available. The curves under different types of underlying true distributions show the similar pattern. The costs of moment-based DRO is always higher than that of the CDF-based DRO and the difference is enlarged when more data is at hand. By merely relying on the information of the first two moments, the moment-base DRO is unable to fully take advantage of the abundance of data, which is in stark contrast to the data-exploiting feature of the CDF-based DRO. Fig. 5 and fig. 6 confirm the foregoing conclusion by observing the violation probability of chance constraint (24) and (25). In this group of test, we have set the $\beta_1 = \beta_2 = 0.05$. Fig. 5 and Fig. 6 show that both methods ensure higher reliability level than required due to their "distributionally robust" nature. As more data is available, the proposed CDF-based DRO gradually and safely reduce the guaranteed reliability level to pursue higher economic efficiency. In contrast, the conservatism of the moment-based DRO remains significant even with 10^6 data. To further reveal the nature of both methods, we then investigate the relationship between β and the minimal τ that ensure $\mathbb{P}[|\xi-\mu| \leq \tau] \geq 1-\beta, \forall \mathbb{P} \in \mathcal{P}$ when $\mathcal{P}$ is the moment-based ambiguity set or CDF-based ambiguity set constructed from the different number of data. It is shown in Fig. 7 that the true relation between β and τ for the specific distribution under study is always upper bounded by those provided by the DRO approaches. The curve given by the moment-based DRO is very far away from the true curve, and the curves provided by the CDF-based DRO approach the true curve as more and more data is available. In summary, the CDF-based DRO extracts much more probabilistic information from data than the moment-based DRO approach, which leads to less conservative operation strategy. In addition, when more data is at hand, the moment-based DRO can merely have a more accurate guess of the mean and covariance, whereas the proposed CDF-based DRO can extract more detailed information about the whole distribution. The CDF-based DRO thus has much stronger ability to exploit data.

TABLE II: Percentage Gap Between the Objective Function and the Costs by Monte Carlo Simulation

data num.	500	1000	2000	5000	10000	50000
normal	0.041%	0.028%	0.020%	0.013%	0.009%	0.004%
laplace	0.494%	0.291%	0.198%	0.120%	0.081%	0.035%
beta	0.907%	0.593%	0.416%	0.252%	0.177%	0.079%
hyperbolic	0.481%	0.288%	0.196%	0.119%	0.080%	0.035%

TABLE III: Solver Time (sec.) of the RO, the moment-based DRO and the CDF-based DRO.

method	RO	M-DRO	CDF-DRO (10^2)	CDF-DRO (10^3)	CDF-DRO (10^4)	CDF-DRO (10^5)
normal	3.8	146.3	6.1	3.6	7.9	11.6
laplace	4.4	132.9	6.2	4.3	7.8	12.3
beta	6.7	187.6	6.2	8.8	10.7	11.7
hyperbolic	3.9	128.5	7.1	3.8	4.3	12.5

The focus is then given to the computational efficiency of the proposed method. Table III lists the solver time of the proposed CDF-based DRO approach using the different number of data points along with the solver time of the RO and the moment-based DRO approaches. The solver time of the proposed method does not necessarily grow with the number of historical data due to the fact that the scale of the proposed MILP model is irreverent to the number of data points. The variation of the solver time with the number of data points is just due to numerical issues of the solver. Compared with the RO approach, the computational burden of the proposed CDF-based DRO approach is slightly more intensive on average due to the larger number of decision variables and constraints. The moment-based DRO, on the other hand, can be solved much slower than the other two approaches due to its SOCP formulation.

Table IV demonstrates the effectiveness of the inactive constraint elimination and convex relaxation procedures discussed in section IV. It is shown that more than 88% of the line capacity constraints are identified to be inactive thus redundant to the optimization model. In all our tests, the average solver time of the proposed method without inactive constraint elimination is around 38.7s. After elimination, the average solver time reduces to 9.6s which is approximately 1/4 of that without redundant constraint elimination. Moreover, the relaxation of constraint (18) to constraint (44) is 100% exact for the energy storages, which indicates the local marginal prices at the storage buses are always beyond the threshold values for exact relaxation [26], [27]. The underlying physical reason for this phenomenon is the relatively adequate transmission and storage capacity of the case under study [27]. Both the elimination of the inactive line capacity constraints and the convex relaxation of constraint (18) significantly contribute to the improvement of numerical tractability.

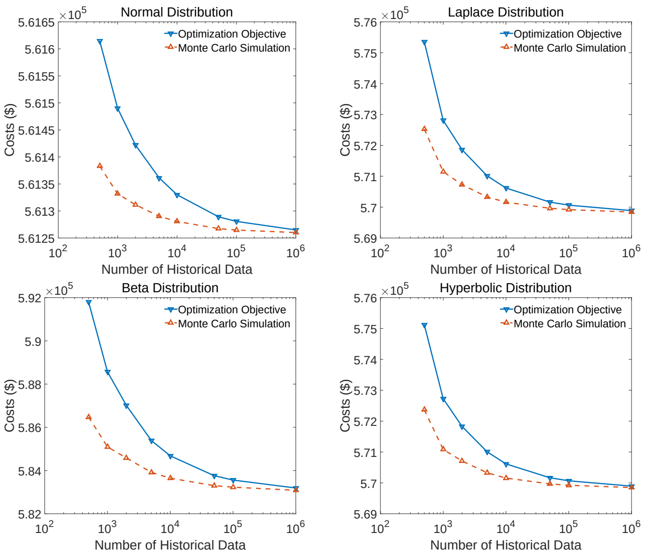
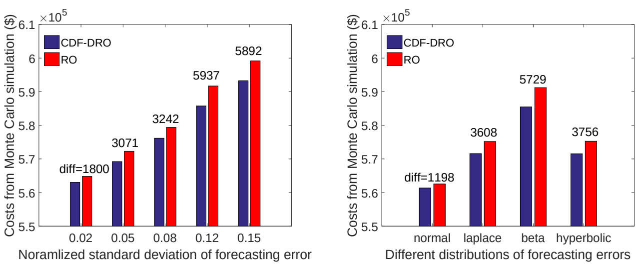


Fig. 2: Evolution of objective function and simulated costs as the increase of available historical data.



(a) different standard deviations (b) different distribuitions

Fig. 3: Cost comparison between CDF-based DRO and RO

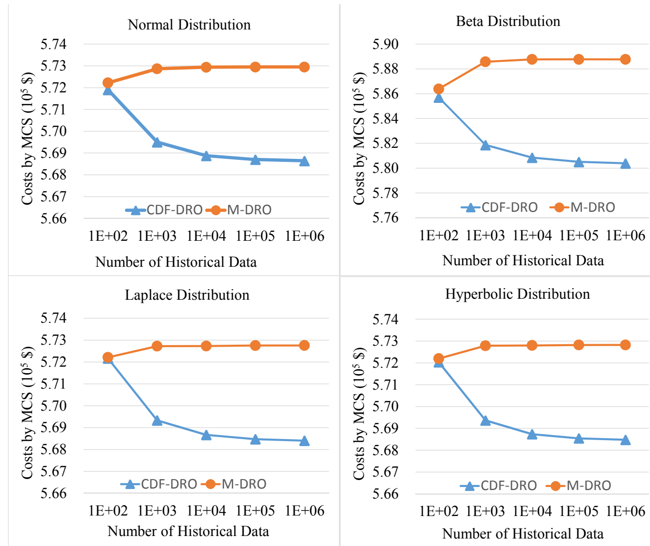


Fig. 4: Comparison between the CDF-based DRO and the moment-based DRO for the operational costs by MCS under different types of distributions.

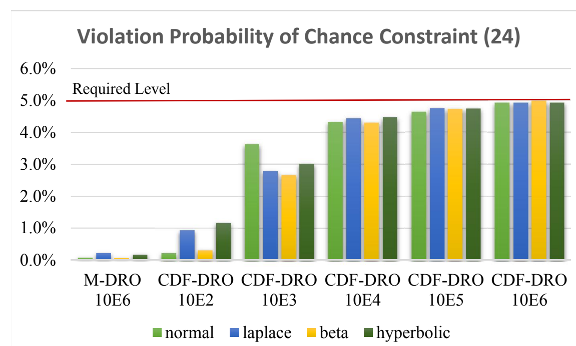


Fig. 5: Comparion of the violation probability of chance constraint (24) between the moment-based DRO and the CDF-based DRO with different number of data.

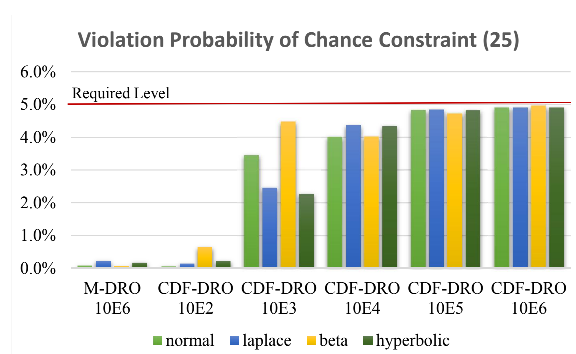


Fig. 6: Comparion of the violation probability of chance constraint (25) between the moment-based DRO and the CDF-based DRO with different number of data.

VI. Conclusion

In this paper, an energy-reserve-storage co-optimization model and a data-driven distributionally robust method are proposed to achieve economical and reliable operation of power systems with variable and uncertain renewable sources. Compared with the SP approach, the proposed method assumes no prior knowledge of the probability distribution of the uncertainties and achieve operational robustness by considering the worst-case distribution consistent with the observed data. Compared with the RO approach, the proposed method exploits detailed probabilistic information learned from historical data and the conservatism of the solution can be reduced by incorporating more historical data. Numerical studies demonstrate the favorable features of the proposed methods.

TABLE IV: Percentages of Identified Inactive Line Capacity Constraints and Exact Convex Relaxations

distribution	normal	laplace	beta	hyperbolic
inactive line constraint (%)	89.57%	88.78%	88.55%	88.82%
relaxation exactness (%)	100%	100%	100%	100%

References

[1] X. Luo, J. Wang, M. Dooner, and J. Clarke, “Overview of current development in electrical energy storage technologies and the application

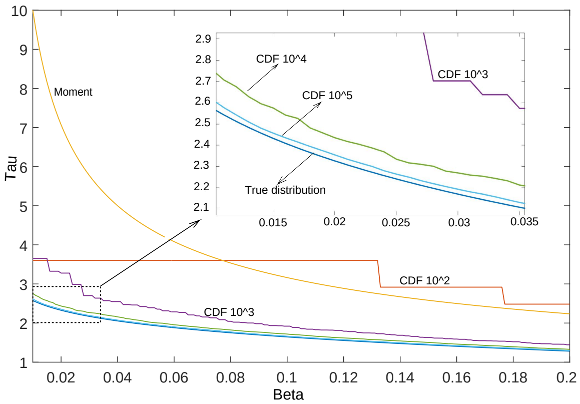


Fig. 7: The relationship between β and the minimal τ which ensure $\mathbb{P}[|\xi - \mu| \leq \tau] \geq 1 - \beta, \forall \mathbb{P} \in \mathcal{P}$ when $\mathcal{P}$ is the moment-based ambiguity set or CDF-based ambiguity set constructed from different number of data.

potential in power system operation," *Applied Energy*, vol. 137, pp. 511–536, 2015.

[2] R. Jabr, S. Karaki, and J. Korbane, "Robust multi-period opf with storage and renewables," *IEEE Trans. Power Syst.*, vol. 30, pp. 2790–2799, Sept 2015.

[3] S. Wogrin and D. Gayme, "Optimizing storage siting, sizing, and technology portfolios in transmission-constrained networks," *IEEE Trans. Power Syst.*, vol. 30, pp. 3304–3313, Nov 2015.

[4] S. Gill, I. Kockar, and G. Ault, "Dynamic optimal power flow for active distribution networks," *IEEE Trans. Power Syst.*, vol. 29, pp. 121–131, Jan 2014.

[5] A. Gabash and P. Li, "Active-reactive optimal power flow in distribution networks with embedded generation and battery storage," *IEEE Trans. Power Syst.*, vol. 27, pp. 2026–2035, Nov 2012.

[6] J. Lavaei and S. Low, "Zero duality gap in optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 27, pp. 92–107, Feb 2012.

[7] D. Gayme and U. Topcu, "Optimal power flow with large-scale storage integration," *IEEE Trans. Power Syst.*, vol. 28, pp. 709–717, May 2013.

[8] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, "A probabilistic framework for reserve scheduling and security assessment of systems with high wind power penetration," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3885–3896, 2013.

[9] 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.

[10] A. Street, A. Moreira, and J. M. Arroyo, "Energy and reserve scheduling under a joint generation and transmission security criterion: An adjustable robust optimization approach," *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 3–14, 2014.

[11] W. Wei, F. Liu, S. Mei, and Y. Hou, "Robust energy and reserve dispatch under variable renewable generation," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 369–380, 2015.

[12] 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.

[13] Z. Li, W. Wu, B. Zhang, and B. Wang, "Adjustable robust real-time power dispatch with large-scale wind power integration," *IEEE Trans. Sustain. Energy*, vol. 6, no. 2, pp. 357–368, 2015.

[14] A. Lorca and X. A. Sun, "Adaptive robust optimization with dynamic uncertainty sets for multi-period economic dispatch under significant wind," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1702–1713, 2015.

[15] Q. Zheng, J. Wang, P. Pardalos, and Y. Guan, "A decomposition approach to the two-stage stochastic unit commitment problem," *Annals of Operations Research*, vol. 210, no. 1, pp. 387–410, 2013.

[16] Q. Wang, Y. Guan, and J. Wang, "A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output," *IEEE Trans. Power Syst.*, vol. 27, pp. 206–215, Feb 2012.

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

[18] T. Ding, Z. Bie, L. Bai, and F. Li, "Adjustable robust optimal power flow with the price of robustness for large-scale power systems," *IET Generation, Transmission & Distribution*, vol. 10, no. 1, pp. 164–174, 2016.

[19] W. Wei, J. Wang, and S. Mei, "Dispatchability maximization for co-optimized energy and reserve dispatch with explicit reliability guarantee," *IEEE Trans. Power Syst.*, vol. 31, pp. 3276–3288, July 2016.

[20] Y. Zhang, S. Shen, and J. 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.

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

[22] M. Goldman and D. M. Kaplan, "Evenly sensitive ks-type inference on distributions," tech. rep., Working paper, available at http://faculty. missouri. edu/˜ kaplandm, 2015.

[23] C. Duan, L. Jiang, W. Fang, and J. Liu, "Data-driven affinely adjustable distributionally robust unit commitment," *IEEE Trans. Power Syst.*, vol. PP, no. 99, pp. 1–1, 2017.

[24] Q. Zhai, X. Guan, J. Cheng, and H. Wu, "Fast identification of inactive security constraints in scuc problems," *IEEE Trans. Power Syst.*, vol. 25, no. 4, pp. 1946–1954, 2010.

[25] H. Ye and Z. Li, "Necessary conditions of line congestions in uncertainty accommodation," *IEEE Trans. Power Syst.*, vol. 31, pp. 4165–4166, Sept 2016.

[26] Z. Li, Q. Guo, H. Sun, and J. Wang, "Sufficient conditions for exact relaxation of complementarity constraints for storage-concerned economic dispatch," *IEEE Trans. Power Syst.*, vol. 31, pp. 1653–1654, March 2016.

[27] C. Duan, L. Jiang, W. Fang, X. Wen, and J. Liu, "Improved sufficient conditions for exact convex relaxation of storage-concerned ed," *arXiv preprint arXiv:1603.07875*, 2016.

[28] H. Yu and W. Rosehart, "An optimal power flow algorithm to achieve robust operation considering load and renewable generation uncertainties," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1808–1817, 2012.

[29] B.-M. Hodge, D. Lew, M. Milligan, H. Holttinen, S. Sillanpaa, E. Gomez Lazaro, R. Scharff, L. Soder, X. Larsen, G. Giebel, D. Flynn, and D. J., "Wind Power Forecasting Error Distributions: An International Comparison," tech. rep., National Renewable Energy Laboratory, 2012.

Chao Duan (S'14) was born in Chongqing, China, in 1989. He received the B.S. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 2012. He is currently pursuing the Ph.D. degree at Xi'an Jiaotong University, Xi'an, China, and the University of Liverpool, Liverpool, U.K. His research interests are in stochastic optimization, stability analysis and robust control of power systems.

Lin Jiang (M'00) received the B.Sc. and M.Sc. degrees from Huazhong University of Science and Technology (HUST), China, in 1992 and 1996; and the Ph.D. degree from the University of Liverpool, UK, in 2001, all in electrical engineering. He is currently a Reader in the University of Liverpool. His current research interests include control and analysis of power system, smart grid, and renewable energy.

Wanliang Fang was born in Henan, China, in 1958. He received the B.S. and M.S. degrees from Xi'an Jiaotong University, Xi'an, China in 1982 and 1988, respectively, and the Ph.D. degree from Hong Kong Polytechnic University, HongKong, in 1999, all in electrical engineering. He is currently a Professor of electrical engineering at Xi'an Jiaotong University. His research interests include power system stability analysis and control, FACTS and HVDC.

Jun Liu (S'09-M'10) received the B.S. and Ph.D. degrees from Xi'an Jiaotong University, Xi'an, China, in 2004 and 2012, respectively, all in electrical engineering. He is currently an associate Professor in the Department of Electrical Engineering, Xi'an Jiaotong University. His research interests include renewable energy integration, power system operation and control, power system stability, HVDC, and FACTS.

Shiming Liu received the Ph.D. degree in electrical engineering from Huazhong University of Science and Technology, Wuhan, China, in June 1999. He is now a Professor in school of electrical engineering of Shandong University. His research is mainly on power system protection and substation automation.