Stochastic Market Operation for Coordinated Transmission and Distribution Systems

Shengfei Yin, *Student Member, IEEE*, Jianhui Wang, *Fellow, IEEE*, and Harsha Gangammanavar

***Abstract*—This paper proposes a three-stage unit commitment model for the market operation of transmission and distribution coordination under the uncertainties of renewable generation and demand variations. The first stage is for the independent system operator (ISO) that determines the commitment decisions of the transmission-level generators and the distribution-level system reconfiguration; the second stage optimizes the transmission economic dispatch; then distribution system operators (DSOs) in the third stage perform their economic dispatch and exchange information with the ISO at the boundary nodes. Between the transmission and distribution networks, not only conventional thermal generators but renewables and variable demands are considered, which are tackled via a multi-stage stochastic programming approach. The model adopts a convexified AC branch flow formulation in the distribution system. We devise a generalized nested L-shaped algorithm to solve the proposed framework in an efficient manner. Numerical experiments on multi-scale test systems corroborate the efficacy of this strategy.**

***Index Terms*—Coordinated transmission and distribution market, uncertainty-based optimization, stochastic programming, nested decomposition method.**

NOMENCLATURE

1) Sets:

T / D	Transmission (T) / distribution (D) system
G^T / G^D	Conventional thermal generator of T / D system
E / I	Renewable generator / connected bus
Q	Network reconfiguration set
L^T / L^D	Load of T / D system
F^T / F^D	Line of T / D system
A^D / A^D_T	Bus of D system / with substation
F^D_Q	Configurable line of D system
S / N	Sending / receiving bus of line
C	Generator or injection mapping with bus
H	Time horizon
O	Iteration

2) Indices:

g^T / g^D	Indices for conventional generators in T/D system
r^T / r^D	Indices for renewable generators in T/D system
ℓ^T / ℓ^D	Indices for loads in T/D system
f^T	Indices for transmission lines
mn	Indices for connected distribution branches
n^T / n^D	Indices for nodes in T/D system
i	Indices for boundary buses
ω^T / ω^D	Indices for operating scenarios in T/D system
o	Indices for iteration counter
h	Indices for time (hours)
c	Indices for multiple D systems
q	Indices for configurable swtich in D system

3) Parameters (in both T and D, otherwise specified):

PC	Penalty cost [$/MWh]
R / X / B	Resistance / reactance / susceptance [p.u.]
RD / RU	Ramp-down / up limits [MW/h]
SD / SU	Shut-down / start-up cost [$]
MO / MD	Minimum ON / OFF time limits [h]
$P_f^{T,\max}$	Active power flow limits [MW]
$Y_{mn}^{D,\max}$	Upper limit of the branch current [A]
$V^{\min}$ / $V^{\max}$	Lower/Upper limit of the nodal voltage [kV]
$P_g^{\min}$ / $P_g^{\max}$	Active power output limits of generators [MW]

4) Uncertainties (in both T and D):

$P_{r,h}^{\max}(\omega)$	Maximum available renewable power [MW]
$L_{\ell,h}(\omega)$	Nodal demand [MW]
Pr_ω	Probability for scenario ω

5) Binary Variables:

$u_{g,h}^T$ / $d_{g,h}^T$	Start-up/shut-down indicator for generators
$i_{g,h}^T$	Unit commitment indicator for generators
$z_{q,h,c}$	Distribution configuration indicator

6) Continuous Variables:

$p_{(g,r)}^{(T,D)}$ / p_i	Active generation output / power injection
$q_{(g,r)}^D$	Reactive generation output
p_n^D / q_n^D	Active / reactive node injection
p_f^T	Active line flow
p_{mn}^D / q_{mn}^D	Active / reactive feeder flow
x_{mn}^D	Line orientation variable
v_n / δ_n^T	Squared nodal voltage magnitude / angle
y_{mn}^D	Squared branch current
s_ℓ	Active load shedding

Manuscript received September 10, 2020; revised January 11, 2021 and March
Q1 21, 2021; accepted April 23, 2021. Paper no. TSTE-00986-2020. *(Corresponding author: Jianhui Wang.)*

Shengfei Yin and Jianhui Wang are with the Department of Electrical and Computer Engineering, Southern Methodist University, Dallas, TX 75275-0221 USA (e-mail: gyin@smu.edu; jianhui@smu.edu).

Harsha Gangammanavar is with the Department of Engineering Management, Information and Systems, Southern Methodist University, Dallas, TX 75205 USA (e-mail: harsha@smu.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSTE.2021.3076037.

Digital Object Identifier 10.1109/TSTE.2021.3076037

I. Introduction

Unit commitment (UC) and economic dispatch (ED) have been of capital importance in the market operations for system operators in the world. The coordination between transmission and distribution systems (T-D coordination) in the market operation has attacked great attention. The primary motivation for T-D coordination is based on the fact that the distribution system has become active due to the increasing deployment of smart grid technologies including distributed energy resources (DERs), electric vehicles, microgrids, *etc*. The system dynamics within the distribution networks can propagate to the transmission side and *vice versa* [1]. A recent report from California Independent System Operator (CAISO) [2] supports this motivation. It indicates that, due to the growing penetration of the DERs, Independent System Operator (ISO), which regulates the upstream market in the transmission level, needs to attain more visibility on the downstream markets to reinforce its decision-making.

In the conventional market hierarchy, ISOs and utilities carry out UC and ED problems within their respective territories. The upstream operators make scheduling decisions according to the estimated information from the downstream aggregators [3]. Without the T-D coordination, the power mismatch at the boundary node between transmission and distribution systems proliferates as distribution systems become more and more active [4]. To cope with the issue of power mismatches, Z. Li *et al.* [5] proposed a distributed paradigm that ISO delivers its optimal locational marginal price (LMP) to distribution system operators (DSOs), whereas DSOs return their optimal energy demand to ISO. The mutually observable information in the T-D coordination should be limited to protect the confidentiality of stakeholders [5].

Recently, based on this decentralized framework, many papers have considered more elements and more efficient algorithms. P. Li *et al.* [6] built a distributionally robust optimization framework considering system uncertainties and adopted an analytical target cascading algorithm to decentralize the problem. R. Nejad *et al.* [7] and J. Zhao *et al.* [8] focused on the integrated system restoration using the alternating direction of multipliers method (ADMM) to decentralize the problem. M. Arpanahi *et al.* [9] proposed a decentralized framework for T-D coordination that considers robust optimization for the uncertainties, while using an enhanced ADMM algorithm to solve. A. Nawaz *et al.* [10] proposed a probabilistic coordination method to select one stochastic scenario to solve the T-D coordinated problem. However, most of the works focus on ED or OPF, instead of the UC, which could not serve as a general reference for electricity market operations. It is also reasonable that the algorithms proposed above lie in the category of the augmented Lagrangian method, which would fail if integer variables are present in the problem.

For the modeling of T-D coordination, recently, research on this topic gains prevalence, and potential modeling frameworks from the perspective of DSOs can be categorized into three types, *i.e.*, centralized, decentralized, and transactive models, which are summarized in [11]. Among the models discussed therein, Model 3, *i.e.*, the distribution node utility model, acts as a transitional model from the centralized operation to the decentralized operation, where the ISO and DSOs take the responsibilities of market operations in the transmission and distribution level, respectively, and they exchange information in the boundary substation. The natural compatibility with the current market practice is the major advantage of this model [11]. Based on this framework, we propose a three-stage stochastic co-optimization for the T-D coordinated UC and ED. In this scheme, the ISO regulates the coordination between the transmission and distribution systems, where the transmission system operation is under control of the ISO and the distribution system operation is under control of the DSO. As in the current industry's practice [12], the ISO carries out the transmission-level market operations including UC and ED with the distribution-level inputs from DSOs. DSOs are utilities that own the distribution network with aggregated DERs. It is also noteworthy that the current research on decentralized or transactive models is fruitful and few are investigating centralized frameworks due to the computational hurdle. However, before the fully decentralized distribution operations, the industry needs a transitional model that allows the ISOs to take the local distribution information into account while regulating the coordination [13]. Hence, in the proposed market hierarchy, the ISO performs the day-ahead transmission-level UC and ED with the aggregated distribution information, and DSOs perform the distribution-level ED. The ISO does not need to attain detailed distribution system information, which protects the local utilities' confidentiality.

Apart from the scheduling, the distribution-level network reconfiguration drastically influences the market operation in the distribution level, and the influence will propagate to the transmission level [14]. Since the ISO monitors the T-D coordination, the topology changes in the active distribution network (ADN) should be taken into consideration in the ISO's operation. Besides, utilities are using an hourly decision-making process to evaluate and perform the reconfiguration [15]. Considering that the reconfiguration cannot be changed frequently in one day and should be determined in a day-ahead manner [15], in this work, we add the network reconfiguration decisions in the first stage together with the UC decisions.

In summary, we propose a market paradigm for the T-D coordinated power system, which is depicted in Fig. 1. Mathematically, the ISO needs to solve a mixed-integer linear programming (MILP) model for transmission-level UC with distribution network reconfigurations, when DSOs need to solve convex programming models for distribution-level ED. In the first stage, the ISO performs the transmission UC and determines the distribution network reconfiguration, then delivers the information to the subsequent stages. Then, the ED operation is carried out in the second stage and delivers the optimal boundary offer to the third stage. The structure of the first and second stages forms a two-stage transmission-level stochastic UC [16]. Finally, DSOs perform their ED operations in the third stage based on the determined configuration from the first stage and the energy offer from the second stage.

For the solution strategy, there has been a significant number of studies towards the efficient solution of large-scale SP problems. One of the most commonly used strategies is the Benders decomposition or the L-shaped method [17]. For multi-stage

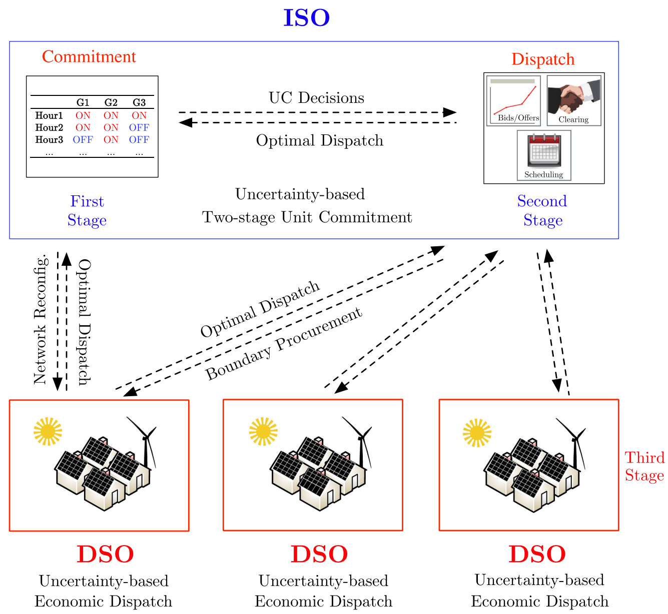


Fig. 1. Proposed coordinated market paradigm.

SP as in this work, scenario-wise decomposition like the progressive hedging [18] and stage-wise decomposition like the approximate dynamic programming [19] are discussed in the literature. Particularly for the dynamic programming category, stochastic dual dynamic integer programming (SDDiP) has been an effective method for multi-stage stochastic optimization [20], with various recent applications on power systems. J. Zhou *et al.* [21] proposed a SDDiP model for multi-stage stochastic UC and enhanced the algorithm by a regularized level approximation for Lagrangian cuts. M. Hjelmeland *et al.* [22] adopted SDDiP solving a hydropower scheduling problem with binary expansion on state variables and found that the strengthened Benders cut presented the highest performance score. T. Ding *et al.* [23] leveraged the SDDiP method in the distribution system planning with a multi-stage nested decomposition algorithm. Though it is well recognized that SDDiP attains fast convergence and accuracy, it could not solve the proposed T-D coordinated market operation. The reasons are twofold: 1) Each time stage in the SDDiP needs to assume the complete information of both transmission and distribution networks, which compromises the stakeholders' privacy; 2) There is no evidence that the SDDiP retains global convergence if recourse problems are nonlinear.

Thus, since our work is both scenario-wise and stage-wise complex, we employ the nested L-shaped method [17] for the proposed T-D coordinated market operation. Furthermore, we enhance the conventional nested L-shaped method and algorithmically extend it to a multi-stage mixed-integer convex programming. To the best of the authors' knowledge, this is the first paper implementing a practical T-D market operation considering accurate network modeling/reconfiguration and uncertainties.

Based on the state-of-the-art research, the main contribution of this work is summarized below.

- We formulate the T-D coordinated UC&ED problem considering the distribution network reconfiguration. System-independent scenario sets capture the uncertainty of renewable generation and elastic demand for each system.
- We propose a market paradigm for the T-D coordinated UC&ED. This paradigm considers the coordination between different market stakeholders while protecting their confidentiality, which ISOs can easily adopt.
- We extend the generalized Benders decomposition method to a nested and stochastic version, which can efficiently solve the proposed T-D coordinated market operation model. We also theoretically prove and analyze the algorithmic convergence.

II. Mathematical Formulation of T-D Coordinated Market Operations

We detail the three-stage stochastic formulation and the convexified AC branch flow for the distribution part in this section. The overall formulation of the stochastic T-D coordinated UC & ED is separated and distributed in the following subsections. Note that the DC power flow is adopted in the transmission level in order to mimic the industrial practice. In contrast, the second-order cone programming (SOCP)-based convexified AC branch flow formulation is adopted in the distribution network to retain accuracy.

A. The First Stage: ISO's UC Problem

The first stage solves a conventional transmission UC problem with the optimal distribution network reconfiguration, which is performed by the ISO.

$$\mathbf{K_1} = \min_{\substack{u^T_{g,h}, d^T_{g,h},\\ i^T_{g,h}, z_{q,h,c}}} \sum_{h}^{H} SU^T_g u^T_{g,h} + SD^T_g d^T_{g,h}$$

$$+\mathbb{E}_{\omega^T}\left\{\mathbf{K_2}(\mathbf{x}_1^*, \omega^T)\right\}, \tag{1a}$$

subject to

$$u^T_{g,h} + d^T_{g,h} \le 1, \qquad \forall g^T \in G^T, \forall h \in H, \tag{1b}$$

$$u^T_{g,h} - d^T_{g,h} = i^T_{g,h} - i^T_{g,h-1}, \qquad \forall g^T \in G^T, \forall h \in H, \tag{1c}$$

$$\sum\nolimits_{h-MO_g+1}^{h} u^T_{g,h} \le i^T_{g,h}, \qquad \forall g^T \in G^T, \forall h \in H, \tag{1d}$$

$$\sum\nolimits_{h-MD_g+1}^{h} d^T_{g,h} \le 1 - i^T_{g,h}, \quad \forall g^T \in G^T, \forall h \in H, \tag{1e}$$

$$u^T_{g,h}, d^T_{g,h}, i^T_{g,h}, z_{q,h,c} \in \{0,1\}, \tag{1f}$$

where $\mathbf{K_2}(\mathbf{x}_1^*, \omega^T)$ denotes the second-stage recourse and $\mathbf{x}_1^* = [u^{T*}_{g,h}, d^{T*}_{g,h}, i^{T*}_{g,h}]$. "*" stands for obtained and fixed decisions. Constraints (1b) and (1c) enforce the binary exclusiveness of the start-up and shut-down indicators; constraints (1d) and (1e) model the minimum ON/OFF time; constraint (1f) indicates the binary nature of the variables. Note that there is no distribution configuration variable, *i.e.*, $z_{q,h,c}$, in the objective function and constraints, which is determined based on the information sent back from the DSO's stage.

B. The Second Stage: ISO's ED Problem

The second stage formulates a transmission-level ED problem based on the UC decisions delivered from the first stage, which is detailed in (2). For each ω^T:

$$\mathbf{K_2}(\mathbf{x}_1^*, \omega^T) = \min_{p^T_{g,h}, s^T_{\ell,h}} \sum_h^H \left\{ \sum_{g^T}^{G^T} f^T(p^T_{g,h}) + \sum_{\ell^T}^{L^T} PC^T_\ell s^T_{\ell,h} \right\} + \sum_c w_c \cdot \mathbb{E}_{\omega_c^D} \left\{ \mathbf{K_3}(\mathbf{x}_2^*, \mathbf{y}^*, \omega_c^D) \right\} \quad \text{(2a)}$$

subject to

$$\sum_{g^T|C(g^T)=n^T}^{G^T} p^T_{g,h} - \sum_{i_c|C(i_c)=n^T}^{I} p_{i_c,h} + \sum_{r^T|C(r)=n^T} p^T_{r,h} - \sum_{f^T|S(f^T)=n^T}^{F^T} p^T_{f,h} + \sum_{f^T|N(f^T)=n^T}^{F^T} p^T_{f,h} = \sum_{\ell^T|C(\ell^T)=n^T}^{L} \{ L^T_{\ell,h}(\omega^T) - s^T_{\ell,h} \}, \quad \forall n^T, \forall h, \quad \text{(2b)}$$

$$p^T_{f,h} = X^{-1}_{f^T}[\delta^T_{n|S(n^T)=f^T,h} - \delta^T_{n|N(n^T)=f^T,h}], \ \forall f^T, \forall h, \quad \text{(2c)}$$

$$-P^{T,\max}_{f,h} \le p^T_{f,h} \le P^{T,\max}_{f,h}, \quad \forall f^T, \forall h, \quad \text{(2d)}$$

$$i^{T,*}_{g,h} P^{T,\min}_{g,h} \le p^T_{g,h} \le i^{T,*}_{g,h} P^{T,\max}_{g,h}, \quad \forall g^T, \forall h, \quad \text{(2e)}$$

$$0 \le p^T_{r,h} \le P^{T,\max}_{r,h}(\omega^T), \quad \forall r^T, \forall h, \quad \text{(2f)}$$

$$p^T_{g,h} - p^T_{g,h-1} \le RU^T_g(1 - u^{T,*}_{g,h}) + u^{T,*}_{g,h} P^{T,\min}_{g,h}, \ \forall g^T, \forall h, \quad \text{(2g)}$$

$$p^T_{g,h-1} - p^T_{g,h} \le RD^T_g(1 - d^{T*}_{g,h}) + d^{T,*}_{g,h} P^{T,\min}_{g,h}, \ \forall g^T, \forall h, \quad \text{(2h)}$$

where $\mathbf{K_3}(\mathbf{x}_2^*, \mathbf{y}^*, \omega_c^D)$ denotes the third-stage recourse, $\mathbf{x}_2^* = [z^*_{q,h,c}]$, and $\mathbf{y}^* = [p^*_{i_c,h}]$. w_c stands for the weight of different distribution systems pertaining to the importance of the corresponding system. For instance, hospitals and military regions require higher priority of electricity supply, which results in higher weight w_c in this model ($w_c \in (0, 1]$ and $\sum w_c = 1$). We also adopt the piecewise-linear thermal generation cost function as $f^T(\cdot)$ and enable load shedding to retain feasibility. Constraint (2b) enforces the active power balance, and constraint (2c) represents the DC power flow equation in the transmission lines. Line flow constraints and active power generation constraints for thermal generators are described in constraints (2d) and (2e), respectively. In constraint (2f), the renewable generation is limited within the uncertain upper bound *w.r.t.* the scenario index ω^T. For brevity, we only give the scenario index to uncertain parameters, but the variables should be associated with the scenario index as well. Besides, constraints (2g) and (2h) represent the ramp-up/down limits, respectively.

C. The Third Stage: DSO's ED Problem

From DSO's perspective, the third stage establishes a distribution-level ED, as shown in (3). Since DER generators are fast-responsive, their UC decisions are not considered. However, the distribution network is reconfigurable [15] and controlled by the ISO in the proposed T-D coordinated market operation. Note that ω_c^D for each c in the third stage is independent of each other and ω^T in the second stage. For each c and ω_c^D:

$$\mathbf{K_3}(\mathbf{x}_2^*, \mathbf{y}^*, \omega_c^D) = \min_{p^D_{g,h,c}, s^D_{\ell,h,c}} \sum_h^H \left\{ \sum_{g^D}^{G^D} f^D(p^D_{g,h,c}) + \sum_{\ell^D}^{L^D} PC^D_{\ell,c} s^D_{\ell,h,c} \right\} \quad \text{(3a)}$$

subject to

$$P^{D,\min}_{g,h,c} \le p^D_{g,h,c} \le P^{D,\max}_{g,h,c}, \quad \forall g^D, \forall h, \quad \text{(3b)}$$

$$-RD^D_{g,c} \le p^D_{g,h+1,c} - p^D_{g,h,c} \le RU^D_{g,c}, \quad \forall g^D, \forall h, \quad \text{(3c)}$$

$$0 \le p^D_{r,h,c} \le P^{D,\max}_{r,h,c}(\omega_c^D), \quad \forall r^D, \forall h, \forall \omega_c^D, \quad \text{(3d)}$$

$$0 \le y^D_{mn,h,c} \le x^D_{mn,h,c} Y^{D,\max}_{mn,h,c}, \quad \forall (m,n) \in F_Q^D, \forall h, \quad \text{(3e)}$$

$$0 \le y^D_{mn,h,c} \le Y^{D,\max}_{mn,h,c}, \quad \forall (m,n) \in F^D \backslash F_Q^D, \forall h, \quad \text{(3f)}$$

$$(p^D_{mn,h,c})^2 + (q^D_{mn,h,c})^2 \le y^D_{mn,h,c} v_{m,h,c}, \ \forall (m,n) \in F^D, \forall h, \quad \text{(3g)}$$

$$x^D_{mn,h,c} \ge 0, \quad \forall (m,n) \in F^D, \forall h, \quad \text{(3h)}$$

$$x^D_{mn,h,c} = 0, \quad \forall n \in A_T^D, \forall h, \quad \text{(3i)}$$

$$x^D_{mn,h,c} + x^D_{nm,h,c} = 1, \quad \forall n \in F^D \backslash F_Q^D, \forall h, \quad \text{(3j)}$$

$$x^D_{mn,h,c} + x^D_{nm,h,c} = z^*_{q,h,c}, \quad \forall n \in F_Q^D, \forall h, \quad \text{(3k)}$$

$$\sum_{n:(m,n)\in F^D} x_{mn,h,c} = 1, \quad \forall m \in A^D \backslash A_T^D, \forall h, \quad \text{(3l)}$$

$\forall n^D, \forall h:$

$$p^D_{n,h,c} = \sum_{k:n\to k} p^D_{nk,h,c} - \sum_{j:j\to n} (p^D_{nj,h,c} - R_{nj} y^D_{nj,h,c}), \quad \text{(3m)}$$

$$q^D_{n,h,c} = \sum_{k:n\to k} q^D_{nk,h,c} - \sum_{j:j\to n} (q^D_{nj,h,c} - X_{nj} y^D_{nj,h,c}), \quad \text{(3n)}$$

$$v_{n,h,c} - v_{m,h,c} \le M(1 - x^D_{nm,h,c}) + (R^2_{mn} + X^2_{mn}) y^D_{mn,h,c} - 2(R_{mn} p^D_{mn,h,c} + X_{mn} q^D_{mn,h,c}), \quad \text{(3o)}$$

$$v_{n,h,c} - v_{m,h,c} \ge -M(1 - x^D_{nm,h,c}) + (R^2_{mn} + X^2_{mn}) y^D_{mn,h,c} - 2(R_{mn} p^D_{mn,h,c} + X_{mn} q^D_{mn,h,c}), \quad \text{(3p)}$$

$$p^D_{n,h,c} = \sum_{g^D|C(g^D)=n} p^D_{g,h,c} + \sum_{i_c|C(i_c)=n} p^*_{i_c,h} + \sum_{r^D|C(r^D)=n} p^D_{r,h,c} - \sum_{\ell^D|C(\ell^D)=n} \left\{ L^D_{\ell,h,c}(\omega_c^D) - s^D_{\ell,h,c} \right\}, \quad \text{(3q)}$$

$$q^D_{n,h,c} = \sum_{g^D|C(g^D)=n} q^D_{g,h,c} + \sum_{r^D D|C(r^D)=n} q^D_{r,h,c} - \sum_{\ell^D|C(\ell^D)=n} q^D_{\ell,h,c}(\omega^D_c), \quad (3r)$$

$$V^{\min}_{n,h,c} \le v_{n,h,c} \le V^{\max}_{n,h,c}, \quad (3s)$$

where $f^D(\cdot)$ represents the piecewise-linear cost function. The constraints for non-reconfigurable branches are identical to the transmission lines in the second-stage problem, while the constraints for reconfigurable branches are with the reconfiguration decisions, as shown in constraints (3e) and (3f).

We cast a SOCP relaxation for power flow modeling. Constraint (3g) is the convex second-order cone (SOC) constraint. Note that we neglect shunt impedances for simplicity. k, j, m are notations for connected buses. We also assume that the distribution networks can well control the voltage and reactive power performance. Hence, we only consider the reactive power in the distribution systems and no reactive power exchange will happen [24]. Future works include that the distribution systems can provide reactive power support to the transmission system, where our proposed model can also be adopted. Also, note that we employ the widely-adopted branch flow model for the SOCP relaxation [25] of the AC power flow. However, other SOCP relaxation techniques, including the bus injection model and its variants [26], could also be adopted in the proposed framework, given that the feasibility region is convex.

To ensure the distribution network's radiality, we also enforce constraints (3h) - (3l). Note that x_{mn} are continuous variables and will be either zero or one to keep arborescence [27], which retains model (3)'s convexity. Upon using this, we restrict the distribution network information in the third stage, which protects the confidentiality of distribution utilities. Other constraints are with the typical branch flow model under SOCP relaxation [28].

III. Decomposition-Based Solution Strategy for the Three-Stage Market Operation

In this section, we provide an efficient solution strategy towards the multi-scale and multi-stage SP problem for T-D coordinated market operation. As the original formulation follows a stage-decomposable structure, it can be readily tackled by the L-shaped method, as adopted in multiple works [17], [29]. However, since the proposed T-D coordinated hierarchy is a three-stage problem, the L-shaped method needs to be nested. In this paper, we tailor a generalized nested decomposition based on the nested L-shaped method to facilitate the solution upon different and individual scenario sets. To be more concise, we use the compact form for equations in this section, and similarly, we only put the scenario indices (ω^T) and (ω^D_c) with the uncertain parameters.

A. ISO's Master Problem

The ISO's UC problem (1), *i.e.*, the first-stage problem, serves as the master problem. The complicating variables in the master problem include the unit commitment decisions, *i.e.* $u_{g,h}, d_{g,h}, i_{g,h}$ (denoted by vector $\mathbf{x}_1$), and the reconfiguration decisions, *i.e.* $z_{q,h,c}$ (denoted by vector $\mathbf{x}_2$). A compact form for the master problem (1) is represented in (4).

$$M = \min_{\mathbf{x}_1,\mathbf{x}_2} \ \mathbf{c}_1^\top \mathbf{x}_1 + \mathbf{0}^\top \mathbf{x}_2 + S^*_t, \quad (4a)$$

subject to

$$\mathbf{A}_1\mathbf{x}_1 + \mathbf{B}_1\mathbf{x}_2 \le \mathbf{b}_1, \quad (4b)$$

$$\mathbf{A}_2\mathbf{x}_1 + \mathbf{B}_2\mathbf{x}_2 = \mathbf{b}_2, \quad (4c)$$

$$S^*_t = \max_{o\in O} \left\{ \alpha^o_t + (\boldsymbol{\beta}^o_{t1})^\top \mathbf{x}_1 + (\boldsymbol{\beta}^o_{t2})^\top \mathbf{x}_2 \right\}, \quad (4d)$$

where S^*_t denotes the maximum of cuts returned from the second-stage subproblem, and constraint (4 d) includes the optimality cuts returned from the second-stage problem with the information of the second and third stages for all iterations.

B. ISO's Subproblem

The second-stage problem (2), *i.e.*, the ISO's subproblem (S_t), models the transmission-level ED, as shown in (5).

$$\forall \omega^T: \quad S_t(\omega^T) = \min_{\mathbf{y}} \ \mathbf{c}_2^\top \mathbf{y} + S^*_d, \quad (5a)$$

subject to

$$\mathbf{K}_1\mathbf{y} = \mathbf{r}_1(\omega^T): \qquad \boldsymbol{\gamma}, \quad (5b)$$

$$\mathbf{H}_1\mathbf{x}^*_1 + \mathbf{A}_3\mathbf{y} \le \mathbf{b}_3: \qquad \boldsymbol{\phi}, \quad (5c)$$

$$S^*_d = \max_{o\in O,c} \left\{ w_c \cdot \left[\alpha^o_d + (\boldsymbol{\beta}^o_{d1})^\top \mathbf{y} + (\boldsymbol{\beta}^o_{d2})^\top \mathbf{x}^*_2 \right] \right\} : \boldsymbol{\mu}, \quad (5d)$$

where $\mathbf{x}^*_1$ and $\mathbf{x}^*_2$ are delivered from the master problem, $\mathbf{y}$ is the second-stage variable vector, including the generation dispatch, load curtailment and line flows, *etc.* $\boldsymbol{\gamma}$, $\boldsymbol{\phi}$ and $\boldsymbol{\mu}$ are the dual vectors, whereas auxiliary variable S^*_d and constraint (5d) formulate the relaxed counterpart of the third-stage problem. To update the optimality cut (4d) in the master problem, we calculate the subgradients as follows.

$$\alpha^o_t = \sum_{\omega^T} \Pr_{\omega^T} \left\{ \boldsymbol{\gamma}^\top \mathbf{r}_1(\omega^T) + \boldsymbol{\phi}^\top \mathbf{b}_3 - \boldsymbol{\mu}^\top [\alpha^o_d + (\boldsymbol{\beta}^o_{d2})^\top \mathbf{x}^*_2 - S^*_d/w_c] \right\}$$

$$\boldsymbol{\beta}^o_{t1} = -\sum_{\omega^T} \Pr_{\omega^T} \mathbf{H}_1^\top \boldsymbol{\phi}; \ \boldsymbol{\beta}^o_{t2} = -\sum_{\omega^T} \Pr_{\omega^T} \boldsymbol{\beta}^\top_{d2} \boldsymbol{\mu}.$$

C. DSO's Subproblem

The third-stage problem (3), *i.e.*, the DSO's subproblem (S_d), which is a convex SOCP model, formulates the distribution-level ED, as shown in the compact form (6). Without loss of generality, we group the affine constraints with the SOC one in constraint (6b), as affine functions are a special case of SOC when $\mathbf{K}_2$ is a zero matrix.

$$\forall c, \forall \omega^D_c: \quad S_d(w_c, \omega^D_c) = \min_{\mathbf{z}} \ \mathbf{c}_3^\top \mathbf{z}, \quad (6a)$$

subject to

$$\|\mathbf{H}_2\mathbf{y}^* + \mathbf{K}_2\mathbf{z} + \mathbf{e}\|_2 \le \mathbf{q}^\top \mathbf{y}^* + \mathbf{p}^\top \mathbf{z} + \mathbf{H}_3\mathbf{x}^*_2 + \mathbf{r_2}(\omega^D_c), \quad (6b)$$

where $\mathbf{x}_2^*$ and $\mathbf{y}^*$ are decisions delivered from the first-stage problem and the second-stage problem, respectively. Then we write the dual form of (6) in (7).

$$S_d^{\text{dual}}(w_c, \omega_c^D) = \max_{\mathbf{u},\mathbf{v}} \ \left[\mathbf{u}^\top(\mathbf{H}_2\mathbf{y}^* + \mathbf{e}) - \mathbf{v}(\mathbf{q}^\top\mathbf{y}^* + \mathbf{H}_3^\top\mathbf{x}_2^* + \mathbf{r}_2(\omega_c^D))\right], \tag{7a}$$

subject to

$$\mathbf{K}_2^\top\mathbf{u} + \mathbf{v}^\top\mathbf{p} = \mathbf{c}_3, \tag{7b}$$

$$\|\mathbf{u}\|_2 \le \mathbf{v}, \tag{7c}$$

where $\mathbf{u}$ and $\mathbf{v}$ are the dual vectors of the Euclidean norm and the SOC, respectively. Note that this dual problem is also a convex SOC and only satisfies the weak duality. However, given the assumption that the nodal voltage constraint is not binding, there exists an interior in the primal and dual feasibility regions and hence Slater's condition holds, which means the strong duality can be retained [30]. Then the subgradients can be calculated for the optimality cut in the ISO's subproblem.

$$\alpha_d^o = \sum_{\omega_c^D} \Pr_{\omega_c^D}[\mathbf{u}^\top\mathbf{e} - \mathbf{v}\cdot\mathbf{r}_2(\omega_c^D)];$$

$$\boldsymbol{\beta}_{d1}^o = \sum_{\omega_c^D} \Pr_{\omega_c^D}[\mathbf{u}^\top\mathbf{H}_2 - \mathbf{v}\cdot\mathbf{q}];\ \boldsymbol{\beta}_{d2}^o = -\sum_{\omega_c^D} \Pr_{\omega_c^D}\mathbf{v}\cdot\mathbf{H}_3^\top;$$

D. Generalized Nested Decomposition Algorithm

To cope with the proposed T-D coordinated hierarchy, we devise a generalized nested decomposition (GND) algorithm. The general procedure is demonstrated in Fig. 2. In particular, with an initial feasible point of the complicating variables determined from the initial master problem, the problem in each stage can deliver the optimal complicating variables back to the previous stage problem via creating the lower bounding affine cuts by subgradients. If any recourse is infeasible, a feasibility cut will be returned to the upper stages and repeat the loop in Fig. 2 [31]. For more details about the feasibility check, please see Appendix A. These cuts formulate the relaxed counterparts of the respective subproblems and hence decompose the overall problem. Each subsystem in the second-stage and third-stage can keep its own scenario set and return its own single cut. For multiple distribution systems in the third stage, they return multiple cuts to the second-stage problem based on the single cut from respective scenario sets. From the third stage to the second stage, it is promising to adopt the cut sharing [32] between each second-stage scenario to facilitate the solution, which does not compromise exactness. Besides, The non-anticipativity in any current stage is implicit since we only have one copy of wait-and-see variables in the previous stage. For the convergence proof, we first give the following assumption.

Assumption 1. *1) The second-stage and third-stage problems are convex and have complete recourse; 2) all recourse stages have finite support; 3) all the scenario sets are stage-wise independent.*

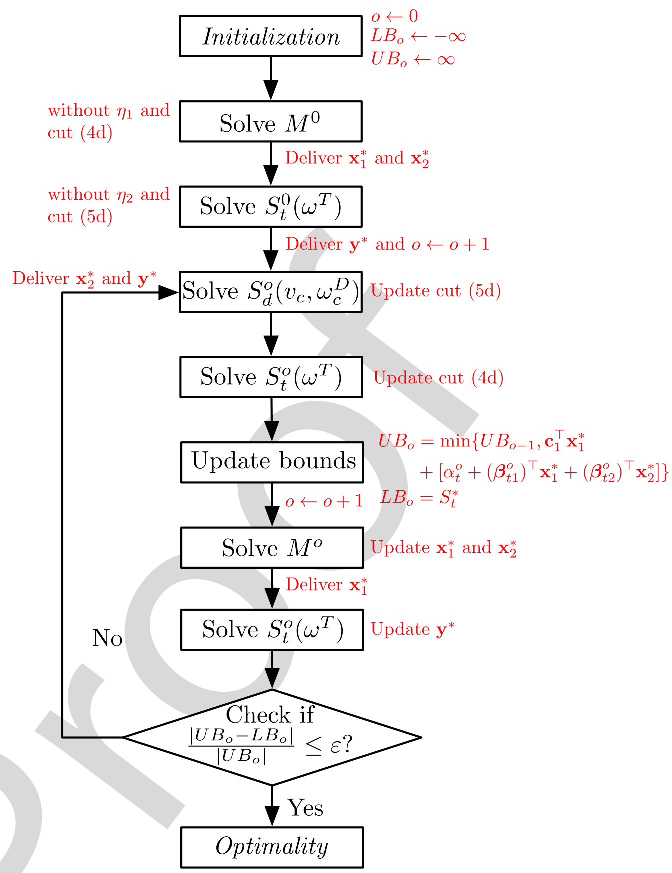


Fig. 2. The GND algorithm workflow.

These are all mild assumptions in the current electricity market in terms of system operations. For the first item, complete recourse means that the subsequent-stage problems are feasible given any first-stage variable. Current industrial ED problems are mostly convex and feasible since the linearized cost curves are adopted and all operating constraints are appropriately relaxed with penalty terms based on the ISO's practice [33]. For the second item, finite support means finite realizations of scenarios for the problems, which is a basic assumption for all scenario-based stochastic programming [34]. For the third item, the transmission system should have an independent and different scenario set with DSOs, as well as scenario sets between DSOs. In conclusion, these assumptions are mild in practical power systems. Then based on [20], consider the following proposition.

Proposition 1: Provided with Assumption 1, constraints (4 d) and (5 d) provide accurate approximations for the second stage and the third stage, respectively.

Proof: From problem S_d to S_t, subproblem (7) is solved for each ω_c^D, yielding corresponding optimal dual solutions. As problem S_d is convex and has complete recourse, the Slater's condition holds. Consider $\mathcal{Q}(\mathbf{x}_2, \mathbf{y}, \omega_c^D)$ as the third-stage problem's dual:

$$\mathcal{Q}(\mathbf{x}_2, \mathbf{y}, \omega_c^D) = \mathbf{u}^\top(\mathbf{H}_2\mathbf{y} + \mathbf{e}) - \mathbf{v}(\mathbf{q}^\top\mathbf{y} + \mathbf{H}_3^\top\mathbf{x}_2^* + \mathbf{r}_2(\omega_c^D)).$$

Based on the convexity of S_d, the subgradients provide:

$$\mathcal{Q}(\mathbf{x}_2, \mathbf{y}, \omega_c^D) \geq \mathbf{u}^\top(\mathbf{H}_2\mathbf{y} + \mathbf{e}) - \mathbf{v}(\mathbf{q}^\top \mathbf{y} + \mathbf{H}_3^\top \mathbf{x}_2^* + \mathbf{r}_2(\omega_c^D)).$$

Taking the expectation according to the distribution of uncertainty ω_d, and using the subgradient notations, we have:

$$\mathcal{Q}(\mathbf{x}_2, \mathbf{y}) \geq \alpha_d + \boldsymbol{\beta}_{d1}^\top \mathbf{y} + \boldsymbol{\beta}_{d2}^\top \mathbf{x}_2^*.$$

Thus, the affine function below for iteration o gives an exact outer linearization for problem S_d and is returned to S_t:

$$\mathcal{F}(\mathbf{y} : \mathbf{y}^o) = \alpha_d^o + (\boldsymbol{\beta}_{d1}^o)^\top \mathbf{y} + (\boldsymbol{\beta}_{d2}^o)^\top \mathbf{x}_2^*$$

When there are multiple S_d in the third stage, each of them returns the corresponding affine cut to S_t. Thus, the optimality of the problem follows from the complete recourse of all subsystems, and the approximation is exact if and only if $\mathcal{Q}^o = \max_o\{\mathcal{F}(\mathbf{y} : \mathbf{y}^o)\}$, *i.e.*, constraint ([illegible]). Proof is similar for the approximation of S_t in M, *i.e.*, constraint ([illegible]). ■

Remark 1: As the approximation of constraints ([illegible]) and ([illegible]) is accurate for recourses, based on Remark 4 and Remark 5 in [20], the cuts from subsequent stages formulate accurate approximations only containing the subgradient information. Hence, provided with Assumption 1, when the cuts provide exact approximations of later-stage problems, after a sufficiently large number of forward and backward iterations, the algorithm will converge to a solution within a sufficiently small gap from the global optimality with probability 1.

Remark 2: The assumption that the nodal voltage constraint is not binding is mild [30] because the voltage deviations in distribution networks should be well controlled by voltage regulators. It is unnecessary to consider the local voltage stability issue in a market clearing problem from the ISO's perspective since there is no such a market product. A Volt/VAR optimization problem could be solved by local utilities to tackle this issue once the hierarchical market is cleared [35].

Remark 3: Note that in our case, M is a MIP. Hence, the final optimal solution satisfies the potential MIP gap without loss of generality as the affine cuts are created from convex subproblems. Practically, we admit that the GND's algorithmic gap is hard to achieve zero, and together with the MIP gap, the final solution might not be optimal. However, the theoretical global optimality remains if convexity holds for recourses.

Remark 4: The first three blocks in Fig. 2 are for the initialization. However, if we can find an excellent warm-start initial point, the algorithm will converge faster. Besides, since the cut herein gives an affine approximation, cut sharing and cut selection can further contribute to the acceleration.

Remark 5: The ISO's master problem receives only the cut information consisting of subgradients from the lower-level problems. With only the subgradients, one cannot recover the full network information, which prevents the ISO from having the network information and thus protects the confidentiality.

IV. Numerical Experiments

To test the efficacy of the proposed T-D coordinated market operation and the solution strategy, we carry out case studies on a modified *Tran6+Dist7+Dist9* test system from [36] and a *Tran118+Dist30×5* from [4] for day-ahead UC & ED problems. We solve all of the experiments by Gurobi on a Windows PC with quad-core Intel i7-6700 CPU and 8 GB of RAM.

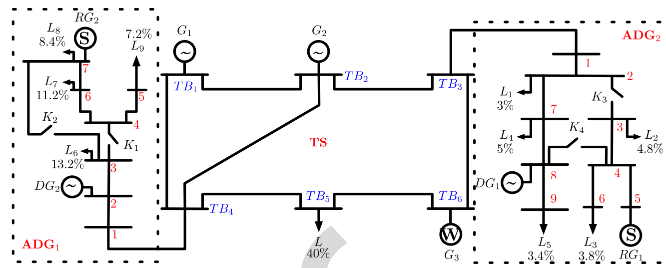


Fig. 3. *Tran6+Dist7+Dist9* system topology [36].

A. Tran6+Dist7+Dist9 Test Case

We depict the system topology in Fig. 3, which consists of one transmission system (TS) and two radial ADNs. The original test system [36] did not consider the resistance and reactive demand/generation limits. In this paper, they are obtained via a fixed R/X ratio (assume all distribution branches are with the same configuration and R/X ratio is 1:2) and power factor (assumed to be V-I 0.95 lagging) using the existing data. We also have four switches K_1˜K_4 installed to enable the network reconfiguration of each ADN while the switch can be reconfigured every 8 hours.

For uncertainties, one wind generator G_3 and two PV units RG_1 in ADN$_1$ and RG_2 in ADN$_2$ are added in the system, whose power outputs are uncertain and sampled from three different historical datasets obtained from [37]. Note that for the renewables, we scale the historical data to a distribution level, with the maximum output as 25MW. We also slightly reduce the capacity of DG_1 and DG_2 in the original data. The hourly load profile is also assumed to be stochastic that respects the sampled demand scenarios and is distributed according to the factor shown in Fig. 3. The penalty cost for the load shedding is \$1000/kWh to make the ADN decisions comparable with the TS.

First, 150, 200, and 200 scenarios for the renewable generators G_3, RG_1, and RG_2 are sampled, and 180 scenarios for the demand are sampled using Monte Carlo sampling from the historical data, respectively. The number of joint renewable and demand scenarios is very high and could jeopardize the computational efficiency. Then, we employ the Fast Forward/Backward approach [38] in the GAMS SCENRED toolbox to reduce the numbers of the joint renewable and load scenarios in the three grids. Fig. 4 shows the sensitivity analysis of scenario number reduction. We have a base scenario reduction Δ that has 5, 10, and 12 scenarios for the TS, ADN_1, and ADN_2, respectively, and we linearly increase the number of the desired scenarios as shown in the x-axis of Fig. 4. It can be seen that using $10 \times \Delta$ scenarios has a solution gap of 2.12% compared with using $50 \times \Delta$ scenarios. Hence, we argue that using $10 \times \Delta$ scenarios achieves the tradeoff between the accuracy and complexity. Note that the scenarios are *i.i.d.* w_c for the two distribution systems

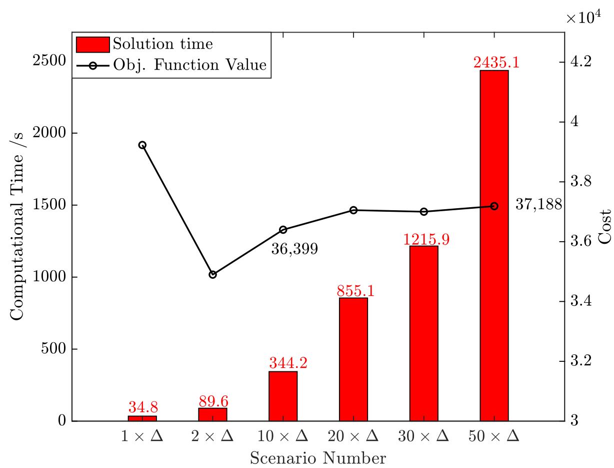


Fig. 4. Sensitivity analysis on the scenario number.

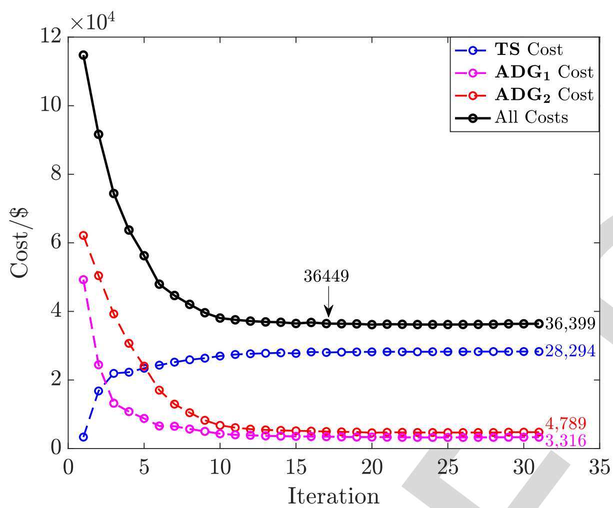


Fig. 5. Algorithmic performance in the *Tran6+Dist7+Dist9* case.

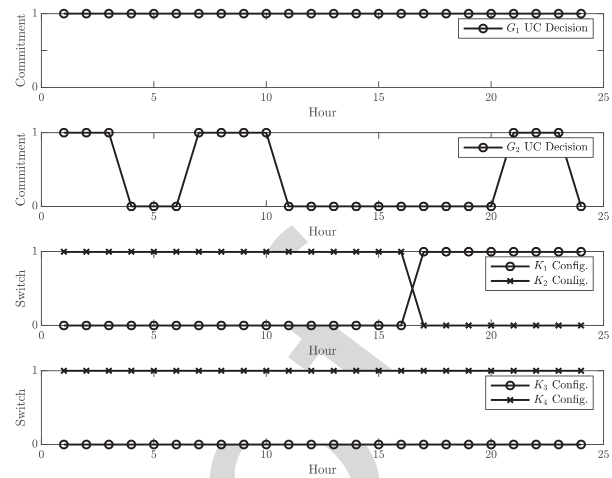


Fig. 6. *Tran6+Dist7+Dist9* commitment and reconfiguration solutions.

TABLE I
COMPARATIVE ANALYSIS ON ISOLATION AND COORDINATION

$IUCED_1$	**TS**	Total Costs:	$38,943.87
		ADN$_1$	**ADN$_2$**
	Power Mismatch	31.1%	37.8%
	Total Costs	$4,324.82	$5,609.71
	Received LMP	$16.83/MWh	$16.83/MWh
$IUCED_2$	**TS**	Total Costs:	$27,835.94
		ADN$_1$	**ADN$_2$**
	Power Mismatch	5.8%	8.4%
	Total Costs	$3,658.62	$5,119.84
	Received LMP	$15.47/MWh	$15.47/MWh
$TDC-UCED$	**TS**	Total Costs:	$28,294.12
		ADN$_1$	**ADN$_2$**
	Power Mismatch	0%	0%
	Total Costs	$3,316.20	$4,789.09
	Received LMP	$14.73/MWh	$14.73/MWh

are set equal in this study, and the convergence criterion ε is 0.01%.

Fig. 5 shows the algorithmic performance in this case. The GND algorithm converges in the 31th iteration. The total load shedding upon convergence is zero. The costs of the two ADNs start at extremely high levels in that the TS draws a large amount of energy from the distribution sides to reduce its power outputs in the first iteration. This phenomenon is because there is no restriction on the power exchange when the two ADNs have to feed the TS even with high load shedding in the local area. However, the injections from the TS to ADNs finally become incentivized as the operation cost of the conventional units in the TS is much lower than the load shedding price in ADNs, and the DER units in ADNs cannot fully support the local demands. Note that the solution reaches $36,449 in iteration 17, and it is within a gap of 0.14%, as shown in Fig. 5. Thus, if we have a slightly higher tolerance on the optimality, we can greatly reduce the computation time.

Fig. 6 provides the unit commitment decisions as well as the network reconfiguration solutions of the master problem upon convergence. G_1 is committed ON throughout the time due to its low generation cost and large capacity. G_2 is mainly used during the twilight time since the PV generation is low at that time, but the energy consumption is high. ADN_1 changes its topology once in this operation based on the optimized switch decisions, whereas ADN_2 remains the original topology. The topology change in ADN_1 is mainly due to the capacity of lateral 3-7 at hour 16 cannot support the increasing demand of load L_8 and L_9, and lateral 3-4 can relieve the congestion pressure by bifurcating the flow to two laterals, i.e., 4-5 and 4-6, towards the demand nodes.

Besides, to illustrate the necessity of the T-D coordination, we carry out the other two comparative experiments on isolated UC&ED and report the results in Table I in comparison with the results of the T-D coordinated case. $IUCED_1$ mode means that TS forecasts the boundary power demand as all demands in the ADN, and $IUCED_2$ mode forecasts it as the maximum forecast demand minus the maximum generation outputs in the ADNs. $TDC-UCED$ stands for the proposed T-D coordinated UC&ED setting. It can be observed from the table that in both isolated modes, there exists a notable power mismatch, even if the $IUCED_2$ mode can fairly approximate the coordination used by some current industry practice [4]. A small power mismatch, however, can still raise severe conditions for system

TABLE II
MAXIMUM SOCP GAPS OF THE TWO ADNS (P.U.)

Network	$\mathbf{ADN}_1$	$\mathbf{ADN}_2$
SOCP gap	0.0064	0.0037

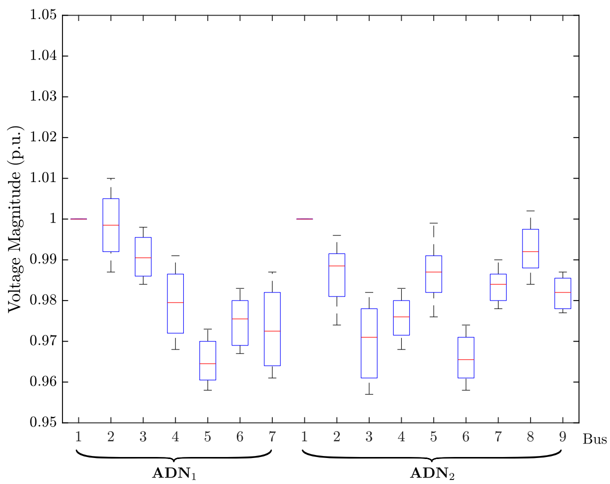


Fig. 7. Boxplot for voltage magnitudes of the two networks.

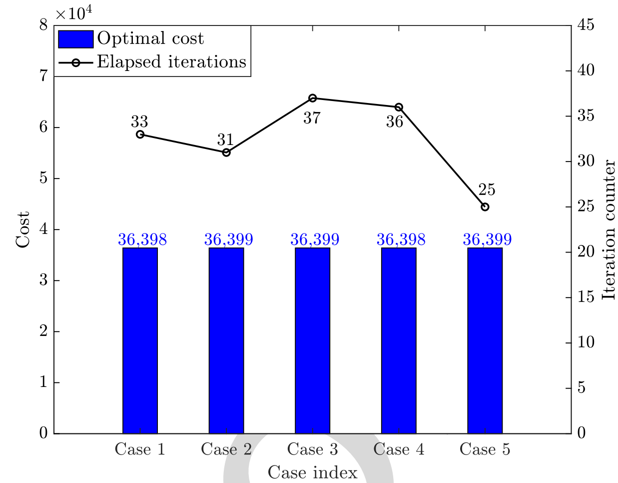


Fig. 8. Sensitivity analysis on the initial values.

operators as it influences the system stability. We also notice that the received LMP in both ADNs is decreasing, which is due to the relieved line congestion by the distribution network reconfiguration considered in the T-D coordinated framework. This observation defends the necessity of adopting the network reconfiguration-embedded T-D coordination in the future market.

We further conduct two additional analyses to illustrate the algorithmic performance.

1). Exactness of the SOCP relaxation

While the branch flow model adopted in this paper has been recognized as a nearly exact relaxation of the actual AC power flow [25], we further demonstrate its exactness by the SOCP gap, which is also a well-known index for evaluating the SOCP performance [39]. Table II tabulates the maximum SOCP gap values throughout scenarios in the two ADNs.

Generally, the smaller the gap is, the higher exactness the formulation achieves [39]. It can be shown from Table II that our proposed model is accurate enough. Though this approximation still cannot fully replace a complete AC power flow model, it gives an excellent starting point for such analyses run by ISOs with superior computational efficiency. Fig. 7 also shows that under normal system conditions, the nodal voltage constraints are non-binding, which supports the strong duality of the SOCP-based subproblem.

2). Sensitivity analysis for initial values

To technically support Remark 4, we conduct a sensitivity analysis for initial values in the algorithm. Note that we are particularly interested in the distribution reconfiguration variables' initial values since the unit commitment variables will render no infeasibility in subsequent stages. We use five cases of initial values for the sensitivity analysis.

- *Case 1.* Throughout the horizon, $K_1 = 1$ and $K_2 = 0$, $K_3 = 1$ and $K_4 = 0$.
- *Case 2.* Throughout the horizon, $K_1 = 0$ and $K_2 = 1$, $K_3 = 0$ and $K_4 = 1$.
- *Case 3.* Throughout the horizon, $K_1 = 1$ and $K_2 = 1$, $K_3 = 0$ and $K_4 = 0$.
- *Case 4.* Throughout the horizon, $K_1 = 0$ and $K_2 = 0$, $K_3 = 1$ and $K_4 = 1$.
- *Case 5.* The optimal configuration settings obtained by the simulation.

It is straightforward that *Case 1* and *Case 2* give a feasible set of initial values, while *Case 3* and *Case 4* give an infeasible one. *Case 5* serves as a comparison with a perfect initial value. Fig. 8 depicts their performance. It is shown that all five cases converge to the optimal solution, whereas *Case 1* and *Case 2* spend fewer iterations than *Case 3* and *Case 4*, but *Case 5* achieves the fastest performance. The additional iterations spent in *Case 3* and *Case 4* are mostly feasibility-check iterations, where distribution systems return feasibility cuts. This experiment shows that a feasible initial value leads to a better performance, whereas a perfect initialization would further save the computational time.

B. *Tran118+Dist34×5 Test Case*

To test the scalability of the proposed method, we carry out experiments on a *Tran118+Dist34×5* test case modified from [4], which consists of one IEEE 118-bus transmission system and 5 IEEE 34-bus distribution systems. A similar data modification such as R/X ratio and power factor is conducted to make the original system suitable for our study. Moreover, we also perform similar scenario generation and reduction for the additional ADNs (only 20 scenarios are considered in each ADN, and 10 scenarios are considered in the TS). The load shedding cost is set as \$1000/kW in ADNs to keep comparability.

Three cases with different stopping criteria ε for the GND algorithm, namely 3%, 1%, and 0.1%, are set up to test the algorithmic performance in a large system. Fig. 9 shows the comparative results. The boundary purchase here is defined by the sum of multiplications between boundary LMPs, which are

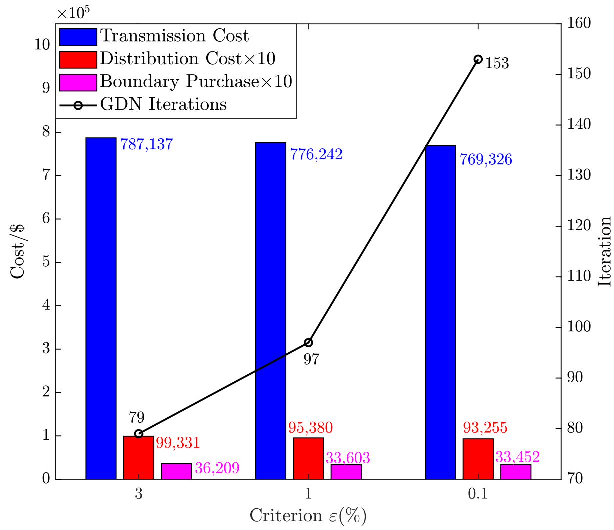


Fig. 9. Comparative analysis with different stopping criteria.

the dual values of the boundary nodal balance constraints in the second stage, and the power injections from the TS. We multiply the total ADN cost and the purchase cost by 10 to make costs comparable. It can be observed that the convergence speed greatly depends on the choice of the stopping criterion. And the algorithm spends a great number of iterations to close the gap between 1% and 0.1%, which is also a feature of optimality conditioned algorithms.

The CPU time for the case when $\varepsilon = 1\%$ is around 4.5 hours, which is acceptable considering such a stochastic, complex, and accurate modeling for both transmission and distribution systems. However, directly running the corresponding deterministic equivalent SP problem drains up the RAM and returns no solution. This observation defends the necessity of using the proposed decomposition technique for acceleration. Note that with a more powerful simulation platform, such as high-performance CPU clusters, solving all the second-stage and third-stage subproblems with each scenario realization could be implemented in parallel, which can significantly reduce the computational time.

V. Conclusion

To help the ISO with tackling the increasing system intermittency from the active distribution system and reduce the boundary power mismatch, we propose a T-D coordinated market paradigm including UC and ED. This work is mainly from the ISO's perspective to optimize its daily operation considering the DSO's performance. A generalized nested decomposition method is tailored and efficiently decomposes the problem, which greatly facilitates the solution. The theoretical convergence proof of this algorithm is also articulated. Numerical experiments corroborate the effectiveness of the proposed strategy and show the necessity of the T-D coordination.

Future studies include exploring more effective cut selection from the convex relaxation of subproblems, *e.g.*, the Lagrangian dual cut, to reduce the computational time. Besides, with the increasing DER installation in distribution systems, similar coordinations between DSOs and microgrids await in-depth investigation.

Appendix A
Feasibility Check

We provide two alternatives for eliminating the infeasibility issue.

A. First Alternative: Feasibility-Check Subproblem

Upon infeasibility, a feasibility-check subproblem for the third-stage problem will need to be solved and return a feasibility cut to the second stage. Note that the second stage will not incur any infeasibility since the transmission ED problem has no integer and the load shedding variables enforce the feasibility of operational constraints. Hence, we only formulate the third-stage feasibility-check problem, as shown in (A-1).

$$\forall c: \quad \min_{j_{mn,h,c}} \sum_{h}^{H} \sum_{mn} j^{+}_{mn,h,c} + j^{-}_{mn,h,c}, \tag{A-1a}$$

subject to

$$\text{Constraint (3b)-(3˜g), (3m-3˜s)}, \tag{A-1b}$$

$$x^{D}_{mn,h,c} \geq 0, \qquad \forall (m,n) \in F^{D}, \forall h, \tag{A-1c}$$

$$x^{D}_{mn,h,c} = 0, \qquad \forall n \in A^{D}_{T}, \forall h, \tag{A-1d}$$

$$x^{D}_{mn,h,c} + x^{D}_{nm,h,c} = 1 + j^{+}_{mn,h,c} - j^{-}_{mn,h,c}, \qquad \forall n \in F^{D} \backslash F^{D}_{Q}, \forall h, \tag{A-1e}$$

$$x^{D}_{mn,h,c} + x^{D}_{nm,h,c} = z^{*}_{q,h,c} + j^{+}_{mn,h,c} - j^{-}_{mn,h,c}, \qquad \forall n \in F^{D}_{Q}, \forall h, \tag{A-1f}$$

$$\sum_{n:(m,n)\in F^{D}} x_{mn,h,c} = 1, \qquad \forall m \in A^{D} \backslash A^{D}_{T}, \forall h, \tag{A-1g}$$

$$j^{+}_{mn,h,c}, j^{-}_{mn,h,c} \geq 0, \qquad \forall (m,n) \in F^{D}, \forall h, \tag{A-1h}$$

If we write in a compact form and for each scenario, we attain (A-2).

$$\forall c, \forall \omega^{D}_{c}: \quad S_{d}(w_{c}, \omega^{D}_{c}) = \min_{\mathbf{j}} \ \mathbf{j}, \tag{A-2a}$$

subject to

$$\|\mathbf{H}_2 \mathbf{y}^{*} + \mathbf{K}_2 \mathbf{z} + \mathbf{e}\|_2 \leq \mathbf{q}^{\top}\mathbf{y}^{*} + \mathbf{p}^{\top}\mathbf{z} + \mathbf{k}^{\top}\mathbf{j} + \mathbf{H}_3 \mathbf{x}^{*}_2 + \mathbf{r_2}(\omega^{D}_{c}), \tag{A-2b}$$

where $j^{+}_{mn,h,c}$ and $j^{-}_{mn,h,c}$ are positive slack variables imposed on the distribution reconfiguration constraints, $\mathbf{j}$ is their vector form. After we solve (A-2), the feasibility cut can be formulated as in (A-3).

$$\mathbf{u}^{\top}(\mathbf{H}_2 \mathbf{y}^{*} + \mathbf{e}) - \mathbf{v}(\mathbf{q}^{\top}\mathbf{y}^{*} + \mathbf{k}^{\top}\mathbf{j} + \mathbf{H}^{\top}_3 \mathbf{x}^{*}_2 + \mathbf{r}_2(\omega^{D}_{c})) \leq 0, \tag{A-3}$$

Note that any infeasible set of distribution reconfigurations will incur infeasibility in all scenarios of this distribution network. Hence, upon encountering an infeasible subproblem instance, a feasibility cut will be generated and all other scenarios will be terminated because (A-3) contains enough feasibility information. The feasibility cut will be added into the second-stage problem. Afterwards, the second-stage problem will be solved and add a cut containing both feasibility and optimality information to the first stage. Then the iteration continues from the first stage. This procedure is called Fast Forward and Fast Back [31], which also coincides with Fig. 2. One of this method's benefits is that distribution networks can return their feasibility cut individually without compromising other distribution systems' optimality.

B. Second Alternative: Always-Feasible Third-Stage Problem

As discussed in [40], we can simply replace the constraints in problem (3) with the ones in problem (A-1) and add a huge penalty for these slack variables in the objective function, which makes the third-stage problem always feasible. Again, as shown in (A-1), not all operational constraints should be relaxed but only the ones with integer variables as they are the potential sources of infeasibility.

Note that, though without a concrete proof, evidence has been found that the Second Alternative appears to be more computationally efficient than the First Alternative [40] and corresponds to the ISO's practice [33], but both methods are able to eliminate the infeasibility issue.

References

[1] S. Yin, J. Wang, and F. Qiu, "Decentralized electricity market with transactive energy - A path forward," *The Electricity J.*, vol. 32, no. 4, pp. 7–13, 2019.

Q2 [2] "Coordination of transmission and distribution operations in high distributed energy resource electric grid," Tech. Rep. 17-IEPR-12, CAISO, Jun. 2017.

Q3 [3] A. Zegers and H. Brunner, "TSO-DSO Interaction: An overview of current interaction between transmission and distribution system operators and an assessment of their cooperation in smart grids," Tech. Rep., ISGAN, Sep. 2014.

[4] Z. Li, Q. Guo, H. Sun, and J. Wang, "Coordinated economic dispatch of coupled transmission and dstribution systems using heterogeneous decomposition," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4817–4830, Nov. 2016.

[5] Z. Li, Q. Guo, H. Sun, and J. Wang, "A new LMP-Sensitivity-Based heterogeneous decomposition for transmission and distribution coordinated economic dispatch," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 931–941, Mar. 2018.

[6] P. Li, Q. Wu, M. Yang, Z. Li, and N. Hatziargyriou, "Distributed distributionally robust dispatch for integrated transmission-distribution systems," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 1193–1205, Mar. 2021.

[7] R. Roofegari nejad, W. Sun, and A. Golshani, "Distributed restoration for integrated transmission and distribution systems with DERs," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4964–4973, Nov. 2019.

[8] J. Zhao, H. Wang, Y. Liu, Q. Wu, Z. Wang, and Y. Liu, "Coordinated restoration of transmission and distribution system using decentralized scheme," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 3428–3442, Sep. 2019.

[9] M. K. Arpanahi, M. E. H. Golshan, and P. Siano, "A comprehensive and efficient decentralized framework for coordinated multiperiod economic dispatch of transmission and distribution systems," *IEEE Syst. J.*, pp. 1–12, 2020, doi: 10.1109/JSYST.2020.3009750.

[10] A. Nawaz and H. Wang, "Stochastically coordinated transmission and distribution system operation with large-scale wind farms," *CSEE J. Power Energy Syst.*, pp. 1–10, 2020.

Q4 [11] "Distribution System Operator (DSO) Models for Utility Stakeholders," Tech. Rep., Black & Veatch Management Consulting, Jan. 2020.

[12] "ERCOT Nodal Protocols Section 4: Day-Ahead Operations," Tech. Rep., ERCOT, 2020.

[13] S. Yin, J. Wang, Z. Li, and X. Fang, "State-of-the-art short-term electricity market operation with solar generation: A review," *Renewable Sustain. Energy Rev.*, vol. 138, 2021, Art. no. 110647.

[14] M. R. Dorostkar-Ghamsari, M. Fotuhi-Firuzabad, M. Lehtonen, A. Safdarian, and A. S. Hoshyarzade, "Stochastic operation framework for distribution networks hosting high wind penetrations," *IEEE Trans. Sustain. Energy*, vol. 10, no. 1, pp. 344–354, Jan. 2019.

[15] Y. Zhang, J. Wang, and Z. Li, "Uncertainty modeling of distributed energy resources: Techniques and challenges," *Current Sustain./Renewable Energy Rep.*, Apr. 2019.

[16] J. Wang, M. Shahidehpour, and Z. Li, "Security-constrained unit commitment with volatile wind power generation," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1319–1327, Aug. 2008.

[17] A. Papavasiliou, Y. Mou, L. Cambier, and D. Scieur, "Application of stochastic dual dynamic programming to the real-time dispatch of storage under renewable supply uncertainty," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 547–558, Apr. 2018.

[18] Y. Liu, R. Sioshansi, and A. Conejo, "Multistage stochastic investment planning with multiscale representation of uncertainties and decisions," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 781–791, Jan 2018.

[19] H. Shuai, J. Fang, X. Ai, Y. Tang, J. Wen, and H. He, "Stochastic optimization of economic dispatch for microgrid based on approximate dynamic programming," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2440–2452, May 2019.

[20] A. Shapiro, "Analysis of stochastic dual dynamic programming method," *Eur. J. Oper. Res.*, vol. 209, no. 1, pp. 63–72, 2011.

[21] J. Zou, S. Ahmed, and X. A. Sun, "Multistage stochastic unit commitment using stochastic dual dynamic integer programming," *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 1814–1823, May 2019.

[22] M. N. Hjelmeland, J. Zou, A. Helseth, and S. Ahmed, "Nonconvex medium-term hydropower scheduling by stochastic dual dynamic integer programming," *IEEE Trans. Sustain. Energy*, vol. 10, no. 1, pp. 481–490, Jan. 2019.

[23] M. Shahidehpour, T. Ding, Q. Ming, C. Huang, Z. Wang, and P. Du, "Multi-period active distribution network planning using multi-stage stochastic programming and nested decomposition by SDDIP," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 2281–2292, May 2021.

[24] M. Bragin, Y. Dvorkin, and A. Darvishi, "Toward coordinated transmission and distribution operations," in *Proc. IEEE Power Energy Soc. General Meeting*, 2018, pp. 1–5.

[25] 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, Aug. 2013.

[26] B. Kocuk, S. S. Dey, and X. A. Sun, "Strong SOCP relaxations for the optimal power flow problem," *Operations Res.*, vol. 64, no. 6, pp. 1177–1196, 2016.

[27] J. A. Taylor and F. S. Hover, "Convex models of distribution system reconfiguration," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1407–1413, Aug. 2012.

[28] M. R. Dorostkar-Ghamsari, M. Fotuhi-Firuzabad, M. Lehtonen, and A. Safdarian, "Value of distribution network reconfiguration in presence of renewable energy resources," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 1879–1888, May 2016.

[29] M. Nick, R. Cherkaoui, and M. Paolone, "Optimal planning of distributed energy storage systems in active distribution networks embedding grid reconfiguration," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1577–1590, Mar. 2018.

[30] H. Zhang, S. J. Moura, Z. Hu, W. Qi, and Y. Song, "Joint PEV charging network and distributed PV generation planning based on accelerated generalized benders decomposition," *IEEE Trans. Transp. Electrific.*, vol. 4, no. 3, pp. 789–803, Sep. 2018.

[31] J. Murphy, "Benders, nested benders and stochastic programming: An intuitive introduction," 2013.

[32] G. Infanger and D. Morton, "Cut sharing for multistage stochastic linear programs with interstage dependency," *Math. Program.*, vol. 75, pp. 241–256, 1996. Q5

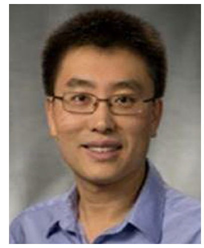
[33] "Business practices manual: Energy and operating reserve markets - attachment B: Day-ahead energy and operating reserve market software formulations and business logic," Tech. Rep. BPM-002-r19, MISO, 2019.

[34] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*. Springer, Incorporated, 2nd ed., 2011.

[35] Y. Zhang, X. Wang, J. Wang, and Y. Zhang, "Deep reinforcement learning based Volt-VAR optimization in smart distribution systems," *IEEE Trans. Smart Grid*, vol. 12, no. 1, pp. 361–371, Jan. 2021.

[36] A. Kargarian and Y. Fu, “System of systems based security-constrained unit commitment incorporating active distribution grids,” *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2489–2498, Sep. 2014.
[37] NREL, “https://www.nrel.gov/analysis/data-products.html.”
[38] J. Dupačová, N. Gröwe-Kuska, and W. Römisch, “Scenario reduction in stochastic programming,” *Math. Program.*, vol. 95, pp. 493–511, Mar. 2003.
[39] 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, pp. 985–995, Jul. 2016.
[40] A. Nasri, S. J. Kazempour, A. J. Conejo, and M. Ghandhari, “Network-constrained AC unit commitment under uncertainty: A. benders’ decomposition approach,” *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 412–422, Jan. 2016.

Shengfei Yin (Student Member, IEEE) received the B.Eng. degree in electrical engineering from Hunan University, Changsha, China, in 2016 and the M.S. degree in electrical engineering from the Illinois Institute of Technology, Chicago, IL, USA, in 2017. He is currently working toward the Ph.D. degree with Southern Methodist University, Dallas, TX, USA. His research interests include operations research and data science with applications on the electricity market. He is a Reviewer of the *IEEE Transactions on Sustainable Energy* and *IEEE Transactions on Vehicular Technology*.



Jianhui Wang (Fellow, IEEE) is currently a Professor with the Department of Electrical and Computer Engineering, Southern Methodist University, Dallas, TX, USA. He has authored and/or coauthored more than 300 journal and conference publications, which have been cited for more than 25000 times by his peers with an H-index of 81. He was invited to give tutorials and keynote speeches at major conferences including IEEE ISGT, IEEE SmartGridComm, IEEE SEGE, IEEE HPSC and IGEC-XI.

He is the past Editor-in-Chief of the *IEEE Transactions on Smart Grid* and an IEEE PES Distinguished Lecturer. He is also the Guest Editor of a *Proceedings of the IEEE* special issue on power grid resilience. He was the recipient of the IEEE PES Power System Operation Committee Prize Paper Award in 2015, the 2018 Premium Award for Best Paper in IET Cyber-Physical Systems: Theory and Applications, and the Best Paper Award in the *IEEE Transactions on Power Systems* in 2020. He is a Clarivate Analytics highly cited Researcher for production of multiple highly cited papers that rank in the top 1% by citations for field and year in Web of Science (2018–2020).

Harsha Gangammanavar received the M.S. degree in electrical engineering and the Ph.D. degree in operations research from The Ohio State University, Columbus, OH, USA. He is currently an Assistant Professor with the Department of Engineering Management, Information, and Systems, Southern Methodist University, Dallas, TX, USA. In the past, he was a Postdoctoral Fellow of industrial engineering with Clemson University, Clemson, SC, USA and Visiting Assistant Professor with the Daniel J. Epstein Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, CA, USA. His research interests include development of stochastic optimization models, algorithms, and their application in power systems planning and operations.