October 29, 2018

Abstract

Nomenclature

Indices

t	a time period (time interval from the beginning through the end of a period)
g	TS network generator index
i	DS network bus number index

Parameters

T	length of the planning horizon
G	total number of generators in the TS
$\overline{R}_g$	TS generator upper ramp rate limit
$\underline{R}_g$	TS generator lower ramp rate limit
$\overline{P}_g$	TS generator generation upper bound
$\underline{P}_g$	TS generator generation lower bound
$Line$	TS line limit
C_g	TS generater g cost
C_g^r	TS generater g reserve cost
C_g^c	TS generator g commitment cost
GSF	TS generation shift factor matrix
W^f	forecasted wind power

W^{up} positive wind deviation from the forecast

W^{dn} negative wind deviation from the forecast

N_b number of buses in the TS

N_b^d number of buses in the DS

L_t TS load power vector in hour t

$\underline{L}_i^d$ DS dispatchable load lower bound at bus i

$\overline{L}_i^d$ DS dispatchable load upper bound at bus i

L_i^{ie} DS inelastic load in time t at bus i

$\overline{P}_i$ DS power flow to bus i upper bound

$\underline{P}_i$ DS power flow to bus i lower bound

C_i^{d1} DS DG generation linear cost at bus i

C_i^{d2} DS DG generation quadratic cost at bus i

C^d DS utility for consuming dispatchable load

C^{dr1} DS linear demand response cost

C^{dr2} DS quadratic demand response cost

$\overline{P}_i^d$ DS DG generation upper bound at bus i

$\underline{P}_i^d$ DS DG generation lower bound at bus i

$\underline{S}^(dr)$ DR scale factor

Variables

$r_{g,t}^{up}$ TS generator g upward reserve in hour t

$r_{g,t}^{dn}$ TS generator g downward reserve in hour t

p_t^{dr} demand response price in hour t

$p_{g,t}$ TS generator g generation in hour t

p_t^{im} DS imported power in hour t

$l_{i,t}^d$ DS dispatchable load in hour t at bus i

dr_t^{dn} DS downward DR in hour t at bus i

$p^{d}_{i,t}$	DS DG power in hour t at bus i
$P_{i,t}$	DS power flow to bus i in hour t
dr^{up}_{t}	DS upward DR in hour t at bus i
c^{im}_{t}	DS energy import price in hour t
p^{inj}_{t}	bus power injection vector in hour t
$w_{g,t}$	TS generator commitment variable in hour t for generator g

Acronyms

TS	transmission system
DL	dispatchable load
WP	wind penetration
DS	distribution system
DR	demand response
DG	distributed generation
DISCO	distribution company
GENCO	generation company
UC	unit commitment

1 Introduction

With the advance in smart grid technologies, the electricity market has been transforming from a centralized market to a deregulated market, from a vertical structure to a horizontal structure due to some driving factors such as renewable generation, demand response and distributed generation [1]. Consequently, the market is becoming more dynamic and competitive [2]. To suit this deregulated market nature, it is paramount for the three main players in the market namely generation, transmission and distribution to optimize their operation strategies by modeling the interactions between themselves and the other market players. This interaction could be essentially viewed from a game-theory perspective and could be modeled through some game structure. Specifically, bilevel optimization which is closely related to the Stackelberg game in economics is widely applied to model the interaction among the three players [3].

Bilevel optimization is defined as a mathematical optimization problem in which one optimization problem (upper level problem) contains another optimization problem (lower level problem) [4]. There are two types of approaches to solve bilevel optimization problems. The first category applies classical methods including single-level reduction [5, 6], descent [7, 8], penalty function [9, 10], and trust-region methods [11, 12]. Most of the classical approaches deal with convex problems. Some strong assumptions such as continuous differentiability and lower semi-continuity are also quite common. Of those methods, single level reduction is the most widely used one when the lower level problem is convex. The single level reduction method is applied in this work. The second category applies evolutionary methods including genetic algorithms [13], particle swarm optimization [14], differential evolution [15], and metamodeling-based methods [16]. They normally require significant computational efforts and may not perform well. For a detailed review of different bilevel optimization methods, please refer to [4, 17].

There has been on-going research of bilevel optimization of in the power systems. In [18–22], the bidding strategies of generation companys(Genco) are studied, in which the Genco's payoff is formulated as the upper level problem and the independent system operator's dispatch is formulated as the lower level problem. The planning and investment of different power system components using a bilevel optimization framework is investigated in [23–27]. [28–32] use bilevel optimization to study power system vulnerability issues which focus on minimizing the system loss under terrorist attacks on some transmission line or generator. The bilevel optimization work closely related to our work which centers on the transmission and distribution interaction appears in [33–35]. In those work, the distribution system optimal dispatch is the upper level problem, while the transmission system optimal dispatch is the lower level problem. In [33], the author uses bilevel optimization for the operational decision making of a distribution company(DS) in a competitive market with many other DSs. A multi-period energy acquisition bilevel optimization model is proposed for a distribution company with distributed generation(DG) and interruptible load(IL) in [34]. The roles of DG and IL to alleviate congestion are also analyzed in the bilevel framework. Trading strategies of DSs with distributed energy resources are examined in the day-ahead market and real time market in [35]. For those work, there is a main problem with having the DS as the upper level problem. The upper level problem needs to know the lower level problem information such as GENCO information as well as other DSs's information, however it is unrealistic to assume the DS has that information. It is more intuitive and reasonable to have the transmission system operated by the independent system operation(ISO) as the upper level problem and the DS as the lower level problem since the ISO receives information from both the GENCO side and DS side and the ISO is at the top level in the power system hierarchy.

Int this work, the ISO or transmission system dispatch is formulated as the upper level problem, the DSs are formulated as the lower level problem.

Under this structure, the upper level problem decides the energy and demand response prices of the DSs, the lower level problem responds by deciding the quantity of DR and energy import. The key contributions of this work are summarized below.

1. Present a bilevel optimization framework for the transmission and distribution system co-optimization.

2. Compare the bilevel optimization approach with the traditional co-optimization approach in different aspects.

3. Reformulate the bilevel problem as a single level problem with KKT conditions and linearize the resulting nonlinear problem with the big-M method and strong duality theorem.

The structure of the paper is as follows; the problem under study is described in section 2, the bilevel optimization model is discussed in section 3. Numerical results for are reported in Section 4. Concluding remarks and future research directions follow in Section **??**.

2 Problem Description

2.1 Transmission System Problem

The TS has a network of transmission lines and buses. Traditional and renewable generation units, loads and DS are connected to different buses in the network. The TS solves a day ahead market unit commitment problem, which is a co-optimization problem of the energy market and ancillary service market. For the energy market optimization, the TS tries to minimize the cost of meeting the system demand with its own generation capacity. For the ancillary service market optimization, the TS minimizes the cost of providing enough reserve to account for the renewable forecast uncertainty. The reserve service could either come from the TS generator's reserve or the DS's DR. The TS objective function minimizes the energy and ancillary service cost at the same time.

2.2 Distribution System Problem

The DS has dispatchable load and non-dispatchable load in a radial network. In the day-ahead market, the DS solves an optimal dispatch problem with power flow. The dispatchable load is optimized at some point in between its upper and lower bound. The difference between the upper/lower bound and its set point could be used to provide upward/downward DR. The objective of the DS is to minimize the cost of meeting its demand either by importing power from the TS and maximize the revenue of DR provision.

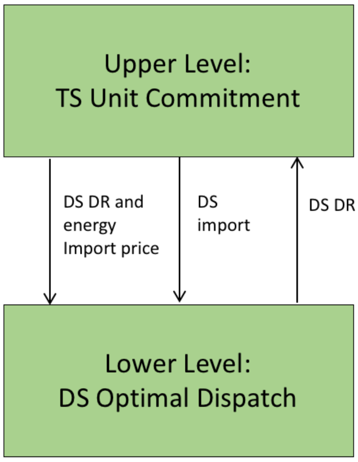

Figure 1: Co-optimization between the TS and MG

2.3 Co-optimization Mode

Under the bilevel co-optimization framework, the TS decides the price of DS energy import as well as the price for purchasing DS DR, the DS responds to those prices by purchasing a certain amount of energy with the TS and selling a certain amount of DR to the TS.

The co-optimization between the TS and DS is illustrated in Fig. 1

2.4 Renewable Forecast Uncertainty Management

In this work, a robust approach is used to manage the uncertainty in renewable forecast. Specifically, for a renewable generation forecast and a set of possible generation scenarios, the upward/downward forecast deviation is calculated by taking the difference of the hourly maximum/minimum generation scenario and the hourly forecast. The downward/upward TS generation reserve and upward/downward MG DR are used to account for the upward/downward renewable forecast deviation.

3 Model Formulation

In this section, the structure of the bi-level optimization problem is described. The general formulation of a bilevel optimization is given below:

$$
\begin{aligned}
&\min_{x\in X, y\in Y} F\left(x,y\right)\\
\text{st: }\ & G_i\left(x,y\right)\le 0, \text{for } i\in\{1,2,...,I\}\\
& H_k\left(x,y\right)=0, \text{for } k\in\{1,2,...,K\}\\
& y\in \underset{y\in Y}{\operatorname{argmin}}\{f(x,y): g_j(x,y)\le 0, \text{for } j\in\{1,2,...,J\},\\
& h_m(x,y)=0, \text{for } m\in\{1,2,...,M\}\}
\end{aligned}
\tag{1}
$$

In the above formulation, $x, F(x,y), (G_i, H_k)$ are the optimization variables, objective function and constraints of the upper level problem. Whereas, $y, f(x,y), (g_i, h_m)$ are the optimization variables, objective function and constraints of the lower level problem. According to this formulation, the bilevel formulation of the TS and DS co-optimization is given in the following sections.

3.1 Upper Level Problem: TS Unit Commitment Problem

The upper level transmission day-ahead unit commitment problem seeks to compute the optimal operation schedule including the generator commitment status $w_{g,t}$, generation output $p_{g,t}$, upward and downward generator's reserve $R^{up}_{g,t}$, $R^{dn}_{g,t}$, DS DR price $P^{dr}_{g,t}$, and the DS energy import c^{im}_t to minimize the total TS operation cost. The optimization variables are denoted by the vector x_t, and include:

$$x_t = [w_{g,t}, p_{g,t}, r^{up}_{g,t}, r^{dn}_{g,t}, p^{dr}_t, c^{im}_t]$$

The objective of the upper level optimization problem is to minimize the TS operation cost including the generator commitment cost, generation cost, reserve cost, the DISCO DR cost, and negative DISCO energy import cost.
Objective function:

$$
\begin{aligned}
F(\{x_t\}_{t=1}^T) \;\; &= \;\; \textstyle\sum_{t=1}^T \sum_{g=1}^G (C^c_{g,t} w_{g,t} + C_{g,t} p_{g,t}\\
&+ \;\; C^r_g (r^{up}_{g,t} + r^{dn}_{g,t}) + p^{dr}_t (dr^{up}_t + dr^{dn}_t)\\
&- \;\; p^{im}_t c^{im}_t)
\end{aligned}
$$

The constraints of the TS are listed below.
Power flow constraints:

$$-\,Line \le GSF * p^{inj}_t \le Line, t\in 1,...,T \tag{2}$$

$$-\,Line \le GSF * p^{inj*}_t \le Line, t\in 1,...,T \tag{3}$$

p^{inj}_t is the DC net power injection vector (ie: generation + wind - demand) for all the buses in each hour. p^{inj*}_t incorporates the wind forecast error, generator reserve and MG DR on top of p_t^{inj}. Eqn (2),(3) bound the transmission line flows within the flow limits.

Generator constraints:

$$\underline{P}_g \leq p_{g,t} \leq \overline{P}_g, t \in 1, ..., T \tag{4}$$

$$(p_{g,t} + r_{g,t}^{up}) - (p_{g,t-1} - r_{g,t-1}^{dn}) \leq \overline{R}_g, t \in 2, ..., T \tag{5}$$

$$\underline{R}_g \leq (p_{g,t} - r_{g,t}^{dn}) - (p_{g,t-1} + r_{g,t-1}^{up}), t \in 2, ..., T \tag{6}$$

Eqn (4) bounds the generator generation within its capacities. Eqn (5),(6) satisfy the generator ramping capability.

Power balance constraint:

$$\sum_{g=1}^{G} p_{g,t} + \mathbf{1}_{1\times N_b} * L_t + W^f = p_t^{im},, t \in 1, ..., T \tag{7}$$

where $\mathbf{1}_{1\times N_b}$ is a vector of length N_b filled with 1's. The dot product $\mathbf{1}_{1\times N_b} * L_t$ gives the total load in the system. Eqn (7) balances the system power supply and demand.

Reserve constraints:

$$W_t^{up} \leq dr_t^{up} + \sum_{g=1}^{G} r_{g,t}^{dn}, t \in 1, ..., T \tag{8}$$

$$W_t^{dn} \leq dr_t^{dn} + \sum_{g=1}^{G} r_{g,t}^{up}, t \in 1, ..., T \tag{9}$$

Eqn (8), (9) ensure enough generator reserve and MG DR to compensate the possible wind forecast deviation.

Finally, the TS unit commitment problem is formulated as:

$$\min_{\{x_t\}_{t=1}^T} F\left(\{x_t\}_{t=1}^T\right)$$

3.2 Lower Level Problem: DS Operation Optimization

The DS considered in this work has a radial network as shown in Fig. 2. There are n buses in the network indexed by $i = 0, 1, ..., n$. A simplified version of power flow that only considers active power according to [36] is considered in this work. The power flow equation at each node i could be expressed as

$$P_{i+1} = P_i - p_{i+1}$$
$$p_i = L^i + l^d$$

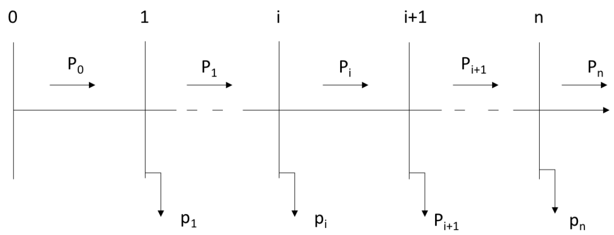

Figure 2: Diagram of a radial distribution network

p_i is the net load at bus i.

The DS could only import energy from the transmission system. The load in the DS is served by the DS DG as well as the TS energy import.

The goal of the DS optimization problem is to compute the energy import p_t^{im} schedule, dispatchable load profile l_t^d, the upward/downward DR dr_t^{up}, dr_t^{dn} provided by the dispachable load. The lower level optimization variables are denoted by the vector y_t, and include:

$$y_t = [p_t^{im}, dr_{i,t}^{up}, dr_{i,t}^{dn}, l_{i,t}^d]$$

The objective of the DS optimization is to minimize the DS operation cost including its DG cost, energy import cost from the TS, and DR cost and maximize its DL utility and DR revenue.

Objective function:

$$\begin{aligned} f(\{y_t\}_{t=1}^T) = &\sum_{t=1}^{T}\sum_{i=1}^{N_b^d}(C_i^{d1}p_{i,t}^d + C_g^{d2}p_{i,t}^d p_{i,t}^d + p_t^{im}c_t^{im} \\ &+C^{dr1}(dr_{i,t}^{up} + dr_{i,t}^{dn}) \\ &+C^{dr2}(dr_{i,t}^{up}dr_{i,t}^{up} + dr_{i,t}^{dn}dr_{i,t}^{dn}) \\ &+C^d(\overline{P}_i^d - p_{i,t}^d)(\overline{P}_i^d - p_{i,t}^d) \\ &-p_t^{dr}(dr_{i,t}^{up} + dr_{i,t}^{dn})) \end{aligned}$$

The constraints for the DS are given below. The λ and μ variables next to the constraints are the dual variables associated with the corresponding inequality and equality constraints.

Power flow constraints:

$$\begin{aligned} &P_{i+1,t} = P_{i,t} - (L_{i,t}^{ie} + l_{i,t}^d - p_{i,t}^d) \\ &t \in 1, ..., T, \mu_{1,i,t}, i \in 1, ..., N_b^d \end{aligned} \tag{10}$$

$$\underline{P}_i \le P_{i,t} \le \overline{P}_i, t \in 1, ..., T, \lambda_{3,i,t}, \lambda_{4,i,t}, i \in 1, ..., N_b^d \tag{11}$$

Eq (10) regulates the power flow between nodes in the distribution system. Eq (11) bounds the power flow within the line limits in the distribution system. Note that $P_{1,t}$ is the DS energy import p_t^{im}.

Dispatchable load constraint:

$$\underline{L}_i^d \leq l_i^d \leq \overline{L}_{i,t}^d, \lambda_{5,i,t}, \lambda_{6,i,t}, t \in 1, ..., T, i \in 1, ..., N_b^d \tag{12}$$

Eq (12) constrains the dispatchable loads within predefined bounds.

DR constraints:

$$0 \leq dr_{i,t}^{up} \leq S^(dr) * l_t^{i,d}, \lambda_{7,i,t}, \lambda_{8,i,t}, t \in 1, ..., T, i \in 1, ..., N_b^d \tag{13}$$

$$0 \leq dr_{i,t}^{dn} \leq S^(dr) * l_t^{i,d}, \lambda_{9,i,t}, \lambda_{10,i,t}, t \in 1, ..., T, i \in 1, ..., N_b^d \tag{14}$$

The amount of DR is set to be within a certain percentage defined by the DR scale factor $S^(dr)$ of the DL, which is reflected in Eq (13), (14).

Finally, the MG optimal dispatch can be formulated as:

$$\begin{aligned} &\min_{\{y_t\}_{t=1}^T} f\left(\{y_t\}_{t=1}^T\right) \\ &\text{s.t. } (10) - (15) \end{aligned}$$

3.3 Reformulation to a Single Level Problem

If the lower level problem is convex and satisfies certain regularity conditions [37], it can be replaced by its Karush-Kuhn-Tucker(KKT) conditions [4], the formulation in (1) becomes a single -level problem reformulation as follows:

$$\begin{aligned} &\min_{x \in X, y \in Y} F\left(x, y\right) \\ \text{st: } &G_i\left(x, y\right) \leq 0, \text{for } i \in \{1, 2, ..., I\} \\ &H_k\left(x, y\right) = 0, \text{for } k \in \{1, 2, ..., K\} \\ &g_i\left(x, y\right) \leq 0, \text{for } j \in \{1, 2, ..., J\} \\ &h_m\left(x, y\right) \leq 0, \text{for } m \in \{1, 2, ..., M\} \\ &\text{dual feasibility: } \lambda_i \geq 0, \text{for } i \in \{1, 2, ..., I\} \\ &\text{complementary slackness:} \\ &\lambda_i * g_i\left(x, y\right) = 0, \text{for } i \in \{1, 2, ..., I\} \\ &\text{stationarity: } \nabla L(x, y, \lambda, \mu) = 0 \\ &where: \\ &L(x, y, \lambda) = f(x, y) + \sum_{i=1}^{J} \lambda_i g_i(x, y) + \sum_{m=1}^{M} \mu_m h_m(x, y) \end{aligned}$$

The lower level problem in this study is a convex optimization problem and satisfies Slater's condition, which is one of the sufficient conditions for strong duality theorem [37]. Therefore, the lower level problem could be replaced by its KKT conditions. The KKT conditions are given below:

Stationarity:
For stationarity, the derivative of the lagrangian function $L(x, y, \lambda)$ is taken with respect to each optimization variable. For example, eq(22) is the derivative of the lagrangian function with respect to $p^d_{i,t}$.

$$2 * C^{d2}_i * p^d_{i,t} + C^{d1}_i + \lambda_{2,i,t} - \lambda_{1,i,t} + \mu_{1,i,t} = 0$$
$$t \in 1, ..., T, i \in \{1, ..., N^d_b\} \tag{16}$$
$$2 * C^{dr2}_i * dr^{up}_{i,t} + C^{dr1}_i - p^{dr}_t + \lambda_{8,i,t} - \lambda_{7,i,t} = 0$$
$$t \in 1, ..., T, i \in \{1, ..., N^d_b\} \tag{17}$$
$$2 * C^{dr2}_i * dr^{up}_{i,t} + C^{dr1}_i - p^{dr}_t + \lambda_{8,i,t} - \lambda_{7,i,t} = 0$$
$$t \in 1, ..., T, i \in \{1, ..., N^d_b\} \tag{18}$$
$$p^{im}_t - \mu_{1,1,t} + \lambda_{4,1,t} - \lambda_{3,1,t} = 0, t \in 1, ..., T, i = 1 \tag{19}$$
$$\mu_{1,i-1,t} - \mu_{1,i,t} + \lambda_{4,i,t} - \lambda_{3,i,t} = 0,$$
$$t \in 1, ..., T, i \in \{2, ..., N^d_b\} \tag{20}$$
$$2 * C^d * l^d_{1,t} + C^d * \overline{L}^d_{i,t} + \lambda_{6,1,t} - \lambda_{5,1,t}$$
$$- \lambda_{8,1,t} - \lambda_{10,1,t} = 0, t \in 1, ..., T, i = 1 \tag{21}$$
$$2 * C^d * l^d_{i,t} + C^d * \overline{L}^d_{i,t} + \lambda_{6,i,t} - \lambda_{5,i,t}$$
$$- \lambda_{8,i,t} - \lambda_{10,i,t} + \mu_{1,i,t} = 0, t \in 1, ..., T, i \in 2, ..., N^d_b \tag{22}$$

Dual feasibility:
For dual feasibility, all dual variables need to be non-negative.

$$\lambda_{1,t}...\lambda_{10,t}, t \in 1, ..., T \geq 0 \tag{23}$$

Complementary slackness:
For complementary slackness, the product of the dual variables and their corresponding inequalities need to be zero.

$$\lambda_{1,i,t} * (p^d_{i,t} - \underline{P}^d_i) = 0, t \in 1, ..., T, i \in 1, ..., N^d_b \tag{24}$$
$$\lambda_{2,i,t} * (p^d_{i,t} - \overline{P}^d_i) = 0, t \in 1, ..., T, i \in 1, ..., N^d_b \tag{25}$$
$$\lambda_{3,i,t} * (P_{i,t} - \overline{P}_i) = 0, t \in 1, ..., T, i \in 1, ..., N^d_b \tag{26}$$
$$\lambda_{4,i,t} * (P_{i,t} - \underline{P}_i) = 0, t \in 1, ..., T, i \in 1, ..., N^d_b \tag{27}$$
$$\lambda_{5,i,t} * (l^d_{i,t} - \overline{L}^d_t) = 0, t \in 1, ..., T, i \in 1, ..., N^d_b \tag{28}$$
$$\lambda_{6,i,t} * (l^d_{i,t} - \underline{L}^d_t) = 0, t \in 1, ..., T, i \in 1, ..., N^d_b \tag{29}$$
$$\lambda_{7,i,t} * dr^{up}_{i,t} = 0, t \in 1, ..., T, i \in 1, ..., N^d_b \tag{30}$$
$$\lambda_{8,i,t} * (dr^{up}_{i,t} - 0.2 * l^d_{i,t}) = 0, t \in 1, ..., T, i \in 1, ..., N^d_b \tag{31}$$
$$\lambda_{9,i,t} * dr^{dn}_{i,t} = 0, t \in 1, ..., T, i \in 1, ..., N^d_b \tag{32}$$
$$\lambda_{10,i,t} * (dr^{dn}_{i,t} - 0.2 * l^d_{i,t}) = 0, t \in 1, ..., T, i \in 1, ..., N^d_b \tag{33}$$

This reformulation is not easy to solve mainly due to the non-convexity in the complementary slackness and bilinear terms in the objective function. The big-M reformulation is used to transform the complementary conditions to mixed integer constraints. The technical details of the Big-M method is given in the appendix.

According to strong duality theorem, the optimal objective function value of the lower level optimization dual problem equals the optimal objective function value of the lower level optimization primal problem. As a result, the bilinear terms $-p_t^{im}c_t^{im} + p_t^{dr}(dr_t^{up} + dr_t^{dn})$ in the lower level objective function, which also appear in the upper level objective function, could be expressed as linear terms using the objective function of the dual problem. The lower level dual problem objective function is as follow:

$$\begin{aligned} D(\{\lambda_t, \mu_t\}_{t=1}^T) &= \textstyle\sum_{t=1}^T \sum_{i=1}^{N_b^d}(-C^{d2}p_{i,t}^d p_{i,t}^d - C^{ld}l_{i,t}^d l_{i,t}^d \\ &- l_{i,t}^d \mu_{i,t} + \lambda_{3,i,t}\underline{P}_{i,t} - \lambda_{4,i,t}\overline{P}_{i,t} \\ &+ \lambda_{5,i,t}\overline{L}_{i,t}^d - \lambda_{6,i,t}\underline{L}_{i,t}^d \\ &+ \lambda_{1,i,t}\overline{P}_i^d - \lambda_{2,i,t}\underline{P}_i^d) \end{aligned}$$

The bilinear terms $-p_t^{im}c_t^{im}+p_t^{dr}(dr_t^{up}+dr_t^{dn})$ in the upper level objective function are represented as:

$$\begin{aligned} y = &\sum_{t=1}^{T}\sum_{i=1}^{N_b^d}(C_i^{d1}p_{i,t}^d + C_g^{d2}p_{i,t}^d p_{i,t}^d \\ &+C^{dr1}(dr_{i,t}^{up} + dr_{i,t}^{dn}) \\ &+C^{dr2}(dr_{i,t}^{up}dr_{i,t}^{up} + dr_{i,t}^{dn}dr_{i,t}^{dn}) \\ &+C^d(\overline{P}_i^d - p_{i,t}^d)(\overline{P}_i^d - p_{i,t}^d)) - D(\{\lambda_{i,t}, \mu_{i,t}\}_{t=1,i=1}^{T,N_b^d}) \end{aligned}$$

The reformulated upper level objective function is as follow:

$$\begin{aligned} F\left(\{x_t, y_t, \lambda_{t,i}, \mu_{t,i}\}_{t=1,i=1}^{T,N_b^d}\right) &= \textstyle\sum_{t=1}^T \sum_{g=1}^G (C_{g,t}^c w_{g,t} + C_{g,t}p_{g,t} \\ &+ C_g^r(r_{g,t}^{up} + r_{g,t}^{dn}) - y \end{aligned}$$

The reformulated single level problem has the following format:

$$\begin{gathered} \min_{\{x_t,y_t,\lambda_t,\mu_t\}_{t=1}^T} F\left(\{x_t, y_t, \lambda_t, \mu_t\}_{t=1}^T\right) \\ \text{st: } (2) - (34) \end{gathered}$$

With the big-M reformulation, a set of binary variables and additional constants are introduced. The bilevel problem becomes a single level mixed integer linear problem and could thus be solved with a wide range of commercial solvers such as Cplex and Gurobi.

4 Numerical Results

In this section, the transmission model described in Section 3 is applied to the IEEE 30-bus system, shown in Fig. 3. Interested readers are referred to [38] for detailed parameters of the TS. The parameters of a 6-bus 18 MW DS are given in the appendix. The schematic of the DS is given in Fig. 4. Bus 8 of the TS hosts a wind farm that provides renewable generation. A DS is also connected at bus 8 which could provide DR to offset the renewable forecast errors.

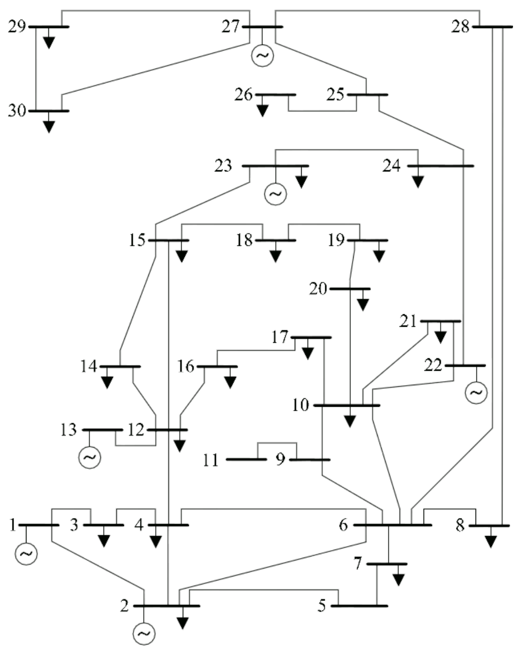

Figure 3: IEEE 30 Bus System

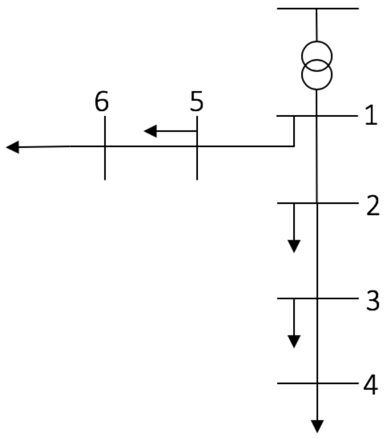

Figure 4: 6-bus Distribution System

The wind data for the wind farms in the TS are selected from the NREL-Eastern Wind Integration Study dataset [39]. Using three years of data, 24-hour trajectories are grouped to identify a set of 54 similar trajectories, with a common initial condition. The set of trajectories was used to represent the realizations of a similar forecast. The central trajectory of the group was selected as the wind forecast, and the remaining were used to estimate the distribution of forecast errors, as described in [40]. From the forecast error distribution, 10000 scenarios are used to generate a robust wind error scenario set.

The objective of this work is to compare and contrast the bilevel optimization approach and the traditional approach adopted by the ISO in various aspects. Traditionally, the ISO solves a single optimization problem for the whole system including the transmission, generation and distribution. The objective functions contains the cost and utility functions functions of all the market participants. Two scenarios are considered for the traditional approach. One considers the distribution network constraints while the other one does not. The energy and DR costs under the traditional approach are both the locational marginal price according to the practice of New York Independent System Operator **??**. To make a fair comparison, the DR price and energy price variables are also set to the same value under the bilevel optimization framework. The formulation of the traditional UC problem is given in the appendix.

First the total system costs of the bilevel optimization approach and the two traditional approach variations are compared in fig. 5. The total cost under the bilevel approach is a posterior cost which is the sum of the upper level and lower level optimal objective function value. The total cost of the traditional approach is just the optimal objective function value. As fig. 5 illustrates, the total cost is the highest under the bilevel optimization approach and is the lowest under the traditional approach without DS network constraints. Since the total cost is not directly optimized under the bilevel approach, its cost is the highest compared to the other two. As the solution of the DS without network constraints might become infeasible with the network constraints, which is what happens in this case. The cost of the DS without network constraints is a lower bound for that of the DS with network constraints. In this case, the cost of the traditional approach without network constraints is lower.

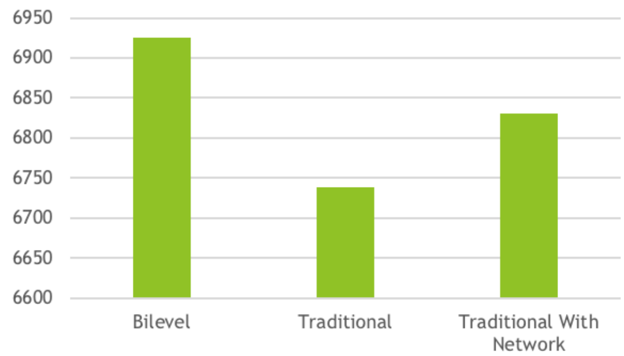

Figure 5: Total System Cost Comparison

The MG costs for the three scenarios are compared in fig. 6. In contrast to the previous case, the MG cost under the bilevel approach is directly optimized in the lower level problem, therefore its cost is the lowest. Whereas the MG cost under the traditional approach is a posterior cost, the two traditional approach costs are higher. Once again the MG cost of the traditional method without DS network constraints is lower due to fewer constraints.

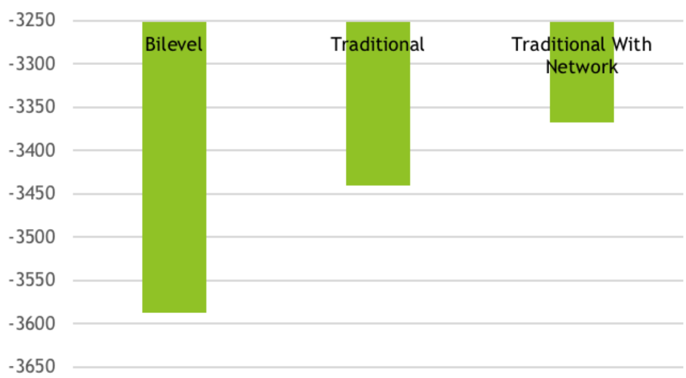

Figure 6: MG Cost Comparison

As it is already shown that the traditional approach without DS network might generate infeasible solutions for the DS, this case is not considered for the subsequent studies. In addition to the cost, the strategies for offsetting wind forecast errors and the environmental implications between the two approaches also exhibit interesting results. For the same set of wind scenarios and same amount of wind forecast errors, the generator reserve and DS DR provisions are given in Table. 1. The main difference between the two approaches is the upward DR and downward generator reserve which are used to handle excess wind generation. The bilevel optimization approach tends to provide more downward generator reserve or generate less power to accommodate more wind generation. On the contrary, the traditional approach tends to provide more upward DR or use more load in this case. Therefore, the bilevel approach has less carbon emission and is thus more environmentally friendly handling excess wind generation. The reason behind that is the lower DR price for the bilevel approach, which corresponds to a smaller quantity on the DR supply curve accordingly to basic economics.

	Drup (MW)	Drdn (MW)	Reserve up (MW)	Reserve dn (MW
Traditional with Network	102	63	5	135
Bilevel	63	63	5	96

Table 1: Results at Different DL Levels

5 Appendix

5.1 Big-M Method

The complementary conditions are set of constraints with the following format:

$$\lambda_i * g_i(x, y) = 0$$

Specifically, given a sufficiently large positive value M_i and binary value ϕ_i, the complementary conditions could be reformulated as below:

$$\begin{aligned} &-(1-\phi_i) * M_i \leq g_i(x, y) \\ &\lambda_i \leq \phi_i * M_i \end{aligned}$$

For a detailed treatment of the Big-M method, please refer to [41].

5.2 Traditional UC Problem

The formulation of the traditional UC problem adopted by the ISOs is based on the UC formulation in [42]. The optimization variables are the sum of the optimization variables from the two levels of the bilevel problem without the energy import price c_t^{im} and DR price p_t^{dr}.

$$x_t = [w_{g,t}, p_{g,t}, r_{g,t}^{up}, r_{g,t}^{dn}, p_t^{dr}, c_t^{im}]$$

Similarly, the objective function is the sum of the upper level and lower level bilevel optimization objective function without the DR and energy import payment terms.

Objective function:

$$\begin{aligned} F(\{x_t\}_{t=1}^T) = & \sum_{t=1}^{T}\sum_{g=1}^{G}(C_{g,t}^{c} w_{g,t} + C_{g,t} p_{g,t} + C_g^r (r_{g,t}^{up} + r_{g,t}^{dn})) \\ & + \sum_{t=1}^{T}\sum_{i=1}^{N_b^d}(C_i^{d1} p_{i,t}^d + C_g^{d2} p_{i,t}^d p_{i,t}^d \\ & + C^{dr1}(dr_{i,t}^{up} + dr_{i,t}^{dn}) \\ & + C^{dr2}(dr_{i,t}^{up} dr_{i,t}^{up} + dr_{i,t}^{dn} dr_{i,t}^{dn}) \\ & + C^d(\overline{P}_i^d - p_{i,t}^d)(\overline{P}_i^d - p_{i,t}^d)) \end{aligned}$$

The constraints of the traditional approach with DS network constraints are the same as the ones used in the bilevel optimization approach. The formulation is as below:

$$\begin{aligned} & \min_{\{x_t\}_{t=1}^T} F\left(\{x_t\}_{t=1}^T\right) \\ & \qquad \text{s.t.}(2) - (15) \end{aligned}$$

If the DS power flow constraints namely Eqn. (10) & (11) are taken out of the constraint set, the problem becomes the tradition UC without DS network constraints version.

5.3 Microgrid Parameters

The parameter values for a 25 MW MG are listed in the table below.

Parameters	Bus 1	Bus 2	Bus 3	Bus 4	Bus 5	Bus 6
$\overline{L}_i^d$	0MW	2MW	2.5MW	0MW	2MW	1MW
$\overline{\underline{L}}_i^d$	0MW	1MW	1.25MW	0MW	1MW	0.5MW
$\underline{L}_i^{ie}$	0MW	2MW	2.5MW	0MW	2MW	1MW
$\overline{\overline{P}}_i$	10MW	10MW	10MW	10MW	10MW	10MW
$\underline{\underline{P}}_i$	0MW	0MW	0MW	0MW	0MW	0MW
C_i^{d1}	0\$/MW	0\$/MW	0\$/MW	0\$/MW	0\$/MW	5\$/MW
C_i^{d2}	0\$/MW	0\$/MW	0\$/MW	0\$/MW	0\$/MW	0.02\$/MW
$\overline{P}_i^d$	0MW	0MW	0MW	0MW	0MW	10MW
$\underline{P}_i^d$	0MW	0MW	0MW	0MW	0MW	0MW

Parameters	Values
C_t^d	3\$/MW
C_t^{dr1}	1\$/MW
C_t^{dr2}	1\$/MW

Table 2: DS parameter values

References

[1] J. J. D. Nesamalar, P. Venkatesh, and S. C. Raja, "Energy management by generator rescheduling in congestive deregulated power system," *Applied energy*, vol. 171, pp. 357–371, 2016.

[2] M. Sun, J. Li, C. Gao, and D. Han, "Identifying regime shifts in the us electricity market based on price fluctuations," *Applied energy*, vol. 194, pp. 658–666, 2017.

[3] B. Colson, P. Marcotte, and G. Savard, "An overview of bilevel optimization," *Annals of operations research*, vol. 153, no. 1, pp. 235–256, 2007.

[4] A. Sinha, P. Malo, and K. Deb, "A review on bilevel optimization: from classical to evolutionary approaches and applications," *IEEE Transactions on Evolutionary Computation*, 2017.

[5] W. F. Bialas and M. H. Karwan, "Two-level linear programming," *Management science*, vol. 30, no. 8, pp. 1004–1020, 1984.

[6] J. F. Bard and J. E. Falk, "An explicit solution to the multi-level programming problem," *Computers & Operations Research*, vol. 9, no. 1, pp. 77–100, 1982.

[7] C. D. Kolstad and L. S. Lasdon, "Derivative evaluation and computational experience with large bilevel mathematical programs," *Journal of optimization theory and applications*, vol. 65, no. 3, pp. 485–499, 1990.

[8] G. Savard and J. Gauvin, "The steepest descent direction for the nonlinear bilevel programming problem," *Operations Research Letters*, vol. 15, no. 5, pp. 265–272, 1994.

[9] Y. Lv, T. Hu, G. Wang, and Z. Wan, "A penalty function method based on kuhn–tucker condition for solving linear bilevel programming," *Applied Mathematics and Computation*, vol. 188, no. 1, pp. 808–813, 2007.

[10] D. J. White and G. Anandalingam, "A penalty function approach for solving bi-level linear programs," *Journal of Global Optimization*, vol. 3, no. 4, pp. 397–419, 1993.

[11] B. Colson, P. Marcotte, and G. Savard, "A trust-region method for nonlinear bilevel programming: algorithm and computational experience," *Computational Optimization and Applications*, vol. 30, no. 3, pp. 211–227, 2005.

[12] P. Marcotte, G. Savard, and D. Zhu, "A trust region algorithm for nonlinear bilevel programming," *Operations research letters*, vol. 29, no. 4, pp. 171–179, 2001.

[13] R. Mathieu, L. Pittard, and G. Anandalingam, "Genetic algorithm based approach to bi-level linear programming," *RAIRO-Operations Research*, vol. 28, no. 1, pp. 1–21, 1994.

[14] X. Li, P. Tian, and X. Min, "A hierarchical particle swarm optimization for solving bilevel programming problems," in *International Conference on Artificial Intelligence and Soft Computing.* Springer, 2006, pp. 1169–1178.

[15] X. Zhu, Q. Yu, and X. Wang, "A hybrid differential evolution algorithm for solving nonlinear bilevel programming with linear constraints," in *Cognitive Informatics, 2006. ICCI 2006. 5th IEEE International Conference on*, vol. 1. IEEE, 2006, pp. 126–131.

[16] G. G. Wang and S. Shan, "Review of metamodeling techniques in support of engineering design optimization," *Journal of Mechanical design*, vol. 129, no. 4, pp. 370–380, 2007.

[17] B. Colson, P. Marcotte, and G. Savard, "An overview of bilevel optimization," *Annals of operations research*, vol. 153, no. 1, pp. 235–256, 2007.

[18] V. Vahidinasab and S. Jadid, "Multiobjective environmental/techno-economic approach for strategic bidding in energy markets," *Applied Energy*, vol. 86, no. 4, pp. 496–504, 2009.

[19] S. Soleymani, A. Ranjbar, and A. Shirani, "New approach to bidding strategies of generating companies in day ahead energy market," *Energy Conversion and Management*, vol. 49, no. 6, pp. 1493–1499, 2008.

[20] G. Kozanidis, E. Kostarelou, P. Andrianesis, and G. Liberopoulos, "Mixed integer parametric bilevel programming for optimal strategic bidding of energy producers in day-ahead electricity markets with indivisibilities," *Optimization*, vol. 62, no. 8, pp. 1045–1068, 2013.

[21] G. Zhang, G. Zhang, Y. Gao, and J. Lu, "Competitive strategic bidding optimization in electricity markets using bilevel programming and swarm technique," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 6, pp. 2138–2146, 2011.

[22] G. K. E. K. P. Andrianesis and G. Liberopoulos, "Mixed integer bilevel programming for optimal bidding strategies in day-ahead electricity markets with indivisibilities," *BALCOR 2011*, p. 72, 2011.

[23] P. Verma, K. Sanyal, D. Srinivsan, and K. Swarup, “Information exchange based clustered differential evolution for constrained generation-transmission expansion planning,” *Swarm and Evolutionary Computation*, 2018.

[24] S. Zolfaghari and T. Akbari, “Bilevel transmission expansion planning using second-order cone programming considering wind investment,” *Energy*, vol. 154, pp. 455–465, 2018.

[25] H. Pandžić, Y. Dvorkin, and M. Carrión, “Investments in merchant energy storage: Trading-off between energy and reserve markets,” *Applied Energy*, vol. 230, pp. 277–286, 2018.

[26] Y. Zhou, L. Wang, and J. D. McCalley, “Designing effective and efficient incentive policies for renewable energy in generation expansion planning,” *Applied Energy*, vol. 88, no. 6, pp. 2201–2209, 2011.

[27] L. Baringo and A. Conejo, “Wind power investment within a market environment,” *Applied Energy*, vol. 88, no. 9, pp. 3239–3247, 2011.

[28] Y. Yuan, Z. Li, and K. Ren, “Modeling load redistribution attacks in power systems,” *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 382–390, 2011.

[29] Y. Xiang, L. Wang, and N. Liu, “Coordinated attacks on electric power systems in a cyber-physical environment,” *Electric Power Systems Research*, vol. 149, pp. 156–168, 2017.

[30] J. M. Arroyo and F. J. Fernández, “A genetic algorithm approach for the analysis of electric grid interdiction with line switching,” in *Intelligent System Applications to Power Systems, 2009. ISAP’09. 15th International Conference on.* IEEE, 2009, pp. 1–6.

[31] A. Pinar, J. Meza, V. Donde, and B. Lesieutre, “Optimization strategies for the vulnerability analysis of the electric power grid,” *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1786–1810, 2010.

[32] J. M. Arroyo, “Bilevel programming applied to power system vulnerability analysis under multiple contingencies,” *IET generation, transmission & distribution*, vol. 4, no. 2, pp. 178–190, 2010.

[33] H. Haghighat and S. W. Kennedy, “A bilevel approach to operational decision making of a distribution company in competitive environments,” *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 1797–1807, 2012.

[34] H. Li, Y. Li, and Z. Li, “A multiperiod energy acquisition model for a distribution company with distributed generation and interruptible

load," *IEEE Transactions on Power Systems*, vol. 22, no. 2, pp. 588–596, 2007.

[35] C. Zhang, Q. Wang, J. Wang, M. Korpås, P. Pinson, J. Østergaard, and M. E. Khodayar, "Trading strategies for distribution company with stochastic distributed energy resources," *Applied energy*, vol. 177, pp. 625–635, 2016.

[36] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Transactions on Power delivery*, vol. 4, no. 2, pp. 1401–1407, 1989.

[37] J. Borwein and A. S. Lewis, *Convex analysis and nonlinear optimization: theory and examples.* Springer Science & Business Media, 2010.

[38] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower's extensible optimal power flow architecture," in *Power & Energy Society General Meeting, 2009. PES'09. IEEE.* IEEE, 2009, pp. 1–7.

[39] G. L. R. W. Energy, "Eastern wind integration and transmission study," *Technical Report, National Renewable Energy Laboratory (NREL)*, 2010.

[40] C. Anderson and R. Zimmerman, "Wind output forecasts and scenario analysis for stochastic multiperiod optimal power flow," *PSERC Webinar*, pp. 1–38, 2011.

[41] W. L. Winston, M. Venkataramanan, and J. B. Goldberg, *Introduction to mathematical programming.* Thomson/Brooks/Cole Duxbury; Pacific Grove, CA, 2003, vol. 1.

[42] A. Papavasiliou and S. S. Oren, "Multiarea stochastic unit commitment for high wind penetration in a transmission constrained network," *Operations Research*, vol. 61, no. 3, pp. 578–592, 2013.