Applied Energy 277 (2020) 115520

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

A stochastic distribution system planning method considering regulation services and energy storage degradation

Xinyi Zhao[a], Xinwei Shen[a,*], Qinglai Guo[a,b], Hongbin Sun[a,b,*], Shmuel S. Oren[a,c]

[a] Tsinghua-Berkeley Shenzhen Institute, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China
[b] Department of Electrical Engineering, Tsinghua University, Beijing 100084, China
[c] Department of Industrial Engineering and Operations Research, University of California at Berkeley, CA 94720, USA

HIGHLIGHTS

- A stochastic distribution system planning with energy storage degradation is solved.
- Energy arbitrage and regulation services are co-optimized for siting and sizing.
- A Gaussian mixture model is adopted to generate stochastic scenarios.
- A modified progressive hedging algorithm which outperforms Gurobi is introduced.

ARTICLE INFO

Keywords:
Energy storage system
Regulation service
Distribution system planning
Two-stage stochastic programming
Progressive hedging

ABSTRACT

With the trend of energy storage participating in ancillary service markets, it is still computationally burdensome to incorporate the rapidly changing real-time signals in the long-run distribution system planning. In this paper, a two-stage stochastic programming is proposed for the distribution system with energy storage, where the storage degradation and ancillary service revenue for frequency regulation are both considered. For this purpose, the problem is formulated as a mixed-integer linear programming optimizing the overall planning cost, including investment and maintenance cost, power transaction cost and revenue from regulation services. A degradation penalty is added in the objective to avoid excessive charge/discharge when providing regulation services, thus further benefiting the economy of the distribution system. The model also considers uncertainties of load demand and electricity prices. A Gaussian mixture model is adopted to characterize these uncertainties and a set of representative scenarios are sampled. To accelerate the optimization, a modified progressive hedging with parallel computing is proposed. It is demonstrated through a 33-bus distribution system that the proposed algorithm has a speed approximately 15 times as fast as the state-of-art commercial software Gurobi when solving the model in 100 scenarios. For this case study, considering degradation penalty has been shown to extend energy storage lifespan by one year.

1. Introduction

The severe peak-valley load difference and distributed renewable energy integration have been two of the most crucial issues in distribution systems. To cope with these challenges, one solution is to install energy storage systems (ESSs) such that they can shift the peak load [1] and benefit the renewable penetration at the same time [2]. Since the price of batteries has decreased significantly [3] and proved to have a startling decline speed in levelized cost of energy [4], ESSs have reached widespread application in distribution systems, as an effective means of energy arbitrage. Moreover, considering that ESS's shorter duration applications (mainly less than 4 h) remain the most cost-effective [5], the potential revenue from added ancillary services can further improve profits of ESS investment [6].

As illustrated in Fig. 1, in today's wholesale market, e.g. the market operated by the California Independent System Operator (CAISO), a large proportion of its revenue consists of energy arbitrage and frequency regulation [5], both of which can be provided by ESSs. Similarly, for a distribution system, these storage units can also play a positive role in enhancing the grid's reliability by providing multiple ancillary services [7]. Besides, ESS's ability of peak load shaving can postpone the upgrades of electric installations [8], thus cutting down

* Corresponding author.
E-mail addresses: sxw.tbsi@sz.tsinghua.edu.cn (X. Shen), shb@mail.tsinghua.edu.cn (H. Sun).

https://doi.org/10.1016/j.apenergy.2020.115520
Received 17 March 2020; Received in revised form 21 June 2020; Accepted 13 July 2020

Nomenclature

Abbreviations

INV	investment
MAT	maintenance
eL	existing lines
nL	candidate new lines
SUB	substation
ESS	energy storage system
PT	power transaction
LMP	locational marginal price
REG	regulation (services)
SOC	state of charge
deg	degradation
cur	load curtailment

Sets

ω^{up}, ω^{dn}	proportion of ESS capacity committed to regulation up/down services
Ψ_{eL}	set of existing lines
Ψ_{nL}	set of candidate new lines
Ω_{SUB}	set of substation candidate nodes
Ω_{ESS}	set of ESS candidate nodes
Ω_D	set of load nodes

Indices

j	index of the jth existing line
k	index of the kth new line
z	index of new line option
b	index of new substation option
e	index of ESS option
i	index of planning year
t	index of hour in a day
l	index of load node
m	index of substation node
n	index of ESS node
α	index of season
s	index of stochastic scenario

Parameters

γ	annual discount rate
C, O	unit cost of investment/maintenance
T	total number of hours
S	total number of scenarios
θ_α, θ_s	probability of season α/scenario s
Γ	total node number of the distribution system
W^{LMP}	locational marginal price [$/MWh]
$W^{\mathrm{up}}_{\mathrm{REG}}$, $W^{\mathrm{dn}}_{\mathrm{REG}}$	regulation up/down service price [$/MWh]
M_{cur}	penalty factor for load curtailment
M_{deg}	penalty factor for ESS degradation
M	penalty factor for the Big-M method
φ	vector of ESS degradation multipliers [MW^{-1}]
Ξ^{eL}	node-branch incidence matrix for existing lines
Ξ^{nL}	node-branch incidence matrix for new lines
D	load demand [MWh]
Z^{eL}	impedance of each existing line [Ω]
Z^{nL}	impedance of each new line [Ω]
σ^{up}, σ^{dn}	standard deviation of regulation up/down signals
F_{max}	maximum of branch current [p.u.]
E^{SUB}_{max}, E^{ESS}_{max}	maximum of substation/ESS capacity [MWh]
P_{max}	maximum of ESS power output [MW]
E_0	initial value of ESS state of charge [MWh]

Decision variables

x	Planning decision on lines, substations and ESSs
y	Operating decision on lines, substations and ESSs
p^{SUB}	Power from the bulk power system [MWh]
r^{up}, r^{dn}	ESS committed capacity for regulation up/down service [MW]
β^{ESS}	Vector of ESS operation decision variables
d^{cur}	Load curtailment [MWh]
f^{eL}	Branch current of existing line [p.u.]
f^{nL}	Branch current of new line [p.u.]
p^{ch}, p^{dis}	ESS charge/discharge power [MW]
ξ^{SOC}	ESS state of charge [MWh]
u	Node voltage [p.u.]
χ^{up}, χ^{dn}	Ancillary variable in the Big-M method

the expenses in the overall planning process [9]. Nevertheless, the feasibility of ESS's profiting from ancillary service products [10] still needs to be further explored in a distribution system scale, among which the battery aging problem is crucial. Taking lithium-ion batteries (LIBs) for example, the degradation process is complex and the corresponding model is nonlinear and typically related to both stress cycles and operating time [11], which poses challenges in the existing distribution system planning.

A large body of literature exists on the modeling of ESS degradation. To estimate ESS's lifespan, some semi-empirical battery degradation models [11,12] have been proposed and applied to predict long-term cycle aging of large-format cells [13]. However, in combination with rainflow uncertainty [14], these models will add excessive computational burden to an optimization scheme. In this regard, there have been several attempts to derive a simplified degradation model and evaluate ESS's profitability. In [6], a data-driven linear penalty term denoting the battery degradation rate is co-optimized in the objective, which maximizes the total revenue of ESS with energy arbitrage and ancillary services. Ref. [15] linearizes the function of storage capacity loss within specific domains and considers battery aging effects by a linear constraint. Ref. [16] embeds the calculation of battery cycle life into a profit maximization model for optimal bidding and operational schedules. Based on case studies of a battery and transformer deployment, [17] co-optimizes four types of services provided by the ESS, namely, energy arbitrage, regulation service, restoration service and transformer load relieving, which are applicable to the utility's transformer capacity planning with batteries. However, the trade-off between the ESS degradation and the income from these ancillary services is ignored in the model. Apart from the above-mentioned literature, the taxonomy of major researches relevant to ESS degradation is listed in Table 1. To the best of our knowledge, distribution system expansion planning (DSEP) which considers both ESS's providing ancillary services and its degradation is rarely studied so far.

Among studies on optimization for DSEP, convex relaxation [18] combined with *DistFlow* [19] has been a popular method although it is not easy to guarantee an exact relaxation. A mixed-integer linear programming (MILP) is an alternative, which can be efficiently solved by current commercial optimization solvers and applicable to bidirectional power flow problems [20]. Besides, considering inherent uncertainties of a realistic distribution system including resource availability, price fluctuations, load change and policy restrictions, MILP can be further coupled with robust or stochastic optimization. Ref. [21] illustrates the uncertainties of PV output and multi-load demand with dual norms to calculate the worst cases directly, then a two-stage robust optimization

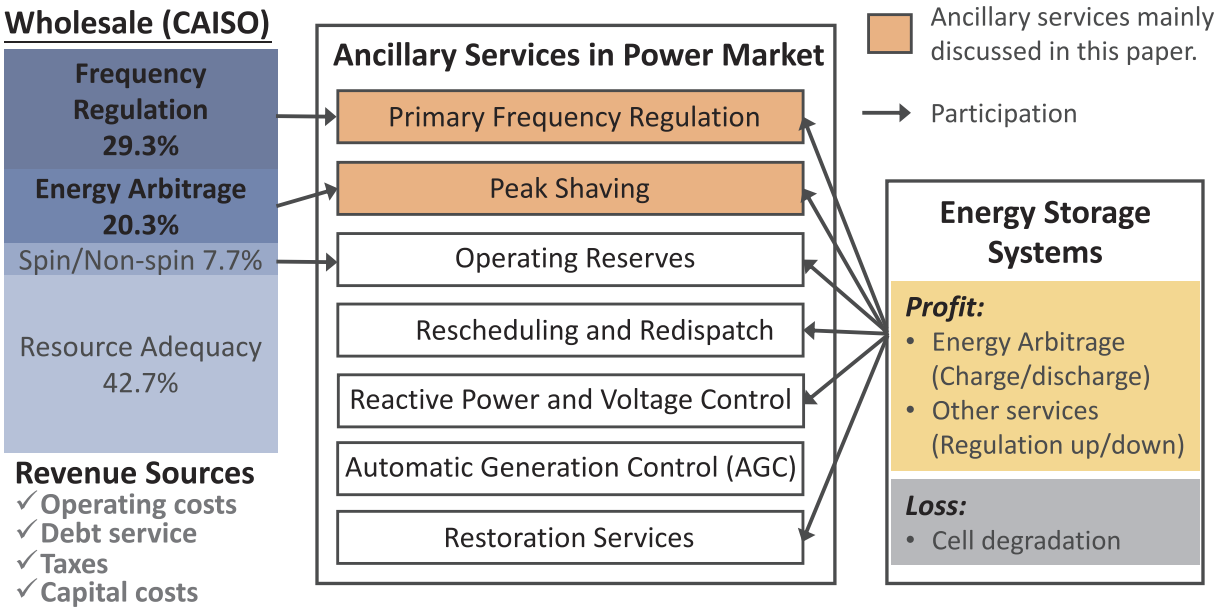

Fig. 1. Overview of the relationship between power ancillary service market and ESSs. Data Source: Lazard's Levelized Cost of Storage Analysis Version 4.0 [5].

Table 1
Review of literature related to ESS degradation (AS: ancillary service).

Motifs	Methods & Algorithms
ESS degradation & AS [6,11,16,33,34]	Semi-empirical battery degradation model; Profit maximization model; Performance-based regulation mechanism; Rainflow cycle-counting algorithm.
DSEP & ESS degradation [21,35–37]	Piecewise linearized battery lifetime model; Planning-operation co-optimization model; Mixed-integer linear programming; Benders decomposition.
DSEP & ESS degradation & AS [17]	Failure threshold of ESS capacity; Stochastic dynamic programming model; Dynamic programming algorithm.

model for energy hub planning and operation is merged into a single stage MILP. Ref. [22] considers both short-term and long-term uncertainty in the DSEP, which are handled by the K-means clustering technique and confidence bounds, respectively. Ref. [23] presents a two-stage stochastic programming model, in which Monte Carlo simulation is utilized for addressing energy demand and supply uncertainty. In [24], uncertainties due to load fluctuations, power generation of wind and solar farm, along with generation deficiency are modeled. Stochastic scenarios are generated by a combined process of Monte Carlo simulation, Roulette wheel mechanism and scenario reduction algorithm. However, for every uncertain parameter in these aforementioned DSEP models, the corresponding scenarios are sampled from a single normal distribution. To deal with uncertainty in this paper, we adopt a Gaussian mixture model (GMM) [25] which can form smooth approximations to arbitrary probability density functions.

Moreover, considering uncertainties may lead to excessive computational burden in DSEP problems. As for stochastic programming framework, decomposition techniques are used to address the tractability issue, and are divided into two categories [26]. One is time-stage-based, known as the L-shaped method, while the other is scenario-based, i.e. progressive hedging (PH).

In fact, the L-shaped method can be considered as a further extension of Benders decomposition for stochastic programming [27]. In 1969, it was introduced as Van Slyke and Wets's method [28], which was 20 years earlier than the PH idea proposed by Rockafellar and Wets in 1991 [29]. The basic rule of the L-shaped method is to replace the nonlinear recourse function in the *master problem* with a lower bound variable, and then approximate the nonlinear term by reformulating the scenario subproblems [30]. This decomposition technique has a broader application in vehicle routing problems [31,32] rather than DSEP.

As for the PH algorithm, it provably converges linearly for stochastic programming with continuous decision variables [38], which is superior to the L-shaped method, since the latter has a significantly increasing difficulty in calculating the *master problem* as the iteration number grows [26]. When comes to power system planning, [39] develops the PH algorithm to solve a scenario-based multi-stage stochastic planning problem. However, a pipeline model is used to avoid binary variables; thus it is not applicable for the siting and sizing of other facilities in DSEP. Ref. [40] designs planning schemes for a transmission system which considers discrete and continuous decision variables denoting transmission and generation investments, respectively. A hedging process is utilized in [41] to resolve decision conflicts in the first stage of large-scale DSEP with scenario uncertainty, but it is solved by an evolutionary algorithm. None of these works have modified the PH algorithm to accelerate its convergence when solving a large-scale MILP, which is typical of the DSEP model.

In this work, the DSEP considering ESSs is formulated as a two-stage stochastic programming model, which is also a MILP one, where ESS degradation is co-optimized in the planning stage. Subsequently, the effects of ancillary service provision by ESSs and uncertainties lying in scenarios are investigated in the solutions. Following are our main contributions:

- A two-stage stochastic DSEP model aims at minimizing the overall planning cost is proposed. To handle the uncertainties, we adopt a GMM based on historical data instead of sampling them from a certain distribution. Besides, the degradation of ESSs and their value on ancillary services are both considered in distribution system planning, which is relatively rare in recent studies.
- To better solve the DSEP model with PH algorithms, the idea of implementing *non-anticipativity* constraints by computing a rational

average solution instead of the mathematical expectation is originally presented. Moreover, with parallel computing process and gap-dependent penalty factors, the modified PH algorithm is further improved in solving the proposed model and has outperformed a well-known commercial solver Gurobi and the L-shaped method.

The remainder of this paper is organized as follows. Section 2 details the mathematical formulation of the MILP model for the distribution system. Based on that, Section 3 reformulates a two-stage stochastic programming to incorporate uncertainties of load demand, electricity prices and regulation signals. In Section 4, representative scenarios are generated via GMM, then the modified PH algorithm is introduced. Section 5 performs case studies of the deterministic MILP and the two-stage stochastic model. The conclusion is drawn in Section 6.

2. Model formulation

The DSEP considering ESS degradation and regulation services is established as a MILP model, illustrated in Fig. 2. The objective includes five expenses within the planning and operation stage. Network configuration, substation expansion and ESS siting and sizing are decided in the planning stage, where binary variable x determines whether to invest the facility or not. As for operation, decision variables can be divided into three categories, which are binary variable y denoting the operating lines, substations and ESSs; continuous variables related to grid operation, i.e. power transmitted by substations $p^{\mathrm{SUB}}_{m,b,t}$; and the continuous vector $\beta^{\mathrm{ESS}}_{n,t}$ related to ESS's behaviors, including the charge/discharge, regulation up/down and state of charge (SOC).

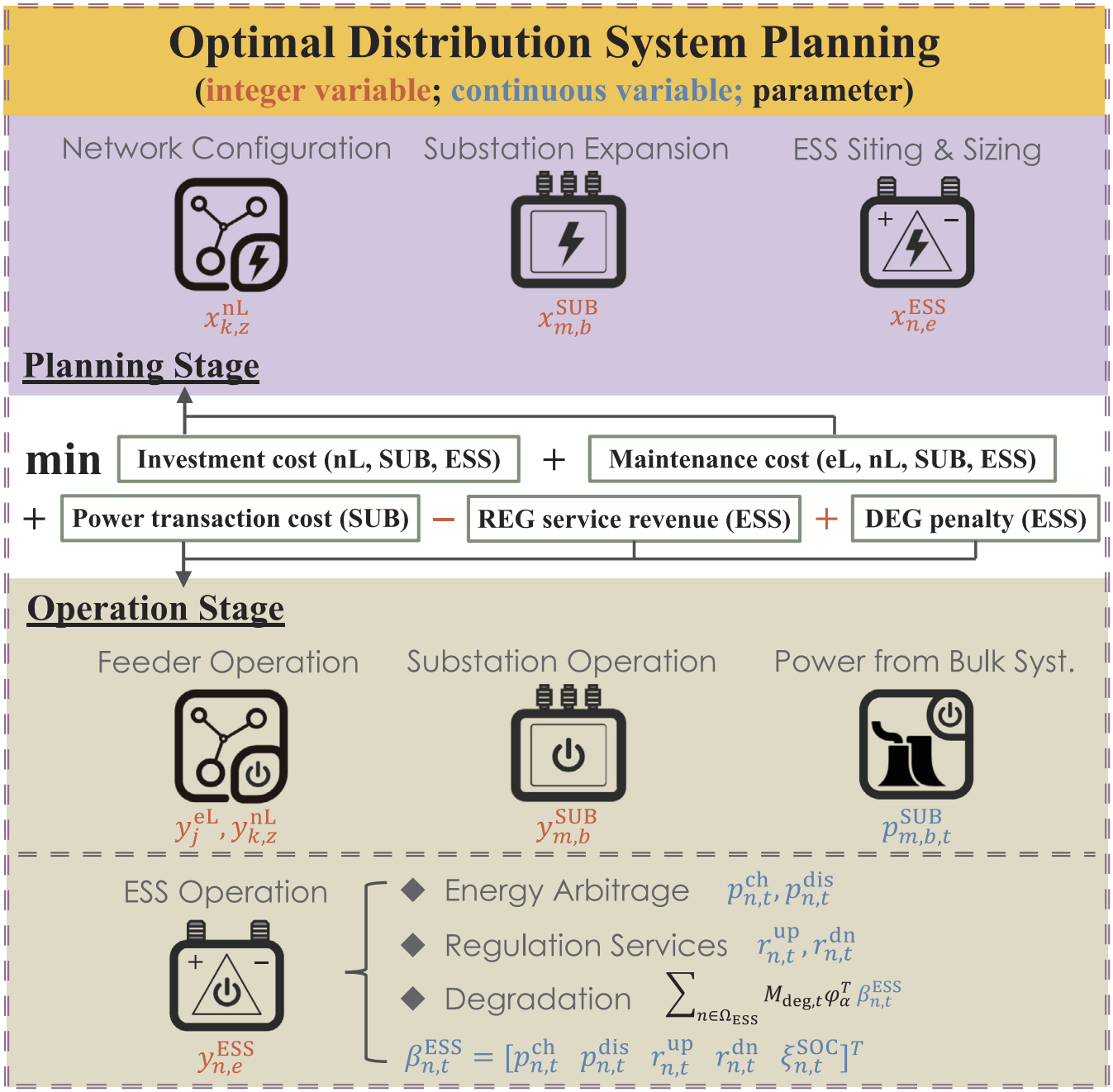

Fig. 2. Overview of the MILP model.

2.1. Objective

In the planning stage, we assume that the distribution system will invest in lines, substations and ESSs, whose maintenance fare is also covered in the overall cost. Besides, electricity needs to be bought from a bulk power system as power transaction cost, which will be affected by ESS operation.

Meanwhile, ESSs will provide regulation services to the bulk power system. In this paper, the overall cost of distribution system is minimized, and the possible revenue of energy arbitrary and frequency regulation services should be considered. To prolong ESS's lifespan, a penalty term relevant to its degradation model is added to the objective function.

2.1.1. Investment cost

The DSEP includes the investment cost of lines, substations and ESSs, denoted in Eq. (1). Where $x = 1$ represents the corresponding facility invested in the distribution system and otherwise $x = 0$. $C^{\mathrm{nL}}_{k,z}$, $C^{\mathrm{SUB}}_{m,b}$ and $C^{\mathrm{ESS}}_{n,e}$ are the unit costs to built feeders, substations and ESSs of certain options.

$$C_{\mathrm{INV}} = \sum_{k\in\Psi_{\mathrm{nL}}}\sum_{z} C^{\mathrm{nL}}_{k,z}x^{\mathrm{nL}}_{k,z} + \sum_{m\in\Omega_{\mathrm{SUB}}}\sum_{b} C^{\mathrm{SUB}}_{m,b}x^{\mathrm{SUB}}_{m,b} + \sum_{n\in\Omega_{\mathrm{ESS}}}\sum_{e} C^{\mathrm{ESS}}_{n,e}x^{\mathrm{ESS}}_{n,e} \tag{1}$$

Note that this work focuses on the expansion planning, hence the investment of existing lines will be precluded. Instead, their maintenance cost is considered in Eq. (2).

2.1.2. Maintenance cost

Similarly, the total operation and maintenance cost needs to involve all the components in the distribution network. $y = 1$ denotes that the facility is in operation and otherwise $y = 0$. O^{eL}_{j}, $O^{\mathrm{nL}}_{k,z}$, $O^{\mathrm{SUB}}_{m,b}$, and $O^{\mathrm{ESS}}_{n,e}$ are the yearly maintenance costs of existing lines, newly-built lines, substations and ESSs.

$$C_{\mathrm{MAT}} = \sum_{t=1}^{T}\left(\sum_{j\in\Psi_{\mathrm{eL}}} O^{\mathrm{eL}}_{j}y^{\mathrm{eL}}_{j} + \sum_{k\in\Psi_{\mathrm{nL}}}\sum_{z} O^{\mathrm{nL}}_{k,z}y^{\mathrm{nL}}_{k,z} + \sum_{m\in\Omega_{\mathrm{SUB}}}\sum_{b} O^{\mathrm{SUB}}_{m,b}y^{\mathrm{SUB}}_{m,b} + \sum_{n\in\Omega_{\mathrm{ESS}}}\sum_{e} O^{\mathrm{ESS}}_{n,e}y^{\mathrm{ESS}}_{n,e}\right) \tag{2}$$

2.1.3. Power transaction cost

To supply the load demand, power is bought from the bulk power system and the cost is denoted by C_{PT} as below:

$$C_{\mathrm{PT}} = \sum_{\alpha}\theta_{\alpha}\sum_{t=1}^{T}\left(\sum_{m\in\Omega_{\mathrm{SUB}}} W^{\mathrm{LMP}}_{\alpha,t}p^{\mathrm{SUB}}_{m,b,t}\right) \tag{3}$$

where $p^{\mathrm{SUB}}_{m,b,t}$ is the power transmitted from substation node m with option b at hour t, θ_{α} is the portion of typical seasonal scenario α. W^{LMP} denotes the locational marginal price (LMP).

2.1.4. Revenue of regulation services

In real time operations, revenue will be earned for ESS's providing regulation services to the bulk system, denoted as W_{REG}:

$$W_{\mathrm{REG}} = \sum_{\alpha}\theta_{\alpha}\sum_{t=1}^{T}\sum_{n\in\Omega_{\mathrm{ESS}}}\left(W^{\mathrm{up}}_{\mathrm{REG},\alpha,t}r^{\mathrm{up}}_{n,t} + W^{\mathrm{dn}}_{\mathrm{REG},\alpha,t}r^{\mathrm{dn}}_{n,t}\right) \tag{4}$$

where nonnegative decision variables $r^{\mathrm{up}}_{n,t}$, $r^{\mathrm{dn}}_{n,t}$ determine how much regulation up/down capacity is committed. $W^{\mathrm{up}}_{\mathrm{REG},\alpha,t}$ and $W^{\mathrm{dn}}_{\mathrm{REG},\alpha,t}$ represent revenue of ESS providing the unit capacity for regulation up and down.

2.1.5. Penalty term of degradation

In order to trade off between profits earned by ESSs and the battery's cycle degradation cost, a penalty is added in the objective, which consists of two important vectors shown in Eq. (5) and Eq. (6). The parameters a_1, a_2 and p_z in (5) are constants relevant to LIB types, which will affect the degradation rates of different ESS's behaviors [6].

$$\varphi_{\alpha} = \begin{bmatrix} \frac{a_2}{4}(1-p_z) \\ \frac{a_2}{4}(1-p_z) \\ \frac{a_1^2p_z}{2a_2}y^{\mathrm{ESS}}_{n,e}P_{e,max}\left(1.5(\sigma^{\mathrm{up}}_{\alpha,t})^2 - 0.5(\sigma^{\mathrm{dn}}_{\alpha,t})^2 - \omega^{\mathrm{up}}_{\alpha,t}\omega^{\mathrm{dn}}_{\alpha,t}\right) \\ \frac{a_1^2p_z}{2a_2}y^{\mathrm{ESS}}_{n,e}P_{e,max}\left(1.5(\sigma^{\mathrm{dn}}_{\alpha,t})^2 - 0.5(\sigma^{\mathrm{up}}_{\alpha,t})^2 - \omega^{\mathrm{up}}_{\alpha,t}\omega^{\mathrm{dn}}_{\alpha,t}\right) \\ 0 \end{bmatrix} \tag{5}$$

$$\beta^{\mathrm{ESS}}_{n,t} = \begin{bmatrix} p^{\mathrm{ch}}_{n,t} & p^{\mathrm{dis}}_{n,t} & r^{\mathrm{up}}_{n,t} & r^{\mathrm{dn}}_{n,t} & \xi^{\mathrm{SOC}}_{n,t} \end{bmatrix}^{T} \tag{6}$$

As illustrated in Fig. 2, five expenses are involved in the objective shown in Eq. (7), with both regulation services and degradation of ESSs taken into account.

$$\min\sum_{i}\gamma^{i}\left(C_{\mathrm{INV}} + C_{\mathrm{MAT}} + C_{\mathrm{PT}} - W_{\mathrm{REG}}, + \sum_{\alpha}\theta_{\alpha}\sum_{t=1}^{T}\sum_{n\in\Omega_{\mathrm{ESS}}} M_{\mathrm{deg},t}\varphi^{T}_{\alpha}\beta^{\mathrm{ESS}}_{n,t}\right) \tag{7}$$

2.2. Constraints

This paper considers major constraints including Kirchhoff's current law (KCL), node voltage limits, feeders' capacity [2], and ESS operation constraints [6], which will be fully listed in Section 3. Furthermore, some extra constraints are supposed to be detailed in the planning.

2.2.1. Construction logical constraints

To avoid building redundant projects on the same node, Eq. (8) denotes that only one option can be chosen among all candidate choices, where $x^{\mathrm{SUB}}_{m,b}$ is the planning decision for substation on node m with option b, $x^{\mathrm{ESS}}_{n,e}$ is the planning decision for ESS on node n with option e. Eq. (9) denotes that the substation and ESS will not be available ($y = 0$) when they are not constructed ($x = 0$). Otherwise, if the facility has been built ($x = 1$), its operating variable can be 1 or 0 ($y = 0/y = 1$).

$$\sum_{b} x^{\mathrm{SUB}}_{m,b} \le 1,\ \sum_{e} x^{\mathrm{ESS}}_{n,e} \le 1 \tag{8}$$

$$y^{\mathrm{SUB}}_{m,b} \leqslant x^{\mathrm{SUB}}_{m,b},\ y^{\mathrm{ESS}}_{n,e} \leqslant x^{\mathrm{ESS}}_{n,e} \tag{9}$$

$$\sum_{j\in\Psi_{\mathrm{eL}}} y^{\mathrm{eL}}_{j} + \sum_{k\in\Psi_{\mathrm{nL}}}\sum_{z} y^{\mathrm{nL}}_{k,z} = \Gamma - N_{\mathrm{SUB}} \tag{10}$$

Eq. (10) guarantees that the number of operating lines are equal to that of load nodes, i.e. distribution network is a spanning tree. N_{SUB} represents the number of operating substations.

2.2.2. Reformulation by Big-M method

Note $\varphi^{T}_{\alpha}\beta^{\mathrm{ESS}}_{n,t}$ in the objective (7) results in some computational difficulties due to the consideration of ESS siting and sizing. That is, the binary decision variable $y^{\mathrm{ESS}}_{n,e}$ is multiplied with the continuous variable $r^{\mathrm{up}}_{n,t}/r^{\mathrm{dn}}_{n,t}$, thus the model is no longer MILP.

To this end, we apply Big-M method to reformulate this term, as illustrated in (11)–(15), where $\chi^{\mathrm{up}}_{n,e,t}$ and $\chi^{\mathrm{dn}}_{n,e,t}$ are two ancillary decision variables and M is a big constant (1e+5). Nonnegative variables $\chi^{\mathrm{up}}_{n,e,t}$ and $\chi^{\mathrm{dn}}_{n,e,t}$ are introduced to replace and relax the product of $y^{\mathrm{ESS}}_{n,e}$ and $r^{\mathrm{up}}_{n,t}/r^{\mathrm{dn}}_{n,t}$.

$$\chi^{up}_{n,e,t} \leqslant r^{up}_{n,t},\ \chi^{dn}_{n,e,t} \leqslant r^{dn}_{n,t} \tag{11}$$

$$r^{up}_{n,t} \leqslant \chi^{up}_{n,e,t} + M\cdot(1 - y^{ESS}_{n,e}) \tag{12}$$

$$0 \leqslant \chi^{up}_{n,e,t} + M\cdot y^{ESS}_{n,e} \tag{13}$$

$$r^{dn}_{n,t} \leqslant \chi^{dn}_{n,e,t} + M\cdot(1 - y^{ESS}_{n,e}) \tag{14}$$

$$0 \leqslant \chi^{dn}_{n,e,t} + M\cdot y^{ESS}_{n,e} \tag{15}$$

3. Two-stage stochastic programming approach

To address the uncertainty issue of load demand and electricity prices in DSEP, we further develop the deterministic MILP model as a two-stage stochastic program. In the first stage, optimal expansion decisions of the *master problem* are obtained, then in the second stage, the *subproblem* is solved to minimize the expected operation cost under previous investments. For each of the typical scenarios considered, the continuous solution of the second-stage economic operation is incorporated into the *master problem* for better planning decisions. When the difference between two adjacent iterations of the binary decision variables becomes negligible, the final optimal expansion planning scheme is attained.

To reduce excessive decision variables as the number of scenarios increases, we assume that in the following context all facilities should go into operation once being built, which means the binary operating variable y equals to the planning decision x. For convenience, we use x to represent all these investment/operating decisions on lines, substations and ESSs.

3.1. Master problem: first-stage expansion decisions

We assume that the substations, ESSs and feeders are invested, owned and operated by the distribution system operator (DSO). The first-stage objective function (16) aims to minimize the overall cost spent during the planning stage, which can be divided into two groups: 1) the investment cost C_{INV} and maintenance cost C_{MAT} for distribution facilities. 2) the mathematical expectation of the operation cost in multiple scenarios, denoted by the recourse function $Q(x, y_s)$. The variable y_s which is continuous herein represents the optimal value of all decision variables in the second stage.

$$\min \sum_i \gamma^i \left(C_{INV} + C_{MAT} + \sum_{s=1}^{S} \theta_s Q\left(x, y_s\right) \right) \tag{16}$$

$$\sum_{j\in\Psi_{eL}} x^{eL}_j + \sum_{k\in\Psi_{nL}} \sum_z x^{nL}_{k,z} = \Gamma - N_{SUB} \tag{17}$$

$$\sum_b x^{SUB}_{m,b} \leq 1,\ \forall\ m \in \Omega_{SUB} \tag{18}$$

$$\sum_e x^{ESS}_{n,e} \leq 1,\ \forall\ n \in \Omega_{ESS} \tag{19}$$

$$x^{eL}_j,\ x^{nL}_{k,z},\ x^{SUB}_{m,b},\ x^{ESS}_{n,e} \in \{0, 1\}$$
$$\forall\ j \in \Psi_{eL},\ k \in \Psi_{nL},\ m \in \Omega_{SUB},\ n \in \Omega_{ESS} \tag{20}$$

Eqs. (17)–(20) are constraints for the *master problem*, which have been explained in Section 2.2.1. Besides, as indicated in Eq. (20), all the first-stage decision variables of the MILP model are binary.

3.2. Subproblem: second-stage operation decisions

After a solution is obtained in the *master problem*, a *subproblem* can be solved where all decision variables are continuous. In the recourse function (21), we add the power transaction cost $W^{LMP}_{s,t} p^{SUB}_{m,b,s,t}$, the penalty due to unserved loads $M_{cur} d^{cur}_{l,s,t}$ and the battery degradation term $M_{deg,t} \varphi^T_s \beta^{ESS}_{n,s,t}$, then minus the revenue earned by ESSs' selling regulation services ($W^{up}_{REG,s,t} r^{up}_{n,s,t} + W^{dn}_{REG,s,t} r^{dn}_{n,s,t}$), where M_{cur} is a large number to avoid load curtailment. Compared with model formulation in Section 2, the subscript s is added to denote that the operation decision variables in $Q(x, y_s)$ are different among scenarios. It is necessary to determine the operation state of feeders, substations and ESSs in each scenario s.

$$Q\left(x, y_s\right) = \min \sum_{t=1}^{T} \left(\sum_{m\in\Omega_{SUB}} W^{LMP}_{s,t} p^{SUB}_{m,b,s,t} + \sum_{l\in\Omega_D} M_{cur} d^{cur}_{l,s,t} + \sum_{n\in\Omega_{ESS}} M_{deg,t} \varphi^T_s \beta^{ESS}_{n,s,t} - \sum_{n\in\Omega_{ESS}} \left(W^{up}_{REG,s,t} r^{up}_{n,s,t} + W^{dn}_{REG,s,t} r^{dn}_{n,s,t} \right) \right) \tag{21}$$

$$\Xi^{eL} \boldsymbol{f^{eL}_{s,t}} + \Xi^{nL} \boldsymbol{f^{nL}_{s,t}} + \sum_{l\in\Omega_D} d^{cur}_{l,s,t} + \sum_{m\in\Omega_{SUB}} p^{SUB}_{m,b,s,t} = \sum_{l\in\Omega_D} D_{l,s,t} - \sum_{n\in\Omega_{ESS}} \left(p^{dis}_{n,s,t} - p^{ch}_{n,s,t} + \omega^{up}_{s,t} r^{up}_{n,s,t} - \omega^{dn}_{s,t} r^{dn}_{n,s,t} \right) \cdot \left(1hr. \right) \tag{22}$$

$$|Z^{eL}_j f^{eL}_{j,s,t} + [\Xi^{eL}]^T_{Row j} u_{s,t}| \leqslant M\cdot(1 - x^{eL}_j),\ \forall\ j \in \Psi_{eL} \tag{23}$$

$$|Z^{nL}_{k,z} f^{nL}_{k,s,t} + [\Xi^{nL}]^T_{Row k} u_{s,t}| \leqslant M\cdot(1 - x^{nL}_{k,z}),\ \forall\ k \in \Psi_{nL} \tag{24}$$

$$0 \leqslant d^{cur}_{l,s,t} \leqslant D_{l,s,t},\ \forall\ l \in \Omega_D \tag{25}$$

$$0 \leqslant p^{SUB}_{m,b,s,t} \leqslant \sum_b x^{SUB}_{m,b} E^{SUB}_{b,max},\ \forall\ m \in \Omega_{SUB} \tag{26}$$

$$\left| f^{eL}_{j,s,t} \right| \leqslant x^{eL}_j F^{eL}_{j,max},\ \left| f^{nL}_{k,s,t} \right| \leqslant \sum_z x^{nL}_{k,z} F^{nL}_{k,z,max},\ \forall\ j \in \Psi_{eL},\ \forall\ k \in \Psi_{nL} \tag{27}$$

$$\xi^{SOC}_{n,s,t+1} = \xi^{SOC}_{n,s,t} - \left(p^{dis}_{n,s,t} - p^{ch}_{n,s,t} + \omega^{up}_{s,t} r^{up}_{n,s,t} - \omega^{dn}_{s,t} r^{dn}_{n,s,t} \right) \cdot (1hr.),\quad \forall\ n \in \Omega_{ESS},\ t = 1, 2, \ldots, 23 \tag{28}$$

$$0 \leqslant \xi^{SOC}_{n,s,t} \leqslant \sum_e x^{ESS}_{n,e} E^{ESS}_{e,max} \tag{29}$$

$$\left(r^{dn}_{n,s,t} + p^{ch}_{n,s,t} \right) \cdot \left(1hr. \right) \leqslant \sum_e x^{ESS}_{n,e} E^{ESS}_{e,max} - \xi^{SOC}_{n,s,t} \tag{30}$$

$$\left(r^{up}_{n,s,t} + p^{dis}_{n,s,t} \right) \cdot (1hr.) \leqslant \xi^{SOC}_{n,s,t} \tag{31}$$

$$\omega^{up}_{s,t} r^{up}_{n,s,t} + p^{dis}_{n,s,t} - \omega^{dn}_{s,t} r^{dn}_{n,s,t} \leqslant \sum_e x^{ESS}_{n,e} P^{ESS}_{e,max} \tag{32}$$

$$\omega^{dn}_{s,t} r^{dn}_{n,s,t} + p^{ch}_{n,s,t} - \omega^{up}_{s,t} r^{up}_{n,s,t} \leqslant \sum_e x^{ESS}_{n,e} P^{ESS}_{e,max} \tag{33}$$

$$r^{up}_{n,s,t} + p^{dis}_{n,s,t} \leqslant \sum_e x^{ESS}_{n,e} P^{ESS}_{e,max} \tag{34}$$

$$r^{dn}_{n,s,t} + p^{ch}_{n,s,t} \leqslant \sum_e x^{ESS}_{n,e} P^{ESS}_{e,max} \tag{35}$$

$$\xi^{SOC}_{n,s,t} = \sum_e x^{ESS}_{n,e} E^{ESS}_{e,0},\ \ t = 1, 24 \tag{36}$$

$$p^{ch}_{n,s,t},\ p^{dis}_{n,s,t},\ r^{up}_{n,s,t},\ r^{dn}_{n,s,t} \geqslant 0 \tag{37}$$

There are two groups of the second-stage constraints: Eqs. (22)–(27) are for distribution system operation, and Eqs. (28)–(37) are for ESS operation. In the first group, we adopt a distribution system power flow model with DC-approximated voltage deviation by introducing KCL in (22), and node voltage limit in (23), (24). Constraints (25)–(27) impose limits on maximal load curtailment, substation's capacity and feeder's capacity, respectively.

By partitioning the battery's capacity into two parts, ESSs can achieve energy arbitrage by charging and discharging, while selling ancillary services by providing the regulation capacity. Eqs. (28)–(35) are physical constraints aligned with the settings of Ref. [6]. Eq. (36) forces all ESSs to maintain the same SOC at the beginning and end hour in a daily scene. The last constraint (37) pertains to the non-negativity of ESS decision variables. It is noteworthy that a Big-M method is also adopted as Eqs. (11)–(15), which is omitted in this section.

4. Solution method

4.1. Modeling of uncertainties

Monte Carlo simulation (MCS) based on single normal distribution may not well describe the diversity in power output, load demand, relevant prices and policies, hence generate scenarios which result in deviation of the optimal solution from the actual economic scheme. In this work, a GMM is introduced with the assumption that the original data follow a linear superposition of K normal distributions [42] instead of a single one. Specifically, we suppose that scenarios are drawn from a joint distribution $P(X|Y)$ where the y_i is a random variable indicating from which normal distribution this scenario x_i is drawn. Therefore, the learning of GMM parameters $\hat{\mu}_k$ and $\hat{\sigma}_k$ lies in the maximal likelihood estimate scheme:

$$\{\hat{\mu}_k, \hat{\sigma}_k\}_1^K = \arg\min \prod_{i=1}^{n} P(x_i|y_i)\cdot P(y_i),\ y_i \in \{1,...,K\} \tag{38}$$

which is tractable through the expectation–maximization algorithm. A scenario is sampled from one of the K normal distributions as:

$$N(\mu_k, \sigma_k),\ \text{for } k = 1,\ ...,K \tag{39}$$

In this work, we set $K = 4$ and for each K generate scenarios according to the weights of different normal distributions. For a given S and K, this sampling process can be described as:

$$\{x_i|y_i = k\}_{i=1}^{S\cdot P(y_i=k)} \sim N(\mu_k, \sigma_k),\ \text{for } k = 1,\ ...,K \tag{40}$$

hence, we obtain S scenarios from these distributions. Fig. 3 illustrates the whole procedure, where scenario x_i is a continuous vector, and the GMM is a weighted sum of four Gaussian distributions, with the corresponding weights $P(y_i = 1), P(y_i = 2), P(y_i = 3), P(y_i = 4)$, respectively. Based on the well estimated parameters of the GMM, we can easily generate adequate scenarios via MCS method.

4.2. Progressive hedging algorithm

To simulate real conditions of distribution systems considering uncertainties in different seasons, days and hours, numerous daily operation scenarios are randomly generated by the MCS method. However, the computation time will become intractable when too many scenarios are included. Hence, we adopt a decomposition technique to address the tractability issue and promote the convergence. To determine a better option between the L-shaped method and the PH algorithm, we compare their basic rules and come to a conclusion.

As for the L-shaped method, the algorithm efficiency relies on iterations between the *master problem* and *subproblem*, i.e. the feasibility cuts and optimality cuts. It is effective for some economic dispatch problems because the feasibility check of the *subproblem* can be omitted. Nevertheless, for a DSEP, the feasibility check of the first-stage planning results is unavoidable. Besides, considering sitting and sizing of facilities like ESSs will add even more computational burden, which probably makes the L-shaped method less efficient in solving DSEP problems.

The significant feature of the PH algorithm, however, lies in the scenario-based decomposition technique. That is, for large-scale MILP model with numerous scenarios, the PH algorithm can solve the *subproblem* in a paralleled way, thus improving the efficiency dramatically. Here, based on the PH algorithm in [38], we decompose the two-stage stochastic model in Section 3 and apply PH to accelerate its convergence.

For notation simplicity, in the following context we define the set of planning decision variables as a new vector X_s in (41) and (42). A modified formulation of $\mathbb{Q}(s)$ in the *subproblem* of scenario s is denoted by (43).

$$\begin{aligned} X_s := \Bigg\{ & x_{j,s}^{\mathrm{eL}}, x_{k,z,s}^{\mathrm{nL}}, x_{m,b,s}^{\mathrm{SUB}}, x_{n,e,s}^{\mathrm{ESS}} \Bigg| \sum_b x_{m,b,s}^{\mathrm{SUB}} \le 1, \\ & \sum_e x_{n,e,s}^{\mathrm{ESS}} \le 1,\ \sum_{j\in\Psi_{\mathrm{eL}}} x_{j,s}^{\mathrm{eL}} + \sum_{k\in\Psi_{\mathrm{nL}}} \sum_z x_{k,z,s}^{\mathrm{nL}} = \Gamma - N_{\mathrm{SUB}}, \\ & x_{j,s}^{\mathrm{eL}}, x_{k,z,s}^{\mathrm{nL}}, x_{m,b,s}^{\mathrm{SUB}}, x_{n,e,s}^{\mathrm{ESS}} \in \{0,\ 1\} \Bigg\} \end{aligned} \tag{41}$$

$$X_s(q) = \begin{cases} x_{j,s}^{\mathrm{eL}} & q = 1 \\ x_{k,z,s}^{\mathrm{nL}} & q = 2 \\ x_{m,b,s}^{\mathrm{SUB}} & q = 3 \\ x_{n,e,s}^{\mathrm{ESS}} & q = 4 \end{cases} \tag{42}$$

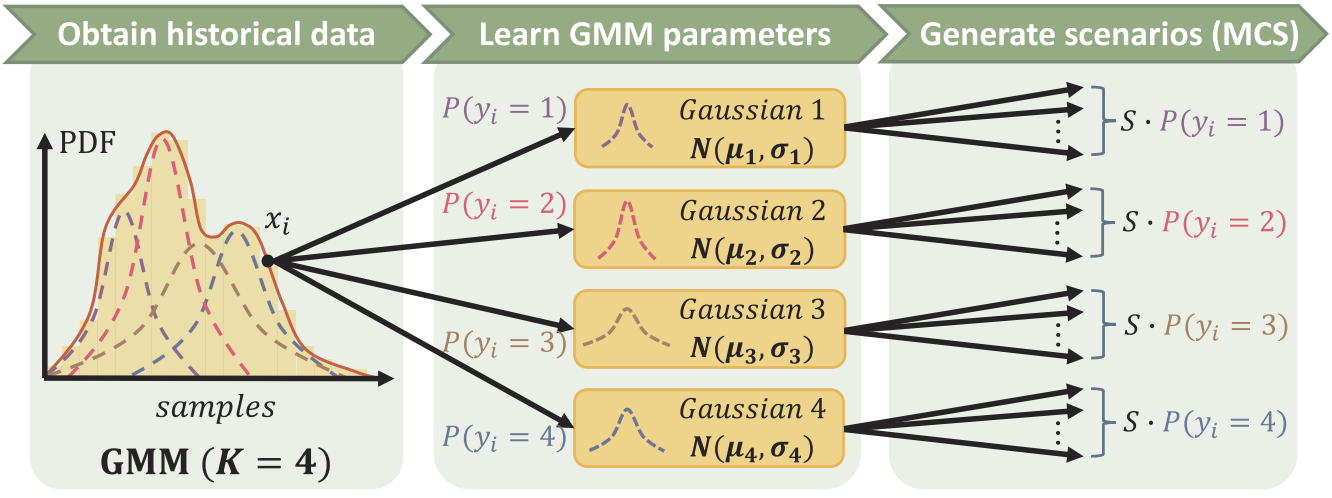

Fig. 3. Procedure of uncertainty modeling (PDF: the probability density function; MCS: the Monto Carlo simulation method based on GMM).

$$\mathbb{Q}(s) = \min_{X_s, y_s} \sum_i \gamma^i \left(C_{\text{INV}} + C_{\text{MAT}} + \sum_{s=1}^{S} \theta_s Q\left(x, y_s\right) \right)$$
$$\text{s.t.} \quad \text{Constraints (22)} - \text{(37)\&(11)} - \text{(15)}. \tag{43}$$

4.2.1. Computing rational vector $\overline{X}$

When making an investment plan, the decision maker may know nothing about which scenario will be realized in the future. To avoid decisions dependent on specific scenarios, the *non-anticipativity* constraints are introduced as Eq. (44).

$$\begin{cases} x_{j,s}^{\text{eL}} = \bar{x}_j^{\text{eL}} & \forall\, j \in \Psi_{\text{eL}} \\ x_{k,z,s}^{\text{nL}} = \bar{x}_{k,z}^{\text{nL}} & \forall\, k \in \Psi_{\text{nL}} \\ x_{m,b,s}^{\text{SUB}} = \bar{x}_{m,b}^{\text{SUB}} & \forall\, m \in \Omega_{\text{SUB}} \\ x_{n,e,s}^{\text{ESS}} = \bar{x}_{n,e}^{\text{ESS}} & \forall\, n \in \Omega_{\text{ESS}} \end{cases} \tag{44}$$

In PH algorithms, these constraints are implicitly implemented. After calculating the average solution $\overline{X}$ of the first-stage decision vector X_s over all scenarios, the deviation $\|X_s - \overline{X}\|$ will be punished in the objective to ensure that all planning decisions in $s \in S$ are prone to approaching a common $\overline{X}$ after finite iterations. In general, $\overline{X}$ is a mathematical expectation derived from $\sum_{s\in S} \theta_s \cdot X_s$, and will be affected by every X_s. However, in DSEP problems, especially those considering large uncertainties, chances are that overload happens in some rare cases where distribution facilities such as feeders, substations and ESSs with larger capacity are necessary. Given this circumstance, if we continue to average planning decision variables over all scenarios, it is more likely to see non-convergence or unacceptable long run-time in the PH algorithm, since $\overline{X}$ will tend to approach the solution with less investment but more frequent occurrence, leading to a large punishment on decision variables of overload scenarios. Nevertheless, the overload scenario has little chance to converge to the average solution, otherwise the *subproblem* will become infeasible.

The solution of a DSEP problem is supposed to be feasible in all scenarios involved. Hence, we give higher priority to the overload scenarios, in other words, the most radical investment solutions and calculate their expectations as a rational vector $\overline{X}$ shown below:

$$\overline{X}^v(q) = \begin{cases} \sum_{s\in S} \theta_s \cdot x_{j,s}^{v,\text{eL}} & q = 1 \\ \sum_{s\in S_{\text{nL}}} \theta_s \cdot x_{k,z,s}^{v,\text{nL}} & q = 2 \\ \sum_{s\in S_{\text{SUB}}} \theta_s \cdot x_{m,b,s}^{v,\text{SUB}} & q = 3 \\ \sum_{s\in S_{\text{ESS}}} \theta_s \cdot x_{n,e,s}^{v,\text{ESS}} & q = 4 \end{cases} \tag{45}$$

where $\overline{X}^v(q)$ is the new average solution of decision variables in iteration v, and S_{nL}, S_{SUB}, S_{ESS} are scenario sets producing planning results with feeders, substations and ESSs of largest capacity, respectively. As for decision variables denoting existing lines' investment, the corresponding $\overline{X}$ is still the mathematical expectation of $x_{j,s}^{v,\text{eL}}$ in all scenarios since the original feeders can not address extreme cases like overload effectively.

As a result, the algorithm converges to a rational solution faster, for the DSEP with more distribution facilities to built is always feasible in normal load scenarios. To avoid over investment in this scheme, sometimes only the average solution of the most crucial and complex decision variables, i.e. ESS investment, will be handled as Eq. (45) does.

4.2.2. Parallel computing process

Since each *subproblem* is independently solved in PH algorithms, the optimization in one scenario does not rely on others. Parallel Computing Toolbox™ with MATLAB interface is adopted here by *parfor*-loops in multiple threads.

Algorithm 1. PH Algorithm with Parallel Computing.

1. **Input:** PH parameters ε, $\rho(q) = \{\rho_{\text{eL}}, \rho_{\text{nL}}, \rho_{\text{SUB}}, \rho_{\text{ESS}}\}$.
2. **Initialization:** $v \leftarrow 1$, $w_s^{v-1} \leftarrow 0$, $g^v \leftarrow 0$, $\forall\, s \in S$.
3. **parfor** $s \in S$ **do:**
4. $\quad X_s^v \leftarrow \text{argmin}_{X_s, y_s} \mathbb{Q}(s)$ s.t. (22)–(37) & (11)–(15).
5. **end parfor**
6. Update: $\bar{X}^v$ according to Eq. (45).
7. **for** $s \in S$ **do:** $w_s^v \leftarrow w_s^{v-1} + \sum \rho(q)(X_s^v - \bar{X}^v)$.
8. **end for**
9. Update: $g^v \leftarrow \sum_{s\in S} \theta_s \cdot \|X_s^v - \bar{X}^v\|$.
10. **while** $g^v \geq \varepsilon$ **do:** $v \leftarrow v + 1$.
11. $\quad$ **parfor** $s \in S$ **do:**
12. $\quad\quad X_s^v \leftarrow \text{argmin}_{X_s, y_s} \left(\mathbb{Q}(s) + w_s^{v-1} \cdot X_s + \frac{\sum \rho(q)}{2} \cdot \left\| X_s - \bar{X}^{v-1} \right\|^2 \right)$.
13. $\quad$ **end parfor**
14. $\quad$ **repeat Step 6–8**.
15. $\quad$ Update: $g(q)^v = \{g_{\text{eL}}^v, g_{\text{nL}}^v, g_{\text{SUB}}^v, g_{\text{ESS}}^v\}$ in **Step 9**.
16. $\quad$ **for** q = 1:4
17. $\quad\quad$ **if** $0 < g(q)^v \leq 1$ **do:** $\rho(q) = (1 + A(q)g(q)^v) \cdot \rho(q)$.
18. $\quad\quad$ **else do:** $\rho(q) = (1 + B(q)(g(q)^v - 1)) \cdot \rho(q)$.
19. $\quad\quad$ **end if**
20. $\quad\quad$ **if** $g(q)^v \geq g(q)^{v-1}$ **do:** $\rho(q) = (2 + A(q)g(q)^v) \cdot \rho(q)$.
21. $\quad\quad$ **end if**
22. $\quad$ **end for**
23. **end while**

Theoretically, if there are enough processors available, the optimization can be solved concurrently with one scenario in a single thread. Consequently, the parallel processing can be applied to the PH designed for Eq. (43) as illustrated in the **Algorithm 1** above.

4.2.3. Gap-dependent penalty factor

Empirically, planning results are found sensitive to the penalty factors. While a large ρ can push to the early termination, it can also harm the economy of facilities to be invested, i.e. the optimality of final solutions. By contrast, a small ρ cannot effectively penalize the deviation of planning variables from their average solution $\overline{X}$. What is worse, for MILP models with binary planning decisions, sometimes the optimization falls into endless loops due to $g(q)^v$ in Step 15 of **Algorithm 1** becoming fixed after several PH iterations. In this case, a belated change of penalty factors can hardly address the situation since the oscillation of some integer variables has occurred, and thus lead to endless cycling.

To deal with it, we adopt gap-dependent $\rho(q)$ values for better convergence rates while ensuring the solution quality. In Step 16–22 of the pseudocode, the gap values $g(q)^v$ of the line, substation and ESS planning decisions are divided into two groups: for those smaller than 1, we perform Step 17 to increase their $\rho(q)$ values; for others larger than 1, Step 18 is conducted. $A(q)$ and $B(q)$ here represent vectors including the tuning parameters for ρ_{eL}^v, ρ_{nL}^v, ρ_{SUB}^v, ρ_{ESS}^v, and can be adjusted for different models while their values are usually around 1.

Both in Step 17 or Step 18, all $\rho(q)$ values are increased to a relatively small extent, and the initial $\rho(q)$ should be consequently smaller than the magnitude of the unit cost (C and O) of different distribution facilities. For early iterations, the PH algorithm will yield large reductions in gaps, then Step 18 will be effective to punish the deviation of binary variables among different scenarios, by increasing the penalty factors with a degree depending on current gaps. Here, $B(q)$ is used to rescale the distance between $g(q)^v$ and 1 for a faster convergence.

However, the majority of PH iterations actually act in narrowing the already tiny gaps which are usually smaller than 1. Then Step 17 comes into play, where we use $(1 + A(q)g(q)^v)$ as the coefficient of $\rho(q)$, to ensure the penalty factors' linearly increasing with the changing gap. Instead of $A(q)g(q)^v$, the coefficient in Step 17 guarantees the continuous growth of $\rho(q)$ values even if $g(q)^v$ becomes quite small. In this case, though, the increment of $\rho(q)$ may not promote a convergence

efficiently. Hence, Step 20–21 act as a soft means of *slamming* compared with that in [38], to avoid a rising $g(q)^v$.

The rapid increase in penalty factors probably results in premature convergence of some decision variables due to excessive punishment in $\|X_s^v - \overline{X}^v\|$. As a result, $B(q)$ tends to be less than $A(q)$, since in early iterations the gaps are large and cause significant growth in the penalty factors of Step 18. In fact, a relatively small $\rho(q)$ can be adopted at first, then increased progressively, yielding less impact on solution quality while ensuring convergence.

5. Case study

5.1. A modified 33-bus distribution system

A 33-bus distribution system [19] has been modified and tested to verify the effectiveness of the proposed method. As shown in Fig. 4, these 5 dotted lines are alternatives for building new feeders, and 32 solid ones represent existing lines. In the planning stage, the topology can be changed with some new feeders built and other existing lines abandoned. As mentioned in Section 2.2.1, no isolated node and loop are allowed in the operation stage, which means only 32 feeders will be operating in the system.

In terms of facility investment, we consider lines with three options varying in impedance, power capacity and unit cost, and relevant parameters are available in [43]. ESSs are considered to be built at the rest 32 nodes except the first one (slack bus), which is the substation node. Options of these facilities in the distribution system are given in Table 2. Besides, data sets for LMP named as $W_{\alpha,t}^{\mathrm{LMP}}$ in Eq. (3) and the regulation price named as $W_{\mathrm{REG},\alpha,t}^{\mathrm{up}}/W_{\mathrm{REG},\alpha,t}^{\mathrm{dn}}$ in Eq. (4) are referred to [44].

5.2. Cases of deterministic MILP model

Assuming that the planning scheme will last for 14 years, four cases are designed with the framework of the deterministic MILP proposed in Section 2 as below:

- Case 1: both regulation service revenue and degradation penalty of ESSs are included in the objective;
- Case 2: ESS degradation penalty is omitted;
- Case 3: regulation services of ESSs are ignored;
- Case 4: no ESS is built in the distribution system.

5.2.1. Network topology

The network topologies of four cases are shown in Fig. 5, where the capacity of Option1<Option2<Option3 for different facilities. For line construction, Case 1 and Case 2 both upgrade one feeder while the number for Case 3 and Case 4 is 2 and 6 respectively. The substation, meanwhile, expands to the highest capacity in Case 3 and Case 4. From the view of ESS deployment, the former two cases build more ESSs with larger capacity than that of the latter two.

These differences demonstrate that without enough storage units, peak shaving in the distribution system will be severely weakened, thus leading to the inevitability of feeder upgrades. Besides, ESSs can discharge to satisfy the increasing load demand and lower the requirements of substation capacity. Apart from Case 4 which does not consider ESSs, less ESSs are constructed in Case 3, which indicates that the revenue from regulation services is crucial to the economy of ESS investment.

5.2.2. Comparison of ESS degradation

To further study the influence of degradation penalty in the objective function, normalization degradation curves of three cases with ESSs are compared in Fig. 6. In practice, the threshold of ESS remaining capacity for the DSO to end its use is set as 80% of the nominal value [11], such that ESSs in Case 2 only work for 7 years before retiring, while the periods for Case 1 and Case 3 are 8 and 11 years, respectively, hence the degradation penalty can prolong ESS lifetime for one year.

On the other hand, without providing regulation services, ESSs tend to have a longer lifespan as the curve of Case 3. Though ESS regulation capacity is much less than that of energy arbitrage, frequent regulation up and down in micro cycles are more harmful to the cell capacity than relatively slow and macro charge–discharge cycles.

To better observe the aging process, ESS's behaviors and its SOC in a typical day of Case 1 are illustrated as Fig. 7. Generally, the number of full charge–discharge cycles is about once per day, which adds up to 3000 to 4000 times in its 8-year lifespan. Besides, the comparison between the LMP and charge/discharge bars demonstrates the energy arbitrage of ESSs, which leverages the electricity price difference between peak and valley load hours.

5.2.3. Economic analysis

In Table 3, all discounted expenses constituting the overall planning cost are listed. Case 1 is the cheapest, while Case 4 is the most expensive one without ESS being built. The second-highest expense is that of Case 3, because no extra revenue can be earned from regulation services, which is crucial to ESS's profitability in a distribution system. And Case 2 is less economical than Case 1 due to ESSs' early retirement.

Considering low ESS profitability in Case 3, higher fare on the substation and feeders is spent, because limited capacity of ESS can not shave the peak load effectively. In Case 1 and Case 2, the ESS investment cost is the same. In other words, once regulation services are considered in the DSEP, the valuation of ESSs is improved significantly. Since Case 2 retires ESSs one year in advance, its total maintenance cost will be reduced. The highest maintenance cost is spent in Case 3, for it invests feeders and the substation with higher capacity. Subsequently, the least cost on purchasing electricity is found in Case 1 since its ESSs can carry out energy arbitrage one year longer than Case 2 and have higher capacity than Case 3. Additionally, the early disposal of ESSs in Case 2 earns less regulation revenue due to the radical charge and discharge behaviors.

In Fig. 8, Case 1–3 are compared with Case 4 (No ESS), where their profits of energy arbitrage is calculated by differences in the power transaction cost shown in Table 3. It is observed from (a) that the investment and maintenance cost of ESSs is always less than their reduction on the upgrade cost of substation and lines, proving the economy of investing ESSs in these cases. In (b), Case 1 with the highest degradation cost also has the highest regulation revenue for a longer storage lifespan. Comparing to energy arbitrage, frequency regulation services provided by ESSs earn more money in the distribution system. This normally happens when the regulation price dominates over the LMP or the LMP has relatively smooth peak and valley.

5.3. Cases of two-stage stochastic MILP model

As for the two-stage stochastic model in Section 3, it is solved in 4, 20, 40 and 100 scenarios by two algorithms mentioned in Section 4.2. Considering uncertainties from load demand, LMP, regulation signals and prices, we try to find a planning scheme of higher reliability and practicability.

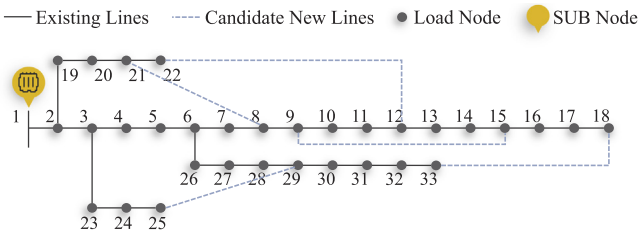

Fig. 4. A modified 33-bus distribution system.

Table 2
Options for substation and ESSs in the distribution system.

Facilities	Different Options			
	Candidate nodes	Capacity (MVA/MWh)	Power (MW)	Construction cost (10^4US$)
SUB	1	10	–	40
		15	–	70
		20	–	110
ESS	2-33	2	0.8	30
		3.5	1.4	50
		5	2	90

5.3.1. Impact of considering uncertainties

In Fig. 9, the topologies of the two-stage stochastic MILP model in 4, 20, 40 and 100 scenarios are illustrated as (a), (b), (c) and (d). When considering more scenarios involving a wider range of uncertainties, the planning scheme tends to become conservative with more feeders built in (c) and (d), which may cope with some extreme cases of overload.

As illustrated in Table 4, planning results in different scenarios are listed and compared with that of deterministic MILP model. As the number of scenarios increases, the overload or unusual regulation signals will be more likely to appear. To ensure the load balancing all the time, more expensive feeders and ESSs need to be built. Consequently, the overall planning result in 100 scenarios is more reliable in real conditions, however, not the most economical one.

Moreover, a positive correlation between the total cost spent in DSEP and the considered scenarios does not exist. Actually, the highest expense is shown in the result of S = 40 as 6.6729 million dollars, since less scenarios generated from normal distributions have a higher randomness which may lead to a deviation from the optimal solution. With more scenarios included in S = 100, the power transaction cost and regulation revenue reflecting ESS's participation in energy arbitrage and regulation services tend to be the closest to real expectation values among all the cases listed here.

5.3.2. Evaluation of algorithm performance

The performance of the L-shaped method and the modified PH algorithm is compared with the Gurobi 8.1.1 in Table 5. With a 2.40 GHz Intel Core i5 processor and 8 GB of memory, the speed of the modified PH is proved to be nearly 2–8 times of Gurobi optimizer, and 3–15 times of the L-shaped method. However, the superiority of the L-shaped method is revealed in its consistent approximation to the optimum solutions, while the modified PH is more heuristic as discussed before.

Frankly, chances are that time savings obtained by PH-based algorithms may come with a larger final optimality gap (S = 100). Repeated experiments on more effective penalty factors are needed but even so the optimal ρ is not easy to find. Considering that, there still lies space in improving the solution quality of PH without harming its efficiency, with which we are trying to come up in Section 4.2. For DSEP with numerous decision variables, the PH algorithm can accelerate the convergence but not ensure the optimality of planning results.

For our model proposed in Section 3, tuning parameters of the modified PH are listed in Table 6. As mentioned in Section 4.2, we develop three algorithmic enhancements to the basic PH proposed in

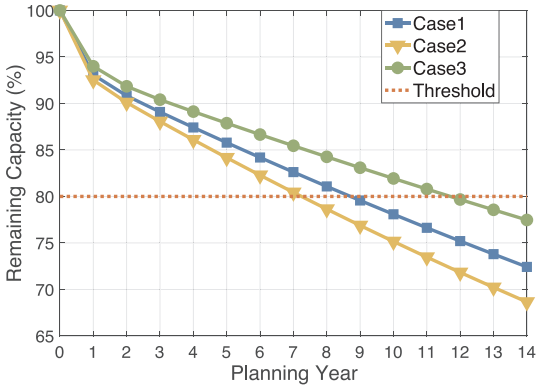

Fig. 6. Capacity degradation behaviors of ESSs in Case 1–3.

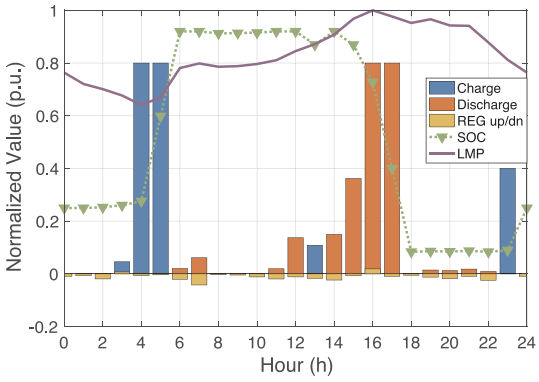

Fig. 7. Relationship between ESS behaviors and the LMP in Case 1.

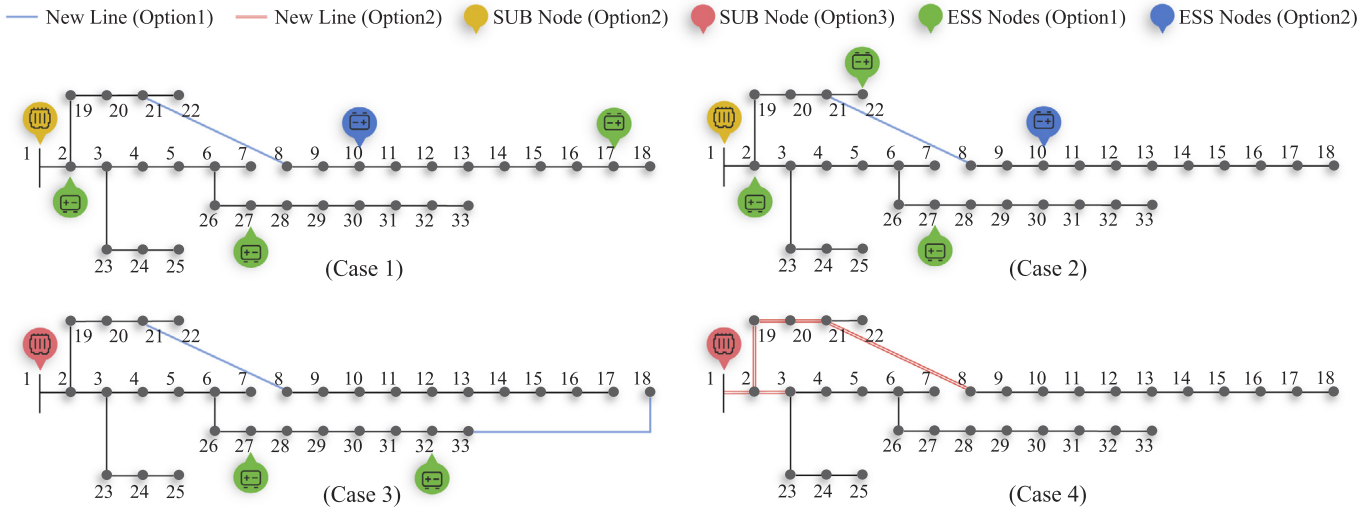

Fig. 5. Final topology of the deterministic MILP in four cases.

Table 3
Discounted planning cost in four cases.

Terms (10^4US$)	Case 1	Case 2	Case 3	Case 4
Regulation service	✓	✓	×	×
ESS degradation	✓	×	✓	×
Total cost	527.14	527.52	535.37	618.37
Line investment	19.80	19.80	39.61	158.43
SUB investment	39.61	39.61	62.24	62.24
ESS investment	79.21	79.21	33.95	**0**
Total maintenance cost	17.11	16.71	18.03	15.53
Power transaction cost	380.51	380.71	381.18	382.18
Regulation revenue	9.74	8.52	**0**	**0**
Degradation penalty	0.63	**0**	0.37	**0**

[38]. To prove their effectiveness, four PH algorithms are defined and compared as follows:

- Basic PH: neither computes a rational average solution in Section 4.2.1 nor adopts gap-dependent penalty factors in Section 4.2.3;
- Average-solution PH: adopts the average solution to implement *non-anticipativity* constraints;
- Gap-dependent PH: adopts gap-dependent penalty factors and uses the same tuning parameters as those in Table 6;
- Modified PH: adopts both algorithmic enhancements and is illustrated as the pseudocode in Section 4.2.2.

Considering that parallel computing technique has been widely used, we develop all above algorithms with *parfor*-loops. Besides, they start with the same initial values of penalty parameters ρ. In Table 7, the performance of four kinds of PH is compared and the superiority of the modified one is reflected in terms of computation time and solution quality.

To save time, we set the maximal iteration as 100. In fact, the running time of one iteration is almost the same among four PH algorithms in single scenario, hence we omit the results of those fail to converge within 100 iterations because their computation time should exceed others significantly. It is noteworthy that both the modified PH and the average-solution PH can converge within 100 iterations in cases with 4–100 scenarios, while the gap-dependent PH only converges in 4 and 20 scenarios. The use of a rational average solution, therefore, is proved superior in stable convergence than computing the mathematical expectation. By cutting down iterations to a certain degree, both average-solution PH and gap-dependent PH are able to accelerate convergence compared with the basic one. That the modified PH reaches the lowest optimality gap with the lowest iteration time among all four algorithms, further demonstrates the effectiveness of combining above-mentioned algorithmic enhancements.

These four PH convergence profiles for the experiments performed in 20 scenarios are given in Fig. 10, in which we provide plots of iteration number versus objective values, to represent the total planning cost in each iteration. The modified PH and average-solution PH have decreasing trends of objective values before convergence, since they both adopt the average solution which reflects more distribution facilities invested to address overload cases. As illustrated in the plot, only basic PH tends to display cycling behaviors within 100 iterations. And the effective use of algorithmic enhancements in Section 4.2.1 and 4.2.3, both yield improvement in convergence rates regarding to objective values of the average-solution PH and gap-dependent one, respectively.

6. Conclusion

In this paper, both regulation services and degradation penalty of energy storage systems are considered to minimize the overall planning cost of the distribution system. The line configuration, substation expansion, storage siting and sizing on the 33-bus distribution network in fixed planning years are optimized via a mixed-integer linear programming model. Subsequently, this deterministic model is reformulated as a two-stage stochastic one when a Gaussian mixture model is adopted to handle with various uncertainties. Then, a modified progressive hedging algorithm is proposed to solve the model with parallel computing process, with which we improve the computation efficiency by 2–8 times compared with the Gurobi optimizer. Moreover, the deterministic planning results demonstrate that with the degradation penalty in the objective, the storage lifetime will be prolonged for one year, and in turn cut down the overall cost during the planning stage.

Numerical results for this method have shown that energy storage systems earn a big potential source of revenue by providing ancillary services, which is comparable with profits of energy arbitrage. Hence, the construction of storage units should be promoted in a distribution network due to their higher profitability. It is noteworthy that considering the storage degradation over a planning horizon is more suitable for real applications and extends lifespan of the expensive energy storage. The use of our modified progressive hedging also allows us to solve the planning problem efficiently in multiple stochastic scenarios, which provides a more reliable investment scheme for practice. For further study, this method will be applied and tested in larger

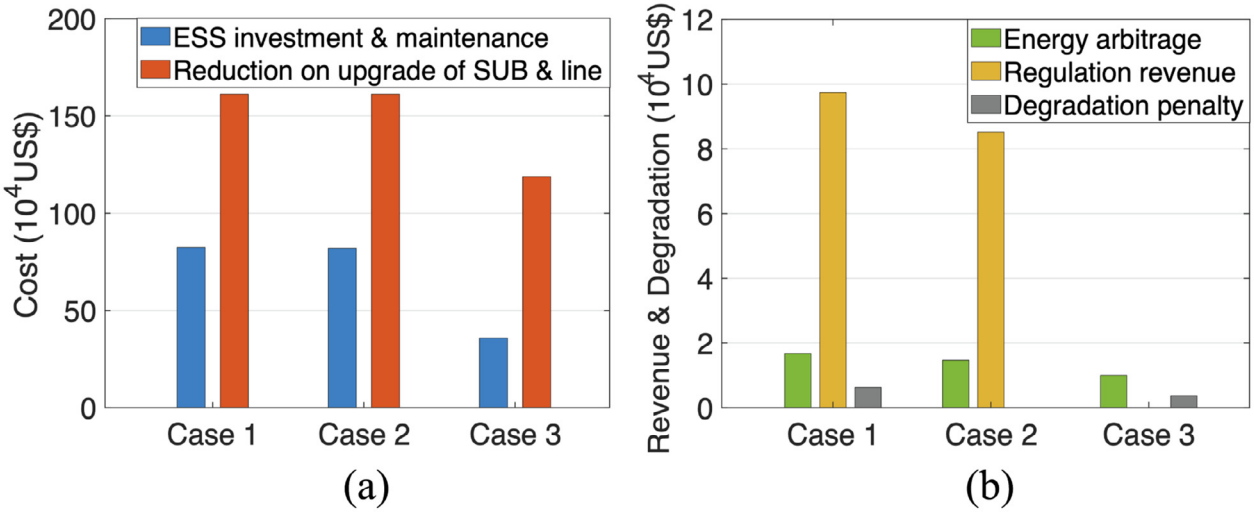

Fig. 8. Profitability of ESSs in Case 1–3.

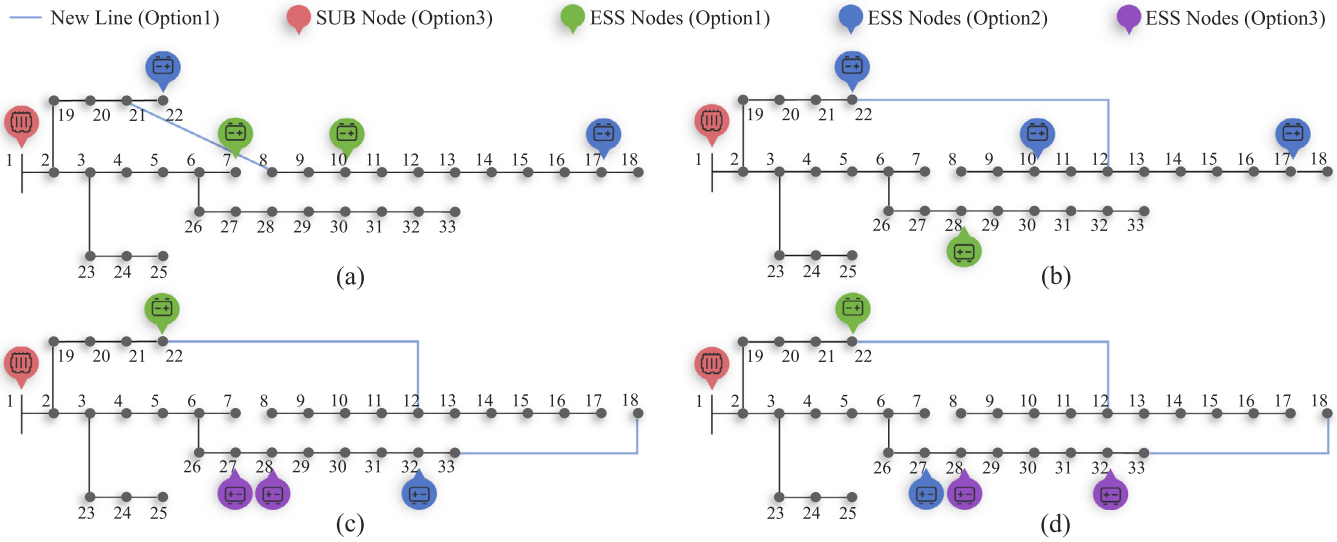

Fig. 9. Network topology of the two-stage stochastic DSEP in scenario 4–100.

Table 4
Comparison of planning results of the deterministic MILP and two-stage stochastic MILP.

DSEP Models	Deterministic	Two-stage stochastic MILP			
(10^4US$)	MILP	S = 4	S = 20	S = 40	S = 100
Total cost	527.14	598.46	603.05	667.29	664.04
Line investment	19.80	19.80	19.80	39.61	39.61
SUB investment	39.61	62.24	62.24	62.24	62.24
ESS investment	79.21	90.53	101.85	147.11	147.11
Total maintenance cost	17.11	19.94	20.39	22.04	22.04
Power transaction cost	380.51	415.78	409.38	409.66	406.26
Regulation revenue	9.74	10.37	11.19	14.07	13.92
Degradation penalty	0.63	0.54	0.58	0.70	0.70

Table 5
Comparison of computation time and solution quality among three solution methods in different scenarios.

Scenario	Method	Time (s)	Objective (10^4US$)	Gap (%)
4	Gurobi	46.213	598.46	0.0027
	L-shaped Method	79.009	598.47	0.0048
	Modified PH	22.112	598.46	0.0023
20	Gurobi	430.607	603.05	0.0122
	L-shaped Method	672.774	603.05	0.0122
	Modified PH	141.251	603.05	0.0122
40	Gurobi	2478.567	667.29	0.0109
	L-shaped Method	2979.775	667.29	0.0109
	Modified PH	312.311	667.29	0.0107
100	Gurobi	4501.718	664.04	0.0416
	L-shaped Method	13916.344	664.04	0.0416
	Modified PH	872.529	672.69	1.3262

Table 6
Tuning parameter values of the modified PH algorithm.

Scenario	Tuning parameters	
	$A(q)$	$B(q)$
4	1, 1, 1, 1	0.4, 0.4, 0.5, 0.8
20	1, 1, 1, 1.2	0.4, 0.4, 0.4, 0.3
40	1.2, 1.2, 1, 1	0.7, 0.8, 0.5, 0.5
100	1.2, 1.5, 1, 1	0.5, 0.8, 0.5, 0.4

Table 7
Comparison of computation time and solution quality among four kinds of PH algorithms in different scenarios.

Scenario	Method	Iteration	Time (s)	Gap (%)
4	Basic PH	17	85.481	0.0023
	Average-sol PH	7	34.913	3.4157
	Gap-dependent PH	5	27.389	0.0023
	Modified PH	4	22.112	0.0023
20	Basic PH	>100	–	–
	Average-sol PH	15	454.266	0.0122
	Gap-dependent PH	13	432.274	0.0122
	Modified PH	6	141.251	0.0122
40	Basic PH	>100	–	–
	Average-sol PH	70	6161.924	0.0107
	Gap-dependent PH	>100	–	–
	Modified PH	5	312.311	0.0107
100	Basic PH	>100	–	–
	Average-sol PH	9	1236.343	1.4447
	Gap-dependent PH	>100	–	–
	Modified PH	5	872.529	1.3262

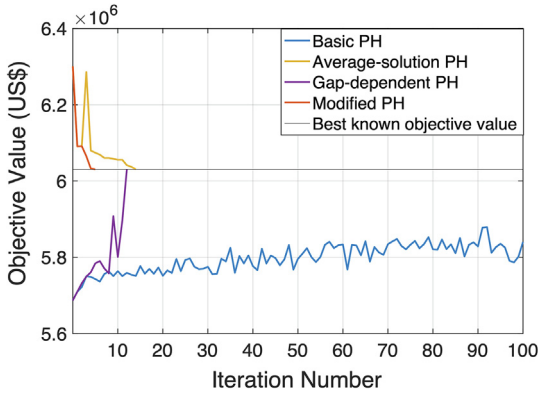

Fig. 10. Objective values throughout iterations of four kinds of PH algorithms.

distribution systems and integrated energy systems.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work is supported by the National Key R&D Program of China (2018YFB0905000), the National Natural Science Foundation of China (NSFC) (51537006), and the Science, Technology and Innovation Commission of Shenzhen Municipality (No. JCYJ20170411152331932).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.apenergy.2020.115520.

References

[1] Parra D, Norman SA, Walker GS, Gillott M. Optimum community energy storage system for demand load shifting. Appl Energy 2016;174:130–43.
[2] Shen X, Shahidehpour M, Han Y, Zhu S, Zheng J. Expansion planning of active distribution networks with centralized and distributed energy storage systems. IEEE Trans Sustainable Energy 2016;8(1):126–34.
[3] Blomgren GE. The development and future of lithium ion batteries. J Electrochem Soc 2016;164(1):A5019.
[4] St John J. Report: Levelized cost of energy for lithium-ion batteries is plummeting. Greentech Media; 2019.
[5] Lazard L. Lazard's levelized cost of storage analysis, version 4.0; 2018.
[6] Foggo B, Yu N. Improved battery storage valuation through degradation reduction. IEEE Trans Smart Grid 2017;9(6):5721–32.
[7] Kumar A, Meena NK, Singh AR, Deng Y, He X, Bansal R, Kumar P. Strategic integration of battery energy storage systems with the provision of distributed ancillary services in active distribution systems. Appl Energy 2019;253. 113503.
[8] Gan W, Ai X, Fang J, Yan M, Yao W, Zuo W, Wen J. Security constrained co-planning of transmission expansion and energy storage. Appl Energy 2019;239:383–94.
[9] Geurin SO, Barnes AK, Balda JC. Smart grid applications of selected energy storage technologies. In: 2012 IEEE PES Innovative Smart Grid Technologies (ISGT), IEEE; 2012. p. 1–8.
[10] Merten M, Olk C, Schoeneberger I, Sauer DU. Bidding strategy for battery storage systems in the secondary control reserve market. Appl Energy 2020;268. 114951.
[11] Xu B, Oudalov A, Ulbig A, Andersson G, Kirschen DS. Modeling of lithium-ion battery degradation for cell life assessment. IEEE Trans Smart Grid 2016;9(2):1131–40.
[12] Choi W-S, Hwang S, Chang W, Shin H-C. Degradation of Co_3O_4 anode in rechargeable lithium-ion battery: a semi-empirical approach to the effect of conducting material content. J Solid State Electrochem 2016;20(2):345–52.
[13] Park J, Appiah WA, Byun S, Jin D, Ryou M-H, Lee YM. Semi-empirical long-term cycle life model coupled with an electrolyte depletion function for large-format graphite/lifepo4 lithium-ion batteries. J Power Sources 2017;365:257–65.
[14] Downing SD, Socie D. Simple rainflow counting algorithms. Int J Fatigue 1982;4(1):31–40.
[15] Cardoso G, Brouhard T, DeForest N, Wang D, Heleno M, Kotzur L. Battery aging in multi-energy microgrid design using mixed integer linear programming. Appl Energy 2018;231:1059–69.
[16] He G, Chen Q, Kang C, Pinson P, Xia Q. Optimal bidding strategy of battery storage in power markets considering performance-based regulation and battery cycle life. IEEE Trans Smart Grid 2015;7(5):2359–67.
[17] Xi X, Sioshansi R. A dynamic programming model of energy storage and transformer deployments to relieve distribution constraints. CMS 2016;13(1):119–46.
[18] Gan L, Li N, Topcu U, Low SH. Exact convex relaxation of optimal power flow in radial networks. IEEE Trans Autom Control 2014;60(1):72–87.
[19] Baran ME, Wu FF. Network reconfiguration in distribution systems for loss reduction and load balancing. IEEE Power Eng Rev 1989;9(4):101–2.
[20] Asensio M, de Quevedo PM, Muñoz-Delgado G, Contreras J. Joint distribution network and renewable energy expansion planning considering demand response and energy storage–Part I: Stochastic programming model. IEEE Trans Smart Grid 2016;9(2):655–66.
[21] Chen C, Sun H, Shen X, Guo Y, Guo Q, Xia T. Two-stage robust planning-operation co-optimization of energy hub considering precise energy storage economic model. Appl Energy 2019;252. 113372.
[22] Baringo L, Boffino L, Oggioni G. Robust expansion planning of a distribution system with electric vehicles, storage and renewable units. Appl Energy 2020;265. 114679.
[23] Zhou Z, Zhang J, Liu P, Li Z, Georgiadis MC, Pistikopoulos EN. A two-stage stochastic programming model for the optimal design of distributed energy systems. Appl Energy 2013;103:135–44.
[24] Rafinia A, Moshtagh J, Rezaei N. Towards an enhanced power system sustainability: An milp under-frequency load shedding scheme considering demand response resources. Sustainable Cities Soc 2020;102168.
[25] Reynolds DA. Gaussian mixture models. Encyclopedia Biomet. 2009;741.
[26] Ryan SM, Wets RJ-B, Woodruff DL, Silva-Monroy C, Watson J-P. Toward scalable, parallel progressive hedging for stochastic unit commitment. In: 2013 IEEE Power & Energy Society General Meeting, IEEE; 2013. p. 1–5.
[27] Crainic TG, Hewitt M, Rei W. Scenario grouping in a progressive hedging-based meta-heuristic for stochastic network design. Comput Oper Res 2014;43:90–9.
[28] Van Slyke RM, Wets R. L-shaped linear programs with applications to optimal control and stochastic programming. SIAM J Appl Math 1969;17(4):638–63.
[29] Rockafellar RT, Wets RJ-B. Scenarios and policy aggregation in optimization under uncertainty. Mathe Oper Res 1991;16(1):119–47.
[30] Birge JR, Louveaux F. Introduction to stochastic programming. Springer Science & Business Media; 2011.
[31] Laporte G, Louveaux FV, Van Hamme L. An integer l-shaped algorithm for the capacitated vehicle routing problem with stochastic demands. Oper Res 2002;50(3):415–23.
[32] Biesinger B, Hu B, Raidl G. An integer l-shaped method for the generalized vehicle routing problem with stochastic demands. Electronic Notes Disc Math 2016;52:245–52.
[33] Berrada A, Loudiyi K, Zorkani I. Valuation of energy storage in energy and regulation markets. Energy 2016;115:1109–18.
[34] González-Garrido A, Gaztañaga H, Saez-de Ibarra A, Milo A, Eguia P. Electricity and reserve market bidding strategy including sizing evaluation and a novel renewable complementarity-based centralized control for storage lifetime enhancement. Appl Energy 2020;262. 114591.
[35] Li Y, Vilathgamuwa M, Farrell TW, Tran NT, Teague J, et al. Development of a degradation-conscious physics-based lithium-ion battery model for use in power system planning studies. Appl Energy 2019;248:512–25.
[36] Dong X, Yuying Z, Tong J. Planning-operation co-optimization model of active distribution network with energy storage considering the lifetime of batteries. IEEE Access 2018;6:59822–32.
[37] Guerrero RC, Angelo M, Pedrasa A. An milp-based model for hybrid renewable energy system planning considering equipment degradation and battery lifetime. In: 2019 IEEE 2nd International Conference on Power and Energy Applications (ICPEA), IEEE; 2019. p. 207–11.
[38] Watson J-P, Woodruff DL. Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. CMS 2011;8(4):355–70.
[39] Liu Y, Sioshansi R, Conejo AJ. Multistage stochastic investment planning with multiscale representation of uncertainties and decisions. IEEE Trans Power Syst 2017;33(1):781–91.
[40] Munoz FD, Watson J-P. A scalable solution framework for stochastic transmission and generation planning problems. CMS 2015;12(4):491–518.
[41] Carvalho PM, Ferreira LA, Barruncho L. Hedging large-scale distribution system investments against uncertainty. In: IEEE Power Engineering Society. 1999 Winter Meeting (Cat. No. 99CH36233), vol. 2, IEEE; 1999. p. 901–06.
[42] Bontemps C, Meddahi N. Testing normality: a gmm approach. J Economet 2005;124(1):149–86.
[43] Zhao X, Shen X. Dataset for stochastic planning of distribution system considering ess regulation services and degradation (Mar 2020). doi:10.6084/m9.figshare.11952531.v1. URL https://figshare.com/articles/Dataset_for_Stochastic_Planning_of_Distribution_System_Considering_ESS_Regulation_Services_and_Degradation/11952531/1.
[44] Foggo B. Replication Data for: Battery Storage Valuation with Optimal Degradation; 2016. doi:10.7910/DVN/KDHAJY. URL https://doi.org/10.7910/DVN/KDHAJY.