# Chance-Constrained Optimization of Storage and PFC Capacity for Railway Electrical Smart Grids Considering Uncertain Traction Load

Yinyu Chen, *Student Member, IEEE,* Minwu Chen, *Member, IEEE,* Lie Xu, *Senior Member, IEEE,* Zongyou Liang

***Abstract*—To foster the utilization of regeneration braking energy and suppress voltage unbalance (VU), a railway electrical smart grid (RESG), intergraded with power flow controller (PFC) and energy storage (ES), is proposed as an important part of next-generation electrified railways. However, under the uncertain traction load, how to design the optimal size of PFC-ES is a challenge during the planning period. Hence, this paper proposes a chance-constrained two-stage programming approach. The first-stage aims to minimising the overall cost of RESG's devices. The second-stage aims to arrange the energy flow of the PFC-ES with the objective of minimising the expected operation cost under the dynamic VU restriction, and the stochastics characteristics of traction load are transformed into a chance constraint by using a scenario approach. Then, traction power predictions are combined with multivariate Gaussian Mixture Model(multi-GMM) model to generate correlated traction power flow scenarios and to assess VU probabilistic metrics distribution with different confidence levels. Finally, a novel algorithm is designed to select the confidence level and violation probability so that the capacity planning results can ensure the high-efficient and high-quality operation of the RESG. Case studies based on an actual electrified railway demonstrate that the proposed PFC-ES sizing approach can reduce the overall cost by up to 13%.**

***Index Terms*—Electrified railways, chance constraint, two-stage programming, probabilistic forecasting.**

## NOMENCLATURE

### A. Abbreviations, Indices and Sets

| | |
|---|---|
| CCP | Chance-constrained programming |
| CI | Confidence interval |
| DoD | Depth of discharge |
| ES | Energy storage |
| GMM | Gaussian mixture model |
| KDE | Ksdensity distribution estimate |
| NZ | Neutral zone |
| O&M | Operation and Maintenance |
| MILP | Mixed-integer linear programming |
| PFC | Power flow controller |
| PSI | Power supply interval |
| UC | Ultracapacitor |
| RBE | Regeneration braking energy |
| RESG | Railways electrical smart grid |
| SoC | State of charge |
| TPSS | Traction power supply system |
| TSS | Traction substation |
| TT | Traction transformer |
| VU | Voltage unbalance |
| VUF | Voltage unbalance factor |
| $t$ | Indices of time intervals during a day |
| $i$ | Indices of chance constraint |
| $\theta$ | Indices of confidence level |
| $\epsilon$ | Indices of violation probability |
| $k$ | Indices of TSS order |

### B. Parameters

| | |
|---|---|
| $\eta_{\mathrm{ch}}^{\mathrm{bt}}$, $\eta_{\mathrm{dis}}^{\mathrm{bt}}$ | Charging and discharging efficiency of battery |
| $\eta_{\mathrm{ch}}^{\mathrm{uc}}$, $\eta_{\mathrm{dis}}^{\mathrm{uc}}$ | Charging and discharging efficiency of UC |
| $\xi^{\mathrm{bt}}$, $\xi^{\mathrm{uc}}$ | Self-discharging coefficient of battery and UC |
| $T_{\mathrm{proj}}$, $T_{\mathrm{day}}$ | Project period (year), operation days of ES within a year |
| $\underline{P^{\mathrm{PFC}}}$, $\overline{P^{\mathrm{PFC}}}$ | Lower and upper bounds of PFC power rating |
| $\underline{P^{\mathrm{bt,uc}}}$, $\overline{P^{\mathrm{bt,uc}}}$ | Lower and upper bounds of battery or UC power rating |
| $\underline{E^{\mathrm{bt,uc}}}$, $\overline{E^{\mathrm{bt,uc}}}$ | Lower and upper bounds of battery or UC capacity rating |
| $\underline{P_t^{\mathrm{Train}}}$, $\overline{P_t^{\mathrm{Train}}}$ | Lower and upper bounds of traction power |
| $c_{\mathrm{pur}}^t$, $c_{\mathrm{ret}}^t$ | Purchased electricity price and penalty charge |
| $c_{\mathrm{dem}}$ | Electricity price of peak demand power |
| $\sigma_l$ | Power loss of converter |
| $c_{\mathrm{om,v}}^{\mathrm{bt,uc}}$, $c_{\mathrm{om,f}}^{\mathrm{bt,uc}}$ | Fixed and variable battery or UC daily O&M cost per unit of power rating |
| $c_{\mathrm{rep}}^{\mathrm{bt,uc}}$ | Battery or UC daily replacement cost |
| $c_{\mathrm{P}}^{\mathrm{bt,uc}}$ | Battery or UC daily capital cost per unit of power rating |
| $c_{\mathrm{P}}^{\mathrm{PFC}}$ | PFC daily capital cost per unit of power rating |

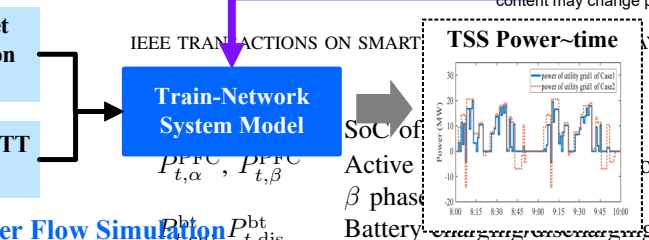
### C. Variables

| | |
|---|---|
| $P_{\mathrm{rate}}^{\mathrm{PFC}}$ | Rated power of PFC |
| $P_{\mathrm{rate}}^{\mathrm{bt}}$, $P_{\mathrm{rate}}^{\mathrm{uc}}$ | Rated power of battery and UC |
| $E_{\mathrm{rate}}^{\mathrm{bt}}$, $E_{\mathrm{rate}}^{\mathrm{uc}}$ | Rated capacity of battery and UC |

This work was supported by the National Natural Science Foundation of China under Grant 52277126.(*Corresponding author: Minwu Chen*)

Y. Chen, M. Chen, Z. Liang are with the School of Electrical Engineering, Southwest Jiaotong University, Chengdu 611756, China (e-mail: yinyuchen@my.swjtu.edu.com; chenminwu@home.swjtu.edu.cn; 2435828243@my.swjtu.edu.cn).

L. Xu is with the Department of Electronic and Electrical Engineering, University of Strathclyde, G1 1XW Glasgow, U.K. (e-mail: lie.xu@strath.ac.uk).




$P^{\mathrm{PFC}}_{t,\alpha}$, $P^{\mathrm{PFC}}_{t,\beta}$ SoC of
Active or $\alpha$ phase and
$\beta$ phase
$P^{\mathrm{bt}}_{t,\mathrm{ch}}$, $P^{\mathrm{bt}}_{t,\mathrm{dis}}$ Battery charging/discharging power
$P^{\mathrm{uc}}_{t,\mathrm{ch}}$, $P^{\mathrm{uc}}_{t,\mathrm{dis}}$ UC charging/discharging power
$E^{\mathrm{bt}}_{t}$, $E^{\mathrm{uc}}_{t}$ Energy power of battery and UC

## I. Introduction

### A. Motivation and Bibliographic Review

ELECTRIFIED railways are considered as one of the strongest contender to revolutionize fossil fuel-based transportation systems, which possess lower carbon emissions and the highest transportation efficiency [1]. The traction power supply systems (TPSSs) are the core component of electrified railways, which provide huge energy to rolling stocks. However, TPSSs pose significant challenges to the power systems such as the severe voltage unbalance (VU) [2], great peak-valley difference and enormous regeneration braking energy (RBE) [3], which interferes with safety and efficient operations.

To cope with this problem, the railways electrical smart grid (RESG) is considered as the next-generation TPSS pa... which utilizes the power flow controller (PFC) and e... storage (ES) to efficiently adjust the power flow distrit... in the transportation system [4], [5]. As shown in Fig. ... PFC installed in the traction substation (TSS) and neutral ... (NZ) can effectively smooth the energy interaction bet... the transportation system and connected power grid. Th... can serve as a flexible source which can charge/dischar... power in different timescales.

As a result, railway energy management has beco... research frontier. Reference [6] proposed energy manage... models in terms of ES, while [7] presented a model pred... control model for the TPSS via a hierarchical structure. ..., a comprehensive railway energy management system w... centralized/decentralized architecture was proposed, wh... [9] presented an energy management model for coordin... energy flow between EV charging stations and urban tra... port system. Reference [10] developed an optimal operati... strategy of distributed energy resources and urban electri... transportation systems. On the other hand, due to the hu... power requirement of electric train (e.g., the rated power ... CRH-380AL is 20 MW) which appears as a single-phase u... balanced variable load in the power system, it is indispensa... for limiting the level of VU emissions from RESG. In [11], ... economic dispatch scheme of RESG was proposed in ter... of dynamic VU constraint, while [12] proposed a two-sta... robust optimization model for RESG considering a stable V... limit as a constraint. In [13], a computation-efficient pow... programming strategy was proposed to regulate power qual... effectively.

Although there are certain benefits of using PFC and ES on railways, their installation and operation costs are high. Therefore, optimizing the size of the PFC and ES to the actual needs of the operation is a critical phase in achieving high-efficient and high-quality operation of RESG. Reference [14] proposed a technical-economic model to design the capacity

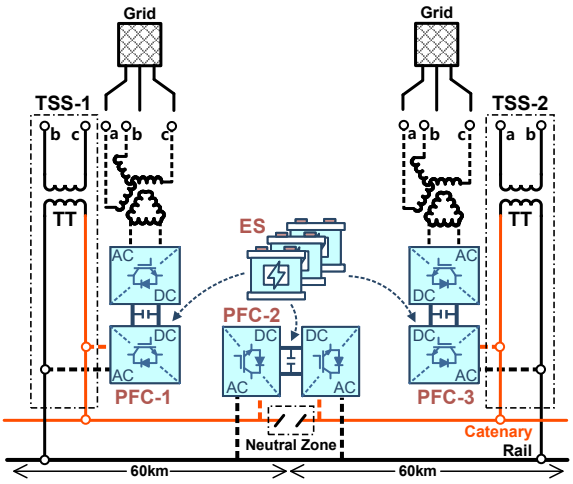
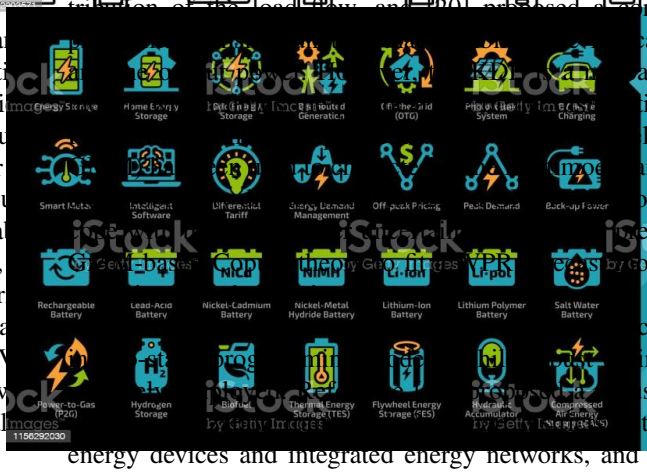



Fig. 1. Structure of RESGs with PFC and ES intergrated.

of the ES, while [15] employed the double-layer optimization model to solve the optimal size of the ES. A mixed-integer linear programming (MILP) model was presented in [16] to determine the ES size utilized by RBE. purpleIn light of this, the optimal capacity models of RESG are usually formulated as two-stage programming problems, where the capacity design and system operation are in different optimization layers. It's worth observing that previous researches are based on deterministic power flow predictions and do not consider the unavoidable prediction errors in practice. In reality, the power flow of network-train interaction is uncertain due to the stochastic nature of drivers' different driving habits and the parameter deviations of the traction force calculation [17]. Therefore, how to accurately model uncertain traction loads is a critical problem. On the one hand, the Ksdensity Distribution Estimate (KDE) is introduced to capture the probabilistic features, including traction power and renewable power [18]–[20]. For example, [18] adopted a forecasted bin to capture the statistical properties of harmonic emissions, [19] presented ... nonparametric probabilistic load flow to determine the dis... tribution of the load flow, and [20] proposed a conditional ... cast error ... arametric ... ion [21]. ... l (multi- ... and integ- ... posed to ... esented a ... odel with ... core step ... imization ... st opera- ... th multi-



energy devices and integrated energy networks, and a robust optimization method was proposed in [25] to address the uncertainties associated with renewable generations and loads. However, this approach solves the best economic operation problem under the worst-case scenario, resulting in high costs for maintaining the scheme's reliability. Alternatively, chance-constrained programming (CCP) provides a promising

TABLE I
TAXONOMY OF THE REVIEWED LITERATURE

| Ref. | Uncertainty | ES cycle life | Power quality | Objective Function | | | Variable | |
|---|---|---|---|---|---|---|---|---|
| | | | | Investment, O&M cost | Replacement cost | Energy Cost | ES sizing | PFC sizing |
| [6] | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | ✗ | ✗ |
| [11] | ✗ | ✓ | ✓ | ✗ | ✗ | ✓ | ✗ | ✗ |
| [12] | ✓ | ✗ | ✓ | ✗ | ✗ | ✓ | ✗ | ✗ |
| [14] | ✗ | ✗ | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ |
| [15] | ✗ | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ | ✗ |
| [16] | ✗ | ✓ | ✗ | ✓ | ✗ | ✗ | ✓ | ✓ |
| Proposed method | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

✓=Considered, ✗=Not considered

alternative that allows the violation of constraints within a confidence level. In [26] , an adjustable chance-constrained approach was proposed to allocate flexible ramping capacity reserves optimally, whereas the total storage power and energy constraints were posed as chance constraints in [27] , and a conservative convex approximations were employed for tractability.

### *B. Aim and Contribution*

Corresponding to the abovementioned literature survey, there is a growing interest in the optimal planning and operating of RESG. However, there are certain gaps from the following perspectives.

1) Uncertain traction load demand is a critical challenge rather than renewable energy in [6] and [12], which will interfere with the capacity decision. It is concluded from Table I that existing approaches based on deterministic numerical calculations cannot obtain a feasible size for RESG's facilities, since they do not illustrate the impacts of uncertainty on the power flow distribution.
2) VU compensation and energy cost are two major concerns of single-phase electrified railways, the serious VU problem will interfere with the economical operation of RESG, different optimization methods in the planning or operating period are summarized in Table I, which is practically incomplete. Notice that the planning approach in [14] can only obtain a feasible solution satisfying the constraints.

In our previous work [11] a deterministic optimal dispatch model considering the dynamic VU constraints was proposed to minimize the operation cost of electrified railways. In this paper, the previous work is extended by programming the optimal capacity and considering the uncertainty of traction load predictions. The two-stage optimization of PFC and ES capacity of RESG aims at minimizing the investment cost and operation and maintenance (O&M) cost under the stochastic power flow during the engineering period. Compared with the existing research, the contributions of this paper are threefold:

1) A complete confidence prediction framework, including a train-RESG interaction simulation and confidence forecast models, is proposed to capture the uncertainty of traction load. With this framework, a better forecast performance in terms of reliability and sharpness can be available compared to other prediction approaches.
2) The statistical model uses multi-GMM theory to build a joint dependence structure between the actual train power output and its prediction, in which the train power output and VU emission feature at different confidence intervals (CIs) can be obtained. This model can characterize the stochastic nature of traction load more accurately than commonly-used distributions, e.g., Gaussian, Gamma and Nonparametric KDE distributions.
3) A chance constraint on power flow is introduced to account for the impact of uncertain traction load on the participation of the PFC-ES, installed in TSS and NZ, in economic dispatch. A scenario approach is adopted to solve chance constraints. More importantly, based on the VU limit and predicted VU indicator's probability distribution, a novel algorithm is designed for selecting the confidence levels and violation probabilities, thus guarantees that all the operation schemes are restricted by power quality standards. Compared with traditional methods, namely specifying confidence intervals, a reasonable trade-off between economical operation and VU regulation can be achieved.
4) The RESG's infrastructure planning problem is formulated as a stochastic two-stage model, and the capacity of PFC and ES in the TSS and NZ is optimized. Investment costs, replacement costs and O&M costs are incorporated into the first-stage objectives, and an ES degradation model is embedded in the first-stage to prolong the ES lifetime. The second-stage takes the PFC-ES settings determined by the first-stage as input parameters to minimize the expected costs including power consumption cost and demand cost. This programming model provides planners with a flexible capacity design approach under different confidence levels.

### *C. Article's Framework*

The remainder of this paper is organized as follows. Section II describes the details of the RESG structure. Section III formulates the first-stage programming problem. The optimal dispatch problem with chance-constraints in the second-stage is developed in Section IV. The confidence prediction framework for traction power flow is presented in Section V. Section

V presents the solution approach and its implementation. Section VI demonstrates and analyzes the numerical results. Conclusions are given in Section VII.

## II. System Configuration and VU Emission Characteristics

### A. Structure of RESG

A schematic diagram of the investigated structure of RESG is shown in Fig.1. In the traction substation (TSS), a single-phase traction transformer (TT) and PFC can provide continuous power with the same voltage amplitued and phase to the train. Generally, when the TSS is powered by high-voltage transmission grid, the power supply intervals (PSI) distance of the TSS is about 60 km. In the neutral zone (NZ) that is used to ensure electrical insulation between different TSSs, a PFC is connected in shunt on both sides of each PSI to effectively control the energy. It should be noted that the acronym NZ-PFC refers to the PFCs located in the NZ, and the acronym TSS-PFC refers to the PFCs located in the TSS. The DC-link in the PFC provides the interface for the integration of ES.

### B. Power Flow Model

Fig. 2 shows the relationship of power flow between the grid, RESG's infrastructures and trains. For example, during in the traction mode, the main power demand of train ($P^{\mathrm{Load}}_{t,\mathrm{trac}}$) can be supported by the nearly TT and TSS-PFC, the remaining power is supplied by the ES and NZ-PFC, the ES and NZ-PFC provide remaining power to reduce the peak power and smooth the power fluctuant; during in the regenerative braking mode, the energy generated by the train ($P^{\mathrm{Load}}_{t,\mathrm{brak}}$) is preferentially stored by ES ($P^{\mathrm{uc}}_{t,\mathrm{ch}}$, $P^{\mathrm{bt}}_{t,\mathrm{ch}}$) or provided to train in another PSI via NZ-PFC, which can effectively reduce the peak valley difference and improve the utilization of energy.

*1) Transformer Mathematical Model:* In the initial scheme, a single-phase TT is employed to realize the co-phase traction power supply for trains. However, the severe VU problem cannot be ignored. As a result, a PFC combined with a Ynd-connection step-down transformer (SDT) is connected in parallel to the TT, the SDT and TT form a balance transformer, which has a minimum VU when the load flow between the two branches is in balance [11]. The mathematical model can be expressed by:

$$\begin{bmatrix} \dot{I}^{\mathrm{grid}}_{a,t} \\ \dot{I}^{\mathrm{grid}}_{b,t} \\ \dot{I}^{\mathrm{grid}}_{c,t} \end{bmatrix} = \underbrace{\begin{bmatrix} 1/N_1 & -1/3N_2 \\ 0 & 2/3N_2 \\ -1/N_1 & 1/3N_2 \end{bmatrix}}_{\mathcal{C}} \begin{bmatrix} \dot{I}^{\mathrm{TT}}_{t} \\ \dot{I}^{\mathrm{PFC}}_{t} \end{bmatrix} \tag{1a}$$

$$\begin{bmatrix} \dot{U}^{\mathrm{TT}}_{t} \\ \dot{U}^{\mathrm{PFC}}_{t} \end{bmatrix} = \underbrace{\begin{bmatrix} 1/N_1 & 0 & -1/N_1 \\ -1/3N_2 & 2/3N_2 & 1/3N_2 \end{bmatrix}}_{\mathcal{V}^{-1}} \begin{bmatrix} \dot{U}^{\mathrm{grid}}_{a,t} \\ \dot{U}^{\mathrm{grid}}_{b,t} \\ \dot{U}^{\mathrm{grid}}_{c,t} \end{bmatrix} \tag{1b}$$

where $N_1$ and $N_2$ represent the transformer ratio of TT and SDT. $\dot{I}^{\mathrm{PFC}}_t$, $\dot{I}^{\mathrm{TT}}_t$, $\dot{U}^{\mathrm{PFC}}_t$ and $\dot{U}^{\mathrm{TT}}_t$ represent the current and voltage of the PFC branch and TT branch, respectively. $\dot{I}^{\mathrm{grid}}_{a,t}$, $\dot{I}^{\mathrm{grid}}_{b,t}$, $\dot{I}^{\mathrm{grid}}_{c,t}$, $\dot{U}^{\mathrm{grid}}_{a,t}$, $\dot{U}^{\mathrm{grid}}_{b,t}$ and $\dot{U}^{\mathrm{grid}}_{c,t}$ represents the phase current and voltage of three-phase in power grid, respectively.

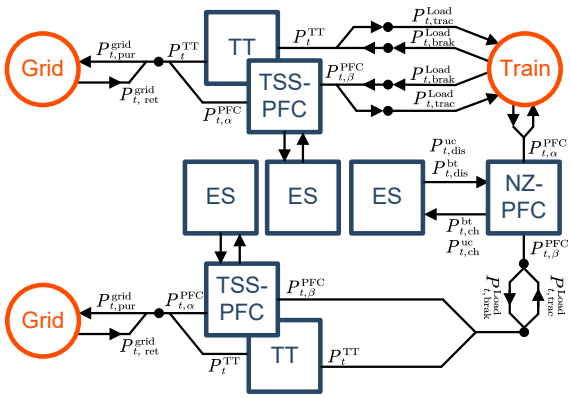



Fig. 2

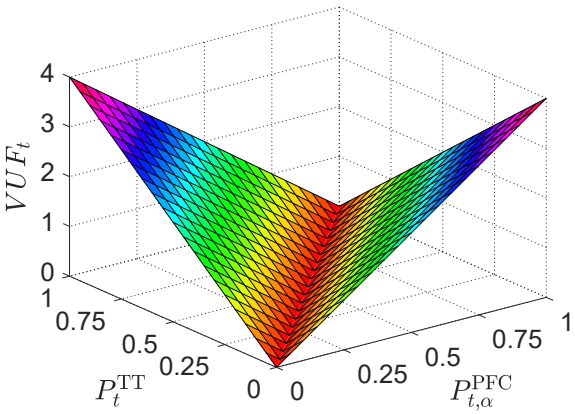



Fig. 3. VU Emission Characteristics of RESG.

$\mathcal{C}$ and $\mathcal{V}$ are defined as current and voltage transformation matrices, respectively. The subscript $t$ represents the time index in a day, and $t \in [0, \mathrm{T}]$.

*2) ES Mathematical Model:* The charging or discharging power and stored energy of the storage are constrained by the physical characteristics of the battery and UC at any time interval. The working status of ES can be evaluated by using the state of charge (SoC):

$$SoC^{\mathrm{bt}}_{t+1} = \left(1 - \xi^{\mathrm{bt}}\right) SoC^{bt}_{t-1} + \eta^{\mathrm{bt}}_{\mathrm{ch}} \frac{P^{\mathrm{bt}}_{t,\mathrm{ch}}}{E^{\mathrm{bt}}_t} \Delta t - \frac{P^{\mathrm{bt}}_{t,\mathrm{dis}}}{\eta^{\mathrm{bt}}_{\mathrm{dis}} E^{\mathrm{bt}}_t} \Delta t \tag{2a}$$

$$SoC^{\mathrm{uc}}_{t+1} = \left(1 - \xi^{\mathrm{uc}}\right) SoC^{\mathrm{uc}}_{t-1} + \eta^{\mathrm{uc}}_{\mathrm{ch}} \frac{P^{\mathrm{uc}}_{t,\mathrm{ch}}}{E^{\mathrm{uc}}_t} \Delta t - \frac{P^{\mathrm{uc}}_{t,\mathrm{dis}}}{\eta^{\mathrm{uc}}_{\mathrm{dis}} E^{\mathrm{uc}}_t} \Delta t \tag{2b}$$

where $SoC^{\mathrm{bt}}_{(\bullet)}$ and $SoC^{\mathrm{uc}}_{(\bullet)}$ represent the SoC of battery and UC, respectively. $P^{\mathrm{bt}}_{t,\mathrm{ch}}$, $P^{\mathrm{uc}}_{t,\mathrm{ch}}$, $P^{\mathrm{bt}}_{t,\mathrm{dis}}$ and $P^{\mathrm{uc}}_{t,\mathrm{dis}}$ represent battery and UC charging/discharging power, respectively. $E^{\mathrm{bt}}_t$ and $E^{\mathrm{uc}}_t$ represent energy of battery and UC, respectively. $\eta^{\mathrm{bt}}_{\mathrm{ch}}$, $\eta^{\mathrm{uc}}_{\mathrm{ch}}$, $\eta^{\mathrm{bt}}_{\mathrm{dis}}$ and $\eta^{\mathrm{uc}}_{\mathrm{dis}}$ represent the charging and discharging efficiency of battery and UC, respectively. $\xi^{\mathrm{bt}}$ and $\xi^{\mathrm{uc}}$ represent self-discharging coefficient of battery and UC. $\Delta t$ represents the sampling interval.

*3) PFC Mathematical Model:* The converters at both TSS and NZ consist of a $\alpha$ phase converter and a $\beta$ phase inverter, as shown in Fig. 1. The active power balance formulation is declared in (3):

$$P^{\mathrm{PFC}}_{t,\alpha} + P^{\mathrm{bt}}_{t,\mathrm{dis}} + P^{\mathrm{uc}}_{t,\mathrm{dis}} = P^{\mathrm{PFC}}_{t,\beta} + P^{\mathrm{bt}}_{t,\mathrm{ch}} + P^{\mathrm{uc}}_{t,\mathrm{ch}} \tag{3}$$

where $P^{\mathrm{PFC}}_{t,\alpha}$ and $P^{\mathrm{PFC}}_{t,\beta}$ represent the input/output active power of converter for $\alpha$ phase and $\beta$ phase, respectively.

### C. VU Emission Characteristics of RESG

To evaluate the VU emission level of TSS, the $VUF_t$ index in this paper is defined as the ratio of the modulus of the negative-sequence to the positive-sequence components of the voltage:

$$VUF_t = \left| \frac{\dot{U}^{\text{grid}}_{n,t}}{\dot{U}^{\text{grid}}_{p,t}} \right| = \frac{\sqrt{3}U^{\text{grid}}_{abc,t}}{S_{sc}} I^{\text{grid}}_{n,t} \tag{4}$$

where $\dot{U}^{\text{grid}}_{p,t}$, $\dot{U}^{\text{grid}}_{n,t}$, and $\dot{I}^{\text{grid}}_{n,t}$ represent the negative- and positive- sequence voltage and current in power grid, respectively. It can be derived from (1) by adopting the Fortescue's transformation matrix. Furthermore, $\dot{I}^{\text{grid}}_{n,t}$ can be denoted by $I^{\text{TT}}_t$ and $I^{\text{PFC}}_t$, which can be calculated as:

$$\begin{aligned} I^{\text{grid}}_{n,t} &= \frac{1}{\sqrt{3}N_1} I^{\text{TT}}_t - \frac{1}{3N_2} I^{\text{PFC}}_t \\ &= \frac{1}{\sqrt{3}N_1} \frac{P^{\text{TT}}_t}{U^{\text{TT}}_t} - \frac{1}{3N_2} \frac{P^{\text{PFC}}_{t,\alpha}}{U^{\text{PFC}}_t} \end{aligned} \tag{5}$$

When the transformer setting and rated voltage are fixed, the VU emission feature can be determined according to (4) and (5) as shown in Fig. 3. It is indicated that the VUF at PCC is restricted by power flow distribution of the two branches.

## III. Programming Problem of PFC and ES Sizing for RESG

### A. Objective Function

Consider a RESG comprising of $N$ storages and $M$ power flow controllers. An objective function of first-stage is to minimize the total cost of RESG during an engineering period, including the capital cost ($C^{\text{ES}}_{\text{cap}}$, $C^{\text{PFC}}_{\text{cap}}$), replacement cost ($C^{\text{ES}}_{\text{rep}}$, $C^{\text{PFC}}_{\text{rep}}$), O&M cost ($C^{\text{ES}}_{\text{om}}$, $C^{\text{PFC}}_{\text{om}}$) and energy costs ($\mathbb{E}[\mathcal{J}_2]$) for PFC and ES.

$$\begin{aligned} \min \mathcal{J}_1 &= C_{tot}\left(P^{\text{PFC}}_{\text{rate}}, P^{\text{bt}}_{\text{rate}}, P^{\text{uc}}_{\text{rate}}, E^{\text{bt}}_{\text{rate}}, E^{\text{uc}}_{\text{rate}}\right) \\ &= \underbrace{C^{\text{ES}}_{\text{cap}} + C^{\text{PFC}}_{\text{cap}} + C^{\text{ES}}_{\text{rep}} + C^{\text{ES}}_{\text{om}} + C^{\text{PFC}}_{\text{om}}}_{First-stage} \\ &+ \underbrace{\mathbb{E}[\mathcal{J}_2]}_{Second-stage} \end{aligned} \tag{6}$$

$$\begin{aligned} C^{\text{ES}}_{\text{cap}} = CRF \cdot \sum_{n\in N} \big(&c^{\text{bt}}_{\text{P}} \cdot P^{\text{bt}}_{\text{rate},n} + c^{\text{bt}}_{\text{E}} \cdot E^{\text{bt}}_{\text{rate},n} \\ &+ c^{\text{uc}}_{\text{P}} \cdot P^{\text{uc}}_{\text{rate},n} + c^{\text{uc}}_{\text{E}} \cdot E^{\text{uc}}_{\text{rate},n}\big) \end{aligned} \tag{7}$$

$$C^{\text{PFC}}_{\text{cap}} = \sum_{m\in \text{M}} \left[c^{\text{PFC}}_{\text{P}} \cdot P^{\text{PFC}}_{\text{rate},m} \cdot (1 + X/D)\right] \tag{8}$$

$$\begin{aligned} C^{\text{ES}}_{\text{rep}} = CRF \cdot \sum_{n\in N} \left[\sum_{\text{Nbt}} \left(c^{\text{bt}}_{\text{rep}} \cdot E^{\text{bt}}_{\text{rate},nbt} \cdot PVF\right)\right] \\ + \sum_{n\in N} \left[\text{Nuc} \cdot c^{\text{uc}}_{\text{rep}} \cdot E^{\text{uc}}_{\text{rate},nuc}\right] \end{aligned} \tag{9}$$

$$\begin{aligned} C^{\text{ES}}_{\text{om}} = &\sum_{n\in N} \left(T^{\text{bt}}_{\text{hr}} \cdot c^{\text{bt}}_{\text{om,v}} \cdot P^{\text{bt}}_{\text{rate}} + T^{\text{uc}}_{\text{hr}} \cdot c^{\text{uc}}_{\text{om,v}} \cdot P^{\text{uc}}_{\text{rate}}\right) \\ &+ \sum_{n\in N} \left(c^{\text{bt}}_{\text{om,f}} \cdot P^{\text{bt}}_{\text{rate}} + c^{\text{uc}}_{\text{om,f}} \cdot P^{\text{uc}}_{\text{rate}}\right) \Big/ T_{\text{day}} \end{aligned} \tag{10}$$

$$C^{\text{PFC}}_{\text{om}} = C^{\text{PFC}}_{\text{cap}} \cdot 5\% \cdot T_{\text{proj}} \tag{11}$$

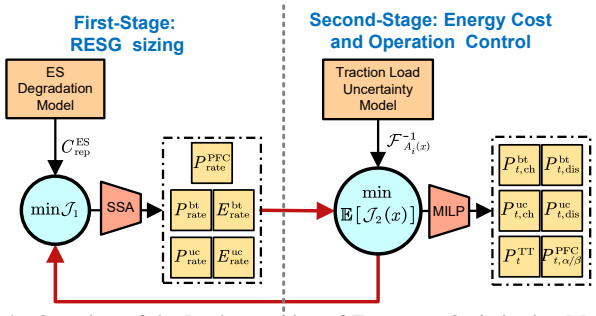



Fig. 4. Overview of the Implementation of Two-stage Optimization Model.

where $P^{\text{PFC}}_{\text{rate}}$ is the rated power of PFC, $P^{\text{bt}}_{\text{rate}}$ and $E^{\text{bt}}_{\text{rate}}$ are the rated power and total capacity of battery respectively, $P^{\text{uc}}_{\text{rate}}$ and $E^{\text{uc}}_{\text{rate}}$ are the rated power and total capacity of UC respectively, $T_{\text{proj}}$, $T_{\text{day}}$, $T^{(\bullet)}_{\text{hr}}$ are the project period (year), operation days of ES within a year (365 in this paper) and daily operating time of ES respectively. $PVF$ is the present value function to discount the future battery replacement cost incurred in $T_{\text{proj}}$ year to time zero, $CRF$ is the capital recovery function associated with the annual discount rate and project period [15]. $D$ and $X$ are the number of major parallel inverters and redundant modules for PFC [29].

### B. Constraints

To minimize the total cost within the time horizon of the project service period, The sizing of the ES involves finding the optimal power rating and energy capacities of batteries and UCs, while the sizing of the PFC involves finding the optimal power rating. Since the equipment sizing requirement of ES ($\text{P}^{\text{bt}}_{\text{cell}}$, $\text{E}^{\text{bt}}_{\text{cell}}$, $\text{P}^{\text{uc}}_{\text{cell}}$, $\text{E}^{\text{uc}}_{\text{cell}}$) and PFC ($\text{P}^{\text{PFC}}_{\text{cell}}$) in engineering applications, the decision variables are formulated as the integer variable modeling the number of minimum units in ES ($q^{\text{bt}}_n$, $q^{\text{uc}}_n$) and PFC ($q^{\text{PFC}}_m$), and Equations (12)–(14) state that these decision variables are limited by upper and lower bounds, as:

$$\text{MinP}^{\text{PFC}}_{\text{cell}} \leqslant P^{\text{PFC}}_{\text{rate},m} \leqslant \text{MaxP}^{\text{PFC}}_{\text{cell}} \tag{12}$$

$$\text{MinP}^{\text{bt}}_{\text{cell}} \leqslant P^{\text{bt}}_{\text{rate},n} \leqslant \text{MaxP}^{\text{bt}}_{\text{cell}}, \text{MinE}^{\text{bt}}_{\text{cell}} \leqslant E^{\text{bt}}_{\text{rate},n} \leqslant \text{MaxE}^{\text{bt}}_{\text{cell}} \tag{13}$$

$$\text{MinP}^{\text{uc}}_{\text{cell}} \leqslant P^{\text{uc}}_{\text{rate}} \leqslant \text{MaxP}^{\text{uc}}_{\text{cell}}, \text{MinE}^{\text{uc}}_{\text{cell}} \leqslant E^{\text{uc}}_{\text{rate}} \leqslant \text{MaxE}^{\text{uc}}_{\text{cell}} \tag{14}$$

where,

$$\begin{aligned} P^{\text{bt}}_{\text{rate},n} &= q^{\text{bt}}_n \cdot \text{P}^{\text{bt}}_{\text{cell}}, E^{\text{bt}}_{\text{rate},n} = q^{\text{bt}}_n \cdot \text{E}^{\text{bt}}_{\text{cell}}, \forall q^{\text{bt}}_n \in \mathcal{Z}^+ \\ P^{\text{uc}}_{\text{rate},n} &= q^{\text{uc}}_n \cdot \text{P}^{\text{uc}}_{\text{cell}}, E^{\text{uc}}_{\text{rate},n} = q^{\text{uc}}_n \cdot \text{E}^{\text{uc}}_{\text{cell}}, \forall q^{\text{uc}}_n \in \mathcal{Z}^+ \\ P^{\text{PFC}}_{\text{rate},m} &= q^{\text{PFC}}_m \cdot \text{P}^{\text{PFC}}_{\text{cell}}, \forall q^{\text{PFC}}_m \in \mathcal{Z}^+ \end{aligned}$$

### C. ES Degradation Analysis

A battery's life is dependent on its calendar ageing and cycle ageing. Calendar ageing is determined by temperature, battery state of life, and calendar time. Cycle ageing is mainly determined by the depth of discharge (DoD) and SoC [30]. To take the battery degradation cost into account, the battery lifetime estimation approach introduced in [31], $T_{\text{bt}}$ is the total times of battery replacement during the project period, which is determined by the battery life. By means of the rain-flow

counting method, the cycle life loss rate ($A_{\mathrm{bt}}(DoD)$) can be extracted from a series of SoC.

$$\mathrm{Nbt} = \lceil T_{\mathrm{proj}} \cdot T_{\mathrm{bt}}^{-1} \rceil = 365 \cdot \lceil T_{\mathrm{proj}} \cdot A_{\mathrm{bt}}(DoD) \rceil \quad (15)$$

where $\lceil\rceil$ represents the rounding up operand.

A UC's life is not limited by the cycle ageing, which assumes that UCs are always working under rated conditions provided by the manufacturing specifications. Therefore, the total times of UC replacement ($T_{\mathrm{uc}}$) is equal to the number given in the manufacturing specification.

### D. Optimization Framework and Solution Methodology

The CC-P model for RESG sizing is illustrated in Fig. 4. The is formulated as a two-stage optimization problem. The rated capacity ($E_{\mathrm{rate}}^{\mathrm{bt}}$, $E_{\mathrm{rate}}^{\mathrm{uc}}$) and rated power ($P_{\mathrm{rate}}^{\mathrm{bt}}$, $P_{\mathrm{rate}}^{\mathrm{uc}}$, $P_{\mathrm{rate}}^{\mathrm{PFC}}$) of the RESG is determined in the first-stage decision, which aims at minimizing the total cost $\mathcal{J}_1$. Especially, in order to prolong the battery's lifetime, a battery degradation model is considered for the replacement cost. The power rating and capacity are regarded as the boundary parameters of the second-stage. The second-stage concentrates on minimizing the expected energy cost $\mathbb{E}[\mathcal{J}_2(x)]$, and the optimized PFC power $P_{t,\alpha/\beta}^{\mathrm{PFC}}$ and transformer power $P_t^{\mathrm{TT}}$, the optimized charging/discharging power $P_t^{\mathrm{TT}}$, $P_{t,\mathrm{dis}}^{\mathrm{bt}}$, $P_{t,\mathrm{ch}}^{\mathrm{uc}}$, $P_{t,\mathrm{dis}}^{\mathrm{uc}}$ for battery and UC are available, respectively. The traction load uncertainty problem is described by multi-GMM theory and converted to a chance constraint solved in second-stage. Note that the objective of second-stage is return to the first-stage for assessment of RESG sizing. Owing to the nonlinear battery lifetime estimation, the first-stage is a nonlinear planning problem, which can be effectively solved by the heuristic search algorithms. For example, the Sparrow Search Algorithm (SSA), one of the effective swarm intelligence optimization algorithms, is employed in this paper, the details of SSA algorithm can be found in [31]. The second-stage can be formulated as a MILP model and addressed by CPLEX solver.

## IV. Chance Constrained Optimal Economic Dispatch Problem for RESG

### A. Objective Function

An objective function of second-stage includes two terms, namely, energy consumption cost and demand cost. In practice, the demand charge is paid by billing month, so the cost term is presented as a daily conversion demand charge of each day by introducing the parameter for operating days in a month, as:

$$\begin{aligned}\mathcal{J}_2 = &\sum_{t=1}^{N_{\mathrm{T}}^{\mathrm{day}}} \left(c_{\mathrm{pur}}^t P_{\mathrm{pur},t}^{\mathrm{grid}} + c_{\mathrm{ret}}^t P_{\mathrm{ret},t}^{\mathrm{grid}}\right) \cdot \Delta t \\ &+ \frac{c_{\mathrm{dem}}}{N_{\mathrm{T}}^{\mathrm{day}}} \cdot \max\left(\sum_{t}^{t+14} P_{\mathrm{pur},t}^{\mathrm{grid}}/15\right)\end{aligned} \quad (16)$$

where $P_{\mathrm{pur},t}^{\mathrm{grid}}$ and $c_{\mathrm{pur}}^t$ represent active power supplied by the utility grid and the purchased electricity price. $P_{\mathrm{ret},t}^{\mathrm{grid}}$ and $c_{\mathrm{ret}}^t$ represent regenerative braking power fed back to utility grid and the penalty charge. $c_{\mathrm{dem}}$ represents the electricity price of peak demand power. $N_{\mathrm{T}}^{\mathrm{day}}$ is the total number of time intervals during a day.

### B. Constraints

*1) Power Balance Constraints:* active power balance constraints for each port of RESG are described in (17a)-(17d) . It states that the active power supplied by the power grid ($P_{\mathrm{pur},t}^{\mathrm{grid}}$), ES discharging ($P_{t,\mathrm{dis}}^{\mathrm{bt}}$, $P_{t,\mathrm{dis}}^{\mathrm{uc}}$), feed superfluous power to grid ($P_{\mathrm{ret},t}^{\mathrm{grid}}$). The power balance for NZ-PFC and TSS-PFC is restricted by applying constraints (17d).

$$P_t^{\mathrm{TT}} + P_{t,\alpha}^{\mathrm{NZ-PFC}} + P_{t,\beta}^{\mathrm{TSS-PFC}} \geqslant \overline{P_t^{\mathrm{Train}}} \quad (17a)$$

$$P_t^{\mathrm{TT}} + P_{t,\alpha}^{\mathrm{NZ-PFC}} + P_{t,\beta}^{\mathrm{TSS-PFC}} \leqslant \underline{P_t^{\mathrm{Train}}} \quad (17b)$$

$$P_t^{\mathrm{TT}} + P_{t,\alpha}^{\mathrm{TSS-PFC}} = P_{\mathrm{pur},t}^{\mathrm{grid}} - P_{\mathrm{ret},t}^{\mathrm{grid}} \quad (17c)$$

$$(1+\sigma_l)\left(P_{t,\beta}^{\mathrm{PFC}} - P_{t,\alpha}^{\mathrm{PFC}}\right) = P_{t,\mathrm{dis}}^{\mathrm{bt}} - P_{t,\mathrm{ch}}^{\mathrm{bt}} + P_{t,\mathrm{dis}}^{\mathrm{uc}} - P_{t,\mathrm{ch}}^{\mathrm{uc}} \quad (17d)$$

where $\sigma_l$ is defined as the power loss of converter.

*2) VU Compensation Constraints:* Based on the previous analysis of the VU emission features of electrified railways in Section II, equations (4) and (5) can be linearized by introducing auxiliary non-negative variable $\kappa_{t,\alpha}^{\mathrm{TSS-PFC}}$, $\kappa_{t,\alpha}^{\mathrm{TT}}$, binary variable $v_t$, and large real number $\mathcal{M}$.

$$\kappa_{t,\alpha}^{\mathrm{TSS-PFC}} + \kappa_t^{\mathrm{TT}} \leqslant \frac{VUF^* \cdot S_{\mathrm{sc}}}{\sqrt{3} U_{\mathrm{grid}}} \quad (18a)$$

$$\frac{P_{t,\alpha}^{\mathrm{TSS-PFC}}}{\sqrt{3} U_t^{\mathrm{PFC}} N_1} - \frac{P_t^{\mathrm{TT}}}{\sqrt{3} U_t^{\mathrm{TT}} N_2} = \kappa_{t,\alpha}^{\mathrm{TSS-PFC}} - \kappa_t^{\mathrm{TT}} \quad (18b)$$

$$0 \leqslant \kappa_{t,\alpha}^{\mathrm{TSS-PFC}} \leqslant v_t \mathcal{M},\ 0 \leqslant \kappa_{t,\alpha}^{\mathrm{TT}} \leqslant (1 - v_t)\mathcal{M} \quad (18c)$$

where $VUF^*$ represents the VU mitigation limitation, which can be generated by dynamic allocation algorithm of VU regulation in [11]. $S_{\mathrm{sc}}$ represents short-circuit capacity of grid.

*3) ES Constraints:* When the ES is connected to the NZ-PFC or TSS-PFC, the limits of discharging and charging power are set up from (19a) to (19d), which means that both the rated power and the available energy variation at time intervals can decide the upper bound of power. The charging state of the ES is restricted by applying constraints (19e) to (19f), which demonstrate the fact that charging status cannot coexist with the discharging state simultaneously.

$$0 \leqslant P_{t,\mathrm{ch}}^{\mathrm{bt}} \leqslant \min\left[\left(\mathrm{S}_{\mathrm{bt}}^{\max} E_{\mathrm{rate}}^{\mathrm{bt}} - E_t^{\mathrm{bt}}\right) \Big/ \eta_{\mathrm{ch}}^{\mathrm{bt}} \Delta t, P_{\mathrm{rate}}^{\mathrm{bt}}\right] \quad (19a)$$

$$0 \leqslant P_{t,\mathrm{dis}}^{\mathrm{bt}} \leqslant \min\left[\left(E_t^{\mathrm{bt}} - \mathrm{S}_{\mathrm{bt}}^{\max} E_{\mathrm{rate}}^{\mathrm{bt}}\right) \Big/ \eta_{\mathrm{ch}}^{\mathrm{bt}} \Delta t, P_{\mathrm{rate}}^{\mathrm{bt}}\right] \quad (19b)$$

$$0 \leqslant P_{t,\mathrm{ch}}^{\mathrm{uc}} \leqslant \min\left[\left(\mathrm{S}_{\mathrm{uc}}^{\max} E_{\mathrm{rate}}^{\mathrm{uc}} - E_t^{\mathrm{uc}}\right) \Big/ \eta_{\mathrm{ch}}^{\mathrm{uc}} \Delta t, P_{\mathrm{rate}}^{\mathrm{uc}}\right] \quad (19c)$$

$$0 \leqslant P_{t,\mathrm{dis}}^{\mathrm{uc}} \leqslant \min\left[\left(E_t^{\mathrm{uc}} - \mathrm{S}_{\mathrm{uc}}^{\max} E_{\mathrm{rate}}^{\mathrm{uc}}\right) \Big/ \eta_{\mathrm{ch}}^{\mathrm{uc}} \Delta t, P_{\mathrm{rate}}^{\mathrm{uc}}\right] \quad (19d)$$

$$P_{t,\mathrm{ch}}^{\mathrm{bt}} \leqslant \kappa_{t,\mathrm{ch}}^{\mathrm{bt}} P_{\mathrm{rate}}^{\mathrm{bt}},\ P_{t,\mathrm{dis}}^{\mathrm{bt}} \leqslant \kappa_{t,\mathrm{dis}}^{\mathrm{bt}} P_{\mathrm{rate}}^{\mathrm{bt}},\ \kappa_{t,\mathrm{ch}}^{\mathrm{bt}} + \kappa_{t,\mathrm{dis}}^{\mathrm{bt}} \leqslant 1 \quad (19e)$$

$$P_{t,\mathrm{ch}}^{\mathrm{uc}} \leqslant \kappa_{t,\mathrm{ch}}^{\mathrm{uc}} P_{\mathrm{rate}}^{\mathrm{uc}},\ P_{t,\mathrm{dis}}^{\mathrm{uc}} \leqslant \kappa_{t,\mathrm{dis}}^{\mathrm{uc}} P_{\mathrm{rate}}^{\mathrm{uc}},\ \kappa_{t,\mathrm{ch}}^{\mathrm{uc}} + \kappa_{t,\mathrm{dis}}^{\mathrm{uc}} \leqslant 1 \quad (19f)$$

where $E_{\mathrm{rate}}^{\mathrm{bt}}$ and $P_{\mathrm{rate}}^{\mathrm{bt}}$ are rated capacity and power of batteries. $E_{\mathrm{rate}}^{\mathrm{uc}}$ and $P_{\mathrm{rate}}^{\mathrm{uc}}$ are rated capacity and power of UCs. $\mathrm{S}_{\mathrm{bt}}^{\max}$ and $\mathrm{S}_{\mathrm{bt}}^{\min}$ are the minimum and maximum state of charge of batteries. $\mathrm{S}_{\mathrm{uc}}^{\max}$ and $\mathrm{S}_{\mathrm{uc}}^{\min}$ are the minimum and maximum state of charge of UCs. $\kappa_{t,\mathrm{ch}}^{\mathrm{bt}}$, $\kappa_{t,\mathrm{dis}}^{\mathrm{bt}}$ , $\kappa_{t,\mathrm{ch}}^{\mathrm{uc}}$ and $\kappa_{t,\mathrm{dis}}^{\mathrm{uc}}$ are binary variables.

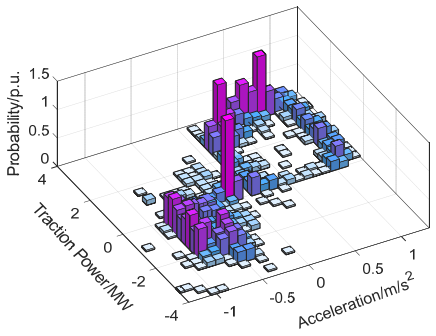



Fig. 5. Scatter Plots of Joint Distributions for Traction Power and Acceleration.

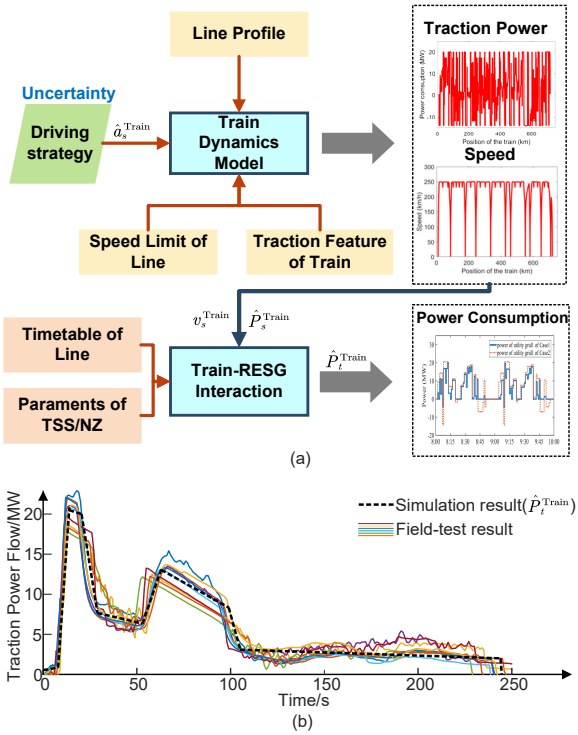



Fig. 6. Train-RESG Interactions Simulation. (a) Train-RESG Interactions Flowchart, (b) Comparison with power flow between prediction and field-test

### C. CC Compact Formulation

*1) General Forms:* The optimal problem with general chance-constraints uses probabilistic constraints and aims to minimize the mean-value of the total daily operating cost of RESG in railways. This problem in a RESG system is modeled in the following manner:

$$\begin{aligned} \min_{x} \quad & \mathbb{E}\left[\mathcal{J}_2\left(x\right)\right] \\ \text{s.t.} \quad & (16)-(18) \\ & \mathbb{P}\left\{A_i\left(x\right)\xi \leqslant b_i\left(x\right)\right\} \geqslant 1-\epsilon_i \ \forall i \end{aligned} \tag{20}$$

$$\begin{aligned} & A_i\left(x\right)\xi \leqslant b_i\left(x\right) = \\ & \left\{\overline{P_t^{\text{Train}}} \geqslant P_t^{\text{TT}} + P_{t,\alpha}^{\text{NZ-PFC}} + P_{t,\beta}^{\text{TSS-PFC}}, \right. \\ & \left. \underline{P_t^{\text{Train}}} \leqslant P_t^{\text{TT}} + P_{t,\alpha}^{\text{NZ-PFC}} + P_{t,\beta}^{\text{TSS-PFC}} \right\} \end{aligned} \tag{21}$$

The letter $x$ represents the decision variable vector (i.e., power output of ES, PFC), and $\xi$ represents the uncertainty variable vector following the distribution $\mathbb{P}$ (i.e., power forecast of train $P_t^{\text{Train}}$). $A_i\left(x\right)$ and $b_i$ are affine functions about the decision variables. The violation probability of the chance constraint is less than $\epsilon_i$, and $i$ represents the number of chance constraints, $i = 1, 2, 3, 4$.

*2) Tractable Forms of CC:* With careful derivations, the chance constraint of (21) is converted into an equivalent form:

$$\begin{aligned} & \mathcal{F}_{A_i(x)}\left[b_i\left(x\right)\right] \geqslant 1-\epsilon_i \\ \Rightarrow \quad & b_i\left(x\right) \geqslant \mathcal{F}_{A_i(x)}^{-1}\left(1-\epsilon_i\right) \\ \Rightarrow \quad & b_i\left(x\right) \leqslant \mathcal{F}_{A_i(x)}^{-1}\left(\epsilon_i\right) \ \ \forall i \end{aligned} \tag{22}$$

where $\mathcal{F}$ is the cumulative probability function of $A_i\left(x\right)$.

Note that $\mathcal{F}_{A_i(x)}^{-1}\left(\epsilon_i\right)$ is a nonlinear function of $\epsilon_i$ , which cannot be directly computed by commercial solvers. For example, the Authors of [17] adopt the non-parametric KDE to capture the uncertainty of traction load, it is hard to give an analytical formulation for function $\mathcal{F}^{-1}\left(\cdot\right)$. In order to replace the chance constraints (20) with the deterministic linear constraints (22), this paper uses as a new decision variable $\mathcal{G}_i$, and obeys the multivariate GMM, to substitute for $\epsilon_i$ [26], [28].

$$\mathcal{G}_i = \mathcal{F}_{A_i(x)}^{-1}\left(1-\epsilon_i\right) \tag{23}$$

Therefore, the chance constraint is converted into a linear inequality:

$$\mathcal{G}_i - b_i\left(x\right) \geqslant 0 \tag{24}$$

## V. Traction Power Forecast and VU Confidence Prediction

### A. Uncertainty Characterization of Train-RESG Interactions

*1) Uncertainty of Train Dynamics Modelling:* For simplicity and from a power engineering perspective, the train can be seen as a single-particle dynamic system [33], the kinematic equations of the train can be described as (25), the acceleration can be derived from force analysis with given mass. Moreover, under the given speed $v_s^{\text{Train}}$ and position $s$ at line, the estimated traction power of the train ($\hat{P}_s^{\text{Train}}$) is determined by its traction/braking force.

$$\hat{P}_s^{\text{Train}} = \hat{F}_{\text{tot}} \cdot v = \hat{a}_s^{\text{Train}} \cdot \mathrm{M}_{\text{tr}} \cdot v_s^{\text{Train}} \tag{25}$$

where $\mathrm{M}_{\text{tr}}$ is the effective mass of the train, $\hat{a}_s^{\text{Train}}$ is the estimated acceleration of train at position $s$, $\hat{F}_{\text{tot}}$ is the total traction/braking force determined by line profile, speed limit and driving strategy.

Under the same railway line conditions, owing to the driver's different driving strategy, the train's acceleration ($a_t^{\text{Train}}$) is different, and the power demand is also different, which reflects the uncertain characterization of traction load. Based on the dates of the Telematics Control Unit (TCU) of train from field-test, Fig. 5 shows the scatter plots of joint distributions of traction power and acceleration. It can be seen that when the train is accelerating, the greater the acceleration, the greater the demand for traction power; and vice versa. Meanwhile, the relationship between traction power and train's acceleration is non-linear and non-unique. Consequently, the fluctuation of traction power is the basic reason that arouses the uncertainty of power flow.

**Algorithm 1** Confidence Prediction of Traction Power Flow

1: Initialization: obtain the timetable, line parameters, train type and TSS parameters;
2: Compute the traction power flow $\hat{P}_t^{\mathrm{Train}}$ by commercial software (e.g. OpenPowerNet);
3: Harvest historical dataset $\boldsymbol{\mathcal{X}} = \left\{x_{[i]}, \hat{x}_{[i]}\right\}_{i=1}^{d}$ for variable in $\boldsymbol{\Gamma}$ in a day $t$;
4: Construct the joint CDF multi-GMM:

$$\mathcal{F}_{\boldsymbol{\mathcal{X}}}\left(\boldsymbol{x}\,|\mathrm{N_G};\boldsymbol{\Gamma}\right) = \int_{-\infty}^{\boldsymbol{x}} \sum_{n=1}^{\mathrm{N_G}} \omega_n \mathcal{N}_n\left(t, \hat{t}\,|\mu_n, \sigma_n\right) dt$$

5: Import $\hat{P}_t^{\mathrm{Train}}$ into $\mathcal{F}_{\boldsymbol{\mathcal{X}}}^{-1}\left(\boldsymbol{x}\,|\mathrm{N_G};\boldsymbol{\Gamma}\right)$:

$$\boldsymbol{x} = \mathcal{F}_{\boldsymbol{\mathcal{X}}}^{-1}\left(r\right)$$

6: Obtain the confidence forecast result $\left[\underline{P_t^{\mathrm{Train}}}, \overline{P_t^{\mathrm{Train}}}\right]$.

$$\left[\underline{P_t^{\mathrm{Train}}}, \overline{P_t^{\mathrm{Train}}}\right] = \hat{P}_t^{\mathrm{Train}} + \left[\hat{x}_t^{\theta_{\mathrm{L}}}, \hat{x}_t^{\theta_{\mathrm{U}}}\right]$$

**Algorithm 2** Confidence Level $\theta$ and Violation Probability $\epsilon$ Selecting for CC-P Solving

1: Harvest the confidence prediction result of traction power flow $\left[\underline{P_t^{\mathrm{Train}}}, \overline{P_t^{\mathrm{Train}}}\right]$ by Algorithm 1;
2: Evaluate the VU distribution at grid for each TSS by equation (4), and count the 95% probability and maximum value set with different confidence level ($\boldsymbol{VUF}_{95\%,\theta}^{\mathrm{TSS}-k}$, $\boldsymbol{VUF}_{\max,\theta}^{\mathrm{TSS}-k}$ $\forall k = 1, 2$);
3: **if** $VUF_{95\%}^{\mathrm{TSS}-k} > \mathrm{E}_{95\%}^{k}$ & $VUF_{\max}^{\mathrm{TSS}-k} > \mathrm{E}_{100\%}^{k}$ **then**
4: $\epsilon = 1 - \max\left\{\mathcal{F}_{95\%}^{\boldsymbol{VUF}}\left(\mathrm{E}_{95\%}^{k}\right), \mathcal{F}_{100\%}^{\boldsymbol{VUF}}\left(\mathrm{E}_{100\%}^{k}\right)\right\}$;
5: **else if** $VUF_{95\%}^{\mathrm{TSS}-k} \leqslant \mathrm{E}_{95\%}^{k}$ & $VUF_{\max}^{\mathrm{TSS}-k} > \mathrm{E}_{100\%}^{k}$ **then**
6: $\epsilon = 1 - \mathcal{F}_{100\%}^{\boldsymbol{VUF}}\left(\mathrm{E}_{100\%}^{k}\right)$;
7: **else if** $VUF_{95\%}^{\mathrm{TSS}-k} > \mathrm{E}_{95\%}^{k}$ & $VUF_{\max}^{\mathrm{TSS}-k} \leqslant \mathrm{E}_{100\%}^{k}$ **then**
8: $\epsilon = 1 - \mathcal{F}_{95\%}^{\boldsymbol{VUF}}\left(\mathrm{E}_{95\%}^{k}\right)$;
9: **else**
10: $\epsilon = 0.999$ and VU compensation constraint is failure;
11: **end if**

*2) Uncertainty of Train-RESG interaction:* Fig. 6(a) shows the deterministic forecast flowchart of the train-RESG. Firstly, the line data, train traction features and speed limits are imported into the train's power simulator, which allows for obtaining the entire speed ($v_s^{\mathrm{Train}}$) and power distribution ($\hat{P}_s^{\mathrm{Train}}$) of a train in each position of line. Secondly, the power dates of a single train, timetable and TSS parameters are input to RESG's power flow simulator [33], and the total power consumption with time ($\hat{P}_t^{\mathrm{Train}}$) and VU distribution of each TSS can be dynamically obtained. In this paper, initial power flow scenarios of two TSSs are generated, assuming that the RESG infrastructures are not operational, which implies that the VU performance of TSS has arrived at the worst level.

Fig. 6 (b) shows the discrepancy between forecasted and measured with time, which is the proxy for the uncertainty. It can be seen that traditional deterministic forecasting only provides single-valued estimation, which misses the significant stochastic distribution information. As result, it is still challenging to model the stochastic dependence analytically. The author of [17] proposes an uncertainty modelling approach based on the nonparametric KDE. Nevertheless, it is hard to obtain an inverse function of cumulative probability density in chance-constrain programming.

### B. Traction Load Uncertainty Modeling

*1) Multivariate GMM Description:* To characterize the uncertainty of traction load features, the joint GMM distribution is used in this paper to accurately model the multimodal probability distributions of traction power forecast. A GMM for a $d$-dimensional random vector $\boldsymbol{\mathcal{X}} = \left\{x_{[i]}, \hat{x}_{[i]}\right\}_{i=1}^{d}$, as shown in (26)-(28), is defined as a convex combination of multi-dimensional Gaussian distribution function with an model parameters set $\boldsymbol{\Gamma} = \omega_n, \mu_n, \sigma_n\,|n = 1, 2, \cdots, \mathrm{N_G}$ [21].

$$f_{\boldsymbol{\mathcal{X}}}\left(\boldsymbol{x}\,|\mathrm{N_G};\boldsymbol{\Gamma}\right) = \sum_{n=1}^{\mathrm{N_G}} \omega_n \mathcal{N}_n\left(x, \hat{x}\,|\mu_n, \sigma_n\right) \tag{26}$$

$$\sum_{n=1}^{\mathrm{N_G}} \omega_n = 1, \omega_n > 0 \tag{27}$$

$$\mathcal{N}_n\left(\boldsymbol{x}\,|\mu_n, \sigma_n\right) = \frac{e^{-\frac{1}{2}\left(\boldsymbol{x}-\mu_n\right)^{\mathrm{T}}\sigma_n^{-1}\left(\boldsymbol{x}-\mu_n\right)}}{\left(2\pi\right)^{d/2} \cdot \det\left(\sigma_n\right)^{1/2}} \tag{28}$$

where $x$ and $\hat{x}$ are the actual power and forecasted power, respectively. $\mathrm{N_G}$ is the number of Gaussian components, $\omega_n$ is the weighting coefficients, $\{\mu_n, \sigma_n\}$ denotes the set of parameters in the $n$th mixture component $\mathcal{N}_n\left(\boldsymbol{x}\,|\mu_n, \sigma_n\right)$. $\mu_n \in \mathbb{R}^{d\times 1}$ and $\sigma_n \in \mathbb{R}^{d\times d}$ are the expectation vector and covariance matrix of Gaussian component, respectively. Standard guidelines for the multi-GMM and the parameter estimation are available in [21].

*2) Analytical CDF of the GMM:* The cumulative distribution $\mathcal{F}_{\boldsymbol{\mathcal{X}}}$ is another essential function to model traction load uncertainty, which is analytically expressed as:

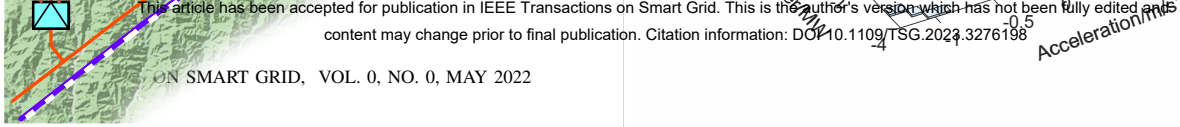
$$\mathcal{F}_{\boldsymbol{\mathcal{X}}}\left(\boldsymbol{x}\,|\mathrm{N_G};\boldsymbol{\Gamma}\right) = \int_{-\infty}^{\boldsymbol{x}} \sum_{n=1}^{\mathrm{N_G}} \omega_n \mathcal{N}_n\left(t, \hat{t}\,|\mu_n, \sigma_n\right) dt \tag{29}$$

To obtain the forecasted power, it can further adopt the inverse transform method by calculating $\boldsymbol{x} = \mathcal{F}_{\boldsymbol{\mathcal{X}}}^{-1}\left(r\right)$ as the sample from $f_{\boldsymbol{\mathcal{X}}}$, where the scalar $r \sim \mathrm{U}\left(0, 1\right)$.

Based on the multi-GMM model, the confidence intervals (CIs) of train load uncertainties $\left[\hat{x}_t^{\theta_{\mathrm{L}}}, \hat{x}_t^{\theta_{\mathrm{U}}}\right]$ can be calculated by using the joint CDF in (29). A CI ($CI_t^{\theta}$) of the forecasted train power with a confidence level $\theta$ can be expressed by:

$$CI_t^{\theta} = \left[\underline{P_t^{\mathrm{Train}}}, \overline{P_t^{\mathrm{Train}}}\right] = \hat{P}_t^{\mathrm{Train}} + \underbrace{\left[\hat{x}_t^{\theta_{\mathrm{L}}}, \hat{x}_t^{\theta_{\mathrm{U}}}\right]}_{\mathrm{UNCERTAINTY}} \tag{30}$$

where the lower and upper confidence coefficient $\theta_{\mathrm{L}}$ and $\theta_{\mathrm{U}}$ equal to $\theta/2$ and $(1 - \theta/2)$, respectively.



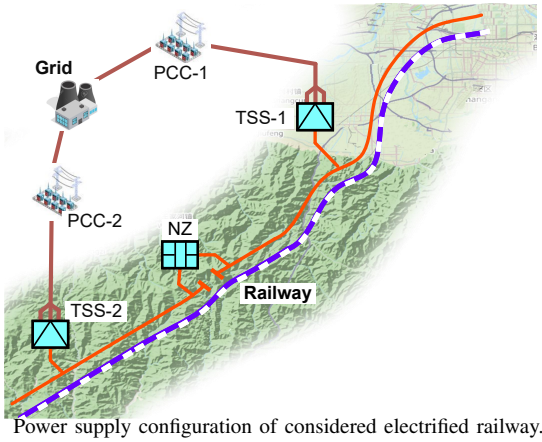




Fig. 7. Power supply configuration of considered electrified railway.

TABLE II
PARAMETERS OF RESG

| Power Grid | | Transformer | | | PFC | |
|---|---|---|---|---|---|---|
| $U_{abc}^{\text{grid}}$ (kV) | $S_{\text{sc}}$ (MVA) | $N_1$ | $N_2$ | $U^{\text{TT}}$ (kV) | $U_{\alpha}^{\text{PFC}}$ (kV) | $U_{\beta}^{\text{PFC}}$ (kV) |
| 220 | 1500 | 4 | $\frac{4}{\sqrt{3}}$ | 27.5 | 27.5 | 27.5 |

TABLE III
PARAMETERS OF ES

| Parameters | Unit | Battery | UC |
|---|---|---|---|
| $c_{\text{P}}$ | kCNY/kW | 2.83 | 2.05 |
| $c_{\text{E}}$ | kCNY/kWh | 4.64 | 19.80 |
| $c_{\text{rep}}$ | kCNY/kWh | 1.29 | 0 |
| $c_{\text{om,f}}$ | CNY/kW/year | 25.5 | 0 |
| $c_{\text{om,v}}$ | CNY/MW/h | 2.78 | 0 |
| $\eta_{\text{dis}}/\eta_{\text{ch}}$ | - | 80%/80% | 95%/95% |
| Initial SoC | - | 50% | 50% |
| SoC range | - | 20-80% | 5-95% |

TABLE IV
TIME-OF-USE TARIFF

| Parameters | Peak | Mid | Valley |
|---|---|---|---|
| Energy (CNY/kWh) | 1.252 | 0.782 | 0.370 |
| Demand (CNY/kW/mon) | 42 | 42 | 42 |
| Time period | 8-11h, 18-21h | 7h, 12-17h | 0-6h, 22-0h |

### C. Standard Processes

*1) Traction Power Forecast:* The confidence prediction approach for the traction power flow is proposed in Algorithm 1. This paper is carried out in the framework of IEC 61000-3-13 and general train-RESG interaction simulation [17], [30], and therefore, the proposed approach can be extended to any railway traction system connected to a power system.

*2) Confidence Level Selecting:* The confidence level and violation probability selecting approaches for solving CC-P are proposed in Algorithm 2. The 95% probability and maximum valve sets of VU distribution are defined as $\boldsymbol{VUF}_{95\%,\theta}^{\text{TSS}-k} = \left\{VUF_{95\%,\theta}^{\text{TSS}-k}, \theta \in [0,1], k = 1,2\right\}$ and $\boldsymbol{VUF}_{100\%,\theta}^{\text{TSS}-k} = \left\{VUF_{100\%,\theta}^{\text{TSS}-k}, \theta \in [0,1], k = 1,2\right\}$, respectively. As a result, the violation probability $\epsilon$ that VUF exceeds the limit under the confidence level $\theta$ is obtained.

Generally, choosing a feasible violation probability and confidence level is an important step in the design process of railways. Higher violation probability means higher compensation cost, and lower confidence level means higher uncertainty of forecasted train power. Single-phase traction load is a significant feature distinguished from the normal power grid load. Consequently, taking the VU limit as a selection standard is one of the available approaches to obtain a compromise solution.

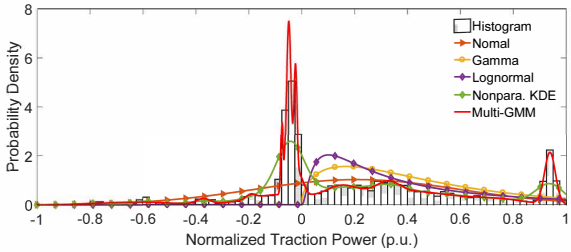



Fig. 8. Multi-GMM and other frequently used distributions for modelling the traction power distributions.

## VI. CASE STUDY

### A. Test System and Data Source

In order to validate the feasibility of the proposed CC-P of RESGs, a busy electrified railway is considered for the case study, as shown in Fig. 7. The detailed network parameters can be found in [12], while the basic parameters of RESG, ES and time of use are listed in Table II-IV. The unit capacity of the battery and UC is 80 kWh and 0.15 kWh, and the unit power of the battery, UC and PFC is 10 kW, 5 kW, and 100 kW, respectively [34].

All the case studies are solved by the MATLAB 2016b interface on a workstation with Intel Xeon Silver 4114 CPU and 64 GB RAM. The first-stage is built as an NLP problem and addressed by the MATLAB Global Optimization Toolbox, the second-stage is formulated as a MILP problem and solved using IBM ILOG CPLEX solver via YALMIP. The entire two-stage optimization framework is shown in Fig. 4.

### B. Result of Modeling Uncertainty

Fig. 8 compares the probability distribution of traction load by using five distributions including multi-GMM, parametric KDE (Gaussian, Gamma, Log-normal) and nonparametric KDE. It can be observed that the multi-GMM accurately match well with empirical distributions than the other four regular distribution. This paper adopts the correlation coefficient and Kolmogorov-Smirnov test integral (KSIPer) to quantify the correlation and statistical similarity between the original CDF and the CDF of fitting distribution, and employs mean absolute error (MAE) to evaluate the uniform fitting errors [21], which shows in Table V. It can be seen that multi-GMM method has the highest correlation and lowest KSIPer and MAE, indicating that the multi-GMM can better represent traction power distribution than the other four approaches.

Fig. 9 (a) and (b) show the lower and upper bounds, with $\theta$ is 20%, of predicted traction power flow for TSS-1 and TSS-2. In order to illustrate the forecast result under the various CIs ($\theta$),the power flow from 12:00 to 13:00 is drawn in Fig. 9 (c) and (d). The is purple curve is measured power flow in real world. It can be seen that the confidence forecasts are not symmetric, and that the measured curve can be completely covered by the CI.

TABLE V
METRICS OF FIVE DISTRIBUTIONS ESTIMATED FOR TRACTION LOAD

| Distributions | Metrics | | |
|---|---|---|---|
| | Correlation coeff. | KSIPer | MAE |
| Gaussian | 0.9598 | 0.3386 | 0.1183 |
| Gramma | 0.9060 | 0.5016 | 0.2178 |
| Log-normal | 0.9117 | 0.5056 | 0.2162 |
| Nonparametric KDE | 0.9626 | 0.3265 | 0.1116 |
| Multi-GMM | 0.9711 | 0.2551 | 0.1105 |

TABLE VI
METRICS OF TWO PREDICTION APPROACH FOR TRACTION POWER

| Prediction Method | $ACE\%$ | $ASV\%$ |
|---|---|---|
| Method in [17] | 5.289 | 28.92 |
| Proposed method | 4.577 | 25.59 |

TABLE VII
OPTIMIZED CAPACITY FOR RESGS

| | TSS1 | TSS2 | NZ |
|---|---|---|---|
| $P^{\mathrm{PFC}}_{\mathrm{rate}}$/kW | 3000 | 3000 | 3300 |
| $P^{\mathrm{bt}}_{\mathrm{rate}}$/kW | 130 | 130 | 130 |
| $E^{\mathrm{bt}}_{\mathrm{rate}}$/kWh | 1040 | 1040 | 1040 |
| $P^{\mathrm{uc}}_{\mathrm{rate}}$/kW | 4150 | 1750 | 1000 |
| $E^{\mathrm{uc}}_{\mathrm{rate}}$/kWh | 12 | 5 | 3 |

TABLE VIII
THE TOTAL DAILY INVESTMENT AND OPERATION COST OF RESGS

| | Without Optimal | With Optimal | | | |
|---|---|---|---|---|---|
| | TSS1&TSS2 | TSS1 | TSS2 | NZ | Total |
| $C^{\mathrm{ES}}_{\mathrm{cap}}$/kCNY | – | 3.07 | 2.34 | 1.55 | 6.96 |
| $C^{\mathrm{PFC}}_{\mathrm{cap}}$/kCNY | – | 0.28 | 0.28 | 0.30 | 0.86 |
| $C^{\mathrm{ES}}_{\mathrm{rep}}$/kCNY | – | 1.24 | 1.24 | 1.24 | 3.71 |
| $C^{\mathrm{ES}}_{\mathrm{om}}$/kCNY | – | 0.01 | 0.01 | 0.01 | 0.03 |
| $C^{\mathrm{PFC}}_{\mathrm{om}}$/kCNY | – | 0.42 | 0.18 | 0.10 | 0.70 |
| $\mathbb{E}[\mathcal{J}_2]$/kCNY | 60.78 | 21.61 | 18.81 | 0.00 | 40.42 |
| $C_t ot$/kCNY | 60.78 | 26.62 | 22.86 | 3.20 | 52.68 |
| Total Cost Savings | – | 13.32% | | | |

To evaluate the accuracy of the proposed model, $ACE$ and $ASV$ are defined, which can evaluate the reliability and sharpness. Reliability indicates the degree of correctness of a forecast model assessed by the hit percentage, while sharpness indicates the uncertainty conveyed by the forecast model [23]. Table VI shows the metrics value for different prediction approaches, compared with the method in [17]. It indicates that the proposed model shows the smallest value, which presents the optimal tradeoff between reliability and sharpness compared to other correlated conditions.Therefore, the smallest $ACE$ and $ASV$ values verify the effectiveness of the proposed prediction model in Section V.

### C. Result of VU Confidence Prediction

Fig. 10 (a)-(b) present the $VUF^{\mathrm{TSS-1}}_{95\%,\theta}$ and $VUF^{\mathrm{TSS-1}}_{100\%,\theta}$ distribution of TSS-1, respectively. Based on the IEC/TR 61000-3-13, 95% probability limitation $\mathrm{E}_{95\%}$ and maximum limitation $\mathrm{E}_{100\%}$ are set as 2% and 4% in this paper. $VUF^{\mathrm{TSS-1}}_{95\%}$ is large than 2% when the confidence level exceeds 0.8, whereas

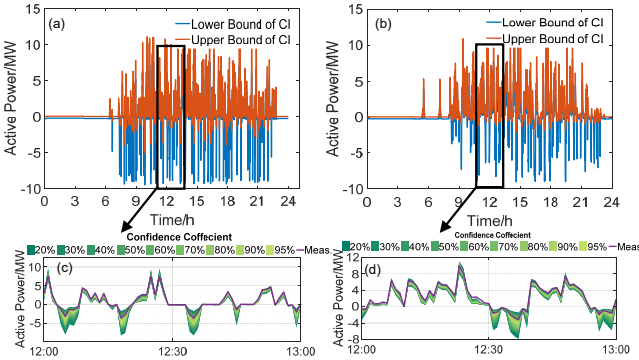



Fig. 9. Prediction result of traction power flow. (a) Lower and upper bounds of power flow for TSS-1, (b) Lower and upper bounds of power flow for TSS-2, (c) Confidence prediction result of TSS-1 with various CIs at 12:00 to 13:00, (d) Confidence prediction result of TSS-2 with various CIs at 12:00 to 13:00.

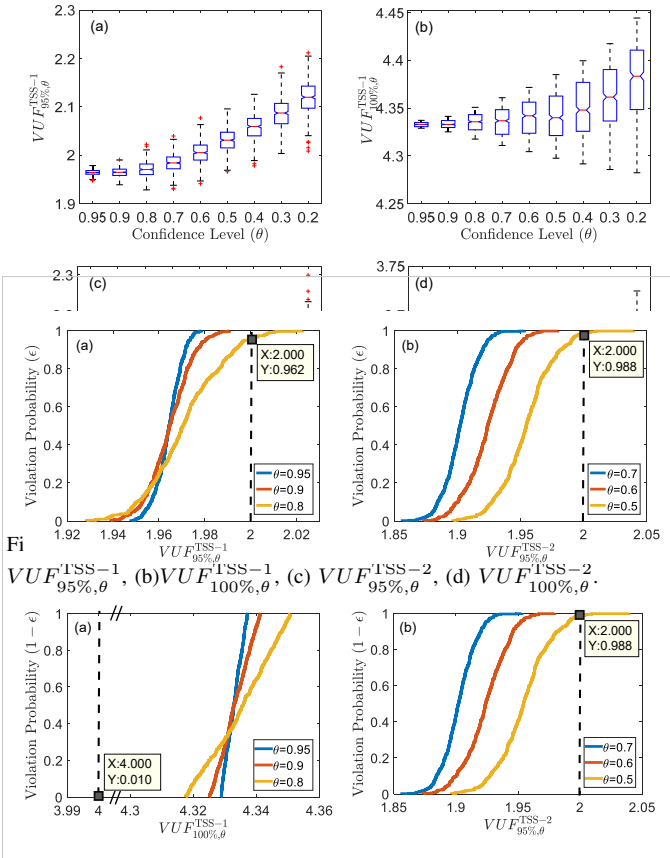



Fig. 11. Violation probability $(1-\epsilon)$ of 95% probability of VUF under different $\theta$. (a) TSS-1, (b) TSS-2.

$VUF^{\mathrm{TSS-1}}_{95\%,\theta}$, (b)$VUF^{\mathrm{TSS-1}}_{100\%,\theta}$, (c) $VUF^{\mathrm{TSS-2}}_{95\%,\theta}$, (d) $VUF^{\mathrm{TSS-2}}_{100\%,\theta}$.

$VUF^{\mathrm{TSS-1}}_{100\%}$ exceeds 4% for all levels. Fig. 10 (c)-(d) present the $VUF^{\mathrm{TSS-2}}_{95\%,\theta}$ and $VUF^{\mathrm{TSS-2}}_{100\%,\theta}$ distribution of TSS-2, respectively. $VUF^{\mathrm{TSS-2}}_{95\%}$ is large than 2% when the confidence level exceeds 0.5, while $VUF^{\mathrm{TSS-2}}_{100\%}$ doesn't exceed 4% for all levels. It can be observed that the of VUF increases as the confidence level $\theta$ increases, since a larger $\theta$ renders the available traction power flow more lenient. According to Algorithm 2, 0.95 and 0.5 are selected as the confidence level of the chance constraints of TSS-1 and TSS-2, respectively.

Fig. 11 shows the selecting result of violation probability for TSS-1 and TSS-2 under different confidence levels. In Fig. 11(a), the violation probability $1-\epsilon = 0.01$ since $VUF^{\mathrm{TSS-1}}_{100\%}$ exceeds 4%. In Fig. 11(b), the the violation probability $1-\epsilon =$

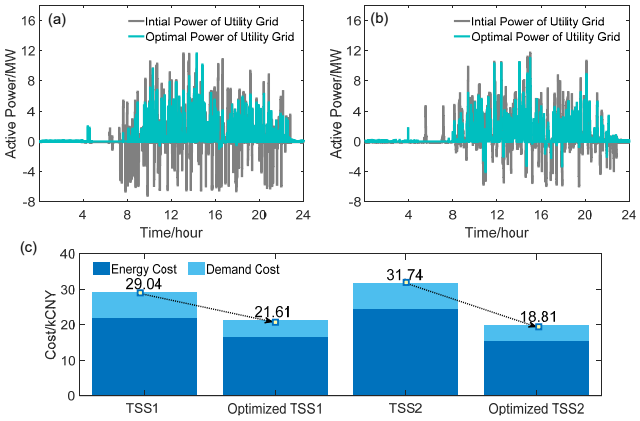



Fig. 12. Optimized result of each TSS. (a)Power flow of TSS-1, (b) Power flow of TSS-2, (c) Comparison with electricity cost and composition in different TSSs.

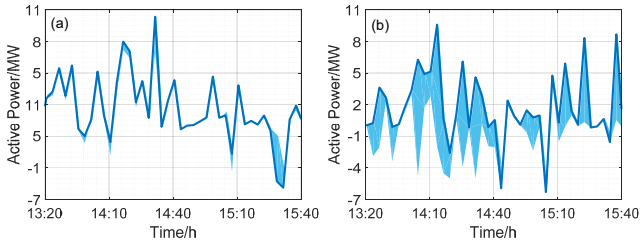



Fig. 13. Planned and actual active power of train. (a) TSS-1 with $1-\epsilon = 0.01, \theta = 0.95$, (b) TSS-2 with $1-\epsilon = 0.98, \theta = 0.5$.

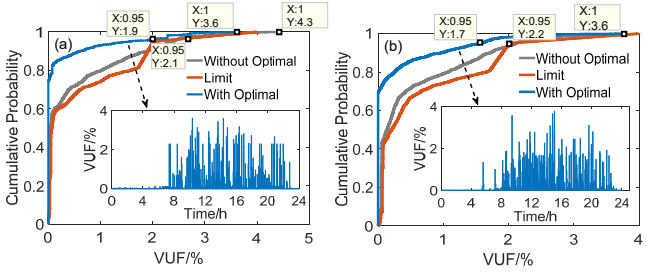



Fig. 14. Optimized result of VUF. (a) TSS-1, (b) TSS-2.

0.98 when the $VUF_{95\%}^{\mathrm{TSS}-2}$ exceeds 2%.

### *D. Result of ES and PFC Planning*

By applying the proposed two-stage model combining ES sizing and daily scheduling, the optimal capacity of the RESGs' facilities is presented in Table VII, the optimal cost is shown in Table VIII. It can be seen that the daily total cost is decreased by 13.32% (from 60.78 kCNY to 52.68 kCNY).

Fig.12 (a) and (b) show the optimized grid power of traction substation compared to the initial grid power. It can be seen that RBE is mostly utilized by the ES and the train.

Fig.12 (c) shows the components of electricity cost, the installation of PFC and ES resulted in a significant decrease in energy and demand costs, with a 25% decrease for TSS-1 (from 29.04 kCNY/pre-day to 21.61 kCNY/per-day), and a 40% decrease for TSS-2 (from 31.74 kCNY/pre-day to 18.81 kCNY/per-day).

### *E. Result of ES and PFC Dispatch Scheme*

Fig.13 (a) and (b) show the planned and actual active power flow of the train by using (21). The light blue area is the feasible region for train power, and the navy blue curve is the optimal solution. The above results show that the proposed chance constraint is available, and it can effectively address the uncertainty of train power.

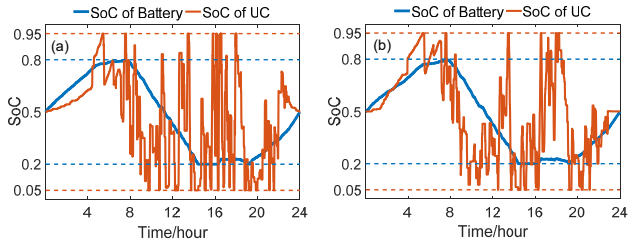



Fig. 15. SoC of ES. (a) TSS-1, (b) TSS-2.

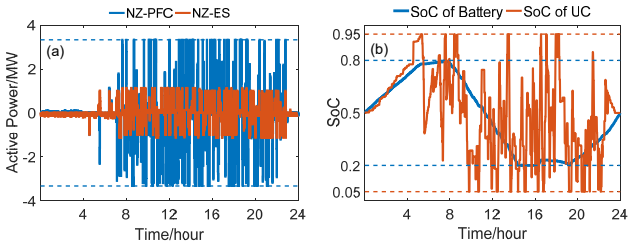



Fig. 16. Optimized result of NZ-PFC. (a) Power flow of PFC and ES, (b) SoC of ES.

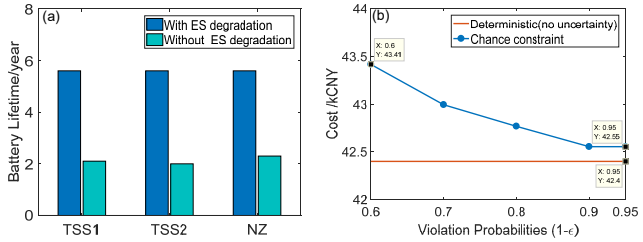



Fig. 17. Comparison result. (a) Comparison of battery lifetime, (b) Comparison of different optimizations in the objective.

Fig.14 (a) and (b) show that the maximum daily emission levels for TSS-1 and TSS-2 are 3.6% and 3.6%, respectively. Similarly, the daily emission levels of 95% probability are 1.9% and 1.7%, respectively. Consequently, the compensation results satisfy the limits of IEC/TR 61000-3-13, and the proposed dynamic VUF constraint is able to regulate the VU effectively via active ES and PFC control.

Fig.15 (a) and (b) depict the SoC for the battery and UC, respectively. The UC is frequently charged and discharged, while the battery has a relatively stable operation curve, which also confirms that the battery responses to long-term energy needs, whereas UC responds to severe power fluctuations. It is interesting that the ES store the energy slowly from 0:00 A.M to 4:00 A.M, due to the state of charging at valley time period with low price.

Fig.16 (a) presents the power of PFC and ES. The SoC of ES is also drawn in Fig.16 (b). The NZ-PFC is able to route the power from one TSS to another TSS where it is required. On the other hand, the RBE is partially absorbed by the NZ-ES, which greatly improves the utilization of RBE. Therefore, the proposed system can dispatch the power, making it a feasible approach for operating railways more cost-efficient and power quality friendly.

### F. Results Comparison

*1) Comparison With ES Ageing:* Fig. 17 (a) shows the batteries' lifetime with and without ageing cost. Considering ES degradation in the O&M cost function, the lifetime thus evaluated in years at three positions is increased from 2.0 years to 5.6 years. As a result, the principal ES lifetime was maximized by considering the storage replacement cost and the ES optimal size was determined economically.

*2) Comparison With Different Violation Probabilities:* Fig. 17 (b) shows a comparison of the objective function of energy cost for different violation probabilistic thresholds $(1-\epsilon_i)$ and the benchmark model. It can be observed that the optimal energy cost $(\mathbb{E}(\mathcal{J}_2))$ decreases as $1-\epsilon_i$ increases from 0.6 to 0.95, since larger $1-\epsilon_i$ renders the chance constraint less lenient. Also, it can be seen that the energy cost under higher violation probabilistic is closer to the deterministic optimization model (the actual power flow is perfectly known).

## VII. Conclusion

This paper proposes a chance-constrained two-stage programming approach to determine the capacity of PFC-ES for RESG, considering the uncertain predictions of the traction power flow. At the first-stage, the size of the RESG infrastructure is optimized by minimizing the investment and operation costs of RESG's devices, including the capital costs, replacement cost and O&M costs. Compared to the initial transformer-based power supply structure, the total cost of RESG can be decreased by 13% over the service cycle of PFC-ES. At the second-stage, the power flow shifting plan of PFC and charging/discharging power control actions of the ES are optimally scheduled with the aim of minimizing the expected electricity cost. Compared with the conventional operation schemes, the energy costs of the two TSSs are significantly reduced by 25% and 40%, respectively. Meanwhile, with the dynamic VU regulation, the 95% probability and the maximum value of both TSSs are less than 2% and 4%, respectively. Further, in order to capture the statistical properties of the traction power flow, probabilistic traction power forecasting incorporating a multi-GMM is proposed. Compared with the traditional uncertain modeling approaches, the model provides significant improvements in reliability and sharpness features.

## Acknowledgments

The author would like to extend their gratitude to the China Scholarship Council for providing support. Additionally, special thanks are due to China Railway First Survey and Design Institute Group Co., Ltd. for generously providing the necessary data for testing the proposed approach.

## References

[1] International Energy Agency, "The Future of Rail Opportunities for Energy and the Environment: 2019," Pairs, French, January, 2019. [Online]. Available: https://www.iea.org/reports/the-future-of-rail

[2] S. M. Mousavi Gazafrudi, A. Tabakhpour Langerudy, E. F. Fuchs and K. Al-Haddad, "Power Quality Issues in Railway Electrification: A Comprehensive Perspective," *IEEE Trans. Ind. Electron.*, vol. 62, no. 5, pp. 3081-3090, May 2015.

[3] J. Chen, H. Hu, Y. Ge, K. Wang, W. Huang and Z. He, "An Energy Storage System for Recycling Regenerative Braking Energy in High-Speed Railway," *IEEE Trans. Power Del.*, vol. 36, no. 1, pp. 320-330, Feb. 2021.

[4] E. Pilo de la Fuente, S. K. Mazumder and I. G. Franco, "Railway Electrical Smart Grids: An Introduction to Next-Generation Railway Power Systems and Their Operation.," *IEEE Electrification Magazine*, vol. 2, no. 3, pp. 49-55, Sept. 2014.

[5] E. Pilo, S. K. Mazumder and I. González-Franco, "Smart Electrical Infrastructure for AC-Fed Railways With Neutral Zones," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 642-652, April 2015.

[6] J. A. Aguado, A. J. Sánchez Racero and S. de la Torre, "Optimal Operation of Electric Railways With Renewable Energy and Electric Storage Systems," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 993-1001, March 2018.

[7] H. Novak, V. Lešić and M. Vašak, "Hierarchical Model Predictive Control for Coordinated Electric Railway Traction System Energy Management," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 7, pp. 2715-2727, July 2019.

[8] S. Khayyam, N. Berr, L. Razik, M. Fleck, F. Ponci and A. Monti, "Railway System Energy Management Optimization Demonstrated at Offline and Online Case Studies," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 11, pp. 3570-3583, Nov. 2018.

[9] Çiçek et al., "Integrated Rail System and EV Parking Lot Operation With Regenerative Braking Energy, Energy Storage System and PV Availability," *IEEE Trans. Smart Grid*, vol. 13, no. 4, pp. 3049-3058, July 2022.

[10] F. Calvillo, Á. Sánchez-Miralles and J. Villar, "Synergies of Electric Urban Transport Systems and Distributed Energy Resources in Smart Cities," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 8, pp. 2445-2453, Aug. 2018.

[11] Y. Chen, M. Chen, Z. Liang and L. Liu, "Dynamic Voltage Unbalance Constrained Economic Dispatch for Electrified Railways Integrated Energy Storage," *IEEE Trans. Ind. Informat.*, vol. 18, no. 11, pp. 8225-8235, Nov. 2022.

[12] Y. Liu, M. Chen, Z. Cheng, Y. Chen and Q. Li, "Robust Energy Management of High-Speed Railway Co-Phase Traction Substation With Uncertain PV Generation and Traction Load," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 6, pp. 5079-5091, June 2022.

[13] J. Lin, S. Hu, Y. Li, L. Luo, J. Zhang, Y. Cao, B. Gao, J. Yu and F. Zhou, "A Novel Power Programming Strategy for Railway Power Regulation With Dynamic Exploration," *IEEE Trans. Smart Grid*, vol. 13, no. 4, pp. 2798-2811, July 2022.

[14] J. Chen, H. Hu, Y. Ge, K. Wang and Z. He, "Techno-Economic Model-Based Capacity Design Approach for Railway Power Conditioner-Based Energy Storage System," *IEEE Trans. Ind. Electron.*, vol. 69, no. 5, pp. 4730-4741, May 2022.

[15] M. Chen, Z. Liang, Z. Cheng, J. Zhao and Z. Tian, "Optimal Scheduling of FTPSS With PV and HESS Considering the Online Degradation of Battery Capacity," *IEEE Trans. Transport. Electrific.*, vol. 8, no. 1, pp. 936-947, March 2022.

[16] S. de la Torre, A. J. Sánchez-Racero, J. A. Aguado, M. Reyes and O. Martínez, "Optimal Sizing of Energy Storage for Regenerative Braking in Electric Railway Systems," *IEEE Trans. Power Syst.*, vol. 30, no. 3, pp. 1492-1500, May 2015.

[17] Y. Chen, M. Chen, M. Cui, W. Lu and Y. Lv, "Voltage Unbalance Probability Pre-Assessment of Electrified Railways With Uncertain Traction Load," *IEEE Trans. Transport. Electrific.*, vol. 9, no. 1, pp. 1509-1520, March 2023.

[18] M. Chen, C. Roberts, P. Weston and S. Hillmansen, "Harmonic Modelling and Prediction of High-Speed Electric Train Based on Non-Parametric Confidence Interval Estimation Method," *Int. J. Elect. Power Energy Syst.*, vol. 87, pp. 176-186, May. 2017.

[19] M. Mohammadi, H. Basirat and A. Kargarian, "Nonparametric Probabilistic Load Flow With Saddle Point Approximation,"*IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4796-4804, Sept. 2018.

[20] N. Zhang, C. Kang, Q. Xia and J. Liang, "Modeling Conditional Forecast Error for Wind Power in Generation Scheduling," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1316-1324, May 2014.

[21] M. Cui, C. Feng, Z. Wang and J. Zhang, "Statistical Representation of Wind Power Ramps Using a Generalized Gaussian Mixture Model," *IEEE Trans. Sustain. Energy*, vol. 9, no. 1, pp. 261-272, Jan. 2018.

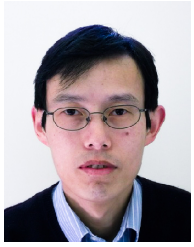
[22] W. Sun, M. Zamani, H. -T. Zhang and Y. Li, "Probabilistic Optimal Power Flow With Correlated Wind Power Uncertainty via Markov Chain Quasi-Monte-Carlo Sampling," *IEEE Trans. Ind. Informat.*, vol. 15, no. 11, pp. 6058-6069, Nov. 2019.

[23] M. Cui, V. Krishnan, B. -M. Hodge and J. Zhang, "A Copula-Based Conditional Probabilistic Forecast Model for Wind Power Ramps," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3870-3882, July 2019.
[24] E. A. Martínez Ceseña and P. Mancarella, "Energy Systems Integration in Smart Districts: Robust Optimisation of Multi-Energy Flows in Integrated Electricity, Heat and Gas Networks," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 1122-1131, Jan. 2019.
[25] Hussain, V. -H. Bui and H. -M. Kim, "Resilience-Oriented Optimal Operation of Networked Hybrid Microgrids," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 204-215, Jan. 2019.
[26] Z. Wang, C. Shen, F. Liu, J. Wang and X. Wu, "An Adjustable Chance-Constrained Approach for Flexible Ramping Capacity Allocation," *IEEE Trans. Sustain. Energy*, vol. 9, no. 4, pp. 1798-1811, Oct. 2018.
[27] N. Yahya Soltani and A. Nasiri, "Chance-Constrained Optimization of Energy Storage Capacity for Microgrids," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 2760-2770, July 2020.
[28] Z. Wang, C. Shen, F. Liu, X. Wu, C. -C. Liu and F. Gao, "Chance-Constrained Economic Dispatch With Non-Gaussian Correlated Wind Power Uncertainty," *IEEE Trans. Power Syst.*, vol. 32, no. 6, pp. 4880-4893, Nov. 2017
[29] M. Chen, M. Wang, D. Zhang, Y. Chen and W. Lu, "Improved Coordinated Control Strategy for Reliability Enhancement of Parallel PFCs With LCC Restriction," *IEEE Trans. Transport. Electrific.*, vol. 8, no. 2, pp. 2093-2105, June 2022.
[30] J. Shen, S. Dusmez and A. Khaligh, "Optimization of Sizing and Battery Cycle Life in Battery/Ultracapacitor Hybrid Energy Storage Systems for Electric Vehicle Applications," *IEEE Trans. Ind. Informat.*, vol. 10, no. 4, pp. 2112-2121, Nov. 2014.
[31] M. Khosravi , S. Afsharnia , S. Farhangi, "Optimal Sizing and Technology Selection of Hybrid Energy Storage System With Novel Dispatching Power for Wind Power integration," *Int. J. Elect. Power Energy Syst.*, vol. 127, pp. 106660, May 2021.
[32] J. Xue and B. Shen, "A Novel Swarm Intelligence Optimization Approach: Sparrow Search Algorithm," *Syst. Sci. & Control Eng.*, vol. 8, no. 1, pp. 22-34, 2020.
[33] B. Mohamed, P. Arboleya, I. El-Sayed and C. González-Morán, "High-Speed 2 × 25 kV Traction System Model and Solver for Extensive Network Simulations," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 3837-3847, Sept. 2019.
[34] X. Liu, K. Li, "Energy Storage Devices in Electrified Railway Systems: A Review," *Transp. Safety & Environ.*, vol. 2, no. 5, pp.183-201, July 2020.

**Yinyu Chen** (Student Member, IEEE) received the B.Eng. degree in electrical engineering from Kunming University of Science and Technology, Kunming, China, and the M.Eng. degrees from Southwest Jiaotong University, Chengdu, China, in 2017 and 2020, respectively. He is currently working toward the Ph.D. degree with electrical engineering from Southwest Jiaotong University. From 2023 to 2024, he also is a Visiting Ph.D. Student at University of Strathclyde under the CSC grant. His current research interests include railway traction systems and energy management.

**Minwu Chen** (Member, IEEE) received the B.Eng. and Ph.D. degrees in electrical engineering from Southwest Jiaotong University, Chengdu, China, in 2004 and 2009, respectively, all in electrical engineering.

Currently, he is a Full Professor at the School of Electrical Engineering, Southwest Jiaotong University, and is the Vice Dean of School of Electrical Engineering. From 2013 to 2018, he was an Associate Professor at the School of Electrical Engineering, Southwest Jiaotong University. From 2014 to 2015, he was a Visiting Scholar at the University of Birmingham, Birmingham, UK. From 2010 to 2012, he undertook postdoctoral researches in the China Railway First Survey and Design Institute Group, Xi'an, China. His research interests include energy management and power quality for railway traction systems.



**Lie Xu** (Senior Member, IEEE) received the B.Sc. degree in mechatronics from Zhejiang University, Hangzhou, China, in 1993, and the Ph.D. degree in electrical engineering from The University of Sheffield, Sheffield, U.K., in 2000.

He worked with Queen's University Belfast, Belfast, U.K., and ALSTOM T&D, Stafford, U.K. He is currently a Professor with the Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, U.K. His current research interests include power electronics, wind energy generation and grid integration, and application of power electronics to power systems such as HVdc and medium voltage direct current (MVdc) systems for power transmission and distribution. Dr. Xu is an Editor of IEEE TRANSACTIONS ON POWER DELIVERY and IEEE TRANSACTIONS ON ENERGY CONVERSION.

**Zongyou Liang** received the B.Eng. and M.Eng. degree in the School of Electrical Engineering from Southwest Jiaotong University, Chengdu, China, in 2020 and 2023, respectively. His primary research interests is the optimal operation and planning for the integration of renewable energy and storage in high-speed railway system.