Capacity Aggregation and Online Control of Clustered Energy Storage Units

Boshen Zheng, *Graduate Student Member, IEEE*, Wei Wei, *Senior Member, IEEE*, Yin Xu, *Senior Member, IEEE*, and Yue Chen, *Member, IEEE*

***Abstract*—With the growing penetration of renewable energy and gradual retirement of thermal generators, energy storage is expected to provide flexibility and regulation services in future power systems. Battery is a major form of energy storage at the demand side. To better exploit the flexibility potential of massive distributed battery energy storage units, they can be aggregated and thus get enough capacity to participate in auxiliary service markets or receive direct orders from the power system operator. This paper proposes an analytical method to determine the aggregate MW-MWh capacity of clustered energy storage units controlled by an aggregator. Upon receiving the gross dispatch order, a capacity-aware water-filling policy is developed to allocate the dispatched power among individual energy storage units, which is called disaggregation. The policy endeavors to track the dispatch order while reducing the difference among the storage level of each battery. It provides effective real-time power control strategies for a particular class of energy management problem without referring to the prediction of dispatch order, although storage operation must respect inter-temporal constraints. Case studies validate the proposed method through applications in wind power ramping alleviation and frequency regulation.**

***Index Terms*—Aggregation, disaggregation, distributed energy storage, flexibility, real-time power control.**

NOMENCLATURE

Main Symbols and Notations Used in This Paper are Defined Below for Quick Reference. Others are Clarified Following Their First Appearances in Case of Need

Indices and Sets

i	Index of storage units $i \in \mathcal{N} = \{1, \ldots, N\}$
t	Index of time periods $t \in \mathcal{T} = \{1, 2, \ldots, T\}$
Ω_i	Feasible set of actions for storage unit i
Ω_{sum}	Minkowski sum of $\Omega_i, i \in \mathcal{N}$
Ω_{ag}	Feasible set of actions for the aggregator.

Parameters

Δ_t	Duration of time period.
η^c/η^d	Charging/Discharging efficiency, i for a small storage unit i, ag for the aggregator.
P_i	Power capacity of small storage unit i
E_i	Energy capacity of small storage unit i
T_i	Rated charging time, $T_i = E_i/P_i$
T_{ag}	$T_{ag} = \max\{T_1, T_2, \ldots, T_N\}$
D_t^c/D_t^d	Charging/Discharging dispatch signal from power system operator in period t
α_i	Occupied level of channel i

Decision Variables

$p_{i,t}^c$	Charging power of storage unit i in period t
$p_{i,t}^d$	Discharging power of storage unit i in period t
$e_{i,t}$	Storage level of storage unit i in period t
$s_{i,t}$	State-of-charge of storage unit i in period t
$p_{ag,t}^c$	Charging power of the aggregator in period t
$p_{ag,t}^d$	Discharging power of the aggregator in period t
$e_{ag,t}$	Storage level of the aggregator in period t
$s_{ag,t}$	State-of-charge of the aggregator in period t
P_{ag}	Aggregate power capacity.
E_{ag}	Aggregate energy capacity.

Manuscript received 6 July 2023; revised 18 December 2023; accepted 14 January 2024. Date of publication 19 January 2024; date of current version 21 June 2024. This work was supported by the National Key Research and Development Program of China under Grant 2022YFB2405500. Paper no. TSTE-00711-2023. *(Corresponding author: Wei Wei.)*

Boshen Zheng and Wei Wei are with the Department of Electrical Engineering, Tsinghua University, Beijing 100084, China (e-mail: zbs21@mails.tsinghua.edu.cn; wei-wei04@mails.tsinghua.edu.cn).

Yin Xu is with the School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China (e-mail: xuyin@bjtu.edu.cn).

Yue Chen is with the Department of Mechanical and Automation Engineering, Chinese University of Hong Kong, Hong Kong, SAR 000000, China (e-mail: yuechen@mae.cuhk.edu.hk).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSTE.2024.3355991.

Digital Object Identifier 10.1109/TSTE.2024.3355991

I. INTRODUCTION

THE power grid is undergoing the transition to a renewable dominated system, with the proliferation of wind and solar plants. The volatility of renewable energy requires more flexible resources to maintain power balance. However, controllable thermal generators are gradually replaced by renewable plants to achieve carbon neutrality, leading to a lack of adequate flexibility resources. Deploying energy storage unit is an effective way to support renewable energy integration. Existing works focus on the coordinated operation of large centralized energy storage units (ESU) with grid-side facilities, such as in unit commitment [1], [2], [3] and economic dispatch [4], [5] problems.

Considering the relatively short lifetime of battery, deployment of large centralized ESU is an expensive option. At the demand side, various consumers own small ESUs, but it is difficult for the power grid operator to dispatch numerous distributed ESUs directly. Hence, aggregation of distributed ESUs offers a hierarchical framework that allows distributed storage resources to act together as a large centralized storage unit [6], [7]. Under

TABLE I
LITERATURE ON FLEXIBILITY RESOURCES AGGREGATION

Category	Ref.	Facilities	Consideration	Model	Solution	Outcome
Envelope	[14]	Flexible load	Flexibility reserve	MPC	Linear programming	Feasible interval
	[15]	DERs	Flexibility reserve	Soft robust optimization	Quadratic programming	Feasible interval
	[16]	Energy hub	Power aggregation	Soft robust optimization	Quadratic programming	Feasible interval
	[17]	DERs	Power aggregation	Adaptive robust optimization	Convex optimization	Feasible interval
	[18]	DERs	Power aggregation	Robust optimization	Mixed integer linear programming	Feasible interval
Set operation	[20]	Thermal controllable loads	Frequency regulation	Virtual battery	Linear programming	Feasible interval
	[21]	Thermal controllable loads	Frequency regulation	Virtual battery	Linear programming	Feasible interval
	[22]	Electric vehicles	Energy arbitrage	Virtual battery	Linear programming	Feasible interval
	[23]	DERs	Energy market	Inner approximation	Linear programming	Feasible zonotope
	[24]	Flexible load	Demand response	Outer approximation	Linear programming	Feasible polytope
	[25]	Flexible load	Demand response	Outer approximation	Second-order cone programming	Feasible polytope
Proposed		ESU	Ancillary service	Inner approximation	Analytical expression	(E_{ag}, P_{ag}, s_{ag})

TABLE II
LITERATURE ON POWER DISAGGREGATION

Category	Ref.	Facilities	Consideration	Uncertainty resources	Model	Solution
Disaggregation feasibility	[16]	Energy hub	Power aggregation	Facilities, price	Soft robust optimization	Quadratic programming
	[17]	DERs	Power aggregation	Facilities	Adaptive robust optimization	Convex optimization
	[27]	Flexible load	Energy-Reserve market	Facilities	Stochastic optimization	Quadratic programming
	[28]	DERs	Optimal power flow	Facilities	Open-loop control	Convex optimization
ESUs management	[26]	ESU	Frequency regulation	None	Cooperative hierarchical control	Optimal control
	[29]	ESU	Frequency regulation	None	Leader-follower consensus control	Proportional-Integral control
	[30]	ESU	Islanded microgrid	None	$P-f$ droop control	Droop control
	[31]	ESU	DC microgrid	None	Consensus-based control	Riccati equation
	[32]	ESU	Frequency control	Dispatch order	Model predictive control	Quadratic programming
	[33]	ESU	Frequency regulation	Dispatch order	Coordinated & priority control	Decision tree
Proposed		ESU	Ancillary service	Dispatch order	Capacity-aware water-filling	Closed-form

some rental agreements, storage owners would like to share some capacity of their ESU, and an aggregator who manages the storage cluster can obtain adequately large capacity to enter auxiliary service markets or receive direct orders from the grid operator [8], [9], [10]. In such a hierarchy architecture, the aggregator submits the aggregate power and energy capacities as well as the storage level to the system operator, which is termed *aggregation*; then from the power grid side, the aggregator is a large ESU. When the aggregator receives a charging or discharging order, it must allocate the dispatched power among small ESUs, which is called *disaggregation*. Existing works focus on either of the two aspects. The literatures are summarized in Tables I and II.

1) The Aggregation Problem: Many works have investigated the flexibility aggregation problem [11], [12], [13], and existing approaches could be roughly categorized into two classes.

The first one aims to find an envelope of the aggregate power profile, such that for any trajectory in this envelope, there exists a disaggregation control strategy that allocates the aggregate power among each device while retaining their operation feasibility. In [14], robust model predictive control (MPC) is adopted to declare the aggregate power reserve flexibility of flexible load in a rolling horizon manner. In [15] and [16], the lower and upper bounds of the aggregate power over multiple periods are calculated from a linear or quadratic program; it is proven that any power trajectory residing in the envelope is implementable via a non-anticipative quantile policy. To reduce the conservativeness of hypercube approximation, an ellipsoidal approximation method is proposed in [17]. In [18], distributed energy resources (DERs) are classified as generator-like and battery-like devices; then a polytope-based bound shrinking approximation method is proposed to further improve accuracy. The outcome is a feasible interval of total power in each period. Such aggregation method is suitable for power aggregation of integrated flexible load, renewable and storage-like device. For the aggregation involving only energy storage devices, the aggregator can be characterized by two parameters (power capacity and energy capacity) and a real-time state (storage level or state-of-charge, SoC for short), which is called the Energy-Power-SoC model. In this work, we want to obtain the Energy-Power-SoC triple. However, the envelope-based methods do not offer the desired information.

The second one is based on polyhedral set operation, mainly Minkowski sum. From a geometric perspective, the feasible operating set of each device can be described via a polyhedral set. The flexibility aggregation can be viewed as calculating the Minkowski sum of all these polyhedral sets. However, this set lacks intuitive interpretation. Given the difficulty of computing the exact Minkowski sum under facet representation [19], an alternative way is to build inner/outer approximations for aggregated power feasible region. For the inner approximation, the key is to find a simple but feature-capturing prototype set to conservatively approximate the exact Minkowski sum polytope.

In [20], [21], [22], virtual battery model is first adopted to describe the aggregate flexibility of thermal controllable loads or electric vehicles; then polyhedral projection is employed to project the augmented polytope to a low-dimensional space and obtain the aggregate power feasible range. In [23], optimal zonotope is calculated to inner-approximate individual polyhedral feasible set; since zonotope is closed under Minkowski sum, the aggregate flexibility can be calculated efficiently by direct summation. The selection of prototype set is vital because an improper prototype will lead to conservative results and loss of flexibility. In [24], an outer approximation of Minkowski sum of polytopes is developed to represent the aggregate capabilities of heterogeneous flexible loads via homothetic polyhedra; one step further, the outer approximation is extended from [24] to convex conic sets with second-order cone and semidefinite constraints in [25]. But the outer approximation may appear to be optimistic and cause infeasibility in disaggregation stage. The above aggregation methods mainly focus on determining an aggregate power feasible interval/polytope for flexible resources, but the aggregation technique for ESUs remains an open problem. The power flexibility of individual ESU is a polytope; thus we can use Minkowski sum method to model the aggregate flexibility of the ESU cluster. Considering operation security, inner approximation is suitable in representing aggregate flexibility. The Energy-Power-SoC model is a natural choice to describe the aggregated flexibility of clustered ESUs. The problem is how to determine the aggregated capacity parameters.

For ESUs aggregation, the aggregated power and energy capacities are acquired by directly summing all individual capacities in [26], but these individual ESUs have the same rated charging time; otherwise, direct summation is not applicable. In addition, the aggregated power and energy capacities should be acquired while the operating feasibility of individual ESUs must be guaranteed, which is affected by how individual ESUs are controlled (i.e., the disaggregation policy).

2) The Disaggregation Problem: When the aggregator receives a charging/discharging order, it needs to decide the real-time power control strategy of each ESU, which affects the flexibility in future periods due to the inter-temporal storage level dynamics. However, the future dispatch signal is not known in advance. If the signal is not properly disaggregated, some ESUs may be fully charged or depleted earlier than others, thus reducing the aggregate flexibility in the next period. The most fundamental requirement of disaggregation policy is *non-anticipativity*, which means no information on future dispatch signal is needed, so the causality of control actions is respected. Existing works on power disaggregation can be classified into two categories based on their topics: Disaggregation feasibility and ESUs management. The first category focuses on the existence of a disaggregation strategy. In [16], disaggregation is implemented by a non-anticipative quantile policy. Soft robust optimization is used to ensure the existence of disaggregate action even in the worst-case scenario. Adaptive robust optimization is adopted in [17], where an affine policy is proposed to ensure the existence of feasible disaggregation strategy. A power disaggregation policy can be obtained by solving a cost minimization problem [23], tracking contracted power with rolling horizon stochastic optimization [27], or following a pre-determined open-loop control policy [28]. But the above literatures only focus on the internal uncertainty originated from its participants, while the external uncertainty from dispatch signal is neglected. The second category is for the management of multiple ESUs, which is closer to the power disaggregation problem, and SoC consistency is a widely adopted philosophy, such as cooperative hierarchical control [26], leader-follower consensus control [29], droop control [30] and consensus-based control [31]. These methods focus on the control scheme design and dynamic response performance; but the uncertainty of control signal received by the aggregator is not considered. In practice, the dispatch signal is received period-by-period, and the aggregator needs to disaggregate the current dispatch signal and control individual ESUs without relying on predicting future dispatch signals. An improper power disaggregation strategy will reduce the flexibility of ESUs cluster. Model predictive control [32] and priority control [33] are also mature techniques, but the former relies on high-quality prediction and the latter leads to over-use of some ESUs.

This paper addresses capacity aggregation and real-time disaggregation control of clustered energy storage in an integrated framework. The contributions include:

1) An analytical solution of capacity aggregation. We adopt the Energy-Power-SoC model to describe the aggregated ESU. In the state-of-the-art geometric approach based on Minkowski sum, approximation is needed and the aggregation problem is finally cast as an optimization problem after complicated reformulation process. Inspired by the SoC consistency idea, we discover a closed-form solution of the MW-MWh capacity of the ESU cluster without solving an optimization problem or calculating the Minkowski sum. The aggregation result is shown to be equivalent to the Minkowski sum based approach but significantly simplified. The analytical expression of aggregated capacity enables various online evaluation, decision and control tasks. For example, when small ESUs are allowed to freely enter or exit the aggregator, the MW-MWh capacity of the ESU cluster should be updated in real time. The proposed method is very convenient to show the aggregated capacity online, while other approaches entail substantial computation effort. It also offers insights on how to preserve flexibility as much as possible through the idea of SoC consistency.
2) A capacity-aware water-filling (CAWF) policy for power disaggregation control. Inspired by the SoC consistency idea, the disaggregation control problem is cast as a convex optimization problem with a resource constraint and two-side bound constraints on charging/discharging power of each ESU. Compared to the standard water filling algorithm [34], where action is not upper bounded, the maximum charging-discharging power is imposed, which captures the practical requirements of storage operation. Compared with SoC consistent control methods, proposed method aims to track uncertain dispatch signal and preserve flexibility for future use. Compared to the water filling methods for distribution system

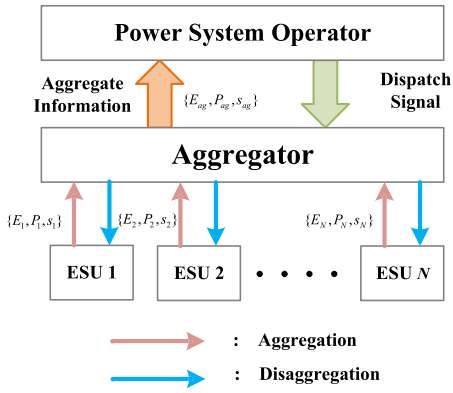


Fig. 1. Framework for ESUs aggregation and disaggregation.

operation which renders a temporal scale energy allocation for load-flattening [35], [36], [37], the proposed method gives spatial-scale coordination for allocating dispatched power among distributed ESUs. Although ESU operation involves inter-temporal constraints on storage dynamics, the CAWF policy is prediction-free and can be applied in an online fashion. We further develop a closed-form of the CAWF policy, which can be implemented with simple hardware without an advanced computational platform.

The rest of this paper is organized as follows: The aggregation and disaggregation problems are formally stated in Section II. The former one is addressed in Section III, in which the aggregated capacity of the ESU cluster is analytically characterized; Section IV resolves the disaggregation problem and develops the CAWF. The holistic framework is presented in Section V. Case study is reported in Section VI; the proposed method is validated through applications in wind power ramping alleviation and frequency regulation. Finally, Section VII concludes the paper.

II. Problem Statement

A. System Configuration

We consider an aggregator that operates a cluster of ESUs indexed by $i \in \mathcal{N} = \{1, 2, \ldots, N\}$. Each ESU submits its power capacity P_i, energy capacity E_i and SoC $s_i = e_i/E_i$ to the aggregator. The power grid dispatches the aggregator as a large ESU whose power and energy capacities are P_{ag} and E_{ag}, respectively. After receiving the dispatch signal, the aggregator needs to disaggregate the power signal among small ESUs. The hierarchical structure is shown in Fig. 1.

In this business model, we assume the aggregator has designed an effective mechanism to incentivize the ESU owners to share their surplus/unused capacity with the aggregator. A simple mechanism is that the aggregator pays rental fee to the ESUs owners. The rental fee may include the opportunity cost and degradation cost. The mechanism design is related to storage sharing and is not the main scope of this work. With a fixed number of small ESUs and rental fees, the aggregator would focus on the interaction with the power system. It has to submit the aggregated MW-MWh capacity and control small ESUs to fulfill the dispatch order. The former is determined in the aggregation problem and the latter is addressed in the disaggregation problem.

B. Capacity Aggregation

Because it is difficult for the power grid operator to manage massive small ESUs individually, the aggregator reduces the complexity by acting as a centralized ESU and submitting aggregated power capacity P_{ag} and energy capacity E_{ag} to the grid operator. The basic requirement for aggregation is that *whenever the aggregated ESU is not fully charged or empty, it is responsible to respond to a charge or discharge order no greater than* P_{ag}.

By disclosing the power capacity P_i and energy capacity E_i, the operation of i-th ESU is constrained by

$$0 \le p^c_{i,t} \le P_i, 0 \le p^d_{i,t} \le P_i, \forall t \tag{1a}$$

$$e_{i,t} = e_{i,t-1} + (\eta^c_i p^c_{i,t} - p^d_{i,t}/\eta^d_i)\Delta_t, \forall t \tag{1b}$$

$$0 \le e_{i,t} \le E_i, \forall t \tag{1c}$$

where η^c_i/η^d_i denotes the charging/discharging efficiency; Δ_t is the duration of period t; $e_{i,t}$ represents the storage level; control action $p_{i,t} = \{p^c_{i,t}, p^d_{i,t}\}$ includes charging/discharging power $p^c_{i,t}/p^d_{i,t}$. Constraint (1a) restricts charging/discharging power within the power capacity. Inter-temporal constraint (1b) describes the dynamics of storage level which must reside in the interval (1c). Expressing $e_{i,t}$ through $p_{i,t}$, constraints (1b) and (1c) reduce to

$$0 \le e_{i,0} + \sum_{t'=1}^{t} (\eta^c_i p^c_{i,t'} - p^d_{i,t'}/\eta^d_i)\Delta_t \le E_i, \forall t \tag{1d}$$

The feasible set of control action is denoted as a polyhedron

$$\Omega_i(P_i, E_i) = \{p_{i,t} | \text{ (1a), (1d)}\} \tag{2}$$

depending on the values of P_i and E_i. It characterizes the individual flexibility of each ESU.

As a large ESU, the power grid only perceives the total power and energy of the aggregator, i.e.

$$p_{ag,t} = \sum_{i \in \mathcal{N}} p_{i,t}, \; e_{ag,t} = \sum_{i \in \mathcal{N}} e_{i,t} \tag{3}$$

where $p_{ag,t} = \{p^c_{ag,t}, p^d_{ag,t}\}$. By the definition of Minkowski sum, the feasible set of aggregate power profile is

$$\Omega_{sum} = \Omega_1(P_1, E_1) \oplus \cdots \oplus \Omega_N(P_N, E_N) \tag{4}$$

where symbol $\oplus$ denotes the Minkowski sum. However, it is complicated to compute the hyperplane representation of Ω_{sum}, and the number of constraints in Ω_{sum} increases exponentially with a large number of ESUs and time periods. Hence, it is less attractive to compute the exact expression of set Ω_{sum} in engineering practice, although it is interesting and important from a theoretical perspective.

In this paper, the ESU cluster is treated as a large ESU with aggregated power capacity P_{ag} and energy capacity E_{ag}. Its

operation is constrained by

$$0 \le p^c_{ag,t} \le P_{ag}, 0 \le p^d_{ag,t} \le P_{ag}, \forall t \tag{5a}$$

$$0 \le e_{ag,0} + \sum_{t'=1}^{t} (\eta^c_{ag} p^c_{ag,t'} - p^d_{ag,t'}/\eta^d_{ag})\Delta_t \le E_{ag}, \forall t \tag{5b}$$

Similar to (2), the feasible set characterized by (P_{ag}, E_{ag}) is

$$\Omega_{ag}(P_{ag}, E_{ag}) = \{p_{ag,t} | (5)\} \tag{6}$$

Although polyhedra Ω_{sum} and $\Omega_{ag}(P_{ag}, E_{ag})$ share the same dimension, they can be different in shape. To ensure the existence of an operation strategy for individual ESUs satisfying constraint (1), the condition

$$\Omega_{ag}(P_{ag}, E_{ag}) \subseteq \Omega_{sum} \tag{7}$$

must hold. The capacity aggregation problem aims to find the optimal tuple (P_{ag}, E_{ag}) such that $\Omega_{ag}(P_{ag}, E_{ag})$ is the largest among all eligible candidates satisfying (7). Clearly, the optimal (P_{ag}, E_{ag}) depends on $(P_1, E_1), \ldots, (P_N, E_N)$. The capacity aggregation problem is summarized below.

Problem 1: Capacity Aggregation Problem.

1: **Input:** Individual ESU parameters $\{E_i, P_i, s_i\}$
2: **Task:** Finding the optimal tuple (P_{ag}, E_{ag}) such that $\Omega_{ag}(P_{ag}, E_{ag})$ is the largest among all eligible candidates satisfying condition (7).
3: **Output:** Aggregate ESU parameters $\{E_{ag}, P_{ag}, s_{ag}\}$

As Ω_{ag} is a subset of Ω_{sum}, dispatching the aggregated ESU is less flexible than dispatching all small ESUs individually. We propose an index LoF to quantify the loss of flexibility during the capacity aggregation process:

$$\text{LoF} = 1 - \sqrt[D]{\text{vol}(\Omega_{ag})/\text{vol}(\Omega_{sum})} \tag{8}$$

where $\text{vol}(\Omega)$ is the volume of a full dimensional polytope Ω; D is its dimension. The $\sqrt[D]{(\cdot)}$ operator normalizes the volume ratio into dimension-1.

Although the aggregation process sacrifices some flexibility, it significantly simplifies the information exchange between the aggregator and grid operator. Since the aggregator submits aggregated information to the operator, from the power grid aspect, the ESU cluster becomes a single ESU, and hence existing results on dispatching centralized ESUs can be applied.

The loss of flexibility is a trade-off between conservativeness and security. On the one hand, the aggregator hopes P_{ag} is as large as possible to provide more flexibility; on the other hand, the aggregation method does not provide an explicit rule to control individual storage unit under which the desired ability (to respond to a charge or discharge order no greater than P_{ag}) is warranted, which inspires the power disaggregation control problem.

C. Power Disaggregation Control

In each period t, the grid operator sends either a charging order D^c_t or a discharging order D^d_t to the aggregator, where $D^c_t, D^d_t \in [0, P_{ag}]$ and $D^c_t \cdot D^d_t = 0$. The aggregator controls each ESU to track the dispatch order. Given the sequence $\{D^c_t, D^d_t\}_{t=1}^T$ of dispatch orders over a certain period, the disaggregation control problem is formulated as follows:

$$\min \sum_t \left[\left(D^c_t - p^c_{ag,t}\right) + \left(D^d_t - p^d_{ag,t}\right) \right] \tag{9a}$$

$$\text{s.t. } (3), \{p^c_{i,t}, p^d_{i,t}\} \in \Omega_i(P_i, E_i), \forall i, \forall t \tag{9b}$$

$$0 \le p^c_{ag,t} \le D^c_t,\ 0 \le p^d_{ag,t} \le D^d_t, \forall t \tag{9c}$$

$$p^c_{i,t} \ge 0, p^d_{i,t} = 0, \forall i \quad \text{if } D^c_t > 0 \tag{9d}$$

$$p^d_{i,t} \ge 0, p^c_{i,t} = 0, \forall i \quad \text{if } D^d_t > 0 \tag{9e}$$

where the objective function (9a) aims to minimize the total deviation from the dispatch order. Upon receiving the dispatch signal, the aggregator has to respond by controlling small ESUs. Because the capacity is limited, the aggregator may not be able to completely fulfill the dispatch order, and also has no incentive for an overreaction, which incurs no benefit but consumes more flexibility potential. So we restrict the charging and discharging actions of aggregator within the intervals $[0, D^c_t]$ and $[0, D^d_t]$, as in (9c) Constraints (9d)–(9e) stipulate the status of each ESU according to the dispatch order, preventing simultaneous charging and discharging. The remaining constraints in (9b) respect the operational feasibility of individual ESUs. Problem (9) is able to model a variety of applications, such as wind power ramping alleviation and frequency regulation, which will be detailed in the case study. The power disaggregation problem is summarized below.

Problem 2: Power Disaggregation Control Problem.

1: **Input:** Dispatch order from power system operator $\{D^c_t, D^d_t\}$
2: **Task:** Allocating $\{D^c_t, D^d_t\}$ among individual ESUs to track the dispatch order, as shown in problem (9).
3: **Output:** Power disaggregation strategy $\{p^c_{i,t}, p^d_{i,t}\}_{i=1}^N$

In practice, however, the dispatch orders over the entire horizon are not sent at the same time. Particularly, the control action $\{p^c_{i,t}, p^d_{i,t}\}$ must be deployed to each ESU after receiving D^c_t/D^d_t in period t without clear information about those orders in future periods. In general, there exist infinitely many feasible disaggregation strategies in a period. However, the disaggregation strategy in period t affects the feasibility of the inter-temporal constraint (5b) and thus the available flexibility in future time periods. In this regard, the disaggregated power control policy should be non-anticipative, which means that $\{p^c_{i,t}, p^d_{i,t}\}$ should not rely on the information about $D^c_{t+1}/D^d_{t+1}, D^c_{t+2}/D^d_{t+2}, \ldots$.

III. Capacity Aggregation

This section presents the analytical expression of aggregated capacity of the ESU cluster. First, the case with ideal ESUs is discussed, and then the result is extended to lossy ESUs.

A. The Case With Ideal ESUs

We consider the case in which all ESUs are lossless, which means $\eta_i^c = \eta_i^d = \eta_{ag}^c = \eta_{ag}^d = 1$. We also assume $e_{ag,0} = 0$, implying $e_{i,0} = 0, \forall i$. This is merely an assumption to derive the aggregated capacity that is independent of the initial SoC. The disaggregation policy allows different and non-zero initial SoCs, and the real value of $e_{ag,0}$ can be easily obtained.

Clearly, the aggregated energy capacity should not exceed the total capacity of all ESUs in the cluster, that is

$$E_{ag} = \sum_{i\in\mathcal{N}} E_i \tag{10}$$

Given the aggregated energy capacity E_{ag}, the SoC of aggregate ESU is expressed as:

$$s_{ag,t} = \frac{e_{ag,t}}{E_{ag}} = \frac{\sum_{i\in\mathcal{N}} e_{i,t}}{E_{ag}}, \forall t \tag{11}$$

where $e_{i,t} = s_{i,t}E_i$. It can be seen that the SoC of aggregated ESU is the weighted average of individual SoCs.

By condition (7) and the lossless assumption on ESUs, P_{ag} solves the following optimization problem:

$$\max\ P_{ag} \tag{12a}$$

$$\text{s.t.}\ -P_i \le p_{i,t} \le P_i, \forall i, \forall t \tag{12b}$$

$$0 \le \sum_{t'=1}^{t} p_{i,t'}\Delta_t \le E_i, \forall i, \forall t \tag{12c}$$

$$p_{ag,t} = \sum_{i\in\mathcal{N}} p_{i,t}, \forall t \tag{12d}$$

$$0 \le \sum_{t'=1}^{t} p_{ag,t'}\Delta_t \le E_{ag}, \forall t \tag{12e}$$

$$P_{ag} \le p_{ag,t}, \forall t \tag{12f}$$

Constraint (12b) restricts the power of individual ESUs; constraint (12c) describes the SoC dynamics and feasible set of individual ESUs; constraint (12d) is same as (3); constraint (12e) prescribes the SoC dynamics and bound of aggregated ESU cluster; constraint (12f) warrants condition (7).

In the following, we will show how to acquire P_{ag} analytically instead of solving problem (12). The insights lead to the development of disaggregated power control policy, in spite of the fact that solving problem (12) is not difficult. Let $T_i = E_i/P_i$ be the rated charging time of storage unit i, and $T_{ag} = \max\{T_1, T_2, \ldots, T_N\}$.

Proposition 1: The aggregate power capacity can be explicitly expressed as

$$P_{ag} = E_{ag}/T_{ag} \tag{13}$$

To see this, for any $t = 1:T$, adding constraints (12c) together for $i \in \mathcal{N}$ and considering (12d) yields:

$$\sum_{t=1}^{T} p_{ag,t}\Delta_t \le \sum_{i\in\mathcal{N}} E_i = E_{ag} \tag{14}$$

Add constraint (12f) from $t = 1:T$

$$\frac{TP_{ag}}{\Delta_t} \le \sum_{t=1}^{T} p_{ag,t} \tag{15}$$

Combining (14) and (15) we obtain

$$P_{ag} \le E_{ag}/T \tag{16}$$

The above inequality is tight, which means equality in (16) can be achieved. To see this, assume all ESUs are empty at the beginning. Because the aggregator wants to charge the aggregated ESU with a constant power that is as much as possible, it uses the SoC consistent policy: the ESU with largest rated charging time T_i is charged at maximum power P_i, and the charging power of ESU j, $j \ne i$ is calculated by

$$P_j = \frac{P_i}{E_i}\cdot E_j = \frac{E_j}{T}$$

such that the SoC of all ESUs are kept at the same level. Under this policy, the maximum charging power is

$$P_{ag} = \sum_{j=1}^{N} \frac{E_j}{T} = \frac{E_{ag}}{T}$$

If a dispatch order exceeds P_{ag}, the storage levels can no longer be maintained at the same level, because the i-th ESU already reaches the maximum charging power. If we allow SoC discrepancy, some ESUs will be fully charged earlier than the remaining ones. Once any ESU is fully charged, the aggregated charging power must be reduced. In this regard, the SoC of the aggregated ESU must follow the ESU with the largest E_i/P_i, which is equal to T_{ag}. So P_{ag} in (13) is optimal.

B. Extension to Non-Ideal Storage

Now consider non-ideal ESUs. When the aggregator receives a charging signal D^c, to keep SoC levels of all ESUs at the same level, we have

$$\begin{cases} p_{ag} = p_1 + \cdots + p_N = D^c \\ \frac{\eta_1^c p_1}{E_1} = \cdots = \frac{\eta_N^c p_N}{E_N} = \frac{\eta_{ag}^c p_{ag}}{E_{ag}} \end{cases}$$

where

$$\eta_{ag}^c = \frac{E_{ag}}{E_1/\eta_1^c + \cdots + E_N/\eta_N^c} \tag{17}$$

represents the aggregated charging efficiency. A similar condition can be given for a discharging signal D^d as follows:

$$\begin{cases} p_{ag} = p_1 + \cdots + p_N = D^d \\ \frac{p_1}{\eta_1^d E_1} = \cdots = \frac{p_N}{\eta_N^d E_N} = \frac{p_{ag}}{\eta_{ag}^d E_{ag}} \end{cases}$$

where

$$\eta_{ag}^d = (E_1\eta_1^d + \cdots + E_N\eta_N^d)/E_{ag} \tag{18}$$

represents the aggregated discharging efficiency.

IV. Online Disaggregation Policy

A. Intuitive Idea of Power Disaggregation

In the disaggregation control problem (9), the difficulty arises from the fact that in period t, the future dispatch order is unknown. An improper disaggregation strategy may either deplete or fully charge some ESUs earlier than others, thus reducing the response ability of aggregator in future periods.

An example is given to demonstrate the above phenomenon. Consider a problem with 2 periods and two ideal ESUs, whose parameters are listed as follows:

$$\begin{cases} P_1 = 10\,\text{MW},\ E_1 = 20\,\text{MWh},\ e_{1,0} = 10\,\text{MWh} \\ P_2 = 10\,\text{MW},\ E_2 = 40\,\text{MWh},\ e_{2,0} = 20\,\text{MWh} \end{cases}$$

So the parameter of the aggregator is

$$E_{ag} = 60\,\text{MWh},\ P_{ag} = 15\,\text{MW},\ s_{ag,0} = 0.5$$

In period t_1, the aggregated charging power is 15 MW. We consider three disaggregation strategies:

$$\begin{aligned} &\text{(S1)} \quad p_{1,t_1} = 10\,\text{MW}, \quad p_{2,t_1} = 5.0\,\text{MW} \\ &\text{(S2)} \quad p_{1,t_1} = 7.5\,\text{MW}, \quad p_{2,t_1} = 7.5\,\text{MW} \\ &\text{(S3)} \quad p_{1,t_1} = 5.0\,\text{MW}, \quad p_{2,t_1} = 10\,\text{MW} \end{aligned}$$

All three strategies are valid responses. At the beginning of period t_2, the respective storage levels of the two ESUs are:

$$\begin{aligned} &\text{Case 1} \quad s_{1,t_1} = 1.000, \quad s_{2,t_1} = 0.6250 \\ &\text{Case 2} \quad s_{1,t_1} = 0.875, \quad s_{2,t_1} = 0.6875 \\ &\text{Case 3} \quad s_{1,t_1} = 0.750, \quad s_{2,t_1} = 0.7500 \end{aligned}$$

Hence, the admissible charging power ranges under the three strategies are:

$$\begin{aligned} &\text{Case 1} \quad p_{1,t_2} \in [0.0, 0.0], \quad p_{2,t_2} \in [0, 10] \\ &\text{Case 2} \quad p_{1,t_2} \in [0.0, 2.5], \quad p_{2,t_2} \in [0, 10] \\ &\text{Case 3} \quad p_{1,t_2} \in [0.0, 5.0], \quad p_{2,t_2} \in [0, 10] \end{aligned}$$

In period t_2, the power grid requests the aggregator continue to charge. The response ability of the aggregator in three cases is analyzed

a) If the charging request is less than 10 MW, then the aggregator can respond to the request in all three cases;
b) If the charging request is between 10 MW and 12.5 MW, then the aggregator is unable to completely follow the request in Case 1 and can still fulfill the charging request in Case 2 and Case 3;
c) If the charging request is between 12.5 MW and 15 MW, then the aggregator is unable to completely follow the request in Case 1 and Case 2, while it can still fulfill the charging request in Case 3.

In practice, when the aggregator makes decision in period t_1, it has no information of the charging request in period t_2. We can see that if a situation where strategy S1 is applied can be handled in period t_2, then the situation where strategy (S2)/(S3) is applied can also be handled in period t_2. So the aggregator is more flexible in Case 2 and Case 3 than it is in Case 1, and hence strategies S2 and S3 dominate strategy S1. For the same reason, strategy S3 dominates strategy S2.

The above observation inspires the SoC consistency principle. If we coordinate all ESUs and keep their storage levels as close as possible, and then the aggregate power flexibility will be reserved to a greater extent. Based on this observation, we develop a capacity-aware water-filling algorithm for disaggregated power control policy.

B. Capacity-Aware Water-Filling Algorithm

Water-filling algorithm is usually applied in communication system for allocating limited power over a set of communication channels, aimed at utilizing maximum channel capacity. In the disaggregated power control, considering the charging order, the capacity-aware water-filling problem is cast as:

$$\min \quad -\sum_{i\in\mathcal{N}} \log(\alpha_i + x_i) \tag{19a}$$

$$\text{s.t. } 0 \le x_i \le \bar{x}_i, \forall i \tag{19b}$$

$$\sum_{i\in\mathcal{N}} x_i = 1 \tag{19c}$$

where constant α_i corresponds to the storage level, and the dispatched power x_i is normalized to be equal to 1, as in (19c); the capacities of small ESUs are taken into account in (19b). The detailed correspondence will be explained later. In the communication system, channel capacity is not a main concern, so x_i is not upper bounded by $\bar{x}_i$; see the example in page 245 [34]. However, in energy storage control, charging power limit must be taken into account in (19b). In this regard, the proposed method is called capacity-aware water-filling algorithm. Although problem (19) remains convex, introducing upper bound constraints complicates the development of an explicit policy, compared to that in [34].

To solve problem (19) without calling a solver, its Karush-Kuhn-Tucker (KKT) optimality condition is analyzed. By introducing Lagrange multipliers λ_i^l and λ_i^u for lower and upper bound constraints in (19b) and a multiplier ν for equality (19c), we can obtain the following optimality condition

$$-(\alpha_i + x_i)^{-1} - \lambda_i^l + \lambda_i^u + \nu = 0 \tag{20a}$$

$$0 \le x_i \le \bar{x}_i,\ \lambda_i^l \ge 0,\ \lambda_i^u \ge 0,\ \forall i \tag{20b}$$

$$(\bar{x}_i - x_i)\lambda_i^u = 0,\ x_i\lambda_i^l = 0,\ \forall i \tag{20c}$$

$$\sum_{i\in\mathcal{N}} x_i = 1 \tag{20d}$$

By (20a), λ_i^l can be expressed using other variables and thus eliminated, then the optimality conditions give rise to

$$0 \le x_i \perp (\lambda_i^u + \nu - (\alpha_i + x_i)^{-1}) \ge 0, \forall i \tag{21a}$$

$$0 \le (\bar{x}_i - x_i) \perp \lambda_i^u \ge 0, \forall i \tag{21b}$$

$$\sum_{i\in\mathcal{N}} x_i = 1 \tag{21c}$$

where $0 \le a \perp b \ge 0$ represents complementary slackness, requiring at least one of two non-negative quantities a and b to be zero. By condition (21), we have four cases:

1) If $\nu \ge 1/\alpha_i$, $\lambda_i^u + \nu - (\alpha_i + x_i)^{-1} > 0$, then $x_i^* = 0$;
2) If $1/(\alpha_i + \bar{x}_i) < \nu < 1/\alpha_i$, suppose that λ_i^u is positive, it follows that $x_i = \bar{x}_i$ due to (21b). But $\lambda_i^u + \nu - (\alpha_i + \bar{x}_i)^{-1}$ must be positive by assumption; consequently, complementary slackness constraint (21a) is violated. Thus $\lambda_i^u = 0$, and from (21a) we have $\nu = (\alpha_i + x_i)^{-1}$, hence $x_i^* = 1/\nu - \alpha_i$;
3) If $\nu = 1/(\alpha_i + \bar{x}_i)$, then constraints (21a)–(21b) can hold if and only if $x_i^* = \bar{x}_i$;
4) If $\nu < 1/(\alpha_i + \bar{x}_i)$, λ_i^u must be positive. According to (21b), we must have $x_i^* = \bar{x}_i$.

In summary, if we have the value of ν, then the optimal solution of problem (19) is

$$x_i^* = \begin{cases} 0 & \nu \in [\alpha_i^{-1}, \infty) \\ 1/\nu - \alpha_i & \nu \in ((\alpha_i + \bar{x}_i)^{-1}, \alpha_i^{-1}) \\ \bar{x}_i & \nu \in (-\infty, (\alpha_i + \bar{x}_i)^{-1}] \end{cases} \tag{22}$$

For convenience, we write x_i^* as $x_i^*(\nu)$, which is a piecewise-linear increasing function in $1/\nu$. Substituting (22) into (21c), we have

$$g(\nu) = \sum_{i \in \mathcal{N}} x_i^*(\nu) = 1 \tag{23}$$

As each $x_i^*(\nu)$ is increasing in $1/\nu$, so is $g(\nu)$. Thus (23) has a unique root ν^*, which can be found by Algorithm 1.

Similarly, considering the discharging order, a water-draining problem is established, which is

$$\min \ -\sum_{i \in \mathcal{N}} \log(A - \alpha_i + x_i) \tag{24a}$$

$$\text{s.t. } 0 \le x_i \le \bar{x}_i, \forall i \tag{24b}$$

$$\sum_{i \in \mathcal{N}} x_i = 1 \tag{24c}$$

where A is a unified upper boundary of capacity. The solution can be acquired in the same way as (19), which is

$$x_i^* = \begin{cases} 0 & \nu \in [(A - \alpha_i)^{-1}, \infty) \\ 1/\nu - \alpha_i & \text{otherwise} \\ \bar{x}_i & \nu \in (-\infty, (A - \alpha_i + \bar{x}_i)^{-1}] \end{cases} \tag{25}$$

The ν value is determined through (24c). The graphic interpretation is shown in Fig. 2.

C. Online Analytical Disaggregation Policy

Based on the closed-form solutions of problems (19) and (24), we develop a disaggregated policy for real-time power control, which can be implemented on microprocessor without an advanced computational platform.

1) Disaggregation of Charging Power: If the aggregator receives a charging order D_t^c, the control actions $p_{i,t}^c$ of individual ESUs solve the following problem

$$\min \ -\sum_{i \in \mathcal{N}} \frac{E_i}{\eta_i^c} \log\left(\frac{e_{i,t-1}}{E_i} + \frac{\eta_i^c p_{i,t}^c \Delta_t}{E_i}\right) \tag{26a}$$

$$\text{s.t. } 0 \le p_{i,t}^c \le \hat{P}_{i,t}^c, \forall i \tag{26b}$$

$$p_{ag,t}^c = \sum_{i \in \mathcal{N}} p_{i,t}^c \le D_t^c \tag{26c}$$

where $\hat{P}_{i,t}^c = \min\{P_i, (E_i - e_{i,t-1})/(\eta_i^c \Delta_t)\}$. The objective function (26a) endeavors to maintain consistent storage levels among all ESUs. Constraint (26b) prescribes the available capacity of each small ESU; constraint (26c) manifests that partial response is allowed but over-response is prohibited. In problem (26), all parameters are known, and no future dispatch order is involved.

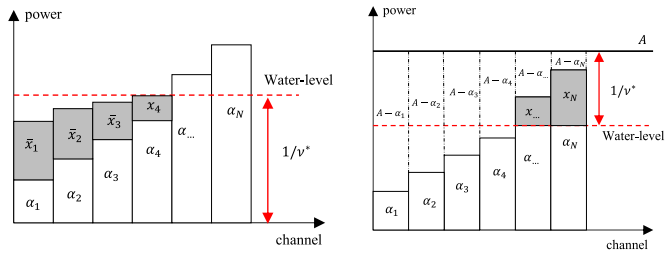
2) Disaggregation of Discharging Power: If the aggregator receives a charging order D_t^d, we solve

$$\min \ -\sum_{i \in \mathcal{N}} E_i \eta_i^d \log\left(1 - \frac{e_{i,t-1}}{E_i} + \frac{p_{i,t}^d \Delta_t}{\eta_i^d E_i}\right) \tag{27a}$$

$$\text{s.t. } 0 \le p_{i,t}^d \le \hat{P}_{i,t}^d, \forall i \tag{27b}$$

$$p_{ag,t}^d = \sum_{i \in \mathcal{N}} p_{i,t}^d \le D_t^d \tag{27c}$$

where $\hat{P}_{i,t}^d = \min\{P_i, (e_{i,t-1})\eta_i^d/\Delta_t\}$. The objective function and constraints have similar interpretations as those in (26).



(a) Water-pouring (b) Water-draining

Fig. 2. Illustration of capacity-aware water-filling method.

Algorithm 1: Capacity-Aware Water-Filling Algorithm.

1: Calculate $1/(\alpha_i + \bar{x}_i)$ and $1/\alpha_i$ for all i. Merge two points with the same value. Denote the total amount of break points as L. The sorted sequence is $\{\nu_1, \nu_2, \ldots, \nu_L\}$. Divide $\nu \in \mathbb{R}$ into $L+1$ intervals.
2: **For** Each interval $[\nu_j, \nu_{j+1}]$
3: Calculate $x_i^*(\nu)$ according to (22)/(25).
4: Solve out ν^* based on equation (23).
5: **if** $\nu^* \in [\nu_j, \nu_{j+1}]$ **then**
6: Report ν^* and $x_i^*(\nu)$; terminate.
7: **else**
8: Update $j = j + 1$ and return to line 3.
9: **end if**
10: **end for**

TABLE III
CORRESPONDENCE IN THE CAWF PROBLEMS

Type		State	Action	Boundary	Water-level
Charging	(19)	α_i	x_i	$[0,\ \bar{x}_i]$	$1/\nu$
	(26)	$e_{i,t-1}/E_i$	$p^c_{i,t}$	$[0,\ \hat{P}^c_{i,t}]$	SoC
Discharging	(24)	$1-\alpha_i$	x_i	$[0,\ \bar{x}_i]$	$1/\nu$
	(27)	$(E_i-e_{i,t-1})/E_i$	$p^d_{i,t}$	$[0,\ \hat{P}^d_{i,t}]$	SoC

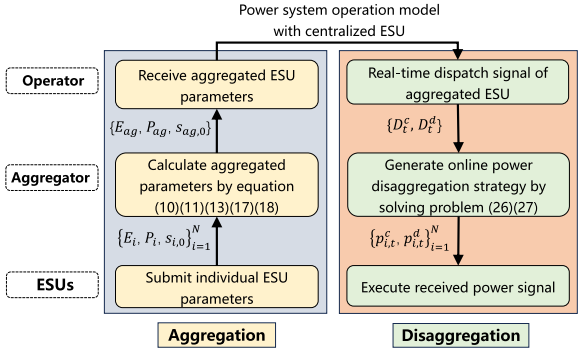


Fig. 3. Flowchart of proposed ESUs aggregation and disaggregation methods.

The relations between online power disaggregation problems (26)/(27) and problems (19)/(24) are summarized in Table III, where $A = 1$ is the unified upper bound. Based on the previous discussions, problems (26) and (27) can be solved by Algorithm 1.

V. HOLISTIC FRAMEWORK

The proposed method is schematically shown in Fig. 3. At the bottom layer, the individual ESU owners submit individual parameters to the aggregator; at the middle layer, the aggregator calculates parameters through (10), (11), (13), (17), and (18) and then submits them to the power system operator at the top layer. This constitutes the aggregation process. For the disaggregation control, the operator treats the aggregator as a large centralized ESU and sends real-time dispatch signal to the aggregator; then the aggregator allocates the dispatch command among small ESUs following the online disaggregation policy obtained from the solution of problems (26) and (27); finally, individual ESUs execute the control signal from the aggregator.

VI. CASE STUDY

Numerical simulation is conducted to verify the performance of the proposed method. The aggregation method is independent of the application scenario, while the disaggregation relies on the concrete implementation of problem (9), and is applied in wind power ramping alleviation and frequency regulation provision. Simulation is done on a laptop with Intel i5-1130G7 CPU and 16 GB RAM. The Minkowski sum is implemented by *Multi-Parametric Toolbox* [38].

TABLE IV
CASE 1: MASSIVE ESUS WITH SAME PARAMETERS

Type	E_i/MWh	P_i/MW	T_i/h	Number
1	0.1	0.05	2.0	50
2	0.2	0.1	2.0	25
3	0.5	0.25	2.0	50
4	0.8	0.4	2.0	25

TABLE V
CASE 1: AGGREGATION RESULTS UNDER DIFFERENT CLUSTERS

Cluster	E_{ag}	P^{ana}_{ag}	P^{MS}_{ag}	P^{Σ}_{ag}	LoF
A = {1,2}	10	5	5	5	0
B = {3,4}	45	22.5	22.5	22.5	0
C = {1,2,3,4}	55	27.5	27.5	27.5	0

TABLE VI
CASE 2: MASSIVE ESUS WITH DIFFERENT PARAMETERS

Type	E_i/MWh	P_i/MW	T_i/h	η_i	Number
1	0.5	0.2	2.5	0.95	10
2	1.2	0.4	3.0	0.90	10
3	1.4	0.5	2.8	0.95	10
4	2.0	0.8	2.5	0.90	10
5	0.2	0.3	0.67	0.95	10
6	0.3	0.4	0.75	0.975	10
7	0.25	0.5	0.5	0.95	10
8	0.4	0.1	4	0.92	10
9	1.5	0.3	5	0.96	10
10	1.2	0.2	6	0.94	10

A. Results of Capacity Aggregation

We test our analytical capacity aggregation method under different cases with various capacity parameter settings of massive distributed ESUs. Besides, the proposed analytical method is compared with Minkowski sum based approach [21] and direct summation method [26]. The aggregated power capacity derived from different methods is noted as P^{ana}_{ag}, P^{MS}_{ag} and P^{Σ}_{ag}, respectively.

1) Case 1: Clustering ESUs With Same Rated Charging Time: A simple case where 4 types and 150 distributed ESUs have same T_i is simulated; parameters are given in Table IV. For brevity, we omit the efficiency term and focus on the capacity term. The aggregation results are summarized in Table V. Clearly, aggregate energy (power) capacity is equal to the sum of energy (power) capacities of individual ESUs, and the aggregation process has no flexibility loss, leading to LoF $= 0$.

2) Case 2: Clustering ESUs With Different Parameters: A case where 10 types and 100 distributed ESUs with parameters in Table VI is studied. We first categorize 10 types of ESUs into 3 clusters (A, B and C) based on their T_i values. ESUs in each cluster share a close T_i value. The aggregation results are summarized in Table VII. It is apparent that $E_{ag} = \sum_i E_i$. However, the aggregated power capacity P_{ag} is different. The direct summation method gives the largest aggregate capacity value $P^{\Sigma}_{ag} = \sum_i P_i$, but this capacity may degrade after some

TABLE VII
CASE 2: AGGREGATION RESULTS UNDER DIFFERENT CLUSTERS

Cluster	E_{ag}	P_{ag}^{ana}	P_{ag}^{MS}	P_{ag}^{Σ}	η_{ag}	LoF
A = {1,2,3,4}	51	17	17	19	0.92	0.1037
B = {5,6,7}	7.5	10	10	12	0.96	0.1608
C = {8,9,10}	31	5.2	5.2	6	0.95	0.1388
{A+B}	58.5	19.5	19.5	31	0.92	0.2290
{A+C}	82	13.7	13.7	25	0.93	0.3485
{B+C}	38.5	6.4	6.4	18	0.95	0.2817
{A+B+C}	89.5	14.9	14.9	37	0.93	0.3591

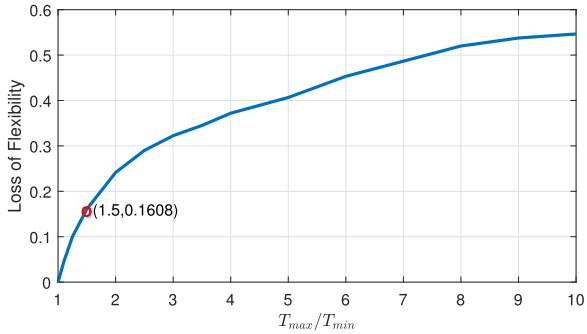


Fig. 4. Loss of flexibility under different γ.

ESUs are fully charged or depleted, which violates the basic requirement for aggregation. The results of proposed method and Minkowski sum method are identical and smaller than P_{ag}^{Σ}, validating the analytical expression in (13).

It is also observed that aggregating all ESUs together leads to the largest LoF as high as 0.3591. In contrast, aggregating ESUs in each cluster can enjoy higher flexibility, reflected by smaller values of LoF ranging between 0.1 and 0.16. The combination of two clusters will get a medium level of flexibility. LoF largely depends on the difference among the ratios T_i of individual ESUs. Therefore, the aggregator should cluster ESUs with similar T_i values to avoid unnecessary loss of flexibility.

Define $T_{\min} = \min_i\{T_i\}$ and $T_{\max} = \max_i\{T_i\}$. We choose 100 ESUs whose T_i is a uniform distribution on the interval $[T_{\min}, T_{\max}]$. We increase $T_{\max}/T_{\min}$ value from 1∼5 and record the corresponding LoF value. The simulation result is plotted in Fig. 4. It can be found that the LoF grows with the increase of $T_{\max}/T_{\min}$. To this end, a recommended principle is

$$\max_i\{T_i\}/\min_i\{T_i\} \le \gamma \tag{28}$$

where γ is a threshold value. A higher value of γ will include a larger number of ESUs in a cluster and incur a higher LoF. In Fig. 4, when $\gamma = 1.5$, the LoF has an acceptable value of 0.1608. So we recommend $\gamma = 1.5$. Based on the above criterion, all ESUs can be categorized into several groups. The group with a smaller average value of T_i can be used to provide frequency support; the group with a larger average value of T_i can be used for energy arbitrage.

B. Application in Wind Power Ramping Alleviation

1) Wind Power Ramping Event: Large fluctuations of wind energy output within a short period, known as wind power ramping event, would threaten the secure operation of power system due to the lack of ramping capacity. A wind ramping event happens when the change of wind output during a certain time period exceeds a threshold. According to the definition in [39], we define upward/downward ramping events as follows:

$$\begin{cases} w_t^{up} = w_t - w_{t-1}, w_t^{dn} = 0 & \text{if } w_t - w_{t-1} > W^{up} \\ w_t^{up} = 0, w_t^{dn} = w_{t-1} - w_t & \text{if } w_{t-1} - w_t > W^{dn} \end{cases}$$

where w_t denotes the wind power output in period t; w_t^{up}/w_t^{dn} represents the upward/downward ramping value; W^{up}/W^{dn} represents the upward/downward ramping threshold. In this test, the duration of each time period is $\Delta_t = 15$ minutes. W^{up}/W^{dn} is set as 10% of the installed wind capacity.

Traditionally, the penetration of renewable energy in power system is moderate, and there are plenty of thermal generation units which are controllable. The volatility of renewable output is compensated by re-dispatching thermal generators. If the penetration of wind power reaches a very high level, thermal generation capacity is replaced by renewable plants and ramping capacity from thermal units becomes insufficient. In an upward ramping event, the excessive wind power may be curtailed; in a downward ramping event, some load may be shed. We assume the power system operator could dispatch an ESU aggregator to alleviate wind ramping event, and the dispatch order is:

$$\{D_t^c, D_t^d\} = \begin{cases} \text{upward}: & D_t^c = w_t^{up}, \quad D_t^d = 0 \\ \text{downward}: & D_t^c = 0, \quad D_t^d = w_t^{dn} \end{cases} \tag{29}$$

In the upward (downward) scenario, the operator sends charging (discharging) signal D_t^c (D_t^d) to the aggregator. Once receiving D_t^c/D_t^d, the aggregator allocates the required power into each ESU. The goal of the aggregator is to track D_t^c/D_t^d as close as possible across all periods $\{1, \ldots, T\}$. So the wind ramping event alleviation problem can be formulated as problem (9) in Section II-C.

In fact, the power grid does not have to completely absorb wind power variation, which means the dispatch signal D_t^c could have a smaller value than w_t^{up}. Equation (29) is just used for testing performance. The real challenge in practice is that in period t, the aggregator has no information on the future dispatch signals $D_{t+1}^c, D_{t+2}^c, \ldots$. The proposed CAWF policy is applied to this problem.

2) Simulation Results: We collect real data of a wind farm in Inner Mongolia, China, whose installed capacity is 100 MW. The time series of wind power output was recorded every 15 minutes. The time series over 3 days with 288 points is shown in Fig. 5. High magnitude upward / downward ramping events are observed. Based on the particular definition of wind power ramping event and the compensation rule in (29), the dispatch order sequence is also given in Fig. 5.

To alleviate the impact of wind power ramping, we assume the aggregator manages group A storage, which consists of type {1, 2, 3, 4} ESUs in Table VI, to compensate for rapid changes in wind power output. As in Table VII, the aggregate

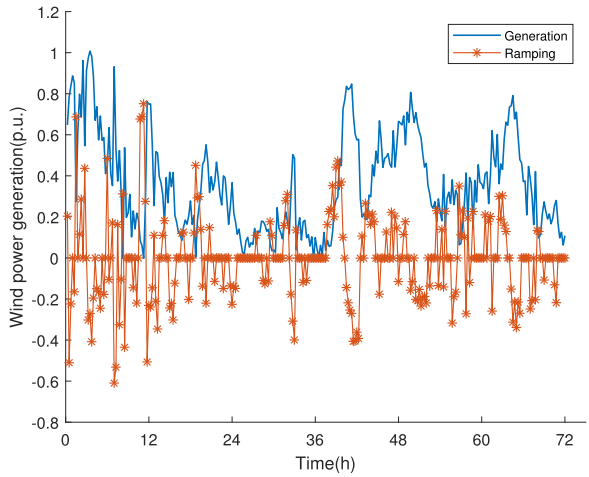


Fig. 5. Wind power generation and wind ramping event.

capacity parameters are $P_{ag} = 17$ MW, $E_{ag} = 51$ MWh, while $P_{ag}^{\Sigma} = 19$ MW. In this case, the dispatch order given by (29) may be greater than P_{ag}, and should be tracked to handle wind ramping event. Particularly, we assume that all ESUs are half charged with an identical storage level of 0.5. For example, if the initial storage levels differ significantly, we can charge some batteries using remaining ones, which is technically viable. Nevertheless, this is not a practical limitation as the initial value is just a parameter in problem (9); any initial value can be used. In the next case, we will assume different initial storage levels.

We compare the performance of the following four policies in alleviating wind power ramping event:

1) *Offline (hindsight) policy:* Provided with the entire sequence of dispatch orders, problem (9) gives the optimal offline solution, which offers a hindsight policy. The hindsight optimum is an ideal result, which serves as a benchmark for comparison but cannot be implemented in practice.
2) The proposed CAWF policy.
3) *MPC policy:* We investigate the performance of MPC under different look-ahead time windows ranging from 4–16 time slots.
4) *Priority control policy:* An intuitive online control policy for multiple ESUs is to charge / discharge them one by one. In the priority control policy, the ESUs are dispatched according to the following rule: The ESU with the highest efficiency will be charged / discharged with the highest priority unless its storage limit is reached. For ESUs with the same efficiency, those with a larger energy capacity will be dispatched first, until the dispatch order is satisfied. A clear disadvantage of this policy is that the utilization rate of each ESU is different. The ESU with a large energy capacity and high efficiency will be dispatched much more frequently compared to the small and inefficient ones. From a long-term perspective, such a policy may harm the health of some high-performance ESUs. However, this long-term effect is not considered in this test.

TABLE VIII
MC SIMULATION RESULTS OF DIFFERENT CONTROL POLICIES

Policy	Optimality gap
Hindsight	benchmark
CAWF	0.2524%
4-slot MPC	0.4303%
8-slot MPC	0.2954%
12-slot MPC	0.2631%
16-slot MPC	0.2516%
Priority Control	0.5358%

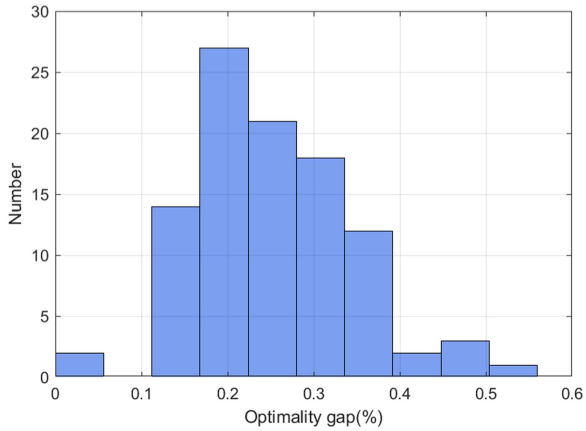


Fig. 6. Optimality gaps of CAWF in MC simulation.

We refer an optimality gap is the difference in the optima between the hindsight policy and a particular policy. The optimality gap clearly shows how the results corresponding to a given policy deviate from the ideal optimum. Monte Carlo (MC) simulations are conducted under different control policies. We randomly choose 100 time series for the tests, each consisting of 288 points with $\Delta_t = 15$ minutes. The MC simulation results are summarized in Table VIII.

The hindsight optimum gives the benchmark of optimality. The average optimality gap of the proposed CAWF policy is only 0.2524% and performs much better than the priority control with 0.5358%. The MPC method is also investigated under different look-ahead time windows. Clearly, the optimality gap decreases when the length of prediction window grows. However, even with exact wind power forecasts in the next 3 hours (12 time periods), MPC is still worse than CAWF in terms of optimality gap. To get a better performance, MPC needs at least exact forecasts in the next 4 hours (16 periods). Such a requirement is still a challenging task even for ultra-short-term wind power forecast, because prediction error can be reduced only if special spatial and temporal correlation patterns in a large area are utilized in the forecast method.

The optimality gaps of the proposed CAWF policy in 100 simulation scenarios are shown in Fig. 6. The maximum optimality gap is only 0.5574%. From Fig. 6, the gap is less than 0.2% (0.4%) in approximately 71% (95%) of the simulated scenarios. Given the 0.2524% average value and distribution

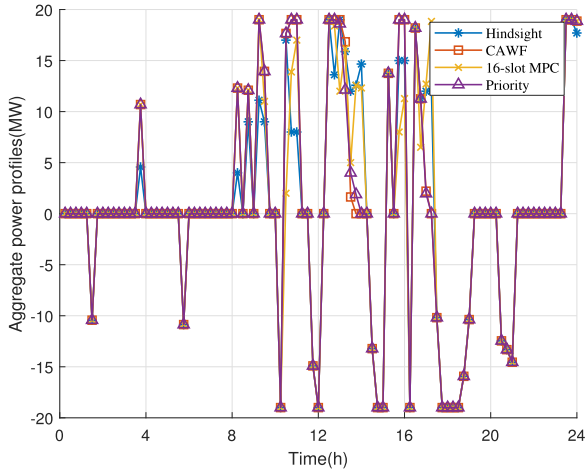
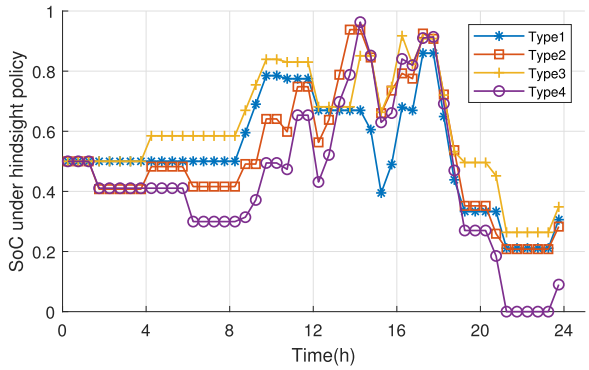
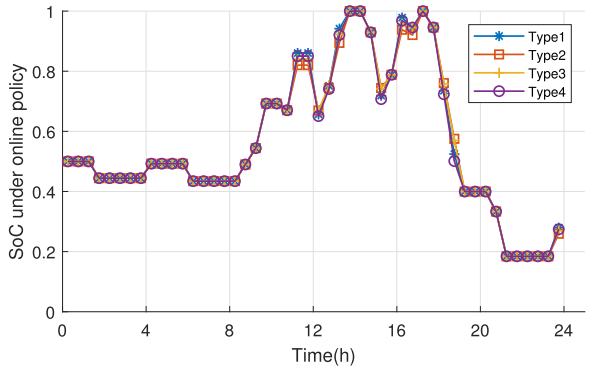


Fig. 7. Simulation results of aggregate power profiles under different policies.



(a) SoC dynamic under hindsight policy



(b) SoC dynamic under online policy

Fig. 8. Simulation results of SoC dynamic.

of the optimality gap, the CAWF policy is competitive and attractive in practical applications because it is prediction-free.

We randomly select a sequence in the MC simulation. In Fig. 7, the aggregated charging / discharging power trajectories (in one day) are plotted. As the ramping power may exceed P_{ag}, the aggregator responds to the dispatch order with its maximum power, which is 19 MW, during periods $10 \sim 18$. The hindsight policy obtains the optimal power profiles for coping with this ramping event. The aggregate response trajectories under 4 control policies differ in some periods. Storage levels of each type of ESU under hindsight and CAWF policies are shown in Fig. 8. Compared with the hindsight result which can utilize full future information, CAWF policy prioritizes SoC consistency, but this does not mean that SoCs must be the same in all periods, especially when the dispatch order exceeds P_{ag} and initial SoCs exhibit significant difference. For example, in period 10, the SoCs are different because $D^c_{10} > P_{ag}$, and keeping SoCs identical will jeopardize the response capability.

We further increase the number and difference of ESUs. We choose 200 ESUs whose T_i obey a uniform distribution in the interval [2,3]; the power capacities uniformly range from [0.05,0.5] MW; initial SoCs follow a uniform distribution in [0.1,0.9]. The aggregated capacities are $P_{ag} = 48.36$ MW and $E_{ag} = 145.08$ MWh. We randomly choose 100 time series for the MC simulation, each consisting of 288 points with $\Delta_t = 15$ minutes. The optimality gaps of the proposed CAWF policy in 100 simulation scenarios are shown in Fig. 9. The average optimality gap is 0.2619%; the maximum optimality gap is 0.7293%; the gap is less than 0.3% (0.5%) in approximately 61% (93%) of the simulated scenarios. Given the 0.2619% average value and distribution of the optimality gap, the CAWF policy is still performing well in large-scale application.

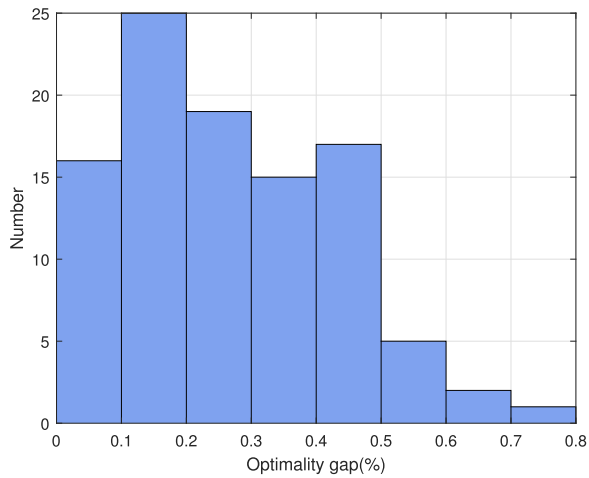


Fig. 9. Optimality gaps of CAWF in largescale MC simulation.

C. Application in Frequency Regulation

Energy storage is expected to provide frequency regulation service because, in future power system, thermal generation capacity will be limited. We assume that the aggregator who manages a cluster of small ESUs has adequate capacity to enter the frequency regulation market.

1) PJM Market for Frequency Regulation Service: In PJM performance-based regulation mechanism [40], the total revenue R^{reg} of frequency regulation service consists of a capacity revenue R^{cap} and a performance revenue R^{perf}, which can be

expressed as follows:

$$R^{cap} = \lambda^{cap} P^r K^{perf}$$
$$R^{perf} = \lambda^{perf} M P^r K^{perf}$$
$$R^{reg} = R^{cap} + R^{perf} = (\lambda^{cap} + \lambda^{perf} M) P^r K^{perf}$$

where P^r is the contracted regulation power capacity with the system operator; λ^{cap} and λ^{perf} are capacity clearing price and performance clearing price, respectively. Due to the relatively small scale, we assume that the aggregator is a price taker; M is mileage ratio between the mileage of the dynamic regulation signal (RegD) and the traditional regulation signal (RegA), whose value is approximately 3 and fixed for each hour [41]; K^{perf} is the performance score during a certain period, which is calculated as the relative error between the assigned AGC signal $P^r \boldsymbol{r}$ and actual response power $\boldsymbol{p}$ [42]:

$$K^{perf} = 1 - \frac{\|P^r \boldsymbol{r} - \boldsymbol{p}\|_1}{P^r \|\bar{\boldsymbol{r}}\|_1}$$

where the normalized AGC signal $\boldsymbol{r} = \{r_t \in [-1, 1]\}$, and $\|P^r \boldsymbol{r} - \boldsymbol{p}\|_1$ represents the absolute error summation of the response. $\bar{\boldsymbol{r}}$ is the average value of AGC signal based on the historical signal. Finally, the R^{reg} can be expressed as

$$R^{reg} = (\lambda^{cap} + \lambda^{perf} M) P^r \left(1 - \frac{\|P^r \boldsymbol{r} - \boldsymbol{p}\|_1}{P^r \|\bar{\boldsymbol{r}}\|_1}\right)$$

We assume the aggregator submits the regulation capacity $P^r = P_{ag}$. Then, maximizing R^{reg} comes down to maximizing the performance score K^{perf}, in other words, minimizing the relative error $\|P^r \boldsymbol{r} - \boldsymbol{p}\|_1$.

In real-time control, the operator sends AGC signal $P^r r_t$ every 2 seconds to the aggregator, which can be rewritten as the dispatch order D_t^c / D_t^d.

$$\{D_t^c, D_t^d\} = \begin{cases} D_t^c = r_t P_{ag}, D_t^d = 0, & \text{if } r_t > 0 \\ D_t^c = 0, D_t^d = -r_t P_{ag}, & \text{if } r_t < 0 \end{cases}$$

The aggregator responds to the dispatch order with $p_t = p_{ag,t}$ accordingly. Thus the relative error minimization problem can be transformed into optimization problem (9) in Section II-C.

From the perspective of the aggregator, future AGC signals are uncertain and unpredictable. Once the aggregator receives a dispatch signal, it must determine the power control strategy of each small ESU without future information. Thus CAWF policy is applied to this problem.

2) Simulation Results: The AGC signals are collected from the PJM market with a time resolution of 2 seconds. The RegD signal requires net zero energy over a certain time horizon. In the following tests, the time span of numerical simulation is 2 hours. The aggregator manages type {5, 6, 7} ESUs in Table VI. As listed in Table VII, the aggregator submits $P^r = P_{ag} = 10$ MW to the market, and the total energy capacity is $E_{ag} = 7.5$ MWh. According to the rule of frequency regulation, $r_t \in [-1, 1]$, so the dispatch order can never exceed P_{ag}.

We compare the performance of following four policies:

1) Offline (hindsight) policy.
2) The proposed CAWF policy.
3) 1-minute exact look-ahead MPC: assuming the aggregator can exactly predict the next 30 AGC commands.
4) Priority control policy.

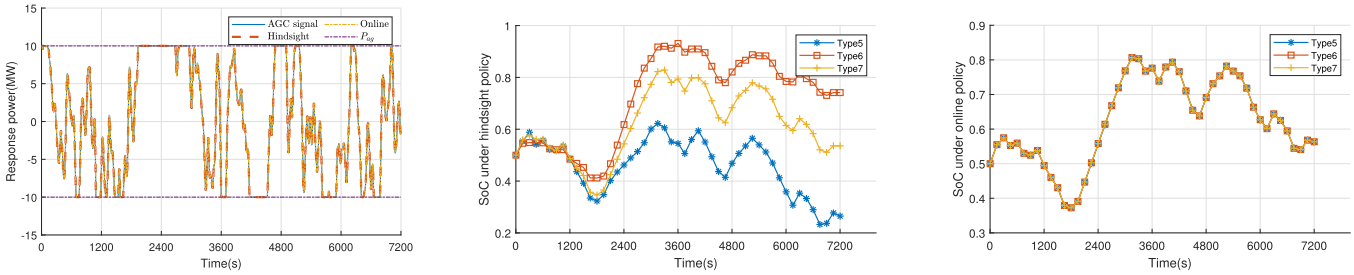
TABLE IX
RESULTS OF 4 POLICIES UNDER DIFFERENT INITIAL SOC LEVELS

Initial SoC	Policy	K^{perf}	Optimality gap
Case 1	Hindsight	1	
$\sigma^2 = 0$	CAWF	1	
$\{0.5, 0.5, 0.5\}$	1-min MPC	1	
	Priority Control	1	
Case 2	Hindsight	1	
$\sigma^2 = 0.0561$	CAWF	1	
$\{0.8, 0.4, 0.38\}$	1-min MPC	0.9958	0.4236%
	Priority Control	1	
Case 3	Hindsight	1	
$\sigma^2 = 0.1092$	CAWF	1	
$\{0.5, 0.8, 0.14\}$	1-min MPC	1	
	Priority Control	1	
Case 4	Hindsight	1	
$\sigma^2 = 0.1685$	CAWF	1	
$\{0.1, 0.9, 0.34\}$	1-min MPC	0.9898	1.0194%
	Priority Control	0.9663	3.3700%
Case 5	Hindsight	1	
$\sigma^2 = 0.1941$	CAWF	1	
$\{0.5, 0.1, 0.98\}$	1-min MPC	0.9868	1.3237%
	Priority Control	0.9910	0.9045%
Case 6	Hindsight	0.9902	
$\sigma^2 = 0.2633$	CAWF	0.9902	
$\{0, 1, 0.3\}$	1-min MPC	0.9761	1.4197%
	Priority Control	0.9603	3.0199%

We consider different initial storage levels in this test. We still assume the aggregated initial SoC is $s_{ag,0} = 0.5$. We use variance σ^2 to measure the degree of individual $s_{i,0}$ deviating from $s_{ag,0}$. We compare four policies under different values of σ^2. Results are listed in Table IX.

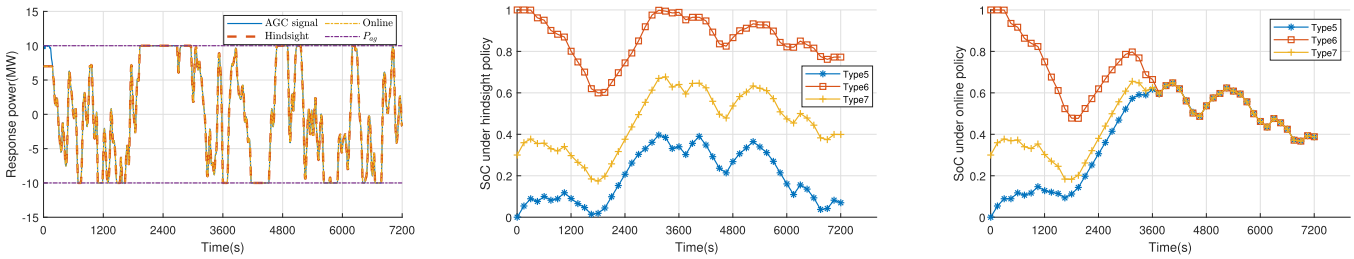
2) Scenario 1: Identical initial SoCs: In this case, all ESUs have the same initial SoCs, so $\sigma^2 = 0, s_{i,0} = s_{ag,0} = 0.5$, which is a perfect initial state. Results in Table IX show that all policies achieve the same and also the highest performance score, which is $K^{perf} = 1$. The aggregated response power profiles and storage level dynamics are shown in Fig. 10. Since AGC signal has a zero mean and $E_{ag}/P_{ag} = 45$ minutes, the energy capacity is sufficient, and the aggregate SoC will never reach the lower or upper limits. Moreover, because the AGC signal cannot exceed the power capacity P_{ag} as required and the initial storage levels are identical, the aggregator can always track the AGC signal accurately to obtain the highest performance score. Compared with the hindsight results in Fig. 10(b), the aggregator can keep individual storage levels much closer under the CAWF policy, as in Fig. 10(c).

2) Scenario 2: Small discrepancy in the initial storage levels: See Case 2–5 in Table IX. We gradually increase the discrepancy of initial storage levels of individual ESUs, causing σ^2 to change from 0.0561 to 0.1941. Under such settings, the performance score of hindsight policy is always equal to 1, which means that



(a) The aggregate response power. (b) SoC dynamic under hindsight policy (c) SoC dynamic under CAWF

Fig. 10. Simulation results of hindsight policy and CAWF with $\sigma^2 = 0$.



(a) The aggregate response power. (b) SoC dynamic under hindsight policy (c) SoC dynamic under CAWF

Fig. 11. Simulation results of hindsight policy and CAWF with $\sigma^2 = 0.2633$.

if the entire sequence of AGC signal is known in advance, the aggregator is still able to faithfully respond to the dispatch order through a comprehensive control of individual ESUs. However, in practice, AGC signal is hardly predictable. The CAWF policy could achieve the same score as the hindsight policy without requiring future AGC signals. MPC fails to get a full performance score in Case 2, Case 4 and Case 5, and the priority control fails to get a full performance score in Case 4 and Case 5. But it is difficult to say which one of MPC and priority control is better. For example, in Case 4, Type 5 ESUs are nearly empty while Type 6 ESUs are almost fully charged; without proper coordination, the empty (full) ESUs can no longer be discharged (charged), thus the regulation capability decreases. Under priority control, Type 6 ESUs with the largest efficiency η is dispatched first; given the initial storage level, Type 6 ESUs quickly become fully charged and cannot store more energy, leading to a poor performance score.

2) Scenario 3: Large discrepancy in the initial storage levels: See Case 6 in Table IX. This is an extreme case, in which $\sigma^2 = 0.2633$. The aggregated response power profiles and storage level dynamics are shown in Fig. 11. At the beginning, some ESUs are empty while others are fully charged, inevitably restricting the capability of charging and discharging. Despite knowing the exact sequence of AGC signals, the hindsight policy cannot perfectly follow dispatch signals. The theoretically optimal bound of performance score is $K^{perf} = 0.9902$. Even under such an extreme condition, CAWF policy could still receive the optimal performance score. In contrast, the optimality gaps of MPC and priority control are as large as 1.4197% and 3.0199%, respectively. As shown in Fig. 11(c),

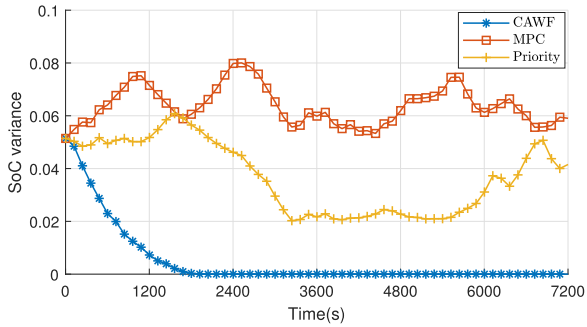


Fig. 12. SoC variance for 2-hour regulation service.

the aggregator aims to keep individual storage levels as close as possible under the CAWF policy. In this way, the charging/discharging flexibility of ESU cluster is preserved for future usage.

We further increase the number and difference of ESUs. We choose 100 ESUs whose T_i obey a uniform distribution on the interval [0.5, 0.75]; the power capacities uniformly range from 0.05 MW to 0.5 MW; the initial SoC is randomly chosen from the interval [0.05, 0.95]. Then, the aggregation results are $P_{ag} = 24.19$ MW and $E_{ag} = 18.14$ MWh. The aggregator submits $P^r = P_{ag}$ and receives frequency regulation signal within $[-P_{ag}, P_{ag}]$. We randomly select 100 AGC signal sequences with 2-hour time span. The optimality gaps of 100 scenarios are equal to zero. Additionally, we plot the SoC variance of 100 ESUs under different policies in Fig. 12. It can be observed that CAWF converges to zero variance, which means that all

SoCs converge to a consistent value. In this way, the aggregated flexibility of ESU cluster is preserved.

In summary, the above case studies show that the proposed CAWF policy can achieve high performance scores in frequency regulation service provision. It outperforms MPC and priority control, exhibiting a promising potential in future frequency regulation markets.

VII. Conclusion

This paper studies the aggregated use of small energy storage units. An analytical expression is derived to determine a reasonable pair of power and energy capacities of the storage cluster, which is shown to be consistent compared to the Minkowski-sum approach. The CAWF policy is developed for real-time power disaggregation control of individual storage units without forecasts on dispatch signals. The key findings in the case studies include:

1) If the E/P values of managed ESUs differ significantly, the aggregator should cluster ESUs into several groups. Each group includes ESUs with close E/P values to avoid loss of flexibility. A recommended criterion for clustering is $\max_{i\in\mathcal{G}}\{T_i\}/\min_i\{T_i\} \leq 1.5$, where $\mathcal{G}$ is the set of ESUs in the cluster. The group with a smaller average value of T_i can be used to provide frequency support; the group with a larger average value of T_i can be used for energy arbitrage.
2) For the task of wind power ramping alleviation, the CAWF and MPC policies perform better than priority control, and CAWF outperforms 12-slot exact lookahead MPC. To get a better performance, MPC needs exact forecasts in the next 4 hours.
3) For frequency regulation provision, if small ESUs have nearly consistent initial SoCs around 0.5, then CAWF, MPC and priority control yield the optimal performance. Otherwise, if the initial SoCs exhibit larger discrepancy, CAWF still achieves optimal performance, while MPC and priority control cannot. In view of the fact that predicting AGC signals is difficult and priority control uses some batteries more frequently, the proposed CAWF is attractive in practical applications. Another salient feature is that CAWF leads to the similar utilization rate of individual ESUs, while other policies do not have such a guarantee, especially priority control.

In the future work, we will consider more flexible resources such as virtual power plants, deferrable loads, thermostatically controlled loads, and so on. We will develop more general online policies that can also utilize imperfect forecasts.

References

[1] S. Wang, N. Zheng, C. D. Bothwell, Q. Xu, S. Kasina, and B. F. Hobbs, "Crediting variable renewable energy and energy storage in capacity markets: Effects of unit commitment and storage operation," *IEEE Trans. Power Syst.*, vol. 37, no. 1, pp. 617–628, Jan. 2022.

[2] Z. Guo, W. Wei, L. Chen, M. Shahidehpour, and S. Mei, "Economic value of energy storages in unit commitment with renewables and its implication on storage sizing," *IEEE Trans. Sustain. Energy*, vol. 12, no. 4, pp. 2219–2229, Oct. 2021.

[3] Y. Zhou, Q. Zhai, and L. Wu, "Multistage transmission-constrained unit commitment with renewable energy and energy storage: Implicit and explicit decision methods," *IEEE Trans. Sustain. Energy*, vol. 12, no. 2, pp. 1032–1043, Apr. 2021.

[4] N. Gu, H. Wang, J. Zhang, and C. Wu, "Bridging chance-constrained and robust optimization in an emission-aware economic dispatch with energy storage," *IEEE Trans. Power Syst.*, vol. 37, no. 2, pp. 1078–1090, Mar. 2022.

[5] A. Azizivahed et al., "Risk-oriented multi-area economic dispatch solution with high penetration of wind power generation and compressed air energy storage system," *IEEE Trans. Sustain. Energy*, vol. 11, no. 3, pp. 1569–1578, Jul. 2020.

[6] N. Zhang et al., "Aggregating distributed energy storage: Cloud-based flexibility services from China," *IEEE Power Energy Mag.*, vol. 19, no. 4, pp. 63–73, Jul./Aug. 2021.

[7] G. C. Gissey, D. Subkhankulova, P. E. Dodds, and M. Barrett, "Value of energy storage aggregation to the electricity system," *Energy Policy*, vol. 128, pp. 685–696, May 2019.

[8] J. E. Contreras-Ocana, M. A. Ortega-Vazquez, and B. Zhang, "Participation of an energy storage aggregator in electricity markets," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1171–1183, Mar. 2019.

[9] B. Vatandoust, A. Ahmadian, M. A. Golkar, A. Elkamel, A. Almansoori, and M. Ghaljehei, "Risk-averse optimal bidding of electric vehicles and energy storage aggregator in day-ahead frequency regulation market," *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 2036–2047, May 2019.

[10] C. Goebel and H.-A. Jacobsen, "Bringing distributed energy storage to market," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 173–186, Jan. 2016.

[11] C. Eid, P. Codani, Y. Perez, J. Reneses, and R. Hakvoort, "Managing electric flexibility from distributed energy resources: A review of incentives for market design," *Renewable Sustain. Energy Rev.*, vol. 64, pp. 237–247, Oct. 2016.

[12] M. I. Alizadeh, M. P. Moghaddam, N. Amjady, P. Siano, and M. K. Sheikh-El-Eslami, "Flexibility in future power systems with high renewable penetration: A review," *Renewable Sustain. Energy Rev.*, vol. 57, pp. 1186–1193, May 2016.

[13] M. Obi, T. Slay, and R. Bass, "Distributed energy resource aggregation using customer-owned equipment: A review of literature and standards," *Energy Rep.*, vol. 6, pp. 2358–2369, Nov. 2020.

[14] W. Mai and C. Chung, "Economic MPC of aggregating commercial buildings for providing flexible power reserve," *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2685–2694, Sep. 2015.

[15] X. Chen, E. Dall'Anese, C. Zhao, and N. Li, "Aggregate power flexibility in unbalanced distribution systems," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 258–269, Jan. 2020.

[16] S. Feng, W. Wei, and Y. Chen, "Day-ahead scheduling and online dispatch of energy hubs: A flexibility envelope approach," *IEEE Trans. Smart Grid*, 2023, early access, doi: 10.1109/TSG.2023.3337629.

[17] B. Cui, A. Zamzam, and A. Bernstein, "Network-cognizant time-coupled aggregate flexibility of distribution systems under uncertainties," *IEEE Control Syst. Lett.*, vol. 5, no. 5, pp. 1723–1728, Nov. 2021.

[18] S. Wang and W. Wu, "Aggregate flexibility of virtual power plants with temporal coupling constraints," *IEEE Trans. Smart Grid*, vol. 12, no. 6, pp. 5043–5051, Nov. 2021.

[19] H. R. Tiwary, "On the hardness of computing intersection, union and Minkowski sum of polytopes," *Discrete Comput. Geometry*, vol. 40, pp. 469–479, Jul. 2008.

[20] H. Hao, B. M. Sanandaji, K. Poolla, and T. L. Vincent, "Aggregate flexibility of thermostatically controlled loads," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 189–198, Jan. 2015.

[21] L. Zhao, W. Zhang, H. Hao, and K. Kalsi, "A geometric approach to aggregate flexibility modeling of thermostatically controlled loads," *IEEE Trans. Power Syst.*, vol. 32, no. 6, pp. 4721–4731, Nov. 2017.

[22] L. Zhao, H. Hao, and W. Zhang, "Extracting flexibility of heterogeneous deferrable loads via polytopic projection approximation," in *Proc. IEEE 55th Conf. Decis. Control*, 2016, pp. 6651–6656.

[23] F. L. Müller, J. Szabó, O. Sundström, and J. Lygeros, "Aggregation and disaggregation of energetic flexibility from distributed energy resources," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1205–1214, Mar. 2019.

[24] S. Barot and J. A. Taylor, "A concise, approximate representation of a collection of loads described by polytopes," *Int. J. Elect. Power Energy Syst.*, vol. 84, pp. 55–63, Jan. 2017.

[25] S. Barot and J. A. Taylor, "An outer approximation of the Minkowski sum of convex conic sets with application to demand response," in *Proc. IEEE 55th Conf. Decis. Control*, 2016, pp. 4233–4238.

[26] J. Zhou, X. Chen, Y. Chen, and J. Wen, "Cooperative hierarchical control of isolated microgrids considering energy storage system aggregation," *IEEE Trans. Power Syst.*, vol. 39, no. 1, pp. 850–862, Jan. 2023.
[27] Z. Yi, Y. Xu, W. Gu, and W. Wu, "A multi-time-scale economic scheduling strategy for virtual power plant based on deferrable loads aggregation and disaggregation," *IEEE Trans. Sustain. Energy*, vol. 11, no. 3, pp. 1332–1346, Jul. 2020.
[28] N. Nazir and M. Almassalkhi, "Grid-aware aggregation and realtime disaggregation of distributed energy resources in radial networks," *IEEE Trans. Power Syst.*, vol. 37, no. 3, pp. 1706–1717, May 2022.
[29] Y. Wang et al., "Aggregated energy storage for power system frequency control: A finite-time consensus approach," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3675–3686, Jul. 2019.
[30] Q. Wu, R. Guan, X. Sun, Y. Wang, and X. Li, "SoC balancing strategy for multiple energy storage units with different capacities in islanded microgrids based on droop control," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 6, no. 4, pp. 1932–1941, Dec. 2018.
[31] B. Wang et al., "Consensus-based control of hybrid energy storage system with a cascaded multiport converter in DC microgrids," *IEEE Trans. Sustain. Energy*, vol. 11, no. 4, pp. 2356–2366, Oct. 2020.
[32] A. Oshnoei, M. Kheradmandi, and S. Muyeen, "Robust control scheme for distributed battery energy storage systems in load frequency control," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4781–4791, Nov. 2020.
[33] D. Zhu and Y.-J. A. Zhang, "Optimal coordinated control of multiple battery energy storage systems for primary frequency regulation," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 555–565, Jan. 2019.
[34] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[35] G. Zhang, S. T. Tan, and G. G. Wang, "Real-time smart charging of electric vehicles for demand charge reduction at non-residential sites," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4027–4037, Sep. 2018.
[36] Y. Xiong, B. Wang, C.-C. Chu, and R. Gadh, "Vehicle grid integration for demand response with mixture user model and decentralized optimization," *Appl. Energy*, vol. 231, pp. 481–493, Dec. 2018.
[37] Z. M. Haider, K. K. Mehmood, M. K. Rafique, S. U. Khan, S.-J. Lee, and C.-H. Kim, "Water-filling algorithm based approach for management of responsive residential loads," *J. Mod. Power Syst. Clean Energy*, vol. 6, no. 1, pp. 118–131, Jan. 2018.
[38] M. Herceg, M. Kvasnica, C. N. Jones, and M. Morari, "Multi-parametric toolbox 3.0," in *Proc. Eur. Control Conf.*, 2013, pp. 502–510.
[39] M. Cui, J. Zhang, A. R. Florita, B.-M. Hodge, D. Ke, and Y. Sun, "An optimized swinging door algorithm for identifying wind ramping events," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 150–162, Jan. 2016.
[40] G. He, Q. Chen, C. Kang, P. Pinson, and Q. Xia, "Optimal bidding strategy of battery storage in power markets considering performance-based regulation and battery cycle life," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2359–2367, Sep. 2016.
[41] G. He, Q. Chen, C. Kang, Q. Xia, and K. Poolla, "Cooperation of wind power and battery storage to provide frequency regulation in power markets," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3559–3568, Sep. 2017.
[42] B. Xu, Y. Shi, D. S. Kirschen, and B. Zhang, "Optimal battery participation in frequency regulation markets," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6715–6725, Nov. 2018.

Boshen Zheng (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering and the B.E. degree in economics in 2021 from Tsinghua University, Beijing, China, where he is currently working toward the Ph.D. degree with the Department of Electrical Engineering. His research interests include online optimization, game theory, transactive energy, and electricity market.

Wei Wei (Senior Member, IEEE) received the bachelor's and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2008 and 2013, respectively. From 2013 to 2015, he was a Postdoctoral Research Associate with Tsinghua University. In 2014, he was a Visiting Scholar with Cornell University, Ithaca, NY, USA, and a Visiting Scholar with Harvard University, Cambridge, MA, USA, in 2015. He is currently a Tenured Associate Professor with Tsinghua University. His research interests include applied optimization, computational decision-making, power system operation, and power market.

Yin Xu (Senior Member, IEEE) received the B.E. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2008 and 2013, respectively. During 2013–2016, he was an Assistant Research Professor with the School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA, USA. He is currently a Professor with Beijing Jiaotong University, Beijing. His research interests include power system resilience, distribution system restoration, and power systems electromagnetic transient simulation. Dr. Xu is currently a Secretary of the Distribution Test Feeder Working Group under the IEEE PES Distribution System Analysis Subcommittee.

Yue Chen (Member, IEEE) received the B.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2015, the B.S. degree in economics from Peking University, Beijing, in 2017, and the Ph.D. degree in electrical engineering from Tsinghua University, in 2020. From 2018 to 2019, she was a Visiting Student with the California Institute of Technology, Pasadena, CA, USA. She is also a Vice-Chancellor Assistant Professor with the Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Hong Kong SAR, China. Her research interests include optimization, game theory, mathematical economics, and their applications in smart grid and integrated energy systems.