Highlights

Distributed Privacy-preserving Optimization for Multi-energy System Considering Carbon Emission

Yujie Zhou, Hongxing Ye, Yinyin Ge

- An optimization model substantially reduces carbon emissions at the demand side.
- Address large-scale nonconvex problems.
- Novel techniques accelerate convergence in distributed optimization.
- An information-exchange mechanism improves privacy and computational performance.

Distributed Privacy-preserving Optimization for Multi-energy System Considering Carbon Emission

Yujie Zhou[a], Hongxing Ye[a,*], Yinyin Ge[b]

[a]*School of Automation Science and Engineering, Xi'an Jiaotong University, Xi'an, 710049, Shaanxi, China*
[b]*School of Cyber Science and Engineering, Xi'an Jiaotong University, Xi'an, 710049, Shaanxi, China*

Abstract

Multi-energy systems (MESs) present great potential for reducing operation costs and carbon emissions. Nodal carbon intensity (NCI) can effectively quantify the carbon emissions at the nodal level. The NCI's potential for carbon emission reduction is yet to be unlocked on the demand side. However, the nonconvexity of the NCI function poses a serious computational challenge, and the corresponding large-size problem is unsolvable even for state-of-the-art solvers. To bridge these gaps, this work introduces an optimization model to reduce carbon emissions via demand management and proposes a novel successive approximation algorithm to address its notorious computational challenge. On the other hand, stakeholders in MESs often have different interests. There thus is a strong demand to manage energy in a distributed fashion and to preserve privacy in MES. A novel information-exchange mechanism and acceleration technique for distributed optimization are proposed based on homomorphic encryption (HE), achieving privacy protection and fast computational performance. Case studies validate the effectiveness and performance of the proposed approaches.

Keywords:
Carbon emission flow, distributed optimization, nonconvex constraints, homomorphic encryption, privacy preserving, multi-energy system

1. Introduction

Multi-energy system (MES) refers to a system that synchronizes the generation, distribution, and consumption of multiple energy sources [1]. MES explicitly considers couplings across different energy carriers. It unlocks flexibility for renewable accommodation and reliability enhancement. However, the integration of diverse energy systems, such as electric power systems, natural gas systems, heating systems, presents inherent challenges associated with their fusion and interaction. The concept of Energy Hub (EH) [2] is introduced to address the integration challenges for MES. An EH serves as a unit that provides the functionality of input, output, conversion, and storage for multiple energy carriers [3]. Classical EH model structures are discussed in [4, 5].

*Corresponding author: Hongxing Ye (yehxing@xjtu.edu.cn)

The continuous drive for energy decarbonization has increased interest in carbon optimization for MES. Traditional approaches to reducing carbon emissions in MES typically focus on the supply side, optimizing the generators' carbon emissions [6]. In the meantime, carbon emissions of the demand side reflect the carbon usage of individual demand nodes. The coordinated regulation of supply and demand sides has significant potential to further reduce total carbon emissions. Carbon Emission Flow (CEF) is increasingly recognized for its capability to quantify demand-side carbon emissions using Nodal Carbon Intensity (NCI). The investigation of MES incorporating CEF has attracted notable interest. The CEF in MES research primarily focuses on two key areas: model development and optimization applications. In terms of modeling, [7] introduces the CEF model based on KCL and KVL laws. Additionally, some classical CEF calculation models for MES systems are introduced by [8, 9]. The optimization applications of CEF in MES primarily include: constraining the upper limit of NCI [10], optimal power scheduling based on NCI [11], and minimizing user-end carbon emission costs [12].

The NCI-constrained optimization is a nonconvex problem, posing significant computational challenges. The hardest part in CEF stems from the nonconvex NCI formulation. Current CEF researches largely rely on data-driven methods to linearize NCI [11, 13]. These methods often suffer from a lack of interpretability. [14] proposes a matrix decomposition method aimed at reducing the computational complexity of CEF models. In optimization literature, convexification approximations are often employed to solve the nonconvex problems [15]. A iterative linearization method [16] is proposed to address the nonconvex constraints in natural gas systems. In [17], a penalty convex-concave procedure-based convexification approach is introduced to handle the nonconvex constraints. However, it is hard to find a feasible solution in convexification approaches unless there is a special structure.

On the other hand, there is a strong demand for distributed optimization in MES considering privacy preservation [18]. First, stakeholders for different energy systems possess various interests in MES. In centralized MES optimization, stakeholders submit energy data to a dispatch center. However, accessing certain energy data is not feasible due to privacy concerns. Also, massive operation data of MES brings great challenges to data processing and exchange [19]. Second, MES consists of different energy systems, leading to a complicated and large-sized problem, which can be decomposed into small subproblems in a distributed approach. Distributed optimization methods not only alleviate the overall computational load, but also significantly reduce the scope of data sharing, thereby enhancing information security.

Distributed optimization algorithms are primarily divided into first-order and second-order methods. Second-order methods generally converge more rapidly than first-order ones. Classic algorithms such as NMDA are introduced in [20], where a distributed Newton method based on average consensus theory is presented. Another type of second-order method includes the primal-dual interior point (PDIP) method [21], which is a Newton-based approach. Based on the PDIP method, [22] proposes a distributed quasi-Newton method. In the meantime, second-order methods typically struggle with asynchronous updates, posing extra difficulties in practical implementations. First-order methods, by contrast, offer advantages in handling asynchronous updates. Gradient tracking (GT) methods [23] represent a classic approach in this category. IDGT, an application of GT methods in economic dispatch for power grids, is introduced in [24]. It can rapidly approximate and track the time-varying optimal solution of the dual problem. A well-established

first-order approach is the distributed (sub)gradient descent (DGD) method, which originates various methods, including EXact firsT-ordeR Algorithm (EXTRA) [25] and the renowned Alternating Direction Method of Multipliers (ADMM) algorithm [26]. Substantial research has employed ADMM to address the Optimal Power Flow (OPF) problems [27, 28].

However, a research gap exists in integrating the CEF into the distributed MES approach due to the nonconvex NCI function, as most distributed optimizations rely on convex optimization theory. Existing methods for nonconvex NCI-constrained optimization commonly suffer from slow computation or convergence issues [29, 30]. In the meantime, ADMM-like distributed optimization algorithms may pose a privacy-leak risk due to data exchange. Many improved ADMM-like frameworks have been proposed based on differential privacy (DP) [31] and homomorphic encryption (HE) [32, 33]. However, they may deteriorate the solution quality or speed.

To fulfill these research gaps, this work proposes a novel distributed privacy-preserving approach. The main contribution of this paper is threefold.

1) A novel Successive Dual Taylor-expansion (SDTE) algorithm is proposed to solve the NCI-constrained optimization for MES. The nonconvex NCI function is notorious to address. When the size increases, NCI-constrained problem easily becomes unsolvable by the existing deterministic approaches, even with state-of-the-art solvers. Therefore, how to solve it with deterministic algorithms remains an open question. We fulfill the research gap by proposing a linearized CEF (LCEF) and dual Taylor approximation.
2) This work develops a distributed optimization for NCI-constrained problem, and propose acceleration techniques, i.e., Dual-regularized ADMM (DR-ADMM). When solving nonconvex problems, the conventional ADMM methods, such as LADMM [29], experience oscillation issue and struggle with slow convergence. To address these challenges, this work presents a novel nested iteration framework. In the meantime, a novel dual variable-based regularization term is introduced in the objective function, substantially boosting both solution quality and speed.
3) A novel information-exchange mechanism is proposed to maintain the privacy. To protect the data to be exchanged in distributed optimization, this work proposes a new privacy-preserving framework, i.e., Dynamic-second-order-term Privacy-preserving (DSOP) algorithm, based on Partially Homomorphic Encryption (PHE) to encrypt and decrypt them. It address the performance challenge due to the encryption [32]. In the proposed algorithm, the second-order term in the algorithm is dynamically updated without compromising privacy. It can effective reduce iterations and communication overhead.

The paper is organized as follows. Section 2 introduces the multi-energy system coupled power flow model based on the EH model. Section 3 introduces the SDTE algorithm to optimize energy scheduling in MESs. Section 4 proposes the DR-ADMM framework and privacy-preserving technique. Section 5 tests the proposed approaches comprehensively by case studies. Section 6 concludes this paper. The schematic overview of the methodology in the paper is shown in Fig. 1.

2. Classic Multi-energy System Model

For the self-contained reason, this section presents a classic MES model first. Power system node set is denoted as $\mathcal{N}$ with a total of N nodes, and the line set is denoted as $\mathcal{E}$ with a total of

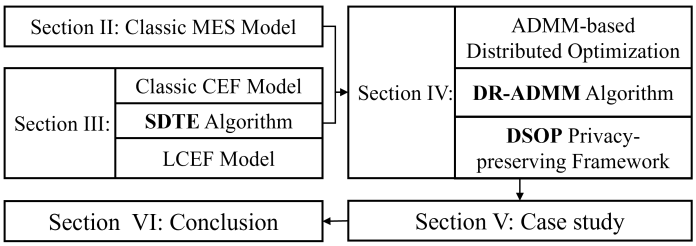


Figure 1: Overview of the scheme and relationship of the main features.

E edges. Let i and j denote the nodes , $\forall i \in \mathcal{N}, \forall (i,j) \in \mathcal{E}$. We formulate the DC optimal power flow (DC-OPF) model that is shown in (1).

$$P_i^{\text{input}} - P_i^{\text{output}} = \sum\nolimits_{j:i\to j} P_{ij} - \sum\nolimits_{j:j\to i} P_{ji}, \forall i \in \mathcal{N} \tag{1a}$$

$$P_{ij} = \left(\theta_i - \theta_j\right)/X_{ij}, \quad \forall (i,j) \in \mathcal{E} \tag{1b}$$

$$P_i^{\min} \le P_i^{\text{input}} \le P_i^{\max} \forall i \in \mathcal{N} \tag{1c}$$

where θ_i denotes the phase angle of node i, and $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_N)^T$. P_i^{input} and P_i^{output} denote the active power input and output at node i, respectively. P_{ij} denotes the active power flow of the line $i \to j$. X_{ij} is the impedance of line $i \to j$. $P_i^{\min}$ and $P_i^{\max}$ are the lower and upper limits of the injected active power at node i, respectively.

We then introduce an MES model, which considers electricity, natural gas, and heat energy, based on the EH model from [4]. The conversion devices utilized in this system include CHP and gas furnaces. The constraints corresponding to this EH model are illustrated in (2).

$$\underbrace{\begin{bmatrix} L_i^{\text{e}} \\ L_i^{\text{t}} \end{bmatrix}}_{\boldsymbol{L_i}} = \underbrace{\begin{bmatrix} \eta^{\text{e}} & \kappa\eta^{\text{e,CHP}} \\ 0 & \kappa\eta^{\text{g,CHP}} + (1-\kappa)\eta^{\text{g,Fur}} \end{bmatrix}}_{\boldsymbol{C_i}} \underbrace{\begin{bmatrix} P_i^{\text{output}} \\ G_i^{\text{output}} \end{bmatrix}}_{\boldsymbol{E_i}} \tag{2}$$

L_i^{e},L_i^{g} are the demands of electricity and heat load, respectively. The allocation of natural gas energy via κ can be noted, which is then used in CHP and gas furnace. Natural gas energy is transformed into heat energy via the heat efficiency of CHP $\eta^{\text{g,CHP}}$ and of gas furnace $\eta^{\text{g,Fur}}$. The electricity load is supplied by the transformer and CHP, whose operating efficiency is η^{e} and $\eta^{\text{e,CHP}}$, respectively.

Following [5], the constraints of the natural gas network are formulated in (3).

$$G_i^{\text{input}} - G_i^{\text{output}} = \sum\nolimits_{j:i\to j} G_{ij}, \quad \forall i \in \mathcal{N} \tag{3a}$$

$$G_i^{\min} \le G_i^{\text{input}} \le G_i^{\max}, \quad \forall i \in \mathcal{N} \tag{3b}$$

$$G_{ij} = \begin{cases} k_{ij}\text{GHV}\sqrt{{S_i}^2 - {S_j}^2}, \text{ if } S_i \ge S_j, \forall (i,j) \in \mathcal{E} \\ -k_{ij}\text{GHV}\sqrt{{S_j}^2 - {S_i}^2}, \text{ if } S_i \le S_j, \forall (i,j) \in \mathcal{E} \end{cases} \tag{3c}$$

where S_i denotes the pressure of node i, and $\boldsymbol{S} = (S_1, S_2, ..., S_N)^T$. G_i^{input} and G_i^{output} respectively denote the energy of gas input and output at node i. G_{ij} denotes the gas energy transferred by line $i \rightarrow j$. $G_i^{\min}$ and $G_i^{\max}$ are the lower and upper limits of the injected gas energy at node i, respectively. k_{ij} and GHV are gas pipeline and energy coefficients, respectively. Due to the nonlinearity of (3c), this constraint can be convex using piecewise linearization.

3. Successive Dual Taylor-expansion (SDTE) Algorithm

3.1. Classic CEF model

Aligning with the CEF for power systems [7], we introduce the CEF model in MES accounting for carbon emissions from both CHP and gas furnaces. NCI formulation is written as

$$E_i = \frac{E_i^P P_i^{\text{input}} + E_i^G \kappa \eta^{e,\text{CHP}} G_i^{\text{output}} + \sum_j P_{ji}^B E_j}{P_i^{\text{input}} + \kappa \eta^{e,\text{CHP}} G_i^{\text{output}} + \sum_j P_{ji}^B}, \forall i, j \in \mathcal{N} \tag{4}$$

$$\boldsymbol{P^B} = \begin{cases} P_{ij}^B = 0, P_{ji}^B = -P_{ij}, P_{ij} \leq 0 \\ P_{ij}^B = P_{ij}, P_{ji}^B = 0, P_{ij} \geq 0 \end{cases} \tag{5}$$

where E_i denotes the NCI of node i, and $\boldsymbol{E} = (E_1, E_2, ..., E_N)^T$. E_i^P denotes the carbon emission intensity from the generator at node i. E_i^G represents the carbon emission intensity from gas at node i. $\boldsymbol{P^B}$ denotes the absolute active power flow distribution matrix, which is defined in (5).

3.2. Proposed LCEF model and SDTE Algorithm

Due to the high nonconvexity, it becomes difficult to handle Equation (4) with off-the-shelf solvers directly. The simulations indicate that the commercial solvers cannot even find a feasible solution when the model size increases significantly. This paper introduces a Successive Dual Taylor-expansion (SDTE) algorithm. The iterative constraint equation is introduced in this section. We begin by rewriting (4) to Equation (6).

$$\begin{aligned} &E_i(P_i^{\text{input}} + \kappa \eta^{e,\text{CHP}} G_i^{\text{output}} + \sum\nolimits_{j \rightarrow i} P_{ji}^B) = \\ &E_i^P P_i^{\text{input}} + E_i^G \kappa \eta^{e,\text{CHP}} G_i^{\text{output}} + \sum\nolimits_{j \rightarrow i} P_{ji}^B E_j \end{aligned} \tag{6}$$

The nonconvex term consists of $\sum_{j \rightarrow i} P_{ij}^B E_j$ and $E_i(P_i^{\text{input}} + \kappa \eta^{e,\text{CHP}} G_i^{\text{output}} + \sum_{j \rightarrow i} P_{ji}^B)$. It can be linearized by the first-order Taylor expansion. However, simply linearizing them can cause oscillation issues, thereby affecting the convergence performance. To address the issue, this paper proposes a novel dual-linearization method. First, we employ the Jacobi iteration method for NCI. The closed-form constraint (6) can be written in the iterative form of (7).

$$\begin{aligned} &E_i^k(P_i^{\text{input},k} + \kappa \eta^{e,\text{CHP}} G_i^{\text{output},k} + \sum\nolimits_{j \rightarrow i} P_{ji}^{B,k}) = \\ &E_i^P P_i^{\text{input},k} + E_i^G \kappa \eta^{e,\text{CHP}} G_i^{\text{output},k} + \sum\nolimits_{j \rightarrow i} P_{ji}^{B,k} E_j^{k-1} \end{aligned} \tag{7}$$

Here, (7) represents the constraint in the kth iteration, where $(\cdot)^k$ denotes the kth iteration solution of the corresponding variable. However, it is still unsolvable with the nonconvex term $E_i^k(P_i^{\text{input},k} + \kappa\eta^{e,\text{CHP}}G_i^{\text{output},k} + \sum_{j\to i} P_{ji}^{B,k})$.

Next, we employ the first-order Taylor expansion to linearize the constraint twice while ensuring the stability of iterative convergence. The kth linearized iteration CEF constraint for node i is shown in (8),

$$\begin{aligned}
&\left(E_i^k - E_i^{k-1}\right)\left(P_i^{\text{input},k-1} + \kappa\eta^{g,\text{CHP}}G_i^{\text{output},k-1} + \sum_{j\to i} P_{ji}^{B,k-1}\right) \\
&+ E_i^{k-1}\left(P_i^{\text{input},k} + \kappa\eta^{e,\text{CHP}}G_i^{\text{output},k} + \sum_{j\to i} P_{ji}^{B,k}\right) = \\
&E_i^P P_i^{\text{input},k} + E_i^G\eta^{e,\text{CHP}}G_i^{\text{output},k} + \sum_{j\to i} P_{ji}^{B,k}E_j^{k-1}
\end{aligned} \tag{8}$$

The objective function consists of electricity costs, natural gas costs, and carbon costs, as shown in (9).

$$\text{cost}^e = \sum\nolimits_{i\in\mathcal{N}} a_i^{\text{e}}(P_i^{\text{input}})^2 + b_i^{\text{e}}P_i^{\text{input}} + c_i^{\text{e}} \tag{9a}$$

$$\text{cost}^g = \sum\nolimits_{i\in\mathcal{N}} b_i^{\text{g}}G_i^{\text{input}} \tag{9b}$$

$$\text{cost}^c = \sum\nolimits_{i\in\mathcal{N}} b_i^{\text{c}}E_iL_i^{\text{e}} \tag{9c}$$

where $a_i^{\text{e}}, b_i^{\text{e}}, c_i^{\text{e}}, b_i^{\text{g}}$ are the coefficient parameters for the node i, and b_i^{c} is the price of demand-side carbon emission for the node i [12]. The optimization problem at iteration $k+1$ is formulated as below

$$\begin{aligned}
(P0)\quad \min\quad & f(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{E}) = (\text{cost}^e + \text{cost}^g + \text{cost}^c) \\
\text{s.t.}\quad & (1)-(3), (5), (8).
\end{aligned} \tag{10a}$$

Using the SDTE algorithm, which is shown in the algorithm 1, we can effectively optimize the problem (P0).

4. Privacy-preserving Distributed Optimization

For self-contained reasons, subsection 4.1 first introduces classic ADMM-based distributed optimization. Then, subsections 4.2 and 4.3 propose new acceleration techniques and a privacy-preserving framework, respectively.

4.1. ADMM-based Distributed Optimization

The basic idea is to partition the original power network into a cluster of subregions, enabling independent problem formulation and solving within each subregion. Assume the original network is partitioned to R regions. Let $\mathcal{R}$ denote the region indices set. To mimic the coupling

Algorithm 1 Successive Dual Taylor-expansion (SDTE) Algorithm

Input: $\boldsymbol{L}^e, \boldsymbol{L}^t$

Output: $\boldsymbol{P}^{\text{input}}, \boldsymbol{G}^{\text{input}}, \boldsymbol{E}, f(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{E})$

Initialisation :$\boldsymbol{E}^0, \boldsymbol{P}^{B,0}, \boldsymbol{P}^{\boldsymbol{input,0}}, \boldsymbol{G}^{\boldsymbol{output,0}}, e^0 = +\infty$

1: **while** $e^k > err$ **do**
2: update CEF constrain by using (8)
3: solve the optimization problem (P0)
4: update $\boldsymbol{E}^k, \boldsymbol{P}^{B,k}, \boldsymbol{P}^{\boldsymbol{input,k}}, \boldsymbol{G}^{\boldsymbol{output,k}}$
5: compute $e^k = \left\|\boldsymbol{E}^k - \boldsymbol{E}^{k-1}\right\|_2^2$
6: $k = k + 1$
7: **end while**

constraints between regions, additional constraints and virtual duplicated nodes are introduced in each subproblem [27]. This work constrains boundary variables for adjacent nodes in the power flow network, natural gas network, and carbon intensity network. Each network forms a large matrix constraint for all regions, as shown in (11)

$$\underbrace{\begin{bmatrix} \boldsymbol{A_1} & \cdots & \boldsymbol{O} \\ \vdots & \ddots & \vdots \\ \boldsymbol{O} & \cdots & \boldsymbol{A_R} \end{bmatrix}}_{\boldsymbol{A}} \cdot \underbrace{\begin{bmatrix} \boldsymbol{X_1} \\ \vdots \\ \boldsymbol{X_R} \end{bmatrix}}_{\boldsymbol{X}} = \underbrace{\begin{bmatrix} \boldsymbol{B_1} \\ \vdots \\ \boldsymbol{B_R} \end{bmatrix}}_{\boldsymbol{B}} \cdot \boldsymbol{Z}, \tag{11}$$

where $\boldsymbol{X_m} \in \mathcal{R}^{N_m \times 1}$ represents the information of all N_m nodes in the mth region. $\boldsymbol{Z} \in \mathcal{R}^{K \times 1}$ represents the K temporary variables of all boundary nodes in every region. $\boldsymbol{A_m} \in \mathcal{R}^{H_m \times N_m}$ represents selection matrix that extracts the information of all the H_m boundary nodes in the mth region. $\boldsymbol{B_m} \in \mathcal{R}^{H_m \times K}$ represents selection matrix that extracts the information of H_m boundary nodes corresponding to the temporary variables in the mth region.

With additional constraints for adjacent nodes, we can recast the Problem (P0) to (P1) as follows.

$$\begin{aligned} (P1) \quad \min \quad & \sum_{m \in \mathcal{R}} f_m(\boldsymbol{X}) \\ \text{s.t.} \quad & \boldsymbol{AX} = \boldsymbol{BZ} && \text{(12a)} \\ & \boldsymbol{X}_m \in \mathcal{X}_m, \ \forall m \in \mathcal{R} && \text{(12b)} \end{aligned}$$

where $\boldsymbol{X} \triangleq (\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{E})$, $\boldsymbol{A} \triangleq (\boldsymbol{A_\theta}, \boldsymbol{A_s}, \boldsymbol{A_c})$, $\boldsymbol{B} \triangleq (\boldsymbol{B_\theta}, \boldsymbol{B_s}, \boldsymbol{B_c})$, and (12a) denotes constraints for neighboring regions in the power system, natural gas system, and carbon network. Equation (12b) denotes the original constraints within region m. The augmented Lagrangian function for the optimization problem $P1$ is denoted as (13)

$$\begin{aligned} \mathcal{L} &= f(\boldsymbol{X}) + \boldsymbol{Y}^T(\boldsymbol{AX} - \boldsymbol{BZ}) + \frac{\rho}{2}\|\boldsymbol{AX} - \boldsymbol{BZ})\|_2^2 \\ &= \sum\nolimits_{m \in \mathcal{R}} \mathcal{L}_m\left(\boldsymbol{X_m}, \boldsymbol{Y_m}, \boldsymbol{Z}, \rho\right) = \sum\nolimits_{m \in \mathcal{R}} \mathcal{L}_m, \end{aligned} \tag{13}$$

where $\boldsymbol{Y} \triangleq (\boldsymbol{Y_\theta}, \boldsymbol{Y_s}, \boldsymbol{Y_c})$ denotes the Lagrange multipliers for boundary constraint in (12b), while $\rho \triangleq (\rho_\theta, \rho_s, \rho_c)$ are the associated second-order penalty coefficients. The symbol with the subscript $(\cdot)_m$ represents the variable corresponding to the mth region.

Following the iterative solving method described in [26], the variables after and before the $k+1$th update are denoted as $(\cdot)^{k+1}$ and $(\cdot)^k$, respectively. The approach consists of three distinct iterative update steps and an update step of the second-order penalty coefficients.

4.1.1. XYZ-updated step

$\boldsymbol{X}$ and $\boldsymbol{Y}$ update can be distributed and solved in each region. (14c) represents the $\boldsymbol{Z}$ update step computed in the dispatch center, which is used to update the intermediate variable, where $\boldsymbol{B}^\dagger$ represents the Moore-Penrose pseudoinverse of matrix $\boldsymbol{B}$.

$$\boldsymbol{X_m^{k+1}} = \underset{\boldsymbol{X_m} \in \mathcal{X}_m}{\operatorname{argmin}} \mathcal{L}_m(\boldsymbol{X_m}, \boldsymbol{Y_m^k}, \boldsymbol{Z^k}, \rho) \tag{14a}$$

$$\boldsymbol{Y_m^{k+1}} = \boldsymbol{Y_m^k} + \rho\left(\boldsymbol{A_m X_m^{k+1}} - \boldsymbol{B_m Z^k}\right) \tag{14b}$$

$$\boldsymbol{Z^{k+1}} = \boldsymbol{B^\dagger A X^{k+1}} + (1/\rho)\boldsymbol{B^\dagger Y^{k+1}} \tag{14c}$$

4.1.2. ρ-updated step

The ρ update step is shown in (15). Typical settings are used $\mu = 10$ and $\tau^{\text{incr}} = \tau^{\text{decr}} = 2$.

$$\rho^{k+1} = \begin{cases} \tau^{\text{incr}}\rho^k, & \text{if } r^k > \mu d^k \\ \rho^k/\tau^{\text{decr}}, & \text{if } d^k > \mu r^k \\ \rho^k, & \text{otherwise} \end{cases} \tag{15}$$

where r^k and d^k are convergence criteria which are shown in (16). r is the primal residual, which could be r_θ, r_s, r_c, and d is the dual residual, which could be d_θ, d_s, d_c.

$$r^k = \left\|\boldsymbol{AX^k} - \boldsymbol{BZ^k}\right\|_2 \tag{16a}$$

$$d^k = \left\|\rho\boldsymbol{A}^T\boldsymbol{B}(\boldsymbol{Z^{k-1}} - \boldsymbol{Z^k})\right\|_2 \tag{16b}$$

4.2. Dual-regularization ADMM (DR-ADMM)

The method mentioned above experiences convergence challenges for real-world problems. Furthermore, due to the iterative updates of nonconvex constraints, it exhibits oscillations. This issue stems from the Lagrange multiplier $\boldsymbol{Y}$ undergoing substantial growth during the iterative process. Therefore, this paper introduces a novel approach by mitigating the growth of $\boldsymbol{Y}$ to accelerate convergence. As shown in (17), we introduce a regularization term $\boldsymbol{Y^{k+1}}$ in the augmented Lagrangian function, effectively curbing the trend of $\boldsymbol{Y}$'s excessive updates.

$$\boldsymbol{X_m^{k+1}} = \underset{\boldsymbol{X_m} \in \mathcal{X}_m}{\operatorname{argmin}} \mathcal{L}_m + \lambda\left\|\boldsymbol{Y_m^{k+1}}\right\|_2^2 \tag{17}$$

By decomposing the acceleration term of (17) into (18), we can illustrate how the acceleration term functions.

$$\begin{aligned} \lambda\left\|\boldsymbol{Y_m^{k+1}}\right\|_2^2 &= \lambda\left\|\boldsymbol{Y_m^k} + \rho(\boldsymbol{A_m X_m^{k+1}} - \boldsymbol{B_m Z^k})\right\|_2^2 \\ &= \lambda\left\|\boldsymbol{Y_m^k}\right\|_2^2 + \lambda\rho^2\left\|\boldsymbol{A_m X_m^{k+1}} - \boldsymbol{B_m Z^k}\right\|_2^2 \\ &\quad + 2\lambda\rho(\boldsymbol{Y_m^k})^T(\boldsymbol{A_m X_m^{k+1}} - \boldsymbol{B_m Z^k}) \end{aligned} \tag{18}$$

It is observed that (17) introduced essentially transforms the Lagrangian multiplier term from $\boldsymbol{Y}^T(\boldsymbol{AX}-\boldsymbol{BZ})$ to (19), and the quadratic term shifts from $\rho\,\|\boldsymbol{AX}-\boldsymbol{BZ}\|_2^2$ to (20).

$$(2\lambda\rho+1)\boldsymbol{Y}^T(\boldsymbol{AX}-\boldsymbol{BZ}) \tag{19}$$

$$(\lambda\rho^2+\rho)\,\|\boldsymbol{AX}-\boldsymbol{BZ}\|_2^2 \tag{20}$$

They serve as the basis for iterations and ensure changes in $\boldsymbol{Y}$ and ρ do not exhibit dramatic fluctuations, thereby promoting a smoother convergence curve. It is observed that this technique is equivalent to introducing dual regularization terms in the Lagrangian function. Therefore, we name it as Dual-regularized ADMM (DR-ADMM).

Next, we briefly discuss the convergence of the proposed DR-ADMM, which is a variant of ADMM. Therefore, we employ the convergence conditions of ADMM for nonconvex problems [34].

1) A1 (Coercivity): This requires that the objective function tends to infinity as decision variables approach infinity. This is easily verified since $cost^c$ is positively correlated with carbon emissions in MES. As demand loads approach infinity, carbon emissions also tend to infinity. Similarly, electricity and gas costs $cost^a$ and $cost^b$ will also diverge to infinity.
2) A2 (Feasibility): This requires $Im(\boldsymbol{A}) \subseteq Im(\boldsymbol{B})$. The construction of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ inherently assumes the existence of boundary nodes. Thus, for any $\boldsymbol{X}$, there always exist corresponding temporary variables Z satisfying the constraint.
3) A3 (Lipschitz Sub-minimization Paths): This requires Lipschitz continuity for $H(\boldsymbol{u})$ and $F(\boldsymbol{u})$. Here, $H(\boldsymbol{u}) = \arg\min_{\boldsymbol{Z}:\boldsymbol{BZ}=\boldsymbol{u}} 0$, trivially satisfies continuity. $F(\boldsymbol{u}) = \arg\min_{\boldsymbol{X}:\boldsymbol{AX}=\boldsymbol{u}} f(\boldsymbol{X})$, where $f(\boldsymbol{X})$ represents operational costs and $\boldsymbol{X} \in \boldsymbol{\mathcal{X}}$. Although variations in u affect power and gas outputs in the MES, physical constraints $\boldsymbol{X} \in \boldsymbol{\mathcal{X}}$ limit the rate of cost changes. Thus, $F(\boldsymbol{u})$ can be reasonably assumed Lipschitz continuous.
4) A4 (Objective Regularity): [34] indicates ADMM convergence when the objective function is piecewise linear. The proposed SDTE method can be interpreted as a piecewise linear function with infinitely many segments. While not strictly piecewise linear, it approximately satisfies this regularity condition.

Hence, the proposed DR-ADMM can converge to the local optima for the problem (P1). It is noted that ADMM does not require precise updates of X and Z in each iteration [26]. The single iteration update process is shown in Fig. 2, in which SDTE and DR-ADMM are updated alternately.

4.3. Privacy-preserving Framework

It can be inferred that $\boldsymbol{X}$ and $\boldsymbol{Y}$ can be updated locally by the agents within their regions from (14). However, $\boldsymbol{Z}$ needs to be updated with global information. Therefore, the agents are required to transmit the signal $\boldsymbol{M}^{\boldsymbol{k+1}} = \boldsymbol{AX}^{\boldsymbol{k+1}} + \boldsymbol{Y}^{\boldsymbol{k+1}}/\rho$ to the dispatch center. The dispatch center then computes the average of $\boldsymbol{M}^{\boldsymbol{k+1}}$ using the $\boldsymbol{B}^{\dagger}$ operator from (14c) to update $\boldsymbol{Z}^{\boldsymbol{k+1}}$. Subsequently, the updated values of $\boldsymbol{Z}^{\boldsymbol{k+1}}$ and ρ^{k+1} are transmitted back to the agents.

In the case of plaintext communication, an attacker could infer the internal load information of regions by collecting the converged $\boldsymbol{Z}$ values. By employing additive homomorphic encryption

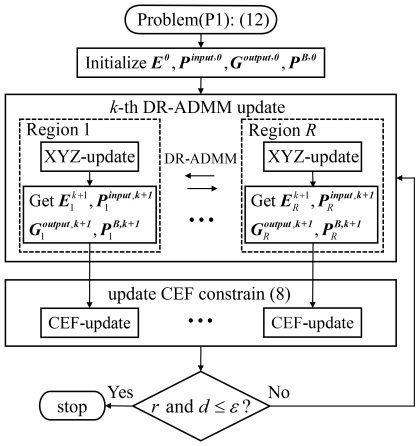


Figure 2: The ADMM optimization with LCEF.

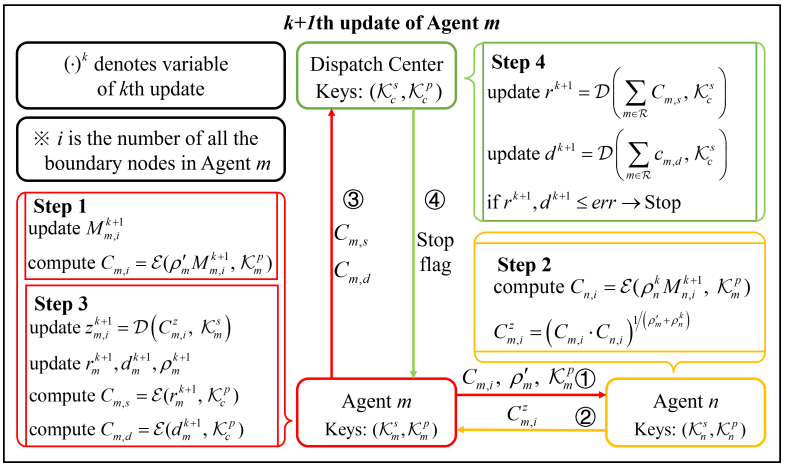


Figure 3: $k+1$th update of Agent m using DSOP.

(AHE) Algorithm Paillier, a type of PHE algorithm, $\boldsymbol{Z}$ can be updated under ciphertext conditions, effectively preventing the leakage of load information privacy.

However, with homomorphic encryption, honest-but-curious neighboring agents could still infer the precise information of the boundary $\boldsymbol{M}$ of adjacent regions from the updated $\boldsymbol{Z}$, thereby exposing load information. Therefore, we propose a new approach in the algorithm 3 to address it.

For convenience, the following notations are established: The public and private keys of region m are denoted as $\mathcal{K}_m^p$ and $\mathcal{K}_m^s$. The encryption function is denoted as $\mathcal{E}(M_i, \mathcal{K}_m^p)$, where M_i represents plaintext of the ith element in $\mathbf{M}$. The decryption function is denoted as $\mathcal{D}(C, \mathcal{K}_m^s)$, where C represents the ciphertext. In addition, here are two properties of Paillier encryption shown below. 1) additive homomorphism: $\mathcal{D}(\mathcal{E}(M_i, \mathcal{K}_m^p) \cdot \mathcal{E}(M_j, \mathcal{K}_m^p), \mathcal{K}_m^s) = M_i + M_j$. 2) homomorphic multiplication of plaintext: $\mathcal{D}((\mathcal{E}(M_i, \mathcal{K}_m^p))^{c_o}, \mathcal{K}_m^s) = c_o \cdot M_i$, where c_o is a constant.

We begin by decentralizing the update steps for $\boldsymbol{Z}$ and ρ, assigning each region m ($m \in \mathcal{R}$) to manage its own local variables $\boldsymbol{Z_m}$ and ρ_m without consulting a central coordinator. By keeping these updates strictly within their own region, non-adjacent regions never exchange boundary data, effectively safeguarding their boundary information.

Taking the $k+1$th iteration of Agent m as an example, the process is shown in Fig. 3. Considering that the $B^\dagger$ operator in (14c) serves to calculate the average, we employ a weighted average to obscure the update of $\boldsymbol{Z}$. Below is an example to illustrate this process. For an adjacent node i that belongs to both the region m and n, the information updated in region m is $M_{m,i}$, and in region n is $M_{n,i}$. Now, region m intends to update its node i's temporary variables $z_{m,i}$, which is the element in $\boldsymbol{Z_m}$, and $C_{m,i}^z$=$\mathcal{E}(z_{m,i}, \mathcal{K}_m^p)$is the ciphertext of $z_{m,i}$. The following steps are involved:

4.3.1. Step 1

Generation of Encrypted Message: Agent m generates a random number ρ'_m based on its region's ρ_m. It then computes the ciphertext $C_{m,i} = \mathcal{E}(\rho'_m M_{m,i}, \mathcal{K}_m^p)$ of $M_{m,i}$ using its public key. Agent m sends its public key $\mathcal{K}_m^p$, the ciphertext $C_{m,i}$, and the random number ρ'_m to region n.

Algorithm 2 Dual-Regularized ADMM (DR-ADMM) for NCI-constrained problem with LCEF

Input: $\boldsymbol{L^e}, \boldsymbol{L^t}, \rho_\theta, \rho_s, \rho_c$

Output: $\boldsymbol{P}^{\text{input}}, \boldsymbol{G}^{\text{input}}, \boldsymbol{E}, f(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{E})$

Initialisation :$\boldsymbol{E^0}, \boldsymbol{Z^0_\theta}, \boldsymbol{Z^0_s}, \boldsymbol{Z^0_c}, \boldsymbol{Y^0_\theta}, \boldsymbol{Y^0_s}, \boldsymbol{Y^0_c}, \rho^0_\theta, \rho^0_s, \rho^0_c$, $\boldsymbol{P}^{B,0}, \boldsymbol{P^{input,0}}, \boldsymbol{G^{output,0}}, r^0 = +\infty, d^0 = +\infty, k = 0$

while $r^k > err_r$ **or** $d^k > err_d$ **do**
  for $m \in \mathcal{R}$ **do**
    update CEF constrain by using (8)
    update $\boldsymbol{\theta^{k+1}_m}, \boldsymbol{S^{k+1}_m}, \boldsymbol{E^{k+1}_m}$ by using (17)
    update $\boldsymbol{Y^{k+1}_{\theta,m}}, \boldsymbol{Y^{k+1}_{s,m}}, \boldsymbol{Y^{k+1}_{c,m}}$ by using (14b)
    update $\boldsymbol{P}^{\text{input},k}_m, \boldsymbol{G}^{\text{output},k}_m, \boldsymbol{E}^k_m, \boldsymbol{P}^{B,k}_m$
  end for
  update $\boldsymbol{Z^{k+1}_\theta}, \boldsymbol{Z^{k+1}_s}, \boldsymbol{Z^{k+1}_c}$ by using (14c)
  update r^{k+1}, d^{k+1} by using (16)
  update $\rho^{k+1}_\theta, \rho^{k+1}_s, \rho^{k+1}_c$ by using (15)
  $k = k + 1$
end while

4.3.2. Step 2

Updating Z by Agent n: After receiving this information from Agent m, Agent n generates the encrypted ciphertext $C_{n,i} = \mathcal{E}(\rho_n M_{n,i}, \mathcal{K}^p_m)$ of $M_{n,i}$ using Agent m's public key. Then it computes $C^z_{m,i}$ by using (21). Finally, Agent n sends $C^z_{m,i}$ back to Agent m.

$$C^z_{m,i} = (C_{m,i} \cdot C_{n,i})^{\frac{1}{(\rho'_m + \rho_n)}} \tag{21}$$

$$z_{m,i} = (\rho'_m M_{m,i} + \rho_n M_{n,i})/(\rho'_m + \rho_n) \tag{22}$$

Due to the properties of homomorphic encryption, the update of $z_{m,i}$ shifts from the original form $z_{m,i} = (M_{m,i} + M_{n,i})/2$ to (22), which is a weighted average using the weights ρ'_m and ρ_n. As a result, Agent m cannot infer the precise node information or Lagrange multiplier information of region n from the updated $z_{m,i}$. It should be noted that to gain more information from $M_{n,i}$, the weight ρ'_m should be set relatively small. In this paper, ρ'_m is set to ρ_m/α, where α is a random number greater than 1.

The weighted average update of $\boldsymbol{Z_m}$ for region m can be summarized in matrix form as shown in (23)

$$\boldsymbol{Z^{k+1}_m} = \boldsymbol{A_m F_m}(\boldsymbol{A X^{k+1}} + \boldsymbol{\mathcal{P}^{-1} Y^{k+1}}), \tag{23}$$

where $\boldsymbol{\mathcal{P}} = diag\,(\rho_1, \rho_2, ..., \rho_R)$, and $\boldsymbol{F}_m$ represents the weight matrix of region m.

$$\boldsymbol{F}_m = (\boldsymbol{B}^T \boldsymbol{\mathcal{P}'_m} \boldsymbol{B})^\dagger \boldsymbol{B}^T \boldsymbol{\mathcal{P}'_m}. \tag{24}$$

where $\boldsymbol{\mathcal{P}'_m} = diag\,(\rho_1, \rho_2, ..., \rho'_m, ..., \rho_R)$. Equation (23) utilizing ρ as the weight has practical significance. When ρ_m for region m increases, it indicates a greater emphasis on the original residual.

Algorithm 3 Dynamic-second-order-term Privacy-preserving (DSOP) Algorithm

Input: $\boldsymbol{L^e}, \boldsymbol{L^t}, \rho_\theta, \rho_s, \rho_c$

Output: $\boldsymbol{P}^{\text{input}}, \boldsymbol{G}^{\text{input}}, \boldsymbol{E}, f(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{E})$

Initialisation :$\boldsymbol{E^0}, \boldsymbol{Z_\theta^0}, \boldsymbol{Z_s^0}, \boldsymbol{Z_c^0}, \boldsymbol{Y_\theta^0}, \boldsymbol{Y_s^0}, \boldsymbol{Y_c^0}, \rho_\theta^0, \rho_s^0, \rho_c^0$, $\boldsymbol{P}^{B,0}, \boldsymbol{P}^{\text{input},\boldsymbol{0}}, \boldsymbol{G}^{\text{output},\boldsymbol{0}}, r^0 = +\infty, d^0 = +\infty, k = 0$

while $r^k > err_r$ **or** $d^k > err_d$ **do**
  for $m \in \mathcal{R}$ **do**
    update CEF constrain by using (8)
    update $\boldsymbol{\theta_m^{k+1}}, \boldsymbol{S_m^{k+1}}, \boldsymbol{E_m^{k+1}}$ by using (17)
    update $\boldsymbol{Y_{\theta,m}^{k+1}}, \boldsymbol{Y_{s,m}^{k+1}}, \boldsymbol{Y_{c,m}^{k+1}}$ by using (14b)
    update $\boldsymbol{P}_m^{\text{input},k}, \boldsymbol{G}_m^{\text{output},k}, \boldsymbol{E}_m^k, \boldsymbol{P}_m^{B,k}$
  end for
  for $m \in \mathcal{R}$ **do**
    update $\boldsymbol{Z_{\theta,m}^{k+1}}, \boldsymbol{Z_{s,m}^{k+1}}, \boldsymbol{Z_{c,m}^{k+1}}$ by using (23) with PHE
    update r_m^{k+1}, d_m^{k+1} by using (25)
    update $\rho_{\theta,m}^{k+1}, \rho_{s,m}^{k+1}, \rho_{c,m}^{k+1}$
  end for
  update r^{k+1}, d^{k+1}
  $k = k + 1$
end while

In this case, the updated node information can be considered more feasible, leading to a higher weight in the weighted average.

4.3.3. Step 3

Decryption by Agent m: Agent m receives $C_{m,i}^z$ and decrypts it using its private key to obtain $z_{m,i} = \mathcal{D}(C_{m,i}^z, \mathcal{K}_m^s)$. After Agent m updates its own $\boldsymbol{Z_m}$, it proceeds to update its s_m and d_m according to (25). Then Agent m updates its ρ_m by using (20). Finally, Agent m encrypts the s_m and d_m into $C_{m,s}$ and $C_{m,d}$ using the dispatch center's public key $\mathcal{K}_c^p$ and uploads the encrypted data to the dispatch center.

$$r_m^{k+1} = \left\| \boldsymbol{A_m X^{k+1}} - \boldsymbol{Z_m^{k+1}} \right\|_2 \tag{25a}$$

$$d_m^{k+1} = \left\| \rho_m (\boldsymbol{Z_m^k} - \boldsymbol{Z_m^{k+1}}) \right\|_2 \tag{25b}$$

4.3.4. Step 4

Convergence check. Verify $r^{k+1} = \sum_{m \in \mathcal{R}} r_m$ and $d^{k+1} = \sum_{m \in \mathcal{R}} d_m$.

Based on these steps, Algorithm 2 can be updated as Algorithm 3.

5. Case study

Numerical case studies are performed with two cases: IEEE 9-bus Power & 8-node Gas System, and IEEE 118-bus Power & 21-node Gas System. Given that Algorithm 3's performance is

Table 1: IEEE 9-bus System - Data for generators

Node i	$P^{\max}$ (MW)	$P^{\min}$ (MW)	a^e ($/MW^2)	b^e ($/MW)	c^e ($)	E^P (kgCO_2/MW)
1	250	10	0.1100	5.0	150.0	0.22
2	300	10	0.0850	1.2	600.0	0.25
3	270	10	0.1225	1.0	335.0	0.28

Table 2: 8-node Gas System - Data for gas supply

Node i	$G^{\max}$ (MW)	$G^{\min}$ (MW)	b^g ($/MW)	E^G (kgCO_2/MW)
1	300	0	1.00	0.15
2	300	0	0.85	0.15
3	300	0	1.25	0.15

Table 3: 8-node Gas System - Data for gas pipeline

Branch	From (i)	To (j)	GHV $\cdot\, k_{ij}$ (p.u.)
1	1	4	3.0
2	4	5	2.5
3	5	6	3.0
4	3	6	3.5
5	6	7	3.0
6	7	8	3.5
7	8	2	3.0

Table 4: Demands of electricity and heat load

node			L^{e} (MW)	L^{t} (MW)
EH	Power	Gas		
1	4	4	0	30
2	5	5	90	90
3	6	6	0	30
4	7	7	100	100
5	8	8	0	90
	9		125	

affected by random variables, Algorithm 2 is selected to evaluate the effectiveness of our method in subsection 5.1 and 5.2. The accuracy and computational efficiency of our linearization method are benchmarked against existing approaches in subsection 5.3. In subsection 5.4, Algorithm 3 is examined to assess the consistency and reliability and to benchmark it against methods from other studies. The case study is performed with a 12th Gen Intel(R) Core(TM) i5-12400 CPU @ 2.5GHz (12 CPUs) and 32G RAM. GUROBI 10.0 is employed to solve the problems.

5.1. IEEE 9-bus Power and 8-node Gas System

A multi-energy system comprising the IEEE 9-bus power network and an 8-node gas network is used for small system testing. The data for generators per node are provided in Table 1. The nodal gas supply data and the gas pipeline data are given in Table 2 and Table 3, respectively. For carbon cost, this test sets the coefficient of NCI $b^c = 10$ \$/kgCO_2.

The location of EH nodes is shown in Fig. 4. Set the EH parameter $\kappa = 0.5$, $\eta^{\mathrm{e}} = 1.0$, $\eta^{\mathrm{e,CHP}} = 0.30$, $\eta^{\mathrm{g,CHP}} = 0.4$, $\eta^{\mathrm{g,Fur}} = 0.7$ according to [5]. The demands of electricity and heat load of different nodes are listed in Table 4. The regions are divided into three areas: $R_1 = \{1, 4, 9\}$, $R_2 = \{2, 7, 8\}$, and $R_3 = \{3, 5, 6\}$. The accuracy for convergence indicators err_r and err_d is set to 10^{-3}.

The initial values for the iteration is shown in Table 5. It is important to ensure the feasibility of the solution when setting the initial values for the NCI constraints. Therefore, $\boldsymbol{P^{B,0}}$ should be initialized according to the line connectivity of each region. To test the generality of the algorithm. The acceleration parameters are set as $\lambda_\theta = 4$, $\lambda_E = 2$, and $\lambda_S = 1.4$.

5.1.1. Convergence performance with synchronous updates

Fig. 5(a) illustrates the convergence performance of Algorithm 2 in the test conducted in this section. It is evident that the convergence performance of the power grid, gas network, and NCI

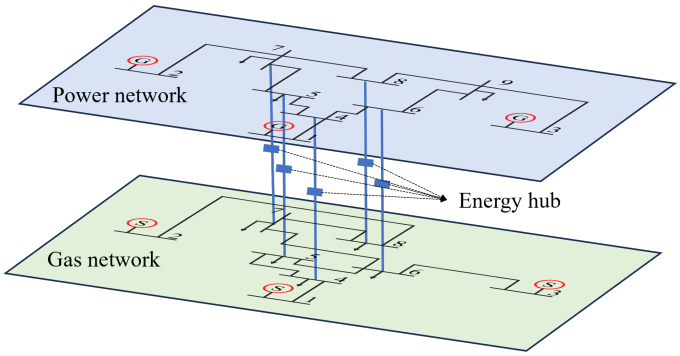


Figure 4: The location of Energy hub nodes.

Table 5: Initial values for the iteration

Variable	Value
$\boldsymbol{P^{B,0}}$	$P_{ij}^{B,0} = 0.2, \forall (i,j) \in \mathcal{E}$ $P_{ij}^{B,0} = 0.0, \forall (i,j) \notin \mathcal{E}$
$\boldsymbol{P}^{\text{input,0}}$	$\boldsymbol{P}^{\min}$
$\boldsymbol{G}^{\text{output,0}}$	$\boldsymbol{G}^{\max}$
$\boldsymbol{E^0}$	$\boldsymbol{E^P}$
$\boldsymbol{Y^0}$	0
$\boldsymbol{Z^0}$	0
$\rho_\theta^0, \rho_s^0, \rho_c^0$	0.1

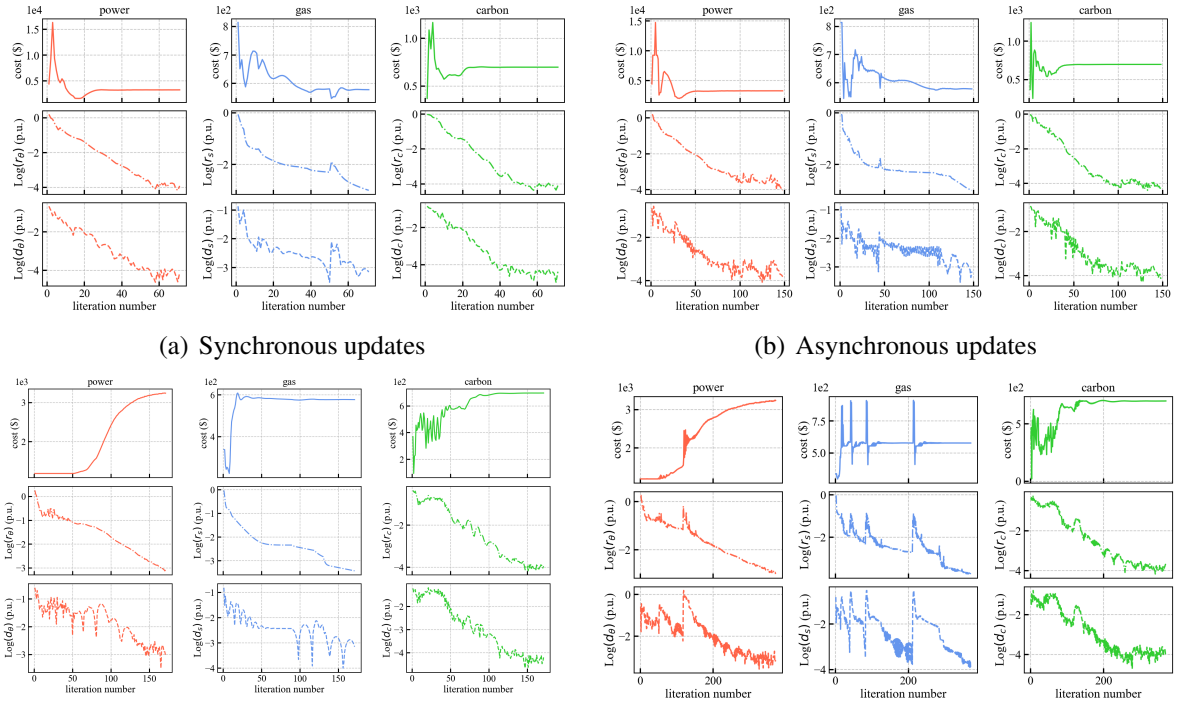
network differ significantly due to the variations in their topological structures and computational methods. Therefore, decoupling the system into three networks can effectively prevent interference between different network systems.

5.1.2. Convergence performance with asynchronous updates

Since the ADMM update allows subproblems to reach an approximate optimal solution, Algorithm 2 is also suitable for asynchronous update environments, which is beneficial for handling communication congestion scenarios. In the experiment, the updates for region $R_1 = \{1, 4, 9\}$ were selected as the main thread, while updates for other regions were allowed to proceed asynchronously, which means regions can update themselves without waiting others. The update results are shown in Fig. 5(b).

5.1.3. Accuracy Comparison

To test the accuracy error compared to the centralized optimization, we compare the results of Algorithm 2 base on synchronous (Syn.) and asynchronous (Asy.) updating with those obtained from a centralized optimization after linearizing the CEF model. The relative error (RE) shown in (26) is used to measure the accuracy between the generation strategy (26a), natural gas strategy (26b), NCI (26c), and objective function (26d). Where $(\cdot)^*$ denotes the result of the centralized optimization. The comparison results are shown in Table 6. It can be observed that the error,



(a) Synchronous updates

(b) Asynchronous updates

(c) Synchronous updates without acceleration items

(d) Asynchronous updates without acceleration items

Figure 5: Convergence result of 9-bus power and 8-node gas system with synchronous updates (a), asynchronous updates (b), synchronous updates without acceleration items (c), and asynchronous updates without acceleration items (d). The iteration convergence performances are divided by columns (different colors), representing the power grid (red), gas network (blue), and NCI network (green), respectively. Divided by rows (different line styles) from top to bottom, they show each network's objective function (solid), primal residual r (dash-dot), and dual residual d (dash) during the iteration convergence process.

whether in synchronous or asynchronous updates, is within an acceptable range.

$$RE_e = \sqrt{\sum\nolimits_{i\in\mathcal{N}} \left[\left(P_i^{\text{input}} - P_i^{\text{input}*}\right)/P_i^{\text{input}*}\right]^2} \tag{26a}$$

$$RE_g = \sqrt{\sum\nolimits_{i\in\mathcal{N}} \left[\left(G_i^{\text{input}} - G_i^{\text{input}*}\right)/G_i^{\text{input}*}\right]^2} \tag{26b}$$

$$RE_c = \sqrt{\sum\nolimits_{i\in\mathcal{N}} \left[\left(E_i - E_i^{*}\right)/E_i^{*}\right]^2} \tag{26c}$$

$$RE_{cost} = \sqrt{\left[\left(cost - cost^{*}\right)/cost^{*}\right]^2} \tag{26d}$$

5.1.4. The impact of computational acceleration

This network is also used to test the impact of computational acceleration on the algorithm's convergence speed. The synchronous and asynchronous updates without the acceleration terms are shown in Fig. 5(c) and 5(d). It can be observed that the acceleration terms deliver a substantial

Table 6: Accuracy Comparison with Centralized Approach

	Centralized	Algorithm 2 (p.u.)		RE (p.u.)	
	methold (p.u.)	Syn.	Asy.	Syn.	Asy.
Generation strategy	0.58615	0.58436	0.58436	0.00107	0.00247
	0.96900	0.97000	0.97000		
	0.66757	0.66964	0.66727		
Natural gas strategy	3.00000	2.97801	2.98854	0.10914	0.05829
	3.00000	3.00000	3.00000		
	0.18182	0.20299	0.19311		
Nodal carbon intensity	0.22000	0.22000	0.22000	0.0041	0.0031
	0.25000	0.25000	0.25000		
	0.28000	0.28000	0.28000		
	0.21143	0.21140	0.21140		
	0.22443	0.22449	0.22448		
	0.26581	0.26585	0.26580		
	0.21728	0.21728	0.21730		
	0.22979	0.22981	0.22981		
	0.22236	0.22237	0.22237		
Cost	4558.045 ($/hr)	4560.702 ($/hr)	4556.427 ($/hr)	0.00058	0.00035

improvement in computational speed. Besides, Fig. 6 shows the impact of different λ parameters on solving accuracy and convergence speed. The convergence speed is measured by the number of iterations. RE_e, RE_g, RE_c are used to measure the accuracy. Analysis of the experimental data reveals that larger parameters accelerate solving speed but may slightly reduce solving accuracy.

5.2. IEEE 118-bus Power and 21-node Gas System

In this section, a larger network is used to validate the model's generalization capability and to verify the optimization effect of the algorithm on NCI. The power grid used is the 118-bus network, witch is shown in Fig. 7. Using the bus numbers of the power grid nodes as references, gas pipelines listed in Table 7 were added to the 8-node gas network. The node numbers in the table correspond to the node numbers at the respective positions in the power grid. In this section's test, the acceleration parameter is set to $\lambda_\theta = 40$, $\lambda_E = 28$, and $\lambda_S = 10$. The regions are divided into four areas: $R_1 = \{1-32, 113-115, 117\}$, $R_2 = \{33-67\}$, $R_3 = \{68-81, 116, 118\}$, and $R_4 = \{82-112\}$.

5.2.1. The impact on NCI

Fig. 8 shows the changes in NCI before and after introducing carbon costs. Through optimization, the NCI of high carbon emission intensity nodes (such as node 34) has been effectively suppressed. Furthermore, by adjusting the values of b^c for different nodes in (9c), it is possible to

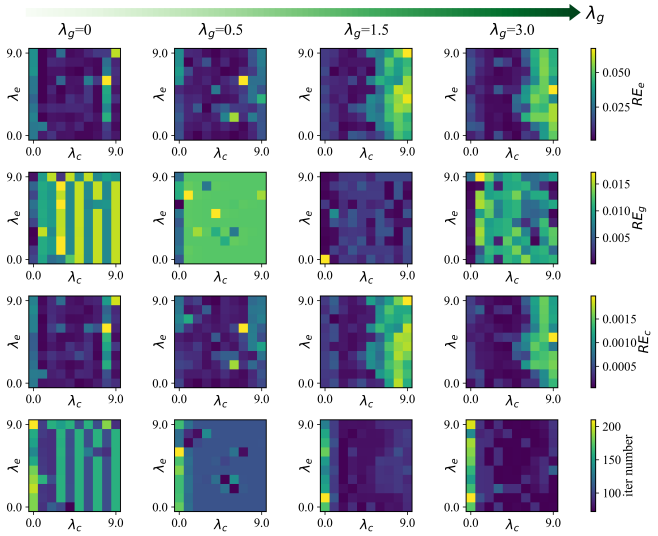


Figure 6: Convergence Speed and RE of Different Acceleration Parameters in 9-bus power and 8-node gas system. Each column of the subplot represents a specific value of λ_g, with each subplot having the horizontal axis as the values of λ_c and the vertical axis as the values of λ_e. Within each row of subplots, the color of each grid cell represents the magnitude of the RE parameter, with darker colors indicating smaller values.

specifically increase costs for high carbon emission users, effectively suppressing their emissions in the carbon market.

5.2.2. Carbon Emission Reduction Performance

To demonstrate the regulatory effect of demand-side carbon emission costs on the carbon emissions of MES, this section introduces demand elastic price curve to simulate demand response. The specific settings are as follows: 1) Electric demand L^e is adjustable within 92%–108% of the original demand. 2) The demand elastic price π^d and the electric demand power L^e follows the relationship: $\pi^d = 0.0005L^{e2} - 0.3L^e + 80$. 3) The objective function is modified to

$$min(\text{cost}^e + \text{cost}^g + \text{cost}^c - \sum_{i \in \mathcal{N}} \pi^d L_i^e).$$

Table 7: 21-node Gas System - Data for gas pipeline added based 8-node Gas System

branch	From (i)	To (j)	GHV $\cdot k_{ij}$ (p.u.)	branch	From (i)	To (j)	GHV $\cdot k_{ij}$ (p.u.)
8	6	11	3.000	14	16	21	3.000
9	2	26	3.500	15	34	16	3.000
10	15	11	2.500	16	14	34	3.500
11	26	37	3.000	17	83	69	3.000
12	4	21	2.000	18	15	69	2.000
13	54	5	2.000				

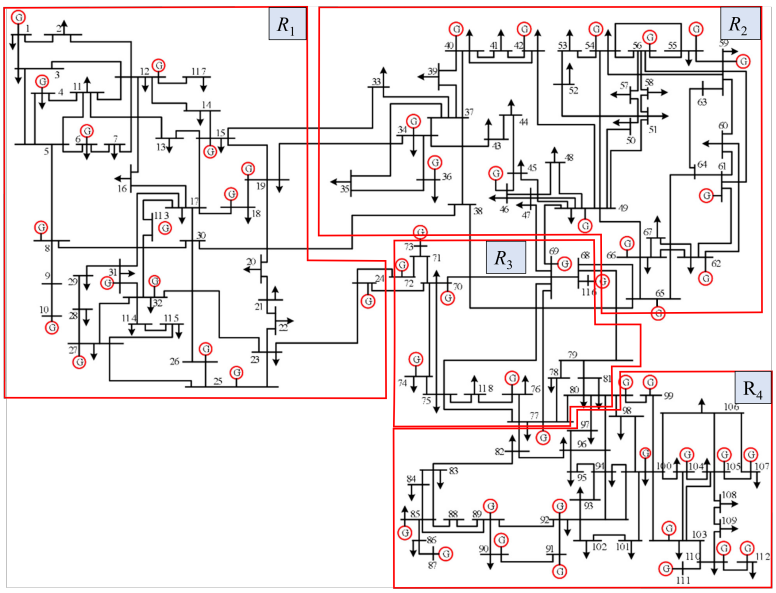


Figure 7: 118-bus power network.

Based on different demand-side carbon emission prices, a comparative analysis is conducted to evaluate the impact of incorporating demand-side carbon emission costs on the overall system carbon emissions, as shown in Table 8. It can be observed that by optimizing the demand-side carbon emission costs, the carbon emissions $\sum_{i\in\mathcal{N}} E_i L_i^e$ can be effectively reduced.

Table 8: Demand-side Carbon Emission Reduction Performance

π^d(\$/MW)	0.00	10.00	50.00	100.00	150.00	200.00
Carbon emission(kg/CO_2)	1073.61	1067.48	1032.40	986.20	964.72	917.03
Carbon emission reduction	0.00%	0.57%	3.99%	8.86%	11.29%	14.58%

5.3. Comparison of the Performance of Linearization Methods

To validate the advantages of the proposed dual linearization method for CEF, we conducted tests on the 118-bus Power and 21-node Gas System (Case 2). We compared the solution times and optimality (optimized cost) under different methods. The 30 test results are shown in Table 9.

In Case 2, which has a large number of nonlinear constraints, our method effectively overcoming the difficulties encountered when solve nonconvex constrain by using solvers or Jacobi iteraion method [14] both in centralized and distributed optimization for CEF in MES. In distributed optimization, compared to the LADMM [29], our method (DR-ADMM) reduces the solution time by 25.64% in Case 2.

Furthermore, to eliminate the influence of coincidence on optimality, The random disturbance with a uniform distribution ranging from $-0.1p.u.$ to $0.1p.u.$ has been added to the load demand of the Case 2 system, generating 50 different load demand scenarios. The comparison of the solution

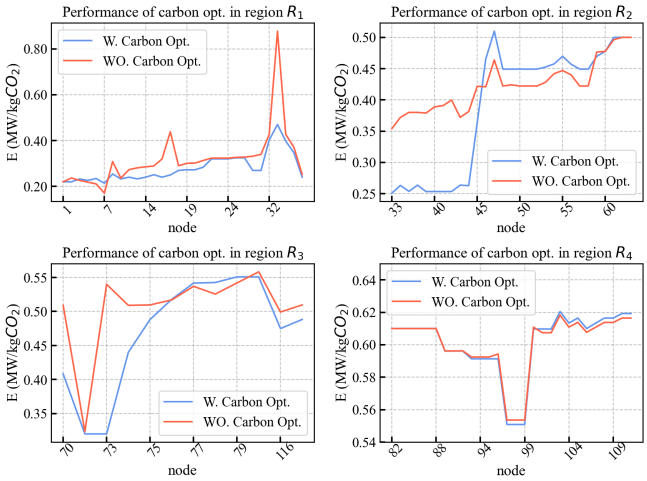


Figure 8: Performance of carbon optimization in all regions.

Table 9: Comparison of diffrent linearization methods in case2

	Method	Time (s)			Cost
		Min.	Aver.	Max.	($/hr)
Cent. Opt.	SDTE	156.21	157.45	158.91	133270.90
	Jacobi Iter.	>3000	>3000	>3000	N/A
	Gurobi	>3000	>3000	>3000	N/A
Dist. Opt.	DR-ADMM	986.59	987.59	989.52	133661.03
	LADMM	1326.65	1328.12	1329.79	133748.19
	Gurobi	>3000	>3000	>3000	N/A

results of the proposed method (DR-ADMM) and LADMM is shown in Fig. 9. It can be observed that our method yields solutions with better optimality.

5.4. Computation Performance of Privacy-preserving Framework

Due to the influence of random numbers in privacy-preserving, this section evaluates the statistical performance of the proposed Algorithm 3 over 30 test runs in the 9-bus Power and 8-node Gas System. The initial value of the quadratic penalty term ρ for each region is set to a random number between 0 and 0.5, and the coefficient α is set to a random number between 1 and 1.2. ρ and α are independently generated for each region.

Finally, a simulation is conducted, comparing the results of Algorithm 3 with Algorithm 2 which also uses random initial values for ρ between 0 and 0.5, and with Algorithm 2 which employs the PPOPF encryption framework [32]. The number of iterations and the optimality gap RE_{cost} from these tests are presented as box plots in Fig. 10.

It can be observed that Algorithm 3, which can dynamically update the coefficient ρ, requires fewer iterations compared to the ADMM algorithm based on the PPOPF encryption framework, while still maintaining an acceptable optimality gap RE_{cost}.

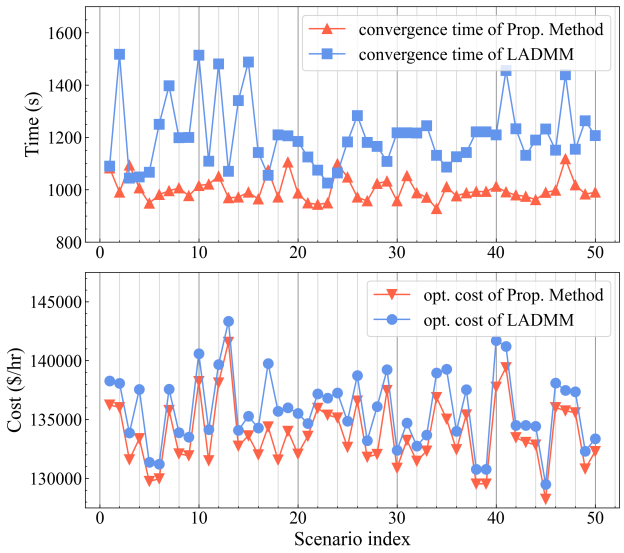


Figure 9: Comparison of the speed and Optimality of different ADMM Methods.

6. Conclusions

This work proposes a distributed approach to reduce carbon emissions in multi-energy systems. Carbon flow and nodal carbon intensity are employed to regulate carbon emissions at the nodal level. To address the nonconvexity constraints, a general linearized iterative method is proposed based on the successful dual Taylor expansion (SDTE). To accelerate the solution performance, this work proposes a general acceleration technique within ADMM framework. To protect privacy for different stakeholders in multi-energy systems, an information-exchange mechanism is proposed based on partially homomorphic encryption. The case studies show that the proposed approach outperforms the commercial solvers, and the privacy can be effectively protected.

In future research, we plan to investigate demand-side management with the proposed LCEF model, focusing on its impact on electricity consumption patterns. Furthermore, we will investigate decentralized optimization problems in the carbon markets.

References

[1] W. Huang, N. Zhang, Y. Cheng, J. Yang, Y. Wang, C. Kang, Multienergy networks analytics: Standardized modeling, optimization, and low carbon analysis, Proceedings of the IEEE 108 (2020) 1411–1436. doi:https://doi.org/10.1109/JPROC.2020.2993787.

[2] P. Favre-Perrod, A vision of future energy networks, in: 2005 IEEE Power Engineering Society Inaugural Conference and Exposition in Africa, Durban, South Africa, 2005, pp. 13–17. doi:https://doi.org/10.1109/PESAFR.2005.1611778.

[3] S. Rehman, M. Ahmed, M. H. Mohamed, F. A. Al-Sulaiman, Energy hub: From a model to a concept - a review, Renewable and Sustainable Energy Reviews 80 (2017) 1512–1527. doi:https://doi.org/10.1016/j.rser.2017.05.218.

[4] X. Zhang, M. Shahidehpour, A. Alabdulwahab, A. Abusorrah, Optimal expansion planning of energy hub with multiple energy infrastructures, IEEE Transactions on Smart Grid 6 (5) (2015) 2302–2311. doi:https://doi.org/10.1109/TSG.2015.2390640.

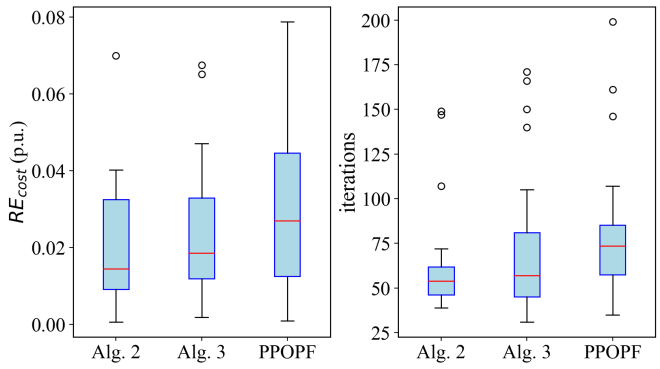


Figure 10: Comparison of optimality gap and number of iterations for different algorithms. In this box plot, the red line represents the median; the blue box's upper and lower edges correspond to the upper and lower quartiles, respectively. The black lines extending above and below the box represent the upper and lower boundaries, which are no more than 1.5 times the the range of the blue box above the upper quartile and below the lower quartile, respectively. The small circles indicate outliers that fall outside these boundaries.

[5] M. Geidl, G. Andersson, Optimal power flow of multiple energy carriers, IEEE Transactions on Power Systems 22 (1) (2007) 145–155. doi:https://doi.org/10.1109/TPWRS.2006.888988.

[6] Y. Xi, L. Zuo, M. Chen, J. Zhang, L. Cheng, I. Lestas, Energy-carbon pricing-guided collaborative optimization for local integrated energy communities with multi-vector electrification consumers, Applied Energy 392 (2025) 125946. doi:https://doi.org/10.1016/j.apenergy.2025.125946.

[7] T. Zhou, C. Kang, Q. Xu, Q. Chen, Preliminary investigation on a method for carbon emission flow calculation of power system, Automation of Electric Power Systems 36 (11) (2012) 44–49. doi:https://doi.org/10.3969/j.issn.1000-1026.2012.11.008.

[8] W. Huang, N. Zhang, Y. Cheng, J. Yang, Y. Wang, C. Kang, Multienergy networks analytics: Standardized modeling, optimization, and low carbon analysis, Proceedings of the IEEE 108 (9) (2020) 1411–1436. doi:https://doi.org/10.1109/JPROC.2020.2993787.

[9] Y. Cheng, N. Zhang, Y. Wang, J. Yang, C. Kang, Q. Xia, Modeling carbon emission flow in multiple energy systems, IEEE Transactions on Smart Grid 10 (4) (2019) 3562–3574. doi:https://doi.org/10.1109/TSG.2018.2830775.

[10] Y. Cheng, N. Zhang, Z. Lu, C. Kang, Planning multiple energy systems toward low-carbon society: A decentralized approach, IEEE Transactions on Smart Grid 10 (5) (2019) 4859–4869. doi:https://doi.org/10.1109/TSG.2018.2870323.

[11] Y. Wang, J. Qiu, Y. Tao, Optimal power scheduling using data-driven carbon emission flow modelling for carbon intensity control, IEEE Transactions on Power Systems 37 (4) (2022) 2894–2905. doi:https://doi.org/10.1109/TPWRS.2021.3126701.

[12] P. Feng, X. He, Mixed neurodynamic optimization for the operation of multiple energy systems considering economic and environmental aspects, Energy 232 (2021) 120965. doi:https://doi.org/10.1016/j.energy.2021.120965.

[13] L. Sang, Y. Xu, H. Sun, Encoding carbon emission flow in energy management: A compact constraint learning approach, IEEE Transactions on Sustainable Energy 15 (1) (2024) 123–135. doi:https://doi.org/10.1109/TSTE.2023.3274735.

[14] Y. Li, Y. Li, Y. Liu, Y. Wang, W. Bao, T. Zhou, C. Kang, Distributed carbon meter system based on iterative calculation method of carbon emission flow (i): theoretical method and analysis, Power System Technology 47 (6) (2023) 2165–2174. doi:https://doi.org/10.13335/j.1000-3673.pst.2023.0476.

[15] L. Zhang, H. Ye, F. Ding, Z. Li, M. Shahidehpour, Increasing pv hosting capacity with an adjustable hybrid power flow model, IEEE Transactions on Sustainable Energy 14 (1) (2023) 409–422. doi:https://doi.org/

10.1109/TSTE.2022.3215287.
[16] L. Yang, Y. Xu, H. Sun, X. Zhao, Two-stage convexification-based optimal electricity-gas flow, IEEE Transactions on Smart Grid 11 (2) (2020) 1465–1475. doi:https://doi.org/10.1109/TSG.2019.2938553.
[17] K. Qu, S. Shi, T. Yu, W. Wang, A convex decentralized optimization for environmental-economic power and gas system considering diversified emission control, Applied Energy 240 (2019) 630–645. doi:https://doi.org/10.1016/j.apenergy.2019.02.038.
[18] D. Xu, Q. Wu, B. Zhou, C. Li, L. Bai, S. Huang, Distributed multi-energy operation of coupled electricity, heating, and natural gas networks, IEEE Transactions on Sustainable Energy 11 (4) (2020) 2457–2469. doi:https://doi.org/10.1109/TSTE.2019.2961432.
[19] T. Ding, W. Jia, M. Shahidehpour, O. Han, Y. Su, Z. Zhang, Review of optimization methods for energy hub planning, operation, trading, and control, IEEE Transactions on Sustainable Energy 13 (3) (2022) 1802–1818. doi:https://doi.org/10.1109/TSTE.2022.3172004.
[20] J. Qin, Y. Wan, X. Yu, Y. Kang, A newton method-based distributed algorithm for multi-area economic dispatch, IEEE Transactions on Power Systems 35 (2) (2020) 986–996. doi:https://doi.org/10.1109/TPWRS.2019.2943344.
[21] S. J. Wright, Primal-Dual Interior-Point Methods, Society for Industrial and Applied Mathematics, 1997. doi:https://doi.org/10.1137/1.9781611971453.
[22] T. Xu, W. Wu, W. Zheng, H. Sun, L. Wang, Fully distributed quasi-newton multi-area dynamic economic dispatch method for active distribution networks, IEEE Transactions on Power Systems 33 (4) (2018) 4253–4263. doi:https://doi.org/10.1109/TPWRS.2017.2771950.
[23] Z. Song, L. Shi, S. Pu, M. Yan, Optimal gradient tracking for decentralized optimization, Mathematical Programming 207 (2024) 1–53. doi:https://doi.org/10.1007/s10107-023-01997-7.
[24] D. Ma, M. Xing, Y. Li, Q. Sun, Distributed economic dispatch with dynamic power demand: An implicit dual gradient tracking algorithm under random-triggered transmission protocol, IEEE Transactions on Power Systems 40 (2) (2025) 1931–1942. doi:https://doi.org/10.1109/TPWRS.2024.3447089.
[25] W. Shi, Q. Ling, G. Wu, W. Yin, Extra: An exact first-order algorithm for decentralized consensus optimization, SIAM Journal on Optimization 25 (2) (2015) 944–966. doi:https://doi.org/10.1137/14096668X.
[26] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Foundations and Trends® in Machine Learning 3 (1) (2011) 1–122. doi:https://doi.org/10.1561/2200000016.
[27] T. Erseghe, Distributed optimal power flow using admm, IEEE Transactions on Power Systems 29 (5) (2014) 2370–2380. doi:https://doi.org/10.1109/TPWRS.2014.2306495.
[28] K. Sun, X. Sun, A two-level admm algorithm for ac opf with global convergence guarantees, IEEE Transactions on Power Systems 36 (6) (2021) 5271–5281. doi:https://doi.org/10.1109/TPWRS.2021.3078001.
[29] X. Yi, S. Zhang, T. Yang, T. Chai, K. Johansson, Sublinear and linear convergence of modified admm for distributed nonconvex optimization, IEEE Transactions on Control of Network Systems 10 (1) (2023) 75–86. doi:https://doi.org/10.1109/TCNS.2022.3186653.
[30] J. Su, S. Yu, B. Li, Y. Ye, Distributed and collective intelligence for computation offloading in aerial edge networks, IEEE Transactions on Intelligent Transportation Systems 24 (7) (2023) 7516–7526. doi:https://doi.org/10.1109/TITS.2022.3160594.
[31] M. Ryu, K. Kim, A privacy-preserving distributed control of optimal power flow, IEEE Transactions on Power Systems 37 (3) (2022) 2042–2051. doi:https://doi.org/10.1109/TPWRS.2021.3120056.
[32] T. Wu, C. Zhao, Y.-J. A. Zhang, Privacy-preserving distributed optimal power flow with partially homomorphic encryption, IEEE Transactions on Smart Grid 12 (5) (2021) 4506–4521. doi:https://doi.org/10.1109/TSG.2021.3084934.
[33] Z.-P. Yuan, P. Li, Z.-L. Li, J. Xia, A fully distributed privacy-preserving energy management system for networked microgrid cluster based on homomorphic encryption, IEEE Transactions on Smart Grid 15 (2) (2024) 1735–1748. doi:https://doi.org/10.1109/TSG.2023.3309405.
[34] Y. Wang, W. Yin, J. Zeng, Global convergence of admm in nonconvex nonsmooth optimization, Journal of Scientific Computing 78 (2019) 29–63. doi:10.1007/s10915-018-0757-z.