Closed-Loop Hierarchical Operation for Optimal Unit Commitment and Dispatch in Microgrids: A Hybrid System Approach

Fangyuan Li, *Student Member, IEEE*, Jiahu Qin, *Senior Member, IEEE*, and Yu Kang, *Senior Member, IEEE*

Abstract—To contribute to a smart grid, the unit commitment and dispatch (UCD) is a very important problem to be revisited in the future microgrid environment. In this paper, the UCD problem is studied by taking account of both traditional thermal units and distributed generation units (DGs) while considering demand response (DR) and carbon emissions trading (CET). Then, the UCD problem is reformulated as the optimal switching problem of a hybrid system. The reformulation allows convenient handling of time-dependent start-up cost for dynamic programming (DP). Afterwards, a closed-loop hierarchical operation (CLHO) algorithm is developed which is capable of producing the optimal schedule in the closed-loop behavior. Theoretical analysis is also presented to show the equivalence of the reformulation and the effectiveness of the CLHO algorithm. Finally, the closed-loop behavior and the effectiveness of the algorithm are further verified by simulation studies. It is also shown that low emission quotas and high emission trading price help reduce the total carbon emissions.

Index Terms—Approximation methods, closed loop systems, optimal control, power generation dispatch.

I. INTRODUCTION

The microgrid is a promising concept to help integrating various distributed generation units (DGs) in smart grid. To cope with the new factors such as DGs and demand response (DR) in microgrid, the unit commitment and dispatch (UCD) is an important problem that should be revisited. The objective of UCD is to schedule the generation units such that the total generation cost over a time horizon is minimized while satisfying power balance constraints and various operational constraints [1].

Traditional methods for solving UCD problem include priority-list (PL) [2], Lagrangian relaxation (LR) [3], and various heuristic search algorithms [4]. In PL method, the shutdown rule or priority-list scheme acts a key role in deciding which unit to commit or drop in order to satisfy the increasing or dropping demand. The PL methods are usually simple and fast but result in relatively high total generation costs. The LR method is based on using Lagrange multipliers to incorporate the constraints onto the total cost, which results in the so-called Lagrangian. The UCD problem is converted to find the primal and dual solution of the Lagrangian. It is relatively easy to add constraints in the Lagrangian for LR method while the main disadvantage is the inherent suboptimality [1]. The

F. Li, J. Qin, and Y. Kang are with the Department of Automation, University of Science and Technology of China, Hefei 230027, China (e-mail:fylimas@mail.ustc.edu.cn; jhqin@ustc.edu.cn;kangduyu@ustc.edu.cn).

dynamic programming (DP) method is also widely employed in solving UCD problem. The distinct advantage of DP method is the ability to find the optimal schedule for the UCD problem. However, it is generally difficult for the DP method to deal with the time-dependent start-up costs and constraints [5]. The heuristic search methods including particle swarm algorithm, evolution programming, and genetic algorithm have been continuously studied to solve UCD problem. The main advantage of these methods is the capability of solving large scale problems under various constraints. However, there is no general rule of tuning the parameters of the algorithms [4]. It is worth noting that though the UCD problem has been extensively studied in various methods, the problem remains challenging due to the nonlinearity and the presence of both integer and continuous variables.

The recent advances in reinforcement learning (RL) [6], [7] and DP [8], [9] areas are providing new promising methods for solving some optimization problems in power systems. A series of successful applications have been made to solve various economic power dispatch problems (ED) which contain no integer variables, see [10]–[13] for examples. However, relatively little works exploit these advances to solve the more challenging UCD problem where integer variables are also involved. In [14], model predictive control is applied to solve the UCD problem with intermittent resources while the start-up and shutdown costs are not considered. Reference [15] employs state-space approximate dynamic programming (ADP) for solving the stochastic UCD under the framework of Markov decision process. However, only the commitment status is treated as decision variables in the UCD problem without taking account of power generations. Recently, reference [16] proposes a consensus based distributed reinforcement learning algorithm to solve the non-convex ED problem considering valve-point effects. The algorithm provided in [16] can also be extended to solve the UCD problem. However, a drawback of [16] is the requirement of the discrete generation output which may introduce discretization errors. Nevertheless, these works have not taken account of the effect of the DR and carbon emissions trading (CET).

In this paper, the deterministic UCD problem is investigated in the microgrid environment consisting of thermal units and DGs while considering DR and CET. The contributions are listed as follows. 1) Besides the traditional thermal units, the DGs, DR, and CET are also considered. With the increasing penetration of renewable generation and pressure of emissions reduction, the UCD problem considered in this paper makes a

more practical scenario than existing works such as [14]–[16]. 2) The UCD problem is firstly reformulated as the optimal switching problem of a hybrid system. In the reformulation, the time-dependent start-up cost is converted to the switching cost of the hybrid system, enabling the dealt of this cost which is generally seen as difficult for DP [5]. The reformulation also allows exploiting advances in related areas [6]–[9] to solve the UCD problem. 3) A closed-loop hierarchical operation (CLHO) algorithm is developed to find the optimal schedule of the hybrid system. The proposed CLHO algorithm is capable of producing the optimal schedule in a closed-loop style, where the trained neural network can deal with different initial conditions and disturbances without retraining. In contrast, most existing methods are usually open-loop, thus cannot guarantee optimality under external disturbances without reoptimization. Theoretical analysis are also provided as opposed to most existing works such as [14], [15], [17]. 4) Simulation results verify that the CLHO algorithm produces the optimal schedule in closed-loop under different initial conditions and external disturbance, which potentially helps the real-time optimal operation of the microgrids. It is also shown in the simulation that low emission quotas and high emission trading price reduce the total amount of carbon emissions.

The paper is organized as follows. Section II provides some preliminaries. The UCD problem is reformulated as an optimal switching problem of a hybrid system in Section III. Section IV develops the CLHO algorithm along with theoretical analysis. In Section V, simulations are conducted to study the behavior and effectiveness of the proposed algorithm. Finally, we draw conclusions in Section VI.

II. PRELIMINARY

In this section, some preliminaries on CET and microgrid environment are presented.

A. Carbon Emissions Trading

CET is a cap-and-trade system where the government sets a cap of carbon emissions and issues quotas for each emitter. The quotas can be seen as allowances for emitters to emit certain amount of greenhouse gas. The emitters can trade quotas in the CET market as needed, i.e., sell surplus or purchase deficit of emission allowances. In the UCD problem in this paper, the emitters are thermal units. As argued in paper [18], we assume that the trade of generation units in UCD does not affect the trading price. The emission cost of thermal units is determined by

$$C_E = \sum_{n=1}^{N} \left(\sum_{t=1}^{T} E_n(P_n[t]) I_n[t] - Q_n \right) \cdot p_e, \qquad (1)$$

$$E_n(P_n[t]) = \alpha_n(P_n[t])^2 + \beta_n P_n[t] + \gamma_n, \qquad (2)$$

where $E_n(P_n[t])$ is the emission of thermal unit n at period t, Q_n is the emission quota of unit n, p_e is the CET price in the market, α_n, β_n, and γ_n are the emission coefficients.

B. Microgrid Environment

1) Thermal Units: The fuel cost of thermal unit n is usually modeled as a quadratic function [19]:

$$C_{F,n}(P_n[t]) = a_n(P_n[t])^2 + b_n P_n[t] + c_n, \qquad (3)$$

where $P_n[t]$ is the active power output of unit n at period t, a_n, b_n, and c_n are the cost coefficients of unit n.

To bring the thermal unit online, a certain amount of energy must be spent, incurring a start-up cost. Assume that the off-line thermal unit n is in banking mode, then the start-up cost is usually modeled as [1]

$$C_{U,n}(\tau_n) = C_{b,n} \cdot \tau_n + C_{f,n}, \qquad (4)$$

where $C_{b,n}$ is the cost for unit n to maintain the operating temperature, τ_n is the period that unit n was in banking mode, and $C_{f,n}$ represents the fixed cost. To decommit a thermal unit, a fixed shutdown cost $C_{D,n}$ is usually incurred.

Define $I_n[t]$ as the commitment decision variable of unit n at time t, where $I_n[t] = 1$ if unit n is committed/online at t and $I_n[t] = 0$ if otherwise. The power output of unit n usually subjects to the following generation capacity constraints:

$$I_n[t]P_n^m \le P_n[t] \le I_n[t]P_n^M, \qquad (5)$$

where P_n^m (P_n^M) is the minimum (maximum) power output of unit n.

The power output variation of thermal units usually subject to the following ramp limit constraint

$$-R_n^D \le (P_n[t] - P_n[t-1])I_n[t]I_n[t-1] \le R_n^U, \qquad (6)$$

where R_n^D and R_n^U are the ramp down and ramp up rate limit.

2) Distributed Generation: Suppose there is an aggregator that manages the DGs such as wind turbines and solar panels, then the aggregator can be treated as a virtual generation unit with cost function [18]

$$C_{DG}(P_{DG}[t]) = a_{DG}(P_{DG}[t])^2 + b_{DG}P_{DG}[t] + c_{DG}, \qquad (7)$$

where a_{DG}, b_{DG}, and c_{DG} are the cost coefficients, and $P_{DG}[t]$ is the aggregated generation.

Since that DGs subject to natural conditions and tend to be intermittent, upper limits on available DGs and their penetration rate are usually imposed as following [18], to ensure a reliable operation

$$0 \le P_{DG}[t] \le P_{DG}^M[t], \qquad (8)$$

and

$$\eta_t = \frac{P_{DG,t}}{\sum_{n=1}^{N} P_n[t] + P_{DG,t}} \le \eta_{\max}. \qquad (9)$$

3) Demand Response: DR captures the relation between the load elasticity and the electricity price. The demand curtailment due to DR can be treated as a virtual generation unit [20]. The cost function of this virtual generation unit can be modeled as follows [20]

$$C_{DR}(P_{DR}[t]) = a_{DR}(P_{DR}[t])^2 + b_{DR}P_{DR}[t] + c_{DR}, \qquad (10)$$

where a_{DR}, b_{DR}, and c_{DR} are the cost coefficients, and $P_{DR}[t]$ is the curtailed demand.

There is an upper limit on the demand curtailment at each period t as follows

$$0 \le P_{DR}[t] \le P_{DR}^M[t]. \qquad (11)$$

4) Operational Constraints: The operation of microgrid subject to the power balance constraints [21]

$$\sum_{n=1}^{N} P_n[t]I_n[t]+P_{\mathrm{DG}}[t]+P_{\mathrm{DR}}[t] = P_{\mathrm{D}}[t], \quad t = 1,...,T. \quad (12)$$

where $P_{\mathrm{D}}[t]$ is the power demand, N is the total number of thermal units, and T is the total number of time periods. Since an isolated microgrid environment is considered in this paper, there is no power exchange with the main grid in (12).

To cope with the fluctuation of renewable energies and demand, the following spinning reserves are usually considered

$$\sum_{n=1}^{N} P_n^m I_n[t]+P_{\mathrm{DG}}[t]+P_{\mathrm{DR}}[t]-P_{\mathrm{D}}[t] \le -R^m[t], \quad (13)$$

$$\sum_{n=1}^{N} P_n^M I_n[t]+P_{\mathrm{DG}}[t]+P_{\mathrm{DR}}[t]-P_{\mathrm{D}}[t] \ge R^M[t], \quad (14)$$

where $R^m[t] > 0$ and $R^M[t] > 0$ are the lower and upper spinning reserves.

5) Power Flow Constraints: Note that the primary focus of this paper is solving the UCD problem, although the optimal power flow problem is also related. Hence, the transmission losses and power flow constraints are not incorporated explicitly in the following UCD model. See [22] for various power flow representations. However, it is worth noting that both the power flow constraints (linear approximations or convex approximations) and the transmission losses can be added. The proposed algorithm is still valid with appropriate decision variables since the lower level problem remains convex.

III. PROBLEM FORMULATION

In this section, the UCD problem is presented in a commonly used form as in literature [5], [18]. Then, the problem is decomposed and reformulated into an optimal switching problem of a hybrid system.

A. Unit Commitment and Dispatch Model

The UCD model considering DR, DG, and CET can be formulated as follows [1], [5]:

$$\min \sum_{t=1}^{T} \Big(\sum_{n=1}^{N} (C_{\mathrm{F},n}(P_n[t])I_n[t]+C_{\mathrm{U},n}[t]+C_{\mathrm{D},n}[t])$$
$$+ C_{\mathrm{DG}}(P_{\mathrm{DG}}[t])+C_{\mathrm{DR}}(P_{\mathrm{DR}}[t]) \Big) + C_{\mathrm{E}} \quad (15)$$

$$\text{s.t.} \sum_{n=1}^{N} P_n[t]I_n[t]+P_{\mathrm{DG}}[t]+P_{\mathrm{DR}}[t] = P_{\mathrm{D}}[t], \quad (16)$$

$$\sum_{n=1}^{N} P_n^m I_n[t]+P_{\mathrm{DG}}[t]+P_{\mathrm{DR}}[t]$$
$$\le P_{\mathrm{D}}[t]-R^m[t], \quad (17)$$

$$\sum_{n=1}^{N} P_n^M I_n[t]+P_{\mathrm{DG}}[t]+P_{\mathrm{DR}}[t]$$
$$\ge P_{\mathrm{D}}[t]+R^M[t], \quad (18)$$

$$I_n[t]P_n^m \le P_n[t] \le I_n[t]P_n^M, \quad (19)$$
$$-R_n^{\mathrm{D}} \le (P_n[t]-P_n[t-1])I_n[t]I_n[t-1] \le R_n^{\mathrm{U}}, \quad (20)$$
$$0 \le P_{\mathrm{DG}}[t] \le P_{\mathrm{DG}}^M[t], \quad (21)$$
$$\eta_t = \frac{P_{\mathrm{DG}}[t]}{\sum_{n=1}^{N} P_n[t]+P_{\mathrm{DG}}[t]} \le \eta_{\max} \quad (22)$$
$$0 \le P_{\mathrm{DR}}[t] \le P_{\mathrm{DR}}^M[t], \quad (23)$$

where the constraints (16)–(23) are imposed for all $n = 1,...,N$ and $t = 1,...,T$. $C_{\mathrm{U},n}[t]$ is computed by (4) if $(1-I_n[t-1])I_n[t] = 1$, otherwise, $C_{\mathrm{U},n}[t] = 0$. $C_{\mathrm{D},n}[t]$ equals $C_{\mathrm{D},n}$ if $I_n[t-1](1-I_n[t]) = 1$, otherwise, $C_{\mathrm{D},n}[t] = 0$. It is assumed that all dispatchable thermal units and nondispatchable DGs belong to a utility microgrid.

B. Model Decomposition

Substituting (1) into objective (15), one obtains

$$\sum_{t=1}^{T} \Big(\sum_{n=1}^{N} \big(I_n[t](C_{\mathrm{F},n}(P_n[t])+E_n(P_n[t]))-Q_n p_e\big)+C_{\mathrm{DG}}(P_{\mathrm{DG}}[t])$$
$$+ C_{\mathrm{DR}}(P_{\mathrm{DR}}[t]) \Big) + \sum_{t=1}^{T}\sum_{n=1}^{N}(C_{\mathrm{U},n}+C_{\mathrm{D},n}). \quad (24)$$

The first item in (24) incorporates all the fuel costs, emissions costs, and the costs of DGs and DR of each time period t, while the second item summarizes all the costs incurred by the commitment or decommitment of thermal units along the time horizon T. With this observation, we denote the first and the second item compactly as

$$\sum_{t=1}^{T} Q(P[t], I[t]) := \sum_{t=1}^{T} \Big(\sum_{n=1}^{N} \big(I_n[t](C_{\mathrm{F},n}(P_n[t])+E_n(P_n[t]))$$
$$- Q_n p_e\big) + C_{\mathrm{DG}}(P_{\mathrm{DG}}[t])+C_{\mathrm{DR}}(P_{\mathrm{DR}}[t]) \Big) \quad (25)$$

and

$$\sum_{t=1}^{T} \kappa(I[t-1], I[t]) := \sum_{t=1}^{T}\sum_{n=1}^{N}(C_{\mathrm{U},n}+C_{\mathrm{D},n}) \quad (26)$$

where $P[t]$ and $I[t]$ are vector $[P_1[t],...,P_N[t],P_{\mathrm{DG}}[t],P_{\mathrm{DR}}[t]]^T$ and $[I_1[t],...,I_N[t]]^T$, respectively. In (26), $I[0]$ represents the initial commitment status of the thermal units. Then, the objective (15) can be denoted as

$$\sum_{t=1}^{T} Q(P[t], I[t]) + \kappa(I[t-1], I[t]). \quad (27)$$

Under certain conditions, the original UCD problem can be seen as a hierarchical optimization problem, where the upper level objective is to minimize (27) along T time horizon and the lower level objective is to minimize (25) subjecting to constraints (16)–(23) at each time t.

C. Hybrid System Model

It can be verified that the start-up and shutdown cost $\kappa(I[t-1], I[t])$ incurred at each time period t can be written as

$$\kappa(I[t-1], I[t]) = \sum_{n=1}^{N} \kappa_n(I[t-1], I[t]) \quad (28)$$

where $\kappa_n(I[t-1], I[t])$ is calculated by

$$\kappa_n(I[t-1], I[t]) = C_{\text{b},n} + (C_{\text{f},n} - C_{\text{b},n} + C_{\text{D},n})I_n[t-1] \\ - (C_{\text{f},n} + C_{\text{D},n})I_n[t-1]I_n[t]. \quad (29)$$

The introduction of (28) and (29) eliminates the need of an implicit variable τ_n to keep track of the banking time as in literature [1], [5]. Since the cost is incurred by switching from $I[t-1]$ to $I[t]$, we call $\kappa(I[t-1], I[t])$ switching cost.

Note that there are two exogenous variables $P_{\text{DG}}^M[t]$ and $P_{\text{D}}[t]$ in the UCD problem. In the day-ahead UC, the short-time renewable power and load forecasting are usually assumed available, which can be denoted as

$$P_{\text{DG}}^M[t] = F_1(\bar{P}_{\text{DG}}^M[t-1]; \theta_1), \quad (30)$$

and

$$P_{\text{D}}[t] = F_2(\bar{P}_{\text{D}}[t-1]; \theta_2), \quad (31)$$

where $F_{(\cdot)}$ are the forecasting models, $\bar{P}_{\text{DG}}^M[t]$ and $\bar{P}_{\text{D}}[t]$ are the regression vectors, and $\theta_{(\cdot)}$ are model parameters. Interested readers are directed to [23] and references therein for details on the establishment of these models. Then, the lower level optimization problem can be rewritten as

$$\begin{cases} P[t] = \arg\min Q(P[t], I[t]) \\ \text{s. t. } (16), (17), (18), (19), (20), (21), (22), (23) \end{cases}. \quad (32)$$

Denote (32) as

$$P[t] = f_I(P[t-1], P_{\text{DG}}^M[t], P_{\text{D}}[t]), \quad (33)$$

where subscript $I \in \{0,1\}^N$ is the index of subsystems. Note that for some $I \in \{0,1\}^N$ with $P_{\text{DG}}^M[t]$, $P_{\text{D}}[t]$, $R^m[t]$, $R^M[t]$, and $P_{\text{DR}}^M[t]$, the lower optimization problem may admit no solutions. To eliminate these exceptions, we use $\mathcal{I}(P_{\text{DG}}^M[t], P_{\text{D}}[t], R^m[t], R^M[t], P_{\text{DR}}^M[t])$, in shorthand as $\mathcal{I}_t$ for given variable values, to denote the set of I when the lower optimization problem admits a solution $P[t] \in \mathcal{P}_t$. $\mathcal{P}_t$ denotes the set of feasible solutions.

Finally, the original UCD model is reformulated as minimizing the objective

$$\sum_{t=1}^{T} Q(P[t], I[t]) + \kappa(I[t-1], I[t]). \quad (34)$$

while subjecting to a hybrid system as follows

$$\begin{cases} P_{\text{DG}}^M[t] = F_1(\bar{P}_{\text{DG}}^M[t-1]; \theta_1), \\ P_{\text{D}}[t] = F_2(\bar{P}_{\text{D}}[t-1]; \theta_2), \\ P[t] = f_I(P[t-1], P_{\text{DG}}^M[t], P_{\text{D}}[t]), \end{cases} \quad (35)$$

Note that the hybrid system (35) has $|\mathcal{I}|$ subsystems where $|\cdot|$ denotes cardinality. The subsystems are indexed by $I \in \mathcal{I}$. At each time t, only one subsystem $I[t] \in \mathcal{I}$ is running. The

objective is to choose a sequence of $I[t]$ such that (34) is minimized. Interested readers are directed to [24], [25] for more details on hybrid systems.

Since all the start-up and shutdown cost are incorporated in (34) without involving any implicit variables, easy handling of these costs is allowed which is generally seen as difficult for DP. In the next section, DP and approximation approach will be employed to solve the reformulated problem. To this end, we show the equivalence between the reformulated problem and the original UCD model. Before that, we have to show that the formulated hybrid system (35) is single-valued and admits a unique trajectory.

Theorem 1. *Given a (feasible) switching schedule $(I[t])_{t=1}^T$, the hybrid system (35) admits a unique trajectory $(P[t])_{t=1}^T$.*

Proof: The existence of a switching schedule $(I[t])_{t=1}^T$ implies that $\mathcal{I}_t \neq \varnothing$ for any $t = 1, ..., T$, which means that the lower level optimization problem (32) admits a solution at all $t = 1, ..., T$. Thus, the hybrid system admits some trajectory $(P[t])_{t=1}^T$. It remains to show that the trajectory is also unique.

Denote $P[0]$ and $I[0]$ the initial condition. Then, $P[1]$ is the optimal solution of problem (32) with $t = 1$ for $I[1]$ and initial condition $P[0]$ and $I[0]$. The objective $Q(P[1], I[1])$ is strictly convex since for all $n = 1, ..., N$, $C_{\text{F},n}(P_n[1])$, $E_n(P_n[1])$, $C_{\text{DG}}(P_{\text{DG}}[1])$, and $C_{\text{DR}}(P_{\text{DR}}[1])$ are strictly convex w.r.t. $P[1]$, and $Q_n p_e$ is constant. The feasible regions of all the constraints (16)–(23) are convex, leading to a convex intersection region. Then, the uniqueness of $P[1]$ follows. Similar argument can be applied to any $t = 1, ..., T$. This completes the proof. ∎

Once the initial $P[0]$ and $I[0]$ are given, and a switching schedule $\{I[t]\}_{t=1}^T$ is selected, the hybrid system can operate from the initial time $t = 1$ to the final time $t = T$, incurring running cost $Q(P[t], I[t])$ and switching cost $\kappa(I[t-1], I[t])$. The objective of the problem is to find an optimal switching schedule that minimizes the total cost (34).

Then we show that the hybrid system (35) minimizing cost (34) is equivalent to the original UCD with objective (15) and constraints (16)–(23) under certain conditions.

Theorem 2. *Assume that the original UCD problem with objective (15) and constraints (16)–(23) admits feasible solutions. Then, the hybrid system (35) w.r.t. cost function (34) admits an optimal switching schedule and a unique corresponding trajectory, denoted by $(I^*[t])_{t=1}^T$ and $(P^*[t])_{t=1}^T$. Moreover, if the ramp limit constraint (20) is relaxed, the optimal switching schedule $(I^*[t])_{t=1}^T$ along with the unique corresponding trajectory $(P^*[t])_{t=1}^T$ minimizes the original UCD problem.*

Proof: Firstly, we show that the hybrid system admits an optimal switching schedule. It is sufficient to show the system admits feasible switching schedules. Then the finiteness leads to the existence of an optimal switching schedule.

Denote the initial conditions by $P[0]$ and $I[0]$, and denote $(P[t])_{t=1}^T$ and $(I[t])_{t=1}^T$ an feasible solution of the original UCD problem. Then, $P[t]$ and $I[t]$ are also a feasible solution of the lower level optimization problem (32) with initial conditions $P[t-1]$ and $I[t-1]$ for any $t = 1, ..., T$. By

definition of hybrid system (35), $(P[t])_{t=1}^T$ is also a trajectory under the switching schedule $(I[t])_{t=1}^T$. Namely, the hybrid system admits switching schedule $(I[t])_{t=1}^T$. Then, it follows that the hybrid system admits an optimal switching schedule.

Denote $(I^*[t])_{t=1}^T$ as the optimal switching schedule. Then, the Theorem 1 indicates that the optimal switching schedule admits a unique trajectory, denoted by $(P^*[t])_{t=1}^T$. The above argument still holds if ramp limit constraint (20) is relaxed.

Secondly, we show that if ramp limit constraint is relaxed, the optimal schedule along with the unique corresponding trajectory minimizes the original UCD problem by contradiction.

Assume that the schedule $(I^*[t])_{t=1}^T$ along with trajectory $(P^*[t])_{t=1}^T$ does not minimize the original UCD problem. Then there exist some $(P'[t])_{t=1}^T$ and $(I'[t])_{t=1}^T$ such that

$$
\sum_{t=1}^T Q(P'[t], I'[t]) + \kappa(I'[t-1], I'[t])
$$
$$
< \sum_{t=1}^T Q(P^*[t], I^*[t]) + \kappa(I^*[t-1], I^*[t]). \quad (36)
$$

Note that $(P'[t])_{t=1}^T$ does not necessarily minimize the original problem nor the reformulated problem under the schedule $(I'[t])_{t=1}^T$. If it holds that $(I^*[t])_{t=1}^T = (I'[t])_{t=1}^T$, then (36) leads to $(P^*[t])_{t=1}^T = (P'[t])_{t=1}^T$ by contradiction. Notice that with (36), $(I^*[t])_{t=1}^T = (I'[t])_{t=1}^T$ leads to

$$
\sum_{t=1}^T Q(P'[t], I'[t]) < \sum_{t=1}^T Q(P^*[t], I^*[t]). \quad (37)
$$

Then, there must exists some $t = \tau \in \{1, ..., T\}$ such that

$$
Q(P'[t], I'[t]) = Q(P'[t], I^*[t]) < Q(P^*[t], I^*[t]). \quad (38)
$$

However, this contradicts the fact that $Q(P^*[t], I^*) \le Q(P'[t], I^*[t])$ for any $P'[t]$ according to (32) and (35) in which the ramp limit constraints have been relaxed. Thus, it holds that $Q(P'[t], I^*[t]) = Q(P^*[t], I^*[t])$ for all $t \in \{1, ..., T\}$, leading to $(P^*[t])_{t=1}^T = (P'[t])_{t=1}^T$ with Theorem 1.

Denote $(P''[t])_{t=1}^T$ as the system trajectory operated under switching schedule $(I'[t])_{t=1}^T$. If otherwise $(I^*[t])_{t=1}^T \ne (I'[t])_{t=1}^T$, then one obtains

$$
\sum_{t=1}^T Q(P''[t], I'[t]) \le \sum_{t=1}^T Q(P'[t], I'[t]) \quad (39)
$$

by observing (32) and (35) where the ramp limit constraints have been relaxed. Thus, it holds that

$$
\sum_{t=1}^T Q(P''[t], I'[t]) + \kappa(I'[t-1], I'[t])
$$
$$
\le \sum_{t=1}^T Q(P'[t], I'[t]) + \kappa(I'[t-1], I'[t])
$$
$$
< \sum_{t=1}^T Q(P^*[t], I^*[t]) + \kappa(I^*[t-1], I^*[t]). \quad (40)
$$

That is, schedule $(I'[t])_{t=1}^T$ with trajectory $(P''[t])_{t=1}^T$ results in lower total cost than schedule $(I^*[t])_{t=1}^T$ with trajectory

$(P^*[t])_{t=1}^T$. Since $(I^*[t])_{t=1}^T$ has been assumed as an optimal switching schedule, contradiction is raised. ∎

Considering the above equivalence, efforts are made to find the optimal switching schedule of hybrid system (35) w.r.t. the cost function (34) in the next section.

IV. PROPOSED ALGORITHM

In this section, the optimal cost-to-go of the hybrid system is presented first, then the CLHO algorithm is developed where neural networks are employed to approximate the optimal cost-to-go.

A. Theory

Denote $J_t^{I[t:T]}(P[t], I[t-1])$ as the cost-to-go associated with $P[t]$ and $I[t-1]$ from time t to the final time T by following the schedule $I[t:T] := (I[t], ..., I[T])$. Similarly, denote the optimal cost-to-go of following the optimal schedule $I^*[t:T] := (I^*[t], ..., I^*[T])$ by $J_t^{I^*[t:T]}(P[t], I[t-1])$, written as $J_t^*(P[t], I[t-1])$ without ambiguity. Then, one has that

$$
J_t^*(P[t], I[t-1]) = \min_{\forall I[t:T]} J_t^{I[t:T]}(P[t], I[t-1]) \quad (41)
$$

for any $P[t] \in \mathcal{P}_t$ and $I[t-1] \in \mathcal{I}_t$.

Considering the cost function (34), the cost-to-go incurred by following schedule $I[t:T]$ is summarized as

$$
J_t^{I[t:T]}(P[t], I[t-1])
$$
$$
= \sum_{\tau=t}^T Q(P[\tau], I[\tau]) + \kappa(I[\tau-1], I[\tau])
$$
$$
= Q(P[t], I[t]) + \kappa(I[t-1], I[t])
$$
$$
+ J_{t+1}^{I[t+1:T]}(f_{I[t]}(P[t]), I[t]) \quad (42)
$$

for all $t = 1, ..., T-1$ and

$$
J_T^{I[T]}(P[T], I[T-1]) = Q(P[T], I[T]) + \kappa(I[T-1], I[T]). \quad (43)
$$

Define $J_{T+1}^{I[T+1:T]}(f_{I[T]}(P[T]), I[T]) = 0$, then the above representation (42) holds for all $t = 1, ..., T-1$ and $t = T$, i.e., (43) is a special form of (42). Similarly, the optimal cost-to-go from time t can be written as

$$
J_t^*(P[t], I[t-1]) = Q(P[t], I^*[t]) + \kappa(I[t-1], I^*[t])
$$
$$
+ J_{t+1}^*(f_{I^*[t]}(P[t]), I^*[t]) \quad (44)
$$

for all $t = 1, ..., T$ where $J_{T+1}^*(f_{I^*[T]}(P[T]), I^*[T]) = 0$.

The Bellman's principle of optimality gives

$$
J_t^*(P[t], I[t-1]) = \min_{\forall I[t]} \Big(Q(P[t], I[t]) + \kappa(I[t-1], I[t])
$$
$$
+ J_{t+1}^*(f_{I[t]}(P[t]), I[t]) \Big) \quad (45)
$$

for any $P[t] \in \mathcal{P}_t$ and $I[t-1] \in \mathcal{I}_t$. Then, the optimal schedule $I^*[t]$ is given by

$$
I_t^*(P[t], I[t-1]) = \arg\min_{\forall I[t]} \Big(Q(P[t], I[t]) + \kappa(I[t-1], I[t])
$$
$$
+ J_{t+1}^*(f_{I[t]}(P[t]), I[t]) \Big). \quad (46)
$$

The above (46) implies that if one has calculated $J_t^*(P[t], I[t-1])$, then the optimal switching schedule $I^*[t] = I_t^*(P[t], I[t-1])$ can be obtained readily for any $P[t] \in \mathcal{P}_t$, $I[t-1] \in \mathcal{I}_t$, and $t = 1, ..., T$.

B. Closed-Loop Hierarchical Operation

Since it is difficult to solve the optimal cost-to-go $J_t^*(P[t], I[t-1])$ in closed form exactly from (45), approximation approach is employed in the following. Use a neural network to approximate the optimal cost-to-go as follows

$$\tilde{J}_t^*(P[t], I[t-1]) = (W_t^{I_{t-1}})^T \phi(P[t]) \approx J_t^*(P[t], I[t-1]),$$
$$\forall t = 1, ..., T, \quad \forall I_t \in \mathcal{I}_t, \quad \forall P[t] \in \mathcal{P}_t \quad (47)$$

where $\phi(\cdot) = [\phi_1(\cdot), ..., \phi_M(\cdot)]^T$ with each element $\phi_i(\cdot) : \mathbb{R}^{N+2} \to \mathbb{R}$ being a continuous basis function such as polynomials. $(W_t^{I_{t-1}})^T \in \mathbb{R}^M$ is the weight matrix, and M is the number of neurons. With approximated optimal cost-to-go, the optimal schedule $I^*[t]$ can be obtained immediately by

$$\tilde{I}_t^*(P[t], I[t-1]) := \arg\min_{\forall I[t]} \Big(Q(P[t], I[t]) + \kappa(I[t-1], I[t])$$
$$+ \tilde{J}_{t+1}^*(f_{I[t]}(P[t]), I[t]) \Big). \quad (48)$$

Then we have the following CLHO algorithm summarized in Algorithm 1. The following Theorem 3 shows that the optimal cost-to-go is continuous w.r.t. $P[t]$. Hence, the neural network approximation in (47) is possible according to Weierstrass's theorem and [26].

Algorithm 1: Closed-Loop Hierarchical Operation (CLHO) Algorithm

---

Data: Renewable power and load forecasting model (30) (31); Initial status of $P[0]$ and $I[0]$.

Result: Optimal schedule $(I^*[t])_{t=1}^T$ of hybrid system (35) w.r.t. objective (34).

For $\tau = 1, ..., T$, if $\tilde{J}_{\tau+1}^*(\cdot, \cdot)$ has been trained, go to *Scheduling Phase*, otherwise go to *Training Phase*. **Training Phase**

> Step 1: For $t = T, ..., \tau + 1$ repeat Step 2 and Step 3;
> Step 2: Select a set of $P[t] \in \mathcal{P}_t$ randomly;
> Step 3: Tune $W_t^{I_{t-1}}$ by using the method of least squares such that
>
> $$\tilde{J}_t^*(P[t], I[t-1]) := (W_t^{I_{t-1}})^T \phi(P[t])$$
> $$\approx \min_{\forall I[t]} \Big(Q(P[t], I[t]) + \kappa(I[t-1], I[t])$$
> $$+ \tilde{J}_{t+1}^*(f_{I[t]}(P[t]), I[t]) \Big). \quad (49)$$

End

Scheduling Phase

> Step 4: Input $P[\tau]$ and $I[\tau-1]$ of current time τ;
> Step 5: Calculate the optimal schedule $I^*[t]$ via
>
> $$I^*[\tau] := \tilde{I}_t^*(P[\tau], I[\tau-1])$$
> $$= \arg\min_{\forall I[\tau]} \Big(Q(P[\tau], I[\tau]) + \kappa(I[\tau-1], I[\tau])$$
> $$+ \tilde{J}_{\tau+1}^*(f_{I[\tau]}(P[\tau]), I[\tau]) \Big). \quad (50)$$

End

Theorem 3. *The optimal cost-to-go $J_t^*(P[t], I[t-1])$ of hybrid system* (35) *minimizing cost function* (34) *is continuous w.r.t. $P[t]$ for any $I[t-1] \in \mathcal{I}_{t-1}$ and $t = 1, ..., T$.*

To prove Theorem 3, we need to prove the following lemma first.

Lemma 1. *The dynamic $P[t] = f_I(P[t-1], P_{DG}^M[t], P_D[t])$ of the hybrid system* (35) *is continuous w.r.t. $P[t-1]$ for any feasible schedule $I \in \mathcal{I}_t$.*

Proof of Lemma 1: Note that if we can obtain the solution of optimization problem (32) in explicit form for any I, then the continuity of dynamic $f_I(\cdot, P_{DG}^M[t], P_D[t]) : \mathbb{R}^{N+2} \mapsto \mathbb{R}^{N+2}$ can be easily obtained by analyzing the explicit solution. However, it is difficult to obtain such an explicit solution. Thus, we turn to the conditions on which the solution (not necessarily in explicit form) can be guaranteed.

For any feasible $I \in \mathcal{I}_t$, function $Q(P[t], I)$ is convex w.r.t. $P[t]$ since $C_{F,n}(P_n[t])$, $E_n(P_n[t])$, $C_{DG}(P_{DG}[t])$, and $C_{DR}(P_{DR}[t])$ are convex for all n. The power balance constraint (16) is linear, and all the other constraints (17)–(23) are affine. Thus, the optimization problem (32) has strong duality. In addition, the function $Q(P[t], I)$ is differentiable w.r.t. $P[t]$ for any feasible $I \in \mathcal{I}_t$. The convex optimization theory [27] gives that a solution of (32) is optimal iff the solution satisfies the Karush-Kuhn-Tucker (KKT) conditions.

Write the optimization problem (32) in general form as

$$\min \ Q(P[t], I[t]) \quad (51)$$
$$\text{s.t. } h(P[t]) = 0, \ g(P[t]) \leq 0, \quad (52)$$

where $h(P[t]) := \sum_{n \in \mathcal{C}} P_n[t] + P_{DG}[t] + P_{DR}[t] - P_D[t]$, $g(P[t])$ is the stack vector of functions $\sum_{n \in \mathcal{C}} P_n^m + P_{DG}[t] + P_{DR}[t] - P_D[t] + R^m[t]$, $P_D[t] + R^M[t] - \sum_{n \in \mathcal{C}} P_n^M - P_{DG}[t] - P_{DR}[t]$, $P_n^m - P_n[t]$, $P_n[t] - P_n^M$, $-P_{DG}[t]$, $-P_{DR}[t]$, $P_{DG}[t] - P_{DG}^M[t]$, $P_{DR}[t] - P_{DR}^M[t]$, and $(1 - \eta_{\max})P_{DG}[t] - \eta_{\max} \sum_{n \in \mathcal{C}[t]} P_n[t]$ for $\forall n \in \mathcal{C}[t] := \{n | I_n[t] = 1 \text{ or } P_n[t] > 0\}$, $-R_n^D - P_n[t] + P_n[t-1]$ and $P_n[t] - P_n[t-1] - R_n^U$ for $\forall n \in \mathcal{C}[t] \cap \mathcal{C}[t-1]$. The KKT conditions are as follows

$$\nabla Q(P[t], I[t]) + \langle \lambda, \nabla h(P[t]) \rangle + \langle \mu, \nabla g(P[t]) \rangle = 0,$$
$$h(P[t]) = 0, \ \mu \geq 0, \ g(P[t]) \leq 0, \ \langle \mu, g(P[t]) \rangle = 0. \quad (53)$$

Define $(P_t^*, \lambda^*, \mu^*)$ the solution of (53) with $P_t^* := P^*[t]$, which is the optimal primal-dual solution of problem (32) as argued previously. Denote the solution of (53) as $(\bar{P}_t, \bar{\lambda}, \bar{\mu})$ when $g(P[t])$ is perturbed by some $\mathbf{c}$ such that $-R_n^D - P_n[t] + P_n[t-1] \leq c$ and $P_n[t] - P_n[t-1] - R_n^U \leq -c$ for $\forall n \in \mathcal{C}[t] \cap \mathcal{C}[t-1]$ while other inequalities remains unchanged. That is to say, $(\bar{P}_t, \bar{\lambda}, \bar{\mu})$ is the solution of perturbed KKT conditions

$$\nabla Q(P[t], I[t]) + \langle \lambda, \nabla h(P[t]) \rangle + \langle \mu, \nabla g(P[t]) \rangle = 0,$$
$$h(P[t]) = 0, \ \mu \geq 0, \ g(P[t]) \leq \mathbf{c}, \ \langle \mu, g(P[t] - \mathbf{c}) \rangle = 0. \quad (54)$$

The upper Lipschitz stability [28] gives that there is a neighborhood of $(P_t^*, \lambda^*, \mu^*)$ such that for any $\mathbf{c}$ close enough to $\mathbf{0}$, it holds

$$\|\bar{P}_t - P_t^*\| + \text{dist}((\lambda^*, \mu^*), (\bar{\lambda}, \bar{\mu})) \leq \|\mathbf{c}\|. \quad (55)$$

Therefore, for any $\epsilon > 0$, there exists $0 < \delta < \epsilon$, if $\|P_n[t-1] - (P_n[t-1] - c)\| = c < \delta$, $\forall n \in \mathcal{C}[t] \cap \mathcal{C}[t-1]$, then $\|\bar{P}_t - P_t^*\| \le \|\mathbf{c}\| < \delta < \epsilon$. Thus, one obtains the continuity of $f_I(P[t-1], P_{\mathrm{DG}}^M[t], P_{\mathrm{D}}[t])$ w.r.t. $P[t-1]$. ∎

Now we are ready to prove the Theorem 3.

Proof of Theorem 3: Firstly, one has the continuity of $J_T^*(P[T], I[T-1])$ w.r.t. $P[T]$ since $Q(P[T], I^*[T])$ is continuous and $\kappa(I[T-1], I^*[T])$ is constant w.r.t. $P[T]$. Secondly, it can be shown by mathematical induction that $J_t^*(P[t], I[t-1])$ is continuous w.r.t. $P[t]$ for any $I[t-1] \in \mathcal{I}_{t-1}$ and $t = 1, ..., T-1$.

Assume that $J_{t+1}^*(P[t+1], I[t])$ is continuous w.r.t. $P[t+1]$. According to Bellman's principle of optimality, one has that

$$J_t^*(P[t], I[t-1]) = \min_{\forall I[t]} \Big(Q(P[t], I[t]) + \kappa(I[t-1], I[t])$$
$$+ J_{t+1}^*(f_{I[t]}(P[t]), I[t]) \Big), \quad (56)$$

which is continuous since $F(P[t], I[t-1], I[t]) := Q(P[t], I[t]) + \kappa(I[t-1], I[t]) + J_{t+1}^*(f_{I[t]}(P[t]), I[t])$ is continuous for any $I[t-1] \in \mathcal{I}_{t-1}$ and $I[t] \in \mathcal{I}_t$ where $|\mathcal{I}_t|$ is finite. The reason is as follows. Assume that $|\mathcal{I}_t| = 2$, then it can be verified that $\min\{F_1, F_2\} = \frac{1}{2}(F_1 + F_2 - |F_1 - F_2|)$ is continuous where $F(P[t], I[t-1], I[t])$ for $I[t] \in \mathcal{I}_t$ are denoted by F_1 and F_2 respectively. By mathematical induction, one has that $\min_{i=1,...,|\mathcal{I}_t|}\{F_i\}$ is also continuous given the continuity of F_i with i in finite set $|\mathcal{I}_t|$. Therefore, $J_t^*(P[t], I[t-1])$ is continuous w.r.t. $P[t]$ for any $I[t-1] \in \mathcal{I}_{t-1}$ given the continuity of $J_{t+1}^*(P[t+1], I[t])$. This completes the proof. ∎

The above Theorem 3 shows that the optimal cost-to-go $J_t^*(P[t], I[t-1])$ is continuous w.r.t. $P[t]$. Hence, $\tilde{J}_t^*(P[t], I[t-1])$ in Algorithm 1 uniformly approximates $J_t^*(P[t], I[t-1])$ by using training methods such as least squares according to Weierstrass's theorem and [26]. That is to say, the CLHO algorithm provides a means of computing the optimal cost-to-go function $J_t^*(P[t], I[t-1])$. Note that the neural networks trained in Algorithm 1 produce the representation of *function* $\tilde{J}_t^*(P[t], I[t-1])$ (rather than *numbers*). The representation of function $\tilde{J}_t^*(P[t], I[t-1])$ implies that for any (disturbed) $P[t]$ and $I[t-1]$, the schedule produced by (50) without retraining is still optimal for the following $[t, T]$. Therefore, we say Algorithm 1 produces optimal solution in closed-loop style.

V. SIMULATION CASES

In this section, simulation cases are presented to study the behavior and effectiveness of the proposed CLHO algorithm.

A. Example 1

In this simulation example, the closed-loop behavior of the CLHO algorithm is studied. A relatively simple UCD problem is selected in this example where the optimal solution can be easily spotted or found with method of exhaustion for ease of studying. Specifically, this example contains only two thermal units while the ramp limit and spinning reserves are not considered. The parameters of the thermal units are shown

in Table I. The demand is shown in table II. The generation costs of the thermal units for serving the demand varying from 100 MW to 800 MW are plotted in Fig. 1.

TABLE I. Parameters of thermal units for example 1

Unit	a_n	b_n	c_n	P_n^m	P_n^M	$C_{\mathrm{b},n}$	$C_{\mathrm{f},n}$	$C_{D,n}$
1	0.00142	7.20	510	150	600	0	0	0
2	0.00194	7.85	310	100	400	0	0	0

TABLE II. Forecasted power demand for example 1

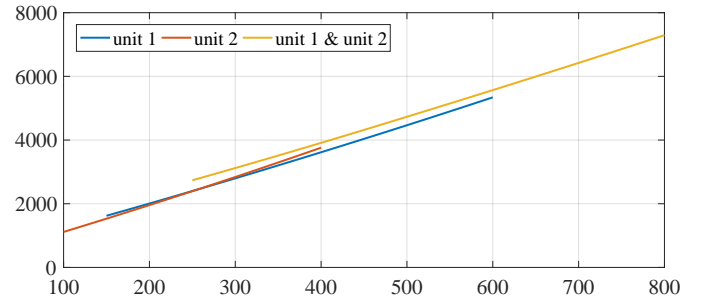| Time (h)	0	1	2	3	4	5	6
Demand (MW)	200	200	350	350	700	700	700

Fig. 1. Generation cost for serving demand 100 – 800 MW.

Training phase: The basis functions are selected as $\phi(P[t]) = [P_1^2[t], P_2^2[t], P_1[t], P_2[t], 1]^T$. At each time step t and schedule I_{t-1}, one hundred possible power output combinations are sampled randomly. The weights $W_t^{I_{t-1}}$ of the neural networks are decided by using least squares. The resulting weights of the neural networks are shown in Fig. 2. The training phase takes 42 s in MATLAB 2016a on a PC with Intel Core i3-4130 CPU and 8 GB of RAM.

Scheduling phase: Once trained, the neural networks are ready to be used to schedule the generation units. In the following, four simulation cases are presented to study the closed-loop behavior of the CLHO algorithm.

1) Case 1: In the first case, the initial conditions are selected as $P[0] = [0, 200]^T$ MW. The scheduling results are shown in Fig. 3. Since the start-up and shutdown costs are zero as shown in Table I, one has that the optimal scheduling at each time step is the optimal power output combinations for that time step. According to Fig. 1, the optimal schedule for demand of 200 MW, 350 MW, and 700 MW are $I = [0, 1]^T$, $I = [1, 0]^T$, and $I = [1, 1]^T$ respectively. It can also be verified that the corresponding optimal generation outputs are $P = [0, 200]^T$ MW, $P = [350, 0]^T$ MW, and $[500.9, 199.1]^T$ MW. Therefore, the scheduling results as shown in Fig. 3 is optimal with total generation cost \$23168.3.

2) Case 2: In the second case, the initial conditions remain the same as $P[0] = [0, 200]^T$ MW. However, the power outputs at time $t = 2$ are $P[2] = [200, 150]^T$ MW due to external disturbance. No re-train is needed in this case. The resulting history of generation outputs are shown in Fig.
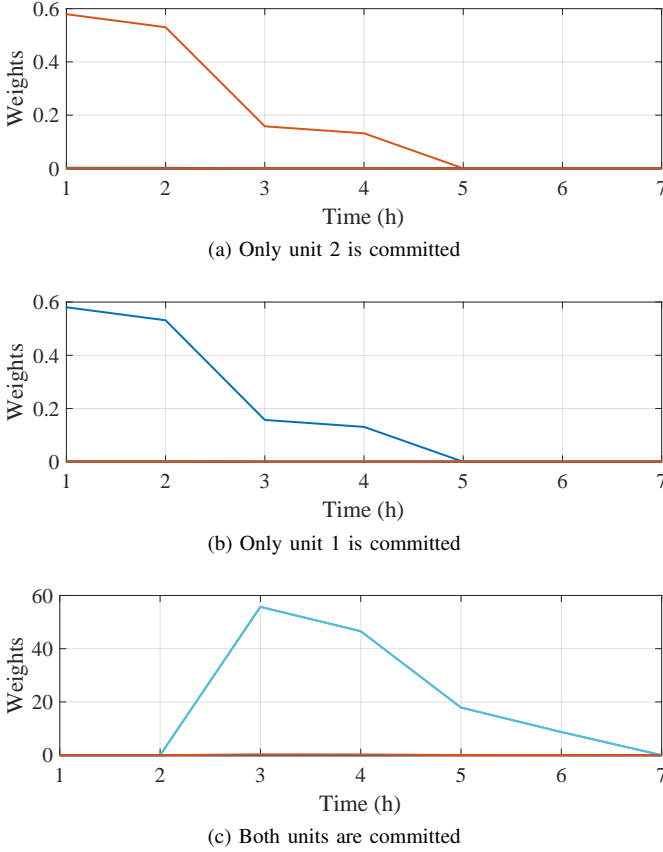
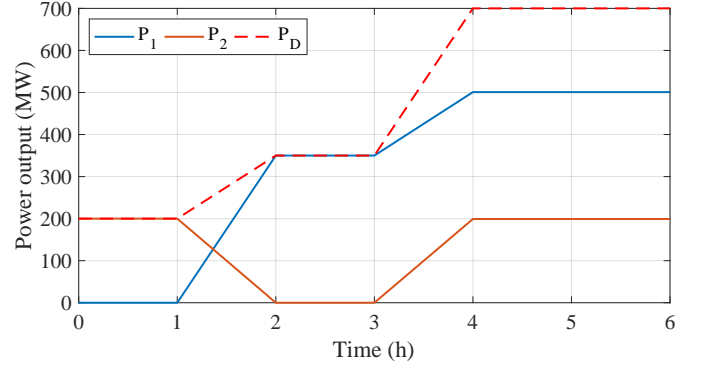Fig. 2. Weight histories of the trained neural networks.
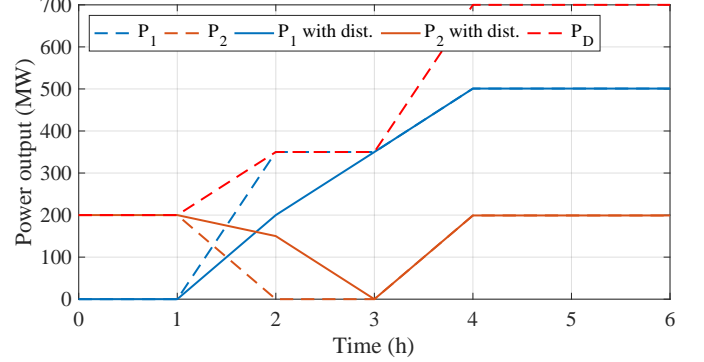

Fig. 3. Scheduling results with initial $P[0] = [0, 200]^T$ MW.

Fig. 4. Scheduling results under disturbance at time $t = 2$.

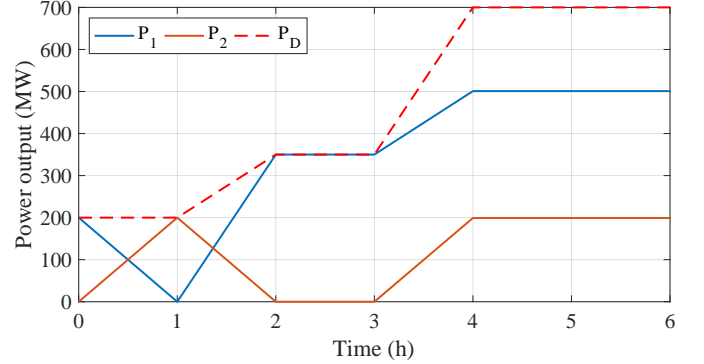Fig. 5. Scheduling results with initial $P[0] = [200, 0]^T$ MW.

4. The dashed lines represent the scheduling results without disturbance while the real lines show the actual power outputs under disturbance. The results show that though the power outputs are $P[2] = [200, 150]^T$ MW at time $t = 2$ due to disturbance, the following power outputs are still optimal comparing with Fig. 3. Thus, the scheduling phase of the CLHO algorithm is closed-loop to external disturbances.

3) Case 3: In the third case, the initial conditions are changed to $P[0] = [200, 0]^T$ MW. No re-train is needed in this case. The scheduling results are shown in Fig. 5. Since there is no start-up and shutdown costs, the generation outputs are scheduled to the optimal one at first time step $t = 1$. In what follows, the scheduling results remain optimal comparing with Fig. 3. Thus, the scheduling phase of the CLHO algorithm is also closed-loop to different initial conditions.

4) Case 4: In the fourth case, the start-up costs are set as $C_{U,1}(\tau_1) = 300 \cdot \tau_1$ and $C_{U,2}(\tau_2) = 200 \cdot \tau_2$. The shutdown costs are set as $C_{D,1} = 600$ and $C_{D,2} = 400$. Due to the changes of switching cost in (49) incurred by the start-up and shutdown costs, the neural networks are re-trained in this case. After training the neural networks, the scheduling phase is carried out. The scheduling results are shown in Fig. 6. Comparing the results with that shown in Fig. 3, one has that the unit 2 is not shutdown at $t = 2$ and start-up again at $t = 3$ due to the shutdown and start-up costs. The total cost from time $t = 0$ to time $t = 6$ incurred by this scheduling results is $\$ 24386.7$, which is less than the total cost $\$ 24568.3$ incurred by the scheduling results in Fig. 3 including non-zero start-

up and shutdown costs. In fact, the scheduling results given in Fig. 6 is optimal, which can be verified in the following with method of exhaustion. All the possible commitment and the corresponding minimal total generation costs for both the first case and this case are shown in Table III. In the table, number 1, 2, and 3 represent the commitment $[0, 1]^T$, $[1, 0]^T$, and $[1, 1]^T$ respectively. According to the results shown in Table III, the optimal commitment for $t = 1$ to $t = 6$ is 133333 and the minimal total generation cost is $\$ 24386.7$, which are the same as the scheduling results shown in Fig. 6 and the corresponding total cost.

Similarly, with the help of Fig. 1, it can be easily verified that all the scheduling results for the previous three simulation cases are also optimal under the corresponding set up. The neural networks are not retrained in the first three cases, revealing an important feature that the proposed algorithm produces scheduling results in closed-loop style. Since the

optimal schedule can be readily produced without retraining, the CLHO algorithm may contribute to the real-time operation of microgrids. As comparison, most existing methods such as [2]–[4] require reoptimization for different initial conditions and external disturbances.

Simulation comparison with traditional methods is not presented due to the following fundamental and clear distinctions. Firstly, the CLHO algorithm can be used differently from traditional methods. Namely, the time-consuming training phase can be conducted day-ahead while in the scheduling phase, no retraining is needed for different initial conditions and disturbances as shown above. Secondly, the CLHO algorithm finds the optimal solution as shown by the theoretical analysis and simulations, while most traditional methods usually produce suboptimal solutions.
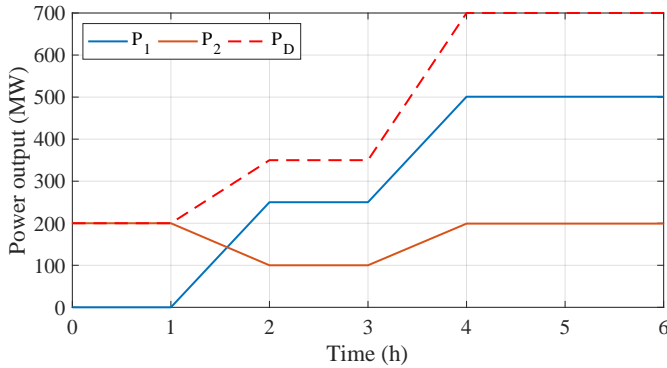

Fig. 6. Scheduling results considering start-up/shutdown costs.

TABLE III. Minimal generation costs for different schedule

$I[1:6]$	111333	112333	113333	121333	122333	123333
Case 1	23350.7	23259.5	23568.7	23259.5	23168.3	23477.5
Case 4	24550.7	24759.5	24468.7	25359.5	24568.3	24677.5
$I[1:6]$	131333	132333	133333	211333	212333	213333
Case 1	23568.7	23477.5	23786.7	23399.9	23308.7	23617.9
Case 4	25068.7	24677.5	24386.7	25499.9	25708.7	25417.9
$I[1:6]$	221333	222333	223333	231333	232333	233333
Case 1	23308.7	23217.5	23526.7	23617.9	23526.7	23835.9
Case 4	25308.7	24517.5	24626.7	25417.9	25026.7	24735.9

B. Example 2

In this simulation example, the IEEE 57-bus system is employed to study the effectiveness of the CLHO algorithm. The forecasted power demand and maximum power generation of DGs over 24 hours is shown in Table IV. Table V and VI show the fuel cost coefficients, generation capacities, start-up and shutdown cost, and emission coefficients of the thermal units. The parameters of the aggregator managing the DGs and virtual unit of DR are also shown in Table V. The penetration rate η_t of the DGs is bounded by $\eta_{\max} = 0.05$. Maximum demand curtailment of the DR is set as 40MW for peak hours ($t = 11\,\mathrm{h}$ and $t = 12\,\mathrm{h}$) and 10MW for non-peak hours. To cope with the renewable energy and demand forecast errors, spinning reserves are also considered. Specifically, the lower and upper spinning reserves are set as

5% of the forecasted demand $P_\mathrm{D}[t]$ for each hour, namely, $R^m[t] = R^M[t] = 0.05 \cdot P_\mathrm{D}[t]$. At the first hour $t = 1$, the power outputs of unit 1 and unit 2 are set as 500 MW and 200 MW respectively, while assuming zero power outputs for other units and the demand curtailment.

1) Case 1: In the first case, the quotas Q_n is set as 0 ton for each thermal unit, thus the operators have to buy all the emission allowance from the market at a price assumed as $p_e = 1$ \$/ton. The scheduling results are shown in Fig. 7. The total emission is 31086 ton.

2) Case 2: In the second case, the quotas Q_n is still set as 0 ton for each thermal unit while the price of emission quotas is increased to $p_e = 10$ \$/ton. The scheduling results are shown in Fig. 8. Numerical simulation shows that the total emission is 28775 ton, which is less than 31086 ton in the first case.

Since the higher price to purchase emission quota results in a higher emission cost, the thermal units are less preferable than in the first case. In consequence, the power output of DGs and curtailment of power demand are larger than in the first case, which reduces the total emission.

3) Case 3: In the third case, the quotas Q_n is set as 90% of the purchased quotas in the first case. The market price is set as $p_e = 1$ \$/ton. Fig. 9 shows the corresponding scheduling results. Simulation shows that the total emission is 31404 ton, which is more than the emission in the first case.

Note that the quotas in the first case are less than the quotas here. The less quotas result in higher emission costs due to the need of buying extra emission allowance. To reduce the total generation cost, the thermal units are less used in the first case which in turn causes a reduction on the carbon emissions.

In summary, the simulation results show that both the low emission quotas and the high emission trading price help reduce the total amount of carbon emissions.

TABLE IV. Forecasted demand and maximum DGs

Time (h)	1	2	3	4	5	6	7	8
Demand (MW)	700	750	850	950	1000	1100	1150	1200
Max. DGs (MW)	0	0	0	0	0	15	35	50
Time (h)	9	10	11	12	13	14	15	16
Demand (MW)	1300	1400	1450	1500	1400	1300	1200	1050
Max. DGs (MW)	72	84	80	88	85	76	36	36
Time (h)	17	18	19	20	21	22	23	24
Demand (MW)	1000	1100	1200	1400	1300	1100	900	800
Max. DGs (MW)	15	2	0	0	0	0	0	0

TABLE V. Parameters of (virtual) units for example 2

Unit	$a_{(.)}$	$b_{(.)}$	$c_{(.)}$	$P^m_{(.)}$	$P^M_{(.)}$	$C_{\mathrm{b},n}$	$C_{\mathrm{f},n}$	$C_{D,n}$
1	0.00048	16.19	990	150	600	550	100	90
2	0.00031	17.26	970	100	500	570	110	100
3	0.00200	16.60	700	20	130	110	85	75
4	0.00211	16.50	680	20	130	120	95	85
5	0.00398	19.70	450	25	162	150	100	90
DG	0.01000	2.60	10	0	–	–	–	–
DR	0.02000	2.20	4	0	–	–	–	–

VI. Conclusion

In this paper, the UCD problem is studied in microgrid environment while taking account of thermal units and DGs and considering DR and CET. Firstly, the UCD problem is reformulated as the optimal switching problem of a hybrid system. The reformulation enables employing recent advances in areas such as RL and DP and allows easy handling of time-dependent costs. Theoretical results show the equivalence between the original UCD problem the reformulated problem. Then, the CLHO algorithm is developed by using neural networks to approximate the optimal cost-to-go and employing convex optimization algorithm to solve the lower problem. The CLHO algorithm is capable of producing the optimal schedule in a closed-loop style, leading to the closed-loop solution for different initial conditions and external disturbances. Theoretical analysis and simulation results are provided to show the effectiveness of the proposed algorithm. It is also shown that the low emission quotas and high CET price help reduce the total carbon emissions.

TABLE VI. Emission coefficients of thermal units

Unit	α_n	β_n	γ_n
1	0.00312	-0.24444	10.33908
2	0.00312	-0.24444	10.33908
3	0.00509	-0.40695	30.03910
4	0.00509	-0.40695	30.03910
5	0.00344	-0.38132	32.00006
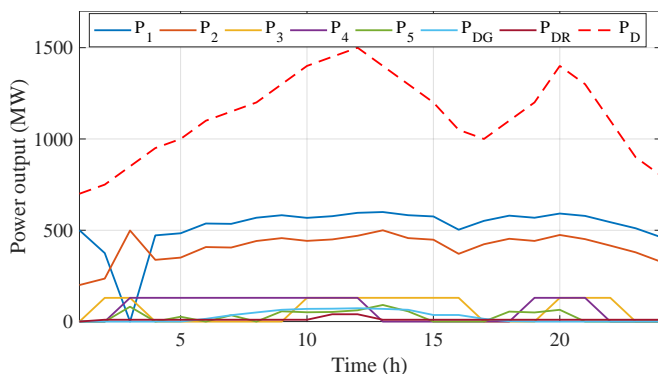
Fig. 7. Scheduling results for case 1 of example 2.
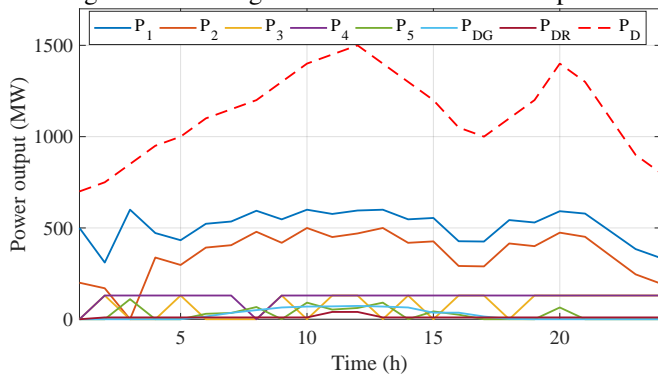
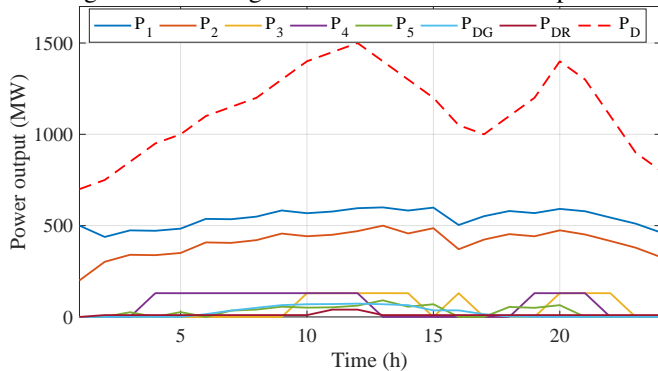Fig. 8. Scheduling results for case 2 of example 2.

Fig. 9. Scheduling results for case 3 of example 2.

References

[1] A. J. Wood, B. F. Wollenberg, and G. B. Sheble, *Power Generation, Operation and Control, 3rd Edition*. Hoboken, NJ: Wiley, Sep. 2013.

[2] T. Senjyu, K. Shimabukuro, K. Uezato, and T. Funabashi, "A fast technique for unit commitment problem by extended priority list," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 882–888, May 2003.

[3] W. Ongsakul and N. Petcharaks, "Unit commitment by enhanced adaptive lagrangian relaxation," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 620–628, Feb. 2004.

[4] D. F. Rahman, A. Viana, and J. P. Pedroso, "Metaheuristic search based methods for unit commitment," *Int. J. Electr. Power Energy Syst.*, vol. 59, pp. 14–22, Jul. 2014.

[5] N. P. Padhy, "Unit commitment—A bibliographical survey," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 1196–1205, May 2004.

[6] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, USA: MIT Press, 1998.

[7] M. L. Littman, "Reinforcement learning improves behaviour from evaluative feedback," *Nature*, vol. 521, no. 7553, pp. 445–451, May 2015.

[8] D. P. Bertsekas, *Abstract Dynamic Programming*. Belmont, USA: Athena Scientific, 2013.

[9] B. Kiumarsi, K. G. Vamvoudakis, H. Modares, and F. L. Lewis, "Optimal and autonomous control using reinforcement learning: A survey," *IEEE Trans. Neural Netw. Learning Syst.*, vol. 29, no. 6, pp. 2042–2062, May 2018.

[10] Q. Wei, D. Liu, F. L. Lewis, Y. Liu, and J. Zhang, "Mixed iterative adaptive dynamic programming for optimal battery energy control in smart residential microgrids," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4110–4120, Mar. 2017.

[11] X. Yang, H. He, and X. Zhong, "Adaptive dynamic programming for robust regulation and its application to power systems," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5722–5732, Feb. 2018.

[12] Y. Zhu, D. Zhao, X. Li, and D. Wang, "Control-limited adaptive dynamic programming for multi-battery energy storage systems," *IEEE Trans. Smart Grid*, vol. PP, no. 99, pp. 1–1, Jul. 2018.

[13] H. Shuai, J. Fang, X. Ai, Y. Tang, J. Wen, and H. He, "Stochastic Optimization of Economic Dispatch for Microgrid Based on Approximate Dynamic Programming," *IEEE Trans. Smart Grid*, pp. 1–1, Jan. 2018.

[14] L. Xie and M. D. Ilic, "Model predictive economic/environmental dispatch of power systems with intermittent resources," in *Energy Society General Meeting (PES)*. IEEE, Jul. 2009, pp. 1–6.

[15] W. Zhang and D. Nikovski, "State-space approximate dynamic programming for stochastic unit commitment," in *2011 North American Power Symposium (NAPS 2011)*. IEEE, Aug. 2011, pp. 1–7.

[16] F. Li, J. Qin, Y. Kang, and W. X. Zheng, "Consensus based distributed reinforcement learning for nonconvex economic power dispatch in microgrids," in *Neural Information Processing*, ser. Lecture Notes in Computer Science. Springer, Cham, Nov. 2017, pp. 831–839.

[17] A. Rong and P. B. Luh, "A Dynamic Regrouping Based Dynamic Programming Approach for Unit Commitment of the Transmission-Constrained Multi-Site Combined Heat and Power System," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 714–722, Dec. 2017.

[18] N. Zhang, Z. Hu, D. Dai, S. Dang, M. Yao, and Y. Zhou, "Unit Commitment Model in Smart Grid Environment Considering Carbon Emissions Trading," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 420–427, Jan. 2016.

[19] J. Qin, Q. Ma, Y. Shi, and L. Wang, "Recent advances in consensus of multi-agent systems: A brief survey," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 6, pp. 4972–4983, Jun. 2017.

[20] J. Aghaei and M.-I. Alizadeh, "Multi-objective self-scheduling of CHP (combined heat and power)-based microgrids considering demand response programs and ESSs (energy storage systems)," *Energy*, vol. 55, pp. 1044–1054, Jun. 2013.

[21] J. Qin, Y. Wan, X. Yu, F. Li, and C. Li, "Consensus-based distributed coordination between economic dispatch and demand response," *IEEE Trans. Smart Grid*, to be published, doi: 10.1109/TSG.2018.2834368.

[22] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, and J. Lavaei, "A Survey of Distributed Optimization and Control Algorithms for Electric Power Systems," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2941–2962, Oct. 2017.

[23] M. Espinoza, J. Suykens, R. Belmans, and B. De Moor, "Electric load forecasting," *IEEE Control Syst. Mag.*, vol. 27, no. 5, pp. 43–57, Sep. 2007.

[24] C. J. Tomlin, J. Lygeros, and S. S. Sastry, "A Game Theoretic Approach to Controller Design for Hybrid Systems," *Proc. IEEE*, vol. 88, no. 7, pp. 949–970, Jul. 2000.

[25] R. Goebel, "Hybrid Dynamical Systems: Modeling, Stability, And Robustness," pp. 1–227, Dec. 2011.

[26] K. Hornik, S. Maxwell, and W. Halbert, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, pp. 359–366, 1989.

[27] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Jan. 2014.

[28] A. F. Izmailov, A. S. Kurennoy, and M. V. Solodov, "A note on upper Lipschitz stability, error bounds, and critical multipliers for Lipschitz-continuous KKT systems," *Mathematical Programming*, vol. 142, no. 1-2, pp. 591–604, Sep. 2012.