An iterative algorithm for regret minimization in flexible demand scheduling problems

Zihang Dong1 | David Angeli1,2 | Antonio De Paola1 | Goran Strbac1

1Department of Electrical and Electronic Engineering, Imperial College London, London, UK
2Dipartimento di Ingegneria dell’Informazione, Università di Firenze, Firenze, Italy

Correspondence
Zihang Dong, Control and Power Group, Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK.
Email: zihang.dong14@imperial.ac.uk

Funding information
Integrated Development of Low-Carbon Energy Systems, Grant/Award Number: EP/R045518/1; Active Building Centre, Grant/Award Number: EP/V012053/1

Abstract
A major challenge to develop optimal strategies for allocation of flexible demand toward the smart grid paradigm is the uncertainty associated with the real-time price and electricity demand. This article presents a regret-based model and a novel iterative algorithm which solves the minimax regret optimization problem. This algorithm has low computational burden compared with traditional linear programming methods and affords iterative convergence through updates of feasible power schedules, thus enabling a scalable parallel implementation for large device populations. Specifically, our approach seeks to minimize the induced worst-case regret over all price scenarios and solves the optimal charging strategy for the electrical devices. The convergence of the method and optimality of the computed solution is justified and some numerical simulations are discussed for the case of a single device operating under different types of price realizations and uncertainty bounds.

KEYWORDS
electric vehicles, flexible demand, minimax regret, robust optimization, uncertainty

1 INTRODUCTION

The electrification of the transportation sector is envisaged as a key step toward a decarbonized and sustainable energy system.1 The increasing penetration of electric vehicles (EVs) will reduce carbon emissions and enable a more efficient integration of renewable generation. The intrinsic flexibility of the EVs’ charging process and the implementation of vehicle-to-grid technologies are expected to support the reliability and the efficiency of the power system by providing a wide range of ancillary services,2-4 such as balancing, voltage support or frequency regulation.

In order to fully achieve these potential benefits while limiting undesired consequences (e.g., network congestion, price volatility), it is crucial to develop efficient and robust control strategies for the charging of the EVs. A significant amount of research has investigated this topic after the seminal papers,5,6 proposing decentralized and game-theory approaches,7-10 adaptive strategies11 or Lagrange relaxation methods12 to achieve an efficient large-scale integration of the EVs in the power system. These analyses have been expanded to consider additional elements such as coupling constraints between the devices,13 network topology14 or the degradation costs of the EVs’ batteries.15

Most of the cited works adopt a deterministic framework that assumes perfect knowledge of all the relevant parameters of the problem, such as electricity prices, network conditions or EVs’ availability. However, these elements are subject to relevant uncertainties, which might have a substantial impact on the efficiency and benefits of the proposed control...
schemes. One possible approach to tackle this issue is to adopt receding-horizon schemes\textsuperscript{16,17} or real-time algorithms\textsuperscript{7} that modify the charging profile of the EVs on the basis of updated system conditions, implicitly accounting for imperfect predictions. Other works have considered the application of competition online algorithms\textsuperscript{18}, stochastic dynamic programming\textsuperscript{19} or robust min–max optimization.\textsuperscript{20}

The analysis presented in this article also adopts a robust optimization framework, considering bounded uncertainties on the electricity prices and determining the cost-minimizing charging profile of a single EV as the solution of a regret minimization problem. In the proposed approach, the EV charging is determined in order to minimize the worst-case regret, that is, the largest additional cost (with respect to the deterministic perfect-information case) due to price uncertainties. The optimal solution is calculated in a constructive manner: a tentative feasible charging profile is iteratively updated through some elementary operations consisting of power swaps between two or multiple time instants, reducing the maximum regret at each iteration and preserving feasibility. Differently from the numerical approach presented in Reference 20, it is analytically demonstrated that the proposed method always converges to the minimized regret solution using Lyapunov techniques, under all the considered price uncertainty hypotheses and for any initial charging profile. Thanks to these features the method is well suited for a large-scale parallel implementation involving asynchronous schedule updates and energy price recalculations. In addition, numerical case studies are presented to (i) demonstrate the increased computational efficiency of the proposed method with respect to standard linear programming approaches while achieving the same optimal control solutions and (ii) qualitatively characterize the optimal power schedules with respect to different uncertainty hypotheses and flexibility parameters of the EV.

The remainder of this article is organized as follows: Section 2 presents the main regret model of a single EV and discusses the impact of correlated price uncertainties on the regret. The description of the iterative optimization algorithm and the proof of its optimality are included in Sections 3.1 and 3.2, respectively. Simulations of the proposed control strategy are presented in Sections 4, and Section 5 concludes this article.

\section{PRELIMINARIES AND SETUP}

Next we introduce our modeling framework and formulate a regret criterion to characterize optimal power allocation strategies in the presence of price uncertainty. In essence, the minimax regret criterion captures the idea that devices are concerned with foregone opportunities. It departs significantly from a worst-case approach where devices optimize the cost incurred for the highest possible electricity price occurrence. The latter is a risk-averse strategy that may result in conservative decisions and is therefore more suited for malicious uncertainties.

This section introduces the minimax regret formulation for the flexible devices under the uncertain electricity market. The impact of correlated uncertainty model on the regret analysis is also discussed.

\subsection{Regret modeling}

Let us denote the discrete time interval by $\mathcal{T} = \{1, \ldots, T\}$, the required total energy and rated power by $E^r$ and $P^r$, respectively. The power consumption schedule is subjected to the availability window $\mathcal{A} \subseteq \mathcal{T}$. If the power consumption profile of the appliance is denoted by $u: \mathcal{T} \rightarrow \mathbb{R}_+$, then a collection of feasible power schedules can be formulated as the following set

$$\mathcal{U} := \{u(\cdot) : \sum_{t=1}^{T} u(t) \Delta t = E^r, \ 0 \leq u(t) \leq P^r I_{\mathcal{A}}(t) \ \forall t \in \mathcal{T} \}.$$  \hfill (1)

where $I_{\mathcal{A}}(t)$ is an indicator function

$$I_{\mathcal{A}}(t) := \begin{cases} 
1 & \text{if } t \in \mathcal{A} \\
0 & \text{otherwise.}
\end{cases}$$  \hfill (2)

While for the case of a single device there is no loss of generality in assuming $\mathcal{A} = \mathcal{T}$, it is meaningful, for multiple devices, to allow possibly distinct availability windows. Therefore, we adopt this more general view point in the description of our modeling framework and algorithms.
**Assumption 1.** The feasible sets of power consumption profile $\mathcal{U}$ characterized by parameters $(E^r, P^r, \mathcal{A})$ are not empty, that is,

$$\sum_{t \in \mathcal{A}} P^r \cdot \Delta t \geq E^r. \quad (3)$$

**Assumption 2.** The required energy is a multiple of its rated power, that is,

$$\exists K \in \mathbb{Z}_+ \text{ such that } E^r = K P^r \Delta t. \quad (4)$$

If Assumption 2 holds and the price function $\Pi : \mathcal{T} \to \mathbb{R}_+$ is convex, at least one optimal power consumption profile will be an element of the following set

$$\mathcal{U}^K := \{ u(\cdot) : \sum_{t=1}^T u(t) \Delta t = E^r, u(t) \in \{0, P^r \mathbb{I}_{\mathcal{A}(t)}\} \; \forall t \in \mathcal{T} \}. \quad (5)$$

In the following, we normalize the discrete time sampling period $\Delta t = 1$ to ease the notation. We assume that energy prices at time $t$, denoted by $\Pi(t)$, are uncertain but known to belong to an a priori assigned interval:

$$\Pi(t) \in [\Pi(t), \overline{\Pi}(t)], \quad \forall t \in \mathcal{T}. \quad (6)$$

**Remark 1.** While in this article we do not model the impact of power demand on energy costs, the inclusion in (6) can be justified by assuming a simple model of power market where prices at any time $t$ are represented an increasing function of total demand (both flexible and inflexible).\(^5\),\(^6\),\(^21\) Although this assumption is only justifiable with 100% traditional generation, it can be easily generalized to more realistic scenarios with high penetration of renewable energy sources by using net demand instead of absolute demand to calculate the electricity price. Since the renewable contributions could vary according to the weather conditions, the forecast electricity price is uncertain within the considered time-window and assumed to belong to the interval in (6). Moreover, the impact of network congestion on the price is also neglected, as commonly assumed in previous research.\(^22\)-\(^24\)

If the signal $\Pi(t)$ was known in advance, one could accurately forecast the cost of energy for a power allocation $u$ as:

$$J(u) := \sum_{t \in \mathcal{T}} \Pi(t) u(t), \quad (7)$$

and an optimal power schedule, could be found by solving the following linear program:

$$\min_{u \in \mathcal{U}} J(u). \quad (8)$$

In particular, one might without loss of generality assume an optimal power schedule in $\mathcal{U}^K$. Thus, the optimal energy cost can be computed identifying the $K$ time instants with the lowest prices which could simply be evaluated according to:

$$U^*(\Pi) := \min_{S \subseteq \mathcal{A} : |S| = K} \sum_{t \in S} \Pi(t) P^r,$$

where $|S|$ denotes the cardinality of set $S$. We call this cost the *utopia* cost, and, for each price profile, it can only be achieved in the absence of uncertainty. In general, each schedule $u(\cdot)$ and each price profile will induce a regret, defined as the difference between the cost incurred by $u(\cdot)$ minus its corresponding utopia cost. Our objective is to schedule $u(\cdot)$ so as to minimize the worst case regret, defined as:

$$R(u) := \max_{\Pi : \Pi(t) \in [\Pi(t), \overline{\Pi}(t)]} \sum_{t \in \mathcal{T}} \Pi(t) u(t) - U^*(\Pi). \quad (9)$$

The objective is to find the optimal power schedule with the minimal worst regret

$$u^*(\cdot) \in \arg\min_{u \in \mathcal{U}} R(u). \quad (10)$$
It can be shown that the considered minimax regret problem in (10) can be reformulated as the linear program detailed below:

\[
\begin{align*}
\min_{u(\cdot), r} & \quad r \\
\text{s.t.} & \quad 0 \leq u(t) \leq P^r, \quad \forall t \in \mathcal{A} \\
& \quad \sum_{t \in \mathcal{A}} u(t) = E^r \\
& \quad \sum_{t \in S} \Pi(t)[u(t) - P^r] + \sum_{t \in S} \Pi(t)u(t) \leq r, \quad \forall S : |S| = K.
\end{align*}
\]

(11)

(11a)

(11b)

(11c)

Note that the number of constraints used to formulate the regret lower bounds (11c) is determined as \( \frac{|\mathcal{A}|}{K} \) a combinatorial function of the size of availability window \( \mathcal{A} \), and of the minimum number of time slots needed for task completion. As a result standard linear optimization algorithms might become computationally inefficient when the cardinality of \( \mathcal{A} \) is large.

To solve the minimax regret optimization problem more efficiently, we propose an iterative algorithm in which the power schedule \( u(\cdot) \) converges to the optimal solution \( u^*(\cdot) \) in (10), and accordingly the regret \( R(u) \) converges to \( R(u^*) \). In order to do so, it is useful to derive a more explicit expression of the regret, as detailed below:

\[
R(u) = \max_{\Pi \in \Pi(t)} \sum_{t \in \mathcal{A}} \Pi(t)[u(t) - P^r(t)] + \sum_{t \in \mathcal{A}} \Pi(t)u(t).
\]

(12)

Now, let us next define a function \( Q(t, u(t)) \) as follows:

\[
Q(t, u(t)) = [\Pi(t) - \Pi(t)]u(t) + \Pi(t)P^r.
\]

(13)

The quantity \( Q \) allows to express the regret in a form which lends itself to explicit formulas for evaluation, without the need of a combinatorial test of all possible subsets of cardinality \( K \) as from the previous expression. In fact:

\[
R(u) = \frac{1}{2} \max_{\Pi \in \Pi(t)} \left[ \sum_{t \in \mathcal{A}} [\Pi(t) + \Pi(t)]u(t) - P^r(t) \right] + \sum_{t \in \mathcal{A}} Q(t, u(t)) - \sum_{t \in \mathcal{A}} Q(t, u(t))
\]

\[
= \frac{1}{2} \left\{ \sum_{t \in \mathcal{A}} [\Pi(t) + \Pi(t)]u(t) - P^r(t) \right\} + \frac{1}{2} \max_{\Pi \in \Pi(t)} \left[ \sum_{t \in \mathcal{A}} Q(t, u(t)) - \sum_{t \in \mathcal{A}} Q(t, u(t)) \right].
\]

(14)

The latter maximization problem can be easily solved after sorting in ascending order the values of \( Q \) and assigning to \( S \) the time instants in \( \mathcal{A} \) with the \( K \) smallest values of \( Q \). More precisely, the \( K \)th smallest value of \( Q \) is defined as:

\[
\overline{Q} = \min_{\Pi \in \Pi(t)} \max_{t \in \mathcal{A}} Q(t, u(t)).
\]

(15a)

Accordingly, the discrete time availability interval \( \mathcal{A} \) is divided into three sets

\[
L_{th} = \{ t \in \mathcal{A} \mid Q(t, u(t)) < \overline{Q} \},
\]

(15b)
\[ \partial L = \{ t \in \mathcal{S} \mid Q(t, u(t)) = \bar{L} \}, \quad (15b) \]
\[ L_{\text{out}} = \{ t \in \mathcal{S} \mid Q(t, u(t)) > \bar{L} \}. \quad (15c) \]

**Remark 2.** Notice that the \( Q \) function is the basis to define the three sets in (15), rather than a evaluation of the energy cost. According to their definition, the sets \( L_{\text{in}} \) and \( L_{\text{out}} \) can be empty, whereas the set \( \partial L \) is always nonempty. Concerning cardinality, it follows that \( |L_{\text{in}}| + |\partial L| \geq K \). Since the support of any power consumption profile in \( \mathcal{Z}^K \) has \( K \) elements, two cases should be taken into account in order to calculate the regret. The simplest case is \( |L_{\text{in}}| + |\partial L| = K \), when every element of \( L_{\text{in}} \) and \( \partial L \) corresponds to a time instant where regret is evaluated costing energy according to the lower bound of the price. For the second case \( |L_{\text{in}}| + |\partial L| > K \), one must choose a subset \( \partial L_{\text{in}} \subseteq \partial L \) with cardinality \( K - |L_{\text{in}}| > 0 \), whose elements corresponds to time instants where energy is costed using its lower bound of the price, while, for the remaining elements in \( \partial L_{\text{out}} = \partial L \setminus \partial L_{\text{in}} \) (of cardinality \( |\partial L| - (K - |L_{\text{in}}|) > 0 \) energy is costed using the price upper bound. The mathematical equation of the regret for both cases is expressed as

\[
R(u) = \begin{cases} 
\sum_{t \in \partial L_{\text{in}}} \Pi(t)[u(t) - P'] + \sum_{t \in \partial L_{\text{out}}} \bar{\Pi}(t)u(t) & \text{if } |L_{\text{in}}| + |\partial L| = K, \\
\sum_{t \in \partial L_{\text{in}}} \Pi(t)[u(t) - P'] + \sum_{t \in \partial L_{\text{out}}} \bar{\Pi}(t)u(t) & \text{if } |L_{\text{in}}| + |\partial L| > K.
\end{cases}
\]

(16)

Note that the set \( \partial L \) is partitioned into two nonempty subsets of suitable cardinality in (16) when \( |L_{\text{in}}| + |\partial L| > K \). While this partition might not be unique, the expression of regret computed in (16) is independent of the choice made. This fact is illustrated by the following lemma:

**Lemma 1.** If \( |L_{\text{in}}| + |\partial L| > K \), the elements in \( \partial L_{\text{in}} \) and \( \partial L_{\text{out}} \) are exchangeable without varying the value of regret computed according to (16), that is,

\[
\Pi(t_1)[u(t_1) - P'] + \bar{\Pi}(t_2)u(t_2) = \Pi(t_2)[u(t_2) - P'] + \bar{\Pi}(t_1)u(t_1), \quad \forall t_1 \in \partial L_{\text{in}}, \forall t_2 \in \partial L_{\text{out}}.
\]

(17)

**Proof.** According to the definition of the set \( \partial L \), for all \( t_1 \in \partial L_{\text{in}} \) and \( t_2 \in \partial L_{\text{out}} \), it holds \( Q(t_1, u(t_1)) = Q(t_2, u(t_2)) \), or explicitly,

\[
[\bar{\Pi}(t_1) - \Pi(t_1)]u(t_1) + \Pi(t_1)P' = [\bar{\Pi}(t_2) - \Pi(t_2)]u(t_2) + \bar{\Pi}(t_2)P',
\]

which implies

\[
\Pi(t_1)[u(t_1) - P'] + \bar{\Pi}(t_2)u(t_2) = \Pi(t_2)[u(t_2) - P'] + \bar{\Pi}(t_1)u(t_1).
\]

Therefore, when \( |L_{\text{in}}| + |\partial L| > K \), we can select arbitrary \( K - |L_{\text{in}}| \) elements in \( \partial L \) as \( \partial L_{\text{in}} \) without changing the regret. \( \square \)

### 2.2 Regrets with correlated uncertainty

The price uncertainty model considered so far can be equivalently formulated by introducing the vector \( \Pi \) of prices,

\[
\Pi = [\Pi(t)]_{t \in \mathcal{F}}
\]

and specifying that \( \Pi \) is known to belong to the set:

\[
D := \prod_{t \in \mathcal{F}} [\Pi(t), \bar{\Pi}(t)].
\]

Since \( D \) is a Cartesian product, this can be interpreted as the deterministic counterpart of independent stochastic prices at each time instant. One might argue that, in practice, this is not often the case. For instance, on a sunny or windy day availability of renewable cheap energy could be uniformly high throughout the day and assuming a strong fluctuation of energy price on a hour to hour basis (or less) might seem not very plausible. We show next how to extend our approach...
to include a more general type of uncertain set \( D \), one which accounts for some form of correlated price uncertainty. In particular, we define \( D \) as follows:

\[
D = \{ \Pi : \exists \Pi_{av} \in [\Pi_{av}, \Pi_{av}] : \Pi(t) - \Pi_{av} \in [\Pi(t), \Pi(t)], \quad \forall t \in T \}.
\]

Notice that \( D \) is no longer a Cartesian product as an uncertain constant price shift is added to each price interval. Alternatively \( D \) can be written as:

\[
D = \bigcup \Pi_{av} \in [\Pi_{av}, \Pi_{av}] \prod_{t \in T} [\Pi(t) + \Pi_{av}, \Pi(t) + \Pi_{av}].
\]

Notice that the following inclusion holds:

\[
D \subseteq \prod_{t \in T} [\Pi(t) + \Pi_{av}, \Pi(t) + \Pi_{av}],
\]

however, this approach tends to overestimate the price fluctuations possible on an hour to hour basis so that using these new bounds on minimal and maximum price at each time \( t \) might result in suboptimal decisions. In the following we show that one may still use the regret expression computed with uncertainty bounds \( \Pi(t) \in [\Pi(t), \Pi(t)] \) in order to devise optimal decisions also in the case of uncertainty set \( D \) as in (18). While the expression for the utopia price is clearly unchanged, a different formula for regret evaluation is needed in this case:

\[
R(u) := \max_{\Pi, \Pi_{av} : \Pi(t) - \Pi_{av} \in [\Pi(t), \Pi(t)]} \sum_{t \in T} \Pi(t)u(t) - U'(\Pi).
\]

More explicitly, by substituting the formula for the utopia price, we obtain:

\[
R(u) = \max_{\Pi_{av} \in [\Pi_{av}, \Pi_{av}], \Pi: \Pi(t) - \Pi_{av} \in [\Pi(t), \Pi(t)], S \subseteq \mathcal{D} : |S| = K} \sum_{t \in S} \Pi(t)[u(t) - P'] - \sum_{t \notin S} \Pi(t)u(t),
\]

\[
= \max_{\Pi_{av} \in [\Pi_{av}, \Pi_{av}], \Pi: \Pi(t) - \Pi_{av} \in [\Pi(t), \Pi(t)], S \subseteq \mathcal{D} : |S| = K} \sum_{t \in S} \Pi(t)[u(t) - P'] + \sum_{t \notin S} \Pi(t)u(t),
\]

\[
= \max_{\Pi_{av} \in [\Pi_{av}, \Pi_{av}], \Pi: \Pi(t) - \Pi_{av} \in [\Pi(t), \Pi(t)], S \subseteq \mathcal{D} : |S| = K} \sum_{t \in S} [\Pi(t) + \Pi_{av}][u(t) - P'] + \sum_{t \notin S} [\Pi(t) + \Pi_{av}]u(t),
\]

\[
= \max_{S \subseteq \mathcal{D} : |S| = K} \sum_{t \in S} [\Pi(t) + \Pi_{av}][u(t) - P'] + \sum_{t \notin S} \Pi(t)u(t),
\]

where the last equality follows since:

\[
\sum_{t \in T} u(t) = E' = KP'.
\]

Notice that \( \Pi_{av} \) and its bounds have no impact at all on the regret and therefore optimization of the regret can proceed along the same lines as with a Cartesian product uncertainty set.

We show next another form of correlated uncertainty which has no impact on the modeling of the problem. In particular we consider an uncertainty set of the following type:

\[
D = \bigcup_{\alpha \in [0, \beta]} \alpha \prod_{t \in T} [\Pi(t), \Pi(t)].
\]
From the physical point of view this corresponds to a correlated deterministic uncertainty that results in a
time-independent rescaling of upper and lower prices. It could be seen as a (time-invariant) uncertainty in the exchange
rate of the currency used to cost energy and the one owned by the agent minimizing its cost. We proceed to the evaluation
of regret for this type of uncertainty set.

\[
R(u) := \max_{\Pi: \Pi(t) \in [\Pi(t), \Pi(t)]} \sum_{t \in T} \Pi(t)u(t) - \mathcal{U}(\Pi).
\]  

(21)

By substituting the formula for the utopia price, we obtain:

\[
R(u) = \max_{\Pi: \Pi(t) \in [\Pi(t), \Pi(t)]} \sum_{t \in S} \Pi(t)u(t) - \sum_{t \in S} \Pi(t)P^*,
\]

\[
= \max_{\Pi: \Pi(t) \in [\Pi(t), \Pi(t)], S \subseteq \mathcal{A}: |S|=K} \alpha \left[ \sum_{t \in S} \Pi(t)[u(t) - P^*] + \sum_{t \in S} \Pi(t)u(t) \right],
\]

where the second equality follows by redefining \(\bar{\Pi} = \Pi/\alpha\) and rewriting the corresponding maximization problem.

In particular, the regret \(R(u)\) equals \(\alpha\) times the regret corresponding to the situation \([\bar{\alpha}, \bar{\alpha}] = \{1\}\). Because of this simple calculation, we conclude that multiplicative constant uncertainty does not affect the choice of optimal regret allocation.

3 | ITERATIVE OPTIMIZATION STRATEGY

An iterative regret minimization algorithm is introduced in this section and its convergence to the optimal power allocation
of minimal worst regret is justified. Before studying the technical details of the algorithm, we first illustrate its overall
structure.

The proposed iterative algorithm is described in Algorithm 1. This updates the power schedule sequentially to reach
the optimal equilibrium. The algorithm is initialized with a feasible power schedule \(u_0 \in \mathcal{U}\) and two variables: \(k\) is the
iteration index whereas \(\text{conv}\) indicates if the optimal equilibrium has been reached. Then, in the main part, a \(\text{while}\) cycle
is performed for the agent to justify if there is any nontrivial power swap \(\delta > 0\) which can potentially reduce it regret.
At each execution of the cycle, one of the three elementary power swap policies, which will be defined in Section 3.1,
is implemented to update the previous power scheduling \(u_{k-1}\). Finally, the indicator \(\text{conv}\) is set to 0 and proceeds to the
next iteration. When there are no power swap \(\delta > 0\) that decrease the regret, the \(\text{while}\) loop terminates and the power
schedule \(u_k\) has reached an optimal solution.

3.1 | Elementary power swap

Given an arbitrary initial power allocation, this is sequentially updated in response to uncertain price signals, with the
objective of reducing the regret function. In particular, this sequential process consists of two types of swap operations:
(i) single time instant to single time instant power swap, that is, \(1 \to 1\); (ii) single time instant (\(n\) multiple time instants)
to \(n\) multiple time instants (single time instant) power swap, that is, \(1 \to n (n \to 1\), respectively).
Algorithm 1. Iterative Scheme - Worst regret optimization

1 Initialization:
\[ k = 0, \quad u_0 \in \mathcal{U}, \quad \text{conv} = 0 \]

2 Iteration procedure:
3 while \( \text{conv} = 0 \) do
4 \[ \text{conv} = 1, \quad k = k + 1 \]
5 if \( \delta > 0 \) exists such that: \( R(u_k + \delta) - R(u_k) < 0 \) then
6 At least one of the following swap policy holds:
   • \( 1 \rightarrow 1 \) power swap : \( \exists \delta > 0 \) and \( t, \tilde{t} \in \mathcal{A} \)
     \[ u_k = u_{k-1} + \delta(t) \vec{e}_t - \delta(t) \vec{e}_{\tilde{t}} \]
   • \( m \rightarrow 1 \) power swap : \( \exists \delta > 0 \) and \( S \subseteq \mathcal{A}, \tilde{t} \in \mathcal{A} \)
     \[ u_k = u_{k-1} + \left( \sum_{t \in S} \delta(t) \right) \vec{e}_\tilde{t} - \sum_{t \in S} \delta(t) \vec{e}_t \]
   • \( 1 \rightarrow n \) power swap : \( \exists \delta > 0 \) and \( t \in \mathcal{A}, S \subseteq \mathcal{A} \)
     \[ u_k = u_{k-1} - \left( \sum_{t \in S} \delta(t) \right) \vec{e}_t + \sum_{t \in S} \delta(t) \vec{e}_t \]
9 \[ \text{conv} = 0 \]

7 Final results: The power schedule is the optimal solution with the minimum worst regret: \( u_k \rightarrow u^* \).

3.1.1 1 \rightarrow 1 power swap

The first type of elementary power swap transfers \( \delta \) power units from time \( t \) to \( \tilde{t} \). Starting from an initial power schedule \( u \in \mathcal{U} \), the updated power profile \( \tilde{u} \) after the swap will have the following expression:

\[
\tilde{u}(t) = \begin{cases} 
  u(t) + \delta & \text{if } t = \tilde{t} \\
  u(t) - \delta & \text{if } t = t' \\
  u(t) & \text{otherwise}
\end{cases}
\]  

(22)

An equivalent compact definition reads as:

\[
\tilde{u} = u + \delta \left( \vec{e}_t - \vec{e}_{\tilde{t}} \right).
\]  

(23)

where \( \vec{e}_s \) is the vector of the canonical basis associated with the \( s \)th element.

For a given power schedule \( u \in \mathcal{U} \), the amount of power \( \delta \) that can be swapped by the single device is limited by the quantity \( \delta \), defined as follows:

\[
\delta(u, t, \tilde{t}) = \min\{a(u, \tilde{t}), b(u, t), c(u, t, \tilde{t})\}.
\]  

(24)

Each of the four terms of the minimum function in (24) is now described in detail:

• Maximum feasible power increase at \( \tilde{t} \):

\[
a(u, \tilde{t}) = P^r - u(\tilde{t}).
\]  

(25)

• Maximum feasible power decrease at \( t' \):

\[
b(u, t') = u(t').
\]  

(26)

The bounds \( a \) and \( b \) ensure the new power schedule \( \tilde{u} \) after the swap fulfills the constraints in (1) for \( \delta \leq a \) and \( \delta \leq b \).
Maximum regret-reducing power swap within region of constant gradient: The last term \(c\) in (24) ensures that \(\delta > 0\) only if the associated power swap reduces the regret and, for the sake of its explicit evaluation, it is computed within a region of constant gradient, which is always guaranteed to exist for a piecewise linear continuous function:

\[
c(u, \vec{\tau}, t) := \max_{\delta \geq 0} \delta
\]

subject to

\[
R(u + \delta (\vec{e}_\tau - \vec{e}_t)) - R(u) = \delta \lim_{\delta \to 0} \frac{R(u + \delta (\vec{e}_\tau - \vec{e}_t)) - R(u)}{\delta} \leq 0.
\]  

(27)

Notice that the feasibility conditions (25) and (26) are easily determined according to the given power allocation, whereas the determination of \(c(u, \vec{\tau}, t)\) in (27) is not straightforward. Since the worst regret \(R(u)\) can be evaluated using the \(Q\) values as expressed in (13), it is worth to analyse the relation between the power swap and the worst regret variation via \(Q\) values. The detailed examination for regret variation as well as the determination of \(c(u, \vec{\tau}, t)\) are elaborated numerically in Appendix A.1.

### 3.1.2 \(1 \to m\) and \(n \to 1\) power swap

The second type of elementary power swap is between the multiple time instants and a single time instant. Let us denote the power units from a single time \(t\) to the set \(S\) of \(m\) time instants by \(\delta_{1 \to m}\), and the proportion of power units into \(t \in S\) by \(\epsilon(t)\delta_{1 \to m}\). The weighting parameter \(\epsilon\) fulfils \(0 \leq \epsilon(t) \leq 1, \forall t \in S\) and \(\sum_{t \in S} \epsilon(t) = 1\). Accordingly, the updated power schedule \(\bar{u}\) is provided:

\[
\bar{u} = \begin{cases} 
  u(t) - \delta_{1 \to m} & \text{if } t = t \\
  u(t) + \epsilon(t)\delta_{1 \to m} & \text{if } t \in S \\
  u(t) & \text{otherwise}
\end{cases}
\]

(28)

and its compact form

\[
\bar{u} = u + \delta_{1 \to m} \left( -\vec{e}_t + \sum_{t \in S} \epsilon(t)\vec{e}_t \right).
\]

(29)

Similarly, the \(n \to 1\) power swap policy from a set \(S\) of \(n\) time instants to a single time \(\vec{\tau}\) with power units \(\delta_{n \to 1}\) can be defined as

\[
\bar{u} = u + \delta_{n \to 1} \left( -\sum_{t \in S} \epsilon(t)\vec{e}_t + \vec{e}_\tau \right).
\]

(30)

where \(0 \leq \epsilon(t) \leq 1, \forall t \in S\) and \(\sum_{t \in S} \epsilon(t) = 1\).

Next, we determine the maximum feasible amount of power \(\delta_{1 \to m}\) and \(\delta_{n \to 1}\) that can be swapped:

\[
\delta_{1 \to m}(u, \vec{\tau}, \vec{\tau}) = \min\{a_{1 \to m}(u, \vec{\tau}, S), b_{1 \to m}(u, \vec{\tau}, c_{1 \to m}(u, \vec{\tau}, \vec{\tau})\},
\]

(31a)

\[
\delta_{n \to 1}(u, \vec{\tau}, \vec{\tau}) = \min\{a_{n \to 1}(u, \vec{\tau}, S), b_{n \to 1}(u, \vec{\tau}, S), c_{n \to 1}(u, \vec{\tau}, \vec{\tau})\}.
\]

(31b)

Each of the terms of the minimum function in (31) is now described in detail:

- Maximum feasible power increase at \(\vec{\tau}\):

\[
a_{1 \to m}(u, \vec{\tau}) = \min_{\epsilon(t) \in S} \frac{p^r - u(t)}{\epsilon(t)},
\]

(32a)
\[a_{n-1}(u, \bar{t}) = P^r - u(\bar{t}). \quad (32b)\]

- Maximum feasible power decrease at \(t\):
  \[b_{1-m}(u, \bar{t}) = u(t), \quad (33a)\]

- Minimum feasible power decrease at \(t\):
  \[b_{n-1}(u, \bar{S}) = \min_{t \in \bar{S}} \frac{u(t)}{t^r}. \quad (33b)\]

- Maximum regret-reducing power swap within region of constant gradient:
  To guarantee that the regret decreases by implementing \(1 \rightarrow m\) or \(n \rightarrow 1\) power swaps, the bounds \(c_{1-m}(u, t, \bar{S})\) in (31a) and \(c_{n-1}(u, \bar{S}, \bar{t})\) in (31b) are imposed in each case. This also guarantees that the rate of decrease as a function of the considered amount \(\delta\) of power swapped is constant. In particular, we let:
  \[
c_{1-m}(u, t, \bar{S}) := \max_{\delta \geq 0} \left\{ \delta \mid \frac{1}{\delta} \sum_{t \in \bar{S}} [\Pi(t) - \Pi(t-1)] \right\}
  \]
  s.t.
  \[
  R\left( u + \delta \left( \sum_{t \in \bar{S}} \varepsilon(t) \bar{e}_t - \bar{e}_t^r \right) \right) - R(u) \leq 0,
  \]
  \[
  \delta \lim_{\delta \to 0^+} \frac{R\left( u + \delta \left( \sum_{t \in \bar{S}} \varepsilon(t) \bar{e}_t - \bar{e}_t^r \right) \right) - R(u)}{\delta} \leq 0,
  \]
  \[
  \varepsilon(t) = \frac{1}{\sum_{t \in \bar{S}} [\Pi(t) - \Pi(t-1)]}, \quad \forall t \in \bar{S}
  \]
  and:
  \[
c_{n-1}(u, \bar{S}, \bar{t}) := \max_{\delta \geq 0} \left\{ \delta \mid \frac{1}{\delta} \sum_{t \in \bar{S}} [\Pi(t) - \Pi(t-1)] \right\}
  \]
  s.t.
  \[
  R\left( u + \delta \left( \bar{e}_t^r - \sum_{t \in \bar{S}} \varepsilon(t) \bar{e}_t \right) \right) - R(u) \leq 0,
  \]
  \[
  \delta \lim_{\delta \to 0^+} \frac{R\left( u + \delta \left( \bar{e}_t^r - \sum_{t \in \bar{S}} \varepsilon(t) \bar{e}_t \right) \right) - R(u)}{\delta} \leq 0,
  \]
  \[
  \varepsilon(t) = \frac{1}{\sum_{t \in \bar{S}} [\Pi(t) - \Pi(t-1)]}, \quad \forall t \in \bar{S}
  \]

Similar to the approach in the \(1 \rightarrow 1\) power swap, \(Q\) values are adopted to investigate the impact of power swap between multiple time instants on the worst regret variation. The calculation of both \(c_{1-m}\) and \(c_{n-1}\) is analytically illustrated in Appendix A.2.

Up until now, based on the described elementary swap policies, the proposed optimization strategy can be summarized as:

1. At the beginning, we initialize a feasible power schedule \(u(\cdot) \in \mathcal{U}\) and determine the associated \(Q\) values over the availability window.
2. By sorting the \(Q\) values, the three sets in (15) are defined and the initial regret is calculated according to (16).
3. Each scenario of power swap as described in Appendix A is tested to explore a feasible swap enabling the worst regret reduction. Specifically:
   (a) The \(1 \rightarrow 1\) power swap is composed of 6 subcases according to which set the considered time instants belong and the swapping direction. The \(Q\) value at the time instant with power swapped out decreases, whereas the other instant, with power swapped in, has an increased \(Q\) value.
   (b) Both \(1 \rightarrow m\) and \(n \rightarrow 1\) power swap are composed of 2 scenarios. In \(1 \rightarrow m\) (\(n \rightarrow 1\)) power swap, the \(Q\) values of \(m\) (\(n\)) time instants increase (decrease) by the same amount.
4. When the amount of power that can be swapped in all scenarios is zero, the device can no longer improve its regret and the algorithm has reached an optimal solution $u^*(\cdot)$.

By adopting this optimization strategy, from the theoretical point of view, only asymptotic convergence to the optimum is guaranteed as analyzed in the next section.

### 3.2 Optimality of the algorithm

This section engages to prove our iterative algorithm asymptotically converges to power allocations with the minimum worst regret by considering all elementary power swapping directions in Section 3.1 that strictly decrease the regret of the agent. For a feasible power allocation we define $\Delta^*(u)$ as follows:

$$\Delta^*(u) := \max_{u' \in F(u)} R(u) - R(u'),$$

where $F(u)$ denotes the set of power allocations resulting from $u$ after a single swap of any kind, which result in a nonincreasing regret. In this respect $\Delta^*(u)$ denotes the maximum regret decrease that can result from $u$ after a single swap.

For technical reasons, and in order to avoid the possibility of swaps that do not lead to a strict regret improvement, we embed our algorithm in the following difference inclusion:

$$u_{k+1} \in F(u_k) \cap \{u^+ : R(u_k) - R(u^+) \geq \varepsilon \Delta^*(u)\} := \tilde{F}(u_k),$$

for some positive $\varepsilon \in (0,1]$. The index $k$ denotes the iteration number and it is increased by one. Taking $\varepsilon = 1$, forces the algorithm to always pick a swap that yields the maximum regret reduction, however, more flexibility can be allowed and arbitrary positive values of $\varepsilon < 1$ are also suitable for convergence analysis. Whenever, for a particular kind of swap, the minimum in (24) or in the subsequent Equation (31) is 0 we include $u$ in $F(u)$. This guarantees the upper semicontinuity of $F(u)$ and consequently upper-semicontinuity of $\tilde{F}(u)$.

Our main result for this subsection is the following:

**Theorem 1.** Under Assumption 1 and 2 any solution of (37) initialized from a feasible power schedule asymptotically converges to a set of power allocations (normally a singleton) of minimal regret.

To prove that the proposed algorithm can always converge to allocations with minimal worst regret, it is necessary to show that the considered set of potential power swap directions is rich enough, and specifically that when no power swaps are found then we necessarily are on a power allocation of minimum regret. The following technical lemmas are presented.

**Lemma 2.** Let $u \in \mathcal{U}$ be a feasible power allocation, $S \subset \mathcal{A}$ and $\bar{S} \subset \mathcal{A}$ be two sets such that $S \cap \bar{S} = \emptyset$. If there exists $\delta > 0$ such that:

$$\sum_{t \in S} \delta(t) - \sum_{t \in \bar{S}} \delta(t) = 0,$$

and $\bar{u} = u + \sum_{t \in S} \delta(t)\mathbf{e}_t - \sum_{t \in \bar{S}} \delta(t)\mathbf{e}_t$ is feasible and fulfils

$$R(\bar{u}) < R(u),$$

then, at least one of the following holds:

1. There exists $t \in S$ and $\bar{\delta} > 0$ such that

$$\bar{u} = u + \sum_{t \in S} \bar{\delta}(t)\mathbf{e}_t - \left(\sum_{t \in \bar{S}} \bar{\delta}(t)\right)\mathbf{e}_t,$$

(38)
is feasible, and

\[ R(\tilde{u}) < R(u). \quad (39) \]

2. There exists \( \tilde{t} \in \bar{S} \) and \( \delta > 0 \) such that

\[
\tilde{u} = u - \sum_{t \in \bar{S}} \delta(t) \vec{e}_t + \left( \sum_{t \in \bar{S}} \delta(t) \right) \vec{e}_{\tilde{t}},
\]

(40)
is feasible, and

\[ R(\tilde{u}) < R(u). \quad (41) \]

Proof. See Appendix B.

Lemma 3. Let \( u \) be a feasible power allocation, \( t \in \mathcal{A} \) and \( \bar{S} \subseteq \mathcal{A} \) such that \( t \notin \bar{S} \). If there exists \( \delta > 0 \) such that:

\[
\delta(t) = \sum_{t \in \bar{S}} \delta(t),
\]

and \( \tilde{u} = u - \delta(t) \vec{e}_t + \sum_{t \in \bar{S}} \delta(t) \vec{e}_t \) is feasible and fulfills

\[ R(\tilde{u}) < R(u). \]

then, at least one of the following holds:

1. There exists \( t \in \bar{S} \) and \( \delta > 0 \) such that

\[
\tilde{u} = u - \delta(t) \vec{e}_t + \delta(t) \vec{e}_{\tilde{t}},
\]

(42)
is feasible, and

\[ R(\tilde{u}) < R(u). \quad (43) \]

2. There exists \( \bar{S} \subseteq \mathcal{A}_{\partial} \subseteq \bar{S} \) and \( \delta > 0 \) such that

\[
\tilde{u} = u - \left( \sum_{t \in \bar{S}} \delta(t) \right) \vec{e}_t + \sum_{t \in \bar{S}} \delta(t) \vec{e}_t,
\]

(44)
is feasible and

\[ R(\tilde{u}) < R(u). \quad (45) \]

Proof. See Appendix C.

Lemma 4. Let \( u \) be a feasible power allocation, \( \tilde{t} \in \mathcal{A} \) and \( \bar{S} \subseteq \mathcal{A} \) such that \( \tilde{t} \notin \bar{S} \). If there exists \( \delta > 0 \) such that:

\[
\delta(\tilde{t}) = \sum_{t \in \bar{S}} \delta(t),
\]

and \( \tilde{u} = u + \delta(\tilde{t}) \vec{e}_t - \sum_{t \in \bar{S}} \delta(t) \vec{e}_t \) is feasible and fulfills

\[ R(\tilde{u}) < R(u), \]
then, at least one of the following holds:

1. There exists \( t \in S \) and \( \hat{\delta} > 0 \) such that

\[
\bar{u} = u + \hat{\delta}(t)\bar{e}_t - \hat{\delta}(t)\bar{e}_t, \tag{46}
\]

is feasible, and

\[
R(\bar{u}) < R(u). \tag{47}
\]

2. There exists \( S^* \subseteq S \) and \( \hat{\delta} > 0 \) such that

\[
\bar{u} = u + \left( \sum_{t \in S^*} \hat{\delta}(t) \right)\bar{e}_t - \sum_{t \notin S^*} \hat{\delta}(t)\bar{e}_t, \tag{48}
\]

is feasible, and

\[
R(\bar{u}) < R(u). \tag{49}
\]

**Proof.** See Appendix D.

**Corollary 1.** According to Lemmas 2, 3, and 4, whenever there exists a feasible swap \( \hat{\delta} > 0 \) such that \( R(u + \delta) < R(u) \), at least one of the following swaps exists and decreases the regret:

i) \( 1 \rightarrow 1 \) swap: \( t \rightarrow \tilde{t} \) where \( t, \tilde{t} \in \mathcal{A}, t \neq \tilde{t} \).

ii) \( n \rightarrow 1 \) swap: \( S \rightarrow t \) where \( S \subseteq \partial L, |S| > |L_{in}| \) and \( t \in \mathcal{A}, t \notin S \).

iii) \( 1 \rightarrow m \) swap: \( t \rightarrow S \) where \( S \subseteq \partial L, |S| > |\partial L| - (K - |L_{in}|) \) and \( t \in \mathcal{A}, t \notin S \).

Corollary 1 enunciates that any power swap can be achieved by employing sequentially the two types of elementary swaps. We can conclude that the application of the iterative algorithm guarantees convergence to the optimal solution \( u^*(\cdot) \) for all feasible initial power schedule under Assumptions 1 and 2. In the final optimal power profile, no elementary power swap is convenient to reduce the regret.

We are now ready to prove Theorem 1.

**Proof.** Consider any solution \( u_k \) of (37). By construction, \( R(u_k) \) is nonincreasing and it admits a finite non-negative limit \( R_\infty \). Since \( u_k \) is bounded, the sequence admits an \( \omega \)-limit set \( \Omega \), and (by continuity of \( R(\cdot) \)) for every element \( \tilde{u} \in \Omega \) it holds \( R(\tilde{u}) = R_\infty \). Since \( F(\cdot) \) is upper-semicontinuous, then \( \Omega \) is weakly invariant, namely for all \( \tilde{u} \in \Omega \), there exists \( \tilde{u}^* \in F(\tilde{u}) \) such that \( \tilde{u}^* \in \Omega \). This, however implies that every element of \( \Omega \) has minimal regret. In fact, \( R(\tilde{u}^*) = R_\infty \) and \( \Delta^*(\tilde{u}) \leq (R(\tilde{u}) - R(\tilde{u}^*)) / \varepsilon = 0 \). By Corollary 1 however, the only possibility for \( \Delta^*(\tilde{u}) \) to be 0 is if \( R(\tilde{u}) \) has minimal regret.

---

**4 | EXAMPLES**

The performance of the proposed iterative algorithm is evaluated in this section through a series of case studies. All simulations are performed under a 24-h time horizon with the resolution of \( \Delta t = 30 \) min. For illustration, a single agent under various types of uncertain prices are considered which is followed by the implementation on a large population of agents.

The required energy and power for the single EV is assumed to be \( E^r = 30 \) kWh and \( P^r = 10 \) kW, respectively. Therefore, the minimum required charging duration for fully charged is 3 h.

- **Constant nominal prices, increasing uncertainties:** This scenario considers a price profile as shown on the left side of Figures 1–3. The red lines represent the price ranges within the availability window which starts at 2:00. As it could
FIGURE 1 Prices (left) and charging profile (right) when availability is 4 h, constant nominal prices, increasing uncertainties

FIGURE 2 Prices (left) and charging profile (right) when availability is 6 h, constant nominal prices, increasing uncertainties

FIGURE 3 Prices (left) and charging profile (right) when availability is 8 h, constant nominal prices, increasing uncertainties
be realistic to have more accurate prediction in the near future, an increasing price interval has been adopted. According to the resulting charging schedules, it is observed that the nature of the optimal power allocation, in this scenario, depends on the ratio between the available charging hours and the minimum required charging duration. In Figure 1, when the available charging time is less than twice of the minimum required charging duration, time instants with lower uncertainty would charge at a higher power. When the available charging time is greater than twice of the minimum required charging duration, the optimal charging schedule is reversed as in Figure 3. Figure 2 shows the optimal charging power is evenly allocated when the available charging time is equal to twice the minimum required charging duration. Notice that minimum regret policies are, in this case, optimal also with respect to nominal (average) prices (this is because such average price is constant and any policy leads to the same cost). On the other hand, neglecting uncertainty in the decision, by simply using nominal cost minimization, could lead to significantly suboptimal energy bills if prices happen to be lower or higher than expected.

Moreover, the computational time at different size of the considered problems are recorded in Table 1, which indicates that our proposed algorithm has no advantages compared to the linear programming approach if the number of decision variables are small and the prices have a constant average value.

- **Varying nominal prices, constant uncertainties:** To consider more diversified patterns of overlapping price intervals, a slowly varying nominal price profile and constant uncertainty interval are exhibited. Based on the results shown in Figures 4–6, it is concluded that the optimal schedule allocates more power to time instants with smaller average prices if the uncertainties are identical. Regarding the computational complexity, we can obtain that the proposed algorithm significantly outperforms the linear programming approach especially when the number of decision variables are large as shown in Table 2. For instance, the number of constraints in the linear OP is $\left(\begin{array}{c}20 \\ 6 \end{array}\right) = 38,760$ when the considered available time window is 10 h.

- **Varying nominal prices, increasing uncertainties:** In reality, this scenario could happen more frequently in applications, for example, the prediction of the price is more accurate at the beginning of the window and less certain afterwards. Figure 7 shows the optimal charging powers are almost flat due to the similar average prices and uncertainties. While the optimal schedule allocates less power at time instants with higher average and uncertain prices as in Figures 8 and 9.

| Available time window | 4 h  | 6 h  | 8 h  | 10 h |
|----------------------|-----|-----|-----|----|
| Iterative method     | 0.1221 | 0.6362 | 4.7387 | 107.9118 |
| Linear OP            | 0.0136 | 0.0285 | 0.7969 | 27.2422 |

**Figure 4** Prices (left) and charging profile (right) when availability is 4 h, varying nominal prices, constant uncertainties
Moreover, our proposed algorithm, in this scenario, also results in significantly higher computational efficiency at larger availability window (Table 3).

- **Constant lower prices, increasing uncertainties**: Figure 10 shows the results under a price realization with identical lower bounds and increasing upper bounds. The uncertainties in this scenario is increasing along the time, so the optimal power profile is to allocate more power at the beginning and less amount of power in the future.

- **Constant upper prices, increasing uncertainties**: While the worst prices are the same as in Figure 11, the optimal minimum regret strategy is to consume more power at time instants with smaller lower prices. Notice that if upper-prices
**FIGURE 7** Prices (left) and charging profile (right) when availability is 4 h, varying nominal prices, increasing uncertainties

**FIGURE 8** Prices (left) and charging profile (right) when availability is 6 h, varying nominal prices, increasing uncertainties

**FIGURE 9** Prices (left) and charging profile (right) when availability is 8 h, varying nominal prices, increasing uncertainties
are virtually constant but in fact slowly increasing, the minimum regret optimal policy will not change significantly. On the other hand, a worst case approach would concentrate all power at the initial times, potentially missing out considerable savings if the worst case does not materialize. Conversely, the cost of the minimum regret strategy when the highest prices occur is virtually the same as that of the optimal worst-case strategy.

5 | CONCLUSION

This article proposes a novel optimization method for a single agent to schedule its energy absorption profile in order to minimize the worst-case regret over all price scenarios. The algorithm guarantees convergence to an equilibrium charging...
schedule whose optimality is justified. Under different uncertainty hypotheses and flexibility parameters, the resulting computational efficiency is compared with the corresponding linear programming approach.

While only a single EV was considered in this article, it is possible to deploy this same algorithm in order to schedule power absorption profiles of EV fleets. However, unlike the case of a single EV which has negligible market power in electricity pricing, a large population of devices could induce price changes and this aspect has not been modeled in the current analysis, where a fixed (albeit uncertain) price profile is considered. Specifically, the worst regret of an individual agent may change if the broadcast price profile is updated by other devices through modifying their operation schedules. The coupling of the regret minimization and price variations can be simulated for large populations, but the overall convergence of the scheme is still an open research question. Future research will extend the algorithm to systems with large device populations and analytically characterize the convergence to the social equilibrium and optimality of the solution. The robustness and disturbance rejection properties will also be investigated. Finally, the performance of the worst regret-based scheme and worst cost method will be compared.

ACKNOWLEDGMENTS

This research was partly supported by two research projects: Integrated Development of Low-Carbon Energy Systems (IDLES, Grant EP/R045518/1) and Active Building Centre (ABC, Grant EP/V012053/1).

ORCID

Zihang Dong https://orcid.org/0000-0003-4541-7259

REFERENCES

1. European Commission A European strategy for low-emission mobility. technical report. European Commission; 2016.
2. Strbac G. Demand side management: benefits and challenges. Energy Policy. 2008;36(12):4419-4426.
3. Qdr Q. Benefits of demand response in electricity markets and recommendations for achieving them. technical report. US Department Energy; 2006.
4. Safdarian A, Fotuhi-Firuzabad M, Lehtonen M. Benefits of demand response operation on operation of distribution networks: a case study. IEEE Syst J. 2016;10(1):189-197.
5. Mohsenian-Rad AH, Wong VW, Jatskevich J, Schober R, Leon-Garcia A. Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid. IEEE Trans Smart Grid. 2010;1(3):320-331.
6. Ma Z, Callaway DS, Hiskens IA. Decentralized charging control of large populations of plug-in electric vehicles. IEEE Trans Control Syst Technol. 2011;21(1):67-78.
7. Fan Z. A distributed demand response algorithm and its application to PHEV charging in smart grids. IEEE Trans Smart Grid. 2012;3(3):1280-1290.
8. Zhang Y, Gatsis N, Giannakis GB. Robust energy management for microgrids with high-penetration renewables. IEEE Trans Sustain Energy. 2013;4(4):944-953.
9. Grammatico S. Dynamic control of agents playing aggregative games with coupling constraints. IEEE Trans Automat Contr. 2017;62(9):4537-4548.
10. Gong X, De Paola A, Angeli D, Strbac G. Distributed coordination of flexible loads using locational marginal prices. IEEE Trans Control Netw Syst. 2019;6(3):1097-1110.
11. Ma Z, Zou S, Ran L, Shi X, Hiskens IA. Efficient decentralized coordination of large-scale plug-in electric vehicle charging. Automatica. 2016;69:35-47. https://doi.org/10.1016/j.automatica.2016.01.035
12. De Paola A, Fele F, Angeli D, Strbac G. Distributed coordination of price-responsive electric loads: a receding horizon approach. Proceedings of the 2018 IEEE Conference on Decision and Control (CDC); 2018:6033-6040; IEEE.
13. Zheng Y, Song Y, Hill DJ, Meng K. Online distributed MPC-based optimal scheduling for EV charging stations in distribution systems. IEEE Trans Ind Inform. 2019;15(2):638-649.
14. Zhao S, Lin X, Chen M. Robust online algorithms for peak-minimizing EV charging under multistage uncertainty. IEEE Trans Automat Contr. 2017;62(11):5739-5754.
20. Korolko N, Sahinoglu Z. Robust optimization of EV charging schedules in unregulated electricity markets. IEEE Trans Smart Grid. 2017;8(1):149-157.
21. De Paola A, Angeli D, Strbac G. Price-based schemes for distributed coordination of flexible demand in the electricity market. IEEE Trans Smart Grid. 2017;8(6):3104-3116.
22. Song L, Xiao Y, Van Der Schaar M. Demand side management in smart grids using a repeated game framework. IEEE J Select Areas Commun. 2014;32(7):1412-1424.
23. Kizilkale AC, Mannor S, Caines PE. Large scale real-time bidding in the smart grid: a mean field framework. Proceedings of the 2012 IEEE 51st IEEE Conference on Decision and Control (CDC); 2012:3680-3687; IEEE.
24. Ma Z, Zou S, Ran L, Shi X, Hiskens IA. Efficient decentralized coordination of large-scale plug-in electric vehicle charging. Automatica. 2016;69:35-47.

AUTHOR BIOGRAPHIES

Zihang Dong received BE degree in Electrical and Electronics Engineering in 2014 from the University of Liverpool, UK. He received MSc degree and the PhD degree in Control Theory from Imperial College London, UK, in 2015 and 2019, respectively. He is currently a research associate at the Department of Electrical and Electronic Engineering of Imperial College London. His research interests include robust economic model predictive control, output feedback and state estimation, flexible demand response and the design of control strategies for the smart grids.

David Angeli was born in Siena, Italy, in 1971. He received the BS degree in Computer Science Engineering and the PhD in Control Theory from University of Florence, Florence, Italy, in 1996 and 2000, respectively. Since 2000 he was an Assistant and Associate Professor (2005) with the Department of Systems and Computer Science, University of Florence. In 2007 he was a visiting Professor with I.N.R.I.A de Rocquencourt, Paris, France, and since 2008, he joined as a Senior Lecturer the Department of Electrical and Electronic Engineering of Imperial College London, where he is currently a Reader in Nonlinear Control and the Director of the MSc in Control Systems. He has been a Fellow of the IEEE since 2015. He has served as an Associate Editor for IEEE Transactions in Automatic Control and Automatica. Overall he authored more than 100 journal papers in the areas of Stability of nonlinear systems, Control of constrained systems (MPC), Chemical Reaction Networks theory and Smart Grids.

Antonio De Paola received the MSc degree in Control Engineering from the University of Rome Tor Vergata, Rome, Italy, in 2011, and the PhD degree in Electrical and Electronic Engineering from Imperial College London, London, UK, in 2015. He is currently an Expert Consultant at REF-E, Milan, Italy and a Visiting Researcher at Imperial College London, UK. His main research interests include optimal control and game theory, with a specific focus on the design of novel control and market solutions for the deployment of a sustainable and decarbonized electricity grid.

Goran Strbac is a Professor of Electrical Energy Systems at Imperial College London, UK. His current research is focused on optimization of energy systems operation and investment, energy infrastructure reliability, and future energy markets.

How to cite this article: Dong Z, Angeli D, De Paola A, Strbac G. An iterative algorithm for regret minimization in flexible demand scheduling problems. Adv Control Appl. 2021;3(4):e92. doi: 10.1002/adc2.92
APPENDIX A. ITERATIVE ALGORITHM SCHEME

For ease of explanation, we also represent the maximum $Q$ value in $L_{\text{in}}$ and the minimum $Q$ value in $L_{\text{out}}$ by

$$L_{\text{in}}^\text{max} = \max_{t \in L_{\text{in}}} Q(t, u(t)), \quad L_{\text{out}}^\text{min} = \min_{t \in L_{\text{out}}} Q(t, u(t)), \quad \text{(A1)}$$

respectively. Then, we do the iterative optimization algorithm according to the following procedures which is mainly divided into two categories.

A.1 Swapping between two time instants

1. $|L_{\text{in}}| \geq 2$, $t_1 \in L_{\text{in}}$, $t_2 \in L_{\text{in}}$, $t_1 \neq t_2$

   We swap power

   $$\Delta = \min \{ P^r - u(\tilde{t}), u(t), \frac{L - Q(\tilde{t}, u(\tilde{t}))}{\Pi(\tilde{t})} \} \quad \text{(A2)}$$

   from $t$ to $\tilde{t}$, where

   $$\tilde{t} = \arg\min_{t \in \{ t_1, t_2 \}} \Pi(t), \quad t = \{ t_1, t_2 \} \setminus \tilde{t}. \quad \text{(A3)}$$

   The first and second term in (A2) correspond to $a(u, \tilde{t})$ and $b(u, t)$ in (24). While the last quantity in (A2), which represents $c(u, t, \tilde{t})$, is obtained by the inequality

   $$Q(\tilde{t}, u(\tilde{t}) + \Delta) = [\Pi(\tilde{t}) - \Pi(t) ][u(\tilde{t}) + \Delta] + \Pi(\tilde{t}) P^r \leq L, \quad \text{(A4)}$$

   where is to preserve the order of $Q$ values and hence the lower price.

   Then, the updated regret is

   $$R(u^+) = R(u) - \Pi(t) \Delta + \Pi(\tilde{t}) \Delta$$

   $$= R(u) - [\Pi(t) - \Pi(\tilde{t})] \Delta. \quad \text{(A5)}$$

   Since the definition in (A3) implies $[\Pi(t) - \Pi(\tilde{t})] \geq 0$, the regret is nonincreasing.

2. $|L_{\text{out}}| \geq 2$, $t_1 \in L_{\text{out}}$, $t_2 \in L_{\text{out}}$, $t_1 \neq t_2$

   We swap power

   $$\Delta = \min \left\{ P^r - u(\tilde{t}), u(t), \frac{Q(t, u(t)) - L}{\Pi(t) - \Pi(t)} \right\} \quad \text{(A6)}$$

   from $t$ to $\tilde{t}$, where

   $$\tilde{t} = \arg\min_{t \in \{ t_1, t_2 \}} \Pi(t), \quad t = \{ t_1, t_2 \} \setminus \tilde{t}. \quad \text{(A7)}$$

   The last quantity in (A6), corresponding to $c(u, t, \tilde{t})$ in (24), is obtained by the inequality

   $$Q(\tilde{t}, u(t) - \Delta) = [\Pi(\tilde{t}) - \Pi(t) ][u(t) - \Delta] + \Pi(\tilde{t}) P^r \geq L. \quad \text{(A8)}$$

   Then, the updated regret is

   $$R(u^+) = R(u) - \Pi(t) \Delta + \Pi(\tilde{t}) \Delta$$

   $$= R(u) - [\Pi(t) - \Pi(\tilde{t})] \Delta. \quad \text{(A9)}$$
Since the definition in (A7) implies \( |\Pi(t) - \Pi(i)| \geq 0 \), the regret is nonincreasing.

3. \(|\partial L| \geq 2, t_1 \in \partial L, t_2 \in \partial L, t_1 \neq t_2\)

Since the adopted prices for \( t \in \partial L \) in the regret evaluation depends on its cardinality, we consider the following two cases:

- \(|L_{\text{in}}| + |\partial L| = K\)

From \( t_1 \) to \( t_2 \), we swap power

\[
\Delta = \min\{P' - u(t_2), u(t_1), c(u, t_1, t_2)\},
\]

where

\[
c(u, t_1, t_2) = \min \left\{ \frac{L - L^\text{max}_{\text{in}}}{\Pi(t_1) - \Pi(t_1)}, \frac{L^\text{min}_{\text{out}} - L}{\Pi(t_2) - \Pi(t_2)} \right\},
\]

is determined by the following inequalities

\[
Q(t_1, u(t_1) - \Delta) = L - [\Pi(t_1) - \Pi(t)]\Delta \geq L^\text{max}_{\text{in}},
\]

\[
Q(t_2, u(t_2) + \Delta) = L + [\Pi(t_2) - \Pi(t)]\Delta \leq L^\text{min}_{\text{out}}.
\]

Then, with the prices \( \Pi(t_1) \) and \( \Pi(t_2) \) as \(|L_{\text{in}}| + |\partial L| = K\), the regret variation is expressed as

\[
R(u^+) = R(u) - \Pi(t_1)\Delta + \Pi(t_2)\Delta = R(u) - [\Pi(t_1) - \Pi(t_2)]\Delta.
\]

- \(|L_{\text{in}}| + |\partial L| > K\)

From \( t_1 \) to \( t_2 \), we swap power

\[
\Delta = \min\{P' - u(t_2), u(t_1), c(u, t_1, t_2)\},
\]

where

\[
c(u, t_1, t_2) = \min \left\{ \frac{L - L^\text{max}_{\text{in}}}{\Pi(t_1) - \Pi(t_1)}, \frac{L^\text{min}_{\text{out}} - L}{\Pi(t_2) - \Pi(t_2)} \right\},
\]

Accordingly, with the prices \( \Pi(t_1) \) and \( \Pi(t_2) \) as \(|L_{\text{in}}| + |\partial L| > K\), the regret is updated by

\[
R(u^+) = R(u) - \Pi(t_1)\Delta + \Pi(t_2)\Delta = R(u) - [\Pi(t_1) - \Pi(t_2)]\Delta.
\]

Therefore, in the case of swapping power between time instants \( t_1 \in \partial L, t_2 \in \partial L \), the policy is to swap \( \Delta \) as in (A10) or (A14) from \( t_1 \) to \( t_2 \), if \(|L_{\text{in}}| + |\partial L| = K\) \( \wedge \) \( |\Pi(t_1) - \Pi(t_2)| \) or \(|L_{\text{in}}| + |\partial L| > K\) \( \wedge \) \( |\Pi(t_1) - \Pi(t_2)| \).

4. \( L_{\text{in}} \neq \emptyset, t_1 \in L_{\text{in}}, t_2 \in \partial L\)

- \(|L_{\text{in}}| + |\partial L| = K\)

- From \( t_1 \) to \( t_2 \), we swap power

\[
\Delta_a = \min\{P' - u(t_2), u(t_1), c(u, t_1, t_2)\},
\]

where \( c(u, t_1, t_2) = \frac{L^\text{min}_{\text{out}} - L}{\Pi(t_2) - \Pi(t_2)} \). The regret is updated by

\[
R(u^+) = R(u) - \Pi(t_1)\Delta_a + \Pi(t_2)\Delta_a = R(u) - [\Pi(t_1) - \Pi(t_2)]\Delta_a.
\]
- From $t_2$ to $t_1$, we swap power

$$\Delta_b = \begin{cases} 
\min\{P' - u(t_1), u(t_2) , \frac{L_{\text{out}} - Q(t_1, u(t_1))}{\Pi(t_1) - \Pi(t_2)} \} & \text{if } |\partial L| = 1 \\
\min\{P' - u(t_1), u(t_2) , \frac{Q(t_1, u(t_1))}{\Pi(t_2) - \Pi(t_1)} \} & \text{if } |\partial L| > 1, 
\end{cases}$$

(A19)

and the regret is updated by

$$R(u^+) = R(u) - \Pi(t_2)\Delta_b + \Pi(t_1)\Delta_b$$

$$= R(u) - [\Pi(t_2) - \Pi(t_1)]\Delta_b.$$  

(A20)

Hence, when $|L_{\text{in}}| + |\partial L| = K$, the swap policy is

$$\Delta = \begin{cases} 
\Delta_a \text{ from } t_1 \text{ to } t_2 & \text{if } \Pi(t_1) > \Pi(t_2) \\
\Delta_b \text{ from } t_2 \text{ to } t_1 & \text{if } \Pi(t_1) < \Pi(t_2). 
\end{cases}$$

(A21)

- $|L_{\text{in}}| + |\partial L| > K$

  - From $t_1$ to $t_2$, we swap power

$$\Delta_a = \min\{P' - u(t_2), u(t_1) , c(u, t_1, t_2) \},$$

(A22)

where $c(u, t_1, t_2) = \frac{L_{\text{out}} - L}{\Pi(t_1) - \Pi(t_2)}$, and the regret is updated by

$$R(u^+) = R(u) - \Pi(t_1)\Delta_a + \Pi(t_2)\Delta_a$$

$$= R(u) - [\Pi(t_1) - \Pi(t_2)]\Delta_a.$$  

(A23)

- From $t_2$ to $t_1$, we swap power

$$\Delta_b = \min\{P' - u(t_1), u(t_2) , c(u, t_2, t_1) \}$$

(A24)

where $c(u, t_2, t_1) = \frac{Q(t_1, u(t_1))}{\Pi(t_2) - \Pi(t_1)}$, and the regret is updated by

$$R(u^+) = R(u) - \Pi(t_2)\Delta_b + \Pi(t_1)\Delta_b$$

$$= R(u) - [\Pi(t_2) - \Pi(t_1)]\Delta_b.$$  

(A25)

Hence, when $|L_{\text{in}}| + |\partial L| > K$, the swap policy is

$$\Delta = \begin{cases} 
\Delta_a \text{ from } t_1 \text{ to } t_2 & \text{if } \Pi(t_1) > \Pi(t_2) \\
\Delta_b \text{ from } t_2 \text{ to } t_1 & \text{if } \Pi(t_1) < \Pi(t_2). 
\end{cases}$$

(A26)

5. $L_{\text{out}} \neq \emptyset$, $t_1 \in L_{\text{out}}$, $t_2 \in \partial L$

- $|L_{\text{in}}| + |\partial L| = K$

  - From $t_1$ to $t_2$, we swap power

$$\Delta_a = \min\{P' - u(t_2), u(t_1) , c(u, t_1, t_2) \},$$

(A27)

where

$$c(u, t_1, t_2) = \min \left\{ \frac{L_{\text{out}}}{\Pi(t_2) - \Pi(t_1)} , \frac{Q(t_1, u(t_1)) - L}{\Pi(t_1) - \Pi(t_2)} \right\}$$

$$= \min \left\{ \frac{L_{\text{out}}}{\Pi(t_2) - \Pi(t_1)} , \frac{Q(t_1, u(t_1)) - L}{\Pi(t_1) - \Pi(t_2)} \right\}$$
is determined according to the inequality
\[ Q(t_2, u(t_2) + \Delta_a) \leq Q(t_1, u(t_1) - \Delta_a). \] (A28)

The regret is updated by
\[ R(u^*) = R(u) - \Pi(t_1)\Delta_a + \Pi(t_2)\Delta_a = R(u) - [\Pi(t_1) - \Pi(t_2)]\Delta_a. \] (A29)

- From \( t_2 \) to \( t_1 \), we swap power

\[ \Delta_b = \min \{ P^r - u(t_1), u(t_2), c(u, t_2, t_1) \} \] (A30)

where \( c(u, t_2, t_1) = \frac{L-L_{\text{max}}}{\Pi(t_2)-\Pi(t_1)} \), and the regret is updated by
\[ R(u^*) = R(u) - \Pi(t_2)\Delta_b + \Pi(t_1)\Delta_b = R(u) - [\Pi(t_2) - \Pi(t_1)]\Delta_b. \] (A31)

Hence, when \(|L_{\text{in}}| + |\partial L| = K\), the swap policy is
\[ \Delta = \begin{cases} \Delta_a & \text{from } t_1 \text{ to } t_2 \text{ if } \Pi(t_1) > \Pi(t_2) \\ \Delta_b & \text{from } t_2 \text{ to } t_1 \text{ if } \Pi(t_1) < \Pi(t_2). \end{cases} \] (A32)

- \(|L_{\text{in}}| + |\partial L| > K\)

- From \( t_1 \) to \( t_2 \), we swap power

\[ \Delta_a = \min \{ P^r - u(t_1), u(t_2), c(u, t_1, t_2) \}, \] (A33)

where
\[ c(u, t_1, t_2) = \min \left\{ \frac{L_{\text{out}} - \bar{L}}{\Pi(t_2) - \Pi(t_1)}, \frac{Q(t_1, u(t_1)) - \bar{L}}{\Pi(t_1) - \Pi(t_2) + \Pi(t_2) - \Pi(t_1)} \right\}. \] (A34)

and the regret is updated by
\[ R(u^*) = R(u) - \Pi(t_1)\Delta_a + \Pi(t_2)\Delta_a = R(u) - [\Pi(t_1) - \Pi(t_2)]\Delta_a. \] (A35)

- From \( t_2 \) to \( t_1 \), we swap power

\[ \Delta_b = \min \{ P^r - u(t_1), u(t_2), c(u, t_2, t_1) \}, \] (A36)

where \( c(u, t_2, t_1) = \frac{L-L_{\text{max}}}{\Pi(t_2)-\Pi(t_1)} \), and the regret is updated by
\[ R(u^*) = R(u) - \Pi(t_2)\Delta_b + \Pi(t_1)\Delta_b = R(u) - [\Pi(t_2) - \Pi(t_1)]\Delta_b. \] (A37)

Hence, when \(|L_{\text{in}}| + |\partial L| > K\), the swap policy is
\[ \Delta = \begin{cases} \Delta_a & \text{from } t_1 \text{ to } t_2 \text{ if } \Pi(t_1) > \Pi(t_2) \\ \Delta_b & \text{from } t_2 \text{ to } t_1 \text{ if } \Pi(t_1) < \Pi(t_2). \end{cases} \] (A38)
6. \( L_{\text{in}} \neq \emptyset, L_{\text{out}} \neq \emptyset, t_1 \in L_{\text{in}}, t_2 \in L_{\text{out}} \)

- From \( t_1 \) to \( t_2 \), we swap power

\[
\Delta_a = \min \{ P^r - u(t_2), u(t_1) \}, \tag{A39}
\]

in which there is no condition on \( c(u, t_1, t_2) \) since the prices at all-time instants will be preserved for all \( \Delta_a > 0 \).

The regret is updated by

\[
R(u^+) = R(u) - \Pi(t_1)\Delta_a + \Pi(t_2)\Delta_a
= R(u) - [\Pi(t_1) - \Pi(t_2)]\Delta_a. \tag{A40}
\]

- From \( t_2 \) to \( t_1 \), we swap power

\[
\Delta_b = \min \{ P^r - u(t_1), u(t_2), c(u, t_2, t_1) \}, \tag{A41}
\]

where

\[
c(u, t_2, t_1) = \min \left\{ \frac{L - Q(t_1, u(t_1))}{\Pi(t_1) - \Pi(t_2)}, \frac{Q(t_2, u(t_2)) - L}{\Pi(t_2) - \Pi(t_1)} \right\}. \tag{A42}
\]

Then, the updated regret is

\[
R(u^+) = R(u) - \Pi(t_2)\Delta_b + \Pi(t_1)\Delta_b
= R(u) - [\Pi(t_2) - \Pi(t_1)]\Delta_b. \tag{A43}
\]

Therefore, in the case of swapping power between time instants \( t_1 \in L_{\text{in}}, t_2 \in L_{\text{out}} \), the policy is defined as follows:

\[
\Delta = \begin{cases} 
\Delta_a \text{ from } t_1 \text{ to } t_2 & \text{if } \Pi(t_2) > \Pi(t_1) \\
\Delta_b \text{ from } t_2 \text{ to } t_1 & \text{if } \Pi(t_2) < \Pi(t_1)
\end{cases} \tag{A44}
\]

### A.2 Power transfer between a single time instant and multiple times

Besides the swaps between two time instants, one needs to consider power swaps simultaneously operating between some instants in \( \partial L \) by preserving alignment of \( Q \) variables. Due to the interest in feasible swaps, we only consider time instants in \( \partial L \) where either upper or lower bounds are not active

\[
\partial L^* = \{ t \in \partial L \mid u(t) > 0 \} \tag{A45a}
\]

\[
\partial L^{**} = \{ t \in \partial L \mid u(t) < P^r \} \tag{A45b}
\]

with the corresponding cardinality \( |\partial L^*| = K_{\partial L^*} \geq 2 \) and \( |\partial L^{**}| = K_{\partial L^{**}} \geq 2 \).

To evaluate the regret after swapping, we need to first partition the set \( \partial L^* \) into two subsets \( \partial L^*_{\text{in}} \) and \( \partial L^*_{\text{out}} \) such that \( \partial L^* = \partial L^*_{\text{in}} \cup \partial L^*_{\text{out}} \), \( \partial L^*_{\text{in}} \cap \partial L^*_{\text{out}} = \emptyset \) and \( |L_{\text{in}}| + |L_{\text{in}}| = K \). The same applies to \( \partial L^{**} \), and the two sets are denoted by \( \partial L^{**}_{\text{in}} \) and \( \partial L^{**}_{\text{out}} \). While this partition is not uniquely defined, the cardinality condition is enough to guarantee that any subsequent regret evaluation is going to be independent from the choice carried out in selecting the partition.

1. From \( \partial L^* \) to \( t_0 \in L_{\text{in}} \)

Let us denote the amount of power added to \( t_0 \) by \( \delta(t_0) \). In order to decrease the corresponding \( Q \) function by the same value for any time in \( \partial L^* \) the power to be swapped out must satisfy the following equations:

\[
\sum_{\tau \in \partial L^*} \delta(\tau) = \delta(t_0)
\]

\[
|\Pi(t_1) - \Pi(t_1)|\delta(t_1) = |\Pi(t_2) - \Pi(t_2)|\delta(t_2) \quad \forall \; \tau_1, \tau_2 \in \partial L^*. \tag{A46}
\]
By solving the above equations, we obtain the swapped out power at \( \bar{\tau} \in \partial L^* \)

\[
\delta(\bar{\tau}) = G(\bar{\tau}, t_0)\delta(t_0),
\]

where

\[
G(\bar{\tau}, t_0) = \frac{1}{\sum_{r \in \partial L^*} \frac{\Pi(\tau) - \Pi(\bar{\tau})}{\Pi(\tau) - \Pi(r)}}.
\]

We swap power

\[
\delta(t_0) = \min\{P - u(t_0), \min_{r \in \partial L^*} \frac{u(r)}{G(r, t_0)}, c_{n-1}(u, \partial L^*, t_0)\},
\]

where the second term corresponds to \( b_{n-1} \) in (31b) and the third term is selected according to

\[
c_{n-1}(u, \partial L^*, t_0) = \min \left\{ \sum_{r \in \partial L^*} \frac{L - L_{\text{in}}^{\text{max}}}{\Pi(\tau) - \Pi(r)}, \frac{L - Q(t_0, u(t_0))}{\Pi(t_0) - \Pi(t_0)} + \frac{1}{\sum_{r \in \partial L^*} \frac{1}{\Pi(\tau) - \Pi(r)}} \right\}
\]

Then, the regret is updated by

\[
R(u^*) = R(u) - \sum_{r \in \partial L^*_m} \Pi(r)\delta(r) - \sum_{r \in \partial L^*_out} \Pi(r)\delta(r) + \Pi(t_0)\delta(t_0)
\]

\[
= R(u) - \left[ \sum_{r \in \partial L^*_m} \Pi(r)G(r, t_0) + \sum_{r \in \partial L^*_out} \Pi(r)G(r, t_0) - \Pi(t_0) \right] \delta(t_0).
\]

2. From \( t_0 \in L_{\text{in}} \cup \partial L^* \) to \( \partial L^{**} \)

Let us denote the amount of power removed from \( t_0 \) by \( \delta(t_0) \). In order to increase the corresponding \( Q \) function by the same value for any time in \( \partial L^{**} \), the power to be swapped out must satisfy the following equations:

\[
\sum_{r \in \partial L^{**}} \delta(r) = \delta(t_0)
\]

\[
[\Pi(\tau_1) - \Pi(\tau_2)]\delta(\tau_1) = [\Pi(\tau_1) - \Pi(\tau_2)]\delta(\tau_2) \quad \forall \ \tau_1, \tau_2 \in \partial L^{**}.
\]

By solving the above equations, we obtain the swapped out power at \( \bar{\tau} \in \partial L^{**} \)

\[
\delta(\bar{\tau}) = G(\bar{\tau}, t_0)\delta(t_0),
\]

where

\[
G(\bar{\tau}, t_0) = \frac{1}{\sum_{r \in \partial L^{**}} \frac{\Pi(\tau) - \Pi(\bar{\tau})}{\Pi(\tau) - \Pi(r)}}.
\]

We swap power

\[
\delta(t_0) = \min \left\{ \min_{r \in \partial L^{**}} \frac{P - u(t_0)}{G(r, t_0)}, u(t_0), c_{1 \rightarrow m}(u, t_0, \partial L^{**}) \right\},
\]

where the first term corresponds to \( a_{1 \rightarrow m} \) in (31a) and the third term

\[
c_{1 \rightarrow m}(u, t_0, \partial L^{**}) = \sum_{r \in \partial L^*} \frac{L_{\text{out}}^{\text{min}} - L}{\Pi(r) - \Pi(r)}.
\]
Then the regret is updated by

\[ R(u^*) = R(u) - \Pi(t_0)\delta(t_0) + \sum_{\tau \in \partial L^*} \Pi(\tau)\delta(\tau) + \sum_{\tau \in \partial L^{**}} \Pi(\tau)\delta(\tau) \]

\[ = R(u) - \left[ \Pi(t_0) - \sum_{\tau \in \partial L^*} \Pi(\tau)G(\tau, t_0) - \sum_{\tau \in \partial L^{**}} \Pi(\tau)G(\tau, t_0) \right] \delta(t_0). \quad (A57) \]

3. From \( \partial L^* \) to \( t_0 \in \partial L^{**} \cup \partial L^* \)

Let us denote the amount of power added to \( t_0 \) by \( \delta(t_0) \). In order to decrease the corresponding \( Q \) function by the same value for any time in \( \partial L^* \), the power to be swapped out must satisfy the following equations:

\[ \sum_{\tau \in \partial L^*} \delta(\tau) = \delta(t_0) \]

\[ [\Pi(\tau_1) - \Pi(\tau_1)]\delta(\tau_1) = [\Pi(\tau_2) - \Pi(\tau_2)]\delta(\tau_2) \quad \forall \tau_1, \tau_2 \in \partial L^*. \quad (A58) \]

By solving the above equations, we obtain the swapped out power at \( \tau \in \partial L^* \)

\[ \delta(\tau) = G(\tau, t_0)\delta(t_0), \quad (A59) \]

where

\[ G(\tau, t_0) = \frac{1}{\sum_{\tau \in \partial L^*} [\Pi(\tau) - \Pi(\tau_1)]}. \quad (A60) \]

We swap power

\[ \delta(t_0) = \min\{P^* - u(t_0), \min_{\tau \in \partial L^*} \frac{u(t)}{G(\tau, t_0)}, c_{n-1}(u, \partial L^*, t_0)\}, \quad (A61) \]

where the second term corresponds to \( b_{n-1} \) in (31b) and the third term

\[ c_{n-1}(u, \partial L^*, t_0) = \sum_{\tau \in \partial L^*} \frac{L - L^{\max}}{[\Pi(\tau) - \Pi(\tau_1)]}. \quad (A62) \]

Then the regret is updated by

\[ R(u^*) = R(u) - \sum_{\tau \in \partial L^*} \Pi(\tau)\delta(\tau) - \sum_{\tau \in \partial L^{**}} \Pi(\tau)\delta(\tau) + \sum_{\tau \in \partial L^*} \Pi(\tau)G(\tau, t_0) - \sum_{\tau \in \partial L^{**}} \Pi(\tau)G(\tau, t_0) \delta(t_0). \quad (A63) \]

4. From \( t_0 \in \partial L^* \) to \( \partial L^{**} \)

Let us denote the amount of power removed from \( t_0 \) by \( \delta(t_0) \). In order to increase the corresponding \( Q \) function by the same value for any time in \( \partial L^{**} \), the power to be swapped out must satisfy the following equations:

\[ \sum_{\tau \in \partial L^{**}} \delta(\tau) = \delta(t_0) \]

\[ [\Pi(\tau_1) - \Pi(\tau_1)]\delta(\tau_1) = [\Pi(\tau_2) - \Pi(\tau_2)]\delta(\tau_2) \quad \forall \tau_1, \tau_2 \in \partial L^{**}. \quad (A64) \]

By solving the above equations, we obtain the swapped out power at \( \tau \in \partial L^{**} \)

\[ \delta(\tau) = G(\tau, t_0)\delta(t_0), \quad (A65) \]
where

\[ G(\bar{\tau}, t_0) = \frac{1}{\sum_{\tau \in \partial L^{**}} \frac{1}{|\Pi(\tau) - \Pi(\bar{\tau})|}}. \]  \hfill (A66)

We swap power

\[
\delta(t_0) = \min \left\{ \min_{\tau \in \partial L^{**}} \frac{\hat{\epsilon}(\tau)}{G(\tau, t_0)}, u(t_0), c_{1 \rightarrow m}(u, t_0, \partial L^{**}) \right\},
\]  \hfill (A67)

where the first term corresponds to \(a_{1 \rightarrow m}\) in (31a) and the third term

\[
c_{1 \rightarrow m}(u, t_0, \partial L^{**}) = \min \left\{ \sum_{\tau \in \partial L^{**}} \frac{L_{\text{out}}^{\min} - \bar{L}}{|\Pi(\tau) - \Pi(\bar{\tau})|}, \frac{Q(t_0, u(t_0)) - \bar{L}}{|\Pi(t_0) - \Pi(\bar{\tau})|} + \frac{1}{\sum_{\tau \in \partial L^{**}} |\Pi(\tau) - \Pi(\bar{\tau})|} \right\}.
\]  \hfill (A68)

Then the regret is updated by

\[
R(u^*) = R(u) - \Pi(t_0)\delta(t_0) + \sum_{\tau \in \partial L^{**}} \Pi(\tau)\delta(\tau) + \sum_{\tau \in \partial L^{**}} \Pi(\tau)\delta(\tau)
\]

\[
= R(u) - \left[ \Pi(t_0) - \sum_{\tau \in \partial L^{**}} \Pi(\tau)G(\tau, t_0) - \sum_{\tau \in \partial L^{**}} \Pi(\tau)G(\tau, t_0) \right] \delta(t_0).
\]  \hfill (A69)

**APPENDIX B. PROOF OF LEMMA 2**

Notice that the function \(R(\cdot)\) is convex and there exists \(\delta > 0\) such that \(R(u + \delta) < R(u)\), then for any \(\varepsilon \in (0, 1)\), the regret of power allocation \(u + \varepsilon \delta\) fulfills

\[ R(u + \varepsilon \delta) = R((1 - \varepsilon)u + \varepsilon(u + \delta)) \leq (1 - \varepsilon)R(u) + \varepsilon R(u + \delta) < R(u). \]

Since \(R(\cdot)\) is piecewise linear, there exists \(\bar{\varepsilon} \in (0, 1)\) with \(\bar{\delta} = \bar{\varepsilon} \delta\) sufficiently small such that for all \(\varepsilon \in (0, \bar{\varepsilon})\) and \(\bar{\delta} = \bar{\varepsilon} \delta\), the sets \(L_{\text{in}}^\delta\), \(L_{\text{out}}^\delta\), and \(L_{\text{out}}^\delta\) are constant, where \(L_{\text{in}}^\delta\), \(L_{\text{out}}^\delta\), and \(L_{\text{out}}^\delta\) corresponds to \(L_{\text{in}}\), \(\partial L\), and \(L_{\text{out}}\) for the schedule \(u + \delta\). Moreover, the power schedule \(u + \bar{\delta}\) is also feasible because \(u + \bar{\delta} = \bar{\varepsilon}(u + \delta) + (1 - \bar{\varepsilon})u\).

Next, based on the power schedule \(u\), let us partition \(S\) and \(\bar{S}\) according to:

\[
S = S_{\text{in}} \cup S_{\text{out}} \cup S_{\text{out}} \quad \text{where} \quad S_{\text{in}} = S \cap L_{\text{in}}, S_{\text{in}} = S \cap \partial L, S_{\text{out}} = S \cap L_{\text{out}}.
\]

\[
\bar{S} = \bar{S}_{\text{in}} \cup \bar{S}_{\text{out}} \cup \bar{S}_{\text{out}} \quad \text{where} \quad \bar{S}_{\text{in}} = \bar{S} \cap L_{\text{in}}, \bar{S}_{\text{in}} = \bar{S} \cap \partial L, \bar{S}_{\text{out}} = \bar{S} \cap L_{\text{out}}.
\]

Swapping power \(\bar{\delta} > 0\) from \(S\) to \(\bar{S}\) only changes the costs associated with the time instants belonging to these two sets, therefore it is essential to evaluate their relation to \(L_{\text{in}}^\delta\), \(L_{\text{out}}^\delta\), and \(L_{\text{out}}^\delta\) which is dependent on the cardinality of \(S_{\text{out}}\) and \(\bar{S}_{\text{out}}\).

Specifically, for \(S_{\text{out}}\), the following two scenarios should be considered:

(i) \(|S_{\text{out}}| < K - |L_{\text{in}}|\):

\[
S_{\text{in}} \subseteq L_{\text{in}}^\delta, \quad S_{\text{out}} \subseteq L_{\text{out}}^\delta
\]

(ii) \(|S_{\text{out}}| \geq K - |L_{\text{in}}|\):

\[
S_{\text{in}} \subseteq L_{\text{in}}^\delta, \quad S_{\text{out}} \supseteq \partial L^\delta, \quad S_{\text{out}} \subseteq L_{\text{out}}^\delta
\]
Similarly, for \( \overline{S}_{\delta t} \), the set \( \overline{S} \) fulfills

(i) \( |\overline{S}_{\delta t}| \leq |\partial L| - (K - |L_{in}|) \)

\[
\overline{S}_{in} \subseteq I_{in}^\delta, \quad \overline{S}_{\delta t} \subseteq I_{out}^\delta, \quad \overline{S}_{out} \subseteq L_{out}^\delta
\]

(ii) \( |\overline{S}_{\delta t}| > |\partial L| - (K - |L_{in}|) \)

\[
\overline{S}_{in} \subseteq I_{in}^\delta, \quad \overline{S}_{\delta t} \supseteq \partial L^\delta, \quad \overline{S}_{out} \subseteq L_{out}^\delta
\]

Note that \( |\overline{S}_{\delta t}| + |\overline{S}_{\delta t}| \leq |\partial L| \), hence at least one of the following inequalities holds:

\[
|\overline{S}_{\delta t}| \leq K - |L_{in}| \quad \text{or} \quad |\overline{S}_{\delta t}| \leq |\partial L| - (K - |L_{in}|).
\]

Thus, the variation of total regret due to \( \delta > 0 \) follows one of the three cases:

1. \( |\overline{S}_{\delta t}| \leq K - |L_{in}| \) and \( |\overline{S}_{\delta t}| \leq |\partial L| - (K - |L_{in}|) \)

For the feasible swap \( \delta > 0 \),

\[
R(u + \delta) - R(u) = - \sum_{t \in \overline{S}_{in}} \Pi(t)\delta(t) + \sum_{t \in \overline{S}_{\delta t}} \Pi(t)\delta(t) + \sum_{t \in \overline{S}_{out}} \Pi(t)\delta(t) + \sum_{t \in \overline{S}_{\delta t}} \Pi(t)\delta(t) + \sum_{t \in \overline{S}_{out}} \Pi(t)\delta(t).
\]

Then, the regret reduction \( R(u + \delta) - R(u) \) can be rewritten in any of the following two forms:

(i) based on \( t \in \overline{S} \):

\[
R(u + \delta) - R(u) = \sum_{t \in \overline{S}_{in}} (-\Pi(t)\delta(t) + \epsilon(t, S)\Delta(S)) + \sum_{t \in \overline{S}_{\delta t}} (-\Pi(t)\delta(t) + \epsilon(t, S)\Delta(S))
\]

\[
+ \sum_{t \in \overline{S}_{out}} (-\Pi(t)\delta(t) + \epsilon(t, S)\Delta(S))
\]

\[
= \sum_{t \in \overline{S}_{in}} \left( R(u + \delta(t), S) - R(u) \right) + \sum_{t \in \overline{S}_{\delta t}} \left( R(u + \delta(t), S) - R(u) \right)
\]

\[
+ \sum_{t \in \overline{S}_{out}} \left( R(u + \delta(t), S) - R(u) \right),
\]

where \( \epsilon(t, S) = \frac{\delta(t)}{\sum_{u \in S} \delta(t)} \) and the power swap \( \delta(t, S) \) is defined as

\[
\delta(t, S) = -\delta(t)\overline{e}_t + \sum_{t \in S} \epsilon(t, S)\delta(t)\overline{e}_t = -\left( \sum_{t \in S} \delta(t) \right)\overline{e}_t + \sum_{t \in S} \delta(t)\overline{e}_t
\]

with \( \delta(t) = \epsilon(t, S)\delta(t), \forall t \in \overline{S} \).

Since by construction \( R(u + \delta) - R(u) < 0 \), there exists \( t \in \overline{S} \) such that the power allocation \( u + \delta(t, S) \) is feasible and the resulting regret variation fulfills \( R(u + \delta(t, S)) - R(u) < 0 \), which implies conclusion 1 in the Lemma.

(ii) based on \( t \in \overline{S} \):

\[
R(u + \delta) - R(u) = \sum_{t \in \overline{S}_{in}} (-\epsilon(t, S)\Delta(S) + \Pi(t)\delta(t)) + \sum_{t \in \overline{S}_{\delta t}} (-\epsilon(t, S)\Delta(S) + \Pi(t)\delta(t)) + \sum_{t \in \overline{S}_{out}} (-\epsilon(t, S)\Delta(S) + \Pi(t)\delta(t))
\]

\[
= \sum_{t \in \overline{S}_{in}} \left( R(u + \delta(t, S)) - R(u) \right) + \sum_{t \in \overline{S}_{\delta t}} \left( R(u + \delta(t, S)) - R(u) \right) + \sum_{t \in \overline{S}_{out}} \left( R(u + \delta(t, S)) - R(u) \right),
\]
where $\epsilon(t, \tilde{S}) = \frac{\delta(t)}{\sum_{t \in S} \delta(t)}$ and the power swap

$$\delta(S, t) = \delta(t) \vec{e}_t - \sum_{t \in S} \epsilon(t, \tilde{S}) \delta(t) \vec{e}_t = \left( \sum_{t \in S} \delta(t) \right) \vec{e}_t - \sum_{t \in S} \delta(t) \vec{e}_t$$

with $\delta(t) = \epsilon(t, \tilde{S}) \delta(t)$, $\forall t \in \tilde{S}$.

Since by construction $R(u + \tilde{\delta}) - R(u) < 0$, there exists $t \in \tilde{S}$ such that the power allocation $u + \tilde{\delta}(S, t)$ is feasible and the resulting regret variation fulfills $R(u + \tilde{\delta}(S, t)) - R(u) < 0$, which implies conclusion 2 in the Lemma.

2. $|S'_{dl}| \leq K - |L_{in}|$ and $|S'_{dl}| > |\partial L| - (K - |L_{in}|)$

For the feasible swap $\tilde{\delta} > 0,$

$$R(u + \tilde{\delta}) - R(u) = - \sum_{t \in S_d'} \sum_{t \in S_{dl}} \Pi(t) \delta(t) + \sum_{t \in S_{dl}} \sum_{t \in S_d} \Pi(t) \delta(t) + \sum_{t \in S_{dl}} \sum_{t \in S_d} \Pi(t) \delta(t)$$

$$+ \sum_{t \in S_d} \sum_{t \in S_{dl}} \Pi(t) \delta(t) + \sum_{t \in S_{dl}} \sum_{t \in S_d} \Pi(t) \delta(t) + \sum_{t \in S_{dl}} \sum_{t \in S_d} \Pi(t) \delta(t),$$

where $S'_{dl} = S'_{dl} \cup S'^{**}_{dl}$ is some set partition such that $|S'_{dl}| = |S'_{dl}| - (|\partial L| - (K - |L_{in}|))$, $|S'^{**}_{dl}| = |\partial L| - (K - |L_{in}|)$, $S'_{dl} \subseteq L_{in} \cup \partial \delta$ and $S'^{**}_{dl} \subseteq L_{out}$. Hence, the regret reduction $R(u + \tilde{\delta}) - R(u)$ is also written as

$$R(u + \tilde{\delta}) - R(u) = \sum_{t \in S_d} \left( R(u + \tilde{\delta}(t, \tilde{S})) - R(u) \right) + \sum_{t \in S_{dl}} \left( R(u + \tilde{\delta}(t, \tilde{S})) - R(u) \right) + \sum_{t \in S_{dl}} \left( R(u + \tilde{\delta}(t, \tilde{S})) - R(u) \right) < 0.$$

with the power swap policy

$$\delta(t, \tilde{S}) = -\delta(t) \vec{e}_t + \epsilon(t, \tilde{S}) \delta(t) \vec{e}_t = - \left( \sum_{t \in \tilde{S}} \delta(t) \right) \vec{e}_t + \sum_{t \in \tilde{S}} \delta(t) \vec{e}_t,$$

where $\delta(t) = \epsilon(t, \tilde{S}) \delta(t)$, $\forall t \in \tilde{S}$. Notice that $\delta(t, \tilde{S})$ feeds energy into $\tilde{S}$ according to the same relative powers as $\tilde{\delta}$, which entails the same relative variations of $Q$ and accordingly the same prices in the expression of the regret variation.

Thus, there exists $t \in S$ such that the power allocation $u + \tilde{\delta}(t, \tilde{S})$ is feasible and the resulting regret variation fulfills $R(u + \tilde{\delta}(t, \tilde{S})) - R(u) < 0$, which implies conclusion 1 in the Lemma.

3. $|S'_{dl}| > K - |L_{in}|$ and $|S'_{dl}| \leq |\partial L| - (K - |L_{in}|)$

For the feasible swap $\tilde{\delta} > 0,$

$$R(u + \tilde{\delta}) - R(u) = - \sum_{t \in S_d} \sum_{t \in S_{dl}} \Pi(t) \delta(t) + \sum_{t \in S_{dl}} \sum_{t \in S_d} \Pi(t) \delta(t) + \sum_{t \in S_{dl}} \sum_{t \in S_d} \Pi(t) \delta(t)$$

$$+ \sum_{t \in S_d} \sum_{t \in S_{dl}} \Pi(t) \delta(t) + \sum_{t \in S_{dl}} \sum_{t \in S_d} \Pi(t) \delta(t) + \sum_{t \in S_{dl}} \sum_{t \in S_d} \Pi(t) \delta(t),$$

where $S'_{dl} = S'_{dl} \cup S'^{**}_{dl}$ is some set partition such that $|S'_{dl}| = K - |L_{in}|$, $|S'^{**}_{dl}| = |S'_{dl}| - (K - |L_{in}|)$, $S'_{dl} \subseteq L_{in} \cup \partial \delta$ and $S'^{**}_{dl} \subseteq L_{out}$. Hence, the regret reduction $R(u + \tilde{\delta}) - R(u)$ is also written as
\[ R(u + \delta) - R(u) = \sum_{t \in S_{in}} (-\varepsilon(t, S) \Delta(S) + \Pi(t) \delta(t)) + \sum_{t \in S_{out}} (-\varepsilon(t, S) \Delta(S) + \Pi(t) \delta(t)) \]
\[ + \sum_{t \in S_{out}} (-\varepsilon(t, S) \Delta(S) + \Pi(t) \delta(t)) \]
\[ = \sum_{t \in S_{in}} (R(u + \delta(t)) - R(u)) + \sum_{t \in S_{out}} (R(u + \delta(t)) - R(u)) \]
\[ + \sum_{t \in S_{out}} (R(u + \delta(t)) - R(u)) < 0 \]

with the power swap policy

\[ \hat{\delta}_{(t)} = \delta(t) e_t - \sum_{t \in S} \varepsilon(t, S) \delta(t) e_t = \left( \sum_{t \in S} \hat{\delta}(t) \right) e_t - \sum_{t \in S} \hat{\delta}(t) e_t, \]

where \( \hat{\delta}(t) = \varepsilon(t, S) \delta(t), \forall t \in S \). Notice that \( \hat{\delta}_{(S)} \) extracts energy from \( S \) according to the same relative powers as \( \delta \), which entails the same relative variations of \( Q \) and accordingly the same prices in the expression of the regret variation.

Thus, there exists \( t \in S \) such that the power swap \( u + \hat{\delta}_t \) is feasible and the resulting regret variation fulfills \( R(u + \hat{\delta}_{(S)}) - R(u) < 0 \), which implies conclusion 2 in the Lemma.

Therefore, if the \( S \to S \) \((n \to m)\) swap is convenient to decrease the total regret, there exists at least a \( 1 \to m \) or \( n \to 1 \) power swapping policy \( \hat{\delta} > 0 \) such that the total regret is reduced.

**APPENDIX C. PROOF OF LEMMA 3**

Following the same lines of proof in Lemma 2, there exists \( \tilde{\delta} = -\tilde{\delta}(t) e_t + \sum_{t \in S} \delta(t) e_t \), sufficiently small such that the power swap \( u + \tilde{\delta} \) is feasible, the sets \( L_{in}^\delta, t_{\partial}^\delta, \) and \( L_{out}^\delta \) are constant, and \( R(u + \tilde{\delta}) - R(u) < 0 \). Specifically, the regret variation is evaluated according to the following three cases:

(i) \(|\partial L| = K - |L_{in}|\):

For the feasible power swap \( \tilde{\delta} > 0 \), it holds

\[ S_{in} \subseteq t_{in}^\delta, \quad \bar{S}_{out} \subseteq t_{\partial}^\delta \cup L_{out}^\delta, \quad S_{out} \subseteq L_{out}^\delta, \]

hence, the total regret variation is calculated as

\[ R(u + \tilde{\delta}) - R(u) = \begin{cases} -\Pi(t) \tilde{\delta}(t) + \sum_{t \in S_{in} \cup S_{out}} \Pi(t) \tilde{\delta}(t) + \sum_{t \in S_{out}} \Pi(t) \tilde{\delta}(t) & \text{if } t \in L_{in} \cup \partial L \\ -\Pi(t) \tilde{\delta}(t) + \sum_{t \in S_{in} \cup S_{out}} \Pi(t) \tilde{\delta}(t) + \sum_{t \in S_{out}} \Pi(t) \tilde{\delta}(t) & \text{if } t \in L_{out} \end{cases} \]

\[ = \sum_{t \in S} (R(u + \tilde{\delta}_{(t)}) - R(u)) < 0, \quad (C1) \]

where the change of power allocation is defined as

\[ \tilde{\delta}_{(t)} = -\tilde{\delta}(t) e_t + \delta(t) e_t, \forall t \in S. \]

Since \( \tilde{\delta}_{(t)} \) transfers energy from \( t \) into \( t \) according to the same relative powers as \( \hat{\delta} \), it preserves the same relative variations of \( Q \) and accordingly the same prices in the expression of the regret variation.

Then, by Equation \((C1)\), there exists \( t \in S \) such that the power allocation \( u + \tilde{\delta}_{(t)} \) is feasible and the resulting regret variation fulfills \( R(u + \tilde{\delta}_{(t)}) - R(u) < 0 \), which implies conclusion 1 of the Lemma.
(ii) $|\partial L| > K - |L_{in}|$ and $|\overline{S}_{at}| \leq |\partial L| - (K - |L_{in}|)$:

For the feasible power swap $\delta > 0$, it holds

$$
\overline{S}_{in} \subseteq L^\delta_{in}, \quad \overline{S}_{at} \subseteq L^\delta_{at}, \quad \overline{S}_{out} \subseteq L^\delta_{out},
$$

hence, the total regret variation is calculated as

$$
R(u + \delta) - R(u) = \begin{cases} 
- \Pi(t)\tilde{\delta}(t) + \sum_{t \in \overline{S}_{in}} \Pi(t)\tilde{\delta}(t) + \sum_{t \in \overline{S}_{at} \cup \overline{S}_{out}} \overline{\Pi}(t)\tilde{\delta}(t) & \text{if } t \in L_{in} \cup \partial L \\
- \overline{\Pi}(t)\tilde{\delta}(t) + \sum_{t \in \overline{S}_{in}} \Pi(t)\tilde{\delta}(t) + \sum_{t \in \overline{S}_{at} \cup \overline{S}_{out}} \overline{\Pi}(t)\tilde{\delta}(t) & \text{if } t \in L_{out}
\end{cases}
$$

where the change of power allocation is defined as

$$
\tilde{\delta}_{(t)} = -\tilde{\delta}(t)\overrightarrow{e}_i + \tilde{\delta}(t)\overrightarrow{e}_i, \forall t \in \overline{S}.
$$

Note that $\tilde{\delta}_{(t)}$ transfers energy from $t$ into $t$ according to the same relative powers as $\tilde{\delta}$, which entails the same relative variations of $Q$ and accordingly the same prices in the expression of the regret variation.

Then, by Equation (C2), there exists $t \in \overline{S}$ such that the power allocation $u + \tilde{\delta}_{(t)}$ is feasible and the resulting regret variation fulfils $R(u + \tilde{\delta}_{(t)}) - R(u) < 0$, which implies conclusion 1 of the Lemma.

(iii) $|\partial L| > K - |L_{in}|$ and $|\overline{S}_{at}| > |\partial L| - (K - |L_{in}|)$

For the feasible power swap $\delta > 0$, it holds

$$
\overline{S}_{in} \subseteq L^\delta_{in}, \quad \overline{S}_{at} \supseteq L^\delta_{at}, \quad \overline{S}_{out} \subseteq L^\delta_{out},
$$

in which case, we need to partition $\overline{S}_{at} = \overline{S}_{at}^+ \cup \overline{S}_{at}^*$ such that $|\overline{S}_{at}^+| = |\overline{S}_{at}^+| - ((\partial L) - (K - |L_{in}|))$, $|\overline{S}_{at}^*| = |\partial L| - (K - |L_{in}|)$, $|\overline{S}_{at}^+| \subseteq L^\delta_{in} \cup L^\delta_{at}$ and $|\overline{S}_{at}^+| \subseteq L^\delta_{at}$.

Then, the total regret variation is calculated as

$$
R(u + \tilde{\delta}) - R(u) = \begin{cases} 
- \Pi(t)\tilde{\delta}(t) + \sum_{t \in \overline{S}_{at} \cup \overline{S}_{out}} \Pi(t)\tilde{\delta}(t) + \sum_{t \in \overline{S}_{in}} \overline{\Pi}(t)\tilde{\delta}(t) & \text{if } t \in L_{in} \cup \partial L \\
- \overline{\Pi}(t)\tilde{\delta}(t) + \sum_{t \in \overline{S}_{at} \cup \overline{S}_{out}} \Pi(t)\tilde{\delta}(t) + \sum_{t \in \overline{S}_{in}} \overline{\Pi}(t)\tilde{\delta}(t) & \text{if } t \in L_{out}
\end{cases}
$$

Notice that for all $t \in \overline{S}_{in} \cup \overline{S}_{out}$, we can define the $1 \rightarrow 1$ power swap $\tilde{\delta}_{(i)} = -\tilde{\delta}(t)\overrightarrow{e}_i + \tilde{\delta}(t)\overrightarrow{e}_i$, such that

$$
R(u + \tilde{\delta}_{(i)}) - R(u) = \begin{cases} 
- \Pi(t)\tilde{\delta}(t) + \Pi(t)\tilde{\delta}(t) & \text{if } t \in L_{in} \cup \partial L, t \in \overline{S}_{in}, \\
- \Pi(t)\tilde{\delta}(t) + \overline{\Pi}(t)\tilde{\delta}(t) & \text{if } t \in L_{in} \cup \partial L, t \in \overline{S}_{out}, \\
- \overline{\Pi}(t)\tilde{\delta}(t) + \Pi(t)\tilde{\delta}(t) & \text{if } t \in L_{out}, t \in \overline{S}_{in}, \\
- \Pi(t)\tilde{\delta}(t) + \overline{\Pi}(t)\tilde{\delta}(t) & \text{if } t \in L_{out}, t \in \overline{S}_{out},
\end{cases}
$$

in which the same prices are taken as in Equation (C3). However, this is not true for $t \in \overline{S}_{at}$ as the prices at $t \in \overline{S}_{at}$ takes their lower values, whereas those $t \in \overline{S}_{at}^*$ takes upper values. In other words, the resulting regret $R(u - \tilde{\delta}(t)\overrightarrow{e}_i + \tilde{\delta}(t)\overrightarrow{e}_i)$ might not be a convex combination of $R(u)$ and $R(u + \tilde{\delta})$. The reason is that the lower prices should always be taken for all $t \in \overline{S}_{at}$, in a $1 \rightarrow 1$ power swap operation, which contradicts with the convenient $1 \rightarrow m$ power swap.

Thus, in the following, we consider the situation where it is not convenient to have $1 \rightarrow 1$ swaps from $t$ to any $t \in \overline{S}_{at}^*$ that $|\overline{S}_{at}^*| > |\partial L| - (K - |L_{in}|)$. In this case, we define some set partition $\overline{S}^* = \overline{S}^{**} \cup \overline{S}^{***}$ $\subseteq \overline{S}_{at}$ satisfying $\overline{S}^{**} \cap \overline{S}^{***} = \emptyset$ and $|\overline{S}^{**}| = |\partial L| - (K - |L_{in}|)$. Moreover, while feeding energy from $t$ to $\overline{S}^*$ according to the same
relative power as $\delta$, it entails the same variation on the $Q$ values for all $t \in \overline{S}^l$ which ensures $\overline{S}^l = \partial L^\delta$ and further preserves the prices in the expression of the regret variation.

Thus, the regret reduction is also written as

$$R(u + \delta) - R(u) = \left( R(u + \hat{\delta}_{(\overline{S}^l)}^t) - R(u) \right) + \sum_{t \in \overline{S}_{in} \cup \overline{S}_{out} \cup (\overline{S}_{in} \setminus \overline{S}^l)} (R(u + \hat{\delta}_{(\overline{S}^l)}^t) - R(u))$$

with the power swap policies

$$\hat{\delta}_{(\overline{S}^l)}^t = -\frac{\tilde{\delta}(t)}{\sum_{t \in \overline{S}^l} \tilde{\delta}(t)} \tilde{e}_t + \tilde{\delta}(t) \tilde{e}_t, \forall t \in \overline{S}_{in} \cup \overline{S}_{out} \cup (\overline{S}_{in} \setminus \overline{S}^l),$$

and

$$\hat{\delta}_{(\overline{S}^l)}^t = -\left( \sum_{t \in \overline{S}^l} \tilde{\delta}(t) \right) \tilde{e}_t + \sum_{t \in \overline{S}^l} \tilde{\delta}(t) \tilde{e}_t.$$ 

Thus, the power swap policies $\hat{\delta}_{(\overline{S}^l)}^t$ and $\hat{\delta}_{(\overline{S}^l)}^t$ are both feasible, and at least one of the regret variation $R(u + \hat{\delta}_{(\overline{S}^l)}^t) - R(u) < 0$ and $R(u + \hat{\delta}_{(\overline{S}^l)}^t) - R(u) < 0$ holds which corresponds to conclusion 1 and 2, respectively, in the Lemma.

Therefore, if the $1 \rightarrow m$ swap is convenient, at least one of the two categories of power swaps decreases the total regret.

**APPENDIX D. PROOF OF LEMMA 4**

The proof follows the similar lines to the proof of Lemma 3, but with the reverse power swapping direction.