A Novel Lagrangian Multiplier Update Algorithm for Short-Term Hydro-Thermal Coordination [†]

P. M. R. Bento [1], **S. J. P. S. Mariano [1,*]**, **M. R. A. Calado [1]** and **L. A. F. M. Ferreira [2]**

[1] IT—Instituto de Telecomunicações, University of Beira Interior, 6201-001 Covilhã, Portugal; pedro.bento@lx.it.pt (P.M.R.B.); rc@ubi.pt (M.R.A.C.)

[2] Instituto Superior Técnico and INESC-ID, University of Lisbon, 1049-001 Lisbon, Portugal; lmf@ist.utl.pt

[*] Correspondence: sm@ubi.pt; Tel.: +351-(12)-7532-9760

[†] This paper is an extended version of our paper published in ICEUBI2019—International Congress on Engineering—Engineering for Evolution, Covilhã, Portugal, 27–29 November 2019; pp. 728–742.

Abstract: The backbone of a conventional electrical power generation system relies on hydro-thermal coordination. Due to its intrinsic complex, large-scale and constrained nature, the feasibility of a direct approach is reduced. With this limitation in mind, decomposition methods, particularly Lagrangian relaxation, constitutes a consolidated choice to "simplify" the problem. Thus, translating a relaxed problem approach indirectly leads to solutions of the primal problem. In turn, the dual problem is solved iteratively, and Lagrange multipliers are updated between each iteration using subgradient methods. However, this class of methods presents a set of sensitive aspects that often require time-consuming tuning tasks or to rely on the dispatchers' own expertise and experience. Hence, to tackle these shortcomings, a novel Lagrangian multiplier update adaptative algorithm is proposed, with the aim of automatically adjust the step-size used to update Lagrange multipliers, therefore avoiding the need to pre-select a set of parameters. A results comparison is made against two traditionally employed step-size update heuristics, using a real hydrothermal scenario derived from the Portuguese power system. The proposed adaptive algorithm managed to obtain improved performances in terms of the dual problem, thereby reducing the duality gap with the optimal primal problem.

Keywords: hydro-thermal coordination; Lagrangian relaxation; Lagrangian dual problem; Lagrange multipliers; subgradient methods; step-size update algorithm

1. Introduction

The objective of short-term hydro-thermal coordination is to optimize electricity generation [1], meaning to find an optimal generation dispatch, or close to ideal, for all the thermal and hydro units available in a system. This ensures the total operation cost is minimized within horizons ranging from one day to one week (168 h), taking into account the entire system and its individual constraints [2–5] and with a planning period (discrete time-step), typically set from hour to hour [5]. In other words, this crucial process is responsible for scheduling the start-up and shutdown of thermal units (binary level decisions), in coordination with hydro plants, to ensure the continuity of electricity supply with appropriate levels of spinning reserve, while minimizing the operating costs [6]. This scheduling constitutes a unit commitment (UC) problem, where the dispatch policy of the thermal units is made in such a way that the total cost (operating cost, starting cost and shut down cost) is minimal over a pre-defined time-horizon. In addition, a series of operational constraints needs to be fulfilled, thus reducing the freedom of choice to turn a thermal unit on or off. In this regard, we are primarily speaking about the load balance constraint, i.e., ensuring that our electric energy demand

is satisfied, yet further constraints include spinning reserve constraints, minimum connected time, minimum time off, generation capacity limits, group restrictions, water restrictions, etc. [7].

Therefore, we can understand why short-term hydro-thermal coordination, in a framework where hydro and thermal power plants are the backbone of conventional power systems (consolidated power generation technologies), is such an important subject for power producers. Moreover, is one of the most complex problems to solve in power system engineering [1,8] due to its inherent large-scale and non-linear and combinatorial nature. This explains why, over the last decades, it has been the subject of intense research in the fields of operation and optimization of electric power systems [1,2,9,10].

Thus, a broad spectrum of methods has tried to solve short-term hydro-thermal coordination, and they can be generally divided into two categories: mathematical methods and deterministic methods [5,7]. In the realm of conventional approaches we can highlight Benders Decomposition [11], Lagrangian Relaxation [1,2,12], Improved LR [13], Dynamic Programming (PD), Nonlinear Programming [14], Dynamic Non-Linear Programming [15], Augmented Lagrangian [16], mixed integer linear programming [17–19], logarithmic size mixed-integer linear programming [20] and nonlinear programming [21], among several others. However, even with this wide array of classical optimization methods a perfectly tailored solution is hard to find, and in general terms the complexity of short-term hydro-thermal coordination has a negative effect on the computing efficiency [19]. Besides, problems of different nature arise with the use of classical approaches, which may impact the performance, mainly scheduling problems [22], slower convergence, premature convergence, computational cost, problems to deal with the nonlinearity and non-convexity, the need to perform problem simplifications, etc. In addition, classical deterministic methods typically rely upon single path search methods, which may help in terms of convergence speed but can be tricky in the presence of non-smooth surfaces [23].

A consolidated trend has been the growing application of evolutionary/heuristic methods and methods of artificial intelligence, in addition to new deterministic heuristics. Hence, we can mention neural networks [24], Cuckoo Search [25], Differential Evolution [26,27], Grey Wolf Optimizer [28], Improved Bacterial Foraging Algorithm [29] and a hybrid approach combining Artificial Bee Colony and the Bat algorithm [22], among others. For example, in [30] by using two different case studies, with and without pumping-storage capability and considering different constraints, the authors tested the effectiveness of different Accelerated Particle Swarm Optimization variants. In another instance, an Improved harmony search algorithm was employed on a non-linear, non-convex, short-term hydrothermal scheduling [31], among many others. However, in general these machine learning and population-based methods require a significant computational effort to solve the problem for an hourly discretized weeklong time-horizon, i.e., for large-scale problems (with a high number of dimensions) its effectiveness drops significantly. In addition, they can frequently end up finding only suboptimal solutions [4,21]. Besides, these metaheuristic methods often rely on a population search to find an optimal solution, turning them into large-scale problems (many dimensions and numbers of search agents), and occasionally several runs are required to find an optimal solution, as premature stagnation or slow convergence may occur.

Due to the presence of multiple sets of constraints, decomposition techniques appear as a natural option to solve this problem [1,10]. Consequently, Lagrangian Relaxation (LR) is one of these preferred decomposition techniques to tackle the short-term hydro-thermal coordination problem [6,12,32,33]. The fundamentals behind LR are to use Lagrange multipliers to relax system constraints such as load demand and reserve requirements. The primal problem is then converted into a two-level structure (subproblems). Given a set of multipliers, all subproblems are resolved at the low level, one for each unit, and the multipliers are updated at the high level. Multipliers are obtained by solving the dual problem, and the feasibility of solving the primal problem is usually obtained based on the dual solution [2,3], i.e., the primal problem is a byproduct of the dual problem solution. Finally, the solution is translated in dispatch generation decisions to meet the demand.

To update Lagrange multipliers, a common approach is to apply subgradient methods, where the step-size update procedure represents a sensible decision and often depends on ad-hoc testing.

This easy-to-implement approach, however, has some limitations, particularly the computational inefficiency and tendency for oscillating solutions. Two other common alternatives are the Bundle method, which in essence is another subgradient method based on a bundle of subgradients from previous iterations, and the Cutting Plane method, which adds new constraints to reduce the size of the feasible region. For example, in [34,35] evolutionary programming with a Gaussian mutation is used to update the multipliers, and in both, parameters are chosen considering convergence criteria. Following the vast existing literature on this subject, and to avoid sensible user dependence concerning the parameter choice, an adaptive algorithm is proposed in this work for an enhanced use of the LR technique for short-term hydro-thermal coordination. Additionally, it is important to refer that this paper is an extended version of work published in [36]; we expand on the formulation and the framework of the short-term hydro-thermal coordination problem, add a second case study, refine the results analysis and improve the figures.

Hence, we can summarize the main contributions of this work as follows:

1. Identify a weakness in the classical update mechanism of the step-size used in the subgradient method.
2. Propose an adaptive Lagrangian multiplier update algorithm that, as its name suggests, dynamically updates the step-size value and subsequently the Lagrange multipliers, so that the dual function converges towards its optimum in a pre-arranged number of iterations.
3. The LR technique is then used with the proposed adaptive update algorithm to solve a real large-scale, short-term hydro-thermal coordination problem, using data from the largest Portuguese electric utility company in two different scenarios.
4. For validation purposes, the algorithm is tested against two traditional step-size update heuristics with different initial parameter values.
5. Finally, the results comparison in both case studies revealed a sizeable advantage in terms of dual problem solution (error reduction) in favor of the proposed adaptive algorithm, therefore reducing the duality gap with the primal problem. Thus, a better allocation of the complex and vast hydro and thermal resources is obtained.

The rest of this paper is organized as follows. Section 2 provides a brief description of the primal problem. Section 3 explains the Lagrangian dual problem. Section 4 introduces subgradient methods and the motivation for the proposed algorithm, introduced in Section 5. Results and discussion for the two case studies are provided in Section 6, and finally, Section 7 presents the conclusion of this work.

2. The Primal Problem

The hydro-thermal coordination problem is a non-linear, large-scale, non-convex and combinatorial problem by nature. It can be understood as the task of establishing a map of feasible operations for each generation unit available in an electrical power system at the lowest cost for a predefined time horizon, in order to satisfy the expected load demand and a set of other system restraints. Typically, the time horizon considered is from one to seven days, and the discrete time-step (in which decisions are made) is a one-hour period. This problem is treated as deterministic, and whenever it is necessary to include stochastic quantities such as load diagram and reservoir inflows, their expected values are used.

In this manner, a primal problem ($\mathcal{P}$) is non-convex and non-linear and can be mathematically formulated as shown in Equations (1)–(6). The total operating cost for all resources (units) and over the entire considered period, K, is defined in Equation (1) and is the problem's objective function, i.e., evaluates the performance of each admissible solution. The cost function, $C_{ik}(x_{i,k-1}, p_{ik}, u_{ik})$, is a measure that evaluates the decision made in each state, since there is an operating cost associated with the state transition (from $x_{i,k-1}$ to x_{ik}), which delivers the power p_{ik}, determined by the control decision u_{ik}, for each unit i at time k. The following three Equations (2)–(4) represent the set of global constraints. Firstly, Equation (2) translates the (global) demand–supply balance restraint, where D_k

is the required load demand that needs to be served by the power output of each resource i in hour k, p_{ik}. Moreover, for simplicity purposes transmission losses were not considered. In turn, Inequality (3) represents all the hourly system capacity requirement constraints, i.e., constraints like the spinning and operating reserve requirements. R_{mi} translates the capacity contribution function associated with resource i for the system capacity requirement of type m, while R_{mk}^{req} is the mth-type system capacity requirement in hour k [5].

Additionally, Inequality (4) represents all the cumulative constraints, such as the limitation on emission by a group of units over the scheduling time horizon or the amount of consumed fuel, where $\mathcal{H}_n$ stands for the set of thermal units on the nth cumulative constraint, H_{ni} is the function which describes a contribution of thermal unit i–nth cumulative constraint, H_n^{req} is the upper bound on nth cumulative constraint and N is the set of cumulative constraints [37].

In turn, Equations (5) and (6) represent the set of local constraints, the state Equation (5) of each resource i at a time k. This equation allows us to obtain the state of each resource x_{ik} and its contribution p_{ik} to satisfy demand, whatever the decision u_{ik}. Last of all, in (6) the domain of admissible values for the control variables, as well as for the initial and final state, are defined for each individual resource i.

$$\mathcal{P} \quad \underset{u}{Min} \quad \sum_{k=1}^{K}\sum_{i=1}^{I} C_{ik}\left(x_{i,k-1}, p_{ik}, u_{ik}\right) \tag{1}$$

Subject to

$$\sum_{i=1}^{I} p_{ik} = D_k \quad i = 1, \ldots, I \ \wedge \ k = 1, \ldots, K \tag{2}$$

$$\sum_{i=1}^{I} R_{mi}(x_{ik}, p_{ik}) \geq R_{mk}^{req} \quad m = 1, \ldots, M \tag{3}$$

$$\sum_{k=1}^{K}\sum_{i \in \mathcal{H}_n} H_{ni}(x_{ik}, p_{ik}, u_{ik}) \geq H_n^{req} \quad n = 1, \ldots, N \tag{4}$$

and wherein

$$(x_{ik}, p_{ik}) = A_{ik}\left(x_{i,k-1}, u_{ik}\right) \quad i = 1, \ldots, I \ \wedge \ k = 1, \ldots, K \tag{5}$$

$$\begin{aligned} u_{ik} \in \mathcal{U}_{ik} \ \ x_{i0} \in X_i^0 \ \ x_{ik} \in X_i^K \\ i = 1, \ldots, I \ \ \wedge \ \ k = 1, \ldots, K \end{aligned} \tag{6}$$

Although the objective function is a separable function in resources and hours, this problem, by its formulation and due to collective constraints, does not allow this separability, providing extreme complexity to the minimization problem. In other words, the optimum value cannot be found by the sum of the various suboptimal (separately) results from each resource. Thus, we are facing a problem of unrestrainable dimension, for which a direct approach is not viable.

The primal problem defined in this study approaches the short-term hydro-thermal coordination considering the generation resources available to the electric utilities company, to match the system-wide load demand over a weekly time-period, while fulfilling a set of other global and local constraints.

3. Lagrangian Dual Problem

As discussed, the primal problem is difficult to solve using conventional nonlinear optimization techniques. A preferable path is to decompose the problem constraints and transfer them to the objective function, i.e., to solve the dual problem, rather than solving the primal problem directly. We know beforehand that the optimal solution of the relaxed problem is a lower bound (good estimate) of the optimal solution of the initial problem [2,10,38].

This is achieved by relaxing the constraints, i.e., weakening of the problem ($\mathcal{P}$), that in practical terms means open the possibility to breach the imposed constraints. However, relaxed restrictions are not completely ignored since its violations are linearly penalized in the Lagrange function (an added cost associated with violating each constraint).

We can write the Lagrange function ($\mathcal{L}$) for problem ($\mathcal{P}$) by relaxing its global constraint, as expressed in Equation (7), where λ, μ and γ are the Lagrange multiplier vectors associated with the load-balance constraint, capacity constraints and cumulative constraints, respectively. These Lagrange multipliers are expressed in units of cost per unit of the parameters of their associated constraint, which in the case of Equation (2) is given in \$/GW.

$$
\begin{aligned}
\mathcal{L}\left(x_{i,k-1}, p_{ik}, u_{ik}, \lambda, \mu, \gamma\right) = &\sum_{k=1}^{K}\sum_{i=1}^{I} C_{ik}\left(x_{i,k-1}, p_{ik}, u_{ik}\right) \\
&+ \sum_{k=1}^{K} \lambda_k\left(D_k - \sum_{i=u}^{I} p_{ik}\right) \\
&+ \sum_{m=1}^{M}\sum_{k=1}^{K} \mu_{mk}\left(R_{mk}^{req} - \sum_{i=1}^{I} R_{mi}(x_{ik}, p_{ik})\right) \\
&+ \sum_{n=1}^{N} \gamma_n\left(H_n^{req} - \sum_{k=1}^{K}\sum_{i \in \mathcal{H}_n} H_{ni}(x_{ik}, p_{ik}, u_{ik})\right)
\end{aligned}
\tag{7}
$$

That is, to now solve the unit commitment problem, $\mathcal{L}$ is minimized, where $\underset{u}{Min}\ \mathcal{L}\left(x_{i,k-1}, p_{ik}, u_{ik}, \lambda, \mu, \gamma\right)$ is subject to local system constraints, i.e., Equations (5) and (6).

Subgradient of the Dual Function

The Lagrangian dual problem is obtained by forming ($\mathcal{L}$), and its solution provides the primal variables as functions of the Lagrange multipliers, which are labeled dual variables. Hence, the new problem is to maximize the objective function with respect to the multipliers under the derived constraints on the dual variables. This implies, by decomposition, that each resource becomes a single entity and is individually optimized, rather than a combined optimal resource allocation. Therefore, the dual function is defined in Equation (8), presenting concave and sub-differentiable traits (resulting in inferiorly limited variables).

$$
q(\lambda, \mu, \gamma) = \underset{u}{Min}\ \mathcal{L}\left(x_{i,k-1}, p_{ik}, u_{ik}, \lambda, \mu, \gamma\right)
\tag{8}
$$

Given that Lagrange's dual function is a concave function with simple bounds on the variables, a local optimum is also the function global optimum. Therefore, our task is to find the Lagrangian multipliers, λ, μ and γ, that maximize the dual function. Nonetheless, this does not mean that solving the dual function is a trivial task (far from it actually), since the function is not smooth and is not given by an easy-to-compute expression [5]. For this purpose, we resort to subgradient methods, which benefit from the fact the subgradients of q are easily derived system constraint deviations. Consequently, we can define the subgradient of the dual function g (9) for each hour k as follows:

$$g = \begin{bmatrix} \cdots \\ D_k - \sum\limits_{i=1}^{I} p_{ik} \\ \cdots \\ - - - - - - - - - - \\ \cdots \\ R_{mk}^{req} - \sum\limits_{i=1}^{I} R_{mi}(x_{ik}, p_{ik}) \\ \cdots \\ - - - - - - - - - - - \\ \cdots \\ H_n^{req} - \sum\limits_{k=1}^{K} \sum\limits_{i \in \mathcal{H}_n} H_{ni}(x_{ik}, p_{ik}, u_{ik}) \\ \cdots \end{bmatrix} \tag{9}$$

Moreover, by the weak duality theorem for a single set of multipliers, the optimal value of the Lagrange dual problem $q(\lambda^*)$ and the optimal value of the primal minimization problem $p(\lambda^*)$ are related by $q(\lambda^*) \leq p(\lambda^*)$, and the difference between the values is called a duality gap. This implies that the dual problem offers a good indirect root to solve the primal one, since the gap in most practical cases is relatively small [6,12,39].

For all the reasons above, this approach to the problem is advantageous since it lessens the computational burden of the primal problem.

4. Subgradient Methods

As we saw earlier, obtaining the Lagrange dual function optimal value goes hand-in-hand with the Lagrange multiplies choice/update method, i.e., at the outset, this choice determines how close we are to the solution of the dual problem and, ultimately, how close we are from reaching the primal problem best solution. To perform this task several methods are described in the literature [5]; however, regarding our problem in particular, subgradient methods prevail as the most fitting solution by achieving higher accuracies. Further benefits include their simplicity as well as the computational ease with which the Lagrange dual function subgradient (solution deviation from the imposed constraints) is calculated.

These methods update the multipliers according to the subgradient direction and in a manner proportional to the violation of the corresponding constraints. Besides, a distinctive trademark of these methods concerns the step-size update heuristic, where again several approaches have been followed [5]. However, the downside of these conventional updating heuristics is that a long-winded trial-and-error procedure as well as a highly specialized operator are frequently required. The simplest and most common subgradient method formulation is given by

$$\lambda^{v+1} = \left[\lambda^v + s^v \frac{g^v}{\|g^v\|} \right]^+ \tag{10}$$

where g^v is the subgradient $g(p_{\lambda^v})$, s^v is a positive scalar that defines the step-size at the current iteration v and, lastly, $[.]^+$ represents the projection in the set of feasible values Λ. Nonetheless, there is no guarantee that after iteration $v + 1$, independently from the chosen step-size, the dual function value will actually improve (walk towards the optimal dual function value), meaning that in some occasions we will have

$$q\left(\left[\lambda^v + s^v \frac{g^v}{\|g^v\|} \right]^+ \right) < q(\lambda^v), \ \forall \quad s > 0 \tag{11}$$

Though, if the step value is sufficiently small, the distance between the obtained point in the current iteration and the optimum value can always be reduced. The following proposition offers an estimate for the step-size domain (range):

Proposition 1. *If λ^v does not lead to the optimum value of the dual function, then for λ^*, which corresponds to the dual function optimum value, the inequality $\|\lambda^{v+1} - \lambda^*\| < \|\lambda^v - \lambda^*\|$ is valid for all step-sizes, $s^v \in \left]0, \frac{2(q(\lambda^*)-q(\lambda^v))}{\|g^v\|}\right[$. Therefore, this suggests a step-size equal to the middle value of the inequality range, i.e., $s^v = \frac{(q(\lambda^*)-q(\lambda^v))}{\|g^v\|}$.*

Since this requires knowledge of the dual function optimal value $q(\lambda^*)$, which is exactly the unknown we want to find, this approach is unviable in our case, and we resort to heuristics that determine the step-size. In this regard, a popular choice is decreasing step-size rule-based approaches, mainly due to its simplicity and effectiveness.

Considering a decrease in step-size, s^v, towards zero, meaning that $\lim_{v \to \infty} s^v = 0 \,\wedge\, s^v > 0$, while at the same time satisfying $\sum_{v=1}^{\infty} s^v = \infty$, the method can "travel" as far as possible (up to infinity) in order to converge to the optimal dual function value. Thus, under these assumptions, we can infer a 2nd proposition, from which we can conclude that it is possible, by appropriately updating the step-size, to reach the dual function maximum value [21].

Proposition 2. *For the sequence of all multiplier values $\{\lambda^v\}$ we have $\lim_{v \to \infty} Max\, q(\lambda^v) = q^*$. However, this analysis does not lead to a finite procedure, pointing to an initial value of the step, as well as a mechanism for decrementing it to zero. As such, for comparison purposes against the proposed new heuristic, the two most widely employed expressions are introduced in Equations (12) and (13), to update the step-size at each iteration v.*

$$s^v = \frac{a_1}{1 + v \times a_2} \tag{12}$$

$$s^v = \frac{a_1}{1 + v^{a_2}} \tag{13}$$

where a_1 and a_2 are control parameters of the heuristic process. Moreover, the chosen initial step is a highly sensitive matter, since small initial steps can prevent the method from reaching the desired optimum value in a reasonable number of iterations. Whereas, using a large initial step may cause the method to oscillate erratically in the early phase, leading to poor convergence. As a result, although the obtained value is stabilized, it could still be improved by running further iterations. This fact is more pronounced in Equation (12) rather than Equation (13), given that $a_2 > 1$ implies a rapid decrease in the step-size. Consequently, selecting the values to assign to parameters a_1 and a_2 is a difficult task, with direct influence on the obtained results. This could be facilitated if a good initial estimate for the dual variable vector λ^0 is available, that is, if $q(\lambda^0)$ is already close to the solution of the dual problem.

Therefore, we can conclude that it is an intrinsically lengthy (experimentation-based) heuristic process that is highly dependent on the user's experience. Precisely to mitigate this scenario, a new algorithm will be proposed next.

5. Proposed Adaptative Algorithm

When applying subgradient methods, the existence of good estimates for the multipliers and careful tune-up of the subgradient step-size are considered essential to improve the computational effectiveness of the method. Subsequently, motivated by the previously exposed shortcomings from the classical subgradient optimization approaches, an adaptative heuristic is proposed in order to automatically update the Lagrange multipliers, thereby removing the need to rely on a user's past experiences or time-consuming trial-and-error tasks. This means that the step-size, s^v, is automatically determined (avoiding lengthy trial-and-error procedures) by the adaptative algorithm when solving

the dual problem with a subgradient method. The different stages of the algorithm that lead to a dual problem solution are illustrated in Figure 1; subsequently, the rationale behind them is detailed below:

(1)   define the initial step-size, $s^0 = 1$, and choose an initial value for the dual variable vector, λ_0; then compute the initial dual function and subgradient values, $q^0(\lambda^0)$ and $g^0(p_{\lambda^0})$, respectively;

(2)   update Lagrange multipliers according to Equation (10);

(3)   determine the new step-size as follows:

if $q^v(\lambda^v) > q^{v-1}(\lambda^{v-1})$ then

$$\alpha \in v_\delta^+(1)$$

else

$$\alpha \in v_\delta^-(1)$$

end, where,

$$v_\delta^+(1) = \{\alpha_1 : 1 < \alpha < 1 + \delta\}$$
$$v_\delta^-(1) = \{\alpha_2 : 1 - \delta < \alpha < 1\}$$

and the step-size is given by

$$s^v = \alpha s^{v-1}$$

(4)   compute the current (iteration) dual function and subgradient values, $q^v(\lambda^v)$ and $g^v(p_{\lambda^v})$;

(5)   if the termination criterion is met:

a.   terminate the algorithm;

else

b.   proceed to the next iteration, $v = v + 1$;

c.   return to (2);

end

Regarding the (above) adaptive algorithm, the following clarifications are made:

In (1) the initial dual variable vector positioning only impacts the convergence speed of the subgradient method; thereby, it can be considered arbitrary. On the contrary, for the step-size update expressed by Equations (12) and (13), this initial positioning benefits heavily from a nearby optimum value (derived from past experiences or other heuristics) in order to guarantee the method's performance, thereby translating an important advantage of the proposed strategy.

In turn, stage (3) depicts the original step-size update mechanism, where the rationale behind it is to dynamically update the step based on the dual function value, i.e., if this value improves then the step should be augmented; in contrast, if this value does not improve then the step should be diminished. Moreover, to prevent a large step-size increasing the distance between the new point and the optimum value, this step-size should be increased smoothly; this fact is less sensitive when reducing step-size. Additionally, it was found that the ideal domains for variables α_1 and α_2 are $[1.01, \, 1.05] \Rightarrow \delta = [0.01, \, 0.05]$ and $[0.83, \, 0.95] \Rightarrow \delta = [0.05, \, 0.17]$, respectively.

Lastly, the stop criterion mentioned in (4) is traditionally run a specific number of iterations, which was also the case in this work.
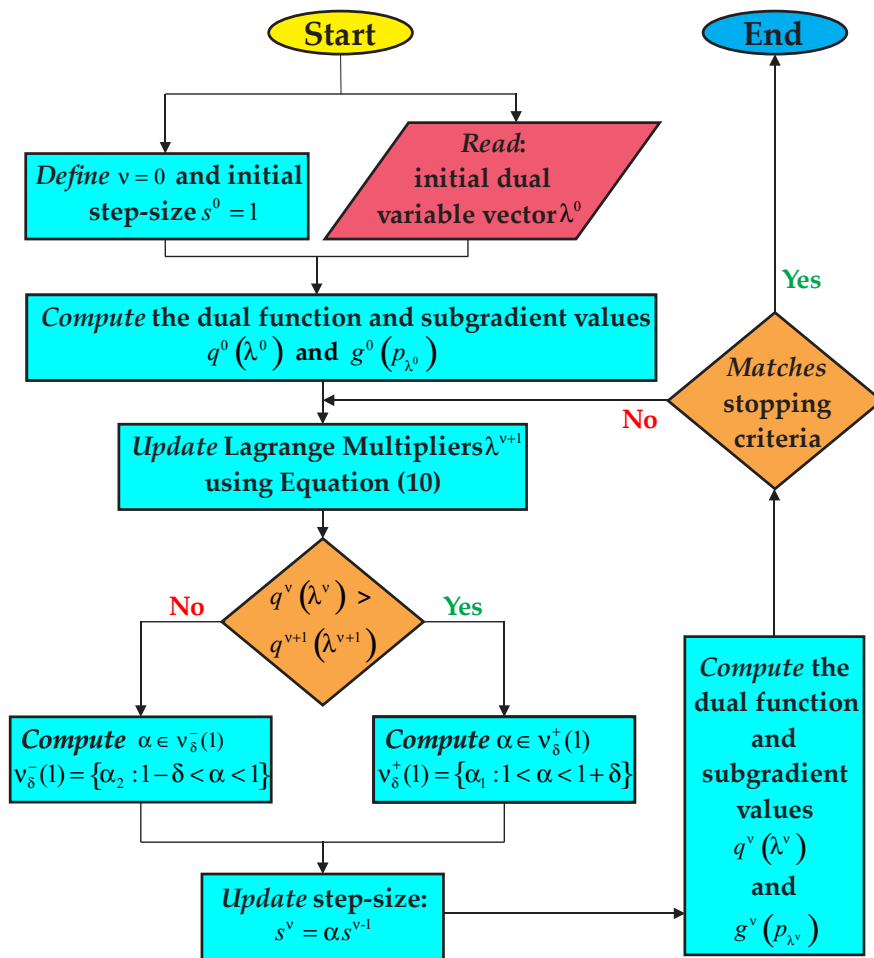
Figure 1. Flowchart of the proposed adaptative algorithm.

6. Numerical Results

The behavior of the subgradient method is analyzed in this section. The step value is updated according to the adaptive algorithm, proposed in Section 5, and then benchmarked against a classical approach. The step-size is updated using Equations (12) and (13) and, consequently, the Lagrange multipliers. The software used in this work was written in Fortran using the development environment (IDE) Microsoft Visual Studio ®.

As previously mentioned, the unit commitment (primal problem) corresponding to the solution of the Lagrange dual problem does not always lead to a feasible solution. As such, the average subgradient norm, $\|g(p_\lambda)\|/K$, is defined as a quantitative metric of how a solution is accurate in terms of the primal problem. This means the lower the value, the closer we will be to a good solution, and typically a value on the order of 0.5% of the peak load typically means that a good solution to the primal problem was found.

The data employed in this work concern the real short-term hydro-thermal coordination problem that the main Portuguese electric utility companies face. Data include all generation parameters and auxiliary variables, i.e., a large-scale study comprising six thermal power plants and 26 hydro power plants (amounting to over 80 individual generation units), which serve the majority of the Portuguese electric power demand. Two different case studies will be considered, diverging over the selected weekly periods, size and characteristics of the system, as well as the economic strategies behind the cost curves. Additionally, the specified parameters will be kept fixed for both case studies.

6.1. Case Study I

For this case study, unit costs are the sole result from the associated generation costs, with no other parallel costs. Consequently, both the evolution of the dual function $q(\lambda)$ value as well the evolution of the average subgradient norm $\|g(p_\lambda)\|/K$ will be evaluated over the course of iterations. Figures 2a,b and 3 show the evolution of the dual function value (left axis) and its step value (right axis). Figures 4a,b and 5 show the evolution of the average subgradient norm. Figures 2a and 4a illustrate the behavior of the subgradient method using a classical step-size update, given by Equation (12). In the same fashion, Figures 2b and 4b illustrate the behavior when using Equation (13). Lastly, the new adaptive algorithm results are shown in Figures 3 and 5.

Figure 2. Evolution of the dual function value (blue plots), $q(\lambda)$, and its step value (red plots), using the heuristics expressed by Equations (12) and (13) for (a) and (b), respectively. The following parameter values are imposed: (**a**) solid line, $a_1 = 20$, $a_2 = 2$; dashed line, $a_1 = 10$, $a_2 = 1.5$; (**b**) solid line, $a_1 = 20$, $a_2 = 1.5$; dashed line, $a_1 = 5.5$, $a_2 = 1.05$.
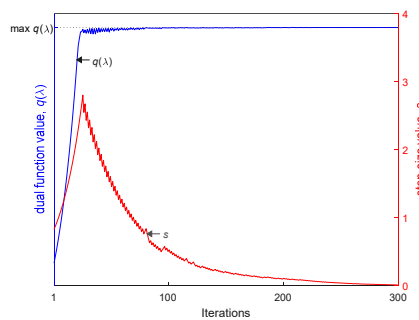

Figure 3. Evolution of the dual function value (blue plot), $q(\lambda)$, and its step value (red plot), using the adaptative algorithm, with the following parameter values: $\alpha_1 = 1.05$, $\alpha_2 = 1.10$.
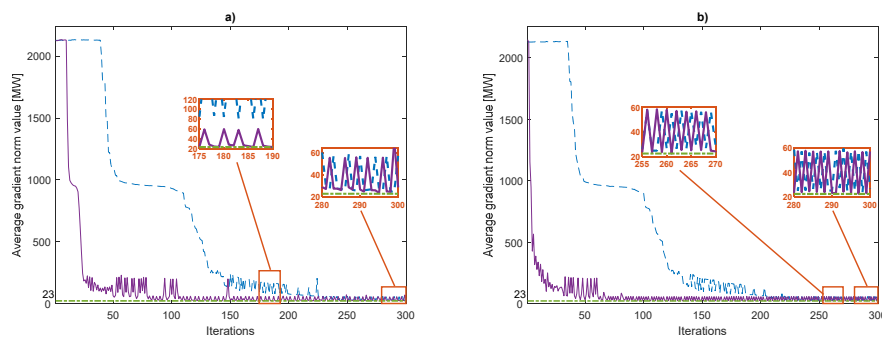

Figure 4. Evolution of the average subgradient norm, $\|g(p_\lambda)\|/K$, corresponding to the values of the dual function represented in Figure 2 (classical step-size equations). (**a**) purple solid line, $a_1 = 20$, $a_2 = 2$; blue dashed line, $a_1 = 10$, $a_2 = 1.5$; (**b**) purple solid line, $a_1 = 20$, $a_2 = 1.5$; blue dashed line, $a_1 = 5.5$, $a_2 = 1.05$. The green dot-dashed line highlights the achieved minimum gradient norm value.
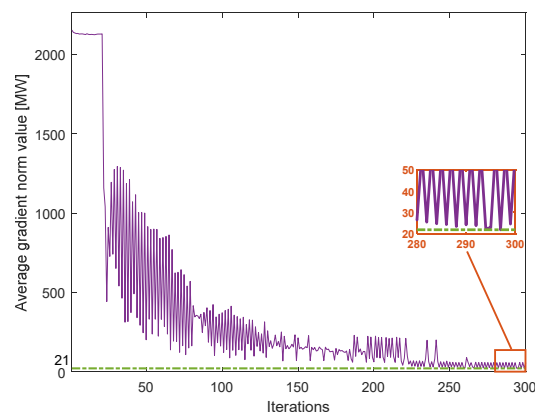
Figure 5. Evolution of the average subgradient norm (purple solid line), $\|g(p_\lambda)\|/K$, corresponding to the values of the dual function represented in Figure 3 (proposed adaptive algorithm). The green dot-dashed line highlights the achieved minimum gradient norm value.

Regarding Figure 2a,b we can observe the following: (i) achieving convergence in more or less iterations is heavily dependent on the choice of the different parameters; (ii) using a smaller initial step-size increased the number of iterations needed to achieve convergence, $\max q(\lambda)$ (dashed line); (iii) the use of a slightly larger initial step leads to some oscillation, still, without compromising convergence, represented by the solid lines; (iv) the step-size evolution is strictly decreasing, and the rate of descent depends on the considered parameters.

With respect to the adaptive algorithm, the evolution of the dual function increases as the value of the step increases, as shown in Figure 3, until a value is reached in the vicinity of the maximum dual function value. From this point onwards, the step value decreases towards zero, but then again it slightly increases whenever the dual function value does not improve compared to the previous iteration. This dynamically adjusted (based on the dual function current value) step-size clearly contrasts with the monotone evolution that occurs with the traditional step update formulation.

We can verify that, in all scenarios, the dual function maximum value was reached, and differences reside in the number of iterations necessary to achieve convergence (which was relatively similar). Nevertheless, the proposed algorithm shows a more robust approach when applying the subgradient method since it did not require educated guesses to converge on an acceptable number of iterations.

Concerning the obtained minimum average subgradient norm, Figures 4 and 5 are presented, where the secondary axis (magnified) provides a greater resolution regarding the convergence of each error curve to its recorded minimum value. Additionally, to summarize the respective minimum values achieved by the different step-size update mechanisms with different initial parameters, Table 1 is also presented.

Table 1. Minimum average subgradient norm, $\|g(p_\lambda)\|/K$, and its corresponding iteration, achieved by the different step-size update mechanisms for case study I.

Min $\frac{\|g(p_\lambda)\|}{K}$ [MW]	Iteration	Step-Size Update Mechanism
22.97	186	Equation (12) with $a_1 = 20$ and $a_2 = 2$
24.59	299	Equation (12) with $a_1 = 10$ and $a_2 = 1.5$
22.49	261	Equation (13) with $a_1 = 20$ and $a_2 = 1.5$
23.74	298	Equation (13) with $a_1 = 5.5$ and $a_2 = 1.05$
20.95	297	proposed adaptive algorithm

As we can see in Figure 4a,b, both processes led to similar final results, with a (best) value close to 23 MW by employing both classical step-size update equations. Nevertheless, the first combination in Table 1 ensured the fastest convergence (186 iterations). As for the proposed Lagrangian multiplier

update algorithm, we noticed an improved (smaller) error of ~21 MW, which in this relatively curtailed error scenario represents an improvement above 8%. These values represent approximately 0.53% of the peak load, which, as mentioned, usually leads to a good solution to the primal problem. Besides, note that once the dual function maximum value or its proximities are reached, the average subgradient norm has not yet reached its minimum value, and it continues to oscillate up and down over the next iterations. This can be explained since small variations in the multipliers can cause large variations in the solutions in terms of the primal problem. Thus, emphasizing the fact that even when reaching the maximum value of the dual function, we may not end up with the best solution in terms of the primal problem.

Moreover, this average subgradient norm oscillation is further accentuated in the adaptative algorithm since, in contrast to the previous two step-size update expressions, it does not have a strictly decreasing behavior. This behavior can lead to a lower convergence speed and, therefore, constitutes a limitation of the proposed adaptive algorithm. Notwithstanding, it is worth noting that a slow convergence is preferred over a premature convergence.

Lastly, regarding the solution in terms of the primal problem (Figure 6), corresponding to the solution of the dual problem for the (achieved) lowest average subgradient norm value. The same was obtained using the adaptive algorithm, since is easy to understand from previous figures that all primary solutions would be similar, so their presentation is redundant. The algorithm used in solving the primal problem based on Lagrangian relaxation, as we saw earlier, does not lead to an optimal solution. The obtained primal solution reveals the existence of Lagrangian duality. That is, we can say that good results were obtained since the generation profile (solid green line) was almost coincident with the desired load demand profile (dashed red line), but it did not match it completely. After solving the dual problem, several methods have been used to look for feasibility [5]. However, if we succeed when solving the dual problem, then we can also get, in terms of the primal problem, a good solution. In fact, in some cases it is enough to carry out an economic dispatch of thermal units to obtain a (close) feasible strategy, which is exactly what happens in the presented case study, where the difference between the maximum generation capacity (dotted orange line) and the allocated thermal unit generation (dashed magenta line) is sufficient to compensate for the mismatch (deviation) between the load profile and the obtained generation profile.
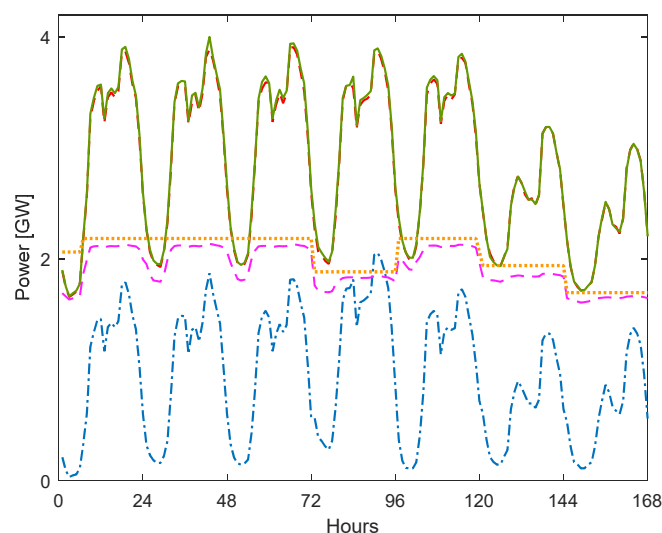

Figure 6. Solution in terms of the primal problem (case study I). In the upper portion the solid green and dashed red lines are almost coincident: the obtained generation profile and the load demand, respectively. Dotted orange line: maximum generation capacity of the affected thermal units. Dashed magenta line: thermal units generation profile. Dash-dot blue line: hydro units generation profile.

6.2. Case Study II

A second case study is considered with the purpose of further validating the proposed adaptive algorithm. For this scenario, the units' cost curves result is not the sole result of the generation costs but is also from additional economic strategies. Furthermore, the peak demand power (delivered to the grid) was 45% higher than in case study I. Thus, it constitutes a case with greater dimensions, with extra hydro and thermal plants considered. However, the parameters required for Equations (12) and (13) and the adaptive algorithm are the same as those specified in case study I, so we can compare the performance between these different methods of updating the step-size value.

In the same manner as in case study I, we started by analyzing the evolution of the dual function value, $q(\lambda)$, and the correspondent step-size values. From looking at Figure 7a,b we can see that the maximum value of the dual function, using both classical step-size update equations, had not been reached. Indeed, when compared to the value obtained using the adaptive algorithm (Figure 8), this value falls short by roughly 0.7% for the parameters illustrated by the solid line, whereas for the ones represented by dashed lines an evident lack of convergence can be spotted.
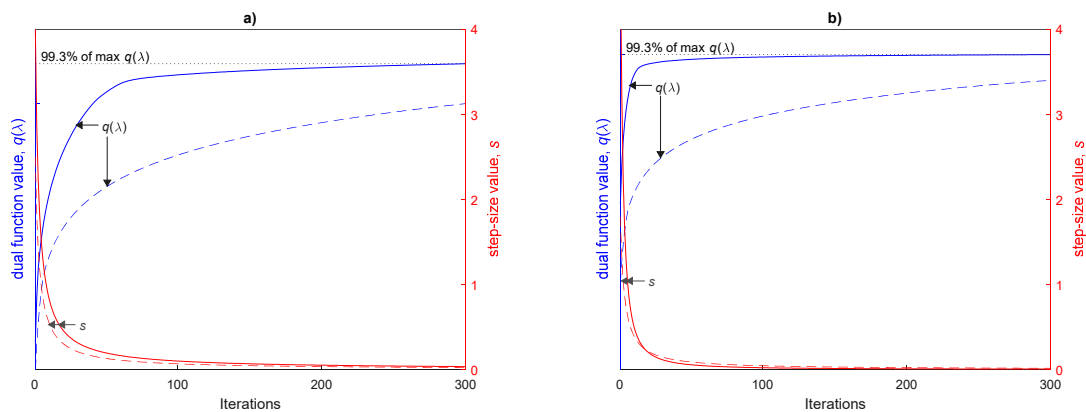

Figure 7. Evolution of the dual function value (blue plots), $q(\lambda)$, and its step value (red plots), using heuristic expressed by Equations (12) and (13) for (a) and (b), respectively. The following parameter values are imposed: (**a**) solid line, $a_1 = 20$, $a_2 = 2$; dashed line, $a_1 = 10$, $a_2 = 1.5$; (**b**) solid line, $a_1 = 20$, $a_2 = 1.5$; dashed line, $a_1 = 5.5$, $a_2 = 1.05$, respectively.
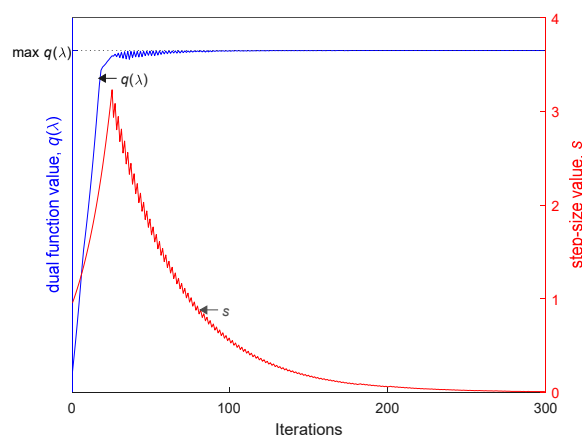

Figure 8. Evolution of the dual function value (blue plot), $q(\lambda)$, and its step value (red plot), using the adaptative algorithm, with the following parameter values: $\alpha_1 = 1.05$, $\alpha_2 = 1.10$.

In turn, the adaptative algorithm updates the step-size value dynamically and adapts to the current dual function value. That is, as mentioned above, if the dual function value improves then the step should be increased; on the contrary, if this value does not improve then the step should be decreased. Thus, in Figure 8 we can see that the step value increased until the dual function value

approached its maximum value; thereafter, the step value was modified accordingly, which enabled convergence with the max$q(\lambda)$ in both case studies, and ultimately led to good primal solutions.

These results anticipate a higher minimum value of the average subgradient norm for the classical update expressions, which inevitably compromises the solution in terms of the primal problem. This effect should be particularly pronounced for the parameters represented by the dashed plots in Figure 7a,b. Therefore, a clear contrast to the results from the first case is established, where the maximum value of the dual function is relatively similar for the different mechanisms used to update the Lagrange multipliers (as shown in Table 1).

As expected, this preliminary assessment is fully backed by examining the evolution of the average subgradient norm, as shown below in Figure 9a,b, as well as in the summary Table 2, which presents the respective minimum values achieved by the different step-size update mechanisms with different initial parameters, represented by each individual error plot in Figures 9 and 10. The results reveal sizeable minimum average subgradient norm values, even for the scenarios where the dual function closed in on its maximum value (~99.3% of the max$q(\lambda)$), presenting values of 318 MW and 317 MW, respectively. These results differ from the ones achieved using the adaptive algorithm (Figure 10), where a much improved $\|g(p_\lambda)\|/K$ minimum value was recorded, 38 MW (~8.3 times smaller), i.e., roughly representing only 0.64% of the peak power demand versus >5% with the classic heuristics.
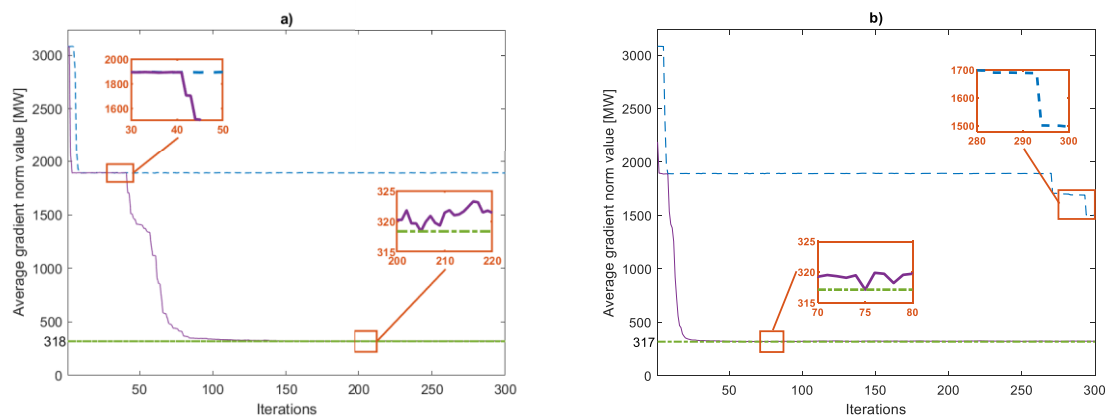

Figure 9. Evolution of the average subgradient norm, $\|g(p_\lambda)\|/K$, corresponding to the values of the dual function represented in Figure 7 (classical step-size equations). (**a**) Purple solid line, $a_1 = 20$, $a_2 = 2$; blue dashed line, $a_1 = 10$, $a_2 = 1.5$; (**b**) purple solid line, $a_1 = 20$, $a_2 = 2$; blue dashed line, $a_1 = 5.5$, $a_2 = 1.05$. The green dot-dashed line highlights the achieved minimum gradient norm value.

Table 2. Minimum average subgradient norm, $\|g(p_\lambda)\|/K$, and its corresponding iteration, achieved by the different step-size update mechanisms for case study II.

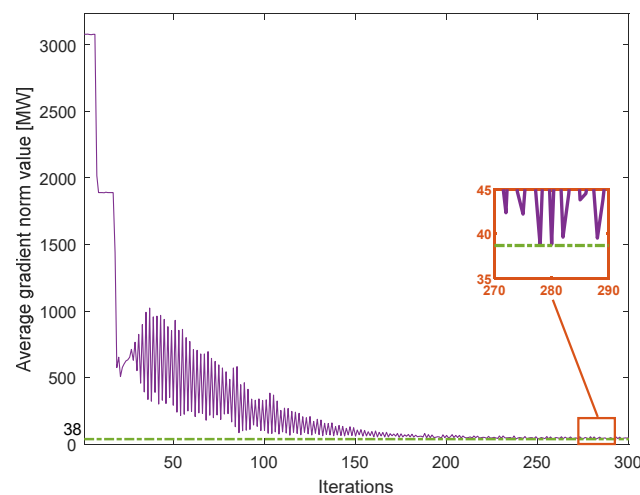| Min $\frac{\|g(p_\lambda)\|}{K}$ [MW]	Iteration	Step-Size Update Mechanism
318.33	205	Equation (12) with $a_1 = 20$ and $a_2 = 2$
1889.21	33	Equation (12) with $a_1 = 10$ and $a_2 = 1.5$
317.17	75	Equation (13) with $a_1 = 20$ and $a_2 = 1.5$
1499.30	299	Equation (13) with $a_1 = 5.5$ and $a_2 = 1.05$
38.37	278	proposed adaptative algorithm

Figure 10. Evolution of the average subgradient norm (purple solid line), $\|g(p_\lambda)\|/K$, corresponding to the values of the dual function represented in Figure 8 (proposed adaptive algorithm). The green dot-dashed line highlights the achieved minimum gradient norm value.

Furthermore, we can also notice that some of the parameters led to a fast but also premature convergence when using traditional heuristics, contrasting with the slower convergence exhibited by the proposed adaptive algorithm due to a more refined step-size update. Thereby, the inferences highlighted for case study I are confirmed. Despite the significant improvement by several orders of magnitude compared to the more modest improvement in case study I, this improved error value is still slightly above the one obtained in the previous case study, which translates to added difficulty when considering additional dispatch strategies. These characteristics will ultimately impact the task of obtaining a feasible solution in terms of the primal problem. Nevertheless, the proposed adaptive algorithm was able to converge to a good solution, and, in comparison, it will result in a smaller duality gap. Besides, it reinforces the mentioned need to thoroughly adjust the control parameters or rely upon educated guesses in order to achieve a good result when using classic heuristics versus the proposed automated algorithm.

Moving on towards the primal problem solution, Figure 11 is presented, revealing a higher peak demand as well as a larger usage of the thermal capacity in direct comparison to its counterpart in case study I, Figure 6. Further, it can be noted that the hydro generation follows the load profile on a smaller scale, and the hydro-thermal generation profile obtained moves further away from the load profile. Thus, the economic dispatch of thermal units is not sufficient for the solution, in terms of the primal problem, to be feasible. In other words, contrary to the first case study, the difference between the maximum generation capacity (dotted orange plot) and the allocated thermal units' generation (dashed magenta plot), illustrated in Figure 11, is insufficient to address the mismatch between the load profile and the obtained generation profile. Additionally, we can see that around 1–3 h and 147–149 h the hydro generation registered a negative value, which is explained by a sporadic period of power consumption (pumping) during an off-peak occurrence.

For this reason, it makes sense to present, for this case study, the mismatch between the load profile and the obtained generation profile as illustrated in Figure 12, where we can observe the non-feasibility after the economic dispatch for a few hours. This behavior was more significant during the last considered day (between 144 and 168 h), which is a weekend day, reaching a maximum above 65 MW at time $k = 149$ h. This translates a mismatch around 1% of the peak demand, and it is justified by the limited availability of thermal generation (orange and magenta lines in Figure 11 almost overlapping during this period). Nonetheless, on average, the primal problem mismatch never exceeded 0.5% of the peak load demand.
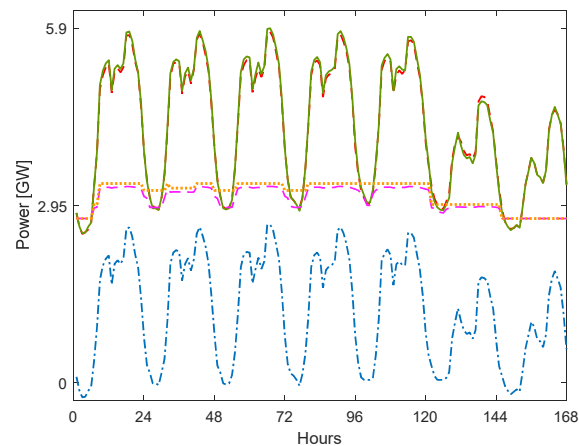
Figure 11. Solution in terms of the primal problem (case study II). In the upper portion are almost coincident solid green and dashed red lines: the obtained generation profile and the load demand, respectively. Dotted orange line: maximum generation capacity of the affected thermal units. Dashed magenta line: thermal units generation profile. Dash-dot blue line: hydro units generation profile.
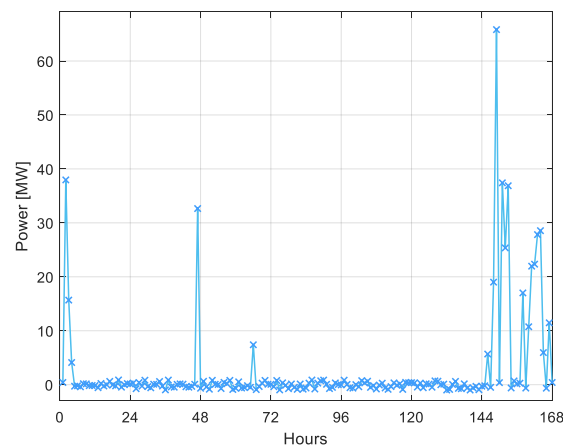

Figure 12. Primal problem solution mismatch: non-feasibility periods of the thermal units' economic dispatch.

7. Conclusions

In this paper, we explored the numerical performance of a novel Lagrangian multiplier update algorithm for short-term hydro-thermal coordination. The step-size update mechanism is a vital component for these methods, and classic approaches are heavily dependent upon a user's experience and fine-tuning procedures, i.e., selecting the appropriate parameters. The proposed algorithm had an important advantage of not requiring parameter choices based on experimentation, and it is subsequently compared against two classical update expressions. After a results assessment of both case studies, we could infer that the adaptive algorithm produced considerably improved dual problem solutions, seen through the percentage gains in terms of average subgradient norm and the respective ratio between the average subgradient norm and peak demand. This ultimately means an improved upper bound for the primal problem solution. Moreover, for most hours it led to feasible primal solutions, which in turn translates into more cost-effective dispatch decisions. The significance of the obtained results is magnified, especially when considering the differences in data prediction, such as the expected inflows.

These improvements proved the algorithm's ability to dynamically adapt the step value according to the dual function value. We can see that during the initial iterations, the step value is incremented

until approaching the vicinities of the dual function maximum. From then onwards, the step evolves dynamically and adapts to the current dual function value, allowing convergence to the maximum dual function value and to the average subgradient norm, which translates to a feasible or a near-feasible primal solution. On the contrary, when using the classical update step-size update equations, it is necessary to rely on educated guesses or perform adjustments over the control parameters (as illustrated by case study II), through a trial-and-error process, in order to obtain a solution similar to the one achieved by the new adaptative algorithm.

Finally, we see that for a high-dimensional optimization problem, the computational burden is not exaggerated and not dependent upon initial guesses, especially considering that a weekly unit commitment is performed.

Author Contributions: Conceptualization, S.J.P.S.M. and L.A.F.M.F.; Methodology, S.J.P.S.M. and L.A.F.M.F.; Investigation, P.M.R.B., M.R.A.C., L.A.F.M.F. and S.J.P.S.M.; Software, S.J.P.S.M. and L.A.F.M.F.; Formal Analysis, P.M.R.B. and S.J.P.S.M.; Visualization, P.M.R.B., S.J.P.S.M. and M.R.A.C.; Supervision, S.J.P.S.M.; Writing—Original Draft Preparation, P.M.R.B., S.J.P.S.M. and M.R.A.C.; Writing—Review & Editing, M.R.A.C., P.M.R.B. and S.J.P.S.M. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

k	discrete time-step (an hour)
i	ith generation resource
K	total number of hours
I	total number of resources
C_{ik}	cost function associated with resource allocation i at time k
x_{ik}	resource state i at time k
p_{ik}	power output by resource i at time k
u_{ik}	control (decision) variable for resource i at time k
D_k	load demand at time k
m	mth-type system capacity requirement
R_{mi}	capacity contribution function associated with resource i for system capacity requirement of type m
R_{mk}^{req}	mth-type system capacity requirement in hour k
M	total number of capacity requirement constraints
n	nth cumulative constraint
$\mathcal{H}_n$	set of all resources constrained by nth cumulative constraint
H_{ni}	function of the contribution of resource i to the nth cumulative constraint
H_n^{req}	lower bound on the nth cumulative constraint
N	number of cumulative constraints
A_{ik}	state function associated with each resource i at time k
$\mathcal{U}_{ik}$	control variables (decision) universe for resource i at time k
X_i^0	resource i initial state
X_i^K	resource i final state
$\mathcal{L}$	Lagrange function
λ	Lagrange multiplier vector associated with the load-balance constraint
μ	Lagrange multiplier vector associated with the capacity constraints
γ	Lagrange multiplier vector associated with the cumulative constraints
$q(\lambda, \mu, \gamma)$	Lagrange dual function
g	subgradient of the dual function
v	current iteration of the subgradient method
s	step-size of the subgradient method
a_1	control parameter of the classic step update heuristic
a_2	control parameter of the classic step update heuristic
α_1	control parameter of the novel adaptative step update heuristic
α_2	control parameter of the novel adaptative step update heuristic
δ	indirect upper and lower bound parameter of the control parameters α_1 and α_2, respectively
$\|g(p_\lambda)\|/K$	average subgradient norm

References

1. Takigawa, F.Y.K.; Finardi, E.C.; da Silva, E.L. A decomposition strategy to solve the short-term hydrothermal scheduling based on Lagrangian relaxation. In Proceedings of the 2010 IEEE/PES Transmission and Distribution Conference and Exposition: Latin America (T&D-LA), Sao Paulo, Brazil, 8–10 November 2010; pp. 681–688.
2. Ye, H.; Zhai, Q.; Ge, Y.; Wu, H. A revised subgradient method for solving the dual problem of hydrothermal scheduling. In Proceedings of the 2011 Asia-Pacific Power and Energy Engineering Conference, Wuhan, China, 25–28 March 2011; pp. 1–6.
3. Zhai, Q.; Wu, H. Several notes on Lagrangian relaxation for unit commitment. In Proceedings of the 29th Chinese Control Conference, Beijing, China, 29–31 July 2010; pp. 1848–1853.
4. Nguyen, T.T.; Vo, D.N. Multi-Objective short-term fixed head hydrothermal scheduling using augmented lagrange hopfield network. *J. Electr. Eng. Technol.* **2014**, *9*, 1882–1890. [CrossRef]
5. Ferreira, L.A.F.M.; Andersson, T.; Imparato, C.F.; Miller, T.E.E.; Pang, C.K.K.; Svoboda, A.; Vojdani, A.F.; Irnparato, C.F.; Miller, T.E.E.; Pang, C.K.K. Short-Term resource scheduling in multi-area hydrothermal power systems. *Int. J. Electr. Power Energy Syst.* **1989**, *11*, 200–212. [CrossRef]
6. Jiménez Redondo, N.; Conejo, A.J. Short-term hydro-thermal coordination by lagrangian relaxation: Solution of the dual problem. *IEEE Trans. Power Syst.* **1999**, *14*, 89–95. [CrossRef]
7. Singhal, P.K. Generation scheduling methodology for thermal units using lagrangian Relaxation. In Proceedings of the 2011 Nirma University International Conference on Engineering, Gujarat, India, 8–10 December 2011; pp. 1–6.
8. Başaran Filik, Ü.; Kurban, M. Solving unit commitment problem using modified subgradient method combined with simulated annealing algorithm. *Math. Probl. Eng.* **2010**, *2010*, 1–15. [CrossRef]
9. Marmolejo-Saucedo, J.A.; Rodríguez-Aguilar, R. A proposed method for design of test cases for economic analysis in power systems. *J. Appl. Res. Technol.* **2015**, *13*, 428–434. [CrossRef]
10. Diniz, A.L.; Sagastizabal, C.; Maceira, M.E.P. Assessment of Lagrangian relaxation with variable splitting for hydrothermal scheduling. In Proceedings of the 2007 IEEE Power Engineering Society General Meeting, Tampa, FL, USA, 24–28 June 2007; pp. 1–8.
11. Lopez-Salgado, C.J.; Ano, O.; Ojeda-Esteybar, D.M. Hydrothermal scheduling with variable head hydroelectric plants: Proposed strategies using benders decomposition and outer approximation. In Proceedings of the 2016 IEEE Power and Energy Conference at Illinois (PECI), Champaign, IL, USA, 19–20 February 2016; pp. 1–8.
12. Rodrigues, R.N.; da Silva, E.L.; Finardi, E.C.; Takigawa, F.Y.K. Solving the short-term scheduling problem of hydrothermal systems via lagrangian relaxation and augmented Lagrangian. *Math. Probl. Eng.* **2012**, *2012*, 1–18. [CrossRef]
13. Che, P.; Tang, Z.; Gong, H.; Zhao, X. An improved Lagrangian relaxation algorithm for the robust generation self-scheduling problem. *Math. Probl. Eng.* **2018**, *2018*, 1–12. [CrossRef]
14. Kumar, R.; Garg, V.; Lal, B. A review paper on hydro-thermal scheduling. *Int. J. Emerg. Technol. Comput. Appl. Sci.* **2013**, *5*, 522–526.
15. Hoseynpour, O.; Mohammadi-Ivatloo, B.; Nazari-Heris, M.; Asadi, S. Application of dynamic non-linear programming technique to non-convex short-term hydrothermal scheduling problem. *Energies* **2017**, *10*, 1–17. [CrossRef]
16. Beltran, C.; Heredia, F.J. Short-Term hydrothermal coordination by augmented Lagrangean relaxation: A new multiplier updating. In Proceedings of the IX Congreso Latino-Iberoamericano de Investigación Operativa, Buenos Aires, Argentina, 31 August–4 September 1998; pp. 1–6.
17. Gil, E.; Araya, J. Short-Term hydrothermal generation scheduling using a parallelized stochastic mixed-integer linear programming algorithm. *Energy Procedia* **2016**, *87*, 77–84. [CrossRef]
18. Fu, B.; Ouyang, C.; Li, C.; Wang, J.; Gul, E. An improved mixed integer linear programming approach based on symmetry diminishing for unit commitment of hybrid power system. *Energies* **2019**, *12*, 833. [CrossRef]
19. Kounalakis, M.E.; Theodorou, P. A hydrothermal coordination model for electricity markets: Theory and practice in the case of the Greek electricity market regulatory framework. *Sustain. Energy Technol. Assess.* **2019**, *34*, 77–86. [CrossRef]
20. Jian, J.; Pan, S.; Yang, L. Solution for short-term hydrothermal scheduling with a logarithmic size mixed-integer linear programming formulation. *Energy* **2019**, *171*, 770–784. [CrossRef]

21. Mariano, S.J.P.S.; Catalão, J.P.S.; Mendes, V.M.F.; Ferreira, L.A.F.M. Optimising power generation efficiency for head-sensitive cascaded reservoirs in a competitive electricity market. *Int. J. Electr. Power Energy Syst.* **2008**, *30*, 125–133. [CrossRef]

22. Ghosh, S.; Kaur, M.; Bhullar, S.; Karar, V. Hybrid ABC-BAT for solving short-term hydrothermal scheduling problems. *Energies* **2019**, *12*, 551. [CrossRef]

23. Farhat, I.A.; El-Hawary, M.E. Optimization methods applied for solving the short-term hydrothermal coordination problem. *Electr. Power Syst. Res.* **2009**, *79*, 1308–1320. [CrossRef]

24. Basu, M. Hopfield neural networks for optimal scheduling of fixed head hydrothermal power systems. *Electr. Power Syst. Res.* **2003**, *64*, 11–15. [CrossRef]

25. Nguyen, T.T.; Vo, D.N. Modified cuckoo search algorithm for short-term hydrothermal scheduling. *Int. J. Electr. Power Energy Syst.* **2015**, *65*, 271–281. [CrossRef]

26. Basu, M. Improved differential evolution for short-term hydrothermal scheduling. *Int. J. Electr. Power Energy Syst.* **2014**, *58*, 91–100. [CrossRef]

27. Zhang, H.; Zhou, J.; Zhang, Y.; Lu, Y.; Wang, Y. Culture belief based multi-objective hybrid differential evolutionary algorithm in short term hydrothermal scheduling. *Energy Convers. Manag.* **2013**, *65*, 173–184. [CrossRef]

28. Das, S.; Bhowmik, D.; Das, S. *A Modified Grey Wolf Optimizer Algorithm for Economic Scheduling of Hydrothermal Systems*; Springer: Cham, Switzerland, 2020; pp. 669–677.

29. Farhat, I.A.; El-Hawary, M.E. Short-Term hydro-thermal scheduling using an improved bacterial foraging algorithm. In Proceedings of the 2009 IEEE Electrical Power & Energy Conference (EPEC), Montreal, QC, Canada, 22–23 October 2009; pp. 1–5.

30. Fakhar, M.S.; Kashif, S.A.R.; Ain, N.U.; Hussain, H.Z.; Rasool, A.; Sajjad, I.A. Statistical performances evaluation of APSO and improved APSO for short term hydrothermal scheduling problem. *Appl. Sci.* **2019**, *9*, 2440. [CrossRef]

31. Nazari-Heris, M.; Babaei, A.F.; Mohammadi-Ivatloo, B.; Asadi, S. Improved harmony search algorithm for the solution of non-linear non-convex short-term hydrothermal scheduling. *Energy* **2018**, *151*, 226–237. [CrossRef]

32. Takigawa, F.Y.K.; Da Silva, E.L.; Finardi, E.C.; Rodrigues, R.N. Solving the hydrothermal scheduling problem considering network constraints. *Electr. Power Syst. Res.* **2012**, *88*, 89–97. [CrossRef]

33. Petcharaks, N.; Ongsakul, W. Hybrid enhanced Lagrangian relaxation and quadratic programming for hydrothermal scheduling. *Electr. Power Compon. Syst.* **2007**, *35*, 19–42. [CrossRef]

34. Padmini, S.; Jegatheesan, R.; Thayyil, D.F. A novel method for solving multi-objective hydrothermal unit commitment and sheduling for GENCO using hybrid LR-EP technique. *Procedia Comput. Sci.* **2015**, *57*, 258–268. [CrossRef]

35. Nayak, C.; Rajan, C.C.A. An evolutionary programming embedded Tabu search method for hydro-thermal scheduling with cooling banking constraints. *J. Eng. Technol. Res.* **2013**, *5*, 21–32. [CrossRef]

36. Bento, P.; Pina, F.; Mariano, S.; do Rosario Calado, M. Short-Term Hydro-Thermal Coordination by Lagrangian Relaxation: A New Algorithm for the Solution of the Dual Problem. Available online: https://knepublishing.com/index.php/KnE-Engineering/article/view/7093/12770 (accessed on 14 September 2020).

37. Catalão, J.P.S.; Mariano, S.J.P.S.; Mendes, V.M.F.; Ferreira, L.A.F.M. A practical approach for profit-based unit commitment with emission limitations. *Int. J. Electr. Power Energy Syst.* **2010**, *32*, 218–224. [CrossRef]

38. Ruiic, S. Optimal distance method for Lagrangian mulitpliers updating in short-term hydro-thermal coordination—Power systems. *IEEE Trans. Power Syst.* **1998**, *13*, 1439–1444.

39. Habibollahzadeh, H.; Brannlund, H.; Bubenko, J.; Sjelvgren, D.; Andersson, N. Optimal Short-Term Planning of Hydro-Thermal Power System—Part II: Solution Techniques. Available online: https://www.sciencedirect.com/science/article/pii/B9780408014687500527 (accessed on 14 September 2020).