Numer Algor (2009) 52:129–149
DOI 10.1007/s11075-008-9261-0

ORIGINAL PAPER

An application of the algorithm of the cyclic coordinate descent in multidimensional optimization problems with constrained speed

L. Bayón · J. M. Grau · M. M. Ruiz · P. M. Suárez

Received: 5 June 2008 / Accepted: 18 December 2008 /
Published online: 13 January 2009

Abstract In this paper we present an application of the algorithm of the cyclic coordinate descent in multidimensional variational problems with constrained speed, the physical motivation of the problem being the optimization of hydrothermal systems. The proof of the convergence of the succession generated by the algorithm was based on the use of an appropriate adaptation of Zangwill's global theorem of convergence. We have also included an algorithm for the formal construction of the descending succession (the solution of an optimum control problem), the approximation of which we carried out using an adaptation of the Euler method in conjunction with a procedure inspired by the shooting method.

Keywords Optimal control · Hydrothermal coordination ·
Coordinate descent · Zangwill's theorem

1 Introduction

A problem that plays a most important role in the safety, reliability and economic operation of electric power systems is the short-term hydrothermal coordination (STHC) problem for power systems. This is a large-scale, multidimensional, nonlinear and constrained problem and there is a vast bibliography describing different formulations and solution methodologies applied to the STHC problem: Linear programming (LP) [1], evolutionary programming

L. Bayón (✉) · J. M. Grau · M. M. Ruiz · P. M. Suárez
Department of Mathematics, University of Oviedo,
E.U.I.T.I. Campus of Viesques, Gijón, 33203, Spain
e-mail: bayon@uniovi.es

(EP) [2], genetic algorithms (GAs) [3, 4], dynamic programming (DP) [5] and Lagrangian relaxation technique [6, 7]. The main drawbacks with the majority of these methods are the difficulty of treating large-scale systems and the fact that these approaches require substantial simplifying assumptions to make the problem computationally tractable. In this paper, we propose using Pontryagin's Minimum Principle (PMP) to solve the STHC problem and an algorithm that we have developed of its numerical resolution prompted by the so-called method of cyclic coordinate descent (CCD).

The coordinate descent method boasts a long-standing history in convex differentiable minimization. Surprisingly, very little is known about the convergence of the iterates generated by this method. Convergence typically requires restrictive assumptions such as that the cost function has bounded level sets and is in some sense strictly convex. Thus, the problem of minimizing a strictly convex (possibly nondifferentiable and nonseparable) function subject to linear constraints is considered in [8]; a convex function of the Legendre type (i.e. a function that is strictly convex, differentiable on an open convex set and whose gradient tends to infinity in norm at the boundary points), subject to linear constraints considered in [9]; and in [10] the author considers that the objective is pseudoconvex in every pair of the coordinate blocks and regular in some natural sense. In the present paper we introduce a relaxation numerical method for our hydrothermal problem and prove its convergence under weak assumptions.

The CCD method has applications in other problems such as in information theory or image reconstruction. Rätsch et al. [11] study Boosting and ensemble learning algorithms and [12] give an unified convergence analysis of ensemble learning methods including e.g. AdaBoost, Logistic Regression and the Least-Square-Boost algorithm for regression. On the other hand, Chambolle and Lions [13] study a classical image denoising technique, namely the constrained minimization of the total variation (TV) of the image and [14] propose a novel bias correction method for magnetic resonance (MR) imaging.

We shall now present our problem from the Electrical Engineering perspective to then go on to resolve the mathematical problem thus formulated.

In prior studies [15, 16], it was proven that the problem of optimization of the fuel cost of a hydrothermal system with several thermal plants may be reduced to the study of a hydrothermal system made up of one single thermal plant, called the *thermal equivalent*. Let us assume that a hydrothermal system accounts for m hydro-plants. The mapping $H : \Omega_H \longrightarrow \mathbb{R}$

$$H(t, z_1(t), \cdots, z_i(t), \cdots, z_m(t), \dot{z}_1(t), \cdots, \dot{z}_i(t), \cdots, \dot{z}_m(t)) = H(t, \mathbf{z}(t), \dot{\mathbf{z}}(t))$$

is called the function of effective hydraulic contribution, and is the power contributed to the system at the instant t by the set of hydro-plants, $z_i(t)$ being the volume that is discharged up to the instant t by the i-th hydro-plant, $\dot{z}_i(t)$ the rate of water discharge at the instant t by the i-th hydro-plant, and $\Omega_H \subset [0, T] \times \mathbb{R}^m \times \mathbb{R}^m$ the domain of definition of H.

In a general model, with hydraulic coupling between the m hydro-plants, we denote by $H_i(t, \mathbf{z}(t), \dot{\mathbf{z}}(t))$ the function of effective hydraulic contribution by the i-th hydro-plant, being

$$H(t, \mathbf{z}(t), \dot{\mathbf{z}}(t)) = \sum_{i=1}^{m} H_i(t, \mathbf{z}(t), \dot{\mathbf{z}}(t))$$

satisfying

$$\frac{\partial^2 H_i(t, \mathbf{z}(t), \dot{\mathbf{z}}(t))}{\partial \dot{z}_i \partial \dot{z}_j} = 0, \, (i \neq j)$$

This condition means that the performance of the i-th hydro-plant is not influenced by the rate of water of the remaining plants, although their volumes may exert an influence.

We say that $\mathbf{z} = (z_1, \cdots, z_m)$ is admissible for H if z_i belong to the class $\widehat{C}^1[0, T]$ (the set of piecewise C^1 functions), and $(t, \mathbf{z}(t), \dot{\mathbf{z}}(t)) \in \Omega_H$, $\forall t \in [0, T]$. The volume b_i that must be discharged up to the instant T by the i-th hydro-plant is called the admissible volume. Let $\mathbf{b} = (b_1, \cdots, b_m) \in \mathbb{R}^m$ be the vector of admissible volumes. The problem consists in minimizing the cost needed to satisfy a certain power demand during the optimization interval $[0, T]$. Said cost may be represented by the functional

$$F(P) = \int_0^T \Psi\,(P(t))\,dt$$

where $\Psi : \mathbb{R} \longrightarrow \mathbb{R}$ is the function of cost of the thermal equivalent and $P(t)$ is the power generated by said plant. Furthermore, the following equilibrium equation of active power will have to be fulfilled

$$P(t) + H(t, \mathbf{z}(t), \dot{\mathbf{z}}(t)) = P_{\mathrm{d}}(t), \;\; \forall t \in [0, T]$$

where $P_{\mathrm{d}}(t)$ is the known power demand (a numerical algorithm for the case with a stochastic noise can be found in [17, 18]), and the following boundary conditions will have to be fulfilled

$$z_i(0) = 0, \, z_i(T) = b_i, \forall i = 1, \cdots, m$$

Besides the previous statement, we consider bounded admissible rates $\dot{z}_i(t)$

$$A_i \leq \dot{z}_i(t) \leq B_i, \;\; \forall i = 1, \cdots, m, \;\; \forall t \in [0, T]$$

Taking into account the equilibrium equation, the thermal component $P(t)$ disappears and our objective functional is

$$J(\mathbf{z}) = \int_0^T L(t, \mathbf{z}(t), \dot{\mathbf{z}}(t))dt \tag{1.1}$$

with $L(t, \mathbf{z}(t), \dot{\mathbf{z}}(t)) = \Psi\left(P_{\mathrm{d}}(t) - H(t, \mathbf{z}(t), \dot{\mathbf{z}}(t))\right)$, on the set

$$\mathbb{D} := \left\{ \mathbf{z} \in \left(\widehat{C}^1[0, T]\right)^m \mid \begin{array}{c} z_i(0) = 0, \, z_i(T) = b_i \\ A_i \leq \dot{z}_i(t) \leq B_i, \;\; \forall t \in [0, T] \end{array}, \forall i = 1, \cdots, m \right\} \tag{1.2}$$

2 Statement of the multidimensional variational problem

At this point, we shall study the above problem, though considering more general constraints for the admissible rates. We shall consider the problem (**Pr**) of minimizing the multidimensional functional

$$J(\mathbf{z}) = \int_0^T L(t, \mathbf{z}(t), \dot{\mathbf{z}}(t))dt \tag{2.1}$$

on $\mathbb{D} := \prod_{i=1}^{m} \mathbb{D}_i$, being

$$\mathbb{D}_i := \{z_i \in (\widehat{C}^1[0,T]) \mid z_i(0) = 0, z_i(T) = b_i \text{ and } \varphi_i(t) \le \dot{z}_i(t) \le \psi_i(t)\} \tag{2.2}$$

Let us assume $\{\varphi_i, \psi_i\}_{i=1}^m \subset C^1[0,T]$ and, with the aim of guaranteeing that $\mathbb{D}_i \neq \varnothing$, we shall also assume that $\varphi_i(t) < \psi_i(t)$, $\forall t \in [0,T]$ and that

$$\int_0^T \varphi_i(s)ds \le b_i \le \int_0^T \psi_i(s)ds$$

We shall assume throughout the paper that

$$L(\cdot,\cdot,\cdot) \in C^2([0,T] \times \mathbb{R}^{2m}) \text{ and } \frac{\partial^2 L(t,\mathbf{z}(t),\dot{\mathbf{z}}(t))}{\partial \dot{z}_i{}^2} > 0, \forall i = 1, \cdots, m$$

We consider $\mathbb{D}$ equipped with the topology induced by the norm

$$||\mathbf{p}||^* := \max\{||\mathbf{p}||_\infty, ||\dot{\mathbf{p}}||_\infty\} = \max\{\max_{i=1,\dots,m} ||p_i||_\infty, \max_{i=1,\dots,m} ||\dot{p}_i||_\infty\}$$

Note that in the topological space $(\mathbb{D}, ||\ ||^*)$:

$\mathbf{p}_n$ converges to $\mathbf{p} \Longleftrightarrow \mathbf{p}_n$ and $\dot{\mathbf{p}}_n$ converge uniformly to $\mathbf{p}$ and $\dot{\mathbf{p}}$ respectively

Definition 1 We define "*i*-th coordination function" of $\mathbf{z} \in \mathbb{D}$ as

$$\mathbb{Y}^i_{\mathbf{z}}(t) := \int_0^t L_{z_i}(s, \mathbf{z}(s), \dot{\mathbf{z}}(s))ds - L_{\dot{z}_i}(t, \mathbf{z}(t), \dot{\mathbf{z}}(t)) \tag{2.3}$$

Observe that $\mathbb{Y}^i_{\mathbf{q}}$ is a constant function when $\forall t \in [0,T]$, $\varphi_i(t) < \dot{q}_i(t) < \psi_i(t)$ q_i, i.e. when is a "free extremal".

Theorem 1 *If* $\mathbf{q} = (q_1, \dots, q_m) \in \mathbb{D}$ *is solution of the problem* Pr*, then there exists* $\{C_i\}_{i=1}^m \subset \mathbb{R}$ *satisfying:*

$$\mathbb{Y}^i_{\mathbf{q}}(t) is \begin{cases} \le C_i & if\ \dot{q}_i(t) = \varphi_i(t) \\ = C_i & if\ \varphi_i(t) < \dot{q}_i(t) < \psi_i(t) \\ \ge C_i & if\ \dot{q}_i(t) = \psi_i(t) \end{cases}$$

Proof To prove the above result, we present the problem considering the state variable to be $\mathbf{z}(t) = (z_1(t), \cdots, z_m(t))$, the control variable $\mathbf{u}(t) = (u_1(t), \cdots, u_m(t))$, and the state equation $\dot{\mathbf{z}}(t) = \mathbf{u}(t)$. The optimal control problem is thus:

$$\min_{\mathbf{u}} \int_0^T L(t, \mathbf{z}(t), \mathbf{u}(t))dt \quad \text{with} \quad \begin{cases} \dot{\mathbf{z}}(t) = \mathbf{u}(t) \\ \mathbf{z}(0) = \mathbf{0}, \quad \mathbf{z}(T) = \mathbf{b} \\ \mathbf{u}(t) \in \Omega(t) = \prod_{i=1}^{m} [\varphi_i(t), \psi_i(t)] \end{cases}$$

We shall term the optimal control $\mathbf{u}^*(t) = (u_1^*(t), \cdots, u_m^*(t)) = \dot{\mathbf{q}}(t)$, therefore the optimal state will be $\mathbf{q}$. Let $\hbar$ be the Hamiltonian associated with the problem

$$\hbar(t, \mathbf{z}, \mathbf{u}, \lambda) = L(t, \mathbf{z}, \mathbf{u}) + \sum_{i=1}^{m} \lambda_i \cdot u_i$$

In virtue of Pontryagin's Principle, there exists a piecewise C^1 function (co-state variable):

$$\lambda^*(t) = (\lambda_i^*(t), \cdots, \lambda_m^*(t))$$

that satisfies the two following conditions:

$$\dot{\lambda}_i^*(t) = -\frac{\partial \hbar(t, \mathbf{q}(t), \mathbf{u}^*(t), \lambda^*(t))}{\partial z_i} = -L_{z_i}(t, \mathbf{q}(t), \mathbf{u}^*(t)) \tag{2.4}$$

$$\hbar(t, \mathbf{q}(t), \mathbf{u}^*(t), \lambda^*(t)) \leq \hbar(t, \mathbf{q}(t), \mathbf{u}, \lambda^*(t)), \ \forall \mathbf{u} \in \Omega(t) \tag{2.5}$$

From (2.4) it follows that

$$\lambda_i^*(t) = -\int_0^t L_{z_i}(s, \mathbf{q}(s), \mathbf{u}^*(s))ds + C_i$$

From (2.5) it follows that for each t, $\mathbf{u}^*(t)$ minimizes the function

$$\Upsilon(\mathbf{u}) := \hbar(t, \mathbf{q}(t), \mathbf{u}, \lambda^*(t)) = L(t, \mathbf{q}(t), \mathbf{u}(t)) + \sum_{i=1}^{m} \lambda_i^*(t) \cdot u_i(t)$$

on

$$\prod_{i=1}^{m} [\varphi_i(t), \psi_i(t)]$$

Bearing in mind that $L_{u_i} = L_{\dot{z}_i}$ and that

$$\frac{\partial \Upsilon(\mathbf{u})}{\partial u_i} = L_{\dot{z}_i}(t, \mathbf{q}(t), \mathbf{u}(t)) + \lambda_i^*(t)$$

we have three possibilities:

1) $\varphi_i(t) < u_i(t) < \psi_i(t) \Longrightarrow 0 = \dfrac{\partial \Upsilon(\mathbf{u}^*)}{\partial u_i} = -\mathbb{Y}^i_{\mathbf{q}}(t) + C_i \Longrightarrow \mathbb{Y}^i_{\mathbf{q}}(t) = C_i$

2) $\varphi_i(t) = u_i(t) \Longrightarrow 0 \leq \dfrac{\partial \Upsilon(\mathbf{u}^*)}{\partial u_i} = -\mathbb{Y}^i_{\mathbf{q}}(t) + C_i \Longrightarrow \mathbb{Y}^i_{\mathbf{q}}(t) \leq C_i$

3) $\psi_i(t) = u_i(t) \Longrightarrow 0 \geq \dfrac{\partial \Upsilon(\mathbf{u}^*)}{\partial u_i} = -\mathbb{Y}^i_{\mathbf{q}}(t) + C_i \Longrightarrow \mathbb{Y}^i_{\mathbf{q}}(t) \geq C_i$ □

3 Definition of the descent algorithm

Let $\mathbf{q} = (q_1, \cdots, q_m) \in \mathbb{D}$. We shall consider

$$L^i_{\mathbf{q}}(t, z_i, \dot{z}_i) := L(q_1(t), \cdots, q_{i-1}(t), z_i, q_{i+1}(t), \cdots, q_m(t), \dot{q}_1(t), \cdots, \dot{z}_i, \cdots, \dot{q}_m(t))$$

and the functional $J^i_{\mathbf{q}} : \mathbb{D}_i \longrightarrow \mathbb{R}$

$$J^i_{\mathbf{q}}(z_i) := J(q_1, \cdots, q_{i-1}, z_i, q_{i+1}, \cdots, q_m) = \int_0^T L^i_{\mathbf{q}}(t, z_i(t), \dot{z}_i(t))dt$$

It is well known that the uniform boundedness of the derivatives, together with the assumed convexity of $L^i_{\mathbf{q}}(t, z_i, \cdot)$, allow us to guarantee that the functional $J^i_{\mathbf{q}}$ reaches its minimum in an absolutely continuous function. However, if $\mathbf{q}$ is of class C^1, we are also in a position to guarantee [19] that the minimum of $L^i_{\mathbf{q}}$ on the set $\mathbb{D}_i$ is of class C^1. We shall also make the additional assumption that for each $\mathbf{q} \in \mathbb{D}$ and for each $i = 1, \ldots m$, the functional $J^i_{\mathbf{q}}$ possesses a unique minimum and hence that the condition of minimality that is derived from Pontryagin's Minimum Principle in the corresponding control problem is a sufficient minimum condition.

Definition 2 We define i-th minimizing map as the map $\Phi_i : \mathbb{D} \longrightarrow \mathbb{D}$ that satisfies:

i) $\mathbf{q} - \Phi_i(\mathbf{q}) = (d_1, \cdots, d_m)$ *with* $d_j(\cdot) = 0, \ \forall j \neq i$
ii) $J(\Phi_i(\mathbf{q})) \leq J(\mathbf{z}), \forall \mathbf{z} \in \{\mathbf{z} \in \mathbb{D} | z_j(\cdot) = q_j(\cdot), \ \forall j \neq i\}$

We shall denote by Φ the map associated with the descent algorithm, which will be the composition of the i-th minimizing map:

$$\Phi = \Phi_m \circ \cdots \circ \Phi_1$$

The following proposition is verified, the demonstration of which is identical to that of Theorem 1.

Proposition 1 *There exists* $\forall \mathbf{q} = (q_1, \ldots, q_m) \in \mathbb{D}$, $C_i \in \mathbb{R}$ *such that if*

$$\Phi_i(\mathbf{q}) = (q_1, \cdots, q_{i-1}, Q, q_{i+1}, \cdots, q_m)$$

then

$$\mathbb{Y}^i_{\Phi_i(\mathbf{q})}(t) \text{ is } \begin{cases} \leq C_i \text{ if } \dot{Q}(t) = \varphi_i(t) \\ = C_i \text{ if } \varphi_i(t) < \dot{Q}(t) < \psi_i(t) \\ \geq C_i \text{ if } \dot{Q}(t) = \psi_i(t) \end{cases}$$

3.1 Formal construction of $\Phi_i(\mathbf{q})$

Given $\mathbf{q} = (q_1, \ldots, q_m) \in \mathbb{D}$, we shall consider, for each $K \in \mathbb{R}$,

$$\mathbf{q}^i_K = (q_1, \cdots, q_{i-1}, Q_K, q_{i+1}, \cdots, q_m)$$

satisfying

$$\begin{cases} \mathbb{Y}^i_{\mathbf{q}^i_K}(t) \leq K \text{ if } \dot{Q}_K(t) = \varphi_i(t) \\ \mathbb{Y}^i_{\mathbf{q}^i_K}(t) = K \text{ if } \varphi_i(t) < \dot{Q}_K(t) < \psi_i(t) \\ \mathbb{Y}^i_{\mathbf{q}^i_K}(t) \geq K \text{ if } \dot{Q}_K(t) = \psi_i(t) \end{cases}$$

$$Q_K(0) = 0$$

That is, Q_K minimizes the functional $J^i_{\mathbf{q}}$ within the set:

$$\{z_i \in (\widehat{C}^1[0, T]) \mid z_i(0) = 0, z_i(T) = Q_K(T) \text{ and } \varphi_i(t) \leq \dot{z}_i(t) \leq \psi_i(t)\}$$

In these terms, the objective is now to find C_i such that $Q_{C_i}(T) = b_i$, and that the following is therefore fulfilled:

$$\Phi_i(\mathbf{q}) = (q_1, \ldots, q_{i-1}, Q_{C_i}, q_{i+1}, \ldots, q_m)$$

In the absence of constraints for the admissible functions, the calculation of C_i could be achieved by means of the shooting method, for example; bear in mind that the goal would be to solve the Euler equation of the functional with the boundary conditions $z(0) = 0$ and $z_i(T) = b_i$. In our case, we shall use the same idea, except that the variation of the initial condition for the derivative, which now need not make sense, is substituted by the variation of the constant K. In short, the problem will consist in finding for each K the function Q_K which satisfies $Q_K(0) = 0$ and the conditions of Proposition 1, and from among these functions, the one which generates an admissible function.

To formally construct the function Q_K, we shall consider

$$0 = t_0 < t_1 < \cdots < t_p = T$$

such that in each (t_{j-1}, t_j) the following is fulfilled:

$$\varphi_i(t) < \dot{Q}_K(t) < \psi_i(t) \text{ or } \varphi_i(t) = \dot{Q}_K(t) \text{ or } \dot{Q}_K(t) = \psi_i(t)$$

We shall carry out p steps, in each of which we shall construct $\omega_j \in C^1[t_{j-1}, t_j]$ such that $\omega_j(t_j) = \omega_{j+1}(t_j)$ and $\dot{\omega}_j(t_j) = \dot{\omega}_{j+1}(t_j)$ and that the function defined from these as

$$Q_K(t) := \omega_j(t) \text{ where } j \text{ is such that } t \in [t_{j-1}, t_j]$$

satisfies the minimality conditions expressed in Proposition 1.

Concatenation of the extremal arcs

Step 1] (the first arc)

i) If $K \geq -\dfrac{\partial L^i_{\mathbf{q}}(t, 0, \varphi(0))}{\partial \dot{z}_i}$ we set $\omega_1(t) := \int_0^t \varphi(s)ds$ in the maximal interval $[0, t_1]$, where

$$K \geq \int_0^t \frac{\partial L^i_{\mathbf{q}}(s, \omega_1(s), \dot{\omega}_1(s))}{\partial z_i} ds - \frac{\partial L^i_{\mathbf{q}}(t, \omega_1(t), \dot{\omega}_1(t))}{\partial \dot{z}_i}$$

ii) If $K \leq -\dfrac{\partial L^i_{\mathbf{q}}(t, 0, \psi(0))}{\partial \dot{z}_i}$ we set $\omega_1(t) := \int_0^t \psi(s)ds$ in the maximal interval $[0, t_1]$, where

$$K \leq \int_0^t \frac{\partial L^i_{\mathbf{q}}(s, \omega_1(s), \dot{\omega}_1(s))}{\partial z_i} ds - \frac{\partial L^i_{\mathbf{q}}(t, \omega_1(t), \dot{\omega}_1(t))}{\partial \dot{z}_i}$$

iii) If $-\dfrac{\partial L^i_{\mathbf{q}}(t, 0, \psi(0))}{\partial \dot{z}_i} < K < -\dfrac{\partial L^i_{\mathbf{q}}(t, 0, \varphi(0))}{\partial \dot{z}_i}$ then $\exists x \in (\varphi(0), \psi(0))$ such that $K = \dfrac{\partial L^i_{\mathbf{q}}(0, 0, x)}{\partial \dot{z}_i}$, and we set $\omega_1(t)$ the arc of the interior extremal de $L^i_{\mathbf{q}}$ (with $\omega_1(0) = 0, \dot{\omega}_1(0) = x$) which satisfies Euler's equation in its maximal domain $[0, t_1]$ and, therefore

$$K = \int_0^t \frac{\partial L^i_{\mathbf{q}}(s, \omega_1(s), \dot{\omega}_1(s))}{\partial z_i} ds - \frac{\partial L^i_{\mathbf{q}}(t, \omega_1(t), \dot{\omega}_1(t))}{\partial \dot{z}_i}$$

j-th Step] (j-th arc)

A) If ω_{j-1} has an interior extremal arc in $[t_{j-2}, t_{j-1}]$, there are two possibilities:

I) If $\dot{\omega}_{j-1}(t_{j-1}) = \varphi_i(t_{j-1})$, we set $\omega_j(t) = \omega_{j-1}(t_{j-1}) + \int_{t_{j-1}}^t \varphi_i(s)ds$ in the maximal interval $[t_{j-1}, t_j]$ such that

$$-\frac{\partial L^i_{\mathbf{q}}(t, \omega_{j-1}(t_{j-1}), \varphi_i(t_{j-1}))}{\partial \dot{z}_i} \geq \int_{t_{j-1}}^t \frac{\partial L^i_{\mathbf{q}}(s, \omega_j(s), \varphi_i(s))}{\partial z_i} ds - \frac{\partial L^i_{\mathbf{q}}(t, \omega_j(t), \varphi_i(t))}{\partial \dot{z}_i}$$

II) If $\dot{\omega}_{j-1}(t_{j-1}) = \psi_i(t_{j-1})$, we set $\omega_j(t) = \omega_{j-1}(t_{j-1}) + \int_{t_{j-1}}^t \psi_i(s)ds$ in the maximal interval $[t_{j-1}, t_j]$ such that

$$-\frac{\partial L^i_{\mathbf{q}}(t, \omega_{j-1}(t_{j-1}), \psi_i(t_{j-1}))}{\partial \dot{z}_i} \leq \int_{t_{j-1}}^t \frac{\partial L^i_{\mathbf{q}}(s, \omega_j(s), \psi_i(s))}{\partial z_i} ds - \frac{\partial L^i_{\mathbf{q}}(t, \omega_j(t), \psi_i(t))}{\partial \dot{z}_i}$$

B) If $[t_{j-2}, t_{j-1}]$ is the boundary interval, we set $\omega_j(t)$ the arc of the interior extremal (with $\omega_j(t_{j-1}) = \omega_{j-1}(t_{j-1}), \dot{\omega}_j(t_{j-1}) = \dot{\omega}_{j-1}(t_{j-1})$) which satisfies Euler's equation in its maximal domain $[t_{j-1}, t_j]$ and, therefore,

$$-\frac{\partial L^i_{\mathbf{q}}(t, \omega_{j-1}(t_{j-1}), \dot{\omega}_{j-1}(t_{j-1}))}{\partial \dot{z}_i} =$$

$$= \int_{t_{j-1}}^{t} \frac{\partial L^i_{\mathbf{q}}(s, \omega_j(s), \dot{\omega}_{j-1}(s))}{\partial z_i} ds - \frac{\partial L^i_{\mathbf{q}}(t, \omega_j(t), \dot{\omega}_{j-1}(t))}{\partial \dot{z}_i}$$

3.2 Approximate construction of $\Phi_i(\mathbf{q})$

The peculiar form of $\Phi_i(\mathbf{q})$, expressed in Proposition 1, allows us to undertake its approximate calculation using similar numerical methods to those used to solve differential equations in combination with an appropriate adaptation of the classical shooting method. More precisely, we shall undertake two processes of approximation:

- Construction of a sequence $\{K_j\}_{j\in\mathbb{N}}$ such that $Q_{K_j}(T)$ converges to b_i (the adapted shooting method).
- Approximate construction of each Q_{K_j} (the adapted Euler method).

The approximate construction of each Q_K, which we shall call $\widetilde{Q}_K$, is carried out by means of polygonals (Euler's method) considering the triple recurring sequence (X_n, Y_n, I_n) with $n = 0, \cdots, N-1$ and $h = \dfrac{T}{N}$ which represents the following approximations:

$$Q_K(t_n) \approx \widetilde{Q}_K(t_n) := X_n$$

$$\dot{Q}_K(t_n) \approx \dot{\widetilde{Q}}_K(t_n) := Y_n$$

$$Q_K(t) \approx \widetilde{Q}_K(t) := X_{n-1} + (t - t_{n-1}) \cdot Y_{n-1} \text{ in } [t_{n-1}, t_n]$$

$$\int_0^{t_n} \frac{\partial L^i_{\mathbf{q}}(s, Q_K(s), \dot{Q}_K(s))}{\partial z_i} ds \approx I_n := \int_0^{t_n} \frac{\partial L^i_{\mathbf{q}}(s, \widetilde{Q}_K(s), \dot{\widetilde{Q}}_K(s))}{\partial z_i} ds$$

and which obeys the following relation of recurrence:

$$X_0 = 0; \;\; I_0 = 0$$

$$Y_n = \begin{cases} \psi_i(t_n) \text{ if } & \chi \geq \psi_i(t_n) \\ \varphi_i(t_n) \text{ if } & \chi \leq \varphi_i(t_n) \\ \quad \chi \quad \text{if } & \varphi_i(t_n) < \chi < \psi_i(t_n) \end{cases}$$

$$\text{with } \chi \text{ solution of: } I_n - \frac{\partial L^i_{\mathbf{q}}(t_n, X_n, \chi)}{\partial \dot{z}_i} = K$$

$$X_{n+1} = X_n + h \cdot Y_n$$

$$I_{n+1} = I_n + \int_{t_n}^{t_{n+1}} \frac{\partial L^i_{\mathbf{q}}(s, X_n + (s - t_n) \cdot Y_n, Y_n)}{\partial z_i} ds$$

The following proposition guarantees that at the nodes $\{t_n\}_{n=0}^{N-1}$ the approximation $\widetilde{\mathbf{q}}^i_K$ satisfies the condition established in Proposition 1.

Proposition 2 *If* $\widetilde{\mathbf{q}}^i_K := (q_1, \cdots, q_{i-1}, \widetilde{Q}_K, q_{i+1}, \cdots, q_m)$ *then in* $\{t_n\}_{n=0}^{N-1}$ *the following is satisfied:*

$$\mathbb{Y}^i_{\widetilde{\mathbf{q}}^i_K}(t_n) \text{ is } \begin{cases} \leq K \text{ if } Y_n = \varphi_i(t_n) \\ = K \text{ if } \varphi_i(t) < Y_n < \psi_i(t_n) \\ \geq K \text{ if } Y_n = \psi_i(t_n) \end{cases}$$

Proof Let us bear in mind that

$$\mathbb{Y}^i_{\mathbf{q}^i_K}(t_n) \approx \mathbb{Y}^i_{\widetilde{\mathbf{q}}^i_K}(t_n) = I_n - \frac{\partial L^i_{\mathbf{q}}(t_n, X_n, Y_n)}{\partial \dot{z}_i}$$

If

$$\varphi_i(t_n) < Y_n < \psi_i(t_n) \Longrightarrow \mathbb{Y}^i_{\widetilde{\mathbf{q}}^i_K}(t_n) = I_n - \frac{\partial L^i_{\mathbf{q}}(t_n, X_n, Y_n)}{\partial \dot{z}} = K$$

Considering now that $\dfrac{\partial L^i_{\mathbf{q}}(t_n, X_n, \cdot)}{\partial \dot{z}_i}$ is increasing and $\dfrac{\partial L^i_{\mathbf{q}}(t_n, X_n, \cdot)}{\partial \dot{z}_i^2} > 0$, we have that:

If $Y_n = \varphi_i(t_n) \Longrightarrow \exists \chi \leq \varphi_i(t_n)$ such that

$$I_n - \frac{\partial L^i_{\mathbf{q}}(t_n, X_n, \chi)}{\partial \dot{z}_i} = K \Longrightarrow \mathbb{Y}^i_{\widetilde{\mathbf{q}}^i_K}(t_n) \leq K$$

If $Y_n = \psi_i(t_n) \Longrightarrow \exists \chi \geq \psi_i(t_n)$ such that

$$I_n - \frac{\partial L^i_{\mathbf{q}}(t_n, X_n, \chi)}{\partial \dot{z}_i} = K \Longrightarrow \mathbb{Y}^i_{\widetilde{\mathbf{q}}^i_K}(t_n) \geq K$$

□

4 Cyclic coordinated descent algorithm

4.1 Extension of Zangwill's global convergence theorem

We now go on to present a topological version of the global convergence theorem of descent algorithms with more general hypotheses that do not affect the correctness of the demonstration given in [20] by Zangwill; specifically, the continuity of the descending function is substituted by sequential continuity and the compactness by relative sequential compactness.

Theorem 2 (Global Convergence, generalized version) *Let Φ be a map on the topological space (X, τ) and $x_0 \in X$. Let us assume that the recurrence sequence $\{x_n\}_{n\in\mathbb{N}}$ defined by*

$$x_{n+1} = \Phi(x_n)$$

verifies the following:

i) $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{K} \subset X$, *where* $\mathbb{K}$ *is relatively sequentially compact.*
ii) *There exists a sequentially continuous function* $F : (\mathbb{X}, \tau) \to (\mathbb{R}, |\quad|)$ *satisfying:*

$$\Phi(x) \neq x \Longrightarrow F(\Phi(x)) < F(x)$$

iii) Φ *is sequentially continuous in* X.

Hence, every convergent subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ *converges to a fixed point on* Φ.

Proof It is identical to that presented by Zangwill for the global convergence theorem of descent algorithms. It need only be pointed out that the transformation Φ associated with the algorithm is pointwise and that the solution set is that of fixed points on Φ. □

Although in measurable topological spaces the sequential character of compactness and continuity is irrelevant, we shall maintain this terminology so as to facilitate the exposition.

4.2 Convergence of the algorithm

Let us now see a series of preparatory results, the demonstrations of which we shall omit in some cases on account of their being simple or well known.

Lemma 1 *Let* (X, τ) *be a topological space with* $K \subset X$ *relatively sequentially compact. If the sequence* $\{x_n\}_{n\in\mathbb{N}} \subset K$ *verifies that all its convergent subsequences have the same limit, then* $\{x_n\}_{n\in\mathbb{N}}$ *converges to this same limit.*

Lemma 2 *Given a family of functions*

$$\mathbb{F} = \{\mathbf{f}_\lambda = (f_{\lambda,1}, \ldots, f_{\lambda,m})\}_{\lambda\in I} \subset \widehat{C}^1[0, T]^m$$

if the family of its derivatives $\{\dot{\mathbf{f}}_\lambda\}_{\lambda\in\mathbb{I}}$ *is uniformly bounded, then* $\mathbb{F}$ *is equicontinuous.*

Lemma 3 *Let* $\Omega \subset [0, T] \times \mathbb{R}^{2m}$ *and* $L : \Omega \longrightarrow \mathbb{R}$ *of class* C^1. *For each* $\mathbf{h} \in C^1([0, T], \mathbb{R}^m)$ *such that* $(t, \mathbf{h}(t), \dot{\mathbf{h}}(t)) \in \Omega\ \forall t \in [0, T]$ *we define*

$$L_{\mathbf{h}}(t) := L(t, \mathbf{h}(t), \dot{\mathbf{h}}(t)) \text{ and } W_h(t) := \int_0^t L(s, \mathbf{h}(s), \dot{\mathbf{h}}(s))ds$$

If $\{\mathbf{q}_n\}_{n\in\mathbb{N}}$ *converges to* $\mathbf{q}$ *in* $(\mathbb{D}, ||\quad||^*)$ *then*

$$\{L_{\mathbf{q}_n}\} \text{ converges uniformly to } L_{\mathbf{q}} \text{ and } \{W_{\mathbf{q}_n}\} \text{ converges pointwise to } W_{\mathbf{q}}$$

Lemma 4 *If* $\{\mathbf{q}_n\}_{n\in\mathbb{N}} \subset \mathbb{D}$ *converges uniformly to* $\mathbf{q}$ *and* $\{\dot{\mathbf{q}}_n\}_{n\in\mathbb{N}}$ *is equicontinuous and uniformly bounded, then:*

$$\{\dot{\mathbf{q}}_n\}_{n\in\mathbb{N}} \text{ converges uniformly to } \dot{\mathbf{q}}$$

We shall next see that if a sequence $\{\mathbf{q}_n\}$ satisfies the thesis of Theorem 1 for a certain sequence $\left\{C_{n,i}\right\}$, then its limit also satisfies said thesis for the limit of the $\left\{C_{n,i}\right\}$.

Proposition 3 *If* L *is of class* C^2 *and* $\{\mathbf{q}_n\}_{n\in\mathbb{N}}$ *converges uniformly to* $\mathbf{q}$ *in* $(\mathbb{D}, ||\ \ ||^*)$, *then* $\{\mathbb{Y}^i_{\mathbf{q}_n}\}_{n\in\mathbb{N}}$ *converges pointwise to* $\mathbb{Y}^i_{\mathbf{q}}$, $\forall i = 1, \ldots m$.

Proof Considering, for each $\mathbf{z} \in \mathbb{D}$

$$\left.\begin{array}{l}\mathbb{L}^i_{\mathbf{z}}(t) := L_{\dot{z}_i}(t, \mathbf{z}(t), \dot{\mathbf{z}}(t)) \\ \mathbb{S}^i_{\mathbf{z}}(t) := L_{z_i}(s, \mathbf{z}(s), \dot{\mathbf{z}}(t)) \\ \mathbb{I}^i_{\mathbf{z}}(t) := \int_0^t \mathbb{S}^i_{\mathbf{z}}(s)ds\end{array}\right\} \Longrightarrow \mathbb{Y}^i_{\mathbf{q}_n}(t) = \mathbb{I}^i_{\mathbf{q}_n}(t) - \mathbb{L}^i_{\mathbf{q}_n}(t)$$

In virtue of Lemma 4, $\{\mathbb{L}^i_{\mathbf{q}_n}\}_{n\in\mathbb{N}}$ converges uniformly to $\mathbb{L}^i_{\mathbf{q}}$ and $\{\mathbb{S}^i_{\mathbf{q}_n}\}_{n\in\mathbb{N}}$ converges uniformly to $\mathbb{S}^i_{\mathbf{q}}$. Thus, $\{\mathbb{I}^i_{\mathbf{q}_n}\}_{n\in\mathbb{N}}$ converges pointwise to $\mathbb{I}^i_{\mathbf{q}}$ and $\{\mathbb{Y}^i_{\mathbf{q}_n}\}_{n\in\mathbb{N}}$ converges pointwise to $\mathbb{Y}^i_{\mathbf{q}}$. □

Corollary 1 *If* $\{\mathbf{q}_n\}_{n\in\mathbb{N}}$ *converges to* $\mathbf{q}$ *in* $(\mathbb{D}, ||\ \ ||^*)$ *verifying*

$$\mathbb{Y}^i_{\mathbf{q}_n}(t) is \begin{cases} \leq C_{n,i} \text{ if } \dot{q}_{n,i}(t) = \varphi_i(t) \\ = C_{n,i} \text{ if } \varphi_i(t) < \dot{q}_{n,i}(t) < \psi_i(t) \\ \geq C_{n,i} \text{ if } \dot{q}_{n,i}(t) = \psi_i(t) \end{cases}$$

then the sequence $\{C_{n,i}\}_{n\in\mathbb{N}}$ *converges and, calling its limit* C_i, *it is verified that*

$$\mathbb{Y}^i_{\mathbf{q}}(t) \text{ is } \begin{cases} \leq C_i \text{ if } \dot{q}_i(t) = \varphi_i(t) \\ = C_i \text{ if } \varphi_i(t) < \dot{q}_i(t) < \psi_i(t) \\ \geq C_i \text{ if } \dot{q}_i(t) = \psi_i(t) \end{cases}$$

Proof It is evident that, $\forall i = 1, \ldots, m$, the sequence $\{C_{n,i}\}_{n\in\mathbb{N}}$ converges, since otherwise $\{\mathbb{Y}^i_{\mathbf{q}_n}\}_{n\in\mathbb{N}}$ would not converge, thus contradicting Proposition 3.

– If $\varphi_i(t) < \dot{q}_i(t) < \psi_i(t)$,there exists $k \in \mathbb{N}$ such that, for every $n > k$, $\varphi_i(t) < \dot{q}_{n,i}(t) < \psi_i(t)$, and therefore

$$\mathbb{Y}^i_{\mathbf{q}}(t) = \lim_{n\to\infty} \mathbb{Y}^i_{\mathbf{q}_n}(t) = \lim_{n\to\infty} C_{n,i} = C_i$$

– If $\varphi_i(t) = \dot{q}_i(t)$ there exists $k \in \mathbb{N}$ such that, for every $n > k$, $\dot{q}_{n,i}(t) < \psi_i(t)$ and therefore

$$\mathbb{Y}^i_{\mathbf{q}_n}(t) \leq C_{i,n} \Longrightarrow \mathbb{Y}^i_{\mathbf{q}}(t) \leq C_i$$

– If $\psi_i(t) = \dot{q}_i(t)$ there exists $k \in \mathbb{N}$ such that, for every $n > k$, $\dot{q}_{n,i}(t) > \varphi_i(t)$ and therefore

$$\mathbb{Y}^i_{\mathbf{q}_n}(t) \geq C_{i_n} \Longrightarrow \mathbb{Y}^i_{\mathbf{q}}(t) \geq C_i$$

□

We shall now show the conservation of convergence by means of the i-th minimizing map Φ_i.

Proposition 4 *If* $\{\mathbf{q}_n\}_{n\in\mathbb{N}}$ *and* $\{\Phi_i(\mathbf{q}_n)\}_{n\in\mathbb{N}}$ *converge in* $(\mathbb{D}, ||\ \ ||^*)$ *then*

$$\{\Phi_i(\mathbf{q}_n)\}_{n\in\mathbb{N}} \text{ converges to } \Phi_i(\lim_{n\to\infty}(\mathbf{q}_n))$$

Proof Let $\mathbf{s}_n := \Phi_i(\mathbf{q}_n)$ which converges uniformly to $\mathbf{s}$ and $\dot{\mathbf{s}}_n$ to $\dot{\mathbf{s}}$. Proposition 1, together with Corollary 1, guarantees that

$$\mathbb{Y}^i_{\mathbf{s}}(t) \text{ is } \begin{cases} \leq C_i \text{ if } \dot{s}_i(t) = \varphi_i(t) \\ = C_i \text{ if } \varphi_i(t) < \dot{s}_i(t) < \psi_i(t) \\ \geq C_i \text{ if } \dot{s}_i(t) = \psi_i(t) \end{cases}$$

and, hence, $\Phi_i(\mathbf{s}) = \mathbf{s}$.

Now note that $\mathbf{q}_n$ and $\mathbf{s}_n = \Phi_i(\mathbf{q}_n)$ differ only in their i-th component, as do their limits. Thus,

$$\Phi_i(\lim_{n\to\infty}(\mathbf{q}_n)) = \Phi_i(\mathbf{s}) = \mathbf{s} = \lim_{n\to\infty}(\Phi_i(\mathbf{q}_n))$$

□

The following corollary extends the above result to the composition of the i-th minimizing map, which constitutes the map associated with the descent algorithm Φ.

Corollary 2 *If* $\{\mathbf{q}_n\}_{n\in\mathbb{N}}$ *and* $\{\Phi(\mathbf{q}_n)\}_{n\in\mathbb{N}}$ *converge in* $(\mathbb{D}, ||\ \ ||^*)$ *then*

$$\{\Phi(\mathbf{q}_n)\}_{n\in\mathbb{N}} \text{ converges to } \Phi(\lim_{n\to\infty}(\mathbf{q}_n))$$

Proof If $\{\mathbf{q}_n = (q_{n,1}, ..., q_{n,m})\}_{n\in\mathbb{N}}$ converges to $\mathbf{q} = (q_1, ..., q_m)$ in $(\mathbb{D}, ||\ \ ||^*)$, we thus have that $\{q_{n,i}\}_{n\in\mathbb{N}}$ and $\{\dot{q}_{n,i}\}_{n\in\mathbb{N}}$ converge uniformly to q_i and $\dot{q}_i$, respectively, $\forall i = 1, ..., m$. As $\{\Phi(\mathbf{q}_n)\}_{n\in\mathbb{N}}$ also converges in $(\mathbb{D}, ||\ \ ||^*)$, we have that $\{\Phi(\mathbf{q}_n) = (q^*_{n,1}, ..., q^*_{n,m})\}_{n\in\mathbb{N}}$ converges to $\mathbf{s}$ and that $\{\dot{\Phi}(\mathbf{q}_n) = (\dot{q}^*_{n,1}, ..., \dot{q}^*_{n,m})\}_{n\in\mathbb{N}}$ converges to $\dot{\mathbf{s}}$. Hence,

$$(\Phi_{m-1} \circ \cdots \circ \Phi_1)(\mathbf{q}_n) = (q^*_{n,1}, ..., q^*_{n,m-1}, q_{n,m})$$

also converges, since $q^*_{n,i}$ converges to s_i, $\forall i = 1, ..., n-1$ and $\{q_{n,m}\}_{n\in\mathbb{N}}$ converges to q_m by hypothesis.

Applying this reasoning reiteratively, we have that the successions

$$\{(\Phi_{m-2} \circ \cdots \circ \Phi_1)\,(\mathbf{q}_n)\}_{n\in\mathbb{N}}, \cdots, \{(\Phi_2 \circ \Phi_1)\,(\mathbf{q}_n)\}_{n\in\mathbb{N}}, \quad \{\Phi_1(\mathbf{q}_n)\}_{n\in\mathbb{N}}$$

are convergent in $(\mathbb{D}, ||\ \ ||^*)$.

Moreover, in virtue of Proposition 4, we have that:

– as $\{\mathbf{q}_n\}_{n\in\mathbb{N}}$ converges to $\mathbf{q}$ and $\{\Phi_1(\mathbf{q}_n)_{n\in\mathbb{N}}\}$ is convergent, then

$$\{\Phi_1(\mathbf{q}_n)_{n\in\mathbb{N}}\} \text{ converges to } \Phi_1(\mathbf{q})$$

– as $\{\Phi_1(\mathbf{q}_n)_{n\in\mathbb{N}}$ converges to $\Phi_1(\mathbf{q})$ and $\left\{(\Phi_2 \circ \Phi_1)\,(\mathbf{q}_n)\right\}_{n\in\mathbb{N}}$ is convergent, then

$$\{(\Phi_2 \circ \Phi_1)\,(\mathbf{q}_n)_{n\in\mathbb{N}}\} \text{ converges to } (\Phi_2 \circ \Phi_1)\,(\mathbf{q})$$

and, reiterating this reasoning, we obtain that

$$\{\Phi(\mathbf{q}_n)\}_{n\in\mathbb{N}} \text{ converges to } \Phi(\mathbf{q})$$

□

In the following corollary, we shall establish the sequential continuity of Φ.

Corollary 3 *If* $\{\mathbf{q}_n\}_{n\in\mathbb{N}}$ *converges to* $\mathbf{q}$ *in* $(\mathbb{D}, ||\ \ ||^*)$ *and* $\{\Phi(\mathbf{q}_n)\}_{n\in\mathbb{N}}$ *and* $\{\Phi(\dot{\mathbf{q}}_n)\}_{n\in\mathbb{N}}$ *are equicontinuous and uniformly bounded then:*

$$\{\Phi(\mathbf{q}_n)\}_{n\in\mathbb{N}} \textit{ converges to } \Phi(\mathbf{q}) \textit{ in } (\mathbb{D}, ||\ \ ||^*)$$

Proof Since $\{\Phi(\mathbf{q}_n)\}_{n\in\mathbb{N}}$ is equicontinuous and uniformly bounded, in virtue of Arzela-Ascoli's Theorem, there exists a subsequence $\{\mathbf{q}_{n_k}\}_{k\in\mathbb{N}}$ such that $\{\Phi(\mathbf{q}_{n_k})\}_{k\in\mathbb{N}}$ converges uniformly and as $\{\Phi(\dot{\mathbf{q}}_{n_k})\}_{k\in\mathbb{N}}$ is uniformly bounded and equicontinuous, by Lemma 4, we may conclude that $\{\Phi(\dot{\mathbf{q}}_{n_k})\}_{k\in\mathbb{N}}$ also converges uniformly.

Hence, in virtue of the Corollary 2 together with Lemma 1, we may state that $\{\Phi(\mathbf{q}_{n_k})\}_{k\in\mathbb{N}}$ converges to $\Phi(\mathbf{q})$. On the other hand, reasoning analogously, any other subsequence of $\{\Phi(\mathbf{q}_n)\}_{n\in\mathbb{N}}$ that converges must also converge to $\Phi(\mathbf{q})$ and, by Lemma 1, it may be concluded that $\{\Phi(\mathbf{q}_n)\}_{n\in\mathbb{N}}$ converges to $\Phi(\mathbf{q})$. □

In the following proposition, we establish the application framework of the extension of Zangwill's Theorem.

Proposition 5 *Let* $\mathbb{U} := \mathbb{D} \cap \widehat{C}^2$. *Then* $\exists M \in \mathbb{R}$ *such that, being* $\mathbb{U}_M := \{\mathbf{z} \in \mathbb{U}\,/\, ||\ddot{\mathbf{z}}||_\infty < M\}$, *it is verified that:*

i) $\Phi(\mathbb{U}_M) \subseteq \mathbb{U}_M$.
ii) $\mathbb{U}_M$ *is relatively sequentially compact in* $(\mathbb{D}, ||\ \ ||^*)$.
iii) $\Phi : (\mathbb{D}, ||\ \ ||^*) \longrightarrow (\mathbb{D}, ||\ \ ||^*)$ *is sequentially continuous.*
iv) $F : (\mathbb{D}, ||\ \ ||^*) \longrightarrow (\mathbb{R}, |\ \ |)$ *is sequentially continuous satisfying*

$$\Phi(\mathbf{x}) \neq \mathbf{x} \Longrightarrow F(\Phi(\mathbf{x})) < F(\mathbf{x})$$

Proof $\forall \mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{D}$ we have that

$$||\dot{p}_i||_\infty \leq M_i := \max\{||\varphi_i||_\infty, ||\psi_i||_\infty\} \Longrightarrow ||p_i||_\infty < N_i := M_i \cdot T$$

Let

$$\mathbb{O} := [0, T] \times \prod_{i=1}^{n} [-N_i, N_i] \times \prod_{i=1}^{n} [-M_i, M_i]$$

and let $f_i : [0, T] \times \mathbb{R}^{2m} \longrightarrow \mathbb{R}$ be the continuous functions associated with the i-th Euler equation of the functional $J^i_{\mathbf{p}}$

$$\ddot{z}_i(t) = f_i(t, p_1(t), \cdots, z_i(t), \cdots, p_m(t), \dot{p}_1(t), \cdots, \dot{z}_i(t), \cdots, \dot{p}_m(t))$$

Let

$$A_i := \max_{\mathbb{O}} f_i(t, x_1, \cdots, x_n, y_1, \cdots, y_n)$$

$$B_i := \max\{A_i, ||\dot{\varphi}_i||_\infty, ||\dot{\psi}_i||_\infty\}$$

$$M := \max_{i=1,\ldots,n} \{B_i\}$$

From all the above, it is clear that $||\{\overset{\bullet\bullet}{\Phi(\mathbf{p})}\}||_\infty \leq M$.

i) If $\mathbf{p} = (p_1, \cdots, p_m) \in \mathbb{U}_{\mathbb{M}} \Longrightarrow \Phi_i(\mathbf{p}) = (p_1, \cdots, p_i^*, \cdots, p_m)$, where p_i^* is $\widehat{C}^2$, since this satisfies the Euler equation of the functional $L^i_{\mathbf{q}}$ in certain intervals and $\dot{p}_i^*(t) = \psi_i(t)$ or $\dot{p}_i^*(t) = \varphi_i(t)$ in others, φ_i and ψ_i being of class C^1. Hence, $\phi_i(\mathbf{p}) \in \mathbb{U}_M$, $\forall i = 1, \cdots, m$, and, obviously, $\Phi(\mathbf{p}) \in \mathbb{U}_M$.

ii) Any sequence $\{\mathbf{p}_n\}_{n\in\mathbb{N}} \subset \mathbb{U}_M$ is uniformly bounded ($||\mathbf{p}_n||_\infty < \max\{N_i\}$) and so are all the sequences $\{\dot{\mathbf{p}}_n\}_{n\in\mathbb{N}}$ ($||\dot{\mathbf{p}}_n||_\infty < \max\{M_i\}$) and $\{\ddot{\mathbf{p}}_n\}_{n\in\mathbb{N}}$ ($||\ddot{\mathbf{p}}_n||_\infty < M$).
Thus, in virtue of Lemma 2, $\{\mathbf{p}_n\}_{n\in\mathbb{N}}$ and $\{\dot{\mathbf{p}}_n\}_{n\in\mathbb{N}}$ are equicontinuous. We are therefore in a situation to use Arzela-Ascoli's Theorem, which guarantees that there exists a subsequence $\{\mathbf{p}_{n_k}\}_{k\in\mathbb{N}}$ that converges uniformly to a certain $\mathbf{p} \in \mathbb{D}$. What's more, in virtue of Lemma 4, $\{\dot{\mathbf{p}}_{n_k}\}_{k\in\mathbb{N}}$ also converges uniformly to $\dot{\mathbf{p}}$ and, in short, $\{\mathbf{p}_{n_k}\}_{k\in\mathbb{N}}$ converges in the topological space $(\mathbb{D}, ||\ \ ||^*)$.

iii) Let the sequence $\{\mathbf{q}_n\}_{n\in\mathbb{N}}$ be convergent in $(\mathbb{D}, ||\ \ ||^*)$, i.e. $\{\mathbf{q}_n\}_{n\in\mathbb{N}}$ converges uniformly to $\mathbf{q}$ and $\{\dot{\mathbf{q}}_n\}_{n\in\mathbb{N}}$ converges uniformly to $\dot{\mathbf{q}}$.
As for each $n \in \mathbb{N}$, $\mathbf{q}_n \in \left(C^1[0, T]\right)^m$, in virtue of Proposition 1, $\Phi_i\left(\mathbf{q}_n\right) \in \left(C^1[0, T]\right)^m$. Thus, it is deduced that $\{\Phi_i\left(\mathbf{q}_n\right)\}_{n\in\mathbb{N}}$ is uniformly bounded and equicontinuous in $\mathbb{D}$.
The reiterative application of this reasoning guarantees that $\{\Phi\left(\mathbf{q}_n\right)\}_{n\in\mathbb{N}}$ is uniformly bounded and equicontinuous in $\mathbb{D}$.
Furthermore, as $\{\overset{\bullet}{\Phi_i(\mathbf{q}_n)}\}_{n\in\mathbb{N}} \in \left(C^0[0, T]\right)^n$, simply by taking the maximum of the bounds for each of these, $\{\overset{\bullet}{\Phi_i(\mathbf{q}_n)}\}_{n\in\mathbb{N}}$ is uniformly bounded.
On the other hand, following the considerations set out at the beginning of this proof, each i-th component, $\ddot{q}^*_{n,i}$ of $\{\overset{\bullet\bullet}{\Phi_i(\mathbf{q}_n)}\}_{n\in\mathbb{N}}$ is bounded. Hence

$\{\Phi(\ddot{\mathbf{q}}_n)\}_{n\in\mathbb{N}}$ is uniformly bounded. Taking into account Lemma 2, we may state that $\{\Phi(\dot{\mathbf{q}}_n)\}_{n\in\mathbb{N}}$ is equicontinuous in $\mathbb{D}$.

As $\{\Phi(\mathbf{q}_n)\}_{n\in\mathbb{N}}$ and $\{\Phi(\dot{\mathbf{q}}_n)\}_{n\in\mathbb{N}}$ are uniformly bounded and equicontinuous in $\mathbb{D}$, Corollary 3 ensures that Φ is sequentially continuous.

iv) If $\{\mathbf{p}_n\}_{n\in\mathbb{N}}$ converges to $\mathbf{p}\in\mathbb{U}$ with the topology $||\ \ ||^*$, making $\Gamma_{\mathbf{z}}(t) := L(t, \mathbf{z}(t), \mathbf{z}(t))$, we are in a situation to use Lemma 3, which guarantees that

$$\{\Gamma_{\mathbf{p}_n}\}_{n\in\mathbb{N}} \text{ converges uniformly to } \Gamma_{\mathbf{p}_n}$$

and, thus

$$\left\{F(\mathbf{p}_n) = \int_a^b \Gamma_{\mathbf{p}_n}(t)dt\right\}_{n\in\mathbb{N}} \text{ converges to } \int_a^b \Gamma_{\mathbf{p}}(t)dt = F(\mathbf{p})$$

□

Theorem 3 *For every $\mathbf{q}_0 \in \mathbb{U}_M$, the sequence generated by the algorithm $\{\mathbf{q}_n = \Phi(\mathbf{q}_{n-1})\}_{n\in\mathbb{N}}$ possesses a subsequence that converges in $(\mathbb{D}, ||\ \ ||^*)$ and the limit is a fixed point of Φ. Moreover, any convergent subsequence of $\{\mathbf{q}_n\}_{n\in\mathbb{N}}$ will converge at a fixed point on Φ.*

Proof It suffices to prove that, in fact, the sequence $\{\mathbf{q}_n\}_{n\in\mathbb{N}} \subset \mathbb{U}_M$ possesses a subsequence that converges uniformly and, as we have guaranteed the verification of the Global Theorem hypothesis, in virtue of Proposition 5, we may conclude that $\{\mathbf{q}_n\}_{n\in\mathbb{N}}$ possesses a subsequence that converges uniformly to a fixed point on Φ. By virtue of the Global Convergence Theorem itself, we have guaranteed that any convergent subsequence of $\{\mathbf{q}_n\}_{n\in\mathbb{N}}$ will converge to a fixed point on Φ.

We see that the sequence $\{\mathbf{q}_n\}_{n\in\mathbb{N}} \subset \mathbb{U}_M$ possesses a subsequence that converges uniformly.

For any $\mathbf{q}_0 \in \mathbb{U}_M$, from (i) in Proposition 5, we know that the sequence $\mathbf{q}_n = \Phi(\mathbf{q}_{n-1})$ is contained in $\mathbb{U}_M$ and following an analogous reasoning to point (ii) in this same proposition, we may conclude that there exists a subsequence $\{\mathbf{q}_{n_k}\}_{k\in\mathbb{N}}$ that converges in $\mathbb{D}$. □

5 Example

A program that resolves the optimization problem was elaborated using the Mathematica package and was then applied to one example of a hydrothermal system made up of 8 thermal plants and 3 hydro-plants. The cost function Ψ that was used is a quadratic model

$$\Psi(P) = \alpha + \beta P + \gamma P^2$$

and we consider Kirchmayer's model for the transmission losses: $l \cdot P^2$, where l is termed the loss coefficient. The units for the coefficients are: α_i in

Table 1 Coefficients of the thermal plants

Plant i	α_i	β_i	γ_i	l_i
1	1,227.83	17.621	0.01325	0.000103
2	743.78	20.842	0.00211	0.000072
3	77.72	21.277	0.00286	0.000172
4	1,615.35	16.676	0.01659	0.000100
5	2,248.16	−7.984	0.17026	0.000353
6	1,459.44	21.569	0.01489	0.000121
7	1,625.43	6.347	0.09803	0.000220
8	2,155.62	17.745	0.01982	0.000097
Equivalent	10,696.1	16.5477	0.00329982	

(Euro/h), β_i in (Euro/h Mw), γ_i in (Euro/h Mw^2), and l_i in (1/Mw). The data of the plants is summarized in Table 1. We construct the equivalent thermal plant as we saw in [15, 16], obtaining the values for α_{eq}, β_{eq} and γ_{eq}.

For the hydro-plants we use a variable head model and the i-th hydro-plant's active power generation P_{hi} is given by

$$P_{hi}(t, z_i(t), \dot{z}_i(t)) := A_i(t)\dot{z}_i(t) - B_i\dot{z}_i(t)\left[z_i(t) - \text{Coup}_i(t)\right]$$

where $A_i(t)$ and B_i are the coefficients

$$A_i(t) = \frac{1}{G_i} B_{y_i}(S_{0i} + t \cdot i_i); \quad B_i = \frac{B_{y_i}}{G_i}$$

and $\text{Coup}_i(t)$ represents the hydraulic coupling between plants. In the variable-head models, the term $-B_i\dot{z}_i(t)\left[z_i(t) - \text{Coup}_i(t)\right]$ represents the negative influence of the consumed volume, and reflects the fact that consuming water lowers the effective height and hence the performance of the plant. We consider that the transmission losses for the hydro-plant are also expressed by Kirchmayer's model. Hence, the function of effective hydraulic generation is

$$H_i(t, z_i(t), \dot{z}_i(t)) := P_{hi}(t, z_i(t), \dot{z}_i(t)) - l_i P_{hi}^2(t, z_i(t), \dot{z}_i(t))$$

The hydro-network is assumed to have two chains of hydro-plants on different rivers. Hydro-plant 1 is a isolated plant, whereas we assume that the rate of discharge at the upstream plant 2 affects the behavior at the downstream plant 3. Thus, $\text{Coup}_1(t) = 0$; $\text{Coup}_2(t) = 0$; $\text{Coup}_3(t) = z_2(t)$. The units for the coefficients of the hydro-plant are: the efficiency G_i in (m^4/h Mw), the constraint on the volume b_i in (10^7 m^3), the loss coefficient l_i in (1/Mw), the natural inflow i_i in (m^3/h), the initial volume S_{0_i} in (m^3) and the coefficient B_{y_i}

Table 2 Hydro-plant coefficients

Plant	G_i	b_i	l_i	i_i	S_{0_i}	B_{y_i}
1	534,660	1.416	0.0	0.0	193.885 10^9	150.1
2	526,315	3.958	0.00022	98.176 10^5	203.904 10^9	149.5
3	570,834	1.912	0.00016	301.952 10^6	407.808 10^8	138.7

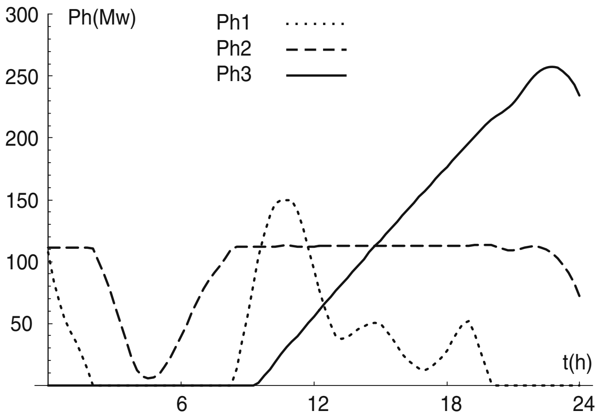

Fig. 1 Optimal power $P_h(t)$ for the hydro-plants

(a parameter that depends on the geometry of the tanks) in (10^{-12} m^{-2}). The data of the hydro-plants is summarized in Table 2.

Moreover, we consider the rates of water to be bound by the following values: $\dot{z}_{i\,\min} = 0$ and $\dot{z}_{i\,\max} = 2.75\ 10^6, \forall i = 1, 2, 3$. We consider a short-term hydrothermal scheduling (24 h) with an optimization interval [0, 24] and we consider a discretization of 96 subintervals. The optimal power for the hydro-plants, $P_h(t)$, is shown in Fig. 1, and the system's power demand, $P_d(t)$, and the optimal power for the equivalent thermal plant, $P_{th}(t)$, in Fig. 2.

We considered very distinct hydraulic models with the aim of achieving an optimum solution for the diverse power stations that is very different depending on their technical characteristics. Thus, for instance, it can be seen that plant 3, whose natural inflow is very high, delays its participation in order to employ the available water more efficiently. At the same time, for one of the power plants (plant 1), we considered a hydraulic model without transmission losses or natural inflow. The presence of plant 1 guarantees constancy in the optimum thermal power. This fact, the proof of which is similar to the one presented in [21] for linear models, enables the aspect of

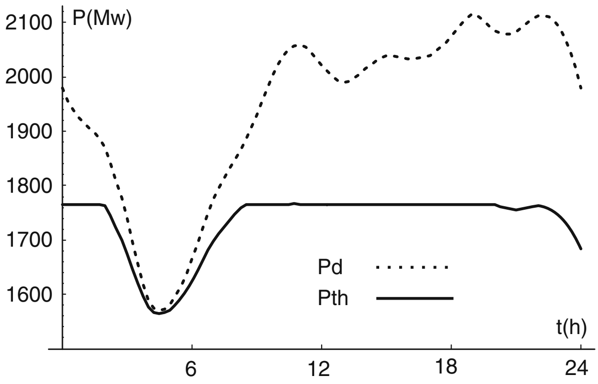

Fig. 2 Power demand $P_d(t)$ and optimal thermal power P_{th}(t)

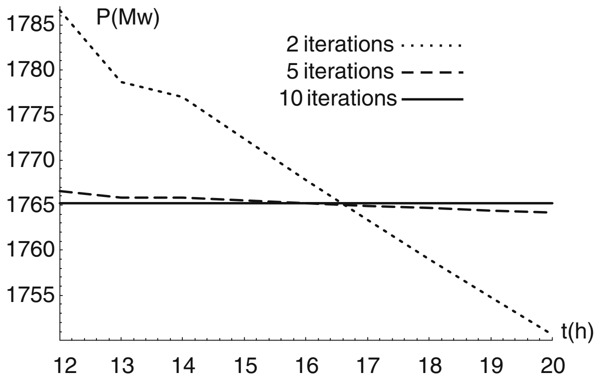

Fig. 3 Optimal thermal power $P_{\mathrm{th}}(t)$

the theoretical solution to be known a priori and facilitates a better analysis of the approximate solution.

It can be seen in Fig. 3 that, in fact, the resulting thermal power after the successive iterations approximates an increasingly more constant performance in the interval in which the constraints for the rates of water are not active. Another two tests were conducted to verify the convergence of the algorithm, considering in both the same eight thermal plants and six and ninehydro-plants situated in two and three river basins, respectively, and repeating for each of these the hydraulic configuration from the above example. Figure 4 presents the obtained results.

The vector $\mathbf{K}^n = (K_1, \ldots, K_m)$ was considered as the stopping criterion for the algorithm in each iteration, the components of which are the coordination constants associated with the different hydro-plants, the tolerance being defined as

$$Tol(n) = \left\| \mathbf{K}^n - \mathbf{K}^{n-1} \right\|$$

We can see how the method presents a rapid convergence. For example, for the case of the 3 hydro-plants, the time required by the program was 300 s on a personal computer (Pentium IV/2GHz).

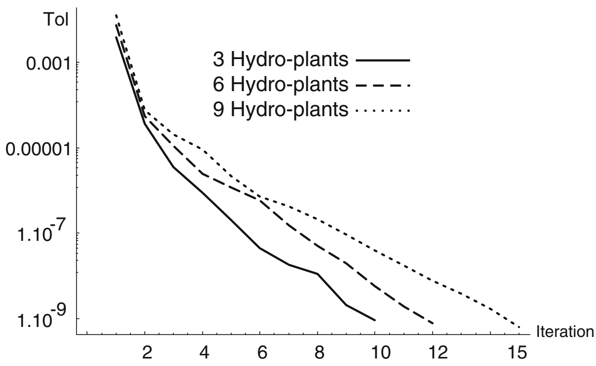

Fig. 4 Convergence of the algorithm

6 Conclusions

In this paper we study a classical Electrical Engineering problem: the short-term hydrothermal coordination (STHC) problem. First, we give a necessary minimum condition using optimal control theory (Pontryagin's Principle). Then we describe a relaxation method for computing the solution, and give a proof of convergence. We solve the optimization problem using a computationally efficient numerical algorithm based on coordinate descent. We show qualitative and quantitative results demonstrating the effectiveness of the proposed method for the STHC problem. Future research includes the generalization to constraints of the type:

$$0 \leq H(t, \mathbf{z}(t), \dot{\mathbf{z}}(t)) \leq P_{\mathrm{d}}(t), \forall t \in [0, T]$$

and we may even consider non-regular Langrangian, formulating the problem within the framework of nonsmooth analysis.

References

1. Mwakabuta, N., Kyaruzi, A.: Short-term scheduling of a hydro-thermal system using linear programming technique. In: Proceedings of the 6th International Power Engineering Conference, pp. 731–736 (2003)
2. Baskar, S., Subbaraj, P., Rao, M.V.C.: Hybrid evolutionary programming solution to short term hydro thermal scheduling problem. J. Inst. Eng. India, Electr. Eng. Div. **82**, 236–242 (2002)
3. Gil, E., Bustos, J., Rudnick, H.: Short-term hydrothermal generation scheduling model using a genetic algorithm. IEEE Trans. Power Syst. **18**(4), 1256–1264 (2003)
4. Zoumas, C.E., Bakirtzis, A.G., Theocharis, J.B., Petridis, V.: A genetic algorithm solution approach to the hydrothermal coordination problem. IEEE Trans. Power Syst. **19**(3), 1356–1364 (2004)
5. Salam, M.S.: Comparison of Lagrangian relaxation and truncated dynamic programming methods for solving hydrothermal coordination problems. In: Proceedings of International Conference on Intelligent Sensing and Information Processing, pp. 265–270 (2004)
6. Salam, M.S., Nor, K.M., Hamdan, A.R.: Hydrothermal scheduling based lagrangian relaxation approach to hydrothermal coordination. IEEE Trans. Power Syst. **13**(1), 226–235 (1998)
7. Redondo, N.J., Conejo, A.J.: Short-term hydro-thermal coordination by Lagrangian relaxation: solution of the dual problem. IEEE Trans. Power Syst. **14**(1), 89–95 (1999)
8. Tseng, P., Bertsekas, D.: Relaxation methods for problems with strictly convex costs and linear constraints. Math. Oper. Res. **16**(3), 462–481 (1991)
9. Luo, Z.Q., Tseng, P.: On the convergence of the coordinate descent method for convex differentiable minimization. J. Optim. Theory Appl. **72**(1), 7–35 (1992)
10. Tseng, P.: Convergence of a block coordinate descent method for nondifferentiable minimization. J. Optim. Theory Appl. **109**(3), 475–494 (2001)
11. Rätsch, G., Mika, S., Warmuth, M.: On the convergence of leveraging. In: Dietterich, T.G., Becker, S., Ghahramani, Z. (eds.), Advances in Neural Information Processing Systems 14. MIT, Cambridge, MA (2002)
12. Rätsch, G., Warmuth, M., Mika, S., Onoda, T., Lemm, S., Müller, K.: Barrier boosting. In: Proceedings of the Thirteenth Annual Conference on Computational Learning Theory, pp. 170–179 (2000)
13. Chambolle, A., Lions, P.L.: Image recovery via total variation minimization and related problems. Numer. Math. **76**(2), 167–188 (1997)

14. Fan, A., Wells, W., Fisher, J., Çetin, M., Haker, S., Mulkern, R., Tempany, C., Willsky, A.: A unified variational approach to denoising and bias correction in MR. Lecture Notes in Computer Science, vol. 2732, pp. 148–159 (2003)
15. Bayón, L., Grau, J.M., Suárez, P.M.: A new formulation of the equivalent thermal in optimization of hydrothermal systems. Math. Probl. Eng. **8**(3), 181–196 (2002)
16. Bayón, L., Grau, J.M., Ruiz, M.M., Suárez, P.M.: New developments on equivalent thermal in hydrothermal optimization: an algorithm of approximation. J. Comput. Appl. Math. **175**(1), 63–75 (2005)
17. Santos, M.S., Vigo-Aguiar, J.: Analysis of a numerical dynamic programming algorithm applied to economic models. Econometrica **66**(2), 409–426 (1998)
18. Santos, M.S., Vigo-Aguiar, J.: Error bounds for a numerical solution for dynamic economic models. Appl. Math. Lett. **9**(4), 41–45 (1996)
19. Bayón, L., Grau, J.M., Suárez, P.M.: A necessary condition for broken extremals in problems involving inequality constraints. Archives Inequalities Appl. **1**(1), 75–84 (2003)
20. Zangwill, W.L.: Nonlinear Programming: a Unified Approach. Prentice Hall, New Jersey (1969)
21. Bayón, L., Grau, J.M., Ruiz, M.M., Suárez, P.M.: Nonsmooth optimization hydrothermal problem. In: Proceedings of the Computational and Mathematical Method on Sciences and Engineering, vol. 1, pp. 116–125 (2004)