JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 117, No. 2, pp. 377–396, May 2003 (© 2003)

On the Application of the Auxiliary Problem Principle[1]

A. Losi[2] and M. Russo[3]

Communicated by M. A. Simaan

Abstract. The auxiliary problem principle (APP) derives from a general theory on decomposition-coordination methods establishing a comprehensive framework for both one-level and two-level methods. In this paper, the results of the two-level methods of APP are specialized for an efficient application to some engineering problems.

Key Words. Optimization algorithms, decomposition/coordination methods, auxiliary problem principle.

1. Introduction

The solution of large-scale optimization problems can be tackled effectively within a parallel/distributed computing environment. The decomposition-coordination approach has proven to be useful to exploit such environment; the methods deriving from such an approach can be classified into one-level or fixed-point methods, in which there is no explicit coordinating action between the subproblems, and two-level methods, in which the subproblems are driven toward the solution thanks to prices which are updated at the coordination level.

The auxiliary problem principle (APP, Refs. 1–2) derives from a general theory on the decomposition-coordination approach. It establishes a comprehensive framework for both one-level and two-level methods; it encompasses many known algorithms as particular cases. For the APP, convergence proofs have been established under reasonable hypotheses, in particular for two-level methods.

[1]This research was supported by MIUR and CNR, Rome, Italy.

[2]Professor, Dipartimento di Ingegneria Industriale, Università di Cassino, Cassino, Frosinone, Italy.

[3]Associate Professor, Dipartimento di Ingegneria Industriale, Università di Cassino, Cassino, Frosinone, Italy.

Using the well-known technique of variable duplication, it is always possible to obtain an optimization problem whose feasibility set can be decomposed easily into the Cartesian product of feasibility subsets, except for some linear equalities. It is then easy to decompose the problem into subproblems coupled only by linear equalities, the perfect candidate for the application of the APP two-level methods. In some applications, such as power system optimization problems, a difficulty arises with the duplication of some variables which appear always in difference form; let them be called Δ-variables. A suitable way for modeling the Δ-variables for the subsequent duplication/decomposition has been envisaged in a previous paper (Ref. 3), and then the straightforward application of the two-level methods of APP is possible.

In this paper, specific results are developed for APP with Δ-variables so as to improve the characteristics of its iterative scheme. We begin by recalling the technique of variable duplication and its use within APP and then show the modeling of Δ-variables. Next, the peculiarities of this modeling are exploited to obtain a specific form of the iterative scheme which has some relevant features, such as an easy formulation of the feasibility constraints of the subproblems with an independent treatment of the Δ-variables. The results obtained for a large-scale power system optimization problem show the effectiveness of the new form of the iterative scheme. The proofs on the sufficient conditions for the convergence of APP methods that include these variables are reported in the Appendix (Section 6).

2. Duplication of Variables and APP

Any optimization problem can be put in the general form

$$\min \quad f(x), \tag{1a}$$

$$\text{s.t.} \quad g(x) = 0, \tag{1b}$$

$$h(x) \geq 0, \tag{1c}$$

where x is the m-vector of variables, $f(x)$ is a scalar function representing the objective of the optimization, $g(x)$ and $h(x)$ are s-vector, $s < m$, and t-vector functions of equality and inequality constraints defining the feasibility constraints for the vector x. In general, $f(x)$ and any component of $g(x)$ or $h(x)$ may be linear or nonlinear and may depend on some or all the components of the vector x.

2.1. Duplication of Variables. Let the problem (1) be put in the form

$$\min \quad f(x_1, x_2, y), \tag{2a}$$

$$\text{s.t.} \quad g_1(x_1, y) = 0, \tag{2b}$$

$$g_2(x_2, y) = 0, \tag{2c}$$

$$h_1(x_1, y) \geq 0, \tag{2d}$$

$$h_2(x_2, y) \geq 0, \tag{2e}$$

where

$$x = [x_1^T, x_2^T, y^T]^T, \qquad g = \{g_1, g_2\}, \qquad h = \{h_1, h_2\},$$

and

$$\dim\{x_1\} = m_1, \qquad \dim\{x_2\} = m_2, \qquad \dim\{y\} = m_y,$$

$$\dim\{g_1\} = s_1, \qquad \dim\{g_2\} = s_2, \qquad \dim\{h_1\} = t_1, \qquad \dim\{h_2\} = t_2.$$

In the relevant case that the subset y is not empty ($m_y \neq 0$), the feasibility constraints subsets $\{g_1, h_1\}$, $\{g_2, h_2\}$ are coupled. In such a case, with the technique of duplication of variables, it is always possible to obtain a new optimization problem, equivalent to (2), with the desirable characteristic that the feasibility constraint subsets are decoupled and the original coupling is enforced via the so-called consistency constraints. It is obtained at the expense of increasing dimensionality: the new optimization problem is defined in a space whose dimensionality is $m + m_y$, and m_y linear consistency constraints are added.

In short, the optimization problem (2), with a consistent new formulation of the objective function, now represented by the function J, can be written as

$$\min \quad J(u_1, u_2) \equiv \min J(x_1, x_2, y_{12}, y_{21}), \tag{3a}$$

$$\text{s.t.} \quad g_1(u_1) \equiv g_1(x_1, y_{12}) = 0, \tag{3b}$$

$$g_2(u_2) \equiv g_2(x_2, y_{21}) = 0, \tag{3c}$$

$$h_1(u_1) \equiv h_1(x_1, y_{12}) \geq 0, \tag{3d}$$

$$h_2(u_2) \equiv h_2(x_2, y_{21}) \geq 0, \tag{3e}$$

$$\Theta(u_1, u_2) \equiv y_{12} - y_{21} = 0, \tag{3f}$$

where u_1 and u_2 are

$$u_1 = \begin{bmatrix} x_1 \\ y_{12} \end{bmatrix}, \qquad u_2 = \begin{bmatrix} x_2 \\ y_{21} \end{bmatrix}, \tag{4}$$

y_{12} is the m_y-vector of the optimization variables that represent the overlapping variables y in the first subset u_1, y_{21} is the m_y-vector of the same variables y in the second subset u_2, and $\Theta(u_1, u_2)$ is the m_y-vector function of the consistency constraints. A compact form for writing problem (3) is

$$\min_{\substack{u_1 \in U_1 \\ u_2 \in U_2}} \quad J(u_1, u_2), \tag{5a}$$

$$\text{s.t.} \quad \Theta(u_1, u_2) = 0, \tag{5b}$$

where

$$U_1 = \{u_1 : g_1(u_1) = 0, h_1(u_1) \geq 0\}, \tag{6a}$$

$$U_2 = \{u_2 : g_2(u_2) = 0, h_2(u_2) \geq 0\}. \tag{6b}$$

In the formulation (5), feasibility constraints $\{g_i, h_i\}$, $i = 1, 2$, of the formulation (3) are seen as implicit to the feasible set U_i, $i = 1, 2$; the only explicit constraints Θ are equalities that represent the consistency of the duplicated variables and have the simple form required by the two-level APP methods for problems with explicit equality constraints (Refs. 1–2).

In the general case of partition of the feasibility constraint set into n subsets, problem (5) can be put in the form

$$\min_{\substack{u_i \in U_i \\ i = 1, \ldots, n}} \quad J(u_1, \ldots, u_n), \tag{7a}$$

$$\text{s.t.} \quad \Theta_{ij}(u_i, u_j) = 0, \qquad i, j = 1, \ldots, n \text{ and } j > i. \tag{7b}$$

2.2. Auxiliary Problem Principle (APP). In the APP, problem (7) with a weakly convex objective function is considered, and the augmented Lagrangian method is used (Ref. 2). Any nonseparable term in the augmented Lagrangian is linearized so as to obtain an approximation which is additive with respect to any decomposition of the variable space; strictly convex terms are added, and auxiliary problems are obtained whose solutions can be driven to that of the original problem. The reader is referred to Refs. 1–2 for more details on the APP.

According to the APP, problem (7) can be solved with the following two-level iterative scheme [$\langle \cdot, \cdot \rangle$ stands for scalar product]:

$$\min_{\substack{u_i \in U_i \\ i = 1, \ldots, n}} \left\{ \begin{array}{l} K(u_1, \ldots, u_n) + \sum\limits_{i=1}^{n} \langle \epsilon J'_{u_i}(u_1^k, \ldots, u_n^k) - K'_{u_i}(u_1^k, \ldots, u_n^k), u_i \rangle \\ + \epsilon \sum\limits_{i=1}^{n} \sum\limits_{j=i+1}^{n} \langle p_{ij}^k + c\Theta_{ij}(u_i^k, u_j^k), \Theta_{ij}(u_i, u_j) \rangle \end{array} \right\}$$

$$\Rightarrow u_1^{k+1}, \ldots, u_n^{k+1}, \tag{8a}$$

$$p_{ij}^{k+1} = p_{ij}^k + \rho \Theta_{ij}(u_i^{k+1}, u_j^{k+1}), \qquad i, j = 1, \ldots, n \text{ and } j > i. \tag{8b}$$

Here, the function $K(u_1, \ldots, u_n)$ is the so-called core function, k is the iteration counter, and p_{ij} is a vector of Lagrange multipliers associated with the consistency constraints Θ_{ij} for the duplicated variables. At the kth iteration, the first-level (auxiliary) problem (8a) is solved: note that it has no explicit constraints; the value of the variables thus obtained is used in the second-level (8b), where the new values of the Lagrange multipliers p_{ij} are obtained.

The (global) convergence of the iterative scheme (8) is guaranteed under the following conditions (Ref. 2):

U_i, $i = 1, \ldots, n$, are closed convex sets;
J is convex and its derivative is Lipschitz with constant A;
Θ is affine, Lipschitz with constant τ;
K is strongly convex, with constant b and with Lipschitz derivative;
$0 < \rho < 2c$;
$0 < \epsilon < b/(A + c\tau^2)$.

In practical nonconvex cases, the iterative scheme can still be used, since the necessary optimality conditions are met in the limit, if convergence results (Refs. 1–2).

Choosing a core function additive with respect to $(u_1, \ldots, u_n)$,

$$K(u_1, \ldots, u_n) = \sum_{i=1}^{n} K_i(u_i), \tag{9}$$

and recalling the form (3f) of Θ, in the kth iteration, the first-level problem (8a) can be split into n subproblems that can be solved in parallel; in fact, (8) can be written as

$$\min_{u_i \in U_i} \left\{ \begin{array}{l} K_i(u_i) + \langle \epsilon J'_{u_i}(u_1^k, \ldots, u_n^k) - K'_{u_i}(u_i^k), u_i \rangle \\ + \epsilon \sum_{j=i+1}^{n} \langle p_{ij}^k + c(y_{ij}^k - y_{ji}^k), y_{ij} \rangle \\ + \epsilon \sum_{j=1}^{i-1} \langle p_{ji}^k + c(y_{ji}^k - y_{ij}^k), -y_{ij} \rangle \end{array} \right\} \Rightarrow u_i^{k+1}, \quad i = 1, \ldots, n, \tag{10a}$$

$$p_{ij}^{k+1} = p_{ij}^k + \rho(y_{ij}^{k+1} - y_{ji}^{k+1}), \qquad i, j = 1, \ldots, n \text{ and } j > i, \tag{10b}$$

where the variables y_{ij} have the above-mentioned meaning [see (3)].

3. Δ-Variables

In many applications (for example, problems which include the elapsed time as a variable, problems in power systems such as the optimal power flow problem and the state estimation problem), there are variables (let us call them Δ-variables) that appear always in difference form. To have a consistent formulation of the optimization problem (1), one of the Δ-variables must be given a constant value; it is usually set equal to zero, which makes the chosen Δ-variable the reference for the others.

An example may help understanding the issue. The feasibility constraints in an electrical network are described in terms of nodal voltages, line currents, and powers injected into the nodes:

$$P_h - V_h \sum_N V_k Y_{hk} \cos(\delta_h - \delta_k - \Phi_{hk}) = 0, \tag{11a}$$

$$Q_h - V_h \sum_N V_k Y_{hk} \sin(\delta_h - \delta_k - \Phi_{hk}) = 0, \tag{11b}$$

$$I_t^M - \sqrt{V_h^2 + V_k^2 - 2V_h V_k \cos(\delta_h - \delta_k)}/z_t \geq 0. \tag{11c}$$

In (11), V_h is the voltage amplitude at the hth node and δ_h is its phase; P_h and Q_h are the active and reactive powers injected into the hth node, respectively; I_t^M is the maximum value of the current that can flow along the tth line (that connects the hth node to the kth node); the constant matrices Y, Φ and the vector z depend on the physical characteristics of the lines, and N is the number of network nodes. The voltage phases δ appear always as differences; one component of δ is given the constant value zero, and the corresponding nodal voltage becomes the phase reference for all the voltages.

3.1. Duplication of Δ-Variables. To apply the iterative scheme (10), it is necessary that each subproblem (10a) be defined completely by its own feasibility constraint set U_i. This means that, in each subproblem, the Δ-variables must be given a reference which is local to that subproblem; in turn, the local reference of each subproblem must assume a coherent value against the global reference (Ref. 3).

Once some Δ-variables have been duplicated, the corresponding consistency constraints are written in the form

$$\Delta_i^r = \Delta_j^s, \tag{12}$$

which indicates that the rth Δ-variable of the ith subproblem and the sth Δ-variable of the jth subproblem are coupled because of the duplication. In (12), the Δ-variables are expressed with respect to the global reference.

Let symbol $\tilde{\Delta}_i$ denote the value of the Δ-variables of the ith subproblem with respect to the local reference; let ω_i represent the value of the ith local reference with respect to the global reference; we have (Ref. 3):

$$\tilde{\Delta}_i = \Delta_i - v\omega_i, \tag{13}$$

where v is a unitary vector of appropriate dimension. Using (13), the consistency constraints (12) can be written as

$$\tilde{\Delta}_i^r + \omega_i = \tilde{\Delta}_j^s + \omega_j. \tag{14}$$

Let us write the subproblems in terms of the local Δ-variables and assume that the nth subproblem is the one whose local reference for the Δ-variables coincides with the global one; problem (7) can be put in the form

$$\min_{\substack{u_i \in U_i,\ \omega_i \in \Re \\ i=1,\dots,n,\ i=1,\dots,n-1}} \quad J(u_1, \dots, u_n), \tag{15a}$$

$$\text{s.t.} \quad \Theta_{ij}(u_i, \omega_i, u_j, \omega_j) = 0, \qquad i,j = 1,\dots,n \text{ and } j > i, \tag{15b}$$

where ω_n is a constant equal to zero and in u_i the Δ-variables are local to the ith subproblem.

Note that, expressing the feasibility constraints for all subproblems $(U_i, i = 1, \dots, n)$ in terms of their local Δ-variables $(\tilde{\Delta}_i)$, the variables ω appear only in the coupling constraints Θ; they do not appear in the objective function nor in the feasibility sets, which are well defined independently of the variables ω.

As an example, consider the nodal power injection balances and the line currents limits (11) for a power systems subproblem. In (11), the voltage phases δ are evaluated against the global reference; due to (13), Eq. (11) can be rewritten as follows, where the absence of the variable ω_i is apparent:

$$\begin{aligned} &P_h - V_h \sum_N V_k Y_{hk} \cos((\tilde{\delta}_h + \omega_i) - (\tilde{\delta}_k + \omega_i) - \Phi_{hk}) \\ &= P_h - V_h \sum_N V_k Y_{hk} \cos(\tilde{\delta}_h - \tilde{\delta}_k - \Phi_{hk}) = 0, \end{aligned} \tag{16a}$$

$$\begin{aligned} &Q_h - V_h \sum_N V_k Y_{hk} \sin((\tilde{\delta}_h + \omega_i) - (\tilde{\delta}_k + \omega_i) - \Phi_{hk}) \\ &= Q_h - V_h \sum_N V_k Y_{hk} \sin(\tilde{\delta}_h - \tilde{\delta}_k - \Phi_{hk}) = 0, \end{aligned} \tag{16b}$$

$$\begin{aligned} &I_t^M - \sqrt{V_h^2 + V_k^2 - 2V_h V_k \cos((\tilde{\delta}_h + \omega_i) - (\tilde{\delta}_k + \omega_i))}/z_t \\ &= I_t^M - \sqrt{V_h^2 + V_k^2 - 2V_h V_k \cos(\tilde{\delta}_h - \tilde{\delta}_k)}/z_t \geq 0. \end{aligned} \tag{16c}$$

3.2. APP and Δ-Variables. In this section, we specialize the iterative scheme (10) to the case of the problem (15). In general, in the problem (15), some coupling constraints represent the consistency of the duplicated Δ-variables, some other constraints represent the consistency of the duplicated non-Δ-variables; for the sake of conciseness, but without loss of generality, in the following we assume that all the coupling constraints express the consistency of the duplicated Δ-variables and have the form

$$\Theta(u_i, \omega_i, u_j, \omega_j) \equiv (u_i + \nu\omega_i) - (u_j + \nu\omega_j) = 0. \tag{17}$$

Suppose that the convexity hypotheses on the feasibility sets $U_i, i = 1, \ldots, n$, and the objective function J are met; note that the constraints Θ are affine (indeed, they are linear). Let us solve problem (15), (17) with the APP and adopt the following core function:

$$K(u, \omega) = K_U(u) + K_\Omega(\omega) = \sum_{i=1}^{n} K_{U_i}(u_i) + \sum_{i=1}^{n-1} K_{\Omega_i}(\omega_i), \tag{18}$$

where

$$u = [u_1^T, u_2^T, \ldots, u_n^T]^T, \quad \omega = [\omega_1, \omega_2, \ldots, \omega_{n-1}]^T. \tag{19}$$

The iterative scheme (10) can be put in the following form:

$$\min_{u_i \in U_i} \left\{ \begin{aligned} &K_{U_i}(u_i) + \langle \epsilon J'_{u_i}(u_1^k, \ldots, u_n^k) - K'_{U_i}(u_i^k), u_i \rangle \\ &+ \epsilon \sum_{j=i+1}^{n} \langle p_{ij}^k + c\,[(y_{ij}^k + \nu\omega_i^k) - (y_{ji}^k \nu\omega_j^k)], y_{ij} \rangle \\ &+ \epsilon \sum_{j=1}^{i-1} \langle p_{ji}^k + c\,[(y_{ji}^k + \nu\omega_j^k) - (y_{ij}^k + \nu\omega_i^k)], -y_{ij} \rangle \end{aligned} \right\}$$

$$\Rightarrow u_i^{k+1}, \qquad i = 1, \ldots, n, \tag{20a}$$

$$\min_{\omega_i \in R} \left\{ \begin{aligned} &K_{\Omega_i}(\omega_i) - \langle K'_{\Omega_i}(\omega_i^k), \omega_i \rangle \\ &+ \epsilon \sum_{j=i+1}^{n} \langle p_{ij}^k + c[(y_{ij}^k + \nu\omega_i^k) - (y_{ji}^k + \nu\omega_j^k)], \nu\omega_i \rangle \\ &+ \epsilon \sum_{j=1}^{i-1} \langle p_{ji}^k + c[(y_{ji}^k + \nu\omega_j^k) - (y_{ij}^k + \nu\omega_i^k)], -\nu\omega_i \rangle \end{aligned} \right\}$$

$$\Rightarrow \omega_i^{k+1}, \qquad i = 1, \ldots, n-1, \tag{20b}$$

$$p_{ij}^{k+1} = p_{ij}^k + \rho[(y_{ij}^{k+1} + \nu\omega_i^{k+1}) - (y_{ji}^{k+1} + \nu\omega_j^{k+1})],$$

$$i, j = 1, \ldots, n \text{ and } j > i. \tag{20c}$$

Again, ω_n is equal to zero.

The formulation (20) has many relevant features. First, the feasibility set of each subproblem (20a) is expressed in terms of local variables [for example, see (16)]; it helps formulating the subproblems, avoiding the intricacies of treating local and global references. Second, in the kth step of the iterative algorithm, subproblems (20a) and (20b) are solved independently;

existing software solving optimization problems without any splitting can be used to solve any subproblem, provided that the objective function is changed. Third, with an appropriate choice of the functions $K_{\Omega_i}(\omega_i)$, the solution of problem (20b) has a closed form; for example, if

$$K_{\Omega}(\omega) = (1/2)\sum_{i=1}^{n-1} \sigma_i \omega_i^2, \tag{21}$$

we have

$$\omega_i^{k+1} = \omega_i^k - (\epsilon/\sigma_i)\begin{bmatrix} \sum_{j=i+1}^{n} \{p_{ij}^k + c[(y_{ij}^k + v\omega_i^k) - (y_{ji}^k + v\omega_j^k)]\} \\ -\sum_{j=1}^{i-1} \{p_{ji}^k + c[(y_{ji}^k + v\omega_j^k) - (y_{ij}^k + v\omega_i^k)]\} \end{bmatrix},$$

$$i = 1, \ldots, n-1. \tag{22}$$

The proof on the sufficient conditions for the convergence of the APP including variables ω is reported in the Appendix (Section 6). There, the general case of the consistency constraints for the duplication of the Δ-variables and non-Δ-variables is considered; it is shown that the conditions for the scheme (20) to converge to a saddle point of the problem (15), (17) are listed below (see Section 2.2):

U_i, $i = 1, \ldots, n$, are closed convex sets;
J is convex, and its derivative is Lipschitz with constant A;
Θ is affine, Lipschitz with constants τ_u and τ_ω related to the variables u and variables ω, respectively;
$K_U(u)$ is strongly convex, with constant β and Lipschitz derivative;
$K_\Omega(\omega)$ is strongly convex, with constant γ and Lipschitz derivative;
$0 < \rho < 2c$;
$0 < \epsilon < \beta/(A + c\tau_u^2)$;
$0 < \epsilon < \gamma/c\tau_\omega^2$.

It is worth nothing that, without the special consideration of the variables ω, the convergence conditions would be the ones reported in Section 2.2; in these conditions, the quantity b would be the strict convexity constant of $K(u, \omega)$ in (18), namely,

$$b = \min\{\beta, \gamma\}, \tag{23}$$

and the Lipschitz constant τ for the constraints Θ would be

$$\tau = \max\{\tau_u, \tau_\omega\}. \tag{24}$$

It is apparent that, if (23) and (24) hold, the convergence conditions of Section 2.2 would be more stringent than those given above.

4. Study Case

We study the case of the optimal power flow (OPF, Ref. 4) solution for a real-size network, the 118-bus IEEE test systems (Ref. 5). Although only the base-case security-constrained OPF is considered, the optimization problem is a large scale problem. The split of the system into two and three subsystems, and thus the split of the optimization problem into two and three subproblems, are considered; the modeling of the Δ-variables requires the introduction of one and two variables ω, respectively. Table 1 shows the characteristic of the basic problem and those of the subproblems in the cases analyzed. The objective function J reflects the power generation costs and is convex; the feasibility constraints are of the same type as (11), besides other inequalities not involving phase angles.

For validating the theoretical results of Section 3, several cases of OPF have been solved, using a software tool (Ref. 6) based on APP and named DistOpt. The tool, for which a specific OPF module has been written (Ref. 3), is useful to model easily and evaluate the solution of optimization problems in a distributed environment; with a suitable modification, it has been made capable of treating Δ-variables according to the scheme (20).

Numerical experiments have been carried out with different values of the parameters of the iterative scheme; in all of them, the following core function [see (18)] has been adopted:

$$K(u, \omega) = (\beta/2)\|u\|^2 + (\gamma/2)\|\omega\|^2. \tag{25}$$

Table 2 reports the values of the constants A and τ^2 (in our problem $\tau_u = \tau_\omega$) and the values of the parameters of the iterative scheme that have been adopted in the numerical experiments. Note that, in Cases 1 and 4, all the parameters are given values inside the bounds established in Section 3.2. In Cases 2 and 5, the parameter ϵ exceeds one of the bounds, while the parameter ρ is given a low value. In Cases 3 and 6, the parameters β and γ are given the same value, as one would do without the special consideration

Table 1. Main characteristics of the OPF problem and subproblems.

	Variables	Equality constraints	Functional inequality constraints	Box constraints	Variables ω	Consistency constraints
Base case	303	236	200	186	0	0
Split into 2 subsystems	129/197	98/146	71/129	72/114	1	16
Split into 3 subsystems	129/78/130	98/54/96	71/50/79	72/38/76	2	24

Table 2. Problem constants and modified APP iterative scheme parameters.

A	No. of subsystems	τ^2	Case	c	ρ	β	ϵ	γ	Results
1.641	2	10	#1	1.00	1.90	1.00	0.080	0.83	Figs. 1–3
1.641	2	10	#2	1.00	0.18	1.00	0.830	8.31	Figs. 1–3
1.641	2	10	#3	1.00	0.18	1.00	0.830	1.00	Failure
1.641	3	9.732	#4	1.00	1.90	1.00	0.083	0.83	Figs. 4–7
1.641	3	9.732	#5	1.00	0.18	1.00	0.830	8.30	Figs. 4–7
1.641	3	9.732	#6	1.00	0.18	1.00	0.830	1.00	Failure

of the variables ω, while the parameters ϵ and ρ have the same values as in Cases 2 and 5. In all the cases, appropriate starting points have been used, as is usually done in OPF studies.

It would be a practical impossibility to show in detail all the results. The quantity E, which is the maximum discrepancy between the values of the duplicated variables,

$$E = \max_i |\Theta_i|, \tag{26}$$

has been used as a convergence indicator. In Fig. 1, the plots of the quantity E versus the iteration counter are reported for Cases 1 and 2; for the same cases, Fig. 2 shows the objective function J, and Fig. 3 shows the variable ω_1. For Cases 4 and 5, Fig. 4 shows the plots of the quantity E, Fig. 5 shows the objective function J, Fig. 6 shows the variable ω_1, and Fig. 7 shows the variable ω_2, all versus the iteration counter.

Supplementary experiments (not reported) have shown that the whole set of convergence conditions established within APP, although only sufficient, is not much restrictive, at least for the study case. Nevertheless, a better convergence performance (see Cases 2 and 5) can result by giving some of the parameters of the iterative scheme a value well inside the bound and others a value exceeding the bound. The possibility of exceeding the bounds and still getting convergence can be explained by the sufficient nature of the convergence conditions.

To summarize, it has been seen that the bounds on the parameters of the iterative scheme provided by the sufficient convergence conditions guarantee the convergence of the scheme, also in the nonconvex example we have treated. Nevertheless, better values, in particular within the modified scheme, can be found in order to achieve a faster convergence; the choice of such values is a matter of successive tuning, starting with the values of the parameters that satisfy the sufficient convergence conditions.

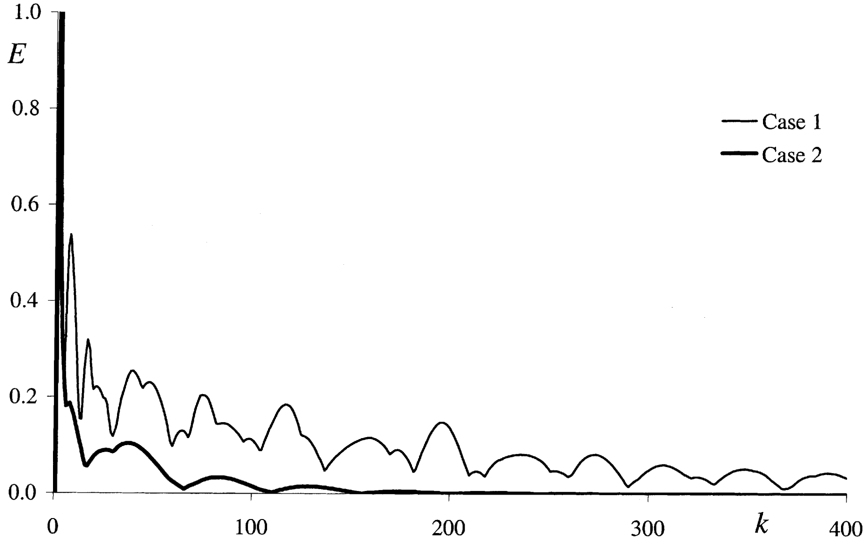

Fig. 1. Maximum discrepancy E in the case of two subproblems.

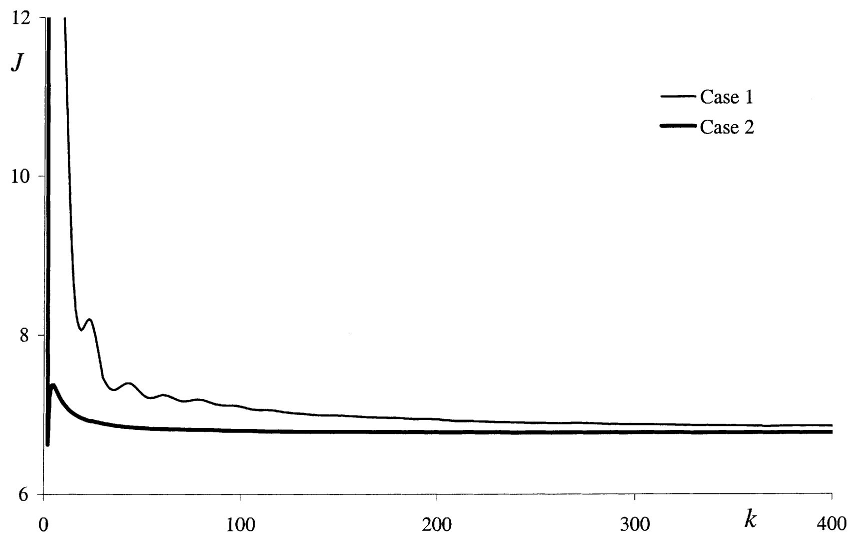

Fig. 2. Objective function J in the case of two subproblems.

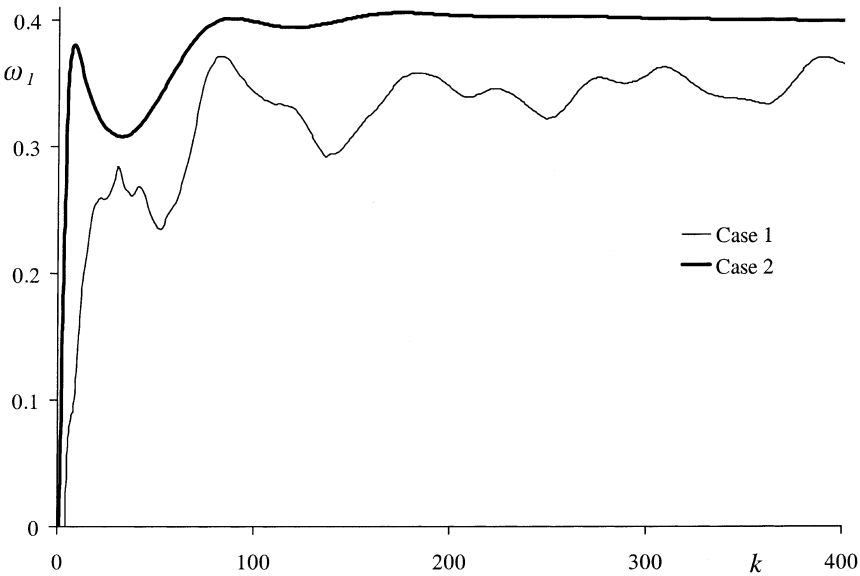

Fig. 3. Phase reference displacement ω_1 in the case of two subproblems.

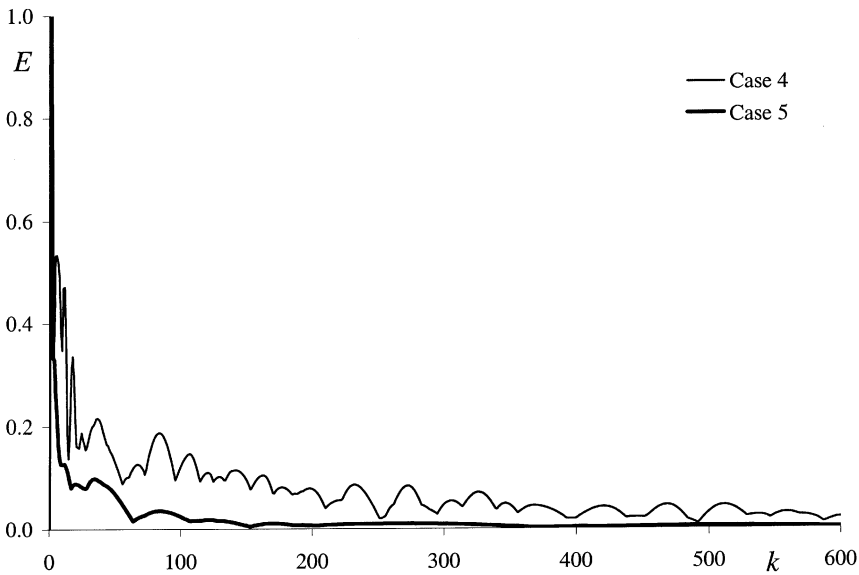

Fig. 4. Maximum discrepancy E in the case of three subproblems.

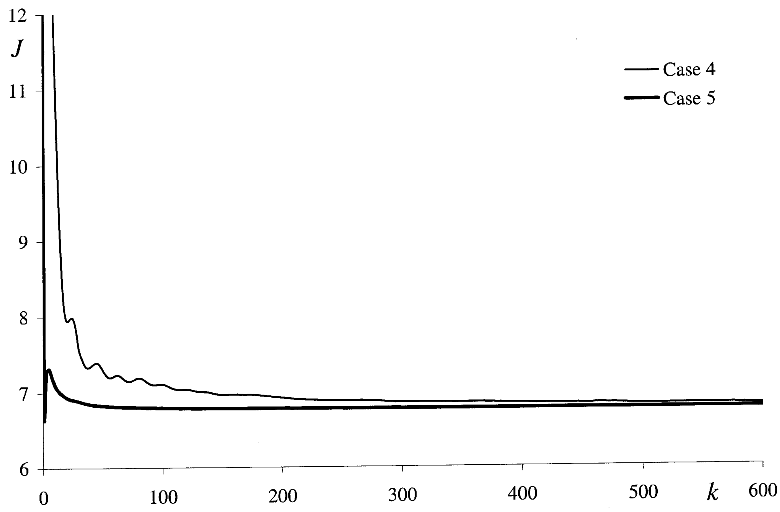

Fig. 5. Objective function J in the case of three subproblems.

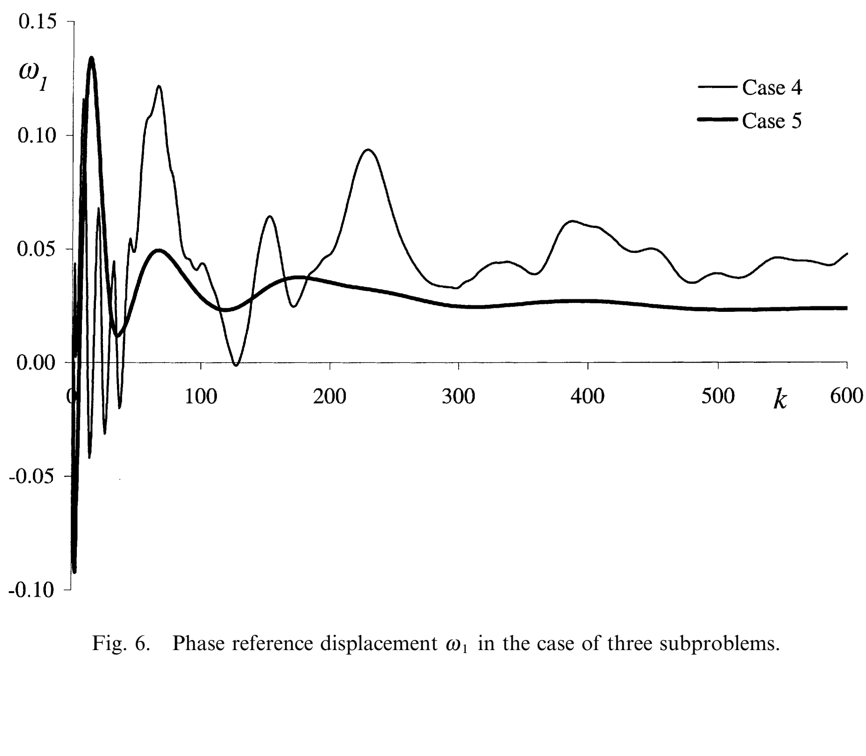

Fig. 6. Phase reference displacement ω_1 in the case of three subproblems.

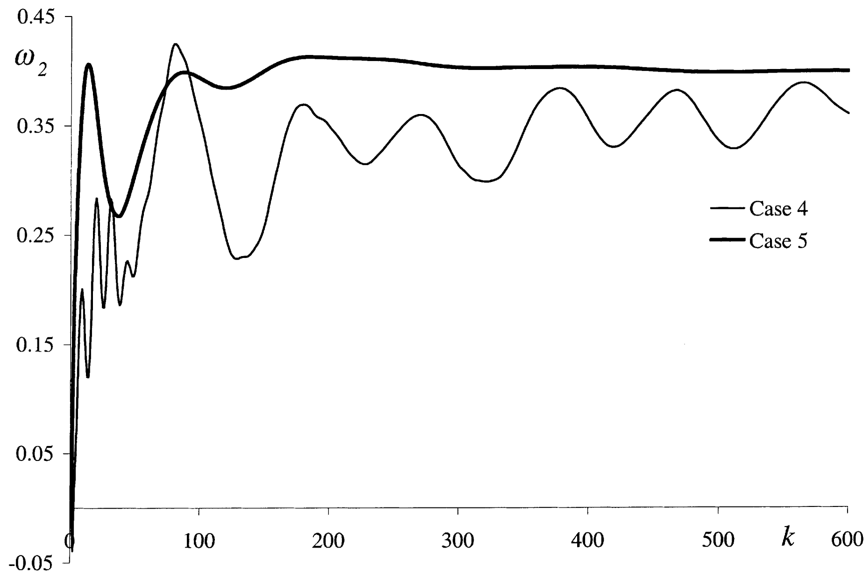

Fig. 7. Phase reference displacement ω_2 in the case of three subproblems.

5. Conclusions

Some results of the auxiliary problem principle pertaining to two-level methods have been specialized for a particular class of optimization problems, namely, the problems with a particular kind of variables, that appear only in the coupling constraints.

The developments have led to a modified form of the iterative APP scheme, which has some relevant features, such as easy formulation of the subproblems, that can be solved using packages that are available for the solution of unsplit problems. Moreover, specific convergence results have been drawn. The usefulness of the development has been proven in a study case that refers to a large-scale power system optimization problem.

6. Appendix

In the following, the proofs of some sufficient conditions for convergence of the APP two-level iterative scheme with Δ-variables (refer to Section 3.2) are outlined. Let us refer to problem (15) reformulated in a more compact nondecomposed form,

$$\min_{u\in U,\ \omega\in\Re^{n-1}} \quad J(u), \tag{27a}$$

$$\text{s.t.} \quad \Theta(u,\omega)=0, \tag{27b}$$

where $J(u)$ is convex and its Gateaux derivative $J'(u)$ is Lipschitz with constant A. Let the coupling constraint set Θ be partitioned into two subsets, Θ_α and Θ_ω, the first representing the consistency of the duplicated variables u_α that are not Δ-variables and the other representing the consistency of the duplicated Δ-variables u_ω,

$$\Theta(u,\omega)\equiv\begin{bmatrix}\Theta_\alpha(u_\alpha)\\ \Theta_\omega(u_\omega,\omega)\end{bmatrix}. \tag{28}$$

Let τ_α and τ_ω represent the Lipschitz constants of Θ_α and Θ_ω, respectively,

$$\|\Theta_\alpha(u_\alpha^k)-\Theta_\alpha(u_\alpha^{k+1})\|^2\le\tau_\alpha^2\|u_\alpha^{k+1}-u_\alpha^k\|^2, \tag{29a}$$

$$\|\Theta_\omega(u_\omega^k,\omega^k)-\Theta_\omega(u_\omega^{k+1},\omega^{k+1})\|^2\le\tau_\omega^2[\|u_\omega^{k+1}-u_\omega^k\|^2+\|\omega^{k+1}-\omega^k\|^2]. \tag{29b}$$

Due to (29), the following inequality holds:

$$\|\Theta(u^k,\omega^k)-\Theta(u^{k+1},\omega^{k+1})\|^2\le\tau_u^2\|u^{k+1}-u^k\|^2+\tau_\omega^2\|\omega^{k+1}-\omega^k\|^2, \tag{30}$$

where

$$\tau_u=\max\{\tau_\alpha,\tau_\omega\}, \tag{31}$$

Let us choose an additive core function

$$K(u,\omega)=K_U(u)+K_\Omega(\omega). \tag{32}$$

Here, $K_U(u)$ is strongly convex with constant β and its Gateaux derivative $K_U'(u)$ is Lipschitz with constant B, and $K_\Omega(\omega)$ is strongly convex with constant γ and its Gateaux derivative $K_\Omega'(\omega)$ is Lipschitz with constant Γ.

Applying the APP to (27), the following two-step algorithm is derived:

$$\min_{u\in U,\ \omega\in\Re^{n-1}}\left\{\begin{aligned}&K_U(u)+\langle\epsilon J'(u^k)-K_U'(u^k),u\rangle\\&+K_\Omega(\omega)-\langle K_\Omega'(\omega^k),\omega\rangle\\&+\epsilon\langle p^k+c\Theta(u^k,\omega^k),\Theta(u,\omega)\rangle\end{aligned}\right\}, \tag{33a}$$

$$p^{k+1}=p^k+\rho\Theta(u^{k+1},\omega^{k+1}). \tag{33b}$$

Since (33a) is a minimization problem with a strongly convex function over a closed convex set, there exists a unique solution (u^{k+1}, ω^{k+1}) that satisfies the following variational inequality:

$$\begin{aligned}&\langle K'_U(u^{k+1}) - K'_U(u^k), u - u^{k+1}\rangle + \langle K'_\Omega(\omega^{k+1}) - K'_\Omega(\omega^k), \omega - \omega^{k+1}\rangle\\ &+ \epsilon\langle J'(u^k), u - u^{k+1}\rangle + \epsilon\langle p^k + c\Theta(u^k, \omega^k), \Theta(u, \omega) - \Theta(u^{k+1}, \omega^{k+1})\rangle \geq 0,\\ &\qquad\qquad \forall u \in U \text{ and } \forall \omega \in \Re^{n-1}. \end{aligned}\tag{34}$$

A necessary and sufficient condition for (u^*, ω^*, p^*) to be a saddle point of problem (27) is the following:

$$J(u) - J(u^*) + \langle p^*, \Theta(u, \omega)\rangle \geq 0, \qquad \forall u \in U \text{ and } \forall \omega \in \Re^{n-1}. \tag{35}$$

By adequately combining (34) for $u = u^*$ and $\omega = \omega^*$ and (35) for $u = u^{k+1}$ and $\omega = \omega^{k+1}$, and upon multiplication by ϵ, the following inequality is obtained:

$$\begin{aligned}&\langle K'_U(u^{k+1}) - K'_U(u^k), u^* - u^{k+1}\rangle + \langle K'_\Omega(\omega^{k+1}) - K'_\Omega(\omega^k), \omega^* - \omega^{k+1}\rangle\\ &+ \epsilon\langle J'(u^k), u^* - u^{k+1}\rangle + \epsilon J(u^{k+1}) - \epsilon J(u^*)\\ &- \epsilon\langle p^k - p^*, \Theta(u^{k+1}, \omega^{k+1})\rangle - \epsilon c\langle \Theta(u^k, \omega^k), \Theta(u^{k+1}, \omega^{k+1})\rangle \geq 0.\end{aligned}\tag{36}$$

Let us now consider in detail each term of (36).

Concerning the first term in (36), thanks to the strong convexity of the core function, the following inequality can be derived easily:

$$\begin{aligned}&\langle K'_U(u^{k+1}) - K'_U(u^k), u^* - u^{k+1}\rangle\\ &\leq \langle K'_U(u^{k+1}), u^* - u^{k+1}\rangle - \langle K'_U(u^k), u^* - u^k\rangle\\ &+ K_U(u^{k+1}) - K_U(u^k) - (1/2)\beta\|u^k - u^{k+1}\|^2.\end{aligned}\tag{37}$$

Similarly, for the second term in (36), the following inequality can be derived:

$$\begin{aligned}&\langle K'_\Omega(\omega^{k+1} - K'_\Omega(\omega^k), \omega^* - \omega^{k+1}\rangle\\ &\leq \langle K'_\Omega(\omega^{k+1}), \omega^* - \omega^{k+1}\rangle - \langle K'_\Omega(\omega^k), \omega^* - \omega^k\rangle\\ &\quad + K_\Omega(\omega^{k+1}) - K_\Omega(\omega^k) - (1/2)\gamma\|\omega^k - \omega^{k+1}\|^2.\end{aligned}\tag{38}$$

Concerning the third, fourth, and fifth terms in (36), using the properties assumed for J, the following inequality holds:

$$\begin{aligned}&\epsilon\langle J'(u^k), u^* - u^{k+1}\rangle + \epsilon J(u^{k+1}) - \epsilon J(u^*)\\ &\leq (1/2)(\epsilon A)\|u^k - u^{k+1}\|^2.\end{aligned}\tag{39}$$

Using (33b), the sixth term in (36) can be written as

$$-\epsilon\langle p^k - p^*, \Theta(u^{k+1}, \omega^{k+1})\rangle$$
$$= (1/2)(\epsilon/\rho)(\|p^k - p^*\|^2 - \|p^{k+1} - p^*\|^2) + (1/2)\epsilon\rho\|\Theta(u^{k+1}, \omega^{k+1})\|^2. \tag{40}$$

Concerning the last terms in (36), thanks to (30), the following inequality holds:

$$\begin{aligned} &-\epsilon c\langle\Theta(u^k, \omega^k), \Theta(u^{k+1}, \omega^{k+1})\rangle \\ &\leq (1/2)\epsilon c\tau_u^2\|u^{k+1} - u^k\|^2 + (1/2)\epsilon c\tau_\omega^2\|\omega^{k+1} - \omega^k\|^2 \\ &\quad - (1/2)\epsilon c\|\Theta(u^k, \omega^k)\|^2 - (1/2)\epsilon c\|\Theta(u^{k+1}, \omega^{k+1})\|^2. \end{aligned} \tag{41}$$

Finally, using (37)–(41), inequality (36) can be rewritten as

$$\begin{aligned} &\varphi(u^k, \omega^k, p^k) - \varphi(u^{k+1}, \omega^{k+1}, p^{k+1}) \\ &\geq (1/2)(\beta - \epsilon A - \epsilon c\tau_u^2)\|u^{k+1} - u^k\|^2 \\ &\quad + (1/2)(\gamma - \epsilon c\tau_\omega^2)\|\omega^{k+1} - \omega^k\|^2 \\ &\quad + (1/2)\epsilon(2c - \rho)\|\Theta(u^{k+1}, \omega^{k+1})\|^2, \end{aligned} \tag{42}$$

where

$$\begin{aligned} \varphi(u, \omega, p) = {} & K_U(u^*) - K_U(u) - \langle K_U'(u), u^* - u\rangle \\ & + K_\Omega(\omega^*) - K_\Omega(\omega) - \langle K_\Omega'(\omega), \omega^* - \omega\rangle \\ & + (1/2)(\epsilon/\rho)\|p - p^*\|^2 - (1/2)\epsilon c\|\Theta(u, \omega)\|^2. \end{aligned} \tag{43}$$

From (42), it is evident that, if

$$0 < \epsilon < \beta/(A + c\tau_u^2), \tag{44a}$$
$$0 < \epsilon < \gamma/c\tau_\omega^2, \tag{44b}$$
$$0 < \rho < 2c, \tag{44c}$$

then

$$\varphi(u, \omega, p) \geq 0, \tag{44d}$$
$$\varphi(u^k, \omega^k, p^k) \geq \varphi(u^{k+1}, \omega^{k+1}, p^{k+1}). \tag{44e}$$

In conclusion, $\varphi(u^k, \omega^k, p^k)$ is a nonincreasing sequence, bounded from below, and thus it converges. Then, the proof of convergence to a saddle point can be carried out as in Refs. 1–2.

In the following, some considerations about the convergence characteristics of the APP two-level iterative scheme with Δ-variables (refer to Section

3.2) are briefly presented. Let us consider inequality (44e); from (43), it can be written as

$$\begin{aligned}
&K_U(u^{k+1}) - K_U(u^k) + \langle K'_U(u^{k+1}), u^* - u^{k+1}\rangle - \langle K'_U(u^k), u^* - u^k\rangle + \\
&+ K_\Omega(\omega^{k+1}) - K_\Omega(\omega^k) + \langle K'_\Omega(\omega^{k+1}), \omega^* - \omega^{k+1}\rangle - \langle K'_\Omega(\omega^k), \omega^* - \omega^k\rangle \\
&+ (1/2)(\epsilon/\rho)[\|p^k - p^*\|^2 - \|p^{k+1} - p^*\|^2] \\
&- (1/2)\epsilon c[\|\Theta(u^k, \omega^k)\|^2 - \|\Theta(u^{k+1}, \omega^{k+1})\|^2] \geq 0. \qquad (45)
\end{aligned}$$

Due to the characteristics of the core function K, we have

$$\begin{aligned}
&K_U(u^{k+1}) - K_U(u^k) + \langle K'_U(u^{k+1}), u^* - u^{k+1}\rangle - \langle K'_U(u^k), u^* - u^k\rangle \\
&\leq (1/2)B\|u^k - u^*\|^2 - (1/2)\beta\|u^{k+1} - u^*\|^2, \qquad (46a) \\
&K_\Omega(\omega^{k+1}) - K_\Omega(\omega^k) + \langle K'_\Omega(\omega^{k+1}), \omega^* - \omega^{k+1}\rangle - \langle K'_\Omega(\omega^k), \omega^* - \omega^k\rangle \\
&\leq (1/2)\Gamma\|\omega^k - \omega^*\|^2 - (1/2)\gamma\|\omega^{k+1} - \omega^*\|^2. \qquad (46b)
\end{aligned}$$

Inequalities (45) and (46) yield

$$\begin{aligned}
&\beta\|u^{k+1} - u^*\|^2 + \gamma\|\omega^{k+1} - \omega^*\|^2 + (\epsilon/\rho)\|p^{k+1} - p^*\|^2 - \epsilon c\|\Theta(u^{k+1}, \omega^{k+1})\|^2 \\
&\leq B\|u^k - u^*\|^2 + \Gamma\|\omega^k - \omega^*\|^2 + (\epsilon/\rho)\|p^k - p^*\|^2 - \epsilon c\|\Theta(u^k, \omega^k)\|^2. \qquad (47)
\end{aligned}$$

Thanks to the characteristics of Θ, for the left-hand side term of (47) we have

$$\begin{aligned}
&(\beta - \epsilon c\tau_u^2)\|u^{k+1} - u^*\|^2 + (\gamma - \epsilon c\tau_\omega^2)\|\omega^{k+1} - \omega^*\|^2 + (\epsilon/\rho)\|p^{k+1} - p^*\|^2 \\
&\leq \beta\|u^{k+1} - u^*\|^2 + \gamma\|\omega^{k+1} - \omega^*\|^2 \\
&\quad + (\epsilon/\rho)\|p^{k+1} - p^*\|^2 - \epsilon c\|\Theta(u^{k+1}, \omega^{k+1})\|^2. \qquad (48)
\end{aligned}$$

If the sufficient conditions (44a)–(44c) stand, the left-hand side term in (48) is nonnegative, and then (47) can be extended as follows:

$$\begin{aligned}
0 &\leq (\beta - \epsilon c\tau_u^2)\|u^{k+1} - u^*\|^2 + (\gamma - \epsilon c\tau_\omega^2)\|\omega^{k+1} - \omega^*\|^2 + (\epsilon/\rho)\|p^{k+1} - p^*\|^2 \\
&\leq B\|u^k - u^*\|^2 + \Gamma\|\omega^k - \omega^*\|^2 + (\epsilon/\rho)\|p^k - p^*\|^2 - \epsilon c\|\Theta(u^k, \omega^k)\|^2. \qquad (49)
\end{aligned}$$

In (49), the left-hand side term is a weighted sum of the a posteriori errors and, due to the sufficient conditions (44), the weights are all positive. From (49), it is apparent that the sum of the a posteriori errors is upper bounded by the sum of the a priori errors on the right-hand side; moreover, the smaller are the values of B and Γ, the smaller is the upper bound.

References

1. Cohen, G., *Auxiliary Problem Principle and Decomposition of Optimization Problems*, Journal of Optimization Theory and Applications, Vol. 37, pp. 287–356, 1980.
2. Cohen, G., and Zhu, D. L., *Decomposition Coordination Methods in Large Scale Optimization Problems: The Nondifferentiable Case and the Use of Augmented Lagrangians*, Advances in Large Scale Systems, Vol. 1, pp. 203–276, 1984.
3. Contreras, J., Losi, A., Russo, M., and Wu, F. F., *Simulation and Evaluation of Optimization Problem Solutions in Distributed Energy Management Systems*, IEEE Transactions on Power Systems, Vol. 17, pp. 57–62, 2002.
4. Wood, J. W., and Wollenberg, B. F., *Power Generation, Operation, and Control*, 2nd Edition, John Wiley and Sons, New York, NY, 1996.
5. *IEEE 118-Bus Test System*, Power System Test Case Archive, available at http://www.ee.washington.edu/research/pstca, 2002.
6. Contreras, J., Losi, A., Russo, M., and Wu, F. F., *DistOpt: A Software Framework for Modeling and Evaluating Optimization Problem Solution in Distributed Environments*, Journal of Parallel and Distributed Computing, Vol. 60, pp. 741–763, 2000.