Working Paper 99-42
Statistics and Econometrics Series 12
June 1999

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34-91) 624-9849

A NEW DECOMPOSITION METHOD APPLIED TO OPTIMIZATION PROBLEMS ARISING IN POWER SYSTEMS: LOCAL AND GLOBAL BEHAVIOR.

Antonio J. Conejo, F.Javier Nogales and Francisco J. Prieto.*

Abstract

In this report a new decomposition methodology for optimization problems is presented. The proposed procedure is general, simple and efficient. It avoids most disadvantages of other common decomposition techniques, such as Lagrangian Relaxation or Augmented Lagrangian Relaxation. The new methodology is applied to a problem coming from interconnected power systems. The application of the new method to this problem allows the computation of an optimal coordinated but decentralized solution. Local and global convergence properties of the proposed decomposition algorithm are described. Numerical results show that the new decentralized methodology has a lower computational cost than other decomposition techniques, and in large-scale cases even lower than a centralized approach.

Keywords: Nonlinear programming, Decomposition methods, Decentralized coordination.

*Conejo, ETSI Industriales Universidad de Castilla-La Mancha, Ciudad Real, e-mail: aconejo@ind-cr.uclm.es; Nogales, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. C/ Madrid, 126 28903 Getafe -Madrid-, Spain, e-mail: fjnm@est-econ.uc3m.es; Prieto, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid, e-mail: fjp@est-econ, Tfno: 91-6249850.

I Introduction

In this report a new decomposition procedure for optimization problems is described and applied to a problem arising in the analysis of interconnected power systems. In such systems, it is often desirable to preserve the autonomy of each utility. This decentralized operation can be carried out using decomposition techniques arising in optimization problems. These techniques are often used in many planning problems. With them, an optimum of the global problem can be achieved in a decentraliced fashion. Also, the application of these techniques allow the realization of potential gains in computational efficiency and useful information is generated as part of the decomposition process.

The proposed decomposition method is applied to the multi-area Optimal Power Flow (OPF) problem. This is an important problem for the secure and economic operation of an interconnected power system. The multi-area OPF determines, in a precise way, the active and reactive power that each interconnected generation unit in the system must generate. This is done to ensure that all demand and security constraints for the system are satisfied, at a minimal cost for all interconnected areas. The resulting multi-area OPF problem is a large-scale non-convex optimization problem [1].

The decomposition methodology proposed in this paper is general, simple and efficient. The procedure allows the companies in each area to operate their systems independently of the others, while obtaining an optimal coordinated but decentralized solution. A central agent in the model is necessary to carry out coordination tasks for the whole system. This agent ensures the coupling of the global system and guarantees transparent transactions.

Several other decomposition techniques, such as Lagrangian Relaxation [2, 3, 4], and relaxation techniques based on Augmented Lagrangian functions [5, 6, 7], have been proposed for the solution of similar problems. An application to solve a multi-area DC OPF is described in [8], while in [9] an Augmented Lagrangian Relaxation procedure is used to solve a distributed OPF. These procedures present several disadvantages. For example, they do not guarantee the computation of an optimal solution for the global system (they obtain a quasi-optimal solution), they require a large computational effort, their convergence is slow, and they make use of several parameters which are difficult to update. Also, a central agent is needed to update information.

The proposed decomposition algorithm avoids the disadvantages from proce-

dures based on Lagrangian methods and presents the following advantages:

- A coordinated solution of the global problem is achieved. The coupling of the system is obtained by the Lagrange multipliers associated with certain constraints.

- The computational efficiency is improved. In Lagrangian-based techniques a subproblem must be solved at each iteration in the process. In the proposed decomposition algorithm, the solution of each subproblem is not necessary. The results from a single iteration are enough, providing a considerable reduction in computing time. This is a consequence of convergence properties of the proposed methodology.

- The implementation is simple and robust. The procedure generates the subproblems as slight modifications of the optimization problems for each area, and the convergence properties of the algorithm do not require any parameter updating.

- The solution process is more transparent. The central agent does no need to update any information, only to distribute it. This information is updated by the different areas of the system.

- It can be implemented in a distributed computation environment. This is due the minimal information interchange between areas that is required.

The report shows local and global convergence behaviour for the proposed Decomposition Algorithm. The main local convergence condition is related to the separable structure of the problem and can be interpreted as a measurement of the coupling between the areas in the global system. It has been verified that this condition holds for all multi-area OPF cases that have been found available to test the procedure. In this regard, these convergence properties seem to be satisfied for most practical cases of interest. It is shown that the local rate of convergence of the proposed Decomposition Algorithm is linear. It must be noted that the rate of convergence of a centralized approach could be quadratic.

Global properties of the method need a merit function in order to measure progress and a line search procedure. An Augmented Lagrangian function has been chosen as this merit function, and the line search procedure can be computed in a appropriate form for the new decomposition algorithm.

The report is organized as follows: in Section II a mathematical formulation suitable to be applied together with decomposition techniques is introduced. This formulation is given for the multi-area OPF model. Section III presents decomposition techniques available to date. Section IV describes the new decomposition methodology. Section V shows an example that illustrates the main ideas behind the procedure. Sections VI and VII present local and global convergence properties, respectively. Section VIII shows computational results and some conclusions are provided in Section IX.

II Problem Formulation

In this section it is discussed the structure that an optimization problem must present so that the application of the new decomposition methodology is advantageously. Decomposition techniques available to date are also based on this structure. The structure is described for a specific class of problems arising in interconnected power systems, specifically for multi-area OPF problems, where the power that each generation unit in the interconnected system must generate is determined. This is done to ensure that all demand and security constraints for the system are satisfied at a minimal cost for all interconnected areas.

Mathematically, a model for this problem can be formulated as

$$\begin{aligned} \text{minimize} \quad & f(z,u) && (1)\\ \text{subject to} \quad & h(z,u) = 0 && (2)\\ & g(z,u) \leq 0 && (3) \end{aligned}$$

where z is the set of dependent variables for the global system, that may contain bus voltage magnitudes, bus phase angles, and the MVAr output of generators performing bus voltage control. The vector u is the set of control variables for the global system. It may include real and reactive power generation, phase-shifter angles, voltage control settings, transformer taps settings, etc.

Function (1) is known as the objective function. Several different objective functions may be of interest, such as: total system operation cost, total power losses from power transmission, total control shifts, or total system emissions, for example.

Equations of type (2) are the Power Flow equations [1], including two equations for each bus of the global system, representing the active and reactive power balance in each bus. Constraints of type (3) are the transmission capacity limits for each line of the global system, and technical limits over dependent and control variables.

The multi-area OPF model (1)-(3) is a non-convex, differentiable, large-scale optimization problem. To simplify the presentation of the new decomposition methodology, problem (1)-(3) will be reformulated. It will be assumed, without loss of generality, that the global system consists of just two areas, X and Y. The variables belonging to area X will be denoted as $x = [z^x, u^x]$, and the variables belonging to area Y will be denoted by $y = [z^y, u^y]$.

The problem equivalent to (1)-(3) presents the structure

$$\begin{aligned} \text{minimize} \quad & f_X(x) + f_Y(y) && (4)\\ \text{subject to} \quad & h_X(x,y) = 0 && (5)\\ & h_Y(x,y) = 0 && (6)\\ & g_X(x) \leqq 0 && (7)\\ & g_Y(y) \leqq 0, && (8) \end{aligned}$$

and is called the *Primal Problem.*

The equations (5)-(6) are commonly known as *complicating equations.* These equations contain variables from both areas and prevent each system from operating independently from the other. If these equations are removed from problem (4)-(8), the resulting problem can be trivially decomposed into one subproblem for each area.

The complicating equations (5) include the power balance equations at the interconnecting buses of system X (the buses from system X connected to buses from system Y). The complicating equations (6) include the power balance equations at the interconnecting buses of system Y. Also, the transmission capacity limits for the interconnecting lines of the global system are complicating equations. These constraints must be introduced into equations (5)-(6) (after slack variables have already been included to convert inequality constraints into equality constraints). Therefore, the only variables appearing in the complicating equations are those corresponding to the interconnecting buses of the global system.

Equations (7) and (8) contain only variables belonging to areas X and Y, respectively. These constraints represent balance equations, transmission limits, and technical constraints for areas X and Y, respectively, and may

correspond to equality or inequality constraints.

III Available Decomposition Techniques

There are several methods that decompose problems of the form (5)-(6). It should be noted that the main goal of these techniques is not to improve computational efficiency. Instead, they seek for a solution of (4)-(8) in a decentraliced fashion, preserving the autonomy of each area.

a) Lagrangian Relaxation

This technique seeks for a solution of the Primal Problem (4)-(8) through another problem called the *Dual Problem.*

In this technique, the equations (5)-(6) must be of the form

$$h_X(x,y) = h_X^1(x) + h_X^2(y)$$
$$h_Y(x,y) = h_Y^1(x) + h_Y^2(y).$$

Equations of the previous form are called *separable equations.*

In order to decompose the Primal Problem (4)-(8), equations (5)-(6) are included in the objective function. Then, the *Relaxed Primal Problem* to be solved is

$$\begin{aligned} \underset{x,y}{\text{minimize}} \quad & \mathcal{L}(x,y,\lambda_X,\lambda_Y) && (9)\\ \text{subject to} \quad & g_X(x) \leqq 0 && (10)\\ & g_Y(y) \leqq 0 && (11) \end{aligned}$$

where the *Lagrangian function,* $\mathcal{L}$, is defined as

$$\mathcal{L}(x,y,\lambda_X,\lambda_Y) = f_X(x) + f_Y(y) + \lambda_X^T\, h_X^1(x) + \\ + \lambda_X^T\, h_X^2(y) + \lambda_Y^T\, h_Y^1(x) + \lambda_Y^T\, h_Y^2(y),$$

and vectors λ_X, λ_Y are known as the *Lagrange multipliers* of equations (5), (6), respectively.

The *Dual Problem* corresponding to (4)-(8) is defined as

$$\sup_{\lambda_X, \lambda_Y} \quad \phi(\lambda_X, \lambda_Y), \tag{12}$$

where the *Lagrangian dual subproblem* is

$$\begin{aligned} \phi(\lambda_X, \lambda_Y) \quad \equiv \quad & \underset{x,y}{\text{minimize}} \quad \mathcal{L}(x, y, \lambda_X, \lambda_Y) \\ & \text{subject to} \quad g_X(x) \leqq 0 \\ & \qquad\qquad\ \ g_Y(y) \leqq 0 \end{aligned} \tag{13}$$

The decomposition algorithm based on the Lagrangian Relaxation is:

Lagrangian Relaxation:

Step 0
Parameters $\overline{\lambda}_X, \overline{\lambda}_Y$ are initialized.

Step 1
System X solves

$$\begin{aligned} &\text{minimize} \quad f_X(x) + \overline{\lambda}_X^T\, h_X^1(x) + \overline{\lambda}_Y^T\, h_Y^1(x) && (14) \\ &\text{subject to} \quad g_X(x) \leqq 0. && (15) \end{aligned}$$

Step 2
System Y solves

$$\begin{aligned} &\text{minimize} \quad f_Y(y) + \overline{\lambda}_X^T\, h_X^2(y) + \overline{\lambda}_Y^T\, h_Y^2(y) && (16) \\ &\text{subject to} \quad g_Y(y) \leqq 0. && (17) \end{aligned}$$

Step 3
Parameters $\overline{\lambda}_X, \overline{\lambda}_Y$ are updated (information exchange).
The algorithm stops if the convergence criteria for the global problem are satisfied. Otherwise, go to Step 1.

The values $\overline{\lambda}_X, \overline{\lambda}_Y$ are kept fixed in Steps 1 and 2.

This algorithm requires a central agent. This agent updates the multipliers corresponding to the interconnected buses. There are several techniques that update the Lagrange multipliers. These techniques seek for a maximum of the dual function. Among these techniques are: Subgradient methods [4],

Cutting Plane methods [3], and Bundle methods [10]. For example, the easiest implementation is the subgradient method:

$$\lambda_X^{k+1} = \lambda_X^k + \alpha_k \frac{h_X^1(x^k) + h_X^2(y^k)}{\|h_X^1(x^k) + h_X^2(y^k)\|} \tag{18}$$

$$\lambda_Y^{k+1} = \lambda_Y^k + \alpha_k \frac{h_Y^1(x^k) + h_Y^2(y^k)}{\|h_Y^1(x^k) + h_Y^2(y^k)\|}. \tag{19}$$

where α_k is a sequence verifying

$$\alpha_k \to 0, \quad \sum_{k=0}^{\infty} \alpha_k = +\infty.$$

The *duality gap* is defined as the difference between the optimum of the Primal and the Dual problems:

$$f_X(x^*) + f_Y(y^*) - \phi(\lambda_X^*, \lambda_Y^*).$$

If problem (4)-(8) is not convex then the duality gap can be greater than 0.

The Lagrangian Relaxation technique presents several disadvantages. For example, it does not guarantee the computation of an optimal solution for the global system (they obtain a quasi-optimal solution because of non-convexity of the global problem), they require a large computational effort, their convergence is slow, and they make use of several parameters which are difficult to update. Also, a central agent is needed to update information.

b) Augmented Lagrangian Relaxation

This technique seeks to mix penalization methods with local duality theory, by including a quadratic penalization term in the objective function. This term makes the Lagrangian function more convex.

Equations (5)-(6) must be separable. In the Primal Problem (4)-(8), the complicating equations (5)-(6) are relaxed and penalized. The new objective function has the form

$$\begin{aligned}\mathcal{LA}(x, y, \lambda_X, \lambda_Y) = f_X(x) + f_Y(y) + \lambda_X^T\, h_X^1(x) + \lambda_X^T\, h_X^2(y) + \\ + \lambda_Y^T\, h_Y^1(x) + \lambda_Y^T\, h_Y^2(y) + \frac{\rho}{2} \left\| \begin{bmatrix} h_X^1(x) + h_X^2(y) \\ h_Y^1(x) + h_Y^2(y) \end{bmatrix} \right\|^2, \end{aligned} \tag{20}$$

Function $\mathcal{LA}$ is called the *Augmented Lagrangian function.* For ρ sufficiently large, the Augmented Lagrangian function is locally convex, and duality theory can be applied.

Function $\mathcal{LA}$ is not separable, because of the quadratic term. In order to decompose the quadratic term it is necessary to fix variables corresponding to neighbouring areas [6]. In area X the quadratic term is given by

$$\frac{\rho}{2}\|h_X(x,\overline{y})\|^2 + \|h_Y(x,\overline{y})\|^2$$

The decomposition algorithm based on the Augmented Lagrangian Relaxation is:

Augmented Lagrangian Relaxation:

Step 0

Variables and parameters $\overline{x}, \overline{y}, \overline{\lambda}_X, \overline{\lambda}_Y$ are initialized.

Step 1

System X solves

$$\begin{aligned} \text{minimize} \quad & f_X(x) + \overline{\lambda}_X^T\, h_X^1(x) + \overline{\lambda}_Y^T\, h_Y^1(x) + \\ & + \frac{\rho}{2}\|h_X(x,\overline{y})\|^2 + \|h_Y(x,\overline{y})\|^2 \qquad (21) \\ \text{subject to} \quad & g_X(x) \leqq 0. \qquad (22) \end{aligned}$$

Step 2

System Y solves

$$\begin{aligned} \text{minimize} \quad & f_Y(y) + \overline{\lambda}_X^T\, h_X^2(y) + \overline{\lambda}_Y^T\, h_Y^2(y) + \\ & + \frac{\rho}{2}\|h_X(\overline{x},y)\|^2 + \|h_Y(\overline{x},y)\|^2 \qquad (23) \\ \text{subject to} \quad & g_Y(y) \leqq 0. \qquad (24) \end{aligned}$$

Step 3

Variables and parameters $\overline{x}, \overline{y}, \overline{\lambda}_X, \overline{\lambda}_Y$ are updated (information exchange).
The algorithm stops if the convergence criteria for the global problem are satisfied. Otherwise, go to Step 1.

This algorithm requires a central agent. This agent updates the Lagrange multipliers corresponding to the interconnected buses. In Step 3, an usual procedure to update these multipliers [2], at iteration k, is

$$\lambda_X^{k+1} = \lambda_X^k + \rho \left(h_X^1(x^k) + h_X^2(y^k)\right) \tag{25}$$

$$\lambda_Y^{k+1} = \lambda_Y^k + \rho \left(h_Y^1(x^k) + h_Y^2(y^k)\right). \tag{26}$$

The values $\overline{x}, \overline{y}, \overline{\lambda}_X, \overline{\lambda}_Y$ are fixed from the solution at previous iteration. At each iteration of the algorithm, it may be necessary to increase ρ. A general procedure is to define

$$\rho^{k+1} = \beta\rho^k \quad \text{with } \beta > 1.$$

The Augmented Lagrangian Relaxation technique presents several disadvantages. For example, it requires a large computational effort, its convergence is slow, and it makes use of several parameters which are difficult to update. Also, a central agent is needed to update information.

IV New Decomposition Methodology

The proposed method is based on the decomposition of the first-order necessary conditions, see [2], for problem (4)-(8). The main characteristic of such an approach is that convergence properties does not require an optimal solution of the subproblems at each iteration of the algorithm. Because the movement directions are computed in a decomposed way, it is enough to compute a single iteration of these subproblems. Therefore, the new method has the potential for large savings in computing times. Other decomposition methods need, in order to attain convergence, to solve the subproblems until the optimum is achieved. Therefore, the new methodology obtains a solution of problem (4)-(8) in a decentralized fashion, and with higher computational efficiency.

The first-order necessary conditions for problem (4)-(8) are

$$\begin{aligned}
\nabla_x f_X(x^*) + J_x^T h_X(x^*, y^*)\ \lambda_X^* + J_x^T h_Y(x^*, y^*)\ \lambda_Y^* + J_x^T g_X(x^*)\ \sigma_X^* &= 0 \\
h_X(x^*, y^*) &= 0 \\
g_X(x^*) &\leq 0 \\
g_X(x^*)^T\ \sigma_X^* &= 0 \\
\sigma_X^* &\geq 0^1 \\
\nabla_y f_Y(y^*) + J_y^T h_x(x^*, y^*)\ \lambda_X^* + J_y^T h_Y(x^*, y^*)\ \lambda_Y^* + J_y^T g_Y(y^*)\ \sigma_Y^* &= 0 \\
h_Y(x^*, y^*) &= 0 \\
g_Y(y^*) &\leq 0 \\
g_Y(y^*)^T\ \sigma_Y^* &= 0 \\
\sigma_Y^* &\geq 0^1.
\end{aligned}$$

These conditions coincide, for the optimal values of $x^*, y^*, \lambda_X^*, \lambda_Y^*$, with the first-order necessary conditions for problems

$$\begin{aligned}
\text{minimize} \quad & f_X(x) + (\lambda_Y^*)^T\ h_Y(x, y^*) && (27) \\
\text{subject to} \quad & h_X(x, y^*) = 0 && (28) \\
& g_X(x) \leqq 0, && (29)
\end{aligned}$$

and

$$\begin{aligned}
\text{minimize} \quad & f_Y(y) + (\lambda_X^*)^T\ h_X(x^*, y) && (30) \\
\text{subject to} \quad & h_Y(x^*, y) = 0 && (31) \\
& g_Y(y) \leqq 0. && (32)
\end{aligned}$$

The values λ_X^*, λ_Y^* are the Lagrange multipliers of equations (5) and (6), respectively. The values σ_X^*, σ_Y^* are the Lagrange multipliers of equations (7) y (8), respectively.

The new approach has the advantage that convergence properties do not require an optimal solution of the subproblems (27)-(29) and (30)-(32) at each iteration of the algorithm. It is enough to perform a single iteration for each subproblem, and then to update these optimality conditions. As a consequence, computation times can be significantly reduced with respect

[1]These conditions are only applicable when constraints (7), (8), respectively, are inequality constraints.

to other methods which require the computation of the optimum for the subproblems in order to attain convergence.

Finally, it must be noted that generalization of the proposed decomposition algorithm for three or more interconnected areas is immediate. An outline of the algorithm is as follows:

Decomposition Algorithm:

Step 0

Variables and parameters $\overline{x}, \overline{y}, \overline{\lambda}_X, \overline{\lambda}_Y$ are initialized.

Step 1

System X carries out one iteration for

$$\begin{aligned} \text{minimize} \quad & f_X(x) + \overline{\lambda}_Y^T \, h_Y(x, \overline{y}) && (33) \\ \text{subject to} \quad & h_X(x, \overline{y}) = 0 && (34) \\ & g_X(x) \leq 0 && (35) \end{aligned}$$

and obtains $\Delta x, \Delta \lambda_X$.

Step 2

System Y carries out one iteration for

$$\begin{aligned} \text{minimize} \quad & f_Y(y) + \overline{\lambda}_X^T \, h_X(\overline{x}, y) && (36) \\ \text{subject to} \quad & h_Y(\overline{x}, y) = 0 && (37) \\ & g_Y(y) \leq 0 && (38) \end{aligned}$$

and obtains $\Delta y, \Delta \lambda_Y$.

Step 3

Variables and parameters $\overline{x}, \overline{y}, \overline{\lambda}_X, \overline{\lambda}_Y$ are updated (information exchange).
The algorithm stops if the convergence criteria for the global problem are satisfied. Otherwise, go to Step 1.

It is relevant to observe that problem (33)-(35) is a modified OPF for area X. The only departure from a standard OPF is the Lagrangian term in the objective function (33). This property preserves the special structure of the problem in the subproblems and it may allow a more robust and simpler implementation of the decomposition procedure, as specialized codes can be used for the subproblems.

The search directions for subproblems (33)-(35) and (36)-(38) can be computed independently of each other which may result in a *parallel implementation.* A modified Newton procedure is used, in conjunction with a nonlinear interior point treatment of the inequality constraints.

Step 3, requires a central agent to coordinate the process; an ISO could perform this role. This agent receives certain information from the areas and returns it to the appropriate areas. This information consists of the values $\overline{x}, \overline{y}, \overline{\lambda}_X, \overline{\lambda}_Y$. The values $\overline{x}, \overline{y}$ represent variables involved in the interconnecting buses. The values $\overline{\lambda}_X, \overline{\lambda}_Y$ are the Lagrange multipliers associated with the complicating equations.

Also, it must be emphasized that in the proposed decomposition algorithm the central agent only distributes information and checks the convergence condition. This central agent needs a minimal amount of information to check this condition, and this information can be exchanged between different areas. This is not the case in other decomposition techniques. In Lagrangian-based procedures, areas must offer the central agent the same minimal information, but the central agent must update this information and return it to the different areas. In the proposed Decomposition Algorithm the central agent does not need to update the information to be exchanged because this information is updated by the areas of the system, implying a more transparent and simpler process.

V Example

Next, a simple example is introduced. This example clarifies how the three considered decomposition techniques, the Lagrangian Relaxation, the Augmented Lagrangian Relaxation, and the proposed Decomposition Algorithm, work.

Assume that the global problem to be solved is

$$\begin{aligned} \text{minimize} \quad & x_1^2 + x_2^2 + y_1^2 + y_2^2 && (39) \\ \text{subject to} \quad & 1 - 4x_1 - y_2 = 0 && (40) \\ & 1 - x_1 - 4y_2 = 0 && (41) \end{aligned}$$

In this example, equations (40) and (41) are complicating equations, and no constraints of the form (7) or (8) have been included, for simplicity sake.

There are only two variables involved in the complicating equations, x_1 and y_2. Equation (40) represents (5) in the general model, and equation (41) represents (6). The solution of this problem is

$$x^* = \begin{bmatrix} 0.2 \\ 0.0 \end{bmatrix}, \quad y^* = \begin{bmatrix} 0.0 \\ 0.2 \end{bmatrix}, \quad \lambda^* = \begin{bmatrix} 0.08 \\ 0.08 \end{bmatrix}.$$

The constraint vector (40)-(41) is denoted by h:

$$h(x, y) = \begin{bmatrix} 1 - 4x_1 - y_2 \\ 1 - x_1 - 4y_2 \end{bmatrix}.$$

This problem is solved below by the three described procedures.

a) Lagrangian Relaxation Solution

Subproblems to be solved in Steps 1 and 2 of the decomposition algorithm based on Lagrangian Relaxation are, respectively,

$$\text{minimize} \quad x_1^2 + x_2^2 - (4\overline{\lambda}_1 + \overline{\lambda}_2)\, x_1 \tag{42}$$

and

$$\text{minimize} \quad y_1^2 + y_2^2 - (\overline{\lambda}_1 + 4\overline{\lambda}_2)\, y_2. \tag{43}$$

Iteration $k = 0$:

Step 0. Variables and multipliers

$$\overline{x} = \begin{bmatrix} 0.4 \\ 0.4 \end{bmatrix}, \quad \overline{y} = \begin{bmatrix} 0.4 \\ 0.4 \end{bmatrix}, \quad \overline{\lambda} = \begin{bmatrix} 0.01 \\ 0.01 \end{bmatrix}$$

are initialized.

Iteration $k = 1$:

Step 1. System X solves (42), beginning at $x = \overline{x}$, and obtains

$$x = \begin{bmatrix} 0.025 \\ 0.0 \end{bmatrix}.$$

Step 2. System Y solves (43), beginning at $y = \overline{y}$, and obtains

$$y = \begin{bmatrix} 0.0 \\ 0.025 \end{bmatrix}.$$

Step 3. Convergence. Multiplier updating. The central agent checks if convergence condition $\|h(x,y)\| < 10^{-4}$ is verified:

$$h = \begin{bmatrix} 0.875 \\ 0.875 \end{bmatrix}, \quad \|h\| = 1.2374 > 10^{-4}.$$

The convergence condition is not verified, and the central agent updates the multipliers by a subgradient procedure:

$$\overline{\lambda} = \overline{\lambda} + \frac{1}{2k} \frac{h(x,y)}{\|h(x,y)\|} =$$

$$= \begin{bmatrix} 0.01 \\ 0.01 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0.875 \\ 0.875 \end{bmatrix} / 1.2374 = \begin{bmatrix} 0.3636 \\ 0.3636 \end{bmatrix}.$$

Next, the variables

$$\overline{x} = x = \begin{bmatrix} 0.025 \\ 0.0 \end{bmatrix}, \quad \overline{y} = y = \begin{bmatrix} 0.0 \\ 0.025 \end{bmatrix}$$

are updated, $k = k+1 = 2$, and Steps 1, 2, and 3 of the algorithm are repeated until convergence is achieved.

The algorithm stops when $k = 53$, $\|h\| = 4.1772 \times 10^{-5}$,

$$x = \begin{bmatrix} 0.2 \\ 0.0 \end{bmatrix}, \quad y = \begin{bmatrix} 0.0 \\ 0.2 \end{bmatrix}, \quad \lambda = \begin{bmatrix} 0.0868 \\ 0.0868 \end{bmatrix}.$$

b) Augmented Lagrangian Solution

Subproblems to be solved in Steps 1 and 2 of the decomposition algorithm based on Augmented Lagrangian Relaxation are, respectively,

$$\text{minimize} \quad x_1^2 + x_2^2 - (4\overline{\lambda}_1 + \overline{\lambda}_2)\, x_1 + \frac{\rho}{2}(1 - 4x_1 - \overline{y}_2)^2 \tag{44}$$

and

$$\text{minimize} \quad y_1^2 + y_2^2 - (\overline{\lambda}_1 + 4\overline{\lambda}_2)\, y_2 + \frac{\rho}{2}(1 - \overline{x}_1 - 4y_2)^2. \tag{45}$$

Iteration $k = 0$:

Step 0. Variables, multipliers, and parameters

$$\overline{x} = \begin{bmatrix} 0.4 \\ 0.4 \end{bmatrix}, \quad \overline{y} = \begin{bmatrix} 0.4 \\ 0.4 \end{bmatrix}, \quad \overline{\lambda} = \begin{bmatrix} 0.01 \\ 0.01 \end{bmatrix}, \quad \rho = 0.1$$

are initialized.

Iteration $k = 1$:

Step 1. System X solves (44), beginning in $x = \overline{x}$, and obtains

$$x = \begin{bmatrix} 0.0806 \\ 0.0 \end{bmatrix}.$$

Step 2. System Y solves (45), beginning in $y = \overline{y}$, and obtains

$$y = \begin{bmatrix} 0.0 \\ 0.0806 \end{bmatrix}.$$

Step 3. Convergence. Multiplier updating. The central agent checks if convergence condition $\|h(x, y)\| < 10^{-4}$ is verified:

$$h = \begin{bmatrix} 0.5972 \\ 0.5972 \end{bmatrix}, \quad \|h\| = 0.8446 > 10^{-4}.$$

The convergence condition is not verified, and the central agent updates the multipliers in the form:

$$\overline{\lambda} = \overline{\lambda} + \rho\, h(x, y) = \begin{bmatrix} 0.01 \\ 0.01 \end{bmatrix} + 0.1 \begin{bmatrix} 0.5972 \\ 0.5972 \end{bmatrix} = \begin{bmatrix} 0.0697 \\ 0.0697 \end{bmatrix}.$$

Next, the parameter ρ is increased

$$\rho^1 = 1.2\rho = 0.12,$$

and the variables

$$\overline{x} = x = \begin{bmatrix} 0.0806 \\ 0.0 \end{bmatrix}, \quad \overline{y} = y = \begin{bmatrix} 0.0 \\ 0.0806 \end{bmatrix},$$

are updated, $k = k+1 = 2$, and Steps 1, 2, and 3 of the algorithm are repeated until convergence is achieved.

The algorithm stops when $k = 16$, $\|h\| = 9.9172 \times 10^{-5}$,

$$x = \begin{bmatrix} 0.2 \\ 0.0 \end{bmatrix}, \quad y = \begin{bmatrix} 0.0 \\ 0.2 \end{bmatrix}, \quad \lambda = \begin{bmatrix} 0.08 \\ 0.08 \end{bmatrix}, \quad \rho = 1.5407.$$

c) Proposed Decomposition Algorithm Solution

Subproblems to be solved in Steps 1 and 2 of the proposed Decomposition Algorithm are, respectively,

$$\begin{array}{ll} \text{minimize} & x_1^2 + x_2^2 - \overline{\lambda}_2\, x_1 \\ \text{subject to} & 1 - 4x_1 - \overline{y}_2 = 0 \end{array} \tag{46}$$

and

$$\begin{array}{ll} \text{minimize} & y_1^2 + y_2^2 - \overline{\lambda}_1\, y_1 \\ \text{subject to} & 1 - \overline{x}_1 - 4y_2 = 0 \end{array} \tag{47}$$

Iteration $k = 0$:

Step 0. Variables and multipliers

$$\overline{x} = \begin{bmatrix} 0.4 \\ 0.4 \end{bmatrix}, \quad \overline{y} = \begin{bmatrix} 0.4 \\ 0.4 \end{bmatrix}, \quad \overline{\lambda} = \begin{bmatrix} 0.01 \\ 0.01 \end{bmatrix}$$

are initialized.

Iteration $k = 1$:

Step 1. System X computes a movement direction for problem (46), by Newton's method, at $x = \overline{x}$:

$$\begin{bmatrix} 2 & 0 & -4 \\ 0 & 2 & 0 \\ -4 & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta \lambda_1 \end{bmatrix} = - \begin{bmatrix} 0.75 \\ 0.80 \\ -1.00 \end{bmatrix},$$

and obtains

$$x = x + \Delta x = \begin{bmatrix} 0.4 \\ 0.4 \end{bmatrix} + \begin{bmatrix} -0.25 \\ -0.40 \end{bmatrix} = \begin{bmatrix} 0.15 \\ 0.00 \end{bmatrix},$$

and

$$\lambda_1 = \lambda_1 + \Delta \lambda_1 = 0.01 + 0.0625 = 0.0725.$$

Step 2. System Y computes a movement direction for problem (47), by Newton's method, at $y = \overline{y}$:

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & -4 \\ 0 & -4 & 0 \end{bmatrix} \begin{bmatrix} \Delta y_1 \\ \Delta y_2 \\ \Delta \lambda_2 \end{bmatrix} = - \begin{bmatrix} 0.80 \\ 0.75 \\ -1.00 \end{bmatrix},$$

and obtains

$$y = y + \Delta y = \begin{bmatrix} 0.4 \\ 0.4 \end{bmatrix} + \begin{bmatrix} -0.40 \\ -0.25 \end{bmatrix} = \begin{bmatrix} 0.00 \\ 0.15 \end{bmatrix},$$

and

$$\lambda_2 = \lambda_2 + \Delta \lambda_2 = 0.01 + 0.0625 = 0.0725.$$

Step 3. Convergence. Multiplier updating. The central agent checks if convergence condition $\|h(x, y)\| < 10^{-4}$ is verified:

$$h = \begin{bmatrix} 0.25 \\ 0.25 \end{bmatrix}, \quad \|h\| = 0.3536 > 10^{-4}.$$

The convergence condition is not verified, and the variables

$$\overline{x} = x = \begin{bmatrix} 0.15 \\ 0.0 \end{bmatrix}, \quad \overline{y} = y = \begin{bmatrix} 0.0 \\ 0.15 \end{bmatrix}, \quad \overline{\lambda} = \lambda = \begin{bmatrix} 0.0725 \\ 0.0725 \end{bmatrix}$$

are fixed, $k = k + 1 = 2$, and Steps 1, 2, and 3 of the algorithm are repeated until convergence is achieved.

The algorithm stops when $k = 8$, $\|h\| = 8.6317 \times 10^{-5}$,

$$x = \begin{bmatrix} 0.2 \\ 0.0 \end{bmatrix}, \quad y = \begin{bmatrix} 0.0 \\ 0.2 \end{bmatrix}, \quad \lambda = \begin{bmatrix} 0.08 \\ 0.08 \end{bmatrix}.$$

The behaviour of the three algorithms is analyzed in the following figures. Fig. 1 shows the evolution of the objective function (39) at each iteration, for each of the three procedures. The dashed line represents the evolution of the objective function evaluated at the iterates for the Lagrangian Relaxation procedure. The dotted line represents the evolution of the objective function evaluated at the iterates for the Augmented Lagrangian Relaxation procedure. Lastly, the solid line represents the evolution of the objective function evaluated at points computed by the proposed decomposition algorithm.

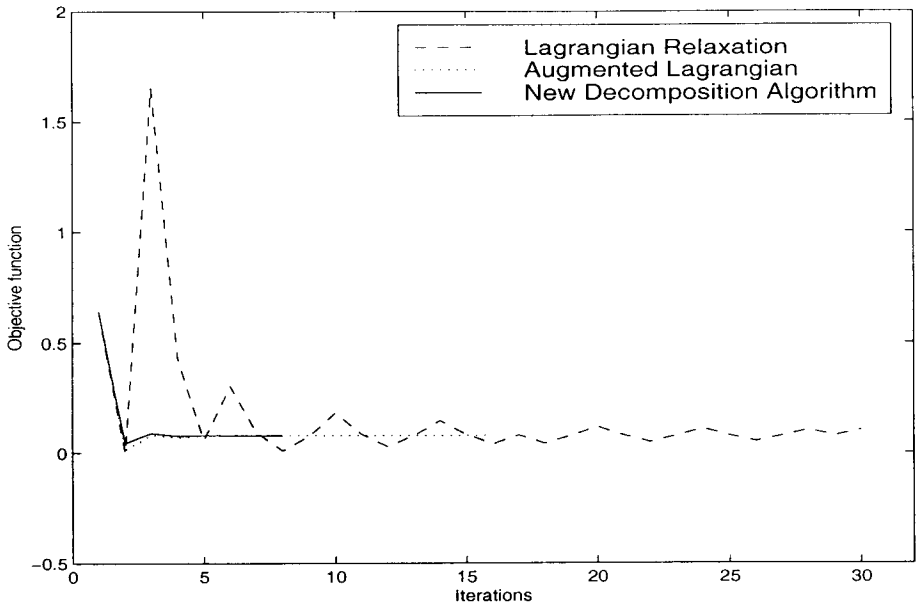


Figure 1: Example. Evolution of the objective function.

Fig. 2 shows the value of multiplier λ_1, at each iteration, for each of the three procedures. The value of multiplier λ_2 coincides, at each iteration, with the value of multiplier λ_1 for each procedure. As in fig. 1, the dashed line represents the values of the multiplier computed by the Lagrangian Relaxation procedure; the dotted line represents the evolution of the multiplier from the Augmented Lagrangian Relaxation procedure; lastly, the solid line represents the evolution of the multiplier as obtained by the proposed decomposition algorithm.

The slow and oscillating behaviour of the Lagrangian Relaxation procedure is apparent. The quadratic penalization term in the Augmented Lagrangian procedure corrects this anomaly, although the convergence is slower than that of the decomposition algorithm.

VI Local Convergence

In this section local convergence properties for the proposed Decomposition Algorithm are analyzed. It is proved that the local rate of convergence of the proposed Decomposition Algorithm is linear. It must be noted that the rate

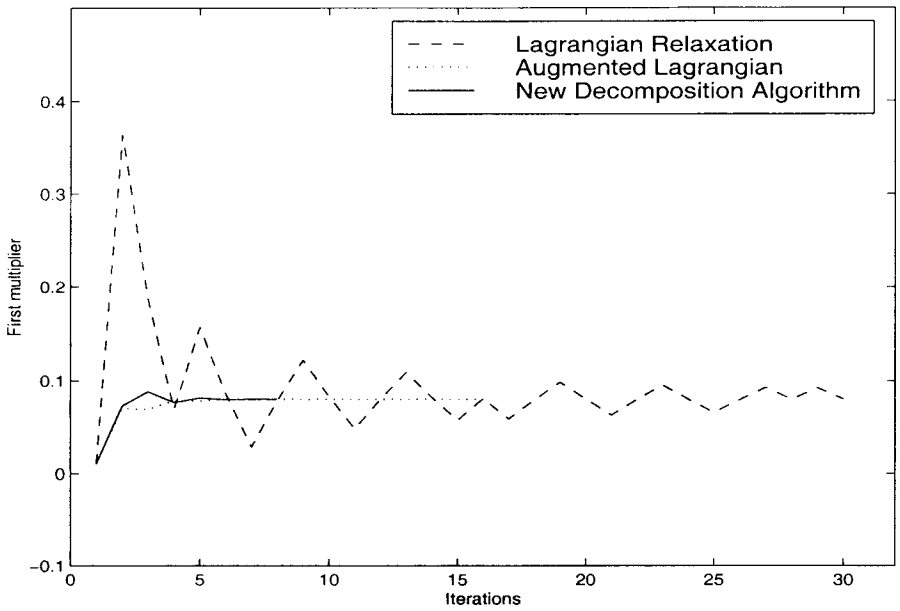


Figure 2: Example. Evolution of first multiplier.

of convergence of a centralized approach could be quadratic.

The notation used in demonstrations is stated below

$\nabla_z g$	gradient, with respect to z, of scalar function g
$\nabla_{wz} g$	hessian, first with respect to w and then with respect to z, of scalar function g
$J_z h$	jacobian, with respect to z, of function h
$\mathcal{L}$	Lagrangian function
$\rho(A)$	spectral radius of matrix A
$\|\|A\|\|$	norm of matrix A
$k(A)$	condition number of matrix A.

Movement directions in Steps 1 and 2 are obtained computing one *decomposed* Newton's iteration for the KKT system of problem (4)-(8), see Appendix A. For the sake of simplicity in this analysis, separable constraints (7) and (8) can be omitted. These constraints can be handled by introducing them into the objective function using an interior point procedure, for example.

The KKT system corresponding to the first-order necessary conditions for

problem (4)-(6), see Appendix A, is

$$\begin{bmatrix} \mathrm{KKT}_X & \mathrm{KKT}_{YX} \\ \mathrm{KKT}_{XY} & \mathrm{KKT}_Y \end{bmatrix} \begin{bmatrix} \Delta_X^N \\ \Delta_Y^N \end{bmatrix} = - \begin{bmatrix} \nabla_{x,\lambda_X} \mathcal{L} \\ \nabla_{y,\lambda_Y} \mathcal{L} \end{bmatrix}, \tag{48}$$

where

$$\mathcal{L}(x, \lambda_X, y, \lambda_Y) = f_X(x) + f_Y(y) + \lambda_X^T \, h_X(x,y) + \lambda_Y^T \, h_Y(x,y)$$

$$\mathrm{KKT}_X = \begin{bmatrix} \nabla_{xx}\mathcal{L} & J_x^T h_X \\ J_x h_X & 0 \end{bmatrix} \quad \mathrm{KKT}_Y = \begin{bmatrix} \nabla_{yy}\mathcal{L} & J_y^T h_Y \\ J_y h_Y & 0 \end{bmatrix}$$
$$\mathrm{KKT}_{YX} = \begin{bmatrix} \nabla_{yx}\mathcal{L} & J_x^T h_Y \\ J_y h_X & 0 \end{bmatrix} \quad \mathrm{KKT}_{XY} = \begin{bmatrix} \nabla_{xy}\mathcal{L} & J_y^T h_X \\ J_x h_Y & 0 \end{bmatrix},$$

and Newton's directions Δ_X^N, Δ_Y^N are

$$\Delta_X^N = \begin{bmatrix} \Delta x \\ \Delta \lambda_X \end{bmatrix}$$
$$\Delta_Y^N = \begin{bmatrix} \Delta y \\ \Delta \lambda_Y \end{bmatrix}.$$

The following notation is useful in the study of convergence. At iteration k of the algorithm, z_k denotes the point $(x, \lambda_X, y, \lambda_Y)$, Δ_k denotes the decomposed direction (Δ_X, Δ_Y), and Δ_k^N denotes Newton's direction (Δ_X^N, Δ_Y^N). Also, matrices

$$\mathrm{KKT}_k = \begin{bmatrix} \mathrm{KKT}_X & \mathrm{KKT}_{YX} \\ \mathrm{KKT}_{XY} & \mathrm{KKT}_Y \end{bmatrix}$$
$$\overline{\mathrm{KKT}}_k = \begin{bmatrix} \mathrm{KKT}_X & 0 \\ 0 & \mathrm{KKT}_Y \end{bmatrix}$$

denote the KKT Newton matrix and KKT decomposed matrix, respectively, at iteration k.

The following theorem shows a sufficient condition for convergence of the Decomposition Algorithm when Steps 1 and 2 are computed in a parallel implementation.

Theorem 1. *Assume that Steps 1 and 2 are solved simultaneously and functions in* (4)-(8) *are supposed to be twice continuously differentiable. If the following condition holds for an optimal solution z^* of problem* (4)-(8) *(where matrix I denotes the identity matrix),*

$$\rho(I - \overline{KKT}^{-1} KKT) < 1 \tag{49}$$

then, the proposed decomposition algorithm converges locally to z^* with linear rate of convergence.

Proof. If Steps 1 and 2 are computed independently of each other, the decomposed system is:

$$\begin{bmatrix} \mathrm{KKT}_X & 0 \\ 0 & \mathrm{KKT}_Y \end{bmatrix} \begin{bmatrix} \Delta_X \\ \Delta_Y \end{bmatrix} = - \begin{bmatrix} \nabla_{x,\lambda_X} \mathcal{L} \\ \nabla_{y,\lambda_Y} \mathcal{L} \end{bmatrix}. \tag{50}$$

Then, the Decomposition Algorithm updates variables at iteration $k+1$ as

$$z_{k+1} = z_k + \Delta_k$$

where

$$\Delta_k = -\overline{\mathrm{KKT}}_k^{-1} \nabla_z \mathcal{L}_k.$$

The error, at iteration k of the algorithm, is denoted by $e_k = z_k - z^*$, where z^* is an optimal solution of problem (4)-(6). Using Taylor's theorem, Appendix A,

$$0 = \nabla \mathcal{L}^* = \nabla \mathcal{L}(z_k - e_k) \simeq \nabla \mathcal{L}_k - \mathrm{KKT}_k \; e_k.$$

If this equation is multiplied by $\overline{\mathrm{KKT}}_k^{-1}$ then,

$$\begin{aligned} 0 &= -\overline{\mathrm{KKT}}_k^{-1} \; \nabla \mathcal{L}_k + \overline{\mathrm{KKT}}_k^{-1} \; \mathrm{KKT}_k \; e_k \\ &= \Delta_k + \overline{\mathrm{KKT}}_k^{-1} \; \mathrm{KKT}_k \; e_k. \end{aligned}$$

Expanding Δ_k the following equation is obtained

$$z_{k+1} = z_k - \overline{\mathrm{KKT}}_k^{-1} \; \mathrm{KKT}_k \; e_k,$$

and

$$e_{k+1} = \left(I - \overline{\mathrm{KKT}}_k^{-1} \; \mathrm{KKT}_k \right) e_k. \tag{51}$$

This result shows that, if the proposed Decomposition Algorithm converges, then convergence is linear. A sufficient condition is shown next for convergence of the algorithm.

Equation (51) shows that $\overline{\mathrm{KKT}}_k$ is a good approximation of KKT_k if the following condition holds

$$\rho(I - \overline{\mathrm{KKT}}_k^{-1}\mathrm{KKT}_k) < 1 \tag{52}$$

for k large enough. But this is true because of (49) and the continuity of the coefficients in systems $\overline{\mathrm{KKT}}$ and KKT. □

Some additional comments on this condition is introduced. Condition (49) is related to the separable structure of the problem and in the multi-area OPF problem it can be interpreted as a measurement of the coupling between the areas in the global system. This measure tends to be smaller for systems with a small number of interconnecting lines.

Matrix KKT_k can be expressed as $\mathrm{KKT}_k = \overline{\mathrm{KKT}}_k + E_k$, and convergence condition (52) is

$$\boxed{\rho(\overline{\mathrm{KKT}}_k^{-1}E_k) < 1} \tag{53}$$

for points near z^*.

Also, condition (53) can be expressed in terms of the condition number of matrix $\overline{\mathrm{KKT}}_k$. The following result is true for any matrix norm that satisfies the submultiplicative property [11]

$$\rho(\overline{\mathrm{KKT}}_k^{-1}E_k) < \|\overline{\mathrm{KKT}}_k^{-1}E_k\|,$$

and then

$$\|\overline{\mathrm{KKT}}_k^{-1}E_k\| \leq \|\overline{\mathrm{KKT}}_k\|\ \|\overline{\mathrm{KKT}}_k^{-1}\|\ \frac{\|E_k\|}{\|\overline{\mathrm{KKT}}_k\|} = k(\overline{\mathrm{KKT}}_k)\ \frac{\|E_k\|}{\|\overline{\mathrm{KKT}}_k\|}.$$

Therefore, if

$$\frac{\|E_k\|}{\|\overline{\mathrm{KKT}}_k\|} < \frac{1}{k(\overline{\mathrm{KKT}}_k)} \tag{54}$$

is verified, then (53) holds. An interpretation of (54) is the following: as the condition numbers of the systems in X and Y increase then the first and second derivatives of complicating equations must become small enough in comparison with the rest of equations and constraints.

Finally, it must be remarked that conditions (52) or (53) are the same ones as those in Jacobi and Gauss-Seidel methods for solving linear systems [11].

Then, these techniques can be used to improve the performance of the proposed Decomposition Algorithm.

Alternatively, Steps 1 and 2 could be implemented one after the other. The following theorem shows a sufficient condition for convergence of the Decomposition Algorithm when Steps 1 and 2 are computed one after the other, *serial implementation.*

Theorem 2. *Assume that Steps 1 and 2 are computed in a serial implementation and functions in* (4)-(8) *are supposed to be twice continuously differentiable. If the following condition holds for an optimal solution* z^* *of problem* (4)-(8) *(where matrix* I *denotes the identity matrix),*

$$\rho(KKT_X^{-1} KKT_{YX} KKT_Y^{-1} KKT_{XY}) < 1 \tag{55}$$

then, the proposed decomposition algorithm converges locally to z^* *with linear rate of convergence.*

Proof. From (48) it follows that

$$\Delta_X^N = -(\mathrm{KKT}_X - \mathrm{KKT}_{YX}\ \mathrm{KKT}_Y^{-1}\ \mathrm{KKT}_{XY})^{-1} (\nabla_x \mathcal{L} - \mathrm{KKT}_{YX}\ \mathrm{KKT}_Y^{-1}\ \nabla_y \mathcal{L}) \tag{56}$$

and

$$\Delta_Y^N = -(\mathrm{KKT}_Y - \mathrm{KKT}_{XY}\ \mathrm{KKT}_X^{-1}\ \mathrm{KKT}_{YX})^{-1} (\nabla_y \mathcal{L} - \mathrm{KKT}_{XY}\ \mathrm{KKT}_X^{-1}\ \nabla_x \mathcal{L}). \tag{57}$$

Then, Newton's direction for system X, if the decomposed direction for system Y is known, is given by

$$\Delta_X^N = -(\mathrm{KKT}_X - \mathrm{KKT}_{YX}\ \mathrm{KKT}_Y^{-1}\ \mathrm{KKT}_{XY})^{-1} (\nabla_x \mathcal{L}\ \mathrm{KKT}_{YX}\ \Delta_Y)$$

and

$$(\mathrm{KKT}_X - \mathrm{KKT}_{YX}\ \mathrm{KKT}_Y^{-1}\ \mathrm{KKT}_{XY})^{-1} \simeq \mathrm{KKT}_X^{-1}$$

is true if

$$\boxed{\rho(\mathrm{KKT}_X^{-1}\mathrm{KKT}_{YX}\mathrm{KKT}_Y^{-1}\mathrm{KKT}_{XY}) < 1} \tag{58}$$

holds. □

Next, the relationship between the convergence condition in a parallel implementation (58) and the convergence condition in a serial implementation (53) is analyzed. If condition (52) is expanded, then

$$\rho(I - \overline{\mathrm{KKT}}_k^{-1}\mathrm{KKT}_k) = \rho\left(\begin{bmatrix} 0 & -\mathrm{KKT}_X^{-1}\mathrm{KKT}_{YX} \\ -\mathrm{KKT}_Y^{-1}\mathrm{KKT}_{XY} & 0 \end{bmatrix}\right). \quad (59)$$

If $\begin{bmatrix} x \\ y \end{bmatrix}$ is an eigenvector of matrix $I - \overline{\mathrm{KKT}}_k^{-1}\mathrm{KKT}_k$ with an associated eigenvalue μ, then

$$\begin{aligned} -\mathrm{KKT}_X^{-1}\ \mathrm{KKT}_{YX}\ y &= \mu x \\ -\mathrm{KKT}_Y^{-1}\ \mathrm{KKT}_{XY}\ x &= \mu y. \end{aligned}$$

Thus,

$$\rho(\mathrm{KKT}_X^{-1}\mathrm{KKT}_{YX}\mathrm{KKT}_Y^{-1}\mathrm{KKT}_{XY})\ x = \mu^2 x. \quad (60)$$

holds and, by the last equation

$$\rho(I - \overline{\mathrm{KKT}}_k^{-1}\mathrm{KKT}_k)^2 = \rho(\mathrm{KKT}_X^{-1}\mathrm{KKT}_{YX}\mathrm{KKT}_Y^{-1}\mathrm{KKT}_{XY}). \quad (61)$$

Therefore, if the Decomposition Algorithm converges, (53) or (58) is verified, then, the ratio of convergence is better in a serial implementation than in a parallel implementation, as (61) shows, and as it would be expected.

VII Global Convergence

Global convergence properties require that new point z_{k+1} be a *better* point than z_k. This is achieved by decreasing a given merit function sufficiently. This reduction is attained in Step 3 of the proposed decomposition algorithm.

The merit function measures progress toward the solution z^*. If $\Phi_m(z_k + \alpha\Delta_k)$ denotes the merit function at the new point $z_k + \alpha\Delta_k$, this function is redefined as $\Phi(\alpha)$ in order to make dependence of α clearer.

In order to reduce the progress toward the solution, it is necessary that $\Phi_m(\alpha) < \Phi_m(0)$. But this condition must be satisfied sufficiently. A *good*

step size should be satisfy

$$\Phi_m(\alpha) \leq \Phi_m(0) + \overline{\sigma}\alpha\Phi'_m(0) \qquad \text{(sufficient-decrease condition)} \qquad (62)$$

$$|\Phi'_m(\alpha)| \leq -\overline{\eta}\Phi'_m(0) \qquad \text{(Wolfe condition)} \qquad (63)$$

where $0 < \overline{\sigma} < \overline{\eta} < 1$. This condition implies that $\Phi'_m(0)$ must be negative.

The line search procedure in Step 3 of the decomposition algorithm is the following:

Line search:

Paso 0
Initial value of α_0 is selected (maximum feasible step).

Paso 1
Condition $\Phi'_m(0) < 0$ must be ensured.

Paso 2
A value $\alpha \in (0, \alpha_0)$ is selected such that (62) or (62)-(63) is satisfied.

An appropriate merit function for problem (4)-(6), based on the Augmented Lagrangian function, is the following

$$\Phi_m = f_X(x) + f_Y(y) + \lambda_X^T\, h_X(x,y) + \lambda_Y^T\, h_Y(x,y) + \\ + \frac{\rho}{2}\|h_X(x,\overline{y})\|^2 + \frac{\rho}{2}\|h_Y(\overline{x},y)\|^2. \quad (64)$$

It can be proved that

$$\Phi'_m(0) = \nabla_x^T\mathcal{L}\ \Delta x + \nabla_y^T\mathcal{L}\ \Delta y + h_X^T\ \Delta\lambda_X + h_Y^T\ \Delta\lambda_Y - \rho\|h_X\|^2 - \rho\|h_Y\|^2.$$

If

$$\theta = \nabla_x^T\mathcal{L}\ \Delta x + \nabla_y^T\mathcal{L}\ \Delta y + h_X^T\ \Delta\lambda_X + h_Y^T\ \Delta\lambda_Y,$$

then

$$\Phi'_m(0) = \theta - \rho\|h\|^2,$$

where $h = \begin{bmatrix} h_X \\ h_Y \end{bmatrix}$.

If $h \neq 0$ (infeasible case), it is possible to ensure that $\Phi'_m(0) < -u$ with $u > 0$. If $\Phi'_m(0) > 0$, then $\theta > 0$ and the penalty parameter can be chosen as

$$\rho > \frac{\theta + u}{\|h\|^2}.$$

If $h = 0$ (feasible case), then

$$\Phi'_m(0) = \nabla_x^T \mathcal{L}\ \Delta x + \nabla_y^T \mathcal{L}\ \Delta y,$$

and as Δx y Δy are the decomposed Newton directions, then

$$\Phi'_m(0) = -(\Delta x)^T\ \nabla_{xx} \mathcal{L}\ \Delta x\ - (\Delta y)^T\ \nabla_{yy} \mathcal{L}\ \Delta y,$$

and this direccional derivative is negative if matrices $\nabla_{xx}\mathcal{L}$ y $\nabla_{yy}\mathcal{L}$ are positive definite.

The line search can be computed in a parallel implementation. In the infeasible case, the penalty parameter can be chosen as $\rho = \rho_X + \rho_Y$ where

$$\rho_X > \frac{\theta_X + 0.5u}{\|h_X\|^2} \tag{65}$$

$$\rho_Y > \frac{\theta_Y + 0.5u}{\|h_Y\|^2} \tag{66}$$

and

$$\theta_X = \nabla_x^T \mathcal{L}\ \Delta x \tag{67}$$

$$\theta_Y = \nabla_y^T \mathcal{L}\ \Delta y, \tag{68}$$

because

$$\frac{\theta_X + 0.5u}{\|h_X\|^2} + \frac{\theta_Y + 0.5u}{\|h_Y\|^2} > \frac{\theta + u}{\|h\|^2}.$$

VIII Numerical Results

This section presents numerical results obtained by applying the proposed decomposition algorithm to several test problems. Cases I and II correspond to the same IEEE-9 system, and only differ in the number of interconnecting lines. Case III is based on the IEEE-30 system; this system has been triplicated with seven interconnecting lines. Cases IV and V are based on the

Table 1: Main characteristics and numerical results for case studies.

Case	Areas	Var.	Constr.	Algor.	Iter	Ratio
I	2	24	27	Global	16	
				Decomp.	23	1.4375
II	2	24	27	Global	16	
				Decomp.	42	2.6250
III	3	72	101	Global	23	
				Decomp.	59	2.5652
IV	2	224	167	Global	27	
				Decomp.	54	2.0000
V	3	336	251	Global	29	
				Decomp.	73	2.5172
VI	3	1032	1344	Global	42	
				Decomp.	43	1.0238

IEEE RTS24 [12]; this system has been duplicated for case IV, and triplicated for case V, based on [13]. Finally, case VI is based on the IEEE-118 system; this system has been triplicated with ten interconnecting lines. In all cases, the objective function (1) is taken to be the total operation cost for the system.

All cases have been solved by a centralized approach and a decentralized one, using the proposed decomposition algorithm. The solutions for both approaches have been computed using a nonlinear interior point procedure, based on a version of [14].

The model and algorithms have been implemented in MATLAB [15]. The two procedures have been initialized using the same starting point.

Table 1 shows the most relevant characteristics and the numerical results for each case study.

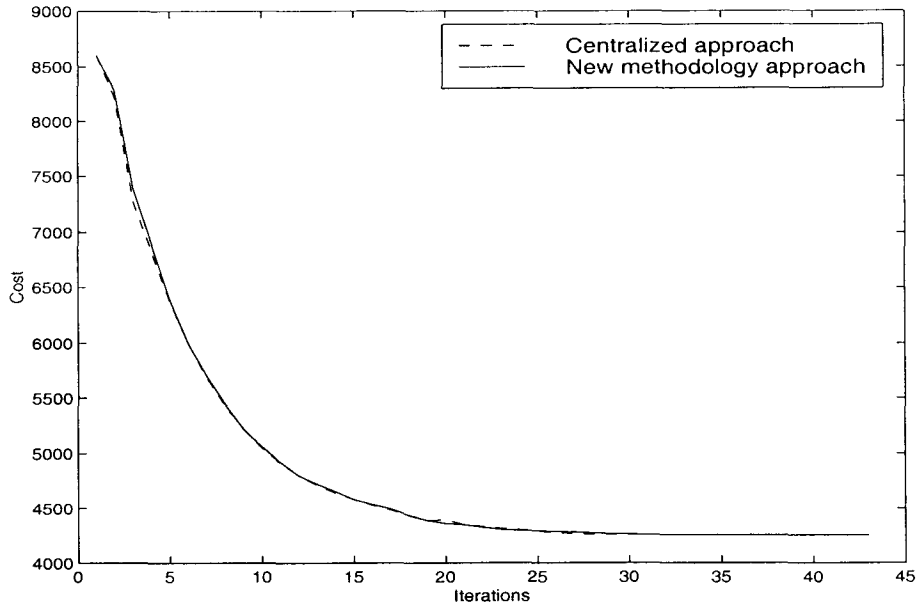


Figure 3: Case VI. Total cost in US $.

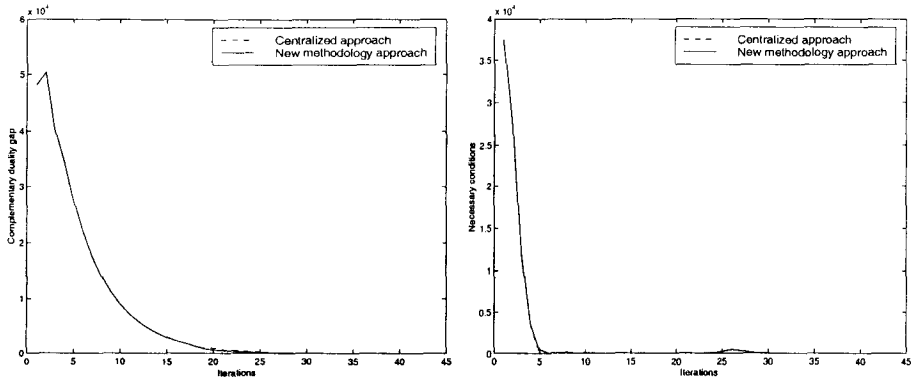
The second column in Table 1 shows the number of areas for the global system. The third and forth columns are total number of variables and constraints, respectively, for each case study. The fifth column indicates the approach used to solved each case, a global procedure or a decomposed one. The sixth column shows the total number of iterations required to reach the optimum for each problem. And the last column shows the ratio between total iterations for the decentralized approach and total iterations for the centralized approach.

To further illustrate the behavior of the decomposition procedure, fig. 3 shows, for Case VI, the evolution of the objective function (1), at each iteration, for each of the two procedures. The dashed line represents the evolution of the objective function evaluated at points obtained by the centralized approach. The solid line represents the evolution of the objective function evaluated at points obtained by the new methodology approach.

Fig. 4 shows, for Case VI, the evolution of the complementarity duality gap [14] and the norm of the first-order necessary conditions, at each iteration, for each of the two procedures.

It must be noted that, in this case, both the centralized and decentralized approach provide nearly the same results.

Figure 4: Case VI.



(a) Complementary duality gap (b) $1^{st.}$ order necessary conditions.

It must be noted that although the total number of iterations for the proposed Decomposition Algorithm is larger than that for the centralized procedure, total computing time may be smaller. This is due to the size of the linear systems that need to be solved by each procedure. A centralized approach should solve systems that are larger by a factor equal to the number of areas than those solved by the decomposition algorithms.

IX Conclusions

In this report a new decomposition methodology is presented and applied to a problem arising in the analysis of interconnected power systems. The new methodology preserves the autonomy of each utility in the global system by means of a coordinated but decentralized procedure.

Numerical results show that the method has less computational cost than a centralized approach and other alternative decomposition techniques.

As a result of these properties, the new methodology is very well-suited for its use in the solution of large-scale multi-area OPF problems, and in general in optimization problems that can be separated by fixing the values of some of their variables.

A Taylor's Series and KKT Systems

This appendix shows some results related to the Taylor's series of a given multivariate function. This is a useful result in optimization, and in particular in the preceding analysis of the proposed Decomposition Analysis convergence. Also, a KKT system of an optimization problem is defined. Notation of Section VI is used.

It is necessary to introduce the following terminology. Given two non-negative sequences, ν_k and e_k, the notation

$$\nu_k = O(e_k)$$

indicates that a real number $C > 0$ exists, such that,

$$\nu_k \leq C\ e_k$$

for all k.

The next theorem is a particular case of Taylor's Theorem.

Theorem 3. *If $x \in \mathbb{R}^n$, and $f(x) \in \mathbb{R}$ is a function with continuous second derivatives, then real number θ exists satisfying $0 \leq \theta \leq 1$, and such that*

$$f(x+p) = f(x) + p^T\ \nabla_x f(x) + \frac{1}{2}\ p^T\ \nabla_{xx} f(x+\theta\ p)\ p. \tag{69}$$

Therefore, it can be concluded that

$$|f(x+p) - f(x) - p^T\ \nabla_x f(x)| = O(\|p\|^2).$$

Conditions

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0 \tag{70}$$
$$h(x^*) = 0 \tag{71}$$

form a nonlinear system of $n+m$ equations and $n+m$ variables for problem

$$\begin{aligned} &\text{minimize} && f(x) && (72)\\ &\text{subject to} && h(x) = 0 && (73) \end{aligned}$$

where n is the dimension of vector x, and m is the dimension of vector $h(x)$. If z denotes the set of variables (x^*, λ^*), and the function $F(z)$ denotes the set of conditions (70)-(71), then

$$|F_j(z+\Delta) - F_j(z) - \Delta^T \nabla_z F_j(z)| = O(\|\Delta\|^2),$$

for $j = 1, \ldots, n+m$. Then, it holds that

$$F_j(z+\Delta) \simeq F_j(z) + \Delta^T \nabla_z F_j(z) \tag{74}$$

for $\|\Delta\|$ small enough.

In optimization algorithms, it is desirable to find a direction Δ such that $F(z+\Delta) = 0$, and so, this value can be approximately computed as

$$F(z) + \Delta^T \nabla_z F(z)$$

where the term $\nabla_z F(z)$ can be expressed in terms of original variables and functions as

$$\nabla_{x,\lambda} F(x,\lambda) = \begin{bmatrix} \nabla_{xx}\mathcal{L}(x,\lambda) & J_x h(x) \\ J_x h(x) & 0 \end{bmatrix}$$

Matrix $\nabla_{x,\lambda} F(x,\lambda)$ is known as the *KKT matrix* [16] for problem (72)-(73). System

$$\begin{bmatrix} \nabla_{xx}\mathcal{L}(x,\lambda) & J_x h(x) \\ J_x h(x) & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \end{bmatrix} = - \begin{bmatrix} \nabla_x \mathcal{L}(x,\lambda) \\ h(x) \end{bmatrix}$$

is the *KKT system* for problem (72)-(73). This system can be formulated in the form

$$\text{KKT}\ \Delta = -\nabla_{(x,\lambda)}\mathcal{L}$$

where

$$\text{KKT} = \begin{bmatrix} \nabla_{xx}\mathcal{L}(x,\lambda) & J_x h(x) \\ J_x h(x) & 0 \end{bmatrix}, \quad \Delta = \begin{bmatrix} \Delta x \\ \Delta\lambda \end{bmatrix}, \quad \nabla_{(x,\lambda)}\mathcal{L} = \begin{bmatrix} \nabla_x \mathcal{L}(x,\lambda) \\ h(x) \end{bmatrix}.$$

References

[1] A. J. Wood and B. F. Wollenberg. *Power Generation, Operation and Control.* John Wiley and Sons, New York, second edition, 1996.

[2] M. S. Bazaraa, H. O. Sherali, and C. M. Shetty. *Nonlinear Programming: Therory and Algorithms.* John Wiley and Sons, New York, second edition, 1993.

[3] J. E. Kelley. The cutting-plane method for solving coonvex programs. *J. Soc. Indust. Appl. Math.*, 8:703–712, 1960.

[4] B. T. Polyak. *Introduction to Optimization.* Optimization Software, Inc., New York, 1987.

[5] P. Carpentier, G. Cohen, J.-C. Culioli, and A. Renaud. Stochastic optimization of unit commitment: a new decomposition framework. *IEEE Transactions on Power Systems*, 11(2):1067–1073, May 1996.

[6] G. Cohen. Auxiliary problem principle and decomposition of optimization problems. *Journal of Optimization Theory and Applications*, 32(3), Nov 1980.

[7] A. Renaud. Daily generation management at electricité de france: From planning towards real time. *IEEE Transactions on Automatic Control*, 38(7):1080–1093, Jul 1993.

[8] A. J. Conejo and J. A. Aguado. Multi-area coordinated decentralized dc optimal power flow. *IEEE Transactions on Power Systems*, 13(4):1272–1278, Nov 1998.

[9] R. Baldick, B. H. Kim, C. Chase, and Y. Luo. A fast distributed implementation of optimal power flow. Accepted for publication in the IEEE Transactions on Power Systems, 1998.

[10] N. Jiménez and A. J. Conejo. Short-term hydro-thermal coordination by lagrangian relaxation: Solution of the dual problem. *IEEE Transactions on Power Systems*, 14(1):89–95, Feb 1999.

[11] G. H. Golub and C. F. Van Loan. *Matrix Computations.* The Johns Hopkins University Press, London, third edition, 1996.

[12] IEEE Committee Report. Ieee reliability test system. *IEEE Transactions on Power Systems*, 98(6):2047–2054, Nov/Dec 1979.

[13] F. Xia and A. P. Sakis Meliopoulos. A methodology for probabilistic simultaneous transfer capability analysis. *IEEE Transactions on Power Systems*, 11(3):1269–1276, Aug 1996.

[14] Y. Wu, A. S. Debs, and R. E. Marsten. A direct nonlinear predictor-corrector primal-dual interior point algorithm for optimal power flow. *IEEE Transactions on Power Systems*, 9(2):876–883, May 1994.

[15] The MathWorks Inc. *Using Matlab. Version 5.* 1998.

[16] D. P. Bertsekas. *Nonlinear Programming.* Athena Scientific, Belmont, Massachusetts, 1995.