

A Bundle Method for Hydrothermal Scheduling

Daoyuan Zhang, Peter B. Luh (Fellow, IEEE), and Yuanhui Zhang
 Department of Electrical Engineering
 University of Connecticut, Storrs, CT 06269-2157

Abstract

Lagrangian relaxation has been widely used in hydrothermal scheduling. Complicating constraints are relaxed by multipliers which are usually updated by a subgradient method (SGM). The SGM suffers from slow convergence caused by the non-differentiable characteristics of dual functions. This paper presents an algorithm that utilizes the Bundle Trust Region Method (BTRM) to update the multipliers within the Lagrangian relaxation framework. The BTRM is shown to converge faster than the SGM as well as other bundle type methods in optimizing non-differentiable dual functions. The application of BTRM for solving hydro subproblems results in greatly improved convergence over the SGM. Comparing BTRM with another bundle type method in updating the high level multipliers shows that better solution can be obtained by BTRM.

Keywords: Hydrothermal Scheduling, Bundle Methods, Lagrangian Relaxation, Identical Units

1. Introduction

Hydrothermal scheduling is concerned with the commitment and dispatch of generating units. The objective is to minimize the total generation cost over a period of up to one week, subject to system-wide demand and reserve requirements and individual unit constraints. This mixed integer programming problem is believed to be NP-hard, and Lagrangian relaxation has been very successful in obtaining solutions with quantifiable quality. The Lagrangian multipliers are usually updated by a subgradient method (SGM) [1, 2], although it may converge slowly, especially for systems with identical units.

Recently, bundle-type methods have been used to update the multipliers, and improvements on scheduling quality have been reported [3, 4, 5]. The major difference between

bundle-type methods and SGMs is the way for calculating searching directions. An SGM uses a subgradient as a direction, whereas bundle-type methods use a convex combination of subgradients accumulated in a "bundle" to form a search direction. Obtaining the combination coefficients usually involves quadratic programming, and this can be quite time consuming as the bundle size increases. In [3], orthogonal projection is used in lieu of quadratic programming to calculate directions to reduce the complexity. Another time-consuming part for a bundle method is the line search which decides how far to move along a direction, and obtaining an accurate step size is generally a difficult task. To overcome these difficulties, the bundle trust region method (BTRM) developed in [6] is used to update the multipliers for hydrothermal scheduling. In BTRM, the line search is integrated with quadratic programming, and is more accurate and efficient. In the implementation, the recursive active set method [7] is used to solve the quadratic programming problem to reduce computation requirements. The method thus effectively overcomes convergence difficulties associated with SGM. Formulation of the hydrothermal scheduling problem is presented in Section 2, and methods for updating multipliers are presented in Section 3. The BTRM and its implementation for hydrothermal scheduling are presented in Section 4.

In Section 5, extensive testing results will be presented to demonstrate the efficiency of BTRM for hydrothermal scheduling. For a problem with identical units, the dual function is shown to have sharp "ridges" making SGM difficult to converge whereas BTRM does not have difficulty at all. In its application to hydro subproblems, BTRM has much faster convergence than SGM. Testing using data sets from Northeast Utilities shows that BTRM can significantly improve scheduling quality in shorter CPU time by comparison with another bundle-type method.

2. Problem Formulation and Solution Methodology

2.1 Problem Formulation

Consider a power system with I thermal units, J hydro units, and K pumped-storage units. The objective is to minimize the total generation cost subject to system-wide demand and reserve requirements, and individual unit constraints.

The **objective function** to be minimized is the sum of thermal generation costs $C_{fi}(p_{fi}(t))$ and start up costs $S_{fi}(t)$, i.e.,

PE-457-PWRS-0-04-1998 A paper recommended and approved by the IEEE Power System Analysis, Computing and Economics Committee of the IEEE Power Engineering Society for publication in the IEEE Transactions on Power Systems. Manuscript submitted August 4, 1997; made available for printing May 18, 1998.

$$\min_{p_{ti}(t), w_{hj}(t), w_{pk}(t)} J, \text{ with } J = \sum_{t=1}^T \{ \sum_{i=1}^I [C_{ti}(p_{ti}(t)) + S_{ti}(t)] \}. \quad (2.1)$$

In the above, $p_{ti}(t)$ is the generation of thermal unit i at time t , $w_{hj}(t)$ the water released of hydro unit j , $w_{pk}(t)$ the water released of pumped-storage unit k (negative for pumping), and T is the time horizon.

The **system demand constraints** require that the sum of all thermal generation $p_{ti}(t)$, hydro generation $p_{hj}(w_{hj}(t))$, and pumped-storage generation $p_{pk}(w_{pk}(t))$ (negative for pumping) should equal the system demand $P_d(t)$ at each hour, i.e.,

$$\sum_{i=1}^I p_{ti}(t) + \sum_{j=1}^J p_{hj}(w_{hj}(t)) + \sum_{k=1}^K p_{pk}(w_{pk}(t)) = P_d(t), \quad t=1, \dots, T, \quad (2.2)$$

where $p_{hj}(w_{hj}(t))$ and $p_{pk}(w_{pk}(t))$ are water-power conversion functions for hydro unit j and pumped-storage unit k , respectively.

Reserve requirements states that the sum of reserve contributions of thermal units $r_{ti}(p_{ti}(t))$, hydro units $r_{hj}(p_{hj}(t))$, and pumped-storage units $r_{pk}(p_{pk}(t))$ should be greater than or equal to the reserve required $P_r(t)$ at each hour, i.e.,

$$\sum_{i=1}^I r_{ti}(p_{ti}(t)) + \sum_{j=1}^J r_{hj}(p_{hj}(t)) + \sum_{k=1}^K r_{pk}(p_{pk}(t)) \geq P_r(t), \quad t=1, \dots, T. \quad (2.3)$$

Individual unit constraints are as described in [9, 10, 11].

2.2 Solution Methodology

2.2.1 The Relaxed Problem

Relaxing system-wide demand (2.2) and reserve requirements (2.3) by using Lagrange multipliers λ and μ , respectively, the following relaxed problem is formed:

$$\begin{aligned} \min_{p_{ti}(t), w_{hj}(t), w_{pk}(t)} L, \text{ with } L = & \sum_{t=1}^T \{ \sum_{i=1}^I [C_{ti}(p_{ti}(t)) + S_{ti}(t)] \} \\ & + \sum_{t=1}^T \lambda(t) \left[P_d(t) - \sum_{i=1}^I p_{ti}(t) - \sum_{j=1}^J p_{hj}(w_{hj}(t)) - \sum_{k=1}^K p_{pk}(w_{pk}(t)) \right] \\ & + \sum_{t=1}^T \mu(t) \left[P_r(t) - \sum_{i=1}^I r_{ti}(p_{ti}(t)) - \sum_{j=1}^J r_{hj}(p_{hj}(t)) - \sum_{k=1}^K r_{pk}(p_{pk}(t)) \right]. \end{aligned} \quad (2.4)$$

2.2.2 Solving Subproblems

After re-grouping relevant terms, individual thermal, hydro and pumped-storage sub-problems are formed, one for each unit. The resolution of thermal, hydro and pumped storage subproblems can be found in [9], [10] and [11], respectively. In solving the subproblems, the ramp rate constraints for thermal subproblems, the pond level limit constraints for pumped storage subproblems, and the available hydro energy constraints for hydro subproblems are each relaxed by another

set of multipliers. Each set of these multipliers is updated at the middle level. The hierarchical structure of the algorithm is shown in Fig. 1. Updating multipliers, which will be further discussed in Section 3, consists of a major part of the algorithm.

3. Updating Multipliers

3.1 Updating the Multipliers at the High Level

The dual problem at the high level is to update multipliers to maximize the dual cost function $\phi(\lambda, \mu)$, i.e.,

$$\begin{aligned} \text{Max}_{\lambda, \mu \geq 0} \phi(\lambda, \mu), \text{ with} \\ \phi(\lambda, \mu) = \min_{p_{ti}(t), w_{hj}(t), w_{pk}(t)} L, \end{aligned} \quad (3.1)$$

where L is defined in (2.4).

The subgradient g of $\phi(\lambda, \mu)$ is a $2T \times 1$ vector consisting of g_λ and g_μ , where g_λ and g_μ are the subgradient of ϕ with respect to λ and μ , respectively. The t -th element of g_λ is

$$g_\lambda(t) = p_d(t) - \sum_{i=1}^I p_{ti}(t) - \sum_{j=1}^J p_{hj}(t) - \sum_{k=1}^K p_{tk}(t), \quad (3.2)$$

and the t -th element of g_μ is

$$g_\mu(t) = p_r(t) - \sum_{i=1}^I r_{ti}(p_{ti}(t)) - \sum_{j=1}^J r_{hj}(p_{hj}(t)) - \sum_{k=1}^K r_{pk}(p_{pk}(t)). \quad (3.3)$$

The dual problem will be solved by the BTRM to be presented in Section 4. Heuristics are performed at the end of each high level iteration, and the feasible solution with the lowest cost is recorded as the primal scheduling decisions.

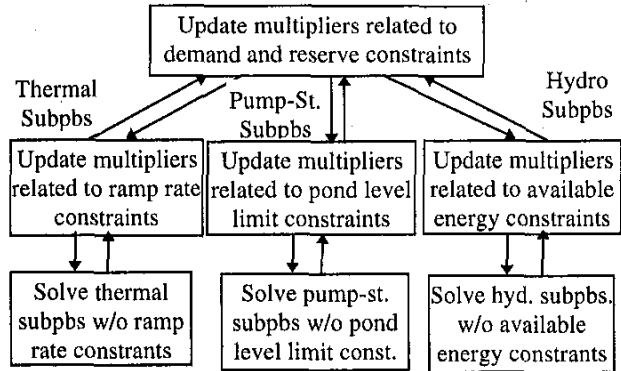


Fig. 1 The Hierarchical Structure of the Algorithm

3.2 Updating the Multipliers at the Middle Level

In the NU system, thermal unit ramp rates can be easily satisfied and the SGM works well at the middle level. The pumped-storage unit in NU has a large pond capacity, and the unit can generate at full capacity continuously for eight hours before the pond level drops from its maximum to minimum. Pond level limits are thus mostly satisfied, and the SGM works well in updating the multipliers for the pumped-storage subproblem. Hydro solutions, however, are very sensitive to the adjustment of middle level multipliers, and are difficult to

converge. This observation supports the discussion in [12]. The convergence of hydro subproblems has direct impact on the convergence of the high level dual problem since the updating of multipliers are based on subproblem solutions.

To improve the convergence of hydro subproblems, the linear power-water relationship is approximated by a quadratic function in [10]. The available energy constraints (total daily generation equals the available daily energy for daily hydro [10]) for hydro unit j are relaxed by a multiplier vector ρ_{hj} . The middle level dual problem for daily hydro unit j is thus written as:

Max $\varphi_{hj}(\rho_{hj})$, with
 ρ_{hj}

$$\varphi_{hj}(\rho_{hj}) \equiv \text{Min}_{w_{hj}(t)} \sum_{t=1}^T [\mu(t) - \lambda(t)] \bar{p}_{hj}(w_{hj}(t)) - \sum_{t=1}^T \mu(t) \bar{p}_{hj}(t) + \sum_{d=1}^D \rho_{hj}(d) \left[\sum_{t=24(d-1)+1}^{24d} w_{hj}(t) - E_{hj}(d) \right], \quad (3.4)$$

where, $\varphi_{hj}(\rho_{hj})$ is the Lagrangian function for hydro unit j and $\bar{p}_{hj}(w_{hj}(t))$ is the quadratic power-water relationship as in [10]; $E_{hj}(d)$ is the available water in day d , and D is the number of days in the scheduling time horizon. The subgradient of $\varphi_{hj}(\rho_{hj})$ with respect to ρ_{hj} is a $D \times 1$ vector with the d -th element being

$$g_{\rho_{hj}}(d) = \sum_{t=24(d-1)+1}^{24d} w_{hj}(t) - E_{hj}(d). \quad (3.5)$$

With given ρ_{hj} , (3.4) can be solved analytically [10].

The solution quality for hydro subproblems is not satisfactory when SGM is used to update the middle level multipliers ρ_{hj} , especially at the first few high level iterations when the high level multipliers are not stable. The BTRM is thus used at the middle level to maximize $\varphi_{hj}(\rho_{hj})$ for hydro units.

4. Implementing BTRM for Hydrothermal Scheduling

4.1 Deriving BTRM from the Cutting Plane Idea

The BTRM is derived from the cutting plane idea [6]. Given a non-differentiable concave function $f(x)$, $x \in \mathbb{R}^n$, the BTRM accumulates subgradients and function values obtained so far in a "bundle." At iteration k , the available information in the bundle includes $f(x_1), f(x_2), \dots, f(x_k)$ and $g(x_1), g(x_2), \dots, g(x_k)$ (will be simplified as g_1, g_2, \dots, g_k). With this bundle, $f(x)$ is approximated with the following cutting plane model:

$$f_{cp}(x) = \min_{1 \leq i \leq k} \{f(x_i) + g_i^T(x - x_i)\}. \quad (4.1)$$

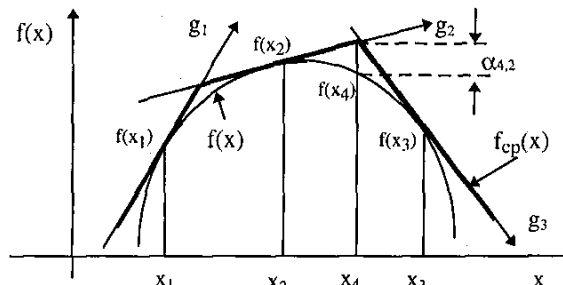


Fig. 2 Cutting Plane Approximation of $f(x)$

Suppose that $f(x)$ is as shown in Fig. 2, and $f(x_1), f(x_2), f(x_3)$, g_1, g_2 and g_3 are in the bundle. The cutting plane method maximizes the piece-wise linear $f_{cp}(x)$ (thick line) to yield the next iterate x_4 . A drawback is that the approximation will be poor if x_{k+1} is far away from x_k . A negative quadratic term is thus added to the cutting plane model, and x_{k+1} is solved by

$$x_{k+1} = \arg \max_{x \in \mathbb{R}^n} \left\{ f_{cp}(x) - \frac{1}{2t_k} (x - x_k)^2 \right\}, \quad (4.2)$$

where t_k is a positive parameter. The magnitude of t_k decides how well the cutting plane model should be "trusted." A larger t_k will decrease the effect of the quadratic term, meaning that the cutting plane model can be trusted in a larger region, and vice versa. A suitable t_k thus has major impact on the convergence of the method.

To simplify (4.2), let

$$\alpha_{k,j} = (f(x_j) + g_j^T(x_k - x_j)) - f(x_k), \text{ and} \quad (4.3)$$

$$d = x - x_k. \quad (4.4)$$

In (4.3), $\alpha_{k,j}$ is the "linearization error" of $f(x)$ at x_k when the linearization is performed at x_j with subgradient g_j , and d is the direction. As an example, Fig. 2 shows how $\alpha_{4,2}$ is calculated. Problem (4.2) can then be rewritten as

$$d = \arg \max_{d \in \mathbb{R}^n} \left\{ v - \frac{1}{2t_k} d^T d \right\}, \quad (4.5)$$

subject to,

$$v \leq g_i^T d + \alpha_k^i, i = 1, 2, \dots, k. \quad (4.6)$$

Solving problem (4.5) and (4.6) for d involves maximization over the $n+1$ dimensional space. For practical problems, n can be quite large, and k (the number of elements in the bundle) is generally much smaller than n . To reduce the computation efforts, the dual form will be solved. By relaxing (4.6) with a k -dimensional multiplier vector γ , the Lagrangian function can be written as

$$L(d, v, \gamma) = v - \frac{1}{2t_k} d^T d + \sum_{i=1}^k \gamma_i (g_i^T d + \alpha_k^i - v). \quad (4.7)$$

From the first order optimality conditions, one gets:

$$\sum_{i=1}^k \gamma_i = 1, \quad (4.8)$$

$$d = t_k \sum_{i=1}^k \gamma_i g_i, \quad (4.9)$$

$$\gamma_i (g_i^T d + \alpha_k^i - v) = 0, \gamma_i \geq 0, i = 1, 2, \dots, k. \quad (4.10)$$

Summing up k equations of (4.10), and making use of (4.8) and (4.9), one gets

$$v = \frac{1}{t_k} \|d\|^2 + \sum_{i=1}^k \gamma_i \alpha_k^i. \quad (4.11)$$

Using (4.9) and (4.11), one gets the dual problem:

$$\gamma = \arg \min_{\gamma \in \mathbb{R}^k} \left\{ \frac{1}{2} \left\| \sum_{j=1}^k \gamma_j g_j \right\|^2 + \frac{1}{t_k} \sum_{j=1}^k \gamma_j \alpha_{k,j} \right\}, \quad (4.12)$$

$$\text{subject to: } \gamma_j \geq 0, j = 1, 2, \dots, k \text{ and } \sum_{j=1}^k \gamma_j = 1. \quad (4.13)$$

Solving for γ is a quadratic programming problem, and the recursive active set strategy in [7] is used. This strategy uses the fact that only one more subgradient is added to the bundle each time, so the solution for the enlarged problem can be efficiently obtained from its previous solution. It can be seen from Fig. 2 that the cutting plane approximation gets better with the enlargement of the bundle. In practical implementations, however, a maximum bundle size is usually imposed to limit memory requirements.

With the vector γ obtained from solving (4.12) and the search direction d calculated from (4.9), a trial iterate x_{k+1}^i can be obtained as

$$x_{k+1}^i = x_k + d. \quad (4.14)$$

The function value $f(x_{k+1}^i)$, a subgradient $g(x_{k+1}^i)$, and the linearization error $\alpha_{k,k+1}^i$ can be readily calculated. The following systematic rule will then be used to adaptively adjust t_k :

- If $f(x_{k+1}^i) \geq f(x_k) + \epsilon$, set $x_{k+1} = x_{k+1}^i$ and increase t_k (trust region enlarged).
- Otherwise, check whether $\alpha_{k,k+1}^i \leq \sigma$. If it is, then $g(x_{k+1}^i)$ can be added to the bundle.
- Otherwise, t_k is too large and the quadratic programming problem is solved again with a reduced t_k .

In the above, ϵ and σ are two threshold values. This process is analytically proved to converge in [6].

The distinct feature of BTRM is that the direction is solved with respect to the trial step size t_k (see (4.9), (4.12) and

(4.14)). Most other bundle-type methods solve for a direction independent of the step size, and a time-consuming line search is necessary to find a suitable step size. In the BTRM, however, the direction is optimal with respect to the trial step size by solving the quadratic programming with t_k considered.

4.2 Implementing BTRM for Hydrothermal Scheduling

The implementation of BTRM for solving the hydro subproblem j with given high level multipliers λ and μ is summarized below.

Step 1: [Initialize Multipliers.] Initialize ρ_{hj} to zero if the subproblem is solved for the first time; otherwise initialize ρ_{hj} to the previous value. Set $x_0 = \rho_{hj}$. Set the initial step size t_0 , the max t_{\max} , and the min t_{\min} . Set the iteration index $k = 0$;

Step 2: [Solve Hydro Subproblem j .] Calculate the hydro Lagrangian function and subgradient using (3.4) and (3.5).

Step 3: [Solve QP.] Solve the quadratic programming problem (4.12) and (4.13) to get γ , and calculate d and x_{k+1}^i using (4.9) and (4.14), respectively.

Step 4: [Test Convergence.] If $\|d\| \leq \delta$ (δ is a pre-specified stopping criterion), stop; else, continue.

Step 5: [Solve Hydro Subproblem j .] Set $\rho_{hj} = x_{k+1}^i$, and calculate the hydro Lagrangian function and the subgradient using (3.4) and (3.5), respectively, and calculate the linearization error $\alpha_{k,k+1}^i$ using (4.3).

Step 6: [Check Function Value.] If $f(x_{k+1}^i) \geq f(x_k) + \epsilon$, go to Step 8; else, continue.

Step 7: [Check Linearization Error.] If $\alpha_{k,k+1}^i \leq \sigma$, go to Step 9; else set $t = (t_{\min} + t)/2$, and go to Step 3.

Step 8: [Update Iterate and Enlarge Bundle.] Add $g(x_{k+1}^i)$ to the bundle, set the current iterate $x_{k+1} = x_{k+1}^i$, update t to $t = \min(t_{\max}, 1.1*t)$, k to $k+1$, and go to Step 3.

Step 9: [Enlarge Bundle.] Add $g(x_{k+1}^i)$ to the bundle, set k to $k+1$, and go to Step 3.

The steps for solving the high level dual problem are essentially the same, and are thus omitted here.

5. Numerical Results

The BTRM and quadratic programming were implemented in C++, and the resolution of individual subproblems and heuristics were implemented in FORTRAN on a SUN Ultra Station. In the following, BTRM is tested for a small problem with identical units in Example 1, and for a hydro subproblem in Example 2. In Example 3, BTRM is tested using data sets of Northeast Utilities system.

5.1 Example 1

There are three units in the system a time horizon of two hours and load $p_d(1) = p_d(2) = 500$ MW. For each unit, the thermal cost function is linear with $C_i(p_i(t)) = q_i p_i(t)$, $i = 1, 2, 3$, and q_i is a constant with respect to unit i . The minimum and maximum generation levels are $\underline{p}_i = 40$ MW, $\bar{p}_i = 200$ MW, $i = 1, 2, 3$, respectively. First, assuming that the three units are different with $q_1 = \$22/\text{MW}$, $q_2 = \$28/\text{MW}$, $q_3 = \$33/\text{MW}$, the dual function is plotted in Fig. 3 where there are 16 facets. The “ridges” are not very sharp, because only one of three units has different decisions between adjacent facets.

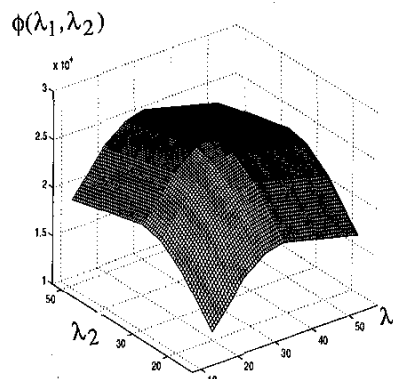


Fig. 3 Dual function for a system with three different units

The three units are then assumed to be identical with $q_i = \$33/\text{MW}$, $i = 1, 2, 3$, and the dual function is plotted in Fig. 4 where there are four facets. Each facet corresponds to a primal decision, e.g., facet 1 corresponds to “all units off” in both hours, and facet 2 corresponds to “all units on at maximum level” at hour 1 and “all units off” at hour 2. It can be seen that identical units are “all-on at maximum level” on one side of a ridge, and “all-off” on the other side. This leads to “sharp ridges” of the dual function.

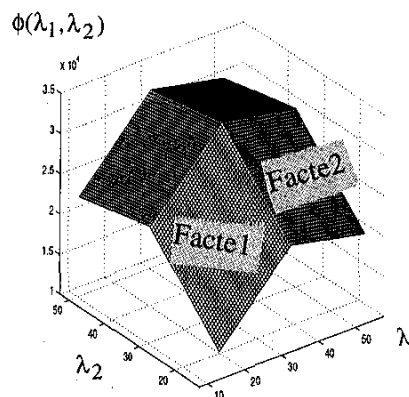


Fig. 4 Dual function for a system with three identical units

The BTRM and SGM are used to maximize the dual function with results summarized in Table 1. It can be seen that when units are different, SGM can converge to the optimal. BTRM, however, converges with less number of function evaluations. It can also be seen that, when units are identical, SGM has difficulty converging to the optimal, while BTRM converges to the optimal without difficulty.

Table 1. Performance Comparison of BTRM and SGM

	D.F.V. BTRM	N.F.E BTRM	D.F.V. SGM	N.F.E SGM	Optimal D.F.V.
Diff. U.	26600	6	26600	10	26600
Iden. U.	33000	5	32990	16	33000

Diff. U.: – Different Units; Iden. U. – Identical Units; D.F.V.– Dual Function Value; N.F.E.– Number of Function Evaluations

5.2 Example 2

The hydro subproblem is one in the scheduling of NU system. The high level multipliers λ and μ are kept the same at their initial values and ρ_{hj} is initialized to zero. The trajectories of ϕ_{hj} with respect to the number of function evaluations are shown in Fig. 5. Although the initial step size for the SGM has been chosen carefully and adjusted adaptively, it still converges slower than BTRM does. To understand the performance difference of BTRM and SGM, the function $\phi_{hj}(\rho_{hj})$ with respect to $\rho_{hj}(1)$ and $\rho_{hj}(2)$ is plotted in Fig. 6. The sharp ridges makes SGM converges slowly.

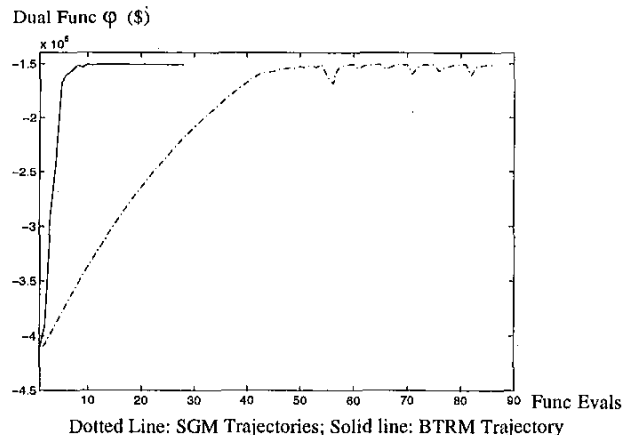


Fig.5 Dual function values vs. number of function evaluations

The number of function evaluations and number of iterations for BTRM to converge are 28 and 26, respectively. They are almost equal, implying that in most of the iterations, only one function evaluation is needed. This shows that most trial step sizes are appropriate and the BTRM adaptive step sizing rule is efficient.

5.3 Example 3

There are 65 thermal units, 7 hydro units, 1 large pumped-storage unit, and 2-10 schedulable contracts. The scheduling horizon is either 7 days (168) or 8 days (192 hours). Results for six NU 1996 data sets are summarized in Tables 3 and 4. It

can be seen that BTRM can obtain better dual solutions and feasible solutions with less CPU time for most cases.

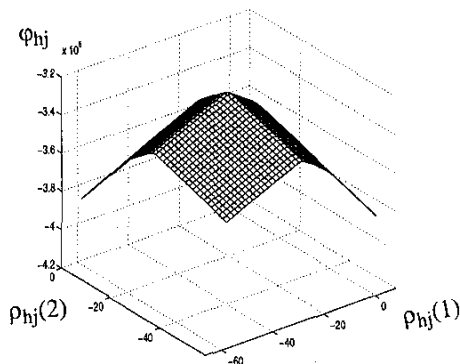


Fig. 6 Dual Function of a Hydro Subproblem

Table 3: Comparison of Dual and Feasible Costs (Dollars)

Cases	D-BTRM	F-BTRM	D-RCBM	F-RCBM
NU-Case1	6650906	6821767	6625105	6835650
NU-Case2	4634620	4723137	4632233	4725172
NU-Case3	7406271	7544156	7360316	7546210
NU-Case4	6801713	6949228	6793646	6967792
NU-Case5	3152792	3255103	3105187	3315255
NU-Case6	5272420	5451045	5271662	5484675

D-BTRM: dual cost of BTRM; D-RCBM: dual cost of RCBM;
F-BTRM: feasible cost of BTRM; F-RCBM: feasible cost of RCBM.

Table 4: Comparison of CPU Time (Seconds)

	CPU-BTRM	CPU-RCBM
NU-Case1	122	120
NU-Case2	108	123
NU-Case3	111	119
NU-Case4	110	125
NU-Case5	93	95
NU-Case6	111	125

5. Conclusion

The BTRM is used to update the multipliers within the Lagrangian relaxation framework to overcome difficulties subgradient methods have. The fast convergence of BTRM is analyzed and supported by extensive testing results.

Acknowledgment

This work is supported in part by the National Science Foundation under Grants ECS-9420972 and DMI-9500037, and by a grant from Northeast Utilities Service Company. Valuable suggestions from Mr. Yajun Wang, Mr. K. Kasiviswanathan of UConn, Mr. Martin B. Gabris, Mr. Jim A. Palmberg and Mr. Carl Larson of Northeast Utilities are greatly appreciated.

References

- [1] Shaw, J., "A Direct Method for Security-Constrained Unit Commitment," *IEEE Transactions on Power Systems*, Vol. 10, No. 3, August 1995, pp. 1329-1339.
- [2] Wang, S. J., S. M. Shahidehpour, D. S. Kirschen, S. Mokhtari, G. D. Irisari, "Short-Term Generation Scheduling with Transmission and

- Environmental Constraints Using an Augmented Lagrangian Relaxation," *IEEE Transactions on Power Systems*, Vol. 10, No. 3, Aug. 1995, pp. 1294-1301.
- [3] Luh, P. B., D. Zhang, R. Tomastic, "An Algorithm for Solving the Dual Problem of Hydrothermal Scheduling," To appear in *IEEE Transactions on Power Systems*. Presented in 1997 IEEE PES Winter Meeting, NY. Paper number: PE-601-PWRS-0-01-1997.
- [4] Pellegrino, F., A. Renaud, T. Socroun, "Bundle and Augmented Lagrangian Method for Short Term Unit Commitment," 12th Power System Computation Conference, 1996, Dresden.
- [5] Mendes, V., L. Ferreira, P. Roldao, R. Pestana, "Optimal Short Term Resource Scheduling By Lagrangian Relaxation: Bundle Type Versus Subgradient Algorithms," 12th Power System Computation Conference, 1996, Dresden.
- [6] Schramm, H., J. Zowe, "A Version of the Bundle Method for Minimizing a nonsmooth Function: Conceptual Idea, Convergence Analysis, Numerical Results", *SIAM J. Optimization*, Vol. 2, No. 1, pp. 121-151, Feb. 1992.
- [7] Kiwiel, K. C., "A Method for Solving Certain Quadratic Programming Problems Arising in Nonsmooth Optimization", *IMA J. Numer. Anal.*, 6(1986), pp. 737-152.
- [8] Hao, S., G. Angelidis, H. Singh and A. Papalexopoulos, "Consumer Payment Minimization in Power Pool Auctions", *Proceedings of the 20th International Conference on Power Industry Computer Applications*, May 1997, Columbus, Ohio, pp. 368-373.
- [9] Guan, X., P. B. Luh, H. Yan, and J. A. Amalfi, "An Optimization-Based Method for Unit Commitment," *Electric Power & Energy Syst.* Vol. 14, 1992, pp. 9-17.
- [10] Guan, X., P. B. Luh and L. Zhang, "Nonlinear Approximation Method in Lagrangian Relaxation-Based Algorithms for Hydrothermal Scheduling," *IEEE Trans. on Power Syst.*, Vol. 10, No. 2, May 1995, pp. 772-778.
- [11] Guan X., P. B. Luh, H. Yan, and P. M. Rogan, "Optimization-based Scheduling of Hydrothermal Power Systems with Pumped-Storage Units," *IEEE Trans. on Power Syst.*, Vol. 9, No. 2, May 1994, 1023-1029.
- [12] Li, C., E. Hsu, A. Svoboda, C. Tseng, R. Johnson, "Hydro Unit Commitment in Hydro-Thermal Optimization", *IEEE Trans. On Power Systems*, Vol. 12, No. 2, May 1997, pp. 764-769.

Peter B. Luh (S'77-M'80-SM'91-F'95) received his B.S. degree in Electrical Engineering from National Taiwan University, Taipei, Taiwan, Republic of China, in 1973, the M.S. degree in Aeronautics and Astronautics Engineering from MIT, Cambridge, Massachusetts, in 1977, and the Ph.D. degree in Applied Mathematics from Harvard University, Cambridge, Massachusetts, in 1980. Since 1980, he has been with the University of Connecticut, and currently is a Professor in the Department of Electrical and Systems Engineering, and Director of Production Systems and Information Technology Program within the Advanced Technology Center for Precision Manufacturing. His major research interests include schedule generation and reconfiguration for manufacturing systems and power systems. Dr. Luh is an Editor of the *IEEE Transactions on Robotics and Automation*, was an Associate Editor of the *IEEE Transactions on Automatic Control*, and has served on program committees and operating committees of many major national, international, and inter-society conferences.

Daoyuan Zhang received his BE degree in Control Engineering from Tsinghua University, P. R. China in 1989. He is a Ph.D. candidate at the Department of Electrical Engineering, University of Connecticut now.

Yuanhui Zhang received his BE and MS in Control Engineering from Tsinghua University, P. R. China in 1993 and 1995, respectively. He is a Ph.D. candidate at the Department of Electrical Engineering, University of Connecticut now.

Discussion

N. Jiménez-Redondo (Universidad de Málaga, Málaga, Spain) **A. J. Conejo** (Universidad de Castilla - La Mancha, Ciudad Real, Spain): We would like to commend the authors for their contribution in using Bundle Methods to solve the dual problem of the short-term hydrothermal coordination problem, and for their clear and well-written paper.

We would like the authors to comment on the following issues:

1. The primal form of the bundle method presented, equation (4.2), is derived from the cutting plane method. In this primal form t_k is the inverse of a penalty parameter. Instead of solving the primal form of the bundle method, the authors solve its dual form [D1]. In this dual form t_k is a stepsize in the direction of a lineal combination of subgradients. According to the results of optimization problem (4.12)-(4.13), the vector of multipliers, t_k , and the *bundle* are updated at each iteration. The method could be categorized as a dual penalty cutting plane method. On the other hand, an optimization procedure is defined as a trust region method when the feasibility region is dynamically modified. This is the case for the method presented in [D2] but not for the method proposed in this paper. Could the authors explain why they categorize their method as a trust region method?
2. In each iteration the ascent direction is computed as a lineal combination of subgradients (an ε -subgradient [D1]). Instead of performing a time consuming line search to compute t_k in each iteration, the authors apply rules to increase or decrease it (section 4.2). Have the authors tried other rules? Could they comment on the behavior of the proposed rules?
3. Could the authors provide further details on the selection of parameters ε and σ , and could they comment on the influence of these parameters on the convergence of their algorithm?
4. The authors use a procedure to keep limited the size of the bundle. Could they provide some details on this procedure? Is it similar to the one presented in [D2]?

Finally, we would like to congratulate again the authors for their relevant paper.

[D1] J. B. Hiriart-Urruty, C. Lemarechal *Convex analysis and minimization algorithms*. Vol II. Springer-Verlag, Berlin 1996.

[D2] N. Jiménez-Redondo, A. J. Conejo. *Short-term hydrothermal coordination by lagrangian relaxation: solution of the dual problem*. Paper PE-333-PWRS-0-12-1997. Approved for publication in the IEEE Transaction on Power Systems.

Manuscript received October 19, 1998.

Daoyuan Zhang –

I would like to thank the discussers for their interests in the paper and their intriguing questions. In the following, the questions will be addressed in the order that appeared.

1. In the Bundle Method presented in the paper, each trial ascent direction is obtained with a fixed t_k that serves as the criteria of “how much” the cutting plane approximation, and thus the bundle, can be “trusted”. This can be seen from equation (4.12). When a large t_k is chosen, the bundle is given more “trust”, and the quadratic penalty plays less roles, and vice verse. In [6], the method is also categorized as a Trust Region Method.
2. In a bundle method, the line search has two purposes: either increasing the dual value by an expected amount (ε), or finding another subgradient that could enhance the bundle. The line search procedures described in the paper serves these purposes well, and we have not tried other methods so far, although we did try quadratic approximation in the context of a subgradient method.
3. The two threshold ε and σ are important to the convergence of the method and are application specific. As can be seen from the line search procedures, a large ε will make it difficult for the algorithm to move to another iterate, and a large σ will make it difficult to find a candidate subgradient that can be added to the bundle. In our implementation for Northeast Utilities, the two thresholds are selected by trial-and-error.
4. The maximum bundle size is also an application specific parameter. The procedure used to control the bundle size is as follows. When the bundle size reaches the maximum, the subgradient with the largest linearization error (α) is removed from the bundle. This is based on the observation that a subgradient with larger linearization error is usually obtained from a point farther away from the current iterate (see Figure 2) and thus carries less important information for searching an ascent direction.

Manuscript received February 3, 1999.