Payment Cost Minimization Auction for Deregulated Electricity Markets Using Surrogate Optimization

Peter B. Luh, *Fellow, IEEE*, William E. Blankson, *Student Member, IEEE*, Ying Chen, *Student Member, IEEE*, Joseph H. Yan, Gary A. Stern, Shi-Chung Chang, *Member, IEEE*, and Feng Zhao, *Student Member, IEEE*

***Abstract*—Deregulated electricity markets use an auction mechanism to select offers and their power levels for energy and ancillary services. A settlement mechanism is then used to determine the payments resulting from the selected offers. Currently, most independent system operators (ISOs) in the United States use an auction mechanism that minimizes the total offer costs but determine payment costs using a settlement mechanism that pays uniform market clearing prices (MCPs) to all selected offers. Under this setup, the auction and settlement mechanisms are inconsistent since minimized costs are different from payment costs. Illustrative examples in the literature have shown that for a given set of offers, if an auction mechanism that directly minimizes the payment costs is used, then payment costs can be significantly reduced as compared to minimizing offer costs. This observation has led to discussions among stakeholders and policymakers in the electricity markets as to which of the two auction mechanisms is more appropriate for ISOs to use. While methods for minimizing offer costs abound, limited approaches for minimization of payment costs have been reported. This paper presents an effective method for directly minimizing payment costs. In view of the specific features of the problem including the nonseparability of its objective function, the discontinuity of offer curves, and the maximum term in defining MCPs, our key idea is to use augmented Lagrangian relaxation and to form and solve offer and MCP subproblems by using the *surrogate optimization framework*. Numerical testing results demonstrate that the method is effective, and the resulting payment costs are significantly lower than what are obtained by minimizing the offer costs for a given set of offers.**

***Index Terms*—Augmented Lagrangian relaxation, deregulated electricity markets, market clearing price (MCP), offer cost minimization, payment cost minimization, surrogate optimization.**

I. Introduction

DEREGULATED wholesale electricity markets (e.g., the day-ahead, hour-ahead, and real-time markets) operated by independent system operators (ISOs) in the United States generally adopt an auction mechanism to select generation offers and demand bids and their levels for energy and ancillary services. A settlement mechanism is then used to determine the corresponding payments for the selected generation offers and demand bids. There are two main auction mechanisms, namely, offer cost minimization where offers are selected to minimize total supply offer costs, and payment cost minimization, where offers are selected to minimize the actual payment costs. The markets are settled by two main settlement mechanisms, namely, the pay-as-offer (also referred to as pay-as-bid) mechanism, where each selected supplier is paid at its offer price, and the pay-at-MCP mechanism, where all selected suppliers are paid at a uniform market clearing price (MCP), usually the price of the most expensive selected offer.[1] In practice, the pay-at-MCP settlement mechanism is widely accepted and used for payments, and so for the rest of this paper, the pay-at-MCP settlement mechanism is assumed. It is also assumed for simplicity that system demand is given, and therefore, demand bids are not considered. The key question is what the objective function should be in the auction to select generation offers and to set MCPs. Note that MCPs are most important to determine settlement costs and have significant impacts on forward transactions outside of the ISO markets.

Currently, most ISOs in the United States adopt the offer cost minimization auction by using the traditional unit commitment approach. The problem is NP hard, but in view of its separability,[2] it can be effectively solved by using the Lagrangian relaxation technique for near-optimal solutions [1]–[4]. Furthermore, existing unit commitment and economic dispatch packages can be readily adapted to solve the problem by simply replacing generator cost functions by offer curves. This approach is good or at least consistent for a market that uses the pay-as-offer mechanism for market settlements. However, since most ISOs in the United States use the pay-at-MCP mechanism for settlement, the auction and settlement mechanisms are inconsistent [5] since the payment costs are usually significantly higher than the minimized auction costs for a given set of offers. Proponents of offer cost minimization indicate that if offer prices represent true production costs, then this mechanism maximizes social welfare [8], [12], [13]. Whether offer prices represent true production costs, however, is subject to much debate.

An alternate mechanism is the payment cost minimization auction [5]–[11], which directly minimizes the payment costs of consumers. Examples in the literature have shown that for a given set of offers, if this auction mechanism is used, then the payment costs can be significantly reduced as compared to what

Manuscript received May 16, 2005; revised November 17, 2005. This work was supported in part by the National Science Foundation under Grant ECS-0323685 and in part by a grant from Southern California Edison. Paper no. TPWRS-00304-2005.

P. B. Luh, W. E. Blankson, Y. Chen, S.-C. Chang, and F. Zhao are with the Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT 06269-2157 USA.

J. H. Yan and G. A. Stern are with the Market Monitoring & Analysis, Southern California Edison, Rosemead, CA 91770 USA.

Digital Object Identifier 10.1109/TPWRS.2006.873104

[1]While markets are moving toward the locational marginal pricing (LMP) of energy where the transmission network is considered, we shall for simplicity consider uniform MCP of the day-ahead energy market.

[2]A problem is separable and can be decomposed into individual subproblems if both the objective function and the constraints that couple the subproblems are additive in terms of subproblem decision variables.

is obtained by minimizing offer costs. This approach is also consistent with the pay-at-MCP mechanism [5] because minimized costs will be the same as settlement costs. However, if this auction mechanism were adopted, suppliers may bid differently, and there is no guarantee that the expected savings will be realized.

The disparate views held by proponents of the two auction mechanisms have led to a discussion among stakeholders of electricity markets as to which of the two auction mechanisms is more appropriate for ISOs to use. While methods for minimizing offer costs abound, limited approaches have been reported for payment cost minimization. Most of the papers on payment cost minimization mathematically formulate the problem and use illustrative examples to demonstrate that it can reduce procurement costs but present no solution methodology [5], [7], [8], [11]. The only paper that presented a solution methodology was based on forward dynamic programming [10], but the author acknowledges that the method is not suited for large-scale problems because of the curse of dimensionality. The objective of this paper is not to discuss the pros and cons of the two auction mechanisms but rather to present a novel and effective solution methodology for solving the payment cost minimization problem for a day-ahead energy market so that the two mechanisms can be effectively compared.

For the rest of this paper, the problem is mathematically formulated in Section II. The objective function to be optimized is the total payment cost that contains cross product terms of MCPs and offer levels as well as startup costs, and the optimization is subject to system demand and individual unit constraints. MCPs are defined as the maximum prices of selected offers and are not known prior to the selection but are the results of optimization. Additional difficulties would arise if the standard Lagrangian relaxation method were used to solve the problem, since the Lagrangian formed would be linear in terms of offer levels.

Our solution methodology presented in Section III consists of using the augmented Lagrangian, which is formed by adding quadratic penalty terms of coupling constraints to the Lagrangian to avoid the difficulties of solution oscillation that would otherwise arise. The surrogate optimization framework is then used to overcome the difficulties caused by inseparability due to the product of MCPs and offer levels and the quadratic penalty terms. The key idea of surrogate optimization is that the relaxed problem does not have to be solved exactly. Rather, approximate optimization is sufficient if the "surrogate optimization conditions" can be satisfied. Once the conditions are satisfied, the "surrogate subgradient" is "good" since it forms an acute angle with the direction toward the optimal multipliers and is used to update multipliers. In view of inseparability, the relaxed problem as a whole is taken as a subproblem and optimized with respect to the decision variables of a particular offer one at a time. To overcome the difficulties caused by the structural change of the augmented Lagrangian as MCP varies when the selection status of the offer under consideration is changed, decision variables of other offers may have to be adjusted to satisfy the surrogate optimization condition.

Numerical testing results presented in Section IV show that the auction and settlement mechanisms are consistent, and the method is effective and yields significantly reduced payment costs as compared to what is obtained by offer cost minimization for a given set of offers.

II. Problem Formulation

In this section, the offer cost minimization and the payment cost minimization auction mechanisms are formulated to highlight their differences. To present the key ideas of the new solution methodology, the following simplifying assumptions are made: system demand is given, system reserve constraints and transmission congestion are not considered, startup costs are fully compensated,[3] and participants submit single block constant price offers with maximum/minimum power levels.

Consider an energy market with I supply offers indexed by $i = 1, 2, \ldots, I$. For offer i, the offer curve (or price) for supplying power $p_i(t)$ at time $t(1 \le t \le T)$ is denoted by $o_i(p_i(t), t)$ (\$/MW), and the cost curve, which is the integral of the offer curve, is denoted by $C_i(p_i(t), t)$ (\$). The startup cost is denoted by $S_i(t)$ (\$/Start) and is incurred if and only if offer i is turned "ON" from an "OFF" state. The objective of the auction is to select offers and their associated power levels to minimize an appropriate objective function while satisfying relevant constraints. The pay-at-MCP mechanism is then used for settlement.

A. Offer Cost Minimization Problem

Currently most ISOs run an auction that minimizes the total offer costs

$$\min_{\{p_i(t)\}} J, \text{ with } J \equiv \sum_{t=1}^{T} \sum_{i=1}^{I} \left\{ \mathrm{C}_i\left(p_i(t), t\right) + S_i(t) \right\}. \tag{1}$$

System demand constraints require that the total power from all selected offers should equal the system demand $P_d(t)$ at each time, i.e.,

$$h(t) = P_d(t) - \sum_{i=1}^{I} p_i(t) = 0, \quad t = 1, 2, \ldots, T. \tag{2}$$

These constraints couple individual offers. Similar to the unit commitment problem where a unit can be on or off, a binary variable $u_i(t)$ is defined to represent the status of an offer: the offer is ON (1) if it is selected and OFF (0) otherwise. Minimum/maximum power levels are represented as

$$p_i(t) = 0, \quad \text{if } u_i(t) = 0 \tag{3}$$

$$p_{i\,min}(t) \le p_i(t) \le p_{i\,max}(t) \quad if\ u_i(t) = 1,\ t = 1, 2, \ldots, T \tag{4}$$

where $p_{imin}(t)$ and $p_{imax}(t)$ are, respectively, the minimum and maximum power levels of offer i at time t. From (3) and (4), the feasible region of $p_i(t)$ is discontiguous. Since both

[3]Full compensation of startup costs is a simplifying assumption. The choice of a startup compensation scheme is a regulator's or ISO's decision. More realistic conditions such as demand bids and partial compensation of startup costs have been considered in our working paper [21].

the objective function and the coupling system demand constraints (2) are additive, the above formulation is separable. By using Lagrange multipliers to relax (2), the problem can be converted to a two-level optimization problem. At the low level, individual offer subproblems are minimized for a given set of multipliers, while multipliers are updated at the high level based on subproblem solutions. This iterative process continues until the system demand constraints are satisfied or some stopping criteria are met.

The MCP for time t is defined as the price for which demand equals supply. Currently under offer cost minimization, the Lagrange multiplier for system demand, which is equivalent to the marginal cost of the last selected offer, is typically used as the MCP. In view that the multipliers may include a price component for the startup costs and *all* suppliers are paid at MCP, this choice of MCP may significantly increase the total payment costs. In general, $MCP(t)$ for a uniform price settlement mechanism such as pay-at-MCP can be defined as the maximum offer price of all selected offers for time t, i.e.,

$$MCP(t) = max\left\{o_i\left(p_i(t), t\right)\right), \forall\, i \text{ such that } p_i(t) > 0\right\}. \tag{5}$$

The MCP thus defined is purely an energy price and is devoid of any component of startup costs. Startup costs can be paid for separately.

B. Payment Cost Minimization Problem

Under the payment cost minimization auction mechanism, procurement costs are directly minimized. The problem as presented in [5] is

$$\min_{\{MCP(t)\},\{p_i(t)\}} J, \text{ with } J \equiv \sum_{t=1}^{T}\sum_{i=1}^{I}\left\{MCP(t)p_i(t) + S_i(t)\right\}. \tag{6}$$

This minimization is subject to (2)–(4). Here $MCP(t)$ is not equivalent to the marginal cost but is defined per (5). Compared to (1), the objective function is now complicated because it is a function of both MCPs and power levels of selected offers, while MCPs themselves are yet to be determined based on selected offer curves per (5). Additionally MCPs as defined in (5) may vary depending on the selected offers at a particular time. Thus, during the iterative process, this can cause "jumps" in the values of MCPs. Furthermore, the existence of cross product terms of MCPs and $\{p_i(t)\}$ in (6) makes the problem inseparable.[4] Consequently, the standard Lagrangian relaxation approach that requires separability cannot be directly applied.

C. MCP-Offer Inequality Constraints

From (5), if offer i is selected with power level $p_i(t) > 0$, then $MCP(t)$ should be greater than or equal to the offer price $o_i(p_i(t), t)$. If not, i.e., $p_i(t) = 0$, then $o_i(p_i(t), t)$ has no role in determining $MCP(t)$. $MCP(t)$ as defined in (5) is thus related to selected offers only. For convenience of derivation, the offer curve is redefined to be zero if no power is awarded, i.e.,

$$o_i^r\left(p_i(t), t\right) \equiv \begin{cases} o_i\left(p_i(t), t\right), & p_i(t) > 0 \\ 0, & p_i(t) = 0. \end{cases} \tag{7}$$

With (7), MCPs and all offers are now related through the following "MCP-offer constraints" in linear inequality form:

$$\begin{aligned} &MCP(t) \geq o_i^r\left(p_i(t), t\right), \forall\, i \text{ and } t, \text{ or equivalently} \\ &o_i^r\left(p_i(t), t\right) - MCP(t) \leq 0,\ \forall\, i \text{ and } t. \end{aligned} \tag{8}$$

The above inequality constraints couple MCPs with all offers[5] and will be relaxed by using a set of Lagrange multipliers. The redefined offer curve (7), however, is now discontinuous at $p_i(t) = 0$.

III. Solution Methodology

When the standard Lagrangian of the problem is formed by relaxing (2) and (8) with multipliers and adding them to (6), the relaxed problem is inseparable because of cross product terms of offers levels and MCPs. Therefore, it cannot be decomposed into individual subproblems. Furthermore, if the relaxed problem is optimized with respect to a particular offer, the linearity of the Lagrangian in terms of offer levels will cause solutions to oscillate. Consequently, direct application of the standard Lagrangian relaxation technique will not be effective. The augmented Lagrangian [15]–[18] as opposed to the standard Lagrangian is thus used to overcome the linearity difficulty [19] and [20], while the surrogate optimization framework of [14] will be used to handle the inseparability difficulty.

The augmented Lagrangian is formed by adding quadratic penalty terms of coupling equality and inequality constraints to the standard Lagrangian, leading to a quadratic relaxed problem with improved convergence [17]. Although this leads to additional inseparability, the surrogate optimization framework is able to efficiently handle the inseparability issues. The key idea of surrogate optimization is that it is not necessary to accurately minimize the relaxed problem. Rather, approximate optimization is sufficient if the surrogate optimization conditions are satisfied. Since the relaxed problem cannot be decomposed, it is optimized as a whole with respect to the decision variables of a particular offer or a particular MCP (an offer subproblem or an MCP subproblem). In solving an offer subproblem, the decision variables of other offers may have to be adjusted to overcome the difficulties caused by the structural change of the augmented Lagrangian as MCP varies when the selection status of the offer under consideration is changed. Once the surrogate optimization conditions are satisfied, multipliers are updated by using the surrogate subgradient obtained.

[4]In this paper, $\{p_i\}$ and $\{MCP(t)\}$ are treated as decision variables. There are other ways to formulate the problem by exploiting special problem features. However, the formulation presented here is generic.

[5]Some ISOs use piecewise linear offer curves. The formulation presented here can be extended to handle such cases by replacing the constant offer curves in (8) with these linear offer curves, which are differentiable except for a finite number of discrete points.

A. Augmented Lagrangian

Let multipliers $\{\lambda(t)\}$ and $\{\eta_i(t)\}$ relax coupling system demand constraints (2) and MCP-offer inequality constraints (8),[6] respectively, then the augmented Lagrangian is formed as

$$\begin{aligned}
&L_c(\lambda,\eta,p,MCP)\\
&\equiv \sum_{i=1}^{I}\sum_{t=1}^{T}\{MCP(t)p_i(t)+S_i(t)\}\\
&+\sum_{t=1}^{T}\left\{\lambda(t)\left(P_d(t)-\sum_{i=1}^{I}p_i(t)\right)+\frac{c}{2}\left(P_d(t)-\sum_{i=1}^{I}p_i(t)\right)^2\right\}\\
&+\sum_{t=1}^{T}\sum_{i=1}^{I}\eta_i(t)\left(o_i^r\left(p_i(t),t\right)-MCP+z_i(t)^2\right)\\
&+\frac{c}{2}\sum_{t=1}^{T}\sum_{i=1}^{I}\left(o_i^r\left(p_i(t),t\right)-MCP(t)+z_i(t)^2\right)^2
\end{aligned}\tag{9}$$

where c is a positive penalty coefficient, and $\{z_i(t)^2\}$ are nonnegative slack variables converting inequality constraints (8) to equality constraints.[7] By minimizing $\{z_i(t)^2\}$ analytically [17, pp. 395–397], the relaxed problem is formed as

$$\min_{\{MCP(t)\},\{p_i(t)\}} L_c(\lambda,\eta,p,MCP)\tag{10}$$

$$\begin{aligned}
&with\ L_c(\lambda,\eta,MCP,p)\\
&=\sum_{i=1}^{I}\sum_{t=1}^{T}\{MCP(t)p_i(t)+S_i(t)\}\\
&+\sum_{t=1}^{T}\left\{\lambda(t)\left(P_d(t)-\sum_{i=1}^{I}p_i(t)\right)+\frac{c}{2}\left(P_d(t)-\sum_{i=1}^{I}p_i(t)\right)^2\right\}\\
&+\sum_{i=1}^{I}\sum_{t=1}^{T}\frac{1}{2c}\left\{max(0,\eta_i(t)+c\left(o_i^r(p_i(t),t)-MCP(t)\right))^2\right.\\
&\qquad\left.-\eta_i(t)^2\right\}.
\end{aligned}\tag{11}$$

The augmented Lagrangian L_c is inseparable because of the cross product terms of MCP and $\{p_i(t)\}$ and between elements of $\{p_i(t)\}$.

B. Surrogate Optimization for Offer Subproblems

To solve a separable problem using the standard Lagrangian relaxation technique, the relaxed problem is decomposed into subproblems, which are optimally solved. The subgradient obtained is then used to update the multipliers since it is a good direction. Because of the inseparability of the original problem (6), the relaxed problem cannot be decomposed. Rather, the relaxed problem is taken as a whole and optimized with respect to a particular offer or a particular MCP. Furthermore for an offer subproblem with a given set of multipliers $\{\lambda^k,\eta^k\}$ at iteration k, it is not necessary to find exact optimized decision variables $\{p_i^k\}$ to obtain a good direction. Rather, approximate optimization is sufficient if the new decision variables $\{p_i^k\}$ satisfy the following *surrogate optimization condition*:

$$\begin{aligned}
&L_c\left(\lambda^k,\eta^k,p_i^k,p_{j,j\neq i}^{k-1},MCP^{k-1}\right)\\
&\qquad< L_c\left(\lambda^k,\eta^k,p_i^{k-1},p_{j,j\neq i}^{k-1},MCP^{k-1}\right)
\end{aligned}\tag{12}$$

while keeping all other variables at their latest available values. If (12) (a special case of equation (28) in [14]) is satisfied, then the resulting "surrogate subgradient" is a good direction in the sense that it forms an acute angle with the direction toward the optimal multipliers. In solving the offer i subproblem, however, (12) may not be satisfied in view of variation in MCPs when the selection status of the offer is changed (to be illustrated later by using a two-offer one-hour problem). To overcome this difficulty, our solution is to appropriately adjust other offers based on the necessary condition for optimizing L_c in (11) with respect to $p_i(t)$ over its differentiable region to satisfy the following modified surrogate optimization conditions (with (12) as a special case):

$$\begin{aligned}
&L_c\left(\lambda^k,\eta^k,p_i^k,p_{j,j\neq i}^{k},MCP^{k-1}\right)\\
&\qquad< L_c\left(\lambda^k,\eta^k,p_i^{k-1},p_{j,j\neq i}^{k-1},MCP^{k-1}\right).
\end{aligned}\tag{13}$$

If (13) is satisfied, then the surrogate subgradient is guaranteed to be a good direction, and multipliers are updated.[8] Otherwise, all decision variables are kept at their old values, and another subproblem is solved until (13) is satisfied.

Solving Offer Subproblems: To solve the subproblem for offer i, the ON/OFF status of each hour must be determined as well as the offer level if ON. To this end, dynamic programming is used where times are stages, ON/OFF statuses for each hour are states, $S_i(t)$ is a state transition cost, and the stage-wise cost $v_i(t)$ is obtained by collecting all terms pertaining to t from L_c in (11) with the exception of $S_i(t)$

$$\begin{aligned}
v_i(t)&=MCP(t)\left(\sum_{j=1}^{I}p_j(t)\right)+\sum_{j=1,j\neq i}^{I}S_j(t)\\
&+\lambda(t)\left(P_d(t)-\sum_{j=1}^{I}p_j(t)\right)+\frac{c}{2}\left(P_d(t)-\sum_{j=1}^{I}p_j(t)\right)^2\\
&+\sum_{i=1}^{I}\frac{1}{2c}\left\{max(0,\eta_i(t)+c(o_i^r(p_i(t),t)-MCP(t)))^2\right.\\
&\qquad\left.-\eta_i(t)^2\right\}.
\end{aligned}\tag{14}$$

To evaluate $\{v_i(t)\}$ in (14), $MCP(t)$ and $(\sum_{j=1}^{I}p_j(t))$ must be determined. According to (5), $MCP(t)$ depends on the set of offers selected. A particular set of offers represents a distinct "sub-region," and is characterized by a unique $MCP(t)$. Therefore the structure of $v_i(t)$ for a sub-region may be different from those of others because of the differences in MCPs, which in turn depend on the set of offers selected. To determine $MCP(t)$ for an ON (or OFF) state in evaluating $\{v_i(t)\}$, the selection

[6]The presence of the MCP-offer inequality constraints in the augmented Lagrangian brings offer-specific cost information into the solution process.

[7]One way to update the Lagrangian multipliers $\{\lambda(t)\}$ and $\{\eta_i(t)\}$ in the augmented Lagrangian is to use the method of multipliers [17]. This approach typically requires the update of the penalty parameter from one iteration to the next. In our implementation, the penalty parameter is chosen to be sufficiently large and then fixed based on testing experience. The Lagrangian multipliers $\{\lambda(t)\}$ and $\{\eta_i(t)\}$ are then updated by using the surrogate subgradient method of [14] as will be presented in Section III-D.

[8]The proof of convergence of this algorithm is similar to that given in [14] and is thus omitted here.

TABLE I
OFFER PARAMETERS

	HOUR 1, System Demand = 60MW			
OFFERS	MIN MW	MAX MW	$/MW	SU. COST
1	5	80	20	0
2	5	45	10	100

status of other offers are assumed unchanged from the previous iteration. $MCP(t)$ in (14) is then obtained from solving MCP subproblems per Section III-C to be presented later if the sub-region under consideration is unchanged from the previous iteration. Otherwise, $MCP(t)$ is determined per (8) to speed up the convergence process.

To determine $(\sum_{j=1}^{I} p_j(t))$, one option is to keep other offers at their latest available values. However, this may lead to the violation of (12) as explained before. Thus, other offers are adjusted based on the necessary condition for optimizing L_c in (11) with respect to $p_i(t)$ over its differentiable region

$$\left(\sum_{j=1}^{I} p_j(t)\right) = \frac{(\lambda(t) - MCP(t))}{c} + P_d(t). \qquad (15)$$

If (15) is satisfied, then the penalty term on system demand in L_c for hour t, i.e., $P_d(t) - \sum_{j=1}^{I} p_j(t) = (\lambda(t) - MCP(t))/c$, is usually small, and (13) is likely to be satisfied. If $(\sum_{j=1}^{I} p_j(t))$ is smaller than the value given by (15), then offers that are currently selected are increased to their maximum offer levels sequentially starting with the least expensive offers in terms of their offer costs. If selected offers at their maximum offer levels cannot satisfy (15), then the least expensive offers in terms of amortized costs [the sum of energy and startup costs divided by the total power (MW)] that are currently off are sequentially selected until (15) is satisfied. If $(\sum_{j=1}^{I} p_j(t))$ is greater than the value given by (15), then the most expensive offers that are currently on are sequentially deselected until (15) is satisfied. In view that (15) is linear in terms of offer levels, the adjustments can be done efficiently.

With $\{v_i(t)\}$ evaluated for both ON and OFF states for all t, the ON/OFF status and the offer level for each hour of offer i are obtained by using dynamic programming following Guan *et al.*, 1992. Multipliers are then updated by using the surrogate subgradient method to be presented in Section III-D after solving one or a few subproblems to satisfy (13).

The rationale for maintaining the Lagrangian relaxation framework to solve one offer subproblem at a time is to reduce the computational complexity as compared to the approach of optimizing all offers at the same time. An example will be presented next to illustrate how an offer subproblem is solved and the need for adjusting other offers.

Illustrative Example of Solving Offer Subproblems: Consider a one-hour two-offer problem with offer parameters presented in Table I and three sub-regions depicted in Fig. 1. The p_1 axis in Fig. 1 represents the level of Offer 1 and is characterized by minimum/maximum power levels and an offer price as given in Table I and similarly for p_2. Three offer selections are possible as represented by three sub-regions: D_1 on the p_1 axis represents the case where Offer 1 is selected and Offer 2 is not; the rectangular D_2 represents the case where both Offers 1 and 2 are selected; and D_3 on the p_2 axis represents the case where Offer 2 is selected and Offer 1 is not. Each sub-region is characterized by a unique MCP.

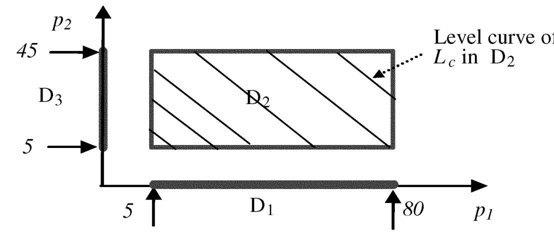

Fig. 1. Sub-regions for the two-unit one-hour example.

For this example, the augmented Lagrangian is given by

$$\begin{aligned} L_c = & MCP(p_1+p_2)+S_1+S_2+\lambda(P_d-p_1-p_2) \\ & +\frac{c}{2}(P_d-p_1-p_2)^2 \\ & +\frac{1}{2c}\left[max(0,\eta_1+c\,(o_1^r(p_1)-MCP))^2-\eta_1^2\right] \\ & +\frac{1}{2c}\left[max(0,\eta_2+c\,(o_2^r(p_2)-MCP))^2-\eta_2^2\right]. \end{aligned} \qquad (16)$$

For a given set of multipliers and MCP, L_c in (16) is quadratic in (p_1+p_2); therefore, level curves in D_2 are parallel lines with slope of negative one as shown in Fig. 1. Assume that in the previous iteration, both Offers 1 and 2 were selected, and an optimal point $(p_{1,D2}^{k-1} = 45, p_{2,D2}^{k-1} = 15)$ in D_2 is obtained with a feasible cost of \$1300. The global optimal solution, which can be easily determined from Table I, is in D_1 with only Offer 1 selected at 60 MW for a feasible cost of \$1200. For a new set of multipliers $\{\lambda^k, \eta^k\}$, consider the subproblem for Offer 2 where L_c in (16) is to be minimized with respect to p_2 to satisfy (13). Two cases with p_2 either ON or OFF should be considered. With p_2 ON, there are an infinite number of equivalent solutions in D_2 satisfying the system demand [e.g., (30, 30), (45, 15), and (15, 45)], and one of them is associated with the latest value of $p_1 = 45$. With p_2 OFF, if p_1 is kept at its previous value, then $p_{1,D2}^{k-1} + p_2^k = p_{1,D2}^{k-1} = 45 \ll P_d = 60$, and L_c in (16) includes a large penalty term causing (13) to be violated. Thus, it is necessary to adjust p_1 to $p_{1,D1}^k \approx 60$ by using (15) to satisfy (13). Although in this particular case the MCPs for D_1 and D_2 are the same, MCP in general might have to be obtained by solving MCP subproblems or by using (8). With the above adjustment, $p_1 \cong P_d$, and a comparison of the costs for the ON and OFF states of p_2 (\$1300 versus \$1200) shows that Offer 2 should be OFF. Without the above adjustment, the algorithm may get stuck in sub-region D_2 without ever finding the optimal set of offers.

C. Formulating and Solving MCP Subproblems

As discussed above, the difficulty for solving offer i subproblem is that $\{p_i\}$, $\{p_{j\neq i}\}$, $\{MCP(t)\}$ are coupled, and $\{MCP(t)\}$ affect the structure of the subproblem objective function depending on the set of selected offers. Therefore, a subproblem for offer i is formed by optimizing the augmented Lagrangian L_c in (11) as a whole, and decisions of other offers may have to be adjusted to satisfy the surrogate optimization conditions. To solve $\{MCP(t)\}$, in view that there is no cross product term of MCPs for different hours and MCP-offer constraints have been relaxed, MCPs for different hours can be solved independently for a given set of $\{p_i(t)\}$. Therefore, by collecting all terms involving $MCP(t)$ from (11) while keeping all other variables at their latest available values, T MCP subproblems are formed, one for each t

$$\min_{MCP(t)} L_{MCP(t)},\ with\ L_{MCP(t)} \equiv \left(\sum_{i=1}^{I} p_i(t)\right) MCP(t) + \frac{1}{2c} \times \left\{ \sum_{i=1}^{I} [max\{0, \eta_i(t) + c(o_i^r(p_i(t),t) - MCP(t))\}]^2 \right\}. \quad (17)$$

Similar to (13) for offer subproblems, a good direction for multiplier updating can be obtained if the following condition is satisfied for MCPs

$$L_c\left(\lambda^k, \eta^k, p_j^{k-1}, MCP^k\right) < L_c\left(\lambda^k, \eta^k, p_j^{k-1}, MCP^{k-1}\right). \quad (18)$$

To solve (17), note that $L_{MCP(t)}$ consists of a linear term $[(\sum p_i(t)) \cdot MCP(t)]$ and multiple quadratic terms depending on the magnitude of $\{\eta_i(t)/c + o_i^r(p_i(t),t)\}$ of individual offers. If $\{\eta_i(t)/c + o_i^r(p_i(t),t)\}$ is positive, then the maximum term is a quadratic function of $MCP(t)$. Otherwise, the maximum term takes zero value. Thus, the second term of $L_{MCP(t)}$ in (17) represents many "half-quadratics" demarcated by $\{\eta_i(t)/c + o_i^r(p_i(t),t)\}$, as depicted in Fig. 2 for a one-hour four-offer problem. In the figure, the $MCP(t)$ axis is divided into four segments (S1–S4) based on the demarcation of $\{\eta_i(t)/c + o_i^r(p_i(t),t)\}$, $i = 1, 2, 3$, and 4. In general, multiple segments are formed and the number of segments is limited by the number of offers. For each segment, the sum of an appropriate number of half-quadratics and the linear term is optimized and a segmental $MCP(t)$ is obtained. The segmental $MCP(t)$ that minimizes $L_{MCP(t)}$ in (17) then gives the overall $MCP(t)$. For example in segment S1 of Fig. 2, the segmental MCP is obtained by minimizing the sum of four half-quadratics and the linear term, while in S4, the segmental MCP is obtained by minimizing one half-quadratic and the linear term. Thus, the MCP subproblem can be solved without major computational requirements. With MCPs solved as above, if the condition in (18) is satisfied, then multipliers are updated by using the multiplier updating formula to be presented next. Otherwise, MCPs are kept at their old values and another subproblem is solved until (18) is satisfied.

D. Updating Multipliers and Stopping Criteria

The surrogate subgradient component with respect to the system demand multiplier $\lambda(t)$ is obtained from L_c in (11) as

$$\tilde{g}_\lambda^k(t) = P_d(t) - \sum_{i=1}^{I} p_i^k(t). \quad (19)$$

Similarly, the surrogate subgradient component with respect to the MCP-offer multiplier $\eta_i(t)$ for offer i is

$$\tilde{g}_{\eta_i}^k(t) = o_i^r\left(p_i^k(t), t\right) - MCP^k(t). \quad (20)$$

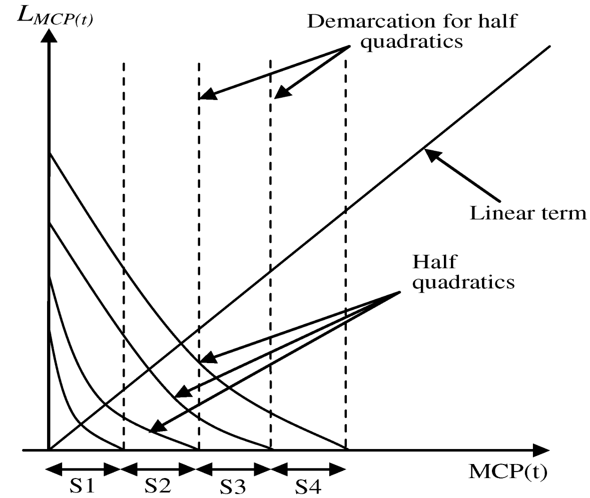

Fig. 2. Cost function components for $L_{MCP(t)}$.

All the multipliers are updated at the high level after solving one or a few subproblems while satisfying (13) or (18) based on the following surrogate subgradient updating formulae:

$$\lambda^{k+1}(t) = \lambda^k(t) + s^k \tilde{g}_\lambda^k(t) \quad (21)$$

$$\eta_i^{k+1}(t) = \max\left(0, \eta_i^k(t) + s^k \tilde{g}_{\eta_i}^k(t)\right) \quad (22)$$

where s^k is the step size at iteration k and is given by

$$0 < s^k < \frac{\left(L^* - L_c^k\right)}{\left(\left\|\tilde{g}_\lambda^k\right\|^2 + \left\|\tilde{g}_\eta^k\right\|^2\right)}. \quad (23)$$

In the above, L^* is the optimal dual value, and $\tilde{g}_\lambda^k$ and $\tilde{g}_\eta^k$ are, respectively, the column vectors of the surrogate subgradient components with respect to λ and $\{\eta_i\}$. Since L^* is generally unknown, the best feasible cost obtained thus far is used as an approximation for L^*.

E. Constraint Violation and Stopping Criteria

The level of constraint violation with respect to the system demand equality constraint (2) at time t is given by the absolute value of $\tilde{g}_\lambda(t)$, i.e., $\|\tilde{g}_\lambda(t)\|$. For MCP-offer inequality constraints (8), if $\tilde{g}_{\eta_i}(t)$ is negative, then the constraint is satisfied and there is no violation. If $\tilde{g}_{\eta_i}(t)$ is positive, then the constraint is violated. Consequently, the level of constraint violation is given by $\max(0, \tilde{g}_{\eta_i}^k(t))$. Let $\max(0, \tilde{g}_\eta^k)$ denote the column vector of $\max(0, \tilde{g}_{\eta_i}^k(t))$ for all i and t, then the level of constraint violation for the entire problem can be measured by the L-2 norm of the following constraint violation vector:

$$g_{cv}^k = \begin{pmatrix} \left|g_\lambda^k\right| \\ max\left(0, g_\eta^k\right) \end{pmatrix}. \quad (24)$$

The iterative process is terminated if the number of iterations is greater than a preset value or if the level of constraint violation is less than a specified small positive number.

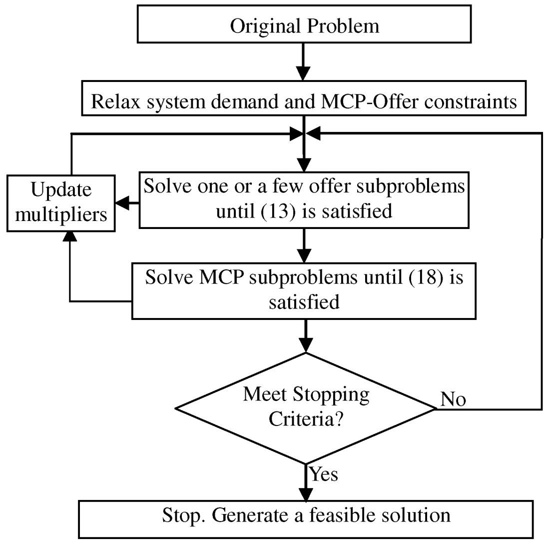

Fig. 3. Flow chart of the algorithm.

F. Generating Feasible Solutions

Since the system demand and MCP-offer constraints have been relaxed, subproblem solutions, when put together, may not constitute a feasible solution. In addition, in view of the linear level curves as depicted in Fig. 1, there are generally an infinite number of equivalent solutions with the same feasible cost J. Simple heuristics are thus used to generate a reasonable feasible solution based on subproblem solutions. In the heuristics, selected offers with prices lower than $\{MCP(t)\}$ are awarded full power, and selected offers with prices equal to $\{MCP(t)\}$ are adjusted to satisfy the system demand constraints (2). If the selected offers cannot meet the system demand, a nonselected offer with the lowest amortized per MW cost is chosen, and this process is repeated until the system demand is met. MCPs are then updated by using definition (8). The flow chart is given in Fig. 3.

IV. Numerical Results

The above algorithm for payment cost minimization was implemented in C++ on a Pentium-IV 1.3-GHz personal computer. For comparison purposes, offer cost minimization was also solved by using augmented Lagrangian relaxation with multipliers updated by using the standard surrogate subgradient method. In this section, three examples are presented. Example 1 illustrates the subtle differences between the two methods and shows that payment cost minimization takes into account the effect of MCPs on actual payment for a given set of offers. Example 2 tests the scalability of the method when applied to a reasonably sized problem and shows that significant savings over offer cost minimization can be achieved. Example 3 establishes that the performance of the payment cost minimization method is significantly better than that of offer cost minimization with New England system demand data averaged over five days in May 1999 and for a set of randomly generated offers.

Example 1: Consider a simple four-offer two-hour problem as described in Table II with all offers in an off state at Hour 0.

The results of offer cost minimization are summarized in Table III, and the results of payment cost minimization are summarized in Table IV. For this small problem, the optimal solution is known, and so the duality gap, which is the relative difference between the optimal feasible cost and the surrogate dual cost, can be evaluated. The duality gaps for both cases are 0%.

Since Offers 1 and 2 are comparatively cheaper than Offers 3 and 4 and have no startup costs, they should be selected and awarded their maximum levels for both hours. Their combined capacity of 90 and 120 MW for hours 1 and 2, respectively, is less than the system demand.

TABLE II
Supply Offer Parameters for Example 2

Hour 1, System Demand = 100 MW				
	Min MW	Max MW	$/MW	Start Up Cost
Offer 1	5	50	10	0
Offer 2	5	40	20	0
Offer 3	0	10	65	50
Offer 4	5	60	30	1800
Hour 2, System Demand = 150 MW				
	Min MW	Max MW	$/MW	Start Up Cost
Offer 1	5	60	15	0
Offer 2	5	60	20	0
Offer 3	0	30	65	50
Offer 4	5	100	30	1800

TABLE III
Results for Example 1, Offer Cost Minimization

Hour 1, MCP = 65$/MW				
	MW	$/MW	Offer Costs	Payment Costs
Offer 1	50	10	500	3250
Offer 2	40	20	800	2600
Offer 3	10	65	700	700
Offer 4	0	30	0	0
Total Cost			**$2,000**	**$6,550**
Hour 2, MCP = 65$/MW				
	MW	$/MW	Offer Costs	Payment Costs
Offer 1	60	15	900	3900
Offer 2	60	20	1200	3900
Offer 3	30	65	1950	1950
Offer 4	0	30	0	0
	Total Cost		**$4,050**	**$9,750**
Total Offer Costs = $6,050				
Total Payment Costs = $16,300				

TABLE IV
Results for Example 1, Payment Cost Minimization

Hour 1, MCP = 30 $/MW				
	MW	$/MW	Offer Costs	Payment Costs
Offer 1	50	10	500	1500
Offer 2	40	20	800	1200
Offer 3	0	65	0	0
Offer 4	10	30	2100	2100
Total Cost			**$3,400**	**$4,800**
Hour 2, MCP = 30 $/MW				
	MW	$/MW	Offer Costs	Payment Costs
Offer 1	60	15	900	1800
Offer 2	60	20	1200	1800
Offer 3	0	65	0	0
Offer 4	30	30	900	900
Total Cost			**$3,000**	**$4,500**
Total Offer Costs = $6,400				
Total Payment Costs = $9,300				

To satisfy the remaining system demand of 10 MW for hour 1 and 30 MW for hour 2, offer cost minimization compares the incremental offer costs of choosing Offer 3 or 4. If Offer 3 is selected for both hours, the incremental offer cost is \$2650 (\$2600 energy charge and \$50 startup cost). If Offer 3 is selected for hour 1 but Offer 4 is selected for hour 2, the incremental cost is \$3400 (\$1550 for energy, \$1850 startup cost). If Offer 4 is selected for both hours, the incremental offer cost is \$3000 (\$1200 energy charge and \$1800 startup cost). Therefore under offer cost minimization, Offer 3 is selected for both hours. However, Offer 3 sets the MCP of \$65/MW for both hours, leading to an actual consumer payment of \$16 300, which is significantly higher than the minimized cost of \$6050. This clearly demonstrates that offer cost minimization minimizes the offer costs without considering the effects of MCPs on the payment by consumers.

To satisfy the remaining system demand of 10 and 30 MW for hours 1 and 2, respectively, payment cost minimization compares the incremental payment cost of Offers 3 or 4. If Offer 3 is selected for both hours, the MCPs will be \$65, and the incremental payment cost will be \$12 100 $[(\$65 - \$20) \cdot 90$ MW $+ (\$65 - \$20) \cdot 120$ MW (incremental cost for Offers 1 and 2)$+ \$65 \cdot 10 + 65 \cdot 30$ (energy charge for last 10 and 30 MW respectively) $+$ 50 (startup)$]$. If Offer 3 is selected for hour 1 but Offer 4 is selected for hour 2, the MCP will be \$65 for hour 1 and \$30 for hour 2, and the incremental payment cost will be \$8650 $[(\$65 - \$20) \cdot 90$ MW $+ (\$30 - \$20) \cdot 120$ MW (incremental cost for Offers 1 and 2)$+ \$65 \cdot 10 + \$30 \cdot 30$ (energy charge for last 10 and 30 MW respectively) $+$ 1850 (startup)$]$. If Offer 4 is selected for both hours, the MCPs will be \$30, and the incremental payment cost will be \$5100$[(\$30 - \$20) \cdot 90$ MW $+ (\$30 - \$20) \cdot 120$ MW (incremental cost for Offers 1 and 2)$+ \$30 \cdot 10 + 30 \cdot 30$ (energy charge for last 10 and 30 MW respectively) $+$ 1800 (startup)$]$. Thus, Offer 4 is selected for both hours, leading to a total payment cost of \$9300 by consumers. It can thus be seen that the payment cost minimization takes into account the effects of MCPs on the actual payment by consumers, and the minimized cost and the actual payment cost are the same at \$9300, which is significantly lower than the current ISO practice with a savings of \$7000 for the given set of offers.

If offers represent true production costs, then the total offer costs obtained under offer cost minimization represents the costs associated with the most efficient selection of offers. The difference in total offer costs between payment cost minimization and offer cost minimization can then be viewed as the cost of "production inefficiency" by using payment cost minimization. From Tables III and IV, the above difference is \$350 (or 5.8% of the total offer costs under offer cost minimization). Thus, even if it is assumed that offers represent true production costs, the loss in production efficiency is only a fraction of the payment savings of \$7000 (or 75.3% of the total payment costs) achieved by using payment cost minimization.

Example 2: Consider a 24-h problem with 25 participants, each submitting a single constant block offer for all the hours. The system demand is given in Table V, and offer parameters are presented in Table VI with a total capacity of 4620 MW. Among the offers, there are four nuclear plants with low offer prices (between \$30 and \$37/MW) but very high startup costs for a total capacity of 1305 MW. There are four base load plants with offer prices between \$40 and \$47/MW and startup costs lower than those for nuclear plants for a total capacity of 1110 MW. Eleven cycling plants make up 1590 MW and have prices between \$55 and \$70/MW. Six gas turbines make up the remaining 615 MW, and among them, three have prices between \$75 and \$80/MW with relatively high startup costs, while the other three have high prices between \$90 and \$95/MW but relatively low startup costs. Assume that initially all nuclear and base load units were on while other units were off.

TABLE V
SYSTEM DEMAND FOR EXAMPLE 2

HOUR	SYSTEM DEMAND	HOUR	SYSTEM DEMAND
1	2500	13	3800
2	2550	14	3900
3	2570	15	4200
4	2530	16	4250
5	2650	17	4400
6	2700	18	4500
7	2680	19	4100
8	2740	20	3850
9	2850	21	3400
10	2900	22	2700
11	3500	23	2300
12	3700	24	2350

TABLE VI
SUPPLY OFFERS FOR EXAMPLE 2

HOUR 1-24				
SUPPLY OFFERS	MIN MW	MAX MW	\$/MW	START UP COST
1	100	455	30	1200
2	60	350	34	1150
3	50	300	35	1100
4	30	200	37	1000
5	40	350	40	350
6	40	320	42	350
7	30	240	45	300
8	30	200	47	300
9	20	190	55	180
10	20	180	57	180
11	20	170	58	175
12	20	160	59	160
13	20	150	60	250
14	20	140	62	200
15	20	130	63	150
16	20	125	65	180
17	20	120	66	140
18	20	115	68	140
19	20	110	70	160
20	20	120	75	250
21	20	115	78	250
22	20	110	80	300
23	20	100	90	50
24	15	90	93	50
25	10	80	95	40

The hourly MCPs for the two methods together with system demand are depicted in Fig. 4. Observe that the MCPs under payment cost minimization are usually less than those under offer cost minimization, and this translates to lower actual payment costs. Table VII summarizes various costs and shows that for this example, \$134 945 (or 2.56%) savings can be achieved. Annual energy costs in the U.S. run in the tens of billions of dollars; therefore, even a 0.1% savings represents significant savings of tens of millions of dollars annually. The difference in total offer costs between payment cost minimization and offer

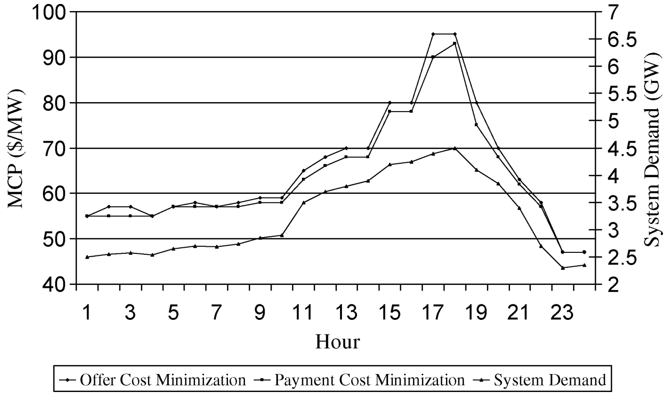

Fig. 4. Hourly MCPs for example 2.

TABLE VII
SUMMARY OF COSTS FOR EXAMPLE 2

Offers	Payment Cost Minimization Payments	Payment Cost Minimization Offer Costs	Offer Cost Minimization Payments	Offer Cost Minimization Offer Costs
1	694620	330620	710487	328530
2	534550	288150	547141	286745
3	458300	254300	469098	253099
4	305800	179400	312995	178597
5	533750	337750	546343	336345
6	488030	324190	499545	322907
7	366060	260460	372452	257349
8	296640	218240	306076	219677
9	257030	215990	239530	192230
10	192045	159645	200666	162257
11	148145	119755	168132	134199
12	129600	104640	140672	110753
13	121600	99850	123529	97230
14	113150	95930	109163	87847
15	95055	80365	102706	84317
16	85680	73805	85283	70293
17	76940	66620	79030	65633
18	61690	53610	64919	54224
19	45700	40440	56422	47622
20	48055	43615	42250	36250
21	36115	33220	36683	32409
22	20430	19800	30217	26917
23	13850	13550	9550	9050
24	4700	4700	3406	3335
25	0	0	6184	6184
Total	**5127535**	**3418645**	**5262480**	**3403999**

cost minimization is $14 646 (or 0.43% of the total offer costs under offer cost minimization). Thus, even if it is assumed that offers were made based on marginal costs, the loss in production efficiency is only a fraction of the payment savings achieved by using payment cost minimization.

To test scalability, the payment cost minimization method was applied to problems with varying number of hours[9] or number of participants. The CPU times needed for obtaining good results are summarized in Table VIII. It can be seen that the method scales well since the increase in computational time is almost linear (at least it is not exponential) with an increase in the number of hours or offers.

[9]CPU times for different hours are presented in Table VIII to illustrate the scalability of the method. In most ISOs in the U.S., the scheduling horizon for the day-ahead market is 24 hours.

TABLE VIII
SCALABILITY RESULTS

Number of Participants	Scheduling Period (in hours)	CPU Time (in secs)
10	12	8.88
10	24	17.5
10	48	35.27
15	12	12.38
15	24	24.36
15	48	48.31
20	12	15.63
20	24	31.69
20	48	69.38
25	12	19.47
25	24	41.04
25	48	89.97
30	12	34.8
30	24	76.53
30	48	163.69

TABLE IX
CHARACTERISTICS OF SUPPLY OFFERS FOR EXAMPLE 3

% of Supply Capacity	Number of Offers	Mean MW/Offer	Std. Dev. MW/Offer	Mean Price $/MW
45 (Nuclear)	5	1745	200	15
15 (Hydro)	4	728	50	21
20 (Thermal)	8	485	30	22
20 (Thermal)	8	485	30	32
% of Supply Capacity	Std. Dev. $/MW	Mean SU. Cost ($)	Std. Dev. of SU. Cost ($)	
45 (Nuclear)	2	0	0	
15 (Hydro)	1	40	10	
20 (Thermal)	1.4	500	100	
20 (Thermal)	1.2	60	15	

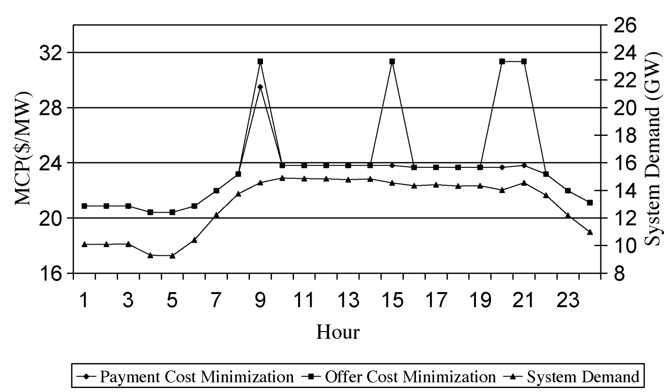

Fig. 5. Average hourly MCPs for example 3.

Example 3: Consider a problem with 25 participants over a 24-h period. The system demand was generated based on the means and standard deviations of the New England load over five days in May 1999. Using these historical values, 35 sets of system demand were randomly generated based on Gaussian distributions. Twenty-five supply offers were also randomly generated with Gaussian distributions based on Table IX. These supply offers were designed so that the average supply capacity is 30% above the maximum mean demand. The offers are made up of nuclear, hydro, and thermal units, each with identical parameters over the day.

The average hourly MCPs obtained for the 35 sets of system demand for both methods are depicted in Fig. 5, together with

TABLE X
SUMMARY OF AVERAGE COSTS FOR EXAMPLE 3

Offers	Payment Cost Minimization Payments	Payment Cost Minimization Offer Costs	Offer Cost Minimization Payments	Offer Cost Minimization Offer Costs
1	893009	681238	931074	671773
2	760979	579223	812818	588329
3	976040	634037	1016453	630954
4	989933	675337	1035509	674440
5	832422	556807	872574	557842
6	7117	7056	4260	4162
7	24249	23168	20981	18895
8	15016	14994	14710	14620
9	18874	18494	12039	11850
10	15466	14622	3556	3501
11	5402	5320	2577	2498
12	4574	4474	2175	2153
13	3187	3121	11374	10975
14	91126	90815	97701	86017
15	157047	152658	186434	170007
16	155814	154172	83719	81026
17	223006	196576	255653	216270
18	205100	190045	209031	183003
19	221510	202872	225929	196320
20	247046	212972	267754	220401
21	194145	182290	213519	190456
22	349938	312698	372780	314795
23	287407	280897	275388	248956
24	310458	279482	322441	271713
25	227253	205801	319996	269782
Total	**7216120**	**5679169**	**7570446**	**5640737**

system demand. Similar to Example 2, the MCPs for payment cost minimization are usually less than those for offer cost minimization. As summarized in Table X, the average savings in payment costs is \$354 326, or 4.7%, demonstrating the consistent performance of the new method.

V. CONCLUSION

Currently most ISOs in the U.S. conduct auctions by using the offer cost minimization mechanism but settle the markets with the pay-at-MCP scheme. Under this system, the auction and settlement mechanisms are inconsistent, and this can lead to higher payment costs for consumers for a given set of offers. A new method that uses augmented Lagrangian relaxation within the surrogate optimization framework has been presented for solving the payment cost minimization problem. To the best of the authors' knowledge, this is the first systematic and practical method for solving the payment cost minimization problem. It can facilitate the discussion on which auction mechanism ISOs should use by providing a tool that can be used for comparative analysis and allows ISOs to adopt this auction if so desired. Numerical results show that the new method is viable and can lead to significant savings for consumers for the given set of offers since it considers the impact of MCPs on total payment costs while offer cost minimization does not. The methodology has been extended to consider more realistic conditions such as demand bids and partial compensation of startup costs in [21]. Research is currently ongoing to incorporate more practical aspects of the deregulated markets, such as ancillary services, and transmission congestion under a location-based market pricing framework. Finally, the method is generic and can be extended to other NP hard problems. In particular, the successful treatment of the inseparability issues encountered with this problem opens the door for the resolution of many inseparable integer or mixed integer problems.

REFERENCES

[1] R. Baldick, "The generalized unit commitment problem," *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 465–475, Feb. 1995.

[2] P. Carpentier, G. Cohen, J. C. Culioli, and A. Renaud, "Stochastic optimization of unit commitment: A new decomposition framework," *IEEE Trans. Power Syst.*, vol. 11, no. 2, pp. 1067–1073, May 1996.

[3] X. Guan, Z. Qiaozhu, and A. Papalexopoulos, "Optimization based methods for unit commitment: Lagrangian relaxation versus general mixed integer programming," in *Proc. IEEE Power Engineering Society General Meeting*, vol. 2, Jul. 2003, pp. 13–17.

[4] N. P. Padhy, "Unit commitment—a biliographical survey," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1196–1205, Aug. 2004.

[5] J. H. Yan and G. A. Stern, "Simultaneous optimal auction and unit commitment for deregulated electricity markets," *Elect. J.*, Nov. 2002.

[6] J. M. Jacobs, "Artificial power networks and unintended consequences," *IEEE Trans. Power Syst.*, vol. 12, no. 2, pp. 968–972, May 1997.

[7] S. Hao, G. A. Angelidis, H. Singh, and A. D. Papalexopoulos, "Consumer payment minimization in power pool auctions," *IEEE Trans. Power Syst.*, vol. 13, no. 3, pp. 986–991, Aug. 1998.

[8] J. Alonso, A. Trias, V. Gaitan, and J. J. Alba, "Thermal plant bids and market clearing in an electricity pool: Minimization of costs vs. minimization of consumer payments," *IEEE Trans. Power Syst.*, vol. 14, no. 4, pp. 1327–1334, Nov. 1999.

[9] C. Vazquez, M. Rivier, and I. J. Perez-Arriaga, "Production cost minimization versus consumer payment minimization in electricity pools," *IEEE Trans. Power Syst.*, vol. 17, no. 1, pp. 119–127, Feb. 2002.

[10] D. P. Mendes, "Resource scheduling and pricing in a centralised energy market," in *Proc. 14th Power System Computation Conf.*, Seville, Spain, Jun. 2002, pp. 1–7.

[11] S. Hao and F. Zhuang, "New models for integrated short-term forward electricity markets," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 478–485, May 2003.

[12] S. Raikar and M. Ilic, "Assessment of transmission congestion for major electricity markets in the US," in *Proc. IEEE Power Engineering Society Summer Meeting*, vol. 2, Jul. 2001, pp. 1152–1156.

[13] J. M. Arroyo and A. J. Conejo, "Multiperiod auction for a pool-based electricity market," *IEEE Trans. Power Syst.*, vol. 17, no. 4, pp. 1225–1231, Nov. 2002.

[14] X. Zhao, P. B. Luh, and J. Wang, "Surrogate gradient algorithm for Lagrangian relaxation," *J. Optim. Theory Appl.*, vol. 100, no. 3, pp. 699–712, Mar. 1999.

[15] G. Cohen and D. L. Zhu, "Decomposition coordination methods in large scale problems: The nondifferentiable case and the use of augmented Lagrangians," *Adv. Large Scale Syst.*, vol. 1, pp. 203–266, 1984.

[16] S. J. Wang and S. M. Shahidehpour, "Short-term generation scheduling with transmission and environmental constraints using an augmented Lagrangian relaxation," *IEEE Trans. Power Syst.*, vol. 10, no. 3, pp. 1294–1301, Aug. 1995.

[17] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.

[18] S. Al-Agtash, "Hydrothermal scheduling by augmented Lagrangian: Consideration of transmission and pumped storage units," *IEEE Trans. Power Syst.*, vol. 16, no. 4, pp. 750–756, Nov. 2001.

[19] H. Yan, P. B. Luh, and L. Zhang, "Scheduling of hydrothermal power systems using the augmented Lagrangian decomposition and coordination technique," in *Proc. American Control Conf.*, 1994.

[20] Q. Zhai, X. Guan, and J. Cui, "Unit commitment with identical units successive subproblem solving method based on Lagrangian relaxation," *IEEE Trans. Power Syst.*, vol. 17, no. 4, pp. 1250–1257, Nov. 2002.

[21] Y. Chen, P. B. Luh, J. Yan, G. Stern, W. E. Blankson, and F. Zhao, "Payment minimization auction with demand bids and partial compensation of startup costs for deregulated electricity markets," in *Proc. IEEE Power Engineering Society General Meeting*, San Francisco, CA, Jun. 2005.

Peter B. Luh (M'80–SM'91–F'95) received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, R.O.C., in 1973, the M.S. degree in aeronautics and astronautics engineering from Massachusetts Institute of Technology, Cambridge, in 1977, and the Ph.D. degree in applied mathematics from Harvard University, Cambridge, in 1980.

Since then, he has been with the University of Connecticut, Storrs, and currently is the SNET Professor of Communications and Information Technologies in the Department of Electrical and Computer Engineering. He is a Visiting Professor at the Center for Intelligent and Networked Systems, Department of Automation, Tsinghua University, Beijing, China. He is interested in planning, scheduling, and coordination of design, manufacturing, and supply chain; configuration and operation of building elevator and HVAC systems; and schedule, auction, portfolio optimization, and load/price forecasting for power systems and decision-making under uncertain or distributed environments.

Dr. Luh is the founding Editor-in-Chief of the new IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING, an Associate Editor of *IIE Transactions on Design and Manufacturing*, an Associate Editor of *Discrete Event Dynamic Systems*, and was the Editor-in-Chief of IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION (1999–2003).

William E. Blankson (S'04) received the B.S. degree in electrical engineering from the Kwame Nkrumah University of Science and Technology, Kumasi, Ghana, in 1997 and the M.S degree in electrical engineering from the University of Connecticut, Storrs, in 2003. He is currently working toward the Ph.D. degree in electrical engineering at the University of Connecticut.

From 1998–2001, he worked as a Controls and Instrumentation Engineer at the Takoradi Thermal Power Plant in Ghana. He worked as an Intern in the Resource Planning and Strategy Department of Southern California Edison Company, Rosemead, CA, in the summer of 2004 and also as a Market Support Intern with the Electric Reliability Council of Texas (ERCOT) during the summer of 2005. His research interests include optimization theory and algorithms, neural networks, power systems, and their application to the restructured electricity markets.

Ying Chen (S'04) received the B.S. degree in automatic control from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1999. She is currently working toward the Ph.D. degree in electrical and computer engineering at the University of Connecticut, Storrs.

Her research interests include optimization, operations, and economics of electricity markets.

Joseph H. Yan received the Ph.D. degree in electrical and systems engineering from the University of Connecticut, Storrs, in 1996.

He is a Project Manager in the Market Monitoring and Analysis Division of Southern California Edison, Rosemead, CA. For the last ten years, he has worked in areas of the electricity system operations, wholesale energy market analysis for both regulated and nonregulated affiliates, market monitoring and market design in California, and California Electricity Refund case of the 2000–2001 crisis. His research interest includes operation research, optimization, unit commitment/scheduling and transaction evaluation, and optimal simultaneous auction in deregulated ISO/RTO markets.

Gary A. Stern received the Ph.D. degree in economics from the University of California at San Diego in 1984.

He is the Director of Market Monitoring and Analysis for Southern California Edison Company (SCE), Rosemead, CA. He reports to the Senior Vice President of the Power Production Business Unit and manages a division responsible for monitoring the wholesale electricity market in California. Gary leads SCE's efforts, working with the California Independent System Operator (CAISO) and the Federal Energy Regulatory Commission (FERC), to ensure that the wholesale electricity market design being implemented by the CAISO results in a reliable and efficient outcome for SCE as a large buyer and seller of electricity on behalf of its customers. His recent endeavors have included managing SCE's efforts to obtain several billion dollars in refunds for market manipulation and market power abuse during the 2000–2001 electricity crisis in California. His participation in the refund settlements negotiations has resulted in several FERC decisions that have refunded over $2 billion to California Utilities. He is also managing the design of a new resource adequacy requirement partnering with the California Public Utilities Commission (CPUC), the California Energy Commission (CEC), and other stakeholders.

Shi-Chung Chang (S'83–M'87) received the B.S.E.E. degree from National Taiwan University, Taiwan, R.O.C., in 1979 and the M.S. and Ph.D. degrees in electrical and systems engineering from the University of Connecticut, Storrs, in 1983 and 1986, respectively.

From 1979 to 1981, he served as an Ensign in the Chinese Navy, Taiwan. In 1987, he was a member of the Technical Staff, decision systems section, ALPHATECH, Inc., Burlington, MA. He has been with the Electrical Engineering Department of National Taiwan University since 1988 and was promoted to Professor in 1994. During 2001–2002, he was the Dean of Student Affairs and a Professor of electrical engineering, National Chi Nan University, Pu-Li, Taiwan. He is now jointly appointed by the Graduate Institute of Industrial Engineering and the Graduate Institute of Communication Engineering, National Taiwan University. He has been a principal investigator and consultant to many industry and government-funded projects. His research interests include optimization theory and algorithms, production scheduling and control, network management, Internet economics, and distributed decision making.

Dr. Chang is a Member of Eta Kappa Nu and Phi Kappa Phi. In 1996, he received the award of outstanding achievements in University-Industry Collaboration from the Ministry of Education for his pioneering research collaborations with Taiwan semiconductor industry on production scheduling and control.

Feng Zhao (S'04) received the B.S. degree in automatic control from Shanghai JiaoTong University, Shanghai, China, in 1998 and the M.S. degree in control theory and control engineering from Tsinghua University, Beijing, China, in 2001. He is currently working toward the Ph.D. degree in electrical and computer engineering at the University of Connecticut, Storrs.

His research interests include optimization, operations, and economics of electricity markets.