A Primal-Relaxed Dual Global Optimization Approach[1]

C.A. FLOUDAS[2] AND V. VISWESWARAN[3]

Communicated by R. Sargent

JOTA, Vol. 78(2), p. 187 (1993).

---

[1]The authors gratefully acknowledge financial support from National Science Foundation Presidential Young Investigator Award CBT-8857013. The authors are also grateful to Drs. F. A. Al-Khayyal, B. Jaumard, P. M. Pardalos and H. D. Sherali for helpful comments on earlier draft of this paper.
[2]Associate Professor, Department of Chemical Engineering, Princeton University, Princeton, New Jersey.
[3]Graduate Student, Department of Chemical Engineering, Princeton University, Princeton, New Jersey.

Abstract. A deterministic global optimization approach is proposed for nonconvex constrained nonlinear programming problems. Partitioning of the variables, along with the introduction of transformation variables, if necessary, convert the original problem into *primal* and *relaxed dual* subproblems that provide *valid* upper and lower bounds respectively on the global optimum. Theoretical properties are presented which allow for a rigorous solution of the relaxed dual problem. Proofs of ϵ-finite convergence and ϵ-global optimality are provided. The approach is shown to be particularly suited to (a) quadratic programming problems, (b) quadratically constrained problems, and (c) unconstrained and constrained optimization of polynomial and rational polynomial functions. The theoretical approach is illustrated through a few example problems. Finally, some further developments in the approach are briefly discussed.

Key Words. Global optimization, quadratic programming, polynomial functions, ϵ-optimal solutions.

1. Introduction

Global optimization of nonconvex programming problems has generated a lot of interest in recent years. Surveys, books and applications for global optimization are available by Dixon and Szego (Refs. 1 and 2), Archetti and Schoen (Ref. 3), Pardalos and Rosen (Refs. 4 and 5), Torn and Zilinskas (Ref. 6), Ratschek and Rokne (Ref. 7), Mockus (Ref. 8), Horst and Tuy (Ref. 9) and Floudas and Pardalos (Refs. 10 and 11). The deterministic approaches for global optimization can be largely classified as : (a) Lipschitzian methods (e.g. Ref. 12); (b) Branch and bound methods (e.g. Refs. 13-15); (c) Cutting Plane Methods (e.g. Ref. 16); (d) Difference of convex (D.C.) and Reverse convex function methods (e.g. Refs. 17 and 18); (e) Outer approximation methods (e.g. Refs. 19 and 20); (f) Primal-Dual methods (e.g. Refs. 21-23); (g) Linearization methods (e.g. Ref. 24); and (h) Interval methods (e.g. Ref. 25). Recent developments in global optimization approaches can be found in Ref. 11.

In this paper, a primal-relaxed dual approach for global optimization is proposed (earlier versions of this work have appeared in Floudas and Visweswaran (Ref. 26) and Visweswaran and Floudas (Ref. 27)). It is related to the work of Geoffrion (Ref. 28) and Wolsey (Ref. 29). It does not require Property (P) stated in Ref. 28, and it differs from the resource decomposition algorithm of Wolsey (Ref. 29) in the way the relaxed dual problem is formulated and solved. A statement of the global optimization problem is given in Section 2, while Section 3 presents the relevant part of duality theory (Extensive discussion of duality theory for decomposition can be found in Flippo (Ref. 30)). Section 4 contains the new theoretical results. Section 5 illustrates the branch-and-bound nature of the proposed algorithm and discusses some properties of the branching that can be used to improve the efficiency of the algorithm. Section 6 describes the global optimization algorithm. Section 7 provides the proofs of finite ϵ-convergence and ϵ-global optimality. The application of the algorithm to two illustrating examples is considered in Sections 8 and 9, while Section 10 contains a geometrical interpretation of the algorithm. Sections 11 and 12 discuss the extensions of Section 4 to quadratically constrained problems and problems with polynomial functions.

2. Statement of the Problem

The global optimization problem addressed in this paper is stated as:

Determine a globally ϵ-optimal solution of the following problem:

$$\min_{x,y} \; f(x,y), \tag{1a}$$

$$s.t. \qquad g(x,y) \;\leq\; 0, \tag{1b}$$

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

$$x \in X, \tag{1d}$$

$$y \in Y, \tag{1e}$$

where $X \in \Re^{n}$ and $Y \in \Re^{n_2}$ are non-empty, compact, convex sets, $g(x,y)$ is an m-vector of inequality constraints and $h(x,y)$ is a p-vector of equality constraints. In this paper, it will be assumed that X consists of bounds on the x variables, and will be incorporated into the constraint set $g(x,y) \leq 0$. It is assumed that the functions $f(x,y)$, $g(x,y)$ and $h(x,y)$, along with any Lagrange function formulated for the problem, are continuous, piecewise differentiable and given in analytical form. The variables y are defined such that the following conditions are satisfied:

Conditions (A)

(a) $f(x,y)$ is convex in x for every fixed y, and convex in y for every fixed x.

(b) $g(x,y)$ is convex in x for every fixed y, and convex in y for every fixed x.

(c) $h(x,y)$ is affine in x for every fixed y, and affine in y for every fixed x.

To identify the classes of mathematical programming problems that can be represented within the framework of (1) and satisfy *Conditions (A)*, the concepts of *partitioning* and *transformations* are used. Using these concepts (see Ref. 22), it can be shown that the proposed approach is applicable to problems involving quadratic and/or polynomial/rational polynomial terms in the objective function and/or constraints. Therefore, the classes of problems addressed by this paper include bilinear programming problems, general quadratic programming problems, quadratic problems with quadratic constraints, polynomial and rational polynomial programming problems among others.

Recently, Hansen and Jaumard (Ref. 31) have proposed an algorithm for the efficient bilinearization of quadratic and polynomial function problems, rational polynomials, and problems involving hyperbolic functions. For a given problem in these classes, the algorithm provides the set of new variables that must be introduced in order to convert the problem into bilinear form. The bilinearization can be achieved with the objective of minimizing either the number of complicating variables (in the context of this paper, this is simply the number of y variables) or the number of variables that must be introduced in order to make the problem completely bilinear. Moreover, given a bilinear problem with variable subsets x and y, the algorithm can also be used to identify any changes in these subsets that will result in a smaller number of either the x or the y variables.

3. Duality Theory

Define the following problem as the **Primal Problem :**

$$\min_x \quad f(x, y^k), \tag{2a}$$

$$s.t. \quad g(x, y^k) \; \leq \; 0, \tag{2b}$$

$$h(x, y^k) \; = \; 0, \tag{2c}$$

where $y^k \in Y$. Here, it is assumed that the bounds on the x variables are incorporated into the first set of constraints. Since this problem is simply (1) solved for fixed values of $y = y^k$, it represents an upper bound on the optimal value of (1).

Using the concept of *projection* (Ref. 28), (1) can be converted to an equivalent formulation, featuring an inner and outer optimization problem :

$$\min_y \qquad v(y), \tag{3a}$$

$$s.t. \quad v(y) \; = \; \min_x \; f(x, y), \tag{3b}$$

$$s.t. \quad h(x, y) = 0, \tag{3c}$$

$$g(x, y) \leq 0, \tag{3d}$$

$$y \; \in \; Y \cap V, \tag{3e}$$

$$V \; \equiv \; \{y \; : \; h(x, y) = 0, \; g(x, y) \leq 0 \quad for \; some \; x\} \, . \tag{3f}$$

From *Conditions (A)* and Slater's constraint qualification, (2) satisfies the conditions of the **Strong Duality Theorem** (Theorem 6.2.4 of Ref. 32). Then, the solution of (2), for any fixed $y = y^k$, is identical to the solution of its corresponding dual problem on $Y \cap V$. That is,

$$\begin{aligned} &\min_{\boldsymbol{x}} f(x, y^k) \, , \quad s.t. \quad g(x, y^k) \leq 0, \quad h(x, , y^k) = 0, \\ &\qquad = \quad \sup_{\substack{\mu \geq 0 \\ \lambda}} \; \inf_{\boldsymbol{x}} \; \left\{ f(x, y^k) + \lambda^T h(x, y^k) + \mu^T g(x, y^k) \right\} \qquad \forall \, y^k \in Y \cap V, \end{aligned}$$

where, λ and μ are the Lagrange multipliers corresponding to the equality and inequality constraints of the primal problem (2). Define

$$v(y) \; = \; \sup_{\substack{\mu \geq 0 \\ \lambda}} \; \inf_x \; \left\{ f(x, y) + \lambda^T h(x, y) + \mu^T g(x, y) \right\} \quad \forall \, y \in Y \cap V.$$

From the definition of supremum, the maximization over λ and μ can be relaxed to :

$$v(y) \;\geq\; \inf_x \;\Big\{\; f(x,y) + \lambda^T h(x,y) + \mu^T g(x,y) \;\Big\} \quad \forall\; \mu \geq 0,\; \lambda.$$

Assuming that there exists a feasible solution to the inner minimization problem (the analysis for infeasible primal problems is presented later on in section 4.2), the dual representation of $v(y)$ leads to the following formulation, equivalent to (3) :

$$\min_y \;\; v(y), \tag{4a}$$

$$s.t. \quad v(y) \;\geq\; \min_x \{\; f(x,y) + \lambda^T h(x,y) + \mu^T g(x,y) \;\}, \quad \forall \mu \geq 0, \lambda \tag{4b}$$

$$y \;\in\; Y \cap V, \tag{4c}$$

$$V \;\equiv\; \{y : h(x,y) = 0, g(x,y) \leq 0 \; for\; some\; x \;\}\,. \tag{4d}$$

By dropping the last two constraints from (4), the **Relaxed Dual** is obtained:

$$\min_{\substack{y \in Y \\ \mu_B}} \;\; \mu_B, \tag{5a}$$

$$s.t. \quad \mu_B \;\geq\; \min_x \{f(x,y) + \lambda^T h(x,y) + \mu^T g(x,y)\}, \quad \forall \mu \geq 0 \;,\; \lambda \;, \tag{5b}$$

where μ_B is a scalar.

The inner minimization problem, denoted as **inner relaxed dual**, is

$$\min_x \quad L(x,y,\lambda^k,\mu^k), \tag{6}$$

$$L(x,y,\lambda^k,\mu^k) = f(x,y) + {\lambda^k}^T h(x,y) + {\mu^k}^T g(x,y), \tag{7}$$

and involves minimzing the Lagrange function $L(x,y,\lambda^k,\mu^k)$ formulated from (2) at the kth iteration.

Remark 3.1. (a) The primal problem (2) represents an upper bound on the original problem (1). The relaxed dual problem (5), on the other hand, contains fewer constraints than (4) and hence provides a valid lower bound for (1).

(b) In the form given by (5) the relaxed dual problem can be very difficult to solve, since it contains the inner relaxed dual problem, which is parametric in y.

4. Mathematical Properties

The mathematical properties will be presented first assuming feasible primal problems. The properties for infeasible primal problems will be considered in Section 4.2.

4.1. Feasible Primal Problems

At iteration k, define x^k to be the solution of the x variables for the kth primal problem (which is solved for $y = y^k$). Also define μ^k and λ^k to be the corresponding optimal Lagrange multipliers for inequality and equality constraints respectively. Then, the following Lemma can be stated.

Lemma 4.1.

(a) The solution of each primal problem $(\mathbf{P})$ is the global solution of that problem.

(b) The Lagrange function of the k^{th} iteration, $L(x, y, \lambda^k, \mu^k)$, is convex in x for every fixed y and convex in y for every fixed x.

(c) The solution of the **inner relaxed dual** problem is its global solution for each fixed $y = y^d$.

(d) $\min_x L(x, y^d, \lambda^k, \mu^k) \geq \min_x L(x, y^d, \lambda^k, \mu^k)|^{lin}_{x^k} \qquad \forall y = y^d,$
where $L(x, y^d, \lambda^k, \mu^k)|^{lin}_{x^k}$ is the linearization of the Lagrange function $L(x, y^d, \lambda^k, \mu^k)$ at x^k, the solution of the k^{th} primal subproblem.

Proof. The proof of these statements follows from the application of *Conditions (A)* of Section 2 and the definition of the Lagrange function. □

Definition 4.1. Define the vector $g^k(y)$ (consisting of elements $g_i^k(y)$) as follows :

$$g^k(y) \;=\; \nabla_x L(x, y, \lambda^k, \mu^k)|_{x^k} \;, \quad \text{and} \qquad g_i^k(y) \;=\; \nabla_{x_i} L(x, y, \lambda^k, \mu^k)|_{x^k},$$

where x_i is the ith x variable, $i = 1, 2, ..., n$. Then, it can be seen that for every fixed y, the linearization of the Lagrange function at x^k is given by

$$L(x, y, \lambda^k, \mu^k)|^{lin}_{x^k} \;=\; L(x^k, y, \lambda^k, \mu^k) + g^k(y) \cdot (x - x^k) \;=\; L(x^k, y, \lambda^k, \mu^k) + \sum_{i=1}^{n} g_i^k(y) \cdot (x_i - x_i^k).$$

Based upon this form of the Lagrange function, the following definition is made:

Definition 4.2. At iteration k, define every variable x_i for which $g_i^k(y)$ is a function of y to be a *connected* variable. Also define I_c^k to be the set of all such *connected* variables.

Property 4.1. The optimal solution of the inner relaxed dual **(IRD)** problem, with the Lagrange function replaced by its linearization at x^k, depends only on those x_i, for which $g_i^k(y)$ is a function of y (i.e. the *connected* x variables.)

Proof. The linearization of the Lagrange function can be written as

$$L(x, y, \lambda^k, \mu^k)|_{x^k}^{lin} = L(x^k, y, \lambda^k, \mu^k) + \sum_{\substack{i=1 \\ i \neq j}}^{n} g_i^k(y) \cdot (x_i - x_i^k) + g_j^k(y)(x_j - x_j^k). \tag{8}$$

From the **KKT** gradient conditions for the kth primal problem,

$$\nabla_{x_i} L(x^k, y^k, \lambda^k, \mu^k) = g_i^k(y^k) = 0 \quad \forall\, i = 1, 2, ..., n\,. \tag{9}$$

Using (9), the **inner relaxed dual** can be written as

$$\begin{aligned} L^*(\overline{x}, y, \lambda^k, \mu^k) \;&=\; \min_x L(x, y, \lambda^k, \mu^k) \geq \min_x L(x, y, \lambda^k, \mu^k)|_{x^k}^{lin} \\ &\geq\; \min_x \{\; L(x^k, y, \lambda^k, \mu^k) + \sum_{\substack{i=1 \\ i \neq j}}^{n} g_i^k(y) \cdot (x_i - x_i^k) + (g_j^k(y) - g_j^k(y^k))(x_j - x_j^k)\;.\; \} \end{aligned}$$

Now, suppose that $g_j^k(y)$ is not a function of y. Then, $g_j^k(y) = g_j^k(y^k)$. Therefore,

$$L^*(\overline{x}, y, \lambda^k, \mu^k) \geq L(x^k, y, \lambda^k, \mu^k) + \min_x\{\; \sum_{\substack{i=1 \\ i \neq j}}^{n} g_i^k(y) \cdot (x_i - x_i^k).\; \}$$

Hence, the linearization of the Lagrange function does not depend on x_j, and the minimization of the Lagrange function in its linearized form with respect to x_j will not have any effect on the solution of the Inner Relaxed Dual problem. □

Remark 4.1. This property is important from the computational point of view. It implies that the inner relaxed dual could be replaced by a problem involving the minimization of the linearization of the Lagrange function over the set of *connected* x variables. This can help reduce the computational requirements by several orders of magnitude.

Property 4.2. Suppose that the optimal solution of the **inner relaxed dual** occurs at $\overline{x}$; that is, for every $y \in Y$,

$$L^*(\overline{x}, y, \lambda^k, \mu^k) = \min_x \; L(x, y, \lambda^k, \mu^k).$$

Then, for every k,

$$L^*(\overline{x}, y, \lambda^k, \mu^k) \geq \min_{B_j \in CB} \left\{ \begin{array}{cc} & L(x^{B_j}, y, \lambda^k, \mu^k)|_{x^k}^{lin}\ , \\ with & \nabla_{x_i} L(x, y, \lambda^k, \mu^k)|_{x^k} \geq 0 \quad \forall x_i^{B_j} = x_i^L\ , \\ & \nabla_{x_i} L(x, y, \lambda^k, \mu^k)|_{x^k} \leq 0 \quad \forall x_i^{B_j} = x_i^U \end{array} \right\} \ \forall\, y. \tag{10}$$

where x_i^L and x_i^U are the lower and upper bounds on the *connected* x variables respectively, B_j indicates a combination of lower/upper bounds of these variables, x^{B_j} is the vector of lower/upper bounds corresponding to the bound combination B_j, and CB is the set of all bound combinations.

Proof. By its definition, $\overline{x}$ must satisfy the following inequality:

$$L^*(\overline{x}, y, \lambda^k, \mu^k) \geq \min_{x}\ L(x, y, \lambda^k, \mu^k)|_{x^k}^{lin} \quad \forall\, y. \tag{11}$$

Using the definition of $L(x, y, \lambda^k, \mu^k)|_{x^k}^{lin}$, the right hand side of (11) is given by

$$\begin{aligned} \min_{x}\ L(x, y, \lambda^k, \mu^k)|_{x^k}^{lin} &= \min_{x}[L(x^k, y, \lambda^k, \mu^k) + \sum_{i \in I_c^k} \nabla_{x_i} L(x, y, \lambda^k, \mu^k)|_{x^k}.(x_i - x_i^k)] \\ &= L(x^k, y, \lambda^k, \mu^k) + \min_{x} \sum_{i \in I_c^k} \nabla_{x_i} L(x, y, \lambda^k, \mu^k)|_{x^k}.(x_i - x_i^k). \end{aligned}$$

For any fixed $y = y^d$, the operators for minimization and summation can be exchanged. Hence,

$$\min_{x}\ L(x, y, \lambda^k, \mu^k)|_{x^k}^{lin} = L(x^k, y, \lambda^k, \mu^k) + \sum_{i \in I_c^k} \min_{x_i} \nabla_{x_i} L(x, y, \lambda^k, \mu^k)|_{x^k}.(x_i - x_i^k). \tag{12}$$

Consider the *i* th component of the second term on the right hand side. It is linear in x_i. Hence, the minimum of this term will lie at a bound of x_i, the specific nature of the bound(lower or upper) being determined by the sign of $\nabla_{x_i} L(x, y, \lambda^k, \mu^k)|_{x^k}$. Two cases are possible:

(a) If $\nabla_{x_i} L(x, y, \lambda^k, \mu^k)|_{x^k} \geq 0$,

$$\min_{x_i} \nabla_{x_i}\ L(x, y, \lambda^k, \mu^k)|_{x^k}.(x_i - x_i^k) \geq \nabla_{x_i} L(x, y, \lambda^k, \mu^k)|_{x^k}.(x_i^L - x_i^k)\ .$$

(b) If $\nabla_{x_i} L(x, y, \lambda^k, \mu^k)|_{x^k} \leq 0$, then

$$\min_{x_i} \nabla_{x_i}\ L(x, y, \lambda^k, \mu^k)|_{x^k}.(x_i - x_i^k) \geq \nabla_{x_i} L(x, y, \lambda^k, \mu^k)|_{x^k}.(x_i^U - x_i^k)\ .$$

These two cases can be combined to yield the following result:

$$\begin{aligned}
&\min_{x_i} \nabla_{x_i} L(x,y,\lambda^k,\mu^k)|_{x^k}.(x_i - x_i^k) \geq \nabla_{x_i} L(x,y,\lambda^k,\mu^k)|_{x^k}.(x_i^{B_j} - x_i^k),\\
&where \\
&\qquad x_i^{B_j} = \begin{cases} x_i^L & \forall\, y:\ \nabla_{x_i} L(x,y,\lambda^k,\mu^k)|_{x^k} \geq 0 \\ x_i^U & \forall\, y:\ \nabla_{x_i} L(x,y,\lambda^k,\mu^k)|_{x^k} \leq 0 \end{cases}
\end{aligned} \tag{13}$$

Combining (11), (12) and (13), it can be seen that

$$L^*(\overline{x},y,\lambda^k,\mu^k) \;\geq\; L(x^k,y,\lambda^k,\mu^k) + \sum_{i\in I_c^k} \nabla_{x_i} L(x,y,\lambda^k,\mu^k)|_{x^k}.(x_i^{B_j} - x_i^k), \tag{14a}$$

$$where$$

$$x_i^{B_j} = x_i^L \quad \forall y:\ \nabla_{x_i} L(x,y,\lambda^k,\mu^k)|_{x^k} \geq 0 \tag{14b}$$

$$x_i^{B_j} = x_i^U \quad \forall y:\ \nabla_{x_i} L(x,y,\lambda^k,\mu^k)|_{x^k} \leq 0 \tag{14c}$$

From this, it is evident that for any fixed $y = y^d$, there exists a combination of bounds B_j for the *connected* x variables such that

$$\begin{aligned}
\min_x L(x,y^d,\lambda^k,\mu^k) \;&\geq\; L(x^k,y^d,\lambda^k,\mu^k) + \sum_{i\in I_c^k} \nabla_{x_i} L(x,y^d,\lambda^k,\mu^k)|_{x^k}.(x_i^{B_j} - x_i^k)\\
&\geq\; L(x^{B_j},y^d,\lambda^k,\mu^k)|_{x^k}^{lin}.
\end{aligned}$$

Hence, for every discretized $y = y^d$, by fixing the values of the x variables at a combination of bounds B_j in the linearized Lagrange function and taking the minimum over all possible combinations of bounds $B_j \in CB$, a lower bound on the value of $L^*(\overline{x},y^d,\lambda^k,\mu^k)$ is obtained. Since this is true for every $y = y^d$, (10) must hold for all y. □

Definition 4.3. The constraints requiring the positivity or negativity of the gradients of a particular Lagrange function w.r.t x_i are called the *qualifying constraints* of that Lagrange function.

Property 4.3. If $g_i^k(y)$ are linear in y $\forall i$, then the qualifying constraints form a linear set of constraints in y.

Proof. It follows from the definition of $g^k(y)$. □

Property 4.4. At the Kth iteration,

(i) Define $(\mu_B^{min})^K$ to be the optimal value of the Relaxed Dual Problem. That is,

$$(\mu_B^{min})^K = \left\{ \begin{array}{lll} & \min\limits_{\substack{y\in Y \\ \mu_B}} \ \mu_B & \\ s.t. & \mu_B \geq \min\limits_{x} L(x, y, \lambda^k, \mu^k) & k = 1, 2....(K-1) \\ & \mu_B \geq \min\limits_{x} L(x, y, \lambda^K, \mu^K) & \end{array} \right\}.$$

(ii) Define $UL(k, K)$ to be the Lagrange function from the kth iteration ($k < K$) whose qualifying constraints are satisfied at y^K, the fixed value of the y variables for the Kth primal problem, and let x^{B_j} be the corresponding combination of bounds of the x variables for this Lagrange function. Note that x^{B_j} can be different for different iterations $k = 1, 2, ..., K-1$.

(iii) Define the following subproblem (15) :

$$\mu_B^{stor}(K, B_l) = \left\{ \begin{array}{l} \min\limits_{\substack{y\in Y \\ \mu_B}} \ \mu_B \\ \\ s.t. \\ \left. \begin{array}{ll} \mu_B \geq & L(x^{B_j}, y, \lambda^k, \mu^k)\big|_{x^k}^{lin} \\ & \nabla_{x_i} L(x, y, \lambda^k, \mu^k)\big|_{x^k} \leq 0 \quad if\ x_i^{B_j} = x_i^U \\ & \nabla_{x_i} L(x, y, \lambda^k, \mu^k)\big|_{x^k} \geq 0 \quad if\ x_i^{B_j} = x_i^L \end{array} \right\} \begin{array}{l} \forall j \in UL(k, K) \\ k = 1, 2....K-1 \end{array} \\ \\ \begin{array}{ll} \mu_B \geq & L(x^{B_l}, y, \lambda^K, \mu^K)\big|_{x^K}^{lin} \\ & \nabla_{x_i} L(x, y, \lambda^K, \mu^K)\big|_{x^K} \leq 0 \quad if\ x_i^{B_l} = x_i^U \\ & \nabla_{x_i} L(x, y, \lambda^K, \mu^K)\big|_{x^K} \geq 0 \quad if\ x_i^{B_l} = x_i^L \end{array} \end{array} \right\} \quad (15)$$

where $\mu_B^{stor}(K, B_l)$ is the stored solution of the above subproblem solved at iteration K with the x variables set to the combination of bounds B_l in the Lagrange function.

(iv) Define ${\mu'}_B^K = \min\limits_{B_l \in CB} \mu_B^{stor}(K, B_l)$ to be the minimum of the stored solutions of all the subproblems of the form (15) solved at the Kth iteration.

Then,

$$(\mu_B^{min})^K \geq MIN[{\mu'}_B^K, \min_{\substack{k=1,2,...,K-1 \\ B_l \in CB}} \mu_B^{stor}(k, B_l)] , \quad (16)$$

where $\mu_{\boldsymbol{B}}^{\boldsymbol{stor}}(k, B_l)$ are the stored solutions from previous iterations ($k < K$).

Proof.

(a) For iteration 1: For $k = 1$, from Property 4.2.,

$$\min_{\boldsymbol{x}} L(x, y, \lambda^1, \mu^1) \geq \min_{B_l \in CB} \begin{Bmatrix} L(x^{\boldsymbol{B}_l}, y, \lambda^1, \mu^1) \\ \nabla_{\boldsymbol{x}_i} L(x, y, \lambda^1, \mu^1)|_{\boldsymbol{x}^1} \leq 0 \quad if\ x_{\boldsymbol{i}}^{\boldsymbol{B}_l} = x_{\boldsymbol{i}}^{\boldsymbol{U}} \\ \nabla_{\boldsymbol{x}_i} L(x, y, \lambda^1, \mu^1)|_{\boldsymbol{x}^1} \geq 0 \quad if\ x_{\boldsymbol{i}}^{\boldsymbol{B}_l} = x_{\boldsymbol{i}}^{\boldsymbol{L}} \end{Bmatrix}, \quad \forall y.$$

Since this holds for all y,

$$\min_{y \in Y} \left\{ \min_{\boldsymbol{x}} L(x, y, \lambda^1, \mu^1) \right\} \geq \min_{y \in Y} \left\{ \min_{B_l \in CB} \begin{Bmatrix} L(x^{\boldsymbol{B}_l}, y, \lambda^1, \mu^1) \\ \nabla_{\boldsymbol{x}_i} L(x, y, \lambda^1, \mu^1)|_{\boldsymbol{x}^1} \leq 0 \quad if\ x_{\boldsymbol{i}}^{\boldsymbol{B}_l} = x_{\boldsymbol{i}}^{\boldsymbol{U}} \\ \nabla_{\boldsymbol{x}_i} L(x, y, \lambda^1, \mu^1)|_{\boldsymbol{x}^1} \geq 0 \quad if\ x_{\boldsymbol{i}}^{\boldsymbol{B}_l} = x_{\boldsymbol{i}}^{\boldsymbol{L}} \end{Bmatrix} \right\}.$$

The operators on the right hand side can be interchanged since $L(x^{\boldsymbol{B}_l}, y, \lambda^1, \mu^1)$ depends only on y. Therefore,

$$\min_{y \in Y} \left\{ \min_{x} L(x, y, \lambda^1, \mu^1) \right\} \geq \min_{B_l \in CB} \left\{ \min_{y \in Y} \begin{Bmatrix} L(x^{\boldsymbol{B}_l}, y, \lambda^1, \mu^1) \\ \nabla_{\boldsymbol{x}_i} L(x, y, \lambda^1, \mu^1)|_{\boldsymbol{x}^1} \leq 0 \quad if\ x_{\boldsymbol{i}}^{\boldsymbol{B}_l} = x_{\boldsymbol{i}}^{\boldsymbol{U}} \\ \nabla_{\boldsymbol{x}_i} L(x, y, \lambda^1, \mu^1)|_{\boldsymbol{x}^1} \geq 0 \quad if\ x_{\boldsymbol{i}}^{\boldsymbol{B}_l} = x_{\boldsymbol{i}}^{\boldsymbol{L}} \end{Bmatrix} \right\}.$$

Equivalently, this can be written as

$$(\mu_{\boldsymbol{B}}^{\boldsymbol{min}})^1 = \begin{Bmatrix} \min\limits_{\substack{y \in Y \\ \mu_B}} \ \mu_{\boldsymbol{B}} \\ s.t. \quad \mu_{\boldsymbol{B}} \geq \min\limits_{x} L(x, y, \lambda^1, \mu^1) \end{Bmatrix} \geq \min_{B_l \in CB} \begin{Bmatrix} \min\limits_{y \in Y, \mu_B} \mu_{\boldsymbol{B}} \\ s.t. \quad \mu_{\boldsymbol{B}} \geq L(x^{\boldsymbol{B}_l}, y, \lambda^1, \mu^1) \\ \nabla_{\boldsymbol{x}_i} L(x, y, \lambda^1, \mu^1)|_{\boldsymbol{x}^1} \leq 0 \quad if\ x_{\boldsymbol{i}}^{\boldsymbol{B}_l} = x_{\boldsymbol{i}}^{\boldsymbol{U}} \\ \nabla_{\boldsymbol{x}_i} L(x, y, \lambda^1, \mu^1)|_{\boldsymbol{x}^1} \geq 0 \quad if\ x_{\boldsymbol{i}}^{\boldsymbol{B}_l} = x_{\boldsymbol{i}}^{\boldsymbol{L}} \end{Bmatrix}.$$

At this stage, there are no stored solutions from previous iterations. Hence, for iteration 1, the property is proved.

(b) For iteration 2: We have

$$(\mu_{\boldsymbol{B}}^{\boldsymbol{min}})^2 = \begin{Bmatrix} & \min\limits_{\substack{y \in Y \\ \mu_B}} \ \mu_{\boldsymbol{B}} \\ s.t. & \mu_{\boldsymbol{B}} \geq \min\limits_{\boldsymbol{x}} L(x, y, \lambda^1, \mu^1) \\ & \mu_{\boldsymbol{B}} \geq \min\limits_{\boldsymbol{x}} L(x, y, \lambda^2, \mu^2) \end{Bmatrix}.$$

From Property 4.2, we have that

$$(\mu_B^{min})^2 \geq MIN \left\{ \min_{\substack{B_p \in CB \\ B_p \neq B_j}} \left\{ \begin{array}{l} \qquad\qquad \min_{y \in Y, \mu_B} \mu_B \\ s.t. \quad \mu_B \geq L(x^{B_p}, y, \lambda^1, \mu^1) \\ \qquad\qquad \nabla_{x_i} L(x, y, \lambda^1, \mu^1)|_{x^1} \leq 0 \quad if\ x_i^{B_p} = x_i^U \\ \qquad\qquad \nabla_{x_i} L(x, y, \lambda^1, \mu^1)|_{x^1} \geq 0 \quad if\ x_i^{B_p} = x_i^L \\ \qquad \mu_B \geq \min_x L(x, y, \lambda^2, \mu^2) \end{array} \right\} , \quad {\mu_B''}^2 \right\} , \tag{17}$$

where B_p represents a combination of bounds of the x variables, B_j is the specific combination of bounds corresponding to the Lagrange function from the first iteration whose *qualifying* constraints are satisfied at $y = y^2$, and

$${\mu_B''}^2 = \left\{ \begin{array}{l} \qquad\qquad \min_{y \in Y, \mu_B} \mu_B \\ s.t. \quad \mu_B \geq L(x^{B_j}, y, \lambda^1, \mu^1) \\ \qquad\qquad \nabla_{x_i} L(x, y, \lambda^1, \mu^1)|_{x^1} \leq 0 \quad if\ x_i^{B_j} = x_i^U \\ \qquad\qquad \nabla_{x_i} L(x, y, \lambda^1, \mu^1)|_{x^1} \geq 0 \quad if\ x_i^{B_j} = x_i^L \\ \qquad \mu_B \geq \min_x L(x, y, \lambda^2, \mu^2) \end{array} \right\} . \tag{18}$$

Due to the presence of an additional set of constraints, it is obvious that

$$\min_{\substack{B_p \in CB \\ B_p \neq B_j}} \left\{ \begin{array}{l} \qquad\qquad \min_{y \in Y, \mu_B} \mu_B \\ s.t. \quad \mu_B \geq L(x^{B_p}, y, \lambda^1, \mu^1) \\ \qquad\qquad \nabla_{x_i} L(x, y, \lambda^1, \mu^1)|_{x^1} \leq 0 \quad if\ x_i^{B_p} = x_i^U \\ \qquad\qquad \nabla_{x_i} L(x, y, \lambda^1, \mu^1)|_{x^1} \geq 0 \quad if\ x_i^{B_p} = x_i^L \\ \qquad \mu_B \geq \min_x L(x, y, \lambda^2, \mu^2) \end{array} \right\} \geq \min_{\substack{B_p \in CB \\ B_p \neq B_j}} \mu_B^{stor}(1, B_p) .$$

Therefore,

$$(\mu_B^{min})^2 \geq MIN[\min_{\substack{B_p \in CB \\ B_p \neq B_j}} \mu_B^{stor}(1, B_p), \ {\mu_B''}^2] .$$

Hence, it only remains to be shown that $\mu''^{\,2}_B \geq \mu'^{\,2}_B$, where $\mu'^{\,2}_B$ is given as

$$\mu'^{\,2}_B = \min_{B_l \in CB} \left\{ \begin{array}{l} \min\limits_{\substack{y \in Y \\ \mu_B}} \quad \mu_B \\ \\ s.t. \\ \quad \mu_B \geq \quad L(x^{B_j}, y, \lambda^1, \mu^1)|^{lin}_{x^1} \\ \qquad\qquad \nabla_{x_i} L(x, y, \lambda^1, \mu^1)|_{x^1} \leq 0 \quad if\ x_i^{B_j} = x_i^U \\ \qquad\qquad \nabla_{x_i} L(x, y, \lambda^1, \mu^1)|_{x^1} \geq 0 \quad if\ x_i^{B_j} = x_i^L \\ \quad \mu_B \geq \quad L(x^{B_l}, y, \lambda^2, \mu^2)|^{lin}_{x^2} \\ \qquad\qquad \nabla_{x_i} L(x, y, \lambda^2, \mu^2)|_{x^2} \leq 0 \quad if\ x_i^{B_l} = x_i^U \\ \qquad\qquad \nabla_{x_i} L(x, y, \lambda^2, \mu^2)|_{x^2} \geq 0 \quad if\ x_i^{B_l} = x_i^L \end{array} \right\}. \tag{19}$$

Now, in the RHS of (19), the Lagrange function from the first iteration has the value of x set to the appropriate bound (x^{B_j}), and is therefore a function only of y. Hence, the $\min\limits_{B_l \in CB}$ operator applies only to the second set of constraints, i.e., those corresponding to the 2^{nd} primal problem. Hence, (19) is equivalent to

$$\mu_B{}'^2 \geq \left\{ \begin{array}{l} \min\limits_{y, \mu_B} \quad \mu_B \\ \\ s.t. \\ \quad \mu_B \geq \quad L(x^{B_j}, y, \lambda^1, \mu^1)|^{lin}_{x^1} \\ \qquad\qquad \nabla_{x_i} L(x, y, \lambda^1, \mu^1)|_{x^1} \leq 0 \quad if\ x_i^{B_j} = x_i^U \\ \qquad\qquad \nabla_{x_i} L(x, y, \lambda^1, \mu^1)|_{x^1} \geq 0 \quad if\ x_i^{B_j} = x_i^L \\ \quad \mu_B \geq \min\limits_{B_l \in CB} \left\{ \begin{array}{l} L(x^{B_l}, y, \lambda^2, \mu^2)|^{lin}_{x^2} \\ \nabla_{x_i} L(x, y, \lambda^2, \mu^2)|_{x^2} \leq 0 \quad if\ x_i^{B_l} = x_i^U \\ \nabla_{x_i} L(x, y, \lambda^2, \mu^2)|_{x^2} \geq 0 \quad if\ x_i^{B_l} = x_i^L \end{array} \right. \end{array} \right\}. \tag{20}$$

The use of Property 4.2 leads to

$$\min_x L(x, y, \lambda^2, \mu^2) \ \geq \ \min_{B_l \in CB} \left\{ \begin{array}{l} L(x^{B_l}, y, \lambda^2, \mu^2)|^{lin}_{x^k} \\ \nabla_{x_i} L(x, y, \lambda^2, \mu^2)|_{x^2} \leq 0 \quad if\ x_i^{B_l} = x_i^U \\ \nabla_{x_i} L(x, y, \lambda^2, \mu^2)|_{x^2} \geq 0 \quad if\ x_i^{B_l} = x_i^L \end{array} \right\}. \tag{21}$$

Hence, from (20) and (21), the second set of constraints on the RHS of (19) is simply a relaxed form of the second set of constraints of the RHS of (18). Hence, for $k = 2$, the property holds. Similarly, by induction, the property can be proved for any k. □

Remark 4.2. Notice that the RHS of (16) represents (a) the solution of a number of subproblems each of which corresponds to a unique combination of bounds B_l of the *connected* x variables, and (b) the selection of the minimum solution from all these subproblems and the stored solutions of previous subproblems that have not already been selected. The solution of the different subproblems correspond to partitioning the y space and solving the problem in each subspace.

Property 4.5. The solution of each subproblem in the form given by (15) is its global solution.

Proof. The Lagrange functions as used in the (15) are convex functions of y. From Property 4.3, the gradients of the Lagrange functions w.r.t x_i are linear in y. Therefore, subproblem (15) satisfies the global optimality conditions (Ref. 32). □

4.2. Infeasible Primal Problems

In cases where the primal problem (2) is infeasible, another problem must be solved for generating the appropriate values of λ and μ. One possible formulation for this problem is

$$\min_{\substack{\beta^+,\beta^-,\alpha \geq 0 \\ x}} \sum_{i=1}^{m} \alpha_i + \sum_{i=1}^{p} (\beta_i^+ + \beta_i^-), \tag{22a}$$

$$h(x,y) + \beta^+ - \beta^- = 0, \tag{22b}$$

$$g(x,y) - \alpha \leq 0, \tag{22c}$$

where α_i , β_i^+ and β_i^- are slack variables that are introduced in order to minimize the sum of the infeasibilities in the constraints. In problem (22), for every fixed y^k, the objective function is linear, the equality constraints are linear and the inequality constraints are convex. If $\theta = \sum_{i=1}^{m} \alpha_i + \sum_{i=1}^{p} (\beta_i^+ + \beta_i^-)$, then, the strong duality theorem provides

$$\begin{aligned} &\min\ \theta \quad s.t. \quad g(x,y) - \alpha \leq 0, \quad h(x,y) + \beta^+ - \beta^- = 0, \\ &= \max_{\substack{\mu_1 \geq 0 \\ \lambda}} \min_{x} \left\{ \theta + \lambda_1^T (h(x,y) + \beta^+ - \beta^-) + \mu_1^T (g(x,y) - \alpha) \right\}, \end{aligned} \tag{23}$$

where λ_1 and μ_1 are the Lagrange multipliers for the equality and inequality constraints for the solution of (22) for fixed $y = y^k$. If $\overline{\theta}$ is the optimal solution of (22), then (23), together with the optimality conditions for this problem, implies that

$$\overline{\theta} = \max_{\substack{\lambda_1 \\ \mu_1 \geq 0}} \min_{x} \{\lambda_1^T h(x,y) + \mu_1^T g(x,y)\}\ .$$

Since we seek to minimize the infeasibilities $\overline{\theta}$, this can be used as a constraint for the relaxed dual problem in the following form :

$$\max_{\substack{\lambda_1 \\ \mu_1 \geq 0}} \min_x \{\lambda_1^T h(x,y) + \mu_1^T g(x,y)\} = 0.$$

A relaxed form of this constraint is

$$\min_x L'(x, y, \lambda_1, \mu_1) \leq 0, \tag{24}$$

where

$$L'(x, y, \lambda_1, \mu_1) = \lambda_1^T h(x,y) + \mu_1^T g(x,y). \tag{25}$$

It can be easily shown that the Properties 4.1-4.5 presented in Section 4.1 are all applicable directly for the case of infeasible primal problems by simply replacing μ_B by 0 and $L(x, y, \lambda^k, \mu^k)$ by $L'(x, y, \lambda_1^k, \mu_1^k)$. The constraints to be added along with the Lagrange function to the Relaxed Dual problem are again of the form

$$\nabla_{x_i} L'(x, y, \lambda_1^k, \mu_1^k)|_{x^k} \leq 0 \ , \qquad or \qquad \nabla_{x_i} L'(x, y, \lambda_1^k, \mu_1^k)|_{x^k} \geq 0 \ ,$$

depending on whether the variable x_i is at its upper or its lower bound respectively.

Remark 4.3. Constraint (24) does not contain μ_B (the objective function for the lower bound problem); at first glance, it does not appear to be a useful cut for the *relaxed dual* problem. However, it can be seen that this constraint is always violated at $y = y^k$ (the fixed value of y for the iteration k at which the relaxed primal problem was solved, leading to (24)). The introduction of (24) ensures that no cycling occurs in the *relaxed dual* problem, i.e. this eliminates the possibility of the *relaxed dual* problem returning the value y^k at any subsequent iteration. Hence, the constraints of the form (24) are useful as *feasibility cuts* for the *relaxed dual* problems.

Remark 4.4. It should be noted that if the implicitly defined set V can be introduced into the dual problem, then solutions for y found by the relaxed dual problem will always be feasible for the primal problems. This is possible for unconstrained problems or for problems where it is possible to incorporate the constraint set explicitly into the relaxed dual problems without destroying the convexity of the problem (for example, general quadratic problems with linear constraints). Hence, for these cases, there will be no infeasible primal problems.

Remark 4.5. The Lagrange functions being introduced into the relaxed dual problems for iterations when the primal is infeasible are of the form given by (24). However, this is not the only possible formulation of the constraint that can be used in the relaxed dual problem. For example, at any iteration, $\mu_B \geq M^{LBD}$, where M^{LBD} is the lower bound obtained from the relaxed dual problems upto that iteration. Using this and (24), another constraint that can be added to the relaxed dual problem is

$$\mu_B \geq M^{LBD} + \min_x L'(x, y, \lambda_1, \mu_1).$$

5. Partition of the y-Space in the Relaxed Dual

Section 4 presented properties, based on which the relaxed dual problem can be reduced to a formulation containing only constraints and not inner optimization problems. However, a number of subproblems corresponding to all combinations of bounds $B_l \in CB$ need to be solved at every iteration. The solutions of the primal problems are used to formulate Lagrange functions that are used in the relaxed dual problems. At any given iteration, therefore, the relaxed dual subproblems will contain a Lagrange function from the current iteration, and one from each of the previous iterations. The criterion of selection of the Lagrange functions from the previous iterations is very important, since it defines the region in which a particular relaxed dual subproblem is solved.

At every iteration K, the primal problem (2) is solved for a fixed value of $y = y^k$. If the primal problem is infeasible, then a relaxed primal problem of the form (22) is solved. In either case, the resulting Lagrange multipliers for the various constraints are stored.

5.1. Selection of Previous Lagrange Functions

Before solving the relaxed dual subproblems, the Lagrangians from all the previous iterations that can be used as constraints for the current relaxed dual subproblems are determined. To achieve this, the *qualifying* constraints of every such Lagrange function (i.e. from iterations $1, 2, \ldots, K-1$) are evaluated at y^K. If the qualifying constraints are satisfied at y^K, then the Lagrange function and its accompanying *qualifying* constraints are selected to be constraints for the current relaxed dual subproblems. This is done even if the primal problem was infeasible for the iteration in question. Exactly one Lagrange function will be selected from each of the previous iterations. In these Lagrange functions, the x variables are then set to the appropriate combination of bounds. Hence, these Lagrange functions, as included in the current relaxed dual subproblems, are functions only of y.

5.2. Partition in the y-space

Once the Lagrange functions from the previous iterations have been selected, the relaxed dual problem is then solved for each combination of x^{B_l} . In each case, the Lagrange function formulated from the *current* primal problem is chosen as a constraint for the relaxed dual problem, with x replaced by x^{B_l}. In addition to this, the corresponding *qualifying* constraints for the Lagrange function are added to the relaxed dual problem.

It should be noted that for each combination of the x variables, there is a corresponding set of *qualifying* constraints in terms of the y variables. Since the relaxed dual subproblems are solved for different combinations of bounds, each *relaxed dual* subproblem has a unique set of *qualifying* constraints associated with it. Therefore, each of these relaxed dual subproblems solved corresponds to a different region in the y-space of the original problem. The actual region for which a particular *relaxed dual* subproblem is solved is determined by the corresponding set of *qualifying* constraints that are associated with that Lagrange function. Therefore, the solution of the *relaxed dual* subproblems can be viewed as a branch-and-bound procedure where the feasible region in terms of the y variables is partitioned into $2^{NI_c^K}$ subregions (where NI_c^K is the number of *connected* x variables at the K^{th} iteraton) and each *relaxed dual* subproblem corresponds to a node of the branch-and-bound tree and is solved for a particular subregion in the y-space. However, the nodes in the tree are generated dynamically through the solution of the various subproblems at every iteration.

After the relaxed dual problem has been solved for every possible combination of the bounds x_l^B, the only remaining task is to determine a new lower bound for the original problem and select a fixed value of y for the next primal problem. This is done by taking the minimum of all the stored values of μ_B as the lower bound, and the corresponding solutions for y as the y^{K+1} for the next primal. Once a particular μ_B and y have been selected, they are no longer considered for future iterations. This is to ensure that the relaxed dual problem will not return the same value of y and μ_B during successive iterations.

Finally, the check for convergence is done. If the lower bound from the relaxed dual problem comes within ϵ of the upper bound from the primal problems, the optimal solution to the original problem has been found and the algorithm terminates. Otherwise, the algorithm continues with the updated stored set of solutions from the *relaxed dual* problems and a new fixed value of y for the primal problem.

5.3. Progress of the lower bound

When a relaxed dual subproblem is solved, it contains Lagrange functions from previous iterations whose *qualifying* constraints are satisfied at the current fixed value of y. The solution of this subproblem lies in a cone defined by the *qualifying* constraints of the selected Lagrange functions. Therefore, in all future iterations, this set of Lagrange functions (and the accompanying *qualifying* constraints) will be present if the fixed value of y for that iteration lies inside this cone. Hence, there is a subsequence in the stored set of solutions of the relaxed dual subproblems that will always lie inside this cone. Therefore, the solution of all the relaxed dual subproblems can be viewed as the union of subsequences lying inside different cones, and the lower bound from the relaxed dual subproblems at any iteration is simply the minimum of these subsequences.

5.4. Properties of the Relaxed Dual

Since the *relaxed dual* problem is based upon the partition of the space of y variables, the position of the current fixed value for y (that is, the value of y^K) plays a very important role in the solution of the subproblems. If, for some iteration K, y^K lies at a boundary of the constraint region, then some of the relaxed dual problems will give redundant solutions, and it is possible to avoid solving these of the subproblems without losing the rigorous nature of the algorithm. To identify such cases, use is made of the structure of the Lagrange function that is used in the *relaxed dual* problem. Based upon this structure, new properties have been developed that help to improve the efficiency of the algorithm by reducing the number of *relaxed dual* subproblems that need to be solved at each iteration. These properties along with the computational improvements arising from their application are discussed in detail in Visweswaran and Floudas (Ref. 33).

6. The Global OPtimization (GOP) Algorithm

The **GOP** algorithm can be formally stated in the following steps:

STEP 0 – *Initialization of parameters:*

Define the storage parameters $\mu_B^{stor}(K^{max}, B_j)$, $y^{stor}(K^{max}, B_j)$ and $y^k(K^{max}, B_j)$ over the set of bounds CB and the maximum expected number of iterations K^{max}. Define P^{UBD} and M^{LBD} as the upper and lower bounds obtained from the primal and relaxed dual problems respectively. Set

$$\mu_B^{stor}(K^{max}, B_j) = U,\ P^{UBD} = U,\ \ and\ M^{LBD} = L.$$

where U is a very large *positive* number and L is a very large *negative* number. Select an initial set of values y^1 for the complicating variables. Set the counter K equal to 1, and sets K^{feas} and K^{infeas} to empty sets. Select a convergence tolerance parameter ϵ.

STEP 1 – *Primal problem:*

Store the value of y^K. Solve the primal problem (2) for $y = y^K$. If the primal problem is feasible, update the set K^{feas} to contain K, and store the optimal Lagrange multipliers λ^K and μ^K. Update the upper bound so that

$$P^{UBD} = MIN(P^{UBD}, P^K(y^K))$$

where $P^K(y^K)$ is the solution of the current (K th) primal problem. If the primal problem is infeasible, update the set K^{infeas} to contain K, and solve the relaxed primal subproblem (22) with $y = y^K$. Store the values of the optimal Lagrange multipliers λ_1^K and μ_1^K.

STEP 2 – *Selection of Lagrange functions from the previous iterations:*

For $k = 1, 2, ..., K - 1$, evaluate all the *qualifying* constraints of every Lagrange function from iteration k, (i.e., corresponding to each set of bounds of x) at y^K. Select the one Lagrange function from every iteration each of whose *qualifying* constraints are satisfied at y^K to be in the set $UL(k, K)$, i.e., to be present in the current relaxed dual problems along with its *qualifying* constraints.

STEP 3 – *Relaxed Dual Problem:*

Formulate the Lagrange functions corresponding to the current primal problem. Add this as a constraint to the relaxed dual problem. Then:

(a) Identify the set of *connected* x variables I_c.

(b) Select a combination of the bounds of the *connected* variables in x, say $B_l = B_1$.

(c) Solve the relaxed dual problem (15). Note that this problem shows the constraints that must be used for iterations when the primal problem is feasible. For iterations when the primal problem is infeasible, substitute 0 for μ_B and L' for L in the constraints corresponding to that iteration. See Section 4.2 for more details on the form of these constraints.

Store the solution in $\mu_B^{stor}(K, B_l)$ and $y^{stor}(K, B_l)$.

(d) Select a new combination of bounds for x, say $B_l = B_2$.

(e) Repeat (c) and (d) until the relaxed dual problem has been solved at each set of bounds of the *connected* variables in x, i.e for every $B_l \in CB$.

STEP 4 – *Selecting a new lower bound and* y^{K+1}*:*

From the stored set μ_B^{stor}, select the minimum μ_B^{min} (including the solutions from the current iteration). Also, select the corresponding stored value of $y^{stor}(k, B_j)$ as y^{min}. Set $M^{LBD} = \mu_B^{min}$, and $y^{K+1} = y^{min}$. Delete μ_B^{min} and y^{min} from the stored set.

STEP 5 – *Check for convergence:*

Check if $M^{LBD} \geq P^{UBD} - \epsilon$. IF yes , STOP. Else, set $K = K + 1$ and return to step 1.

Remark 6.1. In Step 2 of the **GOP** algorithm, one Lagrange function from each of the previous iterations is selected on the basis of satisfaction of its *qualifying* constraints at y^K. If it so happens that for some $j \in \{1, 2, ..., n_1\}$,

$$\nabla_{x_j} L(x, y = y^K, \lambda^k, \mu^k)|_{x^k} = 0$$

then, this implies that the Lagrange functions from the kth iteration with x_j set to *either* its upper or lower bound are eligible constraints for the current *relaxed dual* subproblems. Since the accompanying *qualifying* constraints are also included, this means that

$$\nabla_{x_j} L(x, y, \lambda^k, \mu^k)|_{x^k} = 0$$

for the current *relaxed dual* subproblems. This could potentially lead to some regions of y being unavailable for the *relaxed dual* subproblems. This can be avoided by introducing the *qualifying* constraints in a perturbed form :

$$\nabla_{x_j} L(x, y, \lambda^k, \mu^k)|_{x^k} \;\geq\; \delta \quad if\; x_j^{B_j} = x_j^L \tag{26a}$$

$$\nabla_{x_j} L(x, y, \lambda^k, \mu^k)|_{x^k} \;\leq\; -\delta \quad if\; x_j^{B_j} = x_j^U \tag{26b}$$

where δ is a very small positive number. This ensures that both (26a) and (26b) cannot be simultaneously satisfied at y^K, and consequently exactly one Lagrange function (and its accompanying *qualifying* constraints) will be present from each of the previous iterations for the current *relaxed dual* problems.

7. Finite ϵ-Convergence and ϵ-Global Optimality

This section presents the theoretical proof of the convergence and global optimality of the **GOP** algorithm. The convergence proof is based upon the results of Geoffrion (Ref. 28).

Lemma 7.1. If the sequence of solutions $< y^k >$ of the relaxed dual problem converges to $\overline{y}$, and the corresponding Lagrange multipliers for the primal problem (2) converge to $(\overline{\lambda}, \overline{\mu})$, then

$$L(x^{B_j}, \overline{y}, \overline{\lambda}, \overline{\mu})|_{x^k}^{lin} = P(\overline{y}) \quad for\ any\ x^{B_j}$$

where $P(\overline{y})$ represents the optimal solution of the primal problem at $\overline{y}$.
Proof. The linearization of the Lagrange function evaluated at x^{B_j} can be written as

$$L(x^{B_j}, y, \lambda^k, \mu^k)|_{x^k}^{lin} = L(x^k, y, \lambda^k, \mu^k) + [g^k(y) - g^k(y^k)] \cdot (x^{B_j} - x^k) \qquad \forall B_j \in CB$$

By the Strong Duality Theorem, $L(x^k, y^k, \lambda^k, \mu^k)$ is equal to $P(y^k)$ for every y^k. Then,

$$As \quad y^k \rightarrow \overline{y}, \quad L(x^k, y, \lambda^k, \mu^k)|_{y=\overline{y}} \rightarrow P(\overline{y}). \tag{27}$$

Since $g^k(y)$ is a vector of linear functions of y,

$$As \quad y^k \rightarrow \overline{y}, \quad g^k(y) \rightarrow g(\overline{y}).$$

From this, and (27), we obtain the desired result. □

Lemma 7.2. For a fixed $y = y^k$ for the kth primal problem, let $U(y^k)$ be the set of optimal multipliers. If $y^k \rightarrow \overline{y}$ and $\lambda^k \rightarrow \overline{\lambda}, \mu^k \rightarrow \overline{\mu}$, then $(\overline{\lambda}, \overline{\mu}) \in U(\overline{y})$.
Proof. The proof for this Lemma comes from showing that $U(y)$ is an upper semicontinuous mapping at $\overline{y}$. To do this, we employ the characterization of $U(y)$ as the set of optimal solutions of the dual of the primal problem; that is ,

$$U(y) = \{\mu \geq 0, \lambda : \ L^*(\overline{x}, y, \lambda, \mu) = \max_{\mu^k \geq 0, \lambda^k} L^*(\overline{x}, y, \lambda^k, \mu^k)\}$$

Now, $L^*(\overline{x}, y, \lambda^k, \mu^k) = \min_x L(x, y, \lambda^k, \mu^k)$ is a continuous function, since it is a linear sum of continuous functions of y. Then, application of Theorem 1.5 of Meyer (Ref. 34) proves the desired result. □

Lemma 7.3. Let $P(y)$ be the set of optimal solutions of the primal problem for fixed values of y; then, if $y^k \rightarrow \overline{y}$, $P(y)$ is upper semicontinuous at $\overline{y}$.
Proof. Since the set $U(y)$ is assumed to be nonempty for all y, then from the Strong Duality Theorem, we have that

$$P(y) = \max_{\mu \geq 0, \lambda} L^*(\overline{x}, y, \lambda, \mu) \quad \forall y \in Y$$

Since Y is a compact set, the local uniform boundedness of $U(y)$ implies the uniform boundedness of $U(y)$ on all of Y. Therefore, there exists a compact set U^* such that $U(y) \subseteq U^*$ for all y in Y. The constraint $(\lambda, \mu) \in U^*$ can then be added without disturbing the equality. Since $L^*(\overline{x}, y, \lambda, \mu)$ is a continuous function, $P(y)$ is upper semicontinuous at $\overline{y}$ (Maximum Theorem of Berge, Ref. 35). □

Theorem 7.1. (Finite ϵ-convergence) If the following conditions hold:

(a) X and Y are nonempty compact convex sets, and $Y \subseteq V$,

(b) Conditions (A),

(c) $f(x, y)$, $g(x, y)$ and $h(x, y)$ are continuous on $X \times Y$, and

(d) The set $U(y)$ of optimal multipliers for the primal problem is nonempty for all $y \in Y$ and uniformly bounded in some neighborhood of every such point,

then,
For any given $\epsilon > 0$, the **GOP** algorithm terminates in a finite number of steps.

Proof. Fix $\epsilon > 0$ arbitrarily. Suppose that the **GOP** algorithm does not converge in a finite number of iterations (that is, $(\mu_{\boldsymbol{B}}^{min^k}) < P^{UBD} - \epsilon$). Let $< y^{\boldsymbol{k}}, (\mu_B^{min^k}) >$ be the sequence of optimal solutions to the relaxed dual problem at successive iterations k. Note that any solution from the stored set, if selected as the minimum for a given iteration, is removed from the stored set. Therefore, by taking a subsequence, if necessary, we may assume that $< y^{\boldsymbol{k}}, (\mu_B^{min^k}) >$ converges to $(\overline{y}, \overline{\mu_B^{min}})$ such that $\overline{y} \in Y$. At every iteration, there is an accumulation of constraints from previous iterations. This implies that $\mu_B^{\boldsymbol{min}}$ is a nondecreasing sequence which is bounded above by the optimal value of the original problem. Also, at every iteration, $y^{\boldsymbol{k}}$ is in the compact set Y. Similarly, since $U(y)$ is uniformly bounded for all $y \in Y$, we may assume that the corresponding sequence of multipliers for the primal problems $(\lambda^{\boldsymbol{k}}, \mu^{\boldsymbol{k}})$ converges to $(\overline{\lambda}, \overline{\mu})$, and that the the solutions of the corresponding primal problems $(x^{\boldsymbol{k}}, P(y^{\boldsymbol{k}}))$ converge to $(\overline{x}, P(\overline{y})\ )$. From Lemma 7.2, we have that $(\overline{\lambda}, \overline{\mu}) \in U(\overline{y})$. From Lemma 7.1,

$$L(x^{\boldsymbol{B}_j}, \overline{y}, \overline{\lambda}, \overline{\mu})\big|_{\boldsymbol{x}^k}^{\boldsymbol{lin}} = P(\overline{y}) \quad for\ any\ x^{\boldsymbol{B}_j} \tag{28}$$

Now, at every iteration k, due to accumulation of constraints,

$$\mu_{\boldsymbol{B}}^{min^k} \geq L(x^{\boldsymbol{B}_j}, y^{\boldsymbol{k}+1}, \lambda^{\boldsymbol{k}}, \mu^{\boldsymbol{k}})\big|_{\boldsymbol{x}^k}^{\boldsymbol{lin}}$$

for some combination of bounds $x^{\boldsymbol{B}_j}$. Therefore, by continuity of $L(x^{\boldsymbol{B}_j}, y, \lambda^{\boldsymbol{k}}, \mu^{\boldsymbol{k}})\big|_{\boldsymbol{x}^k}^{\boldsymbol{lin}}$ and (28), we obtain $\overline{\mu_{\boldsymbol{B}}^{min}} \geq P(\overline{y})$. The upper semicontinuity of $P(y)$ at $\overline{y}$ (Lemma 3), then implies that $\mu_{\boldsymbol{B}}^{min^k} \geq P(y^{\boldsymbol{k}}) - \epsilon$

for all k sufficiently large, which contradicts the assumption that the termination criterion in Step 5 is never met. □

Remark 7.1. It should be noted that for the general nonconvex nonlinear problem, condition (d) above may be difficult to prove (or disprove.) However, for the problems considered in this paper (namely problems satisfying *Conditions (A)*), it is easy to show that condition (d) is always satisfied. The proof for this comes from the fact that the set of variables is assumed to be bounded. Hence, the only way a multiplier for a constraint can be unbounded is if that constraint has a variable whose coefficient is zero. However, in such a case, the variable will simply vanish from the constraint, and will not directly affect the multiplier for that constraint. In the case of iterations where the primal problem is infeasible, the *relaxed* primal problem (22) is solved. Again, the same argument (as above) holds for this problem too.

Theorem 7.2. (Global Optimality) If the conditions stated in Theorem 7.1 hold, then

(i) The solution of the Relaxed Dual **(RD)** problem in Step (3) of the algorithm in Section 6 will always be a valid underestimator of the solution of problem (1).

(ii) The **GOP** algorithm will terminate at the global optimum of (1).

Proof.

(i) From Property 4.4, the solution of the relaxed dual problem in Step(3) will underestimate the solution of the relaxed dual problem (5). Since (5) has fewer constraints than the dual of the original problem, it represents a lower bound on the solution of (1). Hence, the solution of the relaxed dual problem in Step (3) will always be a valid underestimator of the optimal solution of (1).

(ii) The primal problem at every iteration represents an upper bound for the original problem (1), while the relaxed dual problem contains fewer constraints than the original problem and thus represents a valid lower bound on the solution of (1). Therefore, since the termination of the algorithm is based on the difference between the lowest upper bound (from the primal problems) and the largest lower bound (from the relaxed dual problems), the algorithm will terminate when these two bounds are both within ϵ of the solution of (1). From Theorem 7.1, the algorithm terminates in a finite number of steps. Hence, the **GOP** algorithm terminates at an ϵ-global optimum of (1). □

Remark 7.2. It has been assumed throughout the theoretical development outlined in this paper that some form of constraint qualification (for example, Slater's qualification) holds for the problem being solved. If such a condition cannot be satisfied, then it is possible that for some fixed values of y, the primal problem will be over-specified, i.e. there are more constraints than variables. Usually, this implies a linear dependency in some of the constraints. It should be noted that for such problems, the **GOP** algorithm cannot be guaranteed to converge to the optimal solution.

8. Illustration for Bilinear Problems

In this section, the **GOP** algorithm is illustrated through application to the following bilinear problem suggested by one of the referees :

$$\min_{\boldsymbol{x},\boldsymbol{y}} \ -y \ , \qquad s.t. \quad xy = 0 \ , \quad -1 \leq x, y \leq 1 \ .$$

The optimal solution is -1, and occurs at $(0,1)$. Consider the starting point of $y = 0$.

Iteration 1 : For $y^1 = 0$, the primal problem can be written as

$$\min_{\boldsymbol{x}} \ 0 \ , \qquad s.t. \quad -1 - x \leq 0 \ , \quad x - 1 \leq 0 \ .$$

The solution of this problem is 0. Since the objective is constant, all the multipliers are zero. The Lagrange function formulated from this problem is $L(x, y, \lambda^1, \mu^1) = -y$. From this, it can be seen that the gradient of the Lagrange function *w.r.t.* x is zero. Hence, the bound used for x in the Lagrange function does not affect the solution, i.e. only one relaxed dual problem needs to be solved. This problem is given below :

$$\min_{\boldsymbol{y}} \ \mu_B \ , \qquad s.t. \quad \mu_B \geq L_1^1(y, \lambda^1, \mu^1) = -y \ , \quad -1 - y \leq 0 \ , \quad y - 1 \leq 0 \ .$$

The solution of this problem is $y = 1, \ \mu_B = -1$. Thus, after the first iteration, the upper bound on the global solution is 0 and the lower bound is -1. The value of $y = 1$ is chosen as the fixed value for the primal problem of the next iteration.

Iteration 2 : For $y^2 = 1$, the primal problem is given below :

$$\min_{\boldsymbol{x}} \ -1 \ , \qquad s.t. \quad x = 0 \ , \quad -1 - x \leq 0 \ , \quad x - 1 \leq 0 \ .$$

The solution of this problem yields $x = 0$, and $\lambda^2 = \mu_1^2 = \mu_2^2 = 0$. The objective function value is -1. This is lower than the solution of the first primal problem, and hence becomes the new upper bound on the global solution.

At this point, the lower and upper bounds from the primal and relaxed dual subproblems are equal. Hence, the algorithm can be terminated, having converged to the global solution.

9. Illustration For Polynomial Problems

Consider the application of the **GOP** algorithm to the following problem :

$$\min_{y} \; -6y + 4.5y^2 - y^3 \qquad s.t. \quad 0 \leq y \leq 3 \; .$$

This problem has a global solution of -4.5 at $y = 3$, and a local solution of -2.5 at $y = 1$.

The introduction of two new variables x_1 and x_2 and two constraints ($x_1 - y = 0$ and $x_2 - x_1 y = 0$) enables the problem to be rewritten in the following equivalent form :

$$\min_{y} \; -6y + 4.5x_2 - x_2 y \; , \qquad s.t. \quad x_1 = y \; , \quad x_2 = x_1 y \; , \quad 0 \leq x_1, y \leq 3 \; , \quad 0 \leq x_2 \leq 9 \; .$$

It should be noted that the bounds on the x variables need not be considered explicitly, since the equivalence relations restrict the values of x_1 and x_2 depending on what values y can take.

Consider the starting point of $y = 2$ for the application of the **(GOP)** algorithm.

Iteration 1 : For $y^1 = 2$, the primal problem can be written as

$$\min_{x} \; -12 + 2.5x_2 \; , \qquad s.t. \quad x_1 - 2 = 0 \; , \quad x_2 - 2x_1 = 0 \; .$$

The solution of this problem is $x_1 = 2$, $x_2 = 4$, $\lambda_1^1 = -5$, and $\lambda_2^1 = -2.5$, where λ_1^1 and λ_2^1 are the Lagrange multipliers corresponding to the two new constraints. The objective function has a value of -2, and provides first upper bound on the global optimum.

The Lagrange function formulated from this problem is

$$L(x, y, \lambda_1^1, \lambda_2^1) = -6y + 4.5x_2 - x_2 y - 5(x_1 - y) - 2.5(x_2 - x_1 y) = (2.5x_1 - x_2) \cdot (y - 2) - y \; .$$

Thus, the gradient of the Lagrange function w.r.t x_1 and x_2 has the form $y - 2 \geq 0$ or $y - 2 \leq 0$. Since there are two x variables, there are four (2^2) subproblems solved in the relaxed dual, as shown below :

	Problem 1	Problem 2	Problem 3	Problem 4
Bounds for x	$x^{B_1} = (0, 0)$	$x^{B_2} = (3, 0)$	$x^{B_3} = (0, 9)$	$x^{B_4} = (3, 9)$
$Qualifying$ Constraint for x_1	$y - 2 \geq 0$	$y - 2 \leq 0$	$y - 2 \geq 0$	$y - 2 \leq 0$
$Qualifying$ Constraint for x_2	$y - 2 \leq 0$	$y - 2 \leq 0$	$y - 2 \geq 0$	$y - 2 \geq 0$

It can be seen that for Problems 1 and 4, the *qualifying* constraints for x_1 and x_2 are simultaneously of the form $y - 2 \geq 0$ and $y - 2 \leq 0$. For these two problems, therefore, the introduction of these constraints is equivalent to fixing $y = 2$. Hence, the solutions of these problems will simply be the point $y = 2, \mu_B = -2$. Hence, it is only necessary to solve problems 2 and 3.

These two *relaxed dual* subproblems to be solved are shown below:

(i) (Problem 2) For $x_1^{B_2} = 3$, $x_2^{B_2} = 0$, $y - 2 \leq 0$.

The relaxed dual subproblem is

$$\min_{y} \ \mu_B \ , \qquad s.t. \quad \mu_B \geq L_1^1(x^{B_2}, y, \lambda_1^1, \lambda_2^1) = 6.5y - 15 \ , \quad y - 2 \leq 0 \ , \quad 0 \leq y \leq 3 \ .$$

The solution of this problem is $y = 0, \ \mu_B = -15$.

(ii) (Problem 3) For $x_1^{B_3} = 0$, $x_2^{B_3} = 9$, $y - 2 \geq 0$.

The relaxed dual subproblem is

$$\min_{y} \ \mu_B \ , \qquad s.t. \quad \mu_B \geq L_2^1(x^{B_3}, y, \lambda_1^1, \lambda_2^1) = -10y + 18 \ , \quad y - 2 \geq 0 \ , \quad 0 \leq y \leq 3 \ .$$

The solution of this problem is $y = 3, \ \mu_B = -12$.

Thus, after the first iteration, there are two solutions of (μ_B, y) in the stored set. From these, the solution corresponding to the minimum μ_B is chosen. In this case, this corresponds to the solution $\mu_B = -15, \ y = 0$. Hence, the fixed value of y for the second iteration is 8 . The selected solution is then deleted from the stored set.

Iteration 2 : For the second iteration, the primal problem, with $y^2 = 0$, is given below :

$$\min_{x} \ 4.5x_2 \ , \quad s.t. \quad x_1 = 0 \ , \quad x_2 = 0 \ .$$

Its solution yields $x_1 = 0$, $x_2 = 0$, $\lambda_1^2 = 0$, $\lambda_2^2 = -4.5$, and objective value of 0.

The Lagrange function formulated from the second primal problem is

$$L(x, y, \lambda_1^2, \lambda_2^2) = -6y + 4.5x_2 - x_2 y + 0(x_1 - y) - 4.5(x_2 - x_1 y) = (4.5x_1 - x_2) \cdot (y - 0) - 6y \ .$$

Again, it can be easily seen that only two relaxed dual subproblems need to be solved, for the combinations of bounds $(3, 0)$ and $(0, 9)$ respectively for x_1 and x_2. Before solving these problems, a Lagrange function needs to be selected from the first iteration. In order to do this, the *qualifying* constraints for the Lagrange functions are checked at $y^2 = 0$. This indicates that the Lagrange function formulated for Problem 3 of the first iteration (with the *qualifying* constraint $y - 2 \leq 0$ for both x_1 and x_2) can be present for the current relaxed dual subproblems.

The two relaxed dual subproblems solved at the second iteration are shown below:

(i) (Problem 2) For $x_1^{B_2} = 3$, $x_2^{B_2} = 0$, $y - 0 \leq 0$.

$$\min_{y} \ \mu_B, \quad s.t. \quad \mu_B \geq L_1^1 = 6.5y - 15, \quad y - 2 \leq 0, \quad \mu_B \geq L_1^2 = 7.5y, \quad y - 0 \leq 0, \quad 0 \leq y \leq 3.$$

The solution of this problem is $y = 0, \ \mu_B = 0$.

(ii) (Problem 3) For $x_1^{B_3} = 0$, $x_2^{B_3} = 9$, $y - 0 \geq 0$.

$$\min_{y} \ \mu_B, \quad s.t. \quad \mu_B \geq L_1^1 = 6.5y - 15, \quad y - 2 \leq 0, \quad \mu_B \geq L_2^2 = -15y, \quad y - 0 \geq 0, \quad 0 \leq y \leq 3.$$

The solution of this problem is $y = 0.6976, \ \mu_B = -10.4651$.

At the end of the second iteration, there are two stored solutions left, namely $y = 3, \mu_B = -12$ and $y = 0.6976, \mu_B = -10.4651$. From these, the first solution is chosen as the one with the smaller value of μ_B. Therefore, the new lower bound for the problem is -10.4651, and the fixed value of y for the next primal problem is $y = 3$.

Iteration 3 : The primal problem is solved for $y^3 = 3$. The solution of this problem yields $x_1 = 3$, $x_2 = 9$, $\lambda_1^3 = -4.5$, and $\lambda_2^3 = -1.5$. The objective function has a value of -4.5. Since this is less than the best solution from previous iterations, the new upper bound for the global solution is -4.5. The Lagrange function formulated from the third primal problem is

$$L(x, y, \lambda_1^3, \lambda_2^3) = -6y + 4.5x_2 - x_2 y - 4.5(x_1 - y) - 1.5(x_2 - x_1 y) = (1.5x_1 - x_2) \cdot (y - 3) - 1.5y$$

Again, two relaxed dual problems are solved, one each for the combinations of bounds $(3, 0)$ and $(0, 9)$ for x_1 and x_2. Before solving these problems, a Lagrange function from each of the previous two iterations is selected. This leads to the selection of L_2^1 and L_2^2 from the first and second iterations respectively, since their *qualifying* constraints are satisfied at $y^3 = 3$.

The two solutions of the two relaxed dual subproblems are $y = 2.423, \mu_B = -6.2308$ and $y = 3, \mu_B = -4.5$. Thus, at the end of the third iteration, there are two solutions less than the upper bound on the global solution (i.e. less than -4.5). From these, the solution of $y = 0.6976, \mu_B = -10.4651$ is selected. Hence, the new lower bound on the global solution is -10.4651 and the fixed value of y for the next iteration is $y^4 = 0.6976$.

The algorithm continues in this fashion until the lower and upper bounds are within ϵ, taking 17 iterations to converge to the global solution.

10. Geometrical Interpretation

The application of the **GOP** algorithm to the second illustrating example (Section 9) can be interpreted geometrically. At every iteration, the solution of the primal problem for a fixed value of $y = y^K$ is simply an evaluation of the objective function at y^K. The two Lagrange functions used for the relaxed dual problems are underestimators of the objective function for $y \leq y^K$ and $y \geq y^K$, and these two Lagrange functions intersect at $y = y^K$. Consider the first relaxed dual problem, for which the Lagrange function is evaluated at $x_1 = 3 \; and \; x_2 = 0$, and the domain of y is restricted to $y \leq y^K$. From a previous iteration k, if $y^k \leq y^K$, then the Lagrange function from that iteration evaluated at $x_1 = 0, x_2 = 9$ will underestimate the objective function for all values of y between y^k and y^K, and therefore will be present in the current relaxed dual problem. Conversely, for an iteration k where $y^k \geq y^K$, the Lagrange function from that iteration evaluated at $x^B = (3, 0)$ will underestimate the objective function for all values of y between y^K and y^k, and therefore will be present in the relaxed dual problem. The converse holds for the other relaxed dual problem for the current iteration. The solutions of these two relaxed dual problems will lie between y^K and the nearest y^k on either side of y^K.

By storing the solutions of each of the relaxed dual problems at the current iteration, we ensure that the algorithm can, if necessary, return a value of y from either side of y^K for the $(K+1)$th iteration. At the same time, the criterion for selecting the Lagrange functions from previous iterations results in the creation of an underestimating function for the objective function that resembles a series of valleys and peaks, with the valleys representing the stored solutions of the relaxed dual problems at different iterations.

For a starting point of $y^1 = 2$, the sequence of points generated by the algorithm is graphically illustrated in Figures 1-4. $f(y)$ is the optimal value of the primal problem for different fixed values of y, and in this case happens to be identical to the original function. For the first iteration (Figure 1), with an optimal value of -2 for the primal problem, L_1^1 and L_2^1 are the Lagrangians evaluated at the two sets of bounds $(3, 0)$ and $(0, 9)$ respectively for (x_1, x_2). As can be seen, each Lagrange function underestimates the objective function for one side of $y^1 = 2$. The solutions of the two relaxed dual problems give (0,-15) and (3,-12) for (y, μ_B). These two values are stored. Then, the solution providing the lower μ_B, i.e. (0,-15) is selected, and deleted from the stored set.

For the second iteration, $y^2 = 0$ for the primal problem, and the optimal solution is 0. From the solution of this problem, two new Lagrange functions are generated, which are indicated by L_1^2 and L_2^2 in Figure 2. For the first relaxed dual problem in this iteration, the Lagrange function L_1^2 is present, along with L_1^1. The solution of this problem is $(y = 0, \mu_B = 0)$. For the second relaxed dual problem, L_2^1 is present from iteration 1, and L_2^2 is present from the current iteration. The solution of this problem is at $(y = 0.698, \mu_B = -10.465)$. These two solutions are stored. From the stored set, the lowest value, which is $\mu_B = -12$, is selected as the new lower bound, and the corresponding $y = 3$ is selected as the fixed value for the next iteration. The selected solution $y = 3, \mu_B = -12$ is deleted from the stored set.

From the third iteration, with an optimal value of -4.5 for the primal problem, the two Lagrange functions obtained are L_1^3 and L_2^3, shown in Figure 3. For this iteration, L_2^1 and L_2^2 are present from the previous iterations as constraints. The solution of the two relaxed dual problems lie at (2.423,-6.24) and (3,-4.5). These are stored, and the least value of all the stored solutions $(\mu_B = -10.465)$ is selected as the new bound. The corresponding $y = 0.698$ is the fixed value of y for the next primal problem. The algorithm continues in this manner until the lower bound comes within ϵ of -4.5 .

Figure 4 shows the underestimating function that is effectively obtained after four iterations. The lower bound from the fourth iteration is simply the lowest *valley* of this underestimating function.

11. Quadratic Problems with Quadratic Constraints

The quadratic programming problem with quadratic constraints has the following form:

$$
\begin{aligned}
& \min_{x} \quad c^T x \;+\; x^T Q x \;, \\
s.t. \quad & x^T A_m x + Bx - b_m \;\leq\; 0 \;, \qquad m = 1, 2, ..., p \;, \\
& x^T A_m x + Cx - b_m \;=\; 0 \;, \qquad m = p+1, p+2, ..., p+q \;, \\
& Dx - d \;\leq\; 0 \;, \\
& Ex - e \;=\; 0 \;,
\end{aligned}
$$

where x an n-vector of variables, and c, d and e are constant vectors. Q, B, C, D, and E are constant matrices. A_m is an $n \times n$ matrix corresponding to the mth quadratic constraint, and b_m is a constant for that constraint. It is assumed that the bounds on the variables are explicitly incorporated into the problem in the third constraint set.

By defining a new set of variables $y = x$, this problem can be rewritten in the following form :

$$
\begin{aligned}
& \min_{x,y} \quad c^T x \;+\; x^T Q y \;, \\
s.t. \quad & x^T A_m x + Bx - b_m \;\leq\; 0 \;, \qquad m = 1, 2, ..., p \;, \\
& x^T A_m x + Cx - b_m \;=\; 0 \;, \qquad m = p+1, p+2, ..., p+q \;, \\
& Dx - d \;\leq\; 0 \;, \\
& Ex - e \;=\; 0 \;, \\
& x - y \;=\; 0 \;.
\end{aligned}
$$

For some problems, it may not be necessary to introduce a y variable for every x variable. The introduction of the variables as shown above is only one of several possible ways of converting the problem to a form satisfying *Conditions (A)*. The reader is referred to Hansen and Jaumard (Ref. 31) for more information on the efficient bilinearization of these problems. The reader is also referred to Visweswaran and Floudas (Ref. 27) for complete details and several examples of application to quadratic problems with linear constraints.

The primal problem is solved for $y = y^k$, and its solution provides the optimal multiplier vectors μ_1^k, λ_1^k, μ_2^k, λ_2^k and ν^k respectively for the five sets of constraints. Using the KKT gradient conditions, the Lagrange function can be formulated as

$$L(x, y, \lambda_1^k, \mu_1^k, \lambda_2^k, \mu_2^k, \nu) = x^T(Q + ({\mu_1^k}^T, {\lambda_1^k}^T)A)(y - y^k)^T - ({\mu_1^k}^T, {\lambda_1^k}^T)b - {\mu_2^k}^T d - {\lambda_2^k}^T e - {\nu^k}^T y.$$

Hence, the *qualifying* constraints to be added to the relaxed dual problem take the form

$$(Q_i + \sum_{m=1}^{p} \mu_{1_m}^k A_{m_i} + \sum_{m=p+1}^{p+q} \lambda_{1_m}^k A_{m_i}).(y - y^k) \ \leq \ 0 \qquad if \quad x_i^B = x_i^U$$
$$(Q_i + \sum_{m=1}^{p} \mu_{1_m}^k A_{m_i} + \sum_{m=p+1}^{p+q} \lambda_{1_m}^k A_{m_i}).(y - y^k) \ \geq \ 0 \qquad if \quad x_i^B = x_i^L$$

where the subscript i myrefers to the ith row of the matrices Q and A_m, i.e., the rows corresponding to the variable x_i in the matrices. Similarly, the *qualifying* constraints for iterations with infeasible primal problems can be generated.

11.1. Example : The Pooling Problem

A complete nonlinear programming **NLP** formulation for the pooling problem (Refs. 12 and 31) is shown below :

$$\min \quad 6A + 13B + 10(C_x + C_y) - 9x - 15y \ ,$$

$$\begin{array}{rrr} s.t. \quad & P_x + P_y - A - B \ = \ 0\ , & p \cdot (P_x + P_y) - 3A - B \ = \ 0\ , \\ & x - P_x - C_x \ = \ 0\ , & y - P_y - C_y \ = \ 0\ , \\ & p \cdot P_x + 2 \cdot C_x - 2.5x \ \leq \ 0\ , & p \cdot P_y + 2 \cdot C_y - 1.5y \ \leq \ 0\ , \\ & 0 \leq x \leq x^U, \quad 0 \leq y \ \leq \ y^U, & A,\ B,\ C_x,\ C_y,\ P_x,\ P_y \ \geq \ 0\ . \end{array}$$

where p is the sulfur quality of the pool; its lower and upper bounds are 1 and 3 respectively. A and B are two input streams to the pool, and P_x and P_y are the two output streams from the pool. These streams are mixed with bypass streams C_x and C_y to produce two final output streams having qualities of x and y respectively.

Projection on the pool quality p makes the primal problem linear in the remaining variables. It can be seen that only the variables P_x and P_y are directly *connected* with the variable p. Hence, the relaxed dual needs to be solved at the bounds of only these two variables.

The **GOP** algorithm was applied to this problem with the upper bounds on x and y being 100 and 200 respectively. The problem exhibits a strong local minimum at $p = 2.5$, with the optimal solution $v(y)$ actually being discontinuous. The algorithm found the global optimum of -750 at $p = 1.5$ from several starting points, required an average of 10 iterations to converge. (Note : An extensive treatment of the application of the **GOP** algorithm to this problem and other quadratically constrained problems is provided in Visweswaran and Floudas (Ref. 27)).

12. Polynomial Functions

A vast number of problems involve the minimization of polynomial functions of one or more variables. This section presents the application of the **GOP** algorithm to polynomial functions of a single variable. The general approach can be easily extended to include functions of more than one variable, and can also be applied for polynomial constraints as well as rational polynomial functions.

Consider the unconstrained minimization problem

$$\min_{y \in Y} \quad f(y) = a_0 + a_1 y + a_2 y^2 + \ldots\ldots + a_s y^s \tag{29}$$

where y is a *single* variable. The presence of either negative coefficients a_i or the presence of odd powers of y in the function can give rise to nonconvexities. Let r be the highest power of y such that $a_r y^r$ is nonconvex in y for $y \in Y$. Then, by introducing $(r-1)$ transformation variables, this problem can be decomposed into convex primal and relaxed dual problems, enabling the application of the **GOP** algorithm.

Consider the following transformations:

$$x_0 = 1, \quad x_1 = y, \quad x_2 = y^2 = x_1 y, \; \ldots \; x_r = y^r = x_{r-1} y \; .$$

Projecting on y, the primal problem for this formulation becomes, for a fixed value of $y = y^K$,

$$\min_x \sum_{i=0}^{r} a_i x_i \; + \sum_{i=r+1}^{s} a_i (y^K)^i,$$

$$s.t. \quad x_i - x_{i-1} y^K = 0 , \quad i = 1, 2, ..., r ,$$
$$x_0 = 1 .$$

The Lagrange function for this problem is given by

$$L(x, y, \lambda^K) = \sum_{i=0}^{r} a_i x_i + \sum_{i=r+1}^{s} a_i y^i + \sum_{i=1}^{r} \lambda_i^K (x_i - x_{i-1} y) ,$$

where λ_i^K are the Lagrange multipliers for the equality constraints. Using the KKT gradient conditions for the primal problem, the Lagrange function can be simplified to

$$L(x, y, \lambda^K) = \sum_{i=0}^{r} \lambda_{i+1}^K (y^K - y) x_i + \sum_{i=r+1}^{s} a_i y^i - \lambda_0^K ,$$

where $\lambda_{r+1}^K = 0$ and λ_0^K is the Lagrange multiplier for the last constraint. Thus, for every x_i, the *qualifying* constraint to be put into the relaxed dual problem is of the form

$$y^K - y \leq 0, \; or \quad y^K - y \geq 0$$

It is therefore sufficient to solve the relaxed dual for these two regions of y, with x_i set to the appropriate bounds, for then there will be a Lagrange function for each of the regions $y \leq y^K$ and $y \geq y^K$, which will underestimate the optimal value of the problem (29) for every y in that region. The bounds for x_i for these two relaxed dual problems can be selected as follows:

(i) For $y^K - y \geq 0$: If $\lambda_{i+1}^K \geq 0$, then $x_i = x_i^L$; otherwise, $x_i = x_i^U$.

(i) For $y^K - y \leq 0$: If $\lambda_{i+1}^K \geq 0$, then $x_i = x_i^U$; otherwise, $x_i = x_i^L$.

Using these combinations of bounds, the two relaxed dual problems can be solved for the appropriate regions of y.

12.1. Example : Rosenbrock's function

This example considers the minimization of Rosenbrock's function, which is given by :

$$f(x_1, y_1) = a(x_1 - y_1^2)^2 + (y_1 - b)^2,$$

where $a = 100$ and $b = 1$. This function has its global minimum of 0 at $x_1 = 1, y_1 = 1$.

The problem can be converted to a form satisfying *Conditions (A)* by the introduction of two transformation variables $x_2 = y_1, \quad x_3 = x_1 - y_1^2 = x_1 - x_2 y_1$. By projecting on y_1, the primal problem becomes convex in x_1, x_2 and x_3, and the relaxed dual problem is a convex problem in y_1 . Since the Lagrange function at iteration K will be convex in x_3, it is linearized around x^K, the solution of the primal problem at that iteration.

When the **GOP** algorithm was applied to the problem in this form, it converged to the global optimum of 0 starting from all starting points. The number of iterations required for convergence depended on the starting point, with an average of 10 iterations.

Remark 12.1. It should be noted that this class of problems has some very interesting features arising from the nature of the transformations that give a direct one-to-one correspondence between the x variable set and the single y variable. Consequently, it is possible to utilize this relationship to improve the bounds for the x variables iteratively. This leads to a very efficient algorithm for solving problems of this class, and is discussed in further detail in Visweswaran and Floudas (Ref. 36).

Conclusions

In this paper, a new deterministic global optimization approach is proposed for the solution of nonconvex programming problems of a specific structure. The proposed approach covers the general quadratic programming problem, quadratic programming problems with quadratic constraints, and problems with polynomial and rational polynomial functions in their objective function and/or constraints. New theoretical properties are proposed that enable the rigorous solution of the *relaxed dual* problem. Based upon these properties, a global optimization algorithm has been developed. The algorithm is shown to have finite ϵ-convergence and ϵ-global optimality. The algorithm has been illustrated geometrically and numerically through a simple example. The application of the algorithm to specific classes of problems is given through the development of the theory, and application to some example problems.

The nature of the solution of the relaxed dual subproblems permits the exploitation of the structure of the Lagrange functions used in these subproblems. By developing new properties based upon this structure, it is possible to eliminate a large number of the relaxed dual subproblems without destroying the rigorous nature of the algorithm. The properties that achieve this result are given in Visweswaran and Floudas (Ref. 33). In addition, it is possible to solve the relaxed dual subproblems simultaneously as a single MILP problem by introducing binary variables representing the sign of the *qualifying constraints*. This is discussed in Floudas *et al* (Ref. 37).

For the case of polynomial functions in one variable, the nature of the transformations provide a direct one-to-one correspondence between the x variable set and the single y variable. This relationship can be used to iteratively improve the bounds for the x variables. This leads to a very efficient algorithm for solving problems of this class, and is discussed in further detail in Visweswaran and Floudas (Ref. 36).

The **GOP** algorithm can be applied to several other classes of problems, most notably bilevel programming problems, linear and nonlinear complementarity problems, and integer quadratic programming problems. Work on these and other classes of problems, as well as work on improving the computational efficiency of the algorithm, is currently in progress, and will be reported in future publications.

References

1. DIXON, L.C.W., and SZEGO, G.P., Editors, Towards Global Optimization, North Holland, Amsterdam, Holland, 1975.

2. DIXON, L.C.W., and SZEGO, G.P., Editors, Towards Global Optimization 2, North-Holland, Amsterdam, Holland, 1978.

3. ARCHETTI, F., and SCHOEN, F., A Survey on the Global Minimization Problem : General Theory and Comp nals of Operations Research, Vol.1, pp. 87-110, 1984.

4. PARDALOS, P. M., and ROSEN, J.B., Methods for Global Concave Minimization : A Bibliographic Survey, SIAM Review, Vol. 28, pp. 367-379, 1986.

5. PARDALOS, P.M., and ROSEN, J.B., Constrained Global Optimization : Algorithms and Applications, Lecture Notes in Computer Science, Vol. 268, Springer-Verlag, Berlin, Germany, 1987.

6. TORN, A., and ZILINSKAS, A., Global Optimization, Lecture Notes in Computer Science, Vol. 350, Springer-Verlag, Berlin, Germany, 1989.

7. RATSCHEK, H., and ROKNE, J., New Computer Methods for Global Optimization, Halsted Press, Chicester, Great Britain, 1988.

8. MOCKUS, J., Bayesian Approach to Global Optimization - Theory and Applications, Kluwer Academic Publishers, Amsterdam, Holland, 1989.

9. HORST, R., and TUY, H., Global Optimization: Deterministic Approaches, Springer-Verlag, Berlin, Germany, 1990.

10. FLOUDAS, C. A., and PARDALOS, P. M., A Collection of Test Problems for Constrained Global Optimization Algorithms, Lecture Notes in Computer Science, Vol. 455, Springer-Verlag, Berlin, Germany, 1990.

11. FLOUDAS, C. A. and PARDALOS, P. M., Recent Advances in Global Optimization, Princeton University Press, Princeton, New Jersey, 1992.

12. PIYAVSKII, S.A., An Algorithm for Finding the Absolute Extremum of a Function, USSR Computational Mathematics and Mathematical Physics, Vol. 12, pp. 57-67, 1972.

13. PARDALOS, P. M., GLICK, J. H., and ROSEN, J. B., Global Minimization of Indefinite Quadratic Problems, Computing, Vol. 39, pp. 281-291, 1987.

14. AL-KHAYYAL, F. A., Jointly Constrained Bilinear Programs and Related Problems : An Overview, Computers in Mathematical Applications, Vol. 19, pp. 53-62, 1990.

15. HANSEN, P., JAUMARD, B., and LU, S-H., An Analytical Approach to Global Optimization, Mathematical Programming, Vol. 52, pp. 227-241, 1991.

16. TUY, H., THIEU, T. V., and THAI, N. Q., A Conical Algorithm for Globally Minimizing a Concave Function over a Closed Convex Set, Mathematics of Operations Research, Vol. 10, pp. 498-514, 1985.

17. TUY, H., Global Minimum of a Difference of Two Convex Functions, Mathematical Programming Study, Vol. 30, pp. 150-182, 1987.

18. TUY, H., Convex Programs With an Additional Reverse Convex Constraint, Journal of Optimization Theory and its Applications, Vol. 52, pp. 463-472, 1987.

19. TUY, H., On Outer Approximation Methods for Solving Concave Minimization Problems, Acta Mathematica Vietnamica, Vol. 8, pp. 3-***, 1983.

20. HORST, R., THOAI, N. V., and De Vries, J., A New Simplicial Cover Technique in Constrained Global Optimization, Journal of Global Optimization, Vol. 2, pp. 1-19, 1992.

21. SHOR, N. Z., Dual Quadratic Estimates in Polynomial and Boolean Programming, Annals of Operations Research, Vol. 25, pp. 163-168, 1990.

22. FLOUDAS, C.A., AGGARWAL, A., and CIRIC, A.R., Global Optimum Search for Nonconvex NLP and MINLP problems, Computers and Chemical Engineering, Vol. 13, pp. 1117-1132, 1989.

23. AGGARWAL, A., and FLOUDAS, C.A., A Decomposition Approach for Global Optimum Search in QP, NLP and MINLP problems, Annals of Operations Research, Vol. 25, pp. 119-146, 1990.

24. SHERALI, H. and TUNCBILEK, C. H., A Global Optimization Algorithm for Polynomial Programming Problems Using a Reformulation-Linearization Technique, Journal of Global Optimization, Vol. 2, pp. 101-112, 1992.

25. HANSEN, E.R., Global Optimization Using Interval Analysis: The Multi-Dimensional Case, Numerische Mathematik, Vol. 34, pp. 247-270, 1980.

26. FLOUDAS, C. A. and VISWESWARAN, V., A Global Optimization Algorithm for Certain Classes of Nonconvex NLPs - I. Theory, Computers and Chemical Engineering, Vol. 14, pp. 1397-1417, 1990.

27. VISWESWARAN, V. and FLOUDAS, C. A., A Global Optimization Algorithm for Certain Classes of Nonconvex NLPs - II. Application of Theory and Test Problems, Computers and Chemical Engineering, Vol. 14, pp. 1419-1434, 1990.

28. GEOFFRION, A.M., Generalized Benders Decomposition, Journal of Optimization Theory and Applications, Vol. 10, pp. 237-260, 1972.

29. WOLSEY, L. A., A Resource Decomposition Algorithm for General Mathematical Programs, Mathematical Programming Study, Vol. 14, pp. 244-257, 1981.

30. FLIPPO, O. E., Stability, Duality and Decomposition in General Mathematical Programming, Erasmus University, PhD Thesis, Rotterdam, 1990.

31. HANSEN, P., and JAUMARD, B., Reduction of Indefinite Quadratic Programs to Bilinear Programs, Journal of Global Optimization, Vol. 2, pp. 41-60, 1992.

32. BAZARAA, M. S. and SHETTY, C. M., Nonlinear Programming : Theory and Algorithms, John Wiley & Sons, New York , New York, 1979.

33. VISWESWARAN, V., and FLOUDAS, C. A., New Properties and Computational Improvement of the GOP Algorithm for Problems with Quadratic Objective Function and Constraints, Submitted for Publication, 1991.

34. MEYER, R., The Validity of a Family of Optimization Methods, SIAM Journal on Control, Vol. 8, pp. 41-54, 1970.

35. BERGE, C., Topological Spaces, The Macmillan Co., New York, New York, 1963.

36. VISWESWARAN, V., and FLOUDAS, C. A., Unconstrained and Constrained Global Optimization of Polynomial Functions in One Variable, Journal of Global Optimization, Vol. 2, pp. 73-100, 1992.

37. FLOUDAS, C. A., JAUMARD, B., and VISWESWARAN, V., Decomposition Techniques for Global Optimization, In Preparation, 1992.