A global optimization method, QBB, for twice-differentiable nonconvex optimization problem

Yushan Zhu* Takahito Kuno

July 30, 2001
October 23, 2001 for Revision

ISE-TR-01-182

Institute of Information Sciences and Electronics, University of Tsukuba, Ibaraki 305-8573, Japan
Institute of Chemical Metallurgy, Chinese Academy of Sciences, Beijing 100080, China
*** Tel: +86-298-536791, Fax: +86-298-535206, e-mail: yszhu@syou.is.tsukuba.ac.jp**

Key words. Global optimization, branch-and-bound algorithm, QBB, simplicial division, interval Hessian matrix, twice-differentiable function.

A global optimization method, QBB, for twice-differentiable nonconvex optimization problem

Yushan Zhu[1] Takahito Kuno

Institute of Information Sciences and Electronics, University of Tsukuba, Ibaraki 305-8573, Japan
Institute of Chemical Metallurgy, Chinese Academy of Sciences, Beijing 100080, China

Abstract

A global optimization algorithm, QBB, for twice-differentiable NLPs (Non-Linear Programming) is developed to operate within a branch-and-bound framework and require the construction of a relaxed convex problem on the basis of the quadratic lower bounding functions for the generic non-convex structures. Within an exhaustive simplicial division of the constrined region, the rigorous quadratic underestimation function is constructed for the generic nonconvex function structures by virtue of the maximal eigenvalue analysis of the interval Hessian matrix. Each valid lower bound of the NLP problem with the division progress is computed by the convex programming of the relaxed optimization problem obtained by preserving the convex or linear terms, replacing the concave term with linear convex envelope, underestimating the special terms and the generic terms by using their customized tight convex lower bounding functions or the valid quadratic lower bounding functions, respectively. The standard convergence properties of the QBB algorithm for nonconvex global optimization problems are guaranteed. The computational studies of the QBB algorithm for a general quadratic programming problem is reported to show the global convergence and the algorithmic efficiency whilst the quadratic coefficients are estimated loosely.

Key words. *Global optimization, branch-and-bound algorithm, QBB, simplicial division, interval Hessian matrix, twice-differentiable function.*

---

[1]Correspondence author. Tel: +86-298-536791, Fax: +86-298-535206, e-mail: yszhu@syou.is.tsukuba.ac.jp

1. Introduction

The vast majority of chemical process design and control problems are attracted by the optimal solutions, however those problems are mainly characterized by the existence of multiple minima and maxima, as well as first, second, and higher order saddle points. Those nonconvex optimization problems always frustrate the chemical engineers in their search to arrive at better designs for novel or existing processes. These problems arise in many sorts of engineering chemistry field, such as heat exchange network design, chemical and phase equilibrium, and separation-reaction sequencing. Despite the importance of identifying the global minimum solution or valid bound on that solution, this can rarely reached rigorously. Contributions from the chemical engineering community to the area of global optimization can be traced to the work of Stephanopous and Westerberg (1975), and Westerberg and Shah (1978). Renewed interest in seeking global solution was motivated from the work of Floudas et al. (1989). Thereofore, in the last decade we have experienced a resurgence of interest in chemical engineering for new methods of global optimization as well as the application of available global optimization algorithms to important engineering field (Grossmann, 1996; Floudas, 1999). This recent surge of interest is attributed to three main reasons. First, a large number of engineering chemistry and computational chemistry problems are indeed global optimization problems (Wales and Scheraga, 1999). Second, the existing local nonlinear optimization approaches may fail to obtain even a feasible solution or are trapped to a local optimum solution, which may differ in value significantly from the global solution. Third, the global optimum solution may have a very different physical interpretation when it is compared to local solution, the chemical and phase equilibrium problem is a very real one since in equilibrium a local solution may provide incorrect prediction of types of phases at equilibrium, as well as the compositions in each phase (Mcdonald and Floudas, 1994; Zhu and Xu, 1999; Zhu and Inoue, 2001).

One of the major difficulties with global optimization problems is the lack of practical criteria, which decide when a local solution is global. Then, many iterative schemes (Horst and Pardalos, 1995) are developed which require some global information in each step. The branch-and-bound framework is one of the most promising methods for solving multiextremal global optimization problems. The main idea of this framework consists of two basic operations: successively refined partitioning of the feasible region and estimation of lower and upper bounds for the optimal value of the objective function over each subset generated by the partitions. Most often, lower bounding procedures are established using suitable types of underestimation of the functions involved in the problem under consideration. As a result, lower bounds are computed by solving relaxed problems in the same space of variables as the original problems. In particular, the GOP algorithm for biconvex problems (Floudas and Visweswaran, 1990, 1993) and the branch-and-bound algorithm for bilinear problem (Al-Khayyal and Falk, 1983) rely on mathematical properties specific to the problem solved in order to obtain the lower bounding problem. Floudas and his colleagues (Floudas, 1999) suggested an approach which necessitates the identification of the minimum eigenvalues of the Hessian matrix of the functions to be convexified over a rectangular domain. The αBB algorithm, based on this technique, converges with mathematical rigor to the class of twice-differentiable nonconvex programs. Recently, with the inspiration of the phase stability analysis problem, a quadratic underestimation function based branch and bound algorithm over simpilicial partition of the constrained region is developed for twice-differentiable NLPs.

The determination of phase stability, i.e. whether or not a given mixture will split into multiple phases, is a key step in separation process. Consequently, it facilitates the search for the true equilibrium solution if a postulated solution is thermodynamically unstable with respect to perturbations in any or all of the phases, which can be evaluated by minimizing the tangent plane distance function (TPDF). Zhu and Xu (1999) developed a novel branch and bound algorithm for TPDF described by UNIQUAC equation on the basis of compact partition of the feasible region, where the separable assumption is no longer needed for the construction of the valid underestimation function. However, the nonconvexity is only generated by the concave function in the D.C. (Difference of two Convex functions) formulation of the TPDF. Further, a quadratic underestimation function based branch and bound, QBB, algorithm (Zhu and Inoue, 2001) is developed for the minimization of the stability analysis problem on the basis of a rigorous underestimator constructed by interval analysis, which is a method to expand the application of the QBB algorithm from the special D.C. structure of the stability analysis problem described by UNIQUAC model to the generic non-convex function structure. However, a systematic investigation of the QBB algorithm for the general twice-differentiable NLPs are indispensable so as to converge asymptotically to the global solution with theoretical guarantee. In this paper, the relaxed convex programming problem is constructed based on the quadratic underestimation function under a branch-and-bound framework. The lower bound computed by solving this relaxed problem is monotonic with the refined division of the optimal region. The algorithm convergences are developed by virtue of the exhaustiveness of the simplicial bisection if the QBB algorithm does not terminate after finitely many iterations, since it generates infinite sequences of feasible and/or infeasible points converging to one of the optimal solutions.

2. The QBB global optimization algorithm

The general nonconvex optimization formulation can be formulated as follows:

$$(P) \qquad \min_{\mathbf{x}} \quad f(\mathbf{x})$$

$$\text{s.t.} \quad \mathbf{g}_i(\mathbf{x}) \le 0 \qquad i = 1, 2, \ldots, m$$

$$\mathbf{x} \in \mathbf{S}^0 \subset \Re^n$$

where f and $\mathbf{g}_i$ belong to C^2, the set of twice-differentiable functions, and $\mathbf{S}^0$ is a simplex defined by

$$\mathbf{S}^0 = \left\{ \mathbf{x} \in \Re^n : \mathbf{x} = \sum_{i=1}^{n+1} \lambda_i \mathbf{V}_i^0, \lambda_i \ge 0, \sum_{i=1}^{n+1} \lambda_i = 1 \right\}$$

where $\mathbf{V}_i^0 \in \mathbf{V} \subset \Re^n$, $i = 1, 2, \ldots, n+1$ are $n+1$ vertices of the simplex $\mathbf{S}^0$, and $\mathbf{V}$ is the set of all its vertices. Let D_g be a subset of $\Re^n$ defined by

$$D_g = \{\mathbf{x} \in \Re^n : \mathbf{g}_i(\mathbf{x}) \le 0, i = 1, 2, \ldots, m\}$$

In general, the set D_g is nonconvex and even disconnected. We assume throughout the paper that Problem (P) has an optimal solution, and at least one feasible point is known. Then, for the general nonconvex optimization problem, i.e. (P), the present algorithm belongs to a branch and bound scheme. At each iteration of this algorithm, a branching step and a bounding step must be finished simultaneously. Then, we start to develop this algorithm with the basic operations needed in this scheme.

2.1 Simplicial Partition

For the branching procedure, the simplex $\mathbf{S}^0$ will be divided into refined subregions by using the well-known simplicial partition often used in global optimization algorithm. For every kind of branching, it is a simple matter to check that for every $i \in I$, the points $\mathbf{V}^1$, ..., $\mathbf{V}^{i-1}$, $\mathbf{U}$, $\mathbf{V}^{i+1}$, ..., $\mathbf{V}^{n+1}$ are vertices of a simplex $\mathbf{S}_i \subset \mathbf{S}$, and that:

$$(\text{int}\,\mathbf{S}_i) \cap (\text{int}\,\mathbf{S}_j) = 0 \quad \forall j \neq i; \quad \bigcup_{i \in I} \mathbf{S}_i = \mathbf{S}$$

Then, the simplexes $\mathbf{S}_i$, $i \in I$, form a subdivision of the simplex $\mathbf{S}$ via $\mathbf{U}$. Each $\mathbf{S}_i$ will be referred to as a subsimplex of $\mathbf{S}$. Clearly, the partition is proper since it consists of at least two members if and only if $\mathbf{U}$ does not coincide with any $\mathbf{V}^i$. An important special case is the bisection where the $\mathbf{U}$ is a point of the longest edge of the simplex $\mathbf{S}$, for example $\mathbf{U} \in [\mathbf{V}^m, \mathbf{V}^n]$, i.e.

$$\left\|\mathbf{V}^m - \mathbf{V}^n\right\| = \max_{\substack{i<j \\ i,j=1,\dots n+1}} \left\{\left\|\mathbf{V}^i - \mathbf{V}^j\right\|\right\}$$

where $\|\cdot\|$ denotes any given norm in $\Re^n$, and $\mathbf{U} = a\mathbf{V}^k + (1-a)\mathbf{V}^h$ with $0 < a \leq 1/2$. It should be noted here that *a* means the simplex V is divided into two subsimplexes such that the ratio of the volume of the smaller subsimplex to that of $\mathbf{S}$ is equal to *a*. Zhu and Inoue (2001) used an exact bisection method since the *a* is equal to 1/2. Obviously, in an infinite filter of simplexes $\mathbf{S}_1 \supset \mathbf{S}_2 \ldots \supset \mathbf{S}_k \supset \ldots$, the diameter of the simplex $\mathbf{S}_k$, i.e. $\delta(\mathbf{S}_k)$, the length of the longest edge of $\mathbf{S}_k$, will monotonically decrease. For the convergence proofs of the branch and bound algorithm, the most useful concept is the exhaustiveness of a partition process (Horst, Pardalos, and Thoai, 1995). A nested subsequence of partition sets $\{S^j\}$, i.e. $S^j \supset S^{j+1}$, $\forall j$, is called exhaustive if S^j shrinks to an unique point, i.e.,

$$\bigcap_{j=1}^{\infty} S^j = \{x\}$$

A partition process in a branch and bound algorithm is called exhaustive if every nested subsequence of partition sets generated throughout the algorithm is exhaustive. Konno, Thach, and Tuy (1997) proved that the above mentioned exact simplicial bisection is exhaustive since $\delta(S_k) \to 0$ as $k \to +\infty$.

2.2 Quadratic underestimation function for general non-convex structures

In the bounding step of a branch and bound algorithm, a lower bound is always obtained by constructing a valid convex underestimation function for the original one appeared in the problem (P), and solving the relaxed convex NLP to global optimality. For current simplex given by

$$S = \left\{ x \in \Re^n : x = \sum_{i=1}^{n+1} \lambda_i V^i, \lambda_i \geq 0, \sum_{i=1}^{n+1} \lambda_i = 1 \right\} \qquad (1)$$

where $V^i \in V \subset \Re^n$, $i = 1, 2, ..., n+1$ are $n+1$ vertices of the current simplex S, and V is the set of all its vertices. Then, we intend to compute a lower bound $\mu(S)$ of the objective function f on $S \cap D_g$. In other words, we compute a lower bound for the optimal value of the problem

$$(\ P(S)\ ) \qquad \min_{x} \quad f(x)$$

$$\text{s.t.} \quad g_i(x) \leq 0 \qquad i = 1, 2, ..., m$$

$$x \in S \subset \Re^n$$

As above mentioned, f and $\mathbf{g}_i$ are generic nonconvex functions belonging to C^2, then the main idea for computing a lower bound $\mu(S)$ is to construct from Problem (P(S)) a convex problem by replacing all those nonconvex functions with their respective convex underestimation functions, then solving the resulting relaxed convex problem. In order to reach this purpose, we see the following Definition:

Definition 2.2.1 *Given any nonconvex function* $f(\mathbf{x}): S \to \Re, \mathbf{x} \in S \subseteq \Re^n$ *belonging to* C^2*, the following quadratic function is defined by*

$$F(\mathbf{x}) = \sum_{i=1}^{n} a_i \mathbf{x}_i^2 + \sum_{i=1}^{n} b_i \mathbf{x}_i + c \qquad (2)$$

where, $\mathbf{x} \in \mathbf{S} \subseteq \Re^n$ *and* $F(\mathbf{x}) = f(\mathbf{x})$ *holds at all vertices of* **S**. a_i *'s are nonnegative scalars and large enough such that* $F(\mathbf{x}) \leq f(\mathbf{x}), \forall \mathbf{x} \in \mathbf{S}$.

It is trivial to see that $F(\mathbf{x})$ is convex since its quadratic coefficients, i.e. a_i's , are nonnegative. And the following Theorem can be used to ensure that it is indeed a rigorous underestimator of $f(\mathbf{x})$, i.e. $F(\mathbf{x}) \leq f(\mathbf{x}), \forall \mathbf{x} \in \mathbf{S}$.

Theorem 2.2.1 $F(\mathbf{x})$ *defined by Definition 2.2.1 is a convex underestimator of* $f(\mathbf{x})$ *if the difference function between them, i.e.* $D(\mathbf{x}) = F(\mathbf{x}) - f(\mathbf{x})$*, is a convex function.*

Proof. Suppose that $\mathbf{x}^1$ and $\mathbf{x}^2$ are two arbitrary points in the current simplex $\mathbf{S}$ defined by Eq.1, then there exists $2(n+1)$ real values, $\alpha_i, \beta_i \in \Re$

satisfying $0 \le \alpha_i, \beta_i \le 1$, $\sum_{i=1}^{n+1} \alpha_i = 1$, $\sum_{i=1}^{n+1} \beta_i = 1$, such that $\mathbf{x}^1 = \sum_{i=1}^{n+1} \alpha_i \mathbf{V}^i$ and $\mathbf{x}^2 = \sum_{i=1}^{n+1} \beta_i \mathbf{V}^i$. Since $D(\mathbf{x}) = F(\mathbf{x}) - f(\mathbf{x})$ is a convex function, we have the following inequality according to the definition of the convex function:

$$D(\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2) \le \lambda D(\mathbf{x}^1) + (1-\lambda) D(\mathbf{x}^2)$$

where, λ is an arbitrary real value, and $0 \le \lambda \le 1$. Substitution of the convex combination of $\mathbf{x}^1$ and $\mathbf{x}^2$ into above equation, and by Jensen's Inequality (Rochafellar, 1972) we have

$$\begin{aligned} D[\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2] &\le \lambda D\left(\sum_{i=1}^{n+1} \alpha_i \mathbf{V}^i\right) + (1-\lambda) D\left(\sum_{i=1}^{n+1} \beta_i \mathbf{V}^i\right) \\ &\le \lambda \sum_{i=1}^{n+1} \alpha_i D(\mathbf{V}^i) + (1-\lambda) \sum_{i=1}^{n+1} \beta_i D(\mathbf{V}^i) \end{aligned}$$

since $\sum_{i=1}^{n+1} \alpha_i = 1$ and $\sum_{i=1}^{n+1} \beta_i = 1$. According to Definition 2.2.1, we know that $F(\mathbf{x}) = f(\mathbf{x})$ holds at all vertices of S, i.e. $F(\mathbf{V}^i) = f(\mathbf{V}^i)$. Then $D(\mathbf{V}^i) = 0$ at each vertex $\mathbf{V}^i$, i=1, ..., $n+1$. Following above inequality, we have

$$D[\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2] \le 0$$

Since $\mathbf{x}^1$ and $\mathbf{x}^2$ are two arbitrary points in simplex $\mathbf{S}$, then let $\mathbf{x} = \lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2$, obviously it is also an arbitrary point in this simplex, and $D(\mathbf{x}) \le 0$. Then, $F(\mathbf{x}) \le f(\mathbf{x}), \forall \mathbf{x} \in \mathbf{S}$. It means that $F(\mathbf{x})$ is a rigorous underestimator of the generic nonconvex function $f(\mathbf{x})$ for any point $\mathbf{x} \in \mathbf{S}$.

It is well known that $D(\mathbf{x})$ is convex if and only if its Hessian matrix $H_D(\mathbf{x})$ is positive semi-definite in the current simplex. A useful convexity condition is derived by noting that $H_D(\mathbf{x})$ is related directly to the Hessian matrix $H_f(\mathbf{x})$ of $f(\mathbf{x})$, $\mathbf{x} \in \mathbf{S}$ by the following equation:

$$\mathbf{H}_D(\mathbf{x}) = 2\Delta - \mathbf{H}_f(\mathbf{x})$$

where Δ is a diagonal matrix whose diagonal elements are a_i's defined in Definition 2.2.1. Δ is refereed as the *diagonal underestimation matrix*, since these parameters guarantee that $F(\mathbf{x})$ defined by Eq.2 is a rigorous underestimator of the generic nonconvex function $f(\mathbf{x})$. Evidently, the following Theorem will help to guarantee that $D(\mathbf{x})$, as defined in Theorem 2.2.1, is convex:

Theorem 2.2.2 *$D(x)$, as defined in Theorem 2.2.1, is convex if and only if $2\Delta - H_f(\mathbf{x}) = 2diag(a_i) - H_f(\mathbf{x})$ is positive semi-definite for all $\mathbf{x} \in \mathbf{S}$.*

In order to simplify the parameter calculation, the underestimator $F(\mathbf{x})$ is reformulated by using a single nonnegative a value, as following

$$F(\mathbf{x}) = a\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} b_i x_i + c \qquad (3)$$

Then, all diagonal elements of the diagonal underestimation matrix Δ are therefore equal to the uniform quadratic coefficient a defined by Eq.3. On the basis of the Theorem 2.2.2, the following theorem can then be used to ensure that $F(\mathbf{x})$ defined by Eq.2 or Eq.3 is indeed a rigorous convex underestimator of $f(\mathbf{x})$:

Theorem 2.2.3. *$F(\mathbf{x})$ as defined by Eq.2 is a rigorous convex underestimator of $f(\mathbf{x})$ if and only if*

$$a_i \geq \max\left\{0, \frac{1}{2}\max_{\mathbf{x}\in\mathbf{S}}\left\{\mathbf{H}_{ii}^f(\mathbf{x}) + \sum_{j\neq i}\left|\mathbf{H}_{ij}^f(\mathbf{x})\right|\right\}\right\} \tag{4}$$

or, if $F(\mathbf{x})$ is defined by Eq.3, we have

$$a \geq \max\left\{0, \frac{1}{2}\max_{i,\mathbf{x}\in\mathbf{S}} \lambda_i(\mathbf{x})\right\} \tag{5}$$

where the $\lambda_i(\mathbf{x})$'s are the eigenvalues of $H_f(x)$, the Hessian matrix of the generic nonconvex function $f(\mathbf{x})$ for $\mathbf{x}\in\mathbf{S}$.

Proof: As $H_f(x)$, the Hessian matrix of the generic nonconvex function $f(\mathbf{x})$, is symmetric, so that all its eigenvalues are real values. According to Theorems 2.2.1 and 2.2.2, $F(\mathbf{x})$ as defined by Eq.2 is a convex (or linear) underestimator of $f(\mathbf{x})$ if and only if $D(\mathbf{x})$ defined in Theorem 2.2.1 is convex. $D(\mathbf{x})$ is convex if and only if for every $\mathbf{x}\in\mathbf{S}$, all eigenvalues $\lambda_i^D(\mathbf{x})$ of $D(\mathbf{x})$ are nonnegative.

In the second case, since the uniform quadratic coefficient is used, the eigenvalue of $D(\mathbf{x})$ can be directly related to that of $f(\mathbf{x})$. The above nonnegative condition is equivalent with requiring the minimum eigenvalue of $D(\mathbf{x})$ over $\boldsymbol{x}$ to be nonnegative:

$$\min_{i,\mathbf{x}\in\mathbf{S}} \lambda_i^D(\mathbf{x}) \geq 0$$

After substituting $\lambda_i^D(\mathbf{x}) = 2a - \lambda_i(\mathbf{x})$ and $a \geq \max\left\{0, \frac{1}{2}\max_{i,\mathbf{x}\in\mathbf{S}} \lambda_i(\mathbf{x})\right\}$, we have

$$\begin{aligned}\min_{i,\mathbf{x}\in\mathbf{S}} \lambda_i^D(\mathbf{x}) &= \min_{i,\mathbf{x}\in\mathbf{S}} (2a - \lambda_i(\mathbf{x})) \\ &\geq \min_{i,\mathbf{x}\in\mathbf{S}} \{\max[0, \max_{i,\mathbf{x}\in\mathbf{S}} \lambda_i(\mathbf{x})] - \lambda_i(\mathbf{x})\} \\ &\geq \min_{i,\mathbf{x}\in\mathbf{S}} \{\max_{i,\mathbf{x}\in\mathbf{S}} [0, \lambda_i(\mathbf{x})] - \lambda_i(\mathbf{x})\}\end{aligned}$$

Obviously, $\max_{\mathbf{x}\in\mathbf{S}}[0, \lambda_i(\mathbf{x})] - \lambda_i(\mathbf{x}) \geq 0$ by considering the two cases for the sign of $\lambda_i(\mathbf{x})$, so $\min_{\mathbf{x}\in\mathbf{S}} \lambda_i^D(\mathbf{x}) \geq 0$, that is, $D(\mathbf{x})$ is convex for $\mathbf{x}\in\mathbf{S}$. Therefore, $F(\mathbf{x})$ as defined by Eq.3 is a rigorous convex underestimator of $f(\mathbf{x})$.

In the first case, by virtue of Gerschgorin's theorem (Gerschgorin, 1931), the eigenvalue lower bound of a real symmetric matrix $A = (a_{ij})$ is given as

$$\min \lambda_i = a_{ii} - \sum_{j\neq i} |a_{ij}|$$

After substituting Eq.4 to $\mathbf{H}_D(\mathbf{x}) = 2\Delta - \mathbf{H}_f(\mathbf{x})$, its lower bound can be given as

$$\begin{aligned}\min_{\mathbf{x}\in\mathbf{S}} \lambda_i^D(\mathbf{x}) &= 2\max\left\{0, \frac{1}{2}\max_{\mathbf{x}\in\mathbf{S}}\left\{\mathbf{H}_{ii}^f(\mathbf{x}) + \sum_{j\neq i}\left|\mathbf{H}_{ij}^f(\mathbf{x})\right|\right\}\right\} - \mathbf{H}_{ii}^f(\mathbf{x}) - \sum_{j\neq i}\left|\mathbf{H}_{ij}^f(\mathbf{x})\right| \\ &\geq \max\left\{0, \max_{\mathbf{x}\in\mathbf{S}}\left\{\mathbf{H}_{ii}^f(\mathbf{x}) + \sum_{j\neq i}\left|\mathbf{H}_{ij}^f(\mathbf{x})\right|\right\}\right\} - \left\{\mathbf{H}_{ii}^f(\mathbf{x}) + \sum_{j\neq i}\left|\mathbf{H}_{ij}^f(\mathbf{x})\right|\right\} \\ &\geq \max_{\mathbf{x}\in\mathbf{S}}\left\{0, \left\{\mathbf{H}_{ii}^f(\mathbf{x}) + \sum_{j\neq i}\left|\mathbf{H}_{ij}^f(\mathbf{x})\right|\right\}\right\} - \left\{\mathbf{H}_{ii}^f(\mathbf{x}) + \sum_{j\neq i}\left|\mathbf{H}_{ij}^f(\mathbf{x})\right|\right\}\end{aligned}$$

Obviously, $\min_{\mathbf{x}\in\mathbf{S}} \lambda_i^D(\mathbf{x}) \geq 0$ by considering the two cases for the sign of

$\mathbf{H}_{ii}^{f}(\mathbf{x})+\sum_{j\neq i}\left|\mathbf{H}_{ij}^{f}(\mathbf{x})\right|$. So $D(\mathbf{x})$ is convex for $\mathbf{x}\in\mathbf{S}$. Therefore, $F(\mathbf{x})$ as defined by Eq.2 is a rigorous convex underestimator of $f(\mathbf{x})$.

The following Proposition shows the relationship between the linear and constant coefficients of $F(\mathbf{x})$ and its quadratic coefficients, and that the former ones can be determined by the latter and all vertices of the current simplex.

Proposition 2.2.1 *The linear and constant coefficients of* $F(\mathbf{x})$ *defined by Eq.2 or 3, i.e.* b_i*'s and* c *can be given by the quadratic coefficients* a_i's *known by Theorem 2.2.3 and the current simplex.*

Proof. In view of the Definition 2.2.1, we know $F(\mathbf{x})=f(\mathbf{x})$ holds at all vertices of ***S***, then the following linear equation group can be obtained as

$$\mathbf{V}^{k^T}\Delta\mathbf{V}+\mathbf{b}^T\mathbf{V}^k+c=f\left(\mathbf{V}^k\right) \qquad k=1,\ldots,n+1$$

where $\Delta\in\Re^{n\times n}$ is the diagonal underestimation matrix whose diagonal elements are the quadratic term coefficients, a_i's defined in Eq.2 or 3. $\mathbf{b}\in\Re^n$ is the linear coefficient matrix whose elements are b_i's defined in Eq.2 or 3, and c is a scalar.

$$\mathbf{b}^T\mathbf{V}^k+c=f\left(\mathbf{V}^k\right)-\mathbf{V}^{k^T}\Delta\mathbf{V} \qquad k=1,\ldots,n+1$$

The matrix $\mathbf{b}\in\Re^n$ is augmented as $(\mathbf{b},c)\in\Re^{n+1}$, in order to include the scalar c . In the same way, the matrix $\mathbf{V}\in\Re^{(n+1)\times n}$ is augmented as $(\mathbf{V},\mathbf{1})\in\Re^{(n+1)\times(n+1)}$, where $\mathbf{1}$ is a column unity matrix of $\Re^n$. $(\mathbf{V},\mathbf{1})\in\Re^{(n+1)\times(n+1)}$ is a regular square matrix since $\mathbf{V} \in \Re^{(n+1)\times n}$ is the coordinate matrix of the simplex which is linearly independent. Then we have

$$(\mathbf{b}, \mathrm{c})^T = (\mathbf{V}, \mathbf{1})^{-1}\left[f(\mathbf{V}) - \mathbf{V}^T \Delta \mathbf{V}\right]$$

where, $\left[f(\mathbf{V}) - \mathbf{V}^T \Delta \mathbf{V}\right] \in \Re^{n+1}$ is a column matrix for the $n+1$ vertices of the current simplex. In virtue of this equation, it is very obvious that the linear and constant coefficients defined by Eq.2 or 3 are determined completely by the quadratic coefficients and the current simplex.

By replacing all the nonconvex functions in Problem (P(S)) with their corresponding quadratic function based convex underestimators described by Eq.3, we have the following relaxed convex programming Problem (QP(S)):

$$(\text{ QP(S) }) \qquad \min_{\mathbf{x}} \quad \mathrm{F}(\mathbf{x})$$

$$\text{s.t.} \quad \mathbf{G}_i(\mathbf{x}) \le 0 \qquad i = 1, 2, \ldots, \mathrm{m}$$

$$x \in \mathbf{S} \subset \Re^n$$

where,

$$F(\mathbf{x}) = \sum_{i=1}^{n} a_i^f x_i^2 + \sum_{i=1}^{n} b_i^f x_i + c^f$$

$$\mathbf{G}_i(\mathbf{x}) = \sum_{i=1}^{n} a_i^{g_i} x_i^2 + \sum_{i=1}^{n} b_i^{g_i} x_i + c^{g_i} \qquad i = 1, 2, \ldots, \mathrm{m}$$

Let D_{G} be a subset of $\Re^n$ defined by

$$D_{\mathrm{G}} = \left\{x \in \Re^n : \mathbf{G}_i(\mathbf{x}) \le 0, \mathrm{i} = 1, 2, \ldots, \mathrm{m}\right\}$$

Obviously, the set D_{G} is convex and compact. Then, the Problem (QP(S)) has an optimal solution according to the well known Weierstrass Theorem.

It should be noted that only additional $m+1$ quadratic parameters, i.e. a^f and a^{g_i} for $i = 1, 2, \ldots, \mathrm{m}$, are introduced during above transforming process if the uniform underestimation function is used, since all other linear and constant coefficients can be calculated by those quadratic parameters and the current simplex consequently. The following Theorem states that the optimal solution F^* of the convex programming Problem (QP(S)) is a valid lower bound of the primal Problem (P(S)).

Theorem 2.2.4 *For each simplex* $\mathbf{S}=\left\{x\in\Re^n : x=\sum_{i=1}^{n+1}\lambda_i\mathbf{V}^i, \lambda_i\geq 0, \sum_{i=1}^{n+1}\lambda_i=1\right\}\subseteq\mathbf{S}^0$, *a lower bound* $\mu(\mathbf{S})$ *of* f *over* $\mathbf{S}\cap D_{\mathrm{g}}$ *can be computed by* $\mu(\mathbf{S})=F^*$, *where* F^* *is the optimal solution of* F *over* $\mathbf{S}\cap D_{\mathrm{G}}$.

Proof. First, we show $\mathbf{S}\cap D_{\mathrm{g}}\subseteq\mathbf{S}\cap D_{\mathrm{G}}$. Since $\mathbf{G}_i(\mathbf{x})$ is a convex underestimator of $\mathbf{g}_i(\mathbf{x})$, i.e. $\mathbf{G}_i(\mathbf{x})\leq\mathbf{g}_i(\mathbf{x})$, we have $\mathbf{G}_i(\mathbf{x})\leq\mathbf{g}_i(\mathbf{x})\leq 0$ for any $\mathbf{x}\in D_{\mathrm{g}}$, then $\mathbf{x}\in D_{\mathrm{G}}$. Finally we have $\mathbf{S}\cap D_{\mathrm{g}}\subseteq\mathbf{S}\cap D_{\mathrm{G}}$ by noting $D_{\mathrm{g}}\subseteq D_{\mathrm{G}}$. Second, by virtue of $F(\mathbf{x})\leq f(\mathbf{x})$ for any $\mathbf{x}\in\mathbf{S}\cap D_{\mathrm{g}}$ and $\mathbf{S}\cap D_{\mathrm{g}}\subseteq\mathbf{S}\cap D_{\mathrm{G}}$, we have

$$F^*=\min\{F(\mathbf{x}),\mathbf{x}\in\mathbf{S}\cap D_{\mathrm{G}}\}\leq F(\mathbf{x})\,\text{for}\,\mathbf{x}\in\mathbf{S}\cap D_{\mathrm{G}}\leq f(\mathbf{x})\,\text{for}\,\mathbf{x}\in\mathbf{S}\cap D_{\mathrm{g}}$$

It shows that $\mu(\mathbf{S})=F^*$ is a valid lower bound of f over $\mathbf{S}\cap D_{\mathrm{g}}$.

In the next Proposition, it shows that the lower bound obtained by Theorem 2.2.4 is always bounded from below and has a monotonic property which is useful within a branch and bound framework.

Proposition 2.2.2.

*(**a**) Let* $\mathbf{S}^1$ *and* $\mathbf{S}^2$ *be two simplexes satisfying* $\mathbf{S}^2\subset\mathbf{S}^1$. *Then,* $\mu(\mathbf{S}^2)\geq\mu(\mathbf{S}^1)$.

*(**b**) If Problem (P) has a feasible solution, then* $\mu(R)>-\infty$ *for each* $\mathbf{S}\subseteq\mathbf{S}^0$.

Proof.

(**a**) Let $F^1(\mathbf{x})$ and $F^2(\mathbf{x})$ be the quadratic underestimation functions of $f(\mathbf{x})$ generated in $\mathbf{S}^1$ and $\mathbf{S}^2$ satisfying $\mathbf{S}^2 \subset \mathbf{S}^1$, respectively. Then, we will show $F^1(\mathbf{x}) \le F^2(\mathbf{x})$ for $\mathbf{x} \in \mathbf{S}^2$. According to Eq.2, we have

$$F^1(\mathbf{x}) = \sum_{i=1}^{n} a_i^1 x_i^2 + \sum_{i=1}^{n} b_i^1 x_i + c^1$$

$$F^2(\mathbf{x}) = \sum_{i=1}^{n} a_i^2 x_i^2 + \sum_{i=1}^{n} b_i^2 x_i + c^2$$

Then,

$$F^1(\mathbf{x}) - F^2(\mathbf{x}) = \sum_{i=1}^{n} \left(a_i^1 - a_i^2\right) x_i^2 + \sum_{i=1}^{n} \left(b_i^1 - b_i^2\right) x_i + c^1 - c^2$$

It follows from the fact, i.e. $\mathbf{S}^2 \subset \mathbf{S}^1$, then $\max_{\mathbf{x}\in\mathbf{S}^1} \lambda_i(\mathbf{x}) \ge \max_{\mathbf{x}\in\mathbf{S}^2} \lambda_i(\mathbf{x})$ or $\max_{\mathbf{x}\in\mathbf{S}^1} \left\{\mathbf{H}_{ii}^f(\mathbf{x}) + \sum_{j\ne i} \left|\mathbf{H}_{ij}^f(\mathbf{x})\right|\right\} \ge \max_{\mathbf{x}\in\mathbf{S}^2} \left\{\mathbf{H}_{ii}^f(\mathbf{x}) + \sum_{j\ne i} \left|\mathbf{H}_{ij}^f(\mathbf{x})\right|\right\}$. In virtue of Theorem 2.2.3, we have $a_i^1 \ge a_i^2$. Then the difference function $D(\mathbf{x}) = F^1(\mathbf{x}) - F^2(\mathbf{x})$ is positive semi-definite.

Since $F^1(\mathbf{x})$ is the underestimation function of $f(\mathbf{x})$, then we have $F^1(\mathbf{x}) \le f(\mathbf{x})$ for all $\mathbf{x} \in \mathbf{S}^1$. According to the Definition 2.2.1, we know $F^2(\mathbf{V}_i^2) = f(\mathbf{V}_i^2)$ for all vertices of simplex $\mathbf{S}^2$, i.e. $\mathbf{V}_i^2$, for $i = 1, 2, \ldots, \mathrm{n}$. Since $\mathbf{S}^2 \subset \mathbf{S}^1$, we have $F^1(\mathbf{V}_i^2) \le F^2(\mathbf{V}_i^2)$ for $i = 1, 2, \ldots, \mathrm{n}$. It means

$$D(\mathbf{V}_i^n) = F^1(\mathbf{V}_i^2) - F^2(\mathbf{V}_i^2) \le 0 \qquad i = 1, 2, \ldots, \mathrm{n}$$

For any $\mathbf{x} \in \mathbf{S}^2$, and $\mathbf{x} = \sum_{i=1}^{n} \lambda_i \mathbf{V}_i^2$ with $\lambda_i \geq 0 \forall \mathrm{i}$ and $\sum_{i=1}^{n} \lambda_i = 1$, by the convex function characteristic of the difference function $D(\mathbf{x})$, we have

$$D(\mathbf{x}) = D\left(\sum_{i=1}^{n} \lambda_i \mathbf{V}_i^2\right) \leq \sum_{i=1}^{n} \lambda_i D(\mathbf{V}_i^2) \leq 0$$

Then, we obtain $F^1(\mathbf{x}) \leq F^2(\mathbf{x})$ for $\mathbf{x} \in \mathbf{S}^2$.

By the same way, for $\mathbf{x} \in \mathbf{S}^2$ we have

$$G_i^1(\mathbf{x}) \leq G_i^2(\mathbf{x}) \qquad \text{for } i = 1, 2, \ldots, \mathrm{m}$$

Then, we have

$$D_G^1 = \{\mathbf{x} \in \Re^n : G_i^1(\mathbf{x}) \leq 0, \mathrm{i} = 1, \ldots, \mathrm{m}\} \cap \mathbf{S}^1 \supseteq D_G^2 = \{\mathbf{x} \in \Re^n : G_i^2(\mathbf{x}) \leq 0, \mathrm{i} = 1, \ldots, \mathrm{m}\} \cap \mathbf{S}^2$$

Since $\mathbf{S}^2 \subset \mathbf{S}^1$, finally we have

$$\mu(\mathbf{S}^2) = \min\{F^2(\mathbf{x}) : \mathbf{x} \in D_G^2 \cap \mathbf{S}^2\} \geq \min\{F^1(\mathbf{x}) : \mathbf{x} \in D_G^1 \cap \mathbf{S}^1\} = \mu(\mathbf{S}^1)$$

(**b**) From (a), we need only to show that $\mu(\mathbf{S}^0) > -\infty$. This bounded property follows from the fact that the relaxed programming problem of Problem (P(S)) over the initial simplex $\mathbf{S}^0$, i.e. Problem (QP($\mathbf{S}^0$)) is convex. Then, this problem has an optimal solution, which implies that $\mu(\mathbf{S}^0) > -\infty$.

2.3 Upper bound

For a simplex $\mathbf{S}$, if the function value of Problem (P(S)) is unbounded from above, i.e. $\mu(\mathbf{S}) = +\infty$, then it follows that

$$f(\mathbf{x}) = +\infty \text{, for all } x \in S.$$

In this case, the partition set $\mathbf{S}$ can be removed from further consideration. Otherwise, one tries to find a set $F(\mathbf{S})$ of feasible solutions in $\mathbf{S}$ and uses it for computing an upper bound of the optimal value of Problem (QP(S)). Throughout the algorithm, more and more feasible solutions can be found, then the upper bound of the optimal value can be improved iteratively. A set $F(\mathbf{S})$ can be obtained by checking a finite set in $\mathbf{S}$ including, e.g. the set of all vertices and the center of the simplex $\mathbf{S}$, or some local solution of the Problem (P) over $\mathbf{S}$ by any convex optimizer. If all of them are infeasible, the current upper bound has to be kept until the new feasible set is found in the further iterations with new branches. It should be noted here, we ever assume that at least one feasible point is known, so that it can always be used until another one with less function value appears within the progress of the algorithm.

2.4 Rigorous calculation of the quadratic coefficients by using interval analysis

For generic nonconvex functions, the elements of its Hessian matrix $H_f(\mathbf{x})$ are likely to be nonlinear and nonconvex functions of variables, so that the derivation of the *diagonal underestimation matrix*, i.e. Δ, valid over the entire simplex is a very challenging task. However, satisfying the convexity condition of Theorem 2.2.2 is essential for the preservation of the guarantee that $F(\mathbf{x})$ defined by Eq.2 is a rigorous convex underestimator of the generic nonconvex function $f(\mathbf{x})$. The complexity arising from the presence of the variables in the convexity condition can become tractable by using the transformation of the exact $\mathbf{x}$-dependent Hessian matrix, i.e. $H_f(\mathbf{x})$, to an interval Hessian matrix $[H_f(\mathbf{x})]$ (Neumaier, 1990, 1996; Hansen, 1992; Kearfott, 1996; Adjiman et al., 1998a, b), such that $H_f(\mathbf{x}) \subseteq \lfloor H_f \rfloor$, $\forall \mathbf{x} \in \mathbf{S}$. $\mathbf{S}$, the current simplex, can be replaced with a more general interval box, described by $\lfloor \mathbf{x}^L, \mathbf{x}^U \rfloor$. $\mathbf{x}^L$ and $\mathbf{x}^U$ are the lower and upper bounds of the current simplex, respectively. Obviously, $\mathbf{S} \subseteq \lfloor \mathbf{x}^L, \mathbf{x}^U \rfloor$. Then the interval Hessian matrix can be calculated in above interval box, which will not influence the rigorousness of the estimation of the Hessian matrix, $H_f(\mathbf{x})$, in the current simplex. The elements of the original Hessian matrix, i.e. $H_f(\mathbf{x})$, are treated as independent when computing their general interval boundaries according to the interval arithmetic. The following Theorem will tell us how to use the interval Hessian matrix family $H_f(\mathbf{x})$ to calculate quadratic parameters a_i's defined by Eq.2:

Theorem 2.4.1*. Consider the generic nonconvex function $f(\mathbf{x})$ with continuous second-order derivatives and its Hessian matrix $H_f(\mathbf{x})$. Let $D(\mathbf{x}) = F(\mathbf{x}) - f(\mathbf{x})$ be defined in Theorem 2.2.1 and $F(\mathbf{x})$ be defined by Eq.2. Let $\lfloor H_f(\mathbf{x}) \rfloor$ be a symmetric interval matrix such that $H_f(\mathbf{x}) \subseteq \lfloor H_f \rfloor$, $\forall x \in \mathbf{S}$. If the matrix $[H_D]$ defined by $[H_D] = 2\Delta - \lfloor H_f \rfloor = 2diag(a_i) - \lfloor H_f \rfloor$ is positive semi-definite, then $D(\mathbf{x})$ is convex over the current simplex encompassed by $\lfloor \mathbf{x}^L, \mathbf{x}^U \rfloor$.*

Since the interval Hessian matrix $\lfloor H_f \rfloor \supseteq H_f(\mathbf{x})$ is obvious, then a valid lower bound of the maximum eigenvalue of $\lfloor H_f(\mathbf{x}) \rfloor$ can be more easily computed by using the interval arithmetic. Then, Eq.5 derived in Theorem 2.2.3 can be replaced with the following interval form, in order to generate a single a value which satisfies the following sufficient condition for that $F(\mathbf{x})$ is indeed a rigorous convex underestimator of $f(\mathbf{x})$:

$$a \geq \max\left\{0, \frac{1}{2}\lambda_{\max}\left(\lfloor H_f \rfloor\right)\right\} \quad (6)$$

where, $\lambda_{\max}\left(\lfloor H_f \rfloor\right)$ is the maximal eigenvalue of the interval matrix family $\lfloor H_f(\mathbf{x}) \rfloor$. For the non-uniform case, the Eq.4 can be transformed into the following equation by replacing the Hessian matrix with its interval form, as

$$a_i \geq \max\left\{0, \frac{1}{2}\left\{\bar{\mathbf{H}}^f_{ii} + \sum_{j \neq i} \left|\mathbf{H}^f\right|_{ij}\right\}\right\} \quad (7)$$

where $\left|\mathbf{H}^f\right|_{ij} = \max\left\{\left|\underline{\mathbf{H}}^f_{ij}\right|, \left|\bar{\mathbf{H}}^f_{ij}\right|\right\}$. Obviously Eq.7 holds since for interval matrix $\lfloor \mathbf{H}_f \rfloor$, we have $\bar{\mathbf{H}}^f_{ii} + \sum_{j \neq i} \left|\mathbf{H}^f\right|_{ij} \geq \lfloor \mathbf{H}^f_{ii} \rfloor + \sum_{j \neq i} \lfloor \mathbf{H}^f_{ij} \rfloor$. In the following sections, some commonly used favorable function structures and the generic nonconvex structure are analyzed in this interval way so as to get the tight convex underestimations for them in the current simplex.

2.4.1 Extended Gerschgorin's theorem for uniform case

For a real symmetric matrix $A = (a_{ij})$, the well-known Gerschgorin's theorem (Gerschgorin, 1931) states that its eigenvalues are bounded, such as $\lambda_{\max}$, by all its elements such that

$$\lambda_{\max} = \max_i \left(a_{ii} + \sum_{j \neq i} \left| a_{ij} \right| \right)$$

In this paper, a straightforward extension of this theorem is presented for interval matrices, as the following Theorem:

Theorem 2.4.2 *For an interval matrix* $[A]=\left(\underline{a}_{ij},\overline{a}_{ij}\right)$, *a lower bound on the maximum eigenvalue is given by*

$$\lambda_{\max} \geq \max_i \left[\overline{a}_{ii} + \sum_{j \neq i} \max\left(\left|\underline{a}_{ij}\right|, \left|\overline{a}_{ij}\right|\right)\right]$$

Proof: By definition of the interval matrix, $\lambda_{\max}([A]) \geq \max_{A\in[A]} \lambda_{\max}(A)$, therefore

$$\lambda_{\max}([A]) \geq \max_{A\in[A]} \max_i \left(a_{ii} + \sum_{j \neq i} \left|a_{ij}\right| \right)$$

$$\geq \max_i \left[\max_{A\in[A]}(a_{ii}) + \max_{A\in[A]}\left(\sum_{j \neq i} \left|a_{ij}\right| \right)\right]$$

$$\geq \max_i \left[\overline{a}_{ii} + \sum_{j \neq i} \max\left(\left|\underline{a}_{ij}\right|, \left|\overline{a}_{ij}\right|\right)\right]$$

Similar with that pointed out by Adjiman et al. (1998a) in their αBB algorithm for the estimation of the minimum eigenvalue of the interval matrix, above computational complexity is $O(n^2)$, then the bound it can provide on the eigenvalue is slightly loose. However, it is still very effective if the problem scale is not too large. For the practical applications, when the generic nonconvex function structures are given in analytical form, their interval Hessian matrix can be obtained by interval analysis, such as some widely used interval calculation packages, as INTLIB, a Portable FORTRAN77 Interval Standard Function Library (Kearfott, 1996), and PROFIL, Programmer's Runtime Optimized Fast Interval Library in C/C++ (Knuppel, 1993).

2.4.2 Underestimator for the convex (linear) function structure

For the convex (linear) function structures, denoted by $f^C(\mathbf{x})$ or $f^L(\mathbf{x})$, obviously their convex envelopes are themselves. Then, they will preserve their original forms in the final underestimators for the objection function or the constraints.

2.4.3 Underestimaor for the concave function structure

For the concave function structure, denoted by $f^{NC}(\mathbf{x})$, whose eigenvalues are all nonpositive, i.e. $\lambda_{i,\mathbf{x}\in\mathbf{S}}(\mathbf{x}) \le 0$. Then, the quadratic coefficient of its underestimator defined by Eq.2 is zero according to the Theorem 2.2.3, so that the valid lower bound of the concave function structure over the current simplex is a linear function whose linear and constant coefficients are given by Proposition 2.2.1. This conclusion is also completely consistent with that presented by Horst, Pardalos, and Thoai (1995, p.19). That is to say, the valid bound constructed by Eq.2 is equivalent to the convex envelope of the concave function over a simplex, which can be constructed by an affine function as given in the following Proposition:

Proposition 2.4.3 *Let* $\mathbf{S}$ *be a simplex generated by the vertices* $\mathbf{V}^0$, $\mathbf{V}^1$, ..., $\mathbf{V}^n$, *i.e.* $\mathbf{S} = \left\{ \mathbf{x} \in \Re^n : \mathbf{x} = \sum_{i=1}^{n+1} \lambda_i \mathbf{V}^i, \lambda_i \ge 0, \sum_{i=1}^{n+1} \lambda_i = 1 \right\}$, *and let* $f^{NC}(\mathbf{x})$ *be a concave function defined on* $\mathbf{S}$. *Then the convex envelope of* $f^{NC}(\mathbf{x})$ *over* $\mathbf{S}$ *is the affine function* $L^{NC}(\mathbf{x}) = \mathbf{b}^T\mathbf{x} + c$ *which is uniquely determined by the system of linear equations* $f^{NC}(\mathbf{V}^i) = \mathbf{b}^T\mathbf{V}^i + c$ *for* $i = 0, \ldots, n$.

For the practical calculation, the matrix $\mathbf{b} \in \Re^n$ is augmented as $(\mathbf{b}, c) \in \Re^{n+1}$, in order to include the scalar c . As that used in Proposition 2.2.1, the matrix $\mathbf{V} \in \Re^{(n+1)\times n}$ is augmented as $(\mathbf{V}, \mathbf{1}) \in \Re^{(n+1)\times(n+1)}$, where $\mathbf{1}$ is a column unity matrix of $\Re^n$. $(\mathbf{V}, \mathbf{1}) \in \Re^{(n+1)\times(n+1)}$ is a regular square matrix since $\mathbf{V} \in \Re^{(n+1)\times n}$ is the coordinate matrix of the simplex which is linearly independent. Then we have

$$(\mathbf{b}, c)^T = (\mathbf{V}, \mathbf{1})^{-1} \mathbf{f}^{NC}(\mathbf{V})$$

where, $\mathbf{f}^{NC}(\mathbf{V}) \in \Re^{n+1}$ is the column matrix for those $n+1$ vertices of above simplex.

2.4.4 Underestimaor for the general quadratic function

The general quadratically-constrained quadratic programming plays an important role in the engineering field. For an arbitrary bilinear function structure, denoted by $\mathbf{x}_i\mathbf{x}_j$ and $i \neq j$, McCormick (1976) and Al-Khayyal and Falk (1983) presented the tightest convex lower bound, i.e. convex envelope, over the rectangular domain $[\mathbf{x}_i^L, \mathbf{x}_i^U] \times [\mathbf{x}_j^L, \mathbf{x}_j^U]$. Here a valid convex underestimation function is easily derived for any general quadratic function, since the eigenvalues of its Hessian matrix is known. The general quadratic function is presented as

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{q}^T \mathbf{x}$$

Obviously, above bilinear structure is just a special case of this general function. Since $\mathbf{H}_f(\mathbf{x}) = \mathbf{Q}$, we have the *diagonal underestimation matrix,* Δ, constructed on the basis of Theorem 2.2.3, as

$$a = \max_i \left\{ 0, \frac{1}{2} \lambda_i^{\mathbf{Q}} \right\}$$

for the uniform case, or for the non-uniform case, we get

$$a_i = \max\left\{0, \frac{1}{2}\left(\mathbf{Q}_{ii} + \sum_{j \neq i} |\mathbf{Q}_{ij}|\right)\right\}$$

Then, we have the quadratic underestimation function as

$$F(\mathbf{x}) = \mathbf{x}^T \Delta \mathbf{x} + \mathbf{b}^T \mathbf{x} + c$$

where, the linear and constant coefficients, i.e. $(\mathbf{b}, c)$, can be computed by virtue of Proposition 2.2.1.

2.4.5 Underestimator for the tri-linear function structures

For the tri-linear function structure, denoted by $\mathbf{x}_i \mathbf{x}_j \mathbf{x}_k$ and $i \neq j \neq k$, Maranas and Floudas (1995) presented a valid convex lower bound over the rectangular domain $[\mathbf{x}_i^L, \mathbf{x}_i^U] \times [\mathbf{x}_j^L, \mathbf{x}_j^U] \times [\mathbf{x}_k^L, \mathbf{x}_k^U]$. Here, we give a valid quadratic convex lower bound over the simplex $\mathbf{S}$. The elements of the whole Hessian matrix of this tri-linear function structure over the simplex $\mathbf{S}$ are zero except for the rows and columns concerning of i, j , or k. Then, the eigenvalues of this Hessian matrix are all zero except for the ith, jth, and kth ones. In fact, these three eigenvalues can be computed on the basis of the following sub-Hessian matrix, as

$$\mathbf{H}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) = \begin{pmatrix} 0 & \mathbf{x}_k & \mathbf{x}_j \\ \mathbf{x}_k & 0 & \mathbf{x}_i \\ \mathbf{x}_j & \mathbf{x}_i & 0 \end{pmatrix}$$

whose interval sub-Hessian matrix can be computed as

$$[\mathbf{H}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k)] = \begin{bmatrix} \mathbf{0} & [\underline{\mathbf{x}}_k, \overline{\mathbf{x}}_k] & [\underline{\mathbf{x}}_j, \overline{\mathbf{x}}_j] \\ [\underline{\mathbf{x}}_k, \overline{\mathbf{x}}_k] & \mathbf{0} & [\underline{\mathbf{x}}_i, \overline{\mathbf{x}}_i] \\ [\underline{\mathbf{x}}_j, \overline{\mathbf{x}}_j] & [\underline{\mathbf{x}}_i, \overline{\mathbf{x}}_i] & \mathbf{0} \end{bmatrix}$$

Then, the lower bounds on the maximum eigenvalues can be computed by using Theorem 2.4.2, as follows

$$\lambda_i = |\mathbf{x}_k| + |\mathbf{x}_j|$$

$$\lambda_j = |\mathbf{x}_k| + |\mathbf{x}_i|$$

$$\lambda_k = |\mathbf{x}_i| + |\mathbf{x}_j|$$

where, $|a| = \max\left(|\underline{a}|, |\overline{a}|\right)$ for any interval $\left[\underline{a}, \overline{a}\right]$. Then, we get the rigorous quadratic coefficient as

$$a = \max(0, \lambda_1, \lambda_2, \lambda_3)$$

and the valid quadratic convex lower bound function can be written as

$$F^{TL}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) = a\mathbf{x}_i^2 + a\mathbf{x}_j^2 + a\mathbf{x}_k^2 + \sum_{l=1}^{n} \mathbf{b}_l \mathbf{x}_l + c$$

In virtue of Proposition 2.2.1, the linear and constant coefficients can be computed easily. It should be noted here, this valid underestimation function concerns of n linear coefficients rather than only ith, j th, and kth.

2.4.6 Underestimaor for the fractional function structures

For the fractional function structure, denoted by $\mathbf{x}_i / \mathbf{x}_j$ and $i \neq j$, Maranas and Floudas (1995) presented a valid convex lower bound over the rectangular domain $[\mathbf{x}_i^L, \mathbf{x}_i^U] \times [\mathbf{x}_j^L, \mathbf{x}_j^U]$. Here, we give a valid quadratic convex lower bound over the simplex $\mathbf{S}$. The sub-Hessian matrix of this fractional function structure is given as follows

$$H(\mathbf{x}_i, \mathbf{x}_j) = \begin{bmatrix} 0 & -\dfrac{1}{\mathbf{x}_j^2} \\ -\dfrac{1}{\mathbf{x}_j^2} & -\dfrac{2\mathbf{x}_i}{\mathbf{x}_j^3} \end{bmatrix}$$

Then, we have the characteristic determinant of this Hessian matrix as

$$|\mathbf{H}-\lambda\mathbf{I}|=\begin{bmatrix} -\lambda & -\dfrac{1}{\mathbf{x}_j^2} \\ -\dfrac{1}{\mathbf{x}_j^2} & -\dfrac{2\mathbf{x}_i}{\mathbf{x}_j^3}-\lambda \end{bmatrix}=\lambda^2+\frac{2\mathbf{x}_i}{\mathbf{x}_j^3}\lambda-\frac{1}{\mathbf{x}_j^4}=0$$

Since the discriminant of this binomial is nonnegative, as following

$$\Delta=\frac{4\left(\mathbf{x}_i^2+\mathbf{x}_j^2\right)}{\mathbf{x}_j^6}\geq 0$$

Then, two eigenvalues of above Hessian matrix are obtained as

$$\lambda=\pm\frac{\sqrt{\mathbf{x}_i^2+\mathbf{x}_j^2}}{|\mathbf{x}_j|^3}-\frac{\mathbf{x}_i}{\mathbf{x}_j^3}$$

Obviously, the maximal eigenvalue is

$$\lambda_{\max}=\frac{\sqrt{\mathbf{x}_i^2+\mathbf{x}_j^2}}{|\mathbf{x}_j|^3}-\frac{\mathbf{x}_i}{\mathbf{x}_j^3}$$

In this paper, we only consider the bounded case of the fractional function structure. Here, we assume either $\underline{\mathbf{x}}_j>0$ or $\bar{\mathbf{x}}_j<0$. For the first condition, i.e. $\underline{\mathbf{x}}_j>0$, we have

$$\lambda_{\max}=\frac{\sqrt{\mathbf{x}_i^2+\mathbf{x}_j^2}-\mathbf{x}_i}{\mathbf{x}_j^3}\geq 0$$

Then, the quadratic coefficient of the valid lower bound defined by Eq.3 over the simplex $\mathbf{S}$ is

$$a = \frac{\sqrt{\bar{\mathbf{x}}_i^2 + \bar{\mathbf{x}}_j^2} - \underline{\mathbf{x}}_i}{\underline{\mathbf{x}}_j^3} \geq 0$$

For the second condition, i.e. $\bar{\mathbf{x}}_j < 0$, we have

$$\lambda_{\max} = \frac{\sqrt{\mathbf{x}_i^2 + \mathbf{x}_j^2} + \mathbf{x}_i}{-\mathbf{x}_j^3} \geq 0$$

Therefore, the quadratic coefficient is computed as

$$a = \frac{\sqrt{\bar{\mathbf{x}}_i^2 + \underline{\mathbf{x}}_j^2} + \bar{\mathbf{x}}_i}{-\bar{\mathbf{x}}_j^3} \geq 0$$

So, the quadratic lower bound function can be written as

$$F^F(\mathbf{x}_i, \mathbf{x}_j) = a\mathbf{x}_i^2 + a\mathbf{x}_j^2 + \sum_{l=1}^{n} \mathbf{b}_l \mathbf{x}_l + c$$

According to Proposition 2.2.1, the n+1 linear and constant coefficients can be computed over the current simplex.

Remarks.

It should be noted here, that the relaxed convex programming Problem (QP(S)) contains not only the quadratic underestimation functions for the generic nonconvex terms, but also the convex function terms which are not necessarily transformed into the quadratic underestimators. Than, the final underestimation strategy of the relaxed Problem (QP(S)) can be slightly revised into the following convex programming formulation, as

$$(\ QP(S)'\ ) \qquad \min_{x} \quad F'(x)$$

$$\text{s.t.} \quad G'_i(x) \le 0 \qquad i = 1, 2, \ldots, m$$

$$x \in S \subset \Re^n$$

where,

$$F'(x) = f^L(x) + f^C(x) + L_f^{CA}(x) + F^{NC}(x)$$

$$G'_i(x) = g_i^L(x) + g_i^C(x) + L_{g_i}^{CA}(x) + G_i^{NC}(x) \qquad i = 1, 2, \ldots, m$$

and $f^L(x)$, $f^C(x)$, $L_f^{CA}(x)$, $g_i^L(x)$, $g_i^C(x)$, $L_{g_i}^{CA}(x)$ represent the linear terms, convex terms, and the linear underestimation functions for the concave terms in the objective function and the constraints, respectively. While $F^{NC}(x)$ and $G_i^{NC}(x)$ represent the quadratic convex underestimation functions for the generic nonconvex terms, and the special function structures, such as bilinear, trilinear, and fractional. Compared with the relaxed problem (QP(S)), the relaxed problem (QP(S)') contains not only quadratic function terms, but also the generic convex terms of the original problem. But, it should be noted here, such kind of relaxation does not affect the monotonicity of the valid convex underestimators given in Proposition 2.2.2, so it will also keep the algorithm convergences presented in the following sections.

2.5 Steps of the global optimization algorithm QBB

At the start of this section, the Problem (P) is formulated in an initial simplex S^0. However, the practical problem does not necessarily give that simplex, then a convenient outer approximation method of obtaining this simplex is presented here on a more broad basis, provided that the linear constrains can be separated from those with nonconvex terms, and the lower and upper bounds of the independent variables are known in a physical way, as follows

$$(\ P'\ ) \qquad \min_{x} \quad f(x)$$

$$\text{s.t.} \quad g_i(x) \le 0 \qquad i = 1, 2, \ldots, m$$

$$\mathbf{A}\mathbf{x} - \mathbf{b} \le \mathbf{0}$$

$$\underline{\mathbf{x}} \le \mathbf{x} \le \bar{\mathbf{x}}$$

where, $\underline{\mathbf{x}}$ and $\bar{\mathbf{x}}$ are the lower and upper bounds of $\mathbf{x}$. The polytope defined by the linear constraints are given as

$$P = \{\mathbf{x} \in \Re^n,\ \mathbf{A}\mathbf{x} - \mathbf{b} \le \mathbf{0}\}$$

In order to incorporate the lower and upper bounds of the variables into this polytope, the matrices $\mathbf{A}$ and $\mathbf{b}$ are expanded respectively as

$$\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{A} \\ \mathbf{1} \\ -\mathbf{1} \end{pmatrix} \text{ and } \tilde{\mathbf{b}} = \begin{pmatrix} \mathbf{b} \\ \bar{\mathbf{x}} \\ \underline{\mathbf{x}} \end{pmatrix}$$

where, $\mathbf{1}$ and $-\mathbf{1}$ are diagonal matrices with 1 and –1 as the diagonal elements, respectively. Then, we get the following polytope as

$$\tilde{P} = \left\{\mathbf{x} \in \Re^n,\ \tilde{\mathbf{A}}\mathbf{x} - \tilde{\mathbf{b}} \le \mathbf{0}\right\}$$

The following linear programming problems will help to generate the initial simplex $\mathbf{S}^0$ as small as possible, as

$$\mu_0 = \max\left\{\sum_{i=1}^{n} \mathbf{x}_j,\quad \mathbf{x} \in \tilde{P}\right\}$$

$$\mu_i = \min\left\{\mathbf{x}_i,\quad \mathbf{x} \in \tilde{P}\right\} \qquad i = 1, \ldots, n$$

Then, all n+1 vertices of the initial simplex can be computed by

$$\mathbf{V}^0 = (\mu_1, \ldots, \mu_n) \qquad (8)$$

$$\mathbf{V}^i = \left(\mu_1, ..., \mu_{i-1}, \mu_0 - \sum_{\substack{j=1 \\ j \neq i}}^{n} \mu_j, \mu_{i+1}, ..., \mu_n \right) \quad i = 1, ..., n \qquad (9)$$

Now, we are in a position to present the proposed algorithm for solving Problem (P) by using the basic operations described in previous sections.

Step 1 - Initialization. A convergence tolerance, ε_c, and a feasibility tolerance, ε_f, are selected and the iteration counter k is set to zero. The initial simplex $\mathbf{S}^0$ is computed by Eqs.8 and 9, as $S^0 = (\mathbf{V}^0, \mathbf{V}^1, ..., \mathbf{V}^n)$, and the current variable bounds $\underline{\mathbf{x}}$ and $\bar{\mathbf{x}}$ for the first iteration are set to be equal to the following linear programming problems, i.e. $\underline{\mathbf{x}}_i = \min\{\mathbf{x}_i, \mathbf{x} \in \mathbf{S}^0\}$ and $\tilde{\mathbf{x}}_i = \max\{\mathbf{x}_i, \mathbf{x} \in \mathbf{S}^0\}$ for $i = 1, ..., n$. The global lower and upper bounds μ_0 and γ_0 on the global minimum of Problem (P) are initialized and an initial current point $\mathbf{x}^{k,c}$ is randomly selected.

Step 2 - Local solution of Problem (P) and update of upper bound

The nonconvex and nonlinear optimization Problem (P) is solved locally within the current simplex $\mathbf{S}$. If the solution f_{local}^k of Problem (P) is ε_f-feasible, the upper bound γ_k is updated as $\gamma_k = \min(\gamma_k, f_{local}^k)$.

Step 3 – Partitioning of current simplex

The current simplex, $\mathbf{S}^k$, is partitioned into the following two simplexes ($r = 1, 2$):

$$\mathbf{S}^{k,1} = \left(\mathbf{V}^{k,0},\ldots,\mathbf{V}^{k,m},\ldots,\frac{\mathbf{V}^{k,m}+\mathbf{V}^{k,l}}{2},\mathbf{V}^{k,n} \right)$$

$$\mathbf{S}^{k,2} = \left(\mathbf{V}^{k,0},\ldots,\frac{\mathbf{V}^{k,m}+\mathbf{V}^{k,l}}{2},\ldots,\mathbf{V}^{k,l},\mathbf{V}^{k,n} \right)$$

where, k,m and k,l correspond to the vertices with the longest edge in the current simplex, i.e. $(k,m),(k,l) = \arg\max_{i<j}\left\{\left\|\mathbf{V}^{k,j}-\mathbf{V}^{k,i}\right\|\right\}$.

Step 4 – Update of $a^{r}_{k,f}$ and $a^{r}_{k,\mathbf{g}_i}$ inside both subsimplexes $r = 1,\ 2$

The nonnegative parameters $a^{r}_{k,f}$ and $a^{r}_{k,\mathbf{g}_i}$ for the general nonconvex terms in the objective function and constraints are updated inside both simplexes $r = 1,\ 2$ according to the methods presented in Section 2.4.

Step 5- Solutions inside both subsimplexes $r = 1,\ 2$

The convex programming Problem (QP(S)$'$) is solved inside both subsimplexes ($r = 1,\ 2$) by using any convex nonlinear solver. If a solution $F^{k,r}_{sol}$ is feasible and less than the current upper bound, γ_k, then it is stored along with the solution point $\mathbf{x}^{k,r}_{sol}$.

Step 6 – Update iteration counter k and lower bound μ_k

The iteration counter is increased by one,

$$k \leftarrow k+1$$

and the lower bound μ_k is updated to the minimum solution over the stored ones from the previous iterations. Furthermore, the selected solution is erased from the stored set.

$$\mu_k = F_{sol}^{k',r'}$$

where, $F_{sol}^{k',r'} = \min_{r,I}\{F_{sol}^{I,r},\ r = 1,2,\ I = 1,\ldots,k-1\}$. If the set I is empty, set $\mu_k = \gamma_k$ and go to **Step 8.**

Step 7 – Update current point $\mathbf{x}^{k,c}$ and current simplex $\mathbf{S}^k$

The current point is selected to be the solution point of the previously found minimum solution in **Step 6,**

$$\mathbf{x}^{k,c} = \mathbf{x}_{sol}^{I',r'}$$

and the current simplex becomes the subsimplex containing the previously found solution,

$$\mathbf{S}^k = \left(\mathbf{V}^{k',0},\ldots,\mathbf{V}^{k',m},\ldots,\frac{\mathbf{V}^{k',m}+\mathbf{V}^{k',l}}{2},\ldots,\mathbf{V}^{k',n}\right), \quad \text{if } r' = 1$$

$$\mathbf{S}^k = \left(\mathbf{V}^{k',0},\ldots,\frac{\mathbf{V}^{k',m}+\mathbf{V}^{k',l}}{2},\ldots,\mathbf{V}^{k',l},\ldots,\mathbf{V}^{k',n}\right), \text{ otherwise}$$

Step 8 – Check for convergence

If $(\gamma_k - \mu_k) > \varepsilon_c$, then return to **Step 2**. Otherwise, ε_c-convergence has been reached. The global minimum solution and solution point are given as:

$$f^* \leftarrow f^{c,k''}$$

$$\mathbf{x}^* \leftarrow \mathbf{x}^{c,k''}$$

where, $k'' = \arg_I \{ f^{c,I} = \gamma_k \}$ $I = 1, \ldots, k$.

Remarks:

(I). It should be noted that the current simplex can be deleted in **Step 5** when either the relaxed Problem (QP(S)') is infeasible or its solution is greater than the current upper bound. The former is obvious since Problem (P) is infeasible too if the relaxed Problem (QP(S)') is infeasible. The latter alternative is valid since the global minimum can not appear in this simplex for the lower bound computed over this simplex is greater than the current upper bound, which states that only local minima or some saddle points can exist there.

(II). In fact, the first condition in (I) for deletion of the infeasible subsimplex is very crucial for the algorithmic efficiency, since we always generate a much large initial simplex by the outer approximation method introduced in above section. But, with the division of this initial simplex, the branching step produces a large number of subsimplexes which are completely infeasible for the relaxed Problem (QP(S)') and then are definitely infeasible for the original problem (P(S)) over these subsimplexes. According to above Remark I, so that these subsimplexes can be removed in the algorithm immediately. This fact will be shown in Section 3 by solving a general quadratic programming problem.

The mathematical proof that the proposed global optimization algorithm QBB converges to the global minimum is presented in the following section.

2.6 Proof of convergence to the global minimum

If the QBB algorithm presented in above section terminates at iteration k, then the point $\mathbf{x}^k$ is an optimal solution of Problem (P). In the case that the algorithm is not finite, it generates at least one infinite sequence of simplexes $\{\mathbf{S}^j\}$ such that $\mathbf{S}^{j+1} \subset \mathbf{S}^j$, for all j. The convergence of the QBB algorithm is stated by means of the following results.

Proposition 2.6.1 *Assume that Problem (P) has a feasible solution. Further,* *assume that the QBB algorithm generates an infinite subsequence of simplexes* $\{\mathbf{S}^j\}$ *such that* $\mathbf{S}^{j+1} \subset \mathbf{S}^j$*, for all* j*, and* $\lim_{j\to\infty} \mathbf{S}^j = \bigcap_{j=1}^{\infty} \mathbf{S}^j = \{\mathbf{x}^*\}$. *Then,* $\mathbf{x}^*$ *is an optimal solution of Problem (P).*

Proof. First, we show that the point $\mathbf{x}^*$ is a feasible point of Problem (P). To do this, for each *j*, let $\mathbf{V}^j$ stand for a vertex of simplex $\mathbf{S}^j$. Further, for each *j*, let $(\mathbf{x}^j)$ be an optimal solution of the relaxed convex programming Problem (QP(S)) with $\mathbf{S} = \mathbf{S}^j$. It should be noted that $(\mathbf{x}^j)$ exists for each *j* as shown in Proposition 2.2.2(b). Since the edges of the simplex $\mathbf{S}^j$ are bounded and

$$\lim_{j\to\infty} \mathbf{S}^j = \bigcap_{j=1}^{\infty} \mathbf{S}^j = \{\mathbf{x}^*\}$$

Then, we also have

$$\lim_{j\to\infty} \mathbf{V}^j = \{\mathbf{x}^*\}$$

We can assume, by passing to subsequence if necessary, that $\mathbf{x}^j \to \mathbf{x}^*$, as $j \to \infty$. From this, we have

$$\mathbf{G}_i(\mathbf{x}^j) \to \mathbf{G}_i(\mathbf{x}^*) \le 0, \text{ for } i = 1, \ldots, m$$

Suppose that $\mathbf{x}^*$ is not a feasible solution of Problem (P); That is to say, there exists a number $\varepsilon > 0$ and for some constraint *k* such that

$$\mathbf{g}_k(\mathbf{x}^*) \ge \varepsilon > 0$$

Since $\mathbf{x}^*$ is a vertex at the limit simplex, then according to Definition 2.2.1, we have

$$\mathbf{g}_k(\mathbf{x}^*) = \mathbf{G}_k(\mathbf{x}^*) \ge \varepsilon > 0$$

which implies that $\mathbf{x}^*$ is not a feasible point to Problem (QP(P)). This contradiction implies that $\mathbf{x}^*$ is a feasible point of Problem (P).

Next, since $\mu(\mathbf{S}^{j+1}) \geq \mu(\mathbf{S}^j) > -\infty$, for all j, by Proposition 2.2.2, there exists a limit μ^* of $\{\mu(\mathbf{S}^j)\}$ bounded by the optimal value of Problem (P). Moreover, in view of the QBB algorithm, we have

$$\lim_{j \to \infty} \mu(\mathbf{S}^j) = \lim_{j \to \infty} \gamma(\mathbf{S}^j) \geq f(\mathbf{x}^*)$$

which implies that $\mathbf{x}^*$ is an optimal solution of Problem (P).

We observe that the accumulation point of the upper bound set also exists because of the compactness of the initial simplex $\mathbf{S}^0$, and is an optimal solution of the Problem (P), then Proposition 2.6.1 trivially leads to the following useful properties of the algorithm.

Proposition 2.6.2 *Assume that Problem (P) has a feasible solution, and that the simplicial partition process of the QBB algorithm presented in Section 2.1 is exhaustive, then the QBB algorithm has the following convergence properties:*

*(**a**) If the QBB algorithm generates an infinite subsequence of simplexes* $\{\mathbf{S}^j\}$ *such that the upper bound set* $F(\mathbf{S}^j) \neq \Phi$ *for each j, then each accumulation point of the corresponding subsequence* $\{\mathbf{x}^j\}$ *is an optimal solution of the Problem (P).*

*(**b**) The QBB algorithm terminates after finitely many iterations whenever the feasible set of Problem (P) is empty.*

For the proof of (a), we can see that a subsequence of the upper bound set exists with the limit as the optimal solution of the Problem (P). Moreover, if the QBB algorithm does not terminate after finitely many iterations, it must generate a subsequence of points converging to an optimal solution of Problem (P) by seeing the argument of Proposition 2.6.1. This contradiction implies that the QBB algorithm terminates finitely.

It is well known that the general nonconvex optimization problem is NP-hard (Vavasis, 1991). Then, we can expect that some large problems are difficult for the QBB algorithm. However, this definitely does not mean that the QBB algorithm is unable to solve the large problem in a reasonable amount of time. As we have described in the QBB algorithmic steps, it is possible to obtain a good feasible solution and show this feasible solution is within a specified tolerance of being optimal, esp. for the problems with some favorable structures. But, if we analyze the branch and bound tree structure generated from the partition process, the finite upper bound on the total number of required iterations for ε-convergence is exponential function of the initial simplex and the global convergence tolerance.

3. Computation studies of QBB algorithm for phase stability analysis

An example consisting of a nonconvex quadratic objective function subject to six inequality constraints all of which are nonconvex quadratic is used here to evaluate the algorithmic efficiency of the QBB. Since the quadratic coefficients of the underestimation function constructed in this paper for any bilinear term are known *a priori*, i.e. 0.5 or 0 which are shown later, then we can use some quadratic coefficients of the underestimation functions for the bilinear terms which are rigorously valid but appointed to be much greater than their accurate values obtained by the strict eigenvalue analysis in order to check the complicated situations where the accurate lower bound of the maximal eigenvalues of their interval Hessian matrix is difficult to be determined by any analytical methods. The problem formulation is shown as follows, where it has 10 inequality constraints representing the bounds on the five variables. This problem is taken from Colville's collection (1970) and also chosen by Floudas and Pardalos (1990) as a typical test for the constrained global optimization problem.

Min

$$37.293239\mathbf{x}_1 + 0.8356891\mathbf{x}_1\mathbf{x}_5 + 5.3578547\mathbf{x}_3^2 - 40792.141$$

s.t.

$$-0.0022053\mathbf{x}_3\mathbf{x}_5 + 0.0056858\mathbf{x}_2\mathbf{x}_5 + 0.0006262\mathbf{x}_1\mathbf{x}_4 - 6.665593 \le 0$$
$$0.0022053\mathbf{x}_3\mathbf{x}_5 - 0.0056858\mathbf{x}_2\mathbf{x}_5 - 0.0006262\mathbf{x}_1\mathbf{x}_4 - 85.334407 \le 0$$
$$0.0071317\mathbf{x}_2\mathbf{x}_5 + 0.00218133\mathbf{x}_3^2 + 0.0029955\mathbf{x}_1\mathbf{x}_2 - 29.48751 \le 0$$
$$-0.0071317\mathbf{x}_2\mathbf{x}_5 - 0.00218133\mathbf{x}_3^2 - 0.0029955\mathbf{x}_1\mathbf{x}_2 + 9.48751 \le 0$$
$$0.0047026\mathbf{x}_3\mathbf{x}_5 + 0.0019085\mathbf{x}_3\mathbf{x}_4 + 0.0012547\mathbf{x}_1\mathbf{x}_3 - 15.699039 \le 0$$
$$-0.0047026\mathbf{x}_3\mathbf{x}_5 - 0.0019085\mathbf{x}_3\mathbf{x}_4 - 0.0012547\mathbf{x}_1\mathbf{x}_3 + 10.699039 \le 0$$
$$78 \le \mathbf{x}_1 \le 102$$
$$33 \le \mathbf{x}_2 \le 45$$
$$27 \le \mathbf{x}_3 \le 45$$
$$27 \le \mathbf{x}_4 \le 45$$
$$27 \le \mathbf{x}_5 \le 45$$

The nonlinearities in above problem arise from the bilinear terms $\pm\mathbf{x}_i\mathbf{x}_j$ for $i \ne j$, and $-\mathbf{x}_i^2$, in the cost and constrained functions. For the latter one, since this bilinear term belongs to the concave function structure, then its convex envelope described in Proposition 2.4.3. being an affine function over the current simplex can be easily constructed. For the former bilinear term, we can see its Hessian matrix over the current simplex is a constant matrix as following:

$$\mathbf{H} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ or } \mathbf{H} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

whose two eigenvalues are 1 and –1, respectively. According to the analysis in Section 2.4.4 for the general quadratic function, we get the unified quadratic underestimation coefficient, i.e. $a = 0.5$. Consequently, the linear and constant quadratic underestimation function can be computed by Proposition 2.2.1 over the current simplex. After all the nonconvex bilinear terms are replaced by their quadratic underestimation functions, the valid underestimation functions for the cost and constrained functions in above problem are obtained. Then, a convex programming problem is obtained which can be solved by any convex optimizer in order to locate a valid lower bound for the original function over the current simplex.

Min

$$37.293239\mathbf{x}_1 + 0.8356891\times\left(0.5\mathbf{x}_1^2 + 0.5\mathbf{x}_5^2 + \sum_{i=1}^{5} b_i^{1,5}\mathbf{x}_i + c^{1,5}\right) + 5.3578547\mathbf{x}_3^2 - 40792.141$$

s.t.

$$0.0022053\times\left(0.5\mathbf{x}_3^2 + 0.5\mathbf{x}_5^2 + \sum_{i-1}^{5} b_i^{-3,5}\mathbf{x}_i + c^{-3,5}\right) + 0.0056858\times\left(0.5\mathbf{x}_2^2 + 0.5\mathbf{x}_5^2 + \sum_{i-1}^{5} b_i^{2,5}\mathbf{x}_i + c^{2,5}\right) +$$
$$0.0006262\times\left(0.5\mathbf{x}_1^2 + 0.5\mathbf{x}_4^2 + \sum_{i-1}^{5} b_i^{1,4}\mathbf{x}_i + c^{1,4}\right) - 6.665593 \le 0$$

$$0.0022053\times\left(0.5\mathbf{x}_3^2 + 0.5\mathbf{x}_5^2 + \sum_{i-1}^{5} b_i^{3,5}\mathbf{x}_i + c^{3,5}\right) + 0.0056858\times\left(0.5\mathbf{x}_2^2 + 0.5\mathbf{x}_5^2 + \sum_{i-1}^{5} b_i^{-2,5}\mathbf{x}_i + c^{-2,5}\right) +$$
$$0.0006262\times\left(0.5\mathbf{x}_1^2 + 0.5\mathbf{x}_4^2 + \sum_{i-1}^{5} b_i^{-1,4}\mathbf{x}_i + c^{-1,4}\right) - 85.334407 \le 0$$

$$0.0071317\times\left(0.5\mathbf{x}_2^2 + 0.5\mathbf{x}_5^2 + \sum_{i=1}^{5} b_i^{2,5}\mathbf{x}_i + c^{2,5}\right) + 0.00218133\mathbf{x}_3^2 +$$
$$0.0029955\times\left(0.5\mathbf{x}_1^2 + 0.5\mathbf{x}_2^2 + \sum_{i=1}^{5} b_i^{1,2}\mathbf{x}_i + c^{1,2}\right) - 29.48751 \le 0$$

$$0.0071317\times\left(0.5\mathbf{x}_2^2 + 0.5\mathbf{x}_5^2 + \sum_{i=1}^{5} b_i^{-2,5}\mathbf{x}_i + c^{-2,5}\right) + 0.00218133\times\left(\sum_{i=1}^{5} b_i^{-3,3}\mathbf{x}_i + c^{-3,3}\right) +$$
$$0.0029955\times\left(0.5\mathbf{x}_1^2 + 0.5\mathbf{x}_2^2 + \sum_{i=1}^{5} b_i^{-1,2}\mathbf{x}_i + c^{-1,2}\right) + 9.48751 \le 0$$

$$0.0047026\times\left(0.5\mathbf{x}_3^2 + 0.5\mathbf{x}_5^2 + \sum_{i=1}^{5} b_i^{3,5}\mathbf{x}_i + c^{3,5}\right) + 0.0019085\times\left(0.5\mathbf{x}_3^2 + 0.5\mathbf{x}_4^2 + \sum_{i=1}^{5} b_i^{3,4}\mathbf{x}_i + c^{3,4}\right) +$$
$$0.0012547\times\left(0.5\mathbf{x}_1^2 + 0.5\mathbf{x}_3^2 + \sum_{i=1}^{5} b_i^{1,3}\mathbf{x}_i + c^{1,3}\right) - 15.699039 \le 0$$

$$0.0047026\times\left(0.5\mathbf{x}_3^2+0.5\mathbf{x}_5^2+\sum_{i=1}^{5}b_i^{-3,5}\mathbf{x}_i+c^{-3,5}\right)+0.0019085\times\left(0.5\mathbf{x}_3^2+0.5\mathbf{x}_4^2+\sum_{i=1}^{5}b_i^{-3,4}\mathbf{x}_i+c^{-3,4}\right)+$$
$$0.0012547\times\left(0.5\mathbf{x}_1^2+0.5\mathbf{x}_3^2+\sum_{i=1}^{5}b_i^{-1,3}\mathbf{x}_i+c^{-1,3}\right)+10.699039\le 0$$

$$\sum_{i=1}^{5}b_i^{k,j}\mathbf{x}_i+c^{k,j}\le 0 \qquad j=1,\dots,6 \tag{10}$$

$$78\le \mathbf{x}_1\le 102$$
$$33\le \mathbf{x}_2\le 45$$
$$27\le \mathbf{x}_3\le 45$$
$$27\le \mathbf{x}_4\le 45$$
$$27\le \mathbf{x}_5\le 45$$

where, $b^{i,j}$ and $c^{i,j}$ represent the linear and constant coefficients of the quadratic underestimation functions generated according to Proposition 2.2.1 for the bilinear terms $\mathbf{x}_i\mathbf{x}_j$ for $i\neq j$, but $b^{-i,j}$ and $c^{-i,j}$ are for those bilinear terms $-\mathbf{x}_i\mathbf{x}_j$ including $i=j$. In Eq.10, the six inequalities are presented to described the current subsimplex, where $b_i^{k,j}$ and $c^{k,j}$ represent the linear and constant coefficients of the *j*th superplane over the subsimplex $\mathbf{S}^k$. Since all vertices of the current simplex are known, then these *N*+1 coefficients can be obtained by solving a linear equation group, and the final signs of these coefficients are determined by the fact that the constrained space lies inside the current simplex (Zhu and Xu, 1999).

In this paper, the NLP optimizer LSGRG2C (Smith and Lasdon, 1992; Lasdon, 2000) is used to solve each convex underestimation problem over the current simplex within the QBB algorithmic framework. The final package cQBB is implemented in C language and also used for some nonconvex optimization problems of chemical and phase equilibria in our form papers (Zhu and Xu, 1999; Zhu and Inoue, 2001). For above mentioned generally quadratical programming problem, all the computational runs by cQBB package were performed on a Pentium III/800 machine. In this paper, all CPU times reported represent the total time taken to solve above problem with different valid quadratic coefficients by QBB algorithm, where the global convergence is

0.001 and the feasible tolerance is 0.001. The initial simplex is calculated by the outer approximation method to the 10 linear inequality constraints representing the bounds on the five variables, given as { {78.0,33.0,27.0,27.0,27.0}, {168.0,33.0,27.0,27.0,27.0}, {78.0,123.0,27.0,27.0,27.0}, {78.0,33.0,117.0, 27.0,27.0}, {78.0,33.0,27.0,117.0,27.0}, {78.0,33.0,27.0,27.0,117.0}}. Obviously, this simplex is much larger than the hypercube for the five variables in the original problem, i.e. { [78,102], [33,45], [27,45], [27,45], [27,45] } which is fully contained in above simplex. The calculation results are shown in Tables 1, 2, and 3 where the quadratic coefficient of the underestimation function for any bilinear terms $\pm \mathbf{x}_i \mathbf{x}_j$ for $i \neq j$ in above problem is assigned to 0.5, i.e. the accurate one, 1.0, and 2.0, respectively. The CPU running time increases and the solution quality deteriorates when the quadratic coefficient is estimated loosely. However, the algorithmic convergence is guaranteed even when the quadratic coefficient is assigned to be four times of the accurate one, see in Table 3. It should be noted that the final number of the unfathomed simplexes is zero for all assigned quadratic coefficients, since this constrained problem has only one global solution and the infeasible subsimplexes and those containing only local minima have been deleted with the algorithm progress, as the Remark I states in Section III.

4. Conclusion

A quadratic underestimation function based branch and bound algorithm, QBB, is developed to solve problems belonging to the broad class of twice-differentiable NLPs. For any such problem, the ability to generate progressively tighter convex lower bounding problems at each iteration guarantees the convergence of the QBB algorithm to within epsilon of the global optimum solution under the exhaustive division framework of the initial simplex. The different methods are presented for the construction of the convex valid underestimators for special function structures and the general nonconvex function structures, and the maximal eigenvalue analysis of the interval Hessian matrix provides the rigorous guarantee for the QBB algorithm to converge to the global solution. The convergence properties of the QBB algorithm for the nonconvex problems are obtained, and some results of the computational experiments for a general quadratic programming problem is reported to show the capacity of the QBB algorithm for the practical applications.

Acknowledgement

The financial support from JSPS for this Project and Y. Zhu is gratefully appreciated. Prof. Nguyen V. Thoai provides Y. Zhu with invaluable suggestions for his further understanding of the optimal theory, and his excellent course of mathematical programming presented during his stay in the University of Tsukuba is an inspiration for Y. Zhu's following work in the field of global optimization.

References

Adjiman, C.S., Dallwig, S., Floudas, C.A., & Neumaier, A. (1998a). A global optimization method, αBB, for general twice-differentiable constrained NLPs - I. Theoretical advances", *Comput.Chem.Eng.*, *22*, 1137-1158.

Adjiman, C.S., Androulakis, I.P., & Floudas, C.A. (1998b). A global optimization method, αBB, for general twice-differentiable constrained NLPs - II. Implementation and computational results, *Comput.Chem.Eng.*, *22*, 1159-1179.

Al-Khayyal, F.A., & Falk, J.E. (1983). Jointly constrained biconvex programming. *Math. of Oper. Res.*, 8, 273-286.

Colville, A.R. (1970). A comparative study of nonlinear programming codes. In Kuhn, H.W., editor, *Princeton Symposium on Mathematical Programming*. Princeton Univ. Press.

Floudas, C.A., Aggarwal, A., & Ciric, A.R. (1989). Global optimum search for non-convex NLP and MINLP problems. *Comput.Chem.Eng.*, 13, 1117-1132.

Floudas, C.A., & Visweswaran, V. (1990). A global optimization algorithm (GOP) for certain classes of nonconvex NLPs: I. Theory. *Comput.Chem.Eng.*, *14*, 1397-1417.

Floudas, C.A. & Pardalos, P.M. (1990). *A collection of test problems for constrained global optimization algorithms*, Lecture Notes in Computer Science 455, Springer-Verlag, Berlin.

Floudas, C.A., & Visweswaran, V. (1993). A primal-relaxed dual global optimization approaches. *J.Optim.Theory Appl.*, *78*, 187-225.

Floudas, C.A. (1999). *Deterministic global optimization: theory, methods and* *applications*. Kluwer Academic Publishers. Dordrecht.

Gerschgorin, S. (1931). Uber die abgrenzung der eigenwerte einer matrix. *Izv.Akad. Nauk SSSR, Ser.fiz.-mat.*, *6*, 749-754.

Grossmann, I.E., editor (1996). *Global optimization in engineering design*, Kluwer book series in nonconvex optimization and its applications, Kluwer Academic Publishers. Dordrecht.

Hansen, E.R. (1992). *Global optimization using interval analysis*. Marcel Dekker, New York.

Horst, R., & Pardalos, P.M. (1995). *Handbook of global optimization*. Kluwer Academic Publishers. New York.

Horst, R., Pardalos, P.M., & Thoai, N.V. (1995). *Introduction to global optimization*. Kluwer Academic Publishers, Dordrecht.

Kearfott, R.B. (1996). *Rigorous global search: Continuous problems. Kluwer Academic Publishers*, Dordrecht, the Netherlands.

Knuppel, O. (1993). *Programmer's Runtime Optimized Fast Interval Library. Technische Informatik III*, Technische Universitat Hamburg-Harburg.

Konno, H., Thach, P.T., & Tuy, H. (1997). *Optimization on low rank nonconvex structures, Nonconvex optimization and its applications*, Kluwer Academic Publishers, Dordrecht.

Lasdon, L. (2000). *LSGRG Version 3.0 Release Notes*. MSIS Department, College of Business Administration, the University of Texas at Austin.

Maranas, C.D., & Floudas, C.A. (1995). Finding all solutions of nonlinearly constrained systems of equations. *J.Global Optim.*, *7*, 143-182.

McCormick, G.P. (1976). Computability of global solutions to factorable nonconvex programs: Part I – Convex underestimating problems. *Math. Programming, 10,* 147-175.

McDonald, C.M., & Floudas, C.A. (1994). Decomposition based and branch and bound global optimization approaches for the phase equilibrium problems. *J.Global Optim.*, *5*, 205-251.

Neumaier, A. (1990). *Interval methods for systems of equations*. Cambridge Univ. Press, Cambridge, England.

Neumaier, A. (1996). Second-order sufficient optimality conditions for local and global nonlinear programming. *J.Global Optim.*, *9*, 141-151.

Rockafellar, R.T. (1972). *Convex analysis*, Princeton Univ. Press, Princeton.

Smith, S. & L. Lasdon, (1992). Solving Large Sparse Nonlinear Programs Using GRG, *ORSA Journal on Computing*, *4(1)*: 2-15.

Stephanopoulos, G., & Westerberg, A.W. (1975). The use of Hestenes' method of multipliers to resolve dual gaps in engineering system optimization. *Journal of Optimization Theory and Application*, 15, 285.

Vavasis, S. (1991). *Nonlinear optimization: Complexity issues*. Oxford University Press. New York, N.Y.

Wales D.J., & Scheraga, H.A. (1999). Global optimization of clusters, crystals, and Biomolecules, *Science*. 285, 1368-1372.

Westerberg, A.W., & Shah, J.V. (1978). Assuring a global optimum by the use of an upper bound on the lower (dual) bound. *Comput. Chem. Engng.*, 2, 83.

Zhu, Y., & Xu, Z. (1999). Calculation of liquid-liquid equilibrium based on the global stability analysis for ternary mixtures by using a novel branch and bound algorithm: Application to UNIQUAC equation. *Ind. Eng. Chem. Res.*, 38, 3549-3556.

Zhu, Y., & Inoue, K. (2001), Calculation of chemical and phase equilibrium based on stability analysis by QBB algorithm: Application to NRTL equation. *Chem. Eng. Sci.*, (in press)

Tables

Table 1. Calculation results of QBB algorithm for Colville's Problem when $a = 0.5$

Variable x_i	Upper bound	Upper bound solution	Lower bound	Lower bound solution	Iteration number	No. of unfathomed subsimplexes	CPU time (s)
1		78.0		78.0			
2		33.0		33.0			
3	-30665.58848	29.99506	-30665.60118	29.99503	1090	0	205.97
4		45.0		45.0			
5		36.77602		36.77601			

Table 2. Calculation results of QBB algorithm for Colville's Problem when $a = 1.0$

Variable x_i	Upper bound	Upper bound solution	Lower bound	Lower bound solution	Iteration number	No. of unfathomed subsimplexes	CPU time (s)
1		78.0		78.0			
2		33.0		33.0			
3	-30665.58848	29.99506	-30665.77273	29.99445	2078	0	422.10
4		45.0		45.0			
5		36.77602		36.77628			

Table 3. Calculation results of QBB algorithm for Colville's Problem when $a = 2.0$

Variable x_i	Upper bound	Upper bound solution	Lower bound	Lower bound solution	Iteration number	No. of unfathomed subsimplexes	CPU time (s)
1		78.0		78.0			
2		33.0		33.0			
3	-30665.58848	29.99506	-30665.68263	29.99471	4472	0	960.54
4		45.0		45.0			
5		36.77602		36.77646			