Global optimization of nonconvex MINLP by a hybrid branch-and-bound and revised general Benders decomposition method

Yushan Zhu* Takahito Kuno

February 4, 2002

ISE-TR-02-186

Institute of Information Sciences and Electronics, University of Tsukuba, Ibaraki 305-8573, Japan
Institute of Process Engineering, Chinese Academy of Sciences, Beijing 100080, China
*** Tel: +81-298-536791, Fax: +81-298-535206, e-mail: yszhu@syou.is.tsukuba.ac.jp**

Key words. Global optimization, MINLP, branch-and-bound, General Benders Decomposition, QBB, Process synthesis and design

Global optimization of nnconvex MINLP by a hybrid branch-and-bound and revised general Benders decomposition approach

Yushan Zhu* Takahito Kuno

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

Abstract

The mixed-integer nonlinear programming, MINLP, has played a crucial role for the chemical process design via superstructure that always involves discrete and continuous variables. In this paper, a global optimization algorithm for the nonconvex MINLP problem is developed by tackling the nonconvexity caused by the nonconvex continuous functions on the basis of the convex quadratic underestimator within a branch and bound framework, as well as the joint one caused by the mixed natures of integer and continuous variables by virtue of a revised General Benders Decomposition (GBD) method, where the latter is designed mainly for three favorable structures, i.e. separable, bilinear, and partly linear between the two domains of continuous and binary variables. The convergence of the revised GBD method on the global solution of the relaxed MINLP subprobelm over each subregion generated in above framework is guaranteed by the convex underestimation functions in terms of the twice differentiable assumptions of the continuous functions and above three favorable joint structures, then the convergence of the proposed hybrid algorithm can be established by the exhaustive partition of the constrained region, the monotonicity of lower bound, and the reliable infeasibility detecting. Finally, a very simple example for process design is used to verify the different implementation aspects of the proposed approach, especially the unique underestimator construction and the infeasibility detecting in each lower bounding problem.

Keywords. *Global optimization, MINLP, branch and bound, General Benders Decomposition, QBB, process synthesis and design.*

Corresponding author. Tel: +81-298-536791, Fax: +81-298-535206, e-mail: yszhu@syou.is.tsukuba.ac.jp, is a JSPS Fellow, and both authors were supported by Grant-in-Aid for Scientific Research of JSPS.

1. Introduction

Companies must design and operate chemical processes effectively and efficiently so they may survive in today's highly competitive world. Providing the methods, tools, and people that allow industry to meet its needs by trying science to engineering is a compelling aspect of process system engineering, which is concerned with the improvement of decision-making processes for the creation and operation of the chemical supply chain ranging from microsystems to industrial-scale continuous and batch systems. It deals with the discovery, design, manufacture, and distribution of chemical products in the context of many conflicting goals (Grossman and Westerberg, 2000). In general, the conventional chemical process engineering has traditionally focused on the macro level. However, with strong economical forces driving the need for product specialization and differentiation, there is a significant incentive for the development of synthesis and optimization tools to aid in the discovery and design of new products. For instance, the decision-making processes that take place during the design of new products or chemical plants can be made more rational and efficient thanks to the use of mathematical model within a global optimization framework (Grossman, 1996; Floudas, 2000).

The most significant constribution of mathematical approaches comes from their ability to incorporate many alternative structures within a single problem. This is achieved through the introduction of integer variables, which leads to the formulation of a mixed-integer nonlinear programming-MINLP (Floudas, 1995). Such an approach has already been used for a wide array of applications, such as the process synthesis of heat exchanger network, distillation-based separation systems, reactor network, and reactor-separator- recycle systems. The solution of many MINLPs relevant to the current science-based chemical processes is made challenging not only by the presence of integer variables but also by the nonconvexities in the models. As a result, the potential contributions of MINLP to above-mentioned problems have not yet been fully realized. The global solutions of the mixed-integer linear programming-MILP (Nemhauser and Wolsey, 1999) problems or convex MINLP problems can be located by the Benders decomposition method (Benders, 1962) or general Benders decomposition method (Geoffrion, 1972). However, those approaches cannot be applied directly into the nonconvex MINLP problems since they always just identify a local optimum owing to the nonconvexity of the nonlinear functions and the joint structure. The earlier endeavors to solve the nonconvex MINLP problems were contributed from Kocis and Grossmann (1988) and Floudas, Aggarwal, and Ciric (1989) stimulated by the MINLP problems encountered in the process synthesis and design. The branch-and-reduce algorithm of Ryoo and Sahinidis (1995), and finally developed into a package named BARON, relies on the existing underestimation techniques, such as those proposed by McCormick (1976), and focuses on the reduction of the size of the solution domain by using the addition of feasibility and optimality tests. The interval analysis algorithm of Vaidyanathan and El-Halwagi (1996) used the interval arithmetic to bound the function values within a branch and bound framework, where the domain size is reduced by partitioning, and fathoming is performed by applying upper bound, infeasibility, monotonicity, nonconvexity, and lower bound tests, as well as the distrust region method. Smith and Pantelides (1997, 1999) designed a reformulation spatial branch and bound algorithm to address functions that involve binary arithmetic operators and concave or convex operators such as logarithms and exponentials. Westerlund et al. (1998) used a extended cutting plane algorithm to tackle problems involving pseudoconvex functions. Zamora and Grossman (1998) proposed more specialized algorithms for certain classes of applications, such as heat-exchanger networks. Adjiman et al. (1998) presented an excellent review of these algorithms, and two broadly applicable global optimization approaches based on the αBB algorithm, i.e. SMIN-αBB and GMIN-αBB, were briefly introduced in Adjiman et al. (1997). A complete description of the theoretical basis of these two algorithms and computational experiments are provided in Floudas (2000) and Adjiman et al. (2000) which enable the determination of the most adequate implementation decisions.

A novel convex underestimation technique developed in the QBB algorithm (Zhu and Xu, 1999; Zhu and Inoue, 20001; Zhu and Kuno, 2001), is applied here to tackle the nonconvexities arisen by the continuous variables, then the resulted convex mixed-integer programming is resolved by a revised General Benders Decomposition-GBD method. The so called hybrid branch and bound and revised GBD algorithm for the nonconvex MINLP problems goes from the simpilical partition of the constrained region of the continuous variables within a branch and bound framework, then the mixed natures of the continuous and the binary variables are treated elaborately by the projection way in the GBD method. The monotonicity of the lowing bounding functions constructed by the quadratic function based underestimators, the infeasibility detecting, and asymptotic convergence are presented to provide a complete theoretical guarantee. Three kinds of mixed-integer function structure widely used in chemical processes are analyzed and applied into above framework and a simple but typical example is illustrated to show the convergence of the proposed hybrid branch and bound and revised GBD algorithm on the global solution of the nonconvex MINLP problems.

2. The hybrid branch and bound and revised general Benders decomposition algorithm

The general nonconvex mixed-integer nonlinear programming, MINLP, problem can be formulated as following:

$$
\begin{aligned}
(\mathrm{P})\quad & \min_{\mathbf{x},\mathbf{y}} \; f(\mathbf{x},\mathbf{y}) \\
& s.t.\;\; \mathbf{g}_i(\mathbf{x},\mathbf{y}) \le \mathbf{0} \quad i = 1, 2, \ldots, m \\
& \qquad \mathbf{x} \in \mathbf{S} \subseteq \Re^n \\
& \qquad \mathbf{y} \in \mathbf{Y} = \{0,1\}^q
\end{aligned}
$$

where, $\mathbf{x}$ represents a vector of n continuous variables, and $\mathbf{y}$ is a vector of q binary variables. $\mathbf{S}$ is a nonempty and convex set, which is a simplex in this paper. And the functions

$$f : \Re^n \times \Re^q \to \Re,$$

$$\mathbf{g} : \Re^n \times \Re^q \to \Re^m,$$

are continuously twice-differentiable functions for each fixed $y \in \mathbf{Y} = \{0,1\}^q$. Let D_g be a subset of $\Re^n$ defined by

$$D_g = \{\mathbf{x} \in \Re^n : \mathbf{g}(\mathbf{x},\mathbf{y}) \le \mathbf{0} \text{ for some } \mathbf{y} \in \mathbf{Y}\},$$

and let V_g be a subset of binary set of B^q defined by

$$V_g = \{\mathbf{y} \in \mathrm{B}^q : \mathbf{G}(\mathbf{x},\mathbf{y}) \le \mathbf{0}, \text{ for some } \mathbf{x} \in \mathbf{S}\}.$$

It should be noted that the above stated formulation for Problem (P) is just a subclass of the problems for which the General Benders Decomposition (GBD) of Geoffrion (1972) can be applied. However, the essential difference between them lies in the conditions for the objective and constrained functions, that is, those functions of Problem (P) in this paper are assumed to be only continuously twice-differentiable, rather than convex in Geoffrion (1972).

The main idea of GBD is that the vector of $\mathbf{y}$ variables is defined as the complicating variables in the sense that Problem (P) is a much easier optimization problem in $\mathbf{x}$ when $\mathbf{y}$ is temporarily held fixed. However, the objective and constrained functions in Problem (P) are assumed to be only twice-differentiable, rather than convex, then we have to handle the nonconvexities arisen not only by the joint $\mathbf{x}-\mathbf{y}$ domain structure, but also the continuous variables $\mathbf{x}$ even after the binary variables, i.e. the complicating variables $\mathbf{y}$, are held fixed. In this paper, a hybrid branch and bound and GBD framework is constructed to treat with above stated complications, in fact, the mixed natures of Problem (P) is resolved by the GBD approach, before that, the nonconvexities caused by the continuous variables are removed by a convex quadratic function underestimation techniques developed in the QBB algorithm (Zhu and Xu, 1999; Zhu and Inoue, 2001; Zhu and Kuno, 2001).

2.1 Convex relaxation of the MINLP Problem

Since the vector of $\mathbf{y}$ variables is defined as the complicating variables in GBD method, then we have the following definition to characterize its use in a branch and bound framework:

Definition 2.1 *Given any function* $f(\mathbf{x},\mathbf{y})$, $f:\Re^n \times \Re^q \to \Re$ *and* $\mathbf{x}$ *represents a vector of n continuous variables and* $\mathbf{y}$ *a vector of q binary variables, is continuously twice-differentiable for each fixed* $y \in \mathbf{Y} = \{0,1\}^q$, *the function*

$F(\mathbf{x},\mathbf{y})$, $F:\Re^n \times \Re^q \to \Re$ *for* $\mathbf{x} \in \mathbf{S} \subseteq \Re^n$ *and* $y \in \mathbf{Y} = \{0,1\}^q$ *is defined as the convex relaxation of* $f(\mathbf{x},\mathbf{y})$ *if* $F(\mathbf{x},\mathbf{y})$ *is convex and* $F(\mathbf{x},\mathbf{y}) \le f(\mathbf{x},\mathbf{y}), \forall \mathbf{x} \in \mathbf{S}$ *for each fixed* $y \in \mathbf{Y} = \{0,1\}^q$.

In above definition, since the relationship between the continuous and binary variables in function $f(\mathbf{x},\mathbf{y})$ is implicit, then the specific structure of its convex relaxation in terms of above definition is unknown. However, for most of the chemical engineering processes, the relevant MINLP problems (Duran and Grossman, 1986; Floudas, 1995) can be formulated as a much explicit form, as

$$
\text{(P(ChE))} \qquad \begin{aligned} &\min_{\mathbf{x},\mathbf{y}} \ f(\mathbf{x}) + c^T\mathbf{y} \\ &s.t. \ \mathbf{g}_i(\mathbf{x}) + \mathbf{C}_i^T\mathbf{y} \le \mathbf{0} \qquad i = 1, \dots, m \\ &\qquad \mathbf{x} \in \mathbf{S} \subseteq \Re^n \\ &\qquad \mathbf{y} \in \mathbf{Y} = \{0, 1\}^q \end{aligned}
$$

where, $f(\mathbf{x})$ and $\mathbf{g}_i(\mathbf{x})$ for $i = 1, \dots, m$ are continuously twice differentiable functions. And c is a constant vector belonging to $\Re^q$ as well as $\mathbf{C}$ being a constant matrix belonging to $\Re^{m\times q}$. Hence, the binary variables appear only in the objective and constrained functions in linear forms without nonlinear terms getting involved in the continuous variables. In fact, most applications described by Grossmann (2001) can be taken as the special cases of above formulation by enforcing the continuous functions being convex. Since binary variables are separable from the continuous one, then the valid relaxation of the objective and constraint functions are only dependent on the valid underestimation construction of the twice differentiable functions appeared in above formulation. First, the following theorem tells that there exists a valid convex underestimation function over a simplex for any twice-differentiable function.

Theorem 2.1 *There exists a convex underestimation function for any continuously twice-differentiable function over a simplex.*

Proof. If, without loss of generality, assume that $f(\mathbf{x})$ is a continuously twice-differentiable function over a simplex $\mathbf{S}$, then all elements of its Hessian matrix are continuous and bounded over $\mathbf{S}$. Let α be a large enough positive scalar such that $\mathbf{H}_f(\mathbf{x})+\alpha\mathbf{I}$ is a positive semi-definite matrix for any $\mathbf{x}\in\mathbf{S}$, where $\mathbf{H}_f(\mathbf{x})$ stands for the Hessian matrix of $f(\mathbf{x})$ at each $\mathbf{x}\in\mathbf{S}$. Then, $f(\mathbf{x})+\alpha\|\mathbf{x}\|^2$ is convex, and the function $f(\mathbf{x})$ can be rewritten as

$$f(\mathbf{x})=f(\mathbf{x})+\alpha\|\mathbf{x}\|^2-\alpha\|\mathbf{x}\|^2$$

Obviously, this is a D.C. (Difference of two Convex functions; Tuy, 1998) formulation of $f(\mathbf{x})$. Since the second term in above formulation, i.e. $-\alpha\|\mathbf{x}\|^2$, is concave over the simplex $\mathbf{S}$, then its convex envelope can be expressed by an affine function $\mathbf{c}^T\mathbf{x}+b$ for all $\mathbf{x}\in\mathbf{S}$ (Horst, Pardalos, and Thoai, 1995). Then we have the following function $F(\mathbf{x})$, as

$$F(\mathbf{x})=f(\mathbf{x})+\alpha\|\mathbf{x}\|^2+\mathbf{c}^T\mathbf{x}+b$$

which is obviously convex, and a valid underestimation function of $f(\mathbf{x})$, i.e. $F(\mathbf{x})\le f(\mathbf{x})$ for all $\mathbf{x}\in\mathbf{S}$ since $\mathbf{c}^T\mathbf{x}+b\le-\alpha\|\mathbf{x}\|^2$ by virtue of the definition of the convex envelope. □

In fact, above theorem implies a way to construct the underestimation function for any continuously twice-differentiable function over a simplex.

Geometrically, above theorem uses a very convex quadratic function to compensate the concave part for any twice-differentiable nonconvex function. However, there is a more straightforward way to do this, in fact, we can directly approximate the convex part by using a convex quadratic function, as that presented in the QBB algorithm for any twice-differentiable nonconvex optimization problem (Zhu and Xu, 1999; Zhu and Inoue, 2001; Zhu and Kuno, 2001). That is to say, there exists a convex quadratic function for any twice differentiable function over a simplex, which is also its valid underestimator, as

Definition 2.2 *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*

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

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* $F(\mathbf{x}) \le f(\mathbf{x}), \forall \mathbf{x} \in \mathbf{S}$.

The determination methods of the quadratic, linear, and constant coefficients of Eq.1 are presented simply in the Appendix of this paper, but a detailed introduction can confer to Zhu and Kuno (2001). Note that a simpler form of above quadratic function uses an uniform quadratic coefficient, then Eq.1 becomes a single parameter underestimator. By replacing all the twice differentiable functions by their corresponding convex underestimators over the simplex in Problem (P(ChE)), we get its relaxed formulation as

$$\begin{aligned} (\text{P(ChE)R}) \quad & \min_{\mathbf{x},\mathbf{y}} F(\mathbf{x}) + c^T \mathbf{y} \\ & s.t. \ \mathbf{G}_i(\mathbf{x}) + \mathbf{C}_i^T \mathbf{y} \le \mathbf{0} \qquad i = 1, \ldots, m \\ & \mathbf{x} \in \mathbf{S} \subseteq \Re^n \\ & \mathbf{y} \in \mathbf{Y} = \{0,1\}^q \end{aligned}$$

where, $F(\mathbf{x})$ and $\mathbf{G}_i(\mathbf{x})$ for $i = 1, \ldots, m$ are convex functions described by a combination of some convex or linear functions presented in the Appendix. We see that the objective and constrained functions are satisfied with the Definition 2.1, then above formulation can be seen as a special case of the following Problem, as

$$
\begin{aligned}
(\text{PR}) \quad & \min_{\mathbf{x},\mathbf{y}} \; F(\mathbf{x},\mathbf{y}) \\
& s.t. \;\; \mathbf{G}_i(\mathbf{x},\mathbf{y}) \le \mathbf{0} \quad i = 1, 2, \ldots, m \\
& \qquad \mathbf{x} \in \mathbf{S} \subseteq \Re^n \\
& \qquad \mathbf{y} \in \mathbf{Y} = \{0, 1\}^q
\end{aligned}
$$

where, the functions $F(\mathbf{x},\mathbf{y})$ and $\mathbf{G}_i(\mathbf{x},\mathbf{y})$ hold the Definition 2.1, and let D_G be a subset of $\Re^n$ defined by

$$D_G = \{\mathbf{x} \in \Re^n : \mathbf{G}(\mathbf{x},\mathbf{y}) \le \mathbf{0} \text{ for some } \mathbf{y} \in \mathbf{Y}\},$$

and let V_G be a subset of binary set B^q defined by

$$V_G = \{\mathbf{y} \in \mathrm{B}^q : \mathbf{G}(\mathbf{x},\mathbf{y}) \le \mathbf{0}, \text{ for some } \mathbf{x} \in \mathbf{S}\}.$$

Then, we observe the following conditions given for a mixed integer nonlinear programming problem (Geoffrion, 1972; Floudas, 1995), in order to apply the GBD approach to handle the binary variables.

Conditions 1. $\mathbf{S}$ *is a nonempty, convex set and the functions*

$$
\begin{aligned}
F &: \Re^n \times \Re^q \to \Re, \\
\mathbf{G} &: \Re^n \times \Re^q \to \Re^m,
\end{aligned}
$$

are convex for each fixed $y \in \mathbf{Y} = \{0,1\}^q$.

Remarks 1. This condition holds trivially in virtue of Definitions 2.1 and 2.2.

Condition 2: *the set*

$$Z_y = \{\mathbf{z} \in \Re^m : G(\mathbf{x},\mathbf{y}) \le \mathbf{z} \text{ for some } \mathbf{x} \in \mathbf{S}\},$$

is closed for each fixed $y \in \mathbf{Y}$.

Remarks 2. This condition holds since the simplex $\mathbf{S}$ is bounded and closed, and $\mathbf{G}(\mathbf{x},\mathbf{y})$ is continuous on $\mathbf{x}$ for each fixed $y \in \mathbf{Y}$.

Condition 3*: For each fixed* $\mathbf{y} \in \mathbf{Y} \cap V_G$, *one of the following two cases holds:*

Case i*: The resulting Problem (PR) has a finite solution and has an optimal multiplier vector for the inequalities.*

Case ii*: The resulting Problem (PR) is unbounded, that is, its objective function value goes to* $-\infty$.

Remarks 3. Problem (PR) is a relaxation of the original Problem (P), which is always overestimated, then only the above two Cases are not enough to include all possibilities since the resulting Problem (PR) may be infeasible for each fixed $\mathbf{y} \in \mathbf{Y} \cap \mathbf{V}$. In fact, later we can see how this case will be used frequently to remove the relaxed but infeasible region of the original problem, but first we introduce this additional case into above condition, as

Case iii*: The resulting Problem (PR) is infeasible.*

For the practical application problems, the objective function is always bounded. Then, in this paper, we consider only the Cases i and iii, that is to say, the resulting Problem (PR) is feasible or infeasible. The following theorem tells the relationship between the optimal solution of the relaxed Problem (PR) and the original Problem (P).

Theorem 2.2 *Assume that Problem (P) is bounded. Then for each simplex* $\mathbf{S}$, *if the resulting relaxed Problem (PR) for any* $\mathbf{y} \in \mathbf{Y} \cap V_G$ *is infeasible, then the* *same is true for the original Problem (P). Otherwise, a lower bound* $\mu(\mathbf{S})$ *of* $f(\mathbf{x},\mathbf{y})$ *over* $\mathbf{S} \cap D_g$ *for any* $\mathbf{y} \in \mathbf{Y} \cap V_g$ *can be computed by* $\mu(\mathbf{S}) = F^*$, *where* F^* *is the optimal solution of* $F(\mathbf{x},\mathbf{y})$ *over* $\mathbf{S} \cap D_G$ *for any* $\mathbf{y} \in \mathbf{Y} \cap V_G$.

Proof. For any $\mathbf{y} \in \mathbf{Y}$, since $\mathbf{G}_i(\mathbf{x},\mathbf{y})$ is a convex underestimator of $\mathbf{g}_i(\mathbf{x},\mathbf{y})$, i.e. $\mathbf{G}_i(\mathbf{x},\mathbf{y}) \leq \mathbf{g}_i(\mathbf{x},\mathbf{y})$, we have $\mathbf{G}_i(\mathbf{x},\mathbf{y}) \leq \mathbf{g}_i(\mathbf{x},\mathbf{y}) \leq 0$ for any $\mathbf{x} \in D_g$, that is to say $\mathbf{x} \in D_G$. Then we get $\mathbf{S} \cap D_g \subseteq \mathbf{S} \cap D_G$ for any $\mathbf{y} \in \mathbf{Y}$ by noting $D_g \subseteq D_G$. By using the same way, we get $\mathbf{Y} \cap V_g \subseteq \mathbf{Y} \cap V_G$ for any $\mathbf{x} \in \mathbf{S}$. If the resulting relaxed Problem (PR) for any $\mathbf{y} \in \mathbf{Y} \cap V_G$ is infeasible, then $\mathbf{S} \cap D_G$ is empty for any $\mathbf{y} \in \mathbf{Y} \cap V_G$. Obviously, we have $\mathbf{S} \cap D_g$ which is empty for any $\mathbf{y} \in \mathbf{Y} \cap V_g$.

For the second claim, by virtue of $F(\mathbf{x},\mathbf{y}) \leq f(\mathbf{x},\mathbf{y})$ at $\mathbf{x} \in \mathbf{S} \cap D_g$ for any $\mathbf{y} \in \mathbf{Y} \cap V_g$, we have

$$\begin{aligned} F^* &= \min_{\mathbf{x},\mathbf{y}} \{F(\mathbf{x},\mathbf{y}), \mathbf{x} \in \mathbf{S} \cap D_G \text{ and } \mathbf{y} \in \mathbf{Y} \cap V_G\} \leq F(\mathbf{x},\mathbf{y}), \mathbf{x} \in \mathbf{S} \cap D_G \text{ and } \mathbf{y} \in \mathbf{Y} \cap V_G \\ &\leq f(\mathbf{x},\mathbf{y}), \mathbf{x} \in \mathbf{S} \cap D_G \text{ and } \mathbf{y} \in \mathbf{Y} \cap V_G \leq f(\mathbf{x},\mathbf{y}), \mathbf{x} \in \mathbf{S} \cap D_g \text{ and } \mathbf{y} \in \mathbf{Y} \cap V_g \end{aligned}$$

It states that $\mu(\mathbf{S}) = F^*$ is a valid lower bound of $f(\mathbf{x},\mathbf{y})$ over $\mathbf{S} \cap D_g$ for any $\mathbf{y} \in \mathbf{Y} \cap V_g$. □

It should be noted that the original Problem (P) may be infeasible even when the relaxed Problem (PR) is feasible since the latter is always overestimated in practical applications. The above theorem only provides a qualitative way of obtaining a valid lower bound of $f(\mathbf{x},\mathbf{y})$ over $\mathbf{S}\cap D_g$ for any $\mathbf{y}\in\mathbf{Y}\cap V_g$. Since the constraint set for the binary variables is implicit, and the complete enumeration is 2^q in worst case, then we can see this number grows exponentially and becomes drastically large with the increase of the number of binary variables. Another difficulty of using above theorem to get a lower bound is that the resulting Problem (PR) is feasible for some $\mathbf{y}\in\mathbf{Y}\cap V_G$, but infeasible for the other. Then, the proposed scheme should be capable of discriminating all feasible cases and find the optimal solution among them. In order to overcome all these difficulties, the projection idea is used by virtue of the dual representation and relaxation in the GBD method. But, first we give the following theorem to ensure that the lower bound obtained by Theorem 2.2 is always bounded from below and has a monotonic property for the continuous variables, which is a necessary condition for the convergence of the branch and bound algorithm on the global solution.

Theorem 2.3

*(**a**) Let $\mathbf{S}^1$ and $\mathbf{S}^2$ be two simplices 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 $\mu(\mathbf{S})>-\infty$ for each $\mathbf{S}\subseteq\mathbf{S}^0$.*

Proof.

(**a**) Let D_G^1 and D_G^2 be subsets of $\Re^n$ defined by

$$D_G^1=\{\mathbf{x}\in\Re^n : \mathbf{G}^1(\mathbf{x},\mathbf{y})\leq\mathbf{0} \text{ for some } \mathbf{y}\in\mathbf{Y}\}$$

$$D_G^2=\{\mathbf{x}\in\Re^n : \mathbf{G}^2(\mathbf{x},\mathbf{y})\leq\mathbf{0} \text{ for some } \mathbf{y}\in\mathbf{Y}\}$$

where, the underestimation functions are generated over two simplices $\mathbf{S}^1$ and $\mathbf{S}^2$, respectively. Let V_G^1 and V_G^2 be two subsets of binary set B^q defined by

$$V_G^1=\{\mathbf{y}\in\mathrm{B}^q : \mathbf{G}^1(\mathbf{x},\mathbf{y})\leq\mathbf{0}, \text{ for some } \mathbf{x}\in\mathbf{S}\}$$

$$V_G^2 = \{\mathbf{y} \in \mathrm{B}^q : \mathbf{G}^2(\mathbf{x},\mathbf{y}) \le \mathbf{0}, \ \text{for some } \mathbf{x} \in \mathbf{S}\}$$

Since $\mathbf{S}^2 \subset \mathbf{S}^1$, and by virtue of the argument of the Proposition 2.2.2 of Zhu and Kuno (2001), for any $y \in \mathbf{Y}$ we have

$$F^1(\mathbf{x},\mathbf{y}) \le F^2(\mathbf{x},\mathbf{y}) \ \text{ and } \ G_i^1(\mathbf{x},\mathbf{y}) \le G_i^2(\mathbf{x},\mathbf{y}) \ \text{ for } \ i = 1,\dots,m$$

Then, we have $D_G^1 \supseteq D_G^2$, and $V_G^1 \supseteq V_G^2$. Since $\mathbf{S}^2 \subset \mathbf{S}^1$, finally we get

$$\mu(\mathbf{S}^2) = \min_{\mathbf{x},\mathbf{y}}\{F^2(\mathbf{x},\mathbf{y}) : \mathbf{x} \in \mathbf{S}^2 \cap D_G^2, \mathbf{y} \in \mathbf{Y} \cap V_G^2\} \ge \min_{\mathbf{x},\mathbf{y}}\{F^1(\mathbf{x},\mathbf{y}) : \mathbf{x} \in \mathbf{S}^1 \cap D_G^1, \mathbf{y} \in \mathbf{Y} \cap V_G^1\} = \mu(\mathbf{S}^1)$$

(**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 (PR($\mathbf{S}^0$)) is convex for each fixed $\mathbf{y} \in \mathbf{Y} \cap V_g$. Then, this problem has an optimal solution, which implies that $\mu(\mathbf{S}^0) > -\infty$. □

2.2 The revised GBD method for the relaxed MINLP Problem (PR)

The relaxed Problem (PR) can provide a lower bound of the original Problem (P) over the current simplex if it is feasible. Otherwise, it can facilitate to remove that simplex with the branch and bound algorithm progress. However, complication arises due to the joint natures of binary and continuous variables in the relaxed Problem (PR). The complete branch and bound algorithm uses continuous relaxation with respect to the integer variables, and then solves the continuous convex NLP to generate the lower bound. But, it is quite inefficient if the integer variable number is slightly big. Then, a more intellectual way is to use the Lagrange relaxation presented by Benders (1962) and Geoffrion (1972) for MILP or MINLP, respectively. In this section, the General Benders

Decomposition, i.e. GBD method, is revised to handle above mentioned difficulty, which consists of two basic operations, the Primal Problem and the Master Problem, to obtain the upper bound and lower bound of the relaxed Problem ($PR(S^k)$), respectively, at each iteration over the current simplex.

2.2.1 The Primal Problem of the relaxed Problem ($PR(S^k)$)

The Primal Problem results from fixing the binary variables $\mathbf{y}$ to a particular 0-1 combination, which is denoted as $\mathbf{y}^t$ where *t* stands for the iteration counter of the GBD method. The formulation of the Primal Problem ($PRS^k(y^t)$), at iteration *t* over subsimplex $\mathbf{S}^k$ is given as

$$
\begin{aligned}
(PRS^k(y^t)) \quad & \min_{\mathbf{x}} \; F(\mathbf{x},\mathbf{y}^t) \\
s.t. \quad & \mathbf{G}_i(\mathbf{x},\mathbf{y}^t) \le \mathbf{0} \quad i = 1,2,\ldots,m \\
& \mathbf{x} \in \mathbf{S}^k \subseteq \Re^n
\end{aligned}
$$

Obviously, this problem is convex due to the Conditions *i* in the former section. However, we have to distinguish two possible cases, i.e. feasible and infeasible, according to the relaxed Problem ($PR(S^k)$). If the relaxed Problem ($PR(S^k)$) is infeasible, of course the Primal Problem ($PRS^k(y^t)$) is infeasible too. But, the relaxed Problem ($PR(S^k)$) is also possible to be infeasible even when the Primal Problem ($PRS^k(y^t)$) is feasible, since unsuitable binary variable vector may be chosen to be fixed. However, these two cases are unable to be distinguished immediately now, then they are treated here together. When the Primal Problem ($PRS^k(y^t)$) is feasible, then its solution provides information on $\mathbf{x}^t$, and $F(\mathbf{x}^t,\mathbf{y}^t)$, which is the upper bound of the relaxed Problem ($PR(S^k)$), and the optimal multiplier vector λ^t for the inequality constraints. Then, the Lagrange function for the feasible case can be constructed as

$$
L(\mathbf{x},\mathbf{y},\lambda) = F(\mathbf{x},\mathbf{y}) + \sum_{i=1}^{m} \lambda_i \mathbf{G}_i(\mathbf{x},\mathbf{y})
$$

If the Primal Problem ($PRS^k(y^t)$) is detected by the NLP solver to be infeasible, then the perturbation theory is used to generate a maximal integer cut to remove this combination of the binary variables. The l_1-minimization problem, i.e. the sum of constraint violation, can be formulated as

$$
\begin{aligned}
(\,\mathrm{FP}(\alpha)\,) \quad & \min_{\mathbf{x},\alpha} \sum_{i=1}^{m} \alpha_i \\
& s.t. \ \ \mathbf{G}_i(\mathbf{x},\mathbf{y}^t) \le \alpha_i \quad i = 1,2,\ldots,m \\
& \quad \mathbf{x} \in \mathbf{S}^k \subseteq \Re^n \\
& \quad \alpha_i \ge 0 \quad i = 1,2,\ldots,m
\end{aligned}
$$

Note, this minimization problem, denoted as ($\mathrm{FP}(\alpha)$), is convex over all variables, and a feasible point has been determined if the minimum of the objective function is zero, i.e. $\sum_{i=1}^{m} \alpha_i = 0$. The solution of above feasibility problem provides information on the Lagrange multiplier vector for the inequality constraints, which are denoted as μ so as to distinguish from the feasible case. Then, the Lagrange function for infeasible case can be constructed as

$$
L(\mathbf{x},\mathbf{y},\mu) = \sum_{i=1}^{m} \mu_i \mathbf{G}_i(\mathbf{x},\mathbf{y})
$$

2.2.2 The Master Problem of the relaxed Problem (PR(S^k))

The derivation of the master problem in the GBD method makes use of the nonlinear duality theory, and can be characterized as minimization of the dual representation of the projection of the relaxed Problem (PR(S^k)) on the $\mathbf{y}$-space over the dual representation of V_G. First, we see the projection of the relaxed Problem (PR(S^k)) onto $\mathbf{y}$-space, let $v(\mathbf{y})$ be defined as

$$
v(\mathbf{y}) = \inf_{\mathbf{x}} F(\mathbf{x},\mathbf{y})
$$

$$s.t.\ \mathbf{G}_i(\mathbf{x},\mathbf{y}) \le \mathbf{0} \quad i = 1, 2, \ldots, m$$
$$\mathbf{x} \in \mathbf{S}^k \subseteq \Re^n$$

where $v(\mathbf{y})$ is parametric with respect to the binary variable vector. Then, the relaxed Problem ($\mathrm{PR(S^k)}$) can be rewritten as

$$(\,\mathrm{PR(y)}\,) \qquad \min_{\mathbf{y}}\ v(\mathbf{y})$$
$$s.t.\ \mathbf{y} \in \mathbf{Y} \cap V_G,$$

which is denoted as Problem (PR(y)). It should be noted that the definition of $v(\mathbf{y})$ is infimum with respect to $\mathbf{x}$ since for given $\mathbf{y}$ the inner optimization problem may be unbounded, and its value corresponds to the optimal value of the relaxed Problem ($\mathrm{PR(S^k)}$) for fixed $\mathbf{y}$. Then, the Problem (PR(y)) is the projection of the relaxed Problem ($\mathrm{PR(S^k)}$) onto the $\mathbf{y}$-space. It can be shown that this projected problem (PR(y)) is equivalent to the relaxed Problem ($\mathrm{PR(S^k)}$) (Geoffrion, 1972; Floudas, 1995). Since we always assume that the Problem (P) has a solution, then the unbounded case is not needed to be considered in this paper according to the Proposition 2.1. Then, the dual representations of V_G and $v(\mathbf{y})$ are presented as follows:

$$\{\mathbf{y} \in \mathbf{Y} \cap V_G\} \equiv \{\mathbf{y} \in \mathbf{Y} : \max_{\mu \ge 0}\ \min_{\mathbf{x} \in \mathbf{S}^k} L(\mathbf{x},\mathbf{y},\mu) \le 0\} \qquad (\ 2\ )$$

and

$$v(\mathbf{y}) = \max_{\lambda \ge 0}\ \min_{\mathbf{x} \in \mathbf{S}^k} L(\mathbf{x},\mathbf{y},\lambda),\ \forall \mathbf{y} \in \mathbf{Y} \cap V_G \qquad (\ 3\ )$$

According to the strong duality theorem (Geoffrion, 1972), those two dual representations are satisfied by virtue of the Conditions 1, 2 and 3. However, the Lagrange functions used in above dual representations involve the maximization over all multipliers. Hence, the relaxation of them will represent only the lower bounds of those Lagrange functions by dropping a number of constraints. For example, the Lagrange multipliers calculated in the Master problem are used here to construct the following relaxations, as

$$\{\mathbf{Y} \cap V_G\} \subseteq \{\mathbf{y} \in \mathbf{Y} \cap V_G^t\} \equiv \{\mathbf{y} \in \mathbf{Y} : \min_{\mathbf{x} \in \mathbf{S}^k} L(\mathbf{x}, \mathbf{y}, \mu^t) \leq 0\} \tag{4}$$

and

$$v(\mathbf{y}) \geq v^t(\mathbf{y}) = \min_{\mathbf{x} \in \mathbf{S}^k} L(\mathbf{x}, \mathbf{y}, \lambda^t), \forall \mathbf{y} \in \mathbf{Y} \cap V_G \tag{5}$$

Then, obviously the following optimal problem, denoted as the Problem ($PR(y^t)$), will produce only a lower bound of the relaxed Problem ($PR(S^k)$), as

$$\begin{aligned} (\,PR(y^t)\,) \quad & \min_{\mathbf{y} \in \mathbf{Y}, y_0} y_0 \\ & s.t. \ \min_{\mathbf{x} \in \mathbf{S}^k} L(\mathbf{x}, \mathbf{y}, \lambda^p) \leq y_0 \quad p = 1, \ldots, p^t \\ & \quad\ \ \min_{\mathbf{x} \in \mathbf{S}^k} L(\mathbf{x}, \mathbf{y}, \mu^l) \leq 0 \quad l = 1, \ldots, l^t \end{aligned},$$

where, y_0 is a scalar introduced to represent the lower bound of the relaxed Problem ($PR(S^k)$), and $p^t + l^t = t$ since the primal Problem ($PRS^k(y^t)$) is possibly either feasible or infeasible. It should be noted that the integer constraint generated in the current iteration, no matter the primal Problem is feasible or infeasible, will be introduced into the next iteration. Since we cannot certainly find a feasible binary combination of the primal Problem at each iteration, or no any feasible one exists at all for a infeasible primal Problem, then the following feasibility problem is introduced to aim at searching a feasible binary variable, or adding a more compact integer cut in the Problem ($PR(y^t)$),

$$\begin{aligned} (\,FP(\beta)\,) \quad & \min_{\mathbf{y} \in \mathbf{Y}, \beta} \beta \\ & s.t. \ \min_{\mathbf{x} \in \mathbf{S}^k} L(\mathbf{x}, \mathbf{y}, \mu^l) \leq 0 \quad l = 1, \ldots, t-1 \\ & \quad\ \ \min_{\mathbf{x} \in \mathbf{S}^k} L(\mathbf{x}, \mathbf{y}, \mu^t) \leq \beta \\ & \quad\ \ \beta \geq 0 \end{aligned}$$

If the minimum objective function value is zero, i.e. $\beta = 0$, then the solution of above optimal problem, denoted as (FP(β)), provides a binary combination for the next iteration. Otherwise, the iteration terminates since it tells that the relaxed Problem (PR(S^k)) is infeasible, then there is no need to fathom the current subsimplex further. The following theorem ensures this relationship, as

Theorem 2.4 *If the feasibility Problem (FP(β)) terminated at some iteration is infeasible, then the same is true for the relaxed Problem (PR(S^k)) and vice versa.*

Proof. If the relaxed Problem (PR(S^k)) is infeasible, then the set V_G is empty. Hence the binary subset $\{\mathbf{y} \in \mathbf{Y} : \max_{\mu \geq 0} \min_{\mathbf{x} \in \mathbf{S}^k} L(\mathbf{x}, \mathbf{y}, \mu) \leq 0\}$ is empty by virtue of the strong dual theory on the basis of the Conditions 1, 2, and 3. If the feasibility Problem is always feasible, then finally we have a multiplier vector μ^* and a binary variable vector $\mathbf{y}^*$ satisfying

$$L(\mathbf{x}^*, \mathbf{y}^*, \mu^*) \leq 0$$

where, $\mathbf{x}^*$ is the optimal solution of $\min_{\mathbf{x} \in \mathbf{S}^k} L(\mathbf{x}, \mathbf{y}^*, \mu^*) \leq 0$. Obviously this contradiction implies that the feasibility Problem (FP(β)) will terminate finitely and be infeasible.

Conversely, if feasibility Problem (FP(β)) terminated at iteration t is infeasible, then the following set

$$\{\mathbf{y} \in \mathbf{Y} \cap V_G^t\} = \{\mathbf{y} \in \mathbf{Y} : \min_{\mathbf{x} \in \mathbf{S}^k} L(\mathbf{x}, \mathbf{y}, \mu^i) \leq 0, i = 1, \ldots, t\}$$

is empty. Then, the relaxed Problem (PR(S^k)) is infeasible since its constrained set, i.e. $\{\mathbf{y} \in \mathbf{Y} \cap V_G \subseteq \mathbf{Y} \cap V_G^t\}$ is empty according to Eq.4. □

2.2.3 Algorithmic procedure of the revised GBD for the relaxed Problem ($PR(S^k)$)

The revised GBD procedure for the relaxed Problem ($PR(S^k)$) can now be stated formally with the consideration that above relaxed Problem is completely infeasible over the current subsimplex. It should be noted here that Geoffrion (1972) did not include the infeasible case, but we assume that the relaxed Problem always has a bounded optimal value if it is feasible.

Procedure of the revised GBD

Step 0. Initialization. Set the current upper bound UBD be a very large positive value, and let the current lower bound LBD be the negative value of UBD. Set the feasible and infeasible counters $p = 0$, $l = 0$, respectively. Then. select the convergence tolerance $\varepsilon^c \geq 0$ and feasibility tolerance $\varepsilon^f \geq 0$. Choose an initial point $\mathbf{y}^1 \in \mathbf{Y}$ and set the counter $t = 1$.

Step 1. Solve the primal Problem. Solve the resulting primal Problem ($PRS^k(y^t)$). If the NLP solver verifies that above Problem is feasible, then set $p \leftarrow p+1$, and the optimal primal solution $\mathbf{x}^p$ and the optimal multiplier vector λ^p are obtained. Compute the current upper bound $UBD = \min\{UBD, f(\mathbf{x}^p, \mathbf{y}^p)\}$; Otherwise, set $l \leftarrow l+1$, and solve the feasibility Problem ($FP(\alpha)$), then obtain the multiplier vector μ^q.

Step 2. Solve the relaxed master Problem. If $p \geq 1$, solve the relaxed master Problem ($PR(y^t)$), as

$$
\begin{aligned}
&\min_{\mathbf{y}\in\mathbf{Y}, y_0} \; y_0 \\
&s.t. \; \min_{\mathbf{x}\in\mathbf{S}^k} L(\mathbf{x}, \mathbf{y}, \lambda^i) \leq y_0 \quad i = 1,\ldots,p \\
&\qquad \min_{\mathbf{x}\in\mathbf{S}^k} L(\mathbf{x}, \mathbf{y}, \mu^j) \leq 0 \quad j = 1,\ldots,l \;,
\end{aligned}
$$

Note here, the second constraint vanishes if $l=0$. Then we get y_0 and $\mathbf{y}^f$, and set the current lower bound $LBD=\max\{LBD, y_0\}$. Check if $UBD-LBD\le\varepsilon$, yes, then terminate the iteration and the solutions of the relaxed Problem (PR(S^k)) are $\{\mathbf{x}^p,\mathbf{y}^p\}$ and UBD. Otherwise, set $t\leftarrow t+1$ and let $\mathbf{y}^t=\mathbf{y}^f$, return to **Step 1**.

If $p=0$, solve the relaxed feasibility Problem (FP(β)), as

$$\begin{aligned}
&\min_{\mathbf{y}\in\mathbf{Y},\beta} \beta \\
&s.t.\ \min_{\mathbf{x}\in\mathbf{S}^k} L(\mathbf{x},\mathbf{y},\mu^i)\le 0 \quad i=1,\dots,l-1 \\
&\quad\ \ \min_{\mathbf{x}\in\mathbf{S}^k} L(\mathbf{x},\mathbf{y},\mu^l)\le \beta \\
&\quad\ \ \beta\ge 0
\end{aligned}$$

We get β and $\mathbf{y}^{if}$. If $\beta\ge\varepsilon^f$, terminate and the relaxed Problem (PR(S^k)) is infeasible over the current subsimplex. Otherwise, set $t\leftarrow t+1$ and let $\mathbf{y}^t=\mathbf{y}^{if}$, return to **Step 1**. □

The following theorem ensures that above revised GBD algorithm converges finitely no matter the relaxed Problem (PR(S^k)) is feasible or not, as

Theorem 2.5 *Assume that* $\mathbf{Y}$ *is a finite binary set, that the representations of* V_G *and* $v(\mathbf{y})$ *are held based on the strong duality theory. Then, the above revised GBD procedure terminates finitely for any given* $\varepsilon^c\ge 0$.

Proof. If the relaxed Problem (PR(S^k)) is infeasible, then the above GBD procedure has finite termination thanks to the integral finiteness of $\mathbf{Y}$ and the fact that no $\mathbf{y}^{if}$ can repeat itself in a solution of the relaxed feasibility Problem (FP(β)), which is manifested by introducing a additional integer cut obtained from the infeasible primal Problem, i.e. the feasibility Problem (FP(α)), into the constraint set of the relaxed feasibility Problem (FP(β)). If the relaxed Problem (PR(S^k)) is feasible, then the feasible $\mathbf{y}^f$ can not be repeated unless the convergence criterion is satisfied (Geoffrion, 1972), which is implied by introducing a additional optimality integer cut generated from the feasible primal Problem into the constraint set of the relaxed master Problem (PR(S^k)). Finally, the worst performance of the above procedure could be the complete enumeration of the integer elements in $\mathbf{Y}$, which is finite. □

Remarks.

In Step 2 of the above GBD procedure, a rather important assumption is that we can always find the solutions of the inner optimization problems for the given multiplier vectors in the relaxed master Problem (PR(y^t)), or the relaxed feasibility Problem (FP(β)). However, the determination of those solutions cannot be achieved in general, since they are always parametric functions of binary variable vector $\mathbf{y}$ obtained from the solutions of the inner optimization problems. Their determination method in general requires a global optimization approach, but there exist a number of special structures for which the solutions of those inner optimization problems can be obtained explicitly as functions of binary variable vector $\mathbf{y}$. For the MINLP problem widely encountered in the chemical engineering processes, as described by Problem (P(ChE)), the inner optimization problems can be explicitly obtained by its relaxed formulation using a further relaxation. Now, if we have got the multiplier vectors being λ or μ for feasible or infeasible primal problems, respectively, then the inner optimization problems generated in the relaxed master problems (feasible or infeasible) for the Problem (P(ChE)R) are represented as

$$\min_{\mathbf{x}} \left\{ \mathrm{F}(\mathbf{x}) + c^T\mathbf{y} + \sum_{i=1}^{m} \lambda_i \left(\mathbf{G}_i(\mathbf{x}) + \mathbf{C}_i^T\mathbf{y} \right) \right\} \le y_0$$

or

$$\min_{\mathbf{x}}\left\{\sum_{i=1}^{m}\mu_i\left(\mathbf{G}_i(\mathbf{x})+\mathbf{C}_i^T\mathbf{y}\right)\right\}\le 0$$

Since the binary variable vectors are separable from the continuous variables, and the optimal multiplier vectors are always nonnegative, then we can reformulate above two inner optimization problems into the following more explicit forms, as

$$\min_{\mathbf{x}}\left\{\mathrm{F}(\mathbf{x})+\sum_{i=1}^{m}\lambda_i\mathbf{G}_i(\mathbf{x})\right\}+c^T\mathbf{y}+\sum_{i=1}^{m}\lambda_i\mathbf{C}_i^T\mathbf{y}\le y_0 \tag{6}$$

or

$$\min_{\mathbf{x}}\left\{\sum_{i=1}^{m}\mu_i\mathbf{G}_i(\mathbf{x})\right\}+\sum_{i=1}^{m}\mu_i\mathbf{C}_i^T\mathbf{y}\le 0 \tag{7}$$

According to the Definition 2.2, we know that all of the relaxed continuous functions are convex, then their minima over the current subsimplex can be calculated by any NLP solver. Hence, by replacing those constraints in the master problems with above further relaxations, we have the explicit formulation with respect to only the binary variable vector. The nonconvexity arisen by the joint continuous and binary variables need to be handled to generate a valid and global integer cut described in Problem (PR(y^t)) for feasible or infeasible case, respectively. Then, the following theorem ensures that the revised GBD method proposed above can identify the global solution of the Problem (P(ChE)R), if it is feasible.

Theorem 2.6 *The revised GBD approach converges on the ε^c-global solution of the Problem (P(ChE)R).*

Proof. Since the Problem (PR(y)) can be equivalently expressed by the strong dual representations, i.e. Eqs.2 and 3, for its objective function $v(\mathbf{y})$ and constrained set V_G, and Eqs.4 and 5 are valid relaxations for above two equations, then we only need to show that the resulting Eqs.4 and 5 are convex or linear with respect to $\mathbf{y}$ for the global convergence of the revised GBD approach of the Problem (P(ChE)R), since the relaxed master problem provides a global underestimation for the Problem (P(ChE)R). In fact, Eqs.6 and 7 are the corresponding formulations of Eqs.4 and 5 for the Problem (P(ChE)R), and it is obvious to observe that these two constraints are linear with respect to $\mathbf{y}$. Then, the revised GBD approach converges on the ε^c-global solution of the Problem (P(ChE)R). □

In fact, the structure of the MINLP problem can be extended from the separable type, i.e. $f(\mathbf{x})+c^T\mathbf{y}$, into the bilinear type, i.e. $\mathbf{x}^T\mathbf{y}$, and partly linear type, i.e. $\mathbf{y}^T\mathbf{f}(\mathbf{x})$, according to above theorem. For the latter two ceases, the inner parametric optimization problems can be relaxed further by noting the nonnegativity of the binary variable. If we can generalize the last two structures as $\sum_{i=1}^{q} f_i(\mathbf{x})y_i$, then its further relaxation is $\sum_{i=1}^{q} \min_{\mathbf{x}}\{F_i(\mathbf{x})\}y_i$, where $F_i(\mathbf{x})$ is the convex underestimation function of $f_i(\mathbf{x})$ over each continuous domain.

2.3 Hybrid branch and bound and GBD procedure for MINLP

Before we present the full procedure of the hybrid branch and bound and revised GBD algorithm for the nonconvex MINLP problems, the two other necessary basic operations in a branch and bound framework should also be illustrated, i.e. the branching procedure for the domain of the continuous variables and the upper bound calculated over each subsimplex generated in above partition process. For the branching procedure, the well known simplicial partition often used in global optimization (Zhu and Xu, 1999; Zhu and Inoue, 2001; Zhu and Kuno, 2001) is applied here. In order to guarantee the convergence of the branch and bound algorithm on the global solution of the nonconvex MINLP problem, the exhaustiveness of the simplicial partition has to be satisfied. In this paper, the typical exhaustive partition process, i.e. the simplicial bisection, is used, where the longest edge of the current simplex is always divided into two parts in terms of its length in order to construct two subsimplices for the potential fathoms. The detailed description of this kind of partition is given in

Zhu and Kuno (2001), which is not given here for the sake of the space limitation.

For the calculation of a rigorous upper bound over the current simplex, the nonconvex MINLP is solved locally for the Problem (P) without any convexification by using the GBD method (Geoffrion, 1972). It should be noted here that it is not certain to obtain an upper bound in this way, especially when the current subsimplex is infeasible for the MINLP. But during the progress of the branch and bound procedure, the obtained upper bounds is updated so as to get a nonincreasing sequence to converge on the global solution of the MINLP, and the concerned discussion about this convergence is presented in the latter part.

Now, we are in a position to describe the proposed branch and bound algorithm for solving the nonconvex MINLP based on above basic operations, especially the convexification techniques embedded in the QBB algorithm, see the Appendix, and the revised general Benders decomposition, i.e. revised GBD, introduced in the above sections, provided that an initial simplex is available for the continuous variables in the original nonconvex MINLP problem. Of course, the last condition is not necessarily given in the MINLP problem, but as we have pointed out in Zhu and Kuno (2001), this simplex can be constructed by using an Out Approximation (OA) method according to the physical or insightful bounds of those continuous variables.

The hybrid Brach and Bound and GBD procedure for nonconvex MINLP problems

Step 1 - Initialization. A convergence tolerance, ε_c, and a feasibility tolerance, ε_f, are selected and the iteration counter k is set to be zero. The initial simplex with respect to the continuous variables is given as $\mathbf{S}^0$, which is known *a prior* or able to be computed by an OA method. The global lower and upper bounds μ_0 and γ_0 on the global minimum of the MINLP Problem (P) are initialized and an initial current point $\left(\mathbf{x}^{k,c}, \mathbf{y}^{k,c}\right)$ is randomly selected.

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

The MINLP Problem (P) is solved locally by the GBD method within the current simplex $\mathbf{S}$. If the solution f_{local}^{k} of the MINLP Problem (P) is ε_f-feasible, then the upper bound γ_k is updated as $\gamma_k = \min(\gamma_{k-1}, f_{local}^{k})$.

Step 3 – Partitioning of current simplex

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

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

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

where, (k,m) and (k,l) correspond to the vertices incident to 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 –Convexify the MINLP inside both subsimplices $r = 1,\ 2$

The nonconvex functions in the objective function and constraints with respect to the continuous variables are convexified to obtain the relaxed MINLP Problem (PR) inside both subsimplices $r = 1,\ 2$ according to the methods presented in the Appendix.

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

The relaxed MINLP Problem (PR) is solved inside both subsimplices ($r = 1,\ 2$) by using the revised GBD method. If a solution $F_{sol}^{k,r}$ is feasible and less than the current upper bound, γ_k, then it is stored along with the solution point $\left(\mathbf{x}_{sol}^{k,r}, \mathbf{y}_{sol}^{k,r}\right)$.

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

Step 7 – Update current point $\left(\mathbf{x}^{k,c}, \mathbf{y}^{k,c}\right)$ 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**,

$$\left(\mathbf{x}^{k,c}, \mathbf{y}^{k,c}\right) = \left(\mathbf{x}_{sol}^{I',r'}, \mathbf{y}_{sol}^{I',r'}\right)$$

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

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:

$$f^{*} \leftarrow f^{c,k''}, \text{ and } \left(\mathbf{x}^{*},\mathbf{y}^{*}\right) \leftarrow \left(\mathbf{x}^{c,k''},\mathbf{y}^{c,k''}\right)$$

where, $k'' = \arg_{I}\left\{f^{c,I} = \gamma_{k}\right\}$, $I = 1,\ldots, k$. □

Remarks: In the above proposed hybrid branch and bound and revised GBD algorithm, the two kinds of nonconvexities are handled separately, i.e. the nonconvexity introduced by the continuous functions is overcome by using relaxations in the branch and bound framework by virtue of the quadratic function based underestimators, and latter caused by the natures of the joint continuous and binary variables is resolved in the revised GBD approach on the basis of the relaxation by using the strong dual theory. Then, the global convergence of above hybrid algorithm depends not only on the construction of the valid underestimators for any twice-differentiable continuous functions, but also the favorable structures of the MINLP problem with respect to the continuous and binary variables. Hence, this algorithm is not universally reliable for any kind of nonconvex MINLP problems, but for those that their special structures can make the revised GBD converge on the global solution of each subproblem within the branch and bound framework, such as the formulation discussed above in the chemical engineering field, i.e. the Problem (P(ChE)R). It should be noted that the current simplex can be deleted in Step 5 when either the relaxed Problem (PR) is infeasible or its solution is greater than the current best upper bound. The former is justified by solving the introduced feasibility problem in the revised GBD approach, then that subsimplex is removed immediately after knowing the infeasibility. The latter depletion is valid since the global minimum can never happen in this simplex for the lower bound computed over this simplex is already greater than the current best upper bound.

If the hybrid algorithm terminates at iteration *k*, then the point $\left(\mathbf{x}^{k},\mathbf{y}^{k}\right)$ is an optimal solution of the MINLP Problem (P(ChE)). In the case that the hybrid algorithm is not finite, it generates at least one infinite sequence of simplices $\left\{\mathbf{S}^{j}\right\}$ for continuous variables such that $\mathbf{S}^{j+1} \subset \mathbf{S}^{j}$, for all j. The convergence of the hybrid branch and bound and GBD algorithm is guaranteed in terms of the following theorem, as

Theorem 2.7 *Assume that Problem (P(ChE)) has a feasible solution. Further, assume that the hybrid branch and bound and GBD algorithm generates an infinite subsequence of simplices* $\{\mathbf{S}^j\}$ *for continuous variables 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,* $(\mathbf{x}^*, \mathbf{y}^*)$ *is an optimal solution of the MINLP Problem (P(ChE)), where* $\mathbf{y}^*$ *is the integer solution of the MINLP Problem (P(ChE)) at the fixed* $\mathbf{x}^*$.

It should be noted that above theorem does not claim for any MINLP problem, since the nonconvexity arisen by the joint natures of the continuous and binary variables always leads to local solution in the GBD step for solving the relaxed problem within the branch and bound framework. However, this difficulty can be avoided by virtue of the favorable structure of the Problem (P(ChE)). The proof of the above theorem can be attained by the classical convergence conditions (Horst, Pardalos, and Thoai, 1995; Zhu and Kuno, 2001) of the branch and bound framework on the basis of the exhaustive partition of the constrained region and the monotonicity of the lower bound stated in Section 2.1. Finally the global integer solution, i.e. $\mathbf{y}^*$, can be manifested by Theorems 2.5 and 2.6 if the three kinds of favorable structures in chemical processes are assumed in the nonconvex MINLP formulation.

3. Computational study of the hybrid branch and bound and GBD algorithm

A very small MINLP problem for process synthesis used here has ever appeared in the literature as a typical test example (Kocis and Grossmann, 1988; Floudas, Aggrwal, & Ciric, 1989; Ryoo and Sahinidis, 1995). To illustrate the global convergence of the proposed algorithm in this paper, let us describe all possible cases possibly happened during the iterations of this problem.

$$
\begin{aligned}
\min_{x,y} \quad & 2x+y \\
s.t. \quad & 1.25-x^2-y \le 0 \\
& x+y \le 1.6 \\
& 0 \le x \le 1.6 \\
& y=\{0,1\}
\end{aligned}
$$

The nonconvexities arise in two aspects from above problem, one is the joint nature of the binary and continuous variables, the other is caused by the continuous concave function, i.e. $-x^2$, appeared in the first constraint. The latter is solved by a continuous relaxation, i.e. to replace this concave function by its convex envelope, see Appendix, over each simplex in a branch and bound framework. And the former difficulty is overcome by the method provided by the revised GBD in terms of the linear joint structure of the continuous and binary variables in the relaxed problem.

Iteration 1. $\varepsilon_c=\varepsilon_f=0.001$, $k=0$, $S=[0.0, 1.6]$, $\mu_0=100$, $\gamma_0=-100$.

Iteration 2. We first fix the binary variable y being 0, we have the following NLP problem, as

$$
\begin{aligned}
\min_{x} \quad & 2x \\
s.t. \quad & 1.25-x^2 \le 0 \\
& x \le 1.6 \\
& 0 \le x \le 1.6
\end{aligned}
$$

Solve this nonconvex NLP, we get the minimum solution at $x=1.118$ with $f=2.236$. Then, the upper bound of the branch and bound algorithm, μ_0, is updated to be 2.236.

Iteration 3. Divide the interval of the continuous variable, i.e. [0.0, 1.6], into two subintervals, i.e. [0.0, 0.8] and [0.8, 1.6].

Iteration 4. The two relaxed problems in above two subintervals are obtained by replacing the concave functions by their convex envelopes in each subinterval, given in the Iteration 5 as well as their solutions solved by the revised GBD method.

Iteration 5. Let us first see the relaxed problem in the subinterval [0.0, 0.8].

Iteration 5.1.

$$\begin{aligned} \min_{x,y} \quad & 2x + y \\ s.t. \quad & 1.25 - 0.8x - y \le 0 \\ & x + y \le 1.6 \\ & 0 \le x \le 0.8 \\ & y = \{0, 1\} \end{aligned}$$

Then, by using the revised GBD method, we can get the global solution of this problem.

Step1. 0, UBD=100, LBD=-100, p=0, l=0, $\varepsilon^c = \varepsilon^f$ =0.001, and t=1.

Step 1.1, fix y=0, we get the resulting convex (linear) problem, as

$$\begin{aligned} \min_{x} \quad & 2x \\ s.t. \quad & 1.25 - 0.8x \le 0 \\ & x \le 1.6 \\ & 0 \le x \le 0.8 \end{aligned}$$

However, this problem is infeasible. Then, l=1, and we solve the following feasibility problem, as

$$\begin{aligned} \min_{x,\alpha} \quad & \alpha_1 + \alpha_2 \\ s.t. \quad & 1.25 - 0.8x - \alpha_1 \le 0 \\ & x - 1.6 - \alpha_2 \le 0 \\ & -\alpha_1 \le 0 \\ & -\alpha_2 \le 0 \\ & -x \le 0 \\ & x - 0.8 \le 0 \end{aligned}$$

By using the KKT condition of above problem, we get the Lagrange multipliers as μ^1=1.0, $\mu^2 = \mu^3 = \mu^4 = \mu^5$=0.0, μ^6=0.8.

Step 1.2. Since p=0, then we solve the following feasibility problem

$$\begin{aligned} \min_{y,\beta} \quad & \beta \\ s.t. \quad & \min_{0 \le x \le 0.8} \{1.25 - 0.8x - y\} \le \beta \\ & \beta \ge 0 \\ & y = \{0, 1\} \end{aligned}$$

We get the solution y=1. Then go to *Step 2.1*, we have the resulting primal problem, as

Step 2.1

$$\begin{aligned} \min_{x} \quad & 2x + 1 \\ s.t. \quad & 0.25 - 0.8x \le 0 \\ & x - 0.6 \le 0 \\ & -x \le 0 \\ & x - 0.8 \le 0 \end{aligned}$$

Solve this convex, in fact linear problem, we get the minimum solution at x=0.3125 with f=1.625. Then, the upper bound of the GBD method, UBD, is updated to be 1.625. And the Lagrange multiplier is λ^1=2.5, $\lambda^2=\lambda^3=\lambda^4$=0.0. And set p=1, go to *Step 2.2*.

Step 2.2. Since p=1 and l=1, then we have the following master problem, as

$$\begin{aligned} \min_{y,y_0} \quad & y_0 \\ s.t. \quad & \min_{0 \le x \le 0.8} \{2x + y + 2.5 \times (1.25 - 0.8x - y)\} \le y_0 \\ & \min_{0 \le x \le 0.8} \{1.25 - 0.8x - y\} \le 0 \\ & y_0 \ge 0 \\ & y = \{0, 1\} \end{aligned}$$

Solve this problem, we get y=1, and y_0=1.625. Then, the lower bound of the GBD method, LBD, is updated to be 1.625. Now, since UBD-LBD<0.001, then the revised GBD approach terminates at {0.3125, 1} with f=1.625.

Iteration 5.2. Now let us see the relaxed problem in the subinterval [0.8, 1.6].

$$\begin{aligned} \min_{x,y} \quad & 2x+y \\ s.t. \quad & 2.53-2.4x-y \le 0 \\ & x+y \le 1.6 \\ & 0.8 \le x \le 1.6 \\ & y=\{0,1\} \end{aligned}$$

By virtue of the same GBD procedures as above, we can get the global solution of this problem at {1.054, 0} with *f*=2.018.

Iteration 6. Now the iteration counter of the branch and bound algorithm increases by one, i.e. k=1, and the lower bound of the branch and bound algorithm, γ_1, is updated to be 1.625 in the subinterval [0.0, 0.8]. Since above lower bound is less than the upper bound by a value greater than 0.001, then this subinterval will be further fathomed by the depth-first branching rule. Here, we neglect some detailed steps of the branch and bound, and directly jump back to the *Step 2* for deeper search. Since the subinterval [0.0, 0.8] will be divided into two subintervals [0.0, 0.4] and [0.4, 0.8], then we describe the procedures in the first subinterval so as to demonstrate how the algorithm can remove the infeasible region in the GBD step.

Step A.5. The MINLP problem in the subinterval [0.0, 0.4] is given as

$$\begin{aligned} \min_{x,y} \quad & 2x+y \\ s.t. \quad & 1.25-x^2-y \le 0 \\ & x+y \le 1.6 \\ & 0 \le x \le 0.4 \\ & y=\{0,1\} \end{aligned}$$

Then its relaxed problem can be obtained as

$$\begin{aligned} \min_{x,y} \quad & 2x + y \\ s.t. \quad & 1.25 - 0.4x - y \le 0 \\ & x + y \le 1.6 \\ & 0 \le x \le 0.4 \\ & y = \{0,1\} \end{aligned}$$

Step A.5.1, fix y=0, we get the resulting convex (linear) problem, as

$$\begin{aligned} \min_{x} \quad & 2x \\ s.t. \quad & 1.25 - 0.4x \le 0 \\ & x \le 1.6 \\ & 0 \le x \le 0.4 \end{aligned}$$

However, this problem is infeasible. Then, l=1, and we solve the following feasibility problem, as

$$\begin{aligned} \min_{x,\alpha} \quad & \alpha_1 + \alpha_2 \\ s.t. \quad & 1.25 - 0.4x - \alpha_1 \le 0 \\ & x - 1.6 - \alpha_2 \le 0 \\ & -\alpha_1 \le 0 \\ & -\alpha_2 \le 0 \\ & -x \le 0 \\ & x - 0.4 \le 0 \end{aligned}$$

By using the KKT condition of above problem, we get the Lagrange multipliers as μ^1=1.0, $\mu^2=\mu^3=\mu^4=\mu^5$=0.0, μ^6=0.4.

Step A.5.2. Since p=0, then we solve the following feasibility problem

$$\begin{aligned} \min_{y,\beta} \quad & \beta \\ s.t. \quad & \min_{0 \le x \le 0.4} \{1.25 - 0.4x - y\} \le \beta \\ & \beta \ge 0 \\ & y = \{0,1\} \end{aligned}$$

We get the solution y=1. Then go to *Step 2.1*, we have the resulting primal problem, as

Step A.5.1.1

$$
\begin{aligned}
\min_{x} \quad & 2x+1 \\
s.t. \quad & 1.25-0.4x-1 \le 0 \\
& x+1 \le 1.6 \\
& 0 \le x \le 0.4
\end{aligned}
$$

However, this problem is infeasible. Then, l=2, and we solve the following feasibility problem, as

$$
\begin{aligned}
\min_{x,\alpha} \quad & \alpha_1+\alpha_2 \\
s.t. \quad & 0.25-0.4x-\alpha_1 \le 0 \\
& x-0.6-\alpha_2 \le 0 \\
& -\alpha_1 \le 0 \\
& -\alpha_2 \le 0 \\
& -x \le 0 \\
& x-0.4 \le 0
\end{aligned}
$$

By using the KKT condition of above problem, we get the Lagrange multipliers as μ^1=1.0, $\mu^2=\mu^3=\mu^4=\mu^5$=0.0, μ^6=0.4.

Step A.5.2.2 Since p=0, and l=2, then we solve the following feasibility problem

$$
\begin{aligned}
\min_{y,\beta} \quad & \beta \\
s.t. \quad & \min_{0\le x\le 0.4} \{1.25-0.4x-y\} \le 0 \\
& \min_{0\le x\le 0.4} \{1.25-0.4x-y\} \le \beta \\
& \beta \ge 0 \\
& y=\{0,1\}
\end{aligned}
$$

But, this problem is infeasible. Then, by virtue of Theorem 2.4, we know that the original MINLP problem over the current subinterval [0.0, 0.4] is infeasible too. Hence, there is no need to further fathom this subinterval, then it is labeled as a fathomed node in the progress of the algorithm. Now, we still have two subintervals over which the further searching is needed, i.e. [0.4, 0.8] and [0.8, 1.6] in order to know the location of the global minimum. In fact, some subintervals in [0.8, 1.6] will be removed in the following iterative steps since they generate lower bounds, which are greater than the incumbent best upper bound in the branch and bound framework. Finally, the algorithm terminates at the global solution {0.5, 1} with the minimal objective function value being 2.

4. Conclusion

A hybrid branch and bound and revised General Benders Decomposition global optimization method is proposed in this paper for some nonconvex MINLP problems. The twice-differentiable condition for the continuous function of the MINLP problems are used to construct the valid convex quadratic underestimation function over a simplex in order to overcome the nonconvexity in the continuous domain. Then, the global solution of the MINLP problems often encountered in chemical processes can be identified provided that the favorable structure of the combinatorial features of the continuous domain and binary domain can ensure the convergence of the revised GBD on the global solution of the relaxed MINLP problem generated over the continuous domain in each iteration of the branch and bound algorithm. In this paper, the separable structure type, i.e. $f(\mathbf{x})+c^T\mathbf{y}$, the bilinear type, i.e. $\mathbf{x}^T\mathbf{y}$, and partly linear type, i.e. $\mathbf{y}^T\mathbf{f}(\mathbf{x})$, are analyzed to resolve the nonconvexity arisen by the joint continuous and binary domains. Hence, the revised GBD method can not only identify the global solution of the relaxed MINLP problem reliably when it is feasible, but also detect the infeasibility over the current subsimplex effectively. Consequently, that subsimplex is removed with the progress of the branch and bound framework efficiently. A very simple, but typical example with concave continuous function and separable combinatorial structure is presented in this paper to demonstrate all possibilities discussed in the hybrid branch and bound and revised GBD algorithm. The efficiency comparison of this hybrid approach with others, especially with the complete branch and bound type, needs the implementations for large process design and synthesis MINLP problems in the chemical engineering field, which are still under development.

Acknowledgement

Y. Zhu gratefully thanks the financial support from the Japan Society for the Promotion of Science (JSPS) and the JSPS Fellow Project numbered 01040, and the always help and encouragement from Prof. W.D.Seider (University of Pennsyvania) and Prof. C.A. Floudas (University of Princeton) for Y. Zhu's research work in the field of the process system engineering are sincerely appreciated.

Appendix. Underestimators for different nonconvex functions

I. Underestimator for 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 and the constraints.

II. Underestimator for concave function structure

For the concave function structures, denoted by $f^{-C}(\mathbf{x})$, and $\mathbf{S}$ being a simplex generated by the vertices $\mathbf{V}^0$, $\mathbf{V}^1$, ..., $\mathbf{V}^n$, i.e. $\mathbf{S}=\left\{\mathbf{x}\in\Re^n:\mathbf{x}=\sum_{i=1}^{n+1}\lambda_i\mathbf{V}^i,\lambda_i\geq 0,\sum_{i=1}^{n+1}\lambda_i=1\right\}$, then the convex envelope of $f^{-C}(\mathbf{x})$ over $\mathbf{S}$ is the affine function $L^{-C}(\mathbf{x})=\mathbf{b}^T\mathbf{x}+c$ which is uniquely determined by the system of linear equations $f^{-C}(\mathbf{V}^i)=\mathbf{b}^T\mathbf{V}^i+c$ for $i=0,\dots,n$.

III. Underestimator for general quadratic function

The general quadratic function is presented as

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

Since $\mathbf{H}_{f}(\mathbf{x}) = \mathbf{Q}$ is constant matrix, then we have the *diagonal underestimation matrix*, Δ, constructed 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

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

Then, we have the quadratic underestimation function as

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

where, the linear and constant coefficients, i.e. $(\mathbf{b}, c)$, can be uniquely determined by the system of linear equations $f^{Q}(\mathbf{V}^{i}) - \mathbf{V}^{i^{T}}\Delta\mathbf{V}^{i} = \mathbf{b}^{T}\mathbf{V}^{i} + c$ for $i = 0, \ldots, n$, and $\mathbf{V}^{0}$, $\mathbf{V}^{1}$, …, $\mathbf{V}^{n}$ are the vertices of the simplex $\mathbf{S}$.

IV. Underestimator for twice-differentiable nonconvex function

For the twice-differentiable nonconvex function, denoted by $f^{NC}(\mathbf{x})$, we have the *diagonal underestimation matrix*, Δ, constructed as

$$a \geq \max\left\{ 0, \frac{1}{2}\max_{i, \mathbf{x} \in \mathbf{S}} \lambda_{i}^{f^{NC}}(\mathbf{x}) \right\}$$

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

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

Then, we have the quadratic underestimation function as

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

where, the linear and constant coefficients, i.e. $(\mathbf{b}, c)$, can be uniquely determined by the system of linear equations $f^{NC}(\mathbf{V}^i) - \mathbf{V}^{i^T} \Delta \mathbf{V}^i = \mathbf{b}^T \mathbf{V}^i + c$ for $i = 0, \ldots, n$, and $\mathbf{V}^0$, $\mathbf{V}^1$, …, $\mathbf{V}^n$ are the vertices of the simplex $\mathbf{S}$.

References

Adjiman, C.S., I.P. Androulakis, and C.A. Floudas. (1997). Global optimization of MINLP problems in process synthesis and design. *Comput. Chem. Eng.*, **21**, S445-S.

Adjiman, C.S., C.A. Schweiger, and C.A. Floudas. (1998). Mixed-integer nonlinear optimization in process synthesis. *Handbook of Combinatorial Optimization*, p. 1-76. D.-Z. Du and P.M. Pardalos, eds., Kluwer Academic Publisher, Dordrecht, The Netherlands.

Adjiman, C.S., I.P. Androulakis, and C.A. Floudas. (2000). Global optimization of mixed-integer nonlinear problems. *AIChE J.*, 46, 1769-1797.

Benders, J.F. (1962). Partitioning procedures for solving mixed-variables programming problems. Numer. Math., **4**, 238-252.

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

Floudas, C.A., (1995). *Nonlinear and mixed-integer optimization, fundamentals and applications*, Oxford Univ. Press, New York.

Floudas, C.A., (2000). *Deterministic global optimization: theory, methods and applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Geoffrion, A.M. (1972). Generalized Benders decomposition. *J. Opt. Theory Appl.*, 10, 237-260.

Grossman, I.E., ed., (1996). *Global optimization in Engineering Design*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Grossman, I.E. and A.W., Westerberg. (2000). Research challenges in process systems engineering. *AIChE J.*, **46**, 1700-1703.

Grossman, I.E. (2001). Review of nonlinear mixed-integer and disjunctive programming techniques for process system engineering.

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

Kocis, G.R., and I.E. Grossman. (1988). Global optimization of nonconvex mixed-integer nonlinear programming (MINLP) problems in process synthesis. *Ind. Eng. Chem. Res.*, **27**, 1407-1421.

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

Nemhauser, G,L. and L.A.wolsey. (1999). *Integer and combinatorial optimization*, John Wiley & Sons, New York.

Ryoo, H.S., and N.V. Sahinidis. (1995). Global optimization of nonconvex NLPs and MINLPs with applications in process design. *Comput. Chem. Eng.*, **19**, 551-566.

Smith, E.M.B., and C.C. Pantelides. (1996). Global optimization of general process models. *Global optimization in Engineering Design*, p. 355-386. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Smith, E.M.B., and C.C. Pantelides. (1999). A symbolic reformulation/spatial branch-and-bound algorithm for the global optimization of nonconvex MINLPs. *Comput. Chem. Eng.*, **23**, 457-478.

Tuy, H. (1998). *Convex analysis and global optimization*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Vaidyanathan, R., and M. El-Halwagi. (1996). Global optimization of nonconvex MINLPs by interval analysis. *Global optimization in Engineering Design*, Chapter 6. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Westerlund, T., H. Akrifvars, I. Harjunkoski, and R. Porn. (1998). An extended cutting plane method for a class of nonconvex MINLP problems. *Comput. Chem. Eng.*, **22**, 357-365.

Zamora, J.M., and I.E. Grossmann. (1998). A global MINLP optimization algorithm for the synthesis of heat exchanger networks with no stream splits. *Comput. Chem. Eng.*, **22**, 367-384.

Zhu, Y., and 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., and Inoue, K. (2001). Calculation of chemical and phase equilibrium based on stability analysis by QBB algorithm: Application to NRTL equation. *Chem. Eng. Sci.*, **56**, 6915-6931.

Zhu, Y. and T., Kuno. (2001). A global optimization method, QBB, for twice-differentiable nonconvex optimization problem. *Journal of Global Optimization*, (Submitted). Technical Report, ISE-TR-01-182, Institute of Information Sciences and Electronics, University of Tsukuba, Japan.