Global Optimization of Nonconvex MINLP by a Hybrid Branch-and-Bound and Revised General Benders Decomposition Approach

Yushan Zhu*,† and Takahito Kuno‡

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

Mixed-integer nonlinear programming, MINLP, has played a crucial role in chemical process design via superstructures that always involve discrete and continuous variables. In this paper, a global optimization algorithm for nonconvex MINLP problems is developed by addressing the nonconvexity caused by the nonconvex continuous functions with a convex quadratic underestimator within a branch-and-bound framework, as well as the joint problem caused by the mixed natures of integer and continuous variables through 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 to the global solution of the relaxed MINLP subprobelm over each subregion generated in the above framework is guaranteed by the convex underestimation functions in terms of the twice-differentiable assumptions of the continuous functions and the 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 the lower bound, and the reliability of the infeasibility detection. 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 detection in each lower-bounding problem.

1. Introduction

The most significant contribution of mathematical approaches comes from their ability to incorporate many alternative structures into a single problem. This is achieved through the introduction of integer variables, which leads to the formulation of a mixed-integer nonlinear programming (MINLP) problems.[7,8,11,12] The global solutions of mixed-integer linear programming (MILP)[16] problems or convex MINLP problems can be located by the Benders decomposition method[5] or general Benders decomposition method.[9] However, these approaches cannot be applied directly to nonconvex MINLP problems because they always identify only local optima owing to the nonconvexity of the nonlinear functions and the joint structure. The early attempts to solve nonconvex MINLP problems contributed by Kocis and Grossmann[14] and Floudas, Aggarwal, and Ciric[6] were stimulated by the MINLP problems encountered in process synthesis and design. The branch-and-reduce algorithm of Ryoo and Sahinidis,[17] eventually developed into a package named BARON, relies on existing underestimation techniques, such as those proposed by McCormick,[15] 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[21] uses interval arithmetic to bound the function values within a branch-and-bound framework, where the domain size is reduced by partitioning and searching is performed by applying upper bound, infeasibility, monotonicity, nonconvexity, and lower bound tests, as well as the distrust region method. Smith and Pantelides[18,19] 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.[22] used an extended cutting plane algorithm to solve problems involving pseudoconvex functions. Zamora and Grossman[23] proposed more specialized algorithms for certain classes of applications, such as heat-exchanger networks. Adjiman et al.[2] presented an excellent review of these algorithms and also briefly introduced two broadly applicable global optimization approaches based on the αBB algorithm, i.e., SMIN-αBB and GMIN-αBB.[1] Complete descriptions of the theoretical basis of these two algorithms and computational experiments are provided in Floudas[8] and Adjiman et al.,[3] which enable the determination of the most adequate implementation decisions.

A novel convex underestimation technique developed in the QBB algorithm framework[24−26] is applied here to address the nonconvexities arising from continuous variables; then, the resulting 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 nonconvex MINLP problems proceeds 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 in detail by the projection approach in the GBD method. The monotonicity of the lower-

* Corresponding author. Tel.: +81-298-536791. Fax: +81-298-535206. mail: yszhu@syou.is.tsukuba.ac.jp.
† University of Tsukuba.
‡ Chinese Academy of Sciences.

10.1021/ie0200813 CCC: $25.00
Published on Web 01/10/2003

bound functions constructed by quadratic-function-based underestimators, infeasibility detection, and asymptotic convergence are presented to provide a complete theoretical guarantee. Three kinds of mixed-integer function structures widely used in chemical processes are analyzed and applied in the 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 to the global solution of nonconvex MINLP problems.

2. Hybrid Branch-and-Bound and Revised General Benders Decomposition Algorithm

A general nonconvex mixed-integer nonlinear programming (MINLP) problem can be formulated as follows

Problem P

$$\min_{\mathbf{x},\mathbf{y}} f(\mathbf{x},\mathbf{y})$$
$$\text{s.t.}\quad g_i(\mathbf{x},\mathbf{y}) \le 0 \qquad i = 1, 2, ..., m$$
$$\mathbf{x} \in \mathbf{S} \subseteq \mathcal{R}^n$$
$$\mathbf{y} \in \mathbf{Y} = \{0, 1\}^q$$

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. In addition, the functions

$$f\colon\ \mathcal{R}^n \times \mathcal{R}^q \to \mathcal{R},$$
$$\mathbf{g}\colon\ \mathcal{R}^n \times \mathcal{R}^q \to \mathcal{R}^m$$

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

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

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

$$V_g = \{\mathbf{y} \in B^q\colon\ \mathbf{g}(\mathbf{x},\mathbf{y}) \le \mathbf{0}, \text{ for some } \mathbf{x} \in \mathbf{S}\}$$

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[9] can be applied. However, the essential difference between problem P and the others 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 as in Floudas.[7]

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 fixed. However, the objective and constrained functions in problem P are assumed to be only twice-differentiable, rather than convex. Thus, we have to consider the nonconvexities that arise not only from the joint $\mathbf{x}-\mathbf{y}$ domain structure, but also from the continuous variables $\mathbf{x}$ even after the binary variables, i.e., the complicating variables $\mathbf{y}$, are fixed. In this paper, a hybrid branch-and-bound and GBD framework is constructed to treat the above-described complications. In fact, the mixed nature 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 technique developed in the QBB algorithm framework.[24-26]

2.1. Convex Relaxation of the MINLP Problem. Because the vector $\mathbf{y}$ of variables is defined as containing the complicating variables in GBD method, we have the following definition to characterize its use in a branch-and-bound framework:

Definition 2.1. Given any function $f(\mathbf{x},\mathbf{y})$, $f\colon\ \mathcal{R}^n \times \mathcal{R}^q \to \mathcal{R}$, where $\mathbf{x}$ represents a vector of n continuous variables and $\mathbf{y}$ represents a vector of q binary variables, that is continuously twice-differentiable for each fixed $y \in \mathbf{Y} = \{0, 1\}^q$, the function $F(\mathbf{x},\mathbf{y})$, $F\colon\ \mathcal{R}^n \times \mathcal{R}^q \to \mathcal{R}$ for $\mathbf{x} \in \mathbf{S} \subseteq \mathcal{R}^n$ 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 the above definition, because the relationship between the continuous and binary variables in function $f(\mathbf{x},\mathbf{y})$ is implicit, the specific structure of the convex relaxation in terms of above definition is unknown. However, for most of chemical engineering processes, the relevant MINLP problems[5,7] can be formulated in a much more explicit form as

Problem P(ChE)

$$\min_{\mathbf{x},\mathbf{y}} f(\mathbf{x}) + \mathbf{c}^{\mathrm{T}}\mathbf{y}$$
$$\text{s.t.}\quad \mathbf{g}_i(\mathbf{x}) + \mathbf{C}_i^{\mathrm{T}}\mathbf{y} \le \mathbf{0} \qquad i = 1, ..., m$$
$$\mathbf{x} \in \mathbf{S} \subseteq \mathcal{R}^n$$
$$\mathbf{y} \in \mathbf{Y} = \{0, 1\}^q$$

where $f(\mathbf{x})$ and $g_i(\mathbf{x})$ for $i = 1, ..., m$ are continuously twice-differentiable functions; $\mathbf{c}$ is a constant vector belonging to $\mathcal{R}^q$, and $\mathbf{C}$ is a constant matrix belonging to $\mathcal{R}^{m\times q}$. Hence, the binary variables appear only in linear form in the objective and constrained functions without nonlinear terms becoming involved with the continuous variables. In fact, most applications described by Grossmann[12] can be taken as some special cases of the above formulation by forcing the continuous functions to be convex. Because the binary variables are separable from the continuous variables, the valid relaxation of the objective and constraint functions is dependent only on the valid underestimation construction of the twice-differentiable functions appearing in the above formulation. First, the following theorem states that there exists a valid convex underestimation function over a simplex for any twice-differentiable function.

Theorem 2.1.[20,13] There exists a convex underestimation function for any continuously twice-differentiable function over a simplex.

Proof. 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 semidefinite matrix for any $\mathbf{x} \in \mathbf{S}$, where $\mathbf{H}_f(\mathbf{x})$ represents 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 DC (difference of two convex functions[20]) formulation of $f(\mathbf{x})$. Because the third term on the RHS in the above formulation, i.e., $-\alpha||\mathbf{x}||^2$, is concave over the simplex $\mathbf{S}$, its convex envelope[13] can be expressed by an affine function $\mathbf{c}^T\mathbf{x} + b$ for all $\mathbf{x} \in \mathbf{S}$. Then, we have the following function $F(\mathbf{x})$

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

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

In fact, the above theorem implies a way to construct the underestimation function for any continuously twice-differentiable function over a simplex. Geometrically, the 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, such as that presented in the QBB algorithm for any twice-differentiable nonconvex optimization problem.[24−26] That is, there exists a convex quadratic function for any twice-differentiable function over a simplex that is also its valid underestimator, as stated in definition 2.2.

Definition 2.2. Given any nonconvex function $f(\mathbf{x})$: $S \rightarrow \mathcal{R}$, with $\mathbf{x} \in S \subseteq \mathcal{R}^n$ belonging to $\mathbf{C}^2$, the quadratic function 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 $\mathbf{x} \in \mathbf{S} \subseteq \mathcal{R}^n$ and $F(\mathbf{x}) = f(\mathbf{x})$, holds at all vertexes of $\mathbf{S}$. The a_i's are nonnegative scalars that are large enough such that $F(\mathbf{x}) \leq f(x)$, $\forall\ \mathbf{x} \in \mathbf{S}$.

The methods for determining the quadratic, linear, and constant coefficients of eq 1 are briefly presented in the Appendix of this paper, but a detailed introduction is offered in Zhu and Kuno.[25,26] Note that a simpler form of the above quadratic function uses a uniform quadratic coefficient, so tha eq 1 becomes a single-parameter underestimator. After replacing all of the twice-differentiable functions by their corresponding convex underestimators over the simplex in problem P(ChE), we obtain its relaxed formulation as

Problem P(ChE)R

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

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

Problem PR

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

where the functions $F(\mathbf{x},\mathbf{y})$ and $\mathbf{G}_i(\mathbf{x},\mathbf{y})$ satisfy definition 2.1. Also, let D_G be a subset of $\mathcal{R}^n$ defined by

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

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

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

Then, for a given mixed-integer nonlinear programming problem,[7,9] the following conditions must hold for the GBD approach to be applied for the binary variables:

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

$$F\text{: } \mathcal{R}^n \times \mathcal{R}^q \rightarrow \mathcal{R}$$

$$\boldsymbol{G}\text{: } \mathcal{R}^n \times \mathcal{R}^q \rightarrow \mathcal{R}^m$$

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

Remark 1. This condition holds trivially by virtue of definitions 2.1 and 2.2.

Condition 2. The set

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

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

Remark 2. This condition holds because 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 $y \in \mathbf{Y} \cap V_G$, one of the following two cases holds:

Case i. The resulting problem PR has a finite solution and an optimal multiplier vector for the inequalities.

Case ii. The resulting problem PR is unbounded; that is, its objective function value goes to $-\infty$.

Remark 3. Problem PR is a relaxation of the original problem P that is always overestimated. Consequently, the above two cases alone are not sufficient to include all possibilities because the resulting problem PR might be infeasible for some fixed $y \in \mathbf{Y} \cap \mathbf{V}$. In fact, later, we can see how this case will be used frequently to remove the relaxed but infeasible region of the original problem. Thus, we introduce this additional case into the above condition as case iii.

Case iii. The resulting problem PR is infeasible.

For practical applications, the objective function is always bounded. Thus, in this paper, we consider only cases i and iii, that is, the resulting problem PR is feasible or infeasible. The following theorem states the relationship between the optimal solution of the relaxed problem PR and the original problem P, which is derived by an approach similar to that of Zhu and Kuno[26] for nonconvex continuous problems.

Theorem 2.2.[26] 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 $y \in \mathbf{Y} \cap V_g$ can be computed according to $\mu(\mathbf{S}) = F^*$, 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 $\mathbf{y} \in \mathbf{Y}$, because $\mathbf{G}_i(\mathbf{x},\mathbf{y})$ is a convex underestimator of $\mathbf{g}_i(\mathbf{x},\mathbf{y})$, i.e., $\mathbf{G}_i(\mathbf{x},\mathbf{y}) \le \mathbf{g}_i(\mathbf{x},\mathbf{y})$, we have $\mathbf{G}_i(\mathbf{x},\mathbf{y}) \le \mathbf{g}_i(\mathbf{x},\mathbf{y}) \le 0$ for any $\mathbf{x} \in D_g$, that is, $\mathbf{x} \in D_G$. Then, we obtain $\mathbf{S} \cap D_g \subseteq \mathbf{S} \cap D_G$ for any $\mathbf{y} \in \mathbf{Y}$ by noting that $D_g \subseteq D_G$. By using the same approach, we obtain $\mathbf{Y} \cap V_g \subseteq \mathbf{Y} \cap V_G$ 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 the fact that $F(\mathbf{x},\mathbf{y}) \le f(\mathbf{x},\mathbf{y})$ for $\mathbf{x} \in \mathbf{S} \cap D_g$ and for any $\mathbf{y} \in \mathbf{Y} \cap V_g$, we have

$$F^* = \min_{\mathbf{x},\mathbf{y}}\{F(\mathbf{x},\mathbf{y}),\ \mathbf{x} \in \mathbf{S} \cap D_G \text{ and } \mathbf{y} \in \mathbf{Y} \cap V_G\} \le F(\mathbf{x},\mathbf{y})$$

$$\mathbf{x} \in \mathbf{S} \cap D_G \text{ and } \mathbf{y} \in \mathbf{Y} \cap V_G \le f(\mathbf{x},\mathbf{y}),\ \mathbf{x} \in \mathbf{S} \cap D_G \text{ and } \mathbf{y} \in \mathbf{Y} \cap V_G \le f(\mathbf{x},\mathbf{y}),\ \mathbf{x} \in \mathbf{S} \cap D_g \text{ and } \mathbf{y} \in \mathbf{Y} \cap V_g$$

This equation 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 might be infeasible even when the relaxed problem PR is feasible because the latter is always overestimated in practical implementations. The above theorem provides only a qualitative method of obtaining a valid lower bound of $f(\mathbf{x},\mathbf{y})$ over $\mathbf{S} \cap D_g$ for any $\mathbf{y} \in \mathbf{Y} \cap V_g$. Because the constraint set for the binary variables is implicit and the complete enumeration is 2^q in the worst case, we can see that this number grows exponentially and becomes drastically large as the number of binary variables increases. Another difficulty of using the above theorem to obtain a lower bound is that the resulting problem PR is feasible for some $\mathbf{y} \in \mathbf{Y} \cap V_G$, but infeasible for others. Thus, the proposed scheme should be capable of discriminating all feasible cases and finding the optimal solution among them. To overcome all of these difficulties, the projection idea is used by virtue of the dual representation and relaxation in the GBD method. First, however, 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, $\mu(\mathbf{S}^2) \ge \mu(\mathbf{S}^1)$.

(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 $\mathscr{R}^n$ defined by

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

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

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

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

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

Because $\mathbf{S}^2 \subset \mathbf{S}^1$, and by virtue of the argument in proposition 2.2.2 of Zhu and Kuno,[26] for any $\mathbf{y} \in \mathbf{Y}$, we have

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

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

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

(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 $\mathbf{y} \in \mathbf{Y} \cap V_g$. Therefore, this problem has an optimal solution, which implies that $\mu(\mathbf{S}^0) > -\infty$.

2.2. 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 the removsl of that simplex with the progress of the branch-and-bound algorithm. However, a complication arises as a result of the joint natures of the 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. However, it is quite inefficient if the integer variable number is somewhat large. Thus, a more intelligible approach is to use the Lagrange relaxations presented by Benders[4] and Geoffrion[9] for MILP and MINLP problems, respectively. In this section, the general Benders decomposition, i.e., the GBD method, is revised to address the above-mentioned difficulty, which consists of two basic operations, the primal problem and the master problem, to obtain the upper and lower bounds, respectively, of the relaxed problem PR($\mathbf{S}^k$) at each iteration over the current simplex.

2.2.1. Primal Problem of the Relaxed Problem PR($\mathbf{S}^k$). 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 PR[$\mathbf{S}^k(\mathbf{y}^t)$], at iteration t over subsimplex $\mathbf{S}^k$ is given by

Problem PR[$\mathbf{S}^k(\mathbf{y}^t)$]

$$\min_{\mathbf{x}} F(\mathbf{x},\mathbf{y}^t)$$

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

$$\mathbf{x} \in \mathbf{S}^k \subseteq \mathscr{R}^n$$

Obviously, this problem is convex because of condition 1 in the former section. However, we have to distinguish two possible cases, i.e., feasible and infeasible, according to the relaxed problem PR($\mathbf{S}^k$). If the relaxed problem PR($\mathbf{S}^k$) is infeasible, of course the primal problem PR[$\mathbf{S}^k(\mathbf{y}^t)$] is infeasible too. However, it is possible that the relaxed problem PR($\mathbf{S}^k$) might be infeasible even when the primal problem PR[$\mathbf{S}^k(\mathbf{y}^t)$] is feasible, because an unsuitable binary variable vector might be chosen to be fixed. However, these two cases cannot immediately be distinguished now, so they are treated together here. When the primal problem PR[$\mathbf{S}^k(\mathbf{y}^t)$] is feasible, its solution provides information on $\mathbf{x}^t$; $F(\mathbf{x}^t,\mathbf{y}^t)$, which is the upper bound of the relaxed problem PR($\mathbf{S}^k$); 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 PR[$\mathbf{S}^k(\mathbf{y}^t)$] is determined by the NLP solver to be infeasible, then perturbation theory is used to generate a maximal integer cut to remove this combination of binary variables. The l_1 minimization problem, i.e., the sum of constraint violation, can be formulated as

Problem FP(α)

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

Note that this minimization problem, denoted as FP-(α), is convex over all variables and that 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 the above feasibility problem provides information on the Lagrange multiplier vector for the inequality constraints, which are denoted as μ so as to distinguish them from the feasible case. Then, the Lagrange function for the 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. Master Problem of the Relaxed Problem PR($\mathbf{S}^k$). 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($\mathbf{S}^k$) onto $\mathbf{y}$ space over the dual representation of V_G. First, we consider the projection of the relaxed problem PR($\mathbf{S}^k$) onto $\mathbf{y}$ space. Let $v(\mathbf{y})$ be defined as

$$\begin{aligned} &v(\mathbf{y}) = \inf_{\mathbf{x}} F(\mathbf{x},\mathbf{y}) \\ &\text{s.t.}\quad \mathbf{G}_i(\mathbf{x},\mathbf{y}) \le \mathbf{0} \qquad i = 1, 2, ..., m \\ &\qquad\ \mathbf{x} \in \mathbf{S}^k \subseteq \mathcal{R}^n \end{aligned}$$

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

Problem PR($\mathbf{y}$)

$$\begin{aligned} &\min_{\mathbf{y}} v(\mathbf{y}) \\ &\text{s.t.}\quad \mathbf{y} \in \mathbf{Y} \cap V_G, \end{aligned}$$

which is denoted as problem PR($\mathbf{y}$). It should be noted that the definition of $v(\mathbf{y})$ is infinite with respect to $\mathbf{x}$ because, for any given $\mathbf{y}$, the inner optimization problem might be unbounded, and its value corresponds to the optimal value of the relaxed problem PR($\mathbf{S}^k$) for fixed $\mathbf{y}$. Thus, problem PR($\mathbf{y}$) is the projection of the relaxed problem PR($\mathbf{S}^k$) onto $\mathbf{y}$ space. It can be shown that this projected problem PR($\mathbf{y}$) is equivalent to the relaxed problem PR($\mathbf{S}^k$).[7,9] Because we always assume that problem P has a solution, the unbounded case need not be considered in this paper according to proposition 2.1. Thus, the dual representations of V_G and $v(\mathbf{y})$ are presented as follows

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

and

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

According to the strong duality theorem,[9] these two dual representations are satisfied by virtue of conditions 1−3. However, the Lagrange functions used in the above dual representations involve maximizations over all multipliers. Hence, their relaxation will represent only the lower bounds of the 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

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

and

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

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

Problem PR($\mathbf{y}^t$)

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

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

Problem FP(β)

$$\min_{\mathbf{y}\in\mathbf{Y},\beta} \beta$$

$$\text{s.t.} \quad \min_{\mathbf{x}\in\mathbf{S}^k} L(\mathbf{x},\mathbf{y},\mu^l) \le 0 \qquad l = 1, ..., t-1$$

$$\min_{\mathbf{x}\in\mathbf{S}^k} L(\mathbf{x},\mathbf{y},\mu^t) \le \beta$$

$$\beta \ge 0$$

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

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($\mathbf{S}^k$) and vice versa.

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

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

where $\mathbf{x}^*$ is the optimal solution of $\min_{\mathbf{x}\in\mathbf{S}^k} L(\mathbf{x},\mathbf{y}^*,\mu^*) \le 0$. 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 set

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

is empty. Then, the relaxed problem PR($\mathbf{S}^k$) is infeasible because 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 for the Revised GBD for the Relaxed Problem PR($\mathbf{S}^k$). The revised GBD procedure for the relaxed problem PR($\mathbf{S}^k$) can now be stated formally with the consideration that the above relaxed problem is completely infeasible over the current subsimplex. It should be noted here that Geoffrion[9] did not include the infeasible case, but we assume that the relaxed problem always has a bounded optimal value if it is feasible.

Procedure for the Revised GBD. Step 0. Initialization. Set the current upper bound (UBD) to 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 to $p = 0$ and $l = 0$, respectively. Then, select the convergence tolerance $\epsilon_c \ge 0$ and feasibility tolerance $\epsilon_f \ge 0$. 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 PR[$\mathbf{S}^k(\mathbf{y}^t)$]. If the NLP solver verifies that the above problem is feasible, then set $p \leftarrow p + 1$, and obtain the optimal primal solution $\mathbf{x}^p$ and the optimal multiplier vector λ^p. 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(α) to obtain the multiplier vector μ^q.

Step 2. Solve the Relaxed Master Problem. If $p \ge 1$, solve the relaxed master problem PR($\mathbf{y}^t$) as

$$\min_{\mathbf{y}\in\mathbf{Y},y_0} y_0$$

$$\text{s.t.} \quad \min_{\mathbf{x}\in\mathbf{S}^k} L(\mathbf{x},\mathbf{y},\lambda^i) \le y_0 \qquad i = 1, ..., p$$

$$\min_{\mathbf{x}\in\mathbf{S}^k} L(\mathbf{x},\mathbf{y},\mu^j) \le 0 \qquad j = 1, ..., l$$

Note here that the second constraint set vanishes if $l = 0$. Then, we obtain y_0 and $\mathbf{y}^f$ and set the current lower bound LBD = max{LBD, y_0}. If UBD − LBD ≤ ϵ, then the iteration is terminated, and the solutions of the relaxed problem PR($\mathbf{S}^k$) are {$\mathbf{x}^p$, $\mathbf{y}^p$} and UBD. Otherwise, set $t \leftarrow t + 1$, let $\mathbf{y}^t = \mathbf{y}^f$, and return to step 1. If $p = 0$, solve the relaxed feasibility problem FP(β) as

$$\min_{\mathbf{y}\in\mathbf{Y},\beta} \beta$$

$$\text{s.t.} \quad \min_{\mathbf{x}\in\mathbf{S}^k} L(\mathbf{x},\mathbf{y},\mu^i) \le 0 \qquad i = 1, ..., l-1$$

$$\min_{\mathbf{x}\in\mathbf{S}^k} L(\mathbf{x},\mathbf{y},\mu^l) \le \beta$$

$$\beta \ge 0$$

We obtain β and $\mathbf{y}^{if}$. If $\beta \ge \epsilon_f$, terminate the procedure, because the relaxed problem PR($\mathbf{S}^k$) is infeasible over the current subsimplex. Otherwise, set $t \leftarrow t + 1$, let $\mathbf{y}^t = \mathbf{y}^{if}$, and return to step 1. □

The following theorem ensures that the above revised GBD algorithm converges finitely no matter whether the relaxed problem PR($\mathbf{S}^k$) is feasible.

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

Proof. If the relaxed problem PR($\mathbf{S}^k$) is infeasible, then the above GBD procedure has a finite termination thanks to the integral finiteness of $\mathbf{Y}$ and the fact that no $\mathbf{y}^{if}$ can repeat itself in the solution of the relaxed feasibility problem FP(β), which is manifested by introducing an additional integer cut obtained by 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($\mathbf{S}^k$) is feasible, then the feasible $\mathbf{y}^f$ cannot be repeated unless the convergence criterion is satisfied,[9] which is implied by introducing an additional optimality integer cut generated from the feasible primal problem into the constraint set of the relaxed master problem PR($\mathbf{S}^k$). Finally, the worst performance of the above procedure could be the complete enumeration of the integer elements in $\mathbf{Y}$, which is finite. Then, the computational complexity of this revised GBD algorithm in terms of the problem input size is exponential, i.e., O(2^p), where p is the total number of the binary variables. □

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($\mathbf{y}^t$) or the relaxed feasibility problem FP(β). However, the determination of those solutions cannot be achieved in general, because they are always parametric functions of binary variable vector $\mathbf{y}$ obtained by the solutions of the inner optimization problems. Their determination method generally requires a global optimization approach, but a number of special structures exist for which the solutions of those inner optimization problems can be obtained explicitly as functions of the binary variable vector $\mathbf{y}$. For the MINLP problems widely encountered in chemical engineering processes, as described by problem P(ChE), the inner optimization problem can be explicitly obtained by its relaxed formulation using a further relaxation. Now, if we have the multiplier vectors 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

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

or

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

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

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

or

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

According to definition 2.2, we know that all of the relaxed continuous functions are convex, so their minima over the current subsimplex can be calculated by any NLP solver. Hence, by replacing those constraints in the master problems with the above further relaxations, we have the explicit formulation with respect to only the binary variable vector. The nonconvexity that arises from the joint continuous and binary variables needs to be handled to generate a valid and global integer cut described in problem PR($\mathbf{y}^t$) for the 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. Because problem PR($\mathbf{y}$) 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 because eqs 4 and 5 are valid relaxations for the above two equations, 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 problem P(ChE)R, because the relaxed master problem provides a global underestimation for problem P(ChE)R. In fact, eqs 6 and 7 are the corresponding formulations of eqs 4 and 5 for problem P(ChE)R, and it is obvious that these two constraints are linear with respect to $\mathbf{y}$. Thus, 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., $f(\mathbf{x}) + \mathbf{c}^{\mathrm{T}}\mathbf{y}$, to the bilinear type, i.e., $\mathbf{x}^{\mathrm{T}}\mathbf{y}$, and partly linear type, i.e., $\mathbf{y}^{\mathrm{T}}\mathbf{f}(\mathbf{x})$, according to the above theorem. For the latter two cases, 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 $\sum_{i=1}^{q} f_i(\mathbf{x})\mathbf{y}_i$, then its further relaxation is $\sum_{i=1}^{q} \min_{x}\{F_i(\mathbf{x})\}\mathbf{y}_i$, 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 for the hybrid branch-and-bound and revised GBD algorithm for 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 continuous variables and the upper-bound calculation over each subsimplex generated in the above partition process. For the branching procedure, the well-known simplicial partition often used in global optimization[24−26] is applied here. To guarantee convergence of the branch-and-bound algorithm to the global solution of 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 to construct two subsimplices for the potential searches. A detailed description of this kind of partition is presented by Zhu and Kuno[25,26] but is not given here because of space limitations.

For the calculation of a rigorous upper bound over the current simplex, the nonconvex MINLP problem is solved locally by the GBD method[9] for the problem P without any convexification. It should be noted here that an upper bound is not necessarily obtained in this way, especially when the current subsimplex is infeasible for the MINLP. However, during the progress of the branch-and-bound procedure, the obtained upper bound is updated so as to generate a nonincreasing sequence to converge on the global solution of the MINLP, and the concerned discussion about this convergence is presented in a latter section.

Now, we are in a position to describe the proposed branch-and-bound algorithm for solving the nonconvex MINLP problem based on the 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 pointed out in Zhu and Kuno,[25,26] this simplex can be constructed by using an outer approximation (OA) method according to the physical or insightful bounds of those continuous variables.

Hybrid Branch-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 priori or can 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 $(\mathbf{x}^{k,c}, \mathbf{y}^{k,c})$ 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 $\gamma_k = \min(\gamma_{k-1}, f_{\text{local}}^{k})$.

Step 3. Partitioning of Current Simplex. The current simplex, $\mathbf{S}^k$, is partitioned into the following two simplices ($r = 1, 2$)

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

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

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

Step 4. Convexify the MINLP inside Both Subsimplices $r = 1, 2$. The nonconvex 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 $F_{\text{sol}}^{k,r}$ is feasible and less than the current upper bound, γ_k, then it is stored along with the solution point $(\mathbf{x}_{\text{sol}}^{k,r}, \mathbf{y}_{\text{sol}}^{k,r})$.

Step 6. Update Iteration Counter k and Lower Bound μ_k. The iteration counter is increased by 1

$$k \leftarrow k + 1$$

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

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

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

Step 7. Update Current Point $(\mathbf{x}^{k,c}, \mathbf{y}^{k,c})$ 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

$$(\mathbf{x}^{k,c}, \mathbf{y}^{k,c}) = (x_{\text{sol}}^{k',r'}, y_{\text{sol}}^{k',r'})$$

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

Step 8. Check for Convergence. If $(\gamma_k - \mu_k) > \epsilon_c$, then return to step 2. Otherwise, ϵ_c convergence has been reached. The global minimum solution and solution point are given as

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

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

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 the nonconvexity caused by the natures of the joint continuous and binary variables is resolved in the revised GBD approach through the relaxation using the strong duality theory. Then, the global convergence of the above hybrid algorithm depends not only on the construction of the valid underestimators for any twice-differentiable continuous functions, but also on the favorable structures of the MINLP problem with respect to the continuous and binary variables. Hence, this algorithm is universally reliable not for any kind of nonconvex MINLP problem, but for those whose special structures 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., 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, so that subsimplex is removed immediately after knowing the infeasibility. The latter deletion is valid because 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 $(\mathbf{x}^k, \mathbf{y}^k)$ 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 $\{\mathbf{S}^j\}$ 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 by the following theorem.

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 $\lim_{j\to\infty} \mathbf{S}^j = \cap_{j=1}^{\infty} \mathbf{S}^j = \{\mathbf{x}^*\}$. 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 value of $\mathbf{x}^*$.

It should be noted that the above theorem does not claim to hold for any MINLP problem, because the nonconvexity that arises from the joint natures of the continuous and binary variables always leads to a 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 problem P(ChE). The proof of the above theorem can be attained by the classical convergence conditions[13,26] of the branch-and-bound framework through 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 obtained 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 often appeared in the literature as a typical test example.[14,6,17] To illustrate the global convergence of the proposed algorithm in this paper, let us describe all cases possibly happened during the iterations of this problem.

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

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

Iteration 1. $\epsilon_c = \epsilon_f = 0.001$, $k = 0$, $S = [0.0, 1.6]$, $\mu_0 = 100$, and $\gamma_0 = -100$.

Iteration 2. First, fix the binary variable y at 0. Then, we have the following NLP problem

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

Solving this nonconvex NLP, we obtain 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 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 the above two subintervals are obtained by replacing the concave functions by their convex envelopes in each subinterval given in iteration 5 as well as their solutions obtained by the revised GBD method.

Iteration 5. First, consider the relaxed problem in the subinterval [0.0, 0.8].

Iteration 5.1.

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

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

Step 1.0. UBD = 100, LBD = −100, $p = 0$, $l = 0$, $\epsilon_c = \epsilon_f = 0.001$, and $t = 1$.

Step 1.1. Fix $y = 0$ to obtain the convex (linear) problem

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

However, this problem is infeasible. Then, let $l = 1$, and solve the feasibility problem

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

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

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

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

We obtain the solution $y = 1$. Then, go to step 2.1 to solve the resulting primal problem.

Step 2.1.

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

Solving this convex (in fact, linear) problem, we obtain the minimum solution at $x = 0.3125$ with $f = 1.625$. Then, the upper bound of the GBD method, UBD, is updated to 1.625, and the Lagrange multipliers are $\lambda^1 = 2.5$ and $\lambda^2 = \lambda^3 = \lambda^4 = 0.0$. Now, set $p = 1$, and go to step 2.2.

Step 2.2. Because $p = 1$ and $l = 1$, we have the following master problem

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

Solving this problem, we obtain $y = 1$ and $y_0 = 1.625$. Then, the lower bound of the GBD method, LBD, is updated to 1.625. Now, because UBD − LBD<0.001, the revised GBD approach terminates at {0.3125, 1} with $f = 1.625$.

Iteration 5.2. Now, consider the relaxed problem in the subinterval [0.8, 1.6].

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

By virtue of the same GBD procedures as above, we can obtain 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 is increased by 1, i.e., $k = 1$, and the lower bound of the branch-and-bound algorithm, γ_1, is updated to 1.625 in the subinterval [0.0, 0.8]. Because the above lower bound is less than the upper bound by a value greater than 0.001, this subinterval will be further explored by the depth-first branching rule. Here, we neglect some detailed steps of the branch-and-bound technique and directly jump back to step 2 for a deeper search. Because the subinterval [0.0, 0.8] is divided into two subintervals [0.0, 0.4] and [0.4, 0.8], 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

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

Then, its relaxed problem can be obtained as

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

Step A.5.1. Fix $y = 0$ to obtain the convex (linear) problem

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

However, this problem is infeasible, so let $l = 1$, and solve the following feasibility problem

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

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

Step A.5.2. Because $p = 0$, we solve the feasibility problem

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

and obtain the solution $y = 1$. Then, we go to step 2.1, where we have the resulting primal problem, as described in step A.5.1.1.

Step A.5.1.1.

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

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

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

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

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

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

However, 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 also infeasible. Hence, there is no need to further examine this subinterval, so it is labeled as a searched node in the progress of the algorithm. For the subinterval [0.4, 0.8], the lower bound is obtained as 1.95 at {0.475, 1} using the revised GBD method. Now, we still have two subintervals over which further searching is needed, i.e., [0.4, 0.8] and [0.8, 1.6], to determine the location of the global minimum. According to the depth-first rule of the branching, the subinterval [0.4, 0.8] is chosen to be the current one, and the MINLP is formulated as

$$\min_{x,y} 2x + y$$
$$\text{s.t.} \quad 1.25 - x^2 - y \le 0$$
$$x + y \le 1.6$$
$$0.4 \le x \le 0.8$$
$$y = \{0, 1\}$$

Solving this problem locally, we obtain an updated upper bound of 2.0 at {0.5, 1} because this upper bound is less than the formerly found one, i.e., 2.236. In fact, the subinterval [0.8, 1.6] can be removed immediately from the following iterative steps because it yields a lower bound, namely, 2.018 calculated earlier, that is greater than the incumbent best upper bound, i.e., 2.0, in the branch-and-bound framework. As we have only one subinterval now, i.e., [0.4, 0.8], and its lower bound is less than the current best upper bound, this interval is further explored. For the two subintervals, i.e., [0.4, 0.6] and [0.6, 0.8], after branching, the latter one can be removed from further investigation because its lower bound calculated by the revised GBD method is 2.2, which is greater than the current best upper bound. However, the lower bound over the subinterval [0.4, 0.6] is 1.98 at {0.49, 1}, so this interval is further divided into the two subintervals [0.4, 0.5] and [0.5, 0.6]. Finally, the algorithm terminates at the global solution {0.5, 1} with the minimal objective function value being 2 over the subinterval [0.4, 0.5], and the underestimation problem over this subinterval is formulated as

$$\min_{x,y} 2x + y$$
$$\text{s.t.} \quad 1.45 - x - y \le 0$$
$$x + y \le 1.6$$
$$0.4 \le x \le 0.5$$
$$y = \{0, 1\}$$

For this simple example, the algorithm explored seven nodes altogether. One of them, i.e., [0.0, 0.4], was found to have an infeasible underestimation problem, and another two, i.e., [0.8, 1.6] and [0.6, 0.8], were found to have lower bounds greater than the current best upper bound.

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 functions of the MINLP problems is used to construct a valid convex quadratic underestimation function over a simplex to overcome the nonconvexity in the continuous domain. Then, the global solution of 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 to 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., $f(\mathbf{x}) + \mathbf{c}^T\mathbf{y}$; 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 that arises from the joint continuous and binary domains. Hence, the revised GBD method not only can identify the global solution of the relaxed MINLP problem reliably when it is feasible, but also can detect the infeasibility over the current subsimplex effectively. Consequently, that subsimplex is removed with the efficient progress of the branch-and-bound framework. A very simple, but typical example with a 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. A comparison of efficiency between this hybrid approach and other techniques, especially the complete branch-and-bound type, requires its implementation for large process design and synthesis MINLP problems in the chemical engineering field, which is still under development.

Acknowledgment

Y.Z. gratefully acknowledges financial support from the Japan Society for the Promotion of Science (JSPS) and JSPS Fellow Project 01040 and sincerely appreciates the constant help and encouragement from Prof. W. D. Seider (University of Pennsyvania) and Prof. C. A. Floudas (University of Princeton) for his research work in the field of process system engineering. Y.Z. is a JSPS Fellow. He is also gratefully supported by a Grant-in-Aid for Scientific Research from JSPS.

Appendix. Underestimators for Different Nonconvex Functions

I. Underestimator for Convex/Linear Function Structure. For convex/linear function structures, denoted by $f^C(\mathbf{x})$ or $f^L(\mathbf{x})$, respectively, obviously, their convex envelopes are themselves. Thus, they will preserve their original forms in the final underestimators for the objective function and the constraints.

II. Underestimator for Concave Function Structure. For concave function structures, denoted by $f^{-C}(\mathbf{x})$, and for $\mathbf{S}$ being a simplex generated by the vertices $\mathbf{V}^0$, $\mathbf{V}^1$, ..., $\mathbf{V}^n$, i.e., $\mathbf{S} = \{\mathbf{x} \in \mathscr{R}^n\colon\ \mathbf{x} = \sum_{i=1}^{n+1} \lambda_i \mathbf{V}^i,\ \lambda_i \ge 0,\ \sum_{i=1}^{n+1}$

$\lambda_i = 1\}$, the convex envelope of $f^{-C}(\mathbf{x})$ over $\mathbf{S}$ is the affine function $L^{-C}(\mathbf{x}) = \mathbf{b}^T\mathbf{x} + c$ whose coefficients are uniquely determined by the system of linear equations $f^{-C}(\mathbf{V}^i) = \mathbf{b}^T\mathbf{V}^i + c$ for $i = 0, ..., n$.

III. Underestimator for General Quadratic Functions. A general quadratic function can be represented as

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

Because $\mathbf{H}_f(\mathbf{x}) = \mathbf{Q}$ is a constant matrix, 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 nonuniform case, we obtain

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

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 $f^Q(\mathbf{V}^i) - \mathbf{V}^{iT}\Delta\mathbf{V}^i = \mathbf{b}^T\mathbf{V}^i + c$ for $i = 0, ..., n$, and the vertices $\mathbf{V}^0, \mathbf{V}^1, ..., \mathbf{V}^n$ of the simplex $\mathbf{S}$.

IV. Underestimator for Twice-Differentiable Nonconvex Functions. 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 nonuniform case, we obtain

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

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 $f^{NC}(\mathbf{V}^i) - \mathbf{V}^{iT}\Delta\mathbf{V}^i = \mathbf{b}^T\mathbf{V}^i + c$ for $i = 0, ..., n$, and the vertices $\mathbf{V}^0, \mathbf{V}^1, ..., \mathbf{V}^n$ of the simplex $\mathbf{S}$.

Literature Cited

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

(2) Adjiman, C. S.; Schweiger, C. A.; Floudas, C. A. Mixed-integer nonlinear optimization in process synthesis. In *Handbook of Combinatorial Optimization*; Du, D.-Z., Pardalos, P. M. Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1998.

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

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

(5) Duran, M. A.; Grossmann, I. E. An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Math. Program.* **1986**, *36*, 307.

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

(7) Floudas, C. A. *Nonlinear and Mixed-Integer Optimization, Fundamentals and Applications*; Oxford University Press: New York, 1995.

(8) Floudas, C. A. *Deterministic Global Optimization: Theory, Methods and Applications*; Kluwer Academic Publishers; Dordrecht, The Netherlands, 2000.

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

(10) Grossman, I. E., Ed. *Global Optimization in Engineering Design*; Kluwer Academic Publishers; Dordrecht, The Netherlands, 1996.

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

(12) Grossman, I. E. Review of nonlinear mixed-integer and disjunctive programming techniques for process system engineering. Chemical Engineering Department, Carnegie Mellon University, Pittsburgh, PA. Unpublished work.

(13) Horst, R.; Pardalos, P. M.; Thoai, N. V. *Introduction to Global Optimization*; Kluwer Academic Publishers; Dordrecht, The Netherlands, 1995.

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

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

(16) Nemhauser, G. L.; Wolsey, L. A. *Integer and Combinatorial Optimization*; John Wiley & Sons: New York, 1999.

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

(18) Smith, E. M. B.; Pantelides, C. C. Global optimization of general process models. In *Global Optimization in Engineering Design*; Grossman, I. E., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1996.

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

(20) Tuy, H. *Convex Analysis and Global Optimization*; Kluwer Academic Publishers; Dordrecht, The Netherlands, 1998.

(21) Vaidyanathan, R.; El-Halwagi, M. Global optimization of nonconvex MINLPs by interval analysis. In *Global Optimization in Engineering Design*; Grossman, I. E., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1996.

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

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

(24) Zhu, Y.; Xu, Z. 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.* **1999**, *38*, 3549.

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

(26) Zhu, Y.; Kuno, T. A global optimization method, QBB, for twice- differentiable nonconvex optimization problem. Technical Report ISE-TR-01-182; Institute of Information Sciences and Electronics, University of Tsukuba: Ibaraki, Japan, 2001 (manuscript also submitted for publication in *J. Global Optim.*).

Received for review January 28, 2002
Revised manuscript received November 25, 2002
Accepted December 4, 2002

IE0200813