# CONVEX UNDERESTIMATION OF $C^2$ CONTINUOUS FUNCTIONS BY PIECEWISE QUADRATIC PERTURBATION

Clifford A. Meyer and Christodoulos A. Floudas*
Department of Chemical Engineering, Princeton University, Princeton, NJ 08544, USA

*Abstract*

This paper presents an efficient branch and bound approach to address the global optimization of constrained optimization problems with twice differentiable functions. A lower bound on the global minimum is determined via a convex nonlinear programming problem in which all nonconvex functions are substituted by their convex underestimators. This work refines the classical $\alpha$BB eigenvalue perturbation method for the convex underestimation of twice differentiable functions. New convex underestimators are proposed based on a smooth, piecewise quadratic, perturbation function. The $\alpha$ parameters, coefficients of the quadratic terms in the perturbation function, are calculated using eigenvalue analysis techniques. Formulae defining the linear coefficients and the constants of the piecewise quadratic perturbations function are derived from continuity, smoothness and end point conditions. The piecewise quadratic form of the perturbation is far more flexible that the quadratic form employed in the classical $\alpha$BB methodology. This flexibility and improved bounds on the $\alpha$ values lead to a vast improvement over the classical aBB method.

*Keywords*

Convex underestimation, $\alpha$BB algorithm, Global optimization

## Introduction

Nonlinear programming and mixed integer nonlinear programming formulations have become ubiquitous in the design and planning of chemical processes, yet are often solved using software that cannot determine the global solution to these formulations. The $\alpha$BB algorithm is a deterministic global optimization algorithm that can be applied to a broad class of nonconvex NLP and MINLP problems [Maranas and Floudas, 1994, Adjiman et al., 1996, 1998a,b, Floudas, 2000]. This algorithm employs a convex relaxation strategy to determine rigorous lower bounds on the global minimum solution. In this algorithm the refinement and convergence of the lower bound to within a predefined $\epsilon$ of the global solution is affected using the branch and bound technique. The tightness of the convex underestimators of the nonconvex functions has a strong influence on the amount of computation needed for convergence of the $\alpha$BB algorithm. This paper extends and refines the convex underestimation approach used in the $\alpha$BB to underestimate general $C^2$ continuous functions.

The convex underestimator, $\phi : \mathbb{R}^n \in x \to \mathbb{R}$, of a general $C^2$ continuous function, $f : \mathbb{R}^n \to \mathbb{R}$, is defined in the $\alpha$BB algorithm as,

$$\phi(x) := f(x) - q(x).$$

where $q : \mathbb{R}^n \to \mathbb{R}$ is a concave quadratic perturbation function. This function has the form

$$q(x) := \sum_{i=1}^{n} \alpha_i (\overline{x}_i - x_i)(x_i - \underline{x}_i).$$

where $\underline{x_i}$ and $\overline{x_i}$ are, respectively, the lower and upper bounds on $x_i$. Notice that by making the $\alpha_i$ parameters sufficiently large the Hessian matrix of the underestimator, $\nabla^2(f(x) - q(x))$, can be forced to be *positive semidefinite*. When the $\alpha$ parameters are all nonnegative $q(x)$ is positive for all $x \in \mathbf{x}$ where $\mathbf{x} \subset \mathbb{R}^n$ denotes the hyperrectangle defined by upper and lower bounds on the elements of $x$. It follows from the nonnegativity of $q$ and the positive definiteness of the Hessian of $\phi$ that $\phi$ is a convex underestimator of $f$ over the domain $\mathbf{x}$.

Adjiman et al. [1998a] proposed the use of the *interval extension* $\mathbf{H^x}$ instead of $\nabla^2 f(x)$ itself to calculate the $\alpha$ parameters. The interval extension of the matrix $\nabla^2 f(x) \in$

* To whom all correspondence should be addressed. Tel:+1(609)2584595; Fax:+1(609)2580211; Email:floudas@titan.princeton.edu.

$\mathbb{R}^{n\times n}$ is a matrix of intervals of $\mathbb{R}$. Each element $\mathbf{H}^{\mathbf{x}}_{ij}$ of the matrix $\mathbf{H}^{\mathbf{x}}$ is defined in such a way that

$$\left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_x \in \mathbf{H}^{\mathbf{x}}_{ij} \text{ for all } x \in \mathbf{x}.$$

In practice an interval extension can be calculated using interval arithmetic [Neumaier, 1990]. Adjiman et al. [1998a] applied the work of Gerschgorin [1931], Kharitonov [1979] and Neumaier [1992] to compute $\alpha$ vectors that guarantee the convexity of the underestimator.

In this paper the form of the $\alpha$BB perturbation function and the way in which it is calculated are reexamined, a novel spline based method for convex underestimation is proposed and an efficient means of computing these tighter underestimators is elucidated.

**A spline-like $\alpha$ underestimator**

The size of the domain $\mathbf{x}$ effects the result of every step in the $\alpha$ calculation and strongly influences the tightness of the resulting convex underestimator. In particular, reducing $\mathbf{x}$ reduces the mismatch between the assumed quadratic functional form and the ideal form; it reduces the overestimation in the interval extension of the Hessian matrix; and the maximum separation distance has been shown to be a quadratic function of interval length [Maranas and Floudas, 1994]. It is therefore useful to construct a convex underestimator using a number of different $\alpha$ vectors, each applying to a *subregion* of the full domain $\mathbf{x}$.

Let $f(x) : \mathbb{R}^n \to \mathbb{R}$ be a $C^2$ continuous function. For each variable $x_i \in \mathbb{R}$, let the interval $[\underline{x}_i, \overline{x}_i]$ be partitioned into $N_i$ subintervals. The endpoints of these subintervals are denoted $x_i^0, x_i^1, \cdots, x_i^{N_i}$ where $\underline{x}_i = x_i^0 < x_i^1 < \cdots < x_i^k < \cdots < x_i^{N_i} = \overline{x}_i$. In this notation the $k^{\text{th}}$ interval is $[x_i^{k-1}, x_i^k]$. A smooth convex underestimator of $f(x)$ over $\mathbf{x}$ is defined by

$$\phi(x) := f(x) - q(x)$$

where

$$\begin{aligned} q(x) &:= \sum_{i=1}^{n} q_i^k(x_i) \quad \text{for } x_i \in [x_i^{k-1}, x_i^k], & (1)\\ q_i^k(x_i) &:= \alpha_i^k (x_i - x_i^{k-1})(x_i^k - x_i) + \beta_i^k x_i + \gamma_i^k. & (2) \end{aligned}$$

In each interval $[x_i^{k-1}, x_i^k]$, $\alpha_i^k \geq 0$ is chosen such that $\nabla^2\phi(x)$, the Hessian matrix of $\phi(x)$, is positive semidefinite for all members of the set $\{x \in \mathbf{x} : x_i \in [x_i^{k-1}, x_i^k]\}$. $q_i^k(x_i)$ is the quadratic function associated with variable $i$ in interval $k$. The function $q(x)$ is a piecewise quadratic function contructed from the functions $q_i^k(x_i)$.

The continuity and smoothness properties of $q(x)$ are produced in a spline-like manner. For $q(x)$ to be smooth the $q_i^k$ functions and their gradients must match at the endpoints $x_i^k$. In addition, we require that $q(x) = 0$ at the vertices of the hyperrectangle $\mathbf{x}$. To satisfy these requirements, the following conditions are imposed for all $i = 1, \ldots, n$:

$$\begin{aligned} q_i^1(x_i^0) &= 0 & (3)\\ q_i^k(x_i^k) &= q_i^{k+1}(x_i^k) \text{ for all } k = 1, \ldots, N_i - 1 & (4)\\ q_i^{N_i}(x_i^{N_i}) &= 0 & (5)\\ \left.\frac{dq_i^k}{dx_i}\right|_{x_i^k} &= \left.\frac{dq_i^{k+1}}{dx_i}\right|_{x_i^k} \text{ for all } k = 1, \ldots, N_i - 1. & (6) \end{aligned}$$

These conditions expand into a set of linear equations with the solution,

$$\beta_i^1 = \left(\sum_{k=1}^{N_i-1} s_i^k (x_i^k - x_i^{N_i})\right) / (x_i^{N_i} - x_i^0) \quad (7)$$

$$\beta_i^k = \beta_i^1 + \sum_{j=1}^{k-1} s_i^j \text{ for all } k = 2, \ldots, N_i \quad (8)$$

$$\gamma_i^k = -\beta_i^1 x_i^0 - \sum_{j=1}^{k-1} s_i^j x_i^j \text{ for all } k = 1, \ldots, N_i. \quad (9)$$

where $s_i^k = -\alpha_i^k (x_i^k - x_i^{k-1}) - \alpha_i^{k+1}(x_i^{k+1} - x_i^k)$.

**Geometrical interpretation**

The construction of the convex underestimator for a nonconvex function

$$f(x) = -2x + 10x^2 - 3x^3 - 5x^4$$

over the domain $x \in [0, 1]$ is illustrated in Figures 1(a) and 1(b). Figure 1(a) shows the nonconvex function $f(x)$ along with underestimators of $f(x)$. A convex underestimator defined using the classical $\alpha$BB approach requires the $\alpha$ value to be large enough to cancel the negative curvature at *all* points in the domain. Noting that the second derivative, $f''(x) = 20 - 18x - 60x^2$, is a monotonically *decreasing* function of $x$, the most negative curvature occurs at $x = 1$, hence the $\alpha$ parameter is defined by $f''(1)$ using the formula,

$$\begin{aligned} \alpha &= -\tfrac{1}{2} f''(1)\\ &= 29. \end{aligned}$$

The classical $\alpha$BB underestimator,

$$\phi(x) = f(x) - 29(1 - x)(x - 0),$$

is shown in Figure 1(b). This underestimator can be improved by partitioning the domain into three subintervals of equal length, $[x^0, x^1]$, $[x^1, x^2]$, $[x^2, x^3]$, where $x^0 = 0$, $x^1 = \frac{1}{3}$, $x^2 = \frac{2}{3}$ and $x^3 = 1$. As $f''(x)$ is a monotonically decreasing function, the $\alpha$ values in each interval are derived from the upper bounds on the respective intervals as follows,

$$\begin{aligned} \alpha^1 &= \max\{0, -\tfrac{1}{2} f''(x^1)\} = 0\\ \alpha^2 &= \max\{0, -\tfrac{1}{2} f''(x^2)\} = 9\tfrac{1}{3}\\ \alpha^3 &= \max\{0, -\tfrac{1}{2} f''(x^3)\} = 29. \end{aligned}$$

The classical $\alpha$BB perturbation functions and underestimators over each of the smaller intervals are depicted in Figures 1(a) and 1(b) ($--$). A convex underestimator over the whole interval is constructed by adding a linear function $\beta^i x + \gamma^i$ to the $\alpha$BB perturbations over each of the subintervals $i = 1, \ldots, 3$. The parameters $\beta^i$ and $\gamma^i$, defining these linear functions are chosen so that the overall perturbation function is smooth and is zero at the end points. These values are calculated using the Equations 7 to 9. The piecewise quadratic perturbation function, shown in Figure 1(a) ($-$) is defined as follows:

$$\begin{array}{ll} q(x) = q^1(x) & \text{for } x \in [0, \frac{1}{3}] \\ q(x) = q^2(x) & \text{for } x \in [\frac{1}{3}, \frac{2}{3}] \\ q(x) = q^3(x) & \text{for } x \in [\frac{2}{3}, 1] \end{array}$$

$$\begin{aligned} q^1(x) = &\; 6.6667x \\ q^2(x) = &\; 9.3333(0.6667 - x)(x - 0.3333) \\ &+3.2221x + 1.0370 \\ q^3(x) = &\; 29(1.0 - x)(x - 0.667) - 9.5552x + 9.5551 \end{aligned}$$

In Figure 1(a) the endpoints of the quadratic pieces are labelled **A**, **B**, **C** and **D**. At the endpoints **A** and **D**, the conditions $q^1(x^0) = 0$ and $q^3(x^3) = 0$, respectively, are enforced. Two conditions are enforced at each of the interior points **B** and **C**, to enforce the smoothness of the piecewise quadratic function. At point **B**, $q^1(x^1) = q^2(x^1)$ and $\left.\frac{dq^1}{dx}\right|_{x^1} = \left.\frac{dq^2}{dx}\right|_{x^1}$ apply, and at point **C**, $q^2(x^2) = q^3(x^2)$ and $\left.\frac{dq^2}{dx}\right|_{x^2} = \left.\frac{dq^3}{dx}\right|_{x^2}$ apply. The convex underestimator, which is the difference, $f(x) - q(x)$, is shown in Figure 1(b) ($-$).

**Nonconcave perturbation**

Consider a function $f(x)$, in which the function is convex in one subdomain and concave in another. In the $\alpha$ spline approach $\phi(x)$ can be convex even if the $\alpha$ values are negative in the regions in which $f(x)$ is strictly convex. The *underestimation* property is guaranteed by the *concavity* of $q(x)$. The concavity of $q(x)$ is, in turn, a result of the *non-negativity* of the $\alpha$ values. In this section we discuss how the underestimation property of $\phi(x)$ can be maintained when some $\alpha$ values are indeed *negative*.

The underestimation property, $\phi(x) \leq f(x)$ for all $x \in \mathbf{x}$, is ensured by the following condition:

$$\min_{x \in \mathbf{x}} q(x) \geq 0.$$

Instead of solving minimization problems, the key idea is to adjust the $\alpha$'s to prevent the creation of local minima at any nonvertex point in $\mathbf{x}$ by prohibiting the occurrence of *stationary points* on convex regions of the perturbation function. A tight convex underestimator is derived by starting with $q(x)$, with non-negative $\alpha$ values as defined in Section , and making the zero $\alpha$'s negative one at a time, while maintaining the convexity of $\phi(x)$ and avoiding the generation of stationary points on the convex portions of $q$.

**Computational Performance**

The Shubert function $f(x) := \sum_{i=1}^{5} i \cos((1+i)x + i)$ was used to construct the objective function in the following minimization problem:

$$\min_{x \in [-10,10]^3} f(x_1)f(x_2) + f(x_2)f(x_3).$$

The lower bounding problem was formulated in two ways which will be referred to as "**A**" and "**B**".

In formulation **A** the function

$$g(x_1, x_2) \quad = f(x_1)f(x_2)$$

is treated as a general $C^2$ continuous function and the lower bounding problem is formulated applying $\alpha$ underestimators to bivariate functions. The following lower bounding formulation results:

$$\min_{x \in [-10,10]^3} \underline{g}(x_1, x_2) + \underline{g}(x_2, x_3). \qquad (\mathbf{A})$$

The lower bounding formulation **B** is as follows:

$$\begin{array}{lrl} & \min_{x \in [-10,10]^3, w \in \mathbb{R}^5} & w_4 + w_5 \\ \text{subject to} & & \\ & \underline{f}(x_1) & \leq w_1 \\ & \underline{f}(x_2) & \leq w_2 \\ & \underline{f}(x_3) & \leq w_3 \\ & \underline{w_2}w_1 + \underline{w_1}w_2 + \underline{w_1}\,\underline{w_2} & \leq w_4 \\ & \overline{w_2}w_1 + \overline{w_1}w_2 + \overline{w_2}\,\overline{w_1} & \leq w_4 \\ & \underline{w_2}w_3 + \underline{w_3}w_2 + \underline{w_2}\,\underline{w_3} & \leq w_5 \\ & \overline{w_2}w_3 + \overline{w_3}w_2 + \overline{w_2}\,\overline{w_3} & \leq w_5 \end{array} \qquad (\mathbf{B})$$

In this formulation the objective function is underestimated through the introduction of auxiliary variables $w_1, \ldots, w_5$, the use of convex envelopes for the underestimation of bilinear terms [McCormick, 1976], and the use of $\alpha$-spline underestimators for the underestimation of the *univariate* Shubert function.

Computational results for both formulations are tabulated in Table 1.

The $\alpha$-spline underestimator performed far better than the classical $\alpha$BB for both formulations. In formulation **A** the classical $\alpha$BB took $341\%$ of the iterations required by the $\alpha$-spline method and in formulation **B** this percentage increased to $51,252\%$. These results can be attributed primarily to the quality of the $\alpha$-spline underestimator being better for *univariate* functions that for *bivariate* ones.

**Conclusions**

The convex underestimator for $C^2$ continuous functions proposed in this paper is a refinement of the classical $\alpha$BB underestimator. The new underestimator is based on a smooth, piecewise quadratic, perturbation function with varying curvature. The perturbation may be nonconcave, yet is guaranteed to form a convex underestimator when subtracted from the nonconvex function. In some cases the new underestimator closely approximates the convex envelope of the nonconvex function. The main computational effort in the calculation of the parameters of the $\alpha$-spline underestimator lies in the evaluation of the interval Hessian matrix in a potentially large number of subregions of the function domain. This effort can be offset by storing the interval Hessian data that are generated at the nodes in the branch and bound tree and reusing this information in other nodes of the tree. Computational results show that the proposed underestimator is indeed more effective than the classical approach.

## References

C. S. Adjiman, I. P. Androulakis, and C. A. Floudas. A global optimization method, $\alpha$BB, for general twice differentiable NLPs-II. implementation and computional results. *Computers & Chemical Engineering*, 22:1159–1179, 1998b.

C. S. Adjiman, I. P. Androulakis, C. D. Maranas, and C. A. Floudas. A global optimization method, $\alpha$BB, for process design. *Computers & Chemical Engineering Supplement*, 20: S419–S424, 1996.

C. S. Adjiman, S. Dallwig, C. A. Floudas, and A. Neumaier. A global optimization method, $\alpha$BB, for general twice differentiable NLPs-I. theoretical advances. *Computers & Chemical Engineering*, 22:1137–1158, 1998a.

C. A. Floudas. *Deterministic Global Optimization: Theory, Algorithms and Applications*. Kluwer Academic Publishers, 2000.

S. Gerschgorin. Über die abgrenzung der eigenwerte einer matrix. *Izv. Akad. Nauk SSSR, Ser. fiz.-mat.*, 6:749–754, 1931.

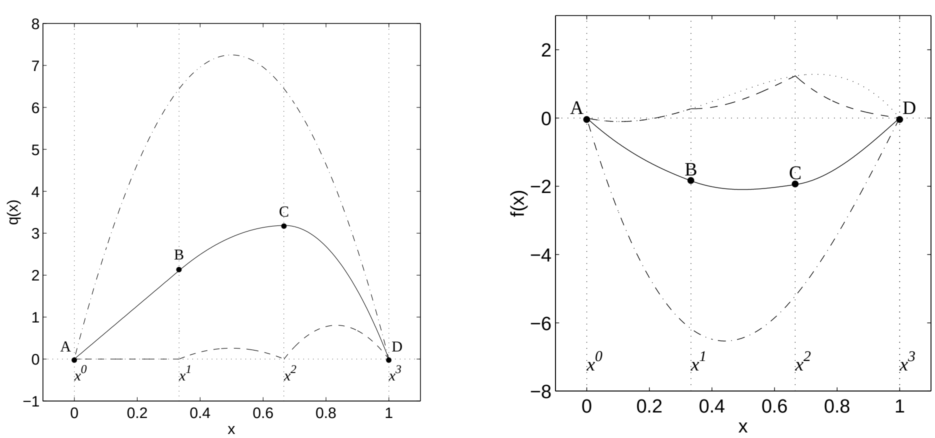
V.L. Kharitonov. Asymptotic stability of an equilibrium position of a family of systems of linear differential equations. *Differential Equations*, 78:1483–1485, 1979.

C. D. Maranas and C.A. Floudas. Global minimum potential energy conformations of small molecules. *Journal of Global Optimization*, 4:135–170, 1994.

G. P. McCormick. Computability of global solutions to factorable nonconvex programs: Part I – convex underestimating problems. *Mathematical Programming*, 10:147–175, 1976.

A. Neumaier. *Interval Methods for Systems of Equations*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1990.

A. Neumaier. An optimality criterion for global quadratic optimization. *Journal of Global Optimization*, 2:201–208, 1992.


(a) Geometric interpretation of conditions on perturbation function A: $q^1(x^0) = 0$, B: $q^1(x^1) = q^2(x^1)$ and $\left.\frac{dq^1}{dx}\right|_{x^1} = \left.\frac{dq^2}{dx}\right|_{x^1}$ C: $q^2(x^2) = q^3(x^2)$ and $\left.\frac{dq^2}{dx}\right|_{x^2} = \left.\frac{dq^3}{dx}\right|_{x^2}$ D: $q^3(x^3) = 0$. $\alpha$BB perturbation: $- \cdot -$, $\alpha$BB perturbations over partial domains: $--$, piecewise quadratic perturbation: $-$.

(b) Geometric interpretation of $\alpha$BB and spline $\alpha$BB underestimators $f(x)$: $\cdots$ and convex underestimators: $\alpha$BB: $- \cdot -$, $\alpha$BB over partial domains: $--$, piecewise underestimator: $-$.

Figure 1: Geometric interpretation of perturbation and underestimation functions

| | method | iters | CPU |
|---|---|---|---|
| A | classical $\alpha$BB | 250952 | 34169 |
| | spline $\alpha$BB | 73427 | 6303 |
| B | classical $\alpha$BB | 51150 | 62021 |
| | spline $\alpha$BB | 998 | 1672 |

Table 1: Computational comparison with classical $\alpha$BB