Slovak University of Technology in Bratislava
Institute of Information Engineering, Automation, and Mathematics

PROCEEDINGS

of the 18$^{\mathrm{th}}$ International Conference on Process Control

Hotel Titris, Tatranská Lomnica, Slovakia, June 14 – 17, 2011

ISBN 978-80-227-3517-9

http://www.kirp.chtf.stuba.sk/pc11

Editors: M. Fikar and M. Kvasnica

Tighter Convex Relaxations for Global Optimization Using αBB Based Approach

R. Paulen * **and M. Fikar** *

* *Institute of Information Engineering, Automation and Mathematics*
Slovak University of Technology in Bratislava
e-mail: {radoslav.paulen, miroslav.fikar}@stuba.sk
url: http://www.kirp.chtf.stuba.sk

Abstract: This paper is devoted to investigation of certain issues that appear in solving of deterministic global optimization problems (GOPs). Basically, we focus ourselves on introducing a procedure which may serve to establish tighter convex relaxations for a certain class of non-convex optimization problems. Tightness of these convex relaxations plays important role in speeding of the convergence of branch-and-bound algorithm which is used as a basic framework of solving GOPs in this study. Two case studies are solved where it is shown how significant improvement can be achieved by considering proposed framework.

Keywords: Global Optimization, Convex Relaxation, αBB Method

1. INTRODUCTION

Global optimization (GO) represents a set of methods which aim to find (global) solution of non-convex optimization problems which may possess multiple suboptimal (local) solutions and are typically encountered in many engineering fields, including chemical engineering, process design, computational biology, and many others. Over past two decades, there was a significant effort dedicated to deterministic GO by many scientists. Efficient algorithms and methods were developed, many new interesting applications were introduced and lot of existing non-convex optimization problems were solved to global optimality. Essence of these can be found in works Floudas and Visweswaran (1990); Adjiman and Floudas (1996); Singer and Barton (2001); Papamichail and Adjiman (2002); Chachuat et al. (2003); Čižniar et al. (2009).

There is a big range of problems addressed by global optimization. Basically non-convex non-linear programs (NLPs) are considered. However, it is popular to convert mixed-integer linear programs (MILPs) and mixed-integer non-linear (MINLPs) to non-convex NLPs. Also the problems of dynamic optimization are usually discretized into NLPs, e.g. by using the method of orthogonal collocation (Biegler, 1984). In all these problems GO plays significant role. The problems are addressed using either stochastic approaches, such as simulated annealing, particle swarm and genetic algorithms, or deterministic ones, such as branch-and-bound or interval analysis methods.

Branch-and-bound (BB) methods are the most popular deterministic GO frameworks. These methods successively partition solution space on which optimization problem is defined into smaller regions. In each region, the upper and lower bounds to the objective function value are generated, by solving the original (non-convex) problem together with its convex relaxation. According to these bounds it is decided whether region is going to be explored further or whether it should be fathomed out of BB tree. Global solution is then obtained once current best (lowest) upper bound (UB) value is close to current best (highest) lower bound (LB) value within specified tolerance ε.

Problems of GO are typically defined over a quite large region of decision variables. However, each deterministic GO algorithm investigates the whole solution space in some manner. This is a critical issue and it is a reason why very tight convex relaxations of non-convex problems are needed. It is the aim of this study to present a technique which involves simple algebraic manipulations but results in a considerable improvement in terms of number of GO algorithm iterations and algorithm run time.

The paper is organized as follows. Section 1 gives the mathematical formulation of the problem, it shows how the solution can be found and points out to some issues which are motivating our research. Section 2 gives procedure of how certain issues revealed in Section 1 may be avoided. Finally in Section 3 selected case studies are solved to prove efficacy of the proposed approach.

2. GLOBAL OPTIMIZATION PROCEDURE

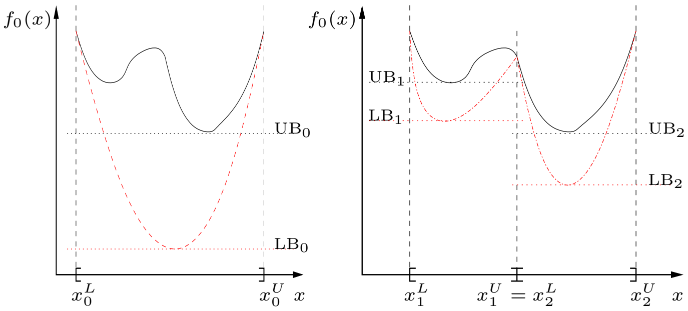
In this section, general procedure is described for solving of non-convex optimization problems to global optimality.

2.1 Problem Formulation

We address an optimization problem in following form

$$\min_x \; f_0(x) \tag{1a}$$
$$\text{s.t.}\; f_i(x) \le 0, \qquad i = 1, \ldots, n_i \tag{1b}$$
$$f_j(x) = 0, \qquad j = 1, \ldots, n_e \tag{1c}$$
$$x \in [x^L, x^U] \tag{1d}$$

where $x \in \mathbb{R}^n$ is a vector of decision variables which values are initially bounded by box constraints (1d). According



(a) node 0 (b) nodes 1&2

Fig. 1. Illustration of branch-and-bound procedure.

to Boyd and Vandenberghe (2004), this is a non-convex optimization problem with n_i inequality and n_e equality constraints, if any of functions in (1a) and (1b) is non-convex, or any of functions in (1c) is not affine. We assume that functions f_k $(k = 0, \dots, n_i + n_e)$ are twice continuously differentiable ($f_k \in \mathbb{C}^2$) and real-valued ($f_k : \mathbb{R}^n \to \mathbb{R}$). Solution to problem (1) provides an upper bound for BB algorithm. Lower bound is found by solving a convex relaxation of (1).

Global minimum of the optimization problem is found employing branch-and-bound framework (Horst and Tuy, 1990). At each branching node, original problem is solved together with its convex relaxation. This is shown in Fig. 1 where first two stages of illustrative BB procedure are shown. In node 0, convex relaxation (red line) of a non-convex problem (black line) is found on a given interval of decision variable x. This interval is then branched creating the first and the second node. Formation of convex relaxations of original problem on this branches follows. It is clear that lower bound in the first node (LB_1) is higher than upper bound in the second node (UB_2) and so the first node is not considered further since global solution does not lie there obviously. The ε-global optimum is found once the UB_i and LB_i meet within specified tolerance ε.

2.2 Convex Relaxation of the Problem

According to Adjiman et al. (1998b), formulation of convex relaxation of original (non-convex) problem requires decomposition of each function f_k $(k = 0, \dots, n_i + n_e)$ in (1) to a sum of linear, convex, special non-convex (univariate concave $\mathrm{UT}(x)$, bilinear $\mathrm{BT}(x)$, ...), and arbitrary non-convex ($\mathrm{NT}(x)$) terms. These terms are then convexly relaxed separately.

Linear terms do not require any convex relaxation since they are already convex. The same applies for convex terms unless they appear in equality constraint functions. In fact, any non-linear equality constraint has to be rewritten such that

$$f_j(x) = 0 \Leftrightarrow \begin{cases} f_j(x) \leq 0 \\ -f_j(x) \leq 0 \end{cases} \quad j = 0, \dots, n_e \tag{2}$$

For this inequality form, the original decomposition into generic terms should be reconsidered. Hence that if f_j is convex then $-f_j$ is concave and vice versa.

Convex relaxation of univariate concave terms $\mathrm{UT}(x)$ is done by linearization around the lower bound of the variable range. Then, every occurrence of such term is replaced by following expression

$$\mathrm{UT}(x^L) + \frac{\mathrm{UT}(x^U) - \mathrm{UT}(x^L)}{x^U - x^L}(x - x^L) \tag{3}$$

Addition of a relaxation function to the arbitrary non-convex term $\mathrm{NT}(x)$ establishes convex relaxation function for such term in a form

$$\mathrm{NT}(x) + \sum_{i=1}^{n} \alpha_i (x_i^L - x_i)(x_i^U - x_i) \tag{4}$$

where values of α_i's are non-negative scalars found such that

$$\alpha_i \geq \max \left\{ 0, -\frac{1}{2} \min_{x_i} \lambda \left(\nabla^2_{x_i} \mathrm{NT}(x) \right) \right\} \tag{5}$$

where λ is eigenvalue of Hessian matrix of non-convex term. Another option is just to concentrate on finding the overall valid α which will guarantee convexity of function

$$\mathrm{NT}(x) + \alpha (x^L - x)^T (x^U - x) \tag{6}$$

Then the value of α is computed such that

$$\alpha \geq \max \left\{ 0, -\frac{1}{2} \min_{x} \lambda \left(\nabla^2_{x} \mathrm{NT}(x) \right) \right\} \tag{7}$$

Problem of minimization of eigenvalue appearing in Eq. (5) and (7) requires solution of non-convex problem in most cases. To avoid this, an interval arithmetic methods can be exploited, e.g. Gerschgorin's theorem for interval matrices (Floudas, 2000). Then, problem of calculation of α_i's values boils down to finding of a minimal eigenvalue of interval family Hessian matrix $[\nabla^2 \mathrm{NT}(x)]$. Interval approaches which can be adopted for this purpose are discussed in Adjiman et al. (1998b) in detail.

As showed in Kearfott (1996), if interval arithmetic operations (multiplication, division, addition, etc.) are composed the interval arithmetic calculations overestimate the range of resulting interval. For example (taken from Kearfott (1996)), if interval function $f(x) = x^2 - x$ over the interval $x = [0, 1]$ is considered, resulting interval calculation is done such that

$$[0, 1]^2 - [0, 1] = [0, 1] - [0, 1] = [-1, 1] \tag{8}$$

This effect is illustrated in Fig. 2 which shows a plot of considered interval function together with its real-valued function equivalent. It can be observed that real-valued function takes values from -0.25 to 0, while its interval extension overestimates this values as it is shown in Eq. (8).

Illustrative Example. Let us consider a simple example to show the effect of the range overestimation of interval arithmetic calculations on convex relaxation of non-convex functions. Here, it is desired to find a convex relaxation of a function $f(x, y) = \cos(x) \sin(y)$ on the interval $x \times y = [-1, 2] \times [-1, 1]$. This convex relaxation is found in form (6). Using of (4) gives the same result. Value of α is found by computing eigenvalues of Hessian matrix

$$\nabla^2 f(x, y) = \begin{pmatrix} -\cos(x)\sin(y) & -\sin(x)\cos(y) \\ -\sin(x)\cos(y) & -\cos(x)\sin(y) \end{pmatrix} \tag{9}$$

Using Gerschgorin's theorem for interval matrices (currently implemented in INTLAB toolbox by Rump (1999)), the value of α was computed according to Eq. (7) to be greater than or equal to 0.92. Exact eigenvalue calculation

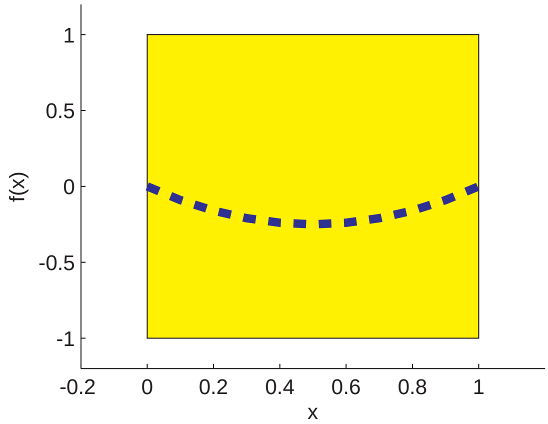
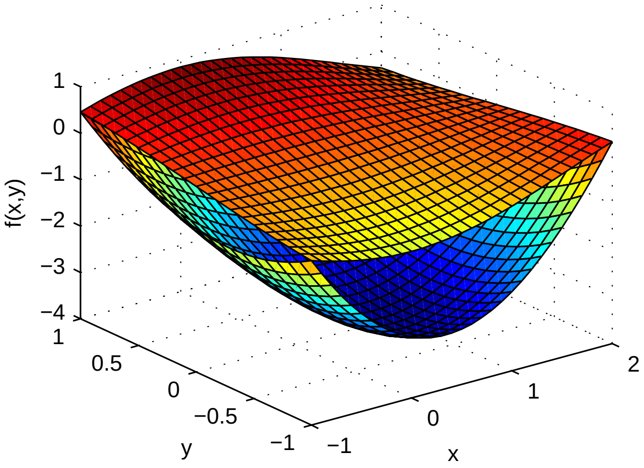
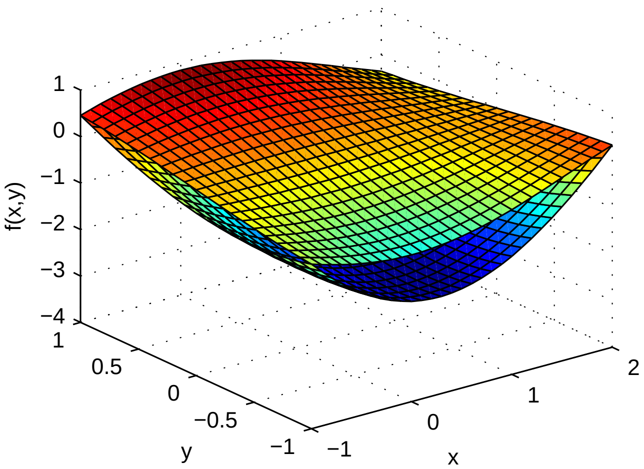


Fig. 2. Real-valued function $f(x)$ (dashed blue line) and its interval function equivalent (yellow box).



(a) $\alpha = 0.92$; found by interval Hessian eigenvalues calculation



(b) $\alpha = 0.5$; found by exact Hessian eigenvalues calculation

Fig. 3. Original function and its convex underestimators.

of Hessian (9) found minimal value of α that guarantee convexity of underestimator (relaxed function) to be 0.5. Fig. 3 compares two convex underestimators obtained by evaluating Eq. (6) with previously computed α values. It is clearly seen that underestimator generated using value of α obtained by an exact Hessian calculation produces convex underestimator tighter almost twice compared to the other one.

Throughout the BB algorithm run, a possibly large number of nodes may appear in BB tree. This happens if loose lower bounds are provided and it results in keeping many nodes where only suboptimal solutions lie. Then, BB algorithm spends a fair amount of time exploring these nodes which is an unwanted feature. It is then straightforward that tighter convex relaxation will result in less iterations needed for a convergence of BB optimization algorithm and less running time of the algorithm as well. In next, we will show how a simple algebraic manipulation can lead to a significant benefit in terms of more efficient algorithm.

3. PROPOSED REFORMULATION PROCEDURE

In previous section we showed how composition of arithmetic operations in interval calculus may result in large overestimation of resulting interval (function). This may be a certain issue if tight convex relaxation functions are wanted to be established for arbitrary non-convex terms. Addition and subtraction operations play just marginal role here since the non-convex term where addition (subtraction) occurs can be rearranged to more non-convex (some possibly convex) terms with no addition (subtraction) operation occurring. Multiplication operations can be decomposed using a simple algebraic transform (Williams, 1993). Suppose that non-convex term in any of functions in (1) is such that $\mathrm{NT}(x) = f_1(x)f_2(x)$. This can be rewritten as

$$f_1(x)f_2(x) = \frac{1}{4}(f_1(x) + f_2(x))^2 - \frac{1}{4}(f_1(x) - f_2(x))^2 \quad (10)$$

Equation (10) can be simplified by considering two new (decision) variables with two equality constraints.

$$f_1(x)f_2(x) = \frac{1}{4}u_1^2 - \frac{1}{4}u_2^2 \quad (11a)$$
$$f_1(x) + f_2(x) = u_1 \quad (11b)$$
$$f_1(x) - f_2(x) = u_2 \quad (11c)$$

Convex relaxation of this rewritten function is now found by the convex relaxation of concave term $-u_2^2$ appearing in Eq. (11a). This is done by a replacement of concave term using (3). Convex relaxation of constraint functions follow in the same manner as described in previous section by rewriting (11b) and (11c) into inequality form and then founding convex relaxations of terms $f_1(x)$, $-f_1(x)$, $f_2(x)$ and $-f_2(x)$. These convex relaxations then produce tighter convex relaxation of problem (1). However, it is needed to provide bounds (box constraints) on new added optimized variables. There are two possibilities. One, is to use interval arithmetic calculations such that u_1^L=min$[f_1(x) + f_2(x)]$ and so on. The second alternative is to consider an optimization problem which minimizes/maximizes u_i having the same constraints as convex relaxation of (1). This approach is similar to variable bound updates approach presented in Adjiman et al. (1998a).

Illustrative Example (Continued) We continue here with illustrative example considered previously. Non-convex term $\cos(x)\sin(y)$ is rewritten to following final form

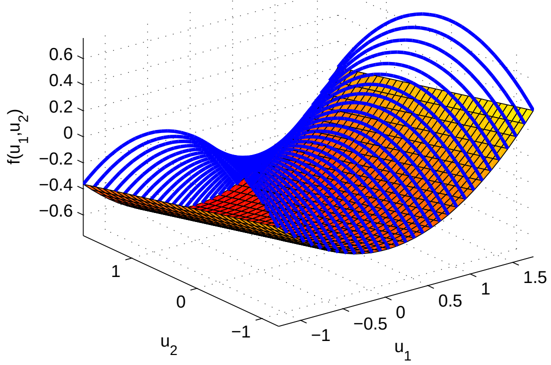


Fig. 4. Function $f(u_1, u_2)$ (depicted in blue lines) and its convex underestimator (red-orange surface).

$$\cos(x)\sin(y) = \frac{1}{4}u_1^2 - \frac{1}{4}u_2^2 \quad (12a)$$
$$\cos(x) + \sin(y) - u_1 \leq 0 \quad (12b)$$
$$-\cos(x) - \sin(y) + u_1 \leq 0 \quad (12c)$$
$$\cos(x) - \sin(y) - u_2 \leq 0 \quad (12d)$$
$$-\cos(x) + \sin(y) + u_2 \leq 0 \quad (12e)$$

Convex relaxation is done by underestimation of terms $-u_2^2$, $\cos(x)$, $-\cos(x)$, $\sin(y)$ and $-\sin(y)$. Box constraints used to bound new added decision variables are found using interval arithmetic calculations. Resulting convex relaxation of the function $f(u_1, u_2)$ is shown in Fig. 4. By transforming this function into original coordinates (by inverting the reformulation) it can be seen that convex relaxation of $\cos(x)\sin(y)$ term using proposed transform is clearly tighter than any of relaxations illustrated in Fig. 3.

4. CASE STUDIES

The global optimization algorithm taken from Papamichail and Adjiman (2004) was implemented in MATLAB 7.11. Solution of NLP problems was found using MATLAB NLP solver *fmincon*. It is an implementation of a general NLP solver, provided by the Optimization Toolbox, uses either a subspace trust region method, based on the interior-reflective Newton method, or a sequential quadratic programming method. The interval calculations needed were performed using INTLAB toolbox by Rump (1999). This toolbox finds the eigenvalues of interval family matrices using Gerschgorin's theorem for interval matrices. All case studies were solved on a workstation with 4.0 GHz Intel® Core™ 2 Duo Processor E8400 with 4GB RAM.

4.1 One dimensional non-convex problem

The first case study considers the problem of minimizing univariate non-convex function over a box constraint of decision variables. This problem was introduced in Pintér (2002) as test problem for GO algorithms. Its objective function is depicted in Fig. 5. Problem takes the form

$$\min_x \; 0.05(x - x_1)^2 + \sin^2(x - x_1) + + \sin^2\left((x - x_1)^2 + (x - x_1)\right) \quad (13a)$$
$$\text{s.t.} \; -10 \leq x \leq 10 \quad (13b)$$

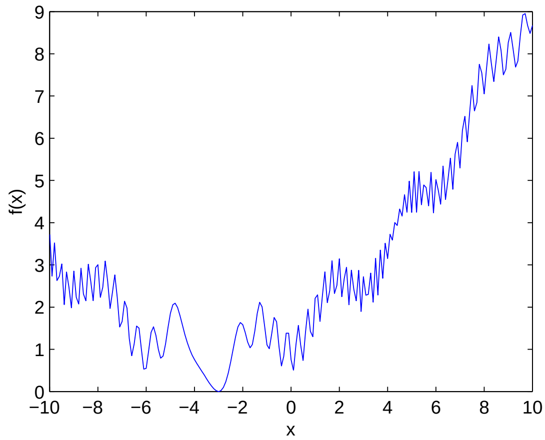


Fig. 5. Plot of the objective function in first case study.

Table 1. Results of the αBB algorithm run with different ranges of box constraints for reformulated problem.

N	No. of iterations	CPU time [s]
1	4	1
10	6	1
100	10	2

Table 2. Results of the αBB algorithm run with different ranges of box constraints for non-reformulated problem.

N	No. of iterations	CPU time [s]
1	41	5
10	497	60
100	8110	1063

where x_1 is a value of minimizer which can be chosen arbitrarily. We choose a minimizer value equal to -3. The first term present in objective function is convex and needs no convex relaxation. The second and the third term are non-convex due to periodicity of a sinus function. Using the proposed procedure, we avoided the squaring of sinus function in both of these terms by introduction of four new decision variables and corresponding eight inequality constraints.

Domain of the problem as it is introduced is arbitrarily set to $[-10, 10]$. In our computations we allow problem domain to be enlarged by a multiplication of box constraint by factor N. If N is equal to 1, originally proposed domain $[-10, 10]$ is considered. When N is set to 10, domain of the problem becomes $[-100, 100]$. This is done to investigate how proposed procedure performs with expanded size of box constraints. Resulting problems are solved to relative global optimality $\varepsilon = 1 \times 10^{-3}$. Results for different values of N are shown in Tab. 1. Comparison with performance of non-reformulated problem αBB algorithm is shown in Tab. 2.

It can be observed that αBB algorithm which exploits reformulation introduced in Section 3 performs significantly better. This feature is most evident if the largest box constraint ($N = 100$) is considered.

Table 3. Results of the αBB algorithm run with different ranges of box constraints for reformulated problem.

N	No. of iterations	CPU time [s]
1	12	16
10	78	90
100	835	929

Table 4. Results of the αBB algorithm run with different ranges of box constraints for non-reformulated problem.

N	No. of iterations	CPU time [s]
1	20	9
10	305	111
100	3083	1243

4.2 Two-dimensional non-convex problem

In this case study, non-convex term $\cos(x)\sin(y)$ appears which was used as an illustrative example for the whole proposed procedure. This optimization problem appears in Adjiman et al. (1996) where it was used as tutorial example to show how the α-based convexification procedure works. The problem is as follows

$$\min_{x,y} \ \cos(x)\sin(y) - \frac{x}{y^2+1} \tag{14a}$$

$$\text{s.t.} \ -1 \le x \le 2 \tag{14b}$$

$$-1 \le y \le 1 \tag{14c}$$

Second non-convex term present in objective function is rewritten in a similar manner to avoid a multiplication between the terms x and $1/(y^2+1)$. By a reformulation procedure, four new decision variables and eight inequality constraints are introduced into a problem. Bounds on new decision variables are found using interval arithmetic calculations. To find out how reformulation procedure performs, we consider not only the original range of decision variables but we multiply the box constraints with factor N equal to 10 and 100. These problems are again solved to relative global optimality $\varepsilon = 1 \times 10^{-3}$. Results are summarized in Tabs. 3 and 4 for the proposed and original approaches, respectively.

This case study again shows that the GO of reformulated problem performs better. However, the overall improvement is not as significant as it is for the first case study. The reduction in number of iterations is satisfying. However, there is a room for improvement if tighter bounds on new added optimized variables are provided. These bounds can be obtained by considering variable bound updates approach (Adjiman et al., 1998a). It can be observed that computational time needed for a single iteration of reformulated problem is almost double than CPU time needed for a single iteration of non-reformulated problem. This can be attributed to a greater size of the reformulated problem.

5. CONCLUSIONS

In this paper, we focused on a problem of finding of a global solution to non-convex non-linear problems. We considered utilization of αBB procedure to solve GO problems in deterministic fashion. Aim of this study was to introduce a simple algebraic reformulation of the non-convex problem to enhance the performance of αBB procedure. Chosen case studies showed that proposed reformulation technique resulted in significant improvement of αBB algorithm. Particularly this was significant if problems were defined over large region of decision variables.

Some issues of this approach were discussed, where one of these is linked to increasing of order of original problem by introducing new optimized variables into it. There is also a possible problem of finding of tight bounds for optimization variables added into reformulated problem. These are the main problems which will be critically addressed in a future work on this promising concept.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the contribution of the Scientific Grant Agency of the Slovak Republic under the grants 1/0071/09, and the Slovak Research and Development Agency under the project APVV-0029-07. The paper is also supported by a grant (No. NIL-I-007-d) from Iceland, Liechtenstein and Norway through the EEA Financial Mechanism and the Norwegian Financial Mechanism. This project is also co-financed from the state budget of the Slovak Republic.

REFERENCES

Adjiman, C.S., Androulakis, I.P., and Floudas, C.A. (1998a). A global optimization method, αBB, for general twice-differentiable constrained NLPs - II. Implementation and computational results. *Computers chem. Engng.*, 22(9), 1159–1179.

Adjiman, C.S., Androulakis, I.P., Maranas, C.D., and Floudas, C.A. (1996). A global optimization method, αBB for process design. *Computers chem. Engng.*, 20, 419–424.

Adjiman, C.S., Dallwig, S., Floudas, C.A., and Neumaier, A. (1998b). A global optimization method, αBB, for general twice-differentiable constrained NLPs - I. Theoretical advances. *Computers chem. Engng.*, 22(9), 1137–1158.

Adjiman, C.S. and Floudas, C.A. (1996). Rigorous convex underestimators for general twice-differentiable problems. *J. Glob. Optimization*, 9, 23–40.

Biegler, L.T. (1984). Solution of dynamic optimization problems by successive quadratic programming and orthogonal collocation. *Computers chem. Engng.*, 8(3/4), 243–248.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization.* Cambridge University Press.

Chachuat, B., Latifi, M.A., and Roche, N. (2003). A new approach in deterministic global optimisation of problems with ordinary differential equations. In C.A. Floudas and P.M. Pardalos (eds.), *Frontiers in Global Optimization*, 83–108. Kluwer Academic Publishers.

Floudas, C.A. (2000). *Deterministic Global Optimization: Theory, Methods and Applications.* Nonconvex Optimization and Its Applications. Kluwer Academic Publishers.

Floudas, C. and Visweswaran, V. (1990). A global optimization algorithm (GOP) for certain classes of nonconvex NLPs – I theory. *Computers chem. Engng.*, 14, 1397–1417.

Horst, R. and Tuy, H. (1990). *Global optimization. Deterministic approaches (3rd ed.).* Springer Verlag, Berlin, Germany.

Kearfott, R.B. (1996). *Rigorous Global Search: Continuous Problems*, volume 13 of *Nonconvex Optimization and Its Applications.* Kluwer Academic Publishers.

Papamichail, I. and Adjiman, C.S. (2002). A rigorous global optimization algorithm for problems with ordinary differential equations. *J. Glob. Optimization*, 24, 1–33.

Papamichail, I. and Adjiman, C.S. (2004). Global optimization of dynamic systems. *Computers chem. Engng.*, 28, 403–415.

Pintér, J.D. (2002). Global optimization: Software, test problems, and applications. In P.M. Pardalos and H.E. Romeijn (eds.), *Handbook of Global Optimization, Volume 2: Heuristic Approaches*, 515–569. Kluwer Academic Publishers.

Rump, S.M. (1999). INTLAB - INTerval LABoratory. In T. Csendes (ed.), *Developments in Reliable Computing*, 77–104. Kluwer Academic Publishers, Dordrecht.

Singer, A.B. and Barton, P.I. (2001). Global solution of linear dynamic embedded optimization problems. *Journal of Optimization Theory and Applications*, 121(3), 613–646.

Čižniar, M., Podmajerský, M., Hirmajer, T., Fikar, M., and Latifi, M.A. (2009). Global optimization for the parameter estimation of differential-algebraic systems. *Chemical Papers*, 63(3), 274–283.

Williams, H.P. (1993). *Model Building in Mathematical Programming.* John Wiley & Sons, third edition.