SOME IDEAS TOWARDS GLOBAL OPTIMIZATION OF IMPROVED EFFICIENCY

LUBOMIR V. KOLEV

Dept. of Theoretical Electrotechnics, Faculty of Automatica,
Technical University of Sofia, 1756 Sofia, Bulgaria,

e-mail: lkolev@tu-sofia.bg

0. Introduction

Scope of the talk. The problem is in a rather general form:

(P) minimize $\varphi_0(x)$

subject to

$$\varphi_i(x) \le 0, \quad i = 1,..., r_1 ,$$
$$\varphi_i(x) = 0, \quad i = r_1 + 1,..., r_2,$$
$$x \in X^0 \subset R^n.$$

We consider a class of deterministic methods for solving (P), namely the interval branch and bound methods, more specifically: only first-order interval method.

Topic of the talk. Presentation of a few ideas on how to improve the numerical efficiency of such methods. Some of these ideas are new while others are less so (and, apparently, not well known to the interval global optimization community).

Toolkit. GICOLAG without AG.

I. Bound-and-branch versus branch-and-bound strategy

A. The standard strategy is **"branch-and-bound".**

Drawbacks:

- branching comes before bounding, so branching is "blind": the current box $X \in X^0$ is split into 2^n (Skelboe), or 2 sub-boxes (Moore), or 8 sub-boxes (Hansen);
- simple (but ineffective) local optimization is done in each X, so good bounding of the global minimizer x^* occurs at late iterations (when X becomes small enough);
- as a result, a lot of computational effort is wasted at the initial stage of the branch-and-bound strategy (before X has become small enough).

B. The new **bound-and-branch strategy**

1. 1. Use of a (cheap) non-interval method for global optimization

Apply it for the initial box X^0. Possible choices:

- multistart of a local optimization method;
- evolutionary method;
- annealing method;
- single run of a local optimization method.

Locate x^* (minimizer – global or local) with low precision by $\tilde{x}^*$.

We could use $\tilde{x}^*$ *in the known way* (constraint propagation, Hansen's techniques etc.) to (hopefully) reduce X^0 to a smaller box. However, *we can do better*.

1.2. Surrounding $\tilde{x}^*$ by a box $\tilde{X}^*$.

The box $\tilde{X}^*$ should be as large as possible and contain $\tilde{x}^*$ but no other minimum. Objective: find such a "sufficiently large" $\tilde{X}^*$ at an "acceptable cost". Why $\tilde{X}^*$ should be large? Two reasons:

- to avoid (reduce) the cluster effect;
- to facilitate removing sub-boxes in X^0 outside $\tilde{X}^*$.

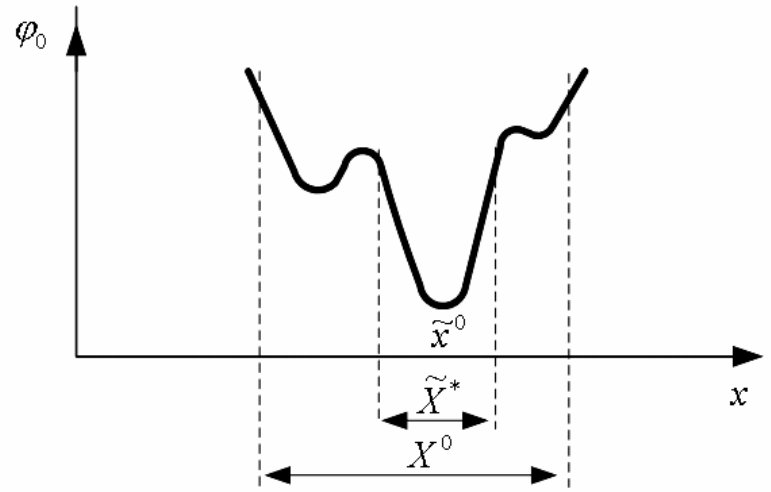


Fig.1

Possible solutions to objective – depending on the type of the problem at hand:

a) no constraints;

b) simple constraints: $x_i \in \lfloor \underline{x}_i , \overline{x}_i \rfloor$;

c) general form constraints given as equality and (or) inequality.

To simplify, we consider only case a).

One possible approach to finding $\tilde{X}^*$ is to form the system

$$f(x) = 0, \quad f : R^n \to R^n \tag{1}$$

where f is the gradient of φ_0. (In the general case c), system (1) will be given by the John-Fritz conditions.) Now $\tilde{x} \equiv \tilde{x}^*$ can be considered as an approximate solution to (1) so a possible candidate for $\tilde{X}^*$ can be a box where it is guaranteed that $\tilde{x}$ is an approximation to x^* (exact solution to (1)) and x^* is unique in $\tilde{X}^*$. This objective can be reached in various ways.

1. Apply the method of Schichl and Neumaier from "Exclusion regions for systems of equations" (2004) SIAM J. Num. An. It is, however, based on the use of first-and second-order slope matrices, a preconditioning matrix and numerous component-wise bounds and therefore seems to be rather costly. Many of the operations have to be repeated for several sub-boxes of an initial box $\tilde{X}$. The choice of $\tilde{X}$ remains unclear.

2. First apply Rump's method (program *verifynlss* in INTLAB) to prove uniqueness of x^* in a small box B. Then make use of

THEOREM 13.6.1 (in Hansen's book 1992). *Let* $f(x)$ *be a continuously differentiable vector function. Let* $J(X)$ *be the Jacobian over a box X. Assume* $J(X)$ *does not contain a single matrix. Then the equation* $f(x) = 0$ *can have no more than one solution in X.*

Hansen's result is very useful in our context since *we have already established the presence of a solution* x^* *and it only remains to check for uniqueness of* x^* *in X.*

Thus, we can *prove uniqueness by checking regularity (strong regularity) of* $J(\tilde{X})$.

A better choice seems to be checking regularity of the slope matrix $S(x, x^*)$, $x \in \tilde{X}$, (not of $J(X)$) by a new method (later in Section 3.2).

Finally, $\tilde{X}^*$ is found by inflating B: let $X_k = (1 + k\,\varepsilon)\,B$, k = 1, 2, ……. and check regularity of X_k until the regularity test fails for some $\bar{k}$ (ε can depend on k).

1.3. Creating boxes adjacent to $\tilde{X}^*$.

Now assume $\tilde{X}^*$ has been determined. The next step is to determine the complement $D = X^{(0)} - \tilde{X}^*$ as a union of boxes. We can use two approaches.

1.3.1. A procedure due to Kalovics and Meszaros (RC 1997): D may consist of 1 to $2n$ adjacent boxes D_i. A graphical illustration:

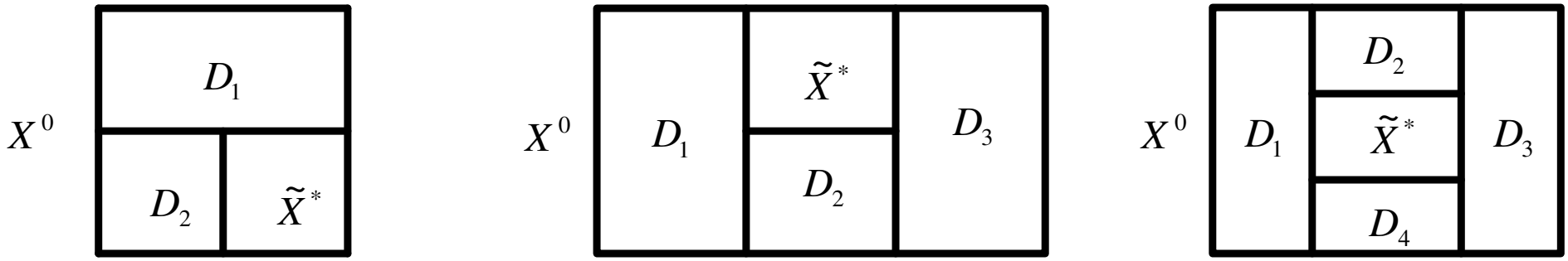


Fig.2

1.3.2. Traditional splitting: split $X^{(0)}$ (along the largest side) into two halves, choose that half which contains x^*, split again until we get near $\tilde{X}^*$ and then adjust splitting in such a way that the last "half" covers $\tilde{X}^*$. Unlike the previous approach, the number of generated adjacent boxes will, generally, be $> 2n$.

1.4. Eliminating adjacent boxes.

Assume x^* is a unique solution to (1) in $X^{(0)}$. Obviously we would like to discard all boxes adjacent to $\tilde{X}^*$.

DEFINITION 1.

Two boxes D_0 and D_i are adjacent if:

$$D_0 \cap D_i \neq \varnothing,$$

$$\text{int}(D_0) \cap \text{int}(D_i) = \varnothing.$$

Thus, the adjacent boxes are not disconnected, but do not overlap; they are just touching.

DEFINITION 2. Let D_0 denote $\tilde{X}^*$. A union of m adjacent boxes D_i, $i = 0, 1, ..,m$, will be called a *box union* if it is a box.

Thus, the union of D_2, D_0 and D_4 in Fig. 2, c) is a box union while the union of D_3 and D_4 is not.

LEMMA 1. *Let $f : D \to R^n$ be a continuously differentiable function where $D \subset R^n$ is a box union of adjacent boxes. Let x^* be a zero of f in D. Further, let*

$J_R(D)$ denote the range in D of the Jacobian matrix J(x) (or slope matrix S(x, x^*)) associated with f. Assume $J_R(D)$ is regular. Then x^* is the unique zero of f in D.

Proof. Assume that there are two solutions, x^* and y, of $f(x) = 0$ in D. Then $f(y) = f(x^*) + J'(y - x^*)$ for some $J' \in J_R(D)$. As $f(y) = f(x^*) = 0$, $J'(y - x^*) = 0$. Since J' is non-singular, it follows that $y = x^*$, i.e. the zero x^* of f in D is unique. □

THEOREM 1. *Let the set $D \in R^n$ be the box union of two adjacent boxes D_0 and D_i, $i \in \{1, ..., n\}$. Let $f : D \to R^n$ be a continuously differentiable function in D. Suppose f has a zero x^* in the interior of D_0. Assume the range matrix $J_R(D_0)$ (associated with f in D_0) is regular. If $J_R(D_i)$ is also regular, then D_i does not contain any zero of f.*

Proof. Since $D = D_0 \bigcup D_i$ and $J_R(D_0)$ and $J_R(D_i)$ are both regular, so is $J_R(D)$. But x^*, being a zero of f in D_0, is also a zero of f in D. Hence,

according to Lemma 1, x^* is the unique zero of f also in D. But $\text{int}(D_0) \cap \text{int}(D_i) = \varnothing$ so D_i cannot contain a zero of f since x^* is in $\text{int}(D_0)$. □

On account of Theorem 1:

THEOREM 2. (Test for enlarging D_0). *Let x^* be a zero of f in D_0. Assume the interval extension $J(D_0)$ of J(x) (of $S(x, x^*)$) in D_0 is a regular matrix. Let D_i be a box adjacent to D_0 and let their union be a box union. If the interval extension $J(D_i)$ is also a regular matrix, then the box D_i can be added to D_0, i.e. D_0 is enlarged to become the union $D = D_0 \bigcup D_i$.*

Remark 1. Thus (unlike the standard approach), testing D_i for "elimination" amounts to testing an interval matrix for regularity.

Remark 2. Initially, $D_0 = \tilde{X}^*$.

Remark 3. It should be stressed that the box union requirement imposes additional splits and a new splitting rule in case of the approach from 1.3.1.

Splitting rule. If the union of D_i and the remaining adjacent boxes is not a box union, split the box D_i in such a way that one of the halves D_i^1 or D_i^2 forms a box union with the remaining adjacent boxes.

Remark 4. Splitting will also be needed in case of the approach from 1.3.2 whenever $J(D_i)$ cannot be shown to be regular. In this case, D_i should be split in such a way as to satisfy the box union requirement.

Expectations for improvements of efficiency - depending on the type of problem:

class A: non-convex but a single minimum,

class B; multiple minima but a single global minimum,

class C: multiple global minima.

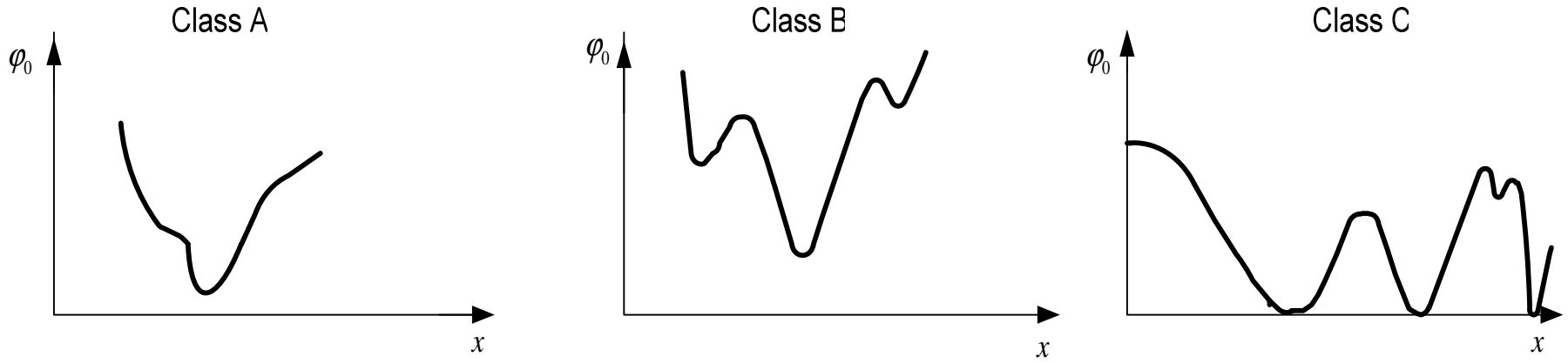


Fig.3

The improvements of efficiency will be less pronounced as we go down the classes, but rather noticeable for A and B (in electronics, about 35 – 40% of the optimization problems are B, the remaining are A).

II. Linear-interval versus interval-linear approximation
(Use of alternative interval linearizations of nonlinear functions)

2.1. Systems of nonlinear equations

A. Standard approach

Consider

$$f(x) = 0\ , \quad f:\ D_i \to R^n, \ \ x \in D_i \subset R^n. \tag{2}$$

We want to bound the solution(s) of (2), or establish their absence, by a first-order interval method. According to the standard approach (interval-Newton method or versions), this is done by repeatedly solving the linear interval system

$$J(X)(y - x) = -f(x)\ , \quad x \in X\ , \quad X \subseteq D_i. \tag{3}$$

It is based on the approximation of f in X:

$$f(y) \in f(x) + J(X)(y - x) \tag{4}$$

referred to here as *interval linear approximation* (ILA). After pre-multiplying (3) becomes

$$\boldsymbol{A}(X)(y-x)=b(x) \tag{5}$$

where $\boldsymbol{A}(X)$ – an interval matrix, $b(x)$ – a real vector (neglecting rounding).

B. A recent idea

An alternative approach to solving (2) has been suggested by the author in a number of publications using the following approximation of $f(x)$ in X :

$$f(x)\in A(X)\,x+\boldsymbol{b}(X),\quad x\in X \tag{6}$$

where $A(X)$ is a real (n x n) matrix while $\boldsymbol{b}(X)$ is an interval vector. The above approximation will be called a *linear interval approximation (LIA)* (just to distinguish it from ILA) or a "sandwich" approximation. In this case solving (2) reduces to repeatedly solve the interval system

$$A(X)\,y=-\boldsymbol{b}(X). \tag{7}$$

Now compare (5) and (7).

Shortcomings of (5):

- applicable only for continuously differentiable functions;
- overestimation of $A_{ij}(X)$ with respect to the range of $a_{ij}(x)$ in X;
- ignoring the dependencies among $x_1,...,x_n$ within each element $a_{ij}(x)$;
- ignoring the dependencies between the elements $a_{ij}(x)$;
- overestimation of an approximate "outer solution" of (5) with respect to the "hull solution" of (5).

In contrast, (7) has the advantages:

- applicable for functions that are only continuous;
- reflects the above mentioned dependencies;
- the solution set of (7) has a much simpler structure than that of (5);
- the hull solution of (7) is straightforward

$$Y = -A^{-1}(X)\boldsymbol{b}(X), \tag{8}$$

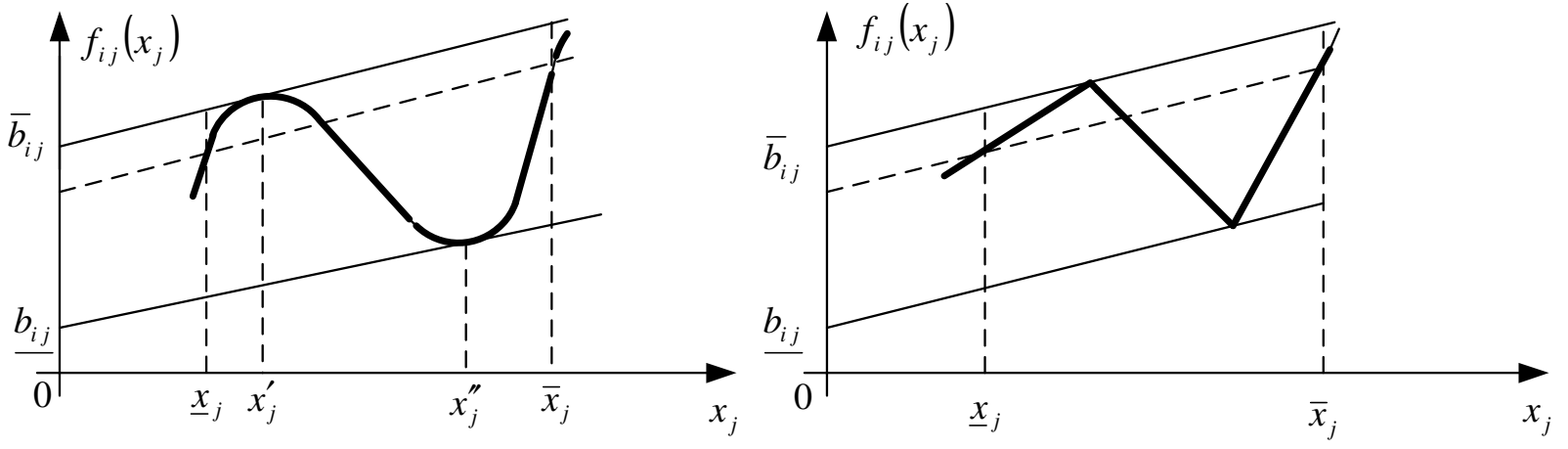
$$X' = Y \cap X. \tag{8'}$$

The approximation of $f(x)$ in (6)

$$L(x) = A\,x + \boldsymbol{b}\,, \quad x \in X \tag{9}$$

was suggested for the first time in K 1997 (European techn. conference) and K 1998 (Intern. techn. journal) in the case where each $f_i(x)$ is separable (see Fig.4):

$$f_i(x) = \sum_j f_{ij}(x). \tag{10}$$



a) b)

Fig. 4

In K 1998 (RC), the interval approximation was extended to non-separable $f(x)$ of n variables at the cost of introducing m auxiliary variables so A in (9) becomes $(n+m)\times(n+m)$. An improved version is proposed in K 1999 (RC): the auxiliary variables are eliminated (thus A is again $n\times n$) and constraint propagation is introduced; see also Kolev 2000 (IEEE CAS). Automatic computation of (9) (in the spirit of automatic computation of slopes) was suggested in K 2001 (RC).

The interval linearization (9) suggested by the author will be referred to as LIA1. Three alternative forms of (9) are also known, denoted for convenience as LIA2, LIA3 and LIA4.

LIA2 was reinvented independently in Buehler and Bart (2001) (intersection algorithm for parametric surfaces) and Braems, Kiefer and Walter (2001) (parameter estimation). In this case, the LIA form is obtained by equivalently transforming the standard ILA form as follows:

$$
\begin{aligned}
& A\,x + \boldsymbol{b} = 0 \\
& A = \boldsymbol{J}(x_0)\,, \quad x_0 = \mathrm{mid}(X) \qquad (7') \\
& \boldsymbol{b} = f(x_0) - A\,x_0 + \left[\boldsymbol{J}(X)\cdot X - A\right](X - x_0)
\end{aligned}
$$

Thus, the similarity of (7′) with (7) is only formal.

LIA3 (suggested in Buehler and Bart (2001) as a second algorithm) is based on affine arithmetic computations. LIA4 is obtained as the first-order case of the Taylor forms (Martin Berz).

Shortcomings of LIA2 to LIA4: *applicable only for continuously differentiable functions, worse approximation* as compared to LIA1 (since LIA1 provides the smallest width of the "sandwich").

To the best of my knowledge, LIA1 (unlike LIA4) is not known by the interval analysis community. Indeed, Hansen, Neumaier, Kearfort, Stadther, Nataraj have never mentioned or used it so far (though A. Neumaier studies the properties of LIA in his Taylor forms paper). On the other hand, numerical examples show that the improvement in efficiency is sometimes spectacular.

Example 1. The system to be solved is

$$g(x_i) + x_1 + x_2 + \ldots + x_n - i = 0\,, \quad i = 1,\ldots n\,,$$

$$g(x_i) = 2.5\,x_i^3 - 10.5\,x_i^2 + 11.8\,x_i,$$

$$X_i^0 = [-10, 10].$$

With accuracy $\varepsilon = 10^{-6}$:

the conventional Krawczyk-Moore method requires

for $n = 12$ about 3h,

for $n = 14$ about 44h

to bound all solutions;

a recent more sophistication method – Yamamura and Igarashi (2004) solves the system

for $n = 700$ in 50h 44 min

to find all the 9 solutions;

the simple method from Kolev 1997 reaches the same solutions

for n = 700 in only 97.16 sec. (number of iterations 101);

for $n = 2500$ it takes only about 1.5h!

Other examples (Walster's example, transistor design problem, etc.) confirm the superiority of LIA1 over the standard approach and the alternative forms LIA2 to LIA4. The better efficiency is due to the better approximation of the non-linear functions involved. To give an idea, consider the above univariate example

$$f(x) = 2.5\,x^3 - 10.5\,x^2 + 11.8\,x, \quad x \in X = [0, 2].$$

The LIA1 of f is determined by the following procedure.

Procedure 1. It reduces, essentially, to finding all real solutions x_k, $k = 1,2,...K$ of the equation

$$d(x) - a = 0 \qquad (*)$$

in X where $d(x)$ denotes the first derivative of f and a is the slope of f in X, i.e.

$$a = (f(\bar{x}) - f(\underline{x}))/(\bar{x} - \underline{x})$$

Using x_k and $x_0 = \underline{x}$, the following quantities are computed

$$b_k = f(x_k) - a\,x_k, \quad k = 0, 1,..., K$$

Finally

$$\underline{b} = \min\{b_k\}, \quad \bar{b} = \max\{b_k\}, \quad k = 0,1,..., K$$

The efficiency of Procedure 1 depends principally on how much computational effort is needed to locate all real solutions of (*) in X. If the latter problem is easy to solve, then the procedure provides a simple and efficient way of determining LIA1. (If (*) is difficult to solve, less accurate, i.e. broader but cheaper approximations are possible – K 2001).

For our example, LIA1 is given by $L(x) = a\,x + B$ with

$a = 0.8$, B = [0, 3.4125].

It is interesting to compare LIA1 to LIA2, which by (7') is determined by

$a = -1.7$, B = [-53.2, 64.2].

Obviously, LIA2 is much worse than LIA1.

2.2. Improved LIA

Finally, an improved linear interval linearization in the form

$$L(x, y) = Ax + By + b\,, \quad x \in \boldsymbol{x}\,, \quad y \in \boldsymbol{y} \tag{11}$$

was suggested in K 2002 (SCAN) and appeared in K2004 (NUMA). It is shown to have a better performance as compared to LIA1.

2.3. Use of linear programming

In the framework of (5), the idea of using LP goes back to Neumaier (1988); may be the latest implementation is in Lin and Stadther (2004).

In the framework of (7), this idea was suggested in Kolev (1978), Kolev and co-author (1998), (2000). Using LP, one can do better than computing

$$X' = Y \cap X . \qquad (8')$$

The improvement is possible if, rather than the hull solution Y, one uses the solution set S of (7):

$$X'' = H(S \bigcap X). \qquad (12)$$

Geometrical illustration:

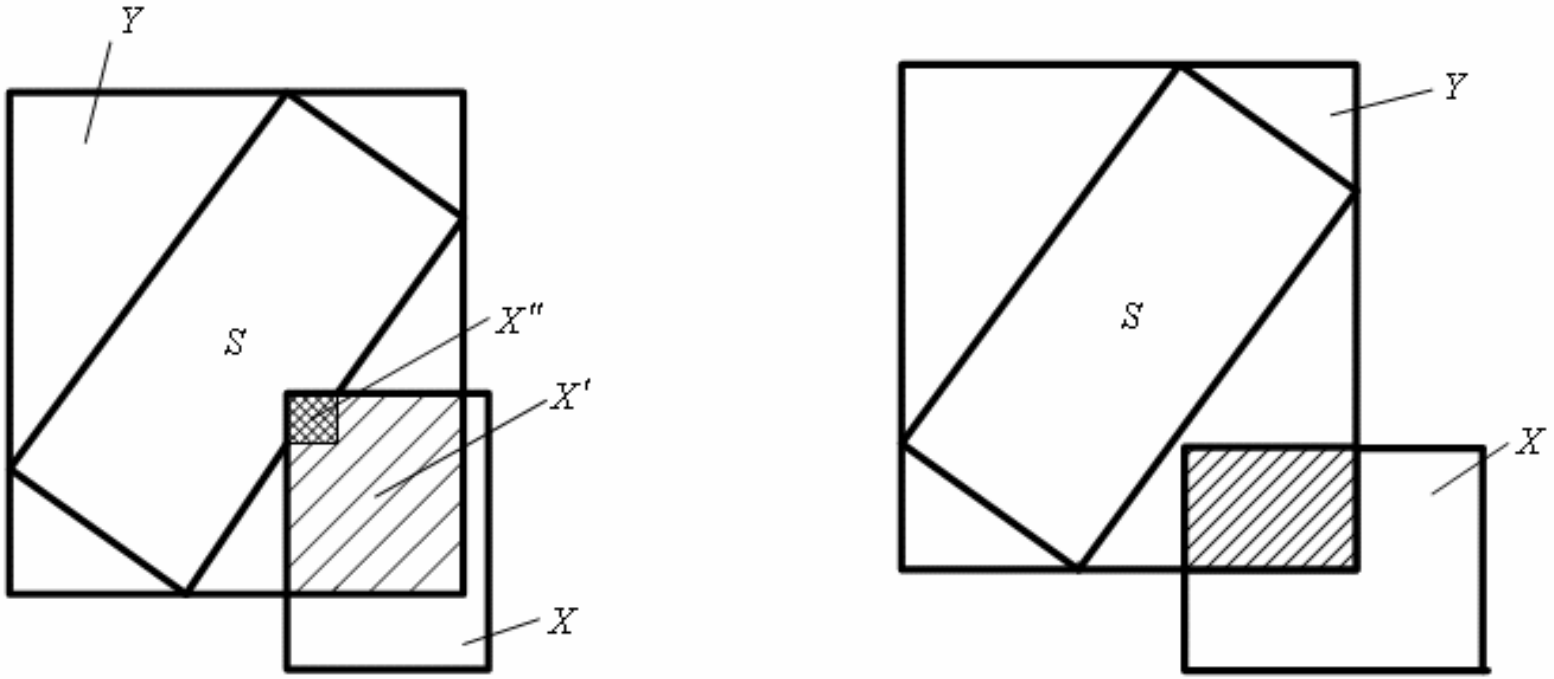


Fig. 5

First advantage - in general

$$X'' \subset X'. \tag{13}$$

Second advantage: using S for discarding is better:

$$S \bigcap X = \varnothing \tag{14}$$

while

$$Y \bigcap X \neq \varnothing. \tag{14'}$$

The box X'' can be determined by solving $2n$ linear programming (LP) problems:

$$\text{(LP)} \qquad \begin{aligned} &\min x_k \text{ or } \max x_k\ , \quad k = 1,...,n \\ &\sum_{j=1}^{n} a_{ij} x_j - x_{i+n} = 0\ , \quad i = 1,...,n\ , \\ &\underline{X}_i \le x_i \le \overline{X}_i\ , \quad i = 1,...,n\ , \\ &\underline{b}_i \le x_{i+n} \le \overline{b}_i\ , \quad i = 1,...,n. \end{aligned} \tag{15}$$

This "brute force" approach was reported in K+M 1998 (International techn. conference): reduced number of iterations but bigger overall computing time.

An improved LP implementation was suggested in Kolev and Nenov (2000) (Intern. Conf. in Poland):

- setting up and solving *one single* LP problem
- combined with *constraint propagation*.

The first algorithm of the second (dual) method for bounded variables from Udin and Goldstein (1963 - in Russian) was used. Our implementation seems to be better than that of L. and Stadtherr since there $2n$ LP problems are solved (each twice as big as (15)) and no constraint propagation is used.

2.4. Direct application of LIA to global optimization.

The problem – the general non-linear programming:

$$\begin{aligned} &\text{minimize } \varphi_0(x) \\ &\quad \text{subject to} \\ &\quad\quad \varphi_i(x) \le 0\,, \quad i = 1,\ldots r_1 \\ &\quad\quad \varphi_i(x) = 0\,, \quad i = r_1 + 1\,,\ldots r_2 \qquad (16a) \\ &\quad\quad x \in X^0 \in R^n \end{aligned}$$

Suppose we have found an upper bound $\tilde{\varphi}_0$ on the global minimum φ_0^*. Adding

$$\varphi_0(x) - \tilde{\varphi}_0 \le 0 \qquad (16b)$$

we obtain the system

$$\psi_i(x_1,...,x_m) \cong 0, \quad i = 1,....., r_3, \tag{17}$$

where the symbol ≅ stands for either the equality or inequality sign.

A. The ILA model of (17)

If we use ILA to enclose (17) (see, for instance Hansen's paper 1980 and his books 1992, 2004), we have to deal with the system (the full Monty in H. 2004)

$$\sum_{j=1}^{n} \boldsymbol{A}_{ij}(X)(y_j - x_j) \cong b_i(x), \quad i = 1,..., r_3 \tag{18}$$

where $\boldsymbol{A}_{ij}(X)$ is the interval extension of a corresponding derivative (or slope) while $b_i(x)$ is (in exact arithmetic) a real number.

Factors limiting the efficiency of such a "standard" approach:

(i) overestimation of $\boldsymbol{A}_{ij}(X)$ with respect to the corresponding range ;

(ii) ignoring the dependencies among $x_1,...,x_n$ within each element $a_{ij}(x)$;

(iii) ignoring the dependencies between the elements a_{ij};

(iv) overestimation of an approximate "solution" ψ of (18) with respect to the "optimal" solution ψ^* of (18) or, still worse, the "hull solution" of (17).

B. The LIA1 model of (17)

This model was suggested in the paper "A new approach to global optimization": (K. 1999, Computational maths. and math. Physics, Moscow – in English). Now, instead of (17), we have

$$\sum a_{ij}(X)\, x_j \cong B_i(X)\,, \quad i = 1,..., r_3 \ . \tag{19}$$

"Solving" (19) is:

- much easier than (17);
- the dependencies (of both kinds) are taken into account.

New algorithms are possible:

1) K + Penev (1999, an article in a tech. book) – (simple + inequality constraints);
2) K - (ISTET 1999) : the algorithm includes *constraint propagation*;
3) K + Penev (2000) ISCAS, IEEE: *LP* + *constraint propagation*.

III. Use of a linear parametric approximation

This is yet another alternative linearization of nonlinear functions. However, it differs so much from the standard approximation (ILA)

$$L(X, x) = f(x) + J(X)(y - x), \tag{20}$$

the "sandwich" approximation (LIA1)

$$L_s(X) = A(X)\,x + \boldsymbol{b}(X) \tag{21}$$

and the improved approximation (ILIA)

$$L_s(x, y) = A\,x + B\,y + b \tag{21a}$$

that I consider it in a separate section.

3.1. Solving non-linear systems

Once again, we want to solve globally

$$f(x) = 0, \tag{22a}$$

$$x \in X^0 \subset R^n. \tag{22b}$$

A. Standard ILA approach

The standard ILA framework, using the interval matrix $J(X)$ (standing for the interval extension of the corresponding derivative or slope matrix), is based on the inclusion

$$f(y) \in f(x) + J(X)(y - x). \tag{23}$$

If y is a zero, $f(x) = 0$ and we get the interval linear system

$$J(X)\, z = b(x). \tag{24}$$

B. The novel approach

It is based on the use of the slope matrix $S(y, x)$ and the equality

$$f(y) = f(x) + S(y,\ x)\ (y - x), \tag{25}$$

where y and x have some fixed values. We will "free" the components y_k of y and consider them as components of a parameter vector p, i.e.

$$p = (y_1, ..., y_n) \in X = (X_1, ..., X_n). \tag{26}$$

Let

$$a_{ij}(p) = S_{ij}(p, x) \tag{27}$$

be the entries of the parametric matrix $A(p)$. On account of (25) to (27)

$$y \in f(x) + A(p)(y - x), \quad p \in X. \tag{28}$$

The right-hand side of (28) is the novel *linear parametric approximation* (LPA) of f in X. If y is a zero

$$A(p)(y - x) = -f(x), \quad p \in P \tag{29a}$$

or, equivalently

$$A(p)\, z = b(x), \quad p \in P. \tag{29b}$$

Thus, using the novel approximation, we obtain the *linear parametric equation* (29). We argue that *(29) is a better way than (24) to bound the solutions of (22).*

Indeed, introduce the solution sets

$$S_J = \{z: \quad J\, z = b, \quad J \in J(X)\}, \tag{30}$$

$$S_p = \{z: \quad A(p)\, z = b, \quad p \in X\}. \tag{31}$$

It is seen that while $J(X)$ depends on n^2 independent entries, there are only n independent elements in $A(p)$. Thus, it follows from (30) and (31) that

$$S_p \subset S_J. \tag{32}$$

The inclusion (32) remains valid even if $J(X)$ is the range of $J(x)$ over X. , i.e. when

$$J_{ij}(X) = \{J_{ij}(x): \quad x \in X\}.$$

In practice, the elements $J_{ij}(X)$ are interval extensions of $J_{ij}(x)$ in X, i.e. they are larger than the ranges and, what is more, they are independent. So the description (29) is much more accurate than (24).

Let $Z^{\mathrm{J}}_{\mathrm{hull}}$ and Z^{p}_{hull} denote the interval hull of S_J and S_p, respectively. From (32)

$$Z^{p}_{\mathrm{hull}} \subset Z^{J}_{\mathrm{hull}}. \tag{33}$$

If Z^{J}_{out} and Z^{p}_{out} denote some outer interval solution of (24) and (29), respectively, then we can expect that also

$$Z^{p}_{\mathrm{out}} \subset Z^{J}_{\mathrm{out}}. \tag{34}$$

However, it all depends how we compute the outer solutions.

As regards Z^{J}_{out} - numerous publications and numerous methods. A recent method due to Hansen (book 2004) (applicable if some sufficient conditions are satisfied) seems to be preferable (pre-multiplication of (25) and then Z^{J}_{out} is found as the hull solution of the modified system).

A simple method for computing Z^{p}_{out} (*in the general case* where the coefficients in (28) are *arbitrary continuous non-linear functions*) was suggested

in K – SCAN 2002, published later in K – NUMA 2004. It was developed as an extension of a method K 2002 (RC) for solving a special kind of linear interval parameters equations when the coefficients in (28) are affine functions. An improvement of the method K2004 was suggested in K 2006 (RC) leading to a narrower Z_{out}^{p}. The method K 2002, K 2006 is based on the LIA1 approximation of $A(p)$:

$$A(p) \in A^{(0)}\ p + \boldsymbol{b}\ , \quad p \in X\ , \tag{35}$$

and reduces computationally to inverting a real $n \times n$ matrix and solving a system of n real linear equations. It is applicable if the solution of the latter system is positive. More precisely we approximate

$$a_{ij}(p) = a_{ij}(p_1, ..., p_n) \tag{36}$$

by

$$\boldsymbol{L}_{ij}(p) = \sum_{k=1}^{n} \alpha_{ijk}\, p_k + \boldsymbol{a}_{ij}\ , \quad p \in X \tag{37}$$

and form the real matrix L^0 whose elements are the centers of (37) i.e.

$$L_{ij}^{0} = a_{ij}^{0} + \sum_{k=1}^{n} \alpha_{ijk}\, p_{k}^{0} . \tag{38}$$

Next, we solve the system

$$L^{0} z^{0} = b. \tag{39}$$

Using z^{0}, α_{ijk} , $B=\left(L^{0}\right)^{-1}$ and $\boldsymbol{a}_{ij}$, we form the real system

$$\left(I - D\right) r = c \tag{40}$$

where I is the n-dimensional identity matrix, D is a non-negative matrix $\left(D_{ij} \geq 0\right)$ and c is a non-negative vector ($c_i \geq 0$). Here

$$c = \left|C\right| r^{p} + \left|B\right| r^{a} , \tag{40a}$$

$$D = \left|B\right| R ,$$

with $C = BA^{u}$,

$$A_{ik}^{u} = \sum_{j=1}^{n} \alpha_{ijk}\, z_{j}^{0} \ , \quad \boldsymbol{a}_{i} = \sum_{\boldsymbol{j=1}}^{n} z_{j}^{0} \boldsymbol{a}_{ij} \ ,$$

$$R_{ij} = \sum_{k=1}^{n} \left|\alpha_{ijk}\right| r_{k}^{p} + R_{ij}^{a} \ ,$$

where r_k^p is the radius of P_k , r^a is a vector whose components are the radii of $\boldsymbol{a}_i$ while R_{ij}^a is the radius of $\boldsymbol{a}_{ij}$.

THEOREM 3 (Theorem 1 from K 2004). *Assume* L^0 *and* $T = I - D$ *are nonsingular. If the solution* r^* *of (40) is positive* ($r_i^* > 0$), *then*

(i) the interval vector

$$\boldsymbol{z} = z^0 + \boldsymbol{h} \tag{41a}$$

where

$$\boldsymbol{h} = \left[- r^* , r^* \right] \tag{41b}$$

is an outer solution to (29);

(ii) *the matrix* $A(p) = S(p, x)$ *is non-singular for each* $p \in X$.

The method can also be used (after obvious simplification) for finding Z_{out}^J and produces an outer solution which is only slightly larger than Hansen's outer solution. It should, however, be stressed that using $\boldsymbol{z}$ rather than Z_{out}^J leads to a better method for solving (22) since $\boldsymbol{z} \subset Z_{\text{out}}^{\text{J}}$.

3.2. Checking regularity of interval parameters matrices

Theorem 3 provides a sufficient test for checking regularity of $S(p,x)$, $p \in X$. Indeed, we can use (29) and choose b in some "nice" way to, hopefully, improve our chances to successfully check regularity, or to simplify computation.

THEOREM 4 (Test for regularity). *Let* $b(x)$ *in*

$$S(p,x)\,z = b(x)\,, \quad p \in X \tag{29}$$

be the vector $b = (1,...,1)$. *Set up the system*

$$T\ r = c\,. \tag{40}$$

If its solution r is positive, then the parametric slope matrix* $S(p,x)$ *is regular for each* $p \in X$.

Theorem 4 can be used as an "elimination" test for D_i. We first need the following result.

LEMMA 2. *Let* $f : D \to R^n$ *be a continuously differentiable function where* $D \subset R^n$ *is a box union of adjacent boxes. Let* x^* *be a zero of* f *in D. Further, let* $S(x, x^*)$ *denote the slope matrix associated with f. Assume* $S(x, x^*)$ *is regular in D. Then* x^* *is the unique zero of* f *in D.*

The proof is similar to that of Lemma 1.

THEOREM 5 (Test for enlarging D_0). *Let* x^* *be a zero of* f *in* D_0 where D_0 equals $\tilde{X}^*$ *a box union of* $\tilde{X}^*$ *and adjacent boxes . Assume the parametric matrix* $S(p, x^*)$, $p \in D_0$ *is regular. Let* D_i *be a box adjacent to* D_0. *Further, let* D_i *and* D_0 *form a box union. If the parametric matrix* $S(p, x^*)$, $p \in D_i$ *is also regular, then* D_i *can beaded to* D_0.

3.3. Checking convexity

We can also apply Theorem 3 to check positive definiteness of interval parameters matrices $A(p)$, $p \in X$. In this case, we have additionally

$$A_{ij}(p) = A_{ji}(p) \qquad (36')$$

which will lead to

$$\boldsymbol{L}_{ij}(p) = \sum_{k=1}^{n} \alpha_{ijk}\, p_k + \boldsymbol{a}_{ij} = \boldsymbol{L}_{ji}(p), \quad p \in X. \qquad (38')$$

We have the following result.

THEOREM 6 (Test for positive definiteness). *Assume the real symmetric matrix* L^0 *whose elements are the centers of (38') is positive definite. Choose the vector b in (29) to be* $b = (1,...,1)$. *Set up system (40). If its solutions r* is positive, then the parameters matrix* $A(p),\ p \in X$ *is positive definite in X.*

Remark 5. We can have a less conservative p.d. test if the symmetry of $A(p)$ is taken into consideration in computing the vector *c* from (40a).

If $A(p)$ stands for the Hessian $H(x)$ of a function $f(x)$ in *X*, then Theorem 6 provides a test for checking if f is convex in X. *Note* that such a test will lead to a much better result (lesser conservatism) as compared to the known criteria where $H(x), x \in X$ is replaced with an interval extension $\boldsymbol{H}$ (using $\boldsymbol{H}$ we ignore the interdependence between the elements of $H(x)$).

IV. Conclusion

In summary: a few ideas for improving the numerical efficiency of a class of interval methods for global optimization. It should, however, be stressed that *the approach from Section I is applicable for any other class of global optimization methods (for instance, the convexification method implemented in BARON).*

Some of the ideas presented – from Section II – have been in existence for a couple of years and have been developed into methods and algorithms; we have already a certain (though rather restricted) experience on their relative efficiency. The others – those of Sections I and III – are quite new (suggested here for the first time). They look promising but need to be implemented into concrete methods, algorithms and codes, and then tested numerically. It is hoped that thus the following open questions will be answered:

- how is the box $\tilde{X}^*$ from Section 1.2 determined in an efficient way;
- how do the linear approximations LIA1 (its best LP implementation), ILIA and LPA compare between each other, which is preferable in what context;
- which of the two approaches 1.3.1 or 1.3.2 is globally more efficient;

- how do the novel techniques fit the best way into the computational schemes of GICOLAG;
- fine tuning of the novel techniques to the remaining techniques of GICOLAG;
- extend the present ideas (suggested mainly in the context of global solution of systems) to the various global optimization problem formulations.

This is a lot of work but it seems to be worth the efforts.

Acknowledgement. The author is grateful to Prof. A. Neumaier for numerous helpful discussions of the topic of the talk.

REFERENCES

I. Braems, M. Kieffer and E. Walter. Set estimation, computation of volumes and data safety, in: Scientific Computing, Validated Numerics, eds. Kramer and Wolf von Gudenberg, Kluwer, Academic/Plenum Publishers, New York, 2001, pp. 267-278.

K. Buhler and W. Barth. A new intersection algorithm for parametric surfaces based on linear interval estimations, in: Scientific Computing, Validated Numerics, eds. Kramer and Wolf von Gudenberg, Kluwer, Academic/Plenum Publishers, New York, 2001, pp. 179-190.

E. Hansen and S. Sengupta. Global constrained optimization using interval analysis, in Nickel, 1980, pp. 25-47.

E. Hansen. Global optimization using interval analysis. Marcel Dekker Inc. (1992).

E. Hansen and W. Walster. Global optimization using interval analysis. Marcel Dekker Inc. (2004).

F. Kalovics and G. Meszaros "Finding global minima of maximin functions by using exclusion functions without derivatives", *Reliable Computing* **3,** 1997, pp. 381-399.

L.V. Kolev. Minimum-fuel minimum-time control of linear discrete systems. *Int. J. Control*, v.27, № 1, 1978, pp. 21-29.

L. Kolev. A general interval method for global nonlinear dc analysis. *Proc. of the 1997 European Conf. on Circuit Theory and Design, ECCD'97*, Technical University of Budapest, 30th August - 3rd Sept., 1997, Budapest, Hungary, volume 3, pp.1460-1462.

L. Kolev. An efficient interval method for global analysis of nonlinear resistive circuits. *Int. J. of Circuit Theory and Appl.* vol. 26, 1998, pp.81-92.

L. Kolev. A new method for global solution of nonlinear equations. *Reliable Computing*, **4** , № 2, 1998, pp. 125-146.

L. Kolev, V. Mladenov. A linear programming implementation of an interval method for global non-linear DC analysis. *IEEE International Conference on Electronics, Circuits and Systems*. Surfing the waves of Science and Technology, Instituto Superior Tecnico Lisboa, Portugal 7-10 September, volume 1, 1998, pp.75-78.

L. Kolev. An Improved Method for Global Solution of Non-Linear Systems. *Reliable Computing*, vol.5, No 2, 1999, pp.103-111.

L. Kolev. A new interval approach to global optimization. *Computational Mathematics and Mathemat. Physics, Russia*, vol. 9,№.5, 1999, pp. 727 – 737.

L. Kolev, D. Penev. An Interval Method for Global Inequality Constraint Optimization Problems. In *"Advances in Intelligent Systems and Computer Science, Word Scientific and Eng. Society Press*, 1999, pp. 53-56.

L. Kolev. A new interval method for global optimization. *ISTET'*99, Magdeburg, Germany, 6-9 Sept., 1999, pp. 625-629.

L. Kolev and I. Nenov. A combined interval method for global solution of nonlinear systems. *Proc. of the XXIII International Conference on Fundamentals of Electronics and Circuit Theory – SPETO 2000*, Gliwice, Poland, 24-27.V 2000, pp. 365-368.

L. Kolev. An Iterval Method for Global Nonlinear Analysis. *IEEE Trans. on Circuits and Systems-I*, May, 2000, vol.47, pp.675 - 683.

L. Kolev and D. Penev. An interval method for global inequality-constraint optimization problems. *ISCAS 2000 – IEEE Intern. Symposium on Circuits and Systems*, May 28-31, 2000, Geneva, Switzeland, pp. IV-617 to IV-620.

L. Kolev. Automatic Computation of a Linear Interval Enclosure. *Reliable Computing*, vol. 7, No.1, 2001, pp. 17-28.

L. V. Kolev. Solving Linear Systems whose elements are nonlinear functions of interval parameters. *Proc. of Intern. Symposium of Scientific Computing, Computer Arithmetic, and Validated Numerics , SCAN 2002*, September 24-27, Paris, France, 2002, p. 45.

L.V. Kolev. An Improved interval linearization for solving non-linear problems. *Proc. of Intern. Symposium of Scientific Computing, Computer Arithmetic, and Validated Numerics , SCAN 2002*, September 24-27, Paris, France, 2002, p.122.

L. Kolev. Outer Solution of Linear Systems Whose Elements are Affine Functions of Interval Parameters. *Reliable Computing*, **8**, 2002, pp. 493-501.

L.V. Kolev. An improved interval linearization for solving non-linear problems. Numerical Algorithms, **37**: pp. 213-224, 2004.

L. V. Kolev. Improvement of a Direct Method for Outer Solution of Linear Parametric Systems**.** Reliable Computing, Volume 12, issue 3, 2006, pp. 193 - 202.

Y. Lin and M. A. Stadtherr, "LP Strategy for Interval-Newton Method in DeterministicGlobal Optimization," Ind. Eng. Chem. Res., 43, pp. 3741-3749 (2004).

A. Neumaier. Enclosure of solutions of parameter-dependent systems of equations. In Realability in Computing, Acad. Press, 1988, pp. 269-286.

A. Neumaier, Taylor forms - use and limits, Reliable Computing 9 (2002), 43-79.

H. Schichl and A. Neumaier “Exclusion regions for systems of equations” (2004) SIAM J. Num. An. 42 (2004), pp. 383-408.

Udin and Goldstein (1963 - in Russian)

K. Yamamura and N. Igarashi. An interval algorithm for finding all solutions of non-linear resistive circuits. *Int. J. Circ. Theor. Appl.*, 2004; **32**; pp. 47-55.