Rigorous Global Search using Taylor Models

Martin Berz
Department of Physics and Astronomy
Michigan State University
East Lansing, Michigan 48824, USA
berz@msu.edu

Kyoko Makino
Department of Physics and Astronomy
Michigan State University
East Lansing, Michigan 48824, USA
makino@msu.edu

ABSTRACT

A Taylor model of a smooth function f over a sufficiently small domain D is a pair (P, I) where P is the Taylor polynomial of f at a point d in D, and I is an interval such that f differs from P by not more than I over D. As such, they represent a hybrid between numerical techniques for the interval and the coefficients of P and algebraic techniques for the manipulation of polynomials. A calculus including addition, multiplication and differentiation/integration is developed to compute Taylor models for code lists, resulting in a method to compute rigorous enclosures of arbitrary computer functions in terms of Taylor models. The methods combine the advantages of numeric methods, namely finite size of representation, speed, and no limitations on the objects on which operations can be carried out with those of symbolic methods, namely the ability to treat functions instead of points and making rigorous statements.

We show how the methods can be used for the problem of rigorous global search based on a branch and bound approach, where Taylor models are used to prune the search space and resolve constraints to high order. Compared to other rigorous global optimizers based on intervals and linearizations, the methods allow the treatment of complicated functions with long code lists and with large amounts of dependency. Furthermore, the underlying polynomial form allows the use of other efficient bounding and pruning techniques, including the linear dominated bounder (LDB) and the quadratic fast bounder (QFB).

Categories and Subject Descriptors

I.1.2 [**Symbolic and Algebraic Manipulation**]: Algorithms; G.1.6 [**Numerical Analysis**]: Optimization—*global optimization, nonlinear programming*

General Terms

Algorithms

Keywords

Taylor model, rigorous global optimization, Taylor polynomial, error bound, interval arithmetic, dynamical system, Henon map, Henon attractor, Beale function

1. INTRODUCTION

An n-th order Taylor model of a multivariate function f that is $(n + 1)$ times continuously partially differentiable on the domain D consists of the n-th order multivariate Taylor polynomial P, expanded around a point $x_0 \in D$ and representing a high order approximation of the function f, and a remainder error bound interval I for verification, the width of which scales in $(n + 1)$-st order that satisfy

$$f(x) \in P(x - x_0) + I \text{ for all } x \in D.$$

The definition of Taylor models, Taylor model binary operations, and Taylor model intrinsic functions were thoroughly developed[6, 4], and a verified implementation of Taylor model arithmetic in the code COSY INFINITY[3] is based on the floating point treatment of the Taylor coefficients. The details about the implementation can be found in [10, 6, 4]. Utilizing the Taylor model concept and arithmetic, various Taylor model based schemes have been developed. For example, see [6, 2, 5, 1].

The exact definition of Taylor models implemented in COSY INFINITY, namely representing the underlying function f by a high order Taylor multivariate polynomial P with floating point coefficients and enclosing the function f around P by a remainder error bound interval I that scales in high order, allows us to develop numerous efficient algorithms for various problems that require practical verification. We have shown in various publications (for example, see [6] and references therein.) the properties of Taylor models, which enable the superb performance of Taylor model based schemes compared to conventional schemes. In short words, those properties are 1) sharpness of remainder intervals, 2) suppression of the dependency problem, and 3) reducing the curse of dimensionality.

Besides those repeatedly demonstrated properties, the structure of Taylor models naturally has a rich resource of information. Namely the coefficients of the polynomial part P of a Taylor model are nothing but the derivatives up to order n. That means when representing a function f by a Taylor model (P, I) on a computer, we also obtain the local slope, Hessian and higher order derivatives free. When a task is focused on range bounding, those pieces of information become particularly useful.

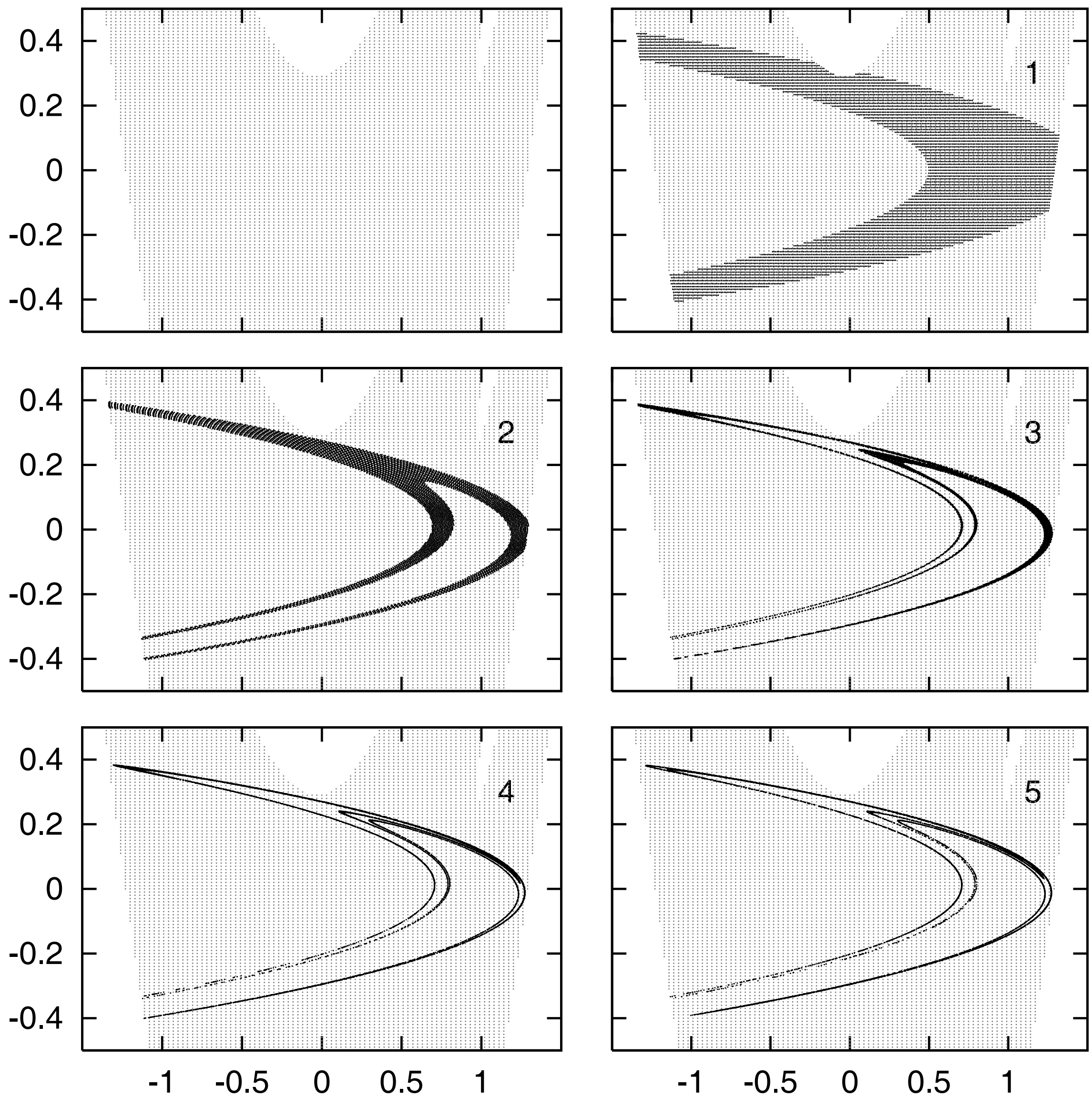

Figure 1: Investigation of a trapping region of the Henon map H.

While naive range bounding of Taylor models, namely merely evaluating each monomial of P using interval arithmetic then summing up all the contributions as well as the remainder interval I [4], already exhibits the superiority over the mere interval arithmetic and the more advanced centered form [6], the active utilization of those additional pieces of information in Taylor models has a lot of potential of developing efficient range bounders. Based on this observation, we have developed various kinds of Taylor model based range bounders, and among them, we will introduce the particularly efficient bounders, namely the linear dominated bounder (LDB) and the quadratic fast bounder (QFB) below, which are the backbones of Taylor model based verified global optimizer COSY-GO[7] that will be discussed afterward as well as some examples.

2. PERFORMANCE OF TAYLOR MODEL METHODS: TIGHT ENCLOSURES FOR THE HENON ATTRACTOR

To illustrate the performance of Taylor model methods and their advantages over other verified methods including the interval approach, we consider an important example from the theory of dynamical systems, the Henon attractor. Specifically, we consider the standard Henon map

$$H(x, y) = (1 - ax^2 + y, bx).$$

We set the parameters $a = 1.4$ and $b = 0.3$, which are the values originally considered by Henon. It is easy to see that the map H has two fixed points

$$\vec{p}_1 \approx (0.63135, 0.18940) \quad \text{and} \quad \vec{p}_2 \approx (-1.13135, -0.33941).$$

Different from other dynamical systems with attractors, the

Henon map has the property that not all points in the plane approach the attractor; rather, many diverge to infinity.

2.1 Non-Rigorous Assessment of the Henon Attractor

We first perform a numerical experiment illustrating this phenomenon. Specifically, we study the region

$$D = [-1.5, 1.5] \times [-0.5, 0.5]$$

and cover it by $101 \times 101 = 10201$ equidistant points. All of the resulting 10201 points are mapped through the Henon map H repeatedly. Any point mapped too far away from D is discarded; as a criterion, we choose a cutoff of $[-1000, 1000] \times [-1000, 1000]$. Through 12 iterations of the Henon map H, 8062 points survived. These points, which give an approximate picture of the basin of attraction of the Henon attractor, are shown in the top left picture of Figure 1.

The set F of 8062 surviving floating point numbers are mapped repeatedly by H, and the mapped points at each iteration are shown in Figure 1 for the iterations numbers from 1 to 5, and finally in Figure 2 for the iteration number 12. Each of these regions provide a non-rigorous estimate of an outer enclosure of the attractor.

2.2 Rigorous Enclosure of the Henon Attractor with Interval Methods

In order to gain a rigorous understanding, we now study the same problem using interval methods. Henon himself already showed that the quadrilateral Q with the corner points

$$q_1 = (-1.060, -0.500),\ q_2 = (+1.245, -0.140)$$
$$q_3 = (+1.320, +0.133),\ q_4 = (-1.330, +0.420)$$

forms a trapping region, i.e. $H(Q) \subset Q$. In the case of the Henon map, this is easily shown by hand by merely studying the images of the curves comprising the four edges of Q. When parameterizing these edges linearly, upon application of H they will form quadratic polynomials, for which it is easy to determine the distances to the edges of Q and in fact show that they lie fully inside Q.

For the study with intervals, we utilize neighboring points in the set F as corner points of a total of 5062 interval boxes which together fully cover the quadrilateral Q representing the trapping region. We denote the set comprising the union of these interval boxes by I. We now propagate the set I through the Henon map H repeatedly with the goal of obtaining a sharper and sharper enclosure of the attractor. Figure 3 shows the situation for iterations 0 to 5. We first observe that the set I is not fully mapped inside of itself, thus not allowing to prove computationally that I is indeed a trapping region for the attractor. Furthermore, subsequent iterations do not lead to a significant refinement of the situation due to the problem of interval overestimation; and the covering of the attractor with smallest area is already obtained in the second iterate.

2.3 Rigorous Enclosure of the Henon Attractor with Taylor Model Methods

In order to utilize Q for the study with Taylor models, we first construct a Taylor model that maps the unit domain box $[-1, +1]^2$ onto Q. This can apparently not be achieved by merely using an affine transformation since Q is not a parallelepiped. However, we observe that any polynomial of

Table 1: Performance of the Henon map iterations using Taylor models

Iteration	1	2	3	4	5
Number of TMs	1	1	1	9	32
Remainder error	10^{-14}	10^{-13}	10^{-11}	10^{-10}	10^{-9}

the form

$$P(x,y) = \begin{pmatrix} P_x^H(x,y) \\ P_y^H(x,y) \end{pmatrix} = \begin{pmatrix} a + a_x x + a_y y + a_{xy} xy \\ b + b_x x + b_y y + b_{xy} xy \end{pmatrix}$$

i.e. missing the purely single variable terms x^2 and y^2 maps the four edges of the box $[-1, +1]^2$ into straight lines. In order to determine the coefficients in P, it is thus sufficient to demand

$$P(-1,-1) = q_1,\ P(+1,-1) = q_2,$$
$$P(+1,+1) = q_3,\ P(-1,+1) = q_4,$$

which leads to the system of equations

$$\begin{aligned} +a - a_x - a_y + a_{xy} &= q_{1x} \\ +a + a_x - a_y - a_{xy} &= q_{2x} \\ +a + a_x + a_y + a_{xy} &= q_{3x} \\ +a - a_x + a_y - a_{xy} &= q_{4x}, \end{aligned}$$

and likewise for the b coefficients. Because of the special structure of the system with coefficients of ± 1, the system is very easy to solve by just suitable addition and subtraction, and we obtain

$$\begin{aligned} a &= \frac{+q_{1x} + q_{2x} + q_{3x} + q_{4x}}{4} \\ a_x &= \frac{-q_{1x} + q_{2x} + q_{3x} - q_{4x}}{4} \\ a_y &= \frac{-q_{1x} - q_{2x} + q_{3x} + q_{4x}}{4} \\ a_{xy} &= \frac{+q_{1x} - q_{2x} + q_{3x} - q_{4x}}{4}. \end{aligned}$$

In passing we note that the method readily generalizes to higher dimensions.

The resulting quadratic polynomial P_0 thus satisfies $P_0([-1, +1]^2) = Q$ and will serve as the starting Taylor model (with vanishing remainder bound). We now iterate P_0 through H repeatedly in Taylor model arithmetic, populating the remainder bound. Furthermore, if in any one step the remainder bound exceeds a certain pre-specified threshold, the step is discarded and the domain of the polynomial split along the dominating direction. This will lead to a decrease of the remainder bound, and overall a dynamic decomposition of the resulting structure into several Taylor models.

Figures 4 and 5 show the resulting structures for iterations 0 to 5, at which point the attractor is enclosed with a sharpness given by printer resolution. The actual sharpness of the remainder errors at the fifth iterate is 10^{-9} on average, so the method has an overestimation that is many orders of magnitude lower than that of the interval method.

For a more quantitative analysis, Table 1 lists the number of Taylor models produced by dynamic domain decomposition as well as the average width of the remainder bounds of these Taylor models as a function of the number of iterations. The Taylor model computation order is 33.

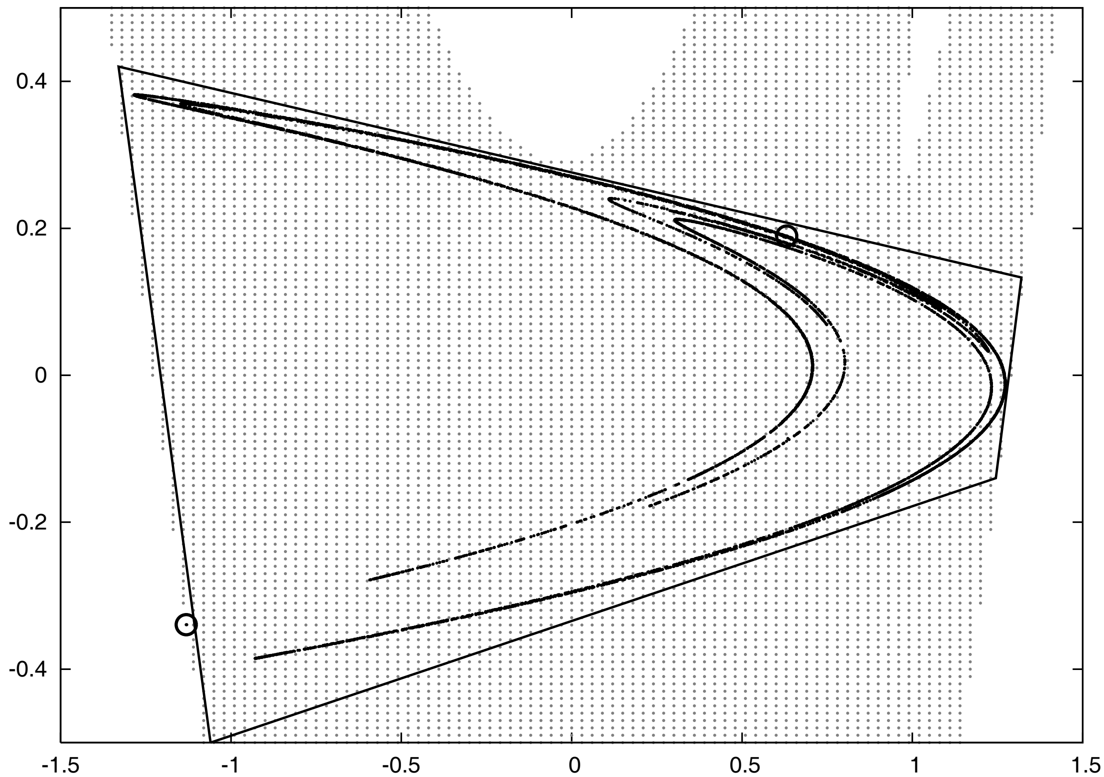

Figure 2: Non-rigorous estimate of an outer enclosure of the attractor of the Henon map H. The two marked points are fixed points, and Henon's trapping region Q is shown.

3. BOUNDERS AND DOMAIN PRUNING FOR GLOBAL OPTIMIZATION

In the following we utilize Taylor model methods for rigorous global optimization based on the conventional branch and bound methods. Specifically, the method discards and prunes regions based on rigorous upper bounds for the minimizer. Discarding and pruning is based on the LDB and QFB algorithms described below.

3.1 The Linear Dominated Bounder LDB

The linear dominated bounder (LDB) introduced in [4] is based on the fact that for Taylor models with sufficiently small remainder bound, the linear part of the Taylor model dominates the behavior, and this is also the case for range bounding. The linear dominated bounder utilizes the linear part as a guideline for iterative domain reduction to bound Taylor models.

LDB Algorithm

Wlog, find the lower bound of minimum of a Taylor model $P + I$ in D.

(1) Re-expand P at the mid-point c of D, call the resulting polynomial P_m and the centered domain D_1.

(2) Turn the linear coefficients L_i's of P_m all positive by suitably flipping coordinate directions, call the resulting polynomial P_+.

(3) Compute the bound of the linear (I_1) and nonlinear (I_h) parts of P_+ in D_n. The minimum is bounded by $[M, M_{in}] := \underline{I}_1 + I_h$. If applicable, lower M_{in} by the left end value and the mid-point value.

(a) If $d = \text{width}([M, M_{in}]) > \varepsilon$, set D_{n+1} such that $\forall i$, if $L_i > 0$ and $\text{width}(D_{n+1,i}) > d/L_i$, then $\overline{D}_{n+1,i} := \underline{D}_{n,i} + d/L_i$. Re-expand P_+ at the mid-point c of D_{n+1}. Prepare the new coefficients L_i's. Go to 3.

(b) Else, M is the lower bound of minimum.

Any errors associated with re-expansion and estimating point values are included in the remainder error bound interval. If f is monotonic, the exact bound is often obtained with high accuracy. If only a threshold cutoff test is needed, the resulting domain reduction or elimination is often superior. The reduction of the domain of interest works multi-dimensionally and automatically, and the observed domain reduction rate is thus often fast. Even when there is no linear part in the original Taylor model, by shifting the expansion point, normally a linear part is introduced.

3.2 The Quadratic Fast Bounder QFB

The natural next idea of Taylor model bounding is to utilize the quadratic part of P, and a preliminary scheme of a quadratic dominated bounder (QDB) is discussed in [4]. For the task of global optimization in practice, an efficient bounding of the quadratic part in the vicinity of interior minimizers is important. Around an isolated interior minimizer, the Hessian of a function f is positive definite, so the purely quadratic part of a Taylor model $P + I$ which locally represents f, has a positive definite Hessian matrix H. The actual definiteness can be tested in a verified way using the common LDL or extended Cholesky decomposition. The quadratic fast bounder (QFB) provides a lower bound

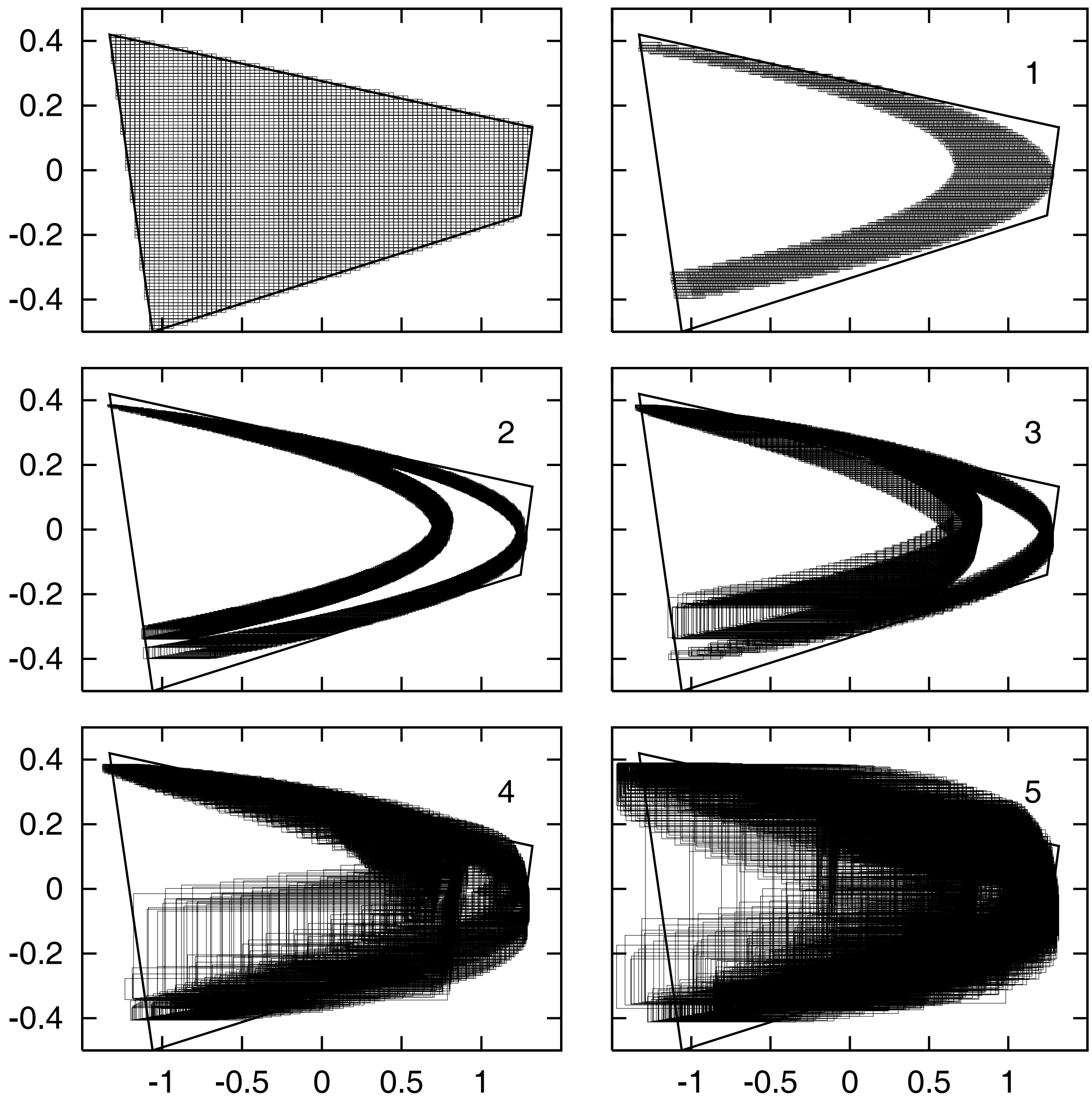

Figure 3: Mapping of a trapping region of the Henon map H using interval enclosures.

of a Taylor model cheaply when the purely quadratic part is positive definite. It is based on the following observation.

Let $P + I$ be a given Taylor model in D, and let H be the Hessian matrix of P. We decompose the polynomial P into two parts via

$$P + I = (P - Q) + I + Q.$$

Then a lower bound for $P + I$ is obtained as

$$l(P + I) = l(P - Q) + l(Q) + l(I).$$

For QFB, we choose

$$Q = Q_{x_0} = \frac{1}{2}(x - x_0)^t H(x - x_0)$$

with any $x_0 \in D$. If H is positive semidefinite, $l(Q_{x_0}) = 0$, and the value 0 is attained. The remaining $P - Q_{x_0}$ does not contain pure quadratic terms anymore, but consists of linear as well third and higher order terms $P_{>2}$. If x_0 is chosen to be the minimizer of the quadratic part P_2 of P in D, then x_0 is also a minimizer of the remaining linear part (a consequence of the Kuhn-Tucker conditions), and so the lower bound estimate is optimally sharp. Thus by choosing x_0 close enough to the minimizer of P_2 in D, a contribution of $P_2 - Q_{x_0}$ to the lower bound can be very small. For a given P_2 in D, x_0 can be determined inexpensively by an iterative scheme to search a series of $x_0^{(i)}$ in the direction of $-\nabla P_2$ while limiting $x_0^{(i)}$ to stay inside D.

4. THE VERIFIED GLOBAL OPTIMIZER COSY-GO

For the example problems of verified global optimization in the next subsections, we apply three branch-and-bound kind methods available in the code COSY-GO[7]. The first

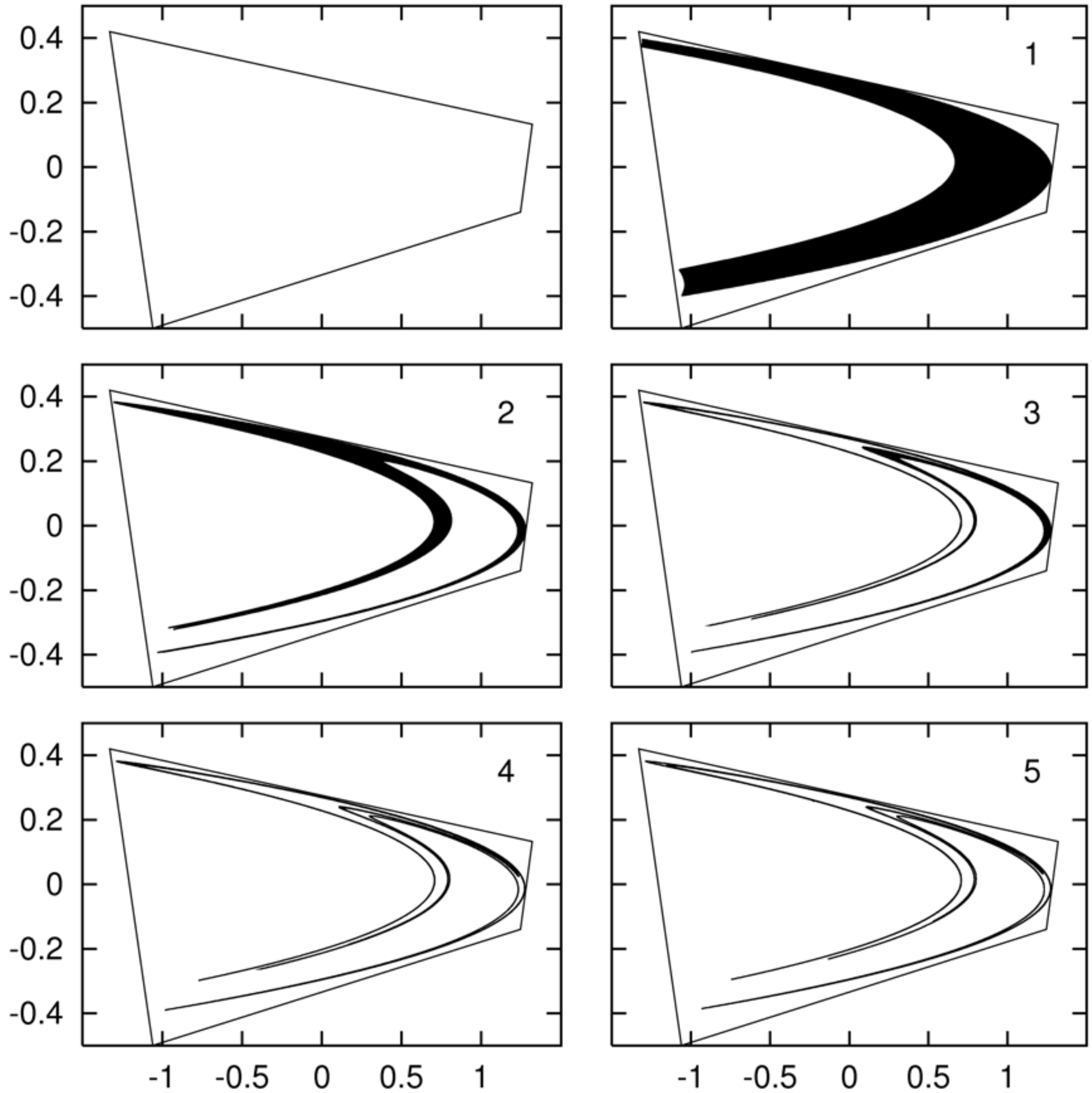

Figure 4: Mapping of a trapping region of the Henon map H using Taylor model enclosures.

one is the Taylor model-based optimizer utilizing the LDB and QFB algorithms (LDB/QFB). We compare the performance with two other optimizers; one based on mere interval bounding (IN) and one based on bounding with centered form (CF). The sub domain box list management is performed in the same way for all three optimizers. At each sub domain box step, the following tasks are performed.

- A function bound is estimated using the tools described below; if the lower bound is above the cutoff value, the box is eliminated; if not, the box is bisected.
- Bounding schemes are applied in a hierarchical manner. The mere interval bounding is estimated for all optimizers. If the interval bound fails to eliminate the box, the centered form bounding is performed for the CF optimizer. Likewise, for the LDB/QFB optimizer, if the interval bound fails, the naive Taylor model bound[6] based on interval evaluation of the Taylor polynomial is determined, and only when it fails, the LDB bound is determined. If it also fails and the quadratic part of the local Taylor model of the function is positive definite, the QFB bounding is performed.
- When the LDB bound fails to eliminate the box, however, often the box can be reduced before bisection.
- The cutoff value is updated. The mid-point value estimate is conducted for all optimizers.
- For the LDB/QFB optimizer, the linear and quadratic parts of the local Taylor model are utilized to guess a candidate for the global minimizer to obtain a better cutoff value estimate.

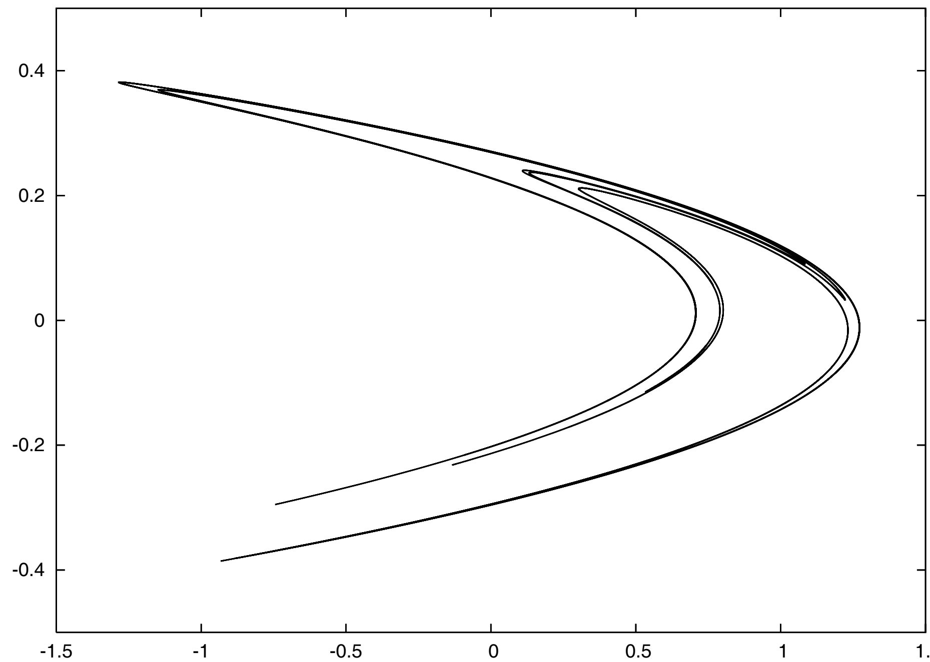

Figure 5: The fifth iterate of mapping of a trapping region of the Henon map H using Taylor model enclosures.

Table 2: Performance of various optimizers.

	$f(x) = 1 + x^5 - x^4$ in $[0, 1]$			Beale function f_B in $[-4.5, 4.5]^2$		
	IN	CF	LDB/QFB	IN	CF	LDB/QFB
Total box processing steps	12471	145	17	3407	3285	353
Max number of active boxes	4044	11	3	236	234	52
Retained small boxes ($< 10^{-6}$)	2591	4	1	25	25	3
LDB domain reduction steps	–	–	8	–	–	108

4.1 A One Dimensional Polynomial

The first example problem is to search the minimum of the polynomial

$$f(x) = 1 + x^5 - x^4$$

in $[0, 1]$, suggested by R. Moore[8]. The function has a shallow minimum at $x = 0.8$, and looks rather innocent as shown in Figure 6, but the dependency problem and the high order of the polynomial prevents the mere interval bounding method from being successful.

Table 2 summarizes the performance of the three optimizers. The Taylor model LDB/QDB optimizer eliminates all sub domain boxes but the one containing the minimum in 17 steps, among which 8 steps are size reductions by LDB. There are at most 3 active boxes kept in the whole optimization process. On the other hand, the interval optimizer requires a total of 12471 steps, and retains 2591 small boxes ($< 10^{-6}$).

The centered form optimizer performs better than the interval optimizer, but cannot reach the performance of the LDB/QFB optimizer. The sub domain boxes active in each step are shown in Figure 6, and an example of LDB domain reduction can be seen in the processing of the parent box in step 5 to yield the bisected boxes appearing in steps 6 and 7. As seen later in Figure 7, the LDB domain reduction works favorably also in multidimensional cases.

4.2 The Beale Function

The next example is the Beale function[9]

$$f_B(x_1, x_2) = (1.5 - x_1(1 - x_2))^2 + \left(2.25 - x_1(1 - x_2^2)\right)^2 + \left(2.625 - x_1(1 - x_2^3)\right)^2 .$$

The problem is to find the minimum in the initial domain $[-4.5, 4.5] \times [-4.5, 4.5]$ with verification. The function has little dependency and the minimum 0 occurs at $(3, 0.5)$, however the very shallow behavior of the function makes a verified global optimization task difficult.

The performance of the optimizers is summarized in Figure 7 and Table 2. Square expressions in f_B are not utilized to simplify the arithmetic. We observe no advantage in the centered form optimizer compared to the interval optimizer. On the other hand, the LDB/QFB optimizer significantly outperformed both others because of more efficient box rejection and LDB domain size reduction.

5. ACKNOWLEDGMENTS

This work was supported by the US Department of Energy, the National Science Foundation, the German National Merit Foundation, and an Alfred P. Sloan Fellowship. We thank Ray Moore for many fruitful discussions.

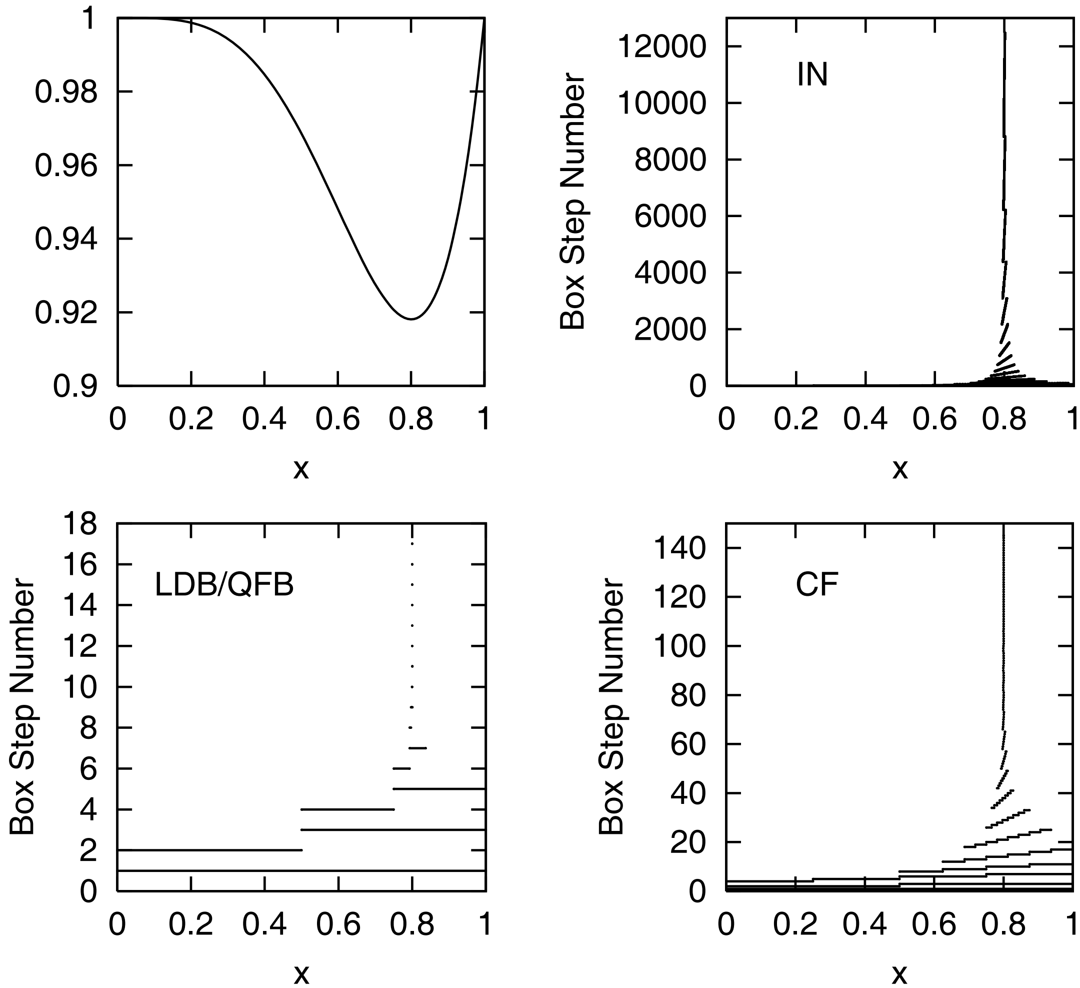

Figure 6: Global optimization of $f(x) = 1 + x^5 - x^4$ in $[0, 1]$ (top left). Sub domain boxes for minimum search are shown at each step: (top right) the interval, (bottom right) the centered form, and (bottom left) the LDB/QFB optimizers.

6. REFERENCES

[1] M. Berz and J. Hoefkens. Verified high-order inversion of functional dependencies and superconvergent interval Newton methods. *Reliable Computing*, 7(5):379–398, 2001.

[2] M. Berz and K. Makino. Verified integration of ODEs and flows using differential algebraic methods on high-order Taylor models. *Reliable Computing*, 4(4):361–369, 1998.

[3] M. Berz and K. Makino. COSY INFINITY Version 9.0 programmer's manual. Technical Report MSUHEP-060803, Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, 2006. See also http://cosyinfinity.org.

[4] K. Makino. *Rigorous Analysis of Nonlinear Motion in Particle Accelerators*. PhD thesis, Michigan State University, East Lansing, Michigan, USA, 1998. Also MSUCL-1093.

[5] K. Makino and M. Berz. Suppression of the wrapping effect by Taylor model-based validated integrators. Technical Report MSUHEP-040910, available at http://bt.pa.msu.edu/pub, Michigan State University, 2003.

[6] K. Makino and M. Berz. Taylor models and other validated functional inclusion methods. *International Journal of Pure and Applied Mathematics*, 6,3:239–316, 2003.

[7] K. Makino and M. Berz. COSY INFINITY version 9. *Nuclear Instruments and Methods*, 558:346–350, 2005.

[8] R. E. Moore. Private communication.

[9] More, Garbow, and Hillstrom. Testing unconstrained optimization software. *ACM Trans. Math. Software*, 7:17–41, 1981.

[10] N. Revol, K. Makino, and M. Berz. Taylor models and floating-point arithmetic: Proof that arithmetic operations are validated in COSY. *Journal of Logic and Algebraic Programming*, 64/1:135–154, 2004.

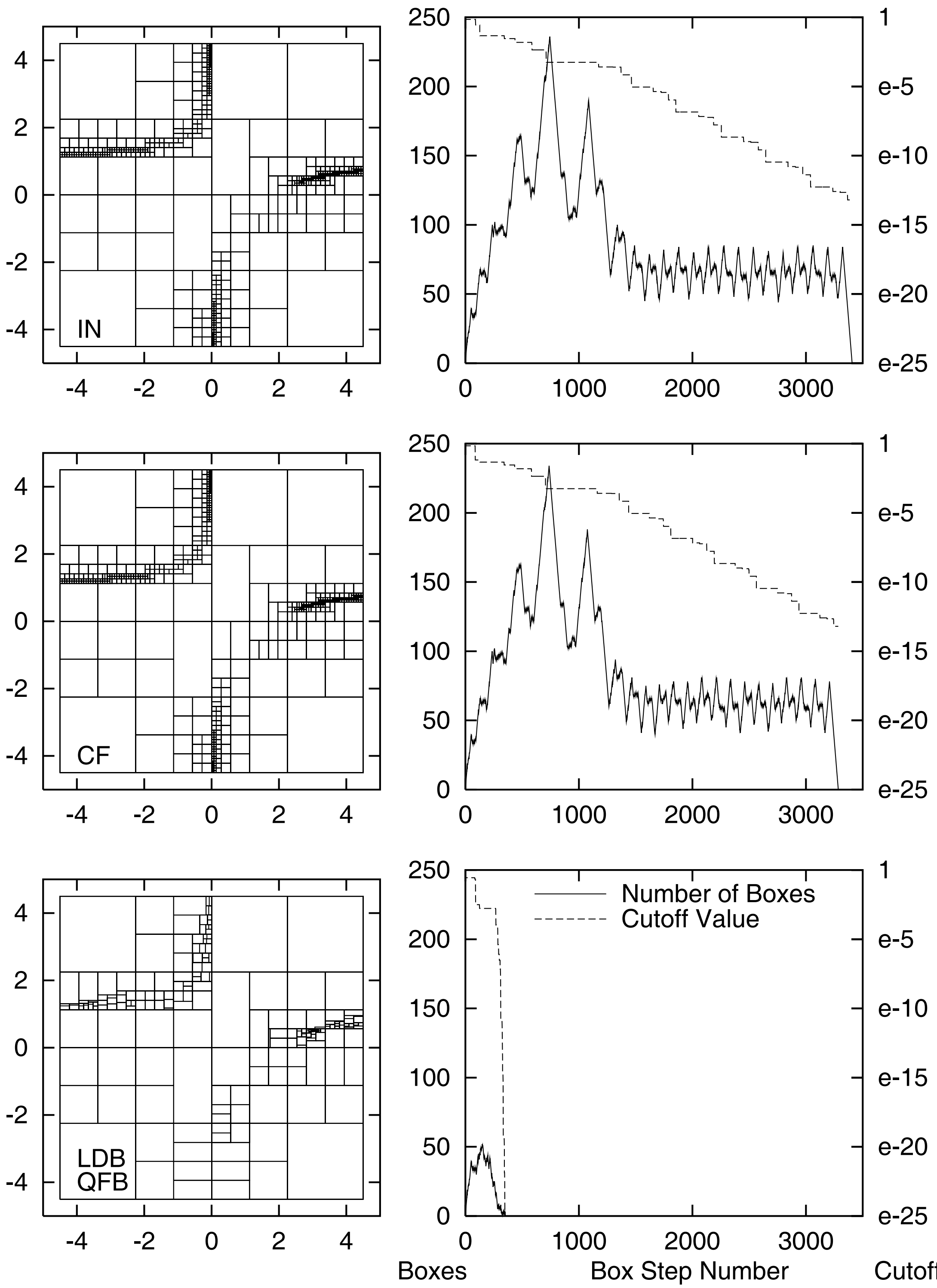

Figure 7: Minimum search for the Beale function in $[-4.5, 4.5]^2$ by the interval, the centered form, and the LDB/QFB optimizers. Left: sub domain boxes. Right: number of active boxes and cutoff value.