

Taylor Model Range Bounding Schemes

I. Kyoko Makino

II. Martin Berz

Department of Physics and Astronomy
Michigan State University

- Introduction
- LDB (Linear Dominated Bounder)
- QDB (Quadratic Dominated Bounder)
- QFB (Quadratic Fast Bounder)
- Validated Global Optimization

Introduction

Taylor model (TM) methods were originally developed for a practical problem from nonlinear dynamics, range bounding of normal form defect functions.

- Functions consist of code lists of 10^4 to 10^5 terms
- Have about the worst imaginable cancellation problem
- Are obtained via validated integration of large initial condition boxes.

Originally nearly universally considered intractable by the community. But ... a small challenge goes a long way towards generating new ideas!

Idea: represent all functional dependencies as a pair of a polynomial P and a remainder bound I , introduce arithmetic, and a new ODE solver. Obtain the following properties:

- The ability to provide enclosures of any function given by a finite computer code list by a Taylor polynomial and a remainder bound with a sharpness that scales with order $(n + 1)$ of the width of the domain.
- The ability to alleviate the dependency problem in the calculation.
- The ability to scale favorably to higher dimensional problems.

One Dimensional TM Range Bounders

(Work with Youn-Kyung Kim at Michigan State Univ.)

It is relatively easy to produce efficient range bounders of up to sixth order

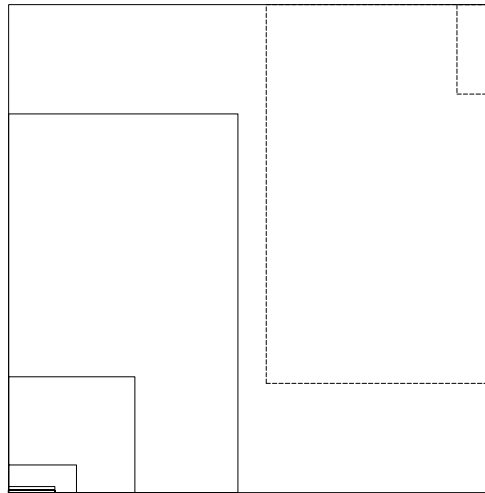
- There are well-known formulas for zeros of polynomials up to order 4
- Apply these to the derivatives and find all real roots
- Yields all critical points of polynomials up to order 5
- Evaluating polynomial at these and boundary points, and take min, max

Care has to be taken about the following aspects:

- Obviously, Evaluate formulas by interval arithmetic
 - Branches in the code because of different sub-cases:
 - follow each one separately, or
 - slightly perturb the original polynomial so that branches disappear
- $$P^*(x) = P(x) + \sum_{i=1}^5 \varepsilon_i x^i, \text{ then } B(P) \subset B(P^*) - B\left(\sum_{i=1}^5 \varepsilon_i x^i\right)$$
- Only interested in real roots: re-write expressions to avoid complex roots
 - Cleverly write formulas to minimize width of enclosures of critical points (cancellation problem)

The Linear Dominated Bounder (LDB)

- The linear part of TM polynomial is the leading part, also for range bounding.
- The idea is easily extended to the multi-dimensional case.
- Use the linear part as a guideline for domain splitting and elimination.
- The reduction of the size of interested box works multi-dimensionally and automatically. Thus, the reduction rate is fast.
- Even there is no linear part in the original TM, by shifting the expansion point, normally the linear part is introduced.
- Exact bound (with rounding) is obtained if monotonic.



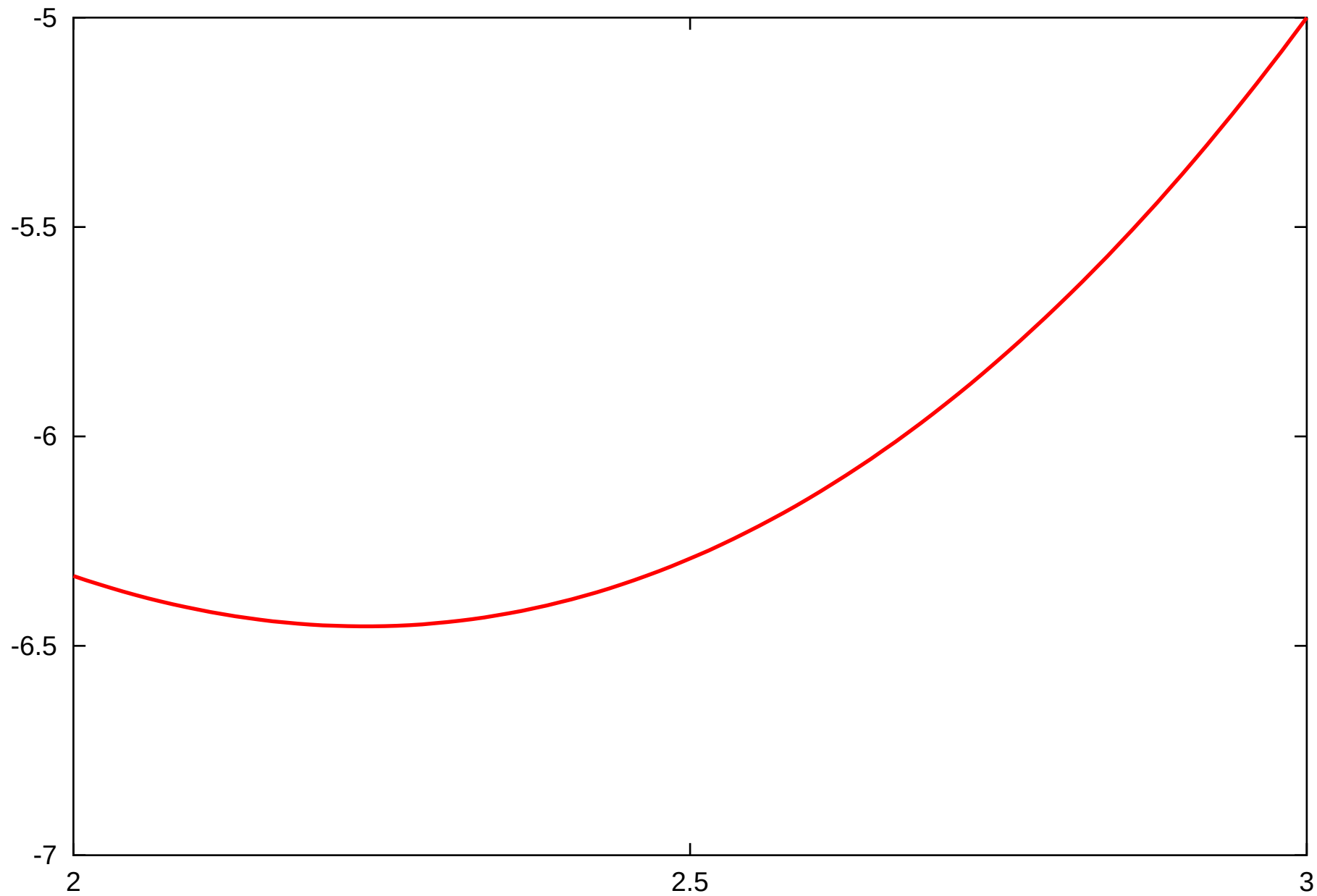
LDB Algorithm

Wlog, find the lower bound of minimum of a polynomial P in B .

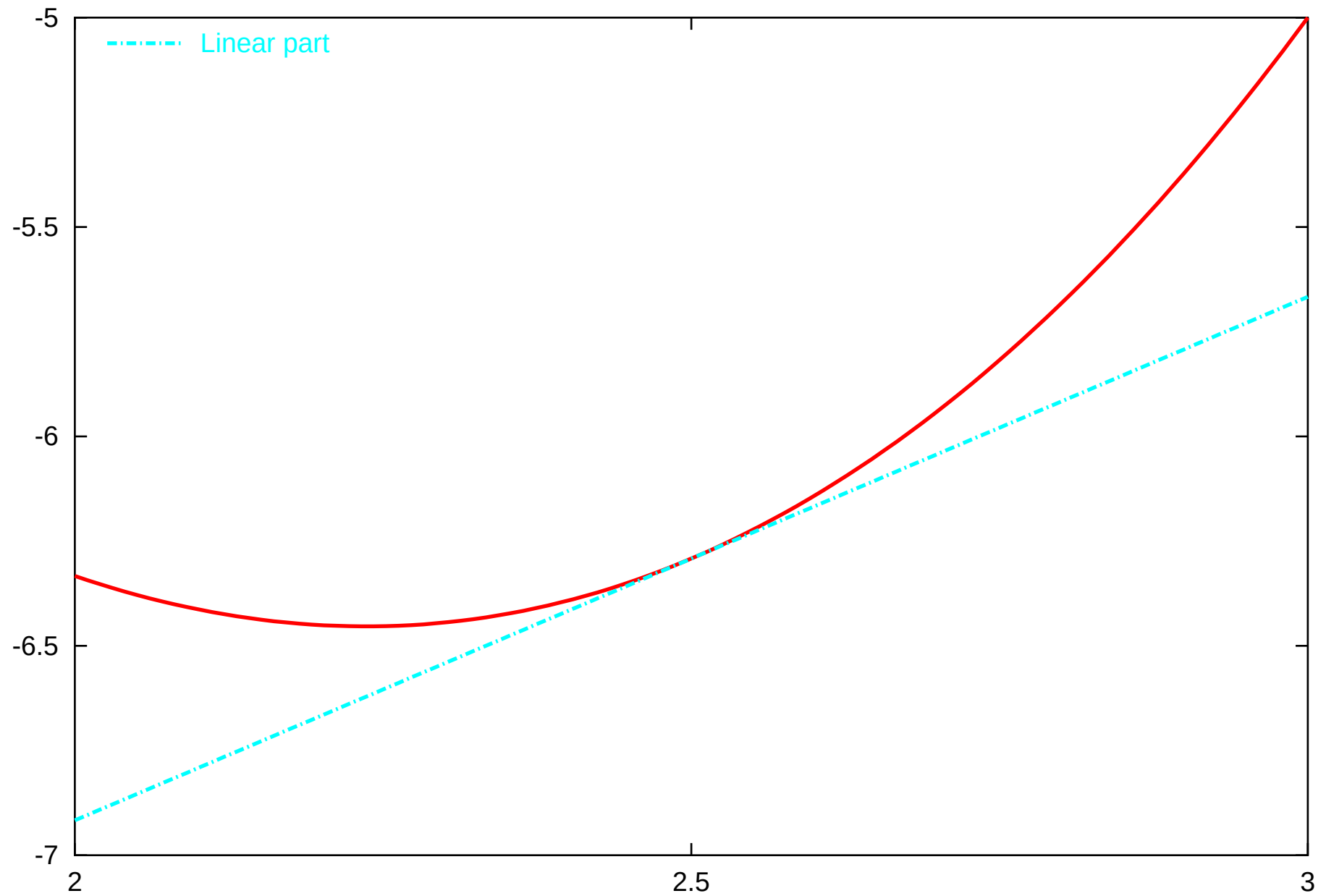
1. Re-expand P at the mid-point m of B to P_m . Center the domain as B_m .
 2. Turn the linear coefficients c_i 's of P_m all positive by a transformation D , with $D_{ii} = \text{sign}(c_i)$, $D_{ij} = 0$ for $i \neq j$.
The polynomial is P_+ in the domain $B_w = B_m$.
 3. Compute the bound of the linear (I_1) and nonlinear (I_h) parts of the polynomial P_+ in B_w . 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.
 4. The refinement iteration
 - (a) If $w = \text{width}([M, M_{in}]) > \varepsilon$, set $B_w : \forall i$, if $c_i > 0$ and $\text{width}(B_{wi}) > w/c_i$, then
 - $\overline{B}_{wi} := \underline{B}_{wi} + w/c_i$.
 - Re-expand P_+ at the mid-point of B_w . c_i 's are the new coefficients.
 - Go to **3**.
 - (b) Else, M is the lower bound of minimum.
- If only a cutoff test is needed, the task is performed more efficiently.

$P(x)=1-5x+x^3/3$ in $[2,3]$

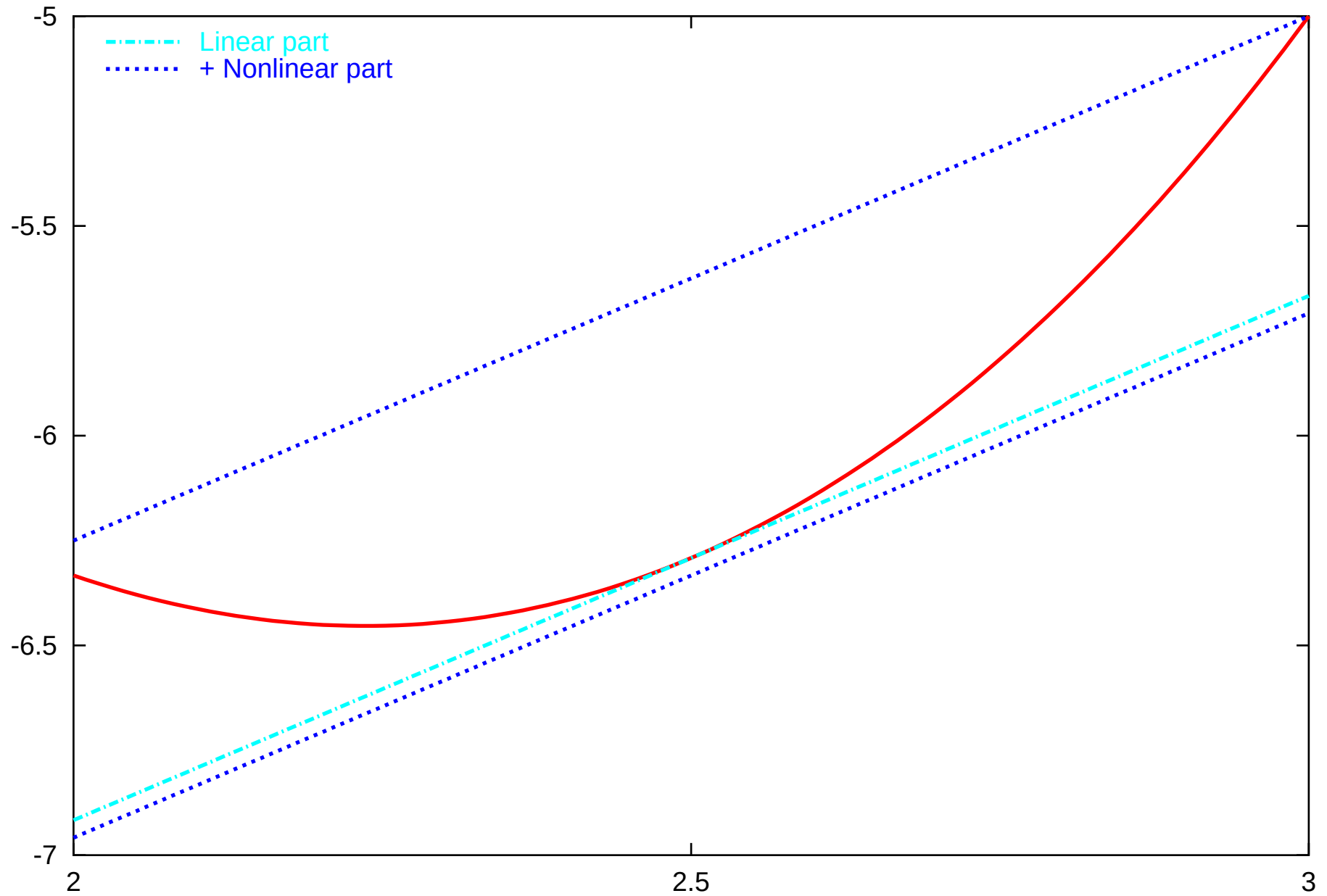
from R. Moore, SIAM 1979



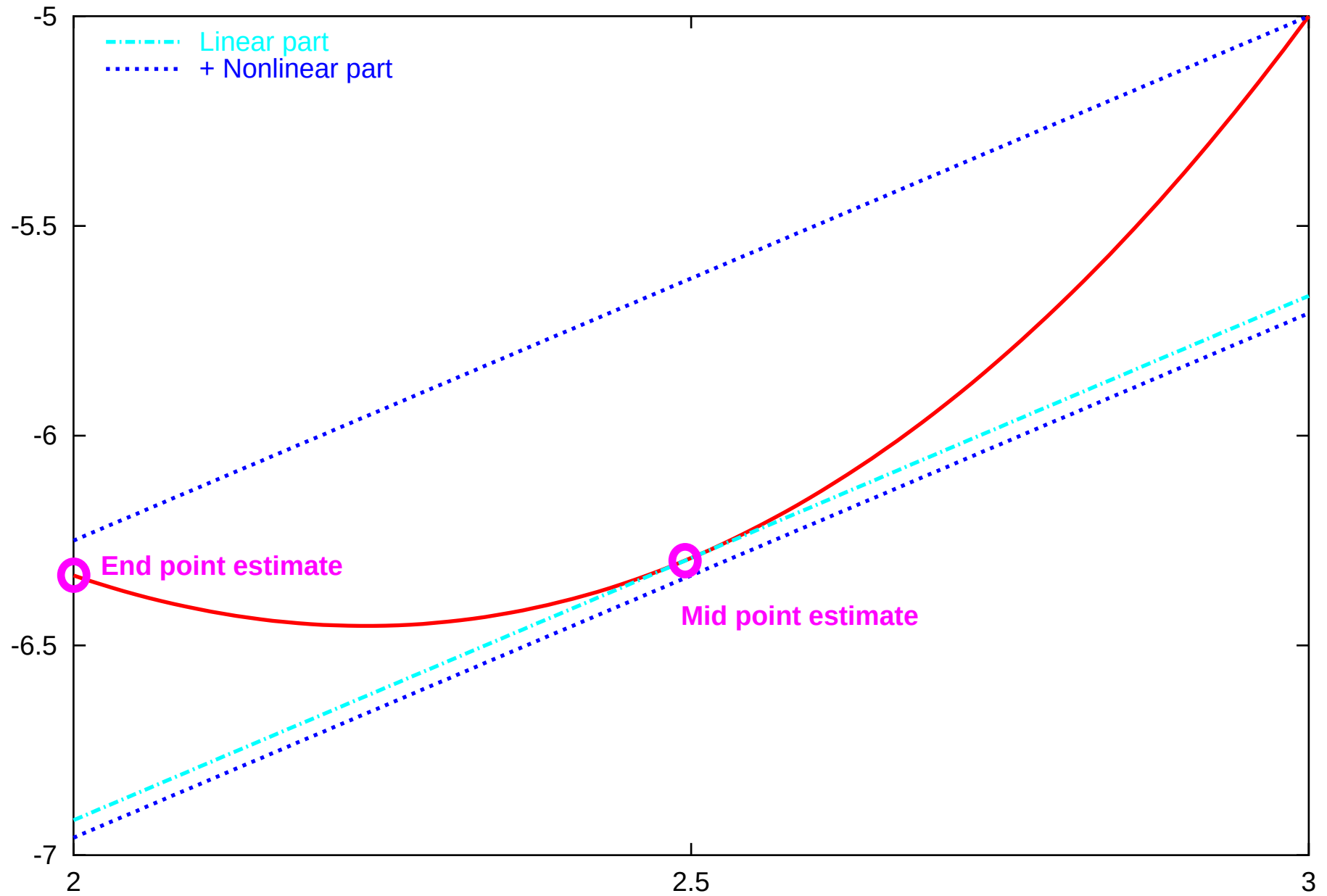
LDBL, $P(x)=1-5x+x^3/3$. Step 0



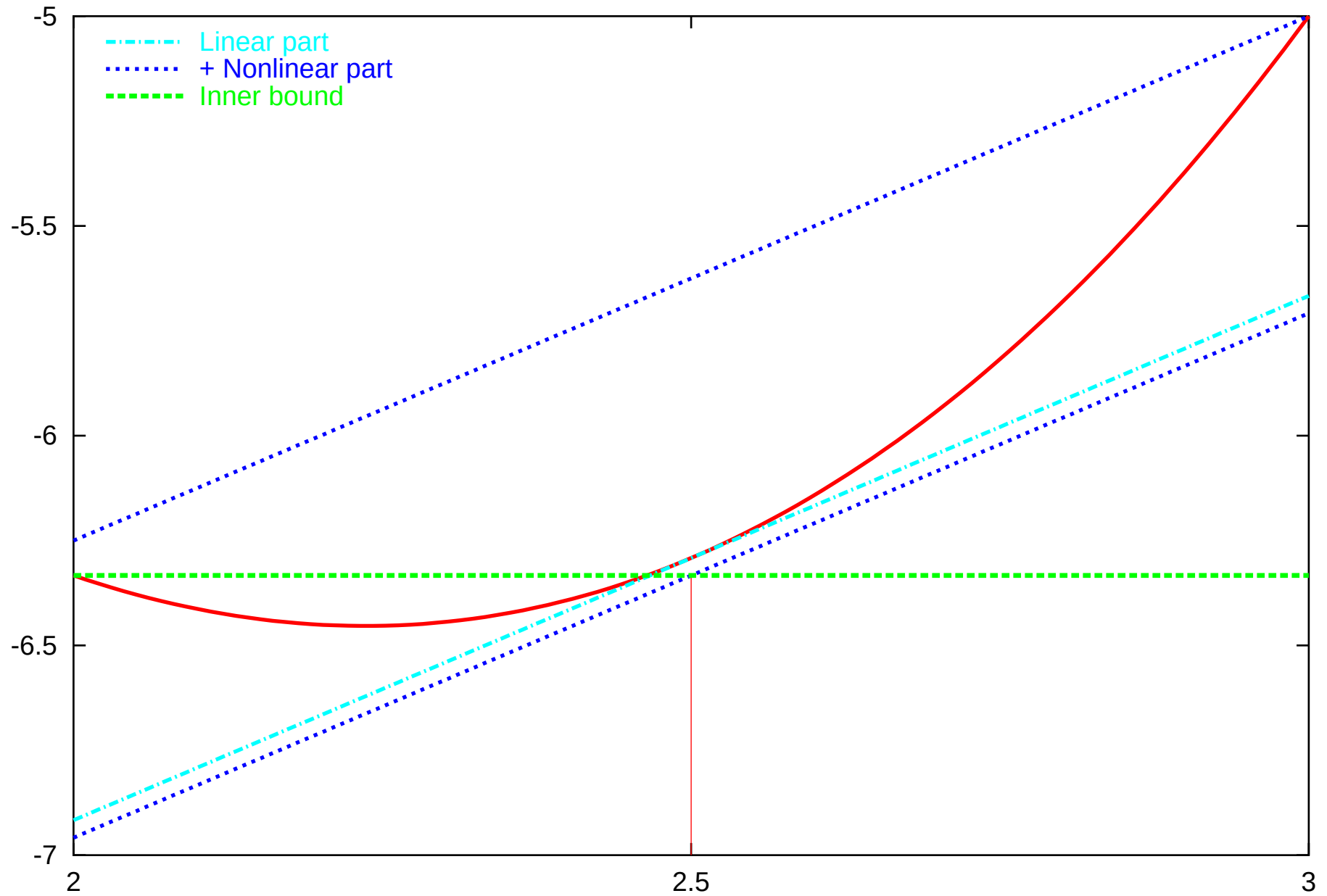
LDBL, $P(x)=1-5x+x^3/3$. Step 0



LDBL, $P(x)=1-5x+x^3/3$. Step 0



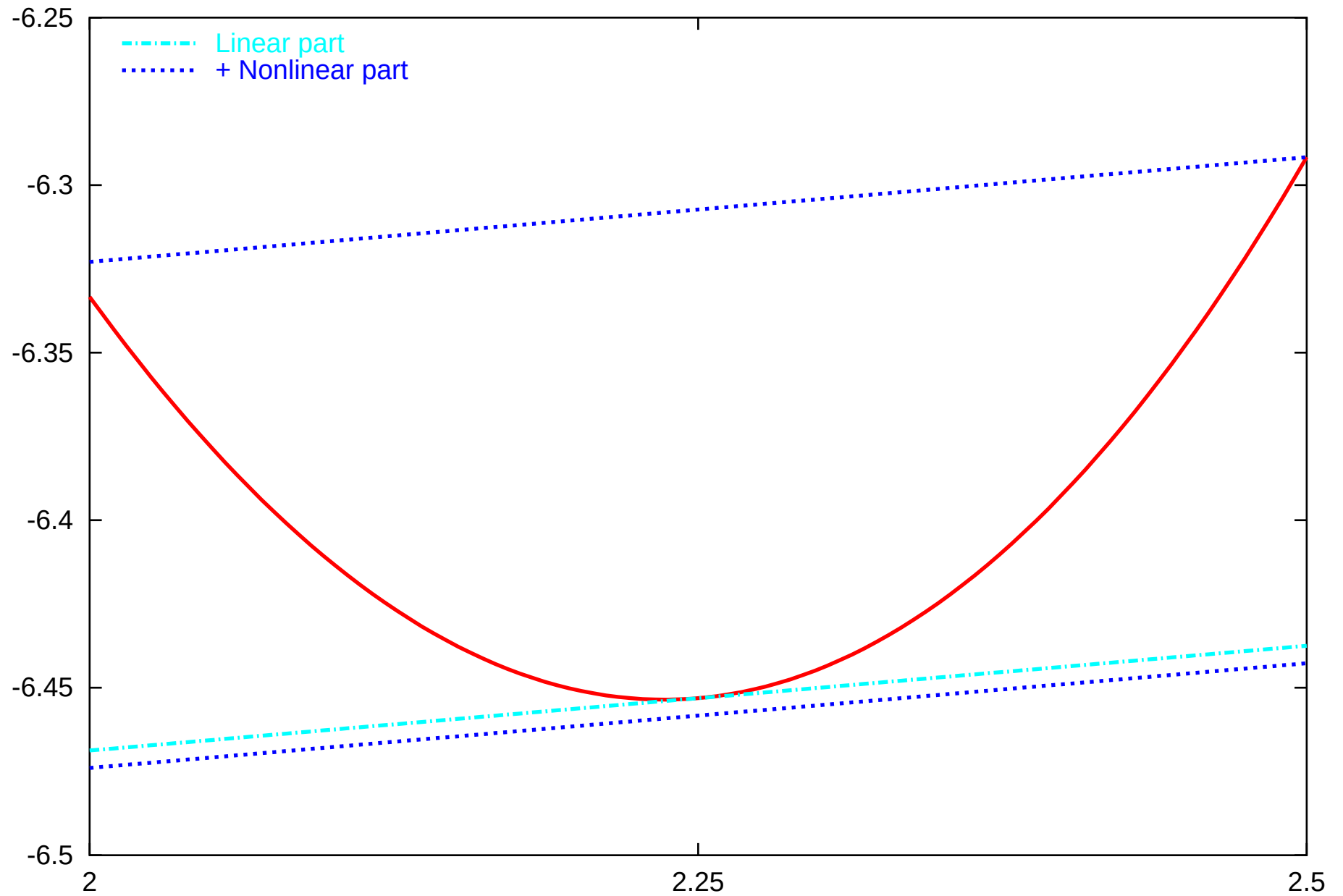
LDBL, $P(x)=1-5x+x^3/3$. Step 0



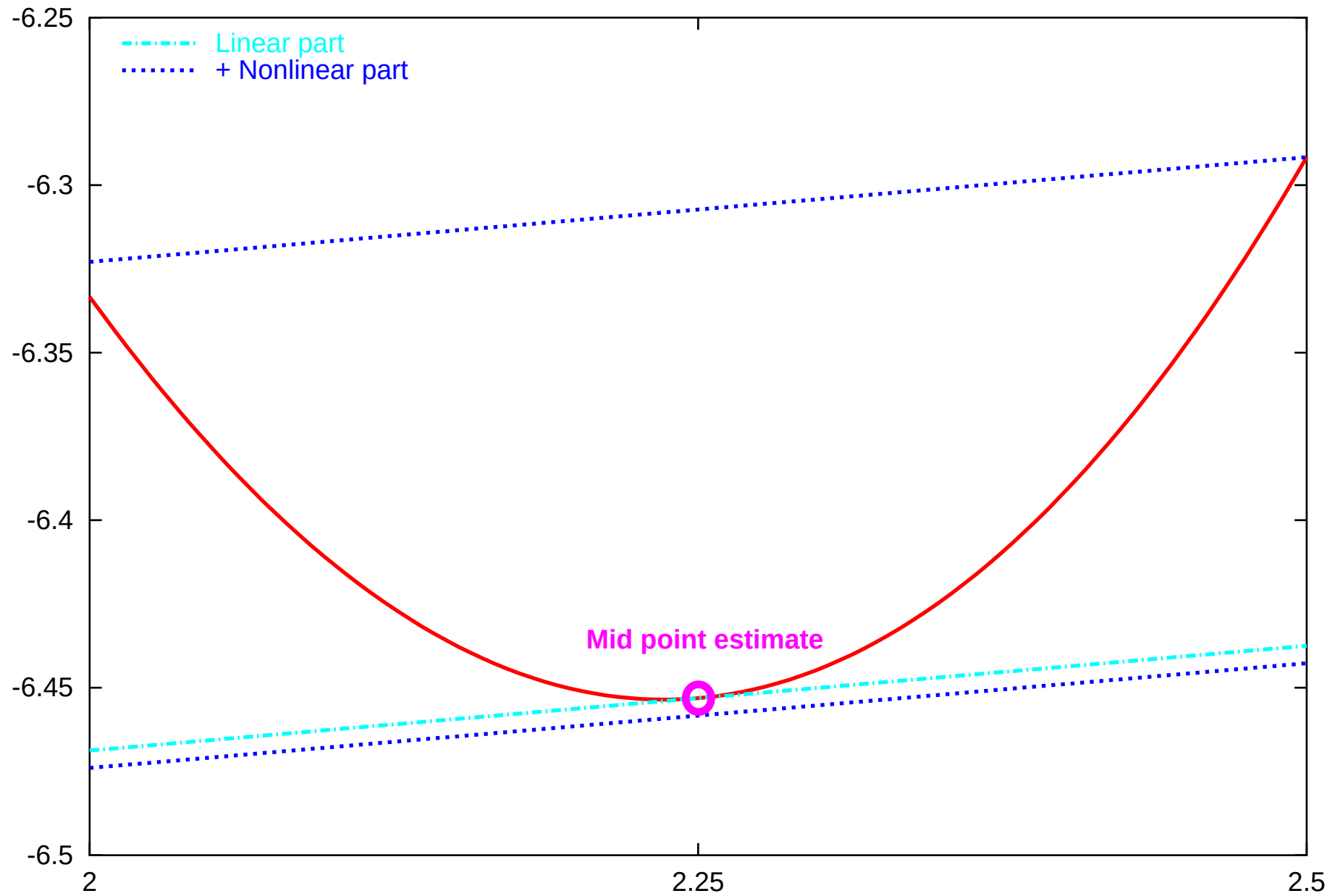
LDBL, $P(x)=1-5x+x^3/3$ in $[2,2.5]$. Step 1



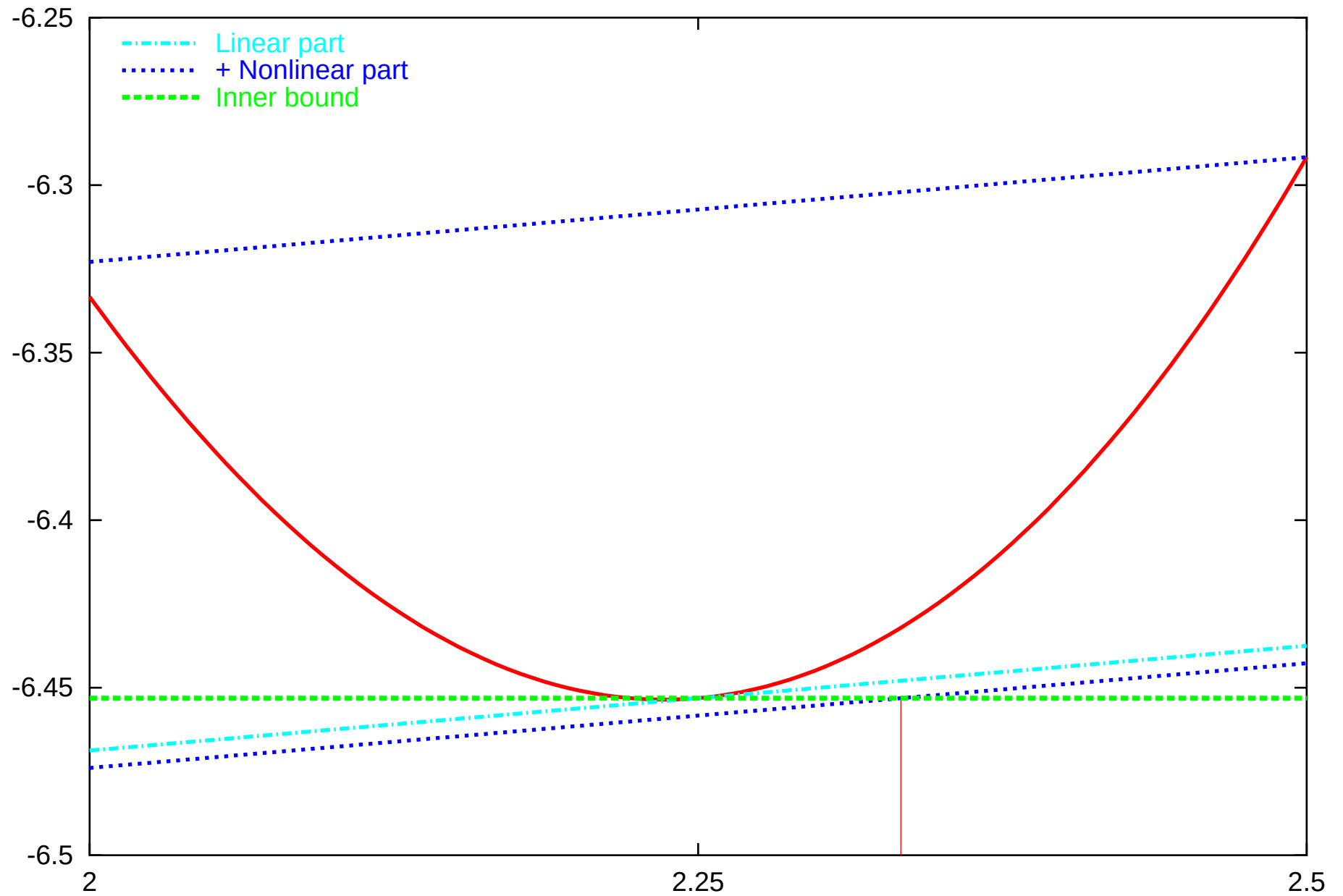
LDBL, $P(x)=1-5x+x^3/3$ in $[2,2.5]$. Step 1



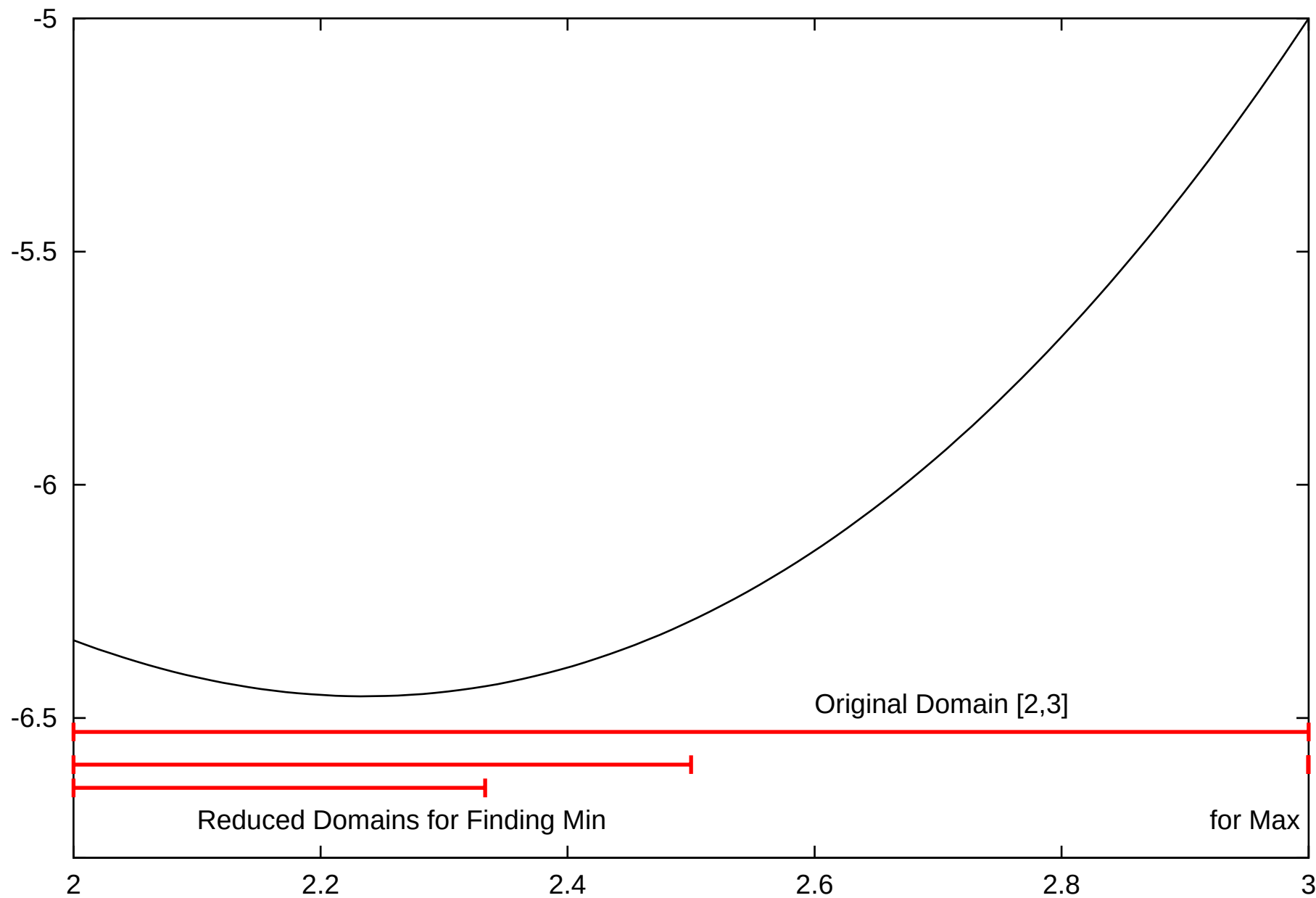
LDBL, $P(x)=1-5x+x^3/3$ in $[2,2.5]$. Step 1



LDBL, $P(x)=1-5x+x^3/3$ in $[2,2.5]$. Step 1



Finding Min and Max using LDBL, $P(x)=1-5*x+x^3/3$ in $[2,3]$



The QDB (Quadratic Dominated Bounder) Algorithm

1. Let u be an external cutoff. Initialize $u = \min(u, Q(\text{center}))$. Initialize list with all 3^n surfaces for study.
2. If no boxes are remaining, terminate. Otherwise select one surface S of highest dimension.
3. On S , apply LDB. If a complete rejection is possible, strike S and all its surfaces from the list and proceed to step 2. If a partial rejection is possible, strike the respective surfaces of S from the list and proceed to step 2.
4. Determine the definiteness of the Hessian of Q which is restricted to S
(Use LDL decomposition.)
5. If the Hessian is not p.d. strike S from the list and proceed to step 2.
6. If the Hessian is p.d., determine the corresponding critical point c .
7. If c is fully inside S , strike S and all surfaces of S from the list, update $u = \min(u, Q(c))$, and proceed to step 2
8. If c is not inside S , strike S and all of its surfaces that are not visible from c and proceed to step 2

The QDB Algorithm - Properties

The QDB algorithm has the following properties.

1. It has the third order approximation property.
2. The effort of finding the minimum requires the study of at most 3^n surfaces.
3. In the p.d. case, the computational effort requires the study of at most 2^n surfaces.
4. For the surfaces studies, parts of the original LDL decomposition can be re-used.
5. Because of extensive box striking, in practice, the numbers of boxes to study is usually much much less.

The QDB Algorithm - Properties

The QDB algorithm has the following properties.

1. It has the third order approximation property.
2. The effort of finding the minimum requires the study of at most 3^n surfaces.
3. In the p.d. case, the computational effort requires the study of at most 2^n surfaces.
4. For the surfaces studies, parts of the original LDL decomposition can be re-used.
5. Because of extensive box striking, in practice, the numbers of boxes to study is usually much much less.

But still, it is desirable to have **something faster**.

The QFB (Quadratic Fast Bounder) Algorithm

Let $P + I$ be a given Taylor model. Idea: Decompose into two parts

$$P + I = (P - Q) + I + Q \text{ and observe}$$

$$l(P + I) \geq l(P - Q) + l(Q) + l(I)$$

Choose Q quadratic such that

1. Q can be easily and sharply bounded from below.
2. $P - Q$ is sufficiently simplified to allow bounding above given cutoff.
3. $l(P + I) \approx l(P - Q) + l(Q) + l(I)$

First possibility: Let H be p.d. part of Hessian of P , set

$$Q = \frac{1}{2}x^t \cdot H \cdot x$$

Then $l(Q) = 0$. Removes **all second order parts** of P (!) Better yet:

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

Allows to manipulate **linear part**. Works for ANY x_0 in domain. Still $l(Q_{x_0}) = 0$.

Which choices for x_0 are good?

The QFB Algorithm - Properties

Most critical case: near local minimizer, so H is the entire purely quadratic part of P .

Theorem: If x_0 is the (unique) minimizer of quadratic part of P on the domain of $P + I$, then x_0 is also the minimizer of linear part of $(P - Q_{x_0})$. Furthermore, the lower bound of $(P - Q_{x_0})$, when evaluated with plain interval evaluation, is accurate to order 3 of the original domain box.

Proof: First part follows readily from Kuhn-Tucker conditions. If x_0 inside, linear part vanishes completely. Otherwise, if i -th component of x_0 is wlog at left end, i -th partial there must be non-negative, so yields smallest contribution obtained at x_0 .

Consequence: If x_0 is the minimizer of quadratic part P , we have

$$l(P + I) = l(P - Q_{x_0}) + l(Q_{x_0}) + l(I)$$

Remark: The closer x_0 is to the minimizer, the closer there is order 3 cutoff.

The QFB Algorithm - Practical Use

Algorithm: (Third Order Cutoff Test). Let $x^{(n)}$ be a sequence of points that converges to the minimum x_0 of the convex quadratic part P_2 . In step n , determine a bound of $(P - Q_{x_n})$ by interval evaluation, and assess whether the bound exceeds the cutoff threshold. If it does, reject the box and terminate; if it does not, proceed to the next point $x^{(n+1)}$.

QMLoc: Tool to generate cheap and efficient sequence $x^{(n)}$. Determine "feasible descent direction"

$$g_i^{(n)} = \begin{cases} -\frac{\partial Q}{\partial x_i} & \text{if } x_i^{(n)} \text{ inside} \\ \min \left(-\frac{\partial Q}{\partial x_i}, 0 \right) & \text{if } x_i^{(n)} \text{ on right} \\ \max \left(-\frac{\partial Q}{\partial x_i}, 0 \right) & \text{if } x_i^{(n)} \text{ on left} \end{cases}$$

Now move in direction of $g^{(n)}$ until we hit box or quadratic minimum along line. Very fast, can cover large ground per step, can change set of active constraints very quickly.

Result: Cheap iterative third order cutoff. Usually requires very few, if any, iterations.

Use of QFB - Example

Let $f_1(x) = \frac{1}{2}x^t \cdot A_v \cdot x - A_v \cdot (a \cdot x) + \frac{1}{2}a^t \cdot A_v \cdot a$ with

$$A_v = \begin{pmatrix} 2 & 3 & \dots & 3 \\ -1 & 2 & \dots & 3 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \dots & 2 \end{pmatrix}$$

known to be p.d. with minimum a . Choose a random vector a , and 5^v boxes around it. Check box rejection with Interval evaluation, Centered Form, QFB. Output average number of QFB iterations.

Use of QFB - Example

Let $f_1(x) = \frac{1}{2}x^t \cdot A_v \cdot x - A_v \cdot (a \cdot x) + \frac{1}{2}a^t \cdot A_v \cdot a$ with

$$A_v = \begin{pmatrix} 2 & 3 & \dots & 3 \\ -1 & 2 & \dots & 3 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \dots & 2 \end{pmatrix}$$

known to be p.d. with minimum a . Choose a random vector a , and 5^v boxes around it. Check box rejection with Interval evaluation, Centered Form, QFB. Output average number of QFB iterations.

v	N=5 ^v	NI	NC	NQFB	Avg. Iter
2	25	25	8	1	1.1
4	625	625	308	1	0.31

Use of QFB - Example

Let $f_1(x) = \frac{1}{2}x^t \cdot A_v \cdot x - A_v \cdot (a \cdot x) + \frac{1}{2}a^t \cdot A_v \cdot a$ with

$$A_v = \begin{pmatrix} 2 & 3 & \dots & 3 \\ -1 & 2 & \dots & 3 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \dots & 2 \end{pmatrix}$$

known to be p.d. with minimum a . Choose a random vector a , and 5^v boxes around it. Check box rejection with Interval evaluation, Centered Form, QFB. Output average number of QFB iterations.

v	N=5 ^v	NI	NC	NQFB	Avg. Iter
2	25	25	8	1	1.1
4	625	625	308	1	0.31
6	15,625	15,625	12,434	1	0.31
8	390,625	390,625	372,376	1	0.43
10	9,765,625	9,765,625	9,622,750	1	0.55

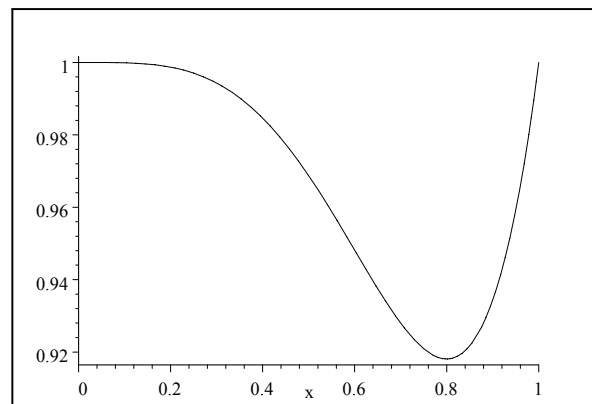
Key Features and Algorithms of COSY-GO

- List management of boxes not yet determined to not contain the global minimizer. Loading a new box. Discarding a box with range above the current threshold value. Splitting a box with range not above the threshold value for further analysis. Storing a box smaller than the specified size.
- Application of a series of bounding schemes, starting from mere interval arithmetic to naive Taylor model bounding, LDB, then QFB. A higher bounding scheme is executed only if all the lower schemes fail.
- Update of the threshold cutoff value via various schemes. It includes upper bound estimates of the local minimum by corresponding bounding schemes, the mid point estimate, global estimates based on local behavior of function using gradient line search and convex quadratic form.
- Box size reduction using LDB.
- Resulting data is available in various levels including graphics output.

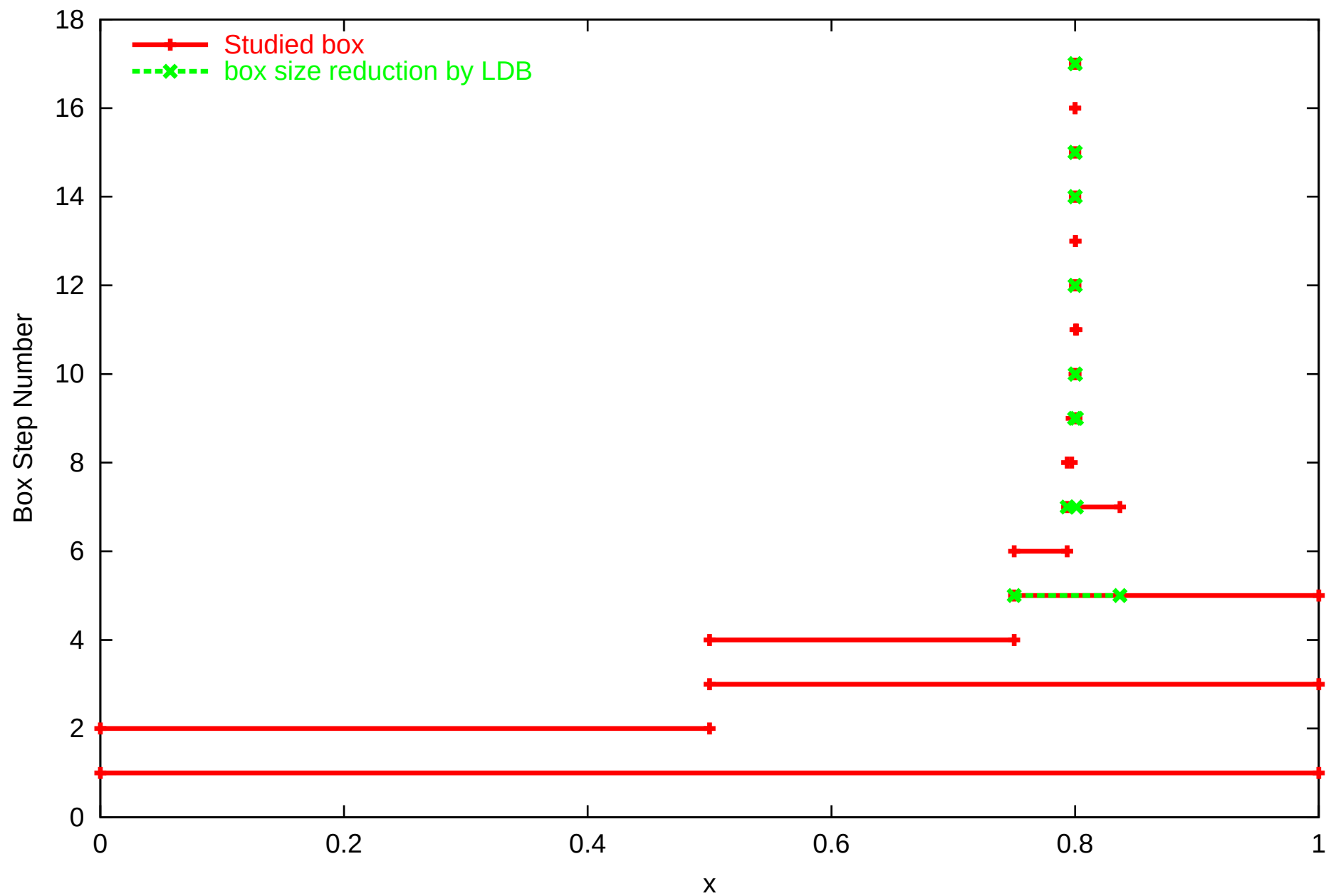
Moore's Simple 1D Function

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

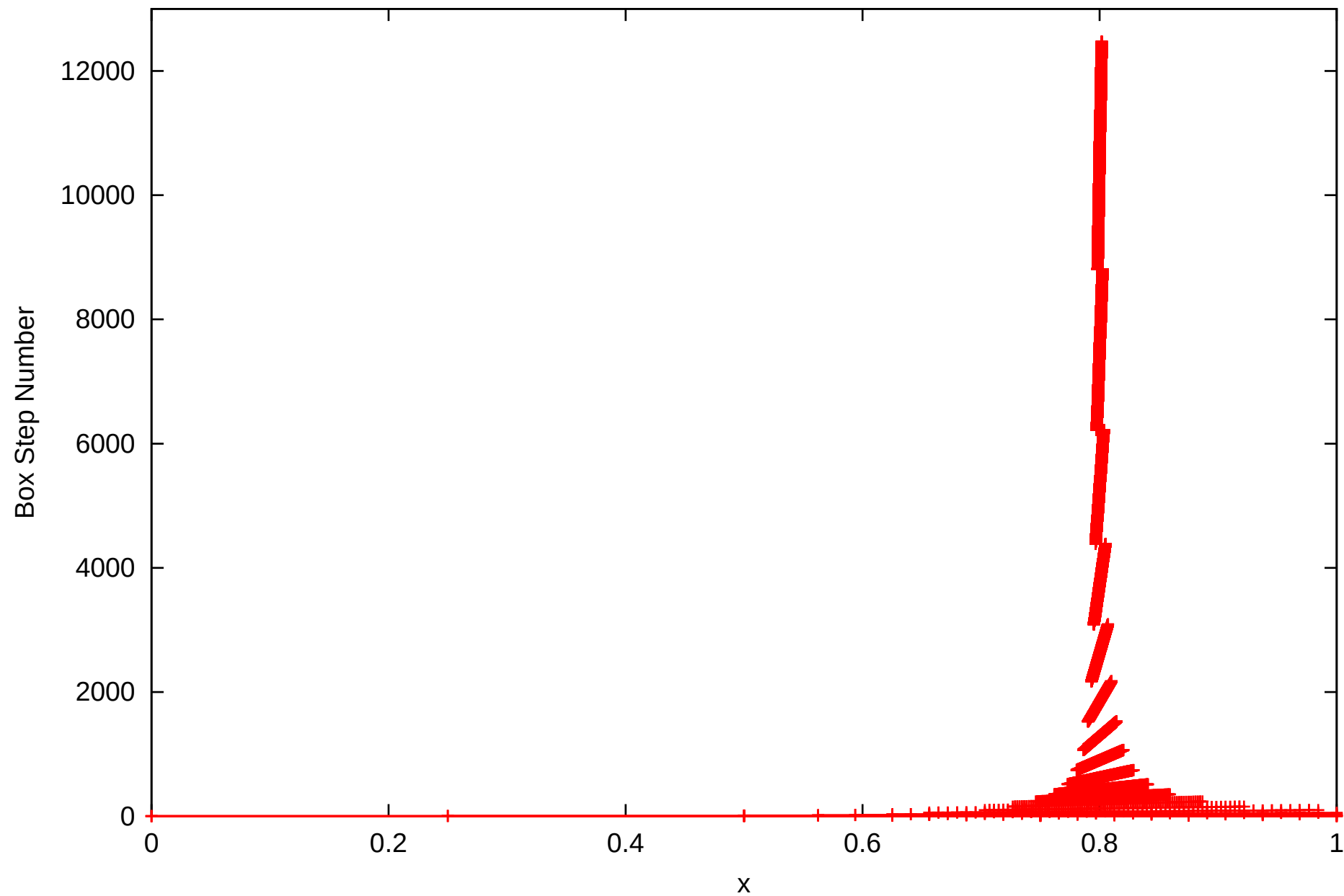
Study on $[0, 1]$. Trivial-looking, but dependency and high order.
Assumes shallow min at 0.8.



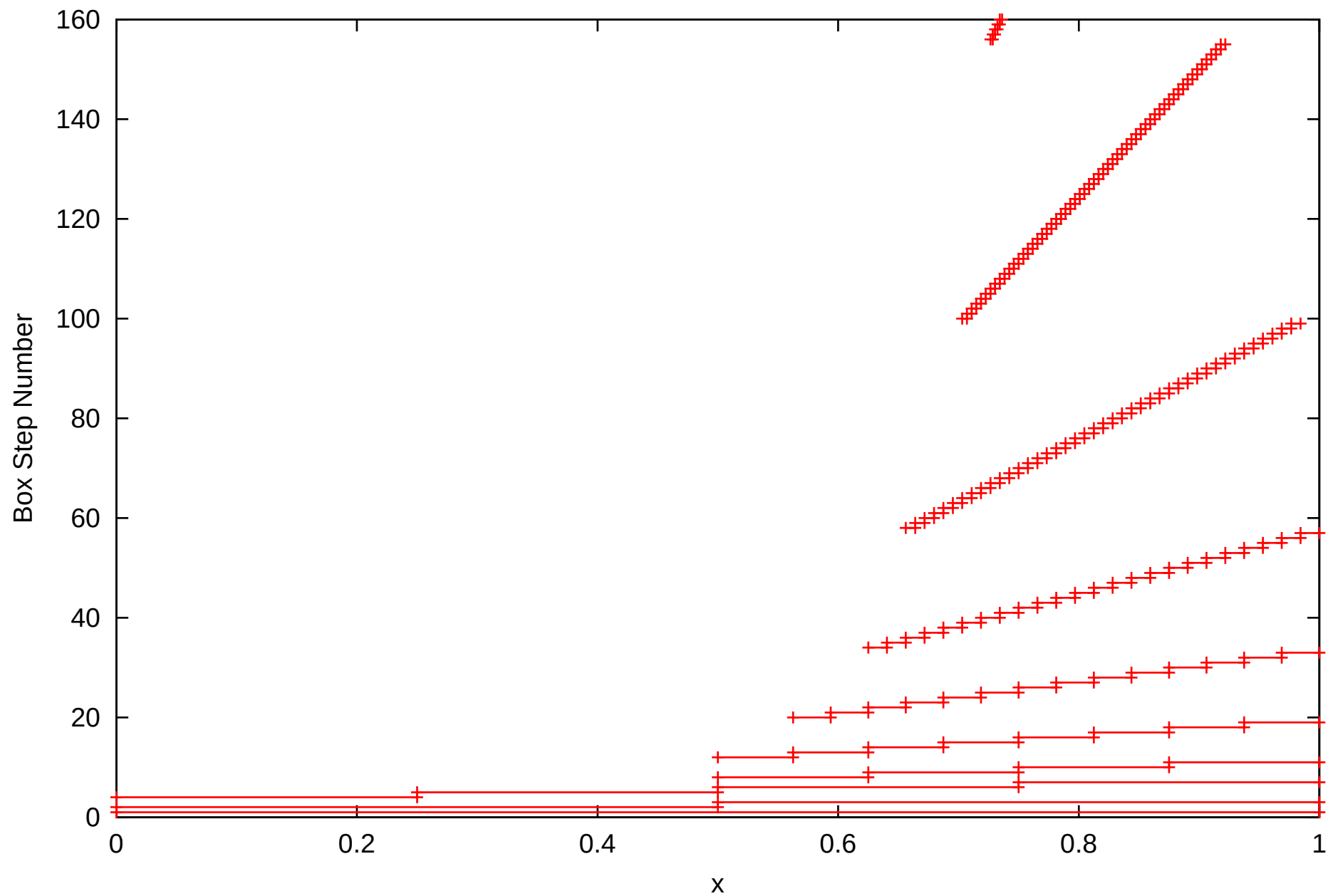
COSY-GO with LDB-QFB (TM 5th). 1D. $f=x^5-x^4+1$



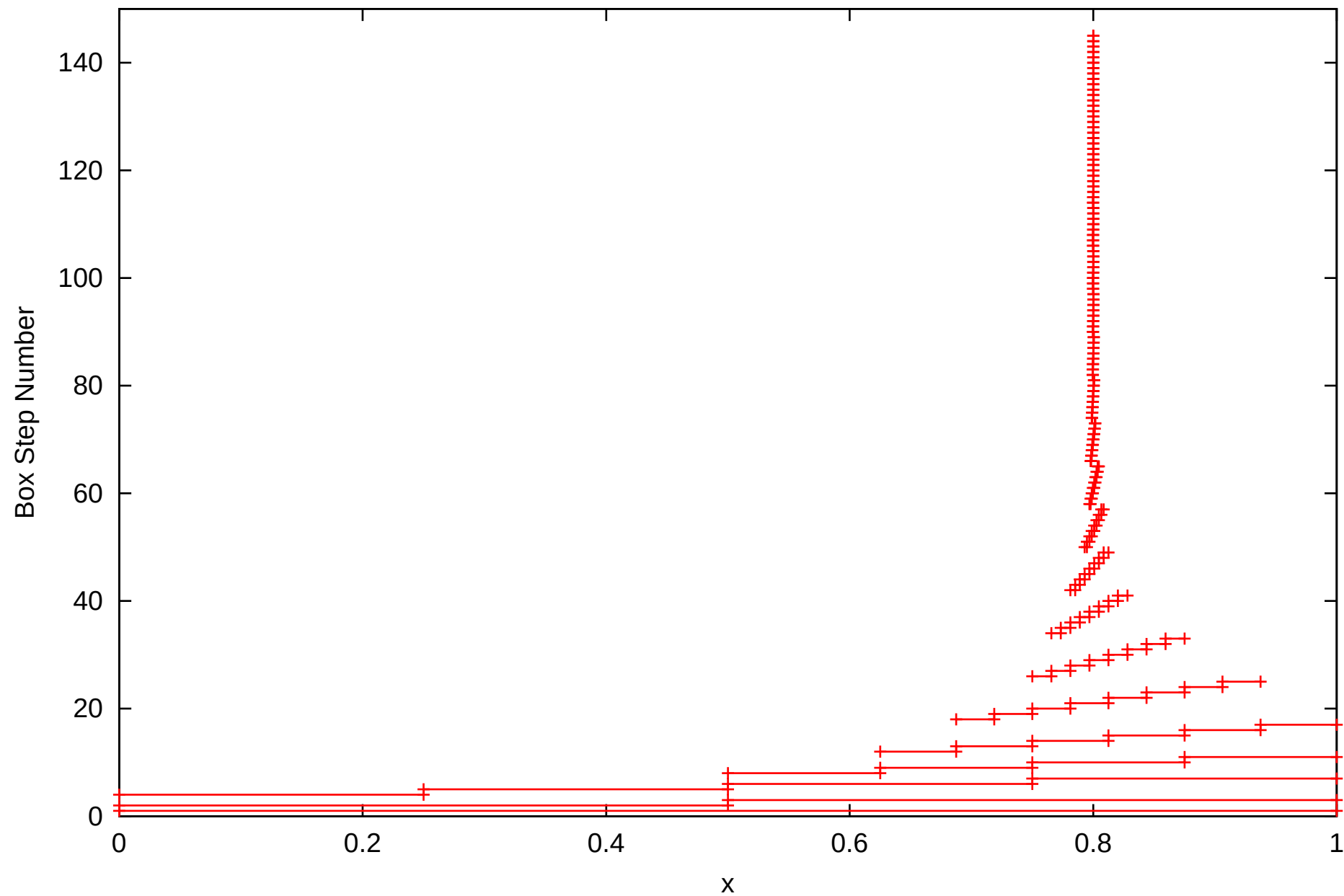
COSY-GO with naive IN with mid point test. 1D. $f=x^5-x^4+1$



COSY-GO with IN. 1D. $f=x^5-x^4+1$. -- Up to the 160th box



COSY-GO with Centered Form with mid point test. 1D. $f=x^5-x^4+1$



Beale's 2D and 4D Function

$$f(x_1, x_2) = (1.5 - x_1(1 - x_2))^2 + (2.25 - x_1(1 - x_2^2))^2 + (2.625 - x_1(1 - x_2^3))^2$$

Domain $[-4.5, 4.5]^2$. Minimum value 0 at $(3, 0.5)$.

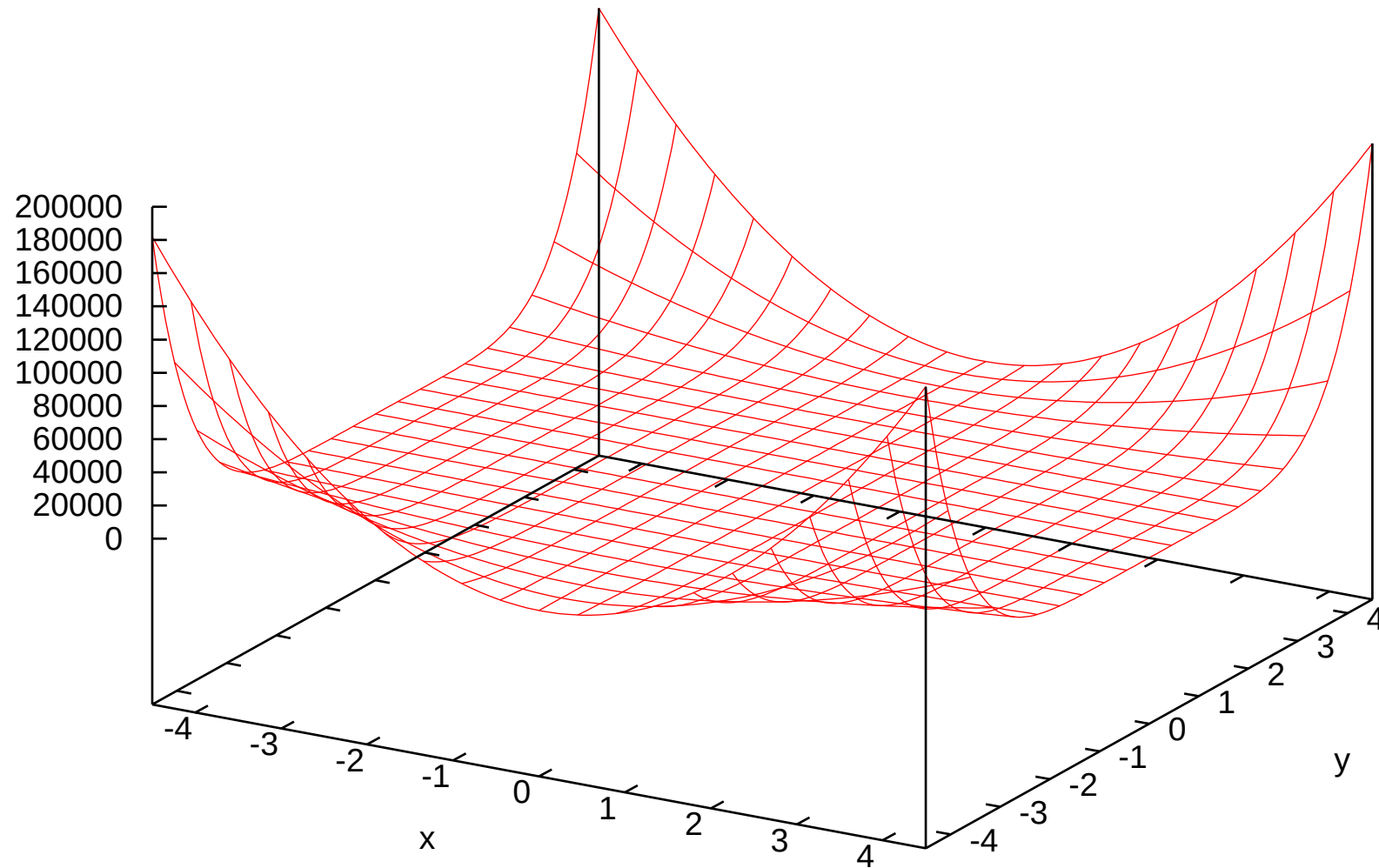
Little dependency, but tricky very shallow behavior.

Generalization to 4D:

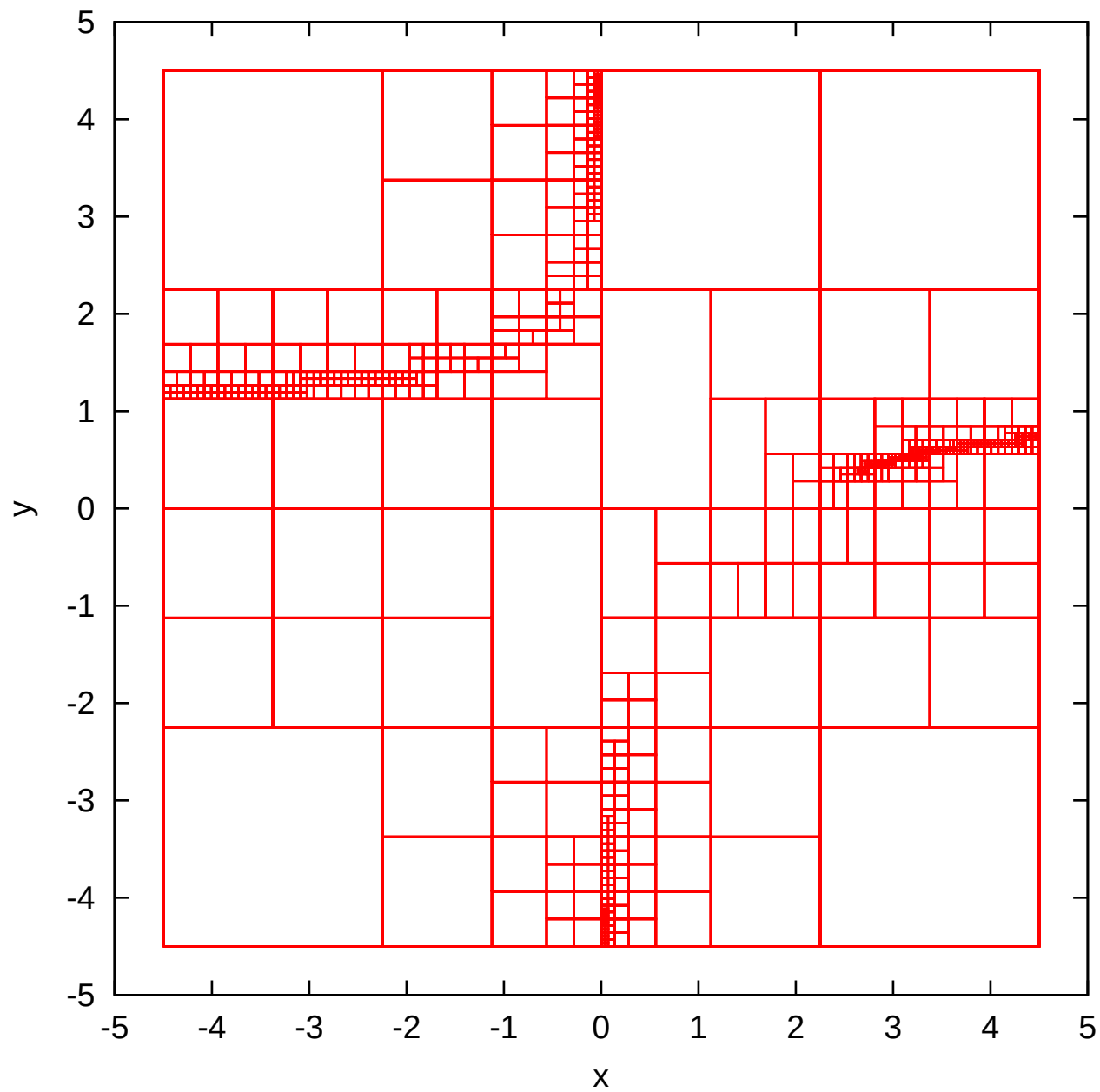
$$\begin{aligned} f(x_1, x_2, x_3, x_4) = & (1.5 - x_1(1 - x_2))^2 + (2.25 - x_1(1 - x_2^2))^2 + (2.625 - x_1(1 - x_2^3))^2 \\ & + (1 + x_3(1 - x_4))^2 + (3 + x_3(1 - x_4^2))^2 + (7 + x_3(1 - x_4^3))^2 \\ & + (3 + x_1(1 - x_4))^2 + (9 + x_1(1 - x_4^2))^2 + (21 + x_1(1 - x_4^3))^2 \\ & + (0.5 - x_3(1 - x_2))^2 + (0.75 - x_3(1 - x_2^2))^2 + (0.875 - x_3(1 - x_2^3))^2 \end{aligned}$$

Domain $[0, 4]^4$. Minimum value 0 at $(3, 0.5, 1, 2)$

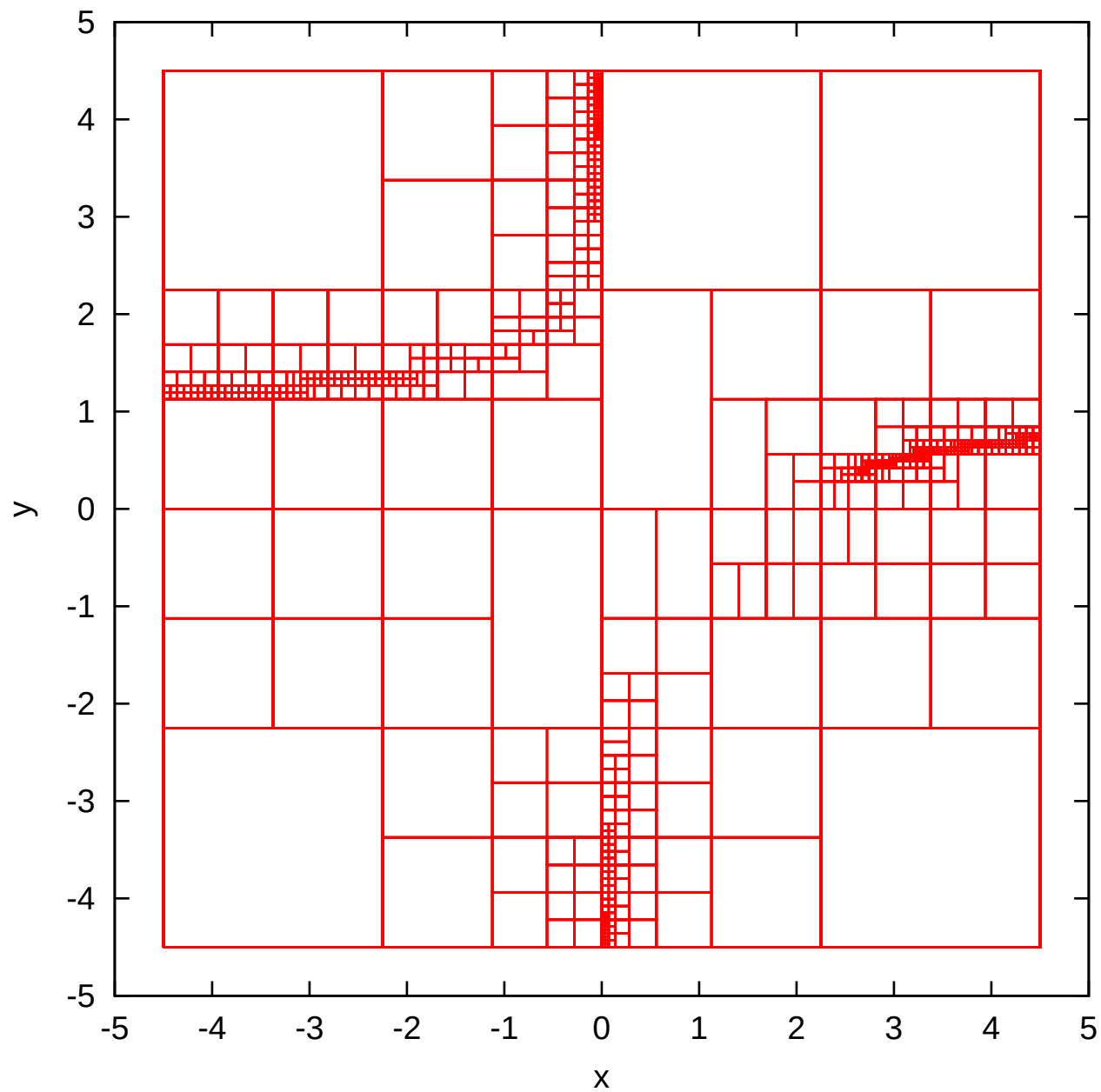
The Beale function. $f = [1.5-x(1-y)]^2 + [2.25-x(1-y^2)]^2 + [2.625-x(1-y^3)]^2$



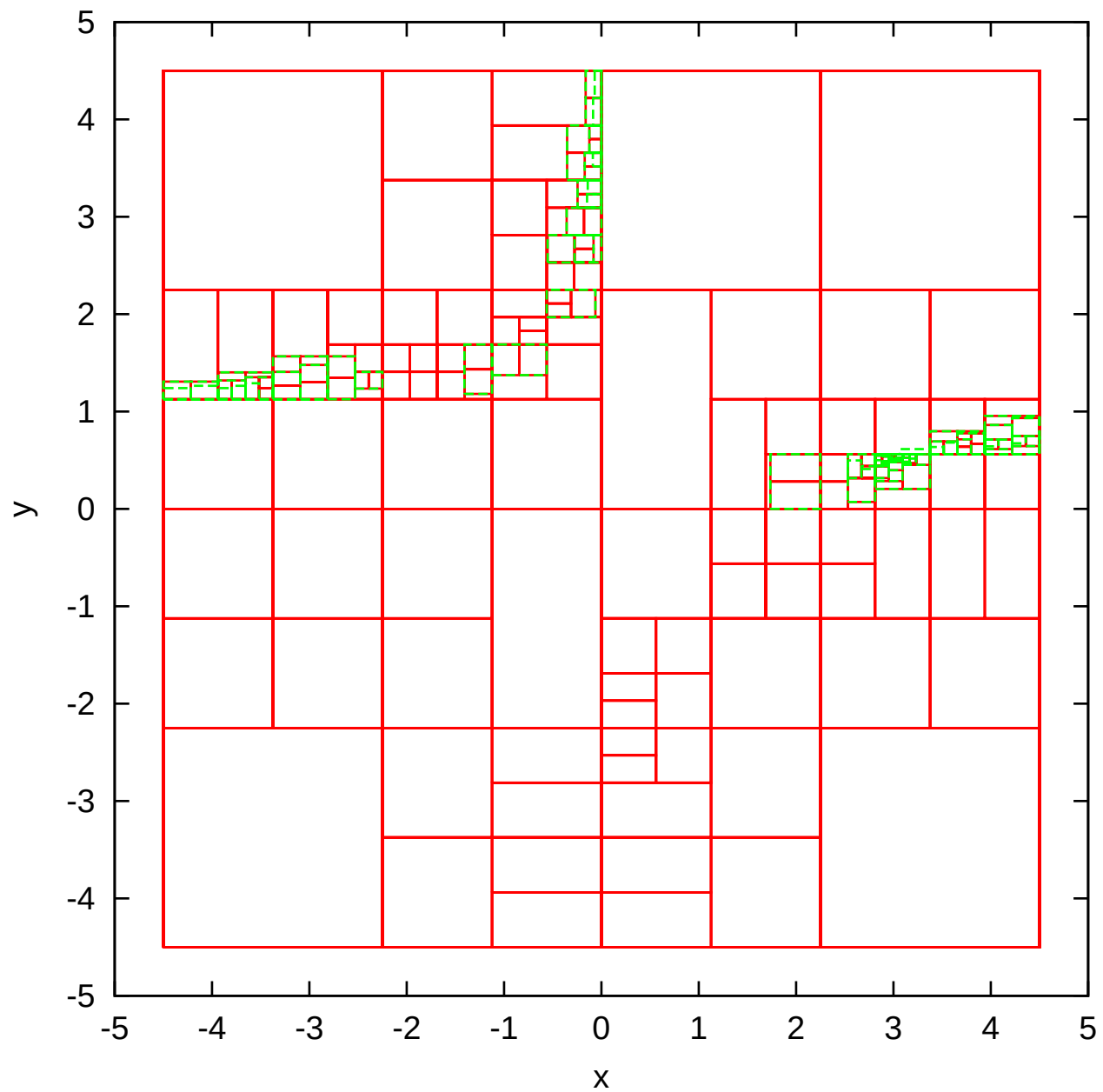
COSY-GO with IN. The Beale function



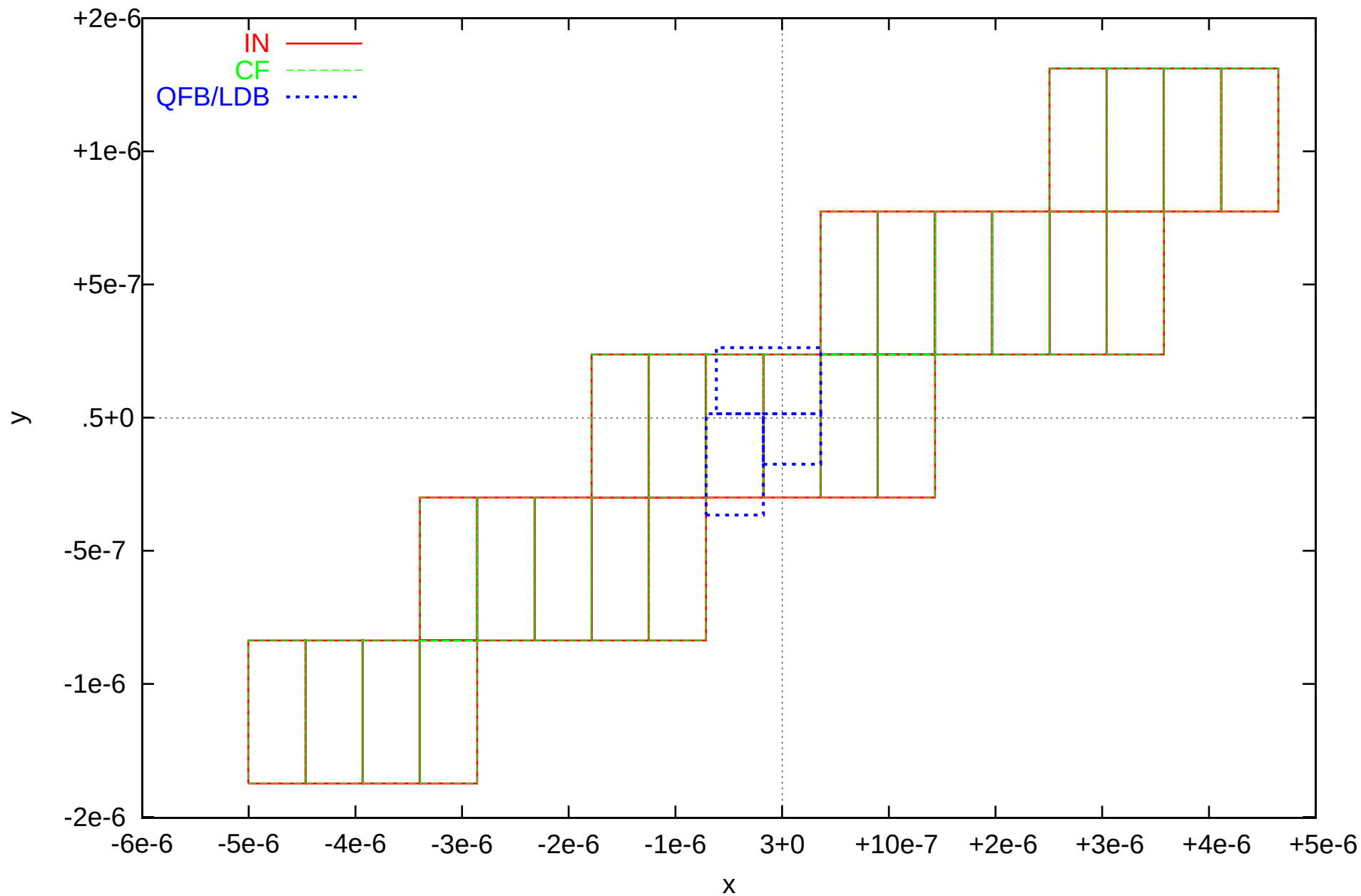
COSY-GO with CF. The Beale function



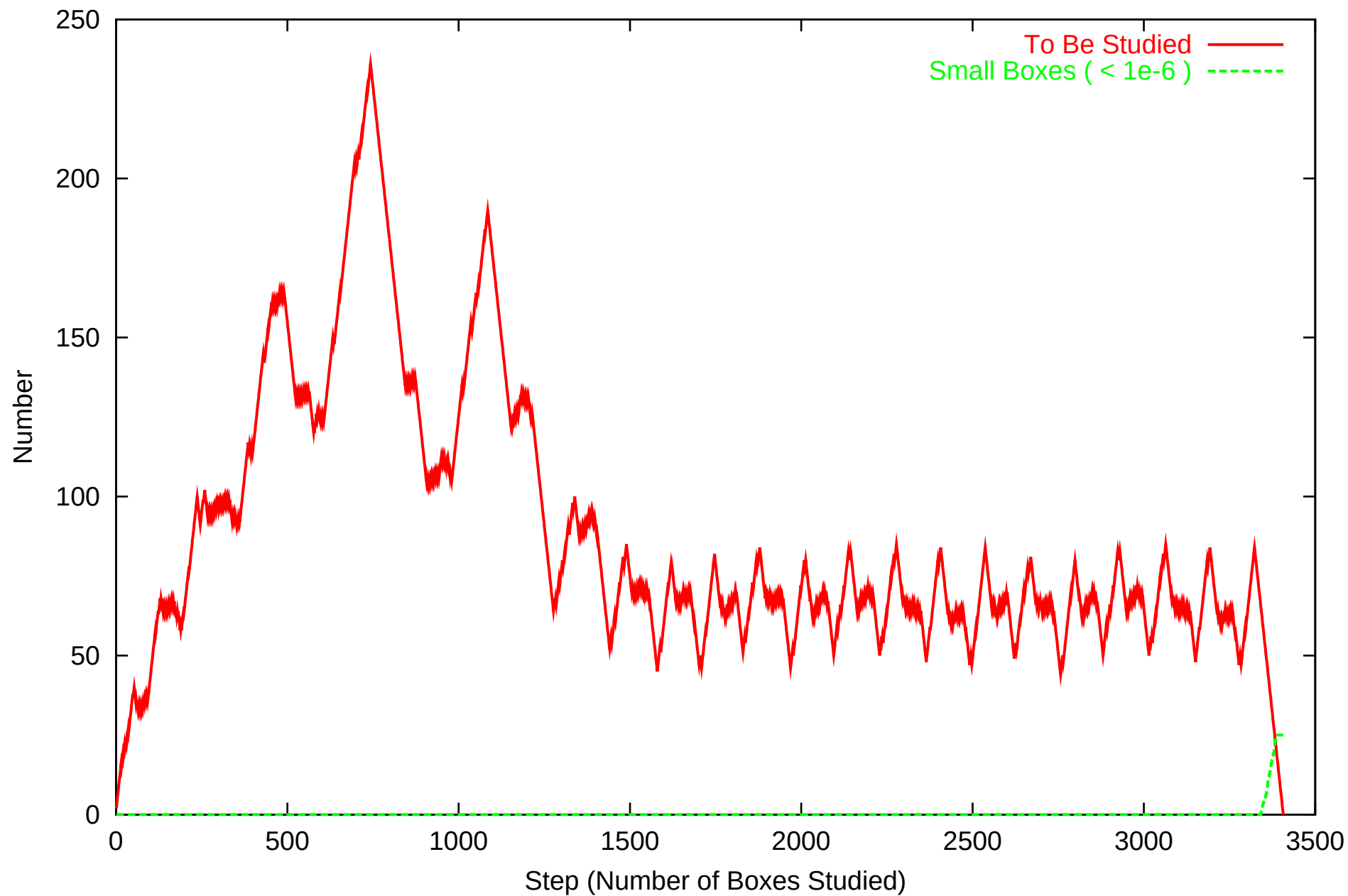
COSY-GO with LDB/QFB. The Beale function



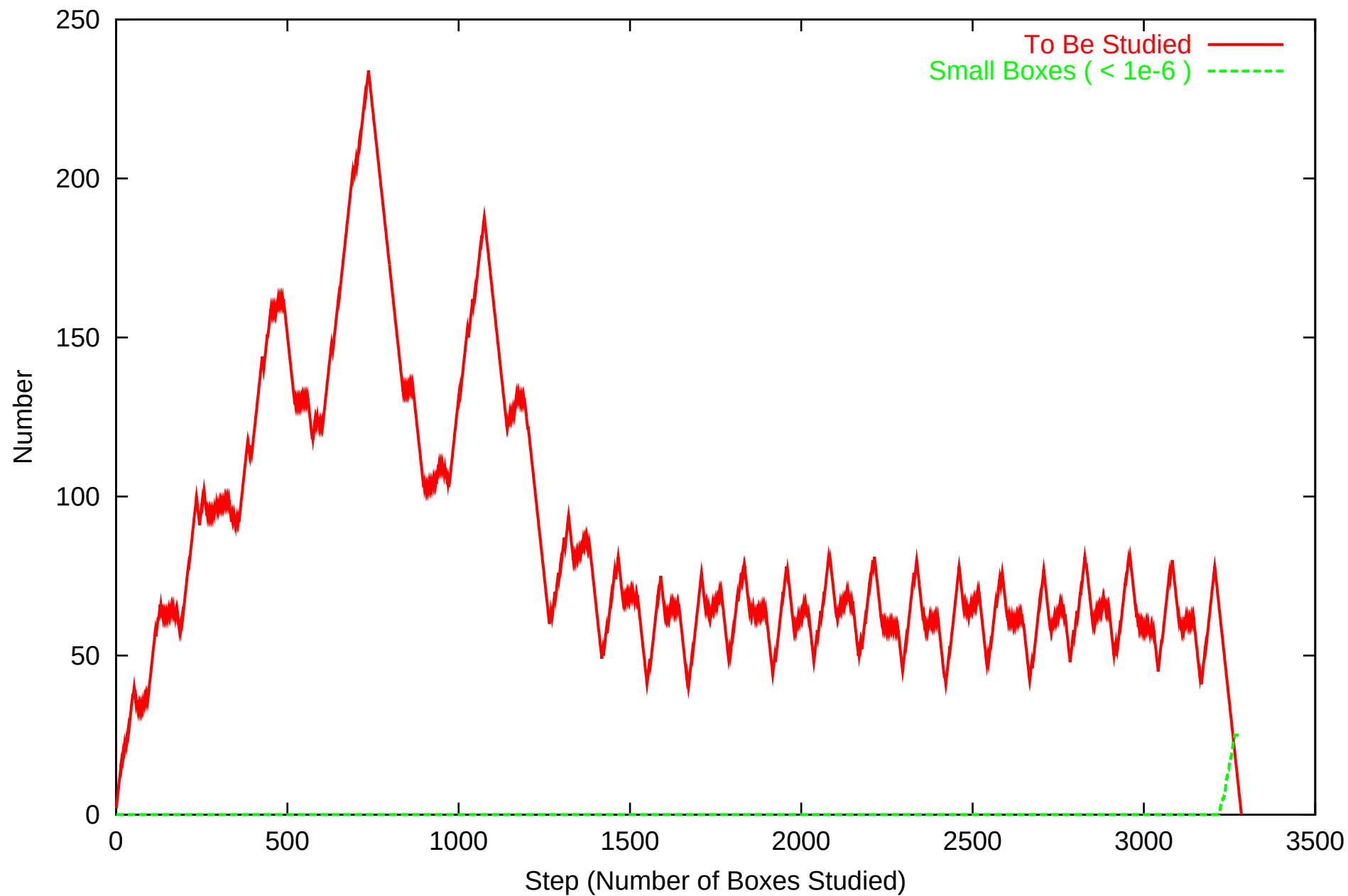
COSY-GO. The Beale function. Remaining Boxes ($< 1e-6$) around (3,0.5)



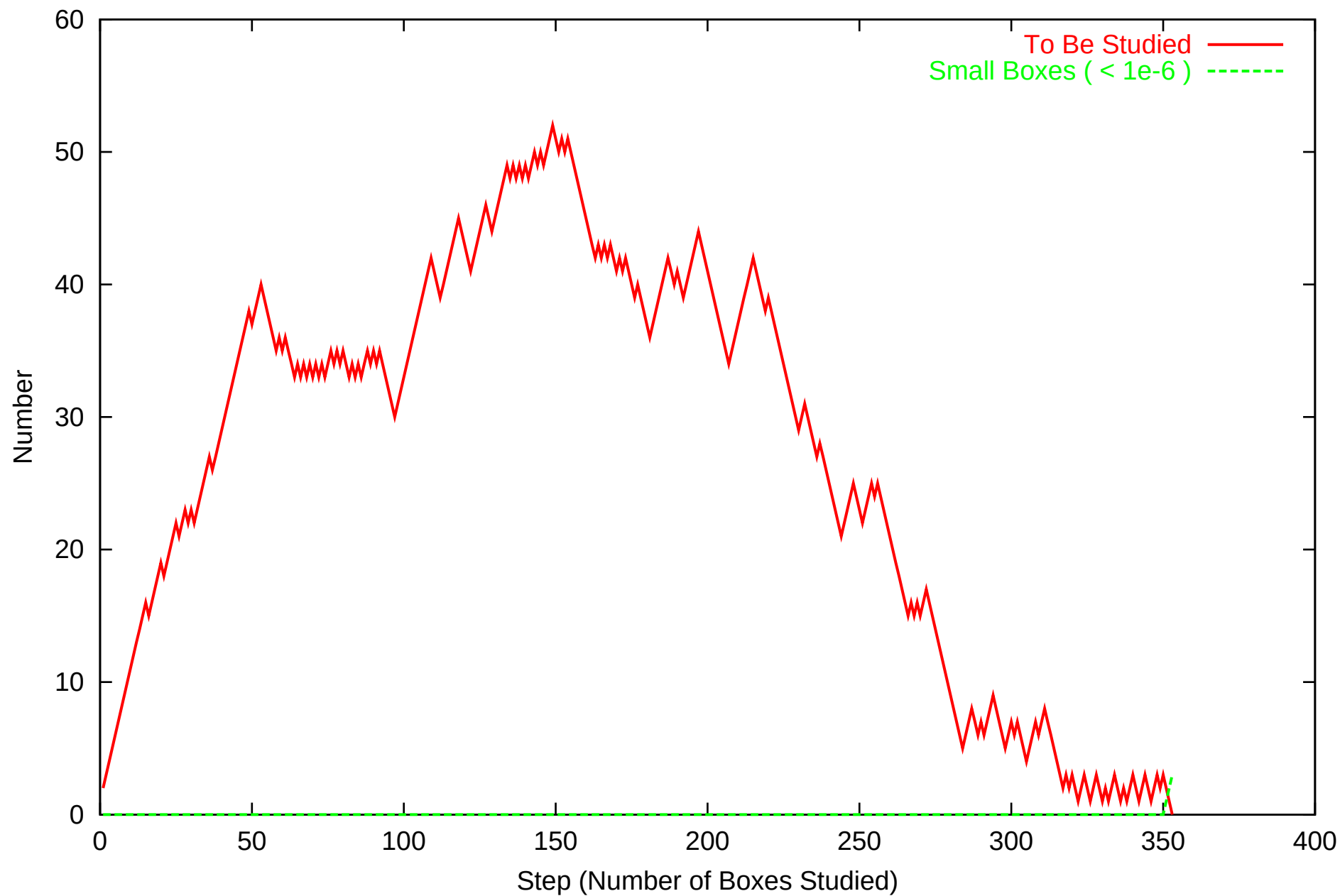
COSY-GO The Beale Function: Number of Boxes -- IN



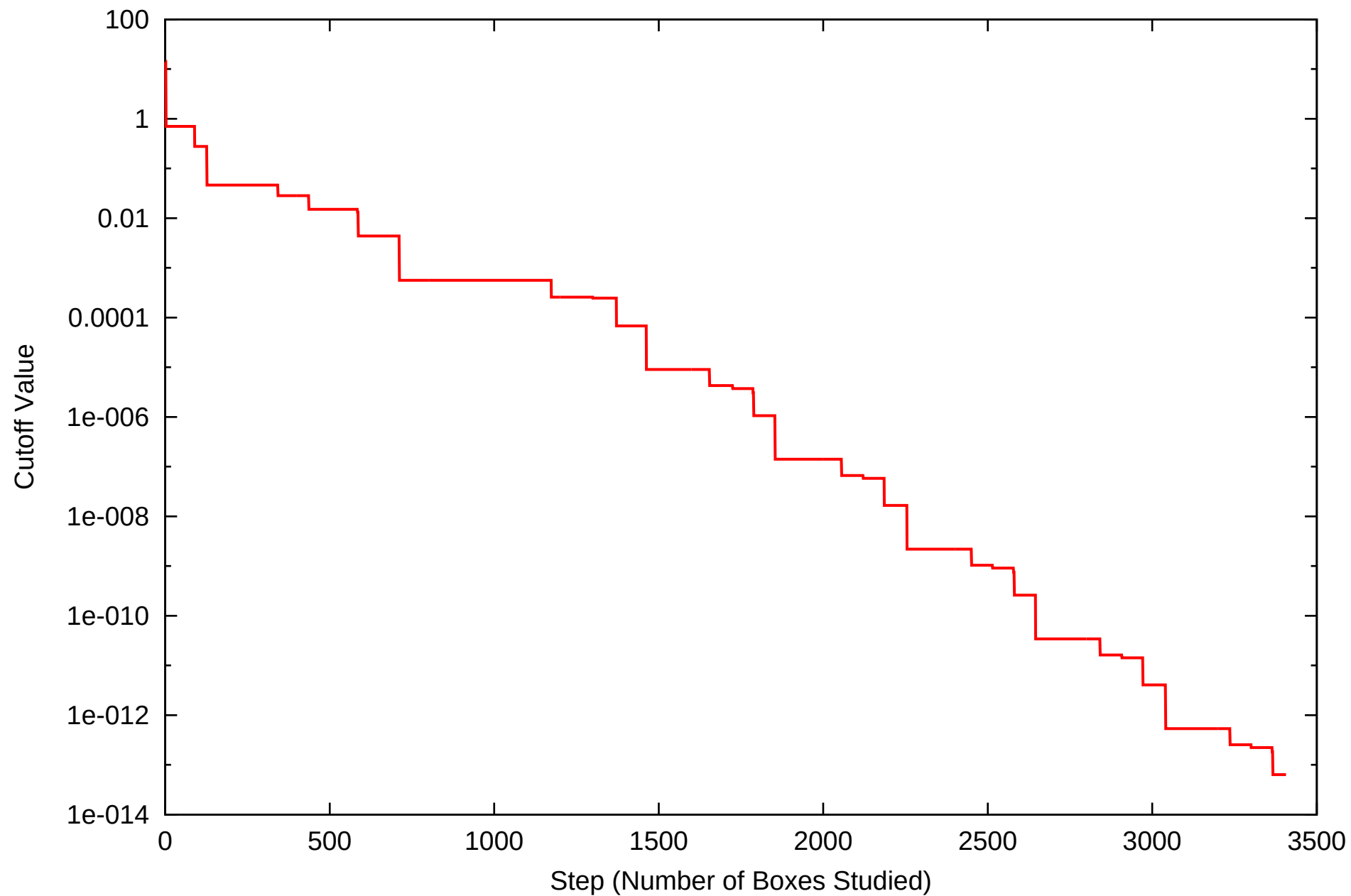
COSY-GO The Beale Function: Number of Boxes -- CF



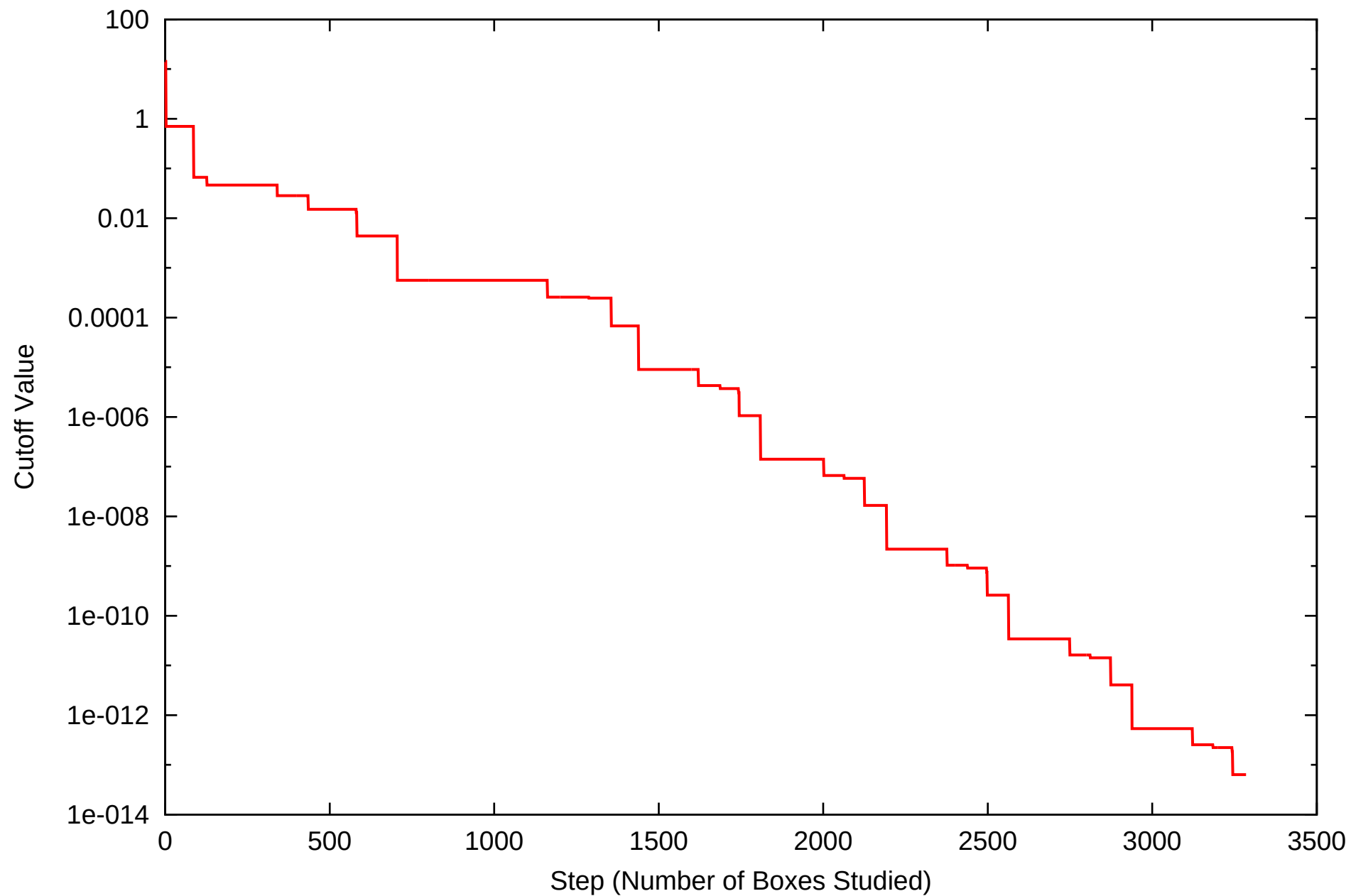
COSY-GO The Beale Function: Number of Boxes -- LDB/QFB



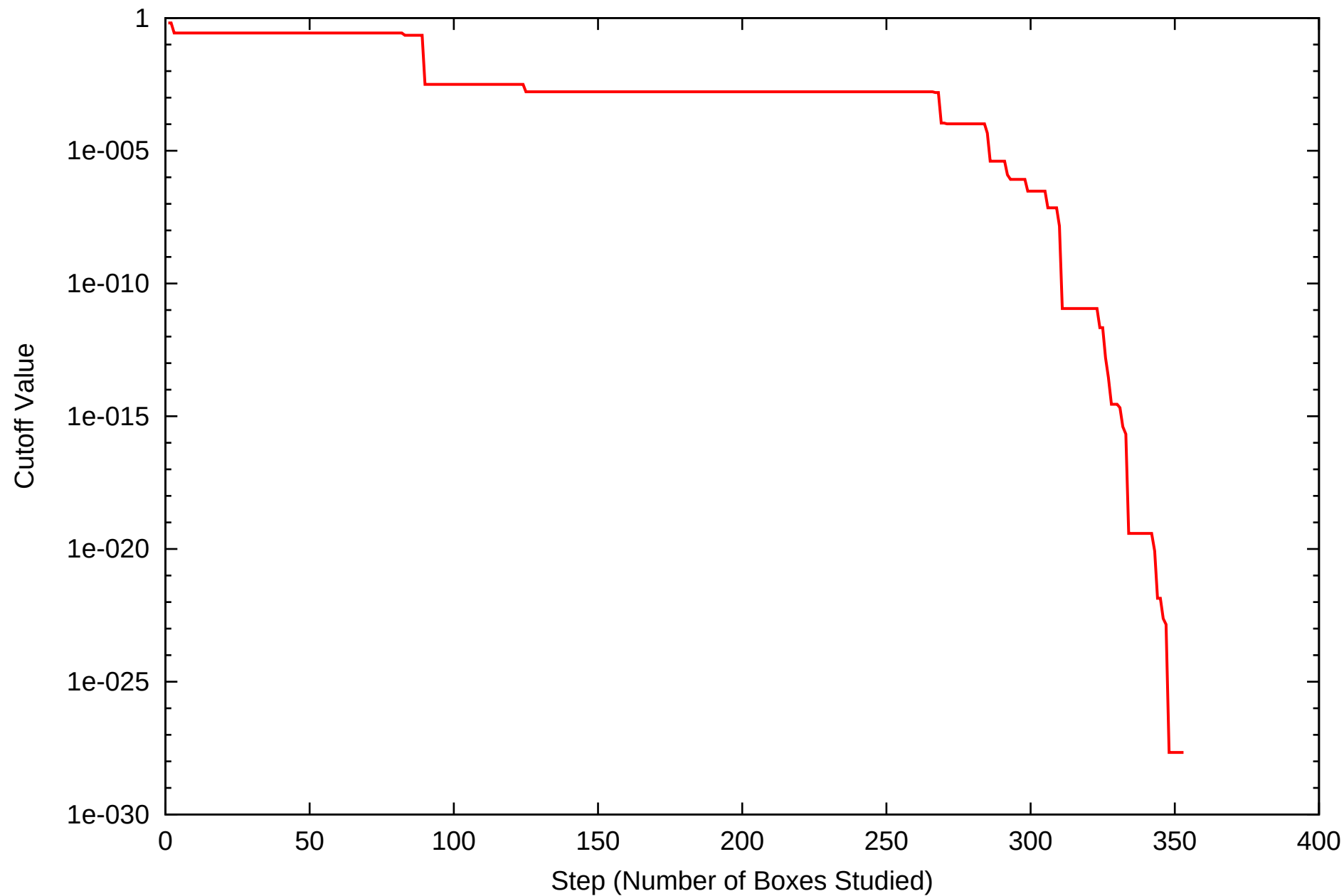
COSY-GO The Beale Function: Cutoff Value -- IN



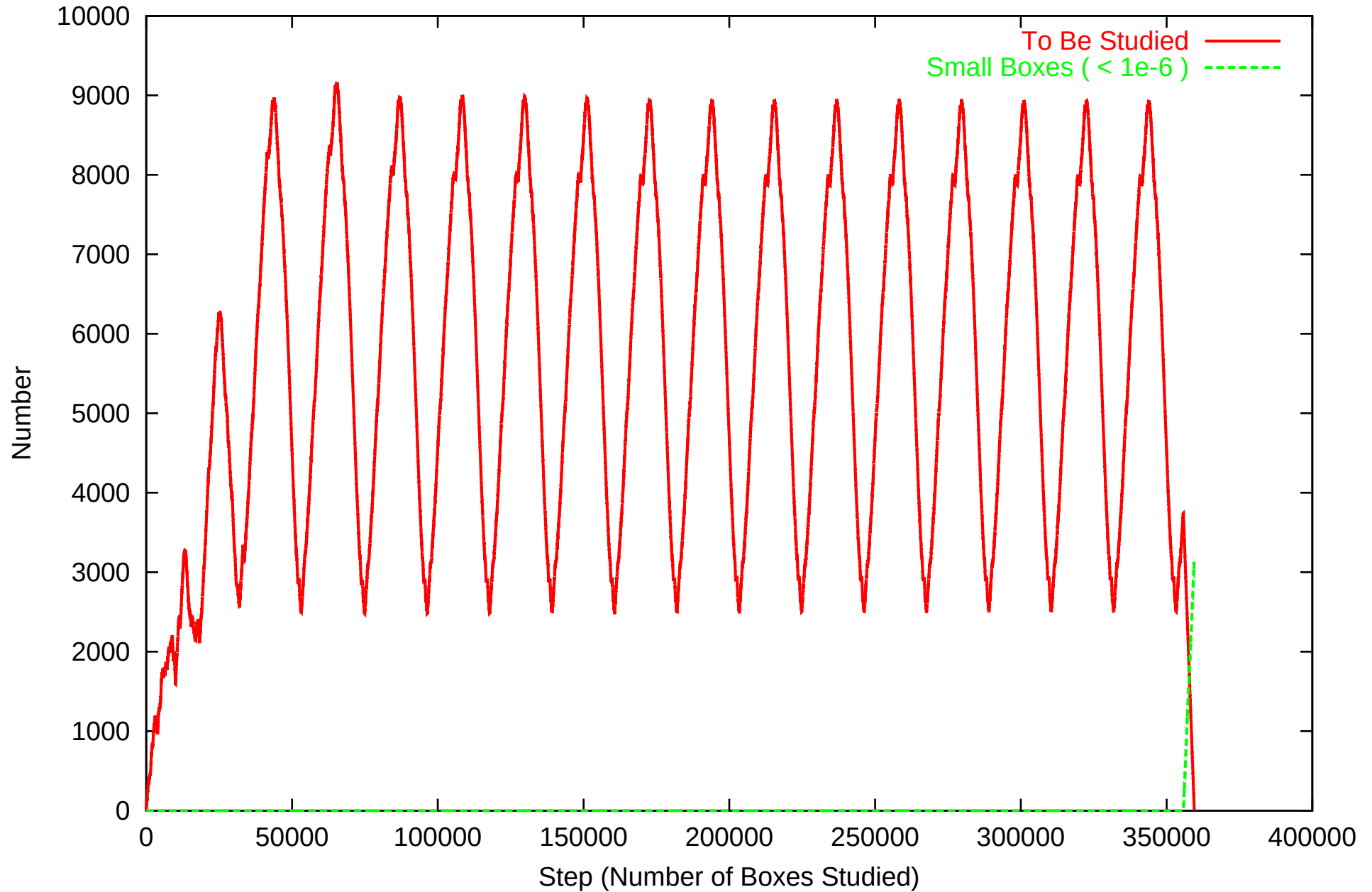
COSY-GO The Beale Function: Cutoff Value -- CF



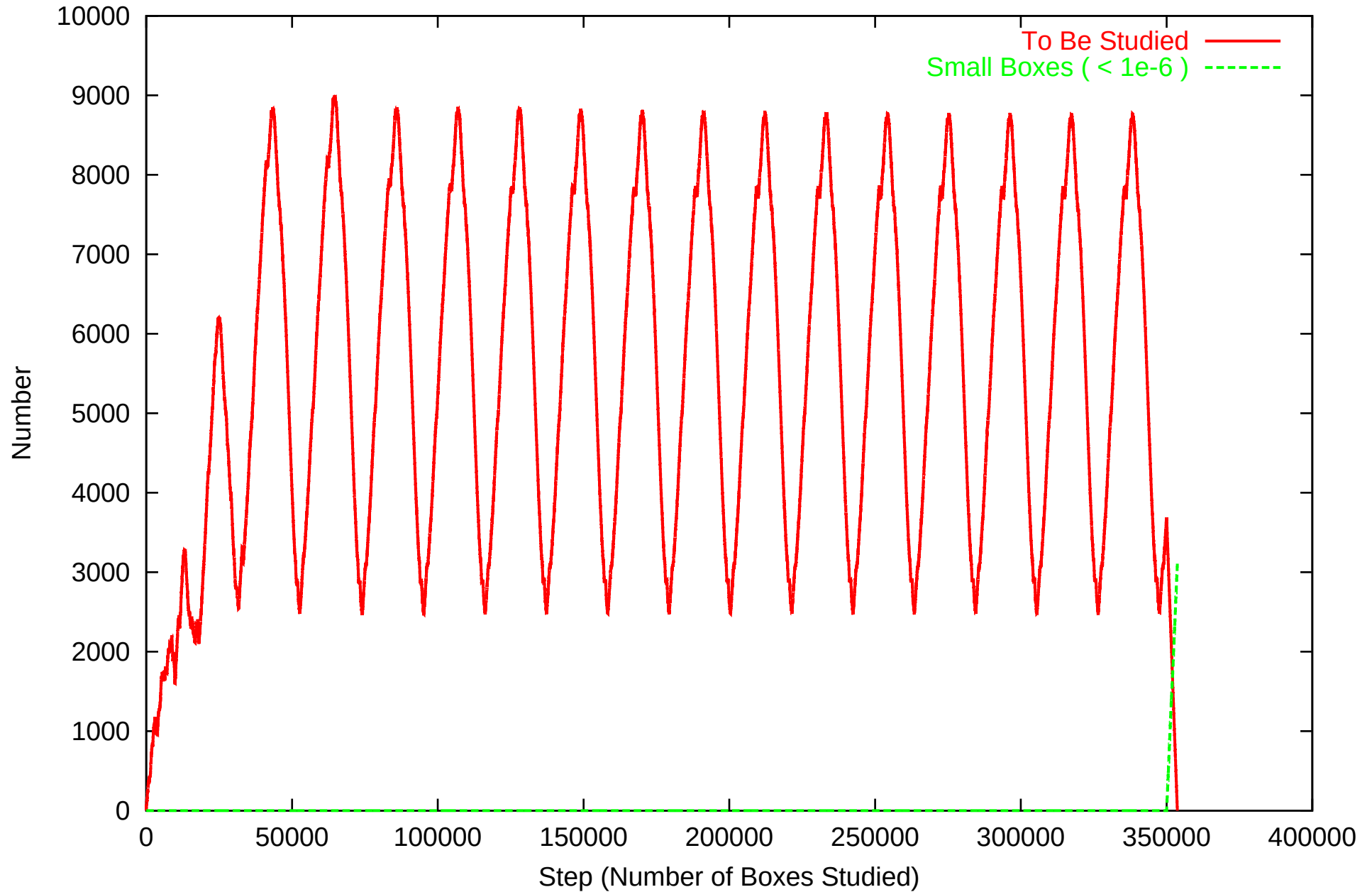
COSY-GO The Beale Function: Cutoff Value -- LDB/QFB



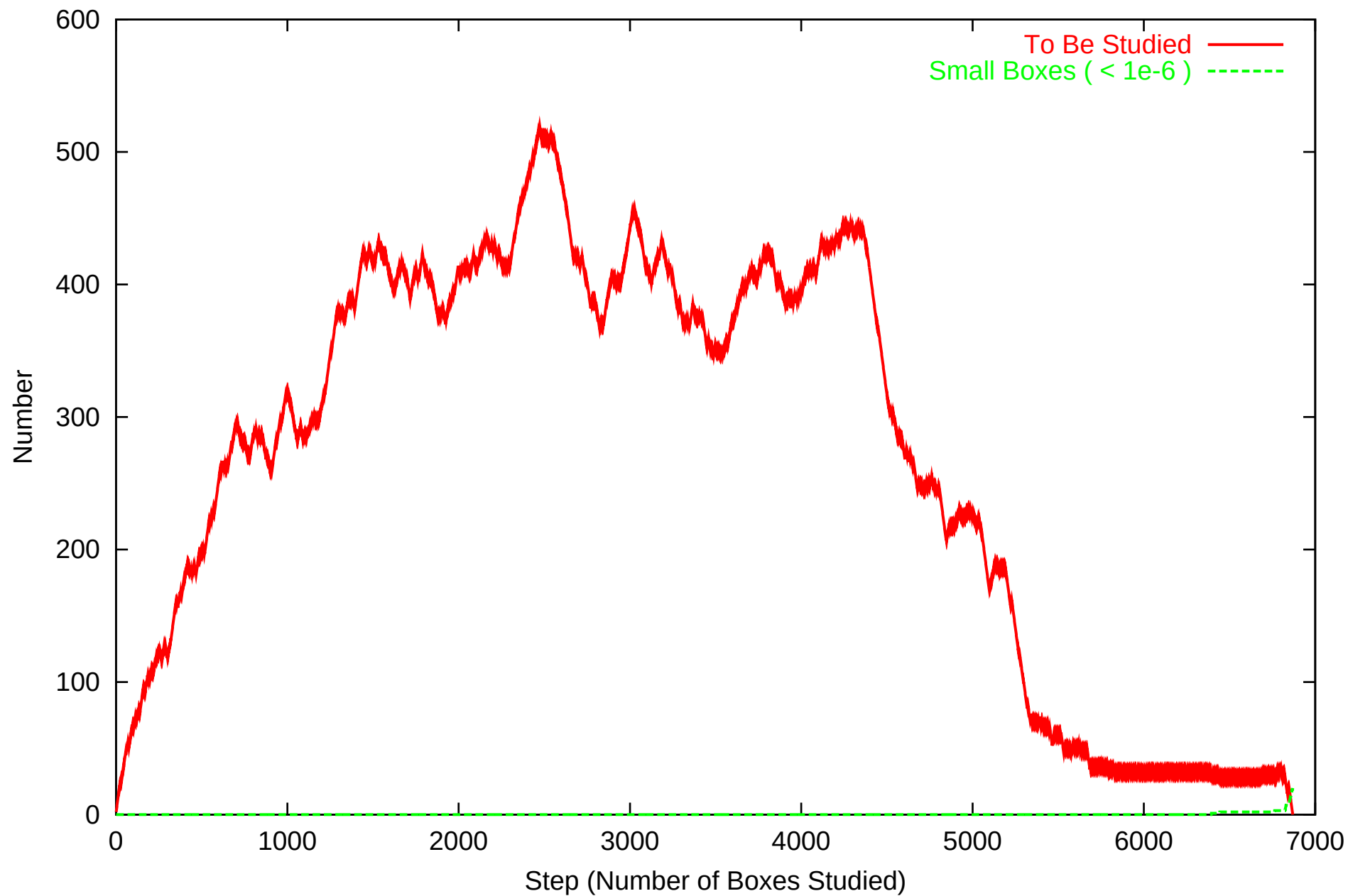
COSY-GO Beale 4D: Number of Boxes -- IN



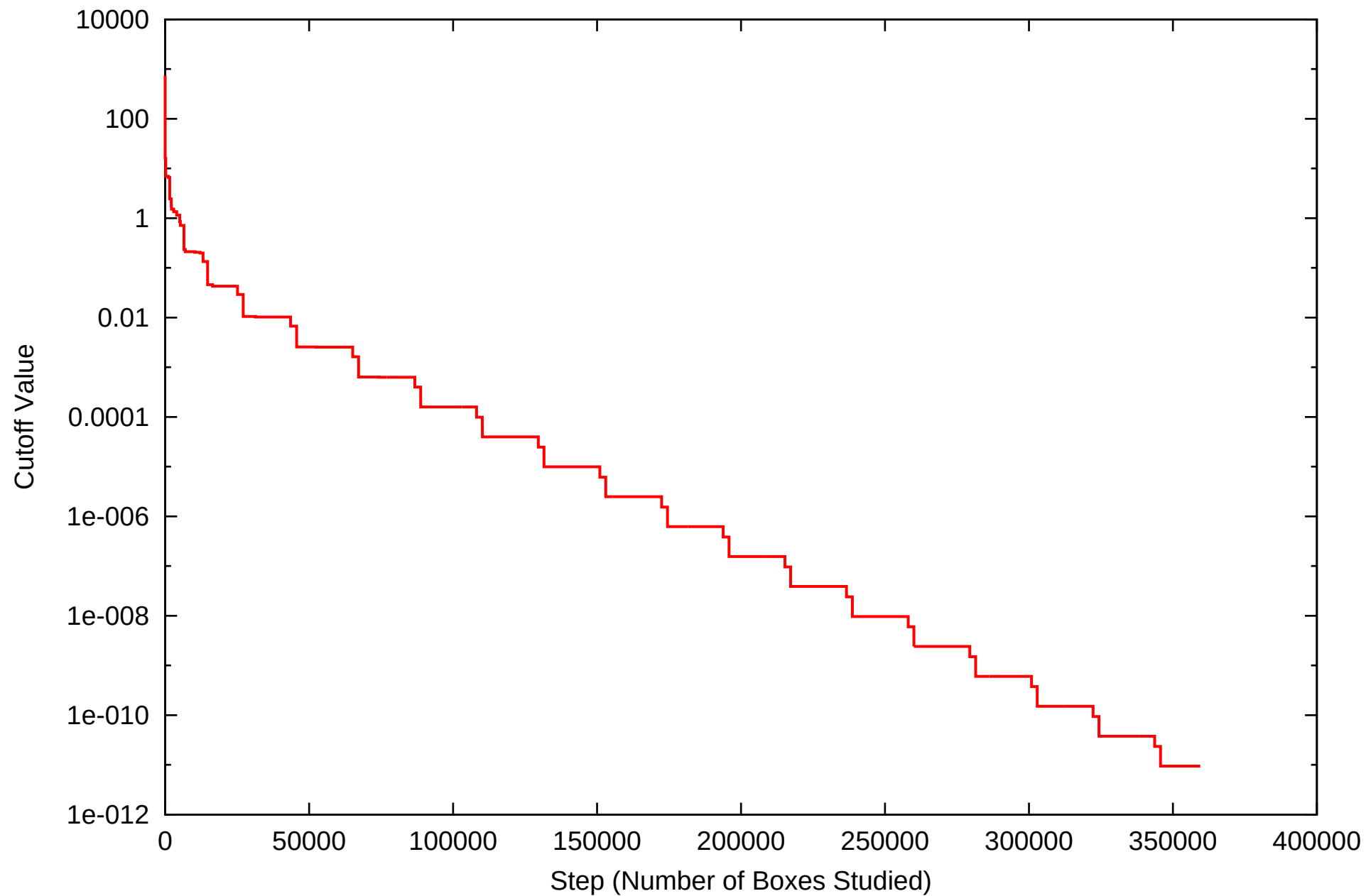
COSY-GO Beale 4D: Number of Boxes -- CF



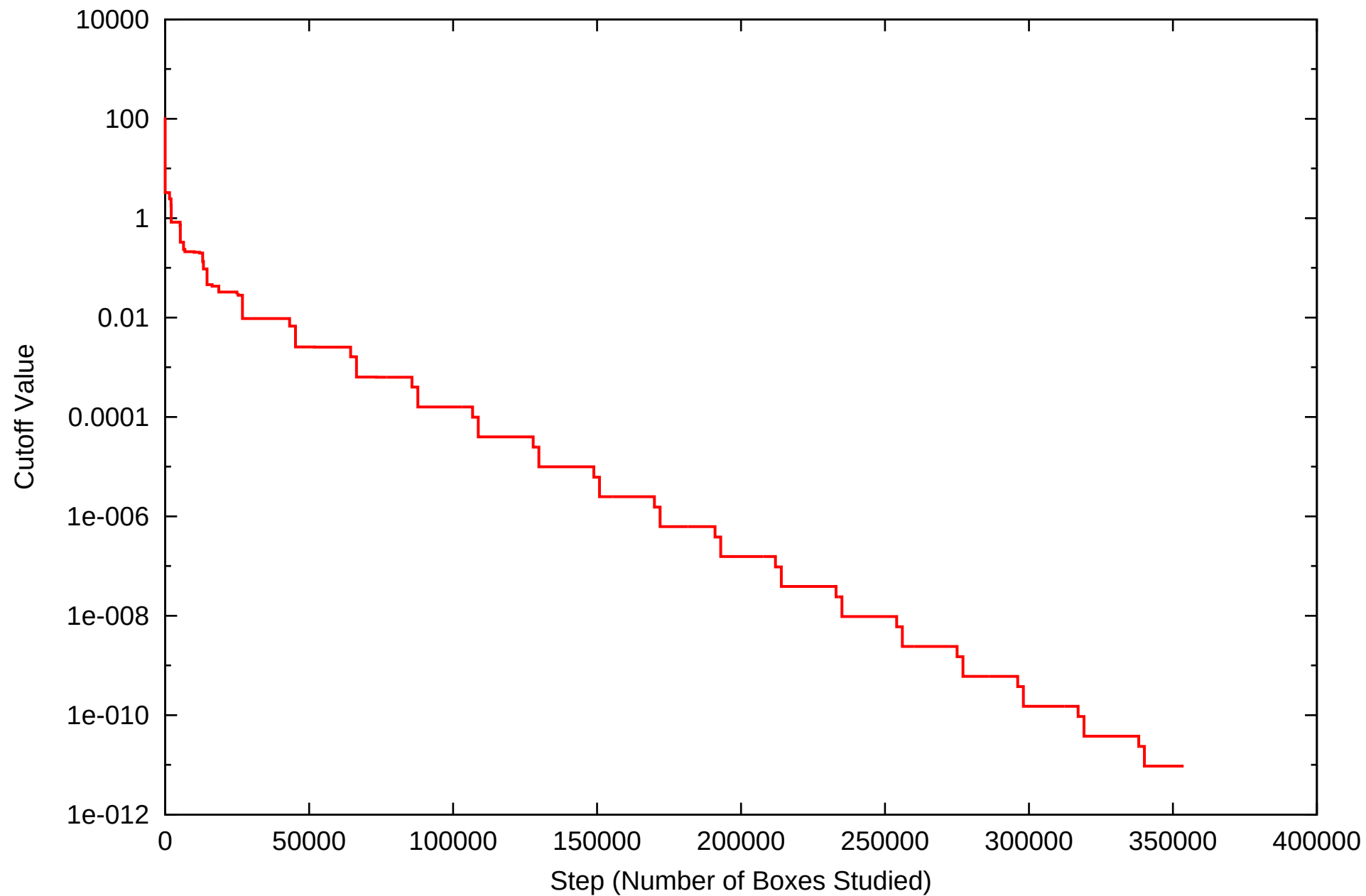
COSY-GO Beale 4D: Number of Boxes -- LDB/QFB



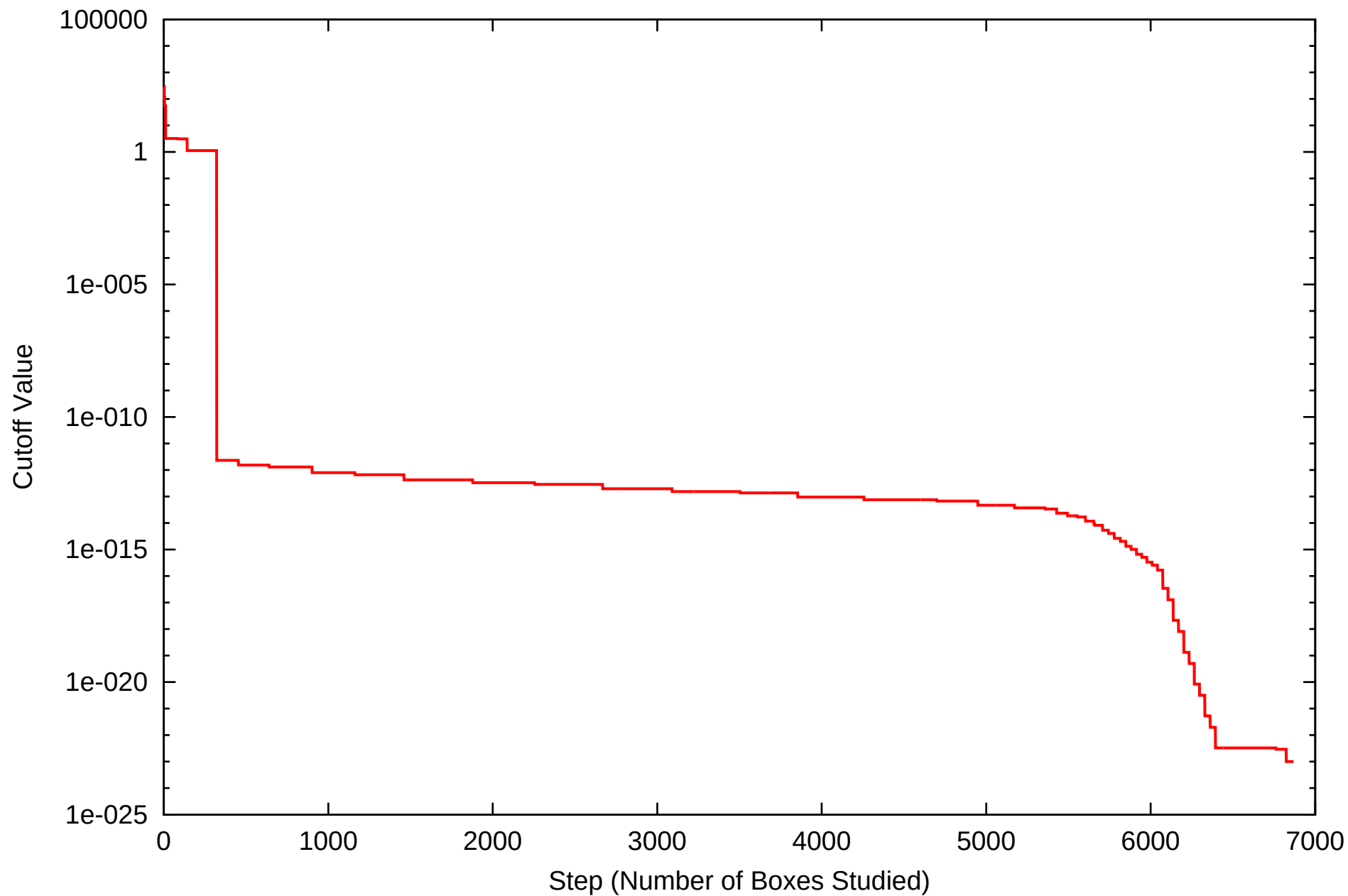
COSY-GO Beale 4D: Cutoff Value -- IN



COSY-GO Beale 4D: Cutoff Value -- CF



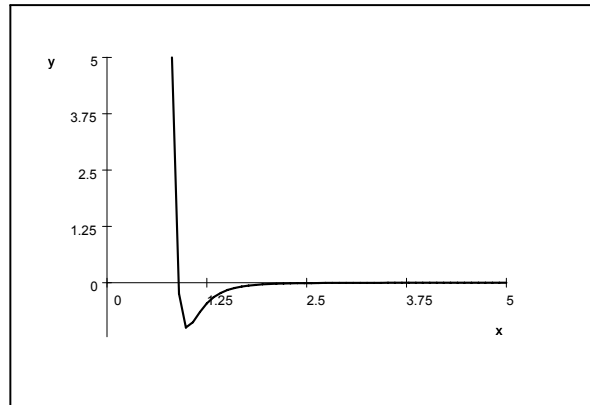
COSY-GO Beale 4D: Cutoff Value -- LDB/QFB



Lennard-Jones Potentials

Ensemble of n particles interacting pointwise with potentials

$$V_{LJ}(r) = \frac{1}{r^{12}} - 2 \cdot \frac{1}{r^6}$$



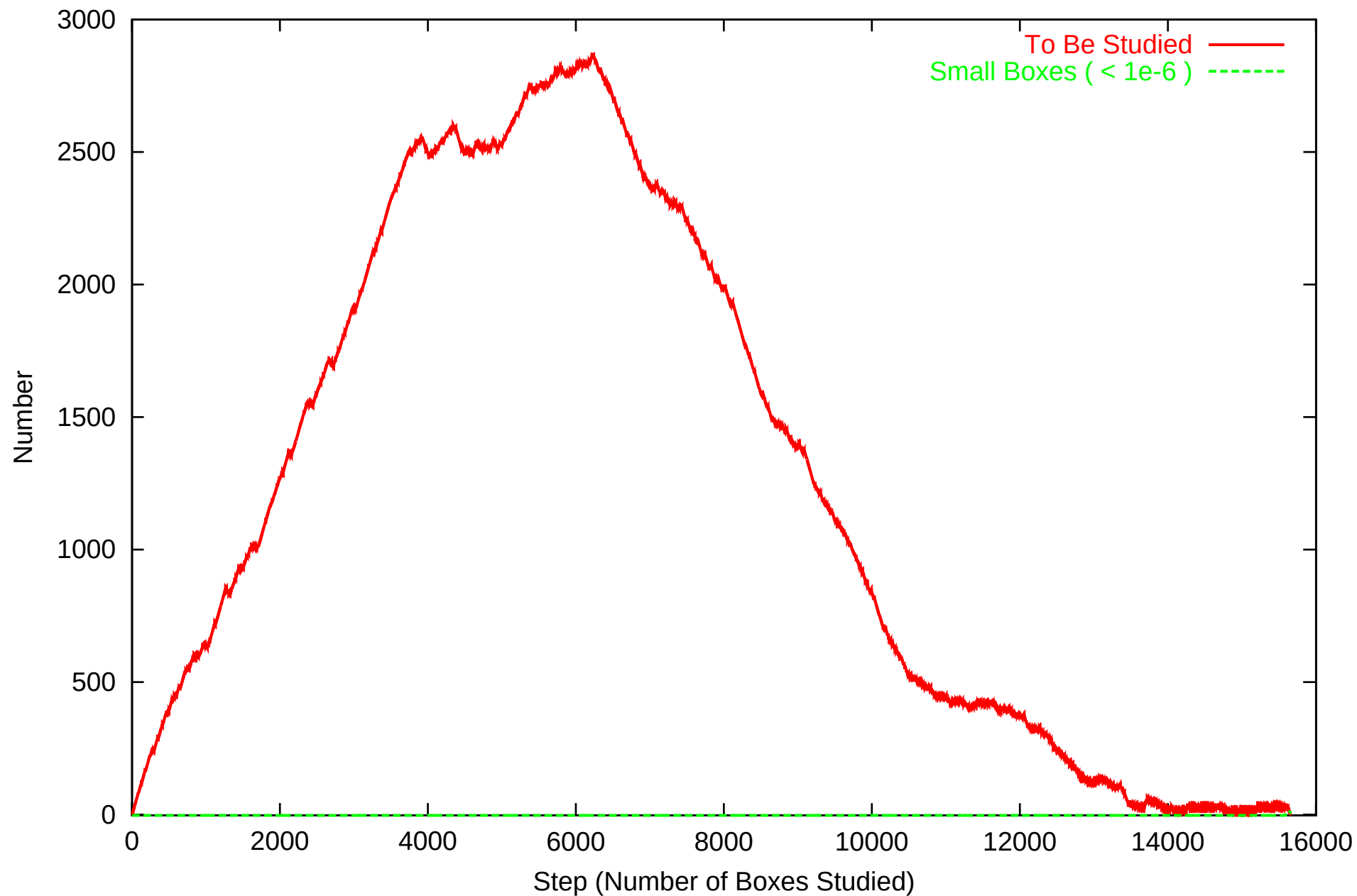
Has very shallow minimum of -1 at $r = 1$. Very hard to Taylor expand.
Extremely wide range of function values: $V_{LJ}(0.5) \approx 4000$, $V_{LJ}(2) \approx 0.03$

$$V = \sum_{i < j}^n V_{LJ}(r_i - r_j)$$

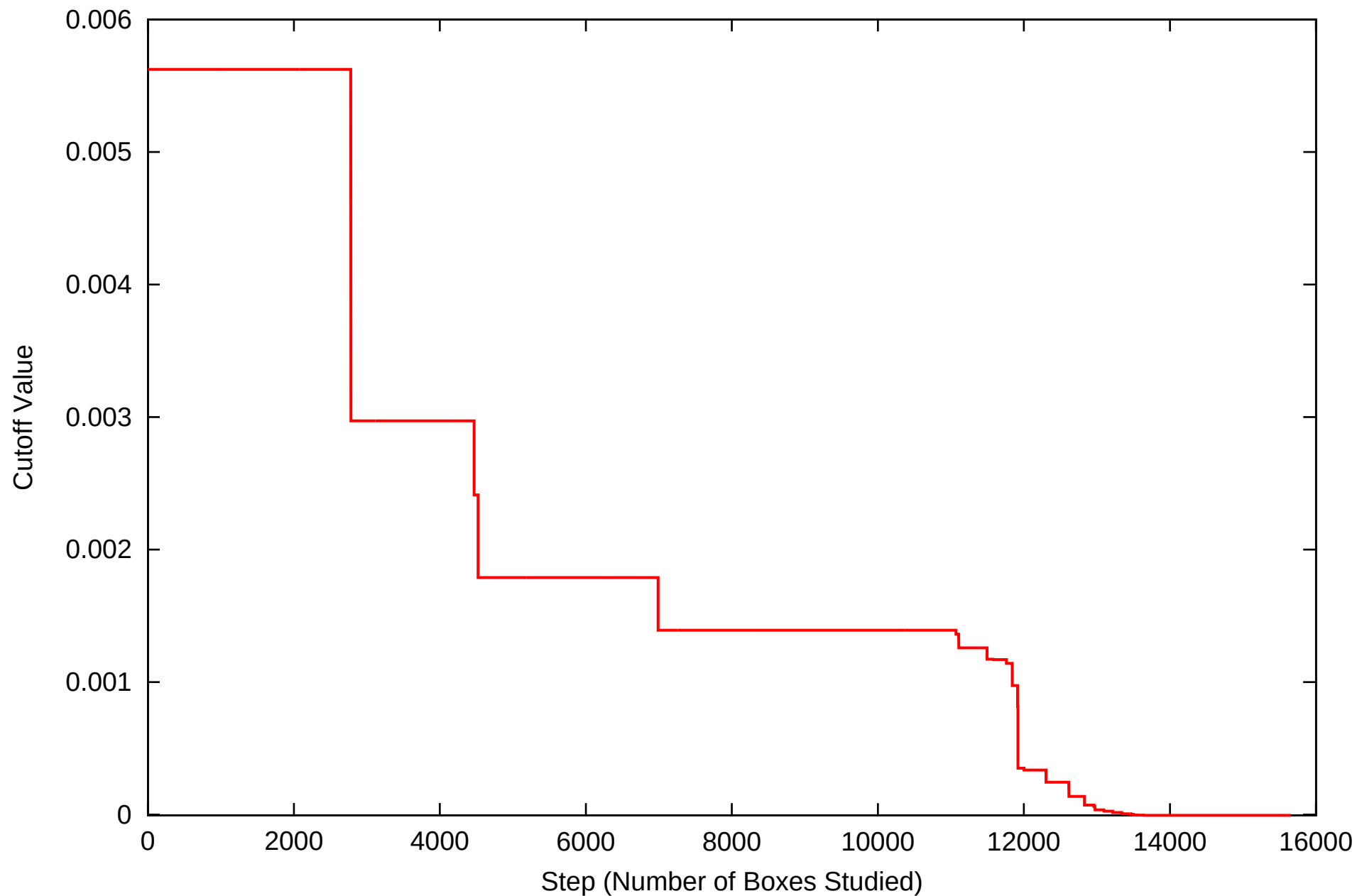
Find $\min f = \sum_{i < j}^n [V_{LJ}(r_i - r_j) + 1]$.

Study $n = 3, 4, 5$. Pop quiz: What do resulting molecules look like?

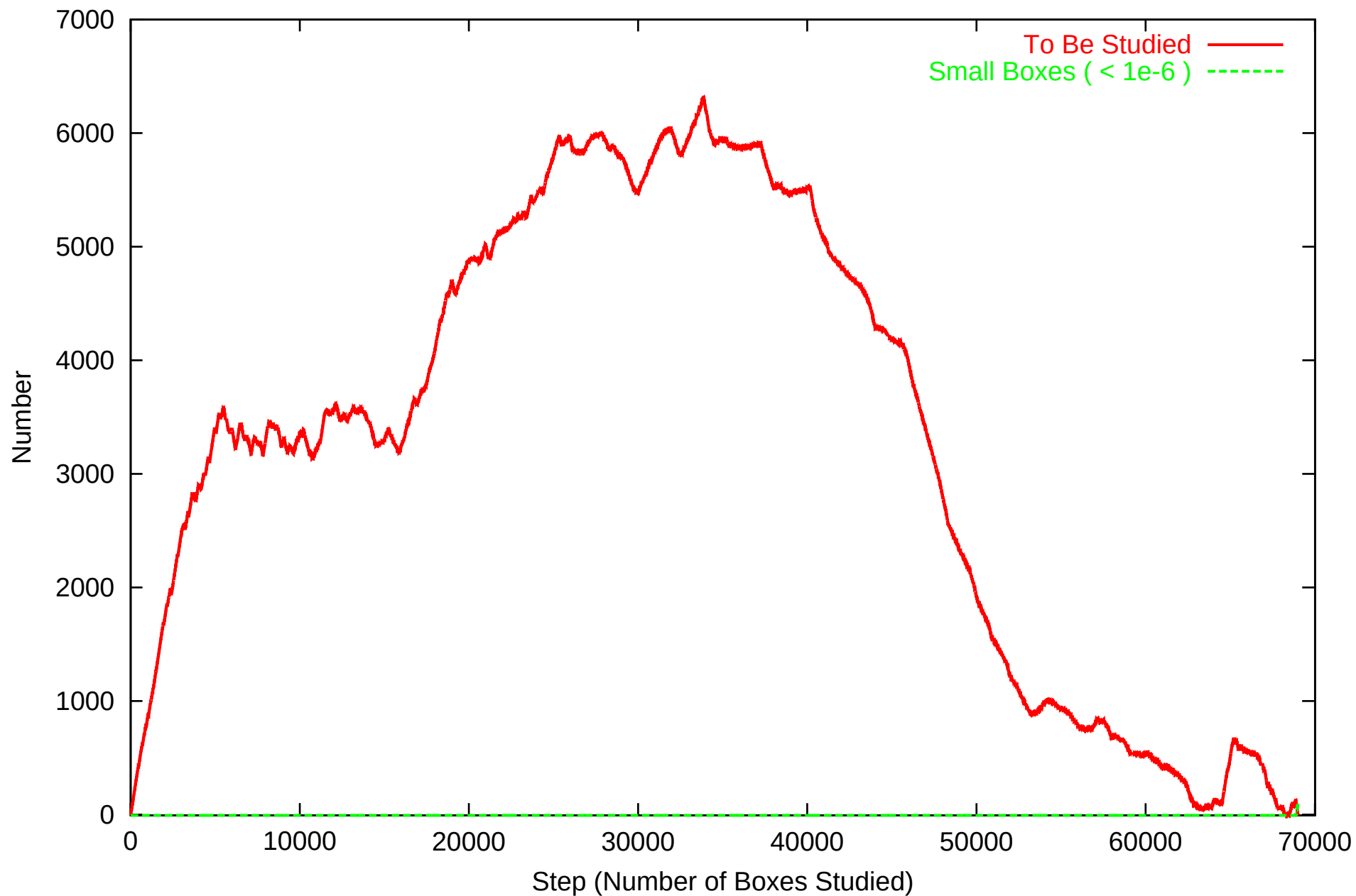
COSY-GO Lennard-Jones potential for 4 atoms: Number of Boxes -- LDB/QFB



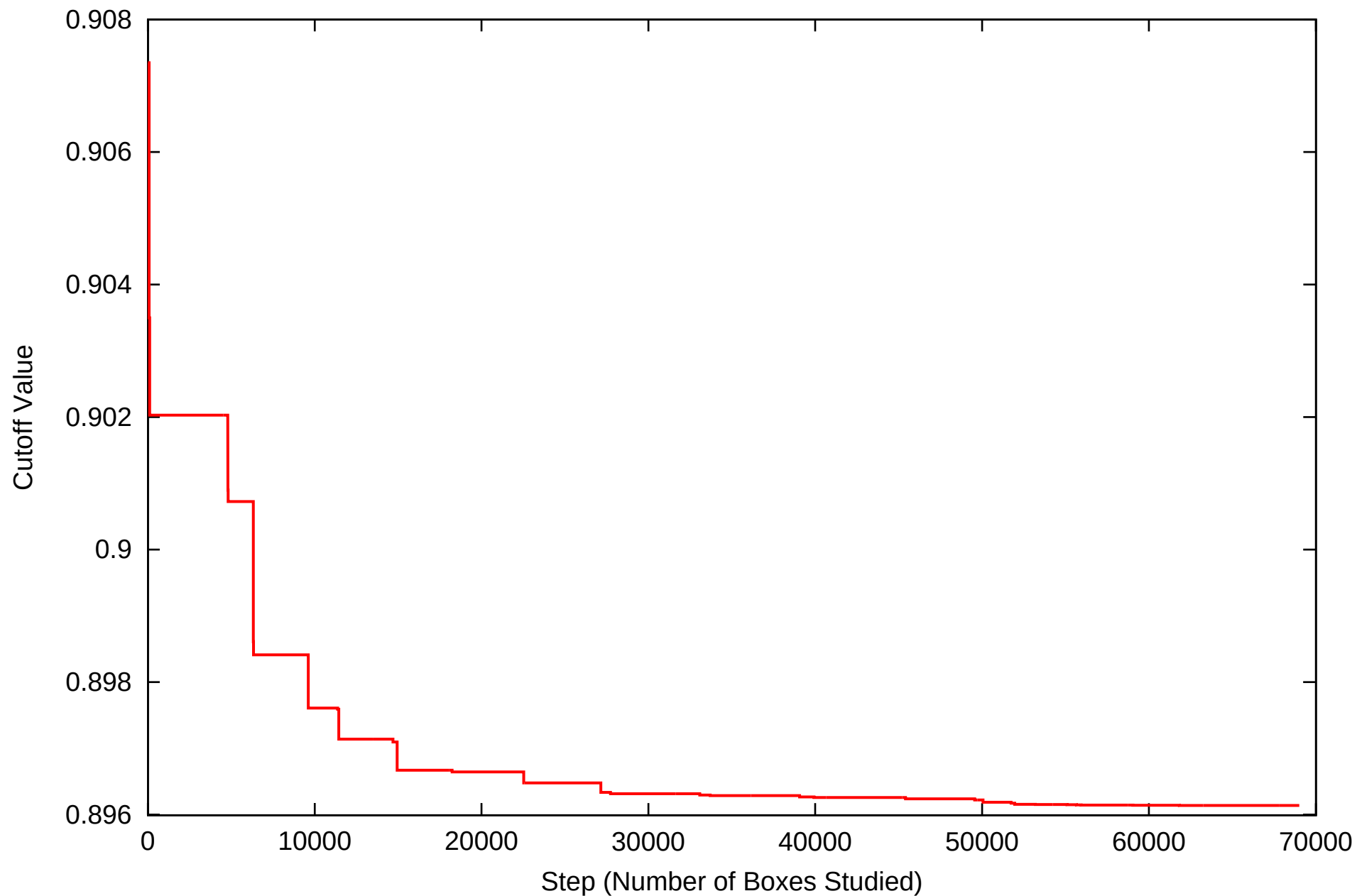
COSY-GO Lennard-Jones potential for 4 atoms: Cutoff Value -- LDB/QFB



COSY-GO Lennard-Jones potential for 5 atoms: Number of Boxes -- LDB/QFB



COSY-GO Lennard-Jones potential for 5 atoms: Cutoff Value -- LDB/QFB



Lennard-Jones Potentials - Results

Find minimum with COSY-GO and Globsol.

Use TMs of Order 5, QFB&LDB.

Use Globsol in default mode.

Problem	CPU-time needed	Max list	Total # of Boxes
n=4, COSY	89 sec	2,866	15,655
n=5, COSY	1,550 sec	6,321	69,001

n=4 atoms: 6D problem, n=5 atoms: 9D problem

Lennard-Jones Potentials - Results

Find minimum with COSY-GO and Globsol.

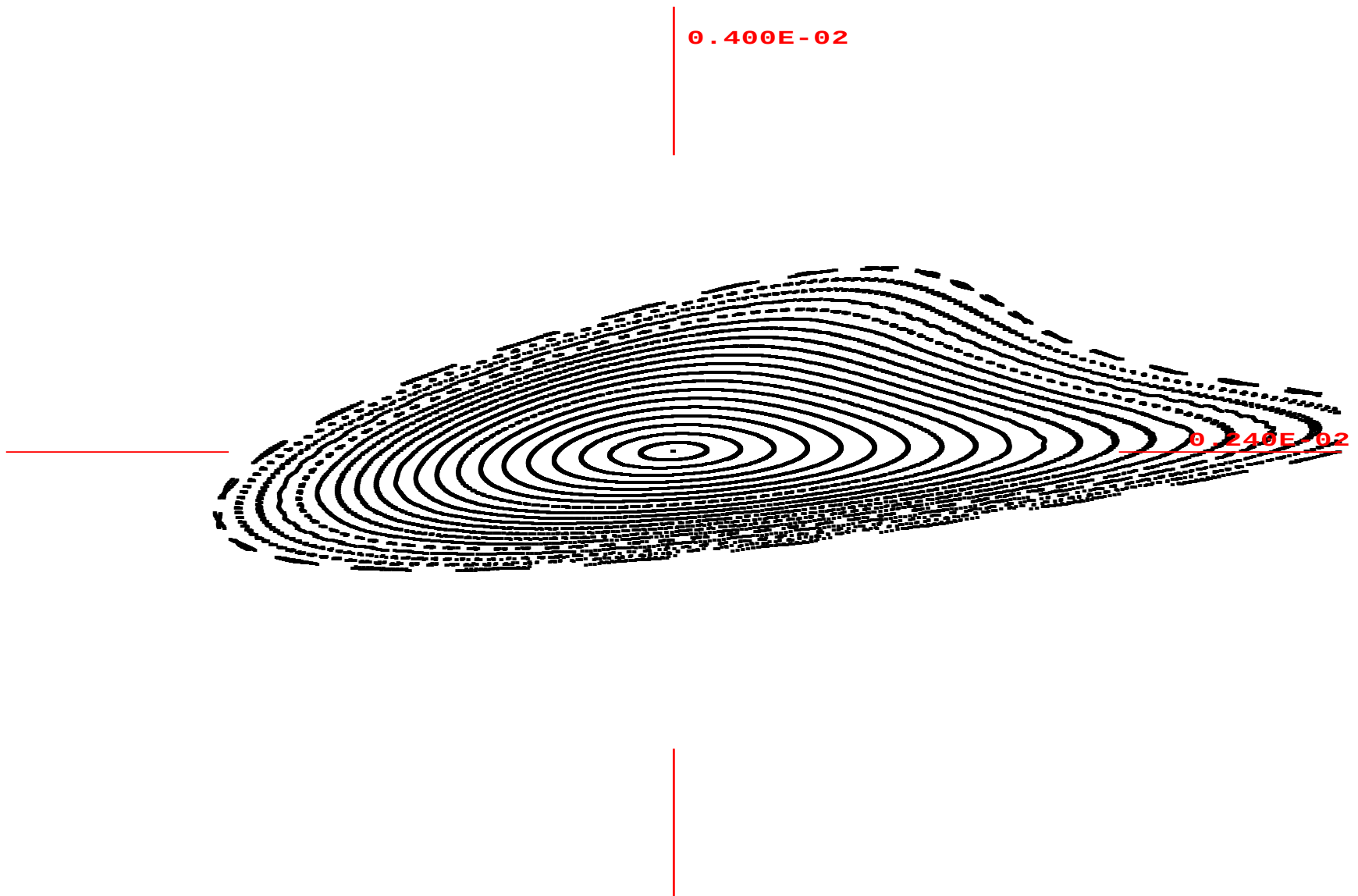
Use TMs of Order 5, QFB&LDB.

Use Globsol in default mode.

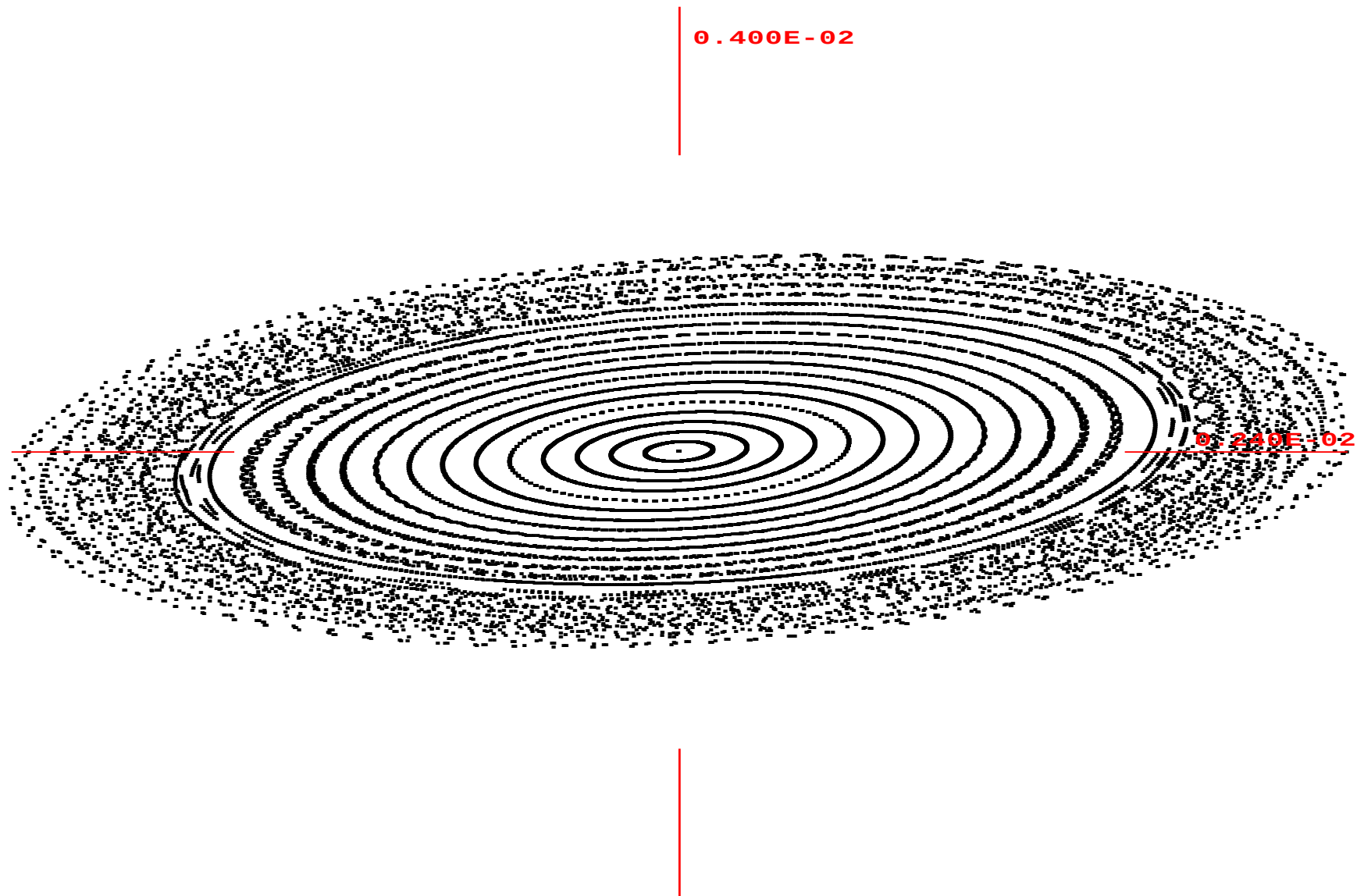
Problem	CPU-time needed	Max list	Total # of Boxes
n=4, COSY	89 sec	2,866	15,655
n=5, COSY	1,550 sec	6,321	69,001
n=4, Globsol	5,833 sec		243,911
n=5, Globsol	>259,200 sec (not finished yet)		

n=4 atoms: 6D problem, n=5 atoms: 9D problem



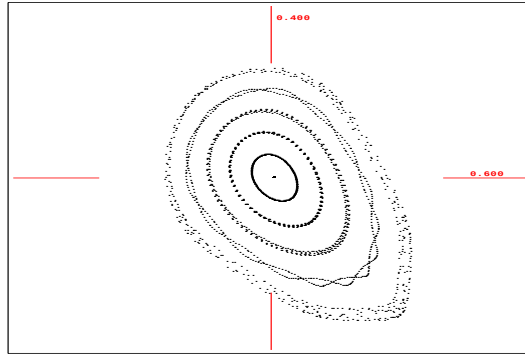


Tracking x-px Phase Space Motion of the Tevatron

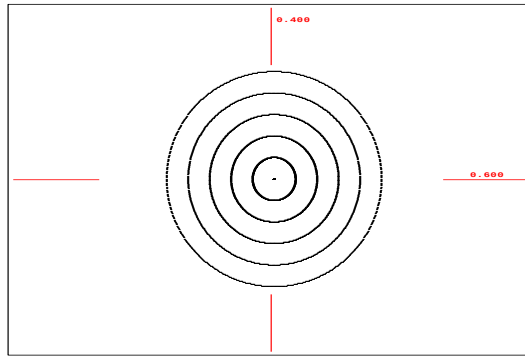


Tracking y - p_y Phase Space Motion of the Tevatron

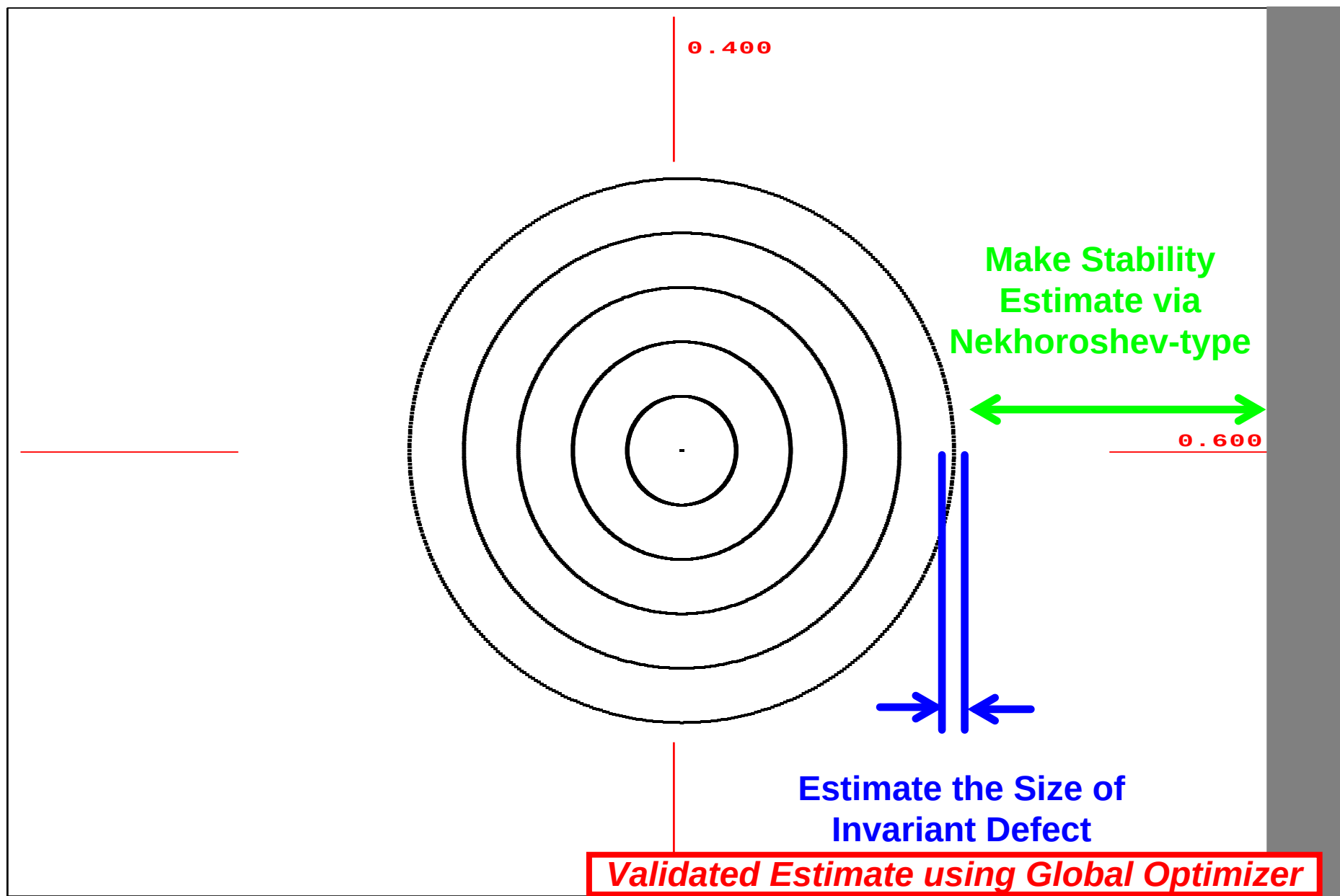
Example of Phase Space Motion



Tracking Phase Space Motion of 5 Particles in Regular Coordinates



Tracking Phase Space Motion of 5 Particles in Normal Form Coordinates



Tracking Phase Space Motion of 5 Particles in Normal Form Coordinates

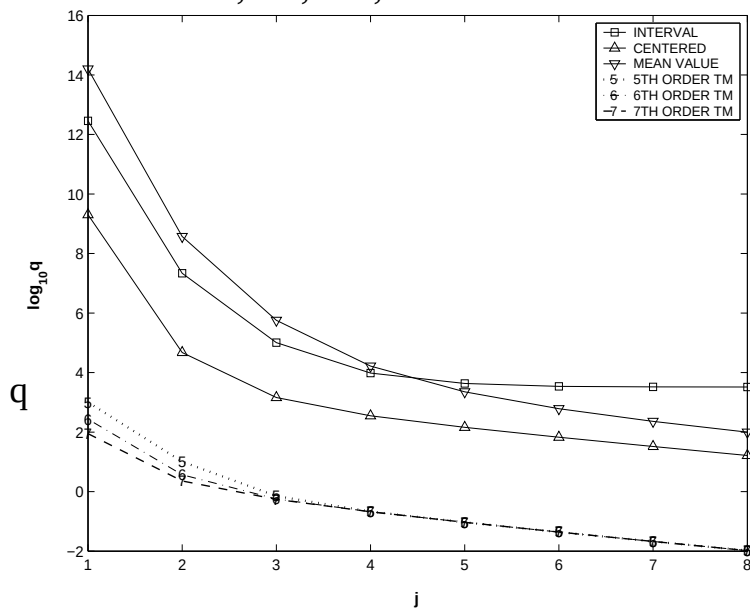
The Normal Form Invariant Defect Function

- **Extreme cancellation**; one of the reasons TM methods were invented
- Six-dimensional problem from dynamical systems theory
- Describes invariance defects of a particle accelerator
- Essentially composition of three tenth order polynomials
- The function vanishes identically to order ten
- Study for $a \cdot (1, 1, 1, 1, 1, 1)$ for $a = .1$ and $a = .2$
- Interesting **Speed observation**: on same machine,
 - * one CF in INTLAB takes 45 minutes (Version 3.1 under Matlab V.6)
 - * one TM of order 7 takes 10 seconds

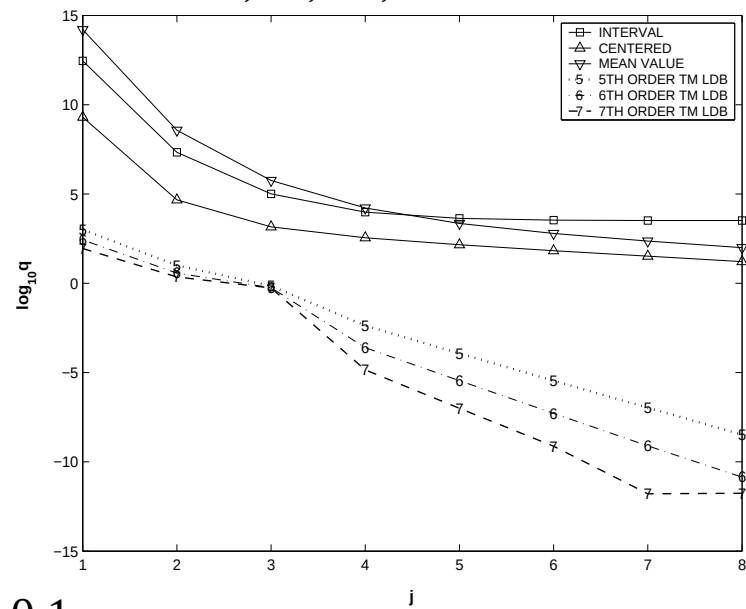
$$f(x_1, \dots, x_6) = \sum_{i=1}^3 \left(\sqrt{y_{2i-1}^2 + y_{2i}^2} - \sqrt{x_{2i-1}^2 + x_{2i}^2} \right)^2$$

where $\vec{y} = \vec{P}_1 \left(\vec{P}_2 \left(\vec{P}_3(\vec{x}) \right) \right)$

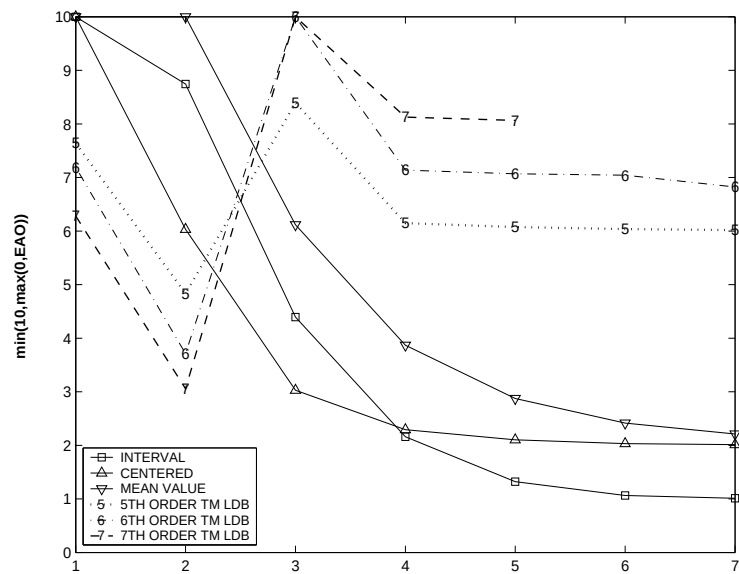
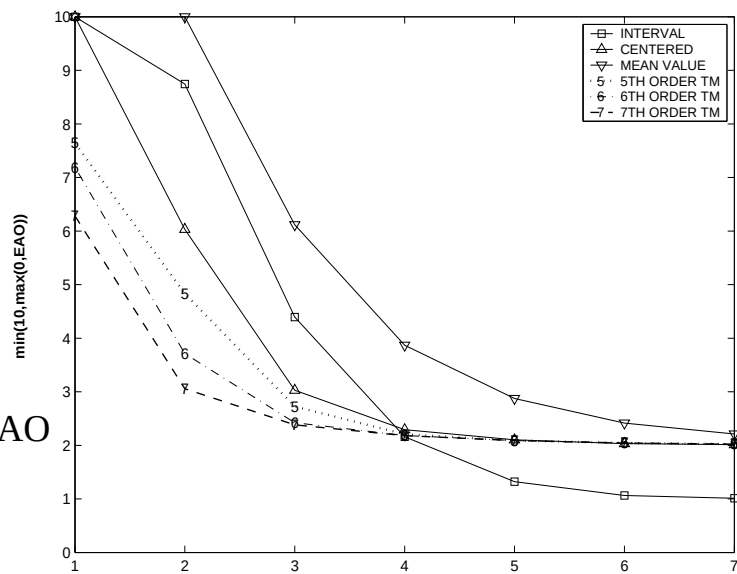
I, CF, MF, naive TM



I, CF, MF, TM-LDB

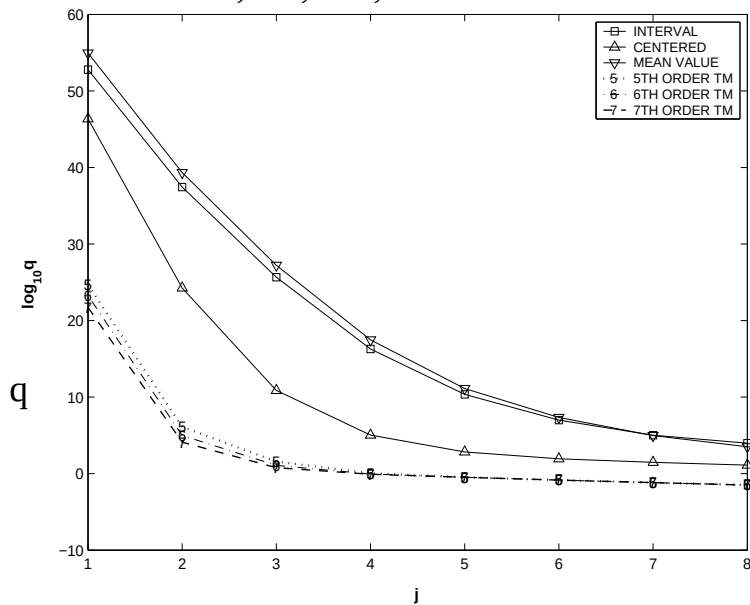


a=0.1

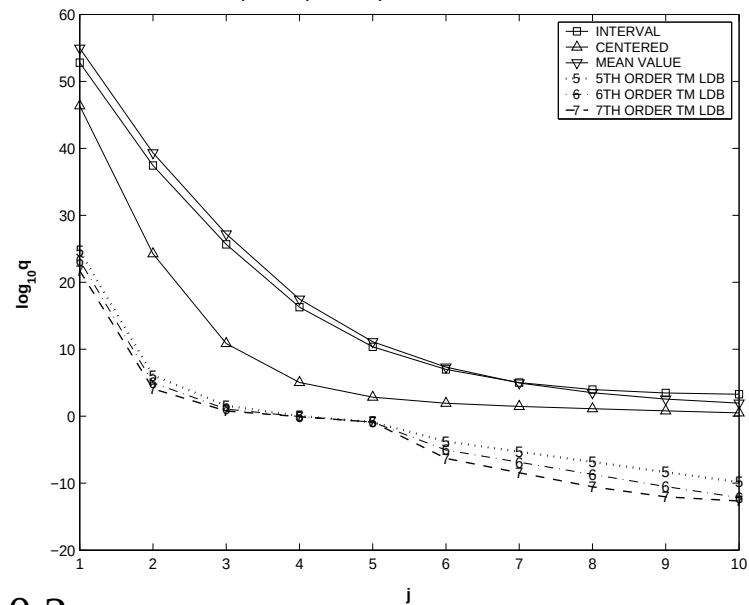


EAO

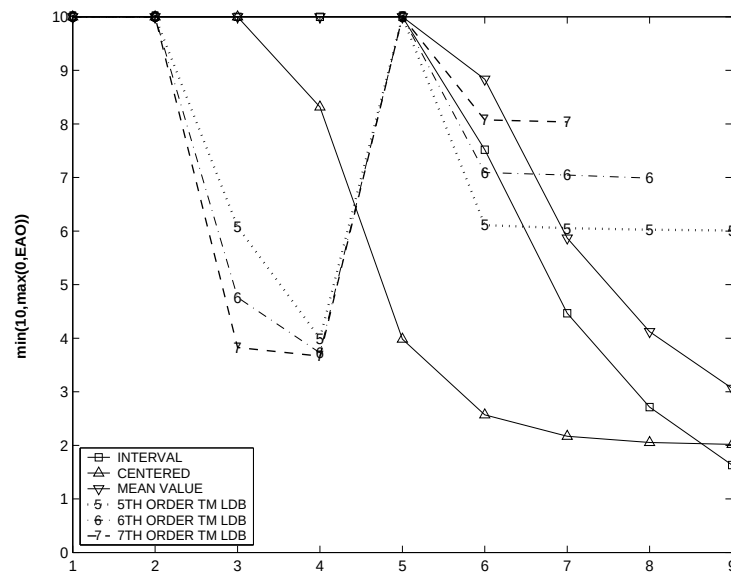
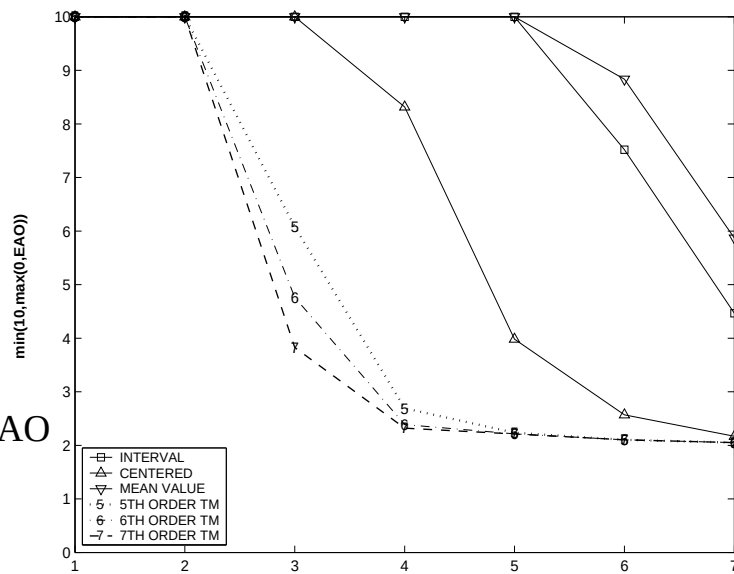
I, CF, MF, naive TM



I, CF, MF, TM-LDB



a=0.2



Normal form deviation function

$\times 10^{-5}$

2

1.5

1

0.5

0

-0.5

-1

-1.5

-2

6

4

2

0

0

1

2

3

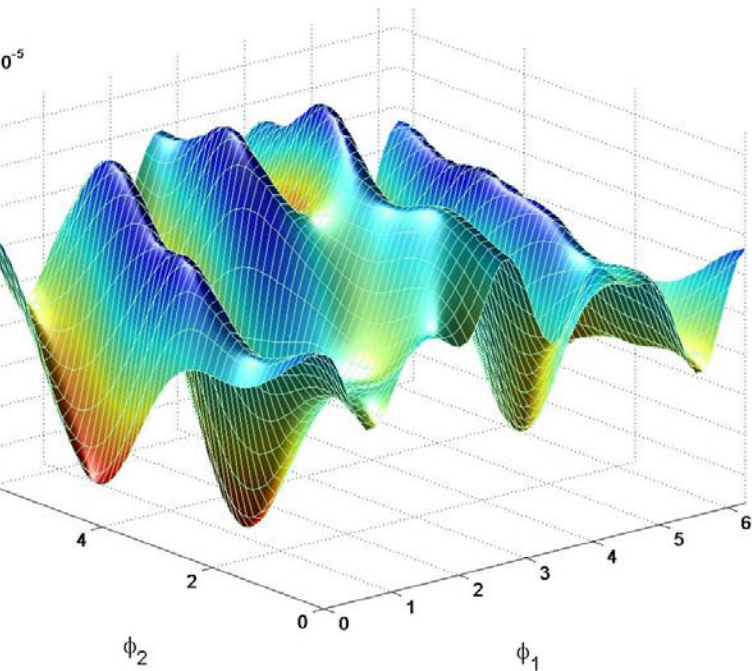
4

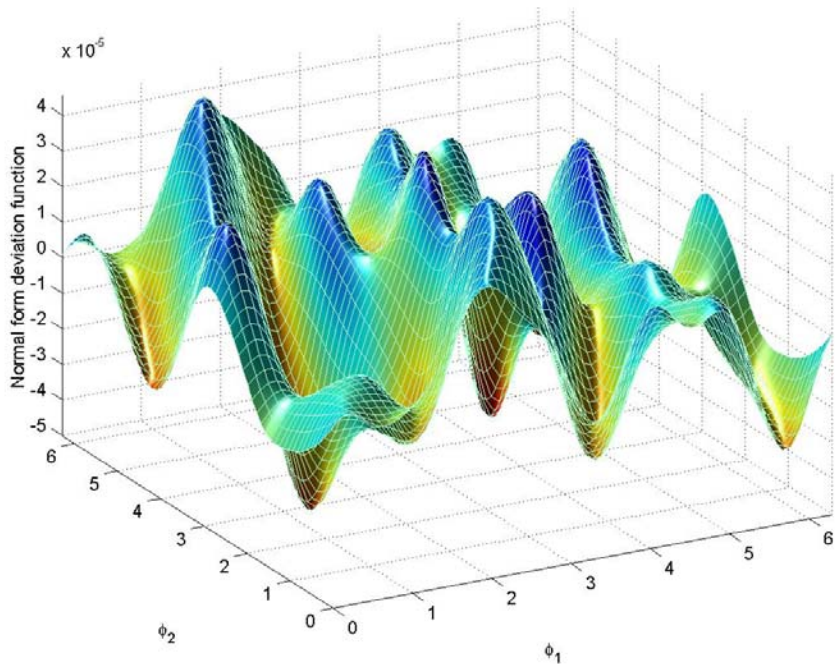
5

6

ϕ_2

ϕ_1





The Tevatron NF Invariant Defect Function

Estimate bound of the defect function over the Tevatron actual emittance (radius r) by **global optimization**. Make the stability estimate via Nekhoroshev-type for $2 \cdot r$.

The Tevatron NF Defect Function - GlobSol Results

For the computations, GlobSol's maximum list size was changed to 10^6 , and the CPU limit was set to 10 days. All other parameters affecting the performance of GlobSol were left at their default values.

Dimension	CPU-time needed	Max list	Total # of Boxes
2	18810 sec		4733
3	>562896 sec (not finished yet)		
4	>259200 sec (could not finish)		
5	> 86400 sec (could not finish)		
6	not attempted		

We observe that in this example, COSY outperforms GlobSol by many orders of magnitude. However, we are not completely sure if a different choice of parameters for GlobSol could result in better performance.

The Tevatron NF Defect Function - COSY-GO Results

Tolerance on the sharpness of the resulting minimum is 10^{-10} . For the evaluation of the objective function, Taylor models of order 5 were used. For the range bounding of the Taylor models, LDB with domain reduction was being used.

Dimension	CPU-time needed	Max list	Total # of Boxes
2	5.747071 sec	11	31
3	38.48828 sec	44	172
4	346.8604 sec	357	989
5	3970.746 sec	2248	6641
6	57841.94 sec	17241	49821

The Tevatron NF Invariant Defect Function

Estimate bound of the defect function over the Tevatron actual emittance (radius r) by **global optimization**. Make the stability estimate via Nekhoroshev-type for $2 \cdot r$.

- The result was very much limited by floating point floor.
- Can guarantee stability for 10^7 turns (emittance: $1.2\pi \cdot 10^{-4}$ mm mrad, normal form radius $R_{NF} = 10^{-5}$).

Conclusion

- Taylor models provide enclosures of functional dependencies by polynomials and a remainder that scales with $n + 1$ st order.
- Range bounding of polynomial is often easier than range bounding of the original function. Thus, the TM range bounding algorithms can lead to a high order method.
- For one dimensional systems, there are bounders up to the sixth order.
- The LDB bounding is cheap, and exact if monotonic. It can be used to assist various other methods. (QDB, Bernstein, ...)
- The QDB bounder provides the third order method.
- If the quadratic part is positive definite, a much faster quadratic bounder, the QFB bounder, can be used for a cutoff test, preventing the problem of cluster effect.
- Combined with various schemes for estimating cutoff values, we demonstrated the efficiency of the Taylor model global optimizer COSY-GO.