Advances in Interval Methods for Deterministic Global Optimization in Chemical Engineering

Youdong Lin and Mark A. Stadtherr

Department of Chemical and Biomolecular Engineering

University of Notre Dame

Notre Dame, IN USA

4th International Conference on Frontiers in Global Optimization

Santorini, Greece

June 8–12, 2003

Outline

- Background
- Interval Methodology
- Solving the Linear Interval Subproblem
- Examples
  - Parameter estimation problems
  - Problems with many local minima
  - Stationary states for triatomic molecules
- Concluding Remarks

Background

- One approach for doing global optimization is the use of interval analysis.
- Interval analysis can:
  - Provide a deterministic methodology for global optimization problems.
  - Deal automatically with rounding error, thus providing both mathematical and computational guarantees.
- Interval methods can be used in various ways in global optimization, e.g.:
  - Interval branch-and-bound

  $\Rightarrow$ Interval-Newton approach
  - As a tool within other methods

Background (Cont'd)

- Interval Newton/Generalized Bisection (IN/GB)
  - Given a system of equations to solve, an initial interval (bounds on all variables), and a solution tolerance:
  - IN/GB can find (enclose) with mathematical and computational certainty either all solutions or determine that no solutions exist.
  - IN/GB can also be extended and employed as a deterministic approach for global optimization problems.
- A general-purpose approach; in general requires no simplifying assumptions or problem reformulations.
- No strong assumptions about functions need to be made.

$\Rightarrow$ Solution of a linear interval equation system is a key subproblem.

Interval Methodology (Cont'd)

Problem: Solve $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ for all roots in interval $\mathbf{X}^{(0)}$.

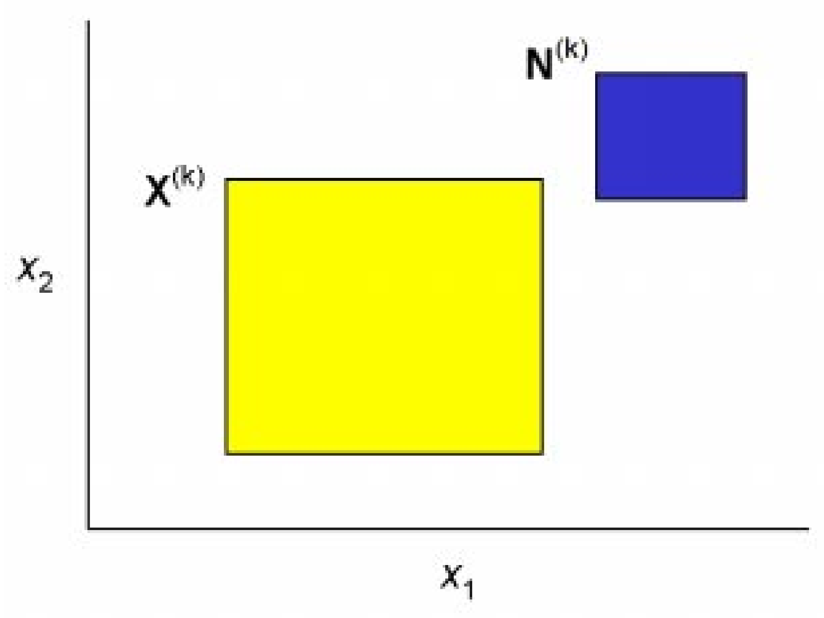
Basic iteration scheme (IN/GB): For a particular subinterval (box), $\mathbf{X}^{(k)}$, perform root inclusion test:

- (Range Test) Compute an interval extension (bounds on range) for each function in the system.
  - If 0 is not an element of any interval extension, delete the box. Otherwise,
- (Interval-Newton Test) Compute the *image*, $\mathbf{N}^{(k)}$, of the box by solving the linear interval equation system

$$\mathbf{F}'(\mathbf{X}^{(k)})(\mathbf{N}^{(k)} - \tilde{\mathbf{x}}^{(k)}) = -\mathbf{f}(\tilde{\mathbf{x}}^{(k)})$$

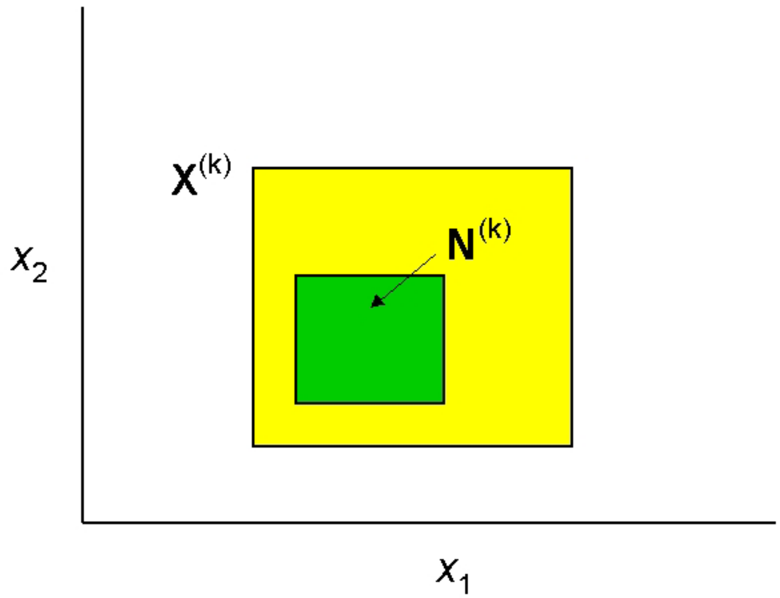
  - $\tilde{\mathbf{x}}^{(k)}$ is some point in $\mathbf{X}^{(k)}$.
  - $\mathbf{F}'\left(\mathbf{X}^{(k)}\right)$ is an interval extension of the Jacobian of $\mathbf{f}(\mathbf{x})$ over the box $\mathbf{X}^{(k)}$.

Interval Methodology (Cont'd)



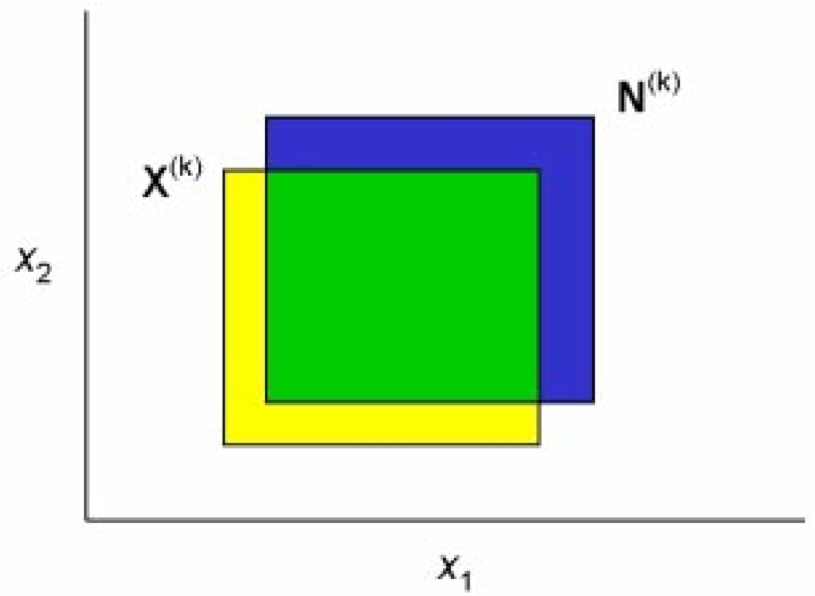
- There is no solution in $\mathbf{X}^{(k)}$.

Interval Methodology (Cont'd)



- There is a *unique* solution in $\mathbf{X}^{(k)}$.
- This solution is in $\mathbf{N}^{(k)}$.
- Additional interval-Newton steps will tightly enclose the solution with quadratic convergence. (Point Newton method will also converge to solution from any point in $\mathbf{N}^{(k)}$.)

Interval Methodology (Cont'd)



- Any solutions in $\mathbf{X}^{(k)}$ are in intersection of $\mathbf{X}^{(k)}$ and $\mathbf{N}^{(k)}$.
- If intersection is sufficiently small, repeat root inclusion test.
- Otherwise, bisect the intersection and apply root inclusion test to each resulting subinterval.

Interval Methodology (Cont'd)

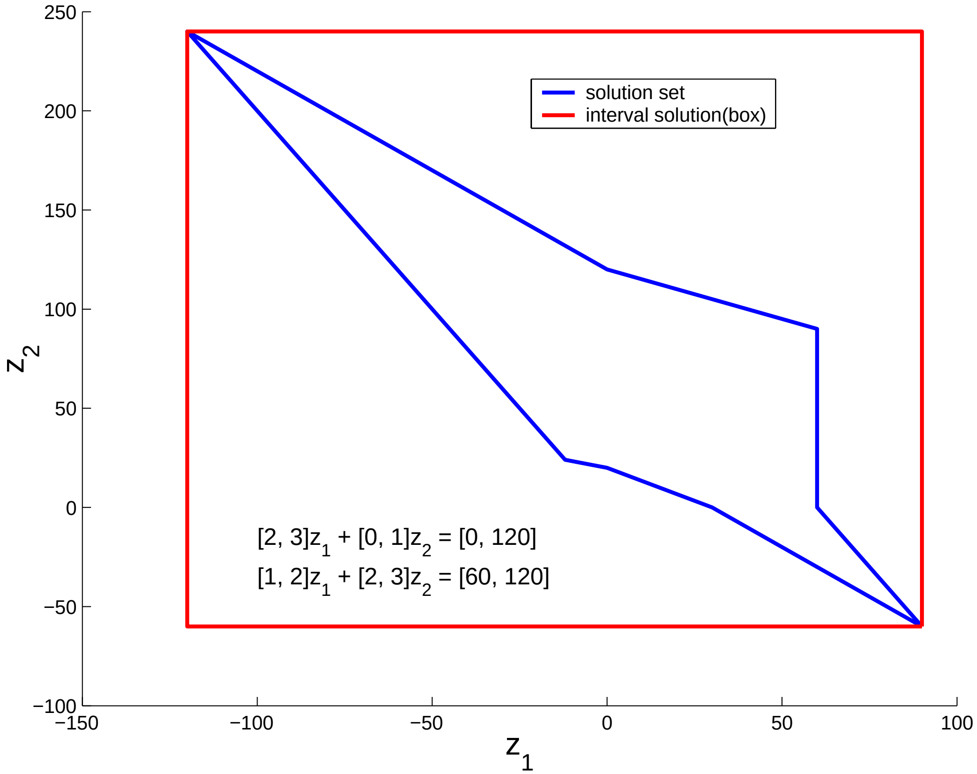
- Easily extended to global optimization problems.
- For unconstrained problems, solve for stationary points.
- For constrained problems, solve for KKT or Fritz-John points.
- Add an additional pruning condition (objective range test):
  - Compute interval extension of objective function.
  - If its lower bound is greater than a known upper bound on the global minimum, prune this subinterval.
- This combines IN/GB with a branch-and-bound scheme.
- Key step, for either optimization or equation solving, is solution of linear interval system

$$\mathbf{F}'(\mathbf{X})(\mathbf{N} - \tilde{\mathbf{x}}) = -\mathbf{f}(\tilde{\mathbf{x}})$$

  Seek tightest possible bounds on solution $(\mathbf{N} - \tilde{\mathbf{x}})$, and thus on $\mathbf{N}$.

Solution Set of Linear Interval System

- Consider linear interval system $\mathbf{A}\mathbf{z} = \mathbf{B}$.
- Solution set is defined: $\mathbf{S} = \{\mathbf{z} \mid \tilde{\mathbf{A}}\mathbf{z} = \mathbf{b}, \tilde{\mathbf{A}} \in \mathbf{A}, \mathbf{b} \in \mathbf{B}\}$.
- Interval solution: An interval $\mathbf{Z}$ containing $\mathbf{S}$.



Solution Set of Linear Interval System (Cont'd)

- Computing the interval hull (tightest interval containing $\mathbf{S}$) is NP-hard (Rohn and Kreinovich, 1995).
- Several methods are available to compute an interval solution $\mathbf{Z}$ that contains $\mathbf{S}$, but that may not give tight bounds.
- Methods used in the context of interval-Newton:
  - Preconditioned (inverse-midpoint) interval Gauss-Seidel
  - Hybrid (pivoting/inverse-midpoint) preconditioner and real point selection (HP/RP) (Gau and Stadtherr, 2002)

$\Rightarrow$ **LP strategy**

LP Strategy for Linear Interval System

- Oettli & Prager(1964) theorem : Solution set $\mathbf{S}$ is defined by the constraints

$$\left|\hat{\mathbf{A}}\mathbf{z} - \hat{\mathbf{B}}\right| \leq \Delta\mathbf{A}\,|\mathbf{z}| + \Delta\mathbf{B}$$

  $\hat{\mathbf{A}}$ – component-wise midpoint matrix of $\mathbf{A}$

  $\Delta\mathbf{A}$– component-wise half width matrix of $\mathbf{A}$

  $\hat{\mathbf{B}}$ – component-wise midpoint vector of $\mathbf{B}$

  $\Delta\mathbf{B}$ – component-wise half width vector of $\mathbf{B}$

- To eliminate absolute value operation on $\mathbf{z}$, the components of $\mathbf{z}$ must keep a constant sign $\longrightarrow$ consider each orthant separately.

LP Strategy for Linear Interval System (Cont'd)

- In each orthant, define D_α, a diagonal matrix whose entries are:

$$(D_\alpha)_{jj} = \begin{cases} 1 & \mathbf{z}_j \geq 0 \\ -1 & \mathbf{z}_j < 0 \end{cases} \quad j = 1, 2, \ldots, n$$

- To determine bounds on $\mathbf{S}$ in each orthant, solve $2n$ linear programming problems:

$$\text{maximize (and minimize)} \quad \mathbf{z}_j, \; j = 1, 2, \ldots, n$$
$$\text{s.t.} \quad \begin{pmatrix} \hat{\mathbf{A}} - \Delta\mathbf{A} D_\alpha \\ -\hat{\mathbf{A}} - \Delta\mathbf{A} D_\alpha \end{pmatrix} \mathbf{z} \leq \begin{pmatrix} \overline{\mathbf{B}} \\ -\underline{\mathbf{B}} \end{pmatrix}$$

- To get optimal solution overall (interval hull), calculate extrema in all orthants (2^n in worst scenario — exponential complexity).

LP Strategy for Linear Interval System (Cont'd)

Application to IN/GB methods:

- Solve linear interval system

$$\mathbf{F}'(\mathbf{X})(\mathbf{N} - \tilde{\mathbf{x}}) = -\mathbf{f}(\tilde{\mathbf{x}})$$

- Only the part of $\mathbf{N}$ that intersects $\mathbf{X}$ needs to be found.
- If $\tilde{\mathbf{x}}$ is selected to be a corner of $\mathbf{X}$, then the part of $\mathbf{N} - \tilde{\mathbf{x}}$ for which $\mathbf{N}$ lies in $\mathbf{X}$ is entirely in one orthant.
- Solution of interval-Newton equation can be sought using LP in only one orthant. Tightest possible solution obtained, while avoiding exponential time complexity.

Numerical Experiments

- LISS_LP(Linear Interval System Solver by Linear Programming) has been developed.
- Option to use sparse linear algebra in solution of LP problem.
- We compare performance results of LISS_LP to HP/RP (Gau and Stadtherr, 2002) on a SUN Blade 1000 model 1600 workstation.
- Performance results include:
  - Number of interval Newton tests performed (I-N tests)
  - CPU time in seconds

Example 1

- Estimation of Van Laar parameters from vapor-liquid equilibrium data using error-in-variables approach (Kim et al., 1990; Esposito and Floudas, 1998):

$$P = \gamma_1 x_1 p_1^0(T) + \gamma_2 (1 - x_1) p_2^0(T)$$

$$y_1 = \frac{\gamma_1 x_1 p_1^0(T)}{\gamma_1 x_1 p_1^0(T) + \gamma_2 (1 - x_1) p_2^0(T)}$$

where

$$p_1^0(T) = \exp\left[18.5875 - \frac{3626.55}{T - 34.29}\right], \quad p_2^0(T) = \exp\left[16.1764 - \frac{2927.17}{T - 50.22}\right]$$

and

$$\gamma_1 = \exp\left[\frac{A}{RT}\left(1 + \frac{A}{B}\frac{x_1}{1 - x_1}\right)^{-2}\right], \quad \gamma_2 = \exp\left[\frac{B}{RT}\left(1 + \frac{B}{A}\frac{1 - x_1}{x_1}\right)^{-2}\right]$$

- There are five data points and four measured variables with two parameters to be determined.

Example 1 (Cont'd)

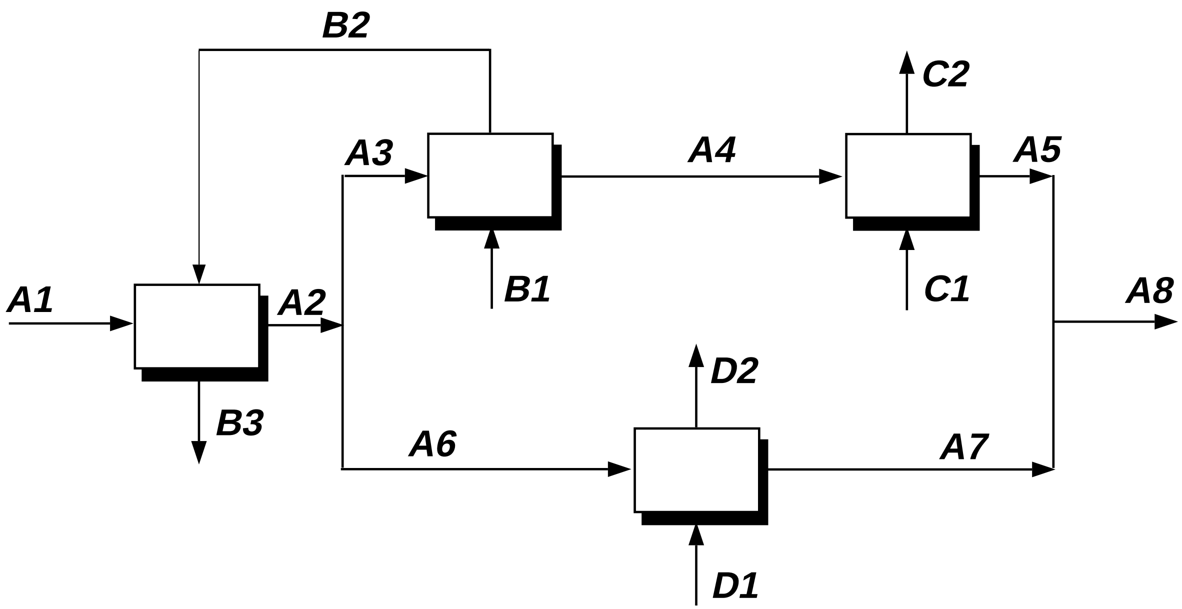
- Formulated as unconstrained global optimization with 2 parameter variables and 10 state variables.
- With standard inverse-midpoint preconditioner in solution of linear interval system, solution time is > 2 CPU days.
- Performance results with hybrid preconditioner (HP/RP) and LISS_LP:

	HP/RP	LISS_LP
I-N tests	303,589	156,182
CPU time (s)	664.4	496.7

- LP solver uses dense linear algebra.

Example 2

- Estimation of parameters in heat exchanger network using error-in-variables approach (Biegler and Tjoa, 1993).
- Network of four exchangers. Estimate the four rating parameters $(UA)_i$.
- Five unconstrained global optimization problems with 4 parameter variables and $13m$ state variables (number of data points $m = 4, 8, 12, 16, 20$).



Example 2 (Cont'd)

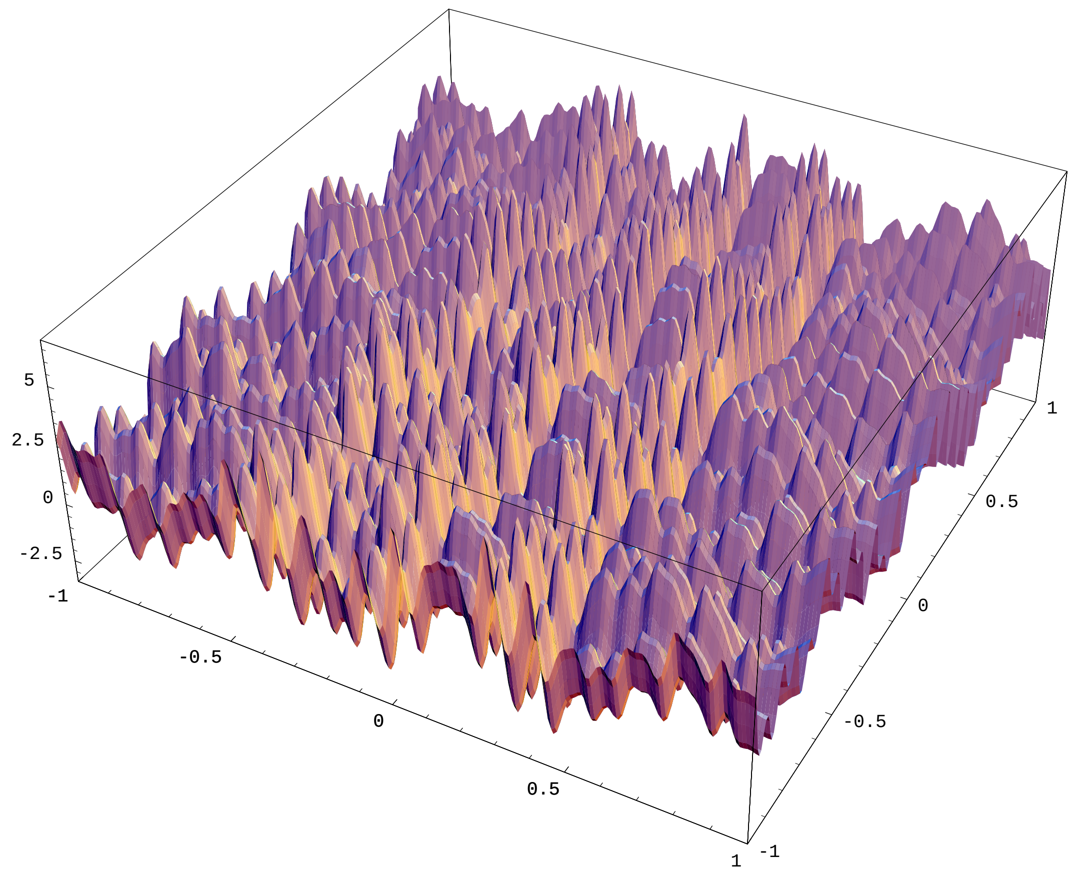
- Performance results:

Data Points	Variables	HP/RP		LISS_LP	
m	n	I-N Tests	CPU Time	I-N Tests	CPU Time
4	56	1	0.12	2	0.27
8	108	375	211.8	44	38.1
12	160	363	498.6	299	346.0
16	212	188	645.8	83	316.8
20	264	220	1357.3	81	504.9

- LP solver uses sparse linear algebra.

Example 3

- Trefethen (2002) Challenge Problem #4 — Find the Global Minimum



$$f(x,y) = \exp(\sin(50x)) + \sin(60\exp(y)) + \sin(70\sin(x)) + \sin(\sin(80y)) - \sin(10(x+y)) + (x^2+y^2)/4; \qquad x \in [-1,1]; \qquad y \in [-1,1]$$

Example 3 (Cont'd)

- Solution

$$x \in [-0.02440307969437517, -0.02440307969437516]$$

$$y \in [0.2106124271553557, 0.2106124271553558]$$

$$f \in [-3.306868647475245, -3.306868647475232]$$

- Global minimum is easily found using interval approach

	HP	LISS_LP
I-N tests	1814	1179
CPU time (s)	0.15	0.16

- On relatively easy problems, LP-based strategy is not needed, but still can be used without significant loss of efficiency due to LP overhead.

Example 4

- Find the global minimum of the function (Siirola et al., 2002):

$$f(\mathbf{x}) = 100 \prod_{i=1}^{N} \sum_{j=1}^{5} \left(\frac{j^5}{4425} \cos(j + j x_i) \right) + \frac{1}{N} \sum_{i=1}^{N} (x_i - x_{0,i})^2$$

where $x_{0,i} = 3, \quad x_i \in [x_{0,i} - 20, x_{0,i} + 20], \quad i = 1, ..., N.$

- Solve for $N = 5$. There are $\approx 10^8$ local optima.

- Performance results:

	Global Minima	I-N tests	CPU time(s)
LISS_LP	5	155,666	389.17
HP	5	171,918	636.76

Example 5

- Find all stationary points (minima, maxima, saddles) on potential energy surface of triatomic molecule ABC.
  - Useful for study of transition states and reaction pathways.
  - We studied the molecules HCN, HSiN, CS_2 and HBO (Westerberg and Floudas, 1999).
- The triatomic molecule geometry is described using the three interatomic distances $R_1 = R_{AB}$, $R_2 = R_{AC}$ and $R_3 = R_{BC}$.
- Murrel-Sorbie analytic potential energy surface $V(R_1, R_2, R_3)$ is used.
- Find all solutions of stationarity condition: $\nabla V = 0$

Example 5 (Cont'd)

- Physical constraints: triangle inequality

$$R_{AB} \leq R_{AC} + R_{BC}$$
$$R_{AC} \leq R_{AB} + R_{BC}$$
$$R_{BC} \leq R_{AB} + R_{AC}$$

Use to tighten interval bounds before application of interval-Newton.

- Collinear case: eliminate one variable, e.g., for $R_{AC} = R_{AB} + R_{BC}$ solve

$$\frac{\partial}{\partial R_{AB}} V(R_{AB}, R_{AB} + R_{BC}, R_{BC}) = 0,$$
$$\frac{\partial}{\partial R_{BC}} V(R_{AB}, R_{AB} + R_{BC}, R_{BC}) = 0$$

- Solve one noncollinear problem and three collinear problems.

- Search in intervals $R_i \in [0.7, 5.0]\text{Å}$.

Example 5 (Cont'd)

- General form of potential energy surface (Aguilar et al., 1992)

$$V(R_1, R_2, R_3) = V_{AB} + V_{AC} + V_{BC} + V_{ABC}$$

- Two-body terms are extended Rydberg functions, e.g.,

$$V_{AB}(R_1) = -D_e(1 + a_1\rho + a_2\rho^2 + a_3\rho^3 + \cdots)e^{-a_1\rho},$$
$$\rho = R_1 - R_e$$

- Three-body term

$$V_{ABC}(R_1, R_2, R_3) = P \times T,$$

$$P = V^0\left(1 + \sum_{i=1}^{3} C_i\rho_i + \sum_{j \geq i=1}^{3} C_{ij}\rho_i\rho_j + \sum_{k \geq j \geq i=1}^{3} C_{ijk}\rho_i\rho_j\rho_k + \cdots\right)$$

$$T = \prod_{i=1}^{3}\left(1 - \tanh\frac{\gamma_i S_i}{2}\right)$$

$$S_i = \sum_{j=1}^{3} b_{ij}\rho_j, \quad \rho_j = R_j - R_j^0$$

Example 5 (Cont'd)

- Stationary states for HCN

Type	Energy (eV)	$R_{CN}(\mathring{A})$	$R_{CH}(\mathring{A})$	$R_{NH}(\mathring{A})$
minimum	-5.548223	—	2.332871	1.038900
saddle	8.094668	—	0.857572	0.806900
minimum	-12.972507	1.159150	—	0.993336
minimum	-13.592215	1.153198	1.065498	—
saddle	-5.249952	2.344235	2.980408	1.044278
saddle	-1.937592	2.311895	1.792854	2.327696
saddle	-3.102483	2.582864	1.081559	2.737335
saddle	-11.444169	1.117973	1.053919	1.387750
saddle	-11.345398	0.929065	1.039138	1.041348
minimum	-11.379410	0.857321	0.980845	0.989052

Example 5 (Cont'd)

- Stationary states for HSiN

Type	Energy (eV)	$R_{SiN}\,(\mathring{A})$	$R_{SiH}\,(\mathring{A})$	$R_{NH}\,(\mathring{A})$
saddle	1.109601	2.778074	2.617596	—
saddle	-3.144738	1.523964	2.426268	—
saddle	-5.148745	2.006322	1.361586	—
minimum	-6.098598	1.529588	1.459586	—
saddle	-5.666608	1.575921	—	2.969229
minimum	-9.358509	1.523293	—	0.998205
maximum	1.720515	2.647092	2.415995	3.498876
saddle	-2.876954	1.501907	2.309780	3.069649
saddle	-4.908995	2.394221	2.137984	0.974496
saddle	-0.728138	2.155741	1.473092	2.044809
saddle	-3.717494	1.461352	1.634575	2.093708

Example 5 (Cont'd)

- Stationary states for CS_2

Type	Energy (eV)	R_{CS} (Å)	$R_{CS'}$ (Å)	$R_{SS'}$ (Å)
saddle	-1.668827	2.761779	—	2.695109
saddle	103.740892	0.949956	—	1.813411
minimum	97.485407	0.909824	—	1.417728
minimum	-12.004548	1.552422	1.552422	—
saddle	-0.049002	4.171034	4.171034	3.978688

Example 5 (Cont'd)

- Stationary states for HBO(PES1)

Type	Energy (eV)	R_{BH} (Å)	R_{BO} (Å)	R_{OH} (Å)
saddle	-7.598281	3.264082	1.187662	—
minimum	-16.678316	1.165505	1.185028	—
minimum	-6.556670	1.162756	—	2.349430
saddle	-0.216647	—	2.554092	3.688901

- Stationary states for HBO(PES2)

Type	Energy (eV)	R_{BH} (Å)	R_{BO} (Å)	R_{OH} (Å)
minimum	-16.678851	1.168947	1.184167	—
minimum	-6.639249	1.154136	—	2.344208
minimum	-11.305022	—	1.192047	2.383483
saddle	-11.134196	2.906575	1.185979	2.398381

Example 5 (Cont'd)

- Summary of triatomic problems

Problem	Stationary points found	CPU time (sec)
HCN	10	6.66
HSiN	11	1.07
CS_2	5	2.18
HBO(PES1)	4	0.88
HBO(PES2)	4	0.56

Concluding Remarks

- An LP-based method can be used to solve the linear interval system arising in the context of the interval-Newton approach for nonlinear equation solving and global optimization.
- The method can obtain tighter bounds on the solution set than standard methods, and thus lead to a large reduction in the number of subintervals that must be tested during the interval-Newton procedure.
- The overhead required to solve the LP subproblems may lead to relatively smaller improvements in overall computation time.
- The interval methodology is a powerful approach for deterministic global optimization.

Acknowledgments

- ACS Petroleum Research Fund 35979-AC9
- Indiana 21st Century Research & Technology Fund
- NSF DMR 00-79647