Reliable Process Modeling and Optimization Using Interval Analysis

Chao-Yang Gau, Robert W. Maier, Gang Xu and Mark A. Stadtherr[1]

Department of Chemical Engineering
University of Notre Dame
Notre Dame, IN 46556 USA

6th SIAM Conference on Optimization
Atlanta, GA, May 10–12, 1999

[1]Author to whom all correspondence should be addressed. Phone: (219)631-9318; Fax: (219)631-8366; E-mail: markst@nd.edu

Outline

- Motivation
- Background
- Examples: Parameter Estimation in Modeling Vapor-Liquid Equilibrium (VLE)
- Results
- Summary and Concluding Remarks

Motivation

- In process modeling, chemical engineers frequently need to solve nonlinear equation systems in which the variables are constrained physically within upper and lower bounds; that is, to solve:

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}$$

$$\mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U$$

- These problems may:
  - Have multiple solutions
  - Have no solution
  - Be difficult to converge to any solution

Motivation (continued)

- There is also frequent interest in globally minimizing a nonlinear function subject to nonlinear equality and/or inequality constraints; that is, to solve (globally):

$$\min_{\mathbf{x}} \phi(\mathbf{x})$$

subject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

$$\mathbf{g}(\mathbf{x}) \geq \mathbf{0}$$

$$\mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U$$

- These problems may:
  - Have multiple local minima (in some cases, it may be desirable to find them all)
  - Have no solution (infeasible NLP)
  - Be difficult to converge to any local minima

Motivation (continued)

- One approach for dealing with these issues is *interval analysis*.

- Interval analysis can
  - Provide the engineer with tools needed to solve modeling and optimization problems with complete certainty.
  - Provide problem-solving reliability not available when using standard local methods.
  - Deal automatically with rounding error, thus providing both mathematical and computational guarantees.

Motivation (continued)

- At Notre Dame, we have successfully applied interval methods for
  - General process modeling problems (Schnepper and Stadtherr, 1996).
  - Phase stability and equilibrium problems using several different thermodynamic models (Stadtherr et al., 1994; Hua et al., 1996,1998,1999; Xu et al., 1998).
  - Computation of azeotropes of multicomponent mixtures (Maier et al., 1998,1999).
  - Computation of mixture critical points (Stradi et al., 1998)

$\Longrightarrow$ Parameter estimation in vapor-liquid equilibrium models (Gau and Stadtherr, 1998,1999).

Background—Interval Analysis

- A real interval $X = [a, b] = \{x \in \Re \mid a \leq x \leq b\}$ is a segment on the real number line and an interval vector $\mathbf{X} = (X_1, X_2, ..., X_n)^T$ is an n-dimensional rectangle or "box".

- Basic interval arithmetic for $X = [a, b]$ and $Y = [c, d]$ is X op $Y = \{x \text{ op } y \mid x \in X,\ y \in Y\}$ where op $\in \{+, -, \times, \div\}$. For example, $X + Y = [a + c, b + d]$.

- Computed endpoints are **rounded out** to guarantee the enclosure.

- Interval elementary functions (e.g. $\exp(X)$, $\log(X)$, etc.) are also available.

- The interval extension $F(\mathbf{X})$ encloses all values of $f(\mathbf{x})$ for $\mathbf{x} \in \mathbf{X}$. That is, $F(\mathbf{X}) \supseteq \{f(\mathbf{x}) \mid \mathbf{x} \in \mathbf{X}\}$.

- Interval extensions can be computed using interval arithmetic (the "natural" interval extension), or with other techniques.

Background—Parameter Estimation

- Observations $y_{\mu i}$ of $i = 1, \ldots, q$ responses from $\mu = 1, \ldots, p$ experiments are available.

- Responses are to be fit to a model $y_{\mu i} = f_i(\mathbf{x}_\mu, \boldsymbol{\theta})$ with independent variables $\mathbf{x}_\mu = (x_{\mu 1}, \ldots, x_{\mu m})^T$ and parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)^T$. Measurement errors in $\mathbf{x}_\mu$ can either be neglected or treated using the "error-in-variable" approach.

- Various objective functions $\phi(\boldsymbol{\theta})$ can be used to determine the parameter values that provide the "best" fit, e.g.

  – Maximum likelihood
  – Relative or weighted least squares

  The latter will be used here.

- Optimization problem to determine parameters can usually be formulated as either a constrained or unconstrained problem. The unconstrained formulation is used here.

Parameter Estimation

- Assuming a relative least squares objective and using an unconstrained formulation, the problem is

$$\min_{\boldsymbol{\theta}} \phi(\boldsymbol{\theta}) = \sum_{i=1}^{q} \sum_{\mu=1}^{p} \left[\frac{y_{\mu i} - f_i(\mathbf{x}_\mu, \boldsymbol{\theta})}{y_{\mu i}} \right]^2$$

- A common approach for solving this problem is to use the gradient of $\phi(\boldsymbol{\theta})$ and to seek the stationary points of $\phi(\boldsymbol{\theta})$ by solving $\mathbf{g}(\boldsymbol{\theta}) \equiv \nabla\phi(\boldsymbol{\theta}) = \mathbf{0}$. This system may have many roots, including local minima, local maxima and saddle points.

- To insure that the global minimum of $\phi(\boldsymbol{\theta})$ is found, the capability to find **all** the roots of $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$ is needed. This is provided by the **interval Newton** technique.

- Interval Newton can be combined with **branch and bound** so that roots of $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$ that cannot be the global minimum need not be found.

Background–Interval Newton Method

- For the system of nonlinear equations $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$, find (enclose) all roots in a given initial interval $\boldsymbol{\Theta}^{(0)}$ or determine that there are none.

- At iteration k, given the interval $\boldsymbol{\Theta}^{(k)}$, if $0 \in \mathbf{G}(\boldsymbol{\Theta}^{(k)})$ solve the linear interval equation system

$$G'(\boldsymbol{\Theta}^{(k)})(\mathbf{N}^{(k)} - \boldsymbol{\theta}^{(k)}) = -\mathbf{g}(\boldsymbol{\theta}^{(k)})$$

  for the "image" $\mathbf{N}^{(k)}$, where $\mathbf{G}(\boldsymbol{\Theta}^{(k)})$ is an interval extension of $\mathbf{g}(\boldsymbol{\theta})$ and $G'(\boldsymbol{\Theta}^{(k)})$ an interval extension of its Jacobian over the current interval $\boldsymbol{\Theta}^{(k)}$, and $\boldsymbol{\theta}^{(k)}$ is a point inside $\boldsymbol{\Theta}^{(k)}$.

- Any root $\boldsymbol{\theta}^* \in \boldsymbol{\Theta}^{(k)}$ is also contained in the image $\mathbf{N}^{(k)}$, suggesting the iteration scheme $\boldsymbol{\Theta}^{(k+1)} = \boldsymbol{\Theta}^{(k)} \cap \mathbf{N}^{(k)}$ (Moore, 1966).

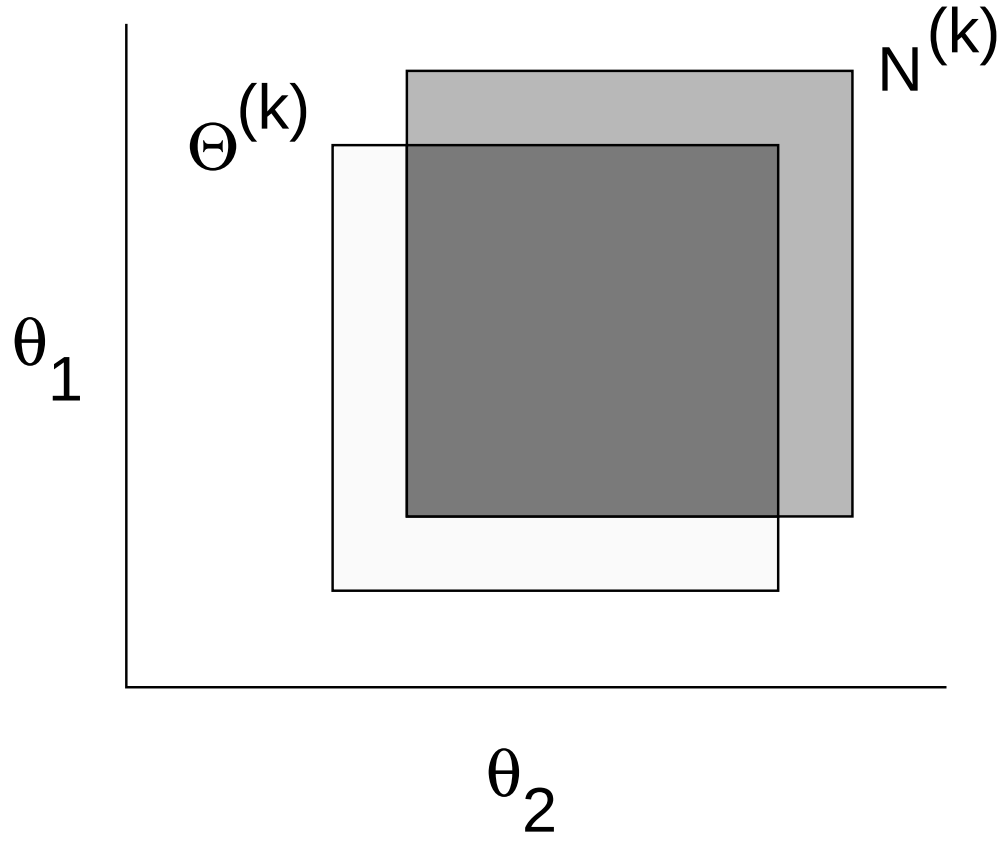
- It follows that if $\boldsymbol{\Theta}^{(k)} \cap \mathbf{N}^{(k)} = \emptyset$, then there is no root in $\boldsymbol{\Theta}^{(k)}$. This is also the conclusion if $0 \notin \mathbf{G}(\boldsymbol{\Theta}^{(k)})$.

Interval Newton Method (continued)

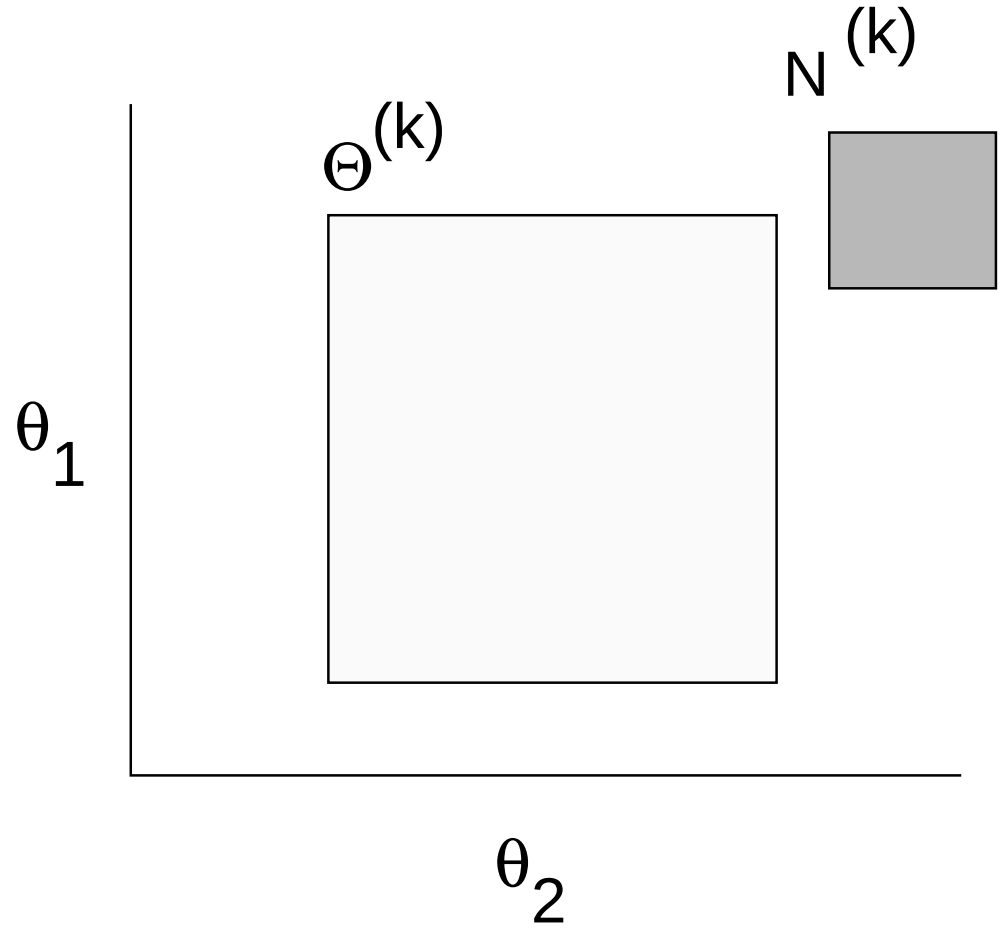
- Interval Newton provides an existence and uniqueness test: If $\mathbf{N}^{(k)} \subset \mathbf{\Theta}^{(k)}$, then:
  - There is a **unique** zero of $\mathrm{g}(\boldsymbol{\theta})$ in $\mathbf{\Theta}^{(k)}$.
  - The interval Newton iteration $\mathbf{\Theta}^{(k+1)} = \mathbf{\Theta}^{(k)} \cap \mathbf{N}^{(k)}$ will converge quadratically to a tight enclosure of the root.
  - The point Newton method will converge quadratically to the root starting from any point in $\mathbf{\Theta}^{(k)}$.

- If a unique root cannot be confirmed ($\mathbf{N}^{(k)} \subset \mathbf{\Theta}^{(k)}$) or ruled out ($\mathbf{\Theta}^{(k)} \cap \mathbf{N}^{(k)} = \emptyset$), then either:
  - Continue with the next iterate $\mathbf{\Theta}^{(k+1)}$ if it is sufficiently smaller than $\mathbf{N}^{(k)}$, or
  - **Bisect** $\mathbf{\Theta}^{(k+1)}$ and perform interval Newton on the resulting intervals.

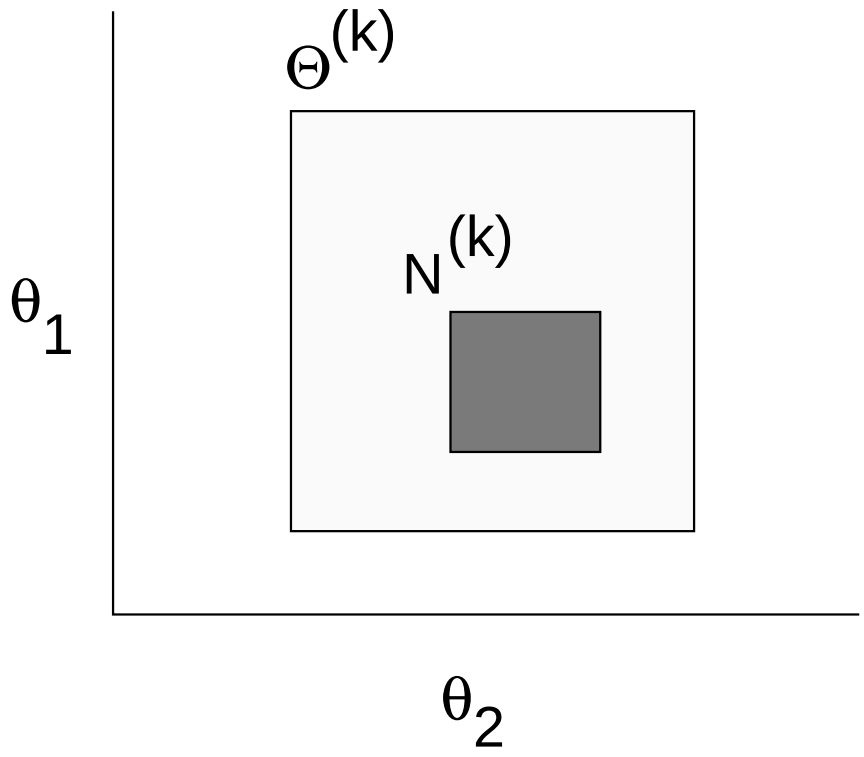
  This is the interval Newton/generalized bisection (IN/GB) approach.



Any solutions in $\Theta^{(k)}$ are also in intersection of $\Theta^{(k)}$ and $N^{(k)}$



There was no solution in $\Theta^{(k)}$



Unique solution in $\Theta^{(k)}$

This solution is in $N^{(k)}$

Point Newton method will converge to it

Interval Newton Method (continued)

- For $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$, this method can enclose **with mathematical and computational certainty** any and all solutions in a given initial interval, or can determine that there are none.

- A preconditioned interval Gauss-Seidel like technique is often used to solve for the image $\mathbf{N}^{(k)}$ (Hansen and coworkers).

- Our implementation is based on modifications of routines taken from the packages INTBIS and INTLIB (Kearfott and coworkers).

- The interval Newton procedure can be performed on multiple intervals independently and in parallel.

- IN/GB was first implemented for process modeling problems by Schnepper and Stadtherr (1990).

Parameter Estimation in VLE Modeling

- Goal: Determine parameter values in liquid phase activity coefficient models (e.g. Wilson, van Laar, NRTL, UNIQUAC):

$$\gamma_{\mu i,\mathrm{calc}} = f_i(\mathbf{x}_\mu, \boldsymbol{\theta})$$

- The relative least squares objective is commonly used:

$$\phi(\boldsymbol{\theta}) = \sum_{i=1}^{n} \sum_{\mu=1}^{p} \left[\frac{\gamma_{\mu i,\mathrm{calc}}(\boldsymbol{\theta}) - \gamma_{\mu i,\mathrm{exp}}}{\gamma_{\mu i,\mathrm{exp}}} \right]^2 .$$

- Experimental values $\gamma_{\mu i,\mathrm{exp}}$ of the activity coefficients are obtained from VLE measurements at compositions $\mathbf{x}_\mu, \mu = 1, \ldots, p$.

- Fit is usually made to binary (sometimes ternary) data. Other types of experimental data may also be used.

Example Problem 1

- The binary system water (1) and formic acid (2) was studied.

- Eleven problems, each a different data set from the DECHEMA VLE Data Collection (Gmehling $et\ al.$, 1977-1990) were considered.

- The model used was the Wilson equation. This has binary interaction parameters

$$\Lambda_{12} = (v_2/v_1)\exp(-\theta_1/RT) \text{ and}$$
$$\Lambda_{21} = (v_1/v_2)\exp(-\theta_2/RT)$$

  where v_1 and v_2 are pure component molar volumes.

- The energy parameters θ_1 and θ_2 must be estimated.

- Parameter estimation results for θ_1 and θ_2 are given in the DECHEMA Collection for all eleven problems.

Results–Example 1

- Each problem was solved using the IN/GB approach to determine the globally optimal values of the θ_1 and θ_2 parameters.

- These results were compared to those presented in the DECHEMA Collection.

- For each problem, the number of local minima in $\phi(\boldsymbol{\theta})$ was also determined (branch and bound steps were turned off).

- Table 1 presents a summary of these results and comparisons. CPU times are on a Sun Ultra 2/1300 workstation.

- Detailed results for one data set will be shown.

TABLE 1: Example 1–IN/GB results vs. DECHEMA values

Data Set	Data points	P (mmHg)	DECHEMA			IN/GB			No. of Minima	CPU time(s)
			θ_1	θ_2	$\phi(\boldsymbol{\theta})$	θ_1	θ_2	$\phi(\boldsymbol{\theta})$		
1	15	760	-195	759	0.0342	-195	759	0.0342	2	14.7
2	15	760	-278	1038	0.0106	-278	1038	0.0106	2	12.9
3	12	760	-310	1181	0.0151	-308	1167	0.0151	2	17.4
4	28	760	-282	985	0.353	-282	984	0.353	2	22.0
5	12	760	-366	1513	0.0257	-365	1509	0.0257	3	16.0
6	15	760	1067	-1122	0.0708	1065	-1120	0.0708	2	21.7
7*	14	200	892	-985	0.1410	**-331**	**1250**	**0.0914**	2	19.5
8*	11	200	370	-608	0.0459	**-340**	**1404**	**0.0342**	3	17.8
9*	16	100	539	-718	0.1650	**-285**	**996**	**0.1114**	2	24.4
10*	19	100	450	-663	0.1510	**-329**	**1394**	**0.0819**	3	27.4
11*	9	70	558	-762	0.0399	**-330**	**1519**	**0.0372**	3	16.4

*:New globally optimal parameters found.

Results–Example 1 (continued)

- Each problem has multiple local minima.
- In five of the problems (data sets 7–11), the result presented in DECHEMA represents a local not global minimum.
- Using the interval approach, the global minimum was found for all problems.
- The parameter estimation results obtained from the global minimization were more consistent than those in DECHEMA.
- There are several other systems for which the results given in the DECHEMA Collection do not represent the globally best fit.

Detailed Results–Data Set 10

- This problem has five stationary points, including three minima and two saddles. Details are shown in Table 2.

- These will not all be found if the branch and bound steps are turned on, so that only the global minimum is actually sought.

- The globally optimal parameters found using the interval approach provide a noticeably better fit to the experimental data. This is shown by the relative deviation plots given in Figures 1 and 2.

TABLE 2: Stationary Points (Roots) for Data Set 10

Root	θ_1, θ_2	Eigenvalues of Hessian	$\phi(\boldsymbol{\theta})$	Status
P1	(1658, -1251)	7.55E-5, 2.58E-7	0.164	minimum
P2	(1165, -1083)	6.83E-5, -1.44E-7	0.178	saddle
P3	(452, -664)	6.97E-5, 9.42E-8	0.151	minimum
P4	(-37.8, 38.5)	9.08E-5, -3.54E-7	0.19	saddle
P5	(-329, 1394)	1.23E-4, 1.47E-7	0.0819	global minimum

FIGURE 1: Example 1–Comparison of Relative Deviation in γ_1

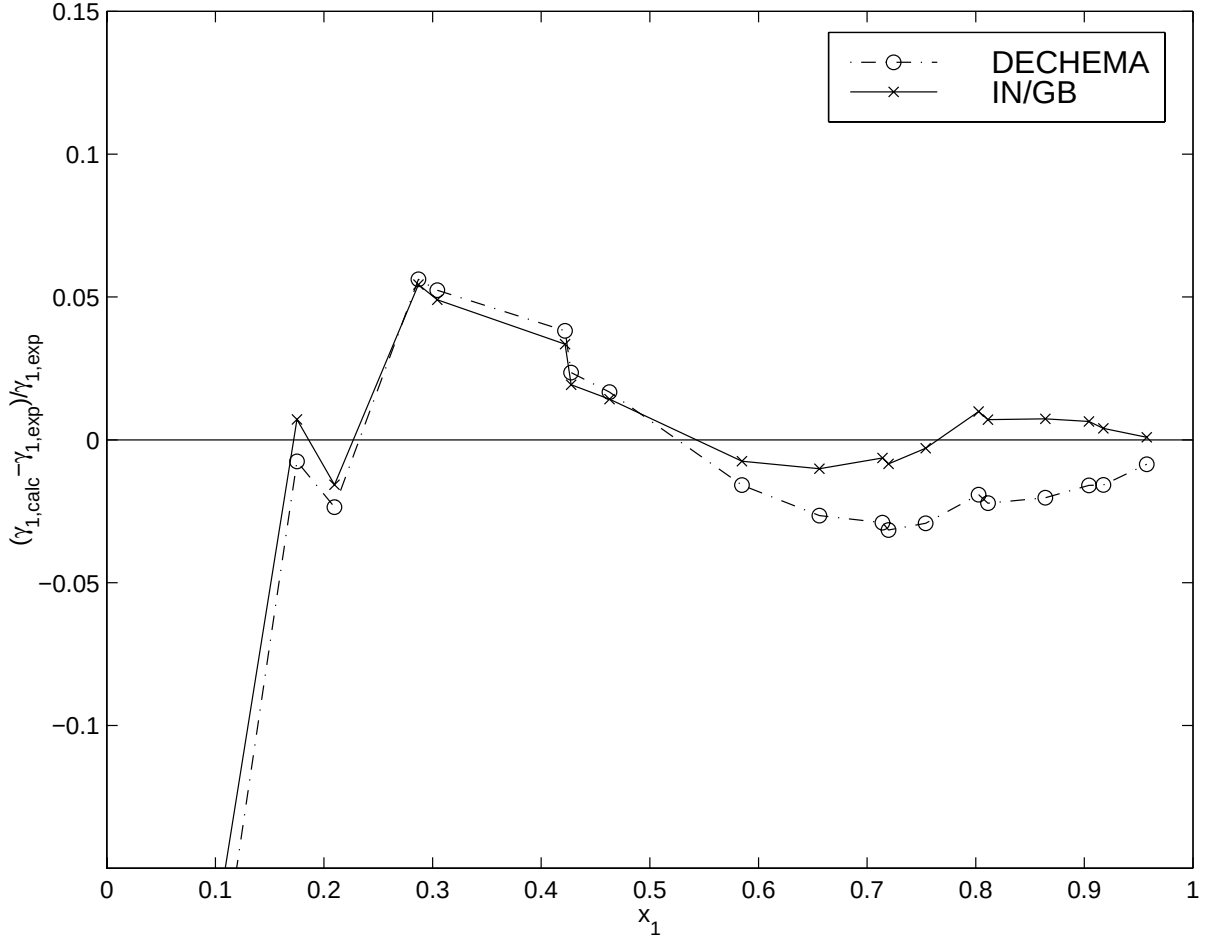
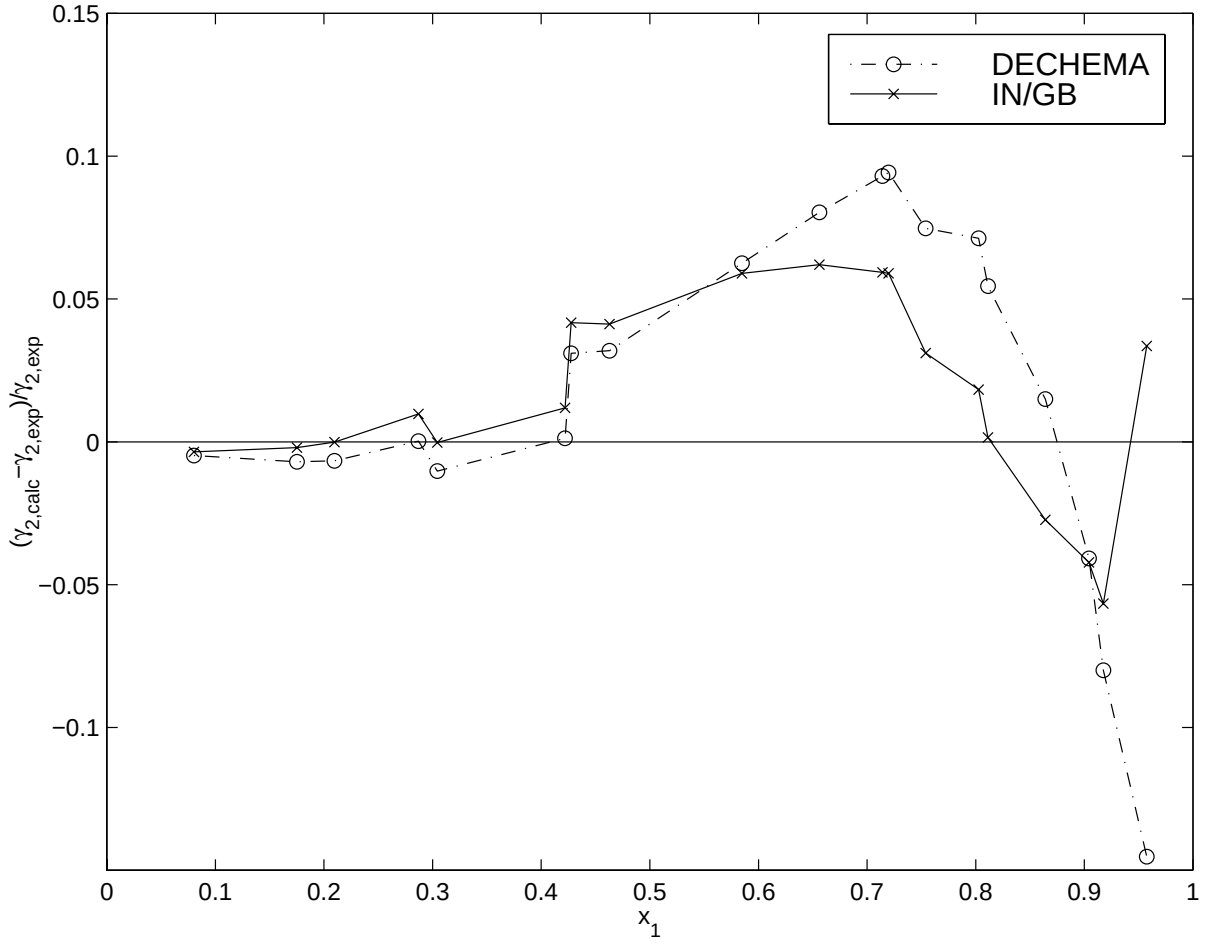


FIGURE 2: Example 1–Comparison of Relative Deviation in γ_2



Example Problem 2

- The binary system $tert$-butanol (1) and 1-butanol (2) was studied.

- Six problems, each a different data set from the DECHEMA VLE Data Collection (Gmehling $et\ al.$, 1977-1990) were considered.

- The model used was the Wilson equation.

- Table 3 compares parameter estimation results for θ_1 and θ_2 with those given in the DECHEMA Collection. New globally optimal parameter values are found in all six cases.

- The globally optimal parameters found using the interval approach provide a noticeably better fit to the experimental data. This is shown by the relative deviation plots given in Figures 3 and 4 for data set 6.

TABLE 3: Example 2–IN/GB results vs. DECHEMA values

Data Set	Data points	P (mmHg)	DECHEMA			IN/GB			No. of Minima	CPU time(s)
			θ_1	θ_2	$\phi(\boldsymbol{\theta})$	θ_1	θ_2	$\phi(\boldsymbol{\theta})$		
1*	9	100	951	-602	0.0136	**-568**	**745**	**0.0103**	2	12.8
2*	9	300	1068	-638	0.0158	**-525**	**626**	**0.0130**	2	11.1
3*	9	500	901	-594	0.0097	**-718**	**1265**	**0.0069**	2	12.4
4*	9	700	801	-561	0.0174	**-734**	**1318**	**0.0137**	2	12.9
5*	17	760	153	-203	0.1300	**-793**	**1757**	**0.1164**	2	16.2
6*	14	760	848	-606	0.0333	**-865**	**2420**	**0.0111**	2	11.0

*:New globally optimal parameters found.

FIGURE 1: Example 2–Comparison of Relative Deviation in γ_1

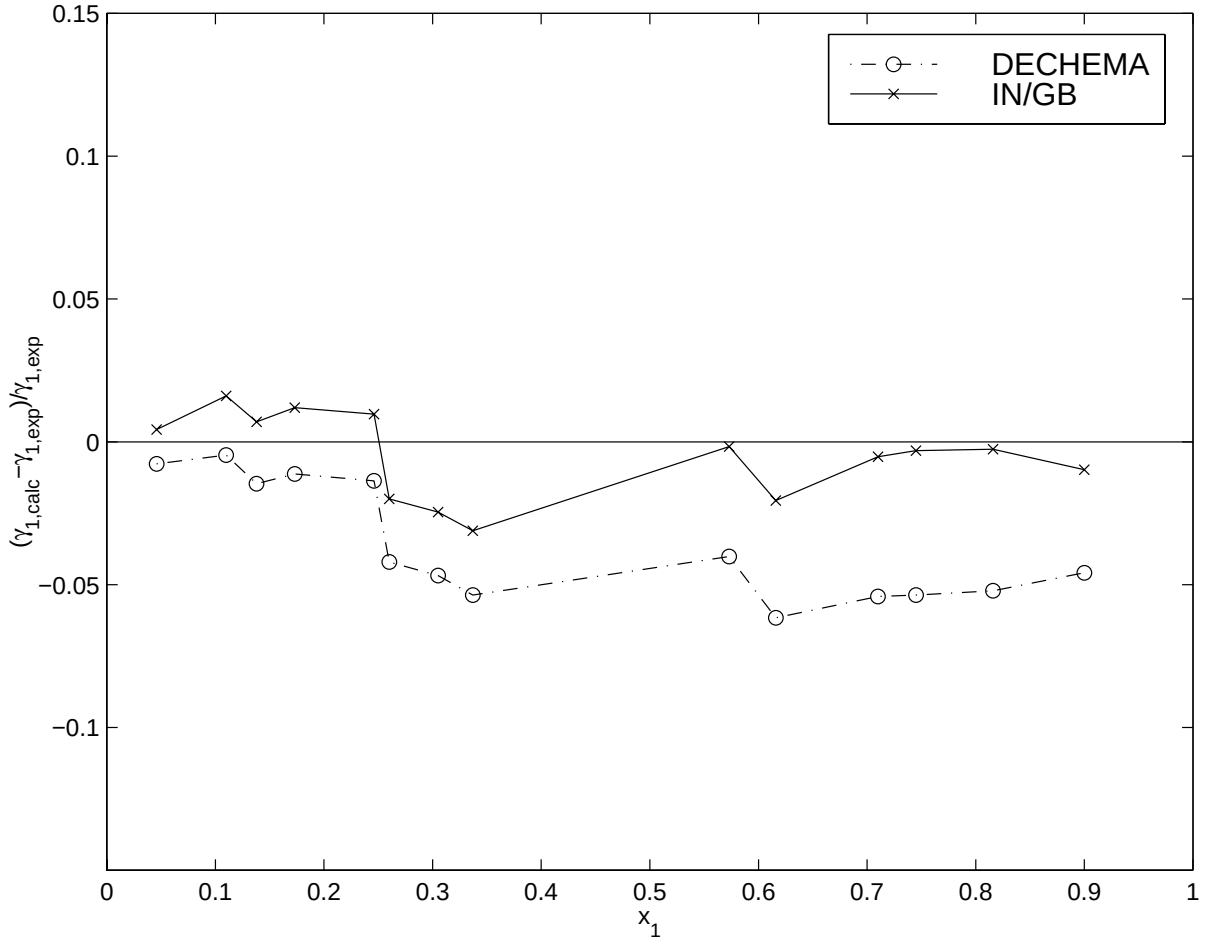
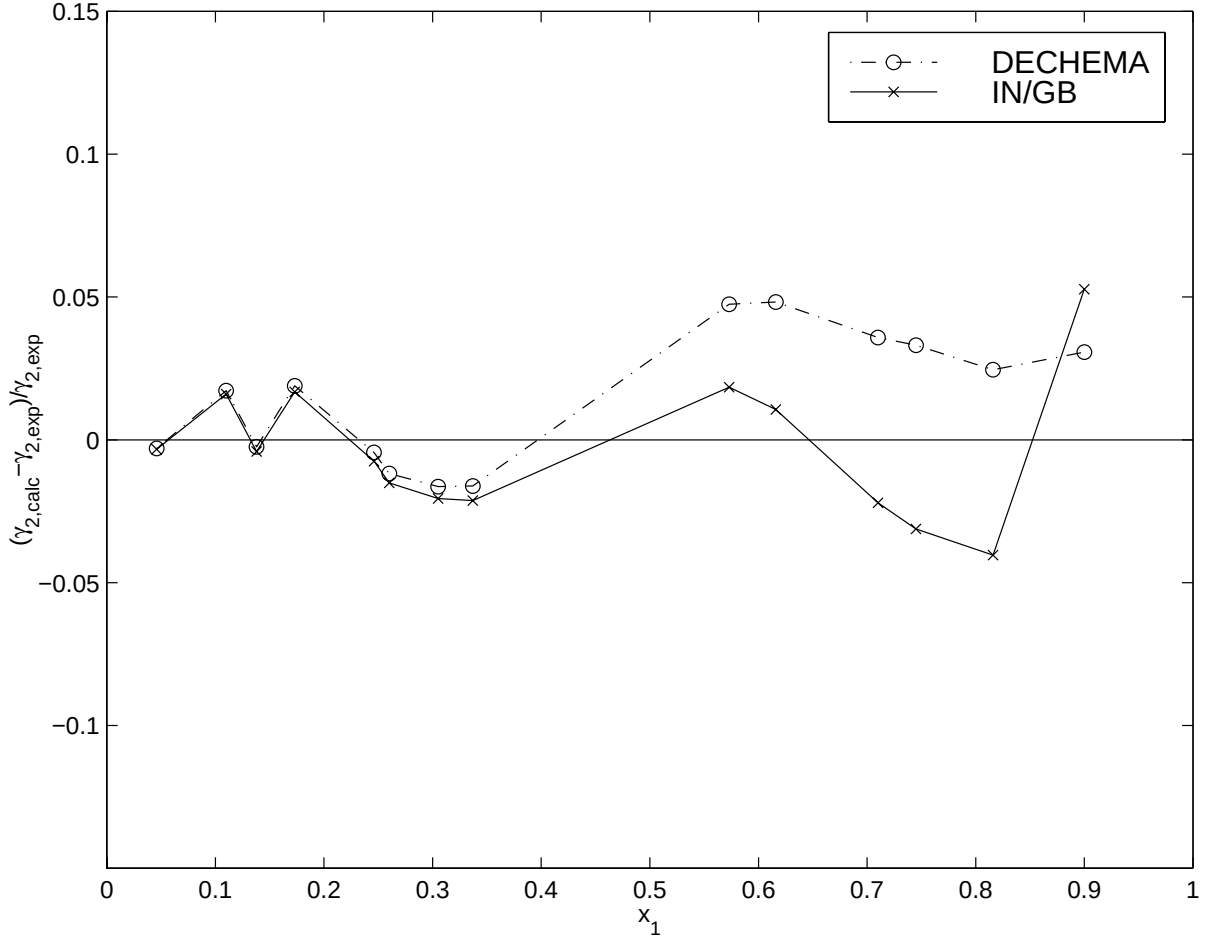


FIGURE 2: Example 2–Comparison of Relative Deviation in γ_2



Example Problem 3

- The binary system benzene (1) and hexafluorobenzene (2) was studied.

- Ten problems, each a different data set from the DECHEMA VLE Data Collection (Gmehling $et\ al.$, 1977-1990) were considered.

- The model used was the Wilson equation.

- Table 4 compares parameter estimation results for θ_1 and θ_2 with those given in the DECHEMA Collection. New globally optimal parameter values are found in five cases.

TABLE 4: Example 3–IN/GB results vs. DECHEMA values

Data Set	Data points	T (^oC)	DECHEMA			IN/GB			No. of Minima	CPU time(s)
			θ_1	θ_2	$\phi(\boldsymbol{\theta})$	θ_1	θ_2	$\phi(\boldsymbol{\theta})$		
1*	10	30	437	-437	0.0382	**-468**	**1314**	**0.0118**	2	15.1
2*	10	40	405	-405	0.0327	**-459**	**1227**	**0.0079**	2	13.7
3*	10	50	374	-374	0.0289	**-449**	**1157**	**0.0058**	2	12.3
4*	11	50	342	-342	0.0428	**-424**	**984**	**0.0089**	2	10.9
5	10	60	-439	1096	0.0047	-439	1094	0.0047	2	9.7
6	9	70	-424	1035	0.0032	-425	1036	0.0032	2	7.9

Data Set	Data points	P (mmHg)	DECHEMA			IN/GB			No. of Minima	CPU time(s)
			θ_1	θ_2	$\phi(\boldsymbol{\theta})$	θ_1	θ_2	$\phi(\boldsymbol{\theta})$		
7*	17	300	344	-347	0.0566	**-432**	**993**	**0.0149**	2	17.4
8	16	500	-405	906	0.0083	-407	912	0.0083	2	14.3
9	17	760	-407	923	0.0057	-399	908	0.0053	1	13.9
10	17	760	-333	702	0.0146	-335	705	0.0146	2	20.5

*:New globally optimal parameters found.

Example Problem 3–Discussion

- Does the use of the globally optimal parameters make a significant difference when the Wilson model is used to predict vapor-liquid equilibrium (VLE)?

- A common test of the predictive power of a model for VLE is its ability to predict azeotropes.

- Using the globally optimal parameters, the Wilson equation predicts two azeotropes, which is correct.

- Using the parameters in DECHEMA that are only locally optimal, the number of azeotropes predicted is zero (three data sets) or one (two data sets).

- Using the globally optimal parameters makes the difference between predicting physical reality or not.

Computational Performance

- With initial parameter intervals of $\Theta_1^{(0)} = \Theta_2^{(0)} = [-8500, 320000]$, the computation times required ranged from roughly 10 to 25 seconds on a Sun Ultra 2/1300 workstation.

- Initial parameter intervals were chosen based on physical knowledge of infinite dilution activity coefficients.

- An inverse midpoint preconditioner was used. Significant improvements in computation time are possible using an improved preconditioner.

- Because of the wide initial interval that can be used, as opposed to an initial point guess, the method is essentially **initialization independent**.

- The additional computation time for the interval approach, as opposed to local methods, is compensated by the guaranteed global reliability of the results.

Concluding Remarks

- Interval analysis is a **general-purpose** and **model-independent** approach for solving parameter estimation problems in modeling VLE, providing a **mathematical and computational guarantee** that the global optimum is found.
  - Other VLE models could be used.
  - Other objective functions (e.g, maximum likelihood) could be used.
  - Error-in-variable approach could be used.
  - Other types of data could be used.

- Interval analysis provides powerful problem solving techniques with many other applications in the modeling of thermodynamics and phase behavior and in other process modeling problems.

- Continuing advances in computing hardware and software (e.g., compiler support for interval arithmetic) will make this approach even more attractive.

- Acknowledgments

  - ACS PRF 30421-AC9
  - NSF CTS95-22835
  - NSF DMI96-96110
  - NSF EEC97-00537-CRCD
  - EPA R824731-01-0
  - DOE DE-FG07-96ER14691
  - Sun Microsystems, Inc.

- For more information:

  - Contact Prof. Stadtherr at markst@nd.edu