Deterministic Global Optimization for Error-in-Variables Parameter Estimation

Chao-Yang Gau and Mark A. Stadtherr
Dept. of Chemical Engineering, University of Notre Dame, Notre Dame, IN 46556

Standard techniques for solving the optimization problem arising in parameter estimation by the error-in-variables (EIV) approach offer no guarantee that the global optimum has been found. It is demonstrated here that the interval-Newton approach can provide a powerful, deterministic global optimization methodology for the reliable solution of EIV parameter estimation problems in chemical process modeling, offering mathematical and computational guarantees that the global optimum has been found. Although this methodology is typically regarded as being applicable only to very small problems, it is successfully applied here to problems with over 200 variables. It is a general-purpose technique and is applied here to a diverse group of problems, including examples in reactor modeling, in modeling vapor-liquid equilibrium, and in modeling a heat exchanger network.

Introduction

Parameter estimation is a central problem in the development of mathematical models that represent the physical phenomena underlying chemical process operations, and is thus an important issue in process systems engineering. In the classical least-squares approach to parameter estimation, it is assumed that there is a set of independent variables not subject to measurement error. The error-in-variables (EIV) approach differs in that it is assumed that there are measurement errors in *all* variables. Accounting for error in all the variables has been demonstrated (e.g., Duever et al., 1987; Patino-Leal and Reilly, 1982), to lead to unbiased estimates of the parameter values, and thus to more accurate models.

Consider the problem of estimating the parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_q)^T$ in a model of the general form $\boldsymbol{f}(\boldsymbol{\theta}, z) = \mathbf{0}$, where z is a vector of n state variables for the system to be modeled, and $\boldsymbol{f}$ is a vector of p model functions. When the EIV approach is used, the optimization problem that must be solved has the form

$$\min_{\theta, \tilde{z}_i} \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left(\tilde{z}_{ij} - z_{ij}\right)^2}{\sigma_j^2} \tag{1}$$

subject to

$$\boldsymbol{f}(\boldsymbol{\theta}, \tilde{z}_i) = \mathbf{0}, \quad i = 1, \ldots, m \tag{2}$$

Here $z_i = (z_{i1}, \ldots, z_{in})^T$ represents measurements of the state variables from $i = 1, \ldots, m$ experiments, $\tilde{z}_i = (\tilde{z}_{i1}, \ldots, \tilde{z}_{in})^T$ represents the unknown "true" values associated with each measurement, and σ_j represents the standard deviation associated with the measurement of state variable j. Details concerning the formulation of this optimization problem are available elsewhere (e.g., Kim et al., 1990; Esposito and Floudas, 1998), as are several good introductions to the general problem of nonlinear parameter estimation (e.g., Bard, 1974; Gallant, 1987; Seber, 1989).

The presence of the true values $\tilde{z}_i$, $i = 1, \ldots, m$, as variables in the optimization problem has a number of practical implications. It means that, in solving the EIV problem, not only are parameter estimation results obtained, but also data reconciliation results. However, this comes at the expense of a substantial increase in the dimensionality of the optimization problem, which at $nm + p$ is now a function of the number of experiments. Furthermore, since the optimization is over both $\boldsymbol{\theta}$ and $\tilde{z}_i$, this is likely to be a nonlinear optimization problem even for models that are linear in the parameters. Thus, in general, the optimization problem is nonlinear

Correspondence concerning this article should be addressed to M. A. Stadtherr.
Current address of C.-Y. Gau: LINDO System, Inc., 1415 North Dayton St., Chicago, IL 60622.

and potentially nonconvex, indicating the need to be concerned about the possible existence of multiple local minima.

Various methods have been used to solve the optimization problem defined by Eqs. 1 and 2. These include gradient-based methods such as Gauss-Newton or Gauss-Marquardt (e.g., Britt and Luecke, 1973; Fabries and Renon, 1975; Anderson et al., 1978; Schwetlick and Tiller, 1985; Valko and Vajda, 1987), generalized reduced gradient (Kim et al., 1990), and successive quadratic programming (e.g., Tjoa and Biegler, 1991, 1992), as well as direct search methods, such as the simplex pattern search (e.g., Vamos and Hass, 1994). However, these are all local methods that offer no assurance that the global minimum in the optimization problem has been found. Towards finding the global optimum, one approach is to introduce a randomized element, either in the selection of multiple initial guesses (e.g., Vamos and Hass, 1994), or in the search procedure itself (e.g., Luus and Hernaez, 2000). These stochastic methods still provide no guarantee that the global optimum has been found. To obtain any such guarantee requires the use of *deterministic* global optimization procedures. Deterministic methods can provide a guarantee that the global optimum within some specified search domain is found. The search domain can be made large enough to enclose all physically feasible and statistically significant behavior.

A powerful deterministic approach is that suggested by Esposito and Floudas (1998), who reformulate the optimization problem in terms of convex underestimating functions and then use a branch-and-bound procedure. This method provides a mathematical guarantee of global optimality. One drawback to this approach is that in general it may be necessary to perform problem reformulations and develop convex underestimators specific to each new application. Also, in principle, branch-and-bound methods implemented in floating point arithmetic may be vulnerable to rounding error problems, and thus lose their mathematical guarantees. Another deterministic approach is that suggested recently by Gau and Stadtherr (2000), who use an interval-Newton approach. This is a general-purpose methodology that provides a mathematical guarantee of global optimality, as well as a computational guarantee, since rounding issues are dealt with through the use of interval arithmetic. In their preliminary study, Gau and Stadtherr (2000) demonstrated the potential of the interval methodology by applying it to some small problems (12 to 32 variables) and finding the approach to compare favorably to the Esposito and Floudas (1998) method in terms of computational efficiency.

In this article, we further explore the feasibility of using the interval methodology to provide a deterministic global optimization tool for solving EIV parameter estimation problems. In particular, we consider some problems that are much larger (up to 264 variables) than considered previously. A diverse group of problems is considered, including examples in reactor modeling, in modeling vapor-liquid equilibrium, and in modeling a heat exchanger network.

Methodology

For many practical problems, the p model equations can be easily solved algebraically for p of the n state variables. Thus, by substitution into the objective function, an unconstrained formulation of the optimization problem can be obtained. The unconstrained problem can be stated

$$\min_{\boldsymbol{\theta},\, \tilde{\boldsymbol{v}}_i} \phi(\boldsymbol{\theta}, \tilde{\boldsymbol{v}}_i) \tag{3}$$

where $\tilde{\boldsymbol{v}}_i$, $i = 1, \ldots, m$, refers to the $n - p$ independent state variables not eliminated using the model equations, and $\phi(\boldsymbol{\theta}, \tilde{\boldsymbol{v}}_i)$ is the objective function in Eq. 1 after the p dependent state variables have been eliminated by substitution. This unconstrained formulation of the problem will be used here. However, it should be noted that the accompanying reduction in the dimensionality of the problem does not necessarily make it any easier to solve, since the objective function in the reduced space of $\boldsymbol{\theta}$ and $\tilde{\boldsymbol{v}}_i$ may be a much more complicated function than the objective function in the original space of $\boldsymbol{\theta}$ and $\tilde{\boldsymbol{z}}_i$.

For the global minimization of $\phi(\boldsymbol{\theta}, \tilde{\boldsymbol{v}}_i)$, an approach based on interval analysis is used. Good introductions to interval analysis, as well as interval arithmetic and computing with intervals, include those of Neumaier (1990), Hansen (1992), and Kearfott (1996). Of particular interest here is the interval-Newton technique. Given a nonlinear equation system with a finite number of real roots in some initial interval, this technique provides the capability to find (or, more precisely, to enclose within a very narrow interval) all the roots of the system within the given initial interval. To apply this technique to the optimization problem of interest here, it is used to seek stationary points; that is, to solve the nonlinear equation system

$$\boldsymbol{g}(\boldsymbol{y}) = \boldsymbol{g}(\boldsymbol{\theta}, \tilde{\boldsymbol{v}}_i) \equiv \nabla\phi(\boldsymbol{\theta}, \tilde{\boldsymbol{v}}_i) = \boldsymbol{0} \tag{4}$$

where for convenience the vector of independent variables has been denoted $\boldsymbol{y} = (\boldsymbol{\theta}, \tilde{\boldsymbol{v}}_i)^T$. The global minimum will be a root of this nonlinear equation system, but there may be many other roots as well, representing local minima and maxima and saddle points. To identify the global minimum, one approach is to simply find all the stationary points and then identify the point with the minimum value of the objective function. Alternatively, by including an objective range test in the solution procedure, as explained by Gau and Stadtherr (2000), one can effectively combine the interval-Newton approach with an interval branch-and-bound technique, so that roots of $\boldsymbol{g}(\boldsymbol{y}) = \boldsymbol{0}$ that cannot be the global minimum need not be found. If the constrained formulation of the problem is used, then instead of applying interval-Newton to solve the stationarity conditions, it is applied to solve the Karush-Kuhn-Tucker (KKT) conditions (or, more generally, the Fritz-John conditions). Note that the search for stationary points (or KKT points), and thus the search for the global optimum, occurs only within the specified initial interval; however, this search domain can be made arbitrarily large.

A summary of the solution algorithm used has been given previously (Gau and Stadtherr, 2000). Applied to nonlinear equation solving, the methodology is basically a branch-and-prune scheme on a binary tree, while for optimization it is a branch-and-bound scheme. It should be noted that recent enhancements (Gau and Stadtherr, 2002) to the methodology, involving the formulation and solution of the interval-Newton equation, and including the use of a new preconditioning

strategy, play an important role in achieving computational efficiency on the problems considered here. A systematic study of the impact of this enhanced methodology is given by Gau and Stadtherr (2002). For most of the large problems considered below, use of this enhanced methodology is essential in order to achieve tractable computation times. It should be emphasized that, when applying this solution procedure, the user must specify an initial interval $\boldsymbol{Y}^{(0)}$ that provides upper and lower bounds on the independent variables. This initial interval can be chosen to be sufficiently large to enclose all physically feasible and statistically significant behavior. This is in contrast to conventional local solution methods in which an initial *point* is needed, often resulting in a highly initialization-dependent procedure. It is assumed here that the global optimum will occur at an interior stationary minimum of $\phi(\boldsymbol{y})$ and not at the boundaries of $\boldsymbol{Y}^{(0)}$. Since the estimator ϕ is derived based on a product of Gaussian distribution functions corresponding to each data point, this is a very reasonable assumption for regression problems of the type considered here. When properly implemented, the interval-Newton method provides a procedure that is mathematically *and* computationally guaranteed to find the global minimum of $\phi(\boldsymbol{y})$, or, if desired (by turning off the objective range test), to enclose *all* of its stationary points.

Catalytic Reactor Model

This problem involves the modeling of an isothermal pseudo-differential reactor for the catalytic hydrogenation of phenol on a palladium catalyst, as described by Rod and Hancil (1980). The experimental kinetic data for this gas-phase reaction involves 28 measured data points of the partial pressure of phenol P_1 (atm), the partial pressure of hydrogen P_2 (atm), and the initial reaction rate r (mol/kg h). It is desired to fit this kinetic data (Rod and Hancil, 1980) to a semi-empirical model of the form

$$r = \frac{\theta_1 \theta_2^2 \theta_3 P_1 P_2^2}{(1 + \theta_1 P_1 + \theta_2 P_2)^3} \tag{5}$$

where θ_1 (atm^{-1}), θ_2 (atm^{-1}), and θ_3 (mol/kg h) are the parameters to be estimated. The vector of state variables is $\boldsymbol{z} = (P_1,\ P_2,\ r)^T$. The standard deviations for the measurements of the state variables are known (Rod and Hancil, 1980), and given by $\boldsymbol{\sigma} = (0.0075,\ 0.0075,\ 2.5)^T$.

In order to formulate the EIV parameter estimation problem as an unconstrained optimization problem, the model Eq. 5 is used to eliminate r. Thus, the vector of independent state variables is $\boldsymbol{v} = (P_1,\ P_2)^T$. In the optimization problem the independent variables are $\boldsymbol{\theta}$ (three variables) and $\tilde{\boldsymbol{v}}_i$, $i = 1, \ldots, 28$ (28 vectors of two variables each), for a total of 59 independent variables. This is roughly twice the size of the largest EIV parameter estimation problem solved previously (Gau and Stadtherr, 2000), using the interval approach. Since, as noted by Rod and Hancil (1980), the values of $\boldsymbol{\theta} = (7.27,\ 0.681,\ 1602)^T$ are suspected of being good parameters, the initial intervals on the parameters were taken here as $\theta_1 \in [6,\ 9]$ atm^{-1}, $\theta_2 \in [0.5,\ 8]$ atm^{-1}, and $\theta_3 \in [1600,\ 1900]$ mol/kg h. The initial intervals on the state variables were set using plus and minus three standard deviations; that is, $\tilde{P}_{1i} \in [P_{1i} - 3\sigma_1,\ P_{1i} + 3\sigma_1]$ and $\tilde{P}_{2i} \in [P_{2i} - 3\sigma_2,\ P_{2i} + 3\sigma_2]$ for $i = 1, \ldots, 28$. Statistically, these initial intervals provide a 99.7% probability of containing the true values of the state variables.

Using the interval methodology discussed above, the globally optimal parameter values obtained were $\theta_1 = 7.39696$ atm^{-1}, $\theta_2 = 0.63782$ atm^{-1}, and $\theta_3 = 1769.71$ mol/kg h with an objective value of 30.3072. These results, along with the results (not shown here) for the $\tilde{\boldsymbol{v}}_i$, are consistent with those of Rod and Hancil (1980). It should be noted that, while point approximations are reported here, and in subsequent examples, for the parameter estimation results, we have actually determined verified enclosures of the corresponding stationary points. Each such enclosure is an extremely narrow interval known to contain a *unique* stationary point, based on the interval-Newton uniqueness test (e.g., Kearfott, 1996).

The CPU time required for this 59-variable global optimization problem was 2,588 s on a Sun UltraServer2/2200 workstation (one processor). Given the size of the problem, and the fact that a rigorous global minimum was obtained, this degree of computational effort is quite reasonable. By turning off the objective range test, thus allowing the technique to enclose *all* the stationary points, not just the global minimum, it was also ascertained that, for this problem, there was only one stationary point (the global minimum) in the specified initial interval. For this problem, finding all the stationary points requires about 5% additional computation time compared to using the objective range test to find the global optimum only; however, the savings provided by using the objective range test can vary significantly from problem to problem.

Heat Exchanger Network Model

This problem involves the modeling of a steady state heat exchanger network as described by Biegler and Tjoa (1993). The network is shown in Figure 1. This system consists of four heat exchangers, with cold stream A1 to be heated using hot streams B1, C1 and D1. It is assumed that all streams have the same constant heat capacity. As the first step in simulation of this system, it is necessary to have an estimate of the rating parameter UA (product of overall heat-transfer coefficient and heat-transfer area) for each exchanger. This parameter vector $\boldsymbol{\theta} = ((UA)_1,\ (UA)_2,\ (UA)_3,\ (UA)_4)^T$ can be estimated from experimental measurements. Measurement

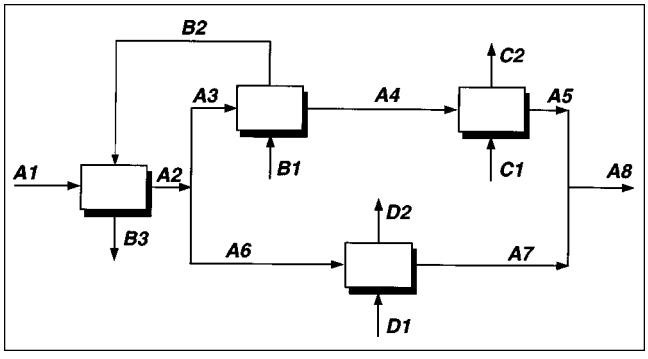

Figure 1. Heat exchanger network example.

data (in arbitrary units) are given by Biegler and Tjoa (1993) for six flow rates and 13 temperatures. In particular, the measured state variable vector is $z = (F_{A1}, F_{B1}, F_{C1}, F_{D1}, F_{A3}, F_{A6}, T_{A1}, T_{A2}, T_{A4}, T_{A5}, T_{A7}, T_{A8}, T_{B1}, T_{B2}, T_{B3}, T_{C1}, T_{C2}, T_{D1}, T_{D2})^T$, with standard deviations of 0.5 and 0.001 for the flow rate and temperature variables, respectively. The measurements were created by Biegler and Tjoa (1993) from the parameter values $\theta^* = (4.85, 4.00, 6.80, 5.35)^T$ with added noise.

The model used here is given by

$$F_{A1}(T_{A2} - T_{A1}) = (UA)_1 \frac{(T_{B2} - T_{A2}) - (T_{B3} - T_{A1})}{\ln\left(\frac{T_{B2} - T_{A2}}{T_{B3} - T_{A1}}\right)} \quad (6)$$

$$F_{B1}(T_{B1} - T_{B2}) = (UA)_2 \frac{(T_{B1} - T_{A4}) - (T_{B2} - T_{A3})}{\ln\left(\frac{T_{B1} - T_{A4}}{T_{B2} - T_{A3}}\right)} \quad (7)$$

$$F_{C1}(T_{C1} - T_{C2}) = (UA)_3 \frac{(T_{C1} - T_{A5}) - (T_{C2} - T_{A4})}{\ln\left(\frac{T_{C1} - T_{A5}}{T_{C2} - T_{A4}}\right)} \quad (8)$$

$$F_{D1}(T_{D1} - T_{D2}) = (UA)_4 \frac{(T_{D1} - T_{A7}) - (T_{D2} - T_{A6})}{\ln\left(\frac{T_{D1} - T_{A7}}{T_{D2} - T_{A6}}\right)} \quad (9)$$

$$F_{A1} - F_{A3} - F_{A6} = 0 \quad (10)$$

$$F_{A1}T_{A8} - F_{A3}T_{A5} - F_{A6}T_{A7} = 0 \quad (11)$$

Additional balance equations are available, and could be included in the model, especially for a case in which data reconciliation was the primary goal.

In order to formulate the EIV parameter estimation problem as an unconstrained optimization problem, the model Eqs. 6–11 is used to solve for the six flow rate variables

$$F_{A1} = (UA)_1 \frac{(T_{B2} - T_{A2}) - (T_{B3} - T_{A1})}{(T_{A2} - T_{A1})\ln\left(\frac{T_{B2} - T_{A2}}{T_{B3} - T_{A1}}\right)} \quad (12)$$

$$F_{B1} = (UA)_2 \frac{(T_{B1} - T_{A4}) - (T_{B2} - T_{A2})}{(T_{B1} - T_{B2})\ln\left(\frac{T_{B1} - T_{A4}}{T_{B2} - T_{A2}}\right)} \quad (13)$$

$$F_{C1} = (UA)_3 \frac{(T_{C1} - T_{A5}) - (T_{C2} - T_{A4})}{(T_{C1} - T_{C2})\ln\left(\frac{T_{C1} - T_{A5}}{T_{C2} - T_{A4}}\right)} \quad (14)$$

$$F_{D1} = (UA)_4 \frac{(T_{D1} - T_{A7}) - (T_{D2} - T_{A2})}{(T_{D1} - T_{D2})\ln\left(\frac{T_{D1} - T_{A7}}{T_{D2} - T_{A2}}\right)} \quad (15)$$

$$F_{A3} = (UA)_1 \left(\frac{T_{A8} - T_{A7}}{T_{A5} - T_{A7}}\right) \frac{(T_{B2} - T_{A2}) - (T_{B3} - T_{A1})}{(T_{A2} - T_{A1})\ln\left(\frac{T_{B2} - T_{A2}}{T_{B3} - T_{A1}}\right)} \quad (16)$$

$$F_{A6} = (UA)_1 \left(\frac{T_{A5} - T_{A8}}{T_{A5} - T_{A7}}\right) \frac{(T_{B2} - T_{A2}) - (T_{B3} - T_{A1})}{(T_{A2} - T_{A1})\ln\left(\frac{T_{B2} - T_{A2}}{T_{B3} - T_{A1}}\right)} \quad (17)$$

which are then eliminated from the objective function. The vector of independent state variables is thus $v = (T_{A1}, T_{A2}, T_{A4}, T_{A5}, T_{A7}, T_{A8}, T_{B1}, T_{B2}, T_{B3}, T_{C1}, T_{C2}, T_{D1}, T_{D2})^T$.

Following Biegler and Tjoa (1993), five versions of the parameter estimation problem were solved, differing in the number of data points, which ranges from $m = 4$ for the smallest problem to $m = 20$ for the largest. In the optimization problem, the number of independent variables $(13\ m + 4)$ thus ranges from 56 to 264 variables. To the best of the authors' knowledge, the 264-variable problem is the largest global optimization problem ever attempted using interval methodology. The initial intervals on the parameters were set at $\theta_i \in [1, 10]$ for $i = 1, \ldots, 4$. The initial intervals on all independent temperature variables were set using plus and minus 1.5 standard deviations.

Results for solving the parameter estimation problem using the interval-based global optimization methodology for each of these five data sets is shown in Table 1. Clearly the results for the parameters are consistent with the values from which the measurement data were created by Biegler and Tjoa (1993). However, for the larger problems, the CPU times required are much larger than that required by the local SQP algorithm used by Biegler and Tjoa (1993). This of course reflects the trade-off between computational speed and the *guarantee* that a global optimum has been found. Considering the very large size of these problems in the context of deterministic global optimization, the computation time requirements are actually quite remarkable.

In our experience, and that of others (e.g., Kearfott and Novoa, III, 1990), the difficulty of solving problems using an

Table 1. Computational Results for Heat Exchanger Network Problem

No. of Data Points	4	8	12	16	20
$(UA)_1$	4.850851	4.837711	4.840830	4.841813	4.842804
$(UA)_2$	3.999357	3.999918	3.999853	4.000018	4.001031
$(UA)_3$	6.796941	6.800698	6.801176	6.801033	6.812213
$(UA)_4$	5.350317	5.350150	5.350694	5.350535	5.350411
Objective Function	0.01225901	4.82968895	5.03713956	5.03980518	5.08279327
No. of Variables	56	108	160	212	264
CPU time (s)*	0.17	440.0	1243.1	1172.3	2157.5
Leaves in Search Tree	1	763	720	380	315
CPU time (s)/Leaf	0.17	0.58	1.73	3.09	6.85

*CPU time is on a Sun UltraServer2/2200 workstation (one processor).

interval-Newton approach does not necessarily correlate well with the number of variables, and can in fact be quite unpredictable, which is not surprising considering the NP-hard nature of the problems being solved. This unpredictability can be seen in the CPU time results for this problem, which indicate a reduction in CPU time in going from 160 variables to 212 variables. To better interpret this result, we also present in Table 1 a measure of the size of the binary search tree for each problem. This measure is the number of leaves in the tree; a leaf represents a subinterval beyond which no further branching occurs, since the subinterval has been shown either to contain no root (stationary point), to contain a unique root, or to not contain the global optimum. We note that the size of the search tree (number of leaves) does not vary predictably with problem size. This occurs in part because, in a branch-and-bound scheme, the size of the search tree will depend significantly on how quickly a good upper bound on the global optimum is found, and this may have little relation to problem size. We also note that, when viewed in terms of CPU time per leaf, the computational effort does increase with problem size as expected.

Vapor-Liquid Equilibrium Model

This problem involves the modeling of vapor-liquid equilibrium (VLE) using the Wilson equation for liquid-phase activity coefficient. Of interest is the binary system benzene(1)/hexafluorobenzene(2). Ten data sets, taken from the DECHEMA VLE Data Collection (Gmehling et al., 1977–1990), are considered, each providing measurements of the state variable vector $z = (x_1, y_1, P, T)^T$, where P is the system pressure (mmHg), T is the system temperature (K), x_1 is the liquid-phase mol fraction of component 1, and y_1 is the vapor-phase mol fraction of component 1. A standard deviation vector of $\boldsymbol{\sigma} = (0.001, 0.01, 0.75, 0.1)^T$ is assumed. Parameter estimation for these ten data sets was considered by Gau et al. (2000) using a simple least-squares approach, and is treated here using the EIV approach.

The model used to describe the VLE can be written as

$$P = \gamma_1 x_1 p_1^0(T) + \gamma_2(1 - x_1) p_2^0(T) \tag{18}$$

$$y_1 = \frac{\gamma_1 x_1 p_1^0(T)}{\gamma_1 x_1 p_1^0(T) + \gamma_2(1 - x_1) p_2^0(T)} \tag{19}$$

where the pure component vapor pressures $p_1^0(T)$ and $p_2^0(T)$ are given by the Antoine relationships

$$p_1^0(T) = \exp\left[15.8412 - \frac{2755.64}{T - 53.99}\right]$$

$$p_2^0(T) = \exp\left[16.1940 - \frac{2827.54}{T - 57.66}\right]$$

and the activity coefficients γ_1 and γ_2 are given by the Wilson equation

$$\ln\gamma_1 = -\ln(x_1 + \Lambda_{12} x_2) + x_2\left[\frac{\Lambda_{12}}{x_1 + \Lambda_{12} x_2} - \frac{\Lambda_{21}}{\Lambda_{21} x_1 + x_2}\right] \tag{20}$$

$$\ln\gamma_2 = -\ln(x_2 + \Lambda_{21} x_1) - x_1\left[\frac{\Lambda_{12}}{x_1 + \Lambda_{12} x_2} - \frac{\Lambda_{21}}{\Lambda_{21} x_1 + x_2}\right] \tag{21}$$

Here the binary parameters Λ_{12} and Λ_{21} are given by

$$\Lambda_{12} = \frac{v_2}{v_1}\exp\left[-\frac{\theta_1}{RT}\right] \tag{22}$$

$$\Lambda_{21} = \frac{v_1}{v_2}\exp\left[-\frac{\theta_2}{RT}\right] \tag{23}$$

where v_1 and v_2 are the pure component liquid molar volumes, and θ_1 and θ_2 are the energy parameters (cal/mol) that need to be estimated.

In order to formulate the EIV parameter estimation problem as an unconstrained optimization problem, the model Eqs. 18–19 are used to eliminate P and y_1 in the objective

Table 2. Parameter Estimation Results for Benzene(1)/Hexafluorobenzene(2) System Using EIV Approach

Data Set	Vol. Pg.†	No. of Data Points	T (°C) or P (mm Hg)	θ_1	θ_2	ϕ	Minima	CPU Time (s)*
1	7:228	10	$T = 30$	−472	1,274	11.899	2	39.4
2	7:229	10	40	−462	1,197	11.170	2	44.9
3	7:230	10	50	−455	1,139	9.9312	1	54.4
4	7:233	11	50	−461	1,116	19.525	1	62.4
5	7:231	10	60	−445	1,086	9.9352	1	58.8
6	7:232	9	70	−424	1,007	8.5034	1	66.0
7	7:234	17	$P = 300$	−478	1,189	37.399	1	109.1
8	7:235	16	500	−444	1,090	13.786	1	234.1
9	7:236	17	760	−435	1,080	5.1490	1	328.4
10‡	7:226	29	760	−421	1,060	16.925	1	9396.6

*CPU time is on a Sun UltraServer2/2200 workstation (one processor).
†Refers to volume and page numbers in DECHEMA VLE Data Collection (Gmehling et al., 1977–1990).
‡Initial intervals on state variables are taken as ±2.5 standard deviations.

function. Consequently, the vector of independent state variables is $\boldsymbol{v} = (x_1, T)^T$. In the unconstrained optimization problem, the independent variables are $\boldsymbol{\theta} = (\theta_1, \theta_2)^T$ and $\tilde{\boldsymbol{v}}_i$, $i = 1, \ldots, m$ (m vectors of two variables each), for a total of $2m + 2$ independent variables. For the ten benzene/hexafluorobenzene data sets considered, the number of data points ranges from $m = 9$ to $m = 29$; thus, the number of independent variables in the global optimization problem ranges from 20 to 60. The initial intervals on the parameters θ_1 and θ_2 were both taken as $[-10{,}000, 200{,}000]$, which covers the range of physical interest as described by Gau et al. (2000). The initial intervals on the independent state variables were set using plus and minus three standard deviations.

The results of solving the EIV parameter estimation problem using the interval methodology for global optimization are shown in Table 2 for each of the ten data sets. By turning off the objective range test, the number of local minima for each case was also determined. Note that, for the first two data sets, there are two local minima. For data set 1, there is a local but not global minimum at $\theta_1 = 407.3$, $\theta_2 = -406.9$, and $\phi = 53.886$. Similarly, for data set 2, there is a local but not global minimum at $\theta_1 = 374.4$, $\theta_2 = -373.8$, and $\phi = 67.799$. In both cases, the *global* optimum was readily found using the interval approach. The presence of multiple local minima emphasizes the importance of using a deterministic global optimization approach to solve the EIV parameter estimation problem. As noted by Gau et al. (2000), convergence to a local, but not global, optimum in parameter estimation problems for VLE models is not an uncommon occurrence. This can lead to the dismissal of a model as inadequate, when in fact the model may be fine, provided that the parameter estimation problem is solved correctly to a global optimum.

Concluding Remarks

We have demonstrated here that the interval-Newton approach is a powerful, deterministic global optimization methodology for the reliable solution of EIV parameter estimation problems in chemical process modeling. The approach provides both mathematical and computational guarantees that the global optimum in the parameter estimation problem has been found. It is a general-purpose approach that has been applied here to a diverse group of problems. Although this methodology is typically regarded as being applicable only to very small problems, we have shown here that it can be successfully applied to problems with over 200 variables. The guaranteed reliability of the interval approach comes at the expense of significantly higher computation time requirements in comparison to local methods that provide no such guarantees. Thus, modelers must consider this trade-off, and ultimately decide how important it is to know for sure that the correct answer has been obtained.

Acknowledgments

This work has been supported in part by the donors of The Petroleum Research Fund, administered by the ACS, under Grant 35979-AC9, by the National Science Foundation Grant EEC97-00537-CRCD, and by the Environmental Protection Agency Grant R826-734-01-0.

Literature Cited

Anderson, T. F., D. S. Abrams, and E. A. Grens., "Evaluation of Parameters for Nonlinear Thermodynamic Models," *AIChE J.*, **24**, 20 (1978).

Bard, Y., *Nonlinear Parameter Estimation*, Academic Press, New York (1974).

Biegler, L. T., and I.-B. Tjoa, "A Parallel Implementation for Parameter Estimation with Implicit Models," *Anns. Opns. Res.*, **42**, 1 (1993).

Britt, H. I., and R. H. Luecke, "The Estimation of Parameters in Nonlinear, Implicit Models," *Technometrics*, **15**, 233 (1973).

Duever, T. A., S. E. Keeler, and P. M. Reilly, "An Application of the Error-in-Variables Model—Parameter Estimation from Van Ness-Type Vapor-Liquid Equilibrium Experiments," *Chem. Eng. Sci.*, **42**, 403 (1987).

Esposito, W. R., and C. A. Floudas, "Global Optimization in Parameter Estimation of Nonlinear Algebraic Models via the Error-in-Variables Approach," *Ind. Eng. Chem. Res.*, **37**, 1841 (1998).

Fabries, J.-F., and H. Renon, "Method for Evaluation and Reduction of Vapor-Liquid Equilibrium Data of Binary Mixtures," *AIChE J.*, **21**, 735 (1975).

Gallant, A. R., *Nonlinear Statistical Models*, Wiley, New York (1987).

Gau, C.-Y., and M. A. Stadtherr, "Reliable Nonlinear Parameter Estimation Using Interval Analysis: Error-in-Variables Approach," *Comput. Chem. Eng.*, **24**, 631 (2000).

Gau, C.-Y., J. F. Brennecke, and M. A. Stadtherr, "Reliable Nonlinear Parameter Estimation in VLE Modeling," *Fluid Phase Equilib.*, **168**, 1 (2000).

Gau, C.-Y., and M. A. Stadtherr, "New Interval Methodologies for Reliable Chemical Process Modeling," *Comput. Chem. Eng.*, in press (2002).

Gmehling, J., U. Onken, and W. Arlt, *Vapor-Liquid Equilibrium Data Collection*, Chemistry Data Series, Vol. I, Parts 1-8, DECHEMA, Frankfurt/Main, Germany (1997–1990).

Hansen, E. R., *Global Optimization Using Interval Analysis*, Marcel Dekkar, New York (1992).

Kearfott, R. B., *Rigorous Global Search: Continuous Problems*, Kluwer, Dordrecht, The Netherlands (1996).

Kearfott, R. B., and M. Novoa III, "Algorithm 681: INTBIS, a Portable Interval-Newton / Bisection Package," *ACM Trans. Math. Soft.*, **16**, 152 (1990).

Kim, I., M. Liebman, and T. F. Edgar, "Robust Error-in-Variables Estimation Using Nonlinear Programming Techniques," *AIChE J.*, **36**, 985 (1990).

Luus, R., and H. Hernaez, "Parameter Estimation for Errors-in-Variables Data," *Hung. J. Ind. Chem.*, **28**, 201 (2000).

Neumaier, A., *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge, U.K. (1990).

Patino-Leal, H., and P. M. Reilly, "Statistical Estimation of Parameters in Vapor-Liquid Equilibrium," *AIChE J.*, **28**, 580 (1982).

Rod, V., and V. Hancil, "Iterative Estimation of Model Parameters When Measurements of All Variables Are Subject to Error," *Comput. Chem. Eng.*, **4**, 33 (1980).

Schwetlick, H., and V. Tiller, "Numerical Methods for Estimating Parameters in Nonlinear Models with Errors in the Variables," *Technometrics*, **27**, 17 (1985).

Seber, G. A. F., *Nonlinear Regression*, Wiley, New York (1989).

Tjoa, I.-B., and L. T. Biegler, "Simultaneous Solution and Optimization Strategies for Parameter Estimation of Differential-Algebraic Equation Systems," *Ind. Eng. Chem. Res.*, **30**, 376 (1991).

Tjoa, I.-B., and L. T. Biegler, "A Reduced Successive Quadratic Programming Strategy for Errors-in-Variable Estimation," *Comput. Chem. Eng.*, **16**, 523 (1992).

Valko, P., and S. Vajda, "An Extended Marquardt-Type Procedure for Fitting Error-in-Variables Models," *Comput. Chem. Eng.*, **11**, 37 (1987).

Vamos, R. J., and C. N. Hass, "Reduction of Ion-Exchange Equilibria Data Using an Error in Variables Approach," *AIChE J.*, **40**, 556 (1994).

Manuscript received June 8, 2001, and revision received Dec. 17, 2001.