



**BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS
FACULTY OF CHEMICAL AND BIOENGINEERING
GEORGE OLAH PHD SCHOOL**

**Application of Interval Methods to
Chemical Engineering Problems**

Extended Abstract of PhD Dissertation

**Author: Baharev Ali, MSc in Chemical Engineering
Supervisor: Rév Endre, DSc of Hung. Acad. Sci.**

Department of Chemical and Environmental Process Engineering

2009

Problem statement and aims

The need of reliably solving large-scale systems of nonlinear equations often arises in the everyday practice of chemical engineering. The following difficulties can occur when performing these calculations with standard methods.

- (a) It remains unknown whether the problem is indeed infeasible or just the initial points were poor if the iteration fails after several attempts due to divergence or oscillation.
- (b) It is hard to prove non-trivial infeasibility of a problem. Such a proof is usually done by exploring the feasibility region with approximate models, if a theoretical methodology exists at all.
- (c) Even if the iteration stops without a warning, the obtained result can be completely wrong due to the finite internal representation of floating point numbers and / or inappropriately chosen stopping criteria.
- (d) Professional simulators can return one solution at a time without indicating possible existence of other solutions. As a consequence, finding multiple solutions / proving uniqueness of a solution requires making a troublesome case study. These difficulties (a)-(d) are known to occur when computing the steady states of azeotropic distillation (especially heterogeneous azeotropic distillation) and reactive distillation columns.

All of the difficulties listed above can be overcome by interval methods. The main disadvantage of the interval methods is the combinatorial explosion because these algorithms involve exhaustive search. In general, their applicability is limited to small problems (say 10 variables and equations). The goal of the research was to identify the bottlenecks and to improve the overall efficiency of the computations while preserving the generality of the interval methods.

Literature review

Design and simulation of phase equilibrium based on separation of chemical components from liquid or gas mixtures such as extraction, absorption, desorption, stripping, and distillation, constitute an important part of the everyday practice in chemical engineering. Computing the steady states of the counter-current multistage processes is equivalent to finding solutions of large scale non-linear equation systems. Although a good deal of effort has been made in constructing efficient and reliable computation techniques, and powerful

results have been achieved, the numerical problems in this field have not yet been perfectly solved. Another problem not mentioned in the previous section arises when computing heteroazeotropic distillation: the isoactivity condition is checked, which is a necessary but not sufficient condition for equilibrium. Therefore an additional phase stability analysis has to be performed in order to avoid false and trivial solutions. The phase stability analysis requires finding the global optimum of a nonconvex function but conventional techniques can only guarantee a local optimum.

The phase stability analysis can be reliably performed by interval methods.¹ Finding the global optimum is also guaranteed by the alpha-BB,² a method based on convex underestimator functions and branch-and-bound technique. The homotopy-continuation methods proved to be robust for phase stability analysis³ although finding the global optimum is not guaranteed. These outstanding results do not change the fact that (i) the solution to be checked for stability is difficult to find when computing steady states of multistage separation processes and (ii) it is hard to prove nontrivial infeasibility if this is the case.

Reliably computing multiple steady states of distillation columns is crucial to their design, simulation, control and operation. Professional simulators can return one solution at a time without indicating possible existence of other solutions. This implies a tiresome and unreliable case study if other possible solutions are needed to be found. The most successful methods in this area seem to be the homotopy-continuation methods.⁴ These methods usually converge and are less sensitive to the initial estimates. Multiple solutions can also be found by them. Their theoretical background is well-established and they have been integrated into professional simulation software (HYSYS®, sparse continuation solver). Numerous examples in which previously unknown multiple steady states were recognized by homotopy-continuation methods are published in the literature.

Despite the above mentioned impressive results, the homotopy-continuation methods cannot guarantee locating all solution or would require drastically increasing computational efforts. Studies on heteroazeotropic distillation report serious numerical difficulties when applying

¹ M. A. Stadtherr, G. Xu, G. I. Burgos-Solorzano, W. D. Haynes. **2007**. *Reliable Computation of Phase Stability and Equilibrium Using Interval Methods*. Int. J. Reliability and Safety, 1, 465–488

² S. T. Harding, C. A. Floudas. **2000**. *Phase Stability with Cubic Equations of State: A Global Optimization Approach*. AIChE Journal, 46 (7), 1422–1440

³ F. Jalali, J. D. Seader, S. Khaleghi. **2008**. *Global solution approaches in equilibrium and stability analysis using homotopy continuation in the complex domain*. Computers and Chemical Engineering, 32, 2333–2345

⁴ D. W. Green, R. H. Perry. **2007**. *Perry's Chemical Engineers' Handbook*, McGraw-Hill, Chapter 13, p. 33

these methods. Nontrivial infeasibility cannot be proved. Homotopy-continuation methods require good understanding of the methods.

Relaxation methods are also known for their robustness. Only one solution can be found with these methods at a time, necessitating a case study to find other possible solutions. These computationally expensive methods tend to be very slow near a solution.

All of the problems mentioned in the *Problem statement and aims* section can be overcome by the interval methods. The interval methods are reliable in the following sense:

- (i) finding all solutions of general nonlinear equations in a given domain is guaranteed; finding the global optimum is guaranteed in case of optimization;
- (ii) nontrivial infeasibility of the problem is verified;
- (iii) the convergence is guaranteed; initial point / estimate to start the iteration is superfluous;
- (iv) problems arising from the finite internal representation of the floating-point numbers (such as round-off errors, under- and overflow) are automatically handled by these methods.

Reliability of the interval methods is assured by their construction. All variables and parameters are enclosed by intervals. All real arithmetic operations and real functions are extended in such a way that the true range of the result is enclosed by the computed interval. This construction makes it possible to discard only those parts of the search space which cannot contain a solution.

Interval methods have been successfully applied to solve a wide variety of chemical engineering problems⁵ such as computation of phase stability with activity coefficient models, cubic equation-of-state (EOS) models, modeling liquid-liquid equilibrium of ionic liquid systems, calculation of critical points from cubic EOS models, location of azeotropes, parameter estimation using standard least squares and error-in-variables. Interval arithmetic can also be applied to compute validated solutions of initial value problems for ODEs, to enclose all solutions of two-point boundary value problems for ODEs, and to deterministic global optimization of nonlinear dynamic systems.

The most expensive and wearing part of the research and development is carrying out the experiments. The importance of the design of experiments is obvious. Algorithms for

⁵ Y. Lin, C. R. Gwaltney, M. A. Stadtherr. **2006.** *Reliable Modeling and Optimization for Chemical Engineering Applications: Interval Analysis Approach.* Reliable Computing, 12, 427–450

computing minimal detectable differences for general ANOVA models are known to return completely incorrect results due to the finite internal representation of floating point numbers. The user is not warned about this kind of errors that makes it a wicked problem. Interval methods are ideal to compute reliable reference values since they automatically perform error analysis, *i.e.* they also return an error bound of the computed result.

The use of the interval methods is limited by their computational demand, often resulting in unacceptable computational time even for small problems. This ineffectiveness of the interval methods is partly due to the so-called dependency problem: dependence between the variables and the computed values are neglected causing overestimated / loose enclosures of the ranges which in turn makes the pruning of the search space unproductive. Decomposition⁶ and consistency techniques⁷ are methods to reduce the unwanted effect of the dependency problem. Another new method to handle the dependency is the affine arithmetic, which automatically keeps track of the first-order correlation between the original variables and the computed results.

Solving nonlinear systems of equations with interval methods usually involves computing a linear enclosure of the original nonlinear system. From now on this linear enclosure is called ‘the linearized system’ for short. The solution set of the linearized system encloses the solution set of the original nonlinear system. The solution set of the linearized system is a complicated nonconvex set if the linearization is computed with the interval Newton method. In contrast, the corresponding solution set is convex if the linearization is performed with affine arithmetic. The convexity of the linear enclosure has advantages, for example linear programming is directly applicable to compute sharp bounds on each variable.

Applied methods

Computation of vapor-liquid equilibrium cascades were realized by simultaneously solving the MESH equations (component **M**aterial balance equations, vapor-liquid **E**quilibrium relationships, **S**ummation equations, **H**eat balance equations) with interval methods. Liquid-liquid phase split calculations were performed by directly solving the isoactivity condition and the component balance equations.

⁶ T. Beelitz, A. Frommer, B. Lang, P. Willems. **2005**. *Symbolic-Numeric Techniques for Solving Non-linear Systems*. PAMM, 5, 705–708

⁷ E. R. Hansen, G. W. Walster. **2004**. *Global Optimization Using Interval Analysis*. Marcel Dekker, Chapter 10.

As discussed in the previous section, the use of the interval methods is limited by their computational demand. Affine arithmetic was chosen from the general and automatable enhancements published in the literature. Affine arithmetic was implemented in C++, following the instructions of the excellent monograph of Stolfi and Figueiredo.⁸ The more recent results of Professor Kolev⁹ were also incorporated. The affine linearization technique was compared with the widely used interval Newton method. The general purpose interval Newton method implemented in *C-XSC* was used for this comparison.

The equations describing the steady state of equilibrium cascades are coupled in a special way by certain variables. Linear programming was chosen in order to automatically cope with this dependence between the equations. The direct application of linear programming is enabled by affine arithmetic. The linear programming subproblems were solved with GNU GLPK written in ANSI C.

Results

Implementation of a new method for solving systems of equations

A general purpose interval method is proposed for solving nonlinear systems of equations. Key elements of the method are the linearization technique based on the hybrid affine and interval arithmetic model⁸ (pp. 75-76) and the pruning technique based on linear programming (LP pruning). Several questions about the implementation had to be answered, especially those about the efficient LP pruning. The conclusions are given at *Thesis 1*.

Computing equilibrium units and cascades

Two linearization techniques are compared by computing vapor-liquid equilibrium cascades and liquid-liquid phase split problems. One of the linearization techniques is the interval Newton method

$$f(x) \in \mathbf{A}(x-z) + f(z), \quad \forall x \in \mathbf{X},$$

where $f(x)$ is the $\mathbf{R}^n \rightarrow \mathbf{R}^n$ function to be linearized, the \mathbf{A} coefficient matrix is an interval matrix, x is an arbitrary real vector in the interval vector \mathbf{X} , z is a fixed inner point of \mathbf{X} . This

⁸ L. H. Figueiredo, J. Stolfi. **1997**. *Self-Validated Numerical Methods and Applications*. Brazilian Mathematics Colloquium monographs, IMPA/CNPq, Rio de Janeiro, Brazil

⁹ L. V. Kolev. **2004**. *An improved interval linearization for solving non-linear problems*. Numerical Algorithms, 37, 213–224

yields a linearized system which has a nonconvex solution set, and the dependence between the computed results is neglected.

The other linearization technique used in this study is based on affine arithmetic

$$f(x) \in A x + \mathbf{b}, \quad \forall x \in \mathbf{X}$$

where the coefficient matrix A is a real matrix in contrast with the interval Newton method, and \mathbf{b} denotes an interval vector. Affine arithmetic automatically takes care of the dependence between the original and computed results, and the linear enclosure is convex. The outcome of the research is summarized at *Thesis 2*.

The equations describing the steady state of the equilibrium cascades involve a special dependence. This dependence is automatically handled during the pruning step by linear programming. The linear programming subproblems are as follows (it seems as if there were $2n$ of them):

$$\min / \max x_i \quad \text{for each } i$$

s.t.

$$A x + \mathbf{b} = 0$$

$$x \in \mathbf{X}.$$

The goal is to study the effectiveness of this LP pruning method for solving the MESH equations if the implementation of the LP pruning is appropriate and the affine linearization technique is used. The results are presented at *Thesis 3*.

Solving the MESH equations of distillation columns with interval methods has not been reported previously, most probably because of the extensive complexity and dimensionality of these problems. Efficiency of the affine linearization technique with LP pruning was examined. The computational experience is given at *Thesis 4*.

Recognizing multiple steady states of distillation columns is critical to their design, simulation, control and operation. In practice, the liquid flow rates are usually specified on a mass or volume basis. The transformation of the mass flow rates to molar flow rates depends on the composition profile of column and is nonlinear. This transformation can become singular causing multiplicity and instability even in the case of ideal two-product distillation. Multiplicity can still occur if the flow rates are given on a molar basis. As shown by the studied examples, this can be the case if molar boilup (which is practically proportional to the heat input to the reboiler) and the molar flow rate of the reflux are specified, and the heat

balance equations are included in the model. Applicability of the proposed nonlinear system solving method to find multiple solutions is studied; this is further discussed at *Thesis 4*.

Reliable reference values for general ANOVA models

The F test statistic used in ANOVA follows non-central F distribution, $F_0 = F_{nc}(v_1, v_2, \lambda)$, where v_1 and v_2 are the degrees of freedom of the nominator and the denominator, respectively, of the F test statistic; λ is the non-centrality parameter. This non-centrality parameter is related to the size of effects. The probability of Type II error β (not detecting an effect) is

$$\beta = P[F_0 \leq F_\alpha(v_1, v_2)] = P[F_{nc}(v_1, v_2, \lambda) \leq F_\alpha(v_1, v_2)]$$

where α is the significance level of the test. The power calculation consists of calculating the power ($Power = 1 - \beta$) for certain of effect sizes. When it is used in the reversed way, the power is fixed (e.g. $Power = 0.9$), and the size of the effect is calculated, this is considered as effect of detectable size.

Detectable differences for a specified β are determined by the noncentrality parameter λ for which the cdf value of the noncentral beta distribution equals β

$$I_{x_\alpha}(a, b, \lambda) = \beta,$$

where x_α is the α quantile of the central beta distribution with shape parameters $a = v_1/2$ and $b = v_2/2$; α denotes the allowed type I error probability.

The following closed formulae are utilized for the computation of the cdf of the central beta distribution

$$I_x(a, b) = x^a \left(1 + \sum_{n=1}^{b-1} \left(\prod_{m=1}^n \frac{a+m-1}{m} \right) (1-x)^n \right),$$

and

$$I_x(a, b; \lambda) = e^{-(\lambda/2)(1-x)} \sum_{i=0}^{b-1} \frac{((\lambda/2)(1-x))^i}{i!} I_x(a+i, b-i)$$

for the noncentral beta distribution. The shape parameter b must be integer. It is interesting to note that the evaluation of the above expressions is possible in finite steps, requiring only the four arithmetic operations, the power function, and the exponential function.

Rigorous enclosures of λ are computed with interval Newton method using automatic differentiation. The interval Newton method is able to bound all zeros of a continuously differentiable function in a given interval (initial interval) with mathematical certainty. *Thesis 5* is about the applicability of the above formulae together with the interval Newton method for computing reliable reference values.

Theses

Thesis 1. A new general-purpose nonlinear system solving method has been proposed [1, 2, 5, 6, 7, 8] based on affine arithmetic and linear programming. The implementation is written in C++ programming language. The proposed method (*a*) do not require an initial point to start the iteration, (*b*) the convergence is guaranteed, (*c*) is able to prove infeasibility, (*d*) is able to find multiple solutions. The method is fairly general, and is applicable to a wide variety of engineering problems.

Thesis 2. Vapor-liquid equilibrium cascades and liquid-liquid phase split problems are solved with the affine linearization technique and with the general-purpose interval Newton method. In the studied cases, the affine linearization, compared to the interval Newton method, proved to be an order of magnitude faster on the simpler problems and orders of magnitude faster on the more complicated problems [2, 5, 6, 7, 8].

Thesis 3. Steady state of the vapor-liquid equilibrium cascades are computed using the pruning technique based on linear programming. This pruning technique is efficient and can be recommended to compute cascades if the implementation is appropriate [2, 5, 6, 7, 8]. To the authors' best knowledge, computation of equilibrium cascades with interval methods has not yet been considered in the literature.

Thesis 4. The MESH equations of an extractive distillation column are successfully solved with the proposed method in acceptable computation time [1, 5, 6, 7]. The convergence is guaranteed, initial estimation of the column profile is not needed. If no solution exists than this information is provided by the solver as a result. It is possible to give (redundant) interval specifications. To the authors' best knowledge, computation of distillation columns with interval methods has not yet been considered in the literature.

Reliably computing multiple steady states of distillation columns is crucial to their design, simulation, control and operation. Even ideal two-product distillation columns can have multiple steady states (output multiplicity). The proposed method finds all of them in one run, without any user intervention [5, 6, 7].

Thesis 5. Numerical difficulties associated with the determination of minimal detectable differences in general ANOVA designs raise the overriding importance of correct values with guaranteed accuracy. A new and easy-to-implement self-verifying algorithm is presented for the computation of intervals containing the true values with mathematical rigor, if the degree of freedom of the denominator of the F test statistic is even. The proposed algorithm makes it possible to systematically test the accuracy of existing scalar algorithms in an automated manner. Numerical examples are presented [3, 4, 9, 10, 11].

Possible applications

The proposed method for solving systems of nonlinear equations is fairly general and is applicable to a wide range of engineering / scientific problems. Hooking the solver to a modeling environment would make its use much easier, especially for those who are not familiar with C++ programming and / or interval methods.

Computing vapor-liquid-liquid equilibrium (VLLE) cascades is drastically more difficult than computing vapor-liquid equilibrium cascades. Applying interval methods in this area is an ongoing international research. This work may also result in a better understanding of the different linearization techniques, *i.e.* which technique is better in what scenario.

Thesis 5 demonstrates that the accuracy of different floating-point algorithms cannot be judged by comparing their returned values to other unreliable ones. The easy implementation of the proposed method makes it an excellent tool for computing reliable reference values and as a consequence for testing other algorithms.

Publications on the subject of the dissertation

Refereed articles

- [1] **A. Baharev**, T. Achterberg, E. Rév; *Computation of an extractive distillation column with affine arithmetic*; AIChE Journal, *in press* (IF: 1.607)
- [2] **A. Baharev**, E. Rév; *Reliable Computation of Equilibrium Cascades with Affine Arithmetic*; AIChE Journal, 2008, **54** (7), 1782–1797 (IF: 1.607)
- [3] **A. Baharev**, E. Rév; *Rigorous enclosures of minimal detectable differences for general ANOVA models*; submitted to Reliable Computing
- [4] **A. Baharev**, S. Kemény; *On the computation of the noncentral F and noncentral beta distribution*; Statistics and Computing, 2008, **18** (3), 333–340 (IF: 1.136)

Presentations

- [5] **Baharev A.**; *Intervallum módszerek alkalmazása vegyészmérnöki számításokban*; az MTA Vegyipari Műveleti Munkabizottságának, a Műszaki Kémiai Komplex Bizottságának és a Magyar Kémikusok Egyesülete Műszaki Kémiai Szakosztályának együttes ülése; Veszprém, 2009. április 23.
- [6] **Baharev A.**; *Intervallum módszerek alkalmazása vegyészmérnöki számításokban*; Oláh György Doktori Iskola VI. konferenciája, Budapest, 2009. február 4.
- [7] **A. Baharev**, E. Rév; *Comparing inclusion techniques on chemical engineering problems*; 13th GAMM - IMACS International Symposium on Scientific Computing, Computer Arithmetic, and Verified Numerical Computations SCAN'2008; El Paso, Texas, USA, Sept 29 - Oct 3, 2008; pp. 17–18.
- [8] **Baharev A.**, Rév E.; *Egyensúlyi egységek és kaszkádok számítása affin aritmetikával*; Műszaki Kémiai Napok'07, Veszprém, 2007. április 25–27. 105–107. o.
- [9] **Baharev A.**, Kemény S.; *Nemcentrális F-eloszlás számításához kapcsolódó numerikus problémák*; IV. Alkalmazott Informatika Konferencia, Kaposvár, 2005. május 27.
- [10] **Baharev A.**; *Számítások nemcentrális F-eloszlással*; XXVII. Országos Tudományos Diákköri Konferencia; FiFöMa szekció, Valószínűségszámítás, statisztika és pénzügyi matematika tagozata; 186. o.; Témavezető: Kemény Sándor; III. helyezés, kiemelt dícséret; Budapest, 2005. március 21–23.
- [11] **A. Baharev**; Conference of MSc Students; *On Computing the noncentrality parameter of the noncentral F-distribution*; Supervisor: S. Kemény; Periodica Polytechnica Ser. Chem. Eng. **48** (2), pp. 119–120, 2004

Presentations not related to the dissertation

- [12] **Baharev A.**, Frits E., Lelkes Z., Rév E.; *Megbízható fázisegyensúlyi számítások*; Műszaki Kémiai Napok'06, Veszprém, 2006. április 25–27. 288–289. o.
- [13] **A. Baharev**, E. R. Frits, Cs. Stéger, Z. Lelkes, E. Rév; *Application of interval arithmetics for exploring feasibility region of extractive distillation*; 10. International Workshop on Chemical Engineering Mathematics; Budapest, Hungary, Aug 18–20, 2005
- [14] E. R. Frits, **A. Baharev**, Z. Lelkes, M. Markót, Z. Fonyó, E. Rév, T. Csendes; *Feasibility Study by interval arithmetics: Application of interval arithmetics for exploring feasibility of extractive distillation variants*; International Workshop on Global Optimization; Almería, Spain, Sep 18–22, 2005 (G05); in Proceedings of the International Workshop on Global Optimization; Ed. I. García et al., pp. 103–108, 2005
- [15] Frits E. R., **Baharev A.**, Rév E., Lelkes Z., Markót M., Csendes T.; *Intervallum-aritmetika alkalmazása vegyipari számítási feladatok megoldására*; Műszaki Kémiai Napok'05, Veszprém, 2005. április 26–28. 216. o.