Numerical Aspects of Phase Equilibrium Calculations with the Cubic and Association Models

Liang, Xiaodong

Published in:
Industrial and Engineering Chemistry Research

Link to article, DOI:
10.1021/acs.iecr.8b03956

Publication date:
2018

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
Numerical aspects of phase equilibrium calculations with the cubic and association models

Xiaodong Liang

Ind. Eng. Chem. Res., Just Accepted Manuscript • DOI: 10.1021/acs.iecr.8b03956 • Publication Date (Web): 28 Sep 2018

Downloaded from http://pubs.acs.org on September 29, 2018
Numerical aspects of phase equilibrium calculations

with the cubic and association models

Xiaodong Liang*

Center for Energy Resources Engineering (CERE), Department of Chemical and Biochemical Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

Keywords: PT Flash, Stability analysis, Successive Substitution, Second-order method, EOS
Abstract

The isobaric-isothermal phase equilibrium (PT Flash) calculation has been an active research topic of thermodynamics for decades. In this work, the conventional framework of the PT Flash calculation, consisting of stability analysis and phase-split calculation, is briefly reviewed by giving the key working equations of the first- and second-order methods. With different type of equations of state, the numerical aspects of the PT Flash calculation have been systematically investigated for various systems over a wide range of conditions: the significance of the first-order methods, volume based versus pressure based second-order methods, a safe-unstable criterion in stability analysis, comparisons of different models and modelling approaches, as well as an iterative reuse of the converged volume as an initial guess in the volume root solver. Moreover, the same numerical algorithm is used in the second-order methods for both volume and pressure based stability analysis as well as pressure based phase-split calculation for fair comparisons to the largest possible extent. The results reveal that a few iterations of the pressure based first-order method will significantly improve the efficiency of stability analysis, and it is not more efficient to use a volume based second-order method from an overall point of view. A volume based second-order method can improve the efficiency of phase-split calculation, of which the extent depends on the systems and models. This study also shows that the efficiency deterioration of using association models compared to cubic ones is moderate.
1. Introduction

The isobaric-isothermal phase equilibrium (PT Flash) calculation is one of the most important one-stage equilibrium problems. A PT Flash calculation is usually needed when the physical properties of a system at given temperature and pressure are required. The PT Flash calculation in principle is a global minimization problem, as the objective is to find the number of phases and their compositions and amounts that represent the minimum Gibbs free energy of the system. The ground-breaking technology of the PT Flash calculation is from Michelsen\textsuperscript{1,2}, of which the framework consists of stability analysis, checking if the Gibbs free energy of the system can be decreased by adding a new phase, and phase-split calculation, determining the phase types, amounts and compositions that give a (local) minimum of the Gibbs free energy. Even though there are some other frameworks for solving the PT Flash problem\textsuperscript{3,4,5}, most of the research on this topic has been following the idea of Michelsen\textsuperscript{6-19}. This will be the framework used in this work as well.

A typical overall procedure of a PT Flash calculation is presented in Figure S1 in the Support Information. The PT Flash calculation may start from either stability analysis or equilibrium K-factors (see the definition in the Theory section 2.2). If the calculation starts from equilibrium K-factors, it is customary to obtain the initial values from the Wilson equation when there is no advanced information available. In this work, however, the calculation starts from stability analysis. If the stability analysis shows that the system is unstable, the output is used to generate the initial estimates of K-factors for phase-split calculation. The phase-split calculation then determines the number, types, fractions and compositions of the phases at equilibrium. When the phase-split calculation finishes, the stability of the new formed system will be analyzed again, and the procedure iterates. In principle, this iteration loop shall continue until a stable system is revealed by stability analysis, while in practice it would be wise to terminate the iteration loop by setting a maximum allowed number of phases in case any numerical failure occurs. Since the works of Michelsen\textsuperscript{1,2}, numerous studies have been devoted to this topic. Some developed local minimization algorithms and some others proposed global minimization methods for stability analysis, phase-split calculation or both\textsuperscript{6-30}. One of the best references on the PT Flash calculation shall be the book from Michelsen and Mollerup\textsuperscript{31}. It is generally considered that local minimization algorithms are more efficient, while the global minimization ones are safer.
According to the previous experience, a well-organized local minimization algorithm will be very robust for general PT Flash problems. Hence, a local minimization approach is used in this work.

The research on using volume instead of pressure as one of the specifications has been becoming a hot topic for stability analysis and/or equilibrium calculations. Michelsen presented the formula and commented the use of volume based method in stability analysis. Nagarajan et al. presented a volume based method in terms of molar densities, and demonstrated its robustness in difficult PT Flash problems. Nichita and co-workers have developed global optimization as well as local minimization methods for volume based stability analysis. Mikyška and co-workers have derived the volume based stability criterion, developed solution methods and applied them into isochoric-isothermal equilibrium problems. Castier presented the use of a combination of a local minimization algorithm and a global optimization method for volume based stability analysis and how the stability analysis results can be used as initial estimates for the subsequent phase-split calculation. It can be seen that the use of volume based method for the PT Flash calculation is still limited. On one hand, as pointed out by Michelsen, the volume based formula is identical to the pressure based formula for determining the stability for a PT (pressure and temperature) specified stability analysis problem. As shown by Paterson et al., on the other hand, a volume based method can be used to solve the PT phase-split calculation. In this work, the usage of volume based second-order (SO) methods for both stability analysis and phase-split calculation is investigated and compared with the traditional approach, i.e. pressure based methods. The computational efficiency and the overall strategy for a general PT Flash calculation are the main focuses. In order to have ‘fair’ comparisons to the largest possible extent, the working equations are formulated in a way that the same numerical algorithm and different type of equation of state (EOS) models can be used for the same systems.

The rest of this work is organized as follows: the key working equations for both stability analysis and phase-split calculation are briefly presented in Section 2; the models and systems are introduced in Section 3; the results are presented and discussed in Section 4, followed by the conclusions in Section 5.
2. Theory

2.1 Stability analysis

The necessary and sufficient condition for the stability of a phase at given temperature and pressure is that the tangent plane distance is non-negative for any trial phase composition.

\[ tpd(T, P, \mathbf{x}) = \sum_i x_i \left[ \frac{\mu_i(T, P, \mathbf{x})}{RT} - \frac{\mu_i(T, P, \mathbf{w})}{RT} \right] \]  \hspace{1cm} (1)

where \( T, P \) and \( R \) are temperature, pressure and gas constant, respectively. \( \mathbf{x} \) and \( \mathbf{w} \) are the mole fractions of the trial and tested phase, respectively. \( \mu_i \) is the chemical potential of component \( i \), which is given by

\[ \mu_i(T, P, \mathbf{x}) = \mu_i^*(T, P^*) + RT \ln \frac{f_i}{\Phi_i} = \mu_i^*(T, P^*) + RT \left[ \ln x_i + \ln \frac{P}{\Phi_i} + \ln \varphi_i(T, P, \mathbf{x}) \right] \]  \hspace{1cm} (2)

where \( \mu_i^* \) is the chemical potential of component \( i \) in the ideal gas state at the reference pressure \( P^* \), \( f_i \) and \( \varphi_i \) are respectively the fugacity and fugacity coefficient of component \( i \). The \( \varphi_i \) is a homogeneous function of degree zero in number of moles and is calculable

\[ \ln \varphi_i(T, P, \mathbf{n}) = \left[ \frac{\partial A^r(T, V, \mathbf{n})}{\partial n_i} \right]_{T,V} \]  \hspace{1cm} (3)

where \( A^r, V \) and \( Z \) are the residual Helmholtz free energy, total volume and compressibility factor, respectively. \( \mathbf{n} \) is the corresponding vector of \( n_i \), which is the number of moles of component \( i \), and apparently has the following relationship to the mole fraction \( x_i \),

\[ x_i = \frac{n_i}{\sum n_i} = \frac{n_i}{n} \]  \hspace{1cm} (4)

where \( n \) is the total number of moles of the given phase.

In terms of composition and fugacity coefficient, equation (1) can be rewritten
\[ tpd(T, P, x) = \sum_{i} x_i \left[ \ln x_i + \ln \phi_i(T, P, x) - \ln w_i - \ln \phi_i(T, P, w) \right] \]
\[ = \sum_{i} x_i \left[ \ln x_i + \ln \phi_i(T, P, x) - d_i \right] \]  

(5)

where \( w_i \) is the mole fraction of component \( i \) in the tested phase, and \( d_i \) is an auxiliary variable.

The computational approach for stability analysis proposed by Michelsen \(^1\,^3\,^1\) is based on the fact that the tangent plane distance is non-negative everywhere if and only if it is non-negative at all of its stationary points, and the practical approach is to locate the local minima by unconstrained optimization methods of the objective function

\[ tm(T, P, n) = 1 + \sum_{i} n_i \left[ \ln n_i + \ln \phi_i(T, P, n) - d_i - 1 \right] \]  

(6)

The stationary points of \( tm \) satisfy

\[ \frac{\partial tm}{\partial n_i} = \ln n_i + \ln \phi_i(T, P, n) - d_i = 0 \]  

(7)

This immediately gives the successive substitution (SS) procedure

\[ \ln n_i = -\ln \phi_i(T, P, n) + d_i \]  

(8)

For a SO method, Michelsen \(^1\,^3\,^1\) proposed to use

\[ \alpha_i = 2 \sqrt{n_i} \]  

(9)

The gradient and Hessian of the objective function equation (6) can then be written

\[ g_i = \frac{\partial tm}{\partial \alpha_i} = \frac{\alpha_i}{2} \left[ \ln n_i + \ln \phi_i(T, P, n) - d_i \right] \]  

(10)

\[ H_{ik} = \frac{\partial^2 tm}{\partial \alpha_i \partial \alpha_k} = \frac{g_i}{\alpha_i} \delta_{ik} + \delta_{ik} + \frac{\alpha_i \alpha_k}{2} \frac{\partial \ln \phi_i}{\partial n_k} \]  

(11)

where \( g_i \) and \( H_{ik} \) are elements of the gradient \( g \) and Hessian \( H \), respectively.
The update of the independent variables $\alpha$ is obtained from

$$\Delta \alpha = -H^{-1} g$$  \hspace{1cm} (12)

More details are given below in Section 2.3.2 how to solve this set of linear equations.

Equivalently, equation (1) can be written

$$tpd(T, V, n) = \sum_i n_i \left[ \frac{\mu_i(T, V, n)}{RT} - \frac{\mu_i(T, v_0, w)}{RT} \right] - V \left[ \frac{P(T, V, n)}{RT} - \frac{P_0}{RT} \right]$$  \hspace{1cm} (13)

where $V$ is the total volume, $v_0$ and $P_0$ are respectively the molar volume and pressure of the tested phase. As pointed out by Michelsen $^{1,31}$ and discussed by Castier $^{36}$, this tangent plane distance is a homogeneous function of degree one in $(V, n)$, and it allows people to replace $V$ by a fixed value and treat $n$ as the only independent variables. Mikyška et al. $^{34}$ and Nichita $^{38}$ discussed that the volume based SS method is neither robust nor efficient to solve equation (13).

By using the same variables defined in equation (9), we obtain

$$g_i = \frac{\alpha_i}{2} \left[ \frac{\mu_i(T, V, n)}{RT} - \frac{\mu_i(T, v_0, w)}{RT} \right]$$  \hspace{1cm} (14)

$$H_{ik} = \frac{g_i}{\alpha_i} \delta_{ik} + \frac{\alpha_i \alpha_k}{2} \left( \frac{\partial \mu_i}{\partial n_k} \right)_{T, V, n_i \neq k}$$  \hspace{1cm} (15)

Equation (12) can be used to obtain the updates of variables $\alpha$ at each iteration step.

The main procedure for stability analysis is given in Figure S2 in the Support Information. Firstly, the value of $d_i$ in equation (5) is setup for each component in the tested phase, and the initial estimates of the mole fractions for a selected trial phase are generated. A certain number of SS steps, equation (8), is conducted and the SO method is invoked if the SS method is not converged. It is possible to use either pressure based equations (10) and (11) or volume based equations (14) and (15) in the SO method, in which the same numerical solution method is adopted. Several trial phases are tried (more details given in the Section 4.2), and the results with the minimum objective function will be chosen for checking the stability.
### 2.2 Phase-split calculation

With a known number of phases, the objective function for phase-split calculation at constant temperature and pressure is the reduced Gibbs free energy. It is commonly formulated

\[
Q_P = \sum_{j=1}^{n} \sum_{i=1}^{c} n_{i,j} \ln f_{i,j}(T, P, n_j)
\]  
(16)

with the mass balance

\[
z_i = \sum_{j=1}^{n} n_{i,j} = \sum_{j=1}^{n} \beta_j x_{i,j}
\]  
(17)

where \(Q_P\) is the objective function for the phase-split calculation expressed in temperature, pressure and number of moles. \(n\) and \(c\) are number of phases and components, respectively. \(\beta_j\) is the phase fraction of phase \(j\). In this work, \(j\) represents the phase index, and it is separated from the component index \(i\) by a comma.

At equilibrium, \(Q_P\) reaches to the minimum state, and we have

\[
\ln f_{i,j}(T, P, n_j) - \ln f_{i,r}(T, P, n_r) = 0
\]  
(18)

where \(r\) is the index of the ‘reference’ phase for component \(i\), and it is commonly to choose the last phase or the phase with the largest amount as the ‘reference’ phase.

This is the equation of fugacity equality at the equilibrium condition and it is the basis for the SS method for phase-split calculation. First, we transfer this equation to

\[
[\ln x_{i,j} + \ln \varphi_{i,j}(T, P, n_j)] - [\ln x_{i,r} + \ln \varphi_{i,r}(T, P, n_r)] = 0
\]  
(19)

The K-factors are defined

\[
K_{i,j} = \frac{x_{i,j}}{x_{i,r}}
\]  
(20)

When the index of the reference phase \(r\) is the same for all components, we apparently have
\[
\sum_{i=1}^{c} (x_{i,j} - x_{i,r}) = 0 \tag{21}
\]

With the assistance of equations (17) and (20), equation (21) can be organized into the famous Rachford-Rice equation, which can be used to solve the phase fractions. And the composition of each phase can be calculated afterwards. In this work, the normal Rachford-Rice equation is used for two phases, while the modified version proposed by Michelsen is adapted for multiple phases. More details of this method can be found in the book of Michelsen and Mollerup \(^3\). A new set of K-factors can then be obtained from fugacity coefficients,

\[
K_{i,j} = \frac{\Phi_{i,r}}{\Phi_{i,j}} \tag{22}
\]

from which new phase fractions and compositions can be calculated. The procedure continues until K-factors do not change or a certain number of iterations is reached.

A SO method may improve the efficiency significantly under certain conditions. With the gradient equation (18) available, the Hessian is given

\[
\frac{\partial^2 Q_p}{\partial n_{i,j} \partial n_{k,m}} = \frac{\partial \ln f_{i,j}}{\partial n_{k,j}} \delta_{jm} + \frac{\partial \ln f_{i,r}}{\partial n_{k,r}} \tag{23}
\]

In order to better scale the problem, as chosen by Michelsen \(^3\), the following independent variables are used

\[
\theta_{i,j} = \frac{n_{i,j}}{z_i} \tag{24}
\]

A similar equation to equation (12) is used to calculate the update of the independent variables \( \theta \), which has a size of \((n - 1) \times c\).

In terms of \( T, V \) and \( \mathbf{n} \), the objective function in equation (16) can be equivalently written

\[
Q_V = \sum_{j=1}^{n} \left[ \sum_{i=1}^{c} n_{i,j} \frac{\mu_{i,j}(T, V_j, \mathbf{n}_j)}{RT} - \frac{P_j - P_0}{RT} V_j \right] \tag{25}
\]
where $Q_\gamma$ is the objective function for the phase-split calculation expressed in temperature, total volume and number of moles, $P_j$ and $P_0$ are the pressure of phase $j$ and the specified pressure, respectively. The mass balance is the same one given in equation (17).

In principle, it is possible to solve the problem following a similar procedure given above as the pressure based method, but Paterson and Michelsen et al. proposed a new approach, which will be followed in this work. At equilibrium, $Q_\gamma$ reaches to the minimum state, and we have

$$\frac{\mu_{i,j}(T, V_j, n_j)}{RT} - \frac{\mu_{i,r}(T, V_r, n_r)}{RT} = \frac{\mu_{i,j}(T, V_j, n_j)}{RT} - \frac{\mu_{i,0}}{RT} = 0$$ (26)

$$\frac{P_j(T, V_j, n_j)}{RT} - \frac{P_0}{RT} = 0$$ (27)

where the equilibrium chemical potential of each component is introduced by the subscript $0$.

Multiplying number of moles to equation (26) and phase volume to equation (27), we obtain

$$\sum_{i=1}^{c} n_{i,j} \frac{\mu_{i,0}}{RT} = \sum_{i=1}^{c} n_{i,j} \frac{\mu_{i,j}(T, V_j, n_j)}{RT} - \frac{P_j - P_0}{RT} V_j$$ (28)

This is the reduced Gibbs free energy of phase $j$.

The changes of the number of moles among different phases in the system have to fulfill

$$\sum_{j=1}^{n} \Delta n_j = 0$$ (29)

In order to obtain $\Delta n_j$, equation (26) is linearized

$$\frac{\mu_j}{RT} + \frac{\partial(\mu_j/RT)}{\partial n_j} \Delta n_j + \frac{\partial(\mu_j/RT)}{\partial V_j} \Delta V_j - \frac{\mu_0}{RT} = 0$$ (30)

in which way we can have

$$\Delta n_j = \left[ \frac{\partial(\mu_j/RT)}{\partial n_j} \right]^{-1} \left( \frac{\mu_0}{RT} - \frac{\mu_j}{RT} \right) + \frac{\Delta V_j}{V_j} n_j$$ (31)
where the homogeneity of Helmholtz free energy (chemical potential) has been used, i.e.

\[
\frac{\partial (\mu_j/RT)}{\partial n_j} n_j + \frac{\partial (\mu_j/RT)}{\partial V_j} V_j = 0
\]  \tag{32}

Substituting equation (31) into equation (29), we obtain

\[
\sum_{j=1}^{n} \left[ \frac{\partial (\mu_j/RT)}{\partial n_j} \right]^{-1} \frac{\mu_0}{RT} + n_j \frac{\Delta V_j}{V_j} = \sum_{j=1}^{n} \left[ \frac{\partial (\mu_j/RT)}{\partial n_j} \right]^{-1} \frac{\mu_j}{RT}
\]  \tag{33}

Now we have \( \Pi \) equations (28) and \( C \) equations (33), and we have \( \Pi + C \) variables \( \Delta V_j/V_j \) and \( \mu_0/RT \). A standard LU decomposition and back substitution method can be used to solve this set of linear equations.

The main procedure of phase-split calculation is given in Figure S3 in the Support Information, which consists of both SS and SO methods. With a given set of K-factors, the phase fractions and compositions are solved, and the fugacity coefficients are calculated for all phases, from which the K-factors are updated. This SS method runs for a number of iterations, and the SO method is activated if the SS method is not converged. One SO method is based on pressure, i.e. equations (18) and (23), for which the same numerical solution method used in stability analysis is adopted. The other SO method is based on volume, i.e. equations (28) and (33), and in this work it switches to the pressure based SO method whenever negative phase volumes or number of moles are obtained.

### 2.3 Numerical aspects

#### 2.3.1 Key working equations

As discussed by Mikyška et al. \(^{34}\) and Nichita \(^{38}\), the volume based SS method does not work for stability analysis and, as pointed out by Paterson and Michelsen et al. \(^{30}\), there is unfortunately no a SS method available for the volume based phase-split calculation. Hence, whenever SS is used in the following discussions, it is referred to the pressure based approach. The key working equations are listed in Table 1. For simpler notations, V-based and P-based are respectively used for the volume and pressure based methods hereafter.
Table 1. Key working equations of P-based SS and V-based and P-based SO methods

<table>
<thead>
<tr>
<th>Order of method</th>
<th>Stability analysis</th>
<th>Phase-split calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>V-based</td>
<td>P-based</td>
</tr>
<tr>
<td>SS</td>
<td>None</td>
<td>Eq. (8)</td>
</tr>
<tr>
<td>SO</td>
<td>Eq. (14) and (15)</td>
<td>Eq. (10) and (11)</td>
</tr>
</tbody>
</table>

2.3.2 The SO numerical method

In the SO methods for stability analysis and the P-based phase-split calculation, the Murray’s method is used to decompose the Hessian matrix when obtaining the searching direction

$$\Delta x = -(H + \eta I)^{-1} g$$  \hspace{1cm} (34)

where $\eta$ is a correction constant when $H$ is not positive definite and $I$ is the unit matrix, and $x$ could be either $\alpha$ in stability analysis or $\theta$ in phase-split calculation. It is a generalized version of equation (12). When the searching direction is ready, the independent variables are updated

$$x^{new} = x^{old} + \lambda \Delta x$$  \hspace{1cm} (35)

When $\eta$ is zero, the full Newton step, i.e. $\lambda = 1$, is tried first. If the objective function cannot be reduced or $\eta$ is not zero, the line search method proposed by Fletcher is used with the full Newton step as the upper bound and $\lambda = 0.5$ as the initial value.

2.3.3 Criteria

Criteria are necessary to check convergence. As a standard choice, on one hand, the Euclidean norm of the gradient is used to check the convergence for the three minimization problems,

$$|g|_2 < e$$  \hspace{1cm} (36)

In stability analysis, the gradient is the right-hand side (RHS) of equations (10) and (14) for the P-based and V-based SO methods, respectively. In the P-based phase-split calculation, the gradient is the RHS of equation (18). On the other hand, the Euclidean norm of the relative change of volume from equations (28) and (33) is used for convergence check in the V-based phase-split calculation. The choice of the value $e$ is discussed below.
It is recommended to monitor whether the objective function has decreased or not during the iteration procedure for a minimization problem. The criterion $10^{-10}$ is used for checking if the objective function is decreased or not for both stability analysis and phase-split calculation. The criterion for instability is $-10^{-8}$. That means, if the minimum value of the tangent plane distance from all trial phases is smaller than $-10^{-8}$, the system is considered unstable and the converged trial phase composition is used to start the phase-split calculation. More discussions are given below regarding the usage of a safe unstable criterion for stability analysis in Section 4.4.

3. Models and systems

The models used in this work are Soave-Redlick-Kwong (SRK)\(^{41}\), Peng-Robinson (PR)\(^{42}\), Cubic Plus Association (CPA), proposed by Kontogeorgis et al.\(^{43}\), and Perturbed-Chain Statistical Associating Fluid Theory (PC-SAFT) EOS\(^ {44,45}\). Two are cubic models and the other two are association ones. The details of these models can be found in the original literature or the recent monograph of Kontogeorgis and Folas\(^ {46}\). All the models are implemented in the form of Helmholtz free energy. It is readily known from Section 2 that the first- and second-order derivatives of the Helmholtz free energy with respect to number of moles are needed for both V-based and P-based SO methods, and the first- and second-order derivatives with respect to volume are also needed for the P-based methods. It is necessary to point out that SRK, PR and CPA share the same code for the cubic part and the volume root solver.

In this work, twelve systems are selected. The first four systems are taken from Nichita\(^ {38}\), which were originally from Mikyška et al.\(^ {34}\). These cases are only used for validating the SO method for stability analysis, and PR is the only model for them. Three models are applied for each of the other systems. As listed in Table 2, these systems cover a wide range of component types: light hydrocarbons, heavy hydrocarbons, (acid) gases, water and large polar molecules (asphaltene). More information of the composition, pure component parameters, binary interaction parameters and conditions is available in the Support Information.
Table 2. Systems investigated in this work

<table>
<thead>
<tr>
<th>System</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1-S4</td>
<td>The examples Ex1, Ex5, Ex6, Ex7 from Nichita $^{38}$</td>
</tr>
<tr>
<td>S5</td>
<td>$C_1$-$nC_{4}$-$nC_{10}$ from Urlic et al. $^{47}$ (gas condensate 2 in the original reference)</td>
</tr>
<tr>
<td>S6</td>
<td>$C_1$-$H_2$-$S$-$CO_2$ from Michelsen $^{1}$</td>
</tr>
<tr>
<td>S7</td>
<td>$N_2$-$C_1$ to $nC_5$ from Avila et al. $^{48}$ (gas 1 in the original reference)</td>
</tr>
<tr>
<td>S8</td>
<td>Normal alkanes from $C_1$ to $nC_8$ plus $nC_{10}$ and $nC_{14}$ (MY10 mixture) with 16.67% CO$_2$ added, from Petitfrere $^{49}$</td>
</tr>
<tr>
<td>S9</td>
<td>$C_1$-$H_2$-$S$-$CO_2$-$H_2$O</td>
</tr>
<tr>
<td>S10</td>
<td>The example 3 from Paterson et al. $^{30}$</td>
</tr>
<tr>
<td>S11</td>
<td>The Fluid-2 from Arya et al. $^{50,51}$</td>
</tr>
<tr>
<td>S12</td>
<td>The Fluid-4 from Arya et al. $^{50,51}$ with 0-20% gas injection</td>
</tr>
</tbody>
</table>

4. Results and discussion

4.1 Algorithm validation

The same numerical algorithm is used in the SO methods for both the V-based and P-based stability analysis and the P-based phase-split calculation. It has been applied to the systems S1 to S4, for which the performance of a V-based SO method is available in details $^{38}$. Using the same initial guess approaches, i.e. assuming the feed as a vapor-like or liquid-like fluid in two separate runs, and the same parameters, i.e. convergence criterion ($\varepsilon=10^{-10}$), maximum iteration number (500), the stability envelopes of the systems S1 to S4 are constructed by the V-based SO method in Figure 1. It can be seen that the results are the same as those reported in the literature $^{38}$.

The convergence behaviors of the V-based and P-based SO methods for the two systems S1 and S4 are compared in Figures 2 and 3, respectively. The same conditions from the literature $^{38}$ have been used in these comparisons. It can be observed that the V-based SO method shows quite similar behavior of the system S1 available in the literature, while the convergence behavior of the system S4 is not easy to see in the literature due to the scale. It is obvious from Figures 2 and 3 that the P-based SO method uses fewer iterations to converge for both systems under all four conditions, which is also consistent with the conclusion from the literature $^{39}$. 

ACS Paragon Plus Environment
It is observed that the maximum iteration numbers in these practices are smaller than 80 and 25 for the V-based and P-based SO methods, respectively. Moreover, it is believed that a convergence criterion of $10^{-7}$ is enough for most practical applications. It can be seen from Figure 4 that the average number of iterations can be decreased 3-9% by changing the convergence criterion from $10^{-10}$ to $10^{-7}$ for these systems. The percentage of decrease or difference is calculated from

$$RD\% = \frac{\Omega_2 - \Omega_1}{\Omega_1} \times 100\%$$

(37)

where $\Omega$ may represent the average number of iterations here or other statistic numbers hereafter, and subscripts 1 and 2 represent the two different scenarios under comparison.

Therefore, the maximum number of iterations and the convergence criterion are respectively set to 150 and $10^{-7}$ in the following discussions, and it needs to be pointed out that these values have been used as default ones in internal applications for quite a long period.
Figure 1. Stability envelopes of the systems S1-S4 from the V-based SO method with no SS steps. The conditions are taken from the literature\textsuperscript{38}.
Figure 2. The convergence behavior of the V-based and P-based $SO$ methods for two conditions of the system S1 at 300K by considering the feed as a liquid-like fluid. The conditions are given in the legend, which are the same ones as those in the literature.$^{38}$
Figure 3. The convergence behavior of the V-based and P-based $SO$ methods for two conditions of the system S4 at 300K by considering the feed as a liquid-like fluid. The conditions are taken from the literature $^{38}$. 

|$\ell_2$ norm of gradient, $|g|_2$ | Iteration number |
|-----------------------------|-----------------|
| $10^{-20}$                  | 2               |
| $10^{-15}$                  | 4               |
| $10^{-10}$                  | 6               |
| $10^{-5}$                   | 8               |
| $10^0$                      | 10              |
| $10^5$                      | 12              |
| $10^{10}$                   | 14              |
| $10^{15}$                   | 16              |
| $10^{20}$                   | 18              |
Figure 4. The percentage of decrease of average number of iterations (one point) by changing the convergence tolerance from $10^{-10}$ to $10^{-7}$ for different methods with different initial guesses for the systems S1 to S4.
4.2 Stability analysis

Stability analysis is in principle a global optimization problem. With local minimization methods, in order to reliably evaluate the stability of a given system as economically as possible, the strategy listed in Table 3 is used for generating initial estimates of the trial phase composition in stability analysis for general applications. It is similar to the procedure proposed by Michelsen\textsuperscript{31}, of which a modified version has also been discussed and used by some researchers\textsuperscript{16,18,19,36}. In this way, there are maximum $C + 3$ trial phases and a good balance between robustness and efficiency is obtained.

Table 3. Strategy for generating initial estimates of the trial phase composition

<table>
<thead>
<tr>
<th>Seq. No. of trial phases</th>
<th>Initial estimates of the trial phase composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Based on the phase type of the tested phase, Wilson K-factors are used: $n_i = w_i \times K_i^W$ or $n_i = w_i/K_i^W$; the mole numbers are normalized.</td>
</tr>
<tr>
<td>2</td>
<td>Based on the phase type of the tested phase, Ideal gas is assumed: $n_i = \exp(d_i)$ or $n_i = 1/\exp(d_i)$; the mole numbers are normalized.</td>
</tr>
<tr>
<td>3</td>
<td>If there are more than one phase, the average phase composition is used: $n_i = \sum_j w_{i,j}/\Pi$; the mole numbers are normalized.</td>
</tr>
<tr>
<td>4 to $C+3$</td>
<td>A near pure liquid phase is used with composition: $n_i = 1 - (C - 1) \times 10^{-8}$, and $n_k = 10^{-8}$ ($k \neq i$)</td>
</tr>
</tbody>
</table>

The stability envelopes for the systems S5 to S12 are presented in the Support Information. They are constructed by running stability analysis at each point over a range of temperature and pressure with steps 0.1K and 0.1bar, respectively, except for the last system S12, in which the temperature is fixed at 363.15K and the pressure and gas inject amount (percentage) are varied. More information of the studied conditions can be found in the Support Information. It is hard to determine if any unstable solution has been missed when people look at a single point only, while there is no single unsmooth behavior observed in any of these stability envelopes. Different combination strategies of SS with both the P-based and V-based $SO$ methods give the same stability envelopes. Therefore, we consider that there are no failures and the algorithms are robust.
It is interesting to point out that on one hand the similarity of the modeling results from different models for a given system depends on the conditions. For instance, different models may give similar bubble or dew points in a narrow range of temperature and pressure. On the other hand, these stability envelopes might provide helpful information in choosing a model and/or tuning model parameters, as the models usually present different stability boundaries over a wide range of temperature and pressure.

It is common practice to apply the \( SS \) method in the PT specified stability analysis \(^{31}\). The efficiency of using different \( SS \) steps has been investigated for the systems S5-S12. The average running time of one point is defined

\[
t = \frac{\text{total running time}}{N_{\text{point}} \times N_{\text{repeat}}} \tag{38}
\]

where \( N_{\text{point}} \) is the number of points and \( N_{\text{repeat}} \) is the repetitions of calculations for each point, \( t \) is the average running time of one point, which includes the running cost from \( SS \) steps and that of the \( SO \) method if there is any.

The results of using 0, 4 and 7 \( SS \) steps before switching to the V-based or P-based \( SO \) method are presented in Figure 5 for the systems S5-S12. As aforementioned, these approaches give the same stability envelopes. The running time statistics depend on the computer, compiler and implementations very much. In this work, the calculations are conducted in a Windows 7 desktop computer with a processor Intel(R) Core™ i5-4590 CPU @ 3.30GHz. The codes are compiled in Intel Fortran 2018 integrated with Visual Studio 2017. There is no third-party library used. The detailed information on temperature, pressure and composition can be found in the Support Information. In all these calculations, \( N_{\text{repeat}} \) is 100, and the statistics were also tried with \( N_{\text{repeat}} \) equal to 200 and 1000. The differences of the average running time of one point among these investigations are within the uncertainty, of which the average value is less than 0.2% with a maximum deviation within 1-2% for stability analysis.

Despite the P-based \( SO \) method converges in fewer iterations as shown in Figures 2 and 3, it can be seen from Figure 5 that the V-based \( SO \) method is actually faster than the P-based \( SO \) method for most cases with only one exception (PC-SAFT in S11). This is because the P-based \( SO \) method needs to solve the volume before calculating the first- and second-order derivatives of
Helmholtz free energy with respect to number of moles. The P-based \( SO \) method becomes more competitive, however, when 4 or 7 \( SS \) steps are introduced before the \( SO \) method. Moreover, there is no doubt that running 4 \( SS \) steps before the \( SO \) method can significantly improve the efficiency. Running 7 \( SS \) steps has almost the same performance for many cases, and it becomes even more pronounced as the number of components increases. Therefore, the combination of 7 \( SS \) steps and the P-based \( SO \) method are used for stability analysis in the following PT Flash calculations.

It is commonly considered that the association models are slower than the cubic ones. Figure 5 shows that the more complex models, CPA and PC-SAFT, cost more in general, while there are some exceptions, for example in the systems S7 and S8. There are no associating fluids in the systems S5-S8, two types of associating sites with CPA-1 and five types of associating sites with CPA-2 in the systems S9 and S10, three types of associating sites with CPA and PC-SAFT in the systems S11 and S12. It is hard to make a completely fair comparison for the computational efficiency among different models over a wide range of systems and conditions, as the stability envelopes are usually different. However, it is meaningful to point out that the running costs of these complex models in stability analysis are on average less than 2 times of those of cubic EOS, as statistically reported in these figures.
Figure 5. Average running time of one point stability analysis versus number of SS steps. The legends are the same for the sub-figures of the systems S5-S8, S9-S10 and S11-S12, respectively. Solid and dash lines are the results with the P-based and V-based SO methods, respectively.
4.3 PT Flash

The phase envelopes of the systems S5 to S12 are presented together with the stability envelopes in the Support Information. These phase envelopes are constructed in the same way as in the stability envelopes. Again, unsmooth behaviors have not been observed in any of these phase envelopes with any of the combination strategies of SS steps and the P-based or V-based SO methods discussed below. Therefore, it is considered that there are no failures and the algorithms are robust. As expected, the phase envelope is the same as the stability envelope when there are no three-phase regions, e.g. the systems S5, S7 and S8. A similar information as revealed from the stability envelopes is obtained: the similarity of the models for a given system largely depends on the conditions. As seen from these figures, different models may give similar phase boundary (dew/bubble points) in a narrow range of temperature and pressures for most of the systems. However, they may show quite different behavior over a wide range of conditions. The most significant cases are S9 and S10, in which SRK and two modeling approaches of CPA are applied. In the first approach, CPA-1, only water is considered as a self-associating fluid, and there is no solvation (cross-association) between water and CO$_2$/H$_2$S. In the second approach, CPA-2, both water and H$_2$S are considered as self-associating fluids, and there are solvation interactions between water and CO$_2$, as well as between water and H$_2$S. The parameters are taken from Paterson et al.\textsuperscript{30}, which were originally developed by Tsivintzelis et al.\textsuperscript{52,53} It is interesting to note that the two CPA approaches present quite similar phase envelopes for the system S10, while they give different stability envelopes and three-phase envelopes respectively in the high temperature and pressure region and in the low to medium temperature and pressure region for the system S9. Moreover, SRK and CPA present significantly different two-phase and three-phase regions in the system S10, in which situation it is necessary to have experimental data in order to select a correct model.

It is common practice to run SS steps for phase-split calculation, of which the purpose could be to solve the problem alone without using a SO method or to provide good initial estimates for a SO method. The influence of the number of SS steps with the P-based SO method is presented in Figure 6. Two-phase and three-phase phase equilibria correspond to the systems S5-S8 and S9-S12, respectively. It turns out, from a statistical point of view, that 4 SS steps give the best performance in most of the investigated cases, and it may not be necessary to take 15 SS steps, as
it deteriorates the performance. Therefore, 4 SS steps are used before switching to the SO method as the default option for phase-split calculation in the following investigations.

The more complex models, CPA and PC-SAFT, cost more time in most cases except the systems S7 and S8 with PC-SAFT and S10 with CPA-1. It is interesting to notice, despite of presenting quite similar results in the system S10, that CPA-2 cost up to 1.5 times of CPA-1, of which the efficiency is comparable to that of SRK. In the asphaltene systems S11 and S12, CPA and PC-SAFT run slower than SRK, but the efficiency deterioration is moderate. As statistically reported in these figures, the running costs of these association models of one point PT Flash calculation are on average less than 2 times of those of cubic models, which is the same behavior observed in stability analysis discussed above.

The V-based and P-based SO methods are compared in Figure 7, which presents the percentage of difference of the average running time of one point. The percentage of difference is calculated via the equation (37), in which the average running time of one point with the P-based SO method is $\Omega_1$, and that with the V-based SO method is $\Omega_2$. As aforementioned, 4 SS steps are used before the SO method in phase-split calculation. Statistically, the figure shows that the V-based SO method is more efficient except for the system S12 with CPA. This is because that the feed composition of the system S12 are changing instead of temperature, and there are failures in the V-based SO method, and these failure points are converged by the P-based SO method. In the other cases, the performance improvement using the V-based SO method depends on the systems and models. The results shown in Figure 9 are not as pronounced as the values reported in the literature $^{30}$, because the average running time presented in this work is the total one from both stability analysis and phase-split calculation. However, the conclusion is consistent with the literature $^{30}$ that the improvement is more pronounced for more complex models, e.g. PC-SAFT, for which 5-10% running time decrease could be expected.
Figure 6. Average running time of one point PT Flash calculation versus number of SS steps. The legends information is the same as in Figure 5.
Figure 7. The percentage of difference of average running time of one point PT Flash calculation with the P-based and V-based $SO$ methods in phase-split calculation. The details of corresponding conditions of each system are given in the Support Information. The lines are just used to connect the points.

4.4 Efficiency improvements

The purpose of stability analysis is to determine if a given fluid is stable or not, and it also provides initial estimates for the following-up calculations. Therefore, people can terminate stability analysis if instability can be secured, i.e. the objective function of stability analysis smaller than a negative enough number. It is proposed to use -0.01 as such a criterion in this work. It is arguable, on one hand, if such a safe unstable criterion will destroy the robustness of the entire PT Flash calculation. There are no failures observed due to this safe unstable criterion in any of the systems investigated in this work as well as all the cases that we have worked previously. Moreover, decrease of the objective function of phase-split calculation can be secured as long as a minimization method is used. On the other hand, it may be arguable if -0.01
is the best value for this safe unstable criterion. A smaller number could be used, which may need systematic investigation for many systems over a wide range of conditions.

The effect of using such a safe unstable criterion on the efficiency can be found in Figures 8 and 9. As suggested above, 7 $SS$ steps and 4 $SS$ steps are respectively used in stability analysis and phase-split calculation before the $SO$ methods in these figures. It shall be mentioned that the P-based $SO$ method is used in stability analysis before phase-split calculation with both the P-based and V-based $SO$ methods. Firstly, the efficiency improvement apparently depends more on the systems and conditions than on the models. As seen from Figure 8, the average running time of one point stability analysis has decreased around 70% for all the models in the systems S9 and S10, which contain the same components CH$_4$, CO$_2$, H$_2$S and H$_2$O, while the decrease is from 10% to 40% in other systems for stability analysis. Regarding the effect of such a safe unstable criterion on PT Flash calculations, as shown in Figure 9, the largest decrease of the running time of one point is seen in the systems S8 and S10, which is around 50%. Secondly, it is seen that this safe unstable criterion shows a comparable impact on the P-base and V-based $SO$ methods for both stability analysis and the entire PT Flash calculation with exception for the system S12. Thirdly, it is interesting to notice that PC-SAFT is slightly more insensitive to this safe unstable than the cubic EOS in both stability analysis and the entire PT Flash calculation with again an exception for the system S12. With these exceptions, we may speculate that it might not be very suitable to make statistics based on composition variations for the PT Flash calculation at one condition. Overall, around 30% running time decrease could be expected in general when such a safe unstable criterion is applied.

If the safe unstable criterion is not met, all the C+3 initial estimates are converged in stability analysis, while it may not be necessary to converge all the trial phases in many real applications. As suggested by Michelsen $^{31}$, it may be enough to converge the most promising candidate trial phase, i.e. with the smallest objective function after a few $SS$ steps. This is investigated in Figure 10 for the entire PT Flash calculation with 7 $SS$ steps to find the most promising candidate in stability analysis. It shows that 10-25% efficiency improvement could be obtained on average.

This approach, converging the most promising candidate trial phase only, gives the same results as the general strategy for all the investigated PT Flash calculations. However, it has to be pointed out that it might affect the robustness when the condition locates in the critical region.
An example is presented in Figure 11 for the PR EOS in the system S5. More discussions on addressing this problem with alternative approaches can be found in the literature\(^1,\)\(^3\)\(^1\).

Figure 8. The effect of a safe unstable criterion on the average running time of one point stability analysis.
Figure 9. The effect of a safe unstable criterion on the average running time of one point PT Flash calculation.
Figure 10. The effect of converging the most promising candidate trial phase against converging all the trial phases on the average running time of one point PT Flash calculation.
Figure 11. Stability envelopes of the system S5 with the PR EOS from the approaches of converging all the trial phases and converging the most promising candidate trial phase. The enlarged critical region has been placed in the center of the plot.
For the P-based SO method, the effect of using the volume from the previous iteration for the consecutive step has been investigated for PC-SAFT, and the result is present in Figure 12. In these calculations, as recommended above, 7 SS steps and 4 SS steps are respectively used in stability analysis and phase-split calculation before entering the P-based SO methods. The stability and phase envelopes have not been affected by this approach. Figure 12 shows that 15-30% running time decrease could be obtained for PC-SAFT.

![Figure 12](image)

Figure 12. The effect of using the converged volume as an initial guess for the volume-root solver of the PC-SAFT EOS in the P-based SO methods on the average running time of one point PT Flash calculation.
5. Conclusion

In this work, the conventional isobaric-isothermal flash framework is briefly reviewed, in a way of presenting the key working equations for their usages in numerical implementations, along with which the computational flowcharts are also given in the Support Information. The V-based and P-based SO methods for stability analysis and the P-based SO method for phase-split calculation are using the same numerical procedure – a modified Newton method for obtaining the searching direction plus a line-search algorithm if the Newton step does not lead to a decrease of the objective function.

In stability analysis, the V-based SO method is found to be more efficient than the P-based SO method when there are no SS steps taken, even though the later method converges in fewer iterations. However, a few SS steps before entering the SO methods can significantly improve the overall efficiency of stability analysis. It is recommended to use 7 SS steps before the SO method for general applications, in which way the P-based and V-based SO methods show comparable performance. In phase-split calculation, 4 SS steps are recommended before the SO method. An appropriate use of the V-based SO method in phase-split calculation can improve the overall efficiency further. By using a safe unstable criterion, e.g. -0.01, in stability analysis, 30% decrease of the running time could be expected for PT Flash calculations on average. By using the volume iteratively in consecutive steps in the volume root solver for the P-based SO methods, the running time can be further decreased, e.g. 15-30% for PC-SAFT. This study also shows that different models may give similar phase boundaries in a certain range of conditions, while they in general present different phase envelopes over a wide range of conditions, and the efficiency deterioration shall not be overlooked when choosing an association model in real applications.

Support Information

The computational flowcharts, details of system composition, model parameters, as well as the results of stability envelopes and PT phase envelopes, along with the average running time statistics are given in the Support Information. Up on request, a demo tool is available for producing some of the presented results.
Acknowledgment

The author thanks all the colleagues and students who have used/been using this tool in their various projects and who have given me feedbacks.

Author Information

Corresponding Author

* Tel.: 045-45252877. Fax: 045-45882258. Email: xlia@kt.dtu.dk
**List of Symbols**

**Abbreviations**

- PT Flash = Isobaric-isothermal flash (equilibrium)
- SS = Successive substitution
- SO = Second-order
- EOS = Equation(s) of state
- SRK = Soave-Redlich-Kwong
- PR = Peng-Robinson
- CPA = Cubic Plus Association
- PC-SAFT = Perturbed Chain Statistical Associating Fluid Theory

**Variables**

- $T$ = Temperature (K)
- $P$ = Pressure (Pa)
- $V$ = Volume ($m^3$)
- $R$ = Ideal gas constant ($J/mol$-$K$)
- $Z$ = Compressibility factor
- $\rho$ = Molar density ($mol/m^3$)
- $A^r$ = Helmholtz free energy ($J$)
- $n_i$ = Number of moles of component $i$ (mol)
- $n$ = Number of moles of components (mol)
- $n$ = Total number of moles (mol)
- $x_i$ = Mole fraction of component $i$
- $x$ = Mole fraction of components
- $z_i$ = Mole fraction of component $i$ in the feed
- $z$ = Mole fraction of components in the feed
- $w_i$ = Mole fraction of component $i$ in the tested phase
- $w$ = Mole fraction of components in the tested phase
- $\mu_i$ = Chemical potential of component $i$ ($J/mol$)
- $f_i$ = Fugacity of component $i$ (Pa)
\( \varphi_i \) = Fugacity coefficient of component \( i \)

\( K_{i,j} \) = Equilibrium factor (K-factor)

\( \alpha_i \) = Auxiliary variable for scaled mole numbers of component \( i \)

\( \theta_{i,j} \) = Auxiliary variable for scaled mole numbers of component \( i \) in phase \( j \)

\( \Pi \) = Number of phases

\( C \) = Number of component

\( \text{tpd} \) = Tangent plane distance

\( \text{tm} \) = Tangent (plane distance) modified

\( Q \) = Objective function symbol

\( g_i \) = Gradient element

\( \mathbf{g} \) = Gradient vector

\( H_{ik} \) = Hessian element

\( \mathbf{H} \) = Hessian matrix

\( \mathbf{I} \) = Identity matrix or unit matrix

\( \eta \) = Correction number for a non-positive definite matrix

\( \lambda \) = Line search length

\( e \) = Convergence criterion

\( t \) = Average running time for one point (\( \mu s \) or \( ms \))

\( \delta_{jm} \) = The Kronecker delta function

\( i, k \) = Subscripts for components

\( j, m, r \) = Subscripts for phases
Reference


