

The prediction multi-phase, multi reactant equilibria by minimizing the Gibbs energy of the system : Review of available techniques and proposal of a new method based on a Monte Carlo technique

Qi Liu, Christophe Proust, François Gomez, Denis Luart, Christophe Len

▶ **To cite this version:**

Qi Liu, Christophe Proust, François Gomez, Denis Luart, Christophe Len. The prediction multi-phase, multi reactant equilibria by minimizing the Gibbs energy of the system : Review of available techniques and proposal of a new method based on a Monte Carlo technique. Chemical Engineering Science, 2020, 216, pp.art. 115433. 10.1016/j.ces.2019.115433 . ineris-03318310

HAL Id: ineris-03318310
https://ineris.hal.science/ineris-03318310

Submitted on 9 Aug 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The prediction multi-phase, multi reactant equilibria by minimizing the Gibbs energy of the system:

Review of available techniques and proposal of a new method based on a Monte Carlo technique

Qi Liu,[a] Christophe Proust,[a,b,] Francois Gomez,[a] Denis Luart,[c] and Christophe Len[a,d]*

a Centre de Recherche Royallieu, Sorbonne Universités, Universite de Technologie de Compiegne, CS60319, F-60203 Compiegne, France

b Dept DRA, parc Technologique Alata, INERIS, P.O Box 2, F-60550 Verneuil-en-Halatte, France

c Ecole Superieure de Chimie Organique et Minerale, 1 Rue du Reseau Jean-Marie Buckmaster, F-60200 Compiègne, France

d UMR 8247 CNRS Chimie ParisTech, Institut de Recherche de Chimie, PSL University, 11 Rue Pierre et Marie Curie, F-75005 Paris, France

Journal Pre-proofs

Abstract:

Since already a few decades, researchers developed tools to predict chemical reactions in the

context of the chemical industry. Numerical tools are now available to predict final chemical equilibriums using the principle of minimizing the Gibbs free energy of the reactions. In this paper, after recalling some basis on thermochemical equilibria, a brief review of the most renown techniques available to find the minimum of the Gibbs energy is presented. From this, limitations are discussed. Mathematically, the equations are always strongly non-linear, and the standard step by step resolution techniques may fail to find the global minimum. When non-mixed phases are present (solids, for instance), the calculations often fail. An example is given with biphenyl-CO_2 mixtures. Especially when many phases are postulated. An alternative resolution method is proposed based on a Monte Carlo method which does not require nor a linearization of the Gibbs equation neither a step by step resolution. The method can solve any multi-phase and any multi-reactant equilibrium but is much more computer resource demanding than the traditional methods. It was implemented in a home-made code (CIRCE) briefly described in Appendix A.

Keywords

Global optimization, Gibbs free energy minimization, Monte-Carlo method, thermodynamic equilibrium calculations.

1. Introduction

In many industrial fields, the spreading of robust and low-cost computer resources is massively used to foster innovation across many technical fields (Council, 1999). It is true also for the chemical industry where numerical tools and models are used to predict phase equilibrium,(Oliveira et al., 2011) solubility (Cañas-Marin et al., 2006), even some complex catalytic reactions (Masel, 2001). A somewhat recent branch is molecular modeling which, based on an improved understanding of microscopic and molecular behaviors, could help in designing new molecule and predicting some of their properties via an ab-initio approach (Edgar, 1999;

Journal Pre-proofs

Fermeglia et al., 2003; Gurkan et al., 2010).

In this work, a new numerical method for chemical engineering is proposed aiming at predicting the chemical equilibrium in a multiphase reactive mixture. The idea is surely not new but being able to predict a "chemical equilibrium in a multiphase reactive mixture" remains a challenge even for very powerful techniques such as the RAND method implemented in ASPEN for instance (Ong'Iro et al., 1995). In this context, very practical applications remain challenging to model such as pyro-gasification, heterogeneous catalysis… In this work, the difficulties are discussed, and a new solution is proposed.

Classically two methods are used to calculate the thermodynamic equilibrium. The "law of mass action" can be implemented according to which the concentrations of the reactants and products are related via a constant, provided the stoichiometric coefficients of the reaction are known. Several reactions can be chained. The method is reasonable, appealing and, at least conceptually, simple. But it is very dependent on the chosen chemical reactions and the values of the equilibrium constants (Paz-García et al., 2013; Reynolds, 1986; Zainal et al., 2001). Besides, it is not suited to very complex problems where the details of the chemistry are not known. The second method involves the well-established concept formulated by Gibbs (Kattner, 1997) according to which the "Gibbs free energy" (free enthalpy) of the reactive system reaches a minimum at the thermodynamic equilibrium. There is no need to know the process of the chemical reactions, only the energy and molecular composition of the reactants and the chemicals likely to be present in the final products.

In the following, the theoretical grounds about this second method are first recalled. Then some existing methods employed to solve the problem are discussed. Lastly, an alternative approach is described.

1.1. General Formulation

The following derivation can be found in textbooks, and only key aspects are briefly recalled here. It is assumed that a chemical equilibrium establishes between *nSp* molecules (each is indexed "*i*") composed of *nEl* atoms (each indexed "*j*"). If a_{ij} is the number of atom *j* in the molecule *i* (available in *ni* moles in the mixture), then the conservation of mass reads (*bj* is the total number of atoms *j* in the reactants):

$$\varphi_j = \sum_{i=1}^{nSp} a_{ij} \times n_i - b_j = 0 \quad for\ 1 < j < nEl \tag{1}$$

In general, the chemical potential μ_i of molecule *i* in a mixture is $\mu_i\ (P,\ T) = \mu_{io}(T) + R \times T \times \ln a_i$ where $\mu_{io}(T)$ is the standard Gibbs energy of formation of the species and a_i stands for its "activity" in the mixture. Often a_i is expressed as a function of the molar fraction x_i. For liquid solutions, $a_i = \gamma_i \times x_i$ where γ_i is the coefficient of activity of the species *i* in the mixture and for real gases $a_i = \phi_i \times (x_i \times P)$, Φ_i, being a ratio of fugacity coefficients and *P* the absolute pressure of the mixture. Those coefficients (γ_i and Φ_i) represent the intermolecular interactions with favor or limit the mobility of the species as compared to an ideal mixture where the intermolecular forces are zero. Note that for solids, it is often assumed that the chemical potential depends mostly on the temperature and not on the other compounds (no "intermolecular effects" and no "mixing" effects). At constant temperature and pressure, the chemical equilibrium is reached when the Gibbs energy of the mixture is minimum. The Gibbs energy of the mixture is most often expressed as the sum of the contributions of the standard Gibbs energies of formation of each species G^0, of the mixing of the species G^{mix} (increases the entropy) and of the non-ideality G^E ("excess Gibbs energy"):

$$G^0 = \sum_{i=1}^{NS} n_i \times \mu_{i0}(T) + \sum_{i=1}^{NL} n_i \times \mu_{i0}(T) + \sum_{i=1}^{NG} n_i \times \mu_{i0}(T) \tag{2}$$

$$G^{mix} = \sum_{i=1}^{NL} n_i \times R \times T \times ln\left(\frac{n_i}{\sum_i^{NL} n_i}\right) + \sum_{i=1}^{NG} n_i \times R \times T \times ln\left(\frac{n_i}{\sum_i^{NG} n_i} \times P\right) \tag{3}$$

$$G^E = \sum_{i=1}^{NL} n_i \times R \times T \times ln(\gamma_i) + \sum_{i=1}^{NG} n_i \times R \times T \times ln(\Phi_i) \quad (4)$$

$$G = G^0 + G^{mix} + G^E \quad (5)$$

Where *NS*, *NL*, and *NG* stand respectively for the number of species respectively in solid, liquid, and gaseous phase (*nSp=NS+NL+NG*). The objective is to minimize *G* while satisfying the conservation of the species. The conservation of the species and G^0 expressions are linear functions of n_i but G^E is not, and G^{mix} is strongly nonlinear. Consider the case where n_i is very small, but not exactly zero. Because of the logarithm, G^{mix} may vary in enormous proportions even for tiny variations of n_i and convergence might rapidly become very difficult.

This mathematical problem is thus very stiff and several algorithms were proposed to solve it.

2. Existing Minimization Methods

2.1 The "Lagrange Multipliers" Method

Most of existing numerical methods were developed on the basis of the Lagrange Multipliers method. This mathematical technique is meant to minimize the Gibbs energy while satisfying the conservation laws. The most renown chemical equilibrium codes, CEA from NASA (Gordon and McBride, 1994) and ASPEN (Plus, 2009) for instance, use it.

The "Lagrange multipliers" method (Bertsekas, 2014) aims at minimizing a function *f* of *n* variables knowing that *p* constraints are applied to the variables. The constraints ($1<j<p$) are represented by:

$$\varphi_j = \sum \alpha_{ij} \times x_i - b_j = 0 \quad (6)$$

The Lagrange function is defined as:

$$La = f(N_1, N_2 \ldots N_n) + \sum_{j=1}^{p} \lambda_j \times \varphi_j \quad (7)$$

Journal Pre-proofs

Where λ_j are the "Lagrange multipliers" which are calculated at the same time than the *n* variable of the problem. The minimum point satisfying the constraints is obtained as (1<*i*<*n*):

$$\frac{\partial La}{\partial N_i} = 0 \Leftrightarrow \frac{\partial f}{\partial N_i} - \sum_{j=1}^{p} \lambda j \times \frac{\partial \varphi_j}{\partial N_i} = 0 \quad (8)$$

In the present context $f=G$, $N_i=n_i$ and the constraints are the conservation of the species. Equations (8) provide *nSp* relationship whereas *nSp*+*nEl* unknowns are looked for (n_i and λ_j). The *nEl* conservation laws (1) need to be solved at the same time. However, the derivatives of *G* depend very significantly on n_i (especially G^{mix}) so that the resolution can only be a very progressive, nonlinear, step-by-step approach. An additional difficulty is that the problem is fully implicit, the researched values n_i being intricate into other variables like *G*. To solve the problem, some explicit formulation needs to be defined. The way of doing so makes differences between the various "Lagrangian methods".

2.1.1 Morley Method (GASEQ)

A technique proposed by Morley (Morley, 2005) (GASEQ software) is a sort of modified Newton-Raphson method in which a first order Taylor development of $\partial La/\partial n_i$ as a function of n_i is applied to approximate the next value of this function. Morley considered only ideal mixtures so that $G^E=0$ containing only gases. Let $F(n_i)$ be the Lagrange function applied to the chemical equilibrium. Note that often $G°$ is replaced by $G°/R.T$.

$$\mathrm{F}(n) = \sum_{i=1}^{nSp} n_i \left(\frac{\mu_i^0}{RT} + \ln \frac{ni}{\sum ni} + \ln P \right) - \sum_{j=1}^{nEl} \lambda_j \sum_{i=1}^{nSp} (a_{ij} n_i - b_j) \quad (9)$$

The solutions correspond to the situations where *F* is a minimum for all *i* from 1 to *nSp* meaning that:

$$\frac{\partial F}{\partial n_i} = \frac{\mu_i^0}{RT} + \ln \frac{n_i}{\sum n_i} + \ln P + n_i \frac{\partial \ln n_i}{\partial n_i} - \sum_{k=1}^{nSp} n_k \frac{\partial \ln \sum n_i}{\partial n_i} - \sum_{j=1}^{nEl} \lambda_j a_{ij} = 0 \quad (10)$$

It can be verified that the terms $n_i\frac{\partial \ln n_i}{\partial n_i} - \sum_{k=1}^{nSp} n_k \frac{\partial \ln \sum n_i}{\partial n_i}$ is zero so that:

$$\frac{\partial F}{\partial n_i} = \frac{\mu_i^0}{RT} + \ln \frac{n_i}{\sum n_i} + \ln P - \sum_{j=1}^{nEl} \lambda_j a_{ij} = 0 \tag{11}$$

Then a first order Taylor expansion of $\partial F/\partial n_i$ is applied. In the following n_i is the searched value of *ni* and *Ni* is the initial value.

$$\frac{\partial F}{\partial n_i} \approx \left(\frac{\partial F}{\partial x_i}\right)_{at\ n_i = N_i} + \sum_{k=1}^{nSp} \left(\frac{\partial}{\partial n_k}\left(\frac{\partial F}{\partial n_i}\right)\right)_{at\ n_i = N_i} (n_i - N_i) \tag{12}$$

Where:

$$\frac{\partial^2 F}{\partial n_k \partial n_i} = -\frac{1}{\sum N_i} \qquad \text{When } i \neq k \tag{13}$$

$$= \frac{1}{N_i} - \frac{1}{\sum N_i} \text{ When } i = k$$

Substituting into (12):

$$\frac{\partial F}{\partial n_i} \approx \frac{\mu_i^0}{RT} + \ln \frac{N_i}{\sum N_i} + \ln P - \sum_{j=1}^{nEl} \lambda_j a_{ij} + \frac{1}{N_i}(n_i - N_i) - \sum_{k=1}^{nSp} \frac{1}{\sum N_k}(n_i - N_i) = 0$$

$$\frac{\partial F}{\partial n_i} \approx \frac{h_i}{N_i} - \sum_{j=1}^{nEl} \lambda_i a_{ij} + \frac{n_i}{N_i} - \frac{\sum n_i}{\sum N_i} = 0 \tag{14}$$

Where:

$$h_i \equiv N_i\left(\frac{\mu_i^0}{RT} + \ln \frac{N_i}{\sum N_i} + \ln P\right) \tag{15}$$

Equation (14) provides *nSp* equations and (1) *nEl* equations whereas the number of unknowns are *nSp* and *nEl* (respectively n_i and λ_j). The system is closed. However, equation (14) is a transcendent function of n_i which may be difficult to solve. The technique is to note that if λ_j and $\sum n_i/\sum N_i$ are chosen as unknowns then n_i is obtained from equation (14):

$$n_i = -h_i + N_i \times \left(\frac{\sum n_i}{\sum N_i} + \sum_j \lambda_j \times a_{ij}\right) \tag{16}$$

Summing up equation (16) for all n_i provides:

Journal Pre-proofs

$$\sum_{i=1}^{nSp} h_i = \sum_{j=1}^{nEl} \lambda_j \sum_{i=1}^{nSp} N_i a_{ij} \quad (17)$$

Substituting (16) into (1) provides a set of *nEl* equations:

$$\sum_{i=1}^{nSp}(-a_{ij}h_i) + \sum a_{ij}x_j + \frac{\sum n_i}{\sum N_i}\sum_{i=1}^{nSp} N_i a_{ij} + \sum_{j=1}^{nEl}\lambda_j \sum_{i=1}^{nSp} a_{ij}a_{ij}N_i - b_j = 0$$

for j= 1 to *nEl* (18)

Equations (17) and (18) have *nEl*+1 unknown: λ_j and $\sum n_i/\sum N_i$. When the latter are known, it is sufficient to replace the obtained values in (16) to find all the value of n_i. Since the equations (17) and (18) are linear functions of the unknown, equations (17) and (18) constitute a matrix (**Figure** 1) involving a constant column vector (containing the chemical potentials and the total number of atoms) equated to the unknown column vector (containing λ_j and $\sum n_i/\sum N_i$) multiplied by a matrix (containing the coefficients calculated as a linear combination of the previously estimated number of moles and of the atomic composition of the products). The solution is obtained by inverting the matrix:

$$\begin{bmatrix} \sum_{i}^{nSp} N_i \times a_{ij} & 0 \\ \sum_{i=1}^{nSp} a_{i,j}a_{i,\,j} \times N_i & \sum_{i=1} N_i \times a_{i,j} \end{bmatrix} \times \begin{bmatrix} \lambda_j \\ \frac{\sum n_i}{\sum N_i} \end{bmatrix} = \begin{bmatrix} \sum_i h_i \\ b_j + \sum_i a_{i,j} \times h_{ij} \end{bmatrix}$$

Figure 1**.** the matrix of resolution of equations (17) and (18).

The method looks simple: from an initial guess of n_i, the derivatives are estimated, new values of n_i are found and the process loops until convergence but:

- as defined, the system of equations does not impose a "realizability" criterion telling that only positive values of n_i are relevant. When negative values occur, logarithms are undefined, and the calculation fails (the term that appears in the calculation of h_i generates an error). A potential solution is to test the values of n_i and when they are negative, to replace them with a value close

to 0. Nevertheless, doing so, convergence problems may arise since the corresponding logarithms may vary in large proportions for tiny variations of n_i;

- as acknowledged by the author, it is difficult to run the method when solids are added into the products, which is a severe limitation.

2.1.2 Gordon and Mc Bride method (CEA code)

A significant improvement of the robustness (and simplicity) of the Morley method was proposed by Gordon and Mc Bride. It is also applicable to perfect gases but incorporates non-miscible condensed materials. They used again the Newton-Raphson principle, but instead of using a first order Taylor expansion of $\partial F/\partial n_i$ as function of n_i, the (Taylor) expansion is performed against $C_i = ln(n_i)$ and $D = ln\left(\sum_i^{NG} n_i\right)$ which fits much better with the evolution of G, especially to tackle the evolutions of G_{mix}. This method is implemented in CEA (Gordon and McBride, 1994), the thermochemical equilibrium code developed by NASA and in many comparable codes. Note also that doing so, the increments in n_i are necessarily positive. A further advantage is that when n_i becomes very small, lnn_i is large so that limiting the increment on lnn_i instead of n_i restricts the risk of divergence.

The starting equation is the same as for equation (11). For the gaseous phase (gaseous components for $1<i<NG$, and condensed materials for $NG+1<i<nSp$), the Lagrange function derivatives reads (given later for solids):

$$\frac{\partial F}{\partial n_i} = \frac{\mu_i^0}{RT} + \ln\frac{n_i}{\sum n_i} + \ln P - \sum_{j=1}^{nEl}\lambda_j a_{ij} = 0 \tag{19}$$

And the first order Taylor expansion in C_k and D around N reads:

$$\frac{\partial F}{\partial n_i} = \frac{\mu^\circ i}{R\cdot T} + ln(N_i) - ln\left(\sum_i^{NG} N_i\right) + \ln(P) - \sum_j \lambda_j \times aij + \sum_k \frac{\partial\left(\frac{\partial F}{\partial Ni}\right)}{\partial c_k} \times \Delta ln(n_k) + \sum_k \frac{\partial\left(\frac{\partial F}{\partial Ni}\right)}{\partial D} \times \Delta ln\left(\sum_i^{NG} n_i\right) = 0 \tag{20}$$

$\frac{\partial\left(\frac{\partial F}{\partial N_i}\right)}{\partial C_k} = 1$ for k=i and zero otherwise and $\frac{\partial\left(\frac{\partial F}{\partial N_i}\right)}{\partial D} = -1$ (21)

Substituting the equation (21) into (20), and setting $\frac{\mu_i}{RT} = \frac{\mu^\circ i}{RT} + ln(N_i) - ln\left(\sum_i^{NG} N_i\right) + \ln(P)$:

$$\frac{\partial F}{\partial n_i} = \mu_i - \sum_j \lambda_j \times a_{ij} + \Delta ln(n_i) + \Delta ln\left(\sum_i^{NG} n_i\right) = 0 \quad (22)$$

The mass conservation equation (1) is reconsidered separating the gases and the condensed part:

$$\sum_{i=1}^{NG} a_{ij} \times N_i + \sum_{i=NG+1}^{nSp} a_{ij} \times N_i = b_j$$

A differential version of this equation can be written (for the gases $N_i \times \Delta\ln(n_i) = \Delta n_i$):

$$\sum_{i=1}^{NG} a_{ij} \times N_i \times \Delta\ln(n_i) + \sum_{i=NG+1}^{nSp} a_{ij} \times \Delta n_i = \Delta b_j \approx b_j - \sum_{i=1}^{nSp} a_{ij} \times N_i \quad (23)$$

The expression for $\Delta\ln n_j$ obtained from equation (22) is substituted into (23) so that:

$$\sum_j^{nEl}\sum_i^{NG} a_{ij} \times a_{ij} \times N_i \times \lambda_j + \sum_{i=NG+1}^{nSp} a_{ij} \times \Delta n_i + \left(\sum_i^{NG} a_{ij} \times N_i\right) \times \Delta\ln\left(\sum n_i\right) = b_j - \sum_{i=1}^{nSp} a_{ij} \times N_i + \sum_{i=1}^{NG} a_i$$

(24)

The last step of the mathematical development consists in differentiating the total number of moles of gas $N = \sum_{i=1}^{NG} N_i$, in the following way:

$$N - \sum_{i=1}^{NG} N_i + dN - \sum_{i=1}^{NG} dN_i = 0$$

As above a transformation of *dN* into *dlnN* is proposed:

$$\sum_{i=1}^{NG} N_i \Delta\ln n_i - N\Delta\ln n = N - \sum_{i=1}^{NG} N_i \quad (25)$$

The expression of $\Delta\ln n_j$ from expression (22) is introduced in (25) to obtain a variational version of the conservation law:

$$\sum_j^{nEl}\sum_i^{NG} a_{ij} \times N_i \times \lambda_j + \sum_{i=1}^{NG}(N_i - N) \times \Delta\ln(N) = N - \sum_{i=1}^{NG} N_i + \sum_{i=1}^{NG} N_i \times \frac{\mu i}{R \times T} \quad (26)$$

Finally, the Lagrange minimization criterion (19) is written as:

$$\sum_j^{nEl} a_{ij} \times \lambda_j = \frac{\mu_i}{R \times T} \quad (27)$$

A system of 4 sets of equations (24), (26) and (27) is obtained which can be presented as a matrix containing as unknown the Lagrange multipliers, the increments of moles of each condensed product and the increment of the logarithm of the total number of moles in the gaseous phase. Note that once the latter parameter and the Lagrange multipliers are known, $\Delta ln\, n_j$ can be calculated using (22).

$$\begin{bmatrix} 0 & a_{ij} & 0 \\ a_{ij} & \sum_i^{NG} a_{ij} \times a_{ij} \times N_i & \sum_{i=1}^{NG} a_{ij} \times N_i \\ 0 & \sum_i^{NG} a_{ij} \times N_i & \sum_i^{NG} \left(N_i - \sum_i^{NG} Ni \right) \end{bmatrix} \times \begin{bmatrix} \Delta n_{icond} \\ \lambda_j \\ \Delta \ln \left(\sum_i^{NG} n_{igaz} \right) \end{bmatrix} = \begin{bmatrix} \frac{\mu i}{R \times T} \\ b_j - \sum_{i=1}^{nSp} a_{ij} \times N_i + \sum_{i=1}^{NG} a_{ij} \times N_i \times \frac{\mu i}{R \times T} \\ N - \sum_{i=1}^{NG} N_i + \sum_{i=1}^{NG} N_i \times \frac{\mu i}{R \times T} \end{bmatrix}$$

Figure 2**.** The matrix presentation of the 'CEA' (Gordon and McBride, 1994) method.

Although this method can be seen as a progress as compared to the Morley one, especially in terms of robustness and computer efficiency, problems remain such as:

- in the case of a phase change, the variations of the Gibbs energy near the equilibrium are such that dG tends towards 0 so that $\partial f / \partial N_i$ and $\partial G / \partial n_i \rightarrow 0$ and consequently the Lagrange multipliers should tend to zero also (see equation (8)). Then, because of the truncation errors, the mathematical problem becomes indeterminate.
- the coexistence of mixtures (especially gases) with condensed materials remains difficult to handle because primarily the variations of G with n_i are linear for the condensed materials and largely logarithmic for gaseous mixtures. If the existence of a condensed material is postulated which should not be present, the minimization process may produce negative values of n_i for the condensed phase and the algorithm fails.

Journal Pre-proofs

Rand and some others tried to improve the convergence of the Lagrange multipliers method.

2.1.3 Rand Method (ASPEN, HSC Code)

In Aspen software, the RGIBBS module minimizes the Gibbs free energy of a system using the Rand technique proposed by Gautam et al. in 1979 (Gautam and Seider, 1979a). The "*N* vector" contains guessed values of the number of moles of compound *i* in phase *l* (n_{il}) at equilibrium. Equation (5) is differentiated analytically with the specific assumption that $\partial\phi_{il}/\partial n_{il} = \partial\gamma_{il}/\partial n_{il} = 0$, and a quadratic 'Taylor development' is used to approximate the Gibbs free energy at the ***n*** vector, a vector of mole numbers in close proximity to the *N* (from the previous iteration) vector. Note that contrary to the QASEQ and CEA code, non-ideal mixtures can be computed (in the Rand formalism non miscible condensed components from 1< *i*< *NS* and miscible species from *NS*+1< *i*< *nSp*):

$$G = \sum_{i=1}^{NS} \mu_i^o n_i^c + \sum_{i=NS+1}^{nSp} \sum_{l=1}^{p} \mu_{il} n_{il} \tag{28}$$

Where *p* is the number of "mixed" phases (i.e. having a contribution in G^{mix} and G^E either in a vapor or liquid phase), and *NS*, non-mixed condensates, usually solids. Parameter μ_{il} is the chemical potential of component *i* in phase *l* (for non-mixed condensates $\mu_i°$ is the Gibbs energy of formation in the conditions of the reaction). Note that the superscript "*c*" is used to identify the number of moles of non-mixed condensed materials in the equations. The second order truncation reads:

$$Q(n_1,n_2,...,n_n) = G(N_1,N_2,...,N_n) + \sum_{i=1}^{NS} \frac{\partial G}{\partial N_i^c}\left(n_i^c - N_i^c\right) + \sum_{l=1}^{p} \sum_{i=NS+1}^{nSp} \frac{\partial G}{\partial N_{il}}\left(n_{il} - N_{il}\right) + \frac{1}{2} \sum_{i=1}^{NS} \frac{\partial^2 G}{\partial N_i^{c^2}}\left(n_i^c - N_i\right)^2 + \frac{1}{2} \sum_{l=1}^{p} \sum_{i=NS+1}^{nSp} \sum_{i'=NS+1}^{nSp} \frac{\partial^2 G}{\partial N_{jl} \partial N_{jl}}\left(n_{il} - N_{il}\right)\left(n_{il} - N_{il}\right) \tag{29}$$

N is computed at minimum ***Q*** subject to the atom balance constraints. An unconstrained objective function, using Lagrange multipliers, λ_j *(j = 1, 2... nEl).*

$$F\{n_1,n_2,...,n_n\} = Q\{N_1,N_2...,N_n\} + RT\sum_{j=1}^{nEl}\lambda_j\left[b_j - \sum_{i=1}^{NS}a_{ij}n_i^c - \sum_{l=1}^{P}\sum_{i=NS+1}^{nSp}a_{ij}n_{il}\right] \quad (30)$$

Is minimized using:

$$\frac{\partial F\{n\}}{\partial n_i^c} = \frac{\partial F\{n\}}{\partial n_{il}} = \frac{\partial F\{n\}}{\partial \lambda_j} = 0 \quad (31)$$

The equation (31) can be written into three equations:

$$\frac{\partial F\{n\}}{\partial n_i^c} = \frac{\partial^2 G}{\partial {n_i^c}^2}\left(n_i^c - N_i^c\right) + \frac{\partial G}{\partial n_i^c} + RT\sum_{j=1}^{nEl}\lambda_j[a_{ij}] = 0 \quad (32)$$

$$\frac{\partial F\{n\}}{\partial n_{il}} = \frac{\partial^2 G}{\partial {n_{il}}^2}(n_{il} - N_{il}) + \frac{\partial G}{\partial n_{il}} + RT\sum_{j=1}^{nEl}\lambda_j[a_{ij}] = 0 \quad (33)$$

$$\frac{\partial F\{n\}}{\partial \lambda_j} = RT\left[b_j - \sum_{i=1}^{NS}a_{ij}n_i^c - \sum_{l=1}^{P}\sum_{i=NS+1}^{nSp}a_{ij}n_{il}\right] = 0 \quad (34)$$

Because, by definition,

$$\frac{\partial^2 G}{\partial {n_i^c}^2} = 0;\ \frac{\partial G}{\partial n_i^c} = \mu_i^0;\ \frac{\partial^2 G}{\partial {n_{il}}^2} = 0;\ \frac{\partial G}{\partial n_{il}} = \mu_{il} \quad (35)$$

This problem, including the species conservation (1), reduces also to a matrix representation:

$$\begin{matrix}(34)\\(32)\\(33)\end{matrix}\begin{bmatrix}\sum_{i=NS+1}^{nSp}a_{ij} & a_{ij} & 0\\ 0 & 0 & a_{ij}\\ 0 & 0 & a_{ij}\end{bmatrix} \times \begin{bmatrix}n_{il}\\ n_i^c\\ \lambda_j\end{bmatrix} = \begin{bmatrix}b_j\\ \mu_i^\circ\\ \mu_{il}\end{bmatrix}$$

Figure 3**.** Coefficient matrix, 'RAND' method.

Solving this system provides the number of moles of the species ("*n* vector") and the Lagrange multipliers. In the Rand method, the new "guessed" value of the *n* vector does not automatically replace *N* in the next iteration because, as shown by Morley, negative values of the number of moles can appear. The new value of *n*, *n'*, is obtained from *N* and *n* as follows. Let *Δ* be the molar gaps between *N* and *n*:

$$\Delta = (n_1,n_2,...,n_n) - (N_1,N_2,...,N_n) \quad (36)$$

n' is obtained as a fraction of Δ added to *n* via a constant coefficient *w*:

Journal Pre-proofs

$$n' = N + w\Delta \quad (37)$$

w is chosen so as to provide a smaller Gibbs energy (than at *N*) while avoiding negative values of n'_{il} and n'^{C}_{i}. This latter condition is reached when w is between w_{max}, (maximum value such that $n' \geq 0$) and w_{min} (minimum value such that $n' \geq 0$). dG/dw is computed as follows:

$$\frac{dG}{dw} = \sum_{i=1}^{NS} \frac{\mu_i^o}{RT} \Delta_{ji}^{C} + \sum_{i=NS+1}^{nSp} \sum_{l=1}^{p} \mu_{il} \Delta_{il} \quad (38)$$

(because $\frac{\partial \gamma_{il}}{\partial w}$ and $\frac{\partial \phi_{il}}{\partial w}$ are zero) and μ_{il} being calculated at *n'*. Beginning with $w = w_{max}$, *w* is reduced by steps of 0.1 w_{max} until a negative slope is obtained. If not found as long as *w*> 0, a similar search is conducted for $w_{min} \leq w < 0$. With the value of w thus estimated, the new guessed value of *n'* is used and the process is looped. Although the Rand method is a significant improvement over the other Lagrangian method especially for phase equilibrium, it is not universally valid:

- When a mixed phase l is postulated at the starting point which cannot exist at the starting conditions, the Rand method decreases rapidly the concentrations of the component of the "illegal phase" to zero. This causes the lines and columns associated with this "illegal phase" in the coefficient matrix to approach zero and singularities to develop which impedes the inversion of the matrix. The technique to avoid this is to eliminate this phase (remove the corresponding lines and columns of the matrix). But when phases are removed, the minimum Gibbs free energy is searched in a restricted domain (phases and products were removed from the initial Gibbs problem) so that the minimum is "constrained" and may not be the absolute one. The software 'HSC' (Dantzig et al., 1958) uses the 'Rand' method and, in some cases, it is recognized by the authors that the minimum is not the good one (constrained).
- For the non-mixed phases (usually solids), the RAND method suffers from the same issue than the Gordon/Mc Bride methods, i.e., if the solid phase is unlikely to exist at equilibrium, values

Journal Pre-proofs

of n_i^c may easily become negative, values of w becomes infinitesimal, and the code does not converge.

2.1.4 Additional Techniques to help finding the Global Minimum

The main drawback of using the Lagrange method to find the global minimum is that, because the calculation has to start from an initially guessed value, there is a risk to go toward a local (constrained) minimum. Gautam et al. (Gautam and Seider, 1979b) devised the 'splitting phase' method to solve this difficulty. The principle is, at a given point in the calculation, to split a phase into two "trial phases" and to use this new configuration in the minimization process if the Gibbs energy was effectively reduced by this splitting operation. Only vapor (*V*) and liquid phases (*L*) can be split (*V* becomes *V*+*L*, *L* becomes *V*+*L* or *L* becomes L_1+L_2). Specific rules, and in particular the same activities of the species in two different phases, are applied to split phases so that they are, before the minimization process, not too unrealistic (otherwise it would be rapidly eliminated by the RAND method). This method is implemented in ASPEN. As compared to the direct use of the Rand method, as in HSC, the "phase splitting" technique seems a significant improvement but is not a panacea. First, as admitted by Gautham, it does not always avoid the constrained minimum and, even, may provide wrong results as for instance when the source phase is, before splitting, already close to the equilibrium composition of the phase.(Meyer, 1996) And Michelsen et al.(Michelsen, 1982a, b) attempted to apply the phase stability criterion initially developed by Gibbs to identify more robustly the situations in which a phase splitting should be applied. But it does not ensure a constrained minimum will not be reached.

Levy and Montalvo (Levy and Gómez, 1985) proposed the "tunneling" method. It is a succession of minimization cycles (with the Rand method for instance) and "tunneling" phases. Suppose a given objective function, f(x) to minimize and having a number of local minima. Starting the first

minimization phase at the point $x_1°$, the first local minimum to be found (**Figure** 4) is x_1*. The Tunneling Function reads:

$$T(x,f(x^*)) = f(x) - f(x^*) \tag{39}$$

x is increased (or decreased) starting from x_1* and T is calculated along the trajectory until T is negative. The point corresponding to T=0 is a new starting point $x_2°$. And the minimization process is restarted to reach a second local minimum x_2*. And so on until T is always positive. The last minimum is the absolute one x_G*. Note that this method is applicable primarily to the minimization problem of the Gibbs functions if the constraints (mass balance) are ignored. It is known that the minimum Gibbs point in a thermochemical problem is not the absolute one because of the mass balance constraints. So, the tunneling technique may be better suited to phase changes where dG=0 as the equilibrium. On that aspect, it may help to decide if phase splitting is required.

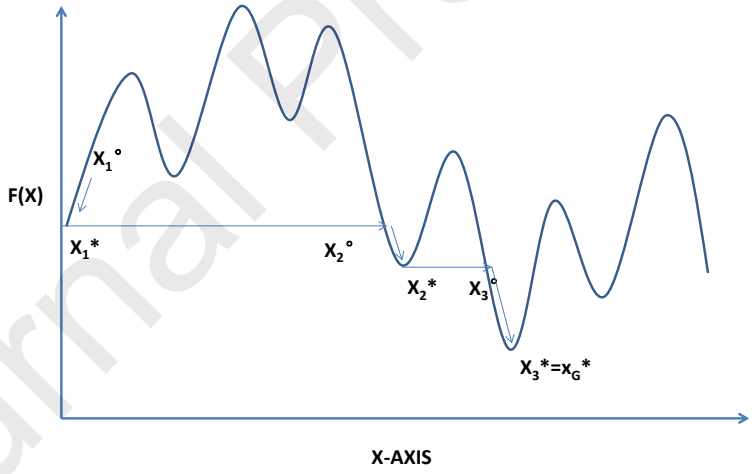


Figure 4. Illustration of the tunneling method.

The development of the Lagrange based algorithm is a significant breakthrough and allowed the development of numerically efficient computer codes to find thermochemical equilibria by minimizing the Gibbs free energy. Nevertheless, many issues remain. The first one is that because it is an iterative process starting from an initial guess of the final composition, the algorithm may

find a local minimum which may not be the absolute one. Although some techniques were proposed to avoid this, this issue is not really solved. The second issue is that the Gibbs function has to be linearized to solve the Lagrange problem. Necessarily the physical representativity is challenged (notably, the influence of the variations of the activity coefficients on G^E is ignored), and the truncation errors may jeopardize the convergence. This is the case when solids are to be considered in the mixture.

Recently, some further developments were proposed (Bonilla-Petriciolet et al., 2011; Burgos-Solórzano et al., 2004; Jalali and Seader, 1999). But these methods are still model-dependent and may require problem reformulation.

2.2 Other Methods

2.2.1 Linear Programming

"Linear programming" has been extensively used in many different field (Pekny et al., 1990; Pertsinidis et al., 1998; Shih et al., 1998) where it is often desired to minimize the cost of production, taking into account various constraints like the fixed and variable costs. Mathematically, the problem is as follows:

$$\begin{aligned} &\text{Min } (C \times X) \\ &A \times X = B \\ &X \geq 0 \end{aligned} \tag{40}$$

Where X is the vector of the variables. A, B and C are matrices (B is a vector) with constant coefficients. $C \times X$ is the objective function to be optimized. The equation $A \times X = B$ defines a set of intersecting hyperplanes in the space (order n) of the variables X. It can be shown that the optimized solution is located on one summit of the polyhedron defined by the intersection of these

hyperplanes. In 1947, Danztig (Dantzig, 1990) introduced an algorithmic method to find this optimum: the well-known "simplex method" aiming at systematically going from one summit to another and calculating the objective function there. Later, other techniques were proposed. The resemblance of (40) is striking with the problem of minimizing the Gibbs free energy minimization problem (5) under the constraints of species conservation (1) and positivity of the number of moles ($n_i \geq 0$). Nevertheless, for the "linear programming" to work, it is essential for the objective function to be a linear function of the variables. For the specific case of ideal mixtures (formulation used in GASEQ and CEA), Dantzig proposed a linearization strategy of G. More recently Rossi et al. (Rossi et al., 2009) used a similar methodology incorporating non-idealities. The authors proposed a methodology which was applied using the software 'GAMS'. 2.5 ("General Algebraic Modeling System"), using the CPLEX solver which can do the linear optimization. This method transforms the non-linear problem into a set of linear problems. Nonetheless, although the authors are not explicit on that aspect, a large number of calculated values of G at different preselected compositions needs to be provided to represent the G function conveniently by linear approximations. The calculations are tedious and long. In practice, it is said that the cost of the calculations becomes prohibitive, especially when the number of species is above 4.

2.2.2 Genetic Algorithms (GA):

Genetic algorithm (Bonilla-Petriciolet et al., 2011) is a search heuristic algorithm for solving optimization problems. It is used in artificial intelligence. It is a kind of evolutionary algorithm. Evolutionary algorithms were firstly developed for several phenomena in biology, including inheritance, mutation, natural selection, and hybridization. The genetic algorithm can be divided into three steps: initialization, iteration, and selection.

In the initialization stage, the problem is translated into "genetic" terms. In the present context

Journal Pre-proofs

of Gibbs Energy minimization where $G(x_i) = \mu_i^{\circ} + RT(\gamma_i \times x_i)$ and x_i is varying between 0 and 1 (a mole fraction for instance), the actual solution space ("phenotype" space) is composed of a set of x_i satisfying the mass conservation. This solution space should be encoded in "chromosomes" to define the "genotype" space where the genetic operations described in the iteration step could be performed. The "genes" describing the "chromosomes" is the binary code representing x_i in the computer language. If a two digits precision after the decimal is expected for x_i, then x_i will be encoded using seven binary numbers. The binary string represents the chromosome. To apply the Genetic Algorithm to phase equilibrium calculations for instance, instead of using n_{il} (for i = 1, 2, …, nSp; l= 1,2, …, p) as solution variables in the optimization, variables x_{il} (for i=1, 2, …, nSp; l=1,2;…; p) varying in the range [0,1] have to be defined and employed as decision variables:

$$n_{il} = x_{il}\left(n_{iT} - \sum_{k=1}^{l-1} n_{ik}\right) \quad l = 1,2,...,p-1 \tag{41}$$

Where n_{il} is the variation range of the mole number of species i in phase l. The mass conservation equation results in:

$$n_{il} = \left(n_{iT} - \sum_{k=1}^{l-1} n_{ik}\right) \tag{42}$$

For i=1, 2, ..., nSp, the benefit of this modification is that all candidate solutions generated will be feasible which promotes an easy and efficient implementation of GA. Since GA can only handle maximization problems, the "fitness value" F is taken as the opposite of the Gibbs energy. Therefore, the minimization of Gibbs energy can be reformulated as:

$$\text{Maximize } F = -G(x_{il}),\ i = 1,2,...,nSp;\ l = 1,...p \tag{43}$$

$$\text{Subject to } 0 \le x_{il} \le 1 \tag{44}$$

The function F is used to evaluate the quality of the points on which a selection will be performed. The initial population is generated by choosing the genes randomly.

The iteration step is performed in three stages: selection, crossover and mutation. There are

Journal Pre-proofs

different types of Genetic algorithms(Lim and Haron, 2013). A first technique is to choose a couple of "parents" randomly, to break their chromosome string as some predefined location and to rearrange them to form the chromosomes of the children (**Figure** 5). In this new population, some random mutations are allowed which consists of exchanging two genes chosen randomly on the chromosome string.

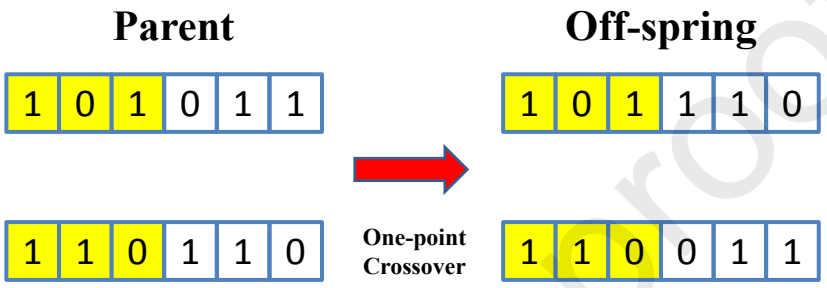


Figure 5**.** Illustration of the crossover stage.



Figure 6**.** Illustration of the mutation stage.

A selection in the parent/child population is operated (so that the total number of individuals is constant) on the basis of the fitness function according to which the "weaker" parents are replaced by the "stronger "children. In another technique, the fittest couple of parents are selected by groups of 4, and their genes are mixed according to arbitrary rules to produce two children. Similarly, for the mutation stage (2 genes changed randomly). The new population replaces the previous one. After many generations of evolution, the population satisfies the optimal requirements, and the chromosomes are the same. The convergence is reached and the iteration stops.

Journal Pre-proofs

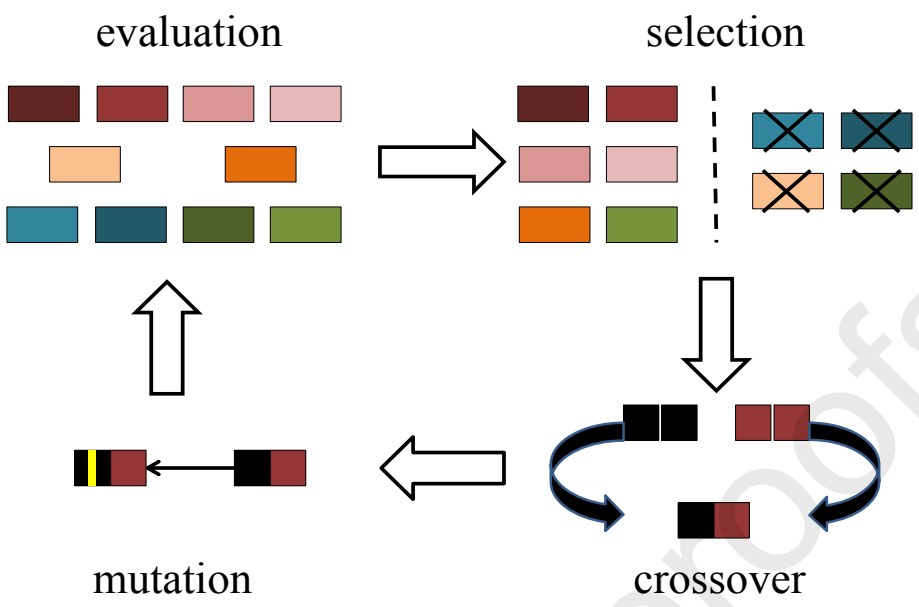


Figure 7**.** Illustration of the acceptance stage.

The example (Rangaiah, 2001) of a binary system of n-butyl acetate and water is one of the most extensively studied examples in the literature. This method is very demanding in computing resources as recognized by the authors. It is thus doubtful this technique can handle a large number of components. Furthermore, there is no guaranty that the global optimum can be found (Bonilla-Petriciolet et al., 2011). Recently, an interesting evolution of this kind of "bio" inspired algorithm was proposed (Moodley et al., 2015), but it is difficult to be sure that the global optimum can be found.

In the following, an alternative minimization technique is proposed avoiding convergence problems and constrained minima difficulties.

3 Towards an alternative based on a Monte Carlo method

The principle of the proposed method is to calculate the Gibbs energy of a "large enough" number of composition vectors chosen in the "realizable" space, i.e., fulfilling the species

Journal Pre-proofs

conservation. Then the global optimum can easily be found using traditional spanning techniques. There is no need to linearize the problem, nor to guess a starting point. The two major drawbacks of the Lagrange/Newton-Raphson methods can thus be avoided. But other difficulties may appear like the number of vectors required to ascertain that the global optimum was found, the accuracy of the results and the costs of the calculation.

3.1 Principle

Once a total number of composition vector N_v is defined (by the user), the species compositions in each vector are randomly selected. This technique is typically derived from the Monte Carlo method (Metropolis and Ulam, 1949) which has very extensive applications in the field of simulation (Moebs, 1974; Zirrahi et al., 2019). One advantage of the Monte Carlo method as compared to a systematic "meshing" of the composition space is that for a given number of composition vectors N_v, two different minimization exercises of the same problem would provide different minima if N_v is not large enough. This is also faster and ensures an equal probability of appearance of each composition vector in the realizable space.

The Monte Carlo method may be used very simply. Composition vectors are produced randomly and then only the "realizable" composition vector, i.e., those satisfying the mass conservation (equation (1)), are conserved. The first level of constraint can be applied to the random choice methodology to make this method more efficient. Consider for instance a particular product "i" with an elementary composition a_{ij} ($1<j<nEl$). The number of moles of this product should satisfy:

$$n_i \leq n_{imax} = min_j\left(\frac{b_j}{a_{ij}}\right) \tag{45}$$

Let Rand() be the mathematical function of the computer generating an arbitrary number between 0 and $RAND_{max}$ ($RAND_{max}$=32767) with equal probabilities, then $Rand\ ()/RAND_{max}$ is

an arbitrary number between 0 and 1. The number of moles of species "*i*" is estimated as:

$$\frac{n_i}{n_{imax}} = \frac{Rand\ ()}{RAND_{max}} \tag{46}$$

Let Δ be the average gap between two successive values of n_i/n_{imax}. Between 0 and 1, there are $1/\Delta$ possible values of n_i/n_{imax} so that for a composition vector containing *nSp* components, a total number of composition vectors produced amounts $(1/\Delta)^{nSp}$ arbitrarily.

But only a fraction of them can satisfy the *nEl* atom conservation laws. Because of the average gap between two consecutive component mole fractions, each conservation law can only be satisfied to within ± Δ at best. Suppose that the conservation laws are satisfied, then the total number of atoms *j* is not different from b_j by more than ± $m \times \Delta$ (with m an integer larger than 1). Starting from a composition vector satisfying the atom conservations laws to within ± $m \times \Delta$, all the composition vectors satisfying this can be derived by reasoning by pairs of components. If the number of atom j of component *i* (a_{ij}) is decreased by δ_j, the number of atom j of component *i*+1 is increased by δ_j. The δ_j is linked with Δ and a_{ij} (also note that since all conservation laws are interlinked, the acceptance criterion for a given atom *j* is automatically valid for the other atoms). Based on this reasoning, the number of composition vectors satisfying the atom conservation laws within the relative accuracy ± $m.\Delta$ reads:

$$2 \times m \times (nSp - 1) \times \left(\frac{1}{2 \cdot \Delta}\right) \tag{47}$$

If a minimum of 1000 "realizable" composition vectors (i.e., satisfying the atom balances) is needed to find the Gibbs energy minimum for *nSp*=3, *m*=1, then (47) equals 1000 and gives Δ=1/500, meaning $(1/\Delta)^{nSp} = 500^3$ randomly chosen initial composition vectors.

So this method is very unproductive and can hardly be applied in practice. Furthermore, the imperfect satisfaction of the atom balance renders the final result of the Gibbs minimization

exercise inaccurate especially for phase changes.

3.2 Monte Carlo and Gaussian Elimination Method（MCGE method）

The idea is to limit the use of the Monte Carlo method to the minimum number of species and to derive the quantities of the remaining species using the atom balance by implementing the "Gaussian elimination method". Consider, for instance, six species *U*, *V*, *W*, *X*, *Y*, *Z* composed of three different atoms *E*1, *E*2, *E*3 (**Table** 1). The unknown variables are the mole's numbers *u*, *v*, *w*, *x*, *y*, *z* whereas the total numbers of atoms *NE*1, *NE*2, *NE*3 are known so as the atomic composition of *U*, *V*, *W*, *X*, *Y*, *Z* (a_i, b_i, c_i are the number of atoms of respective elements *E1*, *E2* and *E3* in molecule *i*).

Table 1**.** "Coefficient" matrix of the products *U*, *V*, *W*, *X*, *Y*, *Z*.

Elements↓/molecules→	U	V	W	X	Y	Z	Totals
E1	a1	a2	a3	a4	a5	a6	NE1
E2	b1	b2	b3	b4	b5	b6	NE2
E3	c1	c2	c3	c4	c5	c6	NE3

This "coefficient" matrix of the products is a problem with six unknowns and three equations. To solve it, *u*, *v* and *w* are randomly chosen, for instance using the Monte Carlo method described above, whereas *x*, *y*, and *z* are calculated using a "Gaussian elimination method" as explained below. Firstly, the atom balance is written as:

$$a_4 \times x + a_5 \times y + a_6 \times z = N_{E1} - a_1 \times u - a_2 \times v - a_3 \times w = R_{E1} \quad (48)$$

$$b_4 \times x + b_5 \times y + b_6 \times z = N_{E2} - b_1 \times u - b_2 \times v - b_3 \times w = R_{E2} \quad (49)$$

$$c_4 \times x + c_5 \times y + c_6 \times z = N_{E3} - c_1 \times u - c_2 \times v - c_3 \times w = R_{E3}$$

(50)

The objective of the method is to eliminate *x* from the second equation and *x*, *y* from the third one so that *z* could be calculated then *y* and lastly *x*. This is done by applying a series of linear combinations of the atom balance equations.

The first equation is kept as such whereas *x* is eliminated from (49) and (50) using a linear combination with (48) ($b_4 \times (48) - a_4 \times (49) => (51)\ and\ \ c_4 \times (48) - a_4 \times (50) => (52)$:

$$a_4 \times x + a_5 \times y + a_6 \times z = R_{E1} \quad (48)$$

$$(a_5 \times b_4 - b_5 \times a_4) \times y + (a_6 \times b_4 - b_6 \times a_4) \times z = b_4 \times R_{E1} - a_4 \times R_{E2}$$

(51)

$$(a_5 \times c_4 - c_5 \times a_4) \cdot y + (a_6 \times c_4 - c_6 \times a_4) \cdot z = c_4 R_{E1} - a_4 \times R_{E3} \quad (52)$$

Lastly, a linear combination of (51) and (52) is applied to eliminate y. The final system reads:

$$a_4 \times x + a_5 \times y + a_6 \times z = R_{E1} \quad (48)$$

$$(a_5 \times b_4 - b_5 \times a_4) \times y + (a_6 \times b_4 - b_6 \times a_4) \cdot z = b_4 \times R_{E1} - a_4 \times R_{E2}$$

(53)

$$[(a_5 \times c_4 - c_5 \times a_4) \cdot (a_6 \times b_4 - b_6 \times a_4) - (a_5 \times b_4 - b_5 \times a_4) \cdot (a_6 \times c_4 - c_6 \times a_4)] \times \mathrm{z} = (a_5 \times c_4 - c_5 \times a_4) \cdot (b_4 \times R_{E1} - a_4 \times R_{E2}) - (a_5 \times b_4 - b_5 \times a_4) \times (c_4 \times R_{E1} - a_4 \times R_{E3}) \quad (54)$$

The mole number *z* is deduced from (54), then *y* is deduced from (53), and then *x* is calculated from (48). The method is straighforward as long as the multipliers of *x*, *y* and *z* are not zero and when the number of species is larger than the number of atom types.

The first situation may arise, for instance, for purely numerical reasons. A simple remedy is to reorganize the "coefficient" matrix so that the multipliers calculated after having applied the Gaussian elimination method are not zero anymore. It may also arise when two chemically

identical compounds (isomers or identical product in two different phases) are present in the mixture. To solve this difficulty, it is essential that one of the "isomers" is part of the components selected using the Monte Carlo method and thus is withdrawn from the Gaussian elimination step.

The second situation is usually derived from the previous one. Consider, for instance, the distillation of a binary mixture (4 products) with three elements (ex: water and ethanol: **Table** 2). If the previous method is applied, *U* is chosen using the Monte Carlo method, and then the Gaussian elimination method is employed.

Table 2**.** "Coefficient" matrix for a binary distillation.

Elements↓/molecules→	*U*	*V*	*W*	*X*	Totals
*E*1	a1	a2	a1	a2	*NE*1
*E*2	b1	b2	b1	b2	*NE*2
*E*3	c1	c2	c1	c2	*NE*3

At the end of the process, the atom balance equations read:

$$a_2 \times v + a_1 \times w + a_2 \times x = R_{E1} \quad (55)$$

$$(a_1 \times b_2 - b_1 \times a_2) \cdot w + (a_2 \times b_2 - b_2 \times a_2) \cdot x = b_2 \times R_{E1} - a_2 \times R_{E2} \quad (56)$$

$$[(a_1 \times c_2 - c_1 \times a_2) \times (a_2 \times b_2 - b_2 \times a_2) - (a_1 \times b_2 - b_1 \times a_2) \times (a_2 \times c_2 - c_2 \times a_2)] \times \mathrm{x} = (a_1 \times c_2 - c_1 \times a_2) \times (b_2 \times R_{E1} - a_2 \times R_{E2}) - (a_1 \times b_2 - b_1 \times a_2) \times (c_2 \times R_{E1} - a_2 \times R_{E3}) \quad (57)$$

In equations (56) and (57), the multipliers of *x* are equal to zero. The system has solutions only if u is chosen so that the right-hand side of (57) is also zero. Then, the value taken by *x* can be selected randomly provided the calculated value of *v* remains positive which complicates somewhat the manipulation of the Gaussian elimination method.

To solve this, the "isomers" are grouped into a single "composite" product of the same elementary composition (*u* with *w* and *v* with *x*) so that the right and left members of the equation

Journal Pre-proofs

are zero which is logical. Then the Gaussian Elimination step is applied, and values for ($u + w$) and ($v + x$) are obtained. Then the Monte Carlo method is used again to generate values of u, w, v, and x the sums ($u + w$) and ($v + x$) obtained during the preceding Gaussian elimination step.

So provided some precautions are taken to organize the "coefficient" matrix, the association of the method of Gaussian elimination makes it possible to limit "naturally" the application of the Monte Carlo method to the field of compositions respecting the conservation of the mass.

4 Results

4.1 Computational Performances

The MCGE method was encoded in CIRCE software, a homemade software applicable to various problems of chemical engineering. It is described briefly in Appendix A.

One of the difficulties is to set the required number of composition vectors to converge towards the global minimum. Two theoretical and "academic" cases were tested: the nearly complete pyrolysis of methane in oxygen and the almost complete condensation of ethanol in water. These are complicated thermodynamic problems for the MCGE method, but the thermodynamic data required to run the simulations are well known and accurate (Appendix A). In both situations, the complexity of the problem was increased gradually by adding more and more components. Thermodynamically, these are transformations at constant pressure and temperature.

The pyrolysis of CH_4 (40%v/v) in O_2 was simulated at 1 atm and 2500 K. Three cases were considered: the first with only 4 final products (CH_4, O_2, CO_2, H_2O), the second with 6 final products (CH_4, O_2, CO_2, H_2O, CO, H_2) and the third with 8 final products (CH_4, O_2, CO_2, H_2O, CO, H_2, C(s), OH). The condensation of ethanol and methanol in water was simulated at 1 bar and

250 K. Four cases were considered : the first with only 2 final products (H_2O_{gas}, H_2O_{liquid}), the second with 4 final products (H_2O_{gas}, $Ethanol_{gas}$, H_2O_{liquid}, $Ethanol_{liquid}$), the third with 6 final products (H_2O_{gas}, H_2O_{2gas}, $Ethanol_{gas}$, H_2O_{liquid}, $H_2O_{2liquid}$, $Ethanol_{liquid}$) and the fourth with 8 final products (H_2O_{gas}, H_2O_{2gas}, $Methanol_{gas}$, $Ethanol_{gas}$, H_2O_{liquid}, $H_2O_{2liquid}$, $Methanol_{liquid}$, $Ethanol_{liquid}$).

In each test case, the number of composition vectors approaching a stable solution was looked for. An example of this procedure is given below. A fixed number of composition vectors is chosen (for instance 10000), and the calculation is performed ten times with the same initial conditions to establish some statistics (**Figure** 8). **Figure** 8-a shows the evolution of the number of moles of the six components-case of the pyrolysis of methane in oxygen (third column of the condensation of ethanol and methanol in water was simulated at 1 bar and 250 K. Four cases were considered: the first with only 2 final products (H_2O_{gas}, H_2O_{liquid}), the second with 4 final products (H_2O_{gas}, $Ethanol_{gas}$, H_2O_{liquid}, $Ethanol_{liquid}$), the third with 6 final products (H_2O_{gas}, H_2O_{2gas}, $Ethanol_{gas}$, H_2O_{liquid}, $H_2O_{2liquid}$, $Ethanol_{liquid}$) and the fourth with 8 final products (H_2O_{gas}, H_2O_{2gas}, $Methanol_{gas}$, $Ethanol_{gas}$, H_2O_{liquid}, $H_2O_{2liquid}$, $Methanol_{liquid}$, $Ethanol_{liquid}$). During ten successive runs, **Figure** 8-b shows the evolution of the vapor fraction for the “eight components” condensation case. Fluctuations of the final result are visible. This makes a difference with the Lagrangian method, for which provided the starting point is the same, the final result will be identical. It does not mean that the latter will be the true minimum as shown before.

Journal Pre-proofs

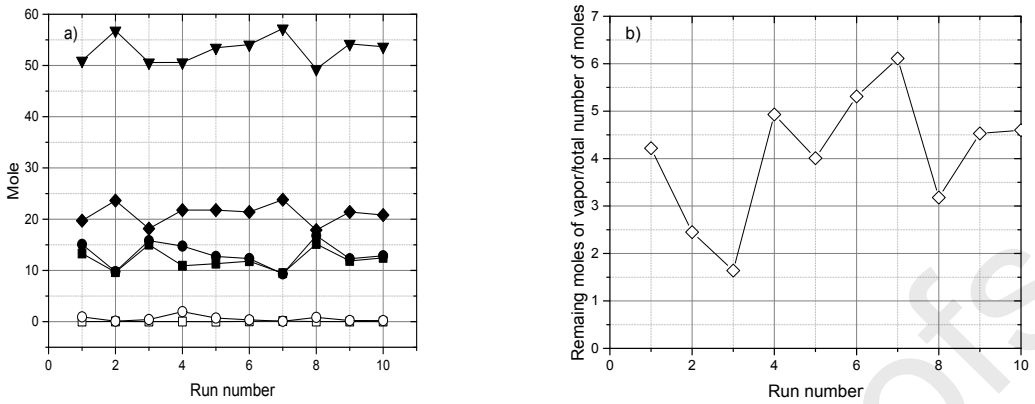


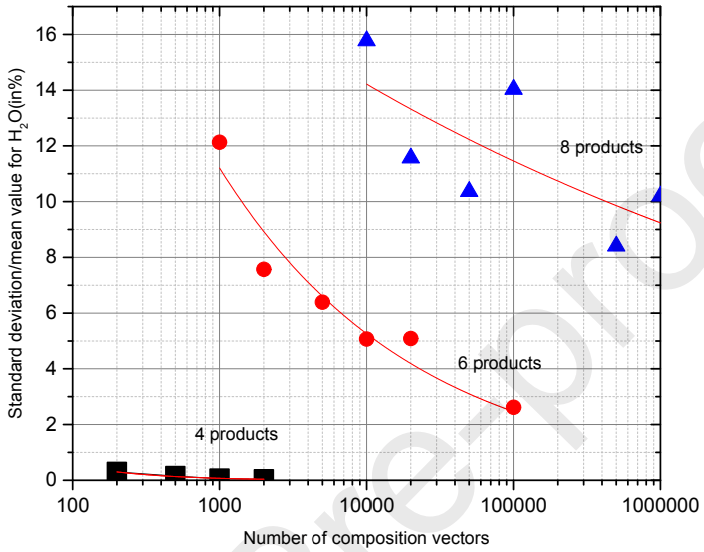
Figure 8**.** a) pyrolysis at 2500 K and 1 bar of methane (40% *v*/*v*) in oxygen containing 100 moles of reactants (case with 6 components, 10000 composition vectors H_2O, ▼; CO, ◆; H_2, ●; CO_2, ■; O_2, ○; CH_4, □); b) condensation at 1 bar – 250 K of an alcoholic solution containing 100 moles of mixture (case with 8 components; 10000 composition vectors) where N_R is remaining moles of vapor, ◇.

Since the Gibbs energy equation is not linearized, those fluctuations may reflect either the presence of many fluctuations in the Gibbs energy close to the minimum or the fact that the Gibbs energy minimum is very shallow but the discretization of one important species is too coarse.

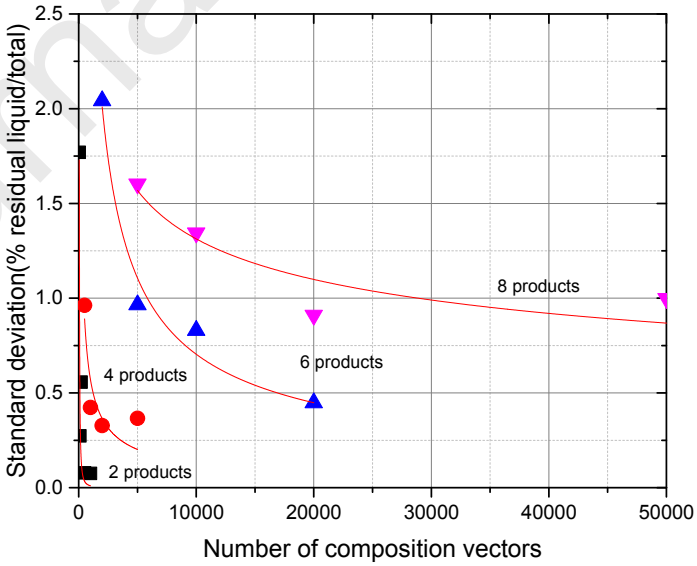
It appears (**Figure** 9) that the larger the number of composition vectors, the smaller the fluctuations from one run to another. This would suggest the second explanation might be more plausible.

Note that whatever the number of compositions vectors those fluctuations always exist even if of decreasing relative value when increasing the number of composition vectors. A direct consequence of this is that with the MCGE method, the final composition can be known only within some margin of uncertainty. This effect is particularly marked in the present case, which is particularly demanding for the MCGE method. Note that the minor species in the products (CH_4

and O_2 for the pyrolysis and % liquid for the condensation) are the major species in the reactants. As a result, the composition space to explore is particularly wide: CH_4 between 0 and 40%, O_2 between 0 and 60%, CO_2 between 0 and 40%, etc.…



(a)



Journal Pre-proofs

(b)

Figure 9**.** (a) standard deviation divided by the mean value calculated on 10 successive runs as a function of the number of composition vectors for the CO_2 concentration in the methane-oxygen pyrolysis case ("from 4 to 8 components case") and (b) for the vapor fraction in the condensation case ("2 to 8 components case").

As shown in these examples, the number of composition vectors needed to reach a given degree of accuracy, for instance 1 or 2% relative deviation for a 6 products reactive case, can be as large as 100000. With a similar complexity, a minimum would be reached using a Lagrangian method by calculating only about 1000 points (composition vectors). The calculation costs are on larger with the MCGE method as compared to the standard Lagrangian based techniques by a factor of 100. For the specific case of the MCGE method, the calculation costs are presented in **Figure** 10 using a standard laptop of an engineer. Note that most of the burden of the calculation is due to the calculation of the intermolecular effects. It can be estimated that without accounting for the intermolecular forces the duration of the calculation will be 100 to 1000 times less for the same number of composition vectors.

Moreover, in case the number of products greatly exceeds the number of elements, many of the composition vectors chosen by the Monte Carlo step will be rejected during the Gaussian Elimination step because negative concentrations will appear. A large number of attempts will be required to obtain the desired number of composition vectors.

Journal Pre-proofs

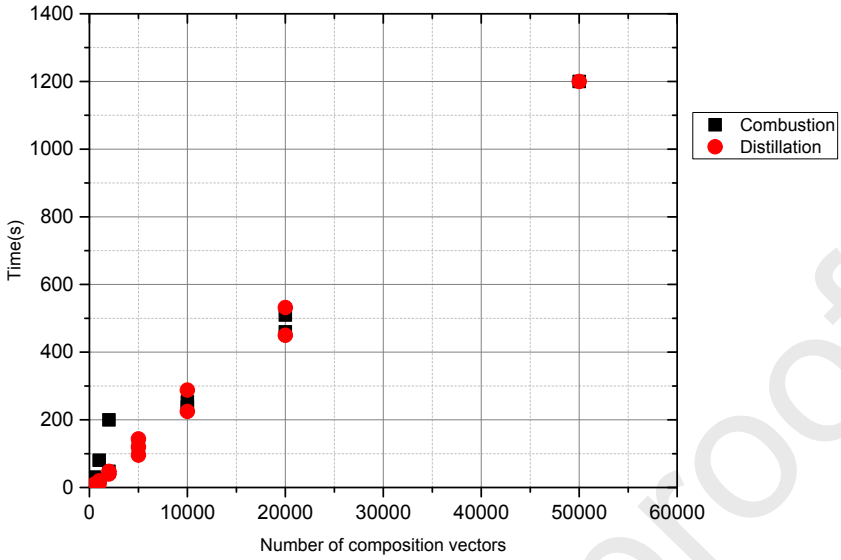


Figure 10**.** duration of the calculation for the 2 to 8 components cases for the distillation and the combustion cases presented in Figure 9 (laptop of Intel(R) Core (TM) i5-4210U CPU - 1.70GHz - 8.00 GB RAM).

4.2 Practical applications

In the following, the capabilities of the various Gibbs energy minimization techniques are compared. The MCGE method (encoded in CIRCE code) is confronted with the Rand method (from ASPEN code, RGIBBS module (Plus, 2009)) and the Gordon and Mc Bride method (from CEA (Gordon and McBride, 1994) code). In the CEA method, only the perfect gas law is implemented both with ASPEN-RGIBBS and CIRCE codes other equation of states can be used (Peng Robinson for ASPEN and LCVM for CIRCE). The chosen thermodynamic data used to run the test cases are presented in Appendix B.

Journal Pre-proofs

4.2.1. Chemical Reactions: Coal Gasification

The experimental data (Yoshida et al., 2008) from Yoshida and al. are used. The reaction is gasification of coal in water vapor inside a continuous flow reactor under ambient pressure. The experimental conditions and results are presented in **Table** 3. The global reaction (unbalanced) reads:

$$Coal + H_2O \rightarrow H_2 + CO + CO_2 + CH_4 + H_2O + C(s) \quad (58)$$

Table 3**.** The experimental conditions of the Texaco entrained-flow gasifiers.

Coal	Illinois no. 6	Illinois no. 6	Wyodak	Illinois	Wyodak	Vaccum Residue
C	71.23	70.96	78.37	69.73	78.06	83.79
H	5.44	5.38	5.81	5.21	5.32	10.5
O	1.97	2.55	3.7	2.7	4.75	0
N	0.74	0.77	0.92	0.8	0.93	0.45
S	1.74	1.63	0	1.38	0.05	5.14
Ash	18.55	18.54	11.05	19.96	10.86	0.12
CFR[c]/kg·h^{-1}	296	342	309	42259	38511	30861
$R_{O2/Coal}$[d]/kg·kg^{-1}	0.812	0.801	0.899	0.803	0.881	1.1
$R_{H2O/Coal}$[e]/kg·kg^{-1}	0.31	0.24	0.318	0.4	0.5	0.35
P[f]/kPa	8280	8280	8280	8280	8280	8280
T[g]/K	1567	1677	1571	1520	1516	1597
Gpr[h]/mol·h^{-1}	30080	33390	34340	4446000	4563000	4302000
Gc[i]/x[a]						
CO	51.88	56.06	53.48	48.53	47.66	45.35
H_2	37.32	37.21	35.72	35.67	34.24	41.37
CO_2	5.24	3.26	4.56	6.18	6.7	4.5
CH_4	0.09	0.05	0.05	0.01	0.01	0.05
H_2O	5.47	3.41	6.2	9.61	11.39	8.73
CC[j]/x[b]	0.981	0.981	0.989	0.994	0.989	0.996

[a] x is the gas composition for each product, x=100x (mole of chosen gas)/sum of mole for all the gas; [b] x= [(N_{CO}+ N_{CO2} + N_{CH4}) $_{outlet}$/ N_{Cfeed}]. [c] CFR is the Coal feed rate; [d] $R_{O2/Coal}$ is the ratio between O_2 and Coal; [e] $R_{H2O/Coal}$ is the ratio between H_2O and Coal; [f] P is pressure; [g] T is temperature; [h] Gpr is the Gas production rate; [i] Gc is gas composition; [j] CC is the carbon conversion rate.

The simulations were performed using the three codes under the assumption of constant pressure

and temperature. The same thermodynamic data and the same equation of state (perfect gas law) were used. The results are presented in **Table** 4.

Table 4**.** Coal gasification under the conditions of **Table** 3 simulated using the Rand method (Aspen-RGIBBS module), Gordon and Mc Bride method (from CEA code) and CIRCE MCGE method.

Gas composition /x[a]	Illinois no. 6			Illinois no. 6			Wyodak			Illinois			Wyodak			Vaccum Residue		
	Aspen - RGIBBS	CEA	CIRCE	Aspen - RGIBBS	CEA	CIRCE	Aspen - RGIBBS	CEA	CIRCE	Aspen - RGIBBS	CEA	CIRCE	Aspen - RGIBBS	CEA	CIRCE	Aspen - RGIBBS	CEA	CIRCE
CO	54.55	54.4	54.2	57.81	57.88	57.14	55.5	55.36	54.43	50.53	50.12	49.86	49.64	49.16	48.95	46.34	46.26	46.35
H_2	35.96	36.11	36.25	36.3	36.22	36.92	34.82	34.95	35.83	34.44	34.85	35.06	32.97	33.46	33.66	40.7	40.79	40.64
CO_2	2.65	2.79	3.01	1.7	1.63	2.37	2.8	2.94	3.87	4.09	4.5	4.77	4.82	5.31	5.52	2.91	3	2.9
CH_4	0	0	0.03	0	0	0.02	0	0	0.02	0	0	0.03	0	0	0.01	0	0	0.02
H_2O	6.84	6.69	6.52	4.19	4.27	3.55	6.89	6.75	5.85	10.93	10.53	10.28	12.56	12.07	11.87	10.04	9.96	10.08
Carbon conversion/ x[b]	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

[a] x is the gas composition for each product, x=100*(mole of chosen gas)/sum of mole for all the gas; [b] $x = [(N_{CO} + N_{CO2} + N_{CH4})_{outlet} / N_{Cfeed}]$.

Globally the results are comparable and rather close to the experimental data. Note that this finding may not be so surprising since the major difficulties associated with minimizing the Gibbs energy may appear when phase changes and mixing intervene as shown below.

4.2.2 Three-Phase Equilibrium

This last case corresponds to a three-phase equilibrium of biphenyl in CO_2 and is particularly challenging. The experiments were performed by McHugh et al. (McHugh and Paulaitis, 1980) investigating the extraction of biphenyl using supercritical CO_2. Biphenyl is an aromatic solid (**Figure** 11).

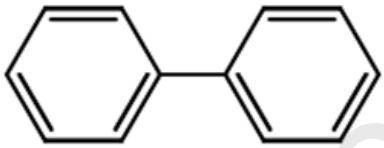
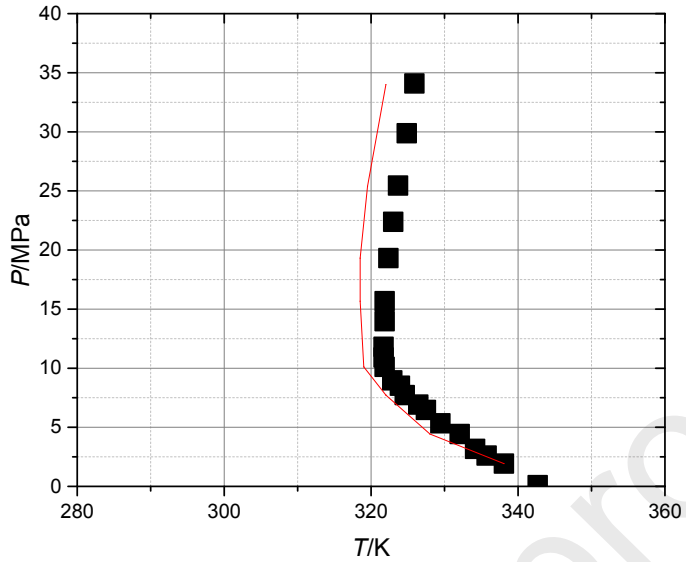


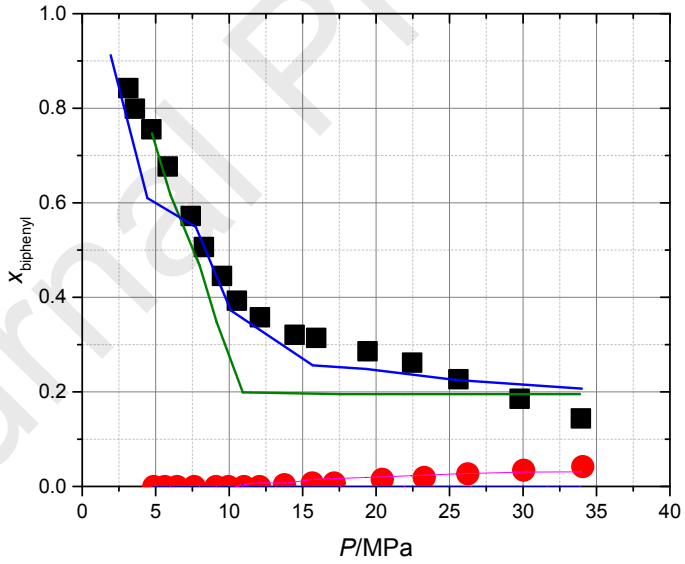
Figure 11. The structure of Biphenyl.

Depending on the pressure and temperature three-phases may coexist: vapor as a mixture of CO_2 and biphenyl, liquid as a mixture of CO_2 and biphenyl and solid with biphenyl only. The experimental P, T curve along which the three phases coexist is shown in **Figure** 12 **(a)**. The particular case at 48°C (321K) and various pressures are shown on **Figure** 12 **(b)** giving the solubility of biphenyl in CO_2 (vapor and liquid).

Journal Pre-proofs



(a)



(b)

Figure 12**.** (a) phase equilibrium diagram from experiments ■ (Cheong et al., 1986) and simulation using CIRCE —. (b) Solubility curve of biphenyl in liquid and vapor phases. ■ experiment-liquid

Journal Pre-proofs

phase, ● experiment-liquid phase, — simulation for the liquid phase CIRCE, — simulation for the vapor phase CIRCE, — simulation for the liquid phase ASPEN, — simulation for the vapor phase ASPEN/RGIBBS.

Again, the same thermodynamic data were used with the three codes, but the equation of state differ: perfect gas law for CEA, Peng Robinson with ASPEN-RGIBBS and LCVM with CIRCE. The results are also shown in **Figure** 12 **(a)** and **Figure** 12 **(b)**. The initial composition of this simulation was set to 0.8 for CO_2 and 0.2 for biphenyl in a pure liquid phase (the amounts are not critical for the calculations provided saturation of the phase is reached). The transformation is operated at constant pressure and temperature. The reader will realize that no point is shown for CEA and ASPEN-RGIBBS in **Figure** 12 **(a)**. The reason is that the Gordon and Mc Bride method (CEA code) systematically diverges and that the Rand method (ASPEN-RGIBBS) is not able to predict a three-phase mixture. In particular, the solid phase never appears. In **Figure** 12 **(b)**, it can be realized that ASPEN-RGIBBS provides a liquid-vapor equilibrium only up to 9 MPa. Above this pressure, the code returns the original mixture (pure liquid). The used database of CIRCE is the data generated by the JOBACK method.

In the same situation, CIRCE does identify the multiphase equilibrium and seems to provide a correct estimate of the VLE. Note however that the cost of the calculation is significant (2 mn for one point….). A comparison of the performance of CIRCE on this test case with refined modelling using directly various equations of state (Hong et al., 2009) is presented in Appendix C. The present method seems at least as accurate without requiring a modification of the coefficients of the LCVM EoS contrary to what was suggested by Hong. It is believed that the variations of the chemical potentials with T and P (particularly that of the solid) accounted for in CIRCE explains this.

5. Conclusion

This paper deals with the calculation of the thermodynamic equilibrium of mixtures submitted to phase changes and chemical transformations. Significant progress was made in the second half of the 20th century when numerical methods attempting to calculate the minimum of the Gibbs energy function were developed. Most of them are based on the Lagrange multipliers method (LM), a powerful mathematical approach and particularly efficient computationally speaking. More exotic techniques were tested like the global optimization (SIMPLEX), genetic algorithm, but which do not really solve the difficulties associated with the use of the Lagrange multiplier technique.

Among them, the necessity to locally linearize the Gibbs function (which is strongly nonlinear) and the step-by-step approach which limits the predictability. In particular, at best, the closest minimum to the initial "guess" is found. For multiphase reactive mixtures in particular, and non-ideal mixtures, the Gibbs function may be strongly nonconvex with a multitude of local minima. Convergence difficulties were also highlighted, especially when solids are present.

Some of these difficulties are illustrated in this work, and another minimization technique is proposed which does not require any approximation of the Gibbs function nor any "initial guess" which releases the two major limitations of the Lagrange multiplier methods.

The principle is to choose some composition vectors randomly in the "composition space" satisfying the atom balance, to calculate at each point the Gibbs energy and then to look for the minimum. To choose the composition vectors, a Monte Carlo technique is used associated with a Gaussian elimination method to ensure the atom balance is entirely satisfied. By performing several runs of the same simulation with the same number of composition vectors, it is possible to estimate the accuracy of the final result. Although computationally much more demanding than

the traditional Lagrange multiplier methods, it seems to be more reliable.

As it stands today, the code in which this method (MCGE) is implemented (CIRCE) is capable of calculating a reactive multiphase equilibrium within a few minutes for about ten components. It is certainly possible to improve the performances either by parallelizing the code or by coupling the MCGE method with some LM method.

An interesting perspective is offered by the fact that the Gibbs function can be as complicated as needed so that supercritical mixtures may be considered for instance or a more complex equation of state including complicated molecular interactions such as SAFT (Statistical Associating Fluid Theory), GCA-EoS (SAFT group contribution model) or *EoS/GE* model (Boukouvalas et al., 1994; Economou and research, 2002; Liang et al., 2014; Pereda et al., 2010).

Corresponding Author

*E-mail: christophe.proust@ineris.fr

ORCID

Christophe. PROUST: 0000-0002-9097-4292

Notes

The authors declare no competing financial interest.

Acknowledgements

The authors thank the Ministère de l'Enseignement Supérieur et de la Recherche for its constant supports.

Journal Pre-proofs

NOTATIONS

A = matrix with constant coefficients

a_{ij} = number of atoms of element j in compound i

a_i = activity of compound i

$a1$, $a2$, $a3$, $b1$, $b2$, $b3$, $c1$, $c2$, $c3$= coefficients in the matrix of elements (MCGE method)

B = matrix with constant coefficients

b_j = total number moles of atoms j in the mixture

C = matrix with constant coefficients

$$C_k = ln(n_i)$$

$$D = ln\left(\sum_{i}^{NG} n_i\right)$$

Ei = name of an atom

F = unconstrained objective function

f_i = fugacity of the species i

f = fugacity

G = Gibbs free energy

$G°$ = standard Gibbs free energy

G_{mix} = mixing Gibbs energy

G^E = excess Gibbs energy

Journal Pre-proofs

h_i = chemical potential of ideal gas *i*

L = liquid phase

La = Lagrange function

N, *n* = number of moles

N_V = total number of composition vectors

Ni = guessed moles of compound *i*

ni = moles of compound *i*

NS = number of moles of solid

NL = number of moles of liquid

NG = number of moles of gas

NEi = total numbers of atoms *Ei*

nEl = number of elements

nSp = number of species (compounds)

p = number of phases (vapor, liquid, solid)

P = absolute pressure

Q = function

R = universal gas constant.

R_{Ei} = residual numbers of atoms *Ei*

T = absolute temperature

Journal Pre-proofs

$T(x,f(x^*))$ = tunneling function

u = number of moles of compound U

v = number of moles of compound V

V = vapor phase

w = number of moles of compound W

w = constant variable

x_{ik} = reduced variable (mole fraction)

x = number of moles of compound X

X = vector of the variables

y = number of moles of compound Y

z = number of moles of compound Z

Greek Symbols

γ_{jl}= activity coefficient of compound j in phase l

$\phi_i =$ fugacity coefficient of compound i

λ_i= Lagrange multipliers

Δ = the average gap between two successive values

δ_j = variation of function

$\varphi_j(N_i)$ = function of the conservation of mass

$\mu°_{jl}$= free energy of formation of species j in the standard state for phase l

Journal Pre-proofs

μ_{jl} = partial molar free energy of compound *j* in phase *l* (chemical potential)

Appendix A: Brief description of CIRCE

CIRCE is proprietary software calculating the final equilibrium of chemical reactions on the basis of the minimum of the Gibbs energy of a list of potential final products like CEA code from NASA.(Gordon and McBride, 1994) The distinctive features of CIRCE are the mathematical method used to find the minimum of G, the way the thermodynamical data are generated into the product database and the method used to foresee the potential final products.

Instead of using the standard Lagrange multiplier method and a Runge Kutta technique to search for the minimum of G, a Monte Carlo method associated to a Gaussian elimination method to preserve the atom conservation laws is implemented (MCGE method). This way, the absolute minimum can be found in any circumstance which is not always the case with the Lagrangian methods.

To improve the internal coherence of the calculations, the thermodynamical data required to calculate G (standard enthalpy of formation, entropy, activity coefficients…) are all calculated using the molecular descriptors of UNIFAC. (Fredenslund et al., 1975) To introduce a new molecule in the database, the user has just to describe it in terms of UNIFAC groups. Then the code calculates the thermodynamical data using correlations from the literature. The LCVM equation of state (Boukouvalas et al., 1994) is used to incorporate the influence of pressure and temperature. Even supercritical transformations can be calculated.

The same group contribution theory is used to predict the potential list of final products. To do

Journal Pre-proofs

this, all the groups which may be formed are extracted from the atom composition of the reactants. They are associated following the rules established by Brignole. (Brignole et al., 1986) Only stable molecules can be predicted.

Note that CIRCE does not only calculate the thermodynamical equilibrium at constant pressure and temperature but also adiabatic equilibrium at constant pressure or constant volume. The internal structure of the code is shown in the following figure. It is coded in C language.

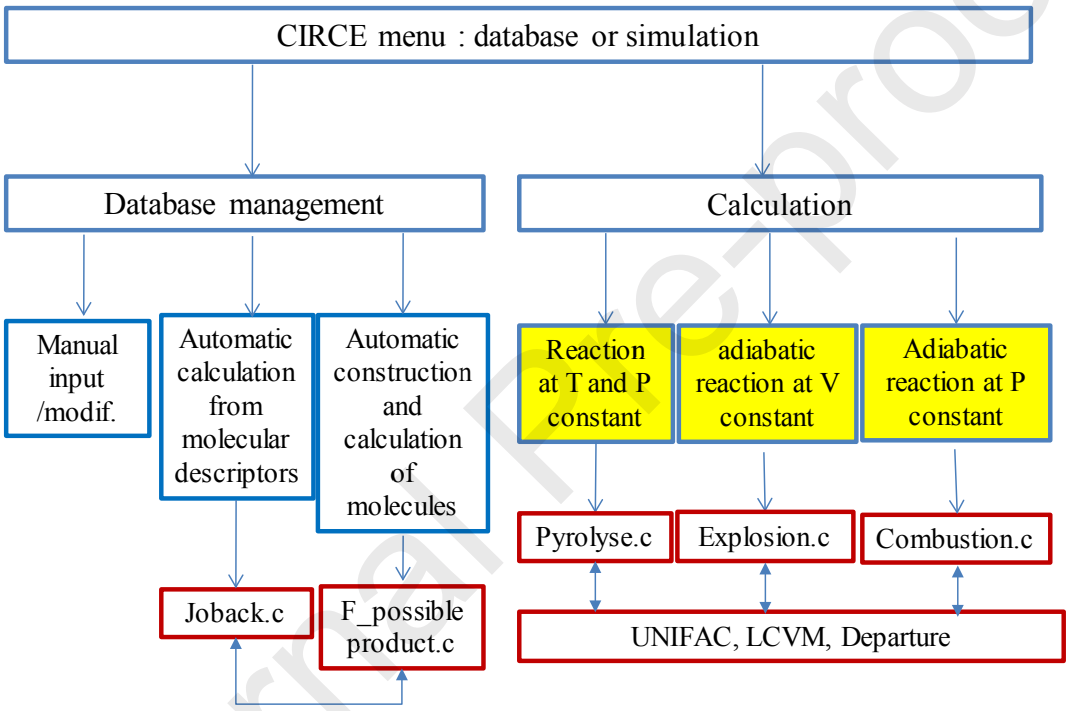


Fig. A.1. Functional structure of the software CIRCE (in red squares: main routines).

Journal Pre-proofs

Appendix B: The data used to compute the test cases

The equations of state used to run the simulations are shown in table B.1, and the main thermodynamic data chosen are shown in table B.2.

Table B.1. The chosen equation of state.

	EOS used		
	ASPEN RGibbs	CEA	CIRCE
Table 4	Perfect gas	Perfect gas	Perfect gas
Figure 12	Peng Robinson	Perfect gas	LCVM

Table B.2. The chosen thermodynamic data.

		ASPEN RGibbs			CEA			CIRCE		
	1bar/298.15K	Enthalpy (kJ/mol)	Entropy (J/mol K)	Cp (J/mol K)	Enthalpy (kJ/mol)	Entropy (J/mol K)	Cp (J/mol K)	Enthalpy (kJ/mol)	Entropy (J/mol K)	Cp (J/mol K)
Table 4	CO (gas)	-110.53	197.453	29.13	-110.535	197.660	29.141	-110.530	197.66	29.15
	H_2(gas)	0	130.684	28.60	0	130.681	28.836	0	130.68	28
	CO_2 (gas)	-393.51	213.622	35.91	-393.510	213.787	37.135	-393.393	213.795	37.12
	CH_4(gas)	-74.52	186.371	34.84	-74.600	186.371	35.691	-74.873	186.251	35.7
	H_2O (gas)	-241.820	188.826	33.94	-241.826	188.829	33.588	-241.826	188.824	33.6
	C(solid)	0	5.6	7.65	0	6.201	8.598	0	5.833	9.251
Figure 12	CO_2 (gas)	-393.51	213.622	35.91	-393.510	213.787	37.135	-393.393	213.795	37.12
	Biphenyl (gas)	178.49	392.564	143.12	182.130	388.944	166.179	180.000	391.24	166.7

Appendix C: The simulation result for the SLG coexistence of biphenyl/CO2.

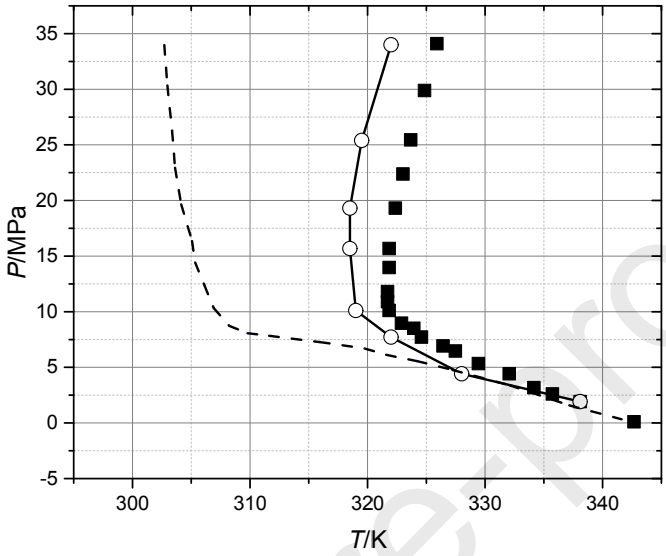


Figure C.1. SLG coexistence lines of biphenyl/CO_2 binary systems. (■) Experiment; (---) CMG with LCVM (λ) 0.36); (-o-) CIRCE.

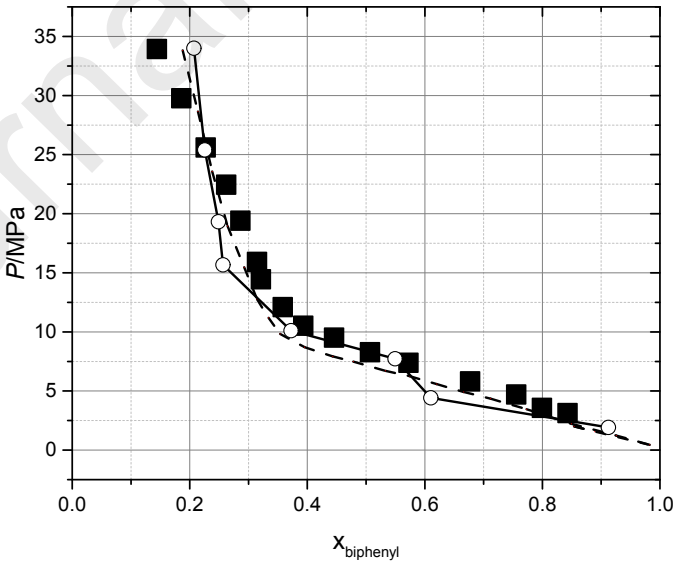


Figure C.2. Mole fractions in liquid phase for biphenyl/CO_2. (■) Experiment; (---) CMG with LCVM (λ) 0.36); (-o-) CIRCE.

Journal Pre-proofs

References:

Bertsekas, D.P., 2014. Constrained optimization and Lagrange multiplier methods. Academic press.

Bonilla-Petriciolet, A., Rangaiah, G.P. and Segovia-Hernández, J.G., 2011. Constrained and unconstrained Gibbs free energy minimization in reactive systems using genetic algorithm and differential evolution with tabu list. *Fluid Phase Equilib.*, **300**, 120-134.

Boukouvalas, C., Spiliotis, N., Coutsikos, P., Tzouvaras, N. and Tassios, D., 1994. Prediction of vapor-liquid equilibrium with the LCVM model: a linear combination of the Vidal and Michelsen mixing rules coupled with the original UNIF. *Fluid Phase Equilib.*, **92**, 75-106.

Burgos-Solórzano, G.I., Brennecke, J.F. and Stadtherr, M.A., 2004. Validated computing approach for high-pressure chemical and multiphase equilibrium. *Fluid Phase Equilib.* **219**, 245-255.

Cañas-Marin, W. A., Ortiz-Arango, J. D., Guerrero-Aconcha, U. E., and Lira-Galeana, C., 2006. Thermodynamics of wax precipitation under the influence of magnetic fields. *AIChE J.*, **52**, 2887-2897.

Cheong, P.L., Zhang, D., Ohgaki, K. and Lu, B.Y., 1986. High pressure phase equilibria for binary systems involving a solid phase. *Fluid Phase Equilib.*, **29**, 555-562.

National Research Council., 1999. Impact of Advances in Computing and Communications Technologies on Chemical Science and Technology: Report of a Workshop. National Academies Press.

Journal Pre-proofs

Dantzig, G., Johnson, S. and White, W., 1958. A linear programming approach to the chemical equilibrium problem. *Manag. Sci.,* **5**, 38-43.

Dantzig, G.B., 1990. Origins of the simplex method. A history of scientific computing (Princeton, NJ, 1987), 141–151.

Economou, I. G., 2002. Statistical associating fluid theory: A successful model for the calculation of thermodynamic and phase equilibrium properties of complex fluid mixtures. *Industrial & engineering chemistry research*, **41**, 953-962.

Edgar, T.F., 1999. Process Engineering in the 21st Century: The Impact of Information Technology. Phillips Lecture.

Fermeglia, M., Pricl, S. and Longo, G., 2003. Molecular modeling and process simulation: Real possibilities and challenges. *Chem. biochem. Eng. Q.*, **17**, 19-30.

Fredenslund, A., Jones, R.L. and Prausnitz, J.M., 1975. Group-contribution estimation of activity coefficients in nonideal liquid mixtures. *AlChE J.* **21**, 1086-1099.

Gautam, R. and Seider, W.D., 1979a. Computation of phase and chemical equilibrium: Part I. Local and constrained minima in Gibbs free energy. *AlChE J.* **25**, 991-999.

Gautam, R. and Seider, W. D., 1979b. Computation of phase and chemical equilibrium: Part II. Phase-splitting. *AlChE J.* **25**, 999-1006.

Gordon, S. and McBride, B.J., 1994. Computer program for calculation of complex chemical equilibrium compositions and applications. Part 1: Analysis.

Gurkan, B., Goodrich, B. F., Mindrup, E. M., et al., 2010. Molecular design of high capacity, low viscosity, chemically tunable ionic liquids for CO_2 capture. *J. phys. Chem. Lett.*, **1**, 3494-3499.

Hong, J., Chen, H., Li, J., Matos, H.A., de Azevedo, E.G.J.I., Research, E.C., 2009. Calculation of solid−liquid−gas equilibrium for binary systems containing CO2. 48, 4579-4586.

Journal Pre-proofs

Jalali, F. and Seader, J., 1999. Homotopy continuation method in multi-phase multi-reaction equilibrium systems. *Comput. Chem. Eng.* **23**, 1319-1331.

Kattner, U. R., 1997. The thermodynamic modeling of multicomponent phase equilibria. *JOM* **49**, 14-19.

Levy, A.V. and Gómez, S., 1985. The tunneling method applied to global optimization. *Numerical optimization*, **1981**, 213-244.

Liang, X., Tsivintzelis, I., Kontogeorgis, G.M.J.I. and Research, E.C., 2014. Modeling water containing systems with the simplified PC-SAFT and CPA equations of state. *Ind. Eng. Chem. Res.*, **53**, 14493-14507.

Lim, S.P. and Haron, H., 2013. Performance of different techniques applied in genetic algorithm towards benchmark functions, *Asian Conference on Intelligent Information and Database Systems*. Springer, pp. 255-264.

Masel, R.I., 2001. Chemical kinetics and catalysis. New York: Wiley-Interscience.

McHugh, M. and Paulaitis, M.E., 1980. Solid solubilities of naphthalene and biphenyl in supercritical carbon dioxide. *J. Chem. Eng. Data*, **25**, 326-329.

Metropolis, N. and Ulam, S., 1949. The monte carlo method. *J. am. Stat. Assoc.*, **44**, 335-341.

Meyer, C.A., 1996. Calculation of chemical and phase equilibria. University of Cape Town.

Michelsen, M.L., 1982a. The isothermal flash problem. Part I. Stability. *Fluid Phase Equilib.*, **9**, 1-19.

Michelsen, M.L., 1982b. The isothermal flash problem. Part II. Phase-split calculation. *Fluid Phase Equilib.,* **9**, 21-40.

Moebs, W.D.J.M.B., 1974. A Monte Carlo simulation of chemical reactions. *Math. Biosci.*, **22**, 113-120.

Moodley, K., Rarey, J. and Ramjugernath, D., 2015. Application of the bio-inspired Krill Herd

Journal Pre-proofs

optimization technique to phase equilibrium calculations. *Comput. Chem. Eng*. **74**, 75-88.

Morley, C., 2005. Gaseq: a chemical equilibrium program for Windows. *Ver. 0.79*. http:// www. gaseq. co. uk.

Oliveira, M.B., Ribeiro, V., Queimada, A.J. and Coutinho, J.A., e.c., 2011. Modeling Phase Equilibria Relevant to Biodiesel Production: A Comparison of g E Models, Cubic EoS, EoS− g E and Association EoS. *Ind. Eng. Chem. Res.*, **50**, 2348-2358.

Ong'Iro, A., Ugursal, V., Al Taweel, A. and Blamire, D., 1995. Simulation of combined cycle power plants using the ASPEN PLUS shell. *Heat Recov. Syst. CHP*, **15**, 105-113.

Paz-García, J.M., Johannesson, B., Ottosen, L.M., Ribeiro, A.B. and Rodríguez-Maroto, J.M., 2013. Computing multi-species chemical equilibrium with an algorithm based on the reaction extents. *Comput. Chem. Eng.* **58**, 135-143.

Pekny, J. F., Miller, D.L. and McRae, G. J., 1990. An exact parallel algorithm for scheduling when production costs depend on consecutive system states. *Comput. Chem. Eng.*, **14**, 1009-1023.

Pereda, S., Brignole, E. and Bottini, S., 2010. Equations of state in chemical reacting systems, *Applied Thermodynamics of Fluids*. RSC Publishing Cambridge, UK, pp. 433-459.

Pertsinidis, A., Grossmann, I. E. and McRae, G. J., 1998. Parametric optimization of MILP programs and a framework for the parametric optimization of MINLPs. *Comput. Chem. Eng.*, **22**, S205-S212.

Plus, A., 2009. Aspen Technology. *Inc., version 11*.

Rangaiah, G.P., 2001. Evaluation of genetic algorithms and simulated annealing for phase equilibrium and stability problems. *Fluid Phase Equilib*. **187**, 83-109.

Reynolds, W., 1986. The element potential method for chemical equilibrium analysis: Implementation in the interactive program STANJAN, version 3. *Technical Rept*.

Journal Pre-proofs

Rossi, C., Cardozo-Filho, L. and Guirardello, R., 2009. Gibbs free energy minimization for the calculation of chemical and phase equilibrium using linear programming. *Fluid Phase Equilib.* **278**, 117-128.

Shih, J.-S., Russell, A.G. and McRae, G.J., 1998. An optimization model for photochemical air pollution control. *Eur. J. Oper.*, **106**, 1-14.

Yoshida, H., Kiyono, F., Tajima, H., Yamasaki, A., Ogasawara, K. and Masuyama, T., 2008. Two-stage equilibrium model for a coal gasifier to predict the accurate carbon conversion in hydrogen production. *Fuel*, **87**, 2186-2193.

Zainal, Z. A., Ali, R., Lean, C. H. and Seetharamu, K. N., 2001. Prediction of performance of a downdraft gasifier using equilibrium modeling for different biomass materials. *Energy Convers. Manag.*, **42**, 1499-1515.

Zirrahi, M., Azinfar, B., Hassanzadeh, H., Abedi, J., Azin, R. and Osfouri, S., 2019. Particles aggregation and fragmentation—A Monte Carlo study. *Chem. Phys.*, **517**, 6-12.

Journal Pre-proofs

Highlights:

The paper :

- provides details, sometimes difficult to extract from the literature, about the various methods used to solve the Gibbs minimization problem.
- gives an analysis of the limits of the (powerful) Lagrange based methods traditional used in renown codes (Aspen-RGIBBS, CEA, Gaseq,…). The two larger, nearly un, are the local minimum limit and the incorporation of solids in the equilibrium.
- proposes a modified Monte Carlo method to solve these two difficulties and demonstrates that it works in situations where Lagrangian methods fail.
- gives an account of the performances of this new method (much more resources demanding but very flexible).
- Shows practical applications in Chemical Engineering via an implementation of the method in a totally new code developed from scratch (CIRCE).

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proofs