A Generic Global Optimization Algorithm for the Chemical and Phase Equilibrium Problem

Ken McKinnon and Marcel Mongeau

September 1996

MS 96-011

Supported by EPSRC research grant RR31817/18

A revised version of MS94-01.

Department of Mathematics and Statistics
University of Edinburgh, The King's Buildings, Edinburgh EH9 3JZ
Tel. (33) 131 650 5075 E-Mail : ken@maths.ed.ac.uk, mongeau@cict.fr

A Generic Global Optimization Algorithm for the Chemical and Phase Equilibrium Problem*

KEN MCKINNON AND MARCEL MONGEAU ken@maths.ed.ac.uk mongeau@cict.fr
Dept of Mathematics and Statistics, University of Edinburgh EH9 3JZ, UK
Labo Approximation & Optimisation, Université Paul Sabatier, 31062 Toulouse cedex, France

Abstract.
This paper addresses the problem of finding the number, K, of phases present at equilibrium and their composition, in a chemical mixture of n_s substances. This corresponds to the global minimum of the Gibbs free energy of the system, subject to constraints representing m_b independent conserved quantities, where $m_b = n_s$ when no reaction is possible and $m_b \leq n_e + 1$ when reaction is possible and n_e is the number of elements present. After surveying previous work in the field and pointing out the main issues, we extend the necessary and sufficient condition for global optimality based on the "reaction tangent-plane criterion", to the case involving different thermodynamical models (multiple phase classes). We then present an algorithmic approach that reduces this global optimization problem (involving a search space of $m_b(n_s - 1)$ dimensions) to a finite sequence of *local* optimization steps in $K(n_s - 1)$-space, $K \leq m_b$, and *global* optimization steps in $(n_s - 1)$-space. The global step uses the tangent-plane criterion to determine whether the current solution is optimal, and, if it is not, it finds an improved feasible solution either with the same number of phases or with one added phase. The global step also determines what *class of phase* (e.g. liquid or vapour) is to be added, if any phase is to be added. Given a local minimization procedure returning a Kuhn-Tucker point and a global optimization procedure (for a lower-dimensional search space) returning a global minimum, the algorithm is proved to converge to a global minimum in a finite number of the above local and global steps. The theory is supported by encouraging computational results.

Keywords: Global optimization, Gibbs free energy, chemical and phase equilibrium, non-convex optimization, tangent-plane criterion, convexity.

1. Introduction

A mixture of substances may separate into different *phases*. The composition, i.e. the proportions of the different substances present, is the same throughout a phase. If no reactions are possible between the different substances the problem is called the phase equilibrium problem (PEP), and if reactions are possible the problem is known as the chemical equilibrium problem (CEP).

The existing techniques for the PEP and CEP divide into two classes [34]. The first concentrates on solving a set of nonlinear equations arising from the stationarity conditions for a thermodynamic function and the material and charge balance equations. In the second approach the thermodynamic function is minimized. When the system temperature and pressure are constant, this function is the Gibbs

* Supported by UK Science and Engineering Research Council grant RR31817/18

free energy (GFE) and the problem is referred to as the *isothermal* chemical and phase equilibrium problem. Furthermore, once an algorithm for this GFE minimization problem is developed, it can be modified to handle other types of equilibrium problems: the *isenthalpic*, *isentropic* and *isochoric* phase equilibrium problems (see e.g. [4]).

In this paper we shall use the minimization approach: we consider a mathematical formulation involving the minimization of a non-convex objective function (the GFE), subject to material and charge balance equality constraints and non-negativity constraints. We assume the pressure and temperature of the system to be fixed. We consider PEPs and CEPs which involve different thermodynamic functions to model the different *classes of phase* that may be present at equilibrium. A phase class, as introduced by Smith et al. [33], is characterized by the chemical potential function used to model it. For example, vapour phase and liquid phase may be modeled as different phase classes. We will extend the proof of the "reaction tangent-plane criterion" of Jiang et al. [17] to the case involving different phase classes. Note however that we shall restrict this paper to class models where each substance can be present.

The tangent-plane criterion (for the PEP and with a single phase class) was introduced by Gibbs [14, 15] and proved by Baker et al. in [2]. Peng [29] illustrated the applicability of a tangent-plane criterion for *binary systems* (i.e. involving two substances and two phases) in the PEP involving multiple-phase-class models. In [33], Smith et al. extended the criterion to the case of the CEP with multiple-phase-class models. Jiang et al. provided in [17], a proof of this result in the case of a single-phase-class model. The formulation used for the CEP in the current paper is different from that used in Jiang et al. in [17]: we shall present in this paper a unified geometric interpretation for the PEP and the CEP.

The generic algorithm we present extends, for example, Michelsen's approach [26, 27] who also treats the PEP in a stepwise manner. Michelsen uses alternately a local method to obtain a stationary point of the GFE corresponding to a given number of phases, and a stability test derived from the tangent-plane-criterion necessary and sufficient global-optimality condition of Baker et al. [2] to decide whether an extra phase should be added. He however uses a *local* equation solving method to find a stationary point of the tangent distance function and makes use of several initial estimates in an attempt to determine whether a phase is stable. Also, Michelsen concentrates on PEPs involving a single phase class. In our approach different thermodynamic models for the different phase classes which may be present at equilibrium are considered. Moreover, the way we make use of the tangent-plane criterion is different. We use the tangent-plane criterion either to add an extra phase, to improve the Gibbs free energy without changing the number of phases, or to establish global optimality. Because of its local nature, Michelsen's algorithm may not succeed in locating the global minimum, whereas the generic algorithm we present is proved to converge to a global minimum in a finite number of local and global steps. The local and global optimization algorithms to be used in the generic algorithm are not specified in this paper: a variety of combinations of methods could be used. At the end of this paper we cite encouraging results from an implementation which uses interval-analysis techniques for the global stability step.

Other work on the minimization of the GFE includes that of Gautam and Seider [12] who compare the performance of the Newton-like Rand and NASA methods, which require removal of phases (and variables) to avoid singular matrices when the amount of a substance in a phase tends towards zero, with Wolfe's quadratic programming algorithm. By making use of the phase stability analysis, Wolfe's algorithm circumvents a problem of Rand and NASA methods: that of getting prematurely trapped with too few phases. Wolfe's quadratic programming algorithm however does not eliminate the possibility of convergence to a local minimum. The algorithm presented in [28] is similar to Michelsen's except for the fact that a "quasi-Newton successive-substitution" method is used for the local-minimization step. Clasen's algorithm [6] parallels Geoffrion's generalized Benders decomposition algorithm [13]. His relaxed *master* problem, used to select values for the number of moles in each phase, is a linear programming problem. The dual of the problem obtained by fixing the number of moles in each phase, is then solved as a subproblem to obtain the compositions. However, Clasen's algorithm is described for the *ideal* GFE minimization problem formulation in which the objective function is convex, and he assumes the number of phases present at equilibrium to be known a priori. The second-order method presented by Trangenstein [36] handles the poorly-scaled minimization problems associated with mixtures near *bubble points*, *dew points* and *critical points* (see [32]). It also addresses the problem of avoiding convergence to *trivial solutions* (i.e. involving a spurious phase: the proportions of the substances of two phases are identical), and it attempts to maximize the accuracy in the solution. More specifically, in order to address the poor scaling of the minimization problem and the indefiniteness of the Hessian, he adapts a modified Newton's method [7] to the particular structure of the phase equilibrium problem and to the stability-analysis optimization subproblem involved in determining the number of phases at equilibrium. He however confines his numerical methods to the search for *local* minima, and he deals with phase equilibrium problems involving at most two phases. In [35], Sun and Seider present a homotopy-continuation algorithm for the phase equilibrium problem. Floudas & Visweswaran's GOP (Global OPtimization) algorithm [10, 11], related to the resource decomposition algorithm of Wolsey [37], is applied by McDonald & Floudas [21] to numerous examples for which the liquid phase is modeled by the NRTL equation and the vapour phase is assumed to be ideal. The GOP algorithm decomposes the original problem into primal and relaxed dual subproblems that provide upper and lower bounds on the global optimum, and makes use of branch and bound. Finally, the introduction of [21] constitutes a good survey of global optimization for the chemical and phase equilibrium problem.

The current paper divides into six parts. Following this introduction, we formulate the Gibbs free energy minimization problem and point out the main issues in this particular global optimization problem. In Section 3, we extend the necessary and sufficient condition for global optimality based on the "reaction tangent-plane criterion", to the case involving multiple-phase-class thermodynamic models, paying special attention to the issue of the number of phases that exist at equilibrium being unknown a priori. In Section 4 we present an algorithmic approach, the GILO method, that reduces the global optimization problem of minimizing the GFE to a finite sequence of *local* optimization steps and *global* optimization steps in a low-dimensional space. The global (*phase stability*) step uses the tangent-plane criterion to determine whether the current solution is optimal, and, if it is not, it finds an improved feasible solution with at most one extra phase. When the global step proves that the solution is not optimal, it also determines what *class of phase* (e.g. liquid or vapour) is to be added or exchanged with an existing phase. We show that the search can be restricted to solutions involving, in the case of the PEP, no more phases than the number of substances in the mixture, and, in the case of the CEP without charge, no more than the number of elements. The algorithm is proved to converge to a global minimum in a finite number of the above local and global steps. In Section 5, we discuss implementation issues and we cite encouraging computational experiments. We draw conclusions in Section 6.

Note that throughout the paper the superscript "+" will denote globally optimal solutions and the superscript "*" will denote local optima or Kuhn-Tucker points.

2. The Gibbs Free Energy Minimization Problem

Assume we have a mixture of n_s substances and that there are a finite number τ of *phase classes* with corresponding thermodynamic models for their GFE. The number of phases, denoted by K, required to reach equilibrium is not known a priori, but it will be shown later not to exceed m_b, where m_b is the number of independent balance constraints modeling conserved quantities.

Let $\{z_{ik}\}_{i=1}^{n_s}$ be the amount of each substance i in each phase k, $1 \leq k \leq K$. Let us denote the proportion of phase k which is substance i by x_{ik}. This is given by

$$x_{ik} \equiv \frac{z_{ik}}{\sum_{j=1}^{n_s} z_{jk}} \quad i = 1, 2, \ldots, n_s;\ k = 1, 2, \ldots, K. \tag{1}$$

The GFE is given by

$$g(K, t, z) \equiv \sum_{k=1}^{K} \sum_{i=1}^{n_s} z_{ik} \mu_{it_k}(x_{1k}, \ldots, x_{n_s k}), \tag{2}$$

where $t_k \in \tau$, $1 \leq k \leq K$, and $\mu_{i\eta} : \mathbb{R}^n \to \mathbb{R}$ is the *chemical potential* of substance i in a phase of class η, a nonlinear function of composition whose form depends on the thermodynamic model chosen to describe a phase of class η. In fact $\mu_{i\eta}$ also depends on the system temperature and pressure but this is not considered here as in this paper we shall be concerned with computing the equilibrium compositions at *specified* pressure and temperature.

The simplest model for the PEP, which is referred to as the *ideal* case, yields a GFE function, $g(K, t, \cdot)$, which is convex [31, lemma 8.7D]. Other models include the Redlich/Kister expansion, the Margules equations and the van Laar equations, which are all special cases of a general treatment based on rational functions (ratios of polynomials), and the Wilson model, the NRTL (Non-Random-Two-Liquid) equation, the UNIQUAC (UNIversal QUAsi-Chemical) equation, and the UNIFAC method [21, 32]. In petroleum reservoir simulation, the Peng-Robinson equation of state [30] is commonly used to describe the behaviour of hydrocarbon phases.

In the PEP there are no reactions so the amount of each substance is conserved. There are therefore $q = n_s$ conserved quantities and the following linear balance constraints hold.

$$\sum_{k=1}^{K} z_{ik} = b'_i, \quad i = 1, 2, \ldots, q, \tag{3}$$

where b'_i denotes the total amount of substance i present. In the CEP, the individual substances are not conserved but may react together to form other substances. However the total amount of each element present is conserved. Also if ionic substances are possible then the constraint that the total charge is conserved is also required. These constraints give rise to linear balance equations of the following form:

$$\sum_{k=1}^{K} \sum_{i=1}^{n_s} a'_{ji} z_{ik} = b'_j. \quad j = 1, 2, \ldots, q. \tag{4}$$

Here $q = n_e$ if there is no charge constraint and $q = n_e + 1$ if there is a charge constraint. If constraint j is an element-balance constraint then the coefficient a'_{ji} denotes the amount of element j per unit amount of substance i, and b'_j is the total amount of element j in the system. If constraint j corresponds to the charge constraint, then a'_{ji} denotes the amount of charge in unit amount of substance i and coefficient b'_j is 0.

Let Z be the vector whose ith component, Z_i, gives the total amount of substance i, i.e. $Z_i = \sum_{k=1}^{K} z_{ik}$. Then (4) can be rewritten equivalently as

$$\sum_{i=1}^{n_s} a'_{ji} Z_i = b'_j, \quad j = 1, 2, \ldots, q.$$

For the CEP, let A' denote the $q \times n_s$ matrix with $(ji)^{th}$ entry a'_{ji}, and for the PEP, let A' be the $n_s \times n_s$ unit matrix. The (global) optimization problem corresponding to either the PEP or CEP can now be written as:
P' :

$$\min_{K,t,z,Z} \quad g(K, t, z)$$

$$
\begin{aligned}
\text{subject to} \quad A'Z &= b', & (5)\\
\sum_{k=1}^{K} z_{ik} &= Z_i, \quad i = 1, 2, \ldots, n_s, & (6)\\
z &\geq 0, & (7)\\
t_k &\in \tau, \quad k = 1, 2, \ldots, K, & (8)\\
K &\in \mathbb{N}^+.
\end{aligned}
$$

The non-negativity constraints (7) are to be satisfied by definition of z_{ik}.

Throughout the paper we assume that the only substances included in the model are those which could occur at a positive level in some feasible solution. By taking a convex combination of such solutions for each substance, it follows that there exists a feasible solution in which all substances are present, i.e.

there exists $\bar{Z} > 0$ such that $A'\bar{Z} = b'$.

The following technique can be used to detect whether a model contains substances which cannot occur in any feasible solution, to eliminate such substances when they occur, and also to find a feasible solution in which all substances are present at a positive level. Solve the following LP problem, P^{feas}, whose objective is to maximize the minimum amount of any substance.

P^{feas} :

$$
\max_{Z,\xi} \ \xi
$$

$$
\begin{aligned}
\text{subject to} \quad \xi &\leq Z_i, \quad i = 1 \ldots, n_s, & (9)\\
A'Z &= b',\\
Z &\geq 0.
\end{aligned}
$$

If P^{feas} is infeasible, then there is no combination of substances in the model which can satisfy the balance constraints, so the problem definition is inconsistent. If a positive objective value is found, then all substances can occur at the positive level given in that solution. Otherwise the optimal objective must be zero. At least one substance will have a non-zero shadow price in (9). All such substances can be eliminated and the LP repeated for the reduced problem until a positive objective is found.

A *globally* optimal solution of P' corresponds to the *true* equilibrium solution (i.e. the one found in nature). Later we shall see that limiting the number of phases K to m_b, where m_b is the rank of A', does not exclude the optimal solution. For P' with K and t fixed, g is a continuous function of z, so the infimum is achieved. Because only a finite number of K and t need be considered, the infimum is achieved for problem P', so the minimum is well defined. For the PEP $m_b = n_s$, for the CEP without charge constraint $m_b \leq n_e$, and for the CEP with charge constraint $m_b \leq n_e + 1$.

There are several difficulties in computing a global minimum of the GFE. First, the number of phases, K^+, and the class, t_k^+, of each of these K^+ phases, at which g achieves a global minimum are usually not known a priori. Moreover, for fixed K and fixed $t \in \tau^K$, since the GFE function may be non-convex in the non-ideal case, a descent algorithm may converge to a local minimum which is not global. Another difficulty is that even with $K = K^+$, a descent algorithm may converge to a point at which the total amount in a phase is zero, or there is a *spurious phase* (i.e. the proportions of the substances in two phases k and l, $k \neq l$ are identical, so $x_{ik} = x_{il}$ for $1 \leq i \leq n_s$). There are numerical difficulties in the neighbourhood of such a point [36]. The chemical potentials are generally assumed to satisfy the Gibbs-Duhem equation [32] and to have a logarithmic singularity when a substance is removed from an existing phase. Also, even the case in which both K^+ is known and the GFE function is convex, numerical difficulties can arise due to the unboundedness of the gradient at points where the amount of some substance in a phase is null. Further numerical difficulties arise for some values of pressure and temperature. For mixtures near bubble points and dew points one of the phases becomes small and the Hessian of the objective function becomes nearly singular. A similar difficulty occurs near critical points (when liquid and gas phases become indistinguishable), as two phases have nearly identical compositions. Trangenstein [36] discusses in detail the poor scaling of the Hessian matrix in the GFE minimization problem and presents a (local optimization) modified Newton's method adapted to the particular structure of the phase equilibrium problem. A major attraction of the approach taken in this paper is that most of these numerical difficulties do not occur in the global optimization part of the algorithm but are dealt with in the local optimization part, for which well-developed techniques are available.

Features that makes the chemical and phase equilibrium problem atypical of many global optimization problems are: firstly, the number of variables z_{ik} present is not known a priori, and secondly, the aim is not so much obtaining a solution with an objective *value* which is close enough to the optimal objective *value* (as is usually the case when the objective function has an economic interpretation), but rather to find a solution which itself is close enough to the optimal solution (so that the number of phases, their class and their composition at equilibrium are correctly determined and thus correspond to the equilibrium found in nature).

Note finally that problems in chemical and phase equilibria often involve instances for which $n_s \leq 10$ and $K^+ \leq 3$. In many of the problems reported in the literature [1, 8, 9, 12, 20, 21, 25, 29, 35, 36, 38], n_s is as small as 2 or 3. Sometimes however these small problems must be solved very rapidly. For instance, an important application of phase equilibria is in the modeling of petroleum reservoir fluid flow [36]. Within a numerical reservoir simulator, phase equilibrium is to be determined at each time step and for each cell in a grid partitioning the reservoir.

3. The Reaction Tangent-Plane Criterion for Multiple-Phase-Class Models

The problem P' will now be transformed to the form P below by changing the variables to the total amounts in each phase and the composition of each phase, and by eliminating redundant constraints. Let y_k denote the total amount in phase k, so

$$y_k \equiv \sum_{i=1}^{n_s} z_{ik}. \tag{10}$$

From (10) and (1) it follows that

$$z_{ik} = y_k x_{ik}. \tag{11}$$

Assume now that there exists a vector w such that

$$\begin{aligned} w^T A' &= e_{n_s} \equiv (1, 1, ..., 1) \in \mathbb{R}^{n_s}, \\ w^T b' &= 1. \end{aligned}$$

Note that for fixed K and fixed compositions x_{ik}, the objective function (2) and the left-hand side of conservation constraints (3) and (4) are linear functions of the z_{ik}, and that the other constraints are all non-negativity restrictions. It follows that if z^*_{ik} and x^*_{ik} give an optimal solution for right-hand side b', then, for any $\rho > 0$, ρz^*_{ik} and x^*_{ik} give an optimal solution for right-hand side $\rho b'$. Hence, provided $w^T b' > 0$ there is no loss in generality in assuming that the total amount present is normalized so that $w^T b' = 1$.

For the PEP, A' is the identity matrix so the vector w consisting entirely of ones has the required property. For the CEP, we choose to measure the amounts of substances and of elements in units of mass. It then follows that the coefficient a'_{ji} corresponding to an element balance constraint j is the mass of the element associated with row j in unit mass of substance i. The sum over a column of all elements corresponding to the element-balance rows is then unity. Hence, a w consisting of ones in all element-balance rows and a zero in the charge balance row, if it exists, has the desired property.

By performing row operations on equations (5) we can therefore obtain the equivalent equations

$$\begin{bmatrix} \hat{A} \\ 1, 1, ..., 1 \\ 0 \end{bmatrix} Z = \begin{bmatrix} \hat{b} \\ 1 \\ 0 \end{bmatrix},$$

where $\hat{A}$ is composed of $m_b - 1$ linearly independent rows of A' and $\hat{b}$ is composed of the corresponding rows of b'.

Now using the equation with unit coefficients to eliminate Z_{n_s} from the first $m_b - 1$ rows gives the equivalent equations

$$\left[\begin{array}{c|c} A & 0 \\ 1,1,...,1 & 1 \\ 0 & 0 \end{array}\right] Z = \left[\begin{array}{c} b \\ 1 \\ 0 \end{array}\right], \tag{12}$$

where A is the $m \times n$ matrix whose i^{th} column is $\hat{A}_i - \hat{A}_{n_s}$, $b = b' - \hat{A}_{n_s}$, $m = m_b - 1$, and $n = n_s - 1$ (when C is a matrix, C_j denotes its jth column). The row operations done to transform equations (5) to equations (12) are rank preserving, so it follows that A is of rank $m_b - 1 = m$, i.e. full rank.

For the development of the theory it is convenient to work in the reduced space obtained by eliminating the variables for one of the substances, say substance n_s.

Equations (12) are equivalent to the equations

$$\sum_{i=1}^{n} A_i Z_i = b, \quad \sum_{i=1}^{n_s} Z_i = 1.$$

Substituting for Z_i using (6) and (11) gives the equivalent equations

$$\sum_{i=1}^{n} A_i \sum_{k=1}^{K} y_k x_{ik} = b, \quad \sum_{i=1}^{n_s} \sum_{k=1}^{K} y_k x_{ik} = 1,$$

which, because by definition (1) $\sum_{i=1}^{n_s} x_{ik} = 1$, are equivalent to

$$\sum_{k=1}^{K} y_k A x_k = b, \quad \sum_{k=1}^{K} y_k = 1, \tag{13}$$

where $x_k \in \bar{X}$, and (recall $n = n_s - 1$)

$$\bar{X} \equiv \{x \in I\!R^n : \sum_{i=1}^{n} x_i \leq 1, \; x_i \geq 0\}.$$

Now for $v_i \geq 0, \; i = 1, ...n_s$ where $\sum_{i=1}^{n_s} v_i = 1$ define

$$\begin{aligned} \hat{f}_\eta(v_1, ..., v_{n_s}) &= \sum_{i=1}^{n_s} v_i \mu_{i\eta}(v_1, ..., v_{n_s}), \\ \text{and} \quad f_\eta(v) &= \hat{f}_\eta(v_1, ..., 1 - \sum_{i=1}^{n} v_i) \quad \text{for } v \in \bar{X}. \end{aligned} \tag{14}$$

Function $\hat{f}_\eta$ defined over the full n_s-space, and f_η defined over the reduced n-space give the GFE of unit amount of a single phase of phase class η. For the true GFE and the common functions used to model it, $\hat{f}_\eta$ is twice continuously differentiable and

$$\lim_{v_i \to 0^+} \frac{\partial \hat{f}_\eta}{\partial v_i} = -\infty, \text{ for all } 1 \leq i \leq n_s.$$

This implies (see [17]) that provided all substances can be present (we are assuming we have eliminated any which cannot by, for example, using problem P^{feas}), any phase that is present at equilibrium (i.e. has $y_k > 0$) has a non-zero amount of every substance in the phase. We can therefore replace $\bar{X}$ with its interior, X, without eliminating the optimal solution. From this observation and (2), (11), (13) and (14), it follows that problem P' is equivalent to the problem
P:

$$\min_{K,t,Y} \quad F(K,t,Y) \equiv \sum_{k=1}^{K} y_k f_{t_k}(x_k)$$

$$\begin{aligned}
\text{subject to} \quad \sum_{k=1}^{K} y_k A x_k &= b & (15)\\
\sum_{k=1}^{K} y_k &= 1 & (16)\\
y_k &\geq 0, \quad k = 1,2,\ldots,K & (17)\\
x_k &\in X, \quad k = 1,2,\ldots,K & (18)\\
t_k &\in \tau, \quad k = 1,2,\ldots,K &\\
K &\in \mathbb{N}^+, &
\end{aligned}$$

where $Y = (y_1, \ldots, y_K, x_1^T, \ldots, x_K^T) \in \mathbb{R}^{K(n+1)}$, A is an $m \times n$ real matrix of rank m, $m = m_b - 1$, $n = n_s - 1$, $b \in \mathbb{R}^m$, τ is a finite index set, and, for $k = 1,2,\ldots,K$, $t_k \in \tau$, $y_k \in \mathbb{R}$.

Note that with this approach problem P represents both the PEP and the CEP. This is in contrast to the work of Jiang et al. [17] who use a different formulation for the CEP: the form used for both the PEP and the CEP in the current paper is the same as that used in [17] for the the PEP.

Let $P_{t^*}^{K^*}$ denote problem P in which K is fixed to a positive integer K^* and $t^* \in \tau^{K^*}$ is fixed, and let $F_{t^*}^{K^*} : \mathbb{R}^{K^*(n+1)} \to \mathbb{R}$ be the corresponding objective function, so $F_{t^*}^{K^*} = F(K^*, t^*, Y^*)$. For a given (K^*, t^*, Y^*) define

$$J^* \equiv \{k : 1 \leq k \leq K^*, y_k^* > 0\}.$$

The Kuhn-Tucker (KT) conditions for $P_{t^*}^{K^*}$ are stated in the following Lemma.

Lemma 1 *Y^* is a Kuhn-Tucker point for $P_{t^*}^{K^*}$ if and only if there exist multipliers $\alpha \in \mathbb{R}^m$ and $\beta \in \mathbb{R}$ such that*
(i) $f_{t_k^*}(x_k^*) \geq \alpha^T A x_k^* + \beta \quad (= \Theta^*(x_k^*))$ for all $1 \leq k \leq K^*$, with equality if $k \in J^*$,
(ii) $\nabla f_{t_k^*}(x_k^*)^T = \alpha^T A \quad (= (\nabla \Theta^*)^T)$ for all $k \in J^*$,
(iii) Y^* is feasible for $P_{t^*}^{K^*}$,
where by definition $\Theta^*(x) = \alpha^T A x + \beta$.
Given any $k \in J^*$, the affine function Θ^* can be written in the form

$$\Theta^*(x) = f_{t_k^*}(x_k^*) + \nabla f_{t_k^*}(x_k^*)^T(x - x_k^*).$$

Proof: Parts (i), (ii) and (iii) are the standard KT conditions. The form of Θ^* is obtained by substituting for $\alpha^T A$ and β using parts (i) and (ii). □

This lemma has the following geometric interpretation. At a KT point Y^* for problem $P_{t^*}^{K^*}$, there exists an n-dimensional hyperplane $\{x : \Theta^*(x) = 0\}$ which is not above $f_{t_k^*}$ at x_k^* for any k, $1 \leq k \leq K^*$, and is tangent to $f_{t_k^*}$ at the points x_k^*, where $y_k^* > 0$.

Lemma 2 *Assume $(\tilde{K}, \tilde{t}, \tilde{Y})$ and (K^*, t^*, Y^*) are feasible for P. Let $\Theta^* : I\!R^n \to I\!R$ be an affine function of the form $\Theta^*(x) = \alpha^T Ax + \beta$, where $\alpha \in I\!R^m$ and $\beta \in I\!R$, and assume that $f_{t_k^*}(x_k^*) = \Theta^*(x_k^*)$ for all $k \in J^*$. Then*

$$F_{t^*}^{K^*}(Y^*) = \sum_{k=1}^{\tilde{K}} \tilde{y}_k \Theta^*(\tilde{x}_k) = \alpha^T b + \beta.$$

In particular the above relation holds if K^* is a KT point for $P_{t^*}^{K^*}$ with Θ^* the corresponding affine function defined in Lemma 1.

Proof: If (K, t, Y) is feasible for P then by (15) and (16)

$$\sum_{k=1}^{K} y_k \Theta^*(x_k) = \sum_{k=1}^{K} y_k(\alpha^T Ax_k + \beta) = \alpha^T(\sum_{k=1}^{K} y_k Ax_k) + (\sum_{k=1}^{K} y_k)\beta = \alpha^T b + \beta.$$

By the definition of Θ^* and fact that $(\tilde{K}, \tilde{t}, \tilde{Y})$ and (K^*, t^*, Y^*) are feasible

$$F_{t^*}^{K^*}(Y^*) = \sum_{k=1}^{K^*} y_k^* f_{t_k^*}(x_k^*) = \sum_{k=1}^{K^*} y_k^* \Theta(x_k^*) = \alpha^T b + \beta = \sum_{k=1}^{\tilde{K}} \tilde{y}_k \Theta^*(\tilde{x}_k).$$

Finally if Y^* is a KT point for $P_{t^*}^{K^*}$ then, by Lemma 1 (i) and (iii), Θ^* and Y^* have the properties required by this lemma. □

The above two lemmas are adaptations of [17, Corollary 1] and [17, Lemma 2] modified to allow for the different phases classes and paying attention to the fact that we may have $K^* \neq \tilde{K}$. The proof of the next theorem can be obtained by adapting that of [17, Theorem 3], where some care has to be given to the class of each phase involved and where [17, Lemma 2] should be replaced by Lemma 2 above in order not only to deal with the case involving multiple-phase-class models but also so as to address explicitly the issue of K being unknown in P (recall that Lemma 1, which is also needed, is valid for *fixed* K). A constraint qualification (see [17, Theorem 1]) is also required to show that all local minima occur at KT points. This relies on A being of full rank, which it is by the construction leading to (12).

Theorem 1 *Let $t^+ \in \tau^{K^+}$ and $Y^+ \in I\!R^{K^+(n+1)}$, for some positive integer K^+. Then, (K^+, t^+, Y^+) is a global minimum for P if and only if Y^+ is a KT point for $P_{t^+}^{K^+}$ and*

$$f_\eta(x) - \Theta^+(x) \geq 0, \quad \text{for all } x \in X \text{ and all } \eta \in \tau, \tag{19}$$

where $\{x : \Theta^+(x) = 0\}$ is the tangent hyperplane corresponding to (K^+, t^+, Y^+), as described by Lemma 1.

Note that to extend the local conditions given in Lemma 1 to the global conditions, we need to specify that Θ^* must not be above f_η for *any* $\eta \in \tau$ or *any* $x \in X$, rather than just at those phase classes t_k^* and compositions x_k^* occurring in the solution (K^*, t^*, Y^*).

We now introduce a lemma and theorem which clarify the geometric interpretation of the problem. They deal with the nonsmooth function $f_M : X \to I\!R$, which gives the minimum GFE of a single phase of given composition,

$$f_M(x) \equiv \min_{\eta \in \tau} f_\eta(x).$$

Lemma 3 *Let (K^+, t^+, Y^+), be a global minimum for P, $K^+ \in I\!N^+$, $t^+ \in \tau^{K^+}$, and $Y^+ \in I\!R^{K^+(n+1)}$, and $\{x : \Theta^+(x) = 0\}$ be the corresponding tangent hyperplane described by Lemma 1. Then we have that, at the points $\{x_k^+ : 1 \leq k \leq K^+ \text{ and } y_k^+ > 0\}$, f_M is differentiable and $\{x : \Theta^+(x) = 0\}$ is tangent to f_M.*

Proof: Let k be such that $1 \leq k \leq K^+$ and $y_k^+ > 0$. We have

$$f_M(x_k^+) = \min_{\eta \in \tau} f_\eta(x_k^+) \leq f_{t_k^+}(x_k^+) = \Theta^+(x_k^+), \tag{20}$$

by Lemma 1. We show that the inequality in (20) is in fact an equality. Suppose otherwise, then there exists a $\eta \in \tau$ such that $f_\eta(x_k^+) < \Theta^+(x_k^+)$. This contradicts Theorem 1. Hence,

$$f_M(x_k^+) = \Theta^+(x_k^+). \tag{21}$$

Furthermore, let $\epsilon > 0$ and $d \in I\!R^n$. We have , by (21),

$$\begin{aligned} f_M(x_k^+ + \epsilon d) - f_M(x_k^+) &= \min_{\eta \in \tau} f_\eta(x_k^+ + \epsilon d) - \Theta^+(x_k^+) \\ &\geq \min_{t \in \tau} \Theta^+(x_k^+ + \epsilon d) - \Theta^+(x_k^+), \end{aligned} \tag{22}$$

since, by Theorem 1, $f_\eta(x) \geq \Theta^+(x)$, for all $x \in X$ and $\eta \in \tau$. On the other hand, since $f_M(x_k^+) = f_{t_k^+}(x_k^+)$ [by (20) and (21)], we have

$$\begin{aligned} f_M(x_k^+ + \epsilon d) - f_M(x_k^+) &= \min_{\eta \in \tau} f_\eta(x_k^+ + \epsilon d) - f_{t_k^+}(x_k^+), \\ &\leq f_{t_k^+}(x_k^+ + \epsilon d) - f_{t_k^+}(x_k^+). \end{aligned} \tag{23}$$

Using (22), (23), dividing by ϵ and passing to the limit, we obtain

$$d^T \nabla \Theta^+ \leq \lim_{\epsilon \to 0^+} \frac{f_M(x_k^+ + \epsilon d) - f_M(x_k^+)}{\epsilon} \leq d^T \nabla f_{t_k^+}(x_k^+),$$

since Θ^+ is affine and $f_{t_k^+}$ is differentiable. Also, by definition of Θ^+,

$$d^T \nabla f_{t_k^+}(x_k^+) = d^T \nabla \Theta^+.$$

Thus, the above limit exists and is equal to $d^T \nabla \Theta^+$. This means that the directional derivative of f_M at x_k^+ in the direction d is given by $d^T \nabla \Theta^+$. Whence, f_M is differentiable at x_k^+ and $\nabla f_M(x_k^+) = \nabla \Theta^+$. $\square$

We thus obtain the following characterization of global minima for P in terms of the function f_M.

Theorem 2 (Tangent-Plane Criterion) *Let $K^+ \in \mathbb{N}^+$, $t^+ \in \tau^{K^+}$, and $Y^+ \in \mathbb{R}^{K^+(n+1)}$.*

If (K^+, t^+, Y^+) is a global minimum for P, then Y^+ is a KT point for $P_{t^+}^{K^+}$ and

$$f_M(x) - \Theta_M^+(x) \geq 0, \quad \textit{for all } x \in X, \tag{24}$$

where $\{x : \Theta_M^+(x) = 0\}$ is the tangent hyperplane corresponding to (K^+, t^+, Y^+), as described by Lemma 3.

Conversely, if Y^+ is a KT point for $P_{t^+}^{K^+}$ for some K^+ and $t^+ \in \tau^{K^+}$ and

$$f_M(x) - \Theta^+(x) \geq 0, \quad \textit{for all } x \in X, \tag{25}$$

where $\{x : \Theta^+(x) = 0\}$ is the tangent hyperplane corresponding to (K^+, t^+, Y^+) as described by Lemma 1, then (K^+, t^+, Y^+) is a global minimum for P and $\Theta^+ = \Theta_M^+$.

Proof: Let (K^+, t^+, Y^+) be a global minimum for P and let $x \in X$. Then by Theorem 1, Y^+ is a KT point for $P_{t^+}^{K^+}$ and $f_\eta(x) - \Theta^+(x) \geq 0$, for all $\eta \in \tau$. By Lemmas 1 and 3, $\Theta_M^+ = \Theta^+$. Hence, we have $f_\eta(x) - \Theta_M^+(x) \geq 0$, for all $\eta \in \tau$. In particular, for $t = \operatorname{argmin}\{f_\eta(x) : \eta \in \tau\}$, we have $f_M(x) - \Theta_M^+(x) \geq 0$.

Now let Y^+ be a KT point for $P_{t^+}^{K^+}$ for some t^+ and K^+, and assume that (25) holds. Let $(\tilde{K}, \tilde{t}, \tilde{Y})$, such that $\tilde{K} \in \mathbb{N}^+$, $\tilde{t} \in \tau^{\tilde{K}}$, and $\tilde{Y} \in \mathbb{R}^{\tilde{K}(n+1)}$, be a feasible point for P. We have

$$\begin{aligned}
F(K^+, t^+, Y^+) \equiv F_{t^+}^{K^+}(Y^+) &= \sum_{k=1}^{\tilde{K}} \tilde{y}_k \Theta^+(\tilde{x}_k), \quad \text{by Lemma 2,} \\
&\leq \sum_{k=1}^{\tilde{K}} \tilde{y}_k f_M(\tilde{x}_k), \quad \text{by (25),} \\
&\leq \sum_{k=1}^{\tilde{K}} \tilde{y}_k f_{\tilde{t}_k}(\tilde{x}_k), \quad \text{by definition of } f_M, \\
&\equiv F_{\tilde{t}}^{\tilde{K}}(\tilde{Y}) \equiv F(\tilde{K}, \tilde{t}, \tilde{Y}).
\end{aligned}$$

Thus, (K^+, t^+, Y^+) is a global minimum for P. Whence, from Lemma 3, $\Theta^+ = \Theta^+_M$.
□

4. Algorithm

A construction based on the following theorem will be used to eliminate unnecessary phases in the GILO algorithm which is introduced below. This will allow the number of phases at any one step to be limited to no more than $m+1$.

Theorem 3 *Let $Y^* \in \mathbb{R}^{K^*(n+1)}$ be a KT point for $P^{K^*}_{t^*}$, for some fixed positive integer K^* and some $t^* \in \tau^{K^*}$, and define $J^* \equiv \{k : 1 \le k \le K^*,\ \text{and}\ y_k > 0\}$. Let M be the dimension of the convex hull of the finite set $\{Ax^*_k : k \in J^*\}$. Then there exists $\tilde{K} \le M+1$ and a feasible solution $\tilde{Y} \in \mathbb{R}^{\tilde{K}(n+1)}$ for $P^{\tilde{K}}_{\tilde{t}}$, for some $\tilde{t} \in \tau^{\tilde{K}}$, such that*
(i) $\quad F^{\tilde{K}}_{\tilde{t}}(\tilde{Y}) = F^{K^*}_{t^*}(Y^*)$,
(ii) $\quad \tilde{y}_k > 0$, *for all* $1 \le k \le \tilde{K}$,
(iii) *for all* $1 \le k \le \tilde{K}$, *there is a* $j \in J^*$ *such that* $\tilde{x}_k = x^*_j$ *and* $\tilde{t}_k = t^*_j$.

Proof:
By (15), (16) and (17) , b is in the convex hull of the vectors $\{Ax^*_k : k \in J^*\}$. By Caratheodory's theorem [18], the vector b can be written as a convex combination of $\tilde{K} \le M+1$ of the vectors $\{Ax^*_k : k \in J^*\}$. Let the scalars $\tilde{y}_k$ for $k = 1, ..., \tilde{K}$ be the coefficients in this convex combination, and let $\tilde{x}_k$ and $\tilde{t}_k$ be the corresponding compositions and phase classes. Without loss of generality assume $\tilde{y}_k > 0$ for all $1 \le k \le \tilde{K}$.

Then $(\tilde{K}, \tilde{t}, \tilde{Y})$ is feasible for P, since by construction

$$\sum_{k=1}^{\tilde{K}} \tilde{y}_k = 1, \quad \sum_{k=1}^{\tilde{K}} \tilde{y}_k A\tilde{x}_k = b \ \text{ and } \ \tilde{t_k} \in \tau, \ \ \tilde{y}_k > 0, \ \ \tilde{x}_k \in X, \quad 1 \le k \le \tilde{K}.$$

Let Θ^* be the affine function corresponding to the KT point Y^* for $P^{K^*}_{t^*}$, as described by Lemma 1.

Since by construction the new feasible solution $(\tilde{K}, \tilde{t}, \tilde{Y})$ consists of phases with the same composition and phase class as some phase in (K^*, t^*, Y^*), for all k, $1 \le k \le \tilde{K}$, there is a $j \in J^*$ such that $\tilde{x}_k = x_j$ and $\tilde{t}_k = t_j$. From this and Lemma 1, using the fact that Y^* is a KT point, it follows that

$$\Theta^*(\tilde{x}_k) = \Theta^*(x^*_j) = f_{t^*_j}(x^*_j) = f_{\tilde{t}_k}(\tilde{x}_k). \tag{26}$$

From Lemma 2, using the fact that $(\tilde{K}, \tilde{t}, \tilde{Y})$ is feasible, and from (26) it follows that

$$F^{K^*}_{t^*}(Y^*) = \sum_{k=1}^{\tilde{K}} \tilde{y}_k \Theta^*(\tilde{x}_k) = \sum_{k=1}^{\tilde{K}} \tilde{y}_k f_{\tilde{t}_k}(\tilde{x}_k) \equiv F^{\tilde{K}}_{\tilde{t}}(\tilde{Y}).$$

This proves part (i). Parts (ii) and (iii) follow from the construction. □

Note that the composition of each phase in the $\tilde{Y}$ is the same as the composition of some non-zero phase in Y^*, and the standard proof of Caratheodory's theorem provides a simple constructive method for eliminating unnecessary phases and calculating the values of $\tilde{y}_k$. We shall use the notation $(\tilde{K}, \tilde{t}, \tilde{Y}) := \mathrm{EL}(K^*, t^*, Y^*)$ to refer to this procedure for eliminating phases from the KT point Y^* of $P_{t^*}^{K^*}$ to give the feasible solution $(\tilde{K}, \tilde{t}, \tilde{Y})$ of P with the same objective value.

Corollary 1 *There is a globally optimal solution to P which has no more than $m+1$ (i.e. m_b) phases.*

Proof: Since the matrix A is of rank m, the convex hull of the set $\{Ax_k^* : k \in J^*\}$ is of dimension at most m. Hence $M \leq m$, and the the result then follows from Theorem 3. □

(Note in passing that it is clear that a feasible solution for P with K phases can be extended to a solution with the same objective value and $\tilde{K} > K$ phases by introducing phases k, $K < k \leq \tilde{K}$, with $y_k = 0$ and arbitrary $x_k \in X$ and $t_k \in \tau$. This and Corollary 1 show that P could be solved by fixing the number of phases to $m+1$. Solving this problem directly as a global optimization problem is however extremely difficult and the following approach is superior.)

For the generic algorithm we are presenting to solve problem P, we assume we have available the elimination procedure EL described above, a Local Optimization procedure, called LO, and a global optimization procedure (for a lower-dimensional search space) called GI (Global Improvement). Given a feasible solution, (K, t, Y), for P, the local optimization procedure must be such that when started from Y it returns a KT point, $Y^* := \mathrm{LO}(K, t, Y) \in I\!R^{K(n+1)}$, for P_t^K, with a function value no higher than at (K, t, Y), i.e. $F_t^K(Y^*) \leq F_t^K(Y)$. Given an affine function, $\Theta^* : I\!R^n \to I\!R$, the global optimization procedure returns a couple, $(t_0, x_0) := \mathrm{GI}(\Theta^*)$, where $x_0 \in X$ and $t_0 \in \tau$ are any feasible point and phase class at which $f_{t_0}(x_0) - \Theta^*(x_0) < 0$, if such exists, and otherwise x_0 is a global minimum for the subproblem

$$\min_{x \in X} \; f_M(x) - \Theta^*(x), \tag{27}$$

and $t_0 \in \tau$ is such that $f_{t_0}(x_0) = f_M(x_0)$.

GILO Algorithm

Step 0: [Initialization:]
Find a feasible solution (K, t, Y) for P with $K \leq m+1$. (For example solve P^{feas}. If infeasible stop – problem is infeasible. If objective is 0, remove substances and repeat. Let Y be feasible solution found, set $K := 1$ and choose any $t \in \tau^1$.)
Step 1: [Local optimization]
$Y^* := \mathrm{LO}(K, t, Y)$, $K^* := K$, $t^* := t$, $J^* := \{k : 1 \leq k \leq K^*, y_k^* > 0\}$.

Select any $j \in J^*$ and define
$\Theta^* : I\!R^n \to I\!R$ by $\Theta^*(x) := f_{t_j^*}(x_j^*) + \nabla f_{t_j^*}(x_j^*)^T(x - x_j^*)$.
$(\tilde{K}, \tilde{t}, \tilde{Y}) := \mathrm{EL}(K^*, t^*, Y^*)$.
Step 2: [Global improvement subproblem]
$(t_0, x_0) := \mathrm{GI}(\Theta^*)$.
Step 3: [Global optimality condition]
If $f_M(x_0) - \Theta^*(x_0) \geq 0$, then stop: [$(\tilde{K}, \tilde{t}, \tilde{Y})$ is a global minimum for P].
Step 4: [Force improvement]
Solve

$$\sum_{k=1}^{\tilde{K}} \alpha_k = 1, \ \sum_{k=1}^{\tilde{K}} \alpha_k A\tilde{x}_k = Ax_0 \tag{28}$$

for $\{\alpha_k\}_{k=1}^{\tilde{K}}$.

- (i) [Phase interchange]
If (28) has a solution, $\{\alpha_k\}_{k=1}^{\tilde{K}}$:

  [A new improved feasible solution $(\hat{K}, \hat{t}, \hat{Y})$ to P is given by]

$$\begin{aligned} \hat{K} &:= \tilde{K} + 1, \\ \hat{y}_{\hat{K}} &:= \min\{\frac{\tilde{y}_k}{\alpha_k} : \alpha_k > 0, 1 \leq k \leq \tilde{K}\}, \ \hat{x}_{\hat{K}} := x_0, \ \hat{t}_{\hat{K}} := t_0, \\ \hat{y}_k &:= \tilde{y}_k - \hat{y}_0\alpha_k, \ \hat{x}_k := \tilde{x}_k, \ \hat{t}_k := \tilde{t}_k, \ 1 \leq k \leq \tilde{K}. \end{aligned}$$

  [At least one $\hat{y}_k$ is zero.] Eliminate from $(\hat{K}, \hat{t}, \hat{Y})$ all phases k such that $\hat{y}_k = 0$ and renumber phases to give solution (K,t,Y). [This will have the same objective function value.]

- (ii) [Phase split]
Otherwise (the system (28) has no solution):

  Select any j such that $1 \leq j \leq \tilde{K}$. Choose $\delta > 0$ to be small enough so that the new solution (K, t, Y) given below is feasible for P and achieves a strict decrease in F. (Such a δ can be found for example by successive halving.)

$$\begin{aligned} K := \tilde{K} + 1 \ &, \ t_K := t_0, \\ y_K := \frac{\delta}{\delta + 1}\tilde{y}_j \ &, \ x_K := x_0, \\ y_j := \frac{1}{\delta + 1}\tilde{y}_j \ &, \ x_j := \delta(\tilde{x}_j - x_0) + \tilde{x}_j, \\ t_k := \tilde{t}_k \ &, \ 1 \leq k \leq \tilde{K}, \\ y_k := \tilde{y}_k \ &, \ x_k := \tilde{x}_k, \quad 1 \leq k \leq \tilde{K}, \ k \neq j. \end{aligned}$$

Go to Step 1.

Note that in the PEP, $A = I$ and so there is a unique single-phase solution given by $y_1 = 1$ and $x_1 = b$. In this case the LO step (Step 1) need not be performed in the first iteration and we can set $(K^*, t^*, Y^*) := (K, t, Y)$. The initial value of the phase class can be chosen arbitrarily or chosen to be the one with the lowest GFE at the initial solution.

Step 4(ii) of the algorithm corresponds physically to *splitting* a phase (phase j) into two phases (the modified phase j and phase $\tilde{K}+1$). Hence, after performing this step, the number of phases is always increased by one. Step 4(i) simply interchanges one phase for another, keeping the number of phases constant or reducing it.

The GILO algorithm is illustrated in Figures 1 and 2 for the case where $m = n = 1$, $A = I$, and $\tau = \{V, L\}$ ('V' for 'vapour', 'L' for 'liquid'). The dotted line represents the nonsmooth function f_M. Values of x_1^*, x_2^*, Θ^*, and x_0 obtained at Steps 1 and 2 are shown. Let superscripts l and $l+1$ refer respectively to iteration number l and iteration number $l+1$. Here ${K^*}^l = {K^*}^{l+1} = 2$. At iteration l (Figure 1), we have two liquid phases (${t_1^*}^l = {t_2^*}^l = L$). The elimination step cannot remove any phases so $(\tilde{K}, \tilde{t}, \tilde{Y}) = (K^*, t^*, Y^*)$. The global step (Step 2) found a vapour phase x_0^l ($t_0^l = V$). Since x_0^l can be written as an affine combination of $\tilde{x}_1^l$ and $\tilde{x}_2^l$ (i.e. (28) has a solution), we do not add x_0^l as a third phase but rather simply *exchange* phase $\tilde{x}_2^l$ for phase x_0^l, as described in Step 4 (i). The local optimization step (Step 1) will then yield the global solution corresponding to ${x_1^*}^{l+1}$, ${x_2^*}^{l+1}$, ${t_1^*}^{l+1} = L$, and ${t_2^*}^{l+1} = V$ shown on Figure 2. Note that at each iteration we have $F(K^*, t^*, Y^*) = \Theta^*(b)$, since $F(K^*, t^*, Y^*) \equiv \sum y_k^* f_{t_k^*}(x_k^*)$, $f_{t_k^*}(x_k^*) = \Theta^*(x_k^*)$, Θ^* is affine, and $b = \sum y_k^* x_k^*$ (or simply by Lemma 2 with $\tilde{K} = 1$, $\tilde{y}_1 = 1$, and $\tilde{x}_1 = b$).

Theorem 4 (Convergence) *The GILO algorithm converges to a global minimum of P in a finite number of steps, provided that P_t^K has a finite number of distinct objective function values at KT points, for any $1 \le K \le m+1$ and any $t \in \tau^K$.*

Proof: We first show that every step of the algorithm is well defined.

As noted earlier P' and hence P attain their infimum as a minimum.

In Step 0, P^{feas} can be used to yield a single-phase feasible solution if one exists, and otherwise prove that the problem is infeasible.

Now consider Step 4. When (28) has a solution, $\{\alpha_k\}_{k=1}^{\tilde{K}}$, Step 4(i) will be taken. Since there must exists $\alpha_k > 0$ with $1 \le k \le \tilde{K}$, and since $\tilde{y}_k > 0$, $1 \le k \le \tilde{K}$ (EL in Step 1 removed redundant phases), $\hat{y}_{\hat{K}}$ is well defined and strictly positive. Next we show that the new solution given by Step 4(i) is feasible for P. Firstly, by (28) and by (16) using the fact that $(\tilde{K}, \tilde{t}, \tilde{Y})$ is feasible,

$$\sum_{k=1}^{\hat{K}} \hat{y}_k = \sum_{k=1}^{\tilde{K}} (\tilde{y}_k - \hat{y}_{\hat{K}} \alpha_k) + \hat{y}_{\hat{K}} = \sum_{k=1}^{\tilde{K}} \tilde{y}_k - \hat{y}_{\hat{K}} \sum_{k=1}^{\tilde{K}} \alpha_k + \hat{y}_{\hat{K}} = 1.$$

Secondly, by (28) and by (15) using the fact that $(\tilde{K}, \tilde{t}, \tilde{Y})$ is feasible,

$$\begin{aligned}
\sum_{k=1}^{\hat{K}} \hat{y}_k A\hat{x}_k &= \sum_{k=1}^{\tilde{K}} (\tilde{y}_k - \hat{y}_{\hat{K}}\alpha_k) A\tilde{x}_k + \hat{y}_{\hat{K}} A\hat{x}_{\hat{K}} \\
&= \sum_{k=1}^{\tilde{K}} \tilde{y}_k A\tilde{x}_k - \hat{y}_{\hat{K}} \sum_{k=1}^{\tilde{K}} \alpha_k A\tilde{x}_k + \hat{y}_{\hat{K}} A\hat{x}_{\hat{K}} \\
&= b - \hat{y}_{\hat{K}} Ax_0 + \hat{y}_{\hat{K}} Ax_0 = b.
\end{aligned}$$

Thirdly, $\hat{y}_k \geq 0$, $1 \leq k \leq \hat{K}$, by construction, so the solution constructed in Step 4(i) is feasible for P, and so (K, t, Y), which is obtained from it by dropping phases with $\hat{y}_k = 0$, is also feasible for P.

Now we show that the new solution, (K, t, Y), given in Step 4(ii) is feasible for P. Firstly it is straightforward to verify that

$$\sum_{k=1}^{K} y_k = \sum_{k=1}^{\tilde{K}} \tilde{y}_k, \quad \text{and} \quad \sum_{k=1}^{K} y_k Ax_k = \sum_{k=1}^{\tilde{K}} \tilde{y}_k A\tilde{x}_k.$$

We then use the feasibility of $(\tilde{K}, \tilde{t}, \tilde{Y})$. Secondly, $x_k \in X$, for δ sufficiently small, since X is open. Finally $y_k > 0$ for $1 \leq k \leq K$.

We now show that if the algorithm terminates in Step 3, then (K^*, t^*, Y^*) and $(\tilde{K}, \tilde{t}, \tilde{Y})$ are global minima for P. Note that Y^* is a KT point for $P_{t^*}^{K^*}$ and that the Θ^* defined in Step 3 is the same as that defined in Lemma 1. Since termination occurred, $f_M(x) - \Theta^*(x) \geq 0$ for all $x \in X$, so by Theorem 2 (K^*, t^*, Y^*) is a global minimum for P. Since, by the EL construction, $(\tilde{K}, \tilde{t}, \tilde{T})$ is feasible, Theorem 3 shows that it is also optimal.

Next we show that Step 4(i) yields a strict decrease in F. Since termination did not occur in step 3, $f_{t_0}(x_0) < \Theta^*(x_0)$, so by definition of phase $\hat{K}$, $f_{\hat{t}_{\hat{K}}}(\hat{x}_{\hat{K}}) < \Theta^*(\hat{x}_{\hat{K}})$. Also by Lemma 1, $f_{t_j^*}(x_j^*) = \Theta^*(x_j^*)$ for all $j \in J^*$. By construction in EL, for all k such that $1 \leq k \leq \tilde{K}$, there is a $j \in J^*$ such that $\tilde{x}_k = x_j$, and by definition $\hat{x}_k = \tilde{x}_k$ and $\hat{t}_k = \tilde{t}_k$. Hence $f_{\hat{t}_k}(\hat{x}_k) = \Theta^*(\hat{x}_k)$ for $1 \leq k \leq \tilde{K}$. Hence the new value for F at the end of Step 4(i) is

$$\begin{aligned}
F(K, t, Y) &= F(\hat{K}, \hat{t}, \hat{Y}), && \text{empty phases eliminated,} \\
&= \sum_{k=1}^{\hat{K}} \hat{y}_k f_{\hat{t}_k}(\hat{x}_k) \\
&< \sum_{k=1}^{\hat{K}} \hat{y}_k \Theta^*(\hat{x}_k), && \text{paragraph above,} \\
&= F(K^*, t^*, Y^*), && \text{Lemma 2.}
\end{aligned}$$

For Step 4(ii), the proof of the strict decrease in F can be found in the modification of the proof of [17, Theorem 3] that we described above for our Theorem 1.

Since A has rank m the maximum value of M in Theorem 3 is m. Hence Theorem 3 shows that the maximum number of phases which can be present after the EL step is $m+1$, and that this only occurs when the the convex hull of $\{Ax_k : k \in J^*\}$ has dimension m. It follows that in this case the vectors $\{Ax_k : k \in J^*\}$ span $I\!R^m$ so there must be a solution to (28). Hence Step 4(i) is taken, which either keeps the same number of phases or decreases it. In the other case, Step 4(ii) increases the number of phases only by one, so the resulting number of phases in either case is no more than $m+1$. The initial feasible solution has no more than $m+1$ phases so every LO is a problem with at most $m+1$ phases, and therefore the KT points found have no more that $m+1$ phases. Step 4 forces a strict reduction in the objective function and the next LO terminates at a KT point with an objective no higher than at the start of the LO search. Hence the sequence of KT points has strictly decreasing objective values. Since we are assuming that there are a finite number of distinct function values at KT points for problems with $m+1$ or fewer phases, it follows that the algorithm converges in a finite number of steps.

Finally we note that since A is of full rank a constraint qualification holds (see [17]), so all local minima occur at KT points. Hence the algorithm converges to a global minimum. □

5. Implementation and Computational Results

In this section we make some observations related to the practical implementation of the GILO algorithm, and we cite encouraging computational experiments.

In the case where one phase class is known to be modelled by a convex function, one can show that no more than one phase of this class needs to occur at a global minimum of P. Moreover the GILO algorithm will not introduce more than one phase of such a class. This is because at the end of the LO step the tangent plane $\{x : \Theta^*(x) = 0\}$ is a tangent to all phase functions corresponding to phases which are present (Lemma 1), so the tangent lies below any such function which is convex. The GI step can therefore never find a point below the tangent for such a phase class, so the GI step can be simplified by ignoring any convex phase class where a phase of that class already exists. Where a convex phase class could be present because there is currently no phase of that class present, the phase class can be dealt with separately within the GI step by a fast local minimization technique.

If the GI procedure provides a lower bound on its optimal objective value when it terminates, then the following theorem can be used to obtain a lower bound on the GFE.

Theorem 5 (Lower Bound) *Let $\Theta : I\!R^n \to I\!R$ be an affine function and let $\underline{\Theta}$ be the optimal objective value of the LP problem*
$P^{\mathrm{lb}}(\Theta)$:

$$\underline{\Theta} = \min_{Z \in I\!R^n} \Theta(Z),$$

subject to $\quad AZ = b,$
$\quad\quad\quad\quad\quad Z \geq 0.$

Assume that when the GI(Θ) procedure terminates, it provides a lower bound $\underline{D}$ on its optimal objective value.
Then for all (K, t, Y) that are feasible for P

$$F(K, t, Y) \geq \underline{\Theta} + \underline{D}.$$

In the PEP, $\underline{\Theta} = \Theta(b)$.

Proof: By the definition of $\underline{D}$

$$f_\eta(x) - \Theta(x) \geq \underline{D} \;\text{ for all } x \in X \;\text{ and }\; \eta \in \tau. \tag{29}$$

Let (K, t, Y) be any feasible solution to P. Then

$$\begin{aligned} F(K,t,Y) &\equiv \sum_{k=1}^{K} y_k f_{t_k}(x_k), \\ &\geq \sum_{k=1}^{K} y_k(\Theta(x_k) + \underline{D}), \quad \text{by (29)}, \\ &= \Theta(\sum_{k=1}^{K} y_k x_k) + \underline{D}, \quad \text{since } \sum_{k=1}^{K} y_k = 1 \text{ and } \Theta \text{ is affine}, \\ &\geq \underline{\Theta} + \underline{D}, \quad \text{since } Z \equiv \sum_{k=1}^{K} y_k x_k \text{ is feasible for } P^{\mathrm{lb}}(\Theta). \end{aligned}$$

For the PEP there is a unique feasible point for $P^{\mathrm{lb}}(\Theta)$ given by $Z = b$, so $\underline{\Theta} = \Theta(b)$. $\square$

Note that the validity of the bounds provided by the Theorem 5 does not rely on any special properties of Θ. In particular it does not rely on the Θ^* provided by LO being exact or indeed being of the form given in Lemma 1 for any α and β. (However the bound is better if the LO step is exact.) The ϵ-global optimality results quoted in the next section use this bound.

Let Θ^+ be the affine function defined in Lemma 1 corresponding to the global optimum (K^+, t^+, Y^+). Assume Z is optimal for $P^{\mathrm{lb}}(\Theta^+)$ and let $K = 1, y_1 = 1, x_1 = Z$ and $t_1 \in \tau$ be arbitrary. Then (K, t, Y) is feasible for P and from Lemma 2 it follows that $\underline{\Theta} = \Theta^+(Z) = y_1\Theta^+(x_1) = F(K^+, t^+, Y^+)$. Also if GI($\Theta^+$) is completed and finds its global minimum, it has a zero objective value, so $\underline{D} = 0$ in this case. Hence in this case the bound given by Theorem 5 is attained.

Figure 3 illustrates the lower bound (LB) implicitly given by the iterate, (K^*, t^*, Y^*), Θ^*, and x_0 of Figure 1. Since LO is exact $\epsilon = -\underline{D}$ in this case.

5.1. Results from Interval-GILO Implementations

In [24], McKinnon et al. describe an implementation of the GILO algorithm, Interval-GILO, where the GI global step uses branch and bound and interval analysis. The GI step is terminated whenever a point is found lying sufficiently below the tangent plane. Encouraging results from this implementation and later improvements on it are summarized below. Full details can be found in [24, 3].

All the results reported for the Interval-GILO implementations give CPU times in seconds for runs on a Sun Sparc 5 70MHz workstation. The number of local-optimization iterations reported refers to the number of iterations needed for LO step to converge. The number of global-optimization iterations corresponds to the total number of times an interval has been divided in two in the branch-and-bound interval-analysis process (see [24]).

In [24], McKinnon et al. describe the behaviour of Interval-GILO on three instances of the PEP which are also studied by McDonald & Floudas in [21]. The NAG E04UCF subroutine (sequential quadratic programming) was used for the LO step. A summary of the results is given below. Table 1 reports numbers of iterations and CPU times required to obtain convergence.

Problem 1 corresponds to the Illustrative Example (n-Butyl-acetate - Water) of [21] ($n_s = 2$ and $\tau = \{L\}$). Started from the only feasible point, which corresponds to a single liquid phase, Interval-GILO performed a GI step which yielded a second liquid phases, then a LO step, and then a second GI which proved that the solution found in the LO step was ϵ-globally optimal, where $\epsilon = 10^{-10}$ (i.e. by Theorem 5, the GFE at the solution is known to be within ϵ of the global minimum objective value.) The global minimum found by the algorithm therefore corresponds to two liquid phases. McDonald & Floudas [21] report 1.23 seconds of CPU time on a Hewlett Packard 9000/730 to converge to the same solution, with precision $\epsilon = 5 \times 10^{-4}$. (In [21] each problem is solved as a single global optimization problem using the GOP method with the optimal number of phases assumed to be known.)

Problems 2a and 2b (Benzene - Water - Acetonitrile) ($n_s = 3$ and $\tau = \{L, V\}$) originate from Castilla and Grossmann [5] and correspond to Example 6 of [21]. These have Liquid-Liquid-Vapour and Liquid-Vapour Equilibria respectively. Solutions were found to ϵ-global optimality with $\epsilon < 1.2 \times 10^{-9}$. These results again compare favourably with those in [21], which reports 766 and 118 seconds of CPU time respectively for Problem 2a and Problem 2b, to converge to the same solution. When McDonald & Floudas assume that it is known that there is one vapour phase and one liquid phase at the global optimum of Problem 2b, their algorithm then requires 0.88 seconds of CPU time to converge with $\epsilon = 10^{-4}$. Note also that in a recent paper [23], they report a CPU time of 0.76 seconds to *verify* the global optimality of the global minimum of Problem 2a.

Results from the Interval-GILO method on more difficult problems are described in [3] and are summarized in Table 2. These problems are modelled using the UNIFAC equation and involve 4, 5 and 6 of the substances (Ethylene Glycol, Do-

Table 1. Results for Problems 1, 2a, and 2b

	Problem 1		Problem 2a		Problem 2b	
Initial phase	liquid		vapour		vapour	
	iterations	sec	iterations	sec	iterations	sec
1^{st} global step	8	0.05	2	0.04	269	0.99
1^{st} local step	9	0.07	10	0.08	11	0.10
2^{nd} global step	46	0.16	2	0.04	297	1.42
2^{nd} local step			33	0.22		
3^{rd} global step			436	3.11		
entire program		0.38		3.65		2.60

† : vapour phase introduced by a local search.

Table 2. Results for Problems 3, 4, and 5

	Problem 3		Problem 4		Problem 5	
Initial phase	liquid		liquid		liquid	
	iterations	CPU	iterations	CPU	iterations	CPU
1^{st} global step†	6	0.03	8	0.06	8	0.08
1^{st} local step	8	0.04	6	0.05	5	0.05
2^{nd} global step	15	0.08	856	7.18	5888	87.20
2^{nd} local step	129	0.82	159	1.78	17	0.32
3^{rd} global step	3016	15.65	57450	667.36	1069179	19983.08
3^{rd} local step	5	0.07				
4^{th} global step	7975	54.22				
entire program		71.05		677.37		20071.41

decanol, Nitromethane, Water, Benzene, n-Butanol), and have up to 4 phases at equilibrium. The problems are solved to ϵ-global optimality with $\epsilon < 10^{-10}$ in all cases. In this implementation the LO step operates in a reduced space and uses an unconstrained modified Newton algorithm. Problem 3 is the most difficult problem described in [23], where it took 1960 seconds to obtain and verify the global optima, compared with 71 seconds by the Interval-GILO method.

Using the current version of the Interval-GILO method the times for Problems 1, 2a, 2b, 3, 4 and 5 are reduced to 0.26, 1.39, 1.16, 53.97, 509.54, 14227.49 seconds respectively.

The interval GI step of the GILO algorithm is suitable for parallel processing, and in [3] a parallel implementation is described which achieves speedups of 10 on a network of workstations for problems 3, 4 and 5.

6. Conclusions

In this paper, we have extended the necessary and sufficient condition for global optimality based on the tangent-plane criterion to the case involving multiple-phase-class models. Moreover we have presented an algorithmic approach that reduces the global optimization problem of minimizing the Gibbs free energy in the multi-phase chemical and phase equilibrium problem into a *finite* sequence of local optimization (LO) steps involving no more that m_b phases, where m_b is the number of independent balance constraints, and global optimization (GI) steps in the smaller space of $(n_s - 1)$ dimensions, where n is the number of substances present. The GI step uses the tangent-plane criterion to determine whether the current solution is optimal, and, if it is not, it finds an improved feasible solution either with the same number or fewer phases, or with one added phase. It also determines what class of phase is to be added.

The major advantages of the method are that the global optimization steps are in a much lower dimensional space than the whole problem, and that the numerical difficulties occur in the local optimization step, for which well developed methods are available.

The generic GILO algorithm can be used with different global optimization methods. A summary has been given in this paper of good numerical results reported in [24, 3] on several instances of the GFE minimization problem with an interval-analysis implementation for the global step of the GILO algorithm. Such computational experiments support the theory.

There is still some scope for improving the current interval-analysis-based GI step, and it would also be of interest to test other global optimization methods for the GI step such as that used for the stability test of McDonald & Floudas' GLOPEQ package [23] and in [19, 22]. Future work should attempt to solve the numerically difficult local optimization subproblems more efficiently, for example along the lines proposed by Trangenstein in [36] for phase equilibrium problems involving two phases. A more accurate solution to the LO problem would produce even more accurate ϵ-global optimal error bounds.

Sometimes a phase-class model may only be valid over part of the composition space. For example, when it is known that a substance can only occur in a phase with very low concentration, it is often a good approximation and is easier numerically to introduce a phase-class model which does not contain that substance. This situation is described by a multiple-phase-class model with special properties [33] and requires further study.

Acknowledgement

The authors are indebted to S. Furnival and B. Morton for interesting discussions on the chemical and phase equilibrium problem, to C. Millar for providing recent results from his Interval-GILO implementation, and also to W. R. Smith and the anonymous referees for comments which strengthened the paper.

References

1. M. N. Ammar and H. Renon. The isothermal flash problem: New methods for phase split calculations. *American Institute of Chemical Engineering Journal*, 33(6):926–939, 1987.
2. L. E. Baker, A. C. Pierce, and K. D. Luks. Gibbs energy analysis of phase equilibria. *Society of Petroleum Engineers Journal*, 22:731–742, 1982.
3. S. Berner, K. I. M. McKinnon, and C. G. Millar. A parallel algorithm for the global minimization of Gibbs free energy. Technical Report MS-96-004, Dept. of Mathematics and Statistics, University of Edinburgh, U.K., 1996.
4. K. M. Brantferger. *Development of a thermodynamically consistent, fully implicit, compositional, equation-of-state, steamflood simulator.* PhD thesis, University of Texas at Austin, 1991.
5. J. Castillo and I. Grossmann. Computation of phase and chemical equilibria. *Computers & Chemical Engineering*, 9:99, 1981.
6. R. J. Clasen. The solution of the chemical equilibrium programming problem with generalized Benders decomposition. *Operations Research*, 32:70–79, 1984.
7. J. E. Dennis and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* Prentice-Hall, 1983.
8. J. Dluzniewski and S. Adler. Calculation of complex reaction and/or phase equilibria problems. In *Institution of Chemical Engineers Symposium Series N° 35*, pages 21–26, London, 1972.
9. C. A. Floudas and P. M. Pardalos. *A Collection of Test Problems for Constrained Global Optimization Algorithms*, volume 455 of *Lecture Notes in Computer Science.* Springer-Verlag, 1987.
10. C. A. Floudas and V. Visweswaran. A global optimization algorithm (GOP) for certain classes of nonconvex NLPs: Theory. *Computers & Chemical Engineering*, 14(12):1397–1417, 1990.
11. C. A. Floudas and V. Visweswaran. Primal-relaxed dual global optimization approach. *Journal of Optimization Theory & Applications*, 78(2):187–225, 1993.
12. R. Gautam and W. D. Seider. Computation of phase and chemical equilibrium. *American Institute of Chemical Engineering Journal*, 25(6):991–1015, 1979.
13. A. M. Geoffrion. Generalized Benders decomposition. *Journal of Optimization Theory & Applications*, 10(4):237–260, 1972.
14. J. W. Gibbs. Graphical methods in thermodynamics of fluids. *Trans. Connecticut Acad*, 2:311, 1873.
15. J. W. Gibbs. A method of geometrical representation of the thermodynamic properties of substances by means of surfaces. *Trans. Connecticut Acad*, 2:382, 1873.
16. A. Griewank, N. R. Nagarajan, and A. S. Cullick. New strategy for phase-equilibrium and critical-point calculations by thermodynamic energy analysis I: stability analysis and flash. *Fluid Phase Equilibria*, 62:191–210, 1991.
17. Y. Jiang, W. R. Smith, and G. R. Chapman. Global optimality conditions and their geometric interpretation for the chemical and phase equilibrium problem. *SIAM Journal on Optimization*, 5(4):813–834, 1995.
18. S. R. Lay. *Convex Sets and Their Applications.* Wiley-Interscience, 1982.

19. C. M. McDonald and C. A. Floudas. Decomposition based and branch and bound global optimization approaches for the phase equilibrium problem. *Journal of Global Optimization*, 5(3):205–251, 1994.
20. C. M. McDonald and C. A. Floudas. Global optimization for the phase equilibrium problem using the NRTL equation. In *Escape 4: 4th European Symposium on Computer Aided Process Engineering. IChemE Symposium Series no 133*, pages 273–280. Institution of Chemical Engineers, 1994.
21. C. M. McDonald and C. A. Floudas. Global optimization for the phase and chemical equilibrium problem: Application to the NRTL equation. *Computers & Chemical Engineering*, 19(11):1111–1141, 1995.
22. C. M. McDonald and C. A. Floudas. Global optimization for the phase stability problem. *American Institute of Chemical Engineering Journal*, 41(7):1798–1814, 1995.
23. C. M. McDonald and C. A. Floudas. GLOPEQ: A new computational tool for the phase and chemical equilibrium problem. *Computers & Chemical Engineering*, 1996. To appear.
24. K. I. M. McKinnon, C. G. Millar, and M. Mongeau. Global optimization for the chemical and phase equilibrium problem using interval analysis. In C. Floudas and P. Pardalos, editors, *State of the Art in Global Optimization*, pages 365–382. Kluwer Academic Publishers, 1995.
25. R. K. Mehra, R. A. Heidemann, and K. Aziz. An accelerated successive substitution algorithm. *Canadian Journal of Chemical Engineering*, 61:590–596, 1983.
26. M. L. Michelsen. The isothermal flash problem. Part I. Stability. *Fluid Phase Equilibria*, 9:1–19, 1982.
27. M. L. Michelsen. The isothermal flash problem. Part II. Phase-split calculation. *Fluid Phase Equilibria*, 9:21–40, 1982.
28. L. X. Nghiem and Y.-K. Li. Computation of multiphase equilibrium phenomena with an equation of state. *Fluid Phase Equilibria*, 17:77–95, 1984.
29. D. Y. Peng. Gibbs energy of mixing from empirical models. *Canadian Journal of Chemical Engineering*, 67(3):462–467, 1989.
30. D. Y. Peng and D. B. Robinson. A new two-constant equation of state. *Industrial and Engineering Chemistry : Fundamentals*, 15:59–64, 1976.
31. N. Z. Shapiro and L. S. Shapley. Mass action laws and Gibbs free energy function. *Journal of the Society for Applied and Industrial Mathematics*, 13:353–375, 1965.
32. J. M. Smith and H. C. Van Ness. *Introduction to Chemical Engineering Thermodynamics*. McGraw-Hill, fourth edition, 1987.
33. J. V. Smith, R. W. Missen, and W. R. Smith. General optimality criteria for multiphase multireaction chemical equilibrium. *American Institute of Chemical Engineering Journal*, 39(4):707–710, 1993.
34. W. R. Smith and R. W. Missen. *Chemical Reaction Equilibrium Analysis: Theory and Algorithms*. Wiley & Sons, 1982.
35. A. C. Sun and W. D. Seider. Homotopy-continuation algorithm for global optimization. In C. A. Floudas and P. M. Pardalos, editors, *Recent Advances in Global Optimization*, pages 561–592. Princeton University Press, 1992.
36. J. A. Trangenstein. Customized minimization techniques for phase equilibrium computations in reservoir simulation. *Chemical Engineering Science*, 42(12):2847–2863, 1987.
37. L. A. Wolsey. A resource decomposition algorithm for general mathematical programs. *Mathematical Programming Study*, 14:244–257, 1981.
38. W. Xiao, K. Zhu, W. Yuan, and H. H. Chien. An algorithm for simultaneous chemical and phase equilibrium calculation. *American Institute of Chemical Engineering Journal*, 35(11):1813–1820, 1989.

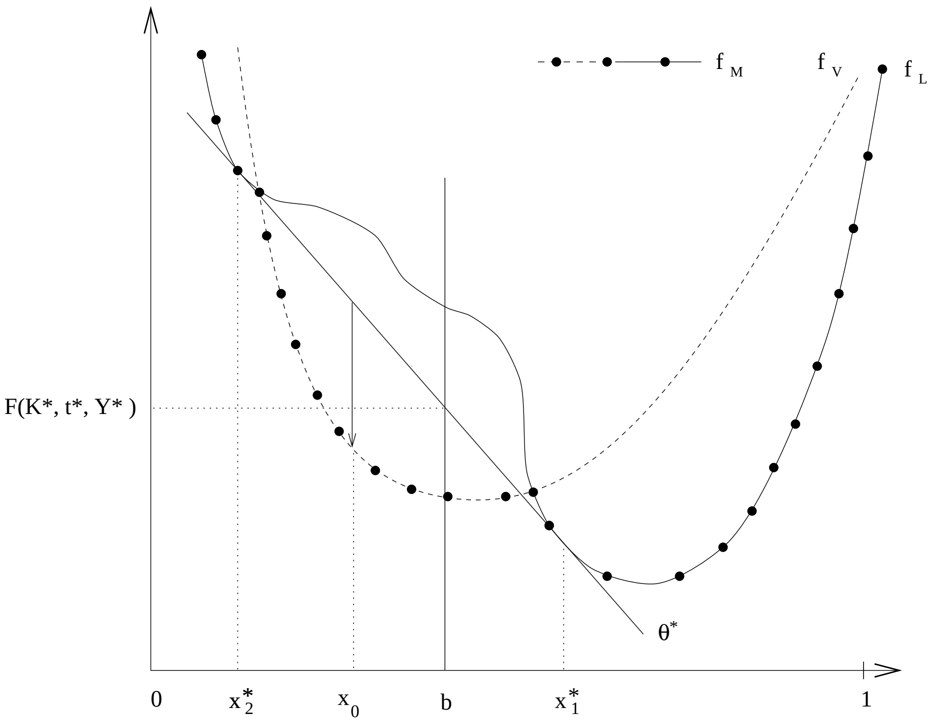

Figure 1. Local optimum and completed GI step at iteration l

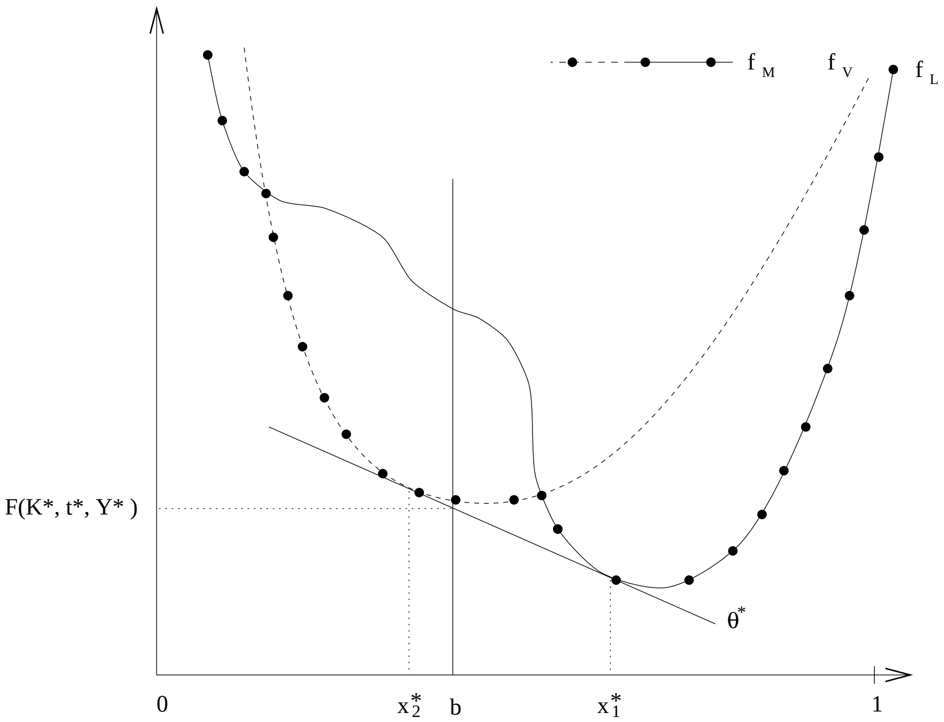

Figure 2. Global optimum at iteration $l + 1$

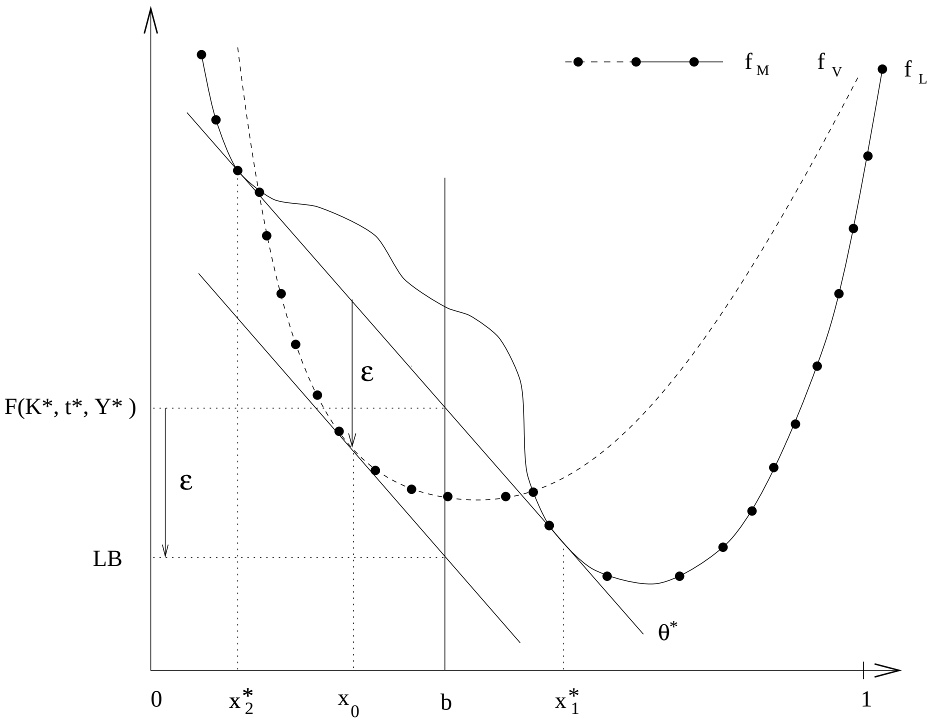

Figure 3. Lower bound yielded by the local optimum of Figure 1