Efficient chemical equilibrium calculations for geochemical speciation and reactive transport modelling

Allan M.M. Leal\textsuperscript{a,}\textsuperscript{*}, Martin J. Blunt\textsuperscript{a}, Tara C. LaForce\textsuperscript{b}

\textsuperscript{a} Imperial College London, QCCSRC and Department of Earth Science and Engineering, London, UK
\textsuperscript{b} CSIRO, Earth Science and Resource Engineering, Australia

Received 12 August 2013; accepted in revised form 28 January 2014; available online 12 February 2014

Abstract

Chemical equilibrium calculations are essential for many environmental problems. It is also a fundamental tool for chemical kinetics and reactive transport modelling, since these applications may require hundreds to billions equilibrium calculations in a single simulation. Therefore, an equilibrium method for such critical applications must be very efficient, robust and accurate. In this work we demonstrate the potential effectiveness of a novel Gibbs energy minimisation algorithm for reactive transport simulations. The algorithm includes strategies to converge from poor initial guesses; capabilities to specify non-linear equilibrium constraints such as pH of an aqueous solution and activity or fugacity of a species; a rigorous phase stability test to determine the unstable phases; and a strategy to boost the convergence speed of the calculations to quadratic rates, requiring only few iterations to converge. We use this equilibrium method to solve geochemical problems relevant to carbon storage in saline aquifers, where aqueous, gaseous and minerals phases are present. The problems are formulated to mimic the ones found in kinetics and transport simulations, where a sequence of equilibrium calculations are performed, each one using the previous solution as the initial guess. The efficiency and convergence rates of the calculations are presented, which require an average of 1–2 iterations. These results indicate that critical applications such as chemical kinetics and reactive transport modelling can potentially benefit by using this multiphase equilibrium algorithm.

1. INTRODUCTION

In a chemical equilibrium state, the forward and reverse rates of the reactions in a system are equal, and therefore no changes in the concentrations of its species are observed with time. It is possible to demonstrate, with the use of the first and second laws of thermodynamics, that a chemical system undergoing an isobaric and isothermal process progress towards a state of minimum Gibbs free energy. Other conditions also apply during this process, such as mass conservation of the chemical elements.

Therefore, an equilibrium problem consists of finding the number of moles of the chemical species that simultaneously minimises the Gibbs free energy of the system and satisfies a system of equilibrium constraints (Smith and Missen, 1982). In addition, a non-negativity constraint for the number of moles is required in order to guarantee a physically feasible molar composition.

The applicability of chemical equilibrium solvers for environmental problems is wide. For instance, speciation modelling of aquatic systems, calculation of solubilities of gases and minerals, analysis of the effect of pH on the dissolution of a mineral, investigation of water-gas-rock effects during carbon storage in geological formations, and radioactive waste disposal modelling are all examples of...
### Nomenclature

#### Greek symbols
- $\zeta_i$: the $i$th species in the chemical system
- $\zeta_i^\pi$: the $i$th species in the $\pi$th phase
- $\dot{\zeta}_r$: the convergence rate of an equilibrium calculation, see Eq. (4.2)
- $\mathbf{z}$: the set of species in the chemical system
- $\mathbf{z}^\pi$: the set of species in the $\pi$th phase
- $\varepsilon$: the set of elements in the chemical system
- $\epsilon_A$: the phase stability tolerance
- $\epsilon_j$: the $j$th element in the chemical system
- $\varphi_{i,i}$: the fugacity coefficient of the gaseous species with index $i$, see Eq. (2.4)
- $\gamma_i$: the activity coefficient of the aqueous species with index $i$, see Eq. (2.3)
- $\Lambda_{\pi}$: the stability index of the $\pi$th phase, see Eq. (3.12)
- $\mu_i$: the constant perturbation parameter used when the watchdog strategy is active
- $\mu_i^\pi$: the perturbation parameter of the interior-point method, see Eq. (3.9)
- $\mu_i^\alpha$: the chemical potential of the $i$th species, see Eq. (2.2)
- $\mu_i^\alpha^\pi$: the standard chemical potential of the $i$th species, see Eq. (2.3)
- $\mu_w$: the threshold used to activate the watchdog strategy
- $\Omega_{\pi}$: the generalised saturation index of the $\pi$th phase, see Eq. (3.12)
- $\Pi$: the number of phases in the chemical system

#### Roman symbols
- $a_{i1}$: the activity of the ionic species $\text{H}^+$, see Eq. (2.13)
- $a_i$: the activity of the $i$th species, see Eq. (2.3)
- $a_i^\ast$: the specified activity of the $i$th species, see Eq. (2.14)
- $b$: the molar abundance vector of the elements, see Eq. (2.5)
- $b_j$: the molar abundance of the $j$th element, see Eq. (2.5)
- $b_j^\ast$: the specified number of moles of the $j$th element, see Eq. (2.8)
- $c$: the equality constraint function in a optimisation problem, see Eq. (3.1)
- $\mathbf{D}$: the diagonal scaling matrix of the primal variable $\mathbf{x}$, see Eq. (3.10)
- $e$: the vector of all ones, see Eq. (3.9).
- $E$: the number of elements in the chemical system
- $F_{\mu}$: the residual function of the perturbed KKT conditions, see Eq. (3.9)
- $f$: the objective function in a optimisation problem, see Eq. (3.1)
- $G$: the Gibbs free energy function of the chemical system, see Eq. (2.1)
- $h$: the vector-valued equilibrium constraint function, see Eq. (2.2)
- $l$: the vector of electrical charges of the aqueous species
- $n$: the molar abundance vector of the species, see Eq. (2.1)
- $N$: the number of species in the chemical system
- $N_{\pi}$: the number of species in the $\pi$th phase
- $n_i$: the number of moles of the $i$th species, see Eq. (2.2)
- $P$: the pressure of the chemical system, see Eq. (2.1)
- $P^\pi$: the reference pressure for the activity of gaseous species, see Eq. (2.4)
- $P_i$: the partial pressure of the $i$th gaseous species, see Eq. (2.16)
- $R$: the universal gas constant, see Eq. (2.3)
- $r_k$: the calculation residual at the $k$th iteration, see Eq. (4.2)
- $T$: the temperature of the chemical system, see Eq. (2.1)
- $w_i$: the $i$th row of the formula matrix $\mathbf{W}$, see Eq. (2.8)
- $W$: the number of watchdog iterations after which the algorithm checks for filter acceptance
- $\mathbf{x}$: the scaled primal variables, see Eq. (3.2)
- $\mathbf{x}$: the primal solution in a previous calculation, see Eq. (3.11)
- $\mathbf{X}$: the diagonal matrix defined by $\mathbf{X} := \text{diag}(\mathbf{x})$, see Eq. (3.6)
- $\mathbf{x}$: the vector of variables in a optimisation problem, see Eq. (3.1)
- $\mathbf{x}^\pi$: the molar fraction of the $i$th species in the $\pi$th phase, see Eq. (3.13)
- $\mathbf{x}^\pi$: the molar fraction of the $i$th species in the $\pi$th phase, see Eq. (2.4)
- $y(y_i, z_i)$: the Lagrange multipliers of a optimisation problem, see Eq. (3.2)
- $y_i$: the vector of electrical charges of the aqueous species

#### Mathematical symbols
- $\nabla$: the gradient operator with respect to the primal variables $\mathbf{x}$ only
- $^\top$: the transpose operator of a matrix or vector
problems that require equilibrium calculations. In addition, it is a fundamental tool for chemical kinetics and subsurface reactive transport modelling. In these applications, some reactions in the system are controlled by thermodynamics instead of kinetics, and consequently equilibrium calculations are necessary (Lichtner, 1985; Steefel and Cappellen, 1990; Steefel and Lasaga, 1994; Steefel and MacQuarrie, 1996; Steefel et al., 2005).

1.1. Existing geochemical equilibrium algorithms

Smith and Misson (1982) classify chemical equilibrium algorithms in two types: stoichiometric and non-stoichiometric. The former solves a system of mass-balance and mass-action equations, while the latter minimises the Gibbs free energy of the system. In Zeleznik and Gordon (1968) and Van Zeggeren and Storey (1970), however, it is shown that both approaches are conceptually equivalent.

The literature contains many geochemical solvers for equilibrium calculations, most of them based on a stoichiometric approach. Examples of stoichiometric based solvers include WATEQ (Truesdell and Jones, 1974), MINEQL (Westall et al., 1976), WATEQF (Ball et al., 1987), MINTEQA2 (Allison and Kevin, 1991), Eq. 3/6 (Wolery et al., 1992), MINEQL+ (Schecher and and McAvoy, 1998), CHESS (der Lee et al., 2002), PHREEQC (Parkhurst and Appelo, 1999, 2013), and The Geochemist’s Workbench (Bethke, 2007).

In general, these solvers rely on equilibrium constants of reactions, which are insufficient to calculate the Gibbs free energy of the geochemical system. As a result, determining the stable equilibrium phase assemblage is a difficult and expensive task. Many heuristic techniques have been developed to resolve this issue for pure mineral phases. For example, Bethke (2007) presents a technique to add the most supersaturated mineral to the calculation and remove the most undersaturated. Although popular among many geochemical solvers, this heuristic is inadequate and unsuitable for aqueous and gaseous phases. Consequently, many of these packages require that the aqueous and gaseous phase always exist at equilibrium, by fixing the amount of liquid water and the fugacity of a gaseous species.

In Leal et al. (2013) we pointed out some issues with the numerical methods frequently used in stoichiometric geochemical solvers. For example, MINTEQA2 (Allison and Kevin, 1991), Eq3/6 (Wolery et al., 1992), CHESS (der Lee et al., 2002), PHREEQC (Parkhurst and Appelo, 1999, 2013), and The Geochemist’s Workbench (Bethke, 2007) adopt an incomplete Newton scheme developed by Morel and Morgan (1972) and further improved by Reed (1982) for solving aqueous speciation. The approach consists of arranging the species in a set of primary species and another of secondary species. The composition of the primary species is calculated by applying Newton’s method to the modified mass-balance equations. The composition of the secondary species, on the other hand, is calculated via a successive substitution approach using the mass-action equations.

This incomplete Newton scheme aims to reduce the dimension of the Jacobian matrix so that less computational effort is spent in solving linear systems. However, combining Newton’s method with a successive substitution approach has the potential of preventing the calculation to converge at quadratic rates near the solution. Moreover, in some chemical equilibrium calculations, the cost of evaluating expensive equations of state and their derivatives could exceed the cost of solving linear systems. See Appendix A for a more detailed description of this stoichiometric formulation, which also presents an in-depth discussion on its convergence rates.

Geochemical packages based on Gibbs energy minimisation algorithms include ChemSage (Eriksson and Hack, 1990), THERIAC (de Capitani and Brown, 1987), HCh (Shvarov, 1999, 2008), FactSage (Bale et al., 2002, 2009), and GEM-Selektor (Karpov et al., 1997, 2001; Karpov, 2002; Kulik et al., 2004, 2013).

In ChemSage (Eriksson and Hack, 1990), a Gibbs energy minimisation algorithm is implemented in which the non-ideal values of the chemical potentials are used, but their derivatives correspond to ideal models. Although convergence might not be an issue with such practice (e.g., quasi-Newton methods converge with artificial derivatives that are considerably different from their real values, Nocedal and Wright, 1999; Fletcher, 2000), the approach cannot converge at a quadratic rate near the solution. In Eriksson and Hack (1990) it is pointed out that a strategy was developed to refine the initial guess to improve the efficiency of the algorithm. However, this improvement does not change the convergence rate of the method.

In GEM-Selektor (Karpov et al., 1997, 2001; Karpov, 2002; Kulik et al., 2004, 2013), an interior-point algorithm is used to minimise the Gibbs free energy of multiphase systems. Their method, however, does not use the logarithmic barrier functions of Fiacco and McCormick (1990) nor the KKT perturbation approach of El-Bakry et al. (1996), which are practices commonly adopted in several non-linear programming packages and algorithms (Ulbrich et al., 2004; Silva et al., 2008; Wächter and Biegler, 2005a,b,c; Vanderbei, 1999; Benson et al., 2000; Vanderbei, 2006; Byrd et al., 1999; Costa and Fernandes, 2008, 2011a,b). GEM-Selektor assumes that at every iteration the minimisation of the Gibbs energy is a convex problem, even though non-ideal phases are assumed in their calculation. As a consequence, the algorithm is limited to linear constraints only, such as mass-balance and charge-balance, otherwise the convexity assumption would not be possible. At every iteration, a convex minimisation algorithm is used to find a feasible descent direction that will be used to find a step length that sufficiently minimises the Gibbs free energy, a procedure similar to a line-search strategy. The algorithm not only maintains feasibility on the bounds at every iteration, but also on the mass-balance constraints. Our approach, in contrast, considers possible non-linear equilibrium constraints, such as the imposition of pH, activities and fugacities. Therefore, these more complex constraints make it impractical to maintain full feasibility at every iteration, though the converged solution satisfies all equality and inequality constraints.

Harvie et al. (1987) presents an equilibrium algorithm for non-ideal multiphase systems based on the minimisation of the Gibbs free energy. Their algorithm consists of
transforming a constrained minimisation problem into an unconstrained one by introducing Lagrange multipliers and quadratic slack variables to circumvent the bound constraints. Moreover, their approach has some similarity to projected gradient active-set minimisation methods, since their calculated directions are always feasible and some species are sometimes made active (i.e., having zero number of moles) during the search for the stable phase assemblage. However, their method presents some strategies for finding the global minimum, an approach we do not attempt in this work, since its computational cost would make our proposed method prohibitive for reactive transport simulations.

In the recent work of Harvey et al. (2013), a Gibbs free energy minimisation method for systems composed of solid solutions was developed. Their method uses exact Hessian expressions, which potentially produce fast convergence rates near the solution. They compared their method with the general-purpose optimisation packages IPOPT (Wächter and Biegler, 2005c), KNITRO (Byrd et al., 1999), and SNOPT (Gill et al., 2002, 2005), identifying the necessity of suitable heuristics to improve the efficiency and robustness of chemical equilibrium calculations. Moreover, they observed that using quasi-Newton approximations for the Hessian matrix, an approach adopted by the general optimisation solver SNOPT, results in severely inefficient chemical equilibrium calculations. In their algorithm, all phases and species are initially assumed to exist at equilibrium, allowing a simplification of the KKT equations by eliminating the complementary equations. However, some phases can be excluded during the calculation if they are detected as unstable. The authors define an unstable phase as a phase that for a certain number of past iterations a full Newton step would bring its molar abundance to a negative value. These phases are then removed because eliminating the complementary conditions from the KKT equations is subjected to all species having non-zero number of moles. In the end of the calculation, phase stability tests are performed to identify any excluded phase that should be added to the system, or any phase at equilibrium that should be removed. Our algorithm, on the other hand, does not exclude or add any phase in the course of the calculation to attain numerical stability. All phases are assumed in the calculation, and the unstable phases are identified in the end of the calculation with a rigorous phase stability test.

1.2. Integration into reactive transport simulators

As discussed by Kulik et al. (2013), a chemical equilibrium solver to be integrated into a reactive transport simulator must be extremely efficient and accurate. Efficiency is paramount in these large scale simulations, since equilibrium calculations must be performed for every grid block of the mesh, at every time step. Accuracy is equally important, since the mass-balance residuals of the calculations can accumulate over time and propagate throughout the reservoir. These accuracy problems can then culminate in an unstable state of the whole transport simulation. Another essential attribute of an equilibrium solver is to accurately and robustly determine the stable phases of a multiphase system. Otherwise, the fronts of the flow, where phases are constantly appearing and disappearing, can be highly inaccurate.

Due to these demanding requirements on equilibrium solvers for reactive transport simulations, Kulik et al. (2013) has recently reviewed their own algorithm. Their new GEMS3K code is an improvement of the previous GEMIPM2K, whose accuracy and stability were not sufficient for transport modelling (Shao et al., 2009). Therefore, this necessity of Kulik et al. (2013) in adapting their numerical method for critical geochemical applications raises the question of whether other approaches in the geochemical literature require some adaptations as well. Since many geochemical equilibrium methods were initially developed with the primary intent for speciation-solubility calculations and plotting phase diagrams, this might indeed be the case.

In principle it may not be very clear how a dedicated equilibrium solver, specially one based on the minimisation of Gibbs free energy, could be integrated into chemical kinetics or reactive transport calculators. This is because the common approach is to directly incorporate the mass-action equations into the system of differential equations, and solve both transport and chemical processes simultaneously. Examples of simulators that adopt this approach are TOUGHREACT (Xu et al., 2004, 2006), PFLOTRAN (Lu and and Lichtner, 2005), and CrunchFlow (Steefel et al., 2009). This approach is also adopted in the works of Saalhink et al. (1998) and Bea et al. (2009).

In Leal et al. (2014b), however, we develop an algorithm for chemical kinetics that uses our Gibbs energy minimisation method. The mathematical formulation of the problem, nonetheless, is derived in a way that any equilibrium solver could be used, which can only be achieved if the equilibrium conditions in the kinetics problem are handled abstractly. For example, without direct substitution of mass-action equations in the kinetics equations. The method is based on an implicit formulation to guarantee higher efficiency and stability in the integration of the differential equations. The quantities that evolve with time are the natural molar amounts of chemical elements, and not the usual total concentration of primary species (Steefel and Cappellen, 1990). The presented systematic approach for kinetics could be straightforwardly applied into the context of reactive transport equations, either using an implicit or an operator splitting approach. For the implicit approach, the partial derivatives of the molar abundance of the species with respect to the molar abundance of the elements are necessary, and its calculation is also shown in Leal et al. (2014b).

1.3. Proposal of a novel geochemical equilibrium algorithm

In this work we propose the solution of the Gibbs energy minimisation problem by casting it into a non-linear programming problem. The advantage of this approach is that it becomes easier to identify the relevant aspects of the problem and to decide which minimisation method, among those existent in the mathematical literature, best solves it in terms of efficiency, accuracy and robustness. Moreover, this approach eliminates the dependence of specific chemical details of the problem on its numerical solution, simplifying its analysis and implementation.
Our numerical methodology for multiphase equilibrium calculations is based on the trust-region primal-dual interior-point algorithm of Ulbrich et al. (2004) and Silva et al. (2008). Their algorithm solves non-linear programming problems with non-convex objective functions containing both equality and inequality constraints. We briefly describe their method here for completeness, together with our modifications tailored to yield more efficient and robust chemical equilibrium calculations. The interested reader, however, is referred to Leal et al. (2014a) for the complete presentation of the method.

This chemical equilibrium method has been implemented in REAKTORO, a scientific library written in the C++ programming language for computational geochemical modelling. Currently, it provides methods for both multiphase chemical equilibrium and kinetics calculations. The code is freely available at http://bitbucket.org/reaktoro, where its licensing information can be found.

2. CHEMICAL EQUILIBRIUM

In this section we present the mathematical formulation for chemical equilibrium problems, where the system is composed of multiple phases and species. The equilibrium problem formulation is done generically, with the chemical system not being restricted in any aspect.

In what follows, we shall assume a chemical system composed of $N$ species, where the $i$th species is denoted by $x_i$, and the set of species by $x = \{x_1, \ldots, x_N\}$. In addition, we consider that there exist $E$ elements from which these species can be formed, where the $j$th element is denoted by $e_j$, and the set of elements by $e = \{e_1, \ldots, e_E\}$. Finally, we assume that the chemical species are partitioned among $\Pi$ phases, where $x^p$ denotes the $i$th species in the $p$th phase, $x^p = \{x^p_1, \ldots, x^p_{N_p}\}$ denotes the set of species in the $p$th phase, $N_p$ denotes the number of species in the $p$th phase, and $T_p$ denotes the set of indices of the species in the $p$th phase.

2.1. Equilibrium problem

Finding the equilibrium point of a multiphase system is a non-linear optimisation problem, where the Gibbs free energy of the system is minimised. The problem, however, is far from being trivial, since it contains both equality and inequality constraints, defining, therefore, a non-linear programming problem.

The equality constraints arise from the need to specify, for example, the number of moles of each chemical element in the system, the pH and the charge balance condition of an aqueous solution, the partial pressure of a gaseous species, and so forth. The inequality constraints, on the other hand, results from the physical condition that the number of moles of the species are non-negative.

From the principle of minimum Gibbs free energy, we calculate the equilibrium state of a chemical system by solving the following constrained minimisation problem:

$$\min_{\mathbf{n}} G(\mathbf{n}; T, P) \text{ subject to } \begin{cases} \mathbf{h}(\mathbf{n}) = 0 \\ \mathbf{n} \geq 0 \end{cases},$$

where $\mathbf{n} \in \mathbb{R}^N$ is the molar composition vector of the system; $T$ and $P$ are, respectively, the temperature and pressure of the system; $\mathbf{h} : \mathbb{R}^N \mapsto \mathbb{R}^M$ is the equilibrium constraint function; $N$ is the number of species; and $M$ is the number of equilibrium constraints. The temperature $T$ and pressure $P$ are assumed given parameters.

The Gibbs energy function $G : \mathbb{R}^{2+1} \mapsto \mathbb{R}$ is defined by:

$$G(T, P, \mathbf{n}) := \sum_{i=1}^{N} n_i \mu_i(T, P, \mathbf{n}),$$

where $n_i$ is the number of moles of the $i$th species; and $\mu_i : \mathbb{R}^{2+1} \mapsto \mathbb{R}$ is the chemical potential function of the $i$th species, given by:

$$\mu_i(T, P, \mathbf{n}) = \mu_i^0(T, P) + RT \ln a_i(T, P, \mathbf{n}).$$

In the previous equation, $\mu_i^0 : \mathbb{R} \mapsto \mathbb{R}$ is the standard chemical potential function of the $i$th species; $a_i : \mathbb{R}^{2+1} \mapsto \mathbb{R}$ is the activity function of the $i$th species; and $R$ is the universal gas constant.

In the literature, the activity of a species is commonly replaced by the product of a concentration quantity and an activity coefficient. For example, the activity of an aqueous species is replaced by:

$$a_i = m_i \gamma_i,$$

where $m_i$ and $\gamma_i$ are its molality and activity coefficient. The activity of a gaseous species, on the other hand, is replaced by:

$$a_i = \varphi_i x_i^p P,$$

where $x_i^p$ and $\varphi_i$ are its molar fraction and fugacity coefficient, and $P$ is a reference pressure, usually equal to 1 bar, though this depends on the standard chemical potential data of the gaseous species.

In our formulation, however, the term $a_i$ representing the activity of a species is included in all equations. The idea is to avoid the explicit dependence of the type of the species every time an equation is written. For example, the Gibbs energy function (2.2) can be written succinctly and generically without the need to write terms that corresponds to the aqueous, gaseous, mineral, and possibly other types of species. This is a fundamental step for the development of a general chemical equilibrium algorithm that can be applied to any type of systems. At the computational level, such abstraction is also possible to be made with programming techniques such as polymorphism.

In general, the standard chemical potential $\mu_i^0$ of a species can be interpolated from temperature versus pressure tables. This is an efficient and sufficiently accurate approach, justified by the fact that the evaluation of their equations of state can be prohibitively expensive. However, the same practice cannot be done with the activity $a_i$ of a species, because this function depends on the composition of the system. Therefore, they must be directly computed from their equations of state. Moreover, fast convergence of the Gibbs energy minimisation problem require at least their first-order partial molar derivatives.

More information about the partial molar derivatives of the species activities, and the calculation of the gradient and
2.2. Equilibrium constraints

Below is a list of quantities that can be constrained at equilibrium:

- number of moles of an element;
- charge-balance of the aqueous solution;
- activity of a species; and
- partial pressure of a gaseous species.

The Gibbs energy minimisation problem, as formulated in Eq. (2.1), already provides the means to enforce these constraints. This can be done by using the equilibrium constraint function \( h \), where each component of this vector function imposes an equilibrium condition.

Observe that our approach for specifying equilibrium constraints is general and flexible. A common constraint used in equilibrium calculations is the mass-balance conservation of elements, which are given by the linear equation:

\[
W n = b,
\]

where \( W \) denotes the \( E \times N \) formula matrix of the system, with its \((j,i)\)th entry denoting the number of atoms of the \( j \)th element in the \( i \)th species; \( E \) the number of elements; and \( b \) the molar abundance vector of the elements, whose component \( b_j \) denotes the number of moles of the \( j \)th element in the system. If the mass-balance constraint is to be specified, then the equilibrium constraint function \( h \) can be defined as:

\[
h(n) := W n - b.
\]

and its gradient as:

\[
\nabla h(n) := W.
\]

Let us now show how some other equilibrium constraints can be constructed individually.

2.2.1. Imposition of number of moles of an element

There are equilibrium problems in which the number of moles of some elements are unknown. Instead of specifying these amounts, they are in fact calculated, which constitutes an inverse problem (Kulik, 2006; Kulik et al., 2013).

As an example, consider the case where one needs to find out the amount of hydrochloric acid that needs to be mixed with 1 kg of water to produce a solution with pH equal to 4.0. In this problem, therefore, the number of moles of elements H and Cl are unknown a priori. However, these amounts can be determined by solving an equilibrium problem where (i) the pH of the solution is imposed together with (ii) the condition of charge-balance of the aqueous solution and (iii) the known amount of element oxygen. Thus, we see that the capability of specifying the number of moles of individual elements is of fundamental importance for chemical equilibrium modelling.

Imposing the number of moles of the \( j \)th element is governed by the following equilibrium constraint function:

\[
h(n) := w^T n - b^*_j,
\]

where \( b^*_j \) denotes the desired number of moles of the \( j \)th element; and \( w_j \in \mathbb{R}^N \) the vector formed from the \( j \)th row of the formula matrix \( W \). The gradient of this function is given by:

\[
\nabla h(n) := w_j.
\]

2.2.2. Imposition of charge-balance

As we just saw, the charge-balance condition is a convenient constraint to be imposed when the amount of a chemical element is unknown. This constraint can be even imposed when the amounts of all elements are known, so that the obtained solution certainly does not violate this physical condition. For example, Kulik et al. (2013) always impose the charge-balance condition, since they use the Lagrange multiplier corresponding to this linear constraint to compute redox potential quantities such as \( \text{pE} \) and \( \text{Eh} \), as demonstrated in Kulik (2006).

By accounting for the balance of charges in an aqueous solution, the following condition must be attained:

\[
z^T n = 0,
\]

where \( z \in \mathbb{R}^N \) is the vector of charges of all species, whose entries are non-zero for the charged aqueous species only. Therefore, the equilibrium constraint function that imposes the charge-balance condition is given by:

\[
h(n) := z^T n,
\]

whose gradient is:

\[
\nabla h(n) = z_n.
\]

2.2.3. Imposition of activity of a species

Imposing the activity of a species can be useful in some cases. For example, the pH of an aqueous solution can be easily measured, and so the activity of species \( H^+ \) is known, given by:

\[
a_{H^+} = \exp(a_0(-\text{pH})).
\]

Therefore, this information can be used towards the calculation of the equilibrium state of the solution.

Imposing the activity of the \( i \)th species requires defining the equilibrium constraint as:

\[
h(n) := a_i(n) - a^*_i,
\]

where \( a^*_i \) is the desired activity for the \( i \)th species. It follows that the gradient of the above function is given by:

\[
\nabla h(n) = \nabla a_i(n)
\]

where \( \nabla a_i(n) \) is the gradient of the activity function \( a_i \). The analytical calculation of these derivatives results in efficient equilibrium calculations, and they are of utmost importance in the minimisation of the Gibbs free energy. For simplicity, note that we have omitted the dependence on temperature \( T \) and pressure \( P \) of the activity function \( a_i \) and its gradient \( \nabla a_i \) in the previous formulation.
2.2.4. Imposition of partial pressure of a gaseous species

Imposing the partial pressure of a gas is a fairly common practice. In Bethke (2007), the equilibrium problem of the dissolution of pyrite (FeS₂) is considered. In this problem, the fugacity of the gaseous species O₂(g) is kept constant in order to simulate the contact of the solution with the atmosphere. Sometimes, however, one might opt to impose the partial pressure of a gas instead of its fugacity, which are equivalent practices when the gases are considered ideal. Thus, for instance, since oxygen makes up about 20% of the atmosphere, one needs to set the partial pressure of O₂ to 0.2 atm, assuming the atmospheric pressure is 1 atm.

The partial pressure of the ith gaseous species is given by:

\[ P_i = x_i^g P \]  
(2.16)

Thus, the equilibrium constraint function can be defined as:

\[ h(n) := P_i - x_i^g P, \]  
(2.17)

and its gradient as:

\[ \nabla h(n) := -\nabla x_i^g P, \]  
(2.18)

where the ith component of the vector \( \nabla x_i^g \), denoted by \( \partial x_i^g / \partial n_i \), should be zero if the ith species does not belong to the gaseous phase. Note that \( \nabla x_i^g \) and \( \nabla x_i^g \) can be easily calculated from \( n \) using the indices of the gaseous species \( T_g \).

3. NUMERICAL METHOD

In this section we briefly describe the developed numerical method for equilibrium calculations of multiphase systems. Its detailed description can be found in Leal et al. (2014a), which adopts an approach of separating pure mathematical concepts from those specific to chemical equilibrium that are irrelevant at the algorithmic level. As mentioned before, the method is based on an adaptation of the interior-point minimisation algorithm of Ulbrich et al. (2004) and Silva et al. (2008).

Let us represent the Gibbs energy minimisation problem (2.1) as a general non-linear programming problem in the standard form:

\[
\min_{x} f(x) \quad \text{subject to} \quad \begin{cases} 
  c(x) = 0 \\
  x \geq 0
\end{cases}, \tag{3.1}
\]

where the objective function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) and the equality constraint function \( c : \mathbb{R}^n \rightarrow \mathbb{R}^m \) are assumed to be twice continuously differentiable. Moreover, let us assume that \( m \leq n \), where \( n \) denotes the number of variables and \( m \) the number of constraints. Finally, we denote the vector of variables in the optimisation problem (3.1) by \( x \in \mathbb{R}^n \), which corresponds to the molar amounts of the chemical species in the context of the Gibbs free energy minimisation.

3.1. First-order optimality conditions

Let us write the necessary first-order optimality conditions for the non-linear programming problem (3.1). These are also known as the Karush–Kuhn–Tucker or KKT conditions (see Nocedal and Wright, 1999), and depends on the following definition of the Lagrange function:

\[
\mathcal{L}(x, y, z) := f(x) + c(x)^T y - x^T z, \tag{3.2}
\]

where \( y \in \mathbb{R}^m \) and \( z \in \mathbb{R}^m \) are Lagrange multipliers. The KKT conditions, and their corresponding names, are then written as:

\[
\nabla_x \mathcal{L}(x, y, z) = 0, \quad \text{optimality} \tag{3.3}
\]

\[
\mathcal{e}(x) = 0, \quad \text{primal—feasibility} \tag{3.4}
\]

\[
xz = 0, \quad \text{complementarity} \tag{3.5}
\]

\[
x, z \geq 0, \quad \text{dual—feasibility} \tag{3.6}
\]

where \( X := \text{diag}(x) \) and \( \nabla_x \mathcal{L} \) denotes the gradient of the Lagrange function with respect to the primal variables \( x \), given by:

\[
\nabla_x \mathcal{L}(x, y, z) = \nabla f(x) + \nabla c(x)^T y - z. \tag{3.7}
\]

Eqs. (3.3)–(3.6) are the requirements that a local solution must satisfy. Note that besides the primal variables \( x \), the Lagrange multipliers \( y \) and \( z \) are also unknowns in the problem. Thus, there are a total of \( 2n + m \) unknowns, which corresponds to two times the number of chemical species plus the number of equilibrium constraints, \( 2N + M \). As discussed in Kulik et al. (2013), it is imperative that the size of the linear systems to be solved are moderate in order to efficiently solve equilibrium problems. Fortunately, the Lagrange multipliers \( z \) can be explicitly written in terms of \( x \) and \( y \), which decreases the size of the linear systems by \( n \), as shown in Leal et al. (2014a).

3.2. Perturbed KKT conditions

Similarly to El-Bakry et al. (1996), the primal-dual interior-point algorithm of Ulbrich et al. (2004) solves the problem (3.1) by suitably perturbing the KKT complementary condition (3.5). The approach results in the following perturbed system of non-linear equations:

\[
F_{\hat{\mu}}(x, y, z) = 0, \tag{3.8}
\]

where \( F_{\hat{\mu}} \) is defined by:

\[
F_{\hat{\mu}}(x, y, z) := \begin{bmatrix} 
\nabla_x \mathcal{L}(x, y, z) \\
\mathcal{e}(x) \\
xz - \hat{\mu} e 
\end{bmatrix}, \tag{3.9}
\]

with \( \hat{\mu} > 0 \) denoting a small perturbation parameter; and \( e \in \mathbb{R}^m \) the vector of all ones. More details on the perturbation parameter \( \hat{\mu} \), and its progress towards zero can be found in Leal et al. (2014a).

Eq. (3.8) is solved using Newton’s method. However, several strategies are adopted to aid convergence from arbitrary or poor initial guesses, since Newton’s method is not guaranteed to converge far from a local solution (Nocedal and Wright, 1999). Other strategies are also adopted to guarantee that the solution is a local minimum, and not a maximum for example.

3.3. Convergence strategies

In this section we list a few convergence strategies adopted in the algorithm to aid convergence from poor initial guesses. A more in-depth description of these strategies
can be found in Ulbrich et al. (2004) and Silva et al. (2008), and their modification for chemical equilibrium calculations is shown in Leal et al. (2014a).

3.3.1. Filter

The filter technique developed by Fletcher and Leyffer (2002) is used. The idea has its origins in multi-criteria optimisation problems, whose adaptation for non-linear programming problems is possible by considering the minimisation of the optimality and feasibility measures as two competing targets. These measures are respectively related to the norm of the optimality and feasibility conditions (3.3) and (3.4).

The filter works as follows. At every iteration, a decision is made about storing or not the current optimality and feasibility measures. These records are then used in subsequent iterations to decide the rejection of poor iterates. A poor iterate is defined as an iterate whose corresponding optimality and feasibility measures are not sufficiently smaller than those stored in the filter. Therefore, the filter guarantees that at every iteration the current iterate is closer to a local optimum solution.

3.3.2. Trust-region

Newton’s method is well known to frequently behave erratically far from the solution. Thus, convergence strategies exist in which the length of the step taken towards a new iterate is somehow controlled. Differently from a line-search strategy, in which the size of Newton’s step is controlled and its direction preserved, a trust-region strategy controls both its size and direction. This is done with the introduction of a decomposition of Newton’s step into a tangential and normal components, whose lengths are restricted so that they fit inside a variable trust-region.

The benefit of this decomposition is the possibility to control how much the iterations should favour the decrease of either the feasibility measure or the optimality measure. By controlling the size of the tangential step, iterates that satisfy more the feasibility condition (3.4) is obtained. For example, the smaller the feasibility measure, the smaller the mass-balance residuals in the minimisation of Gibbs free energy. On the other hand, by controlling the size of the normal step, iterates that conform more to the optimal condition (3.3) are obtained. In addition, the lengths of these components can be appropriately adjusted to enforce positive number of moles at every iteration.

3.3.3. Restoration phase

Whenever the algorithm fails to find an iterate that does not satisfy the filter conditions, some action other than halt its execution must be done. When this happens, the algorithm enters in a restoration phase, in which it focuses in the decrease of an alternative feasibility measure. During the iterations in the restoration phase, some progress towards improved optimality measures is also obtained, though the main result after the end of the restoration phase is in general the decrease of the feasibility measure (e.g., the mass-balance residuals). Once an iterate in the restoration phase satisfies the filter conditions, the normal trust-region algorithm takes back control of the calculation.

3.4. Scaling

As discussed in Nocedal and Wright (1999), it is crucial that proper scaling of the variables is done to improve the performance of the algorithm. Indeed, as we shall see later, the primal-dual interior-point algorithm works more efficiently for sequential optimisation calculations if proper scaling of the variables is performed.

Let us denote $D$ as the diagonal scaling matrix of the primal variable $x$. Then, we define the scaled primal variable $\bar{x}$ as:

$$
\bar{x} = D^{-1}x.
$$

The implication of this scaling on the objective and constraint function, as well as on their derivatives, is detailed in Leal et al. (2014a).

When performing sequential equilibrium calculations, the previous solution serves as an excellent option for scaling the primal variables for the next calculation. Therefore, denoting by $\bar{x}$ the primal solution in the previous calculation, we set:

$$
D = \text{diag}(\bar{x}).
$$

Note, however, that for a standalone calculation, or at the beginning of a sequence of optimisation calculations, a good scaling is not available. This follows from the fact that only a poor initial guess can be provided at that time. Therefore, we always perform the first optimisation calculation without scaling.

The performance boost of the calculation with scaling, as we shall see later, is in part explained by the shift of the scaled variables to the interior of the feasible domain. This is a beneficial condition for interior-point algorithms, which can behave erratically near the boundaries of the feasible domain. Moreover, this solves the issue of some variables having different orders of magnitude, which is very common in chemical equilibrium problems.

3.5. Watchdog strategy

The minimisation algorithm adopts a monotone strategy. In this strategy, every accepted iterate provides a sufficient decrease in either the optimality or feasibility measures. Its main advantage is to aid convergence from arbitrary initial guesses, since an uncontrolled scheme of the step lengths could result in huge increases of these measures.

Nevertheless, a well-known disadvantage of monotone strategies is that it potentially reject iterates that would make good progress towards the solution (Nocedal and Wright, 1999). This is known as the Maratos effect, when the algorithm continuously discards good iterates because they either increase the optimality or feasibility measures.

As a result, convergence to a solution becomes very slow.

The Maratos effect can be circumvented with the use of the watchdog technique of Chamberlain et al. (1982). Our adapted approach consists of applying plain Newton’s method to the relaxed KKT Eq. (3.8) with a constant perturbation parameter $\mu$, rather than the variable $\mu$. In addition, normal and tangential components for the
trust-region approach are not computed, nor is the new iterate supposed to be acceptable by the filter. Therefore, this technique is considered a nonmonotone strategy.

After a sufficient progress has been made in the calculation, our algorithm switches to the watchdog mode. This is done in order to speed up convergence towards the optimum solution. The calculation is said to have progressed sufficiently if \( \mu < \mu_s \), where \( \mu_s \) is the watchdog threshold, whose default value is \( \mu_s = 10^{-1} \). Note, however, that the algorithm leaves a monotone strategy to a nonmonotone one, opening up the possibility of divergence. Hence, it is important to monitor the calculation to detect when no progress is being made towards the solution.

Our monitoring approach consists of checking after \( W \) iterations if the current iterate is acceptable by the filter. If so, then the filter is extended with the current feasibility and optimality measures and a new round of \( W \) iterations under the watchdog strategy is allowed. Otherwise, the algorithm reverts to the monotone trust-region strategy, starting from the last watchdog iterate acceptable by the filter. In addition, we monitor the optimality measure at every watchdog iteration. The idea is to prevent an uncontrolled increase of that measure.

### 3.6. Phase stability test

Once an equilibrium calculation has been performed, it is fundamental to determine the stability of the phases in the chemical system. The presented interior-point minimisation algorithm alone produce a rough estimate of which phases are unstable, since these phases have small number of moles with respect to the total in the system. However, this simplistic phase stability test is not very accurate and helpful, since the small number of moles of the unstable phases are very sensitive to the perturbation parameter \( \mu \) of the interior-point method. In addition, it does not indicate how far from equilibrium the unstable phases are.

Therefore, we adopt a phase stability test similar to the one presented by Kulik et al. (2013), which has been successfully applied in the geochemical package GEM-Selektor. In this test, stability indices for all phases are readily calculated from the Lagrange multipliers of the minimisation calculation. The advantage of this approach over the analysis of the relative number of moles of the phases is that these indices are less sensitive to algorithmic parameters, providing a more accurate indication of which phases are unstable. In addition, it also provides a quantitative measure of how far they are from equilibrium.

In Kulik et al. (2013), the stability index \( \Lambda_p \) of the \( p \)th phase is defined as:

\[
\Lambda_p := \log_{10} \Omega_p,
\]

where \( \Omega_p \) is same phase. Differently from Kulik et al. (2013), however, we define \( \Omega_p \) as:

\[
\Omega_p := \sum_{s \in \mathcal{I}_p} x_{p,s} \exp \left( -\frac{z_s}{RT} \right),
\]

where \( \mathcal{I}_p \) is the set of indices of the species in the \( p \)th phase; \( r^* \) is the local index of the \( s \)th species in its phase; \( x_{p,s} \) is the molar fraction of the \( s \)th species; and \( z_s \) is the \( s \)th component of the vector of Lagrange multipliers \( z \). Note that the minimisation algorithm of Kulik et al. (2013) does not introduce the additional Lagrange multipliers \( z \), and so the calculation of \( \Omega_p \) uses a different, but equivalent, equation.

After the calculation of the phase stability indices, the unstable phases can be identified. For every stability index \( \Lambda_p \) such that \( |\Lambda_p| \geq \epsilon_s \), the \( p \)th phase is classified as unstable, where \( \epsilon_s \) is the phase stability tolerance, whose default value is \( \epsilon_s = 0.01 \). Then, their molar composition is zeroed and the mass-balance of the stable phases is corrected. This procedure ensures that the equilibrium solution does not violate the Gibbs phase rule.

To perform this mass-balance correction procedure, new equilibrium calculations are required. However, at this time, only the stable phases need to be considered. Observe that this correction approach is in general very efficient, because (i) a good initial guess is already known, and (ii) the size of the chemical system has been reduced. Therefore, using a quadratic convergent equilibrium method should require one or two iterations in general.

We have opted to use the equilibrium method based on the law of mass-action presented in Leal et al. (2013) to perform these mass-balance corrections. This is because the linear systems solved in that method have dimension \( N \), instead of \( N + M \) in the interior-point algorithm. In addition, it does not involve any Lagrange multiplier, and so the state of the multipliers \( y \) and \( z \) of the interior-point method is not artificially altered, which could compromise subsequent equilibrium calculations that require these as initial guesses.

### 4. RESULTS AND DISCUSSION

In this section we demonstrate the capability of our Gibbs energy minimisation algorithm to perform efficient chemical equilibrium calculations. Our objective is to show that (i) the algorithm is capable of minimising the Gibbs energy of multiphase systems, (ii) only a few iterations is required to obtain convergence when carrying out sequential chemical equilibrium calculations, and (iii) the calculations converge at quadratic rates near the solution.

The following activity and fugacity coefficient models were used to perform the calculations for the study of carbon dioxide injection in deep saline aquifers:

- HKF extended Debye–Hückel activity coefficient model for solvent water and ionic species, Helgeson and Kirkham (1974a,b, 1976); Helgeson et al. (1981);
- the Setschenow activity coefficient model for neutral aqueous species other than CO\(_2\)(aq); the activity coefficient models of Drummond (1981) or Duan and Sun (2003) for CO\(_2\)(aq);
- the fugacity coefficient models of Spycher and Reed (1988); Spycher et al. (2003) or Duan et al. (2006) for CO\(_2\)(g) and H\(_2\)O(g).

Similarly to Leal et al. (2013), the ideal fugacity coefficient model for H\(_2\)O(g) is used in conjunction with the model of Duan et al. (2006) for CO\(_2\)(g). As shown there, accurate results of CO\(_2\) solubility in brines can be obtained with this simplification. This choice is an alternative to the
use of the expensive equation of state of Duan et al. (1992) for H\textsubscript{2}O(g).

The chemical potential of the species were obtained using the equations of state of Helgeson and Kirkham (1974a); Helgeson et al. (1978); Tanger and Helgeson (1988); Shock and Helgeson (1988) and Shock et al. (1992). The database file slop98.dat from the software SUPCRT92 (Johnson et al., 1992) was used to obtain the parameters for the equations of state. The equation of state of Wagner and Pruss (2002) was used to calculate the density of water and its temperature and pressure derivatives.

4.1. Solubility of CO\textsubscript{2} in NaCl brines

Modelling carbon storage in saline aquifers requires accurate calculations of the solubility of CO\textsubscript{2} in NaCl brines. This is necessary to model its fate after injection and to estimate how much of it could be trapped by solubility mechanisms. In addition, the accuracy of these calculated solubilities have a direct impact on the correctness of the geochemical modelling of the reactions between fluid and rock. This is because these reactions are strongly promoted by the acidification of brine that occurs with the dissolution of CO\textsubscript{2}.

The multiphase system used to model the H\textsubscript{2}O-CO\textsubscript{2}-NaCl system is described in Table 1. The solubility calculations were performed by assuming a mixture containing 1 kg of H\textsubscript{2}O, 10 mol of CO\textsubscript{2}, and 2.5 or 4.0 mol of NaCl. The amount of CO\textsubscript{2} was enough to saturate the aqueous solution and form a gaseous phase at all considered temperatures, pressures and salinities. This is fundamentally necessary since our goal is to calculate the saturated solubility of that gas in brine. Once the amounts of H\textsubscript{2}O, CO\textsubscript{2} and NaCl are specified, one can easily calculate the number of moles of the elements H, O, C, Na, and Cl and perform a chemical equilibrium calculation. From the obtained solution, the solubility of CO\textsubscript{2} is given by the molality of element C in the aqueous phase, which accounts for all aqueous species containing carbon.

Fig. 1 compares our calculated CO\textsubscript{2} solubilities with the experimental measurements of Hou et al. (2013). The calculations were performed at temperatures 323.15, 373.15, and 423.15 K, pressures 20–200 bar, and salinities 2.5 and 4.0 molal. By using the activity coefficient model of Drummond (1981) for CO\textsubscript{2}(aq) and the fugacity coefficient model of Spycher et al. (2003) for CO\textsubscript{2}(g) and H\textsubscript{2}O(g), we obtained solubility results with percentage deviation errors smaller than 5%.

Observe in Fig. 1 the salting out effect on the solubility of CO\textsubscript{2} at different brine salinities. This result shows that the amount of carbon trapped in saline aquifers by solubility mechanisms decreases with the increase in brine salinity. The increase in pressure, on the other hand, causes an increase in the amount of dissolved CO\textsubscript{2}, while the increase in temperature has the opposite effect. Note that the results indicate that our choice of activity and fugacity coefficient models can accurately represent the salting out effect as well.

| Aqueous phase | Gaseous phase |
|---------------|---------------|
| H\textsubscript{2}O(l) | CO\textsubscript{2}(g) |
| H\textsuperscript{+} | H\textsubscript{2}O(g) |
| OH\textsuperscript{-} | |
| Na\textsuperscript{+} | Cl\textsuperscript{-} |
| HCO\textsubscript{3}\textsuperscript{-} | |
| CO\textsubscript{2}(aq) | |

Fig. 1. Comparison of our calculations of CO\textsubscript{2} solubility in 2.5 and 4.0 molal NaCl brine (lines), using the activity coefficient model of Drummond, 1981 for CO\textsubscript{2}(aq) and the fugacity coefficient model of Spycher et al., 2003 for CO\textsubscript{2}(g) and H\textsubscript{2}O(g), with the experimental solubility data of Hou et al., 2013 (points). The calculations assumed a CO\textsubscript{2}-H\textsubscript{2}O-NaCl system composed of an aqueous and gaseous phase.
as compute the solubility of carbon dioxide at high temperatures and pressures.

### 4.2. Sequential chemical equilibrium calculations

As discussed in Section 1, efficient sequential equilibrium calculations are fundamental for some critical applications, such as reactive transport modelling. In this section we show the performance results of our Gibbs energy minimisation algorithm, and we analyse the importance of the watchdog strategy and scaling of the variables on the efficiency of the equilibrium calculations.

Consider the multiphase system in Table 2. This system is of interest for modelling water-gas-rock interactions in CO₂ storage in deep saline aquifers, where the reservoir is mainly composed of carbonate minerals. The formation rock is modelled by the mineral phases calcite, magnesite and dolomite. In order to model a possible precipitation of sodium chloride when the aqueous phase becomes salt saturated, the mineral phase halite is also assumed.

To assess the efficiency of our equilibrium calculations in a carbon injection simulation, we start from a chemical state containing only an aqueous phase and mineral phases. Then, CO₂ is gradually added to the system until a specified amount is reached, which should be high enough to saturate the aqueous phase and produce a gaseous phase. This example is also useful to assess the robustness of the algorithm when handling phase assemblage transitions.

The modelling of the previous problem is now described. Let H₂O, CO₂, NaCl, CaCO₃ and MgCO₃ denote auxiliary components of the chemical system in Table 2. At the initial and final states of the chemical system, we consider their molar abundance given by Table 3. Next, let \( n_i \) denote the vector of molar abundance of the components. Define the following linear path:

\[
n_i(t) = n_i^0 + t(n_i^f - n_i^0),
\]

where \( n_i^0 \) and \( n_i^f \) are the given initial and final molar abundance of the components; and \( t \in [0, 1] \) is a scalar parameter. The inputs of our sequential equilibrium calculations are determined, therefore, by gradually increasing the parameter \( t \) in order to model the addition or removal of components from the system. Note, however, that the molar abundance of the components \( n_i \) are auxiliary inputs for

### Table 2

Description of the chemical system CO₂–H₂O–NaCl–CaCO₃–MgCO₃ with their phases and respective chemical species.

| Aqueous phase | Gaseous phase |
|---------------|--------------|
| H₂O(l)        | H₂O(g)       |
| H⁺            | CO₂(g)       |
| OH⁻           |               |
| HCO₃⁻         | Mineral phase #1 |
| CO₂⁻          | NaCl(s) (Halite) |
| Na⁺           |               |
| Cl⁻           | Mineral phase #2 |
| Ca₂⁺          | CaCO₃(s) (Calcite) |
| Mg₂⁺          |               |
| MgCl⁺         | Mineral phase #3 |
| CaCl⁻         | MgCO₃(s) (Magnesite) |
| CO₂(aq)       |               |
| CaCO₃(aq)     | Mineral phase #4 |
| MgCO₃(aq)     | CaMg(CO₃)₂(s) (Dolomite) |
| CaCl₂(aq)     |               |

### Table 3

The number of moles of each component at both initial and final states.

| Component | Initial state [moles] | Final state [moles] |
|-----------|-----------------------|---------------------|
| H₂O       | 55.507721             | 55.462962           |
| CO₂       | 7.051770 × 10⁻¹⁰      | 1.716270 × 10⁻⁵     |
| NaCl      | 2.791921 × 10⁻⁸       | 1.122305 × 10⁻⁸     |
| CaCO₃     | 9.827989 × 10⁻⁸       | 0.741635            |
| Na⁺       | 1.765120              | 1.767785            |
| Cl⁻       |                       | 1.759351            |
| NaCl(aq)  | 0.234880              | 0.232215            |
| Ca₂⁺      | 4.382456 × 10⁻⁴       | 0.034912            |
| Mg₂⁺      | 1.783294 × 10⁻⁵       | 0.001406            |
| CO₂⁻      | 2.714616 × 10⁻⁸       | 3.501025 × 10⁻⁶     |
| CaCl⁻     | 7.16976 × 10⁻⁵        | 0.006016            |
| MgCl⁺     | 3.341396 × 10⁻⁶       | 2.571146 × 10⁻⁶     |
| MgCO₃(aq) | 3.012652 × 10⁻⁶       | 6.919652 × 10⁻⁶     |
| CaCl₂(aq) | 1.049437 × 10⁻⁷       | 1.035521 × 10⁻⁷     |

### Table 4

Expressions for the number of moles of the chemical elements from the number of moles of the components.

| Element | Expression |
|---------|------------|
| H       | \( \frac{2n_{\text{H}_2\text{O}}}{n_{\text{H}_2\text{O}}} \) |
| O       | \( \frac{n_{\text{H}_2\text{O}} + 2n_{\text{CO}_2} + 3n_{\text{CaCO}_3} + 3n_{\text{MgCO}_3}}{n_{\text{H}_2\text{O}}} \) |
| C       | \( \frac{n_{\text{CO}_2} + n_{\text{CaCO}_3} + n_{\text{MgCO}_3}}{n_{\text{H}_2\text{O}}} \) |
| Na      | \( \frac{n_{\text{NaCl}}}{n_{\text{NaCl}}} \) |
| Cl      | \( \frac{n_{\text{NaCl}}}{n_{\text{NaCl}}} \) |
| Ca      | \( \frac{n_{\text{CaCO}_3}}{n_{\text{CaCO}_3}} \) |
| Mg      | \( \frac{n_{\text{MgCO}_3}}{n_{\text{MgCO}_3}} \) |

### Table 5

Initial and final equilibrium states of the chemical system CO₂–H₂O–NaCl–CaCO₃–MgCO₃ at \( T = 60 \, ^\circ \text{C} \) and \( P = 150 \, \text{bar} \).
the chemical equilibrium calculations. The natural inputs are the number of moles of the chemical elements, which can be determined from the expressions in Table 4.

Using the molar abundance of the components in Table 3, we calculated the initial and final equilibrium states of the system as shown in Table 5. It can be seen that the amount of dissolved carbon was considerably increased from the initial to the final state, while the aqueous solution became more acidic, with initial and final pH as 9.2 and 4.8 respectively. In addition, note that the zeroed number of moles of some species indicates that the phases containing them are not present at equilibrium. For example, from all assumed mineral phases, only calcite and dolomite exist at given conditions of temperature, pressure and molar amounts of the components.

The calculations assumed \( T = 60 \, ^\circ \text{C} \) and \( P = 150 \, \text{bar} \), and a tolerance error of \( 10^{-8} \). In addition, the activity coefficient model of Duan and Sun (2003) was used for the aqueous species \( \text{CO}_2(\text{aq}) \) because of the presence of ions \( \text{Ca}^{2+} \) and \( \text{Mg}^{2+} \) in the aqueous solution. The fugacity coefficient model of Duan et al. (2006) was adopted for the gaseous species \( \text{CO}_2(\text{g}) \). The initial guess for the calculations in the absence of good estimates is shown in Table 6.

Figs. 2 and 3 show the efficiency of the Gibbs energy minimisation algorithm without and with the watchdog strategy. They show the required number of iterations to achieve convergence for each of the 20 sequential equilibrium calculations. Observe that the watchdog strategy is capable of boosting the convergence speed significantly. For example, the first calculation required 3.5 times fewer iterations when adopting this strategy.

Note that the first calculation requires more iterations than the others. This is because only poor initial guesses for the molar abundance of the species and the Lagrange multipliers are available. The second calculation still requires a few more iterations than the others, which is justified by the fact that the initial state is relatively distant from the second, since the former did not contain any \( \text{CO}_2 \). Nevertheless, observe that the calculations achieve convergence in 1–3 iterations after the second calculation.

At about \( n_{\text{CO}_2} = 0.8 \, \text{mol} \), the gaseous phase emerges in the system and the number of iterations increases slightly during the handling of this phase assemblage transition, as seen in Fig. 2. As shown in Fig. 3, however, the number of iterations in this region is just slightly affected when the watchdog strategy is used. This is an important efficiency demonstration of the method, since in multiphase reactive transport simulations the front of the flow is constantly experiencing appearance and disappearance of phases.

Figs. 2 and 3 adopted the scaling procedure presented in Section 3.4. Figs. 4 and 5, on the other hand, show the same calculations without such scaling. Fig. 4 indicates that the trust-region minimisation algorithm alone is highly sensitive to scaling, as discussed in Nocedal and Wright (1999), and its efficiency is severely compromised without it. The watchdog strategy, however, is just slightly affected by the lack of scaling, with its main inefficiency occurring near a phase boundary as shown in Fig. 5. Therefore, we see that scaling of the variables is fundamental for efficient equilibrium calculations, either using or not the watchdog strategy, since any small gain in efficiency can cause expressive performance results in reactive transport simulations.

Figs. 6 and 7 show the efficiency of the method when performing 100 sequential calculations instead of 20. Since these calculations used shorter subintervals, it can be said

### Table 6

| Species | Initial guess |
|---------|---------------|
| \( \text{H}_2\text{O}(\text{l}) \) | \( n_{\text{H}_2\text{O}} \) |
| \( \text{CO}_2(\text{g}) \) | \( n_{\text{CO}_2} \) |
| Halite | \( n_{\text{NaCl}} \) |
| Calcite | \( n_{\text{CaCO}_3} \) |
| Magnesite | \( n_{\text{MgCO}_3} \) |
| Others | 0.01 |

Fig. 2. Performance of 20 sequential chemical equilibrium calculations in terms of number of iterations using the monotone trust-region algorithm.
that the difference between two of its consecutive solutions is smaller than the one when 20 sequential calculations are performed. This is reflected by the decrease in the number of iterations in the region \( \frac{1}{2} \leq \theta \leq 0.8 \) when compared to the same region in Figs. 2 and 3. Finally, increasing the number of subintervals to 1000 yields an average number of iterations of 1.2 with the use of watchdog strategy.

Fig. 8 demonstrate that the sequence of equilibrium calculations are still efficiently calculated assuming simultaneous variations in temperature and pressure. The initial and final temperatures and pressures used were 60–160 °C and 100–300 bar, resulting in a variation of 5 °C and 10 bar for every equilibrium calculation. Note that the number of iterations after the gaseous phase is formed has increased by one. This is because this phase is more sensitive to changes in temperature and pressure, and an extra iteration is necessary to correct the molar abundance of the gaseous species.

These results demonstrate that the use of the watchdog technique is fundamental if this algorithm is to be used for multiphase reactive transport modelling. By adopting this nonmonotone strategy, the calculations will potentially converge in only few iterations, even in regions near the front flow. Moreover, the scaling technique used in the calculations is capable of further decreasing the number of iterations necessary to solve the chemical equilibrium problems. As to the accuracy of the calculations, the mass-balance residuals are either zero or in the range \( 10^{-25} \) to \( 10^{-19} \) for some elements. Thus, the accuracy of the method satisfies the relative threshold of \( 10^{-13} \) discussed in Kulik.
et al. (2013), which is necessary for avoiding accumulation of errors in a reactive transport simulation over time.

In Figs. 2–8, the additional effort in correcting the mass-balance of the stable phases once an equilibrium calculation has finished was not represented. In these sequential calculations, we performed only a single correction iteration when unstable phases were encountered.

Finally, we present Fig. 9 as a side result of all these calculations. It shows the pH of the aqueous solution and the molalities of the species Ca\(^{2+}\) and Mg\(^{2+}\) as carbon dioxide is injected into the system. Observe that the aqueous phase becomes acidic with the injection of CO\(_2\), and promotes further dissolution of the carbonate minerals, as suggested by the increase in the concentrations of the ions Ca\(^{2+}\) and Mg\(^{2+}\). Also note that once the gaseous phase emerges in the system, the solubility of CO\(_2\) in the aqueous phase remains constant, and so does the pH and the concentrations of the ionic species.

### 4.3. Convergence rates of the interior-point method

In Section 1 we argued that superlinear rates of convergence are essential if this algorithm is to be incorporated in critical applications such as chemical kinetics and reactive transport modelling. Because initial guesses in these applications are usually close to the solution, with its proximity being mainly dependent on the used time step, it is of utmost importance that an equilibrium algorithm converges at fast rates locally. In this section we present results that permit us to analyse the convergence rates of the
trust-region interior-point algorithm with and without the watchdog strategy.

Figs. 10–12 present the convergence plot of equilibrium calculations assuming the multiphase system of Table 2. The residual of a calculation is defined as the Euclidean norm of the relaxed KKT function $F(w)$ given in Eq. (3.9). All calculations were performed using poor initial guesses, so that the initial residual would be large. The watchdog threshold $\mu_w = 10^{-6}$ was used instead of $\mu_w = 10^{-3}$ to create two distinct regions in the graphs that show when the watchdog strategy is activated. The solid circles indicate the activation of the watchdog strategy, and the empty circles indicate the maximum residual attained during the non-monotone iterations. Note that all calculations succeeded under the watchdog mode, without returning to the monotone trust-region strategy.

These calculations were performed with different equilibrium conditions. In Fig. 10, for example, the amount of CO$_2$ in the system is not enough to produce a gaseous phase. In Fig. 11, a gaseous phase is about to emerge, and so the calculation was performed near a phase boundary. In Fig. 12, the specified amount of CO$_2$ was sufficient to saturate the aqueous phase and to form a gaseous phase. These results show that the use of the watchdog strategy boosts the convergence rate near the solution. This applies even in critical regions, such as those where a phase boundary exists. By using the nonmonotone watchdog strategy, our calculations could achieve quadratic rates of convergence.
convergence. These rates are superior to the linear ones obtained with the original monotone trust-region interior-point method of Ulbrich et al. (2004) and Silva et al. (2008).

The convergence rates $a_r$ of the calculations were computed using the formula:

$$a_r := \log(r_{k+1}/r_k)/\log(r_k/r_{k-1}),$$

obtained by assuming that at the $k$th iteration the following applies:

$$r_{k+1} = Cr_k^a,$$

where $r_k$ denotes the residual at the $k$th iteration, and $C$ is a positive constant. The values of the convergence rates displayed in Figs. 10–12 are average of the last three iterations.

4.4. Sensitivity of the interior-point method

The description of the interior-point method in the previous sections did not contain any discussion about how sensitive its produced solutions are with respect to the final perturbation parameter $\mu$. Recall that this parameter is used to relax the strict complementary conditions (3.5) of the KKT equations. Therefore, this relaxation is expected to perturb the solution as well, which we shall see in this section at which extent this happens. For our experiments, however, we will allow the watchdog strategy to be used, which uses a small constant perturbation parameter $\bar{\mu}$ until convergence.

Consider the multiphase system of Table 1. Similarly as before, we will fix the molar amounts of components $\text{H}_2\text{O}$...
and NaCl and gradually add CO₂ to the system. Eventually a gaseous phase will be formed, but before this happens we want to investigate the effect of parameter \( \mu \) on the number of moles of the inexistent gaseous species. We expect that these are tiny values in comparison with the number of moles in other phases.

Fig. 13 shows the effect of different values of \( \mu \) on the number of moles of the gaseous species CO₂\(_\text{g}\). The figure shows that the smaller the perturbation parameter \( \mu \) is, the more accurate is the numerical representation of the inexistent gaseous species. For example, using \( \mu = 10^{-15} \) results in the inexistent CO₂\(_\text{g}\) having number of moles in the order of \( 10^{-12} \). Compare this with the case \( \mu = 10^{-9} \), where the number of moles of CO₂\(_\text{g}\) in its inexistent region is about to \( 10^0 \). Moreover, observe that the smaller the perturbation parameter \( \mu \) is, the sharper is the phase assemblage transition curve.

**5. CONCLUDING REMARKS**

The efficiency of a chemical equilibrium method for multiphase systems has been assessed. The method has been developed based on a primal-dual interior-point algorithm for non-convex minimisation problems. It provides flexibility for the imposition of several types of equilibrium constraints besides mass-balance and charge-balance. The trust-region approach aids convergence of the equilibrium calculations starting from poor initial guesses. The restoration algorithm increases the robustness of the method, overcoming convergence problems that would lead the
calculation to a failure. The watchdog strategy produces quadratic rates of convergence locally and substantially reduces the number of iterations. The method comprises a rigorous phase stability test to identify the unstable phases, allowing the enforcement of the Gibbs phase rule and correction of mass distribution among the stable phases.

We have successfully applied the Gibbs free energy minimisation algorithm in a geochemical equilibrium problem relevant to CO2 sequestration in saline aquifers. The method was capable of determining the correct phase assemblage at equilibrium in a system containing many phases. Application of the algorithm for sequential equilibrium calculations showed that the approach is very efficient, requiring 1–3 iterations in most calculations, even in critical regions such as phase boundaries. The results demonstrate that reactive transport simulations can potentially benefit by using this multiphase equilibrium algorithm.

The interested reader can find the implemented Gibbs energy minimisation algorithm in REAKTORO, a C++ scientific library for geochemical modelling, freely available at http://bitbucket.org/reactoro.

ACKNOWLEDGEMENTS

We gratefully acknowledge funding from the Qatar Carbonates and Carbon Storage Research Centre (QCCSRC), provided jointly by Qatar Petroleum, Shell, and Qatar Science & Technology Park.

APPENDIX A. CONVERGENCE RATES OF A STOICHIOMETRIC ALGORITHM

Here we briefly describe a common stoichiometric formulation for equilibrium calculations followed by comments on its convergence rate. Differently from a non-stoichiometric formulation, which is based in the minimisation of the Gibbs free energy, the stoichiometric formulation is based on the solution of a system of mass-action and mass-balance equations (Smith and Missen, 1982).

A.1. Mathematical formulation

In what follows, the chemical system has been partitioned into primary and secondary species. The $i$th primary species is denoted by $x_i$, and the $j$th secondary species by $x_j$. As it is commonly done in the geochemistry literature, the subscripts $i$ and $j$ are enough to tell if a species is primary or secondary (Bethke, 2007). In addition, the number of primary and secondary species are denoted respectively by $N_i$ and $N_j$. For simplicity reasons, only an aqueous phase is assumed in the system.

In the stoichiometric approach, the primary and secondary species are related to each other according to the following linearly independent system of equilibrium reactions:

$$ x_j = \sum_{i=1}^{N_i} v_{ij} x_i \quad (j = 1, \ldots, N_j), \tag{A.1} $$

where $v_{ij}$ is the stoichiometric coefficient of the $i$th primary species in the $j$th secondary species. For example, for the system $\text{H}_2\text{O} - \text{CO}_2$, we can write the following linearly independent reactions:

$$ \text{OH}^- = \text{H}_2\text{O} - \text{H}^+, $$
$$ \text{CO}_2(\text{aq}) = \text{HCO}_3^- + \text{H}^+ - \text{H}_2\text{O}, $$

where the sets of primary and secondary species are $\{\text{H}_2\text{O}, \text{H}^+, \text{HCO}_3^-\}$ and $\{\text{OH}^-, \text{CO}_2(\text{aq})\}$.

Computing the equilibrium state of the system requires solving the following mass-balance equations:

$$ m_i + \sum_{j=1}^{N_j} v_{ij} m_j = M_i \quad (i = 1, \ldots, N_i), \tag{A.2} $$

and the following mass-action equations:

$$ m_j = \frac{1}{K_{j/i}} \prod_{i=1}^{N_i} (m_i x_i)^{v_{ij}} \quad (j = 1, \ldots, N_j), \tag{A.3} $$
where $m_i$ and $m_j$ are the molalities of the $i$th primary species and $j$th secondary species; $M_i$ is the given total molality of the $i$th primary species; $\gamma_i$ and $\gamma_j$ are the activity coefficients of the $i$th primary species and $j$th secondary species; and $K_j$ is the equilibrium constant of the $j$th reaction.

Remark: Using molalities as unknowns has the inconvenience of the need to handle different equations for solvent water and solutes, since the molality of water does not make sense. In this discussion, however, we keep it simple by listing only those mass-balance and mass-action equations that involves only solutes. The presented idea, however, is not compromised by this lack of completeness.

### A.2. Numerical method

The common practice of solving Eqs. (A.2) and (A.3) is to eliminate the dependency of the mass-balance Eq. (A.2) on the molalities of the secondary species $m_j$. This results in the following system of equations:

$$m_i + \sum_{j=1}^{N_j} \frac{v_j}{K_{ji}} \prod_{k=1}^{N_i} (m_i^{x_k})^{y_k} = M_i \quad (i = 1, \ldots, N_i). \quad (A.4)$$

Let $k$ denote the iteration counter of the calculation. Assume that the molalities of the primary and secondary species $m_i$ and $m_j$ are known at iteration $k$. Therefore, the new molalities of the primary species $m_i^{k+1}$ can be calculated by applying Newton’s method to the following system of equations:

$$m_i^{k+1} + \sum_{j=1}^{N_j} \frac{v_j}{K_{ji}} \prod_{k=1}^{N_i} (m_i^{k+1, x_k})^{y_k} = M_i \quad (i = 1, \ldots, N_i). \quad (A.5)$$

Then, using the new values $m_i^{k+1}$, we can update the molalities of the secondary species as:

$$m_j^{k+1} = \frac{1}{K_{ji}} \prod_{k=1}^{N_i} (m_i^{k+1, x_k})^{y_k} \quad (j = 1, \ldots, N_j). \quad (A.6)$$

where it can be observed the resemblance of this update of $m_j^{k+1}$ with a successive substitution method.

Note that the dependence of the activity coefficients $\gamma_i$ and $\gamma_j$ on the molalities $m_i$ and $m_j$ is commonly ignored, and their values are held constant until the end of the iteration. According to Bethke (2007), this is called the soft formulation. The hard formulation, on the other hand, takes into account a partial dependence of the activity coefficients $\gamma_i$ and $\gamma_j$ on the molalities of the primary species $m_i$ only. Their update is done once both $m_i^{k+1}$ and $m_j^{k+1}$ have been obtained, after which the previous two-stage calculations using Eqs. (A.5) and (A.6) are repeated until convergence.

If a full dependence of the activity coefficients on the species molalities were taken into account, then Eq. (A.5) would become:

$$m_i^{k+1} + \sum_{j=1}^{N_j} \frac{v_j}{K_{ji}} \prod_{k=1}^{N_i} (m_i^{k+1, x_k})^{y_k} = M_i \quad (i = 1, \ldots, N_i). \quad (A.7)$$

where $\gamma_i^{k+1}$ and $\gamma_j^{k+1}$ are unknowns that depend on $m_i^{k+1}$ and $m_j^{k+1}$. As a result, it is no longer possible to calculate $m_i^{k+1}$ and $m_j^{k+1}$ in two stages as before. This formulation based on the law of mass-action approach that adopts a complete Newton scheme has been further studied in Leal et al. (2013).

### A.3. Discussion

The motivation of only applying Newton’s method to Eq. (A.5) is to avoid large linear systems. This incomplete Newton’s scheme results in Jacobian matrices that are only $N_i \times N_i$, and since the number of primary species $N_i$ is in general considerably smaller than that of secondary species $N_j$, the size of the linear systems is substantially decreased.

This incomplete scheme is only possible if the non-linear dependence of the activity coefficients on the molalities of the species is completely or partially ignored. Eqs. (A.5) and (A.6) were obtained by completely ignoring this dependence, using a soft formulation, and so the following derivative terms are not included in the Jacobian matrix:

$$\frac{\partial c_j}{\partial m_i}, \frac{\partial c_j}{\partial m_j}, \frac{\partial c_j}{\partial m_i^2}, \frac{\partial c_j}{\partial m_j^2}, \quad (A.8)$$

which are extremely valuable information. Its absence can negatively impact the convergence rate of the calculation, specially in strong non-ideal systems or near phase boundaries, where these values are very sensitive. If, on the other hand, a hard formulation were used to obtain Eq. (A.5), then only the derivative terms:

$$\frac{\partial c_j}{\partial m_i}, \frac{\partial c_j}{\partial m_j}, \quad (A.9)$$

would have been taken into account. Since in general $N_i \ll N_j$, the above terms represent just a small portion of all that should be considered.

Thus, we see that the resulting Jacobian matrix, in both the soft and hard formulations, always lacks important first-order partial derivative information. Since quadratic convergence rates in Newton’s method is subject to exact Jacobian expressions, the presented stoichiometric method should expect slower convergence rates. In fact, it can be said that the incomplete Newton’s method is fundamentally more similar to the family of quasi-Newton methods, which converges $Q$-superlinearly, than standard Newton methods, which converges $Q$-quadratically (see Kelley, 1995; Nocedal and Wright, 1999).

Depending on the kind of chemical system, and its physical and chemical conditions, a soft or hard stoichiometric method might still converge in few iterations. Perhaps even less than an approach that attains quadratic convergence rates, since these optimal rates only happen near the solution. Therefore, it is near the solution that the algorithms should differ in performance, since one that converges with quadratic rates should do so in very few iterations. From this we see that for equilibrium problems in the context of chemical kinetics or reactive transport modelling, where the calculations always start with hot initial guesses, a well implemented quadratic convergent method can deliver optimal performance.
REFERENCES

Allison, J. and Kevin, J. (1991) MINTEQA2/PRODEFA2: A Geochemical Assessment Model for Environmental Systems: Version 3.0 User’s Manual. Technical Report 600/3-91/021.

Bale C., Chartrand P., Dechterov S., Eriksson G., Hack K., Ben Mahmoud R., Melançon J., Pelton A. and Petersen S. (2002) FactSage thermochemical software and databases. *Calphad* **26**(2), 189–228.

Bale C., Bélisle E., Chartrand P., Dechterov S., Eriksson G., Hack K., Jung I.-H., Kang Y.-B., Melançon J., Pelton A., Robelin C. and Petersen S. (2009) FactSage thermochemical software and databases—recent developments. *Calphad* **33**(2), 295–311.

Ball J., Nordstrom D. and Zachmann D. (1987) WATEQ4F—a personal computer FORTRAN translation of the geochemical model WATEQ2 with revised data base. *US Geol. Surv. Open-File Rep. 87*(50), 108.

Bea S., Carrera J., Ayora C., Batlle F. and Saaltink M. (2009) CHEPROO: a Fortran 90 object-oriented module to solve chemical processes in Earth Science models. *Comput. Geosci.* **35**(6), 1098–1112.

Benson, H., Shanno, D. F. and Vanderbei, R. J. (2000) Interior and comparative numerical testing. *Operations Research and Financial Engineering, Princeton University*, ORFE-00-02.

Bethke C. M. (2007) *Geological and Biogeochemical Reaction Modeling, 2* ed. Cambridge University Press, New York.

Byrd R. H., Hribar M. E. and Nocedal J. (1999) An interior point algorithm for large-scale nonlinear programming. *SIAM J. Optim.* **9**(4), 877–900. http://dx.doi.org/10.1137/S1052623497325107.

Chamberlain R., Powell M., Lemarechal C. and Pedersen H. (1982) The watchdog technique for forcing convergence in algorithms for constrained optimization. *Math. Program. Stud.* **16**, 1–17.

Costa M. F. P. and Fernandes E. M. G. P. (2008) Practical implementation of an interior point nonmonotone line search filter method. *Int. J. Comput. Math.* **85**(3-4), 397–409.

Costa M. F. P. and Fernandes E. M. (2011a) Assessing the potential of interior point barrier filter line search methods: nonmonotone versus monotone approach. *Optimization* **60**(10-11), 1251–1268.

Costa, M.F. P. and Fernandes, E. M. G. P. (2011b) On Minimizing Objective and KKT Error in a Filter Line Search Strategy for an Interior Point Method. In *Computational Science and Its Applications - ICCSA 2011, volume 6784 of Lecture Notes in Computer Science* (eds. B. Murgante, O. Gervasi, A. Iglesias, D. Taniar, and B. Apduhan). Springer Berlin Heidelberg. pp. 231–244.

de Capitani C. and Brown T. H. (1987) The computation of chemical equilibrium in complex systems containing non-ideal solutions. *Geochim. Cosmochim. Acta* **51**(10), 2659–2652.

Drummond, S. (1981) Boiling and mixing of hydrothermal fluids: chemical effects on mineral precipitation. Ph. D., Pennsylvania State University.

Duan Z. and Sun R. (2003) An improved model calculating CO$_2$ solubility in pure water and aqueous NaCl solutions from 273 to 533 K and from 0 to 2000 bar. *Chem. Geol.* **193**(3-4), 257–271.

Duan Z., Moller N. and Weare J. H. (1992) An equation of state for the CH$_4$-CO$_2$-H$_2$O system: I. Pure systems from 0 to 1000 °C and 0 to 8000 bar. *Geochim. Cosmochim. Acta* **56**(7), 2605–2617.

Duan Z., Sun R., Zhu C. and Chou I.-M. (2006) An improved model for the calculation of CO$_2$ solubility in aqueous solutions containing Na$^+$, K$^+$, Ca$^{2+}$, Mg$^{2+}$, Cl, and SO$_4^{2-}$. *Mar. Chem.* **98**(2-4), 131–139.

El-Bakry A. S., Tapia R. A., Tsuchiya T. and Zhang Y. (1996) On the formulation and theory of the Newton interior-point method for nonlinear programming. *J. Optim. Theory Appl.* **89**(3), 507–541.

Eriksson G. and Hack K. (1990) ChemSage—a computer program for the calculation of complex chemical equilibria. *Metall. Trans. B* **21**(6), 1013–1023.

Fiacco, A. V. and McCormick, G. P. (1990) Nonlinear programming: sequential unconstrained minimization techniques. *Soc. Ind. Appl. Math.*

Fletcher R. (2000) *Practical Methods of Optimization*, second ed. Wiley.

Fletcher R. and Leyffer S. (2002) Nonlinear programming without a penalty function. *Math. Program.* **91**(2), 239–269.

Gill P. E., Murray W. and Saunders M. a. (2002) SNOPT: an SQP algorithm for large-scale constrained optimization. *SIAM J. Optim.* **12**(4), 979–1006.

Gill, P. E., Murray, W. and Saunders, M. A. (2005) *Users Guide for SNOPT Version 7: A Fortran Package for Large-Scale Linear and Quadratic Programming*. Numerical Analysis Report NA 97-4.

Harvey J.-P., Eriksson G., Orban D. and Chartrand P. (2013) Global minimization of the Gibbs energy of multicomponent systems involving the presence of order/disorder phase transitions. *Am. J. Sci.* **313**(3), 199–241.

Harville C. E., Greenberg J. P. and Weare J. H. (1987) A chemical equilibrium algorithm for highly non-ideal multiphase systems: free energy minimization. *Geochim. Cosmochim. Acta* **51**(5), 1045–1057.

Helgeson H. C. and Kirkham D. H. (1974a) Theoretical prediction of the thermodynamic behavior of aqueous electrolytes at high pressures and temperatures: I. Summary of the thermodynamic/electrostatic properties of the solvent. *Am. J. Sci.* **274**(10), 1089–1198.

Helgeson H. C. and Kirkham D. H. (1974b) Theoretical prediction of the thermodynamic behavior of aqueous electrolytes at high pressures and temperatures: II. Debye-Hückel parameters for activity coefficients and relative partial molar properties. *Am. J. Sci.* **274**(10), 1199–1261.

Helgeson H. C. and Kirkham D. H. (1976) Theoretical prediction of the thermodynamic properties of aqueous electrolytes at high pressures and temperatures: III. Equation of state for aqueous species at infinite dilution. *Am. J. Sci.* **276**(2), 97–240.

Helgeson H. C., Delany J. M., Nesbitt H. W. and Bird D. K. (1978) Summary and critique of the thermodynamic properties of rock-forming minerals. *Am. J. Sci.* **278** *(1)*.

Helgeson H. C., Kirkham D. H. and Flowers G. C. (1981) Theoretical prediction of the thermodynamic behavior of aqueous electrolytes at high pressures and temperatures: IV. Calculation of activity coefficients, osmotic coefficients, and apparent molar and standard and relative partial molar properties to 600 C. *Am. J. Sci.* **281**(10), 1249–1516.

Hou S.-X., Maitland G. C. and Trusler J. M. (2013) Phase equilibria of (CO$_2$)+H$_2$O+NaCl and (CO$_2$)+H$_2$O+KCl: measurements and modeling. *J. Supercrit. Fluids* **78**, 78–88.

Johnson J. W., Oelkers E. H. and Helgeson H. C. (1992) SUPCRT92: a software package for calculating the standard molal thermodynamic properties of minerals, gases, aqueous species, and reactions from 1 to 5000 bar and 0 to 1000 C. *Comput. Geosci.* **18**(7), 899–947.

Karpov I., Chudnenko K., Kulik D. and Bychinskii V. (2002) The convex programming minimization of five thermodynamic potentials other than Gibbs energy in geochemical modeling. *Am. J. Sci.* **302**(4), 281–311.

Karpov I., Chudnenko K. and Kulik D. (1997) Modeling chemical mass transfer in geochemical processes; thermodynamic rela-
sequestration in geologic media. In Shao H., Dmytrieva S. V., Kolditz O., Kulik D. a., Pfingsten W., Saaltink M. W., Ayora C. and Carrera J. (1998) A mathematical equilibrium partitioning with the GEMS/PSI code. Technical report, Paul Scherrer Institut, Villigen, Switzerland.

Kulik D. (2006) Dual-thermodynamic estimation of stoichiometry and stability of solid solution end members in aqueous-solid solution systems. Chem. Geol. 228(3-4), 189–212.

Kulik, D., Berner, U. and Curti, E. (2004) Modelling chemical equilibrium partitioning with the GEMS/PSI code. Technical report, Paul Scherrer Institut, Villigen, Switzerland.

Kulik D. A., Wagner T., Dmytrieva S. V., Kosakovski G., Hingerl F. F., Chudnenko K. V. and Berner U. R. (2013) GEM-Selektor geochemical modeling package: revised algorithm and GEMS3K numerical kernel for coupled simulation codes. Comput. Geosci. 17(1), 1–24. http://dx.doi.org/10.1007/s10096-012-9310-6.

Leal A. M., Blunt M. J. and LaForce T. C. (2013) A robust and efficient numerical method for multiphase equilibrium calculations: application to CO2-brine–rock systems at high temperatures, pressures and salinities. Adv. Water Resour. 62, 409–430.

Leal, A. M., Blunt, M. J. and LaForce, T. C. (2014a) A chemical equilibrium algorithm for geochemical modelling: a trust-region interior-point method. In preparation, Imperial College London, London, UK.

Leal, A. M., Blunt, M. J., and LaForce, T. C. (2014b) A chemical kinetics algorithm for geochemical speciation and reactive transport modelling. In preparation, Imperial College London, London, UK.

Lichtner P. (1985) Continuum model for simultaneous chemical reactions and mass transport in hydrothermal systems. Geochim. Cosmochim. Acta 49(3), 779–800.

Lu, C. and Lichtner, P. C. (2005). PFLOTRAN: massively Parallel 3D simulator for CO2 sequestration in geologic media. In DOE/NETL fourth annual conference on carbon capture and sequestration.

Morel F. and Morgan J. (1972) Numerical method for computing equilibrium in aqueous chemical systems. Environ. Sci. Technol. 6(1), 58-67.

Nocedal J. and Wright S. J. (1999) Numerical Optimization, 2 ed. Springer.

Parkhurst D. and Appelo C. (1999) User’s guide to PHREEQC (Version 2)—a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. USGS Water Resour. Invest. Rep. 99(4259), 326.

Parkhurst, D. and Appelo, C. (2013) Description of input and examples for PHREEQC version 3—A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. In Groundwater Book 6, Modeling Techniques, chapter A43, page 497. U.S. Geological Survey Techniques and Methods.

Reed M. H. (1982) Calculation of multicomponent chemical equilibria and reaction processes in systems involving minerals, gases and an aqueous phase. Geochem. Cosmochim. Acta 46(4), 513–528.

Saaltink M. W., Ayora C. and Carrera J. (1998) A mathematical formulation for reactive transport that eliminates mineral concentrations. Water Resour. Res. 34(7), 1649–1656.

Schecher, W. and McAvoy, D. (1998) MINERL+: A chemical equilibrium modeling system. Environmental Research Software: Hallowell.

Shao H., Dmytrieva S. V., Kolditz O., Kulik D. a., Pfingsten W. and Kosakovski G. (2009) Modeling reactive transport in non-ideal aqueous/solid solution system. Appl. Geochem. 24(7), 1287–1300.

Shock E. and Helgeson H. C. (1988) Calculation of the thermodynamic and transport properties of aqueous species at high pressures and temperatures: correlation algorithms for ionic species and equation of state predictions to 5 kb and 1000 °C. Geochim. Cosmochim. Acta 52(8), 2009–2036.

Shock E. L., Oelkers E. H., Johnson J. W., Sverjensky D. A. and Helgeson H. C. (1992) Calculation of the thermodynamic properties of aqueous species at high pressures and temperatures, effective electrostatic radii, dissociation constants and standard partial molar properties to 1000 °C and 5 kbar. J. Chem. Soc., Faraday Trans. 88(6), 803.

Shvarov Y. V. (1999) Algorithmization of the numeric equilibrium modeling of dynamic geochemical processes. Geochim. Int. 37, 562–570.

Shvarov Y. V. (2008) HCh: new potentialities for the thermodynamic simulation of geochemical systems offered by windows. Geochim. Int. 46(8), 834–839.

Silva, R., Ulbrich, M., Ulbrich, S. and Vicente, L. N. (2008) A globally convergent primal-dual interior-point filter method for nonlinear programming: new filter optimality measures and computational results. Technical report, Department of Mathematics, University of Coimbra.

Smith W. R. and Missen R. W. (1982) Chemical Reaction Equilibrium Analysis: Theory and Algorithms. Wiley Inter-science, New York.

Spycher N. and Reed M. (1988) Fugacity coefficients of H2, CO2, CH4, H2O and of H2O–CO2–CH4 mixtures: a virial equation treatment for moderate pressures and temperatures applicable to calculations of hydrothermal boiling. Geochim. Cosmochim. Acta 52(3), 739–749.

Spycher N., Pruess K. and Ennis-King J. (2003) CO2–H2O mixtures in the geological sequestration of CO2. Assessment and calculation of mutual solubilities from 12 to 100°C and up to 600 bar. Geochim. Cosmochim. Acta 67(16), 3015–3031.

Steefel C. and Lasaga A. (1994) A coupled model for transport of multiple chemical species and kinetic precipitation/dissolution reactions with applications to reactive flow in single phase hydrothermal systems. Am. J. Sci. 294(5), 529–592.

Steefel, C. I. (2009) CrunchFlow: Software for Modeling Multicomponent Reactive Flow and Transport – User’s Manual. Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, USA.

Steefel C. I. and Cappellen P. V. (1990) A new kinetic approach to modeling water-rock interaction: the role of nucleation, precursors, and Ostwald ripening. Geochim. Cosmochim. Acta 54(10), 2657–2677.

Steefel C. I. and MacQuarrie K. T. B. (1996) Approaches to modeling of reactive transport in porous media. Rev. Mineral. Geochem. 34(1), 83–129.

Steefel C., Depaolo D. and Lichtner P. (2005) Reactive transport modeling: an essential tool and a new research approach for the Earth sciences. Earth Planet. Sci. Lett. 240(3–4), 539–558.

Tanger J. C. and Helgeson H. C. (1988) Calculation of the thermodynamic and transport properties of aqueous species at high pressures and temperatures; revised equations of state for the standard partial molar properties of ions and electrolytes. Am. J. Sci. 288(1), 19–98.

Truesdell A. H. and Jones B. F. (1974) WATEQ–A computer program for calculating chemical equilibrium of natural waters. J. Res. US Geol. Sur. 2(2), 233–248.

Ulbrich M., Ulbrich S. and Vicente L. N. (2004) A globally convergent primal-dual interior-point filter method for nonlinear programming. Math. Program. 100(2), 379–410.
van der Lee, J. and Windt, L. D. (2002) CHESS Tutorial and Cookbook, Version 3.0. Technical Report 1-2, Ecole des Mines de Paris, Fontainebleau, France.

Van Zeggeren F. and Storey S. H. (1970) The Computation of Chemical Equilibria. Cambridge University Press, London, England.

Vanderbei R. J. (1999) LOQO: an interior point code for quadratic programming. Optim. Meth. Software 11(1–4), 451–484.

Vanderbei, R. J. (2006) LOQO User's Manual – Version 4.05.

Vächter A. and Biegler L. T. (2005a) Line search filter methods for nonlinear programming: local convergence. SIAM J. Optim. 16(1), 32–48.

Vächter A. and Biegler L. T. (2005b) Line search filter methods for nonlinear programming: motivation and global convergence. SIAM J. Optim. 16(1), 1–31.

Wächter A. and Biegler L. T. (2005c) On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Math. Program. 106(1), 25–57.

Wagner W. and Pruss A. (2002) The IAPWS formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use. J. Phys. Chem. Ref. Data 31(2), 387.

Westall, J., Zachary, J. and Morel, F. (1976) MINEQL: a computer program for the calculation of chemical equilibrium composition of aqueous systems.

Wolery, T. J. (1992) EQ3/6, A Software Package for Geochemical Modeling of Aqueous Systems: Package Overview and Installation Guide. Technical report, Lawrence Livermore National Laboratory, Livermore, California.

Xu, T., Sonnenthal, E., Spycher, N. and Pruess, K. (2004) TOUGHREACT - User’s Guide: A Simulation Program for Non-isothermal Multiphase Reactive geochemical Transport in Variable Saturated Geologic Media. Technical Report 2, Lawrence Berkeley National Laboratory.

Xu T., Sonnenthal E., Spycher N. and Pruess K. (2006) TOUGHREACT - a simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media: Applications to geothermal injectivity and CO2 geological sequestration. Comput. Geosci. 32(2), 145–165.

Zeleznik F. J. and Gordon S. (1968) Calculation of complex chemical equilibria. Ind. Eng. Chem. 60, 27–57.

Associate editor: C. Steefel