Comput Geosci (2017) 21:131–150
DOI 10.1007/s10596-016-9600-5

ORIGINAL PAPER

Comparison of linear solvers for equilibrium geochemistry computations

Hela Machat[1,2,3] · **Jérôme Carrayrou**[1]

Received: 11 February 2015 / Accepted: 24 October 2016 / Published online: 10 December 2016

Abstract Equilibrium chemistry computations and reactive transport modelling require the intensive use of a linear solver under very specific conditions. The systems to be solved are small or very small (4 × 4 to 20 × 20, occasionally larger) and are very ill-conditioned (condition number up to 10^{100}). These specific conditions have never been investigated in terms of the robustness, accuracy, and efficiency of the linear solver. In this work, we present the specificity of the linear system to be solved. Several direct and iterative solvers are compared using a panel of chemical systems, including or excluding the formation of mineral species. We show that direct and iterative solvers can be used for these problems and propose computational keys to improve the chemical solvers.

Keywords Geochemical modelling · Instantaneous equilibrium chemistry · Linear system inversion · Linear solver · Small matrix · Ill-conditioned matrix · Newton-Raphson algorithm

---

Electronic supplementary material The online version of this article (doi:10.1007/s10596-016-9600-5) contains supplementary material, which is available to authorized users.

✉ Jérôme Carrayrou
jerome.carrayrou@unistra.fr

[1] CNRS, ENGEES, LHyGeS UMR 7517, Université de Strasbourg, 67000 Strasbourg, France

[2] Ecole Supérieure des Ingénieur de l'Equipement Rural de Medjez el Bab, Université de Jendouba, Jendouba, Tunisia

[3] Université de Monastir, UR13ES63-Chimie Appliquée et Environnement, 5000 Monastir, Tunisia

1 Introduction

The problem of groundwater management is receiving increasing attention, and many tools have been developed to address this issue. One of these tools, reactive transport models, was first limited to laboratory experiments and was then extended to field problem comprehension. In recent decades, reactive transport models have increased in complexity and efficiency, and they are now used in many fields. Reactive transport models have been used to study the transport of contaminants, such as heavy metals [1, 2] and radioelements [3–5]. Because of the increasing interest in questions related to climate change, many studies on reactive transport have been conducted to examine the possibility of geologic CO_2 sequestration [6–10].

Under the wide variety of models and cases lies a common mathematical description [11–13]. Transport is usually described by an advection-dispersion equation, and the chemistry is formulated under thermodynamic equilibrium. A widely used approach to solve these reactive transport problems is the operator splitting approach [14]. Using this approach, the transport and chemical operators are solved separately at each time step and iteratively for some formulations. As a consequence, the chemistry operator has to be solved at least once per mesh cell per time step. This is one reason for the high computational cost of reactive transport modelling. Some authors have reported that 80 to 90 % of the computation time is dedicated to chemical computation. Many studies have been conducted to reduce the computation time required by reactive transport modelling [15]. Some works have explored parallelization [16], while others have focused on the methods used to solve the transport operator. Nevertheless, improving the resolution of the chemistry operator has been identified as a key point. Some authors have attempted to improve the classic Newton-Raphson method [17], while others have tested other methods, such as Newton-Krylov [16, 18].

In this work, we focus on a specific element of the problem, improving the resolution of the linearized system provided by the Newton-Raphson method. Looking to numerical methods to solve linear systems is not currently a common practice. Indeed, these methods are actually well known [19–23], and all mathematical packages for scientific computation propose several routines for this task. The motivation of this work comes from the specificity of linear systems that have to be solved for equilibrium chemistry computations. Classic tests for the resolution of linear systems [24–30] are performed using systems provided by finite element or finite volume discretization, leading to matrices that are large (at least 10,000 unknowns) and sparse. Moreover, even when ill-conditioned systems have been studied [25, 30, 31], the conditioning of the matrix coming from the chemical system is specific, as underlined by Hoffmann et al. [32]. For example, Soleymani [33] worked with an ill-conditioned system constructed from 10×10 to 20×20 Hilbert matrices. The condition numbers then range from 3.5×10^{13} to 6.2×10^{28}. In this work, we present chemical tests leading to a 7×7 matrix with a condition number of approximately 10^{180}.

We expect to find a method to increase the efficiency of a speciation or reactive transport code. Several properties are required for such a method:

(i) This method should be fast, as the linear system will be solved very often. In the case of reactive transport modelling, the system will be solved at least once per mesh cell per time step.

(ii) The method should be very robust. It should be able to solve the linear system even if it is very poorly conditioned. Because the resolution of the linear system is only part of an iterative Newton step, an accurate solution is not absolutely needed. Thus, some advanced codes (e.g. Linear Algebra PACKage (LAPACK) routine) that check the accuracy of the solution and return an error flag instead of an inaccurate solution are, in this work, less robust than the more rustic routines.

(iii) The method should be able to detect failure and return an error flag to the main program so that a recovery procedure can be initiated. In the case of reactive transport modelling, this procedure could involve rejecting the current time step and recomputing with a smaller one.

(iv) In the initial analysis, the precision of the method is not the key point. Because the linear system resolution is only a part of the Newton-Raphson iterative procedure, *reasonable* error is acceptable for the linear system inversion. If this error is too large, it will slow the convergence speed for the Newton-Raphson method and decrease the efficiency of the reactive transport code. In this work, errors are estimated by comparing the calculated solution to a reference solution.

Because we utilize a markedly small matrix, we did not test parallelization. All the computations were performed on a PC running Windows with 64-bit Fortran 95. Real variables are defined as double-precision real. We prefer double-precision computations because all the chemical codes are, to the best of our knowledge, written as double-precision real and because quadruple-precision computation is much more time consuming. Nevertheless, we have tested one method using quadruple-precision real to determine whether this development could be useful. Reference solutions are also computed using quadruple precision.

We first present the formulation of the equations describing equilibrium reactions and how they are solved using the Newton-Raphson method. This point defines the Jacobian linear system, which is the object of this work. A second part is devoted to the presentation of the chemical tests and the numerical procedures used to perform the tests. Next, we propose a detailed analysis of the structure and properties of the Jacobian matrix. The selected linear solvers are then presented and tested, and the results are compared and discussed. Based on this analysis, we propose an algorithm to optimize the chemical computation in terms of robustness, accuracy, and efficiency. This algorithm is evaluated on the most selective test. By expanding the limits of the currently used methods, we believe that our new algorithm will contribute to enlarging the field of application of reactive transport modelling. As a conclusion, we underline the main advances of this work, the new perspectives and the remaining obstacles.

2 Material and methods

2.1 Geochemical modelling

One efficient formulation for the computation of thermodynamic equilibrium is based on the tableau concept, referred to as Morel's table [34, 35]. N_X components (X_j) are chosen from the N_C species (C_i) and are used to write the formation of each species as a combination of the components. The mass action law for the formation of the C_i species is written with the equilibrium constant (K_i) and the stoichiometric coefficients ($a_{i,k}$) for each component (X_k)

$$\{C_i\} = K_i \prod_{k=1}^{Nx} \{X_k\}^{a_{i,k}} \tag{1}$$

where $\{C_i\}$ and $\{X_k\}$ are the activities of species C_i and component X_k, respectively. In this work, we define X_j as a subset of C_i; then, N_X is N_C minus the number of reactions.

If N_{CP}-precipitated species (Cp_i) are taken into account, the mass action law for the precipitation of Cp_i is written with the precipitation constant (Kp_i) and the stoichiometric coefficients ($ap_{i,k}$). The saturation index (SI_i) of Cp_i is equal to its activity, which is unity for a pure solid phase

$$\mathrm{SI}_i = Kp_i \prod_{k=1}^{Nx} \{X_k\}^{ap_{i,k}} = 1 \tag{2}$$

The conservation of the total concentration $[T_j]$ of the jth component in the system is then written as

$$[T_j] = \sum_{i=1}^{N_C} a_{i,j} \cdot [C_i] + \sum_{i=1}^{N_{CP}} ap_{i,j} \cdot [Cp_i] \tag{3}$$

where $[C_i]$ is the concentration of species (C_i) and $[Cp_i]$ is the amount of precipitated species (Cp_i) per liquid volume unit.

A classic algorithm [17, 36–41] to describe mineral precipitation or dissolution makes an a priori hypothesis about the existence or non-existence of minerals. In this work, we assume that this hypothesis is proposed. The relationships between the activity and concentration are given by activity coefficients (γ_i) calculated using specific models (Davies, Debye-Hückel, etc.)

$$\{C_i\} = \gamma_i\,[C_i] \text{ and } \left\{X_j\right\} = \gamma_j\left[X_j\right] \tag{4}$$

By substituting the mass action law (1) into the mass conservation equation (3), the following relationship, which only depends on the components and the precipitated species concentrations, is obtained:

$$[T_j] = \sum_{i=1}^{N_C} a_{i,j} \cdot \left(\frac{K_i}{\gamma_i} \cdot \prod_{k=1}^{N_X} (\gamma_k\,[X_k])^{a_{i,k}}\right) + \sum_{i=1}^{N_{CP}} ap_{i,j} \cdot [Cp_i] \tag{5}$$

Combining Eqs. 2 and 5 leads to a set of $(N_X + N_{CP})$ nonlinear algebraic equations, which can be numerically solved through iterative methods. The concentrations of component $[X_k]$ and precipitated species $[Cp_i]$ at equilibrium are then determined when the $(N_X + N_{CP})$ objective functions (Y_j) are zero

$$\begin{aligned} Y_j &= -\left[T_j\right] + \sum_{i=1}^{N_C} a_{i,j} \cdot \left(\frac{K_i}{\gamma_i} \cdot \prod_{k=1}^{N_X} (\gamma_k\,[X_k])^{a_{i,k}}\right) + \sum_{i=1}^{N_{CP}} ap_{i,j} \cdot [Cp_i] && \text{for } j = 1 \text{ to } N_X \\ Y_{j=N_X+i} &= -1 + Kp_i \cdot \prod_{k=1}^{N_X} (\gamma_k\,[X_k])^{ap_{i,k}} && \text{for } i = 1 \text{ to } N_{CP} \end{aligned} \tag{6}$$

Using this method, it is possible to include many chemical phenomena, including activity corrections, sorption on a surface using different means (such as ion exchange or surface complexation), and dissolution of gaseous compounds.

According to the criteria typically used for this method [17, 34, 40, 42], the convergence of the Newton-Raphson method is not checked with respect to the norm of the objective function $||Y||$, but the relative error defined as

$$\mathrm{NR}_{\text{relative error}} = \max\left[\left(\frac{\left|Y_j\right|}{\left|T_j\right| + \sum_{i=1}^{Nc}\left|a_{i,j}\,[C_i]\right|}\right)_{j=1,Nx}, \left(\left|Y_j\right|\right)_{j=Nx+1,Nx+NcP}\right] \leq \varepsilon_{N-R} \text{ with } \varepsilon_{N-R} = 10^{-12} \tag{7}$$

The value of the convergence criterion ($\varepsilon_{N-R} = 10^{-12}$) is set according to usual practice.

2.2 The Newton-Raphson method

The historical approach [12, 34, 37, 40, 42–47] involves the resolution of the system (6) with the Newton-Raphson method using $[X_k]$ and $[Cp_i]$ as primary unknowns. This formulation has some weaknesses that are explained later (see Section 3.1).

However, many authors [18, 32, 38, 39, 48] have proposed an alternative approach. Instead of using the component concentrations $\left[X_j\right]$ as the primary variables, they use the logarithm of the component activities ($\xi_j = \ln\left\{X_j\right\}$). According to this convention, the objective functions defined by Eq. 8 become conservation equations

$$Y_j = -\left[T_j\right] + \sum_{i=1}^{N_C} a_{i,j} \cdot \frac{K_i}{\gamma_i} \cdot \exp\left(\sum_{k=1}^{Nx} a_{i,k} \cdot \xi_k\right) + \sum_{i=1}^{N_{CP}} ap_{i,j} \cdot [Cp_i] \quad \text{for } j = 1 \text{ to } N_X \tag{8}$$

In the case of the objective function describing precipitation, it is more interesting to rewrite the mass action law (2) in log form and then define the objective function

$$Y_{Nx+i} = \ln(\mathrm{SI}_i) = \ln(Kp_i) + \sum_{k=1}^{Nx} ap_{i,k} \cdot \xi_k \text{for } i = 1 \text{ to } N_{CP} \tag{9}$$

Equations 8 and 9 are solved at the nth iteration with the Jacobian matrix (Z^n) of the objective functions

$$\begin{aligned} Z^n_{j,k}\Big|_{\substack{j\,=\,1,\,N_X+N_{CP}\\ k\,=\,1,\,N_X}} &= \frac{\partial Y^n_j}{\partial[\xi_k]^n} \\ Z^n_{j,k}\Big|_{\substack{j\,=\,1,\,N_X+N_{CP}\\ k\,=\,N_X+1,\,N_X+N_{CP}}} &= \frac{\partial Y^n_j}{\partial[Cp_{k-N_X}]^n} \end{aligned} \tag{10}$$

Using an analytical computation, we obtain the $(N_X + N_{CP}) \times (N_X + N_{CP})$ values of Z^n by

$$\begin{aligned} Z^n_{j,k}\Big|_{\substack{j\,=\,1,\,N_X\\ k\,=\,1,\,N_X}} &= \sum_{i=1}^{N_C} a_{i,j} \cdot a_{i,k} \cdot [C_i]^n \\ Z^n_{j,k}\Big|_{\substack{j\,=\,1,\,N_X\\ k\,=\,N_X+1,\,N_X+N_{CP}}} &= ap_{k-N_X,j} \\ Z^n_{j,k}\Big|_{\substack{j\,=\,N_X+1,\,N_X+N_{CP}\\ k\,=\,1,\,N_X}} &= ap_{k,j-N_X} \\ Z^n_{j,k}\Big|_{\substack{j\,=\,N_X+1,\,N_X+N_{CP}\\ k\,=\,N_X+1,\,N_X+N_{CP}}} &= 0 \end{aligned} \tag{11}$$

Even if the activity coefficients depend on the component concentrations, they are assumed to be constant during the Newton-Raphson procedure. These activity coefficients are usually actualized by a fixed-point algorithm at each Newton-Raphson loop.

The progress step of the method $(\Delta\xi^n, \Delta Cp^n)$ is achieved by assuming that the objective function Y^{n+1} in Eq. 12 is equal to zero at the $(n + 1)$th iteration. This produces the key equation of this article, the linear system (12), which must be solved to obtain the progress step

$$Z^n \cdot \left(\Delta\xi^n, \Delta Cp^n\right) = Y^{n+1} - Y^n = -Y^n \tag{12}$$

This system yields the values of the component activities and precipitate concentrations at the $(n + 1)$th iteration

$$\begin{aligned} \xi^{n+1} &= \xi^n + \Delta\xi^n \\ [Cp]^{n+1} &= [Cp]^n + \Delta Cp^n \end{aligned} \tag{13}$$

To simplify the notations, ξ is used to denote the full vector of unknowns, including mineral Cp if present.

2.3 Chemical test cases

We choose chemical test cases with various numbers of components. Some of these chemical systems allow the formation of mineral species. Although it is not realistic from a chemical point of view, we test them without minerals and with the maximal possible number of minerals to obtain the largest matrix size. Appendix 1 presents the stoichiometric coefficients, equilibrium constants, and concentrations for these tests.

Appendices are avialable online.

(i) The *gallic acid* test case was presented by Brassard and Bodurtha [49]. It has been recognized as a challenging test for speciation computation [17] (see Appendix 1 (A-1)).

(ii) The *Valocchi* test is from Valocchi et al. [11]. It involves calcium and magnesium ion exchange (see Appendix 1 (A-2)).

(iii) The *pyrite* test case describes the dissolution of a pyrite rock in pure water. It has been used to test speciation algorithms [17]. Because it involves redox reactions, the stoichiometric coefficients cover a wide range, and the equilibrium constants vary over several orders of magnitude. This test is used under the assumption that no mineral phase is present (see Appendix 1 (A-3)).

(iv) The *MoMaS easy* test is the chemical system used for the reactive transport benchmark of MoMaS at the easy level [50]. It has been specifically developed to magnify numerical difficulties in a small system (see Appendix 1 (A-4)).

(v) The *Morel-Morgan* test is the first large chemical system reported in the computational literature. It was used by F. Morel and M. Morgan in 1972 to present the capacities of the computational method they had just developed (and which we still use today). This test includes 52 components (H^+, 20 metals, and 31 ligands), leading to 781 aqueous species (see Appendix 1 (A-5)).

(vi) The *MoMaS medium* test is the chemical system for the medium level of the MoMaS reactive transport benchmark [50] (see Appendix 1 (A-6)).

(vii) The *Fe-Cr* test is an additional redox test that describes the redox reactions between iron and chromium. These types of reactions occur when iron reactive barriers are used to treat chromium-contaminated sites [51, 52]. In this case, we consider only the aqueous phase without minerals (see Appendix 1 (A-7)).

(viii) The *pyrite mineral* test describes the dissolution of a pyrite rock in pure water. We assume that three possible mineral phases are present (see Appendix 1 (A-8)).

(ix) The *MoMaS hard* test is the equilibrium part of the chemical system described in the hard level of the MoMaS reactive transport benchmark. It allows for the formation of two mineral species (see Appendix 1 (A-9)).

(x) The *Fe-Cr mineral* test describes the redox reaction between iron and chromium. We assume the formation of three different mineral phases (see Appendix 1 (A-10)).

2.4 Test procedure

Equation 11 shows that we can obtain multiple linear systems from one chemical problem by changing the activity values of the components. For each chemical system, we select three components and vary their values over a wide range. The concentrations of all minerals are arbitrarily set to 10^{-3} mol L^{-1}. The activity of component H^+ is varied from 10^{-12} to 10^{-2} mol L^{-1} (pH = 12 to pH = 2), while that of component e^- is varied from 10^{-19} to 10^{12}, corresponding to Eh = −0.7 to 1.1 V computed using Eq. 14 at 25 °C

$$\mathrm{Eh} = \ln\left\{e^-\right\}\frac{RT}{F} \tag{14}$$

where T is the temperature (Kelvin), R is the gas constant (8.314 J K mol^{-1}), and F is the Faraday constant (96,487 C mol^{-1}). This range of electrical potential corresponds to the stability of water at pH values between 2 and 12. For the O_2 component, it is not possible to cover the same potential range as e^- because of the computation of the reference solution. The activity is varied from 10^{-70} to 10^4, as computed using Eq. 15 at 25 °C with $E^0 = 1.23$ V and pH varying from 2 to 12. The potential is then varied from −0.5 to 1.1 V

$$\mathrm{Eh} = E^0 + \frac{1}{4}\frac{RT}{F} \times \ln\frac{\{O_2\}\left\{H^+\right\}^4}{\{H_2O\}} \tag{15}$$

The activities of the other components vary from 10^{-12} to 10^{-1} mol L^{-1}. For each of the three selected components, we compute 30 values equally distributed on a log scale over the chosen range, leading to 29,791 different linear systems for each chemical test case. For each of these 29,791 tests, we make only one linear solver (or one Newton step) (except in the last section, Section 4, where the iterative Newton method is performed to solve the non-linear system given by Eqs. 8 and 9).

The matrix norm used in this work is the $||-||_1$ norm, defined as [23]

$$\|Z\|_1 = \max_{1\le j\le n}\left(\sum_{i=1}^{n}\left|Z_{i,j}\right|\right) \tag{16}$$

The condition number of Z is defined [23] as the product of the norm of the matrix per the norm of the inverse matrix (17)

$$\mathrm{cond}\,(Z) = \|Z\|_1 \times \left\|Z^{-1}\right\|_1 \tag{17}$$

To test the numerical methods, we first evaluate the computation time (CPU time) required to solve the linear system. Because we work with a very small matrix, the computations are very fast and we run the same calculation several times to obtain a total computing time of approximately 1 s. The *CPU time* is given in this work in units of seconds per computation (by dividing the total computing time by the number of runs). According to this method, the global computing time for one test case is approximately 6 days.

Many numerical methods, including a *failure indicator*, which indicates the success or failure of the resolution, have been developed. If needed, we include a failure indicator. As *failure*, we include the *crash* of the method, underflow or overflow, non-convergence within the maximum number of iterations (for iterative methods), or excessive inaccuracy for some advanced methods (LAPACK routines) that estimate the accuracy of the proposed solution.

Solving a linear system (13) using a numerical method produces an approximate solution ($d\xi_{\mathrm{method}}$), and the reference method gives ($d\xi_{\mathrm{ref}}$) with accuracy on the same order as the roundoff error. To evaluate the accuracy of the approximate solution, two quantities can be calculated:

1. The *relative error on the norm*, $\mathrm{Err}_{\mathrm{Norm}}$, is obtained by computing the norm of the approximate and reference solution (18)

$$\mathrm{Err}_{\mathrm{Norm}} = \frac{\left|\|d\xi_{\mathrm{method}}\| - \|d\xi_{\mathrm{ref}}\|\right|}{\|d\xi_{\mathrm{ref}}\|} \tag{18}$$

1. The error on the direction is given by $angle_{method}$, the angle (degrees) between the reference and the approximate solution calculated using the scalar product of these two vectors

$$\mathrm{angle}_{\mathrm{method}} = \frac{360}{2\pi}Arc\cos\left(\frac{d\xi_{\mathrm{method}}\cdot d\xi_{\mathrm{ref}}}{\|d\xi_{\mathrm{method}}\|\cdot\|d\xi_{\mathrm{ref}}\|}\right) \tag{19}$$

All of these quantities, namely the failure indicator, relative error on the norm, $\mathrm{angle}_{\mathrm{method}}$, and CPU time, are calculated for the 29,791 linear systems built from each chemical test case for all the tested methods. This enormous amount of data is aggregated in two ways:

(i) For each chemical system and each method, we compute the mean of each quantity.

(ii) For each chemical system and each method, the interval of the condition number is discretized into 100 regular subintervals. For each subinterval, we compute the mean of each quantity.

2.5 Reference solution

Because of the very high condition numbers, it is not possible to directly obtain an exact solution. We equilibrate the rows and columns of the Jacobian matrices to reduce their condition number using the iterative algorithm proposed by Knight et al. [53] because it preserves the symmetry of the Jacobian matrix.

Let $\widetilde{\mathbf{Z}}^k$ be the equilibrated Jacobian matrix at iteration k, with $\widetilde{\mathbf{Z}}^0 = Z$.

These authors defined r_i^k as the vector formed by the ith row of $\widetilde{\mathbf{Z}}^k$ and c_i^k as the vector formed by the ith column. The preconditioning matrices $\mathbf{R}^k$ and $\mathbf{C}^k$ are then defined by

$$\mathbf{R}^k = \mathrm{diag}\left(\frac{1}{\sqrt{\|r_i^k\|_\infty}}\right)_{i=1,Nx+NcP} \quad \text{and} \quad \mathbf{C}^k = \mathrm{diag}\left(\frac{1}{\sqrt{\|c_i^k\|_\infty}}\right)_{i=1,Nx+NcP} \tag{20}$$

The equilibrated matrix is defined at iteration $k+1$ by

$$\widetilde{\mathbf{Z}}^{k+1} = \mathbf{R}^k \cdot \widetilde{\mathbf{Z}}^k \cdot \mathbf{C}^k \tag{21}$$

This procedure is repeated until all $\|r_i^k\|_\infty$ and $\|c_i^k\|_\infty$ are equal to 1 or after 50 iterations. Let $\mathbf{R}$ and $\mathbf{C}$ be the resulting preconditioning matrices and $\widetilde{\mathbf{Z}}$ the equilibrated matrix. Instead of solving the linear system (12), we solve

$$\widetilde{\mathbf{Z}} \cdot \widetilde{\mathbf{x}} = -\widetilde{\mathbf{Y}} \tag{22}$$

where $\widetilde{\mathbf{x}} = \mathbf{C}^{-1} \cdot (\Delta\xi, \Delta Cp)$ and $\widetilde{\mathbf{Y}} = \mathbf{R} \cdot Y$. These procedures are coded using quadruple-precision reals. The linear system (22) is solved by LU decomposition coded with quadruple-precision real.

Even if the condition numbers of the Jacobian matrices (Z) are very high ($10^{213.9}$ for the Fe-Cr mineral test case), the condition numbers of the equilibrated matrices ($\widetilde{\mathbf{Z}}$) are much lower: the maximum condition number obtained after equilibration is $10^{13.4}$. According to Golub and van Loan [54], if the unit roundoff is approximately 10^{-d} and the condition number is approximately 10^q, then the Gaussian elimination gives a solution with approximately $d - q$ correct digits. Because we use quadruple precision, we obtain $d = 32$, leading to $32 - 14 = 18$ correct digits. One can then assume that the reference solution is exact if we compare it to the solutions produced by the tested methods (computed using double-precision real).

2.6 Selected numerical methods for solving linear systems

Studies on linear algebra [19, 23] present methods for solving linear systems as direct or iterative methods. Historically, speciation codes solved linear systems using direct methods, such as Gaussian elimination [34] or LU decomposition [17, 40, 42]. In its actual form, the speciation code SPECY [48] uses unsymmetric multifrontal (UMF) [55] as the linear solver. To the best of our knowledge, no speciation code uses iterative methods to solve linear systems. This point is in accordance with the existing literature, which reports the use of iterative methods for solving large, sparse linear systems [20–22, 24, 26, 28, 29, 56, 57]. Nevertheless, actual developments in speciation codes involve the use of large chemical databases [39, 58, 59], leading to an increase in the size of the chemical systems. The use of iterative methods is also studied in this work.

We select some direct and iterative solvers according to the properties of the linear systems and the speciation computation methods currently in use (Table 1).

For the direct method, we select LU decomposition [60] because it was originally used for speciation computations by Westall [40] and Westall et al. [42]. The UMF method [55] has been implemented in the speciation code SPECY [48] in place of the LU approach [17]. After showing that the Jacobian matrix is symmetric, we test the DSYTRS subroutine from LAPACK [61], which is based on a UDU decomposition. Because the Jacobian matrix is often positive definite, as shown in Table 3, we test the DPOTRS subroutine [61] based on the Cholesky method. Some authors [32] have used iterative QMRCGStab to solve reactive transport under a global approach. Here, we test QR decomposition using the DGELS routine [61].

For the iterative methods, we test the Jacobi [23, 62], Gauss-Seidel [23, 62], and successive over-relaxation (SOR) [23, 62] methods. Barrett et al. [21] proposed an algorithm to select an iterative solver depending on the matrix properties. GMRES was presented as the least selective algorithm. We use a GMRES method developed by HSL [63]. If the matrix is symmetric, Barrett et al. [21] recommend the use of conjugate gradient squared (CGS) or biconjugate gradient stabilized (BiCGStab) methods. CGS and BiCGStab subroutines have been developed by HSL. We test two additional methods devoted to symmetric matrices: SYMMBK [63] and an incomplete Cholesky (Inc. CHOLESKY) factorization [63].

We use the same parameters for all iterative methods: a maximum of 500 iterations and a stopping criterion of 10^{-8}. To determine the influence of the stopping criterion, we test the GMRES method using 50,000 maximum iterations and 10^{-12} as the stopping criterion, denoted by GMRES 10^{-12} in this study. A critical point of the GMRES algorithm is the size of the Hessenberg matrix. In this work, we set it to the max of 8 (Nx + NcP).

The results obtained using the Jacobi and SOR methods are not detailed here. As previously reported [19], the Jacobi method is inefficient, leading to a very high failure ratio (close to 100 %) even for the easiest test cases. For the SOR

Table 1 List of the selected solvers

Name	Source	Method	Matrix properties
Direct			
LU	[60]	LU decomposition	–
DGETRS	[61]	LU decomposition	–
UMF	[55]	Direct multifrontal	–
DSYTRS	[61]	UDU-factored symmetric matrix	Symmetric
DPOTRS	[61]	Cholesky A $= U^{\mathrm{T}} \times U$	Definite positive
DGELS	[61]	QR decomposition	
LU QUAD	[60]	LU decomposition quadruple precision	–
Iterative			
SYMMBK	[63]	Iterative SYMMBK HLS_MI02	Symmetric
Inc. CHOLESKY	[63]	Incomplete Cholesky HSL_MI28	Symmetric
CGS	[63]	Conjugate gradient squared HLS_MI23	–
BiCGStab	[63]	Biconjugate gradient squared stabilized HLS_MI26	–
GMRES	[63]	Flexible GMRES HLS_MI15	–
Gauss-Seidel	[60]	Gauss-Seidel method	–
Preconditioned			
LU Equil	[53–60]	LU and matrix equilibration	–
DGESVX	[61]	LU and optional preconditioning	–
GMRES Equil	[53–63]	GMRES and matrix equilibration	–
GMRES 1.d-12	[63]	GMRES convergence criteria 1.d-12	–

method [23, 26, 56, 62], the over-relaxation parameter is the key factor. Unfortunately, we did not find any efficient relationships to define it. For the same chemical system, the best value varies from 0.097 to 1.91 without apparent order.

We do not extensively test the possibility of using a preconditioner. As stated by Barrett et al. [21]: "Since using a preconditioner in an iterative method incurs some extra cost, both initially for the setup, and per iteration for applying it, there is a trade-off between the cost of constructing and applying the preconditioner, and the gain in convergence speed". In our case, the matrices are very small, leading us to suppose that this trade-off would not be advantageous. Nevertheless, an easy way to test preconditioners is proposed by the LAPACK routine DGESVX, which performs LU decomposition and matrix equilibration depending on the estimated condition number. We implement matrix equilibration according to Knight et al. [53] to obtain a reference solution. We test this preconditioning technique associated with LU decomposition and the GMRES method, denoted by LU Equil and GMRES Equil in this study. The maximum iterations allowed for the equilibration procedure is fixed to 5, according to the recommendations of Knight et al.

Table 2 Structure of the Jacobian matrix

Z	$\frac{\partial}{\partial \xi_k}$	$\frac{\partial}{\partial \left[Cp_{k-N_X} \right]}$
$\frac{\partial Y_j}{\partial}$	$\sum_{i=1}^{N_C} a_{i,j} \cdot a_{i,k} \cdot \left[C_i^n \right]$ $\sum_{i=1}^{N_C} a_{i,j} \cdot a_{i,k} \cdot \left[C_i^n \right]$	$ap_{k-N_X,j}$
$\frac{\partial Y_j}{\partial}$	$ap_{k,j-N_X}$	0

Table 3 Properties of the 10 chemical test cases ranked by increasing the maximal condition number

	Nx	Nc	NcP	Z size	cond(Z) min	cond(Z) max	cond(Z) max after 50 equilibration	%Z diagonal dominant	%Z positive definite
Gallic acid	3	17	0	3	$10^{0.61}$	$10^{12.6}$	$10^{0.95}$	18.4	100
Valocchi	5	7	0	5	$10^{0.49}$	$10^{15.3}$	$10^{0.65}$	67.7	100
Pyrite	4	40	0	4	$10^{4.06}$	$10^{24.9}$	$10^{0.95}$	0.00	100
MoMaS easy	5	12	0	5	$10^{3.44}$	$10^{37.7}$	$10^{1.05}$	0.00	71.1
Morel-Morgan	52	781	0	52	$10^{43.4}$	$10^{60.7}$	$10^{1.13}$	0.00	35.9
MoMaS medium	5	14	0	5	$10^{5.88}$	$10^{103.9}$	$10^{0.95}$	0.00	78.8
Fe-Cr	7	39	0	7	$10^{9.46}$	$10^{113.6}$	$10^{1.05}$	0.00	68.9
Pyrite mineral	4	43	3	7	$10^{1.71}$	$10^{33.1}$	$10^{3.19}$	0.00	0.00
MoMaS hard	6	15	2	8	$10^{5.45}$	$10^{123.9}$	$10^{3.02}$	0.00	0.00
Fe-Cr mineral	7	43	3	10	$10^{8.67}$	$10^{213.9}$	$10^{13.4}$	0.00	0.00

Finally, we test an LU decomposition method compiled as quadruple precision, denoted by LU QUAD. The source of this method is the LU double-precision real of numerical recipes [60], and we adapt it to quadruple precision. Because the usual computations are performed using double precision, the quadruple precision ($d\xi_{\text{QUAD}}$) should be translated in double-precision real. To avoid overflow, we rescale $d\xi_{\text{QUAD}}$ to ensure its validity. If huge $(1.d0)$ is the highest double-precision real represented by the machine, we rescale $d\xi_{\text{QUAD}}$ to obtain the double-precision solution $d\xi_{\text{LU QUAD}}$:

$$d\xi_{\text{LU QUAD}} = \frac{\text{huge}\,(1.d0)}{\max\left|d\xi_{\text{QUAD}}\,(i)\right|} \cdot d\xi_{\text{QUAD}} \tag{23}$$

In this way, we conserve the direction of the Newton step, even if its norm is changed.

3 Results and discussion

3.1 Properties of the Jacobian matrices

As defined by Eq. 11, the Jacobian matrix has several properties:

(i) The matrix is block-structured, as presented in Table 2. A four-block structure is present if precipitation occurs.

(ii) The matrix is symmetric, as shown in Table 2.

(iii) In the case of no precipitation, all the diagonal terms of the matrix are strictly positive because they are the sum of $a_{i,j}^2\,[C_i]$. It is then possible for the matrix to be diagonal dominant. We examine this possibility for the selected test case. Table 3 shows the ratio of diagonal dominant Jacobian matrices for all the chemical tests performed according to the previously defined test procedure. Some matrices in the gallic acid and Valocchi cases are diagonal dominant, but none of the matrices from the other cases are diagonal dominant. By plotting the ratio of diagonal dominant matrices depending on the condition number (see Appendix 2 (B-1)), it appears that only matrices with very low condition numbers can be diagonal dominant.

(iv) Because the Jacobian matrix is real, symmetric, and sometimes diagonal dominant, the question of whether it is positive definite may be posed. In the case of no precipitation, Eq. 11 can be written in matrix form, leading to Eq. 24

$$Z = A^T \cdot \text{diag}\,(C) \cdot A \tag{24}$$

Because the concentrations are positive, the Jacobian matrix is analytically positive definite. Nevertheless, this may not be true numerically. We are not able to propose a general framework, but we can compute the eigenvalues of the Jacobian matrix and test whether they are positive for all test cases. Table 3 shows that for the gallic acid, Valocchi, pyrite, and Morel-Morgan test cases, all the Jacobian matrices are positive definite. For the MoMaS easy, MoMaS medium, and Fe-Cr test cases, a large proportion (66.4 to 74.1 %) of the Jacobian matrices are positive definite. For cases including minerals (pyrite mineral, MoMaS hard, and Fe-Cr), essentially none of the matrices are positive definite (only 0.1 % for the MoMaS hard test). Plotting the ratio of positive definite matrices as a function of the condition number (see Appendix 2 (B-2)) shows that the chemical conditions are more important than the condition number when determining whether the Jacobian matrix is diagonal dominant.

(v) According to the test procedure presented previously, we plot, on the same graph, the logarithm of the norm of ||Y|| and the logarithm of the condition number of the matrix Z (Fig. 1). There is a strong linear relationship between these parameters. Moreover, the linear relationship does not depend on the chemical test, only on the existence of minerals. According to our results, the conditioning of the Z matrix can be evaluated using the following empirical formulas:

$$\begin{aligned} \text{cond}\,(Z)_{\text{no mineral}} &= 10^{5.30\pm0.03} \times \|Y\|^{0.9374\pm0.0008} \\ \text{cond}\,(Z)_{\text{mineral}} &= 10^{-3.23\pm0.08} \times \|Y\|^{1.706\pm0.002} \end{aligned} \qquad (25)$$

The value and uncertainties are obtained through the least squares method over all cond(Z) and $||Y||$. In this way, we propose an estimation of cond(Z) with no computation time cost because the objective function is evaluated during the Newton-Raphson procedure. As shown in Fig. 1, cond(Z) and $||Y||$ are strongly correlated for large condition numbers, and the results are noisier if cond(Z) and $||Y||$ are small. The evolution of this relation for low $||Y||$ can be seen in Appendix 7 (G-11). Therefore, Eq. 25 should not be used for $||Y||$ less than 10^{10}.

Several of these properties are obtained using the logarithm of the component activities as the primary unknown in Eq. 8. The historical approach [34] uses the component concentrations as the primary variable and leads to a less interesting Jacobian matrix. Even if the structure presented in Table 2 exists, the matrix is not symmetric. Moreover, the matrix is worse conditioned (condition number from $10^{11.2}$ to $10^{49.4}$ rather than $10^{4.06}$ to $10^{24.9}$ for the pyrite case). Finally, no specific relation exists between cond(Z) and $||Y||$ for the historical formulation.

As an example, we show one linear system from the Fe-Cr mineral test, corresponding to a condition number of 10^{187}. One can observe the structure of the matrix and the specificity of the linear system (26).

$$\begin{bmatrix} 1.15\cdot10^{94} & 9.09\cdot10^{93} & -5.04\cdot10^{-13} & -11.7 & 3.03\cdot10^{93} & 0 & 1.10\cdot10^{87} & 0 & 5 & -1 \\ & 5.45\cdot10^{93} & 0 & 1.14\cdot10^{10} & 1.82\cdot10^{93} & 0 & 4.11\cdot10^{86} & 2 & 3 & 0 \\ & & 1.00\cdot10^{-6} & 2.37\cdot10^{-15} & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 2.91 & 8.74 & 0 & 1.28\cdot10^{-6} & 1 & 0 & 0.25 \\ & & & & 6.06\cdot10^{92} & 2.23\cdot10^{-2} & 1.37\cdot10^{86} & 0 & 1 & 0.75 \\ & & & & & 2.85\cdot10^{-2} & 1.71\cdot10^{-4} & 0 & 0 & 0 \\ & & & & & & 1.37\cdot10^{86} & 0 & 0 & 0 \\ & & & & & & & 0 & 0 & 0 \\ & & & & & & & & 0 & 0 \\ & & & & & & & & & 0 \end{bmatrix} \cdot (d\xi) = \begin{pmatrix} -3.03\cdot10^{93} \\ -1.82\cdot10^{94} \\ -1.05\cdot10^{-13} \\ 8.99\cdot10^{-3} \\ -6.06\cdot10^{94} \\ -2.25\cdot10^{-2} \\ -1.37\cdot10^{86} \\ -27.6 \\ 180 \\ 3.84 \end{pmatrix} \qquad (26)$$

3.2 Robustness of the methods

Figure 2 presents the failure ratio for each method and each test case. The presence of minerals prevents the DPOTRS, Inc. CHOLESKY, and Gauss-Seidel methods from solving the system. If there are minerals present in the chemical system, a zero-value block appears in the Jacobian matrix, as shown in Table 2 and Eq. 26. This block makes the Inc. CHOLESKY factorization unappropriated. Because the Gauss-Seidel method requires division by each diagonal term, this zero-value block makes the method unadapted. The failure of the DPOTRS routine is explained by the properties of the Jacobian matrix. As shown in Table 3, there is no positive definite matrix in the presence of minerals. In the case of the DPOTRS, Inc. CHOLESKY, and Gauss-Seidel methods, the term failure is ambiguous. These methods are *expected to fail* and should not be used on systems with minerals. If there are no minerals, some matrices are not positive definite in the MoMaS easy, MoMaS medium, Morel-Morgan, and Fe-Cr tests. This explains the failure of the DPOTRS routine.

Some other methods (DGETRS, DSYTRS, DGELS, and DGESVX) present a substantial failure ratio, mainly for high condition number tests (MoMaS easy and Fe-Cr mineral). UMF, SYMMBK, and CGS are robust for the Fe-Cr mineral test but present significant failure ratios for lower-conditioned tests, such as MoMaS easy or pyrite mineral. Some methods adapted to symmetric matrices (DSYTRS and SYMMBK) are included in this class of weak methods.

The BiCGStab method has a very low failure ratio and fails only in the two difficult tests (MoMaS easy and Fe-Cr mineral). GMRES is the only successful iterative method.

Figure 2 shows that some methods are successful for all the test cases. The most successful direct method is LU, while the most successful iterative methods are GMRES and GMRES 10^{-12}. The quadruple-precision method LU QUAD is also successful, which is expected because the double-precision LU method is also successful. The use of an equilibration method as a preconditioner makes LU Equil and GMRES Equil successful.

As stated previously, we focus on the capacity of a method to produce a solution independent of its accuracy.

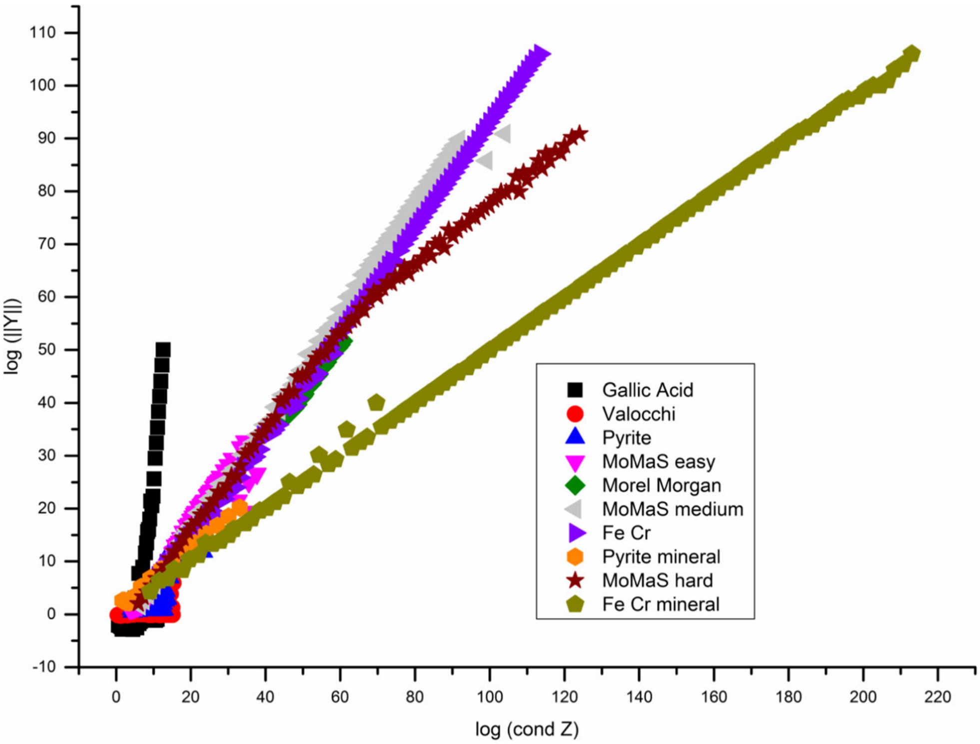


Fig. 1 Relationship between the condition number of Z and the norm of the objective function plotted on a log-log scale

For some advanced methods (e.g. LAPACK methods), a posteriori estimation of the residual and estimation of the condition numbers are performed. If the solution is not sufficiently accurate, no solution is given, leading to a higher failure ratio than for the more rustic methods (LU or Gauss-Seidel). Because the key point of this work—the resolution of a linear system—is included in the iterative Newton procedure, it is preferable to obtain an inaccurate solution (so

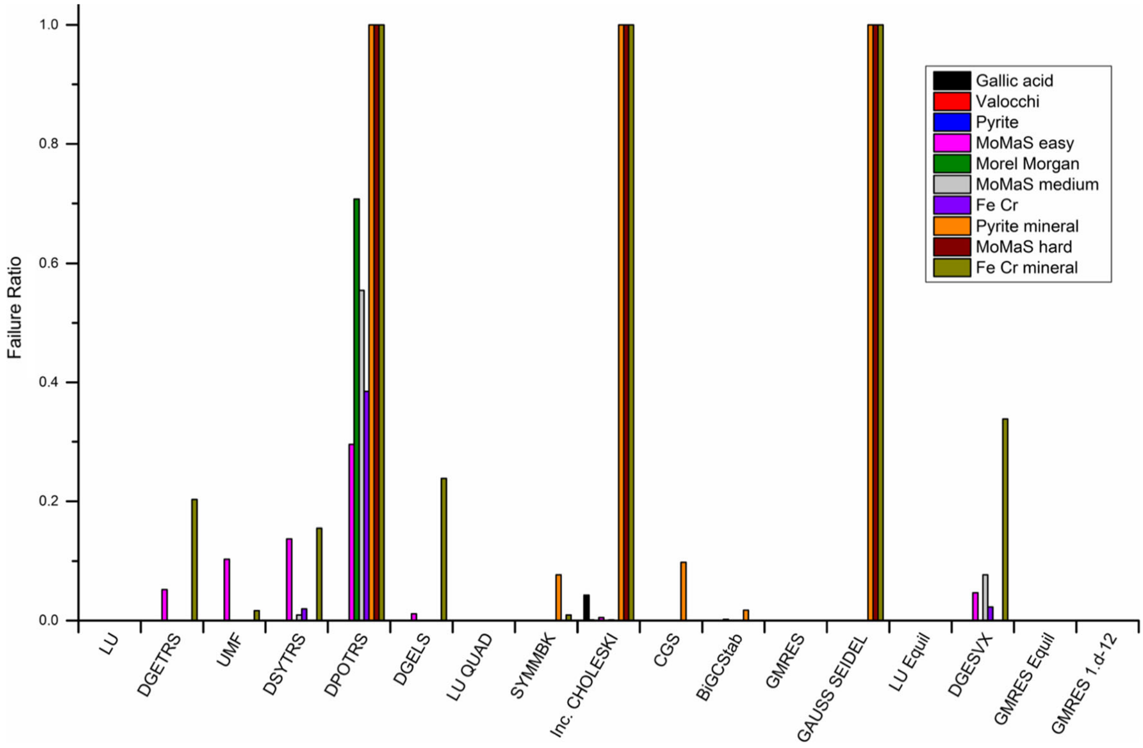


Fig. 2 Mean of the failure ratio for each method and each test case

the iterative procedure can be continued) than no solution (the iterative procedure will be aborted).

Appendix 3 presents the evolution of the failure ratio for each test case and each method depending on the condition number.

For the direct methods (Appendix 3 (C-1 to C-5)), for small condition numbers corresponding to the test cases gallic acid, Valocchi, and pyrite, no failure occurs. As the condition number increases, the failure ratio also increases for some methods. MoMaS easy (Appendix 3 (C-4)), MoMaS medium (Appendix 3 (C-6)), and Fe-Cr (Appendix 3 (C-7)) show that for condition numbers greater than 10^{20}, the failure ratio increases greatly for some of the methods. These methods are DOPTRS and DSYTRS for MoMaS medium and Fe-Cr. DGETRS, UMF, DSYTRS, DOPTRS, and DGELS present some failure for condition numbers greater than 10^{15} for the MoMaS easy case. In the presence of minerals (Appendix 3 (C-8 and C-9)), for low condition numbers (the pyrite mineral case), the methods are either successful (UMF, LU, DSYTRS, DGETRS) or completely unsuccessful (DPOTRS). For very high condition numbers (Fe-Cr mineral case), the success of the method does not depend on the condition number. We suppose that the condition numbers (see Table 3) are too high to exhibit any ordering.

For other iterative methods, the success does not depend on the condition number but on the nature of the matrix and the presence (Appendix 3 (C-18 to C-20)) or absence (Appendix 3 (C-11 to C-17)) of minerals.

3.3 Accuracy of the methods

The accuracy of the methods is evaluated in two ways: (i) the relative error on the norm (18) and (ii) the angle between the reference and the calculated solution (19).

(i) By plotting the mean of the logs of the relative error on the norm of each test case (Fig. 3), some general tendencies are identified. The relative residual tends to increase with the condition number of the system. For direct methods and small condition numbers, the relative residual is small (10^{-10} to 10^{-3}) for the gallic acid, Valocchi, and pyrite test cases. For the iterative methods, the relative residual corresponding to an accurate resolution for tests with small condition numbers is approximately 10^{-4}. This value corresponds to the value of the convergence criteria of the iterative methods. Iterative methods are more sensitive to the condition number than direct methods. Only the Valocchi test case is accurately solved by almost all the iterative methods, whereas the first three tests are accurately solved by all the direct methods. Even in the case of successful resolution (CGS and BiCGStab methods), the relative errors on the norm are high for intermediate cases (pyrite, MoMaS easy, and Morel-Morgan). Nevertheless, the results are better for the iterative methods than for the direct methods for the difficult tests (MoMaS easy, MoMaS medium, MoMaS hard, Fe-Cr mineral). The GMRES and Gauss-Seidel methods have mostly constant mean relative error on the norm, with the same accuracy for all test cases. GMRES and Gauss-Seidel are less efficient than the other methods for the easy tests, but more ill-conditioned tests are better solved by these two methods.

The condition numbers are so high that even LU QUAD cannot provide accurate resolution. For the MoMaS medium and Fe-Cr mineral tests, many of the solutions calculated by the LU QUAD method are rescaled using Eq. 23, leading to excessively high relative error on the norm.

Comparison of the relative error on the norm given by the non-preconditioned (LU, DGETRS, and GMRES) and preconditioned (LU Equil, DGESVX, and GMRES Equil) methods shows that the preconditioned methods lead to lower relative error than the non-preconditioned methods for the direct methods, but the result is more case-dependent for GMRES. The use of preconditioning usually leads to lower relative error on the norm, except for the Morel-Morgan, Fe-Cr, and MoMaS hard cases.

Increasing the maximum number of iterations and reducing the convergence criteria of GMRES leads to less relative error on the norm, but this reduction is not significant.

Nevertheless, the global means of the logs of relative errors on the norm hide the influence of the increasing condition number. Appendix 4 presents the evolution of the relative error on the norm for each test case and each method depending on the condition number. The theoretical behaviour is verified for the direct methods and for all the test cases (except for the Valocchi one, Appendix 4 (D-2)). The relative error on the norm increases regularly with the condition number. It is close to 10^{-16} when the condition number is close to 1 and increases to 1 when the condition number is close to 10^{16}, in accordance with the computation theory presented by Golub and van Loan [54]. For condition numbers greater than 10^{16}, the evolution of the relative error on the norm with the condition number is much noisier. The use of the quadruple-precision LU QUAD method leads to an accurate resolution of a large portion of the tested systems. As expected by computation theory, all the systems with condition numbers less than 10^{32} are solved with a relative error on the norm of approximately 10^{-15}. In some cases (MoMaS medium (Appendix 4 (D-6)), Fe-Cr (Appendix 4 (D-7)), MoMaS hard (Appendix 4 (D-9))), LU QUAD produces an increasing relative error

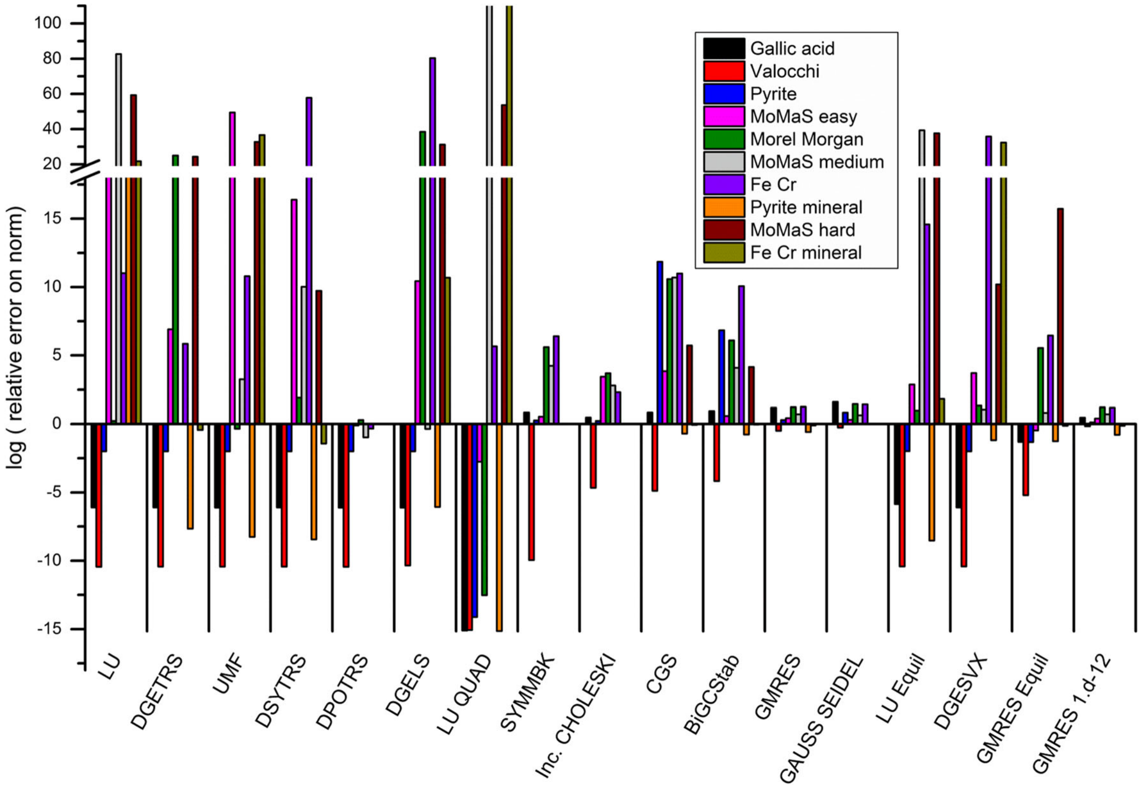


Fig. 3 Mean of the logs of the relative error on norm for each method and each test case

with increasing condition number (if higher than 10^{32}) but not systematically. LU QUAD produces a very low relative error on the norm even if the condition number is very high (Appendix 4 (D-9)). This behaviour can be explained by the fact that the LU QUAD method and/or the reference method is unable to exactly solve such ill-conditioned systems. LU QUAD produces a very high relative error on the norm, one point with 10^{290} error for the MoMaS medium (Appendix 4 (D-6)), and all the values at condition numbers greater than 10^{90} for the Fe-Cr mineral (Appendix 4 (D-10)) test case. These points correspond to the rescaling of the computed quadruple-precision solution to maintain it on the double-precision scale (using Eq. (23)).

Iterative methods present similar behaviour to direct methods, giving very low relative error on the norm (between 10^{-15} and 10^{-8}) when the condition number is less than a critical value. This critical value depends on the method and the test case. It can be set to 10^{8} for SYMMBK CGS, BiCGStab, and GMRES for the gallic acid (Appendix 4 (D-11)) and MoMaS easy (Appendix 4 (D-14)) cases. It can be set to 10^{12} or 10^{15} for Inc. CHOLESKY for the gallic acid and MoMaS easy cases and for SYMMBK, Inc. CHOLESKY, CGS, BiCGStab, and GMRES for the pyrite (Appendix 4 (D-13)), the MoMaS medium (Appendix 4 (D-16)), and MoMaS hard (Appendix 4 (D-19)) tests. Using low convergence criteria (GMRES 1.d-12) leads to lower relative error on the norm for low condition numbers (Appendix 4 (D-21 to D-23, D-26 to D-29)), but no significant improvements are obtained if the condition number increases, as shown in Appendix 4 (D-24 to D-30).

Using preconditioning methods reduces the relative error on the norm for intermediate condition numbers. No gain is obtained for low condition numbers (Appendix 4 (D-21 and D-22)), but the errors given by LU Equil, DGESVX, and GMRES Equil are less than the LU and GMRES errors for higher condition numbers (Appendix 4 (D-24 to D-26)). For very high condition number tests (Appendix 4 (D-27, D-29, and D-30)), the errors given by the preconditioned methods are equivalent to the errors given by the non-preconditioned methods.

(ii) By plotting the angle between the reference solution and the calculated solution, we can compare the methods according to the computed direction (Fig. 4). Because the resolution of the linear system (13) represents one step in the iterative Newton procedure, this information is much more important than the norm of the step. A wrong norm can be corrected using line search methods [64], whereas modifying a wrong direction leads to additional iterations. Small condition number tests (gallic acid, Valocchi, pyrite, and pyrite mineral) are solved using direct methods with the right direction. If the condition number

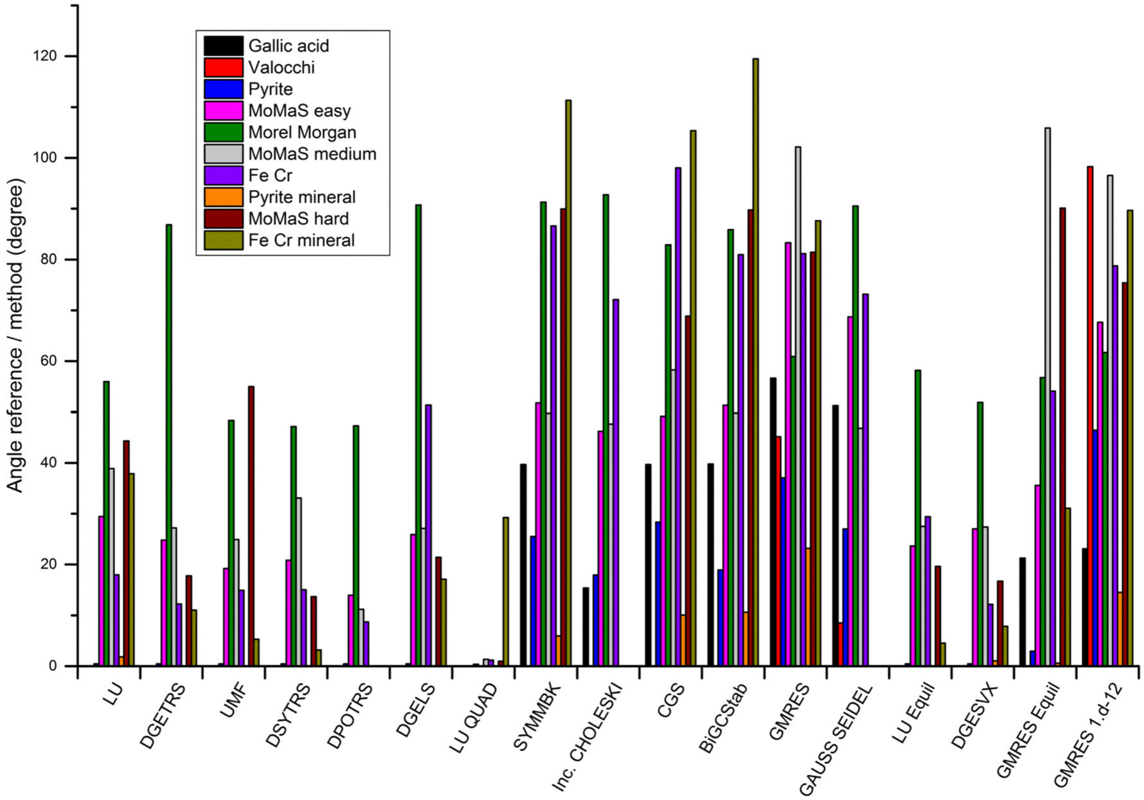


Fig. 4 Mean of the angles between reference and computed solution for each method and each test case

increases, the directions given by the direct methods become inaccurate, but the condition number is not the only governing parameter. Morel-Morgan leads to worse direction than MoMaS medium and Fe-Cr, and MoMaS hard leads to a higher angle than the Fe-Cr mineral test. Iterative methods result in a worse direction than direct methods, and only the Valocchi test case is solved with an accurate direction by all the iterative methods. Imposing lower convergence criteria (10^{-12}) on GMRES leads to a worse direction than using the usual criteria (10^{-8}). Using preconditioning methods leads to a better direction when associated with a direct method (LU Equil and DGESVX), but the conclusion is less clear for the iterative GMRES Equil method. Depending on the test case, the direction can be worse (Valocchi, MoMaS easy, MoMaS medium) or better (gallic acid, pyrite, MoMaS hard, Fe-Cr mineral)

The influence of the condition number on the angle (see Appendix 5) indicates that the direction is correct for direct methods when the condition number is less than 10^{15}. For iterative methods, the limit to obtain an accurate direction is a condition number less than 10^{8}, excepted for the Gauss-Seidel method, which produces wrong directions for low condition numbers. If the condition number increases, the behaviour of the direction becomes noisy. Since the relative error on the norm increases regularly until the condition number reaches the limit of 10^{8} or 10^{15}, the angle is accurately defined until this condition number limit is reached. Using preconditioned methods leads to a better direction for the LU Equil and the GMRES Equil methods when the condition number is higher than 10^{15} for some cases (Appendix 5 (E-21, E-23 to E-25, and E-30)) but to a worse direction for other cases (Appendix 5 (E-26 and E-29)).

We present two successful direct methods, LU and LU QUAD; one iterative method, GMRES (both tested versions, GMRES and GMRES 10^{-12}); and two preconditioned methods, LU Equil and the GMRES Equil. By comparing the relative error on the norm (Appendix 5 (D-21 to D-30)), the successful methods can be ranked from the lowest to highest error: LU QUAD, GMRES 10^{-12}, GMRES Equil, LU Equil, and LU. Ranking these methods according to the angle between the reference and computed solution is more complicated. For all the tests cases (Appendix 5 (E-21 to E-25, E-27, E-28, and E-30)), LU QUAD gives the best direction, followed by LU Equil, LU, GMRES Equil, and GMRES 10^{-12}. The MoMaS medium (Appendix 5 (E-26)) and MoMaS hard (Appendix 5 (E-29)) test cases lead to the same conclusion, except GMRES Equil which gives the worst direction.

3.4 Efficiency of the methods

The speed of the methods is studied by recording the computation time for each test case and plotting the mean CPU time for each test case and each method (see Fig. 5). As expected, the computation times are very short (less than 1 ms) because the systems to solve are small.

Figure 5 shows the influence of the system size. For all methods, the computation time increases with the number of unknowns. The results show that the iterative methods are less sensitive to the system size than the direct methods. For the iterative methods, the number of iterations is important and depends on the first guess and other factors. The slowest method is LU QUAD, for which a large amount of computation time is devoted to the translation of double-precision real to quadruple-precision real and back. Figure 5 also shows the computing time required to obtain the reference solution, which requires more time.

The UMF method is the slowest double-precision direct method, but its multifrontal block strategy becomes interesting for large systems. The resolution of the Morel-Morgan test requires 33 times more CPU time than the resolution of the MoMaS easy test for the UMF method, whereas it takes 190 times more time for the LU method.

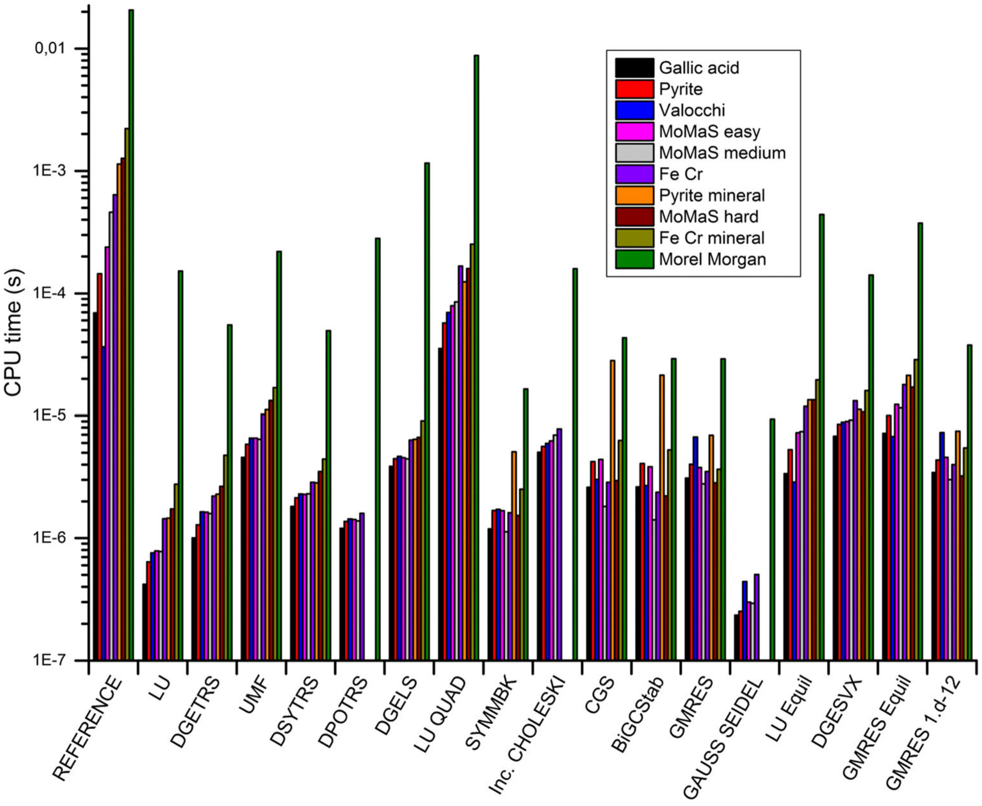
Among the iterative methods, the fastest is the Gauss-Seidel method and the slowest is the Inc. CHOLESKY method. The two most robust iterative methods, BiCGStab and GMRES, are rapid, sometimes more so than the direct robust methods, LU and UMF, especially for large systems (Morel-Morgan test case). GMRES is less case-dependent than BiCGStab, leading to similar computing time, regardless of the test case.

As expected, introducing preconditioning techniques (LU Equil, DGESVX, and GMRES Equil) or decreasing the convergence criteria for an iterative method (GMRES 10^{-12}) leads to increased computing time. The computing time for preconditioning does not depend only on the system's size: the Valocchi, MoMaS easy, and MoMaS medium test cases (system size of 5 ×5) are solved with the same computing time for all the direct methods, but their resolution when using LU QUAD Equil, LU Equil, and GMRES Equil is faster.

Appendix 6 shows the computation time (log scale) for each test case and each method depending on the condition number. Appendix 6 (F-1 to F-10) shows that, as expected, the computation time of the direct methods does not depend on the condition number of the system. The LU method is usually 10 times faster than the UMF method, except for the Morel-Morgan test case, in which LU is only 1.5 times faster.

In Appendix 6 (F-11 to F-20), the general tendency for the iterative methods is to require the same computation time, independent of the condition number. The oscillations presented by the curves seem to be not related to the condition number. For the test case without minerals, the Gauss-Seidel method is efficient. The two most robust methods, BiCGStab and GMRES, are often the third and

Fig. 5 CPU time (s) for each method and each test case



fourth fastest methods (Gauss-Seidel and SYMMBK are the fastest).

4 Proposal of a new algorithm

Based on our results, we propose an algorithm to optimize the resolution of a chemical system using a Newton-Raphson-like method.

Examining the failure ratio results, seven methods are eligible: LU and LU QUAD as direct methods, GMRES and Gauss-Seidel (if no minerals) as iterative methods, LU Equil and GMRES Equil as preconditioned methods, and the reference method (LU QUAD Equil).

Because these methods are included in a Newton minimization procedure, the most important accuracy criterion is the direction of the minimization, i.e. the angle between the reference and the calculated solution. The behaviour of this direction is strongly correlated with the condition number of the system and is correct if the condition number is less than the critical value and wrong if the condition number is greater than the critical value (see Appendix 5). The critical condition number is 10^8 for GMRES, 10^{16} for the double-precision direct methods, 10^{32} for LU QUAD, and case-dependent for preconditioned methods (10^{20} to 10^{60}). Gauss-Seidel leads to wrong directions for very low condition numbers (Appendix 5 (E-11 and E-12)).

Table 4 Algorithm for equilibrium computation

cond(Z)	Inversion method
$>10^{30}$	LU QUAD Equil
$10^{30} \geq \mathrm{cond}(Z) > 10^{14}$	LU QUAD
$10^{14} \geq \mathrm{cond}(Z) > 10^{4}$	LU ($\mathrm{Nx}+\mathrm{NcP}<10$) GMRES ($\mathrm{Nx}+\mathrm{NcP} \geq 10$)
$10^{4} > \mathrm{cond}(Z)$	LU

In terms of efficiency, the most rapid method is Gauss-Seidel when it is available. The second most efficient method is LU for small systems (less than 10 × 10) or GMRES for larger systems (more than 10 × 10), and the slowest method is LU QUAD. For small systems (less than 5 × 5), LU Equil is as fast as GMRES but becomes slower as the system size increases.

We recommend using LU, LU QUAD, GMRES, and the reference method LU QUAD Equil. Gauss-Seidel should be rejected because of its wrong direction, and equilibration does not sufficiently improve the behaviour of double-precision routines.

Using Eq. 25, it is possible to estimate the condition number of the system without additional computation. This estimation enables the selection of the best-adapted method depending on the system size and condition number.

The goal is to use the most robust method (LU QUAD with preconditioning) for high condition number systems (more than 10^{32}) in the first Newton-Raphson iterations. When the condition number is sufficiently decreased, the preconditioning becomes useless and LU QUAD can be used until the condition number is less than 10^{16}. Then, a faster method is used to obtain a coarse approximation of the solution, LU for small systems and GMRES for large systems (more than 10 × 10). To find the exact solution, the LU direct method is used.

We propose the algorithm presented in Table 4 and compare it with several inversion methods in a Newton-Raphson algorithm. The 10 chemical test cases are solved

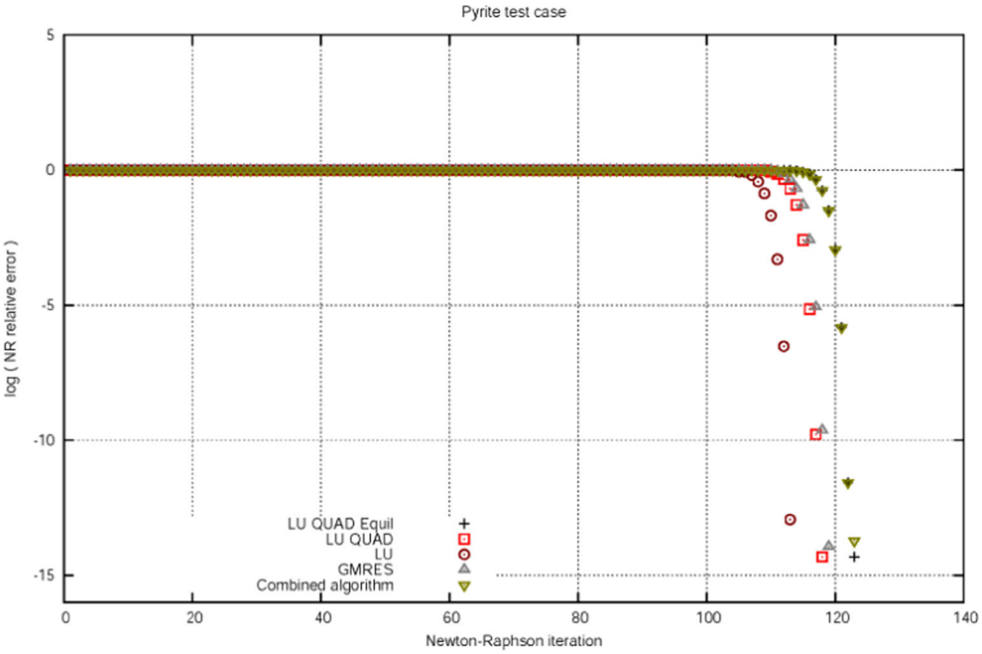
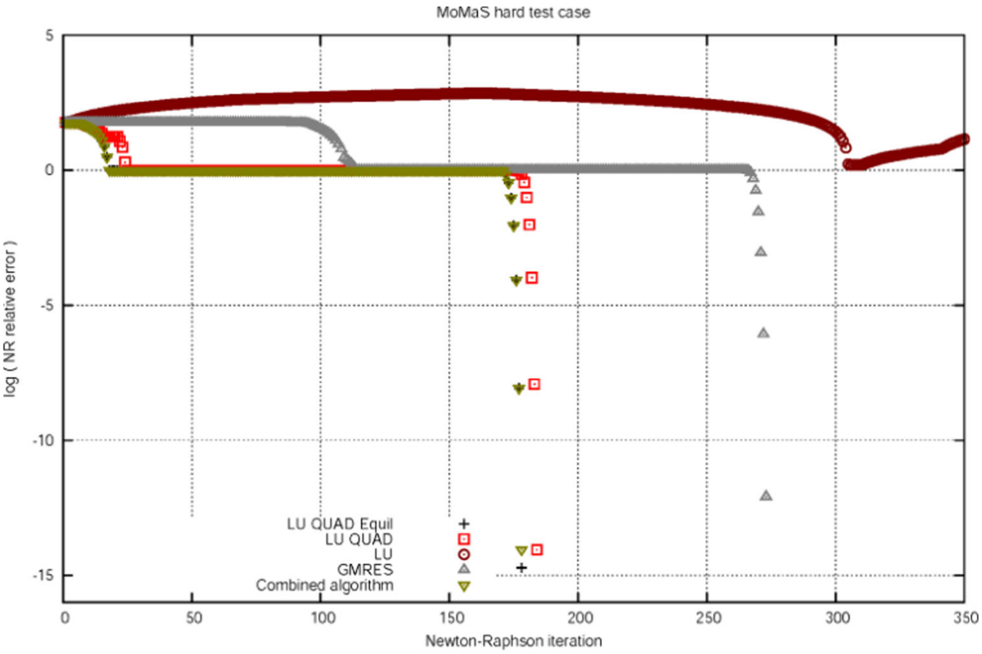


Fig. 6 Evolution of $\mathrm{NR}_{\mathrm{relativeerror}}$ as a function of the Newton-Raphson iteration for the pyrite test case

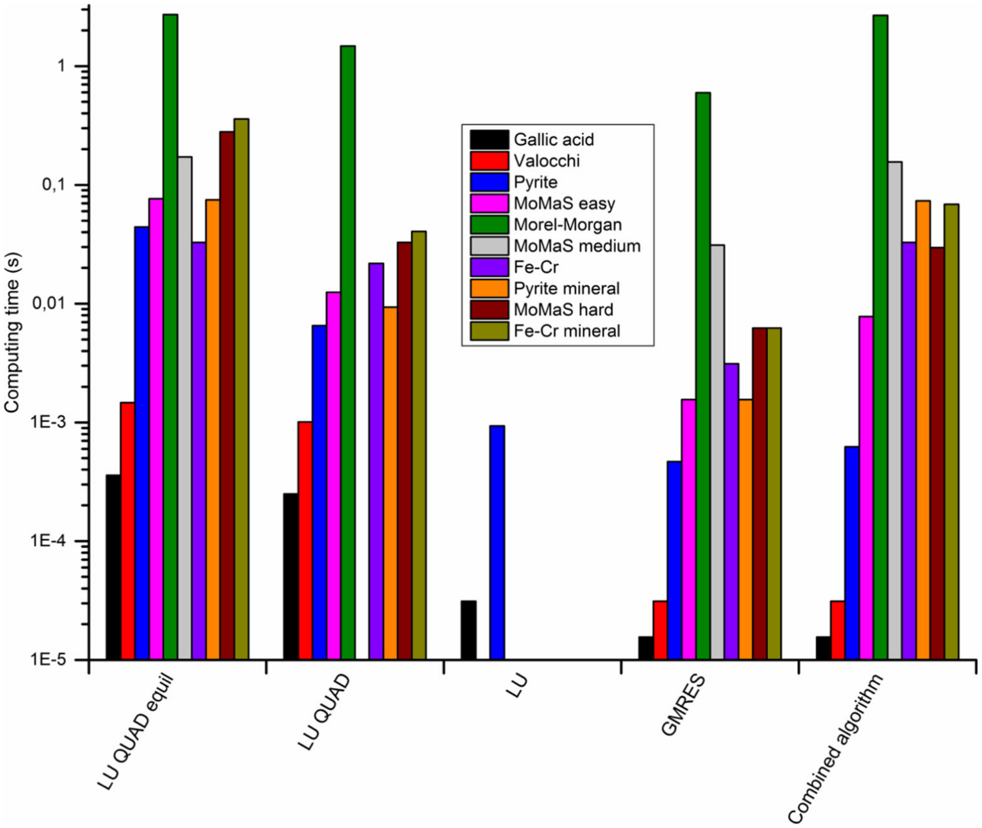
Fig. 7 Evolution of $NR_{relativeerror}$ as a function of the Newton-Raphson iteration for the MoMaS hard case



using the combined algorithm or one of the selected methods: LU QUAD Equil (used as the reference solution), LU QUAD, LU, and GMRES. Appendix 7 shows the evolution of the $NR_{relativeerror}$ (7) as a function of the Newton-Raphson iterations

Figure 6 shows that all the methods are equivalent for easy test cases (see Appendix 7 (G-1 to G-3)). Nevertheless, the use of LU inversion leads to non-convergence, even if the test is easy, as observed for the Valocchi test (Appendix 7 (G-2)). If the difficulty of the test increases, the lower

Fig. 8 Computation time (s) as a function of test case and algorithm



accuracy of GMRES (compared to the quadruple-precision routine used in LU QUAD Equil, LU QUAD, and the combined algorithm) leads to a greater number of Newton iterations, as shown in Fig. 7 for the MoMaS hard case. This point is confirmed for other cases (see Appendix 7 (G-4 to G-9)). For the Fe-Cr mineral case (see Appendix 7 (G-10)), only LU QUAD Equil and the combined algorithm can solve the problem. Other methods lead to non-convergence, due to overflow for the GMRES algorithm (overflow appears in the Newton algorithm and is not due to GMRES itself) and because LU QUAD and LU are unable to give an accurate descent direction.

Appendix 7 (G-11) shows the evolution of the relation between the norm of Y and the condition number of the Jacobian matrix during the minimization process. This figure is similar to Fig. 1, confirming the empirical relation (25). This relation cannot be used close to the solution, and the condition number tends to be a case-dependent limit for very low $||Y||$.

Nevertheless, the number of iterations is not the critical point. Because the time required by one iteration changes depending on the method used, we have to consider the total computation time. By plotting the total computation time required to solve each test case depending on the algorithm used (see Fig. 8), we can see that

(i) LU QUAD Equil, as expected, is the slowest. Nevertheless, this method allows the convergence of the Newton-Raphson method for all test cases.

(ii) LU QUAD is slightly faster. The difference between LU QUAD Equil and LU QUAD gives an indication of the time used for matrix equilibration. This time is greater for pyrite, MoMaS easy, pyrite mineral, MoMaS hard, and Fe-Cr mineral than for the other test cases.

(iii) LU is fast when it leads to convergence, but this method results in a very weak Newton-Raphson algorithm.

(iv) GMRES always results in the fastest Newton-Raphson algorithm. It has been shown (Fig. 7, Appendix 7 (G-8)) that the number of required iterations can be twice the number for other methods, but we show (Fig. 5) that the GMRES method is faster than the other methods.

(v) The proposed combined algorithm leads to intermediate computing times, equivalent to those of LU QUAD Equil and LU QUAD, depending on the case.

According to our results, GMRES should be systematically used because it is fast and usually leads to convergence of the Newton-Raphson algorithm. The combined algorithm should be used for very high condition numbers or for recomputing a failed run.

5 Conclusion

In this work, we focus on the resolution of small linear systems generated using the Newton-Raphson algorithm to solve equilibrium chemistry problems. For the first time, we propose a study of the condition number of such linear systems and find that the range of values covered is unusually large. This characteristic leads to specific numerical problems, with matrices that are quite small (approximately 10 ×10) but very badly conditioned (up to 10^{100}). Ten different chemical systems are studied.

There is a strong linear relationship between the logarithm of the condition number of the matrix and the logarithm of the norm of the objective function. This factor can be exploited to create efficient algorithms. This relation is strictly an empirical one and is not valuable for low condition numbers.

A wide variety of linear solvers have been tested, and several direct and iterative solvers are selected. Some of these solvers are specific for a class of matrix, symmetric or positive definite, while others are generic. A preconditioning method (matrix equilibration) has also been tested to reduce the conditioning of the systems.

According to our selected test cases, only the LU and LU QUAD direct methods, the GMRES iterative method, and LU Equil and GMRES Equil preconditioned methods are sufficiently robust to solve all the tests.

According to the size of the chemical tests, the LU method is faster than the GMRES method. However, our results for the Fe-Cr mineral and Morel-Morgan cases show that GMRES is preferable for larger chemical systems (more than 10 components). Chemical systems with more than 10 components have not been frequently modelled in the past decade. However, the use of geochemical databases makes the construction of large geochemical systems easier, and the increase in computation capacities makes it possible. For very large geochemical systems, we recommend the GMRES method.

We also propose using the linear relationship between the condition number of the Jacobian matrix and the norm of the objective function to develop an efficient algorithm.

The classic LU method is not a good choice. Its weakness is its low robustness for challenging test cases. We recommend using the GMRES method, which is fast and usually leads to convergence of the Newton-Raphson algorithm. For very high condition numbers (more than 10^{100}), we recommend the most robust LU QUAD Equil method. When the Newton-Raphson method is sufficiently near the solution to decrease the condition number, the faster GMRES method can be used. By using the linear relationship between $\mathrm{cond}(Z)$ and $||Y||$, the transition between the two methods can be achieved without computing the condition number (which is very expensive).

This work explores a new research field by studying geochemical computation from a condition number point of view. We attempted to benchmark a wide variety of linear solvers, but it was not possible to explore the flexibility of all the tested solvers. This study will help us to eliminate some solvers so that our future work can focus on the most promising: LU, LU QUAD, GMRES, LU Equil, and GMRES Equil. Some points for future exploration are as follows:

(i) We did not extensively test the robustness and the efficiency of the Newton-Raphson algorithm. Further work should examine the influence of the initial Newton-Raphson guess to confirm our conclusions about the high efficiency of the GMRES method.

(ii) The accuracy of iterative methods depends on the value of the convergence criterion (which we set to 10^{-8}) and on the method used to check the convergence (we used the default method). Moreover, the efficiency can vary depending on the initial guess provided by the user. In this work, we used the easiest initial guess: the residual for the tests from the Newton-Raphson method and the previous Newton-Raphson step for the test in a Newton-Raphson algorithm. We believe that it is possible to make a better choice, markedly enhancing the efficiency of the iterative methods.

(iii) The GMRES method allows the use of left and/or right preconditioners. These preconditioners can increase the robustness, accuracy, and efficiency of the method. More generally, several classes of preconditioners that may reduce the condition number of the linear system can be used [65, 66]. In this work, we explored the use of one preconditioner: matrix equilibration. However, other classes of preconditioners may be more efficient.

(iv) Previous works have addressed the use of methods to solve geochemical equilibria other than the Newton-Raphson method [17, 44, 49, 67]. It has been shown [17] that an efficient algorithm can be obtained by combining a zero-order method with the Newton-Raphson approach.

(v) The size of the chemical tests presented here is representative of the sizes actually used in environmental studies. We have shown that the GMRES method may be efficient for large systems. In anticipation of future needs, it may be useful to test chemical systems larger than the Morel-Morgan system.

(vi) Part of the Newton minimization related to very large condition numbers (far from the solution) can be performed using *random* methods; GMRES is efficient even though its descent direction is not accurate for high condition numbers. Some methods, such as simulated annealing and particle swarm optimization, could be used in future research.

These factors should be explored in light of the results presented in this study. We proposed a large set of chemical tests, a criterion to determine the difficulty of these tests (the condition number), and a panel of numerical methods that should be studied preferentially.

As a more general consideration, the reader should pay particular attention to the old Morel-Morgan test case and the more realistic pyrite test case. The Morel-Morgan test uses Fe^{2+} and Fe^{3+}, Co^{2+} and Co^{3+}, and SO_4^{2-} and S^{2-} as components whereas the pyrite case uses O_2, Fe^{2+}, and SO_4^{2-}. The first studies on geochemical computation avoided redox problems. We show that redox problems lead to higher condition numbers because the stoichiometric coefficients and equilibrium constants cover a wider range. Several geochemical databases avoid the introduction of redox reactions. There is sometimes a good reason to not write redox reactions as equilibria (slow reaction rates, irreversible reactions) as done in Arora et al. [2]. However, the reason is sometimes numeric, and redox reactions are avoided because they lead to non-convergence.

We propose the use of quadruple-precision real for challenging chemical systems. In this work, the core of the geochemical code is conserved as double-precision real, and only the linear system tool is set as quadruple precision. Rewriting an entire geochemical code in a quadruple-precision format will result in robust code but at the cost of an important and rebarbative work as well as efficiency. In this stage of our research, we do not recommend such an effort because implementing LU decomposition using quadruple-precision real is very efficient, requiring only a minor modification of existing code and reducing the computation time.

Acknowledgments Hela Machat has been supported by a grant from the Tunisian Government. This work is supported by the BRGM-CNRS CUBICM project. We thank the reviewers for their helpful comments.

Compliance with Ethical Standards Compliance with ethical standards

Conflict of interests The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Walter, A.L. et al.: Modeling of multicomponent reactive transport in groundwater. 2. Metal mobility in aquifers impacted by acidic mine tailings discharge. Water Resour. Res. **30**(11), 3149–3158 (1994)
2. Arora, B. et al.: A reactive transport benchmark on heavy metal cycling in lake sediments Computational Geosciences (2014)
3. De Windt, L., Leclercq, S., Van der Lee, J.: Assessing the durability of nuclear glass with respect to silica controlling processes in a clayey underground disposal. In: 29th International Symposium on the Scientific Basis for Nuclear Waste Management XXIX. Materials Research Society Symposium Proceedings, Ghent; Belgium (2005)
4. Hoteit, H., Ackerer, P., Mose, R.: Nuclear waste disposal simulations: Couplex test cases. Comput. Geosci. **8**(2), 99–124 (2004)
5. Tompson, A.F.B., et al.: On the evaluation of groundwater contamination from underground nuclear tests. Environ. Geol. **42**(2-3), 235–247 (2002)
6. Andre, L., et al.: Numerical modeling of fluid-rock chemical interactions at the supercritical CO2-liquid interface during CO2 injection into a carbonate reservoir, the Dogger aquifer (Paris Basin, France). Energy Convers. Manag. **48**(6), 1782–1797 (2007)
7. Kang, Q., et al.: Pore scale modeling of reactive transport involved in geologic CO2 sequestration. Transp. Porous Media **82**(1), 197–213 (2010)
8. Navarre-Sitchler, A.K., et al.: Elucidating geochemical response of shallow heterogeneous aquifers to CO2 leakage using high-performance computing: implications for monitoring of CO2 sequestration. Adv. Water Resour. **53**(0), 45–55 (2013)
9. Pruess, K. et al.: Code intercomparison builds confidence in numerical simulation models for geologic disposal of CO2. Energy **29**(9-10), 1431–1444 (2004)
10. Regnault, O., et al.: Etude experimentale de la reactivite du CO2 supercritique vis-a-vis de phases minerales pures. Implications pour la sequestration geologique de CO2. Compt. Rendus Geosci. **337**(15), 1331–1339 (2005)
11. Valocchi, A.J., Street, R.L., Roberts, P.V.: Transport of ion-exchanging solutes in groundwater: chromatographic theory and field simulation. Water Resour. Res. **17**, 1517–1527 (1981)
12. Lichtner, P.C.: Continuum model for simultaneous chemical reactions and mass transport in hydrothermal systems. Geochim. Cosmochim. Acta **49**(3), 779–800 (1985)
13. Appelo, C.A.J.: Hydrogeochemical transport modelling. Proceed. Inf.—Comm. Hydrol. Res. TNO **43**, 81–104 (1990)
14. Yeh, G.T., Tripathi, V.S.: A critical evaluation of recent developments in hydrogeochemical transport models of reactive multi-chemical components. Water Resour. Res. **25**, 93–108 (1989)
15. Carrayrou, J. et al.: Comparison of numerical methods for simulating strongly nonlinear and heterogeneous reactive transport problems—the MoMaS benchmark case. Computational Geosciences **14**(3), 483–502 (2010)
16. Hammond, G.E., Valocchi, A.J., Lichtner, P.C.: Modeling multicomponent reactive transport on parallel computers using Jacobian-Free Newton Krylov with operator-split preconditioning. In: Hassanizadeh, S.M. (ed.) Developments in water science, computational methods in water resources, Proceedings of the XIVth International Conference on Computational Methods in Water Resources (CMWR XIV), pp. 727–734. Elsevier (2002)
17. Carrayrou, J., Mosé, R., Behra, P.: New efficient algorithm for solving thermodynamic chemistry. AIChE J. **48**(4), 894–904 (2002)
18. Amir, L., Kern, M.: A global method for coupling transport with chemistry in heterogeneous porous media. Comput. Geosci. **14**(3), 465–481 (2010)
19. Quarteroni, A., Sacco, R., Saleri, F.: Numerical mathematics. In: Marsden, J.E., Sirovich, L., Antman. S.S. (eds.) Texts in Applied Mathematics. 2nd edn. Springer, Heidelberg (2007)
20. Axelsson, O., et al.: Direct solution and incomplete factorization preconditioned conjugate gradient methods. Comparison of algebraic solution methods on a set of benchmark problems in linear elasticity, in STW report. 2000, Department of Mathematics, Catholic University of Nijmegen: Nijmegen, The Netherlands. pp. 1-36
21. Barrett, R., Berry, M., Chan, T.F., Demmel, J., Donato, J., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C., Van Der Vorst, H. Templates for the solution of linear systems: building blocks for iterative methods, 2nd edn. SIAM, Philadelphia (1994)
22. Gould, N.I.M., Hu, Y., Scott, J.A.: A numerical evaluation of sparse direct solvers for the solution of large sparse, symmetric linear systems of equations. 2005, Council for the Central Laboratory of the Research Councils
23. Allaire, G., Kaber, S.M. In: Marsden, J.E., Sirovich, L., Antman, S.S. (eds.): Numerical linear algebra. Texts in applied mathematics. Springer, New York (2008)
24. Baldwin, C., et al.: Iterative linear solvers in a 2D radiation-hydrodynamics code: methods and performance. J. Comput. Phys. **154**(1), 1–40 (1999)
25. Chao B.T., L.H.L., Scott, E.J.: On the solution of ill-conditioned, simultaneous, linear, algebraic equations by machine computation, in University of Illinois Bulletin. 1961, University of Illinois
26. Hadjidimos, A.: Successive overrelaxation (SOR) and related methods. J. Comput. Appl. Math. **123**(1-2), 177–199 (2000)
27. Kalambi, I.B.: A comparison of three iterative methods for the solution of linear equations. J. Appl. Sci. Environ. Manag. **12**(4), 53–55 (2008)
28. Klisinski, M., Runesson, K.: Improved symmetric and non-symmetric solvers for FE calculations. Adv. Eng. Softw. **18**(1), 41–51 (1993)
29. Schenk, O., Gartner, K.: Solving unsymmetric sparse systems of linear equations with PARDISO. Fut. Gener. Comput. Syst. **20**(3), 475–487 (2004)
30. Xue, X.J. et al.: A direct algorithm for solving ill-conditioned linear algebraic systems. JCPDS-Int. Centre Diffract. Data Adv. X-ray Anal. **42**, 629–633 (2000)
31. Pyzara, A., Bylina, B., Bylina, J.: The influence of a matrix condition number on iterative methods' convergence (2011)
32. Hoffmann, J., Kräsutle, S., Knabner, P.: A parallel global-implicit 2-D solver for reactive transport problems in porous media based on a reduction scheme and its application to the MoMaS benchmark problem. Comput. Geosci. **14**(3), 421–433 (2010)
33. Soleymani, F.: A new method for solving ill-conditioned linear systems. Opuscula Math. **33**(2), 337–344 (2013)
34. Morel, F., Morgan, J.: A numerical method for computing equilibria in aqueous chemical systems. Environ. Sci. Technol. **6**(1), 58–67 (1972)
35. Morel, F.M.M.: Principles of aquatic chemistry. Wiley Interscience, New York (1983)
36. De Windt, L. et al.: Intercomparison of reactive transport models applied to UO2 oxidative dissolution and uranium migration. J. Contam. Hydrol. **61**(1-4), 303–312 (2003)
37. Jauzein, M. et al.: A flexible computer code for modelling transport in porous media: impact. Geoderma **44**(2–3), 95–113 (1989)
38. Parkhurst, D.L., Appelo, C.A.J.: User's guide to PHREEQC (version 2)—a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. Water Resour. Invest., Editor. 1999: Denver. p. 312
39. Van der Lee, J.: CHESS another speciation and surface complexation computer code. E.d.M.d. Paris, Editor. 1993: Fontainebleau. p. 85

40. Westall, J.C.: MICROQL: a chemical equilibrium program in BASIC. Computation of adsorption equilibria in BASIC. S.F.I.o.T. EAWAG, Editor. 1979: Dübandorf. p. 42
41. Westall, J.C.: FITEQL ver. 2.1. 1982: Corvallis
42. Westall, J.C., Zachary, J.L., Morel, F.M.M.: MINEQL: a computer program for the calculation of chemical equilibrium composition of aqueous system. R.M.P. Laboratory, Editor. 1976: Cambridge. p. 91
43. Walter, L.J., Wolery, T.J.: A monotone-sequences algorithm and FORTRAN IV program for calculation of equilibrium distributions of chemical species. Comput. Geosci. **1**, 57–63 (1975)
44. Wigley, T.M.L.: WATSPEC: a computer program for determining the equilibrium speciation of aqueous solutions. B.G.R.G. Tech. Bull., Editor. 1977. p. 49
45. Jennings, A.A., Kirkner, D.J., Theis, T.L.: Multicomponent equilibrium chemistry in groundwater quality models. Water Resour. Res. **18**, 1089–1096 (1982)
46. Cederberg, A., Street, R.L., Leckie, J.O.: A groundwater mass transport and equilibrium chemistry model for multicomponent systems. Water Resour Res. **21**, 1095–1104 (1985)
47. Yeh, G.T., Tripathi, V.S.: A model for simulating transport of reactive multispecies components: model development and demonstration. Water Resour. Res. **27**(12), 3075–3094 (1991)
48. Carrayrou, J.: Looking for some reference solutions for the reactive transport benchmark of MoMaS with SPECY. Comput. Geosci. **14**(3), 393–403 (2010)
49. Brassard, P., Bodurtha, P.: A feasible set for chemical speciation problems. Comput. Geosci. **26**(3), 277–291 (2000)
50. Carrayrou, J., Kern, M., Knabner, P.: Reactive transport benchmark of MoMaS. Comput. Geosci. **14**(3), 385–392 (2010)
51. Fendorf, S.E., Li, G.: Kinetics of chromate reduction by ferrous iron. Environ. Sci. Technol. **30**(5), 1614–1617 (1996)
52. Chilakapati, A. et al.: Groundwater flow, multicomponent transport and biogeochemistry: development and application of a coupled process model. J. Contam. Hydrol. **43**(3-4), 303–325 (2000)
53. Knight, P., Ruiz, D., Ucar, B.: A symmetry preserving algorithm for matrix scaling. SIAM J. Matrix Anal. Appl. **35**(3), 931–955 (2014)
54. Golub, H.V., Van Loan, C.F.: Matrix computations. 3rd ed. The Johns Hopkins University Press, Baltimore (1996)
55. Davis, T.A., Duff, I.S.: A combined unifrontal/multifrontal method for unsymmetric sparse matrices. ACM Trans. Math. Softw. **25**(1), 1–20 (1999)
56. Woźnicki, Z.: On performance of SOR method for solving nonsymmetric linear systems. J. Comput. Appl. Math. **137**(1), 145–176 (2001)
57. Saad, Y., Van Der Vorst, H.A.: Iterative solution of linear systems in the 20th century. J. Comput. Appl. Math. **123**(1-2), 1–33 (2000)
58. Diersch, H.J.G.: FEFLOW reference manual. DHI-WASY GmbH, Berlin (2009)
59. Van der Lee, J., et al.: Presentation and application of the reactive transport code HYTEC. In: Hassanizadeh, S.M. (ed.) Developments in Water Science, Computational Methods in Water Resources, Proceedings of the XIVth International Conference on Computational Methods in Water Resources (CMWR XIV), pp. 599–606. Elsevier (2002)
60. Press, W.H., S.A.T., Vettering, W.T., Flannery, B.P. Numerical recipes in FORTRAN: the art of scientific computation, 2nd edn., pp. 123–124. Cambridge University Press, New Yor (1992)
61. The Linear Algebra Package (LAPACK) can be obtained free of charge from the address listed here: http://www.netlib.org/lapack
62. Kincaid, D., Cheney, W. Numerical analysis: mathematics of scientific computing, 3rd edn. American Mathematical Society (2002)
63. HSL: A collection of Fortran codes for large scale scientific computation. http://www.hsl.rl.ac.uk (2013)
64. Chapter 8 Systems of nonlinear equations. In: Studies in computational mathematics, Claude, B. Editor. 1997, Elsevier. pp. 287–336
65. Soleymani, F.: A rapid numerical algorithm to compute matrix inversion. Int. J. Math. Math. Sci. **2012** (2012)
66. Soleymani, F.: On a fast iterative method for approximate inverse of matrices. Commun. Korean Math. Soc. **28**(2), 407–418 (2013)
67. Morin, K.A.: Simplified explanations and examples of computerized methods for calculating chemical equilibrium in water. Comput. Geosci. **11**, 409–416 (1985)