



HAL
open science

CONVERGENCE ACCELERATION OF ITERATIVE SEQUENCES FOR EQUILIBRIUM CHEMISTRY COMPUTATIONS

Safaa Al Nazer, Mustapha Jazar, Carole Rosier

► **To cite this version:**

Safaa Al Nazer, Mustapha Jazar, Carole Rosier. CONVERGENCE ACCELERATION OF ITERATIVE SEQUENCES FOR EQUILIBRIUM CHEMISTRY COMPUTATIONS. Computational Geosciences, 2021, 25 (5), pp.1509-1538. 10.1007/s10596-021-10054-1 . hal-04121533

HAL Id: hal-04121533

<https://ulco.hal.science/hal-04121533v1>

Submitted on 7 Jun 2023

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

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

CONVERGENCE ACCELERATION OF ITERATIVE SEQUENCES FOR EQUILIBRIUM CHEMISTRY COMPUTATIONS.

SAFAA AL NAZER, MUSTAPHA JAZAR, AND CAROLE ROSIER

ABSTRACT. The modeling of thermodynamic equilibria leads to complex nonlinear chemical systems which are often solved with the Newton-Raphson method. But this resolution can lead to a non convergence or an excessive number of iterations due to the very ill-conditioned nature of the problem. In this work, we combine a particular formulation of the equilibrium system called the Positive Continuous Fraction method with two iterative methods, Anderson Acceleration method and Vector extrapolation methods (namely the reduced rank extrapolation and the minimal polynomial extrapolation). The main advantage of this approach is to avoid forming the Jacobian matrix. In addition, a strategy is used to improve the robustness of the Anderson acceleration method which consists in reducing the condition number of matrix of the least squares problem in the implementation of the Anderson acceleration so that the numerical stability can be guaranteed. We compare our numerical results with those obtained with the Newton-Raphson method on the Acid Gallic test and the 1D MoMas benchmark test case and we show the high efficiency of our approach.

Keywords: Nonlinear systems; Thermodynamic chemistry; Anderson acceleration; Polynomial vector extrapolation.

1. Introduction

In the last decades, reactive transport was considered a major topic in many different fields of science such as combustion, catalysis, fluid mechanics, chemical engineering and geochemistry. Single phase multicomponent reactive flows are modeled by a masse balance law, Darcy's law and equations of state. In the case of equilibrium reactions, mass action laws consist in algebraic equations linking the activities of involved species. The problem of reactive transport is thus modeled by partial differential equations describing the flow coupled with algebraic equations describing chemical reactions. Due to the complexity of systems and the nonlinearity of chemical processes, reactive multicomponent transport results in an important computational requirement. In this context, two numerical strategies are usually used to solve this system : the global implicit algorithm (GIA) and the sequential iterative (and non-iterative) algorithm (SIA), also called operator splitting approach (see for instance references [2, 29, 38, 45]). The global implicit algorithm solves at each time step the complete nonlinear system resulting from the direct substitution of the chemical equations in the transport equations while the operator splitting approach solves sequentially transport equations and biogeochemical reactions. Results of recent comparisons between GIA and SIA obtained by different teams are in good agreement ([14]), such as those given in ([1]) for which a fully implicit finite volume method has been developed and implemented in the framework of the parallel open-source platform Dumu^X ([19, 21]). These different benchmarks have shown that the precision of sequential approaches is comparable with that of global approaches and that global approaches are now more efficient than originally believed (even if in some work, it is mentioned that the global approach is much more expensive in terms of computation time and storage than the operator splitting approach cf [52]). Each of these methods has qualities and drawbacks but regardless of the approach, a nonlinear problem must be solved by a fixed point method and the Newton Raphson method is often used for this numerical resolution. However, the resolution of such nonlinear systems, especially due to chemical processes can yield a non convergence or an excessive number of iterations due to the very ill-conditioned nature of the problem. The goal of this work is to suggest new powerful algorithms (in terms of CPU time et stability) which will allow to deal with these stiff problems.

In thermodynamic terms, a chemical equilibrium calculation, which aims to find the minimum value for the Gibbs free energy, can be carried out through one of the following ways: by minimizing a free energy function or by solving a set of nonlinear equations consisting of equilibrium constants and mass

balance constraints. Note that, in the petroleum industry context, recent alternative approach ([54]) studies phase equilibrium under a fixed volume rather than fixed pressure and minimizes Helmholtz free energy instead of Gibbs free energy. Finally, recent works use efficient deep learning algorithms to estimate the thermodynamic equilibrium states of realistic reservoir fluids with a large number of components thus allowing to accelerate phase equilibrium calculations. More precisely, a simple acceleration strategy reduces the number of components in the fluid mixture improving the efficiency of algorithms without compromising the accuracy of equations of states (see [55, 56]). These methods are thermodynamically equivalent, but the main disadvantage of using a free energy database is that these values are not nearly as reliable as directly measured equilibrium constants. As the accuracy of results of chemical solvers is particularly required, especially if one wants to integrate them in SIA methods, we are going to focus on the numerical resolution of nonlinear equations describing thermodynamic equilibria.

Many mathematical methods were tested to solve the set of nonlinear algebraic equations describing thermodynamic equilibrium: Zero-order methods such as the continuous fractions method [51], the Simplex method ([32]) which do not use the derivative of the objective function. The latter methods converge more slowly [31], but are sometimes considered more robust than first-order methods. The Simplex method is believed to be the most robust and may find the thermodynamic equilibrium when first-order methods are inefficient ([6, 33]). As mentioned above, the Newton-Raphson method is the most used to compute thermodynamic equilibrium or more generally to solve the set of nonlinear equations. For example, let us quote software such as HYDROGEOCHEM [53], DUNE [21], IMPACT [26], CHESS [47], or PHREEQC [33], with the difficulty that the Jacobian matrix has to be computed, stored, factored and is usually very ill-conditioned, which requires preconditioning procedures [15, 16]. This can become problematic for large problems. In addition, the localization of the initial data in any algorithm of Newton type is a recurrent difficulty which slows down and even prevents the convergence of the algorithm. Finally, even small or very small chemical systems (4×4 to 20×20 , occasionally larger) can be very ill-conditioned (condition number up to 10^{100}) as it is shown in [28].

To overcome this problematic, it is more effective to solve the chemical equilibrium problem through other iterative methods not requiring the calculation of the Jacobian matrix, by first transforming it into an appropriated fixed point problem. We especially focused on three iterative acceleration methods: the *Anderson Acceleration* method (AA) originating in [3] and *Vector-Extrapolation methods*, mainly the two polynomial-type methods, which include the *Reduced-Rank Extrapolation* (RRE) of Eddy [17] and MeSina [30], and *Minimal-Polynomial Extrapolation* (MPE) of Cabay and Jackson [9]. To our knowledge, these methods have never been applied to the resolution of thermodynamic equilibria. Moreover, their efficiency is improved by combining them with a particular formulation of the equilibrium system: the positive continuous fractions method PCF. Usually, continuous fractions method is used for preconditioning the Newton-Raphson method for major species (as in the PHREEQC [33]) or to reduce the difficulties due to the lack of global convergence of Newton's method, if the initial condition is not sufficiently close to the solution (see [12]). The direct combination of PCF method with AA, RRE or MPE presented in this work provides very efficient and robust algorithms with a super linear or quadratic convergence from any arbitrary initial data. Let us now briefly describe these three iterative methods which are part of a general framework of Shanks sequence transformations (cf.[7]). Anderson acceleration is related to multiseant methods (extensions of quasi-Newton methods involving multiple secant conditions); actually, Eyert [18] proves that it is equivalent to the so-called "bad" Broyden's method [8], and a similar analysis is done by Fang and Saad [22] and Rohwedder and Schneider [36]. As for linear systems, if $m_k = k$ for each k then Anderson acceleration is essentially equivalent to the generalized minimal residual (GMRES) method [37], as shown by Potra and Engler [34], Rohwedder and Schneider [36], and Walker and Ni [49]. For nonlinear problems Rohwedder and Schneider [36] show that Anderson acceleration is locally linearly convergent under certain conditions. In addition to the previous convergence analysis results, the recent work by Toth and Kelley [46] concerning Anderson acceleration with $m_k = \min(m, k)$, for a fixed m , applied to contractive mappings should be mentioned. Regarding Vector-Extrapolation methods, the aim of such methods is to transform a sequence of vectors generated by some process to a new one with the goal to converge faster than the initial sequence towards the sought limit solution. An example to these vector sequences is those which are obtained from iterative solution of linear and nonlinear systems of equations. These methods can be classified into two main categories: the polynomial methods and

the ϵ -algorithms. There exists many polynomial extrapolation methods but, in this paper, we will be interested in the minimal polynomial extrapolation method (MPE) of Cabay and Jackson [9] as well as the reduced rank extrapolation method (RRE) of Eddy [17] and Mesina [30]. These methods do not require any explicit knowledge of how the sequence is generated, and consequently can be directly applied for solving linear and nonlinear systems. They are especially effective in the nonlinear case.

In the first part of this work, a brief description of the chemical context and the chemical modeling strategy are given. Based on the Positive Continuous Fractions method [12], it is transformed into a fixed point problem. In the second part, a survey of Anderson Acceleration method and of the two most efficient and widely used vector extrapolation methods MPE and RRE is given. By derivating these methods, stable and efficient algorithms are obtained. In the third part, numerical results for solving thermodynamic equilibrium problem by Anderson Acceleration, MPE and RRE methods are detailed. Third part contains a brief description of two chemical tests: Gallic acid test and MoMas easy test case. Using the data of this tests presented by their Morel's Tables, three iterative methods mentioned for solving the fixed point problem of chemical equilibrium are applied and numerical results for each test and each method are given. In the fourth part, a comparison between these results and other results is presented in order to prove the effectiveness of methods used in this work to solve the problem of chemical equilibrium in porous media.

2. Description and modeling of chemistry

In this section, chemical model studied in this work is described. Consider a set of n_e chemical species (\mathcal{E}_j), $j = 1, \dots, n_e$ linked by n_r reactions such that $n^r \leq n^e$

$$\sum_{j=1}^{n_e} \tilde{\mu}_{ij} \mathcal{E}_j \rightleftharpoons 0, \quad i = 1, \dots, n_r, \quad (2.1)$$

where $\tilde{\mu}_{ij}$ is the stoichiometric matrix of species \mathcal{E}_j in the reaction i . (2.1) can be written in matrix form

$$\tilde{\mu} \mathcal{E} \rightleftharpoons 0.$$

After substitution and relabeling, each reaction can be written in a form giving rise a single distinct product per reaction. It is natural to assume that the stoichiometric matrix $\tilde{\mu}$ is of full rank n_r . So $\tilde{\mu} = [-I_{n_r} \quad \mu]$ can be written in the echelon form, where I_{n_r} is the identity matrix of size n_r . The chemical system is then written (after a possible numbering) in the form

$$\mathcal{C}_i \rightleftharpoons \sum_{j=1}^{n_e - n_r} \mu_{ij} \mathcal{X}_j \quad i = 1, \dots, n_r, \quad (2.2)$$

or in matrix form

$$\mathcal{C} \rightleftharpoons \mu^T \mathcal{X}$$

where \mathcal{C} (respectively \mathcal{X}) are called secondary species (respectively component species). Thus, the equation (2.2) show that the formation of secondary species \mathcal{C} is done from the component species \mathcal{X} , in a unique way. The advantage of this approach is that it reduces the size of the chemical system to be solved. *Mobile* and *fixed* species are also distinguished. A species is said to be mobile (m) if it belongs to a mobile phase, fixed (f) if it belongs to the fixed phase and precipitated if it is mineral (π). Using the following notations:

- X : subset of mobile component species of cardinal n_{pm} ,
- S : subset of fixed component species of cardinal n_{pf} ,
- C : subset of mobile secondary species of cardinal n_{sm} ,
- CS : subset of fixed secondary species of cardinal n_{sf} ,
- π : subset of precipitated species of cardinal n_π ,
- $\mu^{(C,X)} \in \mathbb{R}^{n_{sm} \times n_{pm}}$: block of the stoichiometric matrix between C and X ,
- $\mu^{(\pi,X)} \in \mathbb{R}^{n_\pi \times n_{pm}}$: block of the stoichiometric matrix between π and X ,
- $\mu^{(CS,X)} \in \mathbb{R}^{n_{sf} \times n_{pm}}$: block of the stoichiometric matrix between CS and X ,
- $\mu^{(CS,S)} \in \mathbb{R}^{n_{sf} \times n_{pf}}$: block of the stoichiometric matrix between CS and S ,

Chemical system can be synthesized as follows

$$\begin{pmatrix} & \mu^{(C,X)} & 0 \\ -I_{n_r} & \mu^{(CS,X)} & \mu^{(CS,S)} \\ & \mu^{(\pi,X)} & 0 \end{pmatrix} \begin{pmatrix} C \\ CS \\ \pi \\ X \\ S \end{pmatrix} \rightleftharpoons 0,$$

$$\text{with } \mu = \begin{pmatrix} \mu^{(C,X)} & 0 \\ \mu^{(CS,X)} & \mu^{(CS,S)} \\ \mu^{(\pi,X)} & 0 \end{pmatrix}, \mathcal{C} = \begin{pmatrix} C \\ CS \\ \pi \end{pmatrix}, \mathcal{X} = \begin{pmatrix} X \\ S \end{pmatrix} \text{ and } \mathcal{E} = \begin{pmatrix} \mathcal{X} \\ \mathcal{C} \end{pmatrix}$$

Note that the fixed component species do not take part in the homogeneous reactions which only involve the mobile species and that the precipitation reactions do not involve the fixed species. In this work, chemical systems without precipitated species are considered, i.e. ($\pi = \phi$ and $\mu^{(\pi,X)} = 0$). A classic algorithm [12] to describe mineral precipitation or dissolution makes an a priori hypothesis about the existence or non-existence of minerals. In this work, this hypothesis is assumed.

Chemical Reactions In the following, $X = (X_1, \dots, X_{n_{pm}})^T$ denote components where $(X_j)_{j=1}^{n_{pm}}$ are the mobile components species and $S = (S_1, \dots, S_{n_{pf}})^T$ where $(S_j)_{j=1}^{n_{pf}}$ are the fixed components species. In the same way, $C = (C_1, \dots, C_{n_{sm}})^T$ where $(C_i)_{i=1}^{n_{sm}}$ are the mobile secondary species and $CS = (CS_1, \dots, CS_{n_{sf}})^T$ where $(CS_i)_{i=1}^{n_{sf}}$ are the fixed secondary species.

Let $\mu_i, i = 1, 3$ be scalars $\mu_1 = \mu^{(C,X)}$, $\mu_2 = \mu^{(CS,X)}$ and $\mu_3 = \mu^{(CS,S)}$. Using these notations, it becomes easy to distinguish chemical reactions as follows:

- i: Reactions among mobile species: $\sum_{j=1}^{n_{pm}} \mu_{1,i,j} X_j \rightleftharpoons C_i \quad i = 1, \dots, n_{sm};$
- ii: Reactions between mobile and fixed species: $\sum_{j=1}^{n_{pm}} \mu_{2,i,j} X_j + \sum_{j=1}^{n_{pf}} \mu_{3,i,j} S_j \rightleftharpoons CS_i \quad i = 1, \dots, n_{sf}.$

Mass action law The law of mass action describes how to obtain the concentrations of secondary species, given the concentrations of the component species. This law is only valid for a certain type of reaction, including homogeneous reactions. It is assumed during this work that this law is still valid in the case of surface reactions.

Since no precipitation phenomena are considered, for each mobile secondary species C_i , the mass action law is

$$\{C_i\} = K_i^m \prod_{k=1}^{n_{pm}} \{X_k\}^{\mu_{1,i,k}}. \quad (2.3)$$

For each fixed secondary species CS_i , the mass action law is

$$\{CS_i\} = K_i^s \prod_{k=1}^{n_{pm}} \{X_k\}^{\mu_{2,i,k}} \prod_{k=1}^{n_{pf}} \{S_k\}^{\mu_{3,i,k}}. \quad (2.4)$$

where $\{C_i\}$ and $\{CS_i\}$ are the activities of each mobile and fixed secondary species given by the mass action law through the activities of each mobile and fixed component species $\{X_k\}$ and $\{S_k\}$. K^m is the equilibrium constant for reactions among mobile species and K^s is the equilibrium constant for sorption reactions.

The relationship between the activity of a species \mathcal{E}_j and its concentration is given by activity coefficient (γ_j) calculated using specific models (Davies, Debye-Huckel, etc.): $\{\mathcal{E}_j\} = \gamma_j[\mathcal{E}_j]$. A solution is said to be *ideal* when the species does not undergo any interaction. In this case, the activity coefficient (γ) is equal to one. This amounts to confusing activity and concentration. During this work, only the case of ideal solutions will be considered, so $[\mathcal{E}_j] = \mathcal{E}_j$.

Mass conservation law Assuming a closed system (without exchange of matter with outside) and all the reactions at equilibrium, then the total quantity of the species \mathcal{X}_j in the system is invariant. This is expressed in terms of the total concentration T_j^m for an aqueous species and the total concentration T_j^s for a sorbed species. The law of conservation (also called Lavoisier's law) can be expressed by the

following two relations:

$$\begin{aligned} T_j^m &= X_j + \sum_{i=1}^{n_{sm}} \mu_{1,i,j} C_i + \sum_{i=1}^{n_{sf}} \mu_{2,i,j} C S_i \quad j = 1, \dots, n_{pm} \\ T_j^s &= S_j + \sum_{i=1}^{n_{sm}} \mu_{3,i,j} C S_i \quad j = 1, \dots, n_{pf}. \end{aligned} \quad (2.5)$$

The two relationships in (2.5) introduce a distinction between the concentration of the component species and the concentrations of the other secondary species. This distinction is not necessarily necessary. Reactions (2.6) for each component species can be quite considered, which has a stoichiometric coefficient equal to one and an equilibrium constant equal to one,



Then (2.5) is written more simply

$$\begin{aligned} T_j^m &= \sum_{i=1}^{n_{sm}} \mu_{1,i,j} C_i + \sum_{i=1}^{n_{sf}} \mu_{2,i,j} C S_i \quad j = 1, \dots, n_{pm} \\ T_j^s &= \sum_{i=1}^{n_{sm}} \mu_{3,i,j} C S_i \quad j = 1, \dots, n_{pf}. \end{aligned} \quad (2.7)$$

or in matrix form

$$\begin{aligned} T^m &= \mu_1^T \cdot C + \mu_2^T \cdot C S \\ T^s &= \mu_3^T \cdot C S. \end{aligned}$$

2.1. Resolution of the chemical equilibrium

Chemical system By substituting the mass action laws (2.3) and (2.4) into the mass conservation equations (2.7), one can write the equilibrium chemistry like a nonlinear system formed by conservation laws and mass action laws

$$\begin{aligned} T_j^m &= \sum_{i=1}^{n_{sm}} \mu_{1,i,j} \left(K_i^m \prod_{k=1}^{n_{pm}} X_k^{\mu_{1,i,j,k}} \right) + \sum_{i=1}^{n_{sf}} \mu_{2,i,j} \left(K_i^s \prod_{k=1}^{n_{pm}} X_k^{\mu_{2,i,j,k}} \prod_{k=1}^{n_{pf}} S_k^{\mu_{3,i,j,k}} \right) \quad j = 1, \dots, n_{pm} \\ T_j^s &= \sum_{i=1}^{n_{sm}} \mu_{3,i,j} \left(K_i^s \prod_{k=1}^{n_{pm}} X_k^{\mu_{2,i,j,k}} \prod_{k=1}^{n_{pf}} S_k^{\mu_{3,i,j,k}} \right) \quad j = 1, \dots, n_{pf}. \end{aligned} \quad (2.8)$$

This is a system of $(n_{pm} + n_{pf})$ nonlinear algebraic equations with $(n_{pm} + n_{pf})$ unknowns. It is of course not possible (in general) to calculate the exact solution of this system which will be calculated numerically by iterative methods.

A first difficulty in solving (2.8) comes from the fact that the unknowns are concentrations of the component species. These concentrations are likely to vary on several orders of magnitude, and must remain positive to keep their physical significance. These two constraints make numerical resolution difficult. Fortunately, a simple change of variables eliminates these two difficulties, and has been adopted by most computer codes: the logarithms of the concentrations are taken as unknowns. Thus, the concentrations will be automatically positive, and the unknowns of the nonlinear system will keep a reasonable order of magnitude. In this work and computer code, the logarithms at base 10, "log₁₀", of the component concentrations is used as a variable change.

$$\xi_j = \log_{10}(X_j) \quad \text{and} \quad \eta_j = \log_{10}(S_j).$$

We denote by $\mathbf{K}^m = \log_{10}(K^m)$ and $\mathbf{K}^s = \log_{10}(K^s)$.

The consequences of this transformation on the system are limited. The mass action law in the equations (2.3) and (2.4) becomes reformulated, respectively, as

$$C_i = 10^{(\mathbf{K}_i^m + \sum_{k=1}^{n_{pm}} \mu_{1,i,k} \xi_k)} \quad \text{and} \quad C S_i = 10^{(\mathbf{K}_i^s + \sum_{k=1}^{n_{pm}} \mu_{2,i,k} \xi_k + \sum_{k=1}^{n_{pf}} \mu_{3,i,k} \eta_k)}. \quad (2.9)$$

Then, the nonlinear system (2.8) takes the following form

$$T_j^m = \sum_{i=1}^{n_{sm}} \mu_{1,i,j} \cdot 10^{(\mathbf{K}_i^m + \sum_{k=1}^{n_{pm}} \mu_{1,i,k} \xi_k)} + \sum_{i=1}^{n_{sf}} \mu_{2,i,j} \cdot 10^{(\mathbf{K}_i^s + \sum_{k=1}^{n_{pm}} \mu_{2,i,k} \xi_k + \sum_{k=1}^{n_{pf}} \mu_{3,i,k} \eta_k)} \quad j = 1, \dots, n_{pm}$$

$$T_j^s = \sum_{i=1}^{n_{sm}} \mu_{3,i,j} \cdot 10^{(\mathbf{K}_i^s + \sum_{k=1}^{n_{pm}} \mu_{2,i,k} \xi_k + \sum_{k=1}^{n_{pf}} \mu_{3,i,k} \eta_k)} \quad j = 1, \dots, n_{pf}.$$

or the matrix form

$$T^m = \mu_1^T \times 10^{(\mathbf{K}^m + \mu_1 \times \xi)} + \mu_2^T \times 10^{(\mathbf{K}^s + \mu_2 \times \xi + \mu_3 \times \eta)} \quad (2.10)$$

$$T^s = \mu_3^T \times 10^{(\mathbf{K}^s + \mu_2 \times \xi + \mu_3 \times \eta)}.$$

where the symbol of matrix product is denoted by \times . The matrix form (2.10) still can be written in a more reduced manner

$$\mathbf{T} = \mu^T \times 10^{(\mathbf{K} + \mu \times \omega)}, \quad (2.11)$$

where $\mathbf{T} = \begin{pmatrix} T^m \\ T^s \end{pmatrix}$, $\mathbf{K} = \begin{pmatrix} \mathbf{K}^m \\ \mathbf{K}^s \end{pmatrix}$, $\omega = \begin{pmatrix} \xi \\ \eta \end{pmatrix}$ and $\mu = \begin{pmatrix} \mu_1 & 0 \\ \mu_2 & \mu_3 \end{pmatrix}$.

The nonlinear system (2.10) (or (2.11)) corresponds to the *chemical problem* to be solved for ξ and η given T^m and T^s . The concentrations of the secondary species can then be computed from (2.9). In the sequel, we assumed that this problem always has a unique positive solution (ξ^*, η^*) for all feasible values of the data T^m and T^s . This assumption is true due to the fact that the chemical equilibrium problem is the consequence of the Gibbs free energy minimization problem. The existence follows from the convexity of the energy functional [40]. In addition, a condition is given for the uniqueness of the solution, which is particularly verified in the case of a single-phase system, which covers the cases treated here.

In this work, three different iterative numerical methods are applied to solve the chemical equilibrium problem which are the Anderson Acceleration method and the two polynomial vector extrapolation methods (MPE and RRE). According to their definitions, these methods are used to solve a general fixed point problem of the form $G(Y) = (Y)$ where $G: \mathbb{R}^n \rightarrow \mathbb{R}^n$. So it is necessary to write the chemical problem in the form of a fixed point problem.

The numerical method most used in a large number of geochemical codes for the resolution of this nonlinear system is the Newton's method. The thesis [10] by J. Carayrou contains a comparison of the different methods to solve the problem of chemical equilibrium. His recommendation is to use a combination of Newton's method with a fixed-point method on a particular formulation of the equilibrium system (the PCF positive continuous fraction method). This combination makes it possible to reduce the difficulties due to the lack of total convergence of the method of Newton, if the initial point is not sufficiently close to the solution (which is precisely what one seeks to calculate). Furthermore, J. Carayrou limits the risk of overflow or under-filling by forcing the method to search for the solution in a neighborhood of a 'reasonable' value and he defines this reasonable neighborhood as an authorized chemical interval. In this work, we use the PCF method to reformulate the chemical problem as a fixed point problem.

Positive continuous fraction method PCF The continuous fraction method (CF) has been used to solve thermodynamic equilibrium in the computer code WATSPEC [51], or for preconditioning of the Newton-Raphson method for the major species in the PHREEQC code [33]. This method, which only needs one computation of the approximate thermodynamic equilibrium per iteration, is the cheapest zero-order method. Often, the component H^+ has a zero total concentration and is associated with negative stoichiometric coefficients. In the code WATSPEC [51], the pH value must be imposed to find the thermodynamic equilibrium. Hydrogen and oxygen are excluded from the continuous fraction preconditioning in the code PHREEQC [33]. Moreover, it has never been used for non ideal system.

To take into account a component with zero or negative total concentration, and to be more efficient with negative stoichiometric coefficients, a generalization of the (CF) method has been developed, called the positive continues fraction method (PCF) by Carayrou [12]. This new method is an empirical method. Once the equilibrium solution is found, the reactive sum \mathcal{S}^R is equal to the product

sum \mathcal{S}^P . The *reactive sum* (\mathcal{S}^R) and the *product sum* (\mathcal{S}^P) are defined by

$$\mathcal{S}_j^R = \begin{cases} \sum_{\mu_{i,j}>0} \mu_{i,j} \mathcal{C}_i & \text{if } \mathbf{T}_j \geq 0 \\ |\mathbf{T}_j| + \sum_{\mu_{i,j}>0} \mu_{i,j} \mathcal{C}_i & \text{if } \mathbf{T}_j < 0 \end{cases} \quad \text{and} \quad \mathcal{S}_j^P = \begin{cases} \mathbf{T}_j + \sum_{\mu_{i,j}<0} |\mu_{i,j}| \mathcal{C}_i & \text{if } \mathbf{T}_j \geq 0 \\ \sum_{\mu_{i,j}<0} |\mu_{i,j}| \mathcal{C}_i & \text{if } \mathbf{T}_j < 0 \end{cases} \quad (2.12)$$

for $j = 1, \dots, n_p$, with $\mathcal{C} = \begin{pmatrix} C \\ CS \end{pmatrix}$ and $n_p = n_{pm} + n_{pf}$.

Using these two new values, the mass balance (2.7) is written, at equilibrium, as

$$\mathcal{S}_j^R = \mathcal{S}_j^P.$$

The coefficient $\mu_{i_0,j}$ is taken as the smallest value of the strictly positive stoichiometric coefficient in the matrix μ . The mass action laws are written for the reactive sum if $\mu_{i_0,j}$ is positive (respectively, for the product sum if $\mu_{i_0,j}$ is negative) by using component concentrations at iterations n and $n+1$. In particular, the following equality holds

$$(\mathcal{X}_j^{n+1})^{\mu_{i_0,j}} \left[\sum_{\mu_{i,j}>0} \mu_{i,j} \mathbf{K}_i \prod_{k \neq j} (\mathcal{X}_j^n)^{\mu_{i,k}} \cdot (\mathcal{X}_j^n)^{\mu_{i,j} - \mu_{i_0,j}} \right] = \mathbf{T}_j + \sum_{\mu_{i,j}<0} |\mu_{i,j}| \mathcal{C}_i^n, \quad (2.13)$$

where $\mathcal{X} = \begin{pmatrix} X \\ S \end{pmatrix}$, \mathcal{X}_j^n is the concentration of the j th component species \mathcal{X}_j at iteration n and \mathcal{C}_i^n is that of the i th secondary species \mathcal{C}_i at iteration n . After reordering, (2.13) becomes

$$(\mathcal{X}_j^{n+1})^{\mu_{i_0,j}} = \frac{(\mathcal{X}_j^n)^{\mu_{i_0,j}}}{(\mathcal{X}_j^n)^{\mu_{i_0,j}} \sum_{\mu_{i,j}>0} \mu_{i,j} \mathbf{K}_i \prod_{k \neq j} (\mathcal{X}_j^n)^{\mu_{i,k}} \cdot (\mathcal{X}_j^n)^{\mu_{i,j} - \mu_{i_0,j}}} \left(\mathbf{T}_j + \sum_{\mu_{i,j}<0} |\mu_{i,j}| \mathcal{C}_i^n \right).$$

Then, the relationship (2.14) giving \mathcal{X}_j^{n+1} is

$$\mathcal{X}_j^{n+1} = \mathcal{X}_j^n \left(\frac{\mathcal{S}_j^{P,n}}{\mathcal{S}_j^{R,n}} \right)^{\frac{1}{\mu_{i_0,j}}}. \quad (2.14)$$

Since $\omega^{n+1} = \log_{10}(\mathcal{X}^{n+1})$, then, written according to the logarithm of the component species concentrations, the relation (2.14) becomes

$$\omega_j^{n+1} = \omega_j^n + \frac{1}{\mu_{i_0,j}} \left[\log_{10}(\mathcal{S}_j^{P,n}) - \log_{10}(\mathcal{S}_j^{R,n}) \right]. \quad (2.15)$$

This relation is considered to be the conventional fixed point iteration

$$\omega^{n+1} = \mathbf{G}(\omega^n), \quad n = 0, 1, \dots, \quad (2.16)$$

where $\mathbf{G} : \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_p}$ is the fixed point map defined by

$$\mathbf{G}(\omega) = \omega + \frac{1}{\mu_0} \left[\log_{10}(\mathcal{S}^P) - \log_{10}(\mathcal{S}^R) \right]. \quad (2.17)$$

Thus, solving the chemical equilibrium problem (2.8) amounts to solving the fixed point problem

$$\omega = \mathbf{G}(\omega). \quad (2.18)$$

3. Iterative methods

The aim is to solve the previous nonlinear fixed point problem (2.18) whose the solution is denoted by ω^* . Then starting with a suitable vector ω_0 , as an initial approximation to ω^* , the sequence $\{\omega_n\}$ is generated by fixed point iterative (FPI) methods defined by (2.16).

3.1. Anderson Acceleration

To improve the convergence rate of FPI (2.16), Anderson acceleration [3] is applied. It is formulated as follows [49]:

Algorithm 1: ANDERSON ACCELERATION (AA).

Given ω_0 and $m \geq 1$.

Set $x_1 = \mathbf{G}(\omega_0)$ and $f_0 = \mathbf{G}(\omega_0) - \omega_0$

For $k = 0, 1, \dots$

Set $m_k = \min\{m, k\}$.
 Compute $\mathbf{G}(\omega_k)$ and let $f(\omega_k) = \mathbf{G}(\omega_k) - \omega_k$.
 Set $F_k = (f_{k-m_k}, \dots, f_k)$.
 Determine $\alpha_{(k)} = (\alpha_0^{(k)}, \dots, \alpha_{m_k}^{(k)})^T$ that solves

$$\min_{\alpha=(\alpha_0, \dots, \alpha_{m_k})^T} \|F_k \alpha\|_2, \quad \text{s.t.} \quad \sum_{i=0}^{m_k} \alpha_i = 1 \quad (3.1)$$

Set $\omega_{k+1} = (1 - \beta_k) \sum_{i=0}^{m_k} \alpha_i^{(k)} (\omega_{k-m_k+i}) + \beta_k \sum_{i=0}^{m_k} \alpha_i^{(k)} \mathbf{G}(\omega_{k-m_k+i})$

where $\beta_k > 0$ is a relaxation parameter. In [49], it is shown that Anderson acceleration with $\beta_k = 1$ converges when the fixed point map \mathbf{G} is a contraction and that the rate of convergence is comparable to that of the Picard iteration. In this work, as in [49], only the case $\beta_k = 1$ in Algorithm 1 is considered. If $m = 0$, then Anderson acceleration becomes the FPI (2.16).

Form of least-squares problem In practical implementation, the constrained least-squares problem (3.1) is often formulated as the following equivalent unconstrained least-squares problem ([22], [49]): Find $\gamma^{(k)} = (\gamma_0^{(k)}, \dots, \gamma_{m_k-1}^{(k)})^T$ such that

$$\min_{\gamma} \|f_k - \mathcal{F}_k \gamma\|_2 \quad (3.2)$$

where

$$\mathcal{F}_k = (\Delta f_{k-m_k}, \dots, \Delta f_{k-1}) \quad (3.3)$$

with $\Delta f_i = f_{i+1} - f_i$ for $i = k - m_k, \dots, k - 1$. The least-squares coefficient vectors α and γ are related by $\alpha_0 = \gamma_0$, $\alpha_j = \gamma_j - \gamma_{j-1}$ for $1 \leq j \leq m_k - 1$ and $\alpha_{m_k} = 1 - \gamma_{m_k-1}$. The next iterate then becomes $\omega_{k+1} = \mathbf{G}(\omega_k) - \sum_{i=1}^{m_k-1} \gamma_i^{(k)} [\mathbf{G}(\omega_{k-m_k+i+1}) - \mathbf{G}(\omega_{k-m_k+i})] = \mathbf{G}(\omega_k) - \mathcal{G}_k \gamma^{(k)}$, where

$$\mathcal{G}_k = (\Delta \mathbf{G}_{k-m_k}, \dots, \Delta \mathbf{G}_{k-1}) \quad (3.4)$$

with $\Delta \mathbf{G}_i = \mathbf{G}(\omega_{i+1}) - \mathbf{G}(\omega_i)$ for $i = k - m_k, \dots, k - 1$.

Then, a more specific version of the AA algorithm can be given in Algorithm 2.

Algorithm 2: ANDERSON ACCELERATION (AA).

Given ω_0 and $m \geq 1$.

Set $\omega_1 = \mathbf{G}(\omega_0)$ and $f_0 = \mathbf{G}(\omega_0) - \omega_0$.

For $k = 1, 2, \dots$

Set $m_k = \min(m, k)$.

Compute $\mathbf{G}(\omega_k)$ and let $f_k = \mathbf{G}(\omega_k) - \omega_k$.

Update \mathcal{F}_k and \mathcal{G}_k by (3.3) and (3.4).

Determine $\gamma^{(k)} = (\gamma_0^{(k)}, \dots, \gamma_{m_k-1}^{(k)})^T$ that solves $\min_{(\gamma_0, \dots, \gamma_{m_k-1})^T} \|f_k - \mathcal{F}_k \gamma\|_2$.

Set $\omega_{k+1} = \mathbf{G}(\omega_k) - \mathcal{G}_k \gamma^{(k)}$.

The least-squares problem (3.2) is solved by performing the QR factorization of \mathcal{F}_k and using backward substitution to solve the upper triangular system $R_k \gamma = Q_k^T f_k$. This shows that only Q_k and R_k need to be computed. In other words, only the "thin" QR decomposition of \mathcal{F}_k ($\mathcal{F}_k = Q_k R_k$, $Q_k \in \mathbb{R}^{N \times m_k}$ and $R_k \in \mathbb{R}^{m_k \times m_k}$) is necessary.

Since \mathcal{F}_k is obtained from \mathcal{F}_{k-1} by appending a new column on the right and possibly dropping one column from the left, the QR decomposition of \mathcal{F}_k can be efficiently obtained by updating that of \mathcal{F}_{k-1} . For details about this aspect, see [48].

Condition control In practice, there is often a risk that \mathcal{F}_k will become ill-conditioned as iterations go. In this work, the strategy given in [49] is used to monitor the condition number of the matrix \mathcal{F}_k and, if necessary, to modify the matrix to reduce the condition number, as follows: when the condition number of \mathcal{F}_k is larger than a given tolerance, then the left-most columns of \mathcal{F}_k are dropped one by one until the condition number is less than the given tolerance. Note that the l^2 -norm condition number of \mathcal{F}_k is just that of R_k in the QR decomposition of \mathcal{F}_k . Therefore, for the filtering strategy used in this paper, it is only necessary to monitor the condition number of R_k and keep it less than the given tolerance. If the condition number of R_k is larger than the given tolerance, then removing the leftmost column of \mathcal{F}_k involves updating the factors Q_k and R_k (see [48] for details).

3.2. Polynomial Vector Extrapolation Methods MPE and RRE

An important problem that arises in different areas of science and engineering is that of computing the limits of sequences of vectors. Vector sequence arises, for example, in the solution of system of linear or nonlinear equations by fixed-point iterative methods, its limit being simply the required solution.

Let $\{x_k\}_{k \in \mathbb{N}}$ be a sequence of vectors in \mathbb{R}^N , and define the first and second forward differences such that

$$\Delta x_k = x_{k+1} - x_k \quad \text{and} \quad \Delta^2 x_k = \Delta x_{k+1} - \Delta x_k \quad k = 0, 1, \dots$$

When MPE and RRE are applied to the vector sequence $\{x_k\}$, an approximation $t_{k,q}$ is produced of the limit or antilimit of $\{x_k\}_{k \in \mathbb{N}}$ (cf. [41]). It is clear that t_k will be different for each method. Let

$$t_{k,q} = \sum_{j=0}^k \nu_j^{(k)} x_{q+j} \quad (3.5)$$

subject to

$$\sum_{j=0}^k \nu_j^{(k)} = 1 \quad \text{and} \quad \sum_{j=0}^k \tau_{i,j} \nu_j^{(k)} = 0 \quad i = 0, 1, \dots, k-1 \quad (3.6)$$

with the scalars $\tau_{i,j}$ defined by the inner products in \mathbb{R}^N : $\tau_{i,j} = \begin{cases} (\Delta x_{q+i}, \Delta x_{q+j}) & \text{for MPE} \\ (\Delta^2 x_{q+i}, \Delta x_{q+j}) & \text{for RRE} \end{cases}$.

Using (3.6), the transformation (3.5) can also be expressed as a ratio of two determinants as follows

$$t_{k,q} = \frac{\begin{vmatrix} x_q & x_{q+1} & \dots & x_{q+k} \\ \tau_{0,0} & \tau_{0,1} & \dots & \tau_{0,k} \\ \vdots & \vdots & & \vdots \\ \tau_{k-1,0} & \tau_{k-1,1} & \dots & \tau_{k-1,k} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \dots & 1 \\ \tau_{0,0} & \tau_{0,1} & \dots & \tau_{0,k} \\ \vdots & \vdots & & \vdots \\ \tau_{k-1,0} & \tau_{k-1,1} & \dots & \tau_{k-1,k} \end{vmatrix}}.$$

Matrices $\Delta^i S_{k,q} = [\Delta^i x_q, \dots, \Delta^i x_{q+k-1}]$, $i = 1, 2$, are introduced. Using Schur complements, $t_{k,q}$ can be written, for each method as

$$\begin{aligned} t_{k,q}^{\text{MPE}} &= x_q - \Delta S_{k,q} (\Delta S_{k,q}^T \Delta^2 S_{k,q})^{-1} \Delta S_{k,q}^T \Delta x_q \\ t_{k,q}^{\text{RRE}} &= x_q - \Delta S_{k,q} (\Delta^2 S_{k,q}^T \Delta^2 S_{k,q})^{-1} \Delta^2 S_{k,q}^T \Delta x_q \end{aligned}$$

provided that $\det(\Delta S_{k,q}^T \Delta^2 S_{k,q}) \neq 0$ and $\det(\Delta^2 S_{k,q}^T \Delta^2 S_{k,q}) \neq 0$. These two assumptions are assumed in the following. Then $t_{k,q}^{\text{MPE}}$ and $t_{k,q}^{\text{RRE}}$ are well defined and unique. For varying value of k and q , the computation of $t_{k,q}^{\text{MPE}}$ and $t_{k,q}^{\text{RRE}}$ can be done by some of algorithms proposed by Ford and Sidi in [42].

An estimate for the residual norm for nonlinear problems is given. Introduce the new approximation

$$\tilde{t}_{k,q} = \sum_{j=0}^k \nu_j^{(k)} x_{q+j+1}$$

In [25], the generalized residual of $t_{k,q}$ as is defined by

$$\tilde{r}(t_{k,q}) = \tilde{t}_{k,q} - t_{k,q}, \quad (3.7)$$

which can be expressed as

$$\begin{aligned} \tilde{r}(t_{k,q}^{\text{MPE}}) &= \Delta x_q - \Delta^2 S_{k,q} (\Delta S_{k,q}^T \Delta^2 S_{k,q})^{-1} \Delta S_{k,q}^T \Delta x_q \\ \tilde{r}(t_{k,q}^{\text{RRE}}) &= \Delta x_q - \Delta^2 S_{k,q} (\Delta^2 S_{k,q}^T \Delta^2 S_{k,q})^{-1} \Delta^2 S_{k,q}^T \Delta x_q. \end{aligned}$$

Implementation Only the case q kept constant is considered. Without restriction, $q = 0$ is always assumed and $t_{k,0}$ is denoted by t_k ; $\Delta^i S_{k,0}$ by $\Delta^i S_k$. The linear system (3.6) can be written as

$$\begin{aligned} \nu_0^{(k)} &+ \nu_1^{(k)} &+ \dots &+ \nu_k^{(k)} &= 1 \\ \nu_0^{(k)}(u_0, \Delta x_0) &+ \nu_1^{(k)}(u_0, \Delta x_1) &+ \dots &+ \nu_k^{(k)}(u_0, \Delta x_k) &= 0 \\ \nu_0^{(k)}(u_1, \Delta x_0) &+ \nu_1^{(k)}(u_1, \Delta x_1) &+ \dots &+ \nu_k^{(k)}(u_1, \Delta x_k) &= 0 \\ &\vdots &&\vdots &\vdots \\ \nu_0^{(k)}(u_{k-1}, \Delta x_0) &+ \nu_1^{(k)}(u_{k-1}, \Delta x_1) &+ \dots &+ \nu_k^{(k)}(u_{k-1}, \Delta x_k) &= 0 \end{aligned} \quad (3.8)$$

Introduce the scalars $\theta_i^{(k)} = \frac{\nu_i^{(k)}}{\nu_k^{(k)}}$, for $i = 0, \dots, k$. Then, $\nu_i^{(k)} = \frac{\theta_i^{(k)}}{\sum_{i=0}^k \theta_i^{(k)}}$, for $i = 0, \dots, k-1$, and $\theta_k^{(k)} = 1$.

With this new variables, the linear system (3.8) becomes

$$\begin{aligned} \theta_0^{(k)}(u_0, \Delta x_0) &+ \theta_1^{(k)}(u_0, \Delta x_1) &+ \dots &+ \theta_{k-1}^{(k)}(u_0, \Delta x_{k-1}) &= &-(u_0, \Delta x_k) \\ \vdots &&&\vdots &&\vdots \\ \theta_0^{(k)}(u_{k-1}, \Delta x_0) &+ \theta_1^{(k)}(u_{k-1}, \Delta x_1) &+ \dots &+ \theta_{k-1}^{(k)}(u_{k-1}, \Delta x_{k-1}) &= &-(u_{k-1}, \Delta x_k) \end{aligned}$$

This system can be written in the following form

$$(U_k^T \Delta S_k) \theta^{(k)} = -U_k^T \Delta x_k \quad (3.9)$$

where $\theta^{(k)} = (\theta_0^{(k)}, \dots, \theta_{k-1}^{(k)})^T$, $\Delta S_k = (\Delta x_0, \dots, \Delta x_{k-1})$ and $U_k = \begin{cases} \Delta S_k & \text{for the MPE method} \\ \Delta^2 S_k & \text{for the RRE method.} \end{cases}$

Assume now that the coefficients $\nu_0^{(k)}, \dots, \nu_k^{(k)}$ have been calculated and introduce the new variables

$$\sigma_0^{(k)} = 1 - \nu_0^{(k)}, \quad \sigma_j^{(k)} = \sigma_{j-1}^{(k)} - \nu_j^{(k)}, \quad j = 1, \dots, k-1, \quad \text{and} \quad \sigma_{k-1}^{(k)} = \nu_k^{(k)}.$$

Then, for both method, the vector t_k can be expressed as

$$t_k = x_0 + \sum_{j=0}^{k-1} \sigma_j^{(k)} \Delta x_j = x_0 + \Delta S_k \sigma^{(k)} \quad (3.10)$$

where $\sigma = (\sigma_0, \dots, \sigma_{k-1})^T$.

Note that to determine the coefficient $\nu_i^{(k)}$, we must first calculate the $\theta_i^{(k)}$ by solving the linear system of equations (3.9). Using (3.7) and (3.10), the generalized residual $\tilde{r}(t_k)$, for MPE and RRE, can be expressed as

$$\tilde{r}(t_k) = \sum_{i=0}^k \nu_i^{(k)} \Delta x_i = \Delta S_{k+1} \nu^{(k)}.$$

Algorithms for RRE and MPE methods Fast, stable, and storage wise economical algorithms are described in [24]. These algorithms solve least-squares problems by QR factorization. An overview of these algorithms is provided in the following.

ΔS_{k+1} has a full rank, namely $\text{rank}(\Delta S_{k+1}) = k+1$. Then a QR factorization of ΔS_{k+1} can be computed.

For RRE method, this QR decomposition is defined by $\Delta S_{k+1} = Q_k R_k$ where $Q_k = (q_0|q_1|\dots|q_k) \in \mathbb{R}^{N \times (k+1)}$ has orthonormal columns q_j and $R_k \in \mathbb{R}^{(k+1) \times (k+1)}$ is an upper triangular matrix with positive diagonal coefficients. Q_k is obtained from $Q_{k-1} \in \mathbb{R}^{N \times k}$ by adding the column q_k . In the same way, R_k is obtained from $R_{k-1} \in \mathbb{R}^{k \times k}$ by adding a row and a column to R_{k-1} .

For MPE method, $\Delta S_{k+1} = Q_{k+1} R_{k+1}$, where $Q_{k+1} = (q_0|q_1|\dots|q_k) \in \mathbb{R}^{N \times (k+1)}$ is an orthogonal matrix and $R_{k+1} \in \mathbb{R}^{(k+1) \times (k+1)}$ is an upper triangular matrix with positive diagonal coefficients. Q_{k+1} is obtained from $Q_k \in \mathbb{R}^{N \times k}$ by adding the vector column q_k . Similarly, R_{k+1} is obtained from $R_k \in \mathbb{R}^{k \times k}$ by adding a row and a column to R_k .

For both method, the QR factorization of ΔS_{k+1} can be computed inexpensively by applying the modified Gram-Schmidt process (MGS) to the vectors x_0, x_1, \dots, x_{k+1} (cf. MGS algorithm in [43]).

The details of previous algorithms for RRE (resp. MPE) are summarized in Algorithm 3 (resp. Algorithm 4). Note that, in these algorithms, it is only necessary to store the vector x_0 and the

matrix Q_k . The rest can be overwritten as soon as they have been used.

Algorithm 3: RRE METHOD

0. *Inputs:* Vectors x_0, x_1, \dots, x_{k+1} .
1. Compute $v_i = \Delta x_i = x_{i+1} - x_i$, $i = 0, 1, \dots, k$.
 Set $V_j = [v_0 | v_1 | \dots | v_{j-1}]$, $j = 0, 1, \dots$.
 Compute the QR factorization of V_{k+1} , namely $V_{k+1} = Q_k R_k$.
 ($V_k = Q_{k-1} R_{k-1}$ is contained in $V_{k+1} = Q_k R_k$).
2. *Computation of the ν_i :*
 Solve the linear system: $R_k^T R_k d^{(k)} = e$; $d^{(k)} = [d_0^{(k)}, d_1^{(k)}, \dots, d_k^{(k)}]^T$; $e = [1, 1, \dots, 1]^T$.
 (This amounts to solving two upper and lower triangular systems).
 Set $\lambda = (\sum_{i=0}^k d_i^{(k)})^{-1}$, $\lambda \in \mathbb{R}^+$.
 Set $\nu_i^{(k)} = \lambda d_i^{(k)}$, $i = 0, 1, \dots, k$.
3. Compute $\sigma^{(k)} = [\sigma_0^{(k)}, \sigma_1^{(k)}, \dots, \sigma_{k-1}^{(k)}]^T$ by: $\sigma_0^{(k)} = 1 - \nu_0^{(k)}$ and $\sigma_j^{(k)} = \sigma_{j-1}^{(k)} - \nu_j^{(k)}$, $j = 1, \dots, k-1$.
 Compute t_k via: $t_k^{\text{RRE}} = x_0 + Q_{k-1}(R_{k-1}\sigma^{(k)})$

Algorithm 4: MPE METHOD

0. *Inputs:* Vectors x_0, x_1, \dots, x_{k+1} .
1. Compute $v_i = \Delta x_i = x_{i+1} - x_i$, $i = 0, 1, \dots, k$.
 Set $V_j = [v_0 | v_1 | \dots | v_j]$, $j = 0, 1, \dots$.
 Compute the QR factorization of V_{k+1} , namely $V_{k+1} = Q_{k+1} R_{k+1}$.
 ($V_k = Q_k R_k$ is contained in $V_{k+1} = Q_{k+1} R_{k+1}$).
2. *Computation of the ν_i :*
 Solve the upper triangular linear system: $R_k d^{(k)} = -r_k$; $d^{(k)} = [d_0^{(k)}, \dots, d_{k-1}^{(k)}]^T$; $r_k = [r_{0k}, \dots, r_{(k-1)k}]^T$.
 Set $d_k^{(k)} = 1$ and calculate $\lambda = (\sum_{i=0}^k d_i^{(k)})^{-1}$, $\lambda \in \mathbb{R}^+$.
 Set $\nu_i^{(k)} = \lambda d_i^{(k)}$, $i = 0, 1, \dots, k$.
3. Compute $\sigma^{(k)} = [\sigma_0^{(k)}, \sigma_1^{(k)}, \dots, \sigma_{k-1}^{(k)}]^T$ by: $\sigma_0^{(k)} = 1 - \nu_0^{(k)}$ and $\sigma_j^{(k)} = \sigma_{j-1}^{(k)} - \nu_j^{(k)}$, $j = 1, \dots, k-1$.
 Compute t_k via: $t_k^{\text{MPE}} = x_0 + Q_k(R_k\sigma^{(k)})$.

Algorithms 3 and 4 become increasingly expensive as the number of iteration steps k is increasing. Indeed the work requirement grows quadratically with k and the storage requirement grows linearly. A good way to keep the storage requirement and the computation cost low is to periodically restart the RRE and MPE algorithms every c steps, for some integer $c > 1$. Below, a practical strategy of a restarted method is described in Algorithm 5.

Algorithm 5: CYCLIC METHOD EVERY c ITERATIONS

- For $k = 0$, choose an integer c and an initial vector x_0 .
 For $k = 1, 2, \dots$,
 Compute the vectors x_1, \dots, x_c .
 Calculate t_{c-1} using the algorithm of the desired method.
 If t_{c-1} satisfies accuracy test, stop;
 Else, set $x_0 = t_{c-1}$.

Similarly to linear problems [43], it is more useful to run some basic iterations before applying one of the extrapolation methods for solving (2.18):

- Let run some N_0 basic iterations before cycling is started, e.g, before MPE or RRE is applied for the first time (N_0 refers to the size of extrapolation);
- Let run some N basic iterations before MPE or RRE is applied in each cycle after the first cycle.

One way to make the extrapolation process more efficient with high numerical stability is to change (2.16) as follows

$$\omega_{n+1} = \tilde{\mathbf{G}}(\omega_n) \quad n = 0, 1, \dots, \quad (3.11)$$

where

$$\tilde{\mathbf{G}}(\omega) = \omega + \kappa(\mathbf{G}(\omega) - \omega). \quad (3.12)$$

The scalar κ is different than 1 (the sequence generated by taking $\kappa = 1$ is the one generated by (2.16)). Thus ω_{n+1} is now weighted "average" of ω_n and $\mathbf{G}(\omega_n)$, in which weights $1 - \kappa$ and κ do not need to both be positive. By picking κ appropriately, the spectrum of the Jacobian matrix of $(1 - \kappa)\omega + \kappa\mathbf{G}(\omega)$ at $\omega = \omega^*$ can be taken as it is increasingly favorable to $t_{k,q}$ for large values of q ([43]; Section 7). So, with ω_0 , the initial approximation of ω^* , we generate the sequence $\omega_1, \omega_2, \omega_3, \dots$ by the fixed-point iteration (3.11). We consider the following algorithm:

Algorithm 6: EXTRAPOLATION ALGORITHM FOR THE NONLINEAR SYSTEM (2.18)

1. For $k = 0$, choose ω_0 and the integers p and l .
2. Basic iteration:
 - Set $t_0 = \omega_0$.
 - $h_0 = t_0$.
 - $h_{j+1} = \tilde{\mathbf{G}}(h_j)$, $j = 0, \dots, p - 1$.
3. Extrapolation phase:
 - $s_0 = h_p$.
 - If $\|s_1 - s_0\| < \epsilon$, stop,
 - Else $s_{j+1} = \tilde{\mathbf{G}}(s_j)$, $j = 0, \dots, l$.
 - Compute the approximation t_l by RRE or MPE.
4. Set $\omega_0 = t_l$, $k = k + 1$ and go to 2.

4. Numerical experiments

In this section, numerical experiments are reported. Iterative numerical methods cited above are tested to resolve the chemical equilibrium of two different chemical systems. Anderson acceleration is implemented using the approach cited in section 3.1, in its specific version (Algorithms 1-2). Sometimes, (AA) method is applied to the fixed point problem with relaxation (3.12) instead of (2.18), with $\kappa \neq 1$. The two vector extrapolation methods of the polynomial type MPE and RRE are also applied to the problem of the nonlinear chemical equilibrium system thanks to the implementations previously described. These implementations are done by employing the computer program which is provided in [43] using Matlab R2018a (this computer program written in Fortran has been converted in Matlab language). It gives an estimation of the residual norm at each iteration for this nonlinear problem and it allows to stop algorithms without having to compute the true residual which requires an extra evaluation of $\tilde{\mathbf{G}}$.

Some essential numerical parameters for test cases are:

For Anderson Acceleration:

- the iteration is stopped when the residual norm falls below 10^{-10} ;
- the condition-number monitoring is used with a threshold for deleting columns $droptol = 10^{10}$;
- The allowed maximal nonlinear iteration number is $Kmax = 200$ iterations.

For MPE and RRE:

- the maximum number of cycles allowed is $Ncycle = 30$;
- the upper bound of $resc/resp$ used in the stopping criterion is $epsc = 10^{-10}$ (where $resp$ is the l_2 -norm of the residual for $t(Kmax, N_0)$ at the end of the first cycle and $resc$ is the l_2 -norm of the residual at t at the end of each cycle, retrieved at the end of the next cycle). If $resc \leq epsc \times resp$ at the end of some cycle, then one additional cycle is performed, and the corresponding $t(N, Kmax)$ is accepted as the final approximation.
- the upper bound of $res/R(0,0)$, the relative residual for t , used in the stopping criterion is $eps = 0$. Note that $R(0,0) = l_2$ norm of the residual of ω_0 , the initial vector. If, for some k , $res \leq eps \times R(0,0)$ then the corresponding $t(0,k)$ is accepted as the final approximation.

4.1. Chemical tests

Galic acid test This is the simplest test, a system proposed by Brassard and Bodurtha (2000) [6] to illustrate the appearance of problems with numerical methods. The system was originally studied for the speciation of Al(III) in natural waters. It is characterized by the presence of $n_e = 17$ chemical species that can be described through the combination of $n_{pm} = 3$ mobile components species ($n_{sm} = 14$). All reactions describing this chemical system are homogeneous, between the mobile species ($n_{pf} = n_{sf} = 0$). The pH is imposed at 5.8, which gives a problem with two unknowns: concentrations

of the two free components Al^{3+} and H_3L .

The chemical system studied is presented in Table 1 where the initial concentrations of Al^{3+} and H_3L

Species	H^+	Al^{3+}	H_3L	\mathbf{K}^m	C (Equil.)
H^+	1	0	0	0	1.58×10^{-6}
Al^{3+}	0	1	0	0	2.03×10^{-5}
H_3L	0	0	1	0	2.59×10^{-7}
OH^-	-1	0	0	-14	6.31×10^{-9}
H_2L^-	-1	0	1	-4.15	1.16×10^{-5}
HL^{2-}	-2	0	1	-12.59	2.65×10^{-8}
L^{3-}	-3	0	1	-23.67	1.39×10^{-13}
AlHL^+	-2	1	1	-4.93	2.45×10^{-5}
AlL	-3	1	1	-9.43	4.90×10^{-4}
AlL_2^{3-}	-6	1	2	-21.98	8.97×10^{-6}
AlL_3^{6-}	-9	1	3	-37.69	1.14×10^{-10}
$\text{Al}_2(\text{OH})_2(\text{HL})_3^{2-}$	-8	2	3	-22.65	4.01×10^{-6}
$\text{Al}_2(\text{OH})_2(\text{HL})_2\text{L}^{3-}$	-9	2	3	-27.81	1.75×10^{-5}
$\text{Al}_2(\text{OH})_2(\text{HL})\text{L}_2^{4-}$	-10	2	3	-32.87	9.61×10^{-5}
$\text{Al}_2(\text{OH})_2\text{L}_3^{5-}$	-11	2	3	-39.56	1.24×10^{-4}
$\text{Al}_4\text{L}_3^{3+}$	-9	4	3	-20.25	2.61×10^{-7}
$\text{Al}_3(\text{OH})_4(\text{H}_2\text{L})^{4+}$	-5	3	1	-12.52	6.51×10^{-5}
Total T^m (M)	pH=5.8	10^{-3}	10^{-3}		
Initial X_0 (M)	1.58×10^{-6}	variable	variable		
Equil. X^* (M)	1.58×10^{-6}	2.03×10^{-5}	2.59×10^{-7}		

TABLE 1. Morel’s Table for the Gallic Acid Test with pH Fixed to 5.8 ([6], [12])

are variable and thermodynamic values come from Brassard and Bodurtha (2000) [6]. By fixing the pH of the system, we notice that the matrix of the stoichiometric coefficients is reduced to a matrix whose all the coefficients are positive, moreover, total concentrations of Al^{3+} and H_3L are positive, then the relation (3.12) becomes, in matrix form

$$\tilde{\mathbf{G}}(\xi) = \xi + \kappa \frac{I_{n_{pm}}}{\mu_{10}} \cdot \left[\log_{10}(T^m) - \log_{10}(\mu_1^T \cdot 10^{\mathbf{K}^m + \mu_1 \xi}) \right]. \quad (4.1)$$

Methods are tested in 2 cases:

case 1 - Initial X_0 (M): $[\text{Al}^{3+}]_0 = 10^{-11}$ M; $[\text{H}_3\text{L}]_0 = 5 \times 10^{-4}$ M

(i.e Initial ξ_0 : $\log_{10}([\text{Al}^{3+}]_0) = -11$; $\log_{10}([\text{H}_3\text{L}]_0) = -3.3010$)

case 2 - Initial X_0 (M): $[\text{Al}^{3+}]_0 = 5.012 \times 10^{-10}$ M; $[\text{H}_3\text{L}]_0 = 10^{-9}$ M.

(i.e Initial ξ_0 : $\log_{10}([\text{Al}^{3+}]_0) = -9.3$; $\log_{10}([\text{H}_3\text{L}]_0) = -9$)

MoMas Benchmark easy test case The MoMaS Benchmark has been designed to compare numerical methods for reactive transport model in 1D and 2D. Different methods for coupling have been used to solve this benchmark. The definition has been published in [13] and results of participants are compared in the synthesis article [14]. It is composed of three subsequent cases with increasing chemical complexity, named "easy", "medium" and "hard". Systems do not represent real chemical systems: they were devised by [5] to create increasing numerical difficulties.

In this work, only the resolution of the chemical equilibrium of the easy test case will be simulated. For the easy case, the chemical system is composed of $n_e = 12$ chemical species distributed as follows: four mobile component species X_1, X_2, X_3 and X_4 ($n_{pm} = 4$), one fixed component species S ($n_{pf} = 1$), five mobile secondary species C_1, C_2, C_3, C_4 and C_5 ($n_{sm} = 5$), and two fixed secondary species CS_1 and CS_2 ($n_{sf} = 2$).

The geometry of the test case is shown in Figure 1.

For the 1D test case, the domain is heterogeneous and composed of two porous media A and B. Medium A is highly permeable with low porosity and low reactivity in comparison with medium B. In order to be close to realistic cases, boundary and initial conditions are not expressed for fundamental variables, i.e., component concentrations. Indeed, chemical analysis can provide quite easily a measure

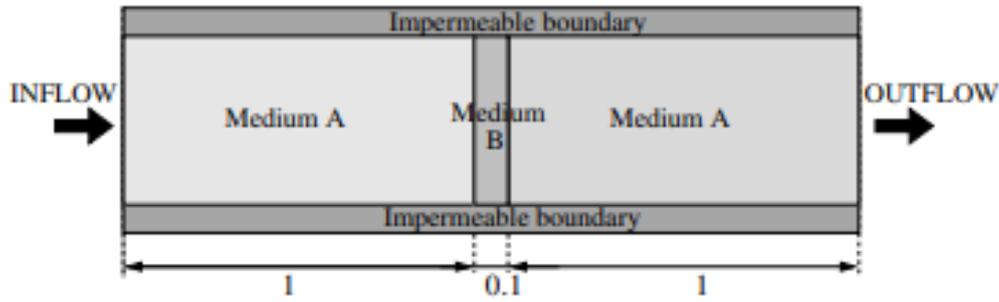


FIGURE 1. Geometry of the domain.

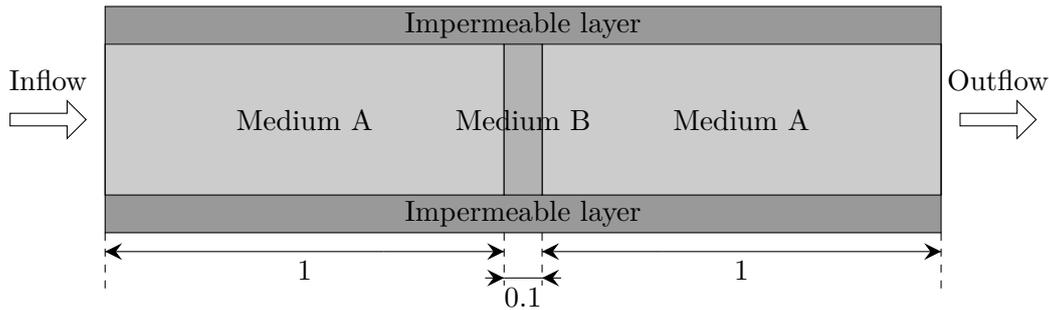


FIGURE 2. Geometry of the domain.

of the total concentration or of the total dissolved concentration for each component. An injection is made on the left side of the domain, followed by leaching on the same side. The injection period corresponds to specific inflow concentrations depending on the MoMas easy test case. All injection periods are 5000 s long. The leaching periods are at least 1000 s long. If needed, leaching period can be extended after 1000 s to reach the following condition: at the end of leaching period, 99,9% of injected pollutant (X_1 , X_3 and S) has been removed from the domain. Imposed concentrations for the inflow boundary are

$$T_j(x=0, t) = T_j^{inj} \quad t < 5000 \text{ s} \quad T_j(x=0, t) = T_j^{leach} \quad t > 5000 \text{ s}$$

These chemical species interact through $n_r = 7$ equilibrium reactions shown in Table 2.

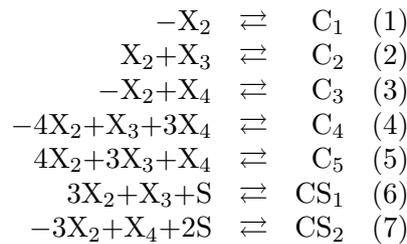


TABLE 2. Chemical reactions for MoMas easy test case

The domain is initially at a local equilibrium with the surface component S in the presence of mobile components X_2 and X_4 . During injection, component X_4 will be removed. Component X_1 is a perfect tracer; X_2 and X_3 will react together, with the surface S and with X_4 still present. During leaching, X_1 and X_3 will be removed. X_2 and X_4 will react with the surface S and with X_3 still present.

In this test case, the equilibrium reaction constant is of the order of 10^{35} which makes the system very rigid and already presents a great challenge. In addition, stoichiometric coefficients are quite large. This allows us to test the robustness of our implementation in the face of such complexity. Table 3 give the stoichiometric coefficients for mass action laws and conservation equations.

The resolution of the thermodynamic equilibrium of this test is carried out at each period and results obtained correspond well to expected results. But first of all it is necessary to put the two medium A

Species	X ₁	X ₂	X ₃	X ₄	S	K
C ₁	0	-1	0	0	0	10 ⁻¹²
C ₂	0	1	1	0	0	1
C ₃	0	-1	0	1	0	1
C ₄	0	-4	1	3	0	0.1
C ₅	0	4	3	1	0	10 ³⁵
CS ₁	0	3	1	0	1	10 ⁶
CS ₂	0	-3	0	1	2	10 ⁻¹
Total concentration T	T ₁	T ₂	T ₃	T ₄	T _S	
Initial conditions						
zone A	0	-2	0	2	1	
zone B	0	-2	0	2	10	
Boundary conditions						
Injection $t \in [0, 5000]$	0.3	0.3	0.3	0	0	
Leaching $t \in [5000, \dots]$	0	-2	0	2	0	

TABLE 3. Equilibrium for MoMas easy test case [13]

and B at local equilibrium, then the chemical equilibrium during injection and leaching periods will be resolved.

4.2. Numerical results and comparisons

Acid Gallic test The thermodynamic chemical equilibrium problem (4.1) is studied without relaxation ($\kappa = 1$).

For each case defined above, we notice the convergence of Anderson Acceleration method, for a maximal depth $m \geq 1$ (cf. Figure 2):

- for $\omega_0 = \log_{10}((10^{-11}, 5 \times 10^{-4})^T)$, i.e in case 1, the convergence of Anderson(m) requires 26 iterations for $m = 1$ and 16 iterations for $m \geq 2$.
- for $\omega_0 = \log_{10}((5.012 \times 10^{-10}, 10^{-9})^T)$, i.e in case 2, the convergence of Anderson(m) requires 109 iterations for $m = 1$ and 15 iterations for $m \geq 2$.

The first iterations performed present disturbances in terms of the variation in concentrations [Al³⁺] et [H₃L], but these disturbances are no complicated and convergence has been obtained without any difficulty. These disturbances result mainly from the choice of the initial concentration of each component. Note that the obtained solution $\omega^* = (-4.6930, -6.5870)^T = \log_{10}((2.028 \times 10^{-5}, 2.6 \times 10^{-7})^T)$ is the same reference solution obtained by J. Carrayrou [12] and cited in Table 1. It is thus numerically

	CPU time (s)
Anderson ($m = 1$)	1.11
Anderson ($m = 2$)	1.3
Anderson ($m = 3$)	1.13

TABLE 4. Gallic acid test, thermodynamic equilibrium by Anderson Acceleration method: CPU time (s)

established that Anderson Acceleration method converges towards the solution in a short computation time (CPU time) (cf. Table 4) for different values of the maximal depth m , $m = 1, 2, 3$.

Since a condition-number monitoring strategy is used in Anderson acceleration, it is not necessary to worry about the condition number becoming problematically large. In this Anderson-acceleration implementation, the condition number is monitored to ensure stability and robustness. The tolerance for the condition-number monitoring is 10^{10} . We can see in Figure 3 that, for $m = 1, 2$, the condition number remains less than 10^{10} and it becomes more than 10^{15} for $m \geq 3$ if there is no condition-number monitoring. This shows the specific effects of condition-number monitoring. With condition-number monitoring, $\mathbf{cond}(\mathcal{F}_3)$, at iteration step 3, is initially greater than 10^{10} (\mathcal{F}_k is defined by (3.3)). However, after using the MATLAB's `qrdelete` function, the condition number is less than 10^{10} , and the convergence succeeds after 3 iterations.

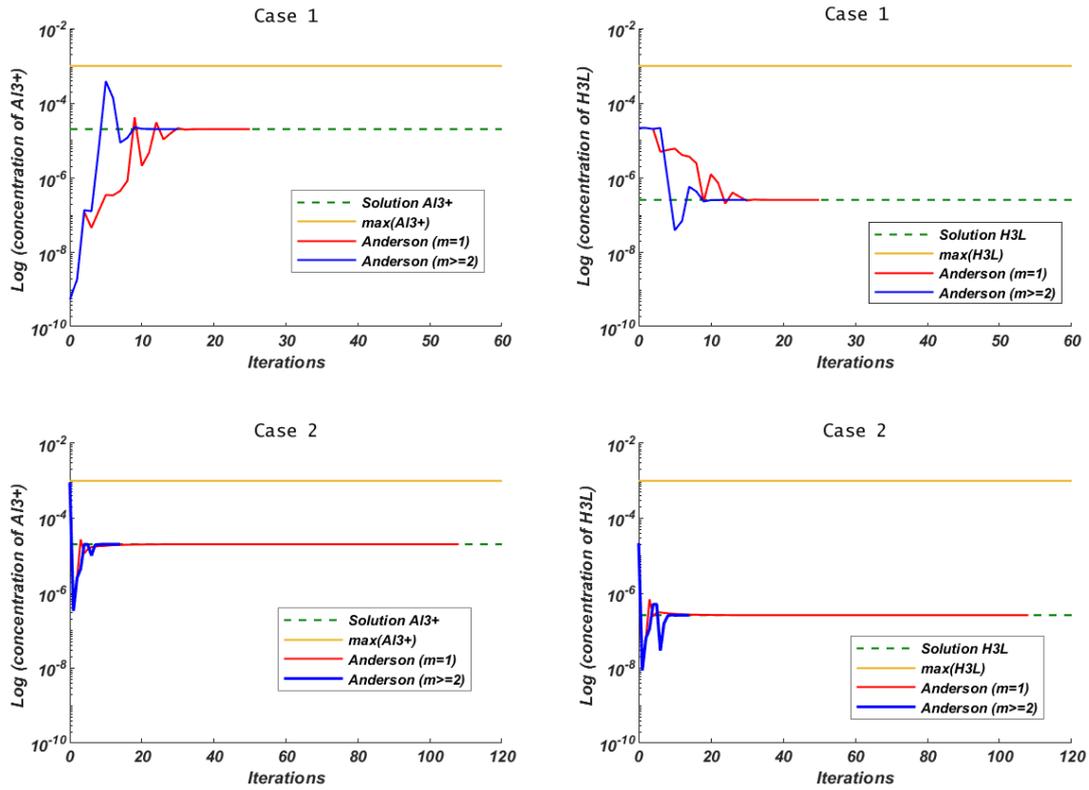
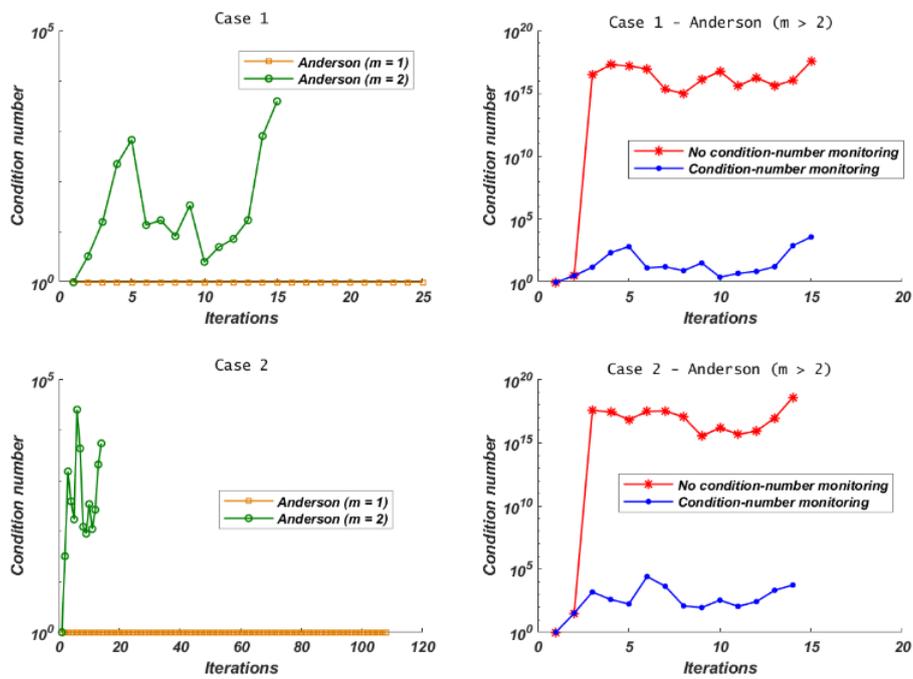


FIGURE 3. Acid Gallic test: Thermodynamic equilibrium for the components H_3L and Al^{3+} with Anderson Acceleration



At
Gc

FIGURE 4. Acid Gallic test: Thermodynamic equilibrium by Anderson Acceleration - Condition number curve.

In addition, from Figure 4 we can see that for $m \geq 2$, Anderson Acceleration method begins to accelerate the convergence of the AA iterates after 12 or 13 iterations in the both case, but it is clear

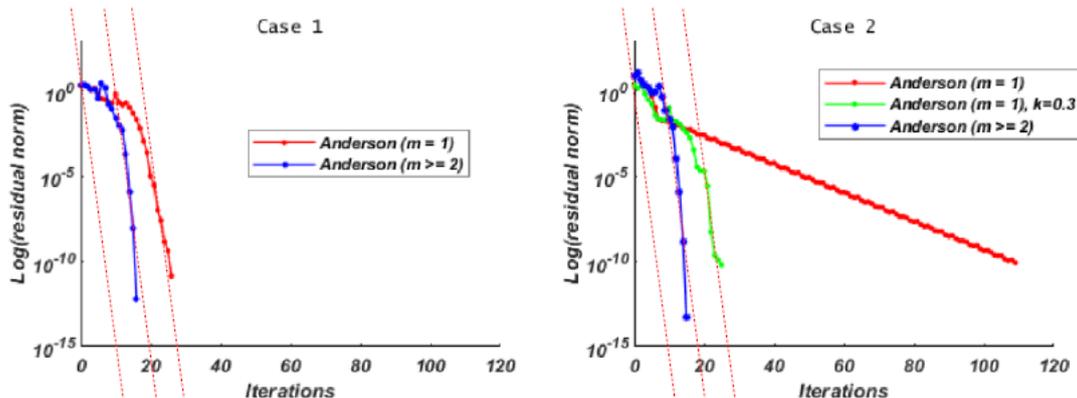


FIGURE 5. Acid Gallic test: Thermodynamic equilibrium by Anderson Acceleration - Residual norm curve.

that results slowly decrease when $m = 1$ especially in the second case where convergence requires more than 100 iterations. More precisely, if we now take $\kappa = 0.3$, Anderson ($m = 1$) converges faster and the residual norm decreases requiring 26 iterations instead of 109, to be less than 10^{-10} . On the other hand, the "theoretical" slopes of Newton are plotted on the Figure 4. These slopes show that a convergence of order 2 is reached, which confirms that the AA method works really well.

To apply the two methods MPE and RRE, we generate the vectors $\omega_1, \omega_2, \dots$ by (3.11) with different values of the parameter κ , $\kappa \in \{0.1, 0.45, 0.5, 0.6\}$. These values of κ are a good choice for this experience, but that does not mean that these parameters cannot take another value

Figures 5 and 6 show the behavior of the residual norm, using a logarithmic scale. It contains some of the residual history obtained by applying RRE and MPE in cycling mode with

- $(Kmax, N_0, N) = (10, 20, 10), (20, 0, 10), (10, 20, 0)$ for $\kappa = 0.1$,
- $(Kmax, N_0, N) = (10, 10, 15), (10, 10, 10)$ for $\kappa = 0.45$,
- $(Kmax, N_0, N) = (10, 5, 15), (20, 5, 15)$ for $\kappa = 0.5$,
- $(Kmax, N_0, N) = (15, 15, 15)$ for $\kappa = 0.6$.

Convergence behavior is overall linear, but a small marginally unstable mode is observed that corresponds to almost a periodic half-sinusoidal oscillation of residuals. With $\kappa = 0.1$, the first choice of the three data is the best because it yields the fastest convergence of the residual error. By comparing the first and third choices of the data $Kmax, N_0$ and N , we notice that performing a few basic iterations before MPE or RRE is applied in each cycle after the first cycle results in a faster and more stable convergence as well as more stable behavior of the residual norm. Such a result is observed again in Figure 6 with $\kappa = 0.45$, $\kappa = 0.5$ and $\kappa = 0.6$. These last three values of κ still give a fast convergence represented by an almost stable decrease of the residual norm.

In addition, the CPU time required for all iterations in all cycles, for each case and each method, is very short, not exceeding 1 second (see Table 5 for case 1). Therefore, for this chemical test, solving

	$(Kmax, N_0, N)$	MPE	RRE
$\kappa = 0.1$	(10,20,10)	0.6406	0.8594
	(20,0,10)	0.6719	0.4688
	(10,20,0)	0.8594	0.3906
$\kappa = 0.45$	(10,10,10)	0.2656	0.2031
	(10,10,15)	0.4531	0.2969
$\kappa = 0.5$	(10,5,15)	0.2656	0.2656
	(20,5,15)	0.3906	0.5

TABLE 5. Acid Gallic test, thermodynamic equilibrium by restarted RRE and MPE: CPU time (s)

the thermodynamic equilibrium using Anderson's acceleration method and the two polynomial methods MPE and RRE works well. The convergence is obtained in a very short computation time and a

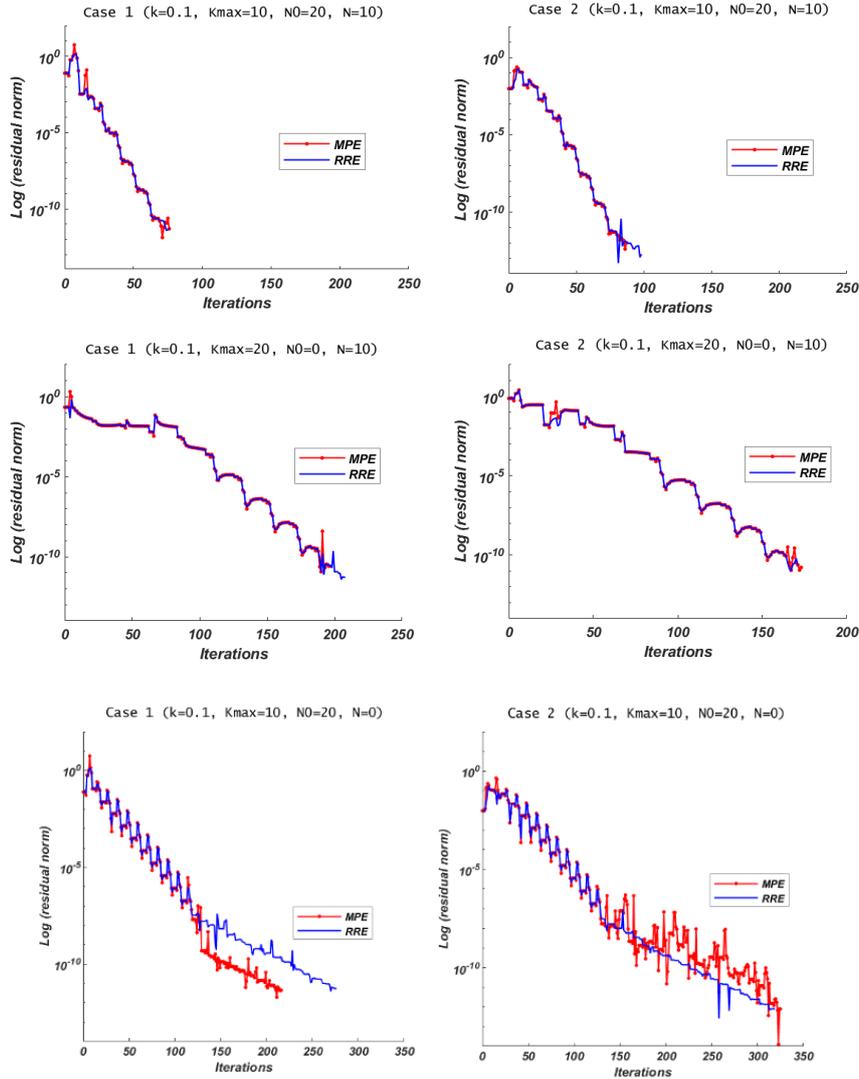


FIGURE 6. Acid Gallic test: Thermodynamic equilibrium by restarted MPE and RRE - Residual norm curve

reasonable number of iterations, but Anderson's method seems to be more stable than the MPE and RRE methods, and requires fewer iterations. This results from the cycling strategy applied to the MPE and RRE methods, while the Anderson Acceleration method does not require such a strategy. After having successfully calculated the equilibrium concentrations of the component species H_3L and Al^{3+} by the methods (AA), (MPE) and (RRE), we can then calculate the concentrations of equilibrium of the other secondary species from equations in (2.9). Equilibrium concentrations are reported in the last column of Table 1.

Benchmark MoMas easy test case In this part, numerical results for the resolution of the thermodynamic equilibrium of the easy MoMas test case are presented. First the chemical equilibrium in each medium (A and B) is solved, then the chemical equilibrium for the injection period of the component X_3 , before looking for the chemical equilibrium of the leaching period. We assume, for example, that the initial concentrations of the component species in each of the two mediums A and B are given by the vector $\omega_{A,B,0} = \log_{10}(\mathcal{X}_{A,B,0})$ where

$$\mathcal{X}_{A,B,0} \text{ (M)} = (X_{1,0}, X_{2,0}, X_{3,0}, X_{4,0}, S_0)^T = (0.3, 0.4, 10^{-11}, 0.21, 0.6)^T.$$

The convergence of Anderson acceleration method is tested for any strictly positive value of the maximum depth m . The fixed point problem (3.11) is implemented without relaxation ($\kappa = 1$). The thermodynamic equilibrium in A is reached after 43 iterations for $m = 1$, 26 iterations for $m = 2$

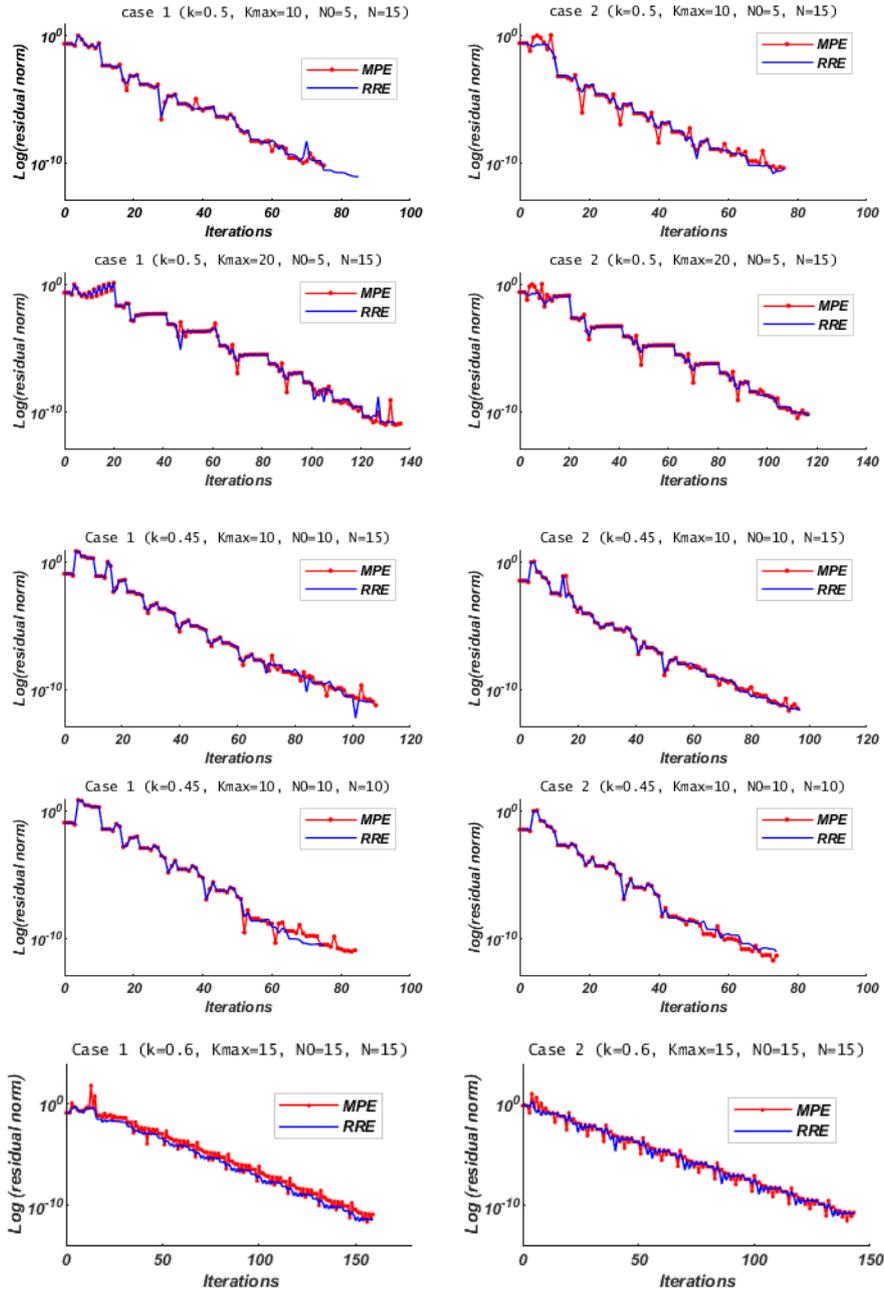


FIGURE 7. Acid Gallic test: Thermodynamic equilibrium by restarted MPE and RRE with $\kappa = 0.45$, $\kappa = 0.5$ and $\kappa = 0.6$ - Residual norm curve

and 21 iterations for $m \geq 3$ (cf. Figure 7). In B, the convergence of Anderson's method requires 51 iterations for $m = 1$, 30 iterations for $m = 2$, 22 iterations for $m = 3$ and 21 iterations for $m \geq 4$ (cf. Figure 8). In Figures 7 and 8, no complicated oscillation phenomenon is observed, therefore, convergence is achieved without difficulty. Note that concentrations of components at thermodynamic equilibrium in the two mediums A and B are defined respectively by the two vectors:

$$\begin{aligned} \mathcal{X}_A^*(M) &= (X_{1,A}^*, X_{2,A}^*, X_{3,A}^*, X_{4,A}^*, S_A^*)^T = (10^{-20}, 0.2597, 10^{-20}, 0.3495, 0.3907)^T \\ \mathcal{X}_B^*(M) &= (X_{1,B}^*, X_{2,B}^*, X_{3,B}^*, X_{4,B}^*, S_B^*)^T = (10^{-20}, 1.5116, 10^{-20}, 0.5756, 7.9128)^T \end{aligned}$$

(or $\omega_A^* = (-20, -0.5855, -20, -0.4565, -0.4081)$ and $\omega_B^* = (-20, 0.1794, -20, -0.2399, 0.8983)$ in \log_{10}). The influence of the most reactive medium B is demonstrated by the higher concentration of S reached at equilibrium.

For the injection period, we consider two cases, injection in medium A (on the left side) and injection in medium B. We previously mentioned that the domain is initially in equilibrium with the species S in the presence of mobile species X_2 and X_4 . Therefore, for this period, we consider as an initial

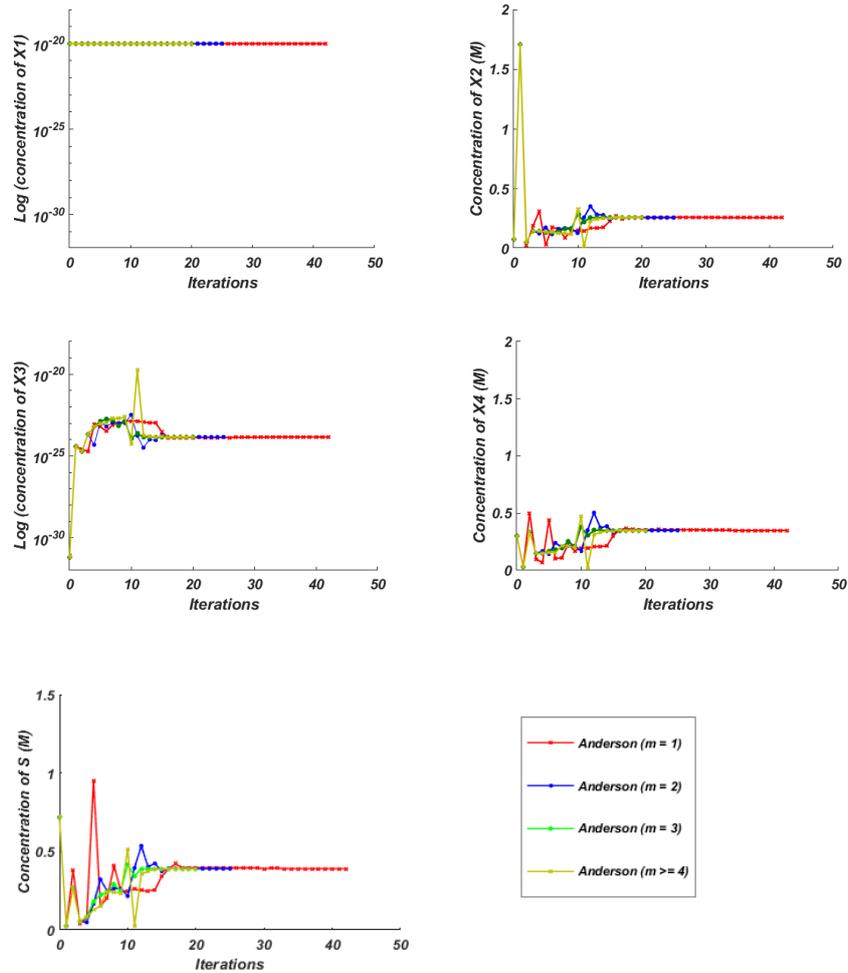


FIGURE 8. MoMas easy test: Thermodynamic equilibrium in medium A by Anderson Acceleration

approximation of the concentrations the thermodynamic equilibrium solution in each medium, namely

$$\omega_{inj,0} = \begin{cases} \omega_A^* & \text{injection in medium A} \\ \omega_B^* & \text{injection in medium B} \end{cases}$$

Figure 9 and 10 present the behavior of the components concentrations until reaching thermodynamic equilibrium for the injection period in the two zones A and B by Anderson(m) for all $m > 0$. A small number of iterations is necessary to achieve convergence. Note that the equilibrium is reached when the vector of components concentrations is

$$\mathcal{X}_{inj}^*(M) = (X_{1,inj}^*, X_{2,inj}^*, X_{3,inj}^*, X_{4,inj}^*, S_{inj}^*)^T = (0.3, 0.2416, 0.2416, 10^{-50}, 10^{-23})^T$$

(or $\omega_{inj}^* = (-0.5229, -0.6169, -0.6169, -50, -23)^T$ in \log_{10}). Once the equilibrium is reached, we notice that the component X_4 is removed to be washed from the domain, X_2 and X_3 remain present in the domain to interact with the surface S and X_4 . X_1 is a tracer, which is why its concentration remains constant at 0.3 M.

Leaching follows injection on the same side once the period of 5000 s has passed. To solve the thermodynamic equilibrium for leaching period, an initial approximation the solution of the thermodynamic equilibrium for the injection period is considered, i.e $\mathcal{X}_{leach,0} = \mathcal{X}_{inj}^*$ (or $\omega_{leach,0} = \omega_{inj}^*$).

Figure 11 shows that, with Anderson's method, the thermodynamic equilibrium of the chemical system describing the leaching period is obtained after 61 iterations by Anderson and $m = 2$ and 39 iterations by Anderson and $m = 3$. However, after some numerical tests, we find that for $m = 1$ and $m \geq 4$, Anderson's method does not converge or sometimes presents convergence difficulties. However, after some numerical tests, we find that for $m = 1$ and $m \geq 4$, Anderson's method does not converge or sometimes presents some convergence difficulties. That comes back, perhaps, to the choice of the initial concentrations. To overcome these difficulties, we consider the fixed point problem (4.1) with

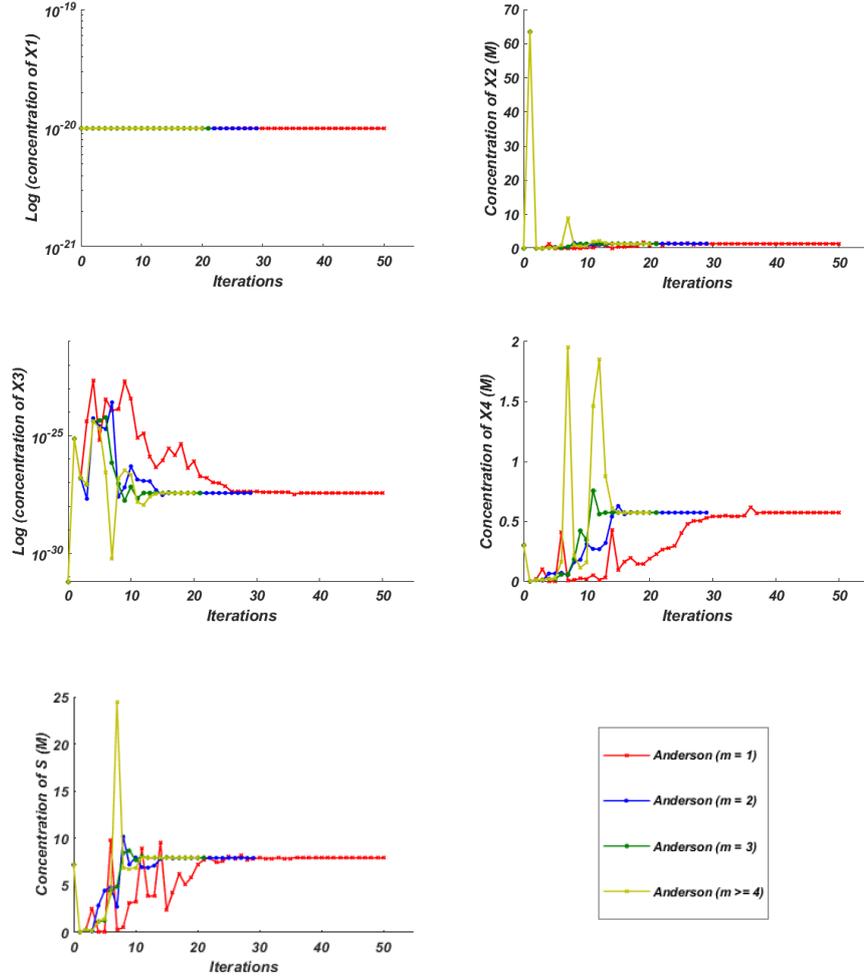


FIGURE 9. MoMas easy test: Thermodynamic equilibrium in medium B by Anderson Acceleration method

relaxation, i.e $\kappa \neq 1$. This parameter is chosen arbitrarily until it is well suited to achieve convergence without difficulty. The convergence of Anderson ($m = 1$), Anderson ($m = 4$), Anderson ($m = 5$) and Anderson($m > 5$) is achieved, in an appropriate number of iterations, taking respectively $\kappa = 0.29$, $\kappa = 0.55$, $\kappa = 0.46$ and $\kappa = 0.42$. The behavior of components concentrations is presented in Figure 11. The solution obtained is

$$\begin{aligned} \mathcal{X}_{leach}^*(M) &= (X_{1,leach}^*, X_{2,leach}^*, X_{3,leach}^*, X_{4,leach}^*, S_{leach}^*)^T \\ &= (10^{-20}, 5, 7735.10^{-7}, 7.223.10^{-27}, 1.1547.10^{-6}, 10^{-20})^T \end{aligned}$$

(or $\omega_{leach}^* = (-20, -6.2382, -26.1413, -5.9372, -20)^T$ in \log_{10}). We notice that, at the equilibrium of leaching period, 99,9% of the injected pollutant are removed from the domain.

These results again show that the AA method works well for every period and every domain. In addition, the convergence is very fast, requiring a very short computation time (CPU time) not exceeding 2 s. The CPU execution time required by Anderson acceleration method, for several values of the maximal depth m ($m = 1, 2, 3, 4$) to solve the thermodynamic equilibrium in each zone A and B is given in Table 6. This time is given during the two periods of injection and leaching.

On other hand, for the equilibrium in medium A and B, we notice that for $1 \leq m \leq 4$, $\mathbf{cond}(\mathcal{F}_k)$ always remains less than 10^{10} , for any k -th iteration, and it becomes more than 10^{15} for $m \geq 5$ at iteration step 6. With the strategy of condition-number monitoring, $\mathbf{cond}(\mathcal{F}_k)$ returns less than 10^{10} for $k \geq 6$ (cf. Figures 12 and 13). Similarly, by solving the equilibrium system for the injection period in medium A using AA ($m = 1, 2, 3$), $\mathbf{cond}(\mathcal{F}_k)$ always remains lower than 10^{10} , for any k -th iteration. However, for $m = 4$, at the 13-th iteration, it becomes greater than 10^{10} (cf. Figure 14). By apply-

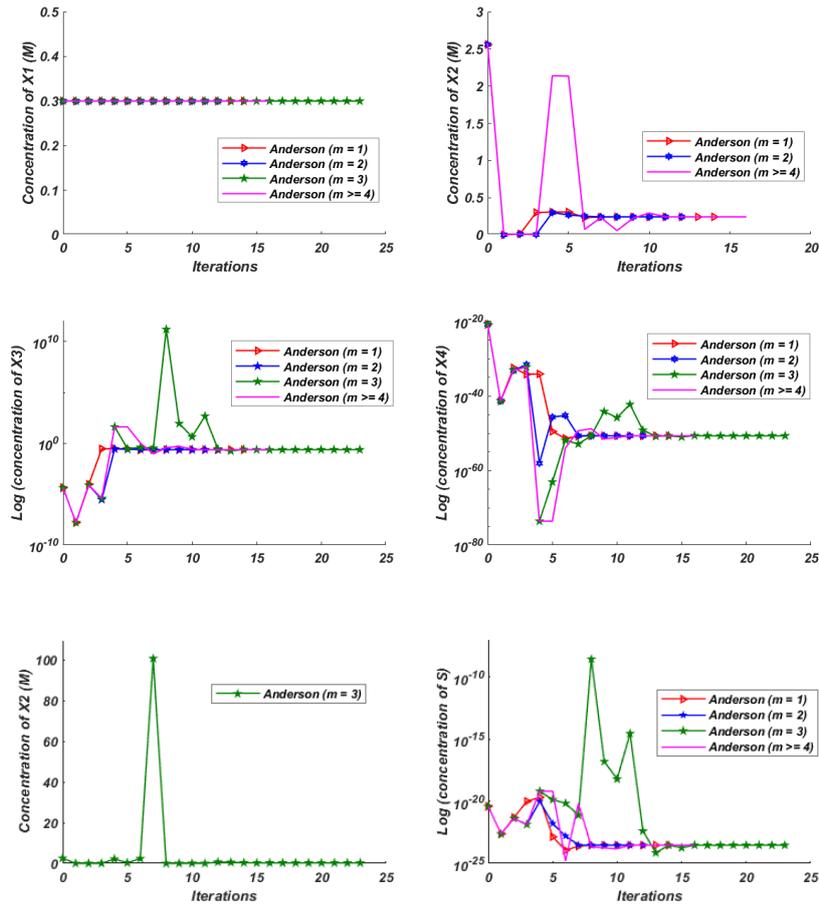


FIGURE 10. MoMas easy test, Injection in medium A: Thermodynamic equilibrium by Anderson Acceleration.

	Anderson (m)	CPU time (s)
Zone A	$m = 1$	1.1875
	$m = 2$	1.0156
	$m = 3$	1.0781
	$m = 4$	1.2031
Zone B	$m = 1$	1.6719
	$m = 2$	1.3281
	$m = 3$	1.4063
	$m = 4$	1.6563
Injection period	$m = 1$	1.0156
	$m = 2$	1.0625
	$m = 3$	1.1094
	$m = 4$	1.9844
Leaching period	$m = 1$	1.3594
	$m = 2$	1.4063
	$m = 3$	1.2656
	$m = 4$	1.0313

TABLE 6. MoMas easy test, thermodynamic equilibrium by Anderson Acceleration method: CPU time.

ing the strategy of condition number monitoring to the matrix \mathcal{F}_{13} , $\text{cond}(\mathcal{F}_k)$ drops below 10^{10} for $k \geq 13$ and the convergence is reached after 13 iterations. Likewise, for $m \geq 5$, the same strategy is applied to the matrix \mathcal{F}_6 . In the same way, for the equilibrium chemical systems during the injection period in medium B and the leaching period, the behavior of the condition number of the matrix

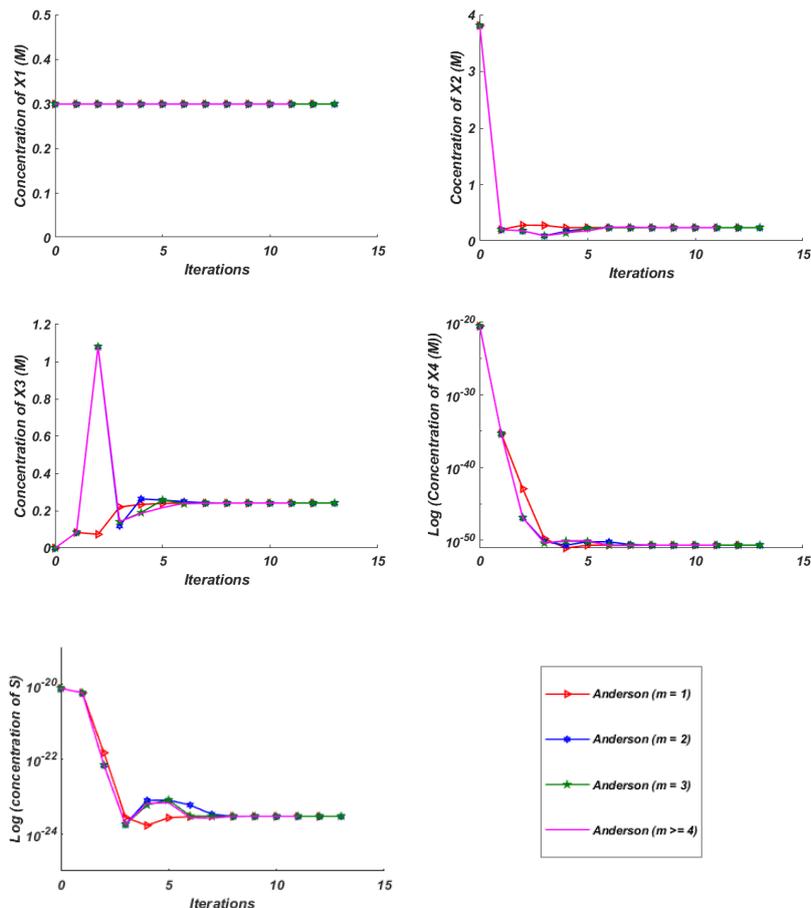


FIGURE 11. MoMas easy test, Injection in medium B: Thermodynamic equilibrium by Anderson Acceleration.

\mathcal{F}_k with the strategy of condition number monitoring is described in Figures 15 and 16. Note that for the leaching period equilibrium system, the strategy of condition number monitoring is applied at several iteration steps k , to matrices \mathcal{F}_k , where $k \in \{11, 12, 13, 15, 17, 18, 22\}$ with Anderson($m=5$) and $6 \leq k \leq 16, 18 \leq k \leq 20$ with Anderson($m \geq 5$). Therefore, this monitoring strategy has largely contributed to increasing the robustness and stability of the Anderson algorithm.

Figure 17 shows convergence plots with the approximate "theoretical" slopes of Newton for all the cases of MoMas easy test with Anderson method ($m=1, 2, 3, 4, 5$). For the equilibrium in zones A and B, Anderson($m=1$) requires twice as many iterations as Anderson($m=3, 4$). Taking these results into account, the slopes prove the order 2 convergence of the AA method and again demonstrate its efficiency.

We also apply MPE and RRE methods in cyclic mode (through their new implementations described in previous sections) to nonlinear system of thermodynamic equilibrium of MoMas easy test case. The vectors $\omega_1, \omega_2, \dots$ is generated by (3.11), where the mixing parameter κ is not the same for all the cases of this test. It is chosen arbitrarily to ensure convergence in the most efficient way.

We solve the chemical equilibrium system in medium A by restarted MPE and RRE methods by taking $\kappa = 0.4$. Note that the initial concentrations of component species, for equilibrium system in zones A and B, is given by the vector $\omega_{A,B,0}$ defined above.

The computer program code was run by taking the maximum number of iterations in each cycle $Kmax = 10$. Several choices for the couple (N, N_0) : $(N, N_0) = (0, 20), (5, 20), (20, 0), (10, 0), (5, 0)$ are considered. We remind that N_0 is the number of iterations performed before cycling is started, e.g, before MPE or RRE is applied for the first time and N is the number of iterations performed before one of this methods is applied in each cycle after the first cycle.

Figure 18 shows the evolution of the nonlinear residual norm, using a logarithmic scale for the restarted MPE and RRE methods. It appears that for $(N, N_0) = (0, 20), (5, 20)$, methods give convergence to

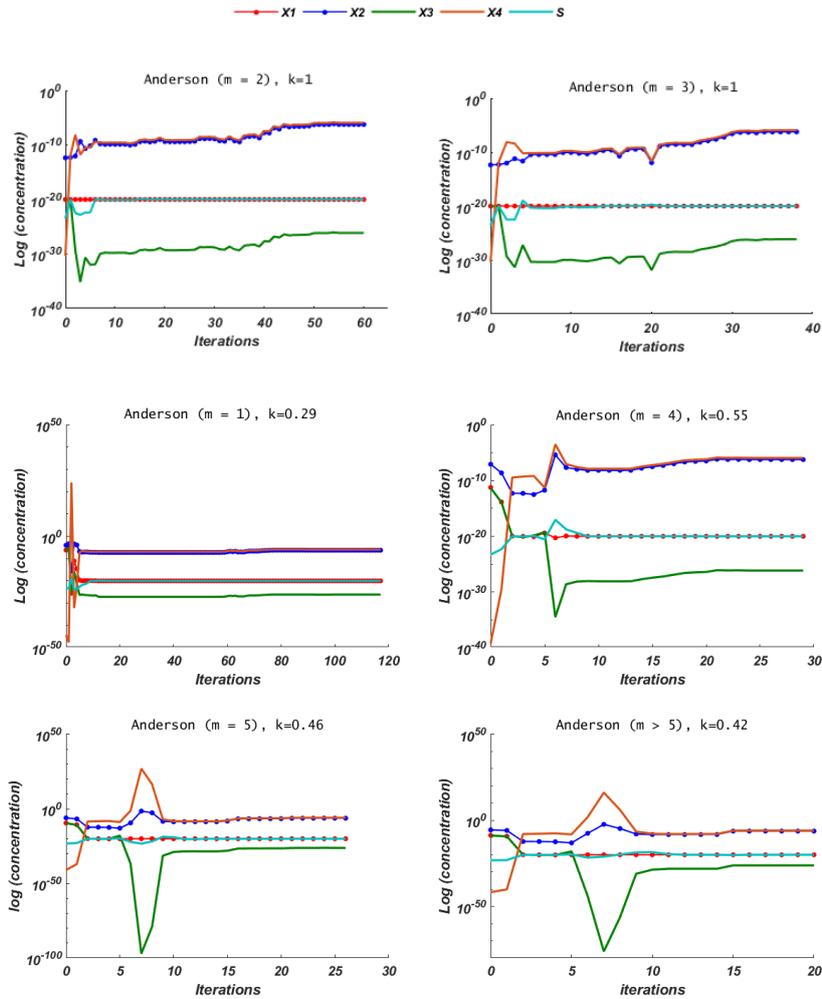


FIGURE 12. MoMas easy test, leaching period: Thermodynamic equilibrium by Anderson Acceleration.

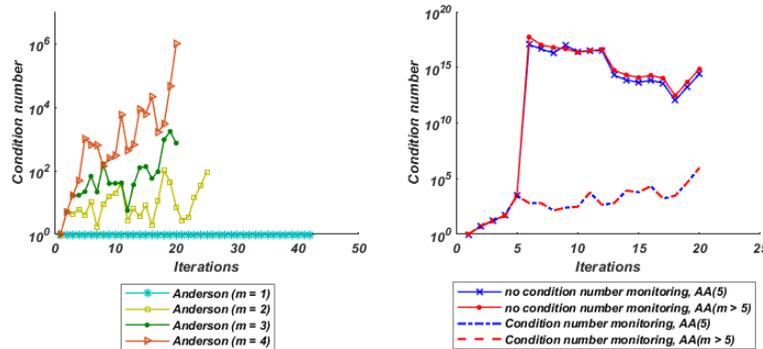


FIGURE 13. MoMas easy test, thermodynamic equilibrium in medium A by Anderson Acceleration method - Condition number curve.

a steady state. Moreover, when performing a certain number of iterations before the application of RRE or MPE to each cycle after the first cycle ($N = 5$), the residual norm decreases more rapidly and convergence is reached in a number reduced iterations. For $(N, N_0) = (20, 0)$, MPE converges faster than RRE with few disturbances described by a residual increase between iterations 15 and 20. Finally, we notice that, in the two last cases, RRE and MPE seem to perform similarly and the convergence seems stable.

Regarding the equilibrium in medium B, we propose to take $\kappa = 0.3$. The maximal number of iterations K_{max} remains constant equal to 10. Figure 19 contains part of the residual history obtained,

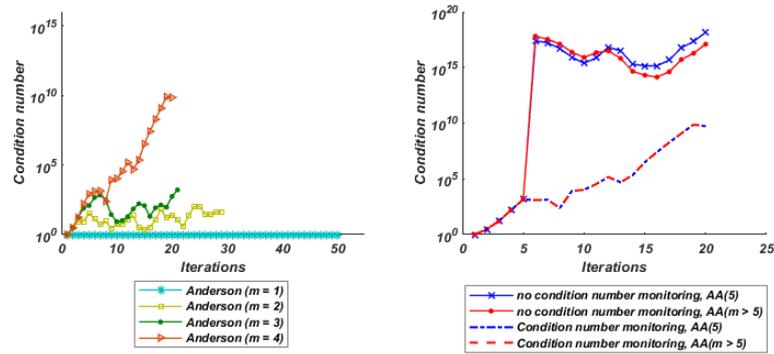


FIGURE 14. MoMas easy test: Thermodynamic equilibrium in medium B by Anderson Acceleration method - Condition number curve.

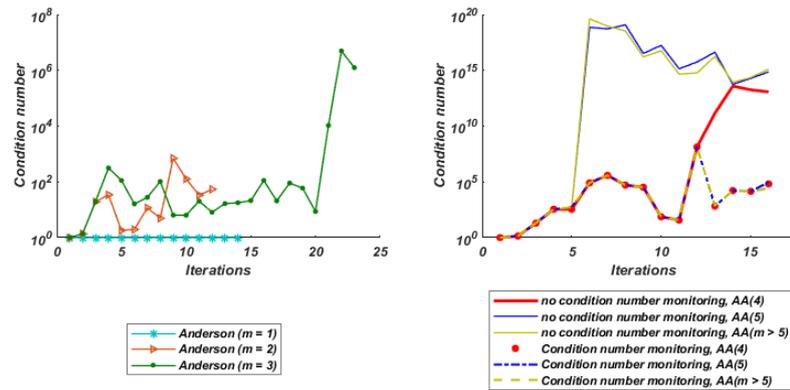


FIGURE 15. MoMas easy test, Injection in medium A: Thermodynamic equilibrium by Anderson Acceleration method - Condition number curve.

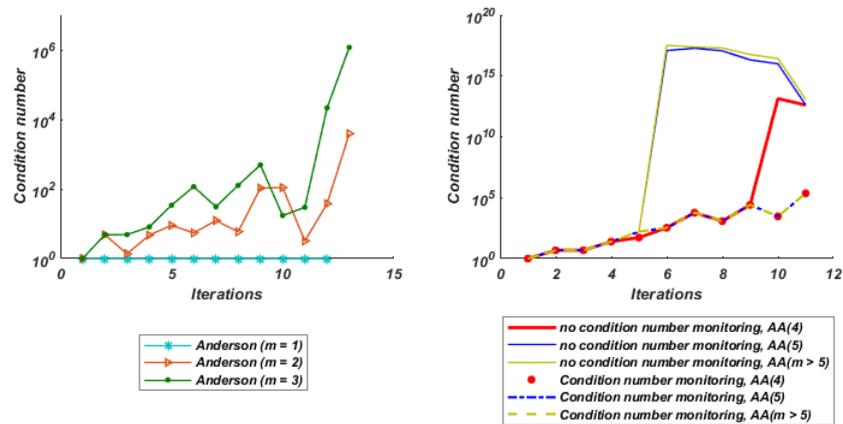


FIGURE 16. MoMas easy test, Injection in medium B: Thermodynamic equilibrium by Anderson Acceleration method - Condition number curve.

for $(N, N_0) = (15, 0)$ and $(N, N_0) = (0, 80)$. It can be seen that, for $(N, N_0) = (0, 80)$, it takes close to 300 iterations to reach the prescribed level of convergence. However, the better case is shown with $(N, N_0) = (15, 0)$ where the RRE and MPE algorithms take approximately 90 and 80 iterations respectively to reach the prescribed levels of residual. Let us emphasize that the numerical experiments carried out for this example show that if one takes $N = 0$, it is necessary to carry out a large number N_0 of iterations before starting the cycling to reach convergence.

Let us now study the thermodynamic equilibrium systems for the injection in media A and B once by taking $\kappa = 0.2$ and once by taking $\kappa = 1$ (i.e. without relaxation). Looking now at Figures 20 and 21, it appears that, with the same $Kmax, N_0$ and N , MPE and RRE seem to work very similarly in this example and achieved the same accuracy with both the injection into medium A and with the

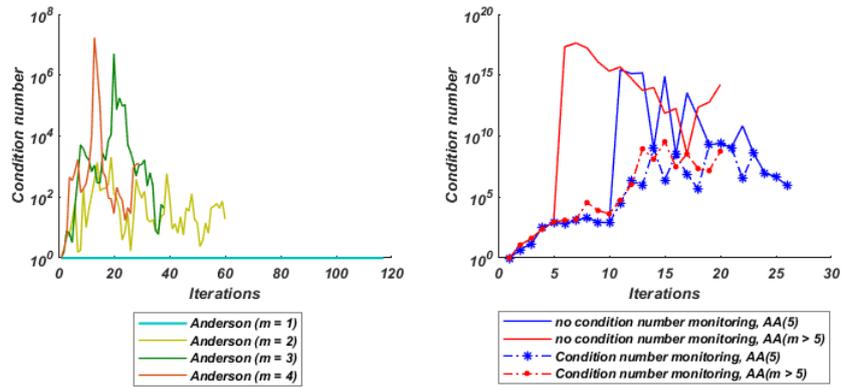


FIGURE 17. MoMas easy test, Leaching period: Thermodynamic equilibrium by Anderson Acceleration method - Condition number curve.

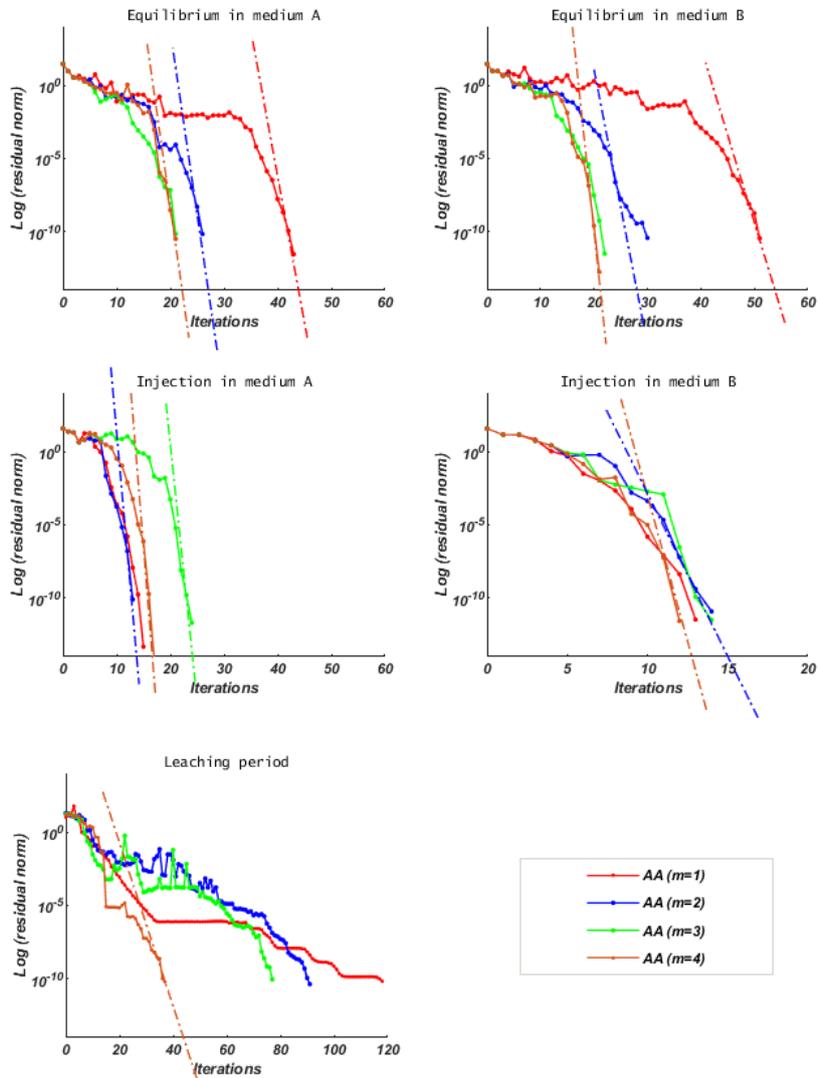


FIGURE 18. MoMas easy test, thermodynamic equilibrium by Anderson Acceleration method -Residual curve.

injection in medium B, except for the choice of $N_0 = 25, N = 0$ in Figure 21. This choice presents a difference between results of two methods: the result of the MPE method decreases compared to the result of RRE, from the fifth cycle and once again resumes its stable and constant trajectory. On the other hand, by taking a relaxation parameter different of 1, we notice that this difference disappears and the two methods again give almost the same results (see Figure 20). Otherwise, keeping $\kappa = 1$,

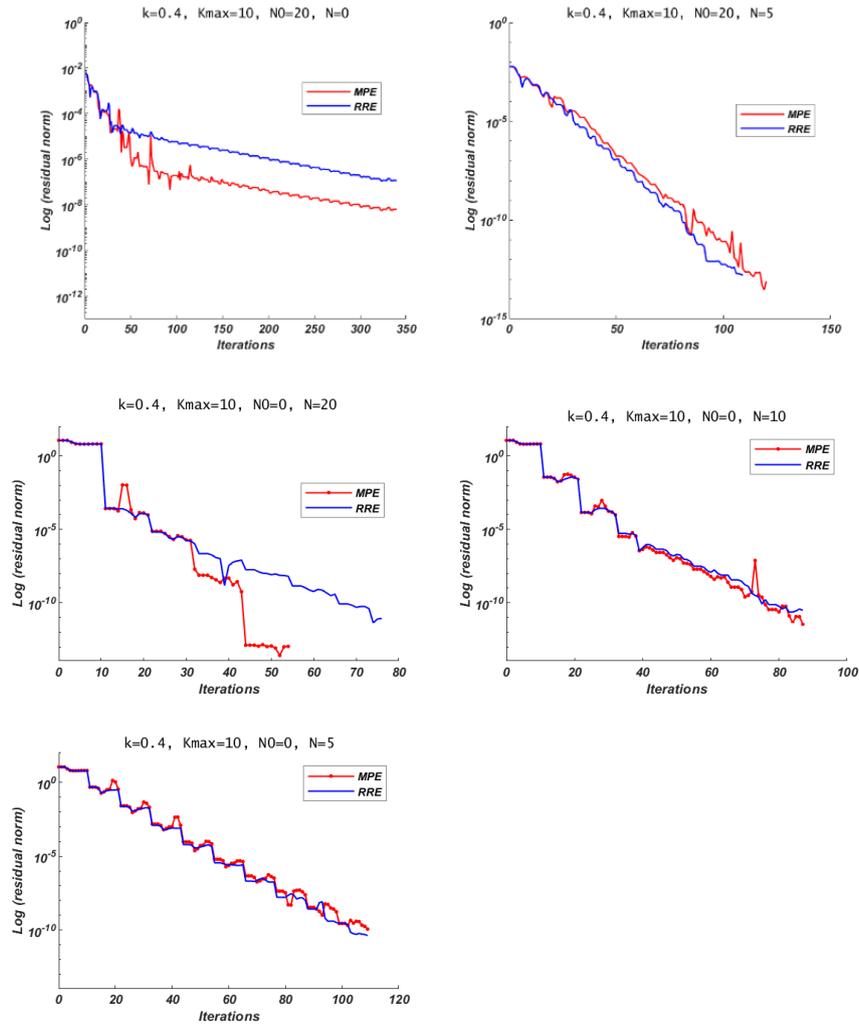


FIGURE 19. MoMas easy test: Thermodynamic equilibrium in the medium A by restarted MPE and RRE, with $\kappa = 0.4$ - Residual norm curve

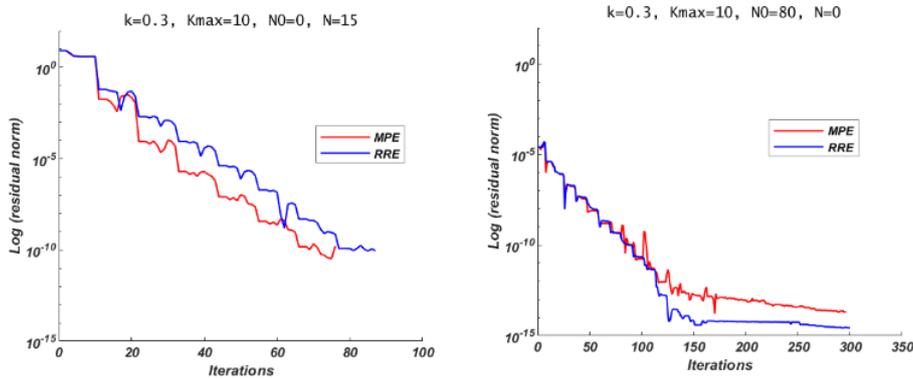


FIGURE 20. MoMas easy test: Thermodynamic equilibrium in the medium B by restarted MPE and RRE, with $\kappa = 0.3$ - Residual norm curve

this is achieved again by taking the parameter N non equal to zero ($N = 2$) (cf. Figure 22). The parameters involved in the numerical computation must therefore be chosen with care so that MPE and RRE give consistent results.

Moreover, the number of cycles is reduced when the value of N is increasing, hence taking the time

overhead of cycling into account saves CPU time. This observation confirms the effectiveness concerning the strategy of cycling [44]. Note that for $\kappa = 1$, the convergence is much faster than for $\kappa = 0.2$. A perfect result is obtained for this experiment with $(N, N_0) = (0, 25), (0, 18)$. The prescribed level of convergence for the injection in media A and B respectively is reached very quickly from the first iteration. In addition, with $(N, N_0) = (0, 28), (0, 24)$, the residual norm sometimes seems to be worth a constant lower than the tolerance indicated at this level.

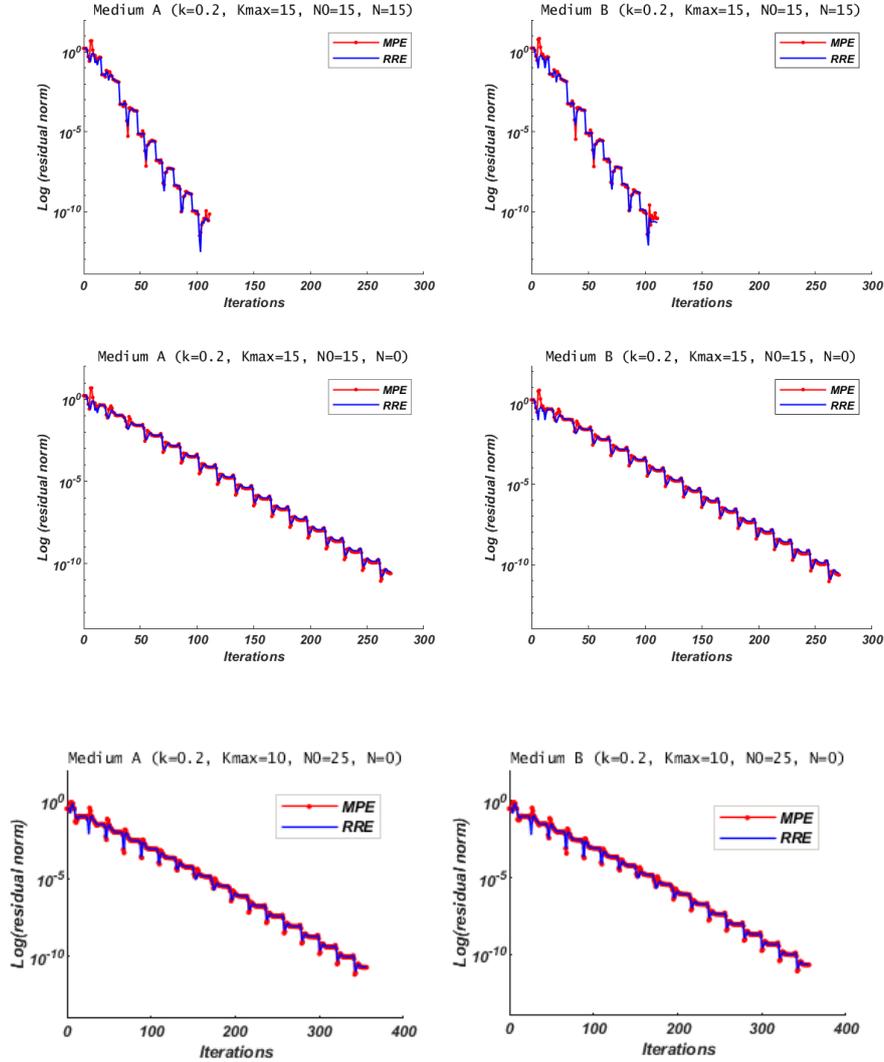


FIGURE 21. MoMas easy test: Thermodynamic equilibrium for injection in media A and B by restarted MPE and RRE, with $\kappa = 0.2$ - Residual norm curve

Finally, it remains to present the resolution of thermodynamic equilibrium by the restarted methods RRE and MPE for the leaching period provided that initial concentrations of components are defined by the solution vector of the equilibrium system for the injection period. We take κ equal to 0.495. This choice for the parameter κ is the best to reach convergence even if it causes difficulties for the convergence. An unstable mode is observed that corresponds to an oscillation of residuals (cf. Figure 23). For $(Kmax, N_0, N) = (10, 0, 10)$, MPE and RRE need the same number of iterations and cycles for convergence, however, for $(Kmax, N_0, N) = (10, 10, 10)$, MPE converges faster than RRE. But, the latter accelerates convergence for $(Kmax, N_0, N) = (10, 5, 18)$ more than MPE. This example appears to be a critical case in that no convergence rule can be deduced by varying the values of N and N_0 .

Consequently, to solve the thermodynamic equilibrium system of the leaching period, Anderson Acceleration method appears to be more efficient than the restarted MPE and RRE methods, in particular for a maximal depth $m = 4$. It succeeds in achieving convergence without difficulty, with a stable mode of residual decrease well observed in Figure 17.

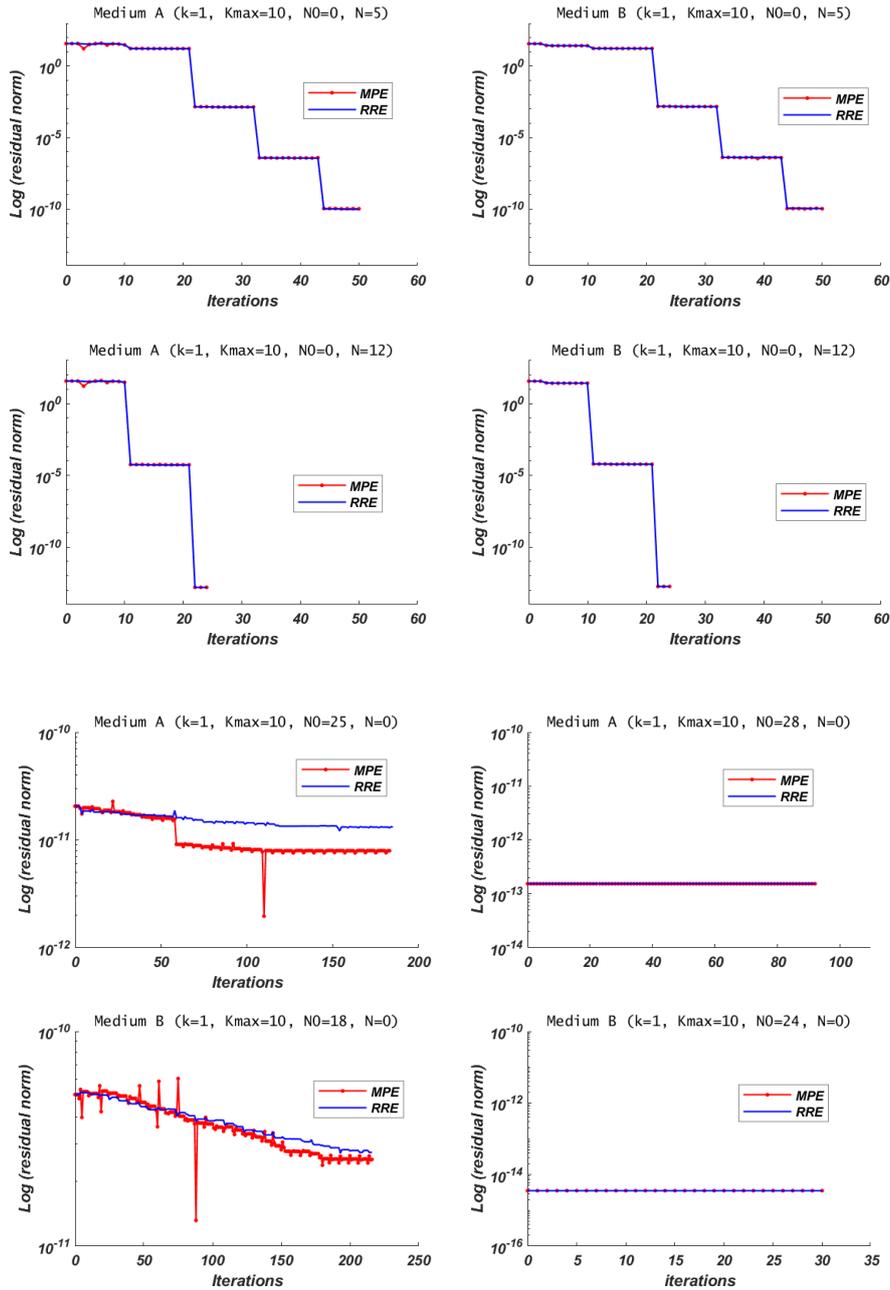


FIGURE 22. MoMas easy test: Thermodynamic equilibrium for injection in media A and B by restarted MPE and RRE, with $\kappa = 1$ - Residual norm curve

All the computation results for the MoMas easy test case are summarized in Table 7. The latter gives the total number of iterations performed $N_{\text{iterations}}$, as well as the number of cycles N_{cycles} and the computation time CPU necessary for performing the $N_{\text{iterations}}$ iterations and reach convergence. We notice that this time is very short in all cases, not exceeding 3s. This illustrates the efficiency and robustness of the MPE and RRE methods, especially in cyclic mode.

After the computation of components concentrations at equilibrium state, for each period and for each medium, thermodynamic equilibrium concentrations of mobile secondary species (C_1, C_2, C_3, C_4, C_5) can be computed as well as those of fixed secondary species (CS_1, CS_2) thanks to equations in (2.9).

Comparison with other results

For the acid Gallic test, the fast convergence observed for the Anderson Acceleration method and for the two polynomial extrapolation methods MPE and RRE is comparable to the results of J. Carrayrou

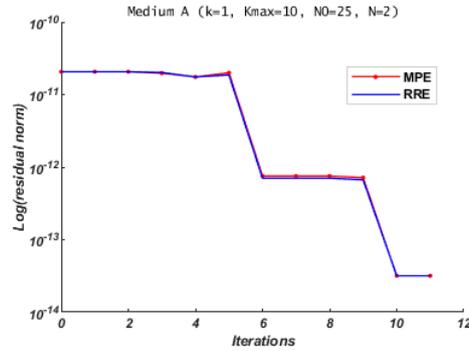


FIGURE 23. MoMas easy test: Thermodynamic equilibrium for injection in medium A by restarted MPE and RRE, $\kappa = 1, N_0 = 25, N = 2, Kmax = 10$ - Residual norm curve

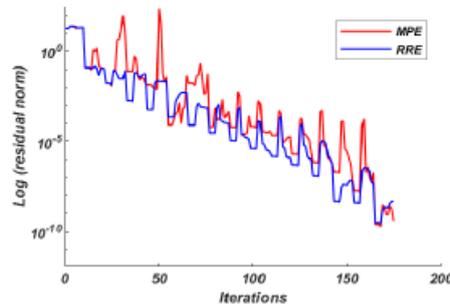


FIGURE 24. MoMas easy test: Thermodynamic equilibrium for leaching period by restarted MPE and RRE - Residual norm curve

	(Kmax, N, N ₀)	N _{iterations}		N _{cycles}		CPU time (s)	
		MPE	RRE	MPE	RRE	MPE	RRE
Zone A , $\kappa = 0.4$	(10, 0, 20)	330	330	30	30	1.0469	1.9531
	(10, 5, 20)	121	110	11	10	1.3906	0.7969
	(10, 20, 0)	55	77	5	7	0.9063	0.7188
	(10, 10, 0)	88	88	8	8	0.9844	1.3594
	(10, 5, 0)	110	110	10	10	1.0625	0.8750
Zone B , $\kappa = 0.3$	(10, 15, 0)	77	88	7	8	1.5469	0.7656
	(10, 0, 80)	291	295	30	30	1.7656	1.3438
Injection period , $\kappa = 0.2$	(15, 15, 15)	112	112	7	7	1.2188	0.7969
	(15, 0, 15)	256	256	17	17	1.5781	1.4688
Injection period , $\kappa = 1$	(10, 5, 0)	51	51	5	5	0.8750	0.8594
	(10, 12, 0)	25	25	3	3	0.5625	0.7813
	(10, 0, 25) in A	178	180	30	30	1.2031	1.1250
	(10, 0, 28) in A	90	90	30	30	1.2031	0.9844
	(10, 0, 18) in B	210	210	30	30	1.3594	1.3281
	(10, 0, 24) in B	30	30	30	30	0.8594	0.6875
Leaching period , $\kappa = 0.495$	(10, 10, 0)	176	176	16	16	1.2813	1.3438
	(10, 10, 10)	154	231	14	21	1.4688	2.2656
	(10, 18, 5)	154	110	14	10	2.1250	1.5313

TABLE 7. MoMas easy test, thermodynamic equilibrium by restarted MPE and RRE methods.

[10]. For the Newton Raphson type methods, the computation of the thermodynamic equilibrium of the Gallic acid test presents difficulties of convergence. By following the evolution of the process to search a solution in case 1 (cf. Figure 2(a) in [12]), we observe a phenomenon of oscillations during the process of finding solution with Newton Raphson method (which means no convergence). On the other hand, we can note that the Simplex and Newton Raphson with PCF methods allow to

obtain an approximation of the solution, without oscillation, but these require a long computation time. The Simplex method requires significant computation times because the search procedure is far from the solution for a long time. The Newton-Raphson method modified by polishing factor makes it possible to quickly obtain the solution but Figure 2(a) in [12] shows that oscillations are located in a neighborhood close to the solution. It is clear that with Newton-Raphson method modified by imposing the CAI, the oscillations responsible for convergence problems are intrinsic to the CAI procedure. Finally, the Newton-Raphson method with the relaxation by the secant method and the SPECY algorithm allow to effectively approach the solution, avoiding the oscillations and reducing the computation time.

If we compare all the Newton type methods mentioned with the Anderson Acceleration method, the search process is not captured with oscillations. It allows a first accurate approximation of the solution to be obtained more quickly, in a short calculation time (cf. Figure 2 and Table 4). In addition, this method converges in both cases and requires a small number of iterations, for all strictly positive values of the maximal depth m (cf. Figure 2). Likewise, we can quickly get a precise approximation of the solution by applying restarted MPE and RRE methods to the sequence $(\omega_n)_{n \geq 0}$, in both cases, without difficulty and in short computation time.

For the MoMas easy test case, a comparison of our results with those obtained in [27] is presented. The reactive transport code HYTEC participated in the realization of the benchmark, when all chemical reactions are solved by the speciation code CHESS [47]. CHESS uses an improved Newton-Raphson scheme to solve the set of nonlinear algebraic equations describing the chemical system. The HYTEC code was applied to the easy MoMas benchmark as such, without any modification to operate more quickly or to improve convergence, taking the precision of the resolution of chemical equations (Newton Raphson) equal to 10^{-8} .

Far from results concerning transport, we can see in [27] that authors give in a table, results of computation of chemical speciation in initial zones A and B independently, obtained with CHESS code. A comparison between results obtained with Anderson Acceleration, MPE and RRE with those ob-

	Medium A	Medium B	Injection	Leaching		zone A	zone B	Injection	Leaching
species					species				
X ₁	1e-20	1e-20	0.3	1e-20	X ₁	-	-	0.3	-
X ₂	0.2597	1.5116	0.2416	5.7734e-07	X ₂	0.25972	1.5116	0.24162	5.7735e-07
X ₃	1.4604e-24	3.6593e-28	0.2416	7.2169e-27	X ₃	-	-	0.24162	-
X ₄	0.3495	0.5756	2.0800e-51	1.1547e-06	X ₄	0.34954	0.57561	-	1.1547e-06
C ₁	3.8503e-12	6.6157e-13	4.1387e-12	1.7321e-06	C ₁	3.8503e-12	6.6157e-13	4.1387e-12	1.7321e-06
C ₂	3.7928e-25	5.5312e-28	0.0584	4.1667e-33	C ₂	-	-	0.05838	-
C ₃	1.3458	0.3808	8.6087e-51	2	C ₃	1.3458	0.38081	-	2
C ₄	1.3707e-24	1.3369e-30	6.3800e-152	1e-20	C ₄	-	-	-	-
C ₅	4.9532e-40	1.4724e-47	1e-20	4.8225e-75	C ₅	-	-	-	-
sites					sites				
S	0.3907	7.9128	2.9332e-24	1e-20	TS	0.39074	7.9128	-	-
CS ₂	0.3046	1.0436	1.2687e-97	6e-29	CS ₂	0.30463	1.0436	-	-
CS ₁	9.9968e-21	1e-20	9.9971e-21	1.3889e-59	CS ₁	-	-	-	-

TABLE 8. Comparison of the chemical speciation in initial zones obtained by Anderson Acceleration, MPE and RRE methods (on the left) with the chemical speciation obtained by CHESS code (on the right).

tained by the CHESS code (i.e Newton Raphson’s method) is summarized in Table 8. One note that the results are the same and in good agreement (all the concentrations lower than 10^{-20} have been represented by "-" in the right part of Table 8).

Four other reactive transport codes also participated in the realization of the benchmark (SPECY, MIN3P, GDAE and *Hoffmann et al*), but the results of chemical equilibrium are not presented independently during transport. All these codes are based on a Newton type method to linearize the chemical system and each uses a specific method to find the solution of the linearized system. A reference solution is given by the calculation of SPECY code [11] and a comparison of the results is carried out in [14], where the simulations are given by coupling transport and chemistry. To make a comparison with our results, results of chemistry (our results) on the right and results of the reactive transport on the left are presented. For example, the simulations in [11] indicate that all codes correctly reproduce the increase and decrease of the concentration front C₂. Chemical equilibrium

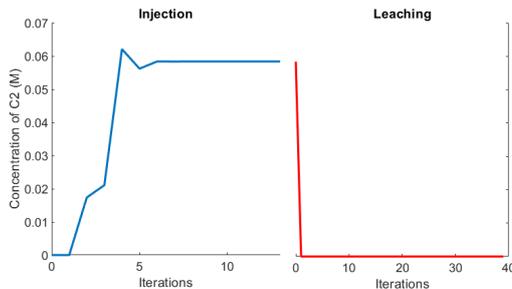


FIGURE 25. MoMas easy test, elution curve for species C2 during injection and leaching periods; chemical equilibrium by Anderson (2), MPE and RRE.

results presented in Figure 24 also exhibit the same behavior for the concentration of C₂ and are in good agreement with those of the reactive transport codes (cf. Fig. 7 in [11]).

On the other hand, chemical equilibrium results obtained for the fixed component S clearly show the influence of the more reactive medium B indicated by the higher concentration of S ($S_B^* = 7.9128$ M). This is in good agreement with the results obtained by the reactive transport codes, in advective case (cf. Fig. 5 in [14]), where at time 10, this high concentration is present in the center of the domain where $1 \leq x \leq 1.1$ (x designates the space), i.e. in B. In addition, we see that in medium A (for $0 \leq x \leq 1$ and $1.1 \leq x \leq 2.1$), the concentration of S converges to the same solution ($S_A^* = 0.39074$ M) that we obtained for chemical equilibrium in A by (AA) (Figure 7) and restarted (MPE) and (RRE) methods.

The most important advantage of the Anderson Acceleration method compared to Newton Raphson type methods is that its algorithm does not require the calculation of the Jacobian matrix. In the resolution of small linear systems using the algorithm of Newton-Raphson, the study of the condition number of Jacobian matrices shows that the range of values covered is unusually large, which leads to specific numerical problems. The matrices are quite small (10×10) but very ill conditioned (up to 10^{200}) (see Table 3 in [28]). This problem is completely overcome with (AA) method when we study the condition-number of the matrix \mathcal{F}_k (or R_k) instead of that of the Jacobian matrix. In addition, the condition-number monitoring strategy used in this method never allows to obtain ill-conditioned matrices. The real $\text{cond}(R_k)$ is always less than 10^{10} with this strategy. In particular, for the Acid Gallic test (respectively MoMas easy test case), with Newton-Raphson type methods, the condition number of the Jacobian matrix varies between $10^{0.61}$ and $10^{12.6}$ (respectively between $10^{3.44}$ and $10^{37.7}$), but with Anderson Acceleration method, the condition-number of matrix \mathcal{F}_k remains less than to 10^{10} , after condition-number monitoring (cf. Figures 3, 12, 13, 14, 15 and 16). Then, with (AA) method, efficiency and (relatively) good conditioning can be obtained through updated QR factorizations. Similarly, for nonlinear problems, the two polynomial-type vector extrapolation methods MPE and RRE do not need the use of the Jacobian of the function \mathbf{G} . Moreover, an important property of these methods is that they can be applied directly to the solution of linear and nonlinear systems. This is because the definitions of these methods do not require explicit knowledge of how the sequence is generated.

5. Conclusion

The aim of this work is to provide a stable and precise chemical solver to be integrated into an iterative sequential algorithm for reactive transport. The methods presented in this article allow to solve thermodynamic equilibria in a completely new way (without using the Newton-Raphson method). To our knowledge, these iterative acceleration algorithms have never been applied to the resolution of thermodynamic equilibria. The numerical results presented in this article improve the existing results (for example those given in [12]). Thus, the direct combination of the method of positive continued fractions with AA, RRE or MPE provides efficient algorithms with quadratic convergence from any initial arbitrary data.

It is planned to apply these methods to other cases constituting the MoMas reference test cases for which the chemical complexity is increasing (ie "medium test case" and "hard test case"). For this work

to bring sufficient novelty, the project consists in coupling our numerical thermodynamic equilibrium resolution to the transport model recently introduced in [4]. In this work, the authors establish a model which describes the water flow in shallow aquifers. The model couples the two dominant flows existing in the aquifer: a vertical 1d-Richards problem is considered in the capillary fringe while a vertical average of the mass conservation law is made in the saturated zone of the aquifer. This study is part of a larger project which aims to model the contamination of groundwater by nitrates.

Obviously, the potential parallelization of the proposed algorithms is an important step in upcoming works, in particular if we want these algorithms to be implemented in the framework of a parallel open-source platform. The parallelization of the MPE and RRE algorithms has already been discussed, in particular in the context of the article [20]. It seems quite possible to adapt these results to our case. AA algorithm mainly generating QR factorizations, its parallelization should not pose any special complications.

Finally and independently, it would be really very interesting to compare on the problem of thermodynamic equilibria, the results obtained by the AA, RRE or MPE approaches with those obtained thanks to the deep learning methods used in [55, 56].

Acknowledgments: We thank referees as well as Jérôme Carrayrou for their helpful comments and interesting suggestions which allowed to improve the actual version of the article. This work is supported by the project NEEDS-NewSolChem of CNRS.

References

- [1] Ahusborde E., Ossmani M. E., Id Moulay M., *A fully implicit finite volume scheme for single phase flow with reactive transport in porous media*, Mathematics and computers in simulation 164 (2019) 3-23.
- [2] Ackerer P., Preface: *Special issue on simulations of reactive transport: Results of the MoMaS benchmarks*, Computat. Geosci. 14(3) (2010), 383.
- [3] Anderson, D. G., *Iterative procedures for nonlinear integral equations*. Journal of the ACM, 12 (1965), pp. 547-560.
- [4] Bourel, C., Choquet, C., Rosier, C., Tsegmid, M., *Modelling of shallow aquifers in interaction with overland water*. Applied Mathematical Modelling, vol. 81, pp 727-751, (2020)
- [5] Bourgeat, A., Bryant, S., Carrayrou, J., Dimier, A., Van Duijn, C.J., Kern, M., Knabner, P., *Benchmark Reactive Transport*. Technical Report GDR MOMAS (2006).
- [6] Brassard, P., Bodurtha, P., *A Feasible Set for Chemical Speciation Problems*. Comput. Geosci., 26, 277 (2000).
- [7] Brezinski, C., Redivo Zaglia, M., Saad, Y., *Shanks sequence transformations and Anderson acceleration*. SIAM Rev., 60 (2018) 646-669.
- [8] Broyden, C. G., *A class of methods for solving nonlinear simultaneous equations*. Math. Comp. 19, 577-593 (1965).
- [9] Cabay, S., Jackson, L.W., *A polynomial extrapolation method for finding limits and antilimits of vector sequences*. SIAM J. Numer. Anal. 13 (1976) 734-752.
- [10] Carrayrou, J., *Modélisation du transport de solutés ractifs en milieu poreux saturé*. Thèse de doctorat, Université Louis Pasteur, Strasbourg, (2001).
- [11] Carrayrou, J., *Looking for some reference solutions for the reactive transport benchmark of MoMaS with SPECY*. Computational Geosciences, 14: 393-403, (2010).
- [12] Carrayrou, J., Mosé, R., Behra, P., *New efficient algorithm for solving thermodynamic chemistry*. AIChE J. 48(4), 894-904 (2002).
- [13] Carrayrou, J., Kern, M., Knabner, P., *Reactive transport benchmark of MoMaS*, Computat. Geosci. 14 (2010), pp. 385-392. 10.1007/s10596-009-9157-7.
- [14] Carrayrou, J., Hoffman, J., Knabner, P., Krautle, S., De Dieuleveult, C., Erhel, J., Van Der Lee, J., Lagneau, V., Mayer, K. U., Macquarrie, K. T. B., *Comparison of numerical methods for simulating strongly nonlinear and heterogeneous reactive transport problems-the MoMaS benchmark case*. Computat. Geosci., 14 (2010), pp. 483-502.
- [15] Marinoni M., Carrayrou, J., Lucas Y., Ackerer P., *Thermodynamic equilibrium solutions through a modified Newton Raphson method*, AIChE Journal (2016).
- [16] Machat H., Carrayrou, J., *Comparisons of linear solvers for equilibrium geochemistry computations*, Comput Geosci (2017) 21: 131-150 DOI 10.1007/s10596-016-9600-5
- [17] Eddy, R.P., *Extrapolating to the limit of a vector sequence*. in: P.C.C. Wang, Ed., Information Linkage between Applied Mathematics and Industry (Academic Press, New York, 1979) 387-396.
- [18] Eyert, V., *A comparative study on methods for convergence acceleration of iterative vector sequences*. J. Comput. Phys. 124(2), 271-285 (1996).
- [19] DuMuX, DUNE for Multi-Phase, Component, Scale, Physics, ..., flow and transport in porous media.
- [20] Duminil, S., Sadok, H.; Silvester, D. *Fast solvers of discretized Navier-Stokes problems using vector extrapolation*, Numer. Algorithmes 66 (2014), no. 1, 89-104.
- [21] DUNE, the Distributed and Unified Numerics Environment, <http://www.dune.project.org>.

- [22] Fang, H., Saad, Y., *Two classes of multisection methods for nonlinear acceleration*. Numer. Linear Algebra Appl. 16 (3) (2009) 197-221.
- [23] Hoffmann J., Kräutle S., Knaber P., *A parallel global-implicit @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 (2010) 421-433.
- [24] Jbilou, K., *A general projection algorithm for solving linear systems of equations*. Numer. Algorithms, 4 (1993), pp. 361-377
- [25] Jbilou, K., Sadok, H., *Vector extrapolation methods. Application and numerical comparison*. J. Comp. Appl. Math, 122 (2000), 149-165.
- [26] Krebs, R., Sardin, M., Schweich, D., *Mineral Dissolution, Precipitation and Ion Exchange in Surfactant Flooding*. AIChE J., 33, 1371 (1987).
- [27] Lagneau, V., Van Der Lee, J., *HYTEC results of the MoMas reactive transport benchmark*
- [28] Machat, H., Carayrou, J., *Comparison of linear solvers for equilibrium geochemistry computations*. Comput. Geosci. 21(1) (2017) 131-150.
- [29] Mayer K.U., MacQuarrie K.T.B., *Solution of the MoMas reactive transport benchmark with MIN3P-model formulation and simulation results*, Comput. Geosci. 14 (2010) 405-419.
- [30] MeSina, M., *Convergence acceleration for the iterative solution of the equations $X = AX + f$* . Comput. Methods Appl. Mech. Engrg. 10 (2) (1977) 165-173.
- [31] Morin, K. A., *Simplified Explanations and Examples of Computerized Methods for Calculating Chemical Equilibrium in Water*. Comput. Geosci., 11, 409 (1985).
- [32] Nelder, J. A., Mead, R., *A Simplex method for function Minimization*. Comput. J., 7, 308 (1965).
- [33] 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. Rep. 99-4259, U.S. Geological Survey, Denver, CO (1999).
- [34] Potra, F. A., Engler, H., *A characterization of the behavior of the Anderson acceleration on linear problems*. Linear Algebra Appl. 438(3), 1002-1011 (2013).
- [35] Pulay, P., *Convergence acceleration of iterative sequences. The case of SCF iteration*. Chem. Phys. Lett. 73(2), 393-398 (1980).
- [36] Rohwedder, T., Schneider, R., *An analysis for the DIIS acceleration method used in quantum chemistry calculations*. J. Math. Chem. 49(9), 1889-1914 (2011).
- [37] Saad, Y., Schultz, M. H., *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*. SIAM J. Sci. Statist. Comput. 7(3), 856-869 (1986).
- [38] Saaltink, MW., Carrera, J., Ayora, C., *Comparison of two approaches for reactive transport modeling*. J. Geochem Explor. (2000); 69:97-101.
- [39] Sadok, H.: *About Henrici's transformation for accelerating vector sequences* J. Comput. Appl. Math. 29 (1990) 101-110.
- [40] Shapiro, N. Z., Shapley, L. S., *Mass action laws and the gibbs free energy function*. J. SOC. Indust. Appl. Math., 13(2): 353-375, (1965).
- [41] Sidi, A., *Convergence and Stability Properties of Minimal Polynomial and Reduced Rank Extrapolation Algorithms*. SIAM J. Numer. Anal. 23, no.1, pp. 197-209, (1986).
- [42] Sidi, A., *Extrapolation vs. projection methods for linear systems of equations*. J. Comput. Appl. Math. 22, pp. 71-88, (1988).
- [43] Sidi, A., *Efficient implementation of minimal polynomial and reduced rank extrapolation methods*. J. Comput. Appl. Math. vol. 36 (1991), p. 305-337 (cf. p. 32-34).
- [44] Sidi, A., *Vector Extrapolation methods with applications to solution of large systems of equations and to PageRank computations*. Computer science department, Technion - Israel institute of technology, Haifa 32000, Israel.
- [45] Steefel, C. I., Appelo, C. A. J., Arora, B., Jacques, D., Kalbacher, T., Kolditz, O., Lagneau, V., Lichtner, P. C., Mayer, K., Meeussen, J. C. L., Molins, S., Moulton, D., Shao, H., Simunek, J., Spycher, N., Yabusaki, S. B., Yeh, G. T., *Reactive transport codes for subsurface environmental simulation*. Comput Geosci. (2015); 19(3): 445-478.
- [46] Toth, A., Kelley, C. T., *Convergence analysis for Anderson acceleration*. SIAM J. Numer. Anal., 53 (2015), pp. 805-819, <https://doi.org/10.1137/130919398>.
- [47] Van Der Lee, J., *Thermodynamic and mathematical concepts of CHESS*. Technical Report LHM/RD/98/39, CIG-Ecole des l'vlines de Paris, Fontainebleau, France, (1998).
- [48] Walker, H.F., *Anderson Acceleration: Algorithms and Implementations*. Research Report, MS-6-15-50, Worcester Polytechnic Institute Mathematical Sciences Department, (2011).
- [49] Walker, H.F., Ni, P., *Anderson acceleration for fixed-point iterations*. SIAM J. Numer. Anal. 49 (4) (2011) 1715-1735.
- [50] Wood, J. R., *Calculation of Fluid-Mineral Equilibria Using the Simplex Algorithm*. Comput. Geosci., 19, 23 (1993).
- [51] Wigley, T. M. L., *WATSPEC: A Computer Program for Determining the Equilibrium Speciation of Aqueous Solutions*. Brit. Geo-morphol. Res. Group Tech. Bull. 20 (1977).
- [52] Yeh, G. T., Tripathi, V. S., *A critical evaluation of recent developments in hydrogeochemical transport models of reactive multichemical components*. Water Resources Res. 25, 93-108, (1989).
- [53] Yeh, G. T., Tripathi, V. S., Gwo J.P., Cheng H.P., Cheng J.R.C., Salvage K.M., Li M.H., Fang Y., Li Y., Sun J.T., Zhang F. and Siegel M.D., *Hydrogeochem: A coupled model of variably saturated flow, thermal transport, and reactive biogeochemical transport*. Groundwater Reactive transport models, (2012) 3-41.
- [54] Zhang, T., Li, Y., Sun, S., *Phase equilibrium calculations in shale gas reservoirs*. Capillarity, 2019, 2(1): 8-16, doi: 10.26804/capi.2019.01.02.

- [55] Zhang, T., Li, Yu, Li Yiteng, Sun, S., Hua B., *A self-adaptive deep learning algorithm for accelerating multi-component flash calculation*, Computer Methods in Applied Mechanics and Engineering (IF 5.763) Volume 369 (2020), DOI: 10.1016/j.cma.2020.113207.
- [56] Zhang, T., Li Yiteng, Sun, S., Gao X., *Accelerating flash calculations in unconventional reservoirs considering capillary pressure using an optimized deep learning algorithm*, Journal of Petroleum Science and Engineering, Vol 195 (2020).

^a UNIV. DU LITTORAL CÔTE D'OPALE, UR 2597, LMPA, LABORATOIRE DE MATHÉMATIQUES PURES ET APPLIQUÉES JOSEPH LIOUVILLE, F-62100 CALAIS, FRANCE. ^b CNRS FR 2037, FRANCE. ^c UNIVERSITÉ LIBANAISE, LAMA-LIBAN, LABORATOIRE DE RECHERCHE EN MATHÉMATIQUES ET APPLICATIONS, P.O. BOX 37 TRIPOLI, LIBAN.
E-mail address: Safaa.Al-Nazer@etu.univ-littoral.fr

^c UNIVERSITÉ LIBANAISE, LAMA-LIBAN, LABORATOIRE DE RECHERCHE EN MATHÉMATIQUES ET APPLICATIONS, P.O. BOX 37 TRIPOLI, LIBAN.
E-mail address: mjazar@laser-lb.org

^a UNIV. DU LITTORAL CÔTE D'OPALE, UR 2597, LMPA, LABORATOIRE DE MATHÉMATIQUES PURES ET APPLIQUÉES JOSEPH LIOUVILLE, F-62100 CALAIS, FRANCE. ^b CNRS FR 2037, FRANCE
E-mail address: rosier@univ-littoral.fr