Computational Geosciences (2021) 25:1063–1082
https://doi.org/10.1007/s10596-021-10039-0

ORIGINAL PAPER

On an efficient time scheme based on a new mathematical representation of thermodynamic equilibrium for multiphase compositional Darcy flows in porous media

J. Coatléven[1] · C. Meiller[1]

Received: 8 June 2020 / Accepted: 21 January 2021 / Published online: 17 March 2021
© The Author(s), under exclusive licence to Springer Nature Switzerland AG part of Springer Nature 2021

Abstract
This paper presents a new mathematical representation of multiphase thermodynamic equilibrium using so-called repartition coefficients. Combined with a global mass formulation of multiphase Darcy flow in porous media, it allows the derivation of a computationally efficient family of time schemes. The model accounts for the mass conservation of an arbitrary number of components flowing through an arbitrary number of phases, coupled with thermodynamic equilibrium and pore volume conservation. By separating the thermodynamic equilibrium part from the flow part through the repartition coefficients, the formulation removes the need for any specific handling of phase appearance and disappearance within the flow solver. Any "black box" thermodynamic equilibrium solver can then be used to compute the repartition coefficients, from EOS based solvers to tabulated representation of the thermodynamic equilibrium, each specific choice of thermodynamic solver leading to a new scheme. Three numerical experiments, from a simple beam to a real case, illustrate the good behavior of the approach.

Keywords Thermodynamic equilibrium model · Multiphase flow modeling · Darcy flow · Porous media

Mathematics Subject Classification (2010) 65M08 · 76T30

1 Introduction

Porous media flows play a major role in many crucial industrial application, from nuclear waste storage to oil and gas production. The interest in such problems is even renewed by environmental concerns such as long term energy or carbon dioxide storage [7]. To ensure operational and economic security, more and more complex chemical interactions must be considered requiring to simulate complex compositional multiphase flows. In particular, modeling thermodynamic equilibrium of such fluids is becoming mandatory in contexts such as CO_2 sequestration where gas dissolution in water is essential to understand and predict the behavior of CO_2 injection. Another major example is the study of biogenic CH_4 which represents around 20% of global resources of conventional gas and that is a great interest for oil and gas companies due to the fact that this gas is generated at shallow depths. The classical approach involves a strong coupling between costly flash calculations [11] and the flow solver which has to deal with phase appearance and disappearance. Those phase changes have in general a strong negative impact on computational time and even worse on the flow solver robustness [19].

Probably the most commonly encountered formulation of the coupling between Darcy flow and thermodynamic equilibrium models is the celebrated Coats' formulation [3], that uses the so-called "natural-variables" as unknowns (pressure, saturations, phase mole-fractions). This formulation maintains two sets of active and inactive unknowns and equations, with only the active ones appearing in the non-linear system. When one phase disappears during a time step or non-linear iteration, the corresponding saturation and phase mole-fractions variables are removed. Conversely, when one phase appears, the corresponding variables are added to the active set. This process is often referred to as

✉ J. Coatléven
julien.coatleven@ifpen.fr

C. Meiller
clementine.meiller@ifpen.fr

[1] IFP Énergies Nouvelles, 1 et 4 avenue de Bois-Préau, 92852 Rueil-Malmaison, France

"active set" or "variable substitution" method, and is the key ingredient of Coats' formulation. However, it generates discontinuities in the non-linear objective function, leading to well documented difficulties (see [1, 12, 20] when attempting to use Newton-Raphson algorithm to solve it. As it is a subject of strong economic importance, many attempts can be found in the literature to improve the management of phase appearance and disappearance through modified thermodynamic formulations. By no means will we try to be exhaustive here, we thus refer the reader to the recent paper of Voskov and Tchelepi [19], where a comprehensive comparison of models of the literature was performed. Let us nevertheless mention in particular the very interesting fugacity formulation of Lauser et al. [10], that was not considered in Voskov and Tchelepi [19]. To solve their fugacity based formulation they handle the phase appearance and disappearance through complementarity conditions rather than the usual variable switch, allowing to use a fixed set of variables. Regardless of the chosen formulation, let us also mention the very interesting operator based linearization (OBL) technique of [18], used to circumvent the difficulty of solving the complex non-linear flow problem. Its principle is to replace the non-linear operators and functions appearing in the numerical scheme by a multilinear interpolation in unknown space, thus reducing the effective non-linearity of the problem to be solved through Newton-Raphson's method. It was first applied to flow with thermodynamic equilibrium in [9, 18] and later extended to both thermodynamic and chemical equilibrium in [8], using what we call here a global mole formulation. Notice that the OBL approach could be applied to the scheme we present here, potentially further improving computational efficiency.

We introduce a new mathematical representation of thermodynamic equilibrium using so-called repartition coefficients. Combined with a global mass formulation of multiphase Darcy flow in porous media, it allows the derivation of a computationally efficient time scheme. The model accounts for the mass conservation of an arbitrary number of components flowing through an arbitrary number of phases, coupled with thermodynamic equilibrium and pore volume conservation. The approach naturally covers phase appearance and disappearance without impacting the structure of the flow solver, by keeping the same set of unknowns and equations whatever the thermodynamic phase state is.

The paper is organized as follows. In a first section, we recall the basics of thermodynamic equilibrium modeling, and explain through the two most encountered models in reservoir simulation how we can reformulate them as an abstract function. Section 3 introduces the repartition coefficients that allows a new mathematical description of this thermodynamic equilibrium function. Section 4 presents a global mass formulation of compositional multiphase Darcy flows that naturally accounts for phase appearance and disappearance. Section 5 presents a family of schemes that uses time explicit repartition coefficients to enhance robustness and computational efficiency and that is in our opinion the main contribution of the present work. Section 6 is devoted to numerical exploration of the proposed approach, showing 3 illustration cases with increasing complexity.

2 General representation of thermodynamic equilibrium and repartition coefficients

In this section, we first explain how most of the usual thermodynamic equilibrium models can be gathered under a unified abstract vision of thermodynamics. The idea in itself is relatively classical, as it basically consists in formalizing the fact that thermodynamics can be considered as a "black box" solver with fixed inputs and outputs. The identification of a very general family of inputs and outputs is the true contribution of this section. As obvious as it may seem, by hiding the details of each particular model this vision of thermodynamics forces to reinvent the way to couple thermodynamics and flow, in the sense that we cannot rely on the particular properties of any model and we have to use only what is shared by all the usual models.

2.1 Notations

A chemical component is defined by its chemical composition, while a chemical species is defined by its chemical composition and the phase under which it exists. For instance, the component CO_2 can be present in both the aqueous and the gaseous phases, while aqueous CO_2 and gaseous CO_2 are two different species. For a mixture with N_{ph} phases labeled from 0 to $N_{ph}-1$, and N_{comp} components labeled from 0 to $N_{comp}-1$, corresponding to N_{spec} species labeled from 0 to $N_{spec}-1$ and defined as couples (i,α), we denote:

- T the temperature of the system
- P_α the pressure of phase α, $\boldsymbol{P}_{th} = (P_\alpha)_{0\leq\alpha\leq N_{ph}-1}$ being the vector of phase pressures
- P_{ref} is a reference pressure, equal to the pressure of a reference phase (water in our numerical experiments)
- S_α the saturation (or porous volume fraction) of phase α
- ρ_α the mass density of phase α
- μ_α the viscosity of phase α
- θ_α the total molar fraction of α
- n_α the mole number of phase α per unit volume
- n_i^α the mole number of species (i,α) per unit volume
- x_i^α the molar fraction of species (i,α) in phase α

- μ_i^α the chemical potential of species (i,α) in phase α
- n_i the total mole number of component i per unit volume
- z_i the total molar fraction of component i
- Θ_α the total mass fraction of phase α
- m_α the mass of phase α per unit volume
- m_i^α the mass of species (i,α) per unit volume
- X_i^α the mass fraction of species (i,α) in phase α
- η_i^α the mass repartition coefficient of component i in phase α
- f_i^α the fugacity of of species (i,α) in phase α
- k_i^α the equilibrium coefficient of component i in phase α, defined as the inverse of the fugacity coefficient ϕ_i^α
- δ_i^α the presence index of component i in phase α
- Z_i the total mass fraction of component i
- m_i the mass of component i per unit volume
- M_{m_i} the molar mass of component i
- M_{m_α} the molar mass of phase α
- $\boldsymbol{V}_\alpha$ the Darcy velocity of phase α
- P_{c_i} the critical pressure of component i
- T_{c_i} the critical temperature of component i
- ω_i the acentric factor of component i
- Φ the porosity
- Λ the permeability tensor
- R the perfect gas constant

We use mole numbers and masses per unit volume as we always describe the thermodynamic equilibrium inside a fixed volume. This allows to get rid of the volume dependency and thus avoids using a factor V everywhere in order to ease the reading. As a consequence we will often speak of mass and mole number instead of mass per unit volume and mole number per unit volume with a slight abuse of language. Notice that this has no consequences on the results themselves provided all the involved quantities remain coherent (i.e. all absolute or all per unit volume).

We define the presence index δ_i^α of a component i in phase α by setting $\delta_i^\alpha = 1$ if species (i,α) exists and 0 otherwise. The molar mass M_{m_i} of each component i is assumed to be a constant depending only on the chemical composition, thus all species (i,α) share the same molar mass M_{m_i}. Of course, we assume that each component i exists at least in one phase α under the form of the species (i,α), in order that:

$$\sum_{\alpha=0}^{N_{ph}-1} \delta_i^\alpha \geq 1 \quad \text{for all } 0 \leq i \leq N_{comp}-1,$$

and thus each phase is represented by at least one species:

$$\sum_{i=0}^{N_{comp}-1} \delta_i^\alpha \geq 1 \quad \text{for all } 0 \leq \alpha \leq N_{ph}-1.$$

With a slight abuse of notations this allows to speak for instance of a species (i,α) that does not truly exist, by setting $x_i^\alpha = X_i^\alpha = 0, n_i^\alpha = m_i^\alpha = 0$ if $\delta_i^\alpha = 0$.

2.2 Some remarks on thermodynamic equilibrium

Denote N_{ph}^a the number of non-existing (or absent) phases, i.e. phases α for which $\theta_\alpha = 0$, and thus $N_{ph} - N_{ph}^a$ the number of actually existing phases, i.e. phases α for which $\theta_\alpha > 0$. Denote $\mathcal{A}_{ph}$ the set of the N_{ph}^a absent phases and $\mathcal{P}_{ph}$ the set of the $N_{ph} - N_{ph}^a$ existing phases. Roughly speaking, finding the thermodynamic equilibrium always involve solving for each couple of phases $(\alpha,\beta) \in \mathcal{P}_{ph}^2$ (i.e. both phases are truly present in the system):

$$\mu_i^\alpha = \mu_i^\beta \text{ for all } 0 \leq i \leq N_{comp}-1 \text{ with } \delta_i^\alpha = \delta_i^\beta = 1, \quad (1)$$

where μ_i^α is the chemical potential of species (i,α). However the chemical potential, despite the fact that it is the suitable quantity for describing the equilibrium state of a system, is not the ideal quantity in practice as it depends on some integration constant (that is only partially constant as it may depend on pressure, temperature and even composition, depending on the reference state). With only equilibrium reactions, fortunately the chemical potential can be replaced by fugacities or related quantities, that are fully determined and much easier to manipulate in practice.

To recall the definition of fugacity and the aforementioned related quantities, let us denote $\mu_i^{\bullet,*}(P_{ref},T)$ the chemical potential of the component i alone in perfect gas state, at pressure P_{ref} and temperature T. By definition, fugacity f_i^α of species (i,α) is given by:

$$\mu_i^\alpha = \mu_i^{\bullet,*}(P_{ref},T) + RT \ln\left(\frac{f_i^\alpha}{P_{ref}}\right),$$

where R is the perfect gas constant. From Gibbs theorem, a perfect gas mixture being an ideal solution, the chemical potential $\mu_i^\bullet(P_{ref},T,\boldsymbol{x})$ of component i at pressure P_{ref} and temperature T in a perfect gas mixture of molar composition $\boldsymbol{x} = (\boldsymbol{x}^\alpha)_{0\leq\alpha\leq N_{ph}-1}$ is given by, if $\boldsymbol{x}^\alpha = (x_i^\alpha)_{0\leq i\leq N_{comp}-1}$ is the composition of phase α:

$$\mu_i^\bullet(P_{ref},T,\boldsymbol{x}^\alpha) = \mu_i^{\bullet,*}(P_{ref},T) + RT\ln\left(x_i^\alpha\right),$$

and thus:

$$\mu_i^\alpha = \mu_i^\bullet(P_{ref},T,\boldsymbol{x}^\alpha) + RT\ln\left(\frac{f_i^\alpha}{x_i^\alpha P_{ref}}\right).$$

The fugacity coefficient ϕ_i^α is then naturally defined as:

$$\phi_i^\alpha = \frac{f_i^\alpha}{x_i^\alpha P_{ref}}.$$

For practical reasons, we introduce the equilibrium coefficient k_i^α defined as the inverse of the fugacity coefficient ϕ_i^α, i.e.:

$$k_i^\alpha = \frac{1}{\phi_i^\alpha} = \frac{x_i^\alpha P_{ref}}{f_i^\alpha}.$$

Notice that Eq. 1 implies the equality of fugacities f_i^α and f_i^β, which gives in terms of fugacity coefficients:

$$\phi_i^\alpha x_i^\alpha = \phi_i^\beta x_i^\beta,$$

and then rewrites:

$$\frac{x_i^\alpha}{k_i^\alpha} = \frac{x_i^\beta}{k_i^\beta}.$$

The more usual equilibrium constant (sometimes also called partition coefficient) $k_i^{\alpha\beta}$ is then defined as:

$$k_i^{\alpha\beta} = \frac{k_i^\alpha}{k_i^\beta} = \frac{\phi_i^\beta}{\phi_i^\alpha} = \frac{1}{k_i^{\beta\alpha}}.$$

This equilibrium constant is the most frequently encountered quantity in the literature to describe the equilibrium between only two phases. However, as we consider more general situations here and in particular the three-phase flow case where $N_{ph} = 3$, it seems much easier to use either the k_i^α's or the ϕ_i^α's, to keep some symmetry in the thermodynamic description.

Relation (1) is in fact one of the optimality conditions corresponding to the minimization of the Gibbs free energy

$$G = \sum_{i=0}^{N_{comp}-1} \sum_{\alpha=0}^{N_{ph}-1} \delta_i^\alpha n_i^\alpha \mu_i^\alpha.$$

This minimization problem can be considered as the most general description of thermodynamic equilibrium. The other optimality conditions correspond to the constraints under which the minimization is performed, the most important being that the quantity of each component remains invariant and that relation (1) is only valid for phases whose total fraction is strictly positive (which also means phases with a physically admissible composition). The treatment of those constraints is the origin of the major difficulties arising in solving the thermodynamic equilibrium problem and many of its combinatorial aspects. Model specialization basically consists in choosing a particular form for the chemical potentials μ_i^α, or equivalently for the fugacity, fugacity coefficient or equilibrium coefficients. Because of the inherent difficulty of solving the Gibbs energy minimization problem, specialized solvers associated with each particular model have been developed, taking advantage of model specificity to accelerate calculations.

The inputs of those minimization problems are those corresponding to the definition of the constraints and the thermodynamic state parameters used in the considered model. In general, most chemical potential models require P_{ref} and T, while the constraints (in particular the conservation of mass or matter) require at least the total fractions of each component, either in moles (denoted $\boldsymbol{z}_{comp} = (z_i)_{0\le i\le N_{comp}-1}$) or in mass (denoted $\boldsymbol{Z}_{comp} = (Z_i)_{0\le i\le N_{comp}-1}$) defined by:

$$z_i = \frac{n_i}{\sum_{j=0}^{N_{comp}-1} n_j} \quad \text{and} \quad Z_i = \frac{m_i}{\sum_{j=0}^{N_{comp}-1} m_j},$$

or directly the mole numbers $(n_i)_{0\le i\le N_{comp}-1}$ or the masses $(m_i)_{0\le i\le N_{comp}-1}$. The expected outputs are the phase molar fractions at equilibrium $\boldsymbol{\theta}_{ph}^{eq}$ and the molar composition of each phase at equilibrium $\boldsymbol{x}^{eq}$. For the reader's better understanding, the two families of models that are mostly encountered in the literature for hydrocarbon phases equilibrium are recalled.

Equation of state models Equation of state (EOS) based models are the most classical and probably the most used models in the industry to describe hydrocarbon phases. The most common equations of state for hydrocarbon phases are the cubic EOS of Redlich-Kwong [14], Soave-Redlich-Kwong [15] and Peng-Robinson [13], but more advanced models can be encountered, especially when water vaporization or specific component dissolution must be accurately modeled (e.g. the Søreide and Whitson EOS [16]. For each phase α, we assume that we are given an equation of state relating the thermodynamic state parameters (P_{ref}, T, V_m^α), where V_m^α denotes the molar volume of phase α, and the phase composition $\boldsymbol{x}^\alpha = (x_i^\alpha)_{0\le i\le N_{comp}-1}$. Generally speaking, it involves a relation of the form:

$$E_\alpha(P_{ref}, T, \mathcal{Z}_\alpha, \boldsymbol{x}^\alpha) = 0,$$

where the compressibility factor $\mathcal{Z}_\alpha$ of phase α is defined by:

$$\mathcal{Z}_\alpha = \frac{P_{ref} V_m^\alpha}{RT}.$$

From the solution of this equation, one deduces many properties of phase α and in particular its fugacity and fugacity coefficient, given by:

$$\ln\left(\frac{f_i^{\alpha,EOS}}{x_i^\alpha P_{ref}}\right) = \ln \phi_i^{\alpha,EOS} = \int_0^P \frac{\mathcal{Z}_\alpha - 1}{P_{ref}} dP_{ref} + \int_0^{P_{ref}} \frac{1}{P_{ref}} \left(\frac{\partial \mathcal{Z}_\alpha}{\partial x_i^\alpha} - \sum_{j=0}^{N_{comp}-1} \delta_j^\alpha x_j^\alpha \frac{\partial \mathcal{Z}_\alpha}{\partial x_j^\alpha}\right) dP_{ref}. \quad (2)$$

Solving the equation of state formally means that we are able to compute pointwise the compressibility factor $\mathcal{Z}_\alpha$ as a function of $(P_{ref}, T, \boldsymbol{x}^\alpha)$, and thus at least implicitly we define through this equation and Eq. 2 a law for the equilibrium coefficients $k_i^{\alpha,EOS}(P_{ref}, T, \boldsymbol{x}^\alpha)$ (or equivalently the fugacity coefficients).

Thus, in the case of EOS models, the thermodynamic equilibrium problem can be reformulated as finding the sets $\mathcal{A}_{ph}$ and $\mathcal{P}_{ph}$ and the associated composition vector $\boldsymbol{x}$ and phase fraction vector $\boldsymbol{\theta}_{ph} = (\theta_\alpha)_{0\leq\alpha\leq N_{ph}-1}$ such that for all $(\alpha, \beta) \in \mathcal{P}_{ph}^2$ and all $0 \leq i \leq N_{comp} - 1$ such that $\delta_i^\alpha = \delta_i^\beta = 1$:

$$\frac{x_i^\alpha}{k_i^{\alpha,EOS}(P_{ref}, T, \boldsymbol{x}^\alpha)} = \frac{x_i^\beta}{k_i^{\beta,EOS}(P_{ref}, T, \boldsymbol{x}^\beta)},$$

that minimizes the Gibbs free energy and satisfies the conservation of matter:

$$z_i = \sum_{\alpha\in\mathcal{P}_{ph}} \theta_\alpha \delta_i^\alpha x_i^\alpha. \tag{3}$$

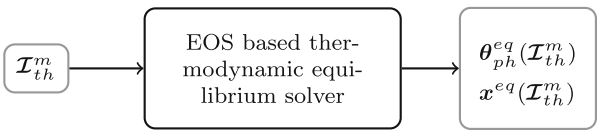
Then, if one denotes $\mathcal{A}_{ph}^{eq}$, $\mathcal{P}_{ph}^{eq}$, $\boldsymbol{\theta}_{ph}^{eq}$ and $\boldsymbol{x}^{eq}$ the optimal solution, it is clear that it only depends on the vector $\boldsymbol{z}_{comp}$ and the thermodynamic state parameters (P_{ref}, T) (and of the model specific constant parameters involved in the definition of the EOS, such as acentric factors, mixing rules, etc...). Thus, denoting

$$\boldsymbol{\mathcal{I}}_{th}^m = (P_{ref}, \boldsymbol{P}_{ph}, T, z_{comp}),$$

it seems clear that solving this optimization problem implicitly defines functions

$$\boldsymbol{\theta}_{ph}^{eq}(\boldsymbol{\mathcal{I}}_{th}^m) \quad \text{and} \quad \boldsymbol{x}^{eq}(\boldsymbol{\mathcal{I}}_{th}^m),$$

which are thus the outputs of our EOS model.



As a by-product, this also defines the laws of equilibrium coefficients taken at their equilibrium value, i.e.:

$$k_i^{\alpha,eq}(\boldsymbol{\mathcal{I}}_{th}^m) = k_i^{\alpha,EOS}(P_{ref}, T, \boldsymbol{x}^{\alpha,eq}(\boldsymbol{\mathcal{I}}_{th}^m)), \tag{4}$$

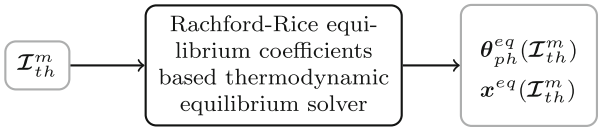
and thus we can consider if needed that the functions $k_i^{\alpha,eq}(\boldsymbol{\mathcal{I}}_{th}^m)$ are also the outputs of our EOS based model. To conclude this paragraph on EOS-based models, let us mention that it is common practice to use a single pressure P_{ref} for all phases in the thermodynamic equilibrium calculations as we have done here (it is generally taken equal to the pressure of one of the phases, common choices are the aqueous phase or the oil phase). The full pressure vector of the flow calculations and which takes into accounts capillary pressures is used mostly for equilibrium calculations when one wants to fully couple thermodynamics and geomechanics. This is most probably due to the fact that many capillary pressure models commonly used are phenomenological models optimized for modeling the flow but not so well-suited for thermodynamic equilibrium calculations.

Equilibrium constants based models Those models are much simpler than equation of state models from which they are in fact often derived. The simplification relies in the fact that they directly assume that we know the values taken by the equilibrium coefficients when the equilibrium is indeed reached, under the form of functions $k_i^{\alpha,eq}$. Notice that those functions are different from the functions $k_i^{\alpha,EOS}$ defining the equilibrium coefficients themselves in EOS models, as they correspond to the value taken by these equilibrium coefficients when the equilibrium is indeed reached (4). In particular, they do not depend on intermediate variables such as phase compositions. The most common model simply considers them as functions of (P_{ref}, T). In particular, this is the case of classical liquid-vapor (LV) equilibrium models for hydrocarbon phases, which are described in an even more compact way by directly using the equilibrium constants $k_i^{\alpha\beta,eq}$. In the literature, the most commonly encountered correlations are analytic or tabulated laws providing $k_i^{LV,eq}$ as a function of either (P_{ref}, T) or $(\ln P_{ref}, T^{-1})$ [6, 17]. Those models are fully justified in situations where the mixture total composition varies slowly over time and space, the dependency of the equilibrium coefficients (or fugacity coefficients) being in general relatively small, at least for the most usual hydrocarbon phases. For reservoir simulation and in particular production simulation where the composition of the fluids in place is well-known, such models are particularly well-suited. Similarly to the case of EOS models, the thermodynamic equilibrium problem can be reformulated as finding the sets $\mathcal{A}_{ph}$ and $\mathcal{P}_{ph}$ and the associated composition vector $\boldsymbol{x}$ and phase fraction vector $\boldsymbol{\theta}_{ph} = (\theta_\alpha)_{0\leq\alpha\leq N_{ph}-1}$ such that for all $(\alpha, \beta) \in \mathcal{P}_{ph}^2$ and all $0 \leq i \leq N_{comp} - 1$ such that $\delta_i^\alpha = \delta_i^\beta = 1$:

$$\frac{x_i^\alpha}{k_i^{\alpha,eq}(\boldsymbol{\mathcal{I}}_{th}^m)} = \frac{x_i^\beta}{k_i^{\beta,eq}(\boldsymbol{\mathcal{I}}_{th}^m)}, \tag{5}$$

that minimizes the Gibbs free energy and satisfies the conservation of matter (3), which again implicitly defines

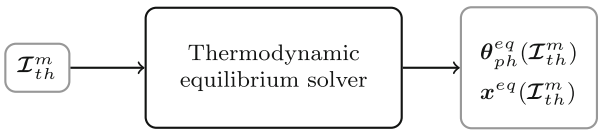
functions $\boldsymbol{\theta}_{ph}^{eq}(\boldsymbol{\mathcal{I}}_{th}^{m})$ and $\boldsymbol{x}^{eq}(\boldsymbol{\mathcal{I}}_{th}^{m})$, if one denotes $\mathcal{A}_{ph}^{eq}$, $\mathcal{P}_{ph}^{eq}$, $\boldsymbol{\theta}_{ph}^{eq}$ and $\boldsymbol{x}^{eq}$ the optimal solution.



This is nothing but a generalized Rachford-Rice problem, extended to more than two phases. As they are part of the model, if needed the functions $k_i^{\alpha,eq}(\boldsymbol{\mathcal{I}}_{th}^{m})$ can again be considered as outputs of the equilibrium coefficient based model.

2.3 Thermodynamic equilibrium as state laws

From the general perspective, and as the two above families of models emphasize, it seems legitimate to consider the thermodynamic equilibrium calculation as a "black box" solver of the form:



If such a "black box" solver corresponds to a physically acceptable model of thermodynamic equilibrium, the output functions $\boldsymbol{\theta}_{ph}^{eq}$ and $\boldsymbol{x}^{eq}$ must satisfy some simple properties. Indeed, the functions $\boldsymbol{\theta}_{ph}^{eq}$ must obviously satisfy:

$$\sum_{\alpha=0}^{N_{ph}-1} \theta_{\alpha}^{eq}(\boldsymbol{\mathcal{I}}_{th}^{m}) = 1,$$

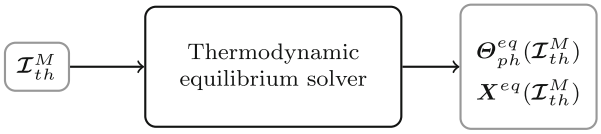
and

$$\theta_{\alpha}^{eq}(\boldsymbol{\mathcal{I}}_{th}^{m}) \geq 0 \quad \text{for all} \quad 0 \leq \alpha \leq N_{ph} - 1.$$

In the same way, the functions $\boldsymbol{x}^{eq}$ must satisfy for any phase α

$$\sum_{i=0}^{N_{comp}-1} x_i^{\alpha,eq}(\boldsymbol{\mathcal{I}}_{th}^{m}) = 1.$$

Thermodynamic equilibria are usually defined in terms of molar quantities. However from the Darcy-flow perspective, it is more natural to express mass balance rather than molar amounts balance, thus it is interesting to convert those molar state functions into mass state functions (see Appendix A for details). This immediately provides the equivalent mass "black box" solver:



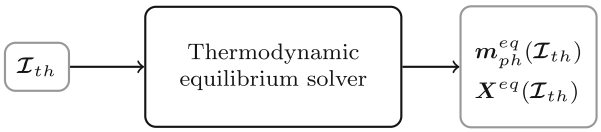
where the mass input is given by

$$\boldsymbol{\mathcal{I}}_{th}^{M} = (P_{ref}, \boldsymbol{P}_{ph}, T, \boldsymbol{Z}_{comp}).$$

In the case where one prefers to directly use masses instead of mass fractions, which is what we favor in practice, denoting:

$$\boldsymbol{\mathcal{I}}_{th} = (P_{ref}, \boldsymbol{P}_{ph}, T, \boldsymbol{m}_{comp}),$$

we can equivalently see such a mass thermodynamic equilibrium as the alternative "black box" solver:



the $\boldsymbol{Z}_{comp}$ being directly computable from the $\boldsymbol{m}_{comp}$ by (the total mass being assumed to always be non-zero):

$$Z_i(\boldsymbol{\mathcal{I}}_{th}) = \frac{m_i}{\sum_{j=0}^{N_{comp}-1} m_j}.$$

As in the molar case, the output functions satisfy the constraints:

$$\sum_{\alpha=0}^{N_{ph}-1} m_{\alpha}^{eq}(\boldsymbol{\mathcal{I}}_{th}) = \sum_{i=0}^{N_{comp}-1} m_i, \tag{6}$$

and

$$m_{\alpha}^{eq}(\boldsymbol{\mathcal{I}}_{th}) \geq 0 \quad \text{for all} \quad 0 \leq \alpha \leq N_{ph} - 1. \tag{7}$$

The functions $\boldsymbol{X}^{eq}$ must still satisfy for any phase α

$$\sum_{i=0}^{N_{comp}-1} X_i^{\alpha,eq}(\boldsymbol{\mathcal{I}}_{th}) = 1. \tag{8}$$

3 Repartition coefficients

Many thermodynamic models underlying the above abstract thermodynamic state functions require solving costly non-linear problems. To completely avoid any solver call when coupling thermodynamics with flow, it is tempting to look for correlations to directly compute θ_{α} and x_i^{α}. However,

contrary to equilibrium coefficients, for which one can check that they generally evolve slowly with composition, this cannot hold for fractions, in particular for phase fractions. Indeed, even in the simplest case of one phase with only one component, we have $\theta_\alpha = z_i$, and it is clear that when considering equilibrium the dependency in z_i will still be very strong as well as more non-linear. If one wants to use tables to represent this kind of correlation, very fine tabulations in terms of composition would be required to keep reasonable precision on the thermodynamic behavior.

This reasoning is the basis from which the idea of repartition coefficients was born, even if those coefficients will ultimately have a much more general meaning, completely disconnected of the practical way chosen to compute them. The idea is to try to extract the major part of the composition dependency of the thermodynamic state functions, which can ultimately ease the derivation of precise enough correlations.

We define the mass repartition coefficient η_i^α of component i in phase α by setting:

$$m_i^\alpha = \eta_i^\alpha m_i. \tag{9}$$

From this definition, we deduce:

$$m_\alpha = \sum_{i=0}^{N_{comp}-1} m_i^\alpha = \sum_{i=0}^{N_{comp}-1} \eta_i^\alpha m_i.$$

Dividing by the total mass, we obtain:

$$\Theta_\alpha = \sum_{i=0}^{N_{comp}-1} \eta_i^\alpha Z_i.$$

In the same way, as soon as $\Theta_\alpha > 0$, we get:

$$X_i^\alpha = \frac{m_i^\alpha}{m_\alpha} = \frac{\eta_i^\alpha Z_i}{\Theta_\alpha} = \frac{\eta_i^\alpha Z_i}{\sum_{j=0}^{N_{comp}-1} \eta_j^\alpha Z_j} = \frac{\eta_i^\alpha m_i}{\sum_{j=0}^{N_{comp}-1} \eta_j^\alpha m_j},$$

with the two previous formulae being still valid when the mass of phase α is zero. Finally notice that:

$$m_i = \sum_\alpha \delta_i^\alpha m_i^\alpha = \sum_\alpha \delta_i^\alpha \eta_i^\alpha m_i,$$

and thus

$$\sum_\alpha \delta_i^\alpha \eta_i^\alpha = 1. \tag{10}$$

It should be clear that starting from any model of thermodynamic equilibrium (EOS, etc...) one can always define the corresponding equilibrium repartition coefficients law by setting for any $0 \le \alpha \le N_{ph} - 1$ and any $0 \le i \le N_{comp} - 1$:

$$\eta_i^{\alpha,eq}(\mathcal{I}_{th}) = \begin{cases} \dfrac{X_i^{\alpha,eq}(\mathcal{I}_{th}) m_\alpha^{eq}(\mathcal{I}_{th})}{m_i} & \text{if } m_i > 0 \\ \dfrac{\delta_i^\alpha}{\sum_{\beta=0}^{N_{ph}-1} \delta_i^\beta} & \text{otherwise.} \end{cases} \tag{11}$$

Conversely, given an equilibrium repartition coefficient law $\eta_i^{\alpha,eq}(\mathcal{I}_{th})$, we define the corresponding thermodynamic equilibrium state law by setting $0 \le \alpha \le N_{ph} - 1$:

$$m_\alpha^{eq}(\mathcal{I}_{th}) = \sum_{i=0}^{N_{comp}-1} \eta_i^{\alpha,eq}(\mathcal{I}_{th}) m_i,$$

and for any $0 \le \alpha \le N_{ph} - 1$ and any $0 \le i \le N_{comp} - 1$, if $m_\alpha^{eq}(\mathcal{I}_{th}) > 0$

$$X_i^{\alpha,eq}(\mathcal{I}_{th}) = \frac{\eta_i^{\alpha,eq}(\mathcal{I}_{th}) m_i}{\sum_{j=0}^{N_{comp}-1} \eta_j^{\alpha,eq}(\mathcal{I}_{th}) m_j},$$

and otherwise:

$$X_i^{\alpha,eq}(\mathcal{I}_{th}) = \frac{\delta_i^\alpha}{\sum_{j=0}^{N_{comp}-1} \delta_j^\alpha}.$$

Provided that Eq. 10 is satisfied, it is immediate to check that Eqs. 6, 7 and 8 will automatically also be satisfied, leading to an admissible thermodynamic model. If needed, the same can be done in terms of moles, defining the molar repartition coefficient $\eta_i^{\alpha,m}$ of component i in phase α by setting:

$$n_i^\alpha = \eta_i^{\alpha,m} n_i,$$

and we get exactly the same relations as in the case of mass quantities:

$$n_\alpha = \sum_{i=0}^{N_{comp}-1} \eta_i^{\alpha,m} n_i \quad \text{and} \quad \theta_\alpha = \sum_{i=0}^{N_{comp}-1} \eta_i^{\alpha,m} z_i,$$

and if $\theta_\alpha > 0$:

$$x_i^\alpha = \frac{n_i^\alpha}{n_\alpha} = \frac{\eta_i^{\alpha,m} z_i}{\theta_\alpha} = \frac{\eta_i^{\alpha,m} z_i}{\sum_{j=0}^{N_{comp}-1} \eta_j^{\alpha,m} z_j} = \frac{\eta_i^{\alpha,m} n_i}{\sum_{j=0}^{N_{comp}-1} \eta_j^{\alpha,m} n_j},$$

along with

$$\sum_\alpha \delta_i^\alpha \eta_i^{\alpha,m} = 1.$$

4 A global mass formulation for compositional multiphase Darcy flow in porous media

4.1 Reference species equations

The molar balance of each chemical species (i, α) is given by:

$$\frac{\partial n_i^\alpha}{\partial t} + div\left(\frac{\rho_\alpha x_i^\alpha}{M_{m_\alpha}} \boldsymbol{V}_\alpha\right) = Q_i^\alpha + R_i^{eq,\alpha}, \tag{12}$$

where:

- $\boldsymbol{V}_\alpha$ is the Darcy velocity of phase α, given by:

$$\boldsymbol{V}_\alpha = -\Lambda \frac{kr_\alpha}{\mu_\alpha} \left(\nabla P_\alpha + \rho_\alpha g \boldsymbol{e}_z\right), \tag{13}$$

  where Λ is the permeability tensor, kr_α the relative permeability of phase α, g is the gravity acceleration, the axis $\boldsymbol{e}_z$ being directed upwards (we recall that μ_α is the viscosity of phase α and ρ_α is the density of phase α).
- Q_i^α is the source term for species (i, α)
- $R_i^{eq,\alpha}$ is the reaction rate corresponding to the thermodynamic equilibrium for species (i, α)

Those equations are complemented by the saturation closure relation:

$$\sum_{\alpha=0}^{N_{ph}-1} S_\alpha = 1,$$

and a pressure equation for each phase

$$P_\alpha = P_{ref} + Pc_\alpha, \tag{14}$$

where P_{ref} is the reference pressure, equal in general to the pressure of one of the liquid phases (here we use the water phase but the oil phase is another classical choice), and Pc_α is the capillary pressure law for phase α, describing the pressure difference between phase α and the reference phase (thus if the reference phase is denoted α_{ref}, we have by definition $Pc_{\alpha_{ref}} = 0$). Of course the system is finally complemented by a thermodynamic equilibrium model, as described in Section 2.

4.2 Component equations

As the components are conserved quantities for thermodynamic equilibrium, it is natural to rewrite the full conservation system in terms of those components only. Summing the species equations on all phases, we obtain:

$$\frac{\partial n_i}{\partial t} = \sum_{\alpha=0}^{N_{ph}-1} \frac{\partial n_i^\alpha}{\partial t}$$

$$= -\sum_{\alpha=0}^{N_{ph}-1} div\left(\frac{\rho_\alpha x_i^\alpha}{M_{m_\alpha}} \boldsymbol{V}_\alpha\right) + \sum_{\alpha=0}^{N_{ph}-1} Q_i^\alpha + \sum_{\alpha=0}^{N_{ph}-1} R_i^{eq,\alpha}.$$

As the total mole number of components remains unchanged by the equilibrium reactions, we have by construction:

$$\sum_{\alpha=0}^{N_{ph}-1} R_i^{eq,\alpha} = 0.$$

The overall system can thus be rewritten $\forall\ 0 \leq i \leq N_{comp} - 1$:

$$\sum_{\alpha=0}^{N_{ph}-1} \frac{\partial n_i^\alpha}{\partial t} + div\left(\frac{\rho_\alpha x_i^\alpha}{M_{m_\alpha}} \boldsymbol{V}_\alpha\right) - Q_i^\alpha = 0. \tag{15}$$

We denote Q_i^m the molar source of component i defined by:

$$Q_i^m = \sum_{\alpha=0}^{N_{ph}-1} Q_i^\alpha,$$

and

$$\rho_{m,\alpha} = \frac{\rho_\alpha}{M_{m_\alpha}},$$

the molar density of phase α. With those notations, we naturally obtain for $0 \leq i \leq N_{comp} - 1$:

$$\frac{\partial n_i}{\partial t} + \sum_{\alpha=0}^{N_{ph}-1} div\left(\rho_{m,\alpha} x_i^\alpha \boldsymbol{V}_\alpha\right) = Q_i^m. \tag{16}$$

Notice that:

$$\frac{x_i^\alpha M_{m_i}}{M_{m_\alpha}} = \frac{n_i^\alpha M_{m_i}}{n_\alpha M_{m_\alpha}} = \frac{m_i^\alpha}{m_\alpha} = X_i^\alpha,$$

thus we get

$$\frac{\rho_\alpha x_i^\alpha}{M_{m_\alpha}} = \frac{\rho_\alpha x_i^\alpha M_{m_i}}{M_{m_i} M_{m_\alpha}} = \frac{\rho_\alpha X_i^\alpha}{M_{m_i}},$$

and thus multiplying (16) by M_{m_i}, we get the mass balance equation of component i:

$$\frac{\partial m_i}{\partial t} + \sum_{\alpha=0}^{N_{ph}-1} div\left(\rho_\alpha X_i^\alpha \boldsymbol{V}_\alpha\right) = Q_i, \tag{17}$$

where we have denoted

$$Q_i = Q_i^m M_{m_i}$$

the mass source term for component i.

4.3 A global mass formulation

The key idea of the global mass formulation is to combine the mass balance equation of the components with the abstract functional rewriting of the thermodynamic equilibrium introduced in Section 2. In other words, we consider that the thermodynamic equilibrium model is given as a function with inputs

$$\mathcal{I}_{th} = (P_{ref}, \boldsymbol{P}_{ph}, T, \boldsymbol{m}_{comp}),$$

and outputs:

$$(X_i^\alpha)_{0\leq i\leq N_{comp}-1, 0\leq\alpha\leq N_{ph}-1} \quad \text{and} \quad (m_\alpha)_{0\leq\alpha\leq N_{ph}-1}.$$

Thus, we assume that we have at our disposal the following functions, $\forall\ 0 \leq i \leq N_{comp}-1$ and $\forall\ 0 \leq \alpha \leq N_{ph}-1$:

$$\left|\begin{array}{l} \mathcal{I}_{th} \longrightarrow X_i^{\alpha,eq}(\mathcal{I}_{th}) \\ \mathcal{I}_{th} \longrightarrow m_\alpha^{eq}(\mathcal{I}_{th}). \end{array}\right. \tag{18}$$

Then, we have to rewrite the flow model in such a way that the thermodynamic equilibrium only appears through the above functions. To this end, it suffices to remark that for any phase α, if Φ denotes the porosity, then:

$$m_\alpha = \rho_\alpha \Phi S_\alpha,$$

and thus the global mass formulation for Darcy flow in porous media is given by:

$$\frac{\partial m_i}{\partial t} + \sum_{\alpha=0}^{N_{ph}-1} div\left(\rho_\alpha X_i^{\alpha,eq}(\mathcal{I}_{th})\boldsymbol{V}_\alpha\right) = Q_i, \tag{19}$$

$$\rho_\alpha \Phi S_\alpha = m_\alpha^{eq}(\mathcal{I}_{th}), \tag{20}$$

$$\sum_{\alpha=0}^{N_{ph}-1} S_\alpha = 1, \tag{21}$$

$$P_\alpha = P_{ref} + Pc_\alpha \quad \forall\ 0 \leq \alpha \leq N_{ph}-1, \tag{22}$$

with $\boldsymbol{V}_\alpha$ is the Darcy velocity of phase α:

$$\boldsymbol{V}_\alpha = -\Lambda \frac{kr_\alpha}{\mu_\alpha}\left(\nabla P_\alpha + \rho_\alpha g \boldsymbol{e}_z\right),$$

and of course the usual laws providing the densities, viscosities, relative permeabilities and capillary pressures of each phase, along with the source terms.

The functions (18) are assumed to be given by a thermodynamic module, and can take in practice almost any form. Without any impact on the global mass formulation, they could be computed through Gibbs energy minimization, equation of state models, Rachford-Rice simplified models, or given explicitly as analytic functions, tabulated functions, response surfaces, etc...

One of the major advantages of this so-called global mass formulation compared with the classical Coats formulation [3] is that the equation system is independent of the appearance/disappearance of phases, that are simply handled by the values of the m_α^{eq} functions. The disappearance of a phase α indeed simply corresponds to the fact that m_α^{eq} takes the value zero. Thus, the set of equations is always the same, whatever the thermodynamic context.

5 A robust and efficient time scheme for coupling thermodynamics and Darcy flows

5.1 Time semi-discrete global mass formulation

We assume that the time interval $[T_0, T]$ is subdivided into $N_T - 1 > 0$ sub-intervals $[t_n, t_{n+1}]$, with:

$$T_0 = t_0 < t_1 < \cdots < t_n < \cdots < t_{N_T-2} < t_{N_T-1} = T.$$

Our time scheme is based on the classical Euler implicit scheme, except for the way we evaluate the functions $\boldsymbol{m}_{ph}^{eq}$ and $\boldsymbol{X}^{eq}$. More precisely, we consider, denoting: $\Delta t^n = t_{n+1} - t_n$:

$$\frac{m_i^{n+1} - m_i^n}{\Delta t^n}$$

$$+ \sum_{\alpha=0}^{N_{ph}-1} div\left(\rho_\alpha^{n+1} X_i^{\alpha,eq}(\mathcal{I}_{th}^n, \mathcal{I}_{th}^{n+1})\boldsymbol{V}_\alpha^{n+1}\right) = Q_i^{n+1}, \tag{23}$$

where we have denoted

$$\boldsymbol{V}_\alpha^{n+1} = -\Lambda \frac{kr_\alpha^{n+1}}{\mu_\alpha^{n+1}}\left(\nabla P_\alpha^{n+1} + \rho_\alpha^{n+1} g \boldsymbol{e}_z\right),$$

$$\rho_\alpha^{n+1} \Phi S_\alpha^{n+1} = m_\alpha^{eq}(\mathcal{I}_{th}^n, \mathcal{I}_{th}^{n+1}), \tag{24}$$

$$\sum_{\alpha=0}^{N_{ph}-1} S_\alpha^{n+1} = 1, \tag{25}$$

$$P_\alpha^{n+1} = P_{ref}^{n+1} + Pc_\alpha^{n+1} \quad \forall\ 0 \leq \alpha \leq N_{ph}-1. \tag{26}$$

The slight abuse of notation on $\boldsymbol{m}_{ph}^{eq}$ and $\boldsymbol{X}^{eq}$ is intended to suggest that we evaluate the thermodynamic functions $\boldsymbol{m}_{ph}^{eq}$ and $\boldsymbol{X}^{eq}$ with a mixing of the explicit $\mathcal{I}_{th}^n$ and implicit $\mathcal{I}_{th}^{n+1}$ thermodynamic inputs. Correctly choosing which

part of $\mathcal{I}_{th}$ is explicit or implicit depends on which thermodynamic solver underlines the functions $\boldsymbol{m}_{ph}^{eq}$ and $\boldsymbol{X}^{eq}$, and is the key ingredient to obtain robust and efficient schemes. An essential condition is that:

$$\sum_{\alpha=0}^{N_{ph}-1} m_\alpha^{eq}(\mathcal{I}_{th}^n, \mathcal{I}_{th}^{n+1}) = \sum_{i=0}^{N_{comp}-1} m_i^{n+1}.$$

Indeed, all the semi-implicit choices satisfying the above constraint will lead to schemes preserving the expected physical bounds by construction (this is a direct consequence of the positivity of the m_α^{eq} functions). Completely defining a scheme now consists in explaining how we define and evaluate the functions $\boldsymbol{m}_{ph}^{eq}$ and $\boldsymbol{X}^{eq}$. Using the repartition coefficients based formulation, we will see that one can derive an original familiy of schemes which remain robust and efficient independently of the chosen thermodynamic equilibrium model (EOS, equilibrium coefficients, etc...).

Once the functions $\boldsymbol{m}_{ph}^{eq}$ and $\boldsymbol{X}^{eq}$ are properly defined, the numerical scheme consists in solving Eqs. 23 to 26 for the pressure unknowns

$$(P_{ref}^{n+1}, (P_\alpha^{n+1})_{0\le\alpha\le N_{ph}-1}),$$

the saturation unknowns

$$(S_\alpha^{n+1})_{0\le\alpha\le N_{ph}-1},$$

and the component mass unknowns:

$$(m_i^{n+1})_{0\le i\le N_{comp}-1},$$

i.e. $N_{comp} + 2N_{ph} + 1$ unknowns (the full set being denoted $\boldsymbol{Y}$), corresponding to the N_{comp} mass balance equations (23) , the N_{ph} (24) phase conservation equations, the saturation closure equation (25) and the N_{ph} capillary pressure equations (26). We of course assume that densities, relative permeabilities, capillary pressures and source terms are given through functions ρ_α^{dat}, kr_α^{dat}, Pc_α^{dat} and Q_i^{dat} of the problem unknowns $\boldsymbol{Y}$, and can correspond to any usual model (for instance Brooks and Corey models for the relative permeabilities and capillary pressures, mixing laws for the phase densities, Peaceman's well model for the source term, etc...).

We are now going to detail the different schemes we have obtained considering different options for computing the thermodynamic functions $\boldsymbol{m}_{ph}^{eq}$ and $\boldsymbol{X}^{eq}$. Some will require tabulations as functions of $(P_{ref}, T, \boldsymbol{Z}_{comp})$, and to this end we will consider two kinds of interpolations: type I consists in a fully linear interpolation in all variables, while type II is a linear interpolation in (P_{ref}, T) and piecewise constant in $\boldsymbol{Z}_{comp}$. This second kind of interpolation leads to substantial computational and memory savings, at the expense of a decreased precision. On Table 1, we display a summary of the tested schemes.

Table 1 Definition of the tested schemes

	Tabulated data	Method for $\boldsymbol{\eta}^{eq}(\mathcal{I}_{th}^n)$
$S1$	None	EOS on $\mathcal{I}_{th}^n$
$S2$	$\boldsymbol{K}^{eq}(\mathcal{I}_{th}^n)$, type I	Rachford-Rice on $\boldsymbol{K}^{eq}(\mathcal{I}_{th}^n)$ and $\boldsymbol{Z}_{comp}^n$
$S3$	$\boldsymbol{K}^{eq}(\mathcal{I}_{th}^n)$, type II	Rachford-Rice on $\boldsymbol{K}^{eq}(\mathcal{I}_{th}^n)$ and $\boldsymbol{Z}_{comp}^n$
$S4$	$\boldsymbol{\eta}^{eq}(\mathcal{I}_{th}^n)$, type I	Interpolation
$S5$	$\boldsymbol{\eta}^{eq}(\mathcal{I}_{th}^n)$, type II	Interpolation

5.2 Repartition coefficients based time semi-explicit scheme

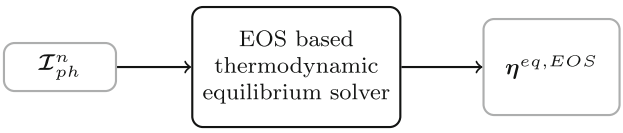
This family of schemes is in our opinion the main contribution of the present work, as it leads to robust and efficient numerical schemes in practice. The key idea is very simple: we use explicit repartition coefficients and implicit component masses, and reconstruct the phase masses and compositions using formulae of Section 3. In other words, we define for any $0 \le \alpha \le N_{ph} - 1$:

$$m_\alpha^{eq}(\mathcal{I}_{th}^n, \mathcal{I}_{th}^{n+1}) = \sum_{i=0}^{N_{comp}-1} \eta_i^{\alpha,eq}(\mathcal{I}_{ph}^n) m_i^{n+1}, \tag{27}$$

and for any $0 \le \alpha \le N_{ph} - 1$ and any $0 \le i \le N_{comp} - 1$:

$$X_i^{\alpha,eq}(\mathcal{I}_{th}^n, \mathcal{I}_{th}^{n+1}) = \frac{\eta_i^{\alpha,eq}(\mathcal{I}_{ph}^n) m_i^{n+1}}{\sum_{j=0}^{N_{comp}-1} \eta_j^{\alpha,eq}(\mathcal{I}_{ph}^n) m_j^{n+1}}. \tag{28}$$

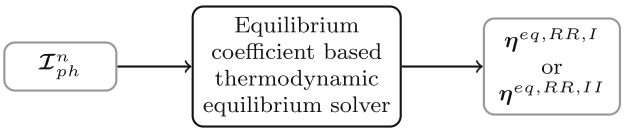
The above formula defines in fact a family of schemes: given the values $\mathcal{I}_{ph}^n$, one still has to choose a thermodynamic model like the ones described in Section 2 to fully define the way the repartition coefficients are computed in practice. Notice however that as we use only the explicit thermodynamic input for the η_i^α's, then the only non zero derivatives with respect to $\boldsymbol{Y}^{n+1}$ of the functions $\boldsymbol{m}_{ph}^{eq}$ and $\boldsymbol{X}^{eq}$ defined through Eqs. 27 and 28 come from the dependency in m_i^{n+1} of the above formulae Eqs. 27 and 28 and not from the thermodynamic solver, from which only values are requested. Thus, it can truly be used as a black box solver. We consider here several ways to compute the explicit values of the repartition coefficients. At first, we consider a solver based on EOS based thermodynamic equilibrium sub-problems:



Scheme $S1$: (27)-(28) with $\eta_i^{\alpha,eq}(\mathcal{I}_{th}^n) = \eta_i^{\alpha,EOS}(\mathcal{I}_{th}^n)$,

where $\boldsymbol{\eta}^{eq,EOS}$ denotes the repartition coefficients obtained by applying Eq. 11 to the solution of the EOS based thermodynamic equilibrium problem. Scheme $S1$ requires solving an EOS based sub-problem each time the repartition coefficients $\eta_i^{\alpha,eq}$ are evaluated. However, as they are evaluated for the explicit input $\mathcal{I}_{th}^n$, this can be done once per time step of the flow, and not at each iteration of the flow non-linear solver. Using those pre-computed values for the repartition coefficients, one then uses the analytic relations (27)–(28) to evaluate the functions $\boldsymbol{m}_{ph}^{eq}$ and $\boldsymbol{X}^{eq}$ during the flow non-linear solver iterations, without any further call to a thermodynamic solver. No tabulated data is needed for this scheme, however the usual properties of components (molar mass, critical temperature, critical pressure, acentric factor, etc...) are required for computing the associated EOS. We do not elaborate any further on the algorithm used in practice to solve the EOS based sub-problem, as we use a classical accelerated successive substitution [11]. Since it is EOS-based, scheme $S1$ should be considered here as the most complete and accurate version of the use of repartition coefficients, and in our numerical experiments we will consider it as our reference scheme. The next four schemes we introduce now are in fact attempts to speed up the online evaluation of the repartition coefficients, regarding the expensive call to the EOS solver of $S1$. To derive them we gradually simplify and in principle accelerate the handling of the thermodynamic sub-problems underlying the evaluation of the explicit repartition coefficients $\eta_i^{\alpha,eq}$.

The next two schemes replace the EOS based thermodynamic sub-problem by an equilibrium coefficient based one, with tabulated equilibrium coefficients



where $\boldsymbol{\eta}^{eq,RR,I}$ and $\boldsymbol{\eta}^{eq,RR,II}$ denote the repartition coefficients obtained by applying (11) to the solution of the equilibrium coefficient based thermodynamic model with tabulated equilibrium coefficients of type I or II. The explicit values $K_i^{\alpha,eq,I}(P_{ref}^n, T^n, \boldsymbol{Z}_{comp}^n)$ for type I or $K_i^{\alpha,eq,II}(P_{ref}^n, T^n, \boldsymbol{Z}_{comp}^n)$ for type II of the equilibrium coefficients are used in the formally identical mass counterpart of Eq. 5 (see Appendix A), while the explicit fractions $\boldsymbol{Z}_{comp}^n$ computed from $\boldsymbol{m}_{comp}^n$ are used in the mass counterpart of Eq. 3 (again, see Appendix A). Scheme $S2$ uses tabulated equilibrium coefficients of type I

$$Scheme\ S2\ :\ (27)\text{-}(28)\ \text{with}\ \eta_i^{\alpha,eq}(\mathcal{I}_{th}^n) = \eta_i^{\alpha,RR,I}(\mathcal{I}_{th}^n),$$

while scheme $S3$ uses tabulated equilibrium coefficients of type II

$$Scheme\ S3\ :\ (27)\text{-}(28)\ \text{with}\ \eta_i^{\alpha,eq}(\mathcal{I}_{th}^n) = \eta_i^{\alpha,RR,II}(\mathcal{I}_{th}^n).$$

In the same way as scheme $S1$, schemes $S2$ and $S3$ require solving a thermodynamic equilibrium sub-problem each time the repartition coefficients $\eta_i^{\alpha,eq}$ are evaluated. Again, as they are only evaluated on the explicit input $\mathcal{I}_{th}^n$ this can be done once per time step and the stored values are finally used during the non-linear flow solver iterations to compute the functions $\boldsymbol{m}_{ph}^{eq}$ and $\boldsymbol{X}^{eq}$ through Eqs. 27–28. The next two schemes directly use tabulated repartition coefficients, of type I (denoted $\eta_i^{\alpha,Tab,I}$) for scheme $S4$

$$Scheme\ S4\ :\ (27)\text{-}(28)\ \text{with}$$
$$\eta_i^{\alpha,eq}(\mathcal{I}_{th}^n) = \eta_i^{\alpha,Tab,I}(P_{ref}^n, T^n, \boldsymbol{Z}_{comp}^n),$$

and of type II (denoted $\eta_i^{\alpha,Tab,II}$) for scheme $S5$

$$Scheme\ S5\ :\ (27)\text{-}(28)\ \text{with}$$
$$\eta_i^{\alpha,eq}(\mathcal{I}_{th}^n) = \eta_i^{\alpha,Tab,II}(P_{ref}^n, T^n, \boldsymbol{Z}_{comp}^n).$$

As the functions $\eta_i^{\alpha,eq}$ are tabulated for schemes $S5$ and $S1$, no call to a thermodynamic solver occurs during the flow simulation.

As our numerical experiments on those five schemes confirm, the underlying black box thermodynamic equilibrium model has no major impact on the behavior of the flow non-linear solver. Indeed, at each time step, the repartition of the components among phases is fixed. Thus, phase appearance and disappearance is in some sense fixed at the beginning of the time step (at least from the flow solver perspective). If the global mass formulation underlying our schemes might be surprising for readers more familiar with Coats' formulation, the idea underlying it is not particularly new in the literature. First traces of such a formulation date back to 1983 [4, 21], and its molar counterpart is used for instance in [9, 18]. The true originality of our formulation relies in fact on one hand on the way we have chosen to represent the thermodynamic equilibrium through repartition coefficients and on the other hand on the fact that we use Eq. 20 to define the saturations. Indeed classical global formulations, using molar as well as mass unknowns, that can be found in the classical literature directly impose the saturations by transforming Eqs. 20 into laws S_α^{eq} fixing the saturations, or by an equivalent process. We have of course experimented with this classical version for the definition of the saturations, and unsurprisingly we have recovered the results of Voskov and Tchelepi [19], that is the non convergence of the Newton solver as soon as the petrophysical model becomes quite complex. This is the reason why, even if both versions are formally identical, we have chosen to keep the saturations as unknowns in the system and not defined them as laws. This

has indeed a tremendous impact on the non-linear solver behavior, and is the key to achieve robustness for the global mass formulation. This can be easily explained by the fact that the saturations variations control the relative permeability and capillary pressure variations for most common petrophysical models, thus they consequently control the main non-linearities of Darcy flux. Keeping the saturations as unknowns and controlling them with the usual relaxation strategies allows to control the flux variations throughout the non-linear solver iterations and thus greatly improves the convergence properties of the overall method.

6 Numerical exploration

In the following, we consider thermodynamic systems with three phases, the aqueous phase (or water phase), the liquid hydrocarbon phase (or oil phase) and the gaseous hydrocarbon phase (or gas phase). We also assume that Peng-Robinson's EOS is a suitable model for each phase. In all the considered test cases, the source term is given by classical Peaceman's wells.

When tabulated values for either $\boldsymbol{K}^{eq}$ or $\boldsymbol{\eta}^{eq}$ are required by a scheme, they are of course generated in a preprocessing step by solving the EOS model on a wide range of inputs $(P_{ref}, T, \boldsymbol{Z}_{comp})$. We use a regular grid that ranges from 0.0 to 1.0 for the $\boldsymbol{Z}_{comp}$, from 274.15 Kto 625.15 K for the temperature and from 1.0 MPa to 161.0 MPa for the reference pressure, with 11 points in each Z_i direction, 8 for the temperature and 33 for the reference pressure. As we keep only the cells of the underlying Cartesian tabulation mesh that intersect the hyperplane $\sum_i Z_i = 1$, with 4 components this gives tables with 640 992 entries for type I and 229 944 for type II, while with 5 components this leads to 3 003 792 entries for type I and 758 208 for type II. If one discards the tabulated data loading time that will be negligible on test cases of industrial interest, we naturally expect scheme $S1$ to be the most computationally expensive as it requires solving the EOS sub-problem once per time step, while schemes $S2$ and $S3$ only require to solve the simpler Rachford-Rice non-linear sub-problem once per time step. Finally, schemes $S4$ and $S5$ do not require any non-linear solver call to compute the thermodynamic equilibrium functions and are thus expected to be the most computationally efficient ones. However, we will see that this is only true if the test case is large enough.

After choosing one or the other of the above time discretizations, it is necessary to correctly define a space discretization. As the goal of the present paper is to describe a new way to write the coupling between thermodynamics and flow and the corresponding time scheme, we have chosen to limit ourselves to the TPFA finite volume scheme [5], that is only valid on permeability tensor orthogonal meshes. To ease the numerical experiments, we thus consider only diagonal and isotropic permeability tensors, thus the mesh only need to be orthogonal in the classical sense and we can resort to Cartesian meshes. Remark that this is by no means a restriction of the presented approach, our new time schemes could be combined with any advanced space scheme for diffusion-convection operators, without any major difference. However as we mainly aim at exploring the ability of the new scheme to efficiently cope with thermodynamic complexities such as phase appearance/disappearance and species exchanges between phases, we felt that we could limit ourselves to such a simplified mesh setting without loosing generality. For both the reader's convenience and the sake of completeness, we recall in Appendix B the details of the TPFA space discretization of our global mass formulation.

6.1 Beam test case

Initial state The first illustration case is a simple beam with a 1D behavior. It measures 100 meter-long with 20 cells and 30 meter-depth with 3 cells. This system mimics a sandstone reservoir whose porosity is 0.33 and permeability 10^{-11} m^2, covered with a shale layer that plays the role of a cap-rock, whose porosity is 0.4 and permeability 10^{-14} m^2. This beam is buried 800 meter-depth. The thermal regime is defined through a simple geothermal gradient of 0.025 K/m with a surface temperature of 291.15 K (18 °C), moreover at initial state the system is at hydrostatic pressure. Boundary conditions are no flow for lateral and top boundaries and hydrostatic pressure for the bottom.

The relative permeabilities of the system are defined with the following relation:

$$kr_\alpha(S_\alpha) = \left(\frac{S_\alpha - S^{kr}_{min}}{S^{kr}_{max} - S^{kr}_{min}} \right)^{n_{kr}}.$$

The capillary pressures of the system are defined with the following relation:

$$Pc_\alpha(S_\alpha) = Pc_e + \Delta Pc \left(\frac{S_\alpha - S^{Pc}_{min}}{S^{Pc}_{max} - S^{Pc}_{min}} \right)^{n_{Pc}}.$$

The different parameters (S^{kr}_{min}, S^{kr}_{max} and n_{kr} for relative permeabilities and Pc_e, ΔPc, S^{Pc}_{min}, S^{Pc}_{max} and n_{Pc} for capillary pressures) for both facies and the three phases can be found on Table 2. Notice that we use a primary drainage capillary pressure curve, typical of basin modeling, to assess the behavior of the scheme with such a stiff datum. Please note that, water phase being in our case the reference phase, $Pc_W = 0$.

An injection well is located at the left of the beam and a production well is located at the right of the beam. The geometry, mesh and wells positions can be seen on Fig. 1.

Table 2 Parameters for petrophysical laws. W, O, G stand respectively for water phase, oil phase and gas phase

	Sandstone			Shale		
	W	O	G	W	O	G
S_{min}^{kr}	0.2	0.02	0.02	0.2	0.02	0.02
S_{max}^{kr}	0.98	0.8	0.8	0.98	0.8	0.8
n_{kr}	2	2	2	2	2	2
$Pc_e\ (MPa)$	/	0	0	/	5	6
$\Delta Pc\ (MPa)$	/	4	4	/	30	30
S_{min}^{Pc}	/	0	0	/	0	0
S_{max}^{Pc}	/	0.8	0.8	/	0.8	0.8
n_{Pc}	/	10	10	/	10	10

Both wellheads reach the surface. The initial state is defined through hydrostatic equilibration and a water-oil contact located 826 m depth. All the thermodynamic properties of the components are described on Table 3. In this system, the component H_2O is only allowed to go in the water phase, and all the other components are allowed to go in both gas and oil phases.

Dynamic state From the beginning and during the whole simulation (300 days), the injection well injects a gas constituted of 100% of CH_4 at a pressure of 10.05 bar, the production well producing all fluids arriving at the right of the beam.

On Fig. 2, the three fluid saturations are plotted as a function of time for scheme $S1$. From the beginning of the simulation to 75 days, the well produces water and oil, the oil fraction decreases while the water fraction increases. At 75 days, a gas phase appears and is then produced leading to a drop of oil saturation. At the end of the simulation, the three fluids are produced together. This result will be the reference in order to compare the other schemes.

Figure 3 shows the importance of the interpolation model: in all the cases, a smoothing of the results can

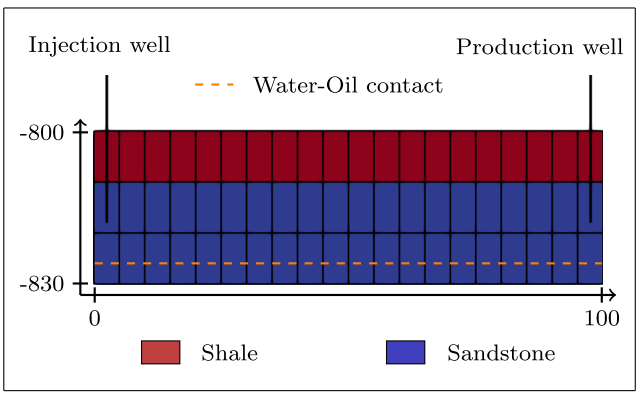


Fig. 1 First illustration case geometry, mesh and wells positions

Table 3 Thermodynamic properties of the 4 components

	$T_{c_i}[K]$	$P_{c_i}[MPa]$	ω_i	$M_{m_i}[g/mol]$
H_2O	647.096	22.064	0.344	18
CH_4	190.56	4.6	0.0111	16.04
C2-C5	353.6	4.432	0.139	41.9
C6+	823	1.8	0.55	300

be observed when type I is used, avoiding the presence of the small peaks that appear for type II interpolation, which decreases the accuracy of the results. Moreover, scheme $S5$, for which type II interpolation is used for representing $\boldsymbol{\eta}^{eq}(\mathcal{I}_{th}^n)$, shows bigger peaks than scheme $S3$ where this is the $\boldsymbol{K}^{eq}(\mathcal{I}_{th}^n)$ which are tabulated. This is consistent with the fact that $\boldsymbol{K}^{eq}(\mathcal{I}_{th}^n)$ are less dependent on the composition than $\boldsymbol{\eta}^{eq}(\mathcal{I}_{th}^n)$. Nevertheless, for all the schemes, the same global trend is observed.

Figure 4 shows the evolution of computational time for the simple beam test, taking scheme $S1$ as the reference. It is important to notice that due to the very small computational time, the comparison might be biased: in fact, the loading of the tabulated data, in particular in the case of type I the interpolation, is not negligible regarding the effective computational time for such a small test case. This is most probably the reason why scheme $S1$ is the most efficient one on this small test case, despite the fact that the four other schemes were introduced as attempts to improve the computational efficiency. Comparing types I and II interpolation, for scheme $S2$ and $S3$ we obtain the expected result, in

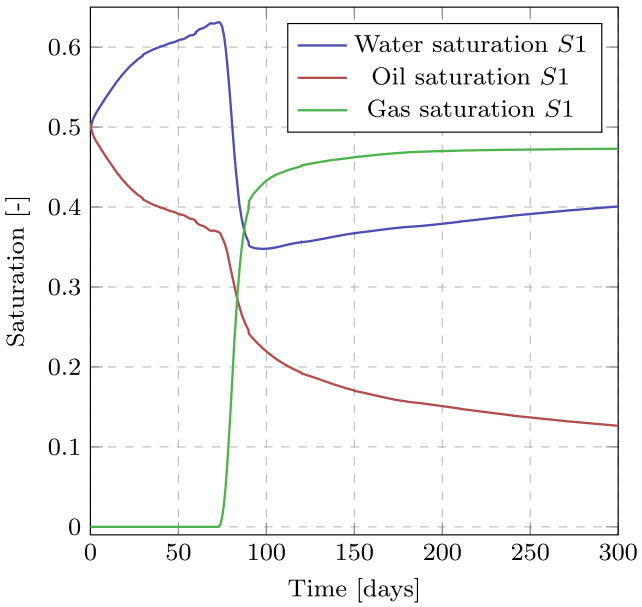


Fig. 2 Fluid saturations in the production well for scheme $S1$ as a function of time

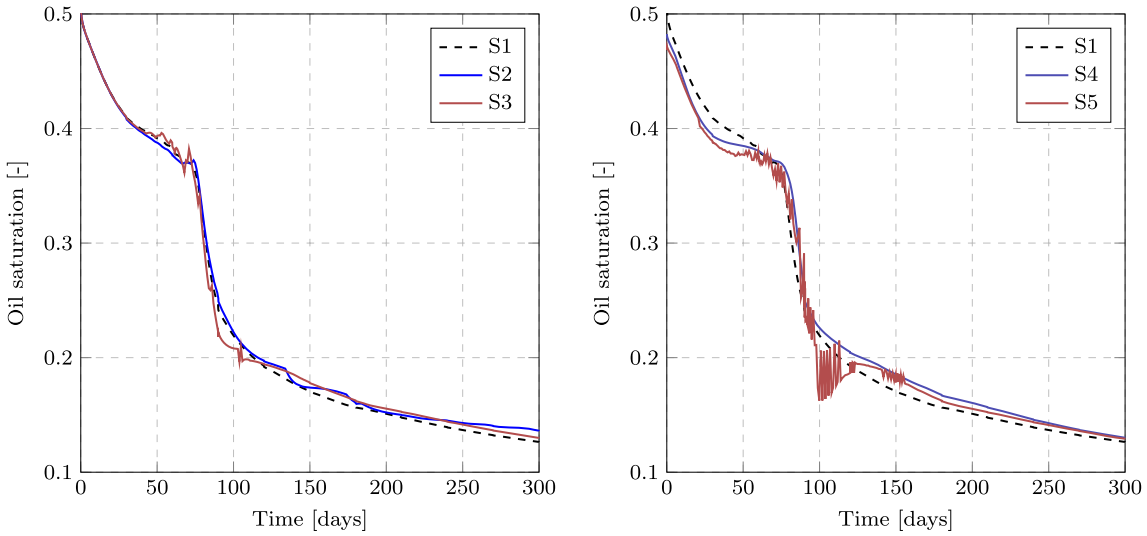


Fig. 3 Oil saturation as a function of time in the production well. In dashed black, for scheme $S1$; in blue, schemes using type I interpolation (schemes $S2$ and $S4$); in red, schemes using type II interpolation (schemes $S3$ and $S5$). In all the situations, the schemes using type I interpolation give smoother results than the schemes using type II interpolation

the sense that scheme $S3$ which uses the cheaper type II interpolation is faster that scheme $S2$ that uses type I which is a full interpolation in all the variables. However, we do not retrieve this result for schemes $S4$ and $S5$, as scheme $S4$ is faster than scheme $S5$. This is counter-intuitive as type I consists in a full interpolation in all the variables, and probably reveals the impact of the oscillations observed on Fig. 3 on the computational time. Illustrating the impact of the interpolation quality on the behavior of the solution is in fact the main interest of this very simple test case.

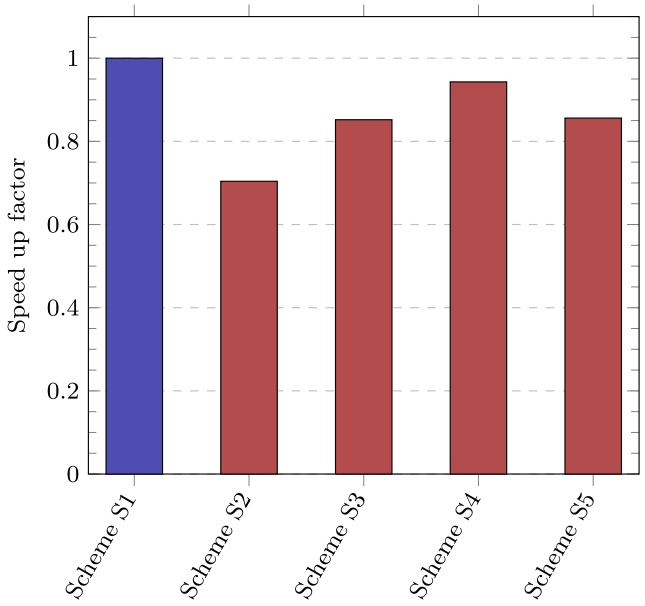


Fig. 4 Speed up factor for the 5 schemes compared to scheme $S1$: a value below 1 means a longer run, a value higher than 1 means a speed up

6.2 Synthetic trap test case

Initial state The second illustration case is a system with a truly 2D behavior. It is 300-meter long and 200-meter deep with 20 cells in each direction. The top of the model is buried at 800 m depth. It is constituted of the 2 same facies as the previous illustration case to mimic two traps: a sandstone reservoir whose porosity is 0.33 and permeability 10^{-11} m^2 with a shale layer whose porosity is 0.4 and permeability 10^{-14} m^2, but this time with a more complex geometry that mimics two anticlinals. The petrophysical properties of both facies can be found on Table 2. Figure 5 illustrates the geometry of the system showing the location of both facies. Both wellheads reach the surface.

An injection well is located at the top of deepest anticlinal (in the first sandstone layer) and it injects pure methane. A production well is located at the top of the shallowest anticlinal (in the first sandstone layer) and produces all fluids arriving. Their positions are shown on Fig. 5. The thermal regime is the same as for the previous illustration case. The initial state is defined through hydrostatic equilibration and a water-oil contact located at 916 m depth. The same composition is considered (see Table 3).

Dynamic state From the beginning and during the whole simulation (600 days), the injection well injects a gas constituted of 100% of CH_4 with a pressure of 10.05 bar, the production well producing all kind of fluids that arrive in the shallowest trap. The gas saturations at the

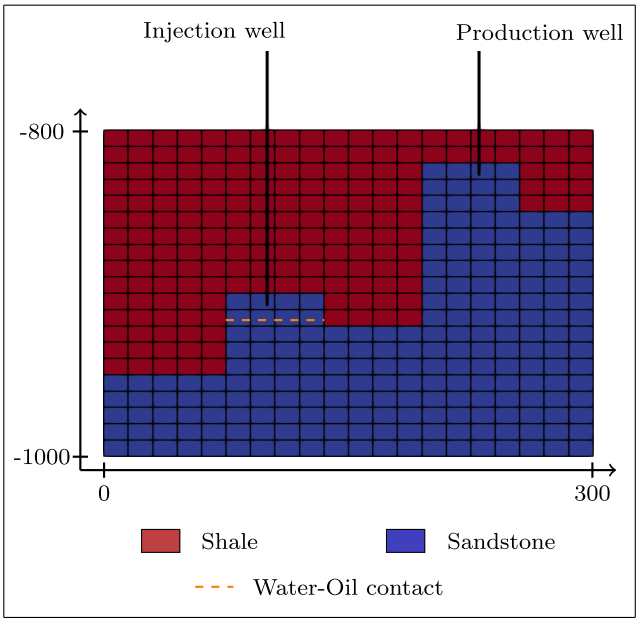


Fig. 5 Second illustratrion case geometry, mesh and wells positions

end of the simulation for all schemes are represented on Fig. 6. All the results remain quite close to each other. In particular schemes $S4$ and $S5$, that uses tabulated repartition coefficients, remain accurate despite the fact tabulations are coarse. The study of computational times (Fig. 7) shows that this time schemes $S3$ and $S4$ do provide a speed up

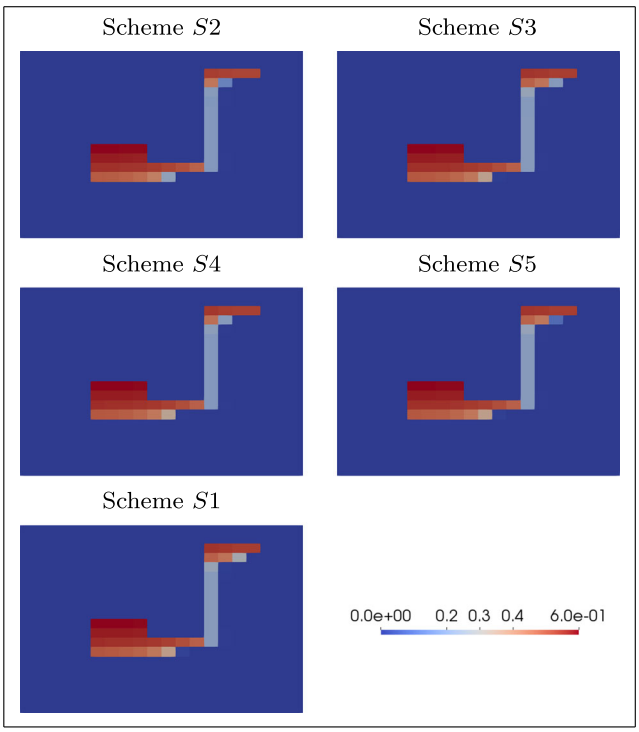


Fig. 6 Gas saturations ar the end of the simulation for the second illustration case

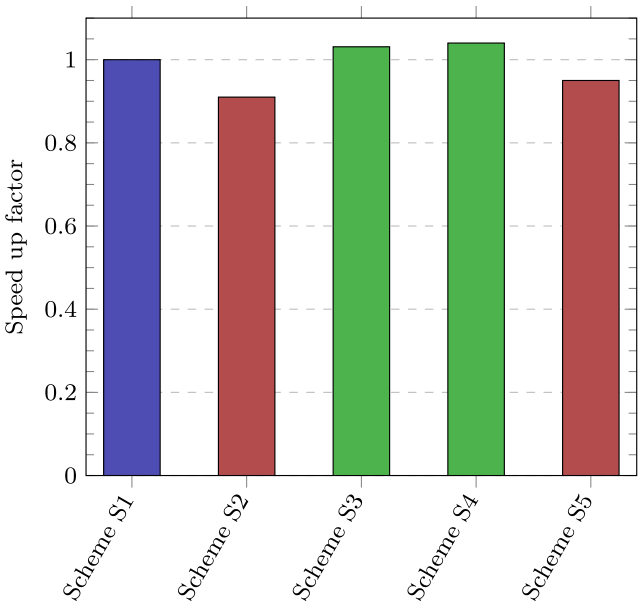


Fig. 7 Speed up factor for the 5 schemes compared to scheme $S1$: a value below 1 means a longer run, a value higher than 1 means a speed up

regarding scheme $S1$, however schemes $S2$ and $S5$ still take more time than scheme $S1$. Comparing type I and type II interpolation, we obtain the same as for the previous illustration case : type II is indeed faster when it is used for tabulating equilibrium coefficients, but not when it is directly used for tabulating repartition coefficients.

6.3 SPE10 test case

Initial state The last illustration case is based on the upper layer of SPE model 2 which consists of part of a Brent sequence. This model is widely used in literature and the reader can easily find a description of this case [2, 19]. The mesh is similar to the model provided by SPE but it is coarsened with a factor 2 in both directions, giving 110 × 30 cells. The upscaling of porosities is an arithmetic mean and the upscaling of permeabilities is a geometric mean. The thickness of the layer is 70 m, such as in Voskov and Tchelepi [19]. Five wells are located on the model: one injection well at (255 m, 255 m) and one production well in each corner. On Table 4, we display the petrophysical parameters used for this test case, while Fig. 8 shows the permeability distribution and the location of the five wells.

Dynamic state From the beginning and during the whole simulation (400 days), the injection well injects 40 vol% of water and 60 vol% of gaseous CO_2 with an imposed bottom hole pressure of 68.948 MPa. The production wells produce all fluids that arrive. The water-oil contact is located at the reservoir bottom, i.e. 3778.8064 m. The same PVT as in

Table 4 Parameters for petrophysical laws. W, O, G stand respectively for water phase, oil phase and gas phase

SPE10			
	W	O	G
S_{min}^{kr}	0.2	0.02	0.02
S_{max}^{kr}	0.98	0.8	0.8
n_{kr}	2	2	2
$Pc_e\ (MPa)$	/	0	0
$\Delta Pc\ (MPa)$	/	4	4
S_{min}^{Pc}	/	0.02	0.02
S_{max}^{Pc}	/	0.8	0.8
n_{Pc}	/	4	4

previous cases (see Table 3) is considered with the addition of CO_2. Thermodynamic properties of CO_2 are:

- $T_{cCO_2} = 304.2$ K,
- $P_{cCO_2} = 7.383$ MPa,
- $\omega_{CO_2} = 0.2236$,
- $M_{mCO_2} = 44.01$ g/mol.

At the beginning of the simulation, the model is homogeneously filled with 34 vol% of oil and 66 vol% of water. On Fig. 9, the gas saturations for all schemes at the end of the simulation are shown. We can see that around the injection well, where gas and water are injected, the gas saturation highly increases, consequently water and oil saturations decrease. The gas never reaches the production wells, so they produce only water and oil. With the exception of scheme $S5$, the same behavior is observed for all schemes (see Fig. 9), with some differences near the gas saturation front reflecting the fact that repartition coefficients of scheme $S1$, the most precise ones, are approximated for all other schemes. Scheme $S5$ however shows relatively large differences. In particular, the oil region of the other schemes is here replaced by a mixture of oil and gas phases. Comparing it to scheme $S4$, which only differs by the quality of the interpolation confirms that the use of both a coarse tabulation and a coarse interpolation is the source of the problem.

Looking at computational times on Fig. 10, on this larger test case we clearly see that tabulated repartition coefficients of schemes $S4$ and $S5$ provide a clear speed up compared to online calls to the EOS solver of $S1$ or the Rachford-Rice solver of $S2$ and $S3$ for computing them. Scheme $S3$ again provides a speed up, while scheme $S2$ still fail to achieve this, which reveals that $S2$ is definitively not a good option. However, the results of Fig. 9, strongly emphasize that some care should be taken when choosing the table's refinement, as here we have some noticeable difference between schemes $S4$-$S5$ and scheme $S1$, contrary to our previous tests. Nevertheless, refining the tables or using more advanced way to represent the repartition coefficients (for instance response surfaces or neural networks) undoubtedly allows to recover precise results while maintaining the computational speed up.

6.4 Concluding remarks on numerical experiments

To conclude this numerical exploration, let us mention that a comprehensive performance study of the proposed approach would require a comparison with a state-of-the-art

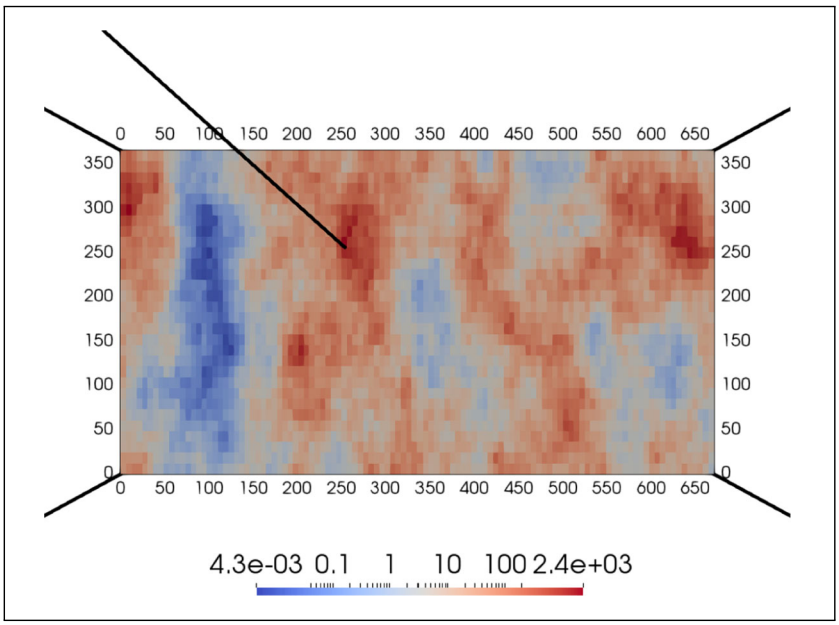


Fig. 8 Top view of the third illustration case permeability distribution with a log scale and the wells positions: 4 production wells are located at the corner of the model and 1 injection well is located at the center

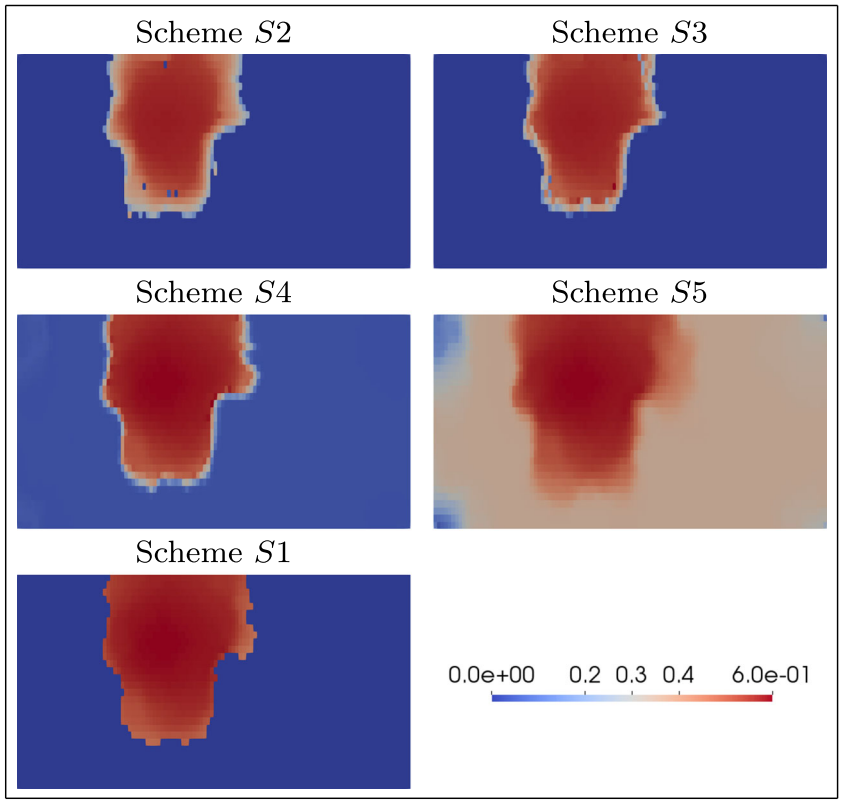


Fig. 9 Gas saturations ar the end of the simulation for the SPE10 test case

solver using Coats' formulation. Such a solver being quite complex to implement and optimize, contrary to our approach that is much simpler in this regard, we would have to resort to available advanced reservoir simulators. Those software being in general also highly optimized, for such a comparison to be fair we need to implement our new scheme with as much optimization ideally within one of those software, rather than using the basic prototype code used here to perform our numerical experiments. Otherwise, most of the performance differences could arguably be considered as a consequence of differences in implementation, programming language, etc... We felt that the available results as well as their simpler formulation will be enough to motivate an implementation of the new repartition coefficients based scheme in an existing reservoir simulator, and this is the reason why we have chosen to postpone such a performance comparison to once this implementation will have been carried on. With any chance, a comparison with alternative formulations such as the fugacity formulation of Lauser et al. [10] could then also be performed. Nevertheless, if one chooses to use the repartition coefficient based semi-explicit time scheme, combined with some kind of fast interpolation for the $\boldsymbol{\eta}^{eq,n}$, we expect a considerable speed up when compared to the usual Coats' formulation involving numerous on the fly EOS solver calls and complex phase appearance/disappearance handling.

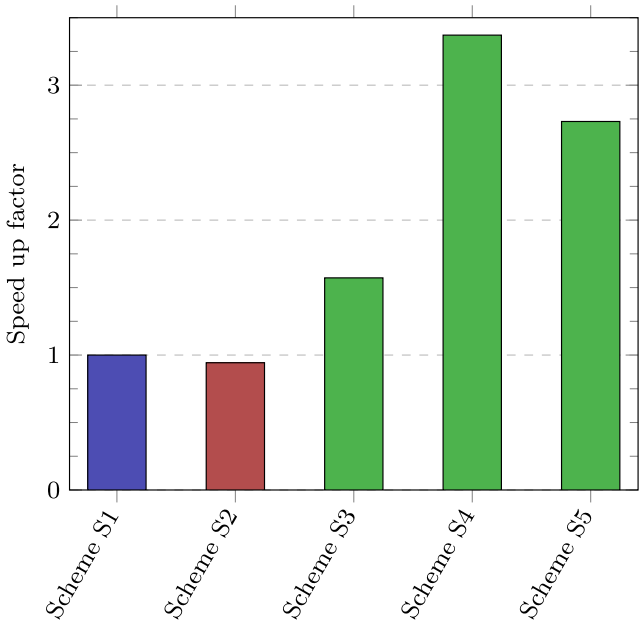


Fig. 10 Speed up factor for the 5 schemes compared to scheme $S1$: a value below 1 means a longer run, a value higher than 1 means a speed up

7 Conclusion

After observing that generic thermodynamic equilibrium models admit a reformulation leading to a global mass

formulation of their coupling with Darcy flow, we have introduced the notion of repartition coefficients. They allow the derivation of a new and efficient numerical scheme whose thermodynamic precision seems to only depend on the precision of the repartition coefficients themselves. Numerical experiments illustrate the robustness of the approach, as well as the potential speed up of using coarse representations of the repartition coefficients rather than online calls to EOS solvers. Future works concern the extension of the repartition coefficients and associated numerical scheme to the more challenging case of reactive equilibrium, as well as a rigorous numerical analysis of the method. Replacing the very basic tabulations used in the present paper to speed up computations by more advanced representations of the repartition coefficients such as response surfaces or neural networks is also the subject of active research.

Appendix A: Thermodynamic conversions

By definition, one has

$$m_i^\alpha = n_i^\alpha M_{m_i}.$$

Then

$$n_i^\alpha = \frac{m_i^\alpha}{M_{m_i}} = m_\alpha \frac{X_i^\alpha}{M_{m_i}},$$

and consequently

$$n_\alpha = \sum_{j=0}^{N_{comp}-1} \frac{m_j^\alpha}{M_{m_j}} = m_\alpha \sum_{j=0}^{N_{comp}-1} \frac{X_j^\alpha}{M_{m_j}},$$

and thus $x_i^\alpha = \frac{n_i^\alpha}{n_\alpha}$ becomes

$$x_i^\alpha = \frac{\dfrac{X_i^\alpha}{M_{m_i}}}{\displaystyle\sum_{j=0}^{N_{comp}-1} \frac{X_j^\alpha}{M_{m_j}}}.$$

In the same way, starting from

$$m_i^\alpha = n_i^\alpha M_{m_i} = n_\alpha x_i^\alpha M_{m_i},$$

we obtain

$$m_\alpha = \sum_{j=0}^{N_{comp}-1} n_j^\alpha M_{m_j} = n_\alpha \sum_{j=0}^{N_{comp}-1} x_j^\alpha M_{m_j},$$

and thus $X_i^\alpha = \frac{m_i^\alpha}{m_\alpha}$ becomes

$$X_i^\alpha = \frac{x_i^\alpha M_{m_i}}{\displaystyle\sum_{j=0}^{N_{comp}-1} x_j^\alpha M_{m_j}}.$$

Thus for the molar masses of phases, depending whether we have molar fractions or mass fractions at our disposal, we can use either:

$$M_{m_\alpha} = \sum_{i=0}^{N_{comp}-1} x_i^\alpha M_{m_i},$$

or

$$M_{m_\alpha} = \frac{\displaystyle\sum_{i=0}^{N_{comp}-1} X_i^\alpha}{\displaystyle\sum_{j=0}^{N_{comp}-1} \frac{X_j^\alpha}{M_{m_j}}} = \frac{1}{\displaystyle\sum_{j=0}^{N_{comp}-1} \frac{X_j^\alpha}{M_{m_j}}},$$

as $\sum_{i=0}^{N_{comp}-1} X_i^\alpha = 1$.

In the same way, for phases we have, if m_{tot} denotes the total mass and n_{tot} the total mole number of the system

$$n_{tot} = \sum_{\alpha=0}^{N_{ph}-1} n_\alpha \quad \text{and} \quad m_{tot} = \sum_{\alpha=0}^{N_{ph}-1} m_\alpha,$$

and $n_\alpha = \frac{m_\alpha}{M_{m_\alpha}}$ give for $\theta_\alpha = \frac{n_\alpha}{n_{tot}}$ and $\Theta_\alpha = \frac{m_\alpha}{m_{tot}}$:

$$\theta_\alpha = \frac{\dfrac{\Theta_\alpha}{M_{m_\alpha}}}{\displaystyle\sum_{\beta=0}^{N_{ph}-1} \frac{\Theta_\beta}{M_{m_\beta}}} \quad \text{and} \quad \Theta_\alpha = \frac{\theta_\alpha M_{m_\alpha}}{\displaystyle\sum_{\beta=0}^{N_{ph}-1} \theta_\beta M_{m_\beta}}.$$

For the total fractions of components,

$$n_{tot} = \sum_{j=0}^{N_{comp}-1} n_j \quad \text{and} \quad m_{tot} = \sum_{j=0}^{N_{comp}-1} m_j,$$

and $m_i = n_i M_{m_i}$ lead for $z_i = \frac{n_i}{n_{tot}}$ et $Z_i = \frac{m_i}{m_{tot}}$ to:

$$z_i = \frac{\dfrac{Z_i}{M_{m_i}}}{\displaystyle\sum_{j=0}^{N_{comp}-1} \frac{Z_j}{M_{m_j}}} \quad \text{and} \quad Z_i = \frac{z_i M_{m_i}}{\displaystyle\sum_{j=0}^{N_{comp}-1} z_j M_{m_j}},$$

finally, the mass equilibrium coefficients are defined through:

$$\frac{X_i^\alpha}{K_i^\alpha} = \frac{X_i^\beta}{K_i^\beta}.$$

Indeed, as:

$$x_i^\alpha = \frac{\frac{X_i^\alpha}{M_{m_i}}}{\sum_{j=0}^{N_{comp}-1} \frac{X_j^\alpha}{M_{m_j}}} = \frac{\frac{X_i^\alpha}{M_{m_i}}}{\sum_{j=0}^{N_{comp}-1} \frac{x_j^\alpha}{\sum_{k=0}^{N_{comp}-1} x_k^\alpha M_{m_k}}}$$
$$= X_i^\alpha \sum_{k=0}^{N_{comp}-1} x_k^\alpha \frac{M_{m_k}}{M_{m_i}},$$

we deduce that we can convert from molar to mass by setting (their is no uniqueness as only the coefficient ratios play a role in the above identity):

$$K_i^\alpha = \frac{k_i^\alpha}{\sum_{j=0}^{N_{comp}-1} x_j^\alpha \frac{M_{m_j}}{M_{m_i}}}.$$

Notice that by construction:

$$\frac{x_i^\alpha}{k_i^\alpha} = \frac{X_i^\alpha}{K_i^\alpha},$$

which means that mass equilibrium is formally identical to molar equilibrium.

Appendix B: Full discretization using TPFA finite volumes

Meshes and notations We assume that the computational domain Ω is an open polygonal subset of $\mathbb{R}^d$, $d = 2$ or 3, such that

$$\overline{\Omega} = \bigcup_{i=0}^{N_{layer}-1} \overline{\Omega_i} \quad \text{where} \quad \Omega_i \cap \Omega_j = \emptyset \text{ if } i \neq j,$$

where the sets $(\Omega_i)_{0\leq i\leq N_{layer}-1}$ are also open polygonal subsets of $\mathbb{R}^d$, in which the geological properties are assumed to evolve continuously (in general, they correspond to geological layers). We recall the usual notations describing a mesh $\mathcal{M} = (\mathcal{T}, \mathcal{F})$ of Ω. $\mathcal{T}$ is a finite family of connected open disjoint polygonal subsets of Ω (the cells of the mesh), such that $\overline{\Omega} = \cup_{K\in\mathcal{T}} \overline{K}$. For any $K \in \mathcal{T}$, we denote by $|K|$ the measure of $|K|$ and by $\partial K = \overline{K} \setminus K$ the boundary of K. $\mathcal{F}$ is a finite family of disjoint subsets of hyperplanes of $\mathbb{R}^d$ included in $\overline{\Omega}$ (the faces of the mesh) such that, for all $\sigma \in \mathcal{F}$, its measure is denoted $|\sigma|$. For any $K \in \mathcal{T}$, there exists a subset $\mathcal{F}_K$ of $\mathcal{F}$ such that $\partial K = \cup_{\sigma\in\mathcal{F}_K} \overline{\sigma}$. Then, for any $\sigma \in \mathcal{F}$, we denote by $\mathcal{T}_\sigma = \{K \in \mathcal{T} \mid \sigma \in \mathcal{F}_K\}$. Next, for all $K \in \mathcal{T}$ and all $\sigma \in \mathcal{F}_K$, we denote by $\boldsymbol{n}_{K,\sigma}$ the unit normal vector to σ outward to K. The set of boundary faces is denoted $\mathcal{F}_{ext}$, while interior faces are denoted $\mathcal{F}_{int}$. We complement the mesh by a family of points $\mathcal{P} = \left((\boldsymbol{x}_K)_{K\in\mathcal{T}}, (\boldsymbol{x}_\sigma)_{\sigma\in\mathcal{F}_{ext}}\right)$ indexed by the cells and boundary faces such that $\boldsymbol{x}_K \in \mathring{K}$ for any $K \in \mathcal{T}$ and $\boldsymbol{x}_\sigma \in \sigma$ and any $\sigma \in \mathcal{F}_{ext}$. If $\mathcal{T}_\sigma = \{K, L\}$, we assume that $\boldsymbol{x}_K \neq \boldsymbol{x}_L$. If $\sigma \in \mathcal{F}_K$, we denote $d_{K,\sigma}$ the distance between $\boldsymbol{x}_K$ and σ. Finally, we assume that for any $0 \leq i \leq N_{layer} - 1$, there exists $\mathcal{F}_i \subset \mathcal{F}$ such that:

$$\partial\Omega_i = \bigcup_{\sigma\in\mathcal{F}_i} \overline{\sigma},$$

thus the mesh is assumed adapted to the geological discontinuities.

To any continuous variable $p(\boldsymbol{x}, t)$, we associate a family of discrete variables $(p_K^n)_{K\in\mathcal{T},0\leq n\leq N_T}$ such that p_K^n is in principle an approximation of $p(\boldsymbol{x}_K, t_n)$. As we consider the TPFA finite volume approximation, to ensure this approximation property we assume that the mesh is $(\Lambda_K)_{K\in\mathcal{T}}$- orthogonal, with $(\Lambda_K)_{K\in\mathcal{T}}$ is the discrete permeability tensor. More precisely, there exists a family of straight lines $(\mathcal{D}_{K,\sigma})_{\sigma\in\mathcal{F}_K}$, with $\mathcal{D}_{K,\sigma}$ orthogonal to σ with respect to the scalar product induced by Λ_K^{-1}, such that

- For any $K \in \mathcal{T}$, $\bigcap_{\sigma\in\mathcal{F}_K} \mathcal{D}_{K,\sigma} = \boldsymbol{x}_K$
- For any $\sigma \in \mathcal{F}_{int}$ with $\mathcal{T}_\sigma = \{K, L\}$, $\mathcal{D}_{K,\sigma} \cap \sigma = \mathcal{D}_{L,\sigma} \cap \sigma \neq \emptyset$
- For any $\sigma \in \mathcal{F}_{ext}$ with $\mathcal{T}_\sigma = \{K\}$, $\mathcal{D}_{K,\sigma} \cap \sigma \neq \emptyset$.

TPFA Finite volume scheme for porous media flow For simplicity, we assume that the boundary conditions are homogeneous Neumann boundary conditions everywhere, i.e. no flow can leave the computational domain. In this case, the discrete mass balance equations of each component $0 \leq i \leq N_{comp} - 1$ become, for any $K \in \mathcal{T}$ and any $0 \leq n \leq N_T - 1$:

$$Q_{i,K}^{n+1} = |K| \frac{m_{i,K}^{n+1} - m_{i,K}^n}{\Delta t^n} + \sum_{\sigma\in\mathcal{F}_K\cap\mathcal{F}_{int}} \sum_{\alpha=0}^{N_{ph}-1} |\sigma| \rho_{\alpha,\sigma}^{n+1} X_{i,\sigma}^{\alpha,eq,n,n+1} V_{\alpha,K,\sigma}^{n+1} \quad (29)$$

where we have denoted:

$$V_{\alpha,K,\sigma}^{n+1} = \frac{\lambda_{K,\sigma}\lambda_{L,\sigma}}{\lambda_{K,\sigma} d_{L,\sigma} + \lambda_{L,\sigma} d_{K,\sigma}} \frac{k r_{\alpha,\sigma}^{n+1}}{\mu_{\alpha,\sigma}^{n+1}} \Delta_{P,K,\sigma}^{n+1},$$

with

$$\lambda_{K,\sigma} = \Lambda_K \boldsymbol{n}_{K,\sigma} \cdot \boldsymbol{n}_{K,\sigma},$$

and

$$\Delta_{P,K,\sigma}^{n+1} = P_{\alpha,K}^{n+1} - P_{\alpha,L}^{n+1} + \rho_{\alpha,KL}^{n+1} g(z_K - z_L),$$

and also

$$\rho_{\alpha,KL}^{n+1} = \frac{1}{2}(\rho_{\alpha,K}^{n+1} + \rho_{\alpha,L}^{n+1}).$$

The upwind relative permeabilities are given by:

$$kr_{\alpha,\sigma}^{n+1} = \begin{vmatrix} kr_\alpha(S_{\alpha,K}^{n+1}) & \text{if } \Delta_{P,K,\sigma}^{n+1} \geq 0 \\ kr_\alpha(S_{\alpha,L}^{n+1}) & \text{if } \Delta_{P,K,\sigma}^{n+1} < 0. \end{vmatrix}$$

In the same way, the upwind mass fractions are defined by:

$$X_{i,\sigma}^{\alpha,eq,n,n+1} = \begin{vmatrix} X_i^{eq,\alpha}(\mathcal{I}_{th,K}^n, \mathcal{I}_{th,K}^{n+1}) & \text{if } \Delta_{P,K,\sigma}^{n+1} \geq 0 \\ X_i^{eq,\alpha}(\mathcal{I}_{th,L}^n, \mathcal{I}_{th,L}^{n+1}) & \text{if } \Delta_{P,K,\sigma}^{n+1} < 0, \end{vmatrix}$$

as are $\rho_{\alpha,\sigma}^{n+1}$ and $\mu_{\alpha,\sigma}^{n+1}$. Finally $Q_{i,K}^{n+1}$ is given by the usual Peaceman's well source term. The discretization of the remaining equations is immediate. We have, for all $K \in \mathcal{T}$

$$\rho_{\alpha,K}^{n+1} \Phi_K S_{\alpha,K}^{n+1} = m_\alpha^{eq}(\mathcal{I}_{th,K}^n, \mathcal{I}_{th,K}^{n+1}), \tag{30}$$

$$\sum_{\alpha=0}^{N_{ph}-1} S_{\alpha,K}^{n+1} = 1, \tag{31}$$

$$P_{\alpha,K}^{n+1} = P_{ref,K}^{n+1} + Pc_{\alpha,K}^{n+1} \quad \forall\, 0 \leq \alpha \leq N_{ph} - 1. \tag{32}$$

References

1. Aziz, K., Settari, A.: Petroleum Reservoir Simulation. Elsevier, London (1979)
2. Christie, M., Blunt, M.: Tenth spe comparative solution project: a comparison of upscaling techniques. In: SPE Reservoir Simulation Symposium, 11–14 February, Houston (2001)
3. Coats, K.H.: An equation of state compositional model. SPE J. **20**(5), 363–376 (1980)
4. Collins, D., Nghiem, L., Li, Y.-K., Grabonstotter, J.E.: An efficient approach to adaptive implicit compositional simulation with an equation of state. SPE Reserv. Eng. **7**(2), 259–264 (1992)
5. Eymard, R., Gallouët, T., Herbin, R.: Finite volume methods. In: Ciarlet, P.G., Lions, J.-L. (eds.) Techniques of Scientific Computing, Part III, Handbook of Numerical Analysis, pp. 713–1020. North-Holland, Amsterdam (2000)
6. de Hemptinne, J.C., Ledanois, J.M., Mougin, P., Barreau, A.: Select thermodynamic models for process simulation—a practical guide using a three steps methodology. Edition Technip (2012)
7. IPCC: Special report on carbon dioxide capture and storage. Technical Report (2005)
8. Kala, K., Voskov, D.: Element balance formulation in reactive compositional flow and transport with parameterization technique. Comput. Geosci. **24**, 609–624 (2020)
9. Khait, M., Voskov, D.: Operator-based linearization for general purpose reservoir simulation. J. Pet. Sci. Eng. **157**, 990–998 (2017)
10. Lauser, A., Hager, C., Helmig, R., Wohlmuth, B.: A new approach for phase transitions in miscible multi-phase flow in porous media. Adv. Water Resour. **34**, 957–966 (2011)
11. Mehra, R.K., Heidmann, R.A., Aziz, K.: An accelerated successive substitution algorithm. Can. J. Chem. Eng. **61**, 590–596 (1983)
12. Jenny P., Tchelepi, H.A., Lee, S.H.: Unconditionally convergent nonlinear solver for hyperbolic conservation laws with s-shaped flux functions. J. Comput. Phys. **228**, 7497–7512 (2009)
13. Peng, D., Robinson, D.: A new two constant equation of state. Ind. Eng. Chem. Fundam. **15**(1), 59–64 (1976)
14. Redlich, O., Kwong, J.: On the thermodynamics of solutions. V. An equation of state. Fugacities of gaseous solutions. Chem. Rev. **44**(1), 233–244 (1949)
15. Soave, G.: Equilibrium constants from a modified Redlich-Kwong equation of state. Chem. Eng. Sci. **27**(6), 1197–1203 (1972)
16. Søreide, I., Whitson, C.: Peng-Robinson predictions for hydrocarbons, CO_2, N_2 and H_2S with pure water and NaCl brine. Fluid Phase Equilib. **77**, 217–240 (1992)
17. Vidal, J.: Thermodynamique, Application au génie chimique et à l'industrie pétrolière. Edition Technip (1997)
18. Voskov, D.: Operator-based linearization approach for modeling of multiphase multi-component flow in porous media. J. Comput. Phys. **337**, 275–288 (2017)
19. Voskov, D.V., Tchelepi, H.A.: Comparison of nonlinear formulations for two-phase multi-component eos based simulation. J. Pet. Sci. Eng. **82–83**, 101–111 (2012)
20. Wang, X., Tchelepi, H.A.: Trust-region based solver for nonlinear transport in heterogeneous porous media. J. Comput. Phys. **253**, 114–137 (2013)
21. Young, L.C., Stephenson, R.E.: A Generalized Compositional Approach for Reservoir Simulation, pp. 727–742. Society of Petroleum Engineers (1983)

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.