Advances in Geo-Energy Research

Vol. 5, No. 3, p. 239-251, 2021

Original article

An exploratory multi-scale framework to reservoir digital twin

Tao Zhang, Shuyu Sun*

Computational Transport Phenomena Laboratory (CTPL), Physical Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

Keywords:
Digital twin
reservoir simulation
multi-scale framework
oil-water separation

Cited as:
Zhang, T., Sun, S. An exploratory multi-scale framework to reservoir digital twin. Advances in Geo-Energy Research, 2021, 5(3): 239-251, doi: 10.46690/ager.2021.03.02

Abstract:
In order to make full use of the information provided in the physical reservoirs, including the production history and environmental conditions, the whole life cycle of reservoir discovery and recovery should be considered when mapping in the virtual space. A new concept of reservoir digital twin and the exploratory multi-scale framework is proposed in this paper, covering a wide range of engineering processes related with the reservoirs, including the drainage, sorption and phase change in the reservoirs, as well as extended processes like injection, transportation and on-field processing. The mathematical tool package for constructing the numerical description in the digital space for various engineering processes in the physical space is equipped with certain advanced models and algorithms developed by ourselves. For a macroscopic flow problem, we can model it either in the Navier-Stokes scheme, suitable for the injection, transportation and oil-water separation processes, or in the Darcy scheme, suitable for the drainage and sorption processes. Lattice Boltzmann method can also be developed as a special discretization of the Navier-Stokes scheme, which is easy to be coupled with multiple distributions, for example, temperature field, and a rigorous Chapman-Enskog expansion is performed to show the equivalence between the lattice Bhatnagar-Gross-Krook formulation and the corresponding Navier-Stokes equations and other macroscopic models. Based on the mathematical toolpackage, for various practical applications in petroleum engineering related with reservoirs, we can always find the suitable numerical tools to construct a digital twin to simulate the operations, design the facilities and optimize the processes.

1. Introduction

Reservoir simulation has been recognized as an efficient approach to study the reserve mechanisms and to enhance the resource discovery/recovery using a number of models and algorithms, and the significant superiority on the flexibility and efficiency compared with experimental studies has made numerical studies more popular and accepted in the current energy industry. The limited space and time scale challenging experimental studies can be resolved well in well-designed numerical schemes, and the restricted laboratory and field experimental conditions are replaced by the performance of computational devices. Moreover, the reliability of numerical methods should be verified by comparing with the experimental results, while the value of the key parameters involved in the mathematical formulations should be determined by experimental analysis. The improved models and algorithms after continual optimization on the numerical structure and parameter selection can only be considered as "realistic" if meeting well with experimental data, and then trusted to describe the complex phenomena in a reasonable scale with practical meanings. Digital twin (DT), a comprehensive and integrated concept proposed in this century, has been extensively recognized as an effective cutting-edge approach in this intelligent and digitizing era. Numerous publications have been reported in the establishment and application of digital twins in various fields, including complete-cycle management, automatic production and predictive maintenance (Boschert and Rosen, 2016; Kritzinger et al., 2018). Multi-discipline, multi-physics and multi-scale numerical methods have been integrated in digital twins to establish a simulated system and process that reflects the corresponding realistic system and process in the physical world, using the data obtained from various engineering sensors and the model constructed based on theoretical analysis (Shao and Helu, 2020). In order to make full use of the information provided in the physical side,

Yandy Scientific Press

*Corresponding author.
E-mail address: tao.zhang.1@kaust.edu.sa (T. Zhang); Shuyu.Sun@kaust.edu.sa (S. Sun).

Received May 19, 2021; revised May 31, 2021; accepted June 1, 2021; available online June 4, 2021.

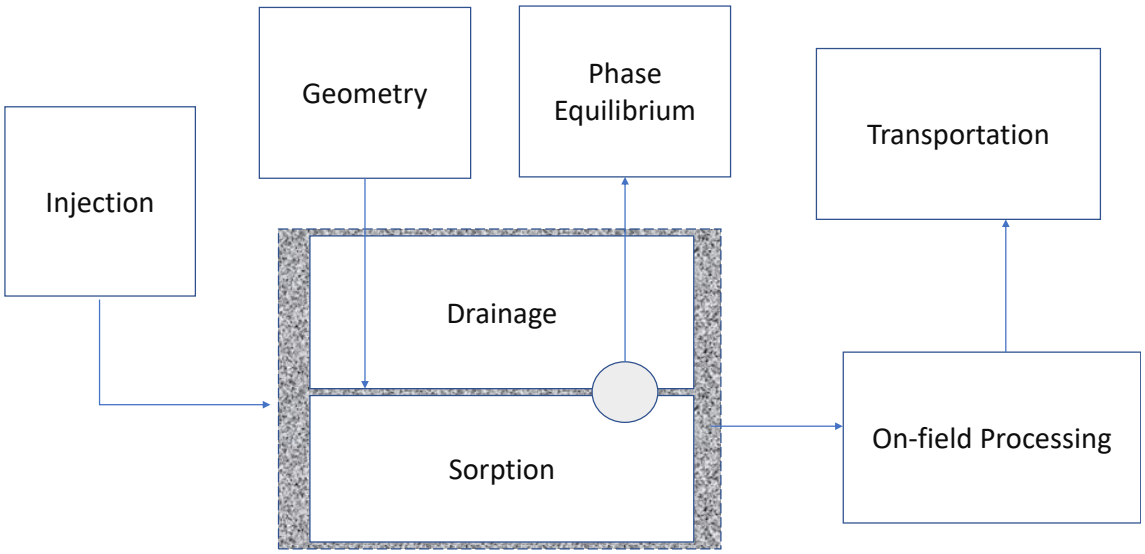


Fig. 1. A schematic diagram of the digital twin designed for reservoirs.

including the production history and environmental conditions, the whole life cycle should be considered when mapping in the virtual space. It should also be noted that the concept of DT was first proposed to benefit intelligent manufacturing using the most advanced information technologies emerged in the intelligent era (Tao et al., 2018), and this concept has been extended to a wider range involving designing, manufacturing, marketing and whole-life service thanks to the sufficient data supply in the current big data era (Qi and Tao, 2018). The data provided from the physical world and the information provided from the digital space can be transferred and exchanged between the two sides to feed the simulation and support the practical manufacturing. Sun and Zhang (2020) pointed out an another advantage of digital twins that the isolated and fragmented data obtained from the physical reality can be organized thoroughly in the digital twin to exhaust the underneath chemical or physical correlations among the various data. Such a high-level consistency between intelligent manufacturing and fundamental natural theories can enhance the industrial upgrading in an efficient and sustainable way (Zhuang et al., 2021). Thus, the intelligence and efficiency of industrial manufacturing are constantly improved to achieve intelligent production and management (Zhou et al., 2020; He and Bai, 2021).

The concept of digital twin is believed to be easily extended to reservoir studies in which data of the reservoir geometry, fluid properties, environmental and operation conditions can all feed into the model in virtual space, and we already have a mature understanding of the underneath physical and chemical rules. By simulating the mathematically-described assets in virtual space, the petroleum enterprises can gain more operating data about routine and abnormal situations (Mayani et al., 2020), analyze how configuration differences and operating choices translate to key performance indicator in the oil and gas production (Khasanov and Krasnov, 2019), reduce lag between project conceptualization and implementation in reservoir exploitation (Thoresen et al., 2019; Wanasinghe et al., 2020), run trials of loss events to learn how the systems might respond to the real conditions in the complex reservoir environment and identify the most viable strategies for real-life pilot tests (Rao, 2020). However, as a newly-popular approach with a short history, there has not been a widely-accepted definition of reservoir digital twin, as well as a common-recognized range of processes needing to be included in the digital platform. In this dissertation, our main focus is to construct a digital twin for reservoirs, constituting of the modeling and simulation of certain key processes including injection, drainage, sorption, on-field processing and transportation. As shown in Fig. 1, geometry of the reservoirs is also needed to define the domain details and boundary conditions of the simulation, and phase equ ilibrium studies are also needed for a physically-consistent description of multi-component multi-phase fluid flow.

This paper starts with the generation of mathematical tools, i.e., the multi-scale, multi-domain, multi-physics and multi-numerics numerical modeling and simulation of multi-component and multi-phase fluid flow, regarding three schemes in Section 2. In Navier-Stokes scheme, the unconditionally stable numerical scheme solving incompressible Navier-Stokes equations is constructed using finite different methods on cell-centered staggered grids with detailed discretization on both space and time. We will go through the unconditional stability estimates of the three main discretization schemes solving incompressible Navier-Stokes equations with both time and space discretizations in this section and show the effect of mathematical stability on simulation accuracy and reliability. In LBM scheme, A compact LBM scheme will be generated in this chapter, using Boussinesq approximation method (Spiegel and Veronis, 1960) to couple the velocity and temperature distribution function evolution formulated by two different lattice Bhatnagar-Gross-Krook (LBGK) equations discretized in various lattice systems. In the Darcy scheme, we will focus on the development of improved implicit pressure explicit saturation (IMPES) schemes solving immiscible and

incompressible multi-phase fluid flow in subsurface reservoirs considering gravity. Based on that, example digital twins are established in Section 3.

2. Mathematical tools in the digital space

The main numerical framework in three schemes, namely Navier-Stokes schemes, LBM schemes and Darcy schemes, will be formulated in this section. Representative models and algorithms will be presented with the key procedures, and we will state the conservation analysis in each scheme. Navier-Stokes schemes are the most common computational fluid dynamic methods used in engineering computations, and they are more preferred in a single pore or in pipelines/facilities. LBM schemes can be viewed as a special discretization of Navier-Stokes equations, and the easy implementation and incorporation of a large number of different models in the distribution function makes it more popular in engineering computations. Darcy schemes are especially suitable for flow and transport studies in the representative element volume scale and field scale porous media.

2.1 Navier-Stokes schemes

The numerical scheme investigated in this section is mainly focused on the time-dependent Navier-Stokes equations for incompressible single-phase flow coupled with mass conservation and boundary/initial conditions, which can be formulated as:

$$\rho\left(\frac{\partial u}{\partial t}+(u\cdot\nabla)u\right)=\eta\Delta u-\nabla p+\rho g \quad \text{in } \Omega\times(0,T] \tag{1}$$

$$\nabla\cdot u=0 \quad \text{in } \Omega\times(0,T] \tag{2}$$

$$u|_{\partial\Omega}=0 \tag{3}$$

$$u|_{t=0}=u_0 \tag{4}$$

where ρ and η represent fluid density and the kinematic viscosity coefficient respectively, g denotes the gravity acceleration, u and p denote the flow velocity and pressure respectively, and $\int_\Omega p\,dx=0$. For incompressible single phase flow, ρ is constant and we could also assume a constant η in the computation, where the domain is designed as $\Omega\subset\mathbb{R}^2$ and the boundary condition is selected as the common-used homogeneous Dirichlet boundary condition. The past decades have witnessed numerous efforts paid in the numerical solution to the previous scheme, which can be originated from the projection methods proposed in 1960s (Chorin, 1968), which decouples the pressure and velocity computations. Further developments on projection methods had been critically reviewed and discussed in (Guermond et al., 2006), and we will go through the details of two representative projection methods, i.e., pressure-correction method and pressure-stabilization method respectively in this chapter. In this section, we will start with a compact and efficient numerical implementation for solving Stokes equations using a matrix-vector method called the Shift-Matrix method, and then extend to the energy stability analysis of fully discrete numerical schemes for incompressible Navier-Stokes equations. The three fully discretized numerical schemes using finite difference method on staggered grids mentioned above, i.e., the linear implicit scheme for time discretization, the pressure-correction injection scheme and pressure-stabilization injection scheme. The energy stability estimates are carried out for each scheme, and the upwind technique is applied in discretizing the convection term, which is proved to be a critical role in establishing unconditionally stable numerical schemes solving Navier-Stokes equations.

The difference operations have been vectorized in our scheme, thereby eliminating the conventional iterations needing for convergence. This is particularly important when coding in computer languages that require interpretations, e.g., R, Python and Matlab. The x-direction velocity term is defined as:

$$\begin{aligned}&\begin{bmatrix} u^{x,h}_{1,\frac{1}{2}} & u^{x,h}_{\frac{1}{2},\frac{3}{2}} & \cdots & u^{x,h}_{\frac{1}{2},n-\frac{1}{2}} \\ u^{x,h}_{2,\frac{1}{2}} & u^{x,h}_{2,\frac{3}{2}} & \cdots & u^{x,h}_{2,n-\frac{1}{2}} \\ \vdots & & & \vdots \\ u^{x,h}_{m,\frac{1}{2}} & u^{x,h}_{m,\frac{3}{2}} & \cdots & u^{x,h}_{m,n-\frac{1}{2}} \end{bmatrix} \\ &=\left[u^{x,h}_{i,j-\frac{1}{2}}\left(\begin{bmatrix} i=1,2\cdots m \\ j=1,2\cdots n \end{bmatrix}\right)\right] \\ &=\left[u^{x,h}_{i,j-\frac{1}{2}}\right] \\ &=\left[u^{x,h}\right]\end{aligned} \tag{5}$$

with the similar definition of the y-direction velocity term and pressure, the Navier-Stokes equation can be written into the matrix variation form as:

$$\begin{bmatrix} Dsq & 0 & -D^T_{cx} \\ 0 & Dsq & -D^T_{cy} \\ -Dcx & -Dcy & 0 \end{bmatrix}\begin{bmatrix} u_x \\ u_y \\ p \end{bmatrix}=\begin{bmatrix} \rho g_x \\ \rho g_y \\ 0 \end{bmatrix} \tag{6}$$

where the operators D are defined in details in Zhang et al. (2015). The proposed Shift-Matrix method in this section can significantly reduce the length of the implementation code and easily extended to other numerical schemes constituting of more kinds of partial difference equations, such as Cahn-Hilliard equation common used in phase field models.

The linear implicit scheme discretizing Navier-Stokes equation reads:

$$\rho\left(\frac{u^{k+1}-u^k}{\triangle t}+\left(u^k\cdot\nabla\right)u^{k+1}\right)=\eta\Delta u^{k+1}-\nabla p^{k+1}+\rho g \tag{7}$$

$$\nabla\cdot u^{k+1}=0 \tag{8}$$

where the energy stability can be stated as:

$$E^{k+1}-E^k\le\frac{C\triangle t}{\eta}\|\rho g\|^2_{L^2(\Omega)} \tag{9}$$

The upwind scheme is applied to ensure the unconditional energy stability, which can be formulated as (for the x-direction momentum equation as an example):

$$\begin{aligned}
&\rho\left(\frac{u_{i,j+1/2}^{k+1}-u_{i,j+1/2}^{k}}{\triangle t}\right)+\frac{\rho}{\triangle x}(X+Y)+\frac{\rho}{\triangle y}(Z+W)\\
&=\frac{\eta}{\triangle x}\left(\triangle_x^{+}u_{i,j+1/2}^{k+1}-\triangle_x^{-}u_{i,j+1/2}^{k+1}\right)\\
&+\frac{\eta}{\triangle y}\left(\triangle_y^{+}u_{i,j+1/2}^{k+1}-\triangle_y^{-}u_{i,j+1/2}^{k+1}\right)\\
&-\triangle_x^{-}p_{i+1/2,j+1/2}^{k+1}+\rho g_{i,j+1/2}^{u,k+1}
\end{aligned}\tag{10}$$

$$\begin{cases}
X=\left(\beta_u^k\cdot n\right)_{i+1/2,j+1/2}\cdot u_{i+1/2,j+1/2}^{*,k+1}\\
Y=\left(\beta_u^k\cdot n\right)_{i-1/2,j+1/2}\cdot u_{i-1/2,j+1/2}^{*,k+1}\\
Z=\left(\beta_u^k\cdot n\right)_{i,j+1}\cdot u_{i,j+1}^{*,k+1}\\
W=\left(\beta_u^k\cdot n\right)_{i,j}\cdot u_{i,j}^{*,k+1}
\end{cases}$$

where the upwind velocity is defined as:

$$u_{i+1/2,j+1/2}^{*,k+1}=\begin{cases}
u_{i,j+1/2}^{k+1} & \text{if } \beta_u^k\cdot n=\overline{u}_{i+1/2,j+1/2}^{k}\geqslant 0\\
u_{i+1,j+1/2}^{k+1} & \text{if } \beta_u^k\cdot n=\overline{u}_{i+1/2,j+1/2}^{k}<0
\end{cases}\tag{11}$$

It is easy to state that $E^{(k+1)}\leqslant E^k$ if the external force g is zero, for example, when gravity is neglected. C is a constant depending on the domain discretization and shape regularity only.

The first-order pressure-correction scheme can be described as: to find $(u^{k+1},\ p^{k+1})$ using the following formulations:

$$\begin{cases}
\rho\dfrac{u^{k+1}-\widetilde{u}^{k+1}}{\triangle t}+\nabla\left(p^{k+1}-p^{k}\right)=0\\
\nabla\cdot u^{k+1}=0\\
u^{k+1}\cdot n=0\quad\text{on the boundary}
\end{cases}\tag{12}$$

where the auxiliary velocity $\widetilde{u}^{k+1}$ is calculated by

$$\begin{cases}
\rho\left(\dfrac{\widetilde{u}^{k+1}-u^{k}}{\triangle t}+\left(u^{k}\cdot\nabla\right)\widetilde{u}^{k+1}\right)=\eta\Delta\widetilde{u}^{k+1}-\nabla p^{k}+\rho g^{k+1}\\
\widetilde{u}^{k+1}=0\quad\text{on the boundary}
\end{cases}\tag{13}$$

The energy stability estimate can be stated as:

$$\begin{aligned}
E^{n+1}\leq E^{0}&+\frac{(\triangle t)^2}{2\rho}\left(\parallel\triangle_x^{-}p_h^{0}\parallel_{\mathscr{T}_h}^2+\parallel\triangle_y^{-}p_h^{0}\parallel_{\mathscr{T}_h}^2\right)\\
&+\frac{C\rho^2}{\eta}\sum_{k=0}^{n}\triangle t\left(\parallel g^{u,k+1}\parallel_{\mathscr{T}_h^u}^2+\parallel g^{v,k+1}\parallel_{\mathscr{T}_h^v}^2\right)
\end{aligned}\tag{14}$$

The incremental second-order pressure-correction scheme solving Navier-Stokes equations can be described as: to find $(u^{k+1},\ p^{k+1})$ using the following formulations:

$$\begin{cases}
\rho\dfrac{3u^{k+1}-3\widetilde{u}^{k+1}}{\triangle t}+\nabla\left(p^{k+1}-p^{k}\right)=0\\
\nabla\cdot u^{k+1}=0\\
u^{k+1}\cdot n=0\quad\text{on the boundary}
\end{cases}\tag{15}$$

where the auxiliary velocity $\widetilde{u}^{k+1}$ is calculated by

$$\begin{cases}
\rho\left(\frac{3\widetilde{u}^{k+1}-4u^{k}+u^{k-1}}{\triangle t}+\left(\left(2u^{k}-u^{k-1}\right)\cdot\nabla\right)\widetilde{u}^{k+1}\right)\\
\quad=\eta\Delta\widetilde{u}^{k+1}-\nabla p^{k}+\rho g^{k+1}\\
\widetilde{u}^{k+1}=0\quad\text{on the boundary}
\end{cases}\tag{16}$$

The first-order pressure-stabilization scheme can be formulated as: to find p^{k+1} using the following formulations:

$$\begin{cases}
\Delta p^{k+1}=\dfrac{\rho}{\triangle t}\nabla\cdot u^{k+1}\\
\dfrac{\partial p^{k+1}}{\partial n}\mid_{\partial\Omega}=0\quad\text{on the boundary}
\end{cases}\tag{17}$$

where the velocity is calculated by

$$\begin{cases}
\rho\frac{u^{k+1}-u^{k}}{\triangle t}+\rho\left(u^{k}\cdot\nabla\right)u^{k+1}+\frac{\rho}{2}\left(\nabla\cdot u^{k}\right)u^{k+1}\\
\quad=\eta\Delta u^{k+1}-\nabla p^{k}+\rho g^{k+1}\\
u^{k+1}=0\quad\text{on the boundary}
\end{cases}\tag{18}$$

The energy stability estimate can be stated as:

$$E^{n+1}\leq E^{0}+\frac{C\rho^2}{\eta}\sum_{k=0}^{n}\triangle t\left(\parallel g^{u,k+1}\parallel_{\mathscr{T}_h^u}^2+\parallel g^{v,k+1}\parallel_{\mathscr{T}_h^v}^2\right)\tag{19}$$

It should also be noted that in fully discrete pressure-stabilization schemes (with both time and space discretization) the averaged velocity $\overline{u}_{i+1/2,j+1/2}^{k+1}$, $\overline{u}_{i,j+1}^{k+1}$, $\widehat{v}_{i,j}^{k}$, $\widehat{v}_{i-1/2,j+1/2}^{k+1}$ can be used instead of upwind schemes as presented in Eq. (11).

2.2 Lattice Boltzmann schemes

Currently, the lattice Boltzmann method is often constructed based on the BGK approximation of Boltzmann equation as:

$$\partial_t f+\xi\cdot\nabla_x f=\Omega(f)=-\frac{1}{\tau_c}\left[f(x,\xi,t)-f^{eq}(x,\xi,t)\right]\tag{20}$$

where $f(x,\xi,t)$ denotes the particle distribution function with physical space x and particle velocity ξ. Such approximation can be viewed as a special discretization which deserves a merit of easy implementation as it is a linear scheme. The relaxation time, τ_c, is introduced to model the time interval between each collision, and the collision frequency can be calculated easily as $1/\tau_c$. The correlation between macroscopic properties including velocity u, fluid density ρ and the internal energy e and the distribution function f can be formulated as:

$$\rho=\int f d\xi,\rho u=\int\xi f d\xi,\rho e=\int(\xi-u)^2 f d\xi\tag{21}$$

where the total energy E can be calculated as $E=\rho e+\frac{1}{2}\rho u^2$. The distribution function at equilibrium in the standard LBGK

equation as the most popular Lattice Boltzmann scheme in current engineering computations can be formulated as:

$$f^{eq}=\frac{\rho}{(2\pi RT)^{D/2}}\exp\left[-\frac{(\xi-u)^2}{2RT}\right] \tag{22}$$

where denotes the gas constant $R=k_B/m$, and k_B denotes the Boltzmann constant. Eq. (22) can be further transformed into the following discretized equation using Taylor expansion regarding macroscopic flow velocity u as:

$$f^{eq}=\frac{\rho}{(2\pi RT)^{D/2}}\exp\left(-\frac{\xi^2}{2RT}\right)\left[1+\frac{\xi\cdot u}{RT}+\frac{(\xi\cdot u)^2}{2(RT)^2}-\frac{u^2}{2RT}\right]+O\left(u^3\right) \tag{23}$$

The DnQb model is common used to implement the discretized LBGK approximation (Wanasinghe et al., 2020), $f_i(x+c_i\delta t,t+\delta t)-f_i(x,t)=-1/\tau\left[f_i(x,t)-f_i^{eq}(x,t)\right]$, as:

$$f_i^{eq}=\rho\omega_i\left[1+\frac{ce_i\cdot u}{c_s^2}+\frac{uu:\left(c^2e_ie_i-c_s^2I\right)^2}{2c_s^4}\right] \tag{24}$$

where ω_i denotes the weight coefficient in each direction i, $c_s=\sqrt{RT}$ denotes the sound speed, $c_i=ce_i$ denotes the particle velocity. The parameters in D2Q9 model have derived in Qian et al. (1992) as:

$$\begin{cases}\dfrac{c_s^2}{c^2}=\dfrac{1}{3}\\ \omega_0=\dfrac{4}{9}\\ \omega_1=\dfrac{1}{9}=\omega_2=\omega_3=\omega_4\\ \omega_5=\dfrac{1}{36}=\omega_6=\omega_7=\omega_8.\end{cases} \tag{25}$$

In the Lattice Boltzmann method's point of view on fluid flow simulation, the macroscopic fluid dynamics can be seen as the collective representation of microscopic particles motion behaviors within the system, so that the LBGK approximation is a special discretization of Navier-Stokes equations. The Chapman-Enskog expansion (CE expansion), a common used multi-scale expansion analysis method is applied to derive the macroscopic Navier-Stokes equations from the Lattice Boltzmann equation. In this section, a CE expansion is carried out for a D2Q9 model. We first expand the distribution function as:

$$f_i=f_i^{(0)}+\varepsilon f_i^{(1)}+\varepsilon^2 f_i^{(2)}+\cdots \tag{26}$$

while the space and time can be scaled into

$$\mathrm{x}=\varepsilon^{-1}x,\ t_1=\varepsilon t,\ t_2=\varepsilon^2 t \tag{27}$$

In the above scheme, t_1 and t_2 denote the convective and diffusive time scale respectively. Thus, the differential operators can be formulated into

$$\frac{\partial}{\partial t}=\varepsilon\frac{\partial}{\partial t_1}+\varepsilon^2\frac{\partial}{\partial t_2},\ \nabla=\varepsilon\nabla_1 \tag{28}$$

which can further transform the 3rd-order lattice tensor $\sum_i c_ic_ic_if_i^{eq}=c_s^2\rho\triangle\cdot u$ into

$$\begin{aligned}&\left(\varepsilon\partial t_1+\varepsilon^2\partial t_2+\varepsilon c_i\cdot\nabla_1\right)\left(f_i^{(0)}+\varepsilon f_i^{(1)}+\varepsilon^2 f_i^{(2)}+\cdots\right)\\&+\frac{\delta t}{2}\left(\varepsilon\partial t_1+\varepsilon^2\partial t_2+\varepsilon c_i\cdot\nabla_1\right)^2\left(f_i^{(0)}+\varepsilon f_i^{(1)}+\varepsilon^2 f_i^{(2)}+\cdots\right)\\&=-\frac{1}{\tau\delta t}\left(f_i^{(0)}-f_i^{(eq)}+\varepsilon f_i^{(1)}+\varepsilon^2 f_i^{(2)}+\cdots\right)\end{aligned} \tag{29}$$

In each order of ε, we can obtain that:

$$\mathrm{O}\left(\varepsilon^0\right):\ f_i^{(0)}=f_i^{(eq)} \tag{30}$$

$$\mathrm{O}\left(\varepsilon^1\right):\ \left(\partial t_1+c_i\cdot\nabla_1\right)f_i^{(0)}=-\frac{1}{\tau\delta t}f_i^{(1)} \tag{31}$$

$$\mathrm{O}\left(\varepsilon^2\right):\ \partial t_2 f_i^{(eq)}+\left(\partial t_1+c_i\cdot\nabla_1\right)\left(1-\frac{1}{2\tau}\right)f_i^{(1)}=-\frac{1}{\tau\delta t}f_i^{(2)} \tag{32}$$

Taking summation over i, we can get

$$\partial_t\rho+\nabla\cdot(\rho u)=0 \tag{33}$$

which recovers the mass conservation equation in the macroscopic scale. Similarly, we can obtain in the first and second order of ε as:

$$\mathrm{O}\left(\varepsilon^1\right):\ \left(\partial t_1+c_i\cdot\nabla_1\right)f_i^{(eq)}=-\frac{1}{\tau\delta t}f_i^{(1)} \tag{34}$$

$$\mathrm{O}\left(\varepsilon^2\right):\ \partial t_2\rho u=\nabla_1\cdot\left[\nu\rho\left(\nabla_1 u+(\nabla_1 u)^T\right)\right] \tag{35}$$

which recovers the Navier-Stokes momentum equation using viscosity definition $\mathrm{v}=c_s^2\left(\tau-\frac{1}{2}\right)\delta t$ as:

$$\partial_t(\rho u)+\nabla\cdot(\rho uu)=-\nabla p+\nabla\cdot\left[\nu\rho\left(\nabla_1 u+(\nabla_1 u)^T\right)\right] \tag{36}$$

In the original LBM schemes, momentum and mass conservation are taken into account in evaluating the scheme consistency (Abe, 1997). However, many other variables are needing modeled in plenty of applications, for example, the temperature field in thermal fluid flows. LBGK models for thermal fluid flows have been developed by many previous papers, but in general, these thermal LBGK models fall into two categories: the multi-speed approach and the multi-distribution function (MDF) approach. An improved coupled LBGK scheme was proposed in Zhang and Sun (2018) to simulate thermal fluid flows using the idea of MDF model. The Boussinesq approximation approach is introduced to couple the additional LBGK equation generated to model temperature with the original distribution function modeling velocity distribution. The LBGK equation discretizing the general heat transfer equation $\partial T/\partial t+\nabla\cdot(uT)=\mathfrak{D}\nabla^2 T$ can be formulated as:

$$T_i(x+c_i\delta t,t+\delta t)-T_i(x,t)=-\frac{1}{\tau'}\left(T_i(x,t)-T_i^{eq}(x,t)\right) \tag{37}$$

where the relaxation time τ' could either be the same or different with that in the velocity LBGK equation. The temperature at equilibrium term in Eq. (37) can be calculated as:

$$T_i^{(eq)} = \frac{T}{4}\left(1+2\frac{e_i\cdot u}{c}\right) \tag{38}$$

The coupled LBGK approach is similar to the classical MDF model, but there remains a difference that the DnQb model for temperature and velocity distribution can be different. For example, D2Q5 model can be used for temperature distribution while D2Q9 model can be used for velocity distribution to improve the flexible computation efficiency as well as preserving reasonable consistency. In D2Q5 model, the macroscopic temperature can be recovered as:

$$T = \sum_{i=1}^{4} T_i \tag{39}$$

A similar Chapman-Enskog expansion can be carried out for the recovering of macroscopic heat transfer equation from the LBGK scheme, starting from the Taylor expansion of temperature distribution function T_i as:

$$T_i = T_i^{(0)} + \varepsilon T_i^{(1)} + \varepsilon^2 T_i^{(2)} + \cdots \tag{40}$$

where the 0-th order temperature is defined as $T_i^{(0)} = T_i^{(eq)}$. The lattice BGK formulation Eq. (37) could be similarly Taylor expanded into

$$D_i T_i + \frac{\Delta t}{2} D_i^2 T_i + O(\Delta t)^2 = -\frac{1}{\tau'\cdot t}\left(T_i - T_i^{(0)}\right) \tag{41}$$

where the partial derivative notations are defined as: $\partial/\partial t = \varepsilon\partial/\partial t_1 + \varepsilon^2\partial/\partial t_2, \nabla = \varepsilon\nabla_1, D_i = (\partial/\partial t + ce_i\cdot\nabla)$. Summing over i in the $t_1 = \varepsilon t$ and $t_2 = \varepsilon^2 t$ two scales, we can get

$$\frac{\partial T}{\partial t_1} + \nabla_1\cdot(uT) = 0 \tag{42}$$

$$\frac{\partial T}{\partial t_2} + \left(1-\frac{1}{2\tau'}\right)\nabla_1\cdot\Pi^{(1)} = 0 \tag{43}$$

where $\Pi^{(1)} = \sum_{i=1}^{4} ce_i T_i^{(1)}$. Then, the heat transfer equation could be recovered to the $O\left(\Delta t^2\right)$ order if we define the heat diffusivity $\mathfrak{D}$ as $\mathfrak{D} = \left(2\tau' - 1\right)/4\cdot\Delta x^2/\Delta t$.

2.3 Darcy schemes

The immiscible and incompressible two phase flow in subsurface porous media can be modeled based on the Dacry equation as:

$$\phi\frac{\partial S_\alpha}{\partial t} + \nabla\cdot\boldsymbol{u}_\alpha = F_\alpha \quad \text{in } \Omega, \alpha = w,n \tag{44}$$

$$\boldsymbol{u}_\alpha = -\frac{k_{r\alpha}}{\mu_\alpha}\boldsymbol{K}\left(\nabla p_\alpha + \rho_\alpha g\nabla z\right) \quad \text{in } \Omega, \alpha = w,n \tag{45}$$

with saturation conservation constraint $S_n + S_w = 1$. In the above scheme, Ω denotes the porous media, α denotes the phase index, w denotes the wetting phase, n denotes the non-wetting phase, ϕ denotes the porosity and $\boldsymbol{K}$ denotes the absolute permeability tensor. S_α, $\boldsymbol{u}_\alpha$, p_α, F_α denote the saturation of each phase, Darcy's velocity of each phase, pressure of each phase and the sink/source term of each phase respectively.

As mentioned in Chen et al. (2019), both the original IMPES scheme and the other improved IMPES schems are biased and only mass-conservative for one phase, thus the fully mass-conservative IMPES schemes are essentially required by petroleum industry to accurately describe the change of saturation field of both phases in the subsurface reservoirs with bound-preserving, unbiased and well-proposed properties. The physical conservation properties are directly relevant with the simulation reliability and practical meanings. In this chapter, we will construct two fully mass-conservative IMPES schemes with the total conservation equation obtained by summing the discretized conservation equation for each phase. Based on the energy stability estimates for Navier-Stokes equation in Section 2.1, the upwind technique is also applied in this chapter to construct the conservative scheme solving the velocity-pressure system formulated above discretized by the mixed finite element method. The constraints on phase saturation of each phase and the capillary pressure equation should be incorporated in the total system and an auxiliary velocity is introduced to simply the implementation and mathematical proof. It should be noted that although a better treatment on capillary pressure was developed in Kou and Sun (2010) by coupling the pressures of both phases, the computational cost is not reduced, and even increased compared with original IMPES schemes. Two fully mass-conservative IMPES schemes will be presented in this section solving the immiscible and incompressible two-phase flow in porous media. Using the standard definition and notations for Sobolev spaces, we can define $(\psi,\phi)_D = \int_D \psi\phi$ and $(\psi,\phi)_D = \int_D \psi\cdot\phi$. The lowest-order Raviart-Thomas mixed finite element approximation method is introduced in the spacial discretizations of the coupled flow model to simplify the implementation. RT_0 is defined as the lowest-order Raviart-Thomas mixed finite element space, which could be written as $RT_0 = [P_0]^d + \boldsymbol{x}P_0$ for simplicial mesh and P_0 is defined as the piecewise constant space.

The first fully mass-conservative IMPES scheme, also known as FC-IMPES-I, can be formulated as the following.

We first calculate $p_w^{h,n+1}$ and $\boldsymbol{w}_w^{h,n+1}$ by

$$B_t\left(\boldsymbol{w}_w^{h,n+1}, q; S_w^{h,n}\right) = (F_t, q) - B_n\left(\boldsymbol{w}_c^{h,n+1}, q; S_w^{h,n}\right),\ q\in Q_h \tag{46}$$

$$\begin{aligned}&\left(\boldsymbol{K}^{-1}\boldsymbol{w}_w^{h,n+1}, \boldsymbol{v}\right) - \left(p_w^{h,n+1}, \nabla\cdot\boldsymbol{v}\right) = \\ &-\int_{\Gamma_D} p_w^B\boldsymbol{v}\cdot\boldsymbol{n} - (\rho_w g\nabla z, \boldsymbol{v}),\ \boldsymbol{v}\in\boldsymbol{U}_h^0\end{aligned} \tag{47}$$

We can then update the saturation of both the wetting and non-wetting phase as:

$$\left(\phi\frac{S_w^{h,n+1} - S_w^{h,n}}{t_{n+1} - t_n}, q\right) + B_w\left(\boldsymbol{w}_w^{h,n+1}, q; S_w^{h,n}\right) = (F_w, q) \tag{48}$$

$$\left(\phi\frac{S_n^{h,n+1} - S_n^{h,n}}{t_{n+1} - t_n}, q\right) + B_n\left(\boldsymbol{w}_n^{h,n+1}, q; S_w^{h,n}\right) = (F_n, q) \tag{49}$$

It can be proved that the solution of Eq. (49) is equivalent to that of $S_w^{h,n+1}+S_n^{h,n+1}=1$, which validated the full conservation property of both the two phases.

The second fully-conservative IMPES scheme, also known as FC-IMPES-II, can be formulated as the following.

We first calculate $p_w^{h,n+1}$ and $\boldsymbol{\xi}_w^{h,n+1}$ by

$$\widetilde{B}_t\left(\boldsymbol{\xi}_w^{h,n+1},q;S_w^{h,n}\right)=(F_t,q)-\widetilde{B}_n\left(\boldsymbol{\xi}_c^{h,n+1},q;S_w^{h,n}\right),\ q\in Q_h \tag{50}$$

$$\begin{aligned}&\left((\lambda_t\boldsymbol{K})^{-1}\boldsymbol{\xi}_w^{h,n+1},\boldsymbol{v}\right)-\left(p_w^{h,n+1},\nabla\cdot\boldsymbol{v}\right)=\\&-\int_{\Gamma_D}p_w^B\boldsymbol{v}\cdot\boldsymbol{n}-(\rho_w g\nabla z,\boldsymbol{v}),\ \boldsymbol{v}\in\boldsymbol{U}_h^0\end{aligned} \tag{51}$$

$\boldsymbol{\xi}_\alpha$ is defined as $\boldsymbol{\xi}_\alpha=\lambda_t\boldsymbol{w}_\alpha$, where λ_t denotes the total mobility $\lambda_t=\lambda_w+\lambda_n$. The wetting and non-wetting phase saturation can be calculated as:

$$\left(\phi\frac{S_w^{h,n+1}-S_w^{h,n}}{t_{n+1}-t_n},q\right)+\widetilde{B}_w\left(\boldsymbol{\xi}_w^{h,n+1},q;S_w^{h,n}\right)=(F_w,q) \tag{52}$$

$$\left(\phi\frac{S_n^{h,n+1}-S_n^{h,n}}{t_{n+1}-t_n},q\right)+\widetilde{B}_n\left(\boldsymbol{\xi}_n^{h,n+1},q;S_w^{h,n}\right)=(F_n,q) \tag{53}$$

It can be easily found that $\left(\phi\frac{S_w^{h,n+1}-S_w^{h,n}}{t_{n+1}-t_n},q\right)+\left(\phi\frac{S_n^{h,n+1}-S_n^{h,n}}{t_{n+1}-t_n},q\right)=0$, which indicates the fully mass conservation as $S_w^{h,n+1}+S_n^{h,n+1}=1$.

3. Digital twin examples

Based on the mathematical tool package, for various practical applications in petroleum engineering related with reservoirs, we can always find the suitable numerical tools to construct a digital twin to simulate the operations, design the facilities and optimize the processes. It should be noted that for a specific application, the construction of digital twin may incorporate additional numerical tools in the digital space, for example, thermodynamic analysis for the phase equilibrium estimates in a single pore and phase field model for the oil-water separation in a separator.

3.1 Digital twin regarding phase behaviors in a single pore

Understanding the flow and transport behaviors at the pore-scale is of great importance in many fields such as hydrology, contaminant cleanup and petroleum engineering. In the pore-network modeling method, the pore bodies are connected by pore throats in the microscopic scale, and the flow and transport behaviors in the pore bodies and throats are determining the macroscopic fluid properties including Darcy's velocity as modeled in Section 2. Multi-phase multi-component flows are the key problems needing to be solved in the study of subsurface geological formation and fluid flows, which are essentially required in the understanding and description of complicated heat and mass transfer behaviors commonly seen in oil and gas reservoirs. A large number of chemical species have been detected in the reservoir fluids, which challenges the conventional computational multi-phase fluid dynamic simulation using empirical formulas. The number of phases existing in the fluid mixture, as well as the phase partitioning information of each component, play an important role in the multi-component multi-phase model and simulation to keep the thermodynamic consistency and physical meaningfulness. Flash calculation, the main approach to obtain these information, including overall density, chemical composition and the total phase numbers at equilibrium, has shown its inevitability in energy discovery and recovery, especially when the concept of enhanced oil recovery (EOR) is discussed. As a promising approach to bring more economical productions to recharge the social development from subsurface reservoirs, the main EOR techniques include carbon dioxide sequestration and steam-assisted gravity drainage, in which the underneath complex thermodynamic process controlling the increased oil productivity. Thus, information of the phase behaviors and surface properties including capillary pressure and surface tension is essentially needed as a basic knowledge of large scale reservoir simulation, while the thermodynamic consistency should always be preserved to ensure the reliability and stability of the feedback provided by our digital twin.

To understanding physical phenomena involving multiple phases, such as liquid droplets, gas bubbles and phase change and separation, it is necessary to model and simulate the interface between phases. Understanding and modeling the interface between phases have been approached by at least three main methodologies in different scales. Molecular Dynamics and Monte Carlo method are the main microscopic approaches, which are computationally more challenging as well as more accurate. In the second methodology, known as diffuse interface models, gradient theory, or phase field theory, the interface is described as a continuum three-dimensional entity separating the two bulk single-phase fluid regions. Another approach in macroscopic level is known as the sharp interface model, which utilizes a zero-thickness two-dimensional entity to model the interface, where the molar density experiences a jump across the interface. Typical sharp interface models, including the VOF and level-set method, cannot provide details within the interface itself, but they can be used to predict the shape and the dynamics of the interface once the interface tension is given. In the meantime, modeling approaches combining different scales have also been used, such as a multiscale simulation including MD and DI. Diffuse interface models have been extensively studied in the previous literatures, especially in recent years. In these works, a simple double-well potential is usually employed for the interface tension modeling. However, it is not possible to use quantitatively meaningful parameters of a simple double-well potential for simulation realistic hydrocarbon species in an oil-gas two-phase system. Here, a more realistic equation of state (e.g., Peng-Robinson equation of state), which is more accurate, is of a good choice. Meanwhile, such methods can also make the system consistent with the second laws of thermodynamics. Other methods include the dynamic van der Waals theory, which had been viewed as a combination of equation of state for thermodynamic processes and general numerical solution to Navier-Stokes equations for fluid flow.

Here, Lattice Boltzmann Method and other N-S equation solvers have already applied. The scalar auxiliary variable (SAV) approach is developed in this section, to construct a linear time-marching system solving multi-phase flow problems with a constant coefficient, which is an advantage compared with the previous invariant energy quadratization method to benefit the implementation and accelerate the convergence.

We first define the total Helmholtz free energy of the equation as:

$$F = \int_\Omega \left(\frac{C}{2} |\nabla n|^2 + f \right) dx \tag{54}$$

which can be further transformed into

$$F = \int_\Omega \frac{C}{2} |\nabla n|^2 \, dx + r^2 - C_0 \tag{55}$$

if we define $r(t) = \sqrt{E_p + C_0}$ as a variable changing with time, C_0 is a constant to ensure the positive definition of $E_p + C_0$ and the homogeneous term of the free energy density as $E_p = \int_\Omega f dx$. n is the molar density, which represents the particle number per unit volume for the single component as given by equation $n = N/V$. Then, the original problem (also called gradient flow in the art) for calculating the flow in the subsurface, which is described by equation:

$$n_t + c\Delta^2 n = \Delta \mu_0 \tag{56}$$

which can be written as:

$$\begin{aligned} n_t &= \Delta \mu \\ \mu &= -C\Delta n + \frac{r}{\sqrt{E_p + C_0}} \mu_0(n) \\ r_t &= \frac{1}{2\sqrt{E_p + C_0}} \int_\Omega \mu_0(n) n_t dx \end{aligned} \tag{57}$$

where μ is the chemical potential, $n_t = \partial n / \partial t$, $r_t = \partial r / \partial t$, and $\mu_0 = \partial f_0 / \partial n$. Finally, the single-component, multi-phase system can be rewritten as:

$$n_t + C\Delta^2 n - \frac{r}{\sqrt{E_p + C_0}} \Delta \mu_0(n) = 0 \tag{58}$$

$$r_t = \frac{1}{2\sqrt{E_p + C_0}} \int_\Omega \mu_0(n) n_t dx \tag{59}$$

The above scheme can be discretized as:

$$\frac{n^{k+1} - n^k}{\Delta t} + C\Delta_h^2 n^{k+1} - \frac{r^{k+1}}{\sqrt{E_p(k^n) + C_0}} \Delta_h \mu_0 \left(n^k \right) = 0 \tag{60}$$

$$r^{k+1} - r^k = \frac{1}{2\sqrt{E_p(n_k) + C_0}} \int_\Omega \mu_0(n^n) \left(n^{k+1} - n^k \right) dx \tag{61}$$

where h is the approximation, and k is an integer that describes successive values of an associated parameter (e.g., r or n) as this parameter changes in time and/or space. The evolution scheme for the molar density n can then be calculated as:

$$\begin{aligned} n^{k+1} &= \left(1 + \Delta t C \Delta_h^2\right)^{-1} n^k + \\ & dt \frac{r^{k+1}}{\sqrt{E_p(n^k) + C_0}} \left(1 + \Delta t C \Delta_h^2\right)^{-1} \Delta_h \mu_0 \left(n^k\right) \end{aligned} \tag{62}$$

$$r^{k+1} = \frac{r^k + \frac{1}{2\sqrt{E_p(n^k) + C_0}} \int_\Omega \mu_0(n^k) \left[\left(1 + \Delta t C \Delta_h^2\right)^{-1} - 1 \right] n^k dx}{1 - \frac{\Delta t}{\sqrt{E_p(n^k) + C_0}} \int_\Omega \mu_0(n^k) \left[\left(1 + \Delta t C \Delta_h^2\right)^{-1} - 1 \right] n^k dx} \tag{63}$$

It is noted that Eqs. (62)-(63) are now linear and thus, they do not require intensive CPU resources for solving them.

The robustness and efficiency of this scheme can be proved by a 3D simulation of droplet motions in a single pore. The whole domain can be represented by $\Omega = (0 \times L^3)$, which is constructed by a uniform mesh of 200 × 200 × 200 grids and the time step is 0.0001 s. The initial condition is to impose the liquid phase at 350 K in the region of $(0.3L, 0.7L)^3$, and the rest domain is full of the nC4 in gas phase. Four cube droplets are designed at the beginning. Here the initial value of the liquid molar density and gas molar density are 8878.893849 mol/m^3 and 403.172584 mol/m^3.

Fig. 2 shows the evolution history of the 3D simulation which describes the cube liquid droplet will merge together and turns its shape into a perfect ball shape after certain time steps by the force of the surface tension. A very short CPU time is needed in the total 3D calculation using this scheme, in the level of a few minutes reaching the steady state. The Fig. 3 shows the evolution history of the total Helmholtz free energy and total mass of the fourth-order equation in the same process. It's easy to see that an obvious dissipative trend can be captured during the evolution history of the total Helmholtz free energy, with a sharp decline at the start and gradual flat in the later time. It's a significant symbol to show that the solution has reached the stable state. This energy decay trend meets well with the general Second Law in Thermodynamics, which indicates the energy conservation property. In the meantime, the mass conservation property has also been maintained during the whole process, where the mass is represented by molar density.

3.2 Digital twin regarding oil-water separation

The water presence in oil flow is hard to avoid in realistic petroleum engineering, but the risk of chemical or electrochemical corrosion in the transportation and storage facilities along the production cycle from oil well to the market is significantly increased. Damage to the facilities structure may cause further serious problems to the human society including environmental pollution, economic loss and endangering the public health. Thus, oil-water separation is urgently required in the petroleum industry along the whole cycle, from on-field processing to oil refineries. In this section, the well-designed digital twin equipped with the Cahn-Hilliard-Navier-Stokes type phase field model can advise the design of separation facilities and processes by providing a more accurate description of oil/water droplet motion tracks and velocities.

In order to identify the regions occupied by the two fluids, we introduce a phase-field variable ϕ such that:

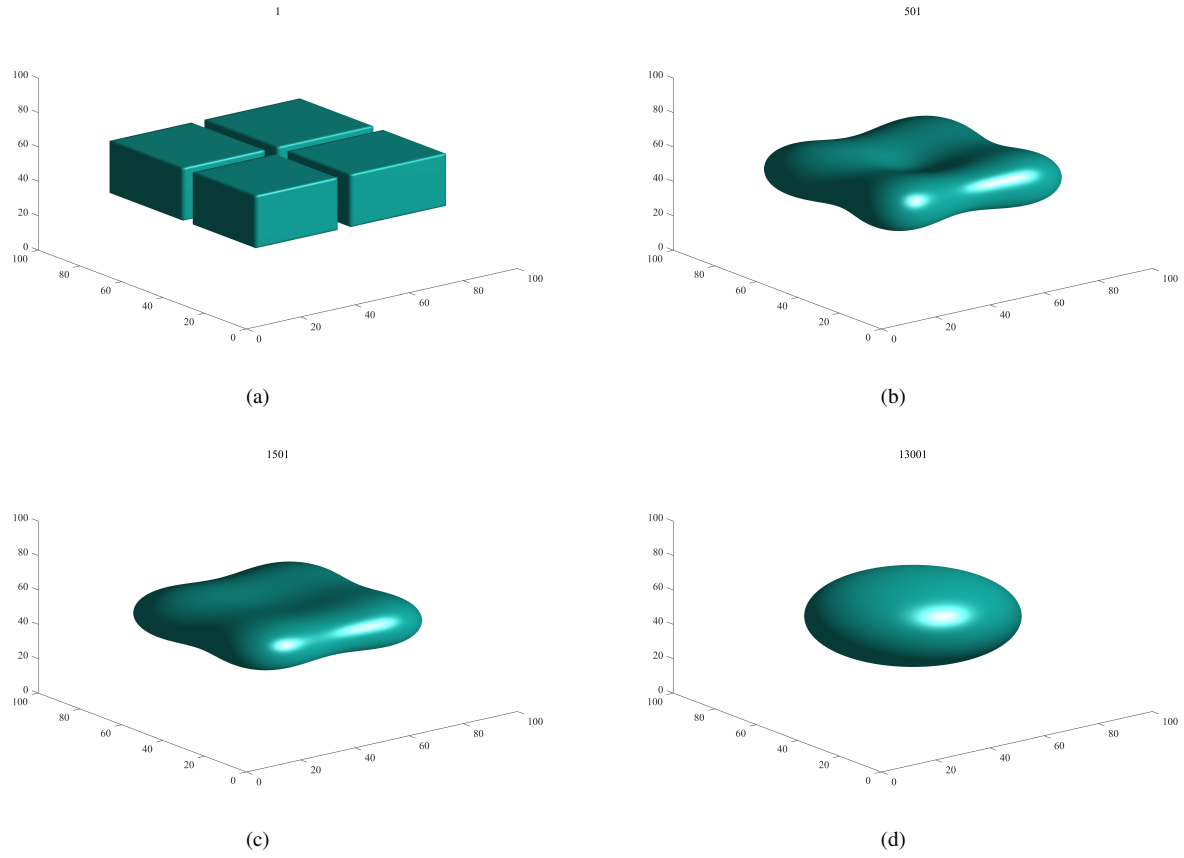


Fig. 2. 2D cross profile of the phase diagram of the four droplets.

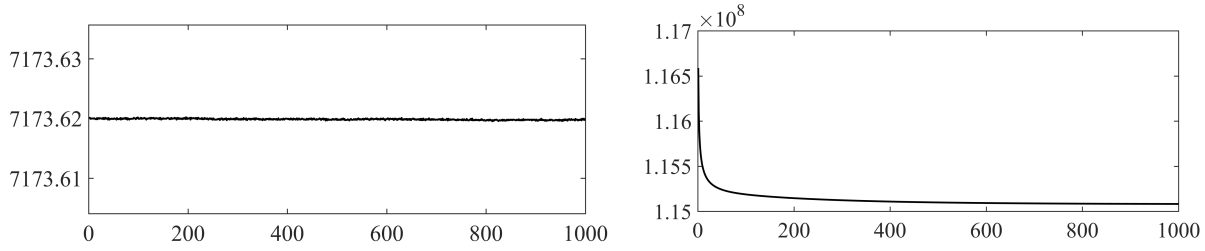


Fig. 3. The evolution history of the total mass (left) and total energy (right).

$$\phi(\boldsymbol{x},t) = \begin{cases} 1 & \text{fluid 1} \\ -1 & \text{fluid 2} \end{cases} \tag{64}$$

The phase-field variable $\phi(\boldsymbol{x},t)$ can be either 1 or -1 to indicate different bulk phases of the system, but a value between $(-1,1)$ within the interface, modeled as a thin (but nonzero-thickness) transient layer. The (free) energy of the system (considering only spatial variation, not time change for now) is $F(\phi) = F_b(\phi) + F_\nabla(\phi)$, where the bulk free energy $F_b(\phi)$ can be modeled by the double-well potential $F_b(\phi) = \int_\Omega f_b(\boldsymbol{x})d\boldsymbol{x}$, where $f_b(\phi) = c_b/4(\phi^2-1)^2$. The general form of Cahn-Hilliard equation is as follows:

$$\frac{\partial\phi}{\partial t} = \nabla\left(M\nabla\frac{\partial f(\phi)}{\partial\phi}\right) \tag{65}$$

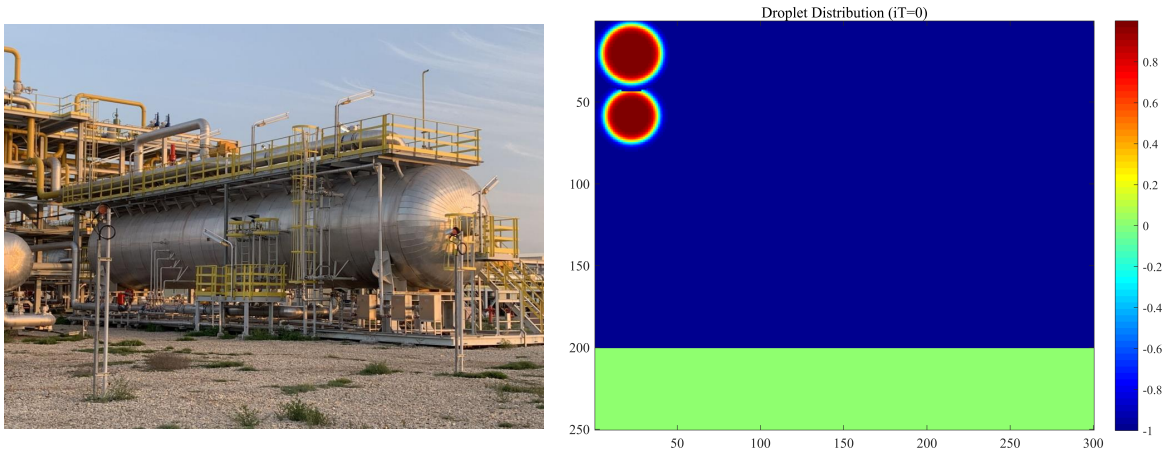
where M is mobility and ϕ represents order parameter, which can be calculated by

$$\phi = \frac{\rho_2-\rho_1}{\rho_2+\rho_1} \tag{66}$$

where ρ_1, ρ_2 represents the local density of Fluid 1 and 2 individually and apparently, $\phi = 0$ at the interface, which is a common standard to check our algorithm. The term $f(\phi)$ in Eq. (65) is free energy, calculated by

$$f(\phi) = \pm\left(\frac{1-\phi^2}{2}\right)^2 \tag{67}$$

The space derivative of order parameter ϕ is



(a) (b)

Fig. 4. Realistic oil-water separator on the field (left) and the corresponding computational scenario in the digital space (right).

$$\phi' = \frac{1-\phi^2}{\sqrt{2}\varepsilon} \tag{68}$$

and it's easy to get that:

$$\frac{1}{2}\left(\frac{1}{1-\phi}+\frac{1}{1+\phi}\right)\phi' = \frac{1}{\sqrt{2}\varepsilon} \tag{69}$$

After integration, we can get

$$\frac{1}{2}\ln\frac{1+\phi}{1-\phi} = \frac{x}{\sqrt{2}\varepsilon} \tag{70}$$

So the order parameter is finally transferred to

$$\phi = \tanh\left(\frac{x}{\sqrt{2}\varepsilon}\right) \tag{71}$$

If the free energy is calculated by

$$f(\phi) = \frac{u\phi^4}{4} - \frac{r\phi^2}{2} + \frac{r^2}{4u} \tag{72}$$

where κ, r, u are interface parameters. Finally, the chemical potential can be written as:

$$\mu = -\kappa\Delta\phi - r\phi + u\phi^3 \tag{73}$$

The two-phase Navier-Stokes equation can be formulated by introducing the phase index into the Navier-Stokes schemes described in Section 2.1 as:

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v}\right) = \rho\mathbf{g} - \nabla p + \eta\nabla^2\mathbf{v} + \mu\nabla\phi \quad \text{in } \Omega \tag{74}$$

Using the above established numerical scheme, oil-water separations in a realistic-scale separator can be simulated in the digital space, with the computational scenario designed as illustrated in Fig. 4, in order to recover the realistic oil-water separation process on the filed (plotted in the left), a computational scenario is designed in the right. The oil flow (marked as blue) with water droplets (marked as red) is injected from the left-top inlet in the horizontal oil-water separator, where a water-pass membrane is placed at the right-bottom corner. A gathering pipeline is placed under the separator to collect the separated water for further usage, for example, purified and then injected back into the reservoir for recovery.

The simulated separation process is illustrated in Fig. 5. The two droplets will drop down under the effect of gravity and merge together under the effect of surface tension. Moving with the oil fluid, the merged water droplet will flow to the right-bottom corner and then pass through the membrane into the gathering pipeline. Using this algorithm with proved effectiveness, the digital twin can provide reliable information suggesting practical oil-water separations, for example, optimized injection speed for a better separation performance.

The simulated separation process is illustrated in Fig. 5. The two droplets will drop down under the effect of gravity and merge together under the effect of surface tension. Moving with the oil fluid, the merged water droplet will flow to the right-bottom corner and then pass through the membrane into the gathering pipeline. Using this algorithm with proved effectiveness, the digital twin can provide reliable information suggesting practical oil-water separations, for example, optimized injection speed for a better separation performance.

4. Concluding remarks

A mathematical tool package is generated in this paper, equipped with advanced numerical models and algorithms with the "6M" properties (multi-scale, multi-domain, multi-physics and multi-numerics numerical modeling and simulation of multi-component and multi-phase fluid flow). As illustrated in Fig. 6, for a macroscopic flow problem, we can model it either in the Navier-Stokes scheme, suitable for the injection, transportation and oil-water separation processes, or in the Darcy scheme, suitable for the drainage and sorption processes. A numerical scheme is implemented to solve Navier-Stokes equations based on cell-centered finite difference over staggered grid. In this scheme, all the difference operations

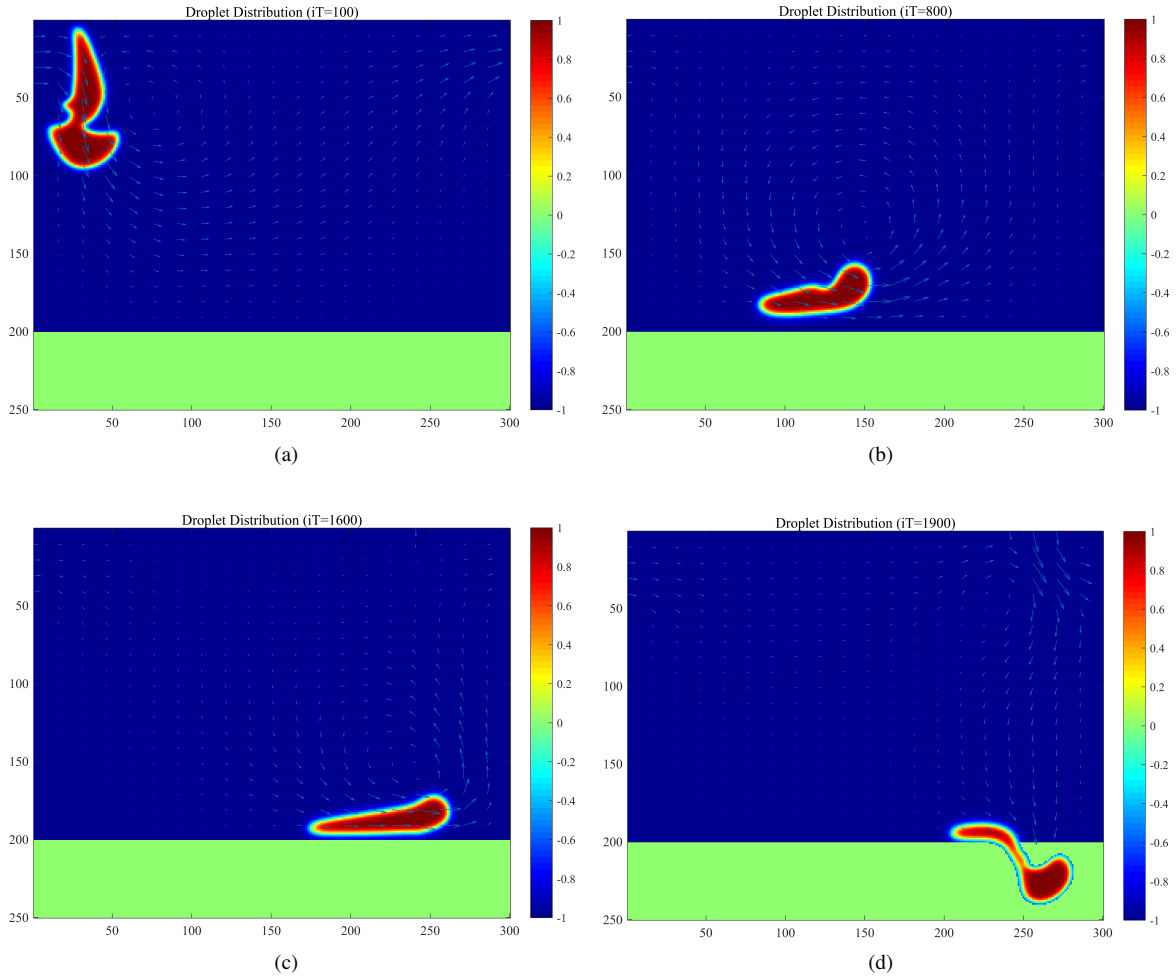


Fig. 5. Simulation results of the oil-water separation process.

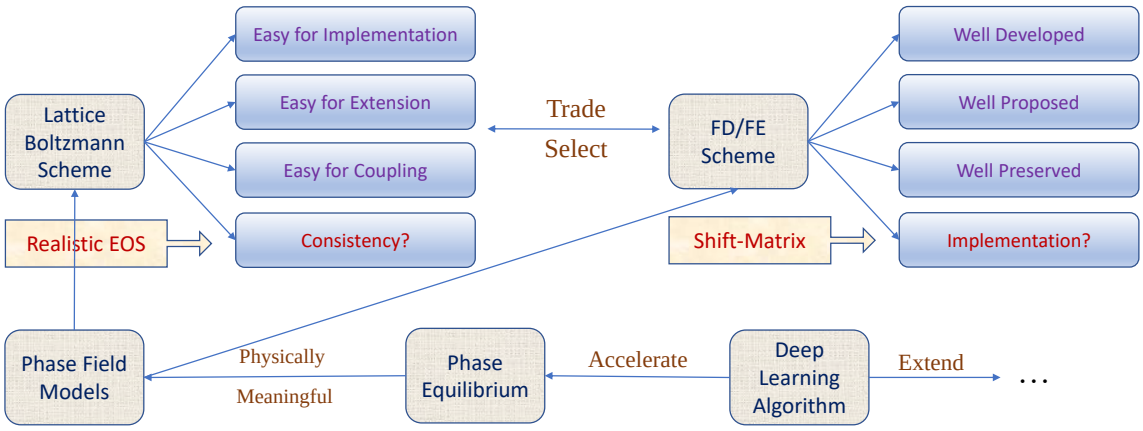


Fig. 6. A schematic diagram of the mathematical tool package.

have been vectorized thereby eliminating loops. The length of the code was also obviously reduced with this shift-matrix approach, and this technique can be easily extended to other numerical schemes constituting of more kinds of partial difference equations, such as Cahn-Hilliard equation common used in phase field models. The energy stability estimates are obtained for three common used discrete schemes, i.e., the linear implicit scheme, the pressure correction scheme and the pressure correction schemes, and an upwind scheme is applied in the discretization of the convection term which plays an important role in the design of unconditionally stable discrete schemes. Lattice Boltzmann method can also be developed as a special discretization of the Navier-Stokes scheme, which is easy to be coupled with multiple distributions, for example,

temperature field, and a rigorous Chapman-Enskog expansion is performed to show the equivalence between the LBGK formulation and the corresponding macroscopic Navier-Stokes equations and generalized convection-diffusion equation. Two kinds of fully conservative IMPES schemes are proposed to investigate the oil-water displacement in porous reservoirs. In order to ensure our model and simulation thermodynamically consistency, the phase equilibrium studies are needed to provide a physically-consistent initial state for the multiphase flow systems, while the deep learning algorithm can accelerate the conventional iterative flash calculations estimating phase equilibrium states.

Based on that, for various practical applications in petroleum engineering related with reservoirs, we can always find the suitable numerical tools to construct a digital twin to simulate the operations, design the facilities and optimize the processes. For example, we can model and simulate the oil-water separation process in on-field processing using a Navier-Stokes scheme, coupled with phase field models to describe the two-phase distribution and phase equilibrium studies to evaluate the effect of temperature and electric forces on the droplet motion and phase change. The droplet motion track and velocities simulated in the digital twin can help the design and optimization of the oil-water separator on the field in the physical space. The multi-phase flow can be simulated in a single pore using an efficient SAV scheme to ensure an acceptable CPU time for 3D cases, which can provide the feedback on the capillary pressure and surface tension information for the upper scale description. The recent popularity on deep learning techniques has reformed many conventional numerical schemes for better computational efficiency and robustness. Especially, the very recent development on physics-informed neural network has enabled more accurate description of physical rules in the neural network structures and the trained model is believed to adhere to general physical and thermodynamic laws better compared with data-driven deep learning techniques. Thus, deep learning techniques with a strong physical basis is expected to be the first choice of engineering computing.

This exploratory study has the following highlights:

1) The new concept: A new concept of reservoir digital twin is proposed in this dissertation, covering a wide range of engineering processes related with the reservoirs, including the drainage, sorption and phase change in the reservoirs, as well as extended processes like injection, transportation and on-field processing.
2) The advanced method: The mathematical tool package for constructing the numerical description in the digital space for various engineering processes in the physical space is equipped with certain advanced models and algorithms developed by ourselves.
3) The open extensions: The "6M" property of our digital twin can ensure a quick extension to other mechanisms or processes, and the preparations in many numerical schemes can help us select the most suitable combination for various applications. For example, we have shown the coupling of phase field models to the Navier-Stokes schemes, and the multiple distribution LBM scheme is also easy for extensions.
4) The practical applications: In order to bridge the gap between the numerical investigations in digital space and the engineering needs in physical space, the efficient implementation is our first goal. For example, the mass conservation and thermodynamic stability properties in our numerical approach can ensure a large time step to benefit the study in real time scales, and a large space scale for practical applications.

Acknowledgement

This work was supported by funding from the National Natural Scientific Foundation of China (Grants No. 51874262) and King Abdullah University of Science and Technology (KAUST) through the grants BAS/1/1351-01-01.

Conflict of interest

The authors declare no competing interest.

References

Abe, T. Derivation of the lattice Boltzmann method by means of the discrete ordinate method for the Boltzmann equation. Journal of Computational Physics, 1997, 131(1): 241-246.

Boschert, S., Rosen, R. Mechatronic Futures, in Digital Twin—the Simulation Aspect, edited by H. Peter and B. David, Springer, Cham, pp. 59-74, 2016.

Chen, H., Kou, J., Sun, S., et al. Fully mass-conservative IMPES schemes for incompressible two-phase flow in porous media. Computer Methods in Applied Mechanics and Engineering, 2019, 350: 641-663.

Chorin, A. J. Numerical solution of the Navier-Stokes equations. Mathematics of Computation, 1968, 22(104): 745-762.

Guermond, J. L., Minev, P., Shen, J. An overview of projection methods for incompressible flows. Computer Methods in Applied Mechanics and Engineering, 2006, 195(44-47): 6011-6045.

He, B., Bai, K. -J. Digital twin-based sustainable intelligent manufacturing: A review. Advances in Manufacturing, 2021, 9(1): 1-21.

Khasanov, M., Krasnov, F. Digital twin of a research organization: Approaches and methods. Paper SPE 198372 Presented at SPE Annual Caspian Technical Conference, Baku, Azerbaijan, 16-18 October, 2019.

Kou, J., Sun, S. A new treatment of capillarity to improve the stability of IMPES two-phase flow formulation. Computers and Fluids, 2010, 39(10): 1923-1931.

Kritzinger, W., Karner, M., Traar, G., et al. Digital Twin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine, 2018, 51(11): 1016-1022.

Mayani, M. G., Baybolov, T., Rommetveit, R., et al. Optimizing drilling wells and increasing the operation

efficiency using digital twin technology. Paper SPE 199566 Presented at IADC/SPE International Drilling Conference and Exhibition, Galveston, Texas, 3-5 March, 2020.

Qi, Q., Tao, F. Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison. IEEE Access, 2018, 6: 3585-3593.

Qian, Y., d'Humières, D., Lallemand, P. Lattice BGK models for Navier-Stokes equation. Europhysics Letters, 1992, 17(6): 479-484.

Rao, S. V. Using a digital twin in predictive maintenance. Journal of Petroleum Technology, 2020, 72(8): 42-44.

Shao, G., Helu, M. Framework for a digital twin in manufacturing: Scope and requirements. Manufacturing Letters, 2020, 24: 105-107.

Spiegel, E. A., Veronis, G. On the Boussinesq approximation for a compressible fluid. The Astrophysical Journal, 1960, 131: 442-447.

Sun, S., Zhang, T. A 6M digital twin for modeling and simulation in subsurface reservoirs. Advances in Geo-Energy Research, 2020, 4(4): 349-351.

Tao, F., Cheng, J., Qi, Q., et al. Digital twin-driven product design, manufacturing and service with big data. The International Journal of Advanced Manufacturing Technology, 2018, 94(9): 3563-3576.

Thoresen, K. E., Kyllingstad, Å., Hovland, S., et al. Using an advanced digital twin to improve downhole pressure control. Paper SPE 194088 Presented at SPE/IADC International Drilling Conference and Exhibition, The Hague, The Netherlands, 5-7 March, 2019.

Wanasinghe, T. R., Wroblewski, L., Petersen, B. K., et al. Digital twin for the oil and gas industry: Overview, research trends, opportunities, and challenges. IEEE Access, 2020, 8: 104175-104197.

Zhang, T., Salama, A., Sun, S., et al. A compact numerical implementation for solving Stokes equations using matrix-vector operations. Procedia Computer Science, 2015, 51: 1208-1218.

Zhang, T., Sun, S. A Compact and Efficient Lattice Boltzmann Scheme to Simulate Complex Thermal Fluid Flows, in Computational Science-ICCS 2018, edited by Y. Shi, H. Fu, Y. Tian, et al., Springer, Cham, pp. 149-162, 2018.

Zhou, G., Zhang, C., Li, Z., et al. Knowledge-driven digital twin manufacturing cell towards intelligent manufacturing. International Journal of Production Research, 2020, 58(4): 1034-1051.

Zhuang, C., Gong, J., Liu, J. Digital twin-based assembly data management and process traceability for complex products. Journal of Manufacturing Systems, 2021, 58: 118-131.