Application of Deep Neural Networks to the Operator Space of Nonlinear PDE for Physics-Based Proxy Modelling

Hadjisotiriou, George; Pour, Kiarash Mansour; Voskov, Denis

DOI
10.2118/212217-MS

Publication date
2023

Document Version
Final published version

Published in
SPE Reservoir Simulation Conference 2023 Proceedings Papers

SPE-212217-MS

Application of Deep Neural Networks to the Operator Space of Nonlinear PDE for Physics-Based Proxy Modelling

George Hadjisotiriou and Kiarash Mansour Pour, Delft University of Technology; Denis Voskov, Delft University of Technology Stanford University

 DOI 10.2118/212217-MS

This paper was prepared for presentation at the SPE Reservoir Simulation Conference held in Galveston, Texas, USA, 28–30 March 2023.

Abstract

In this study, we utilize deep neural networks to approximate operators of a nonlinear partial differential equation (PDE), within the Operator-Based Linearization (OBL) simulation framework, and discover the physical space for a physics-based proxy model with reduced degrees of freedom. In our methodology, observations from a high-fidelity model are utilized within a supervised learning scheme to directly train the PDE operators and improve the predictive accuracy of a proxy model. The governing operators of a pseudo-binary gas vaporization problem are trained with a transfer learning scheme. In this two-stage methodology, labeled data from an analytical physics-based approximation of the operator space are used to train the network at the first stage. In the second stage, a Lebesgue integration of the shocks in space and time is used in the loss function by the inclusion of a fully implicit PDE solver directly in the neural network's loss function. The Lebesgue integral is used as a regularization function and allows the neural network to discover the operator space for which the difference in shock estimation is minimal. Our Physics-Informed Machine Learning (PIML) methodology is demonstrated for an isothermal, compressible, two-phase multicomponent gas-injection problem. Traditionally, neural networks are used to discover hidden parameters within the nonlinear operator of a PDE. In our approach, the neural network is trained to match the shocks of the full-compositional model in a 1D homogeneous model. This training allows us to significantly improve the prediction of the reduced-order proxy model for multi-dimensional highly heterogeneous reservoirs. With a relatively small amount of training, the neural network can learn the operator space and decrease the error of the phase-state classification of the compositional transport problem. Furthermore, the accuracy of the breakthrough time prediction is increased therefore improving the usability of the proxy model for more complex cases with more nonlinear physics.

Introduction

Accurate reservoir simulation of gas injection problems for hydrocarbon production or carbon dioxide sequestration requires compositional simulation in order to account for complex compositional effects including the development of miscibility. However, compositional simulation is computationally intensive due to nonlinear thermodynamics based on equilibrium relations and a complex coupling of flow, transport

and mass transfer between the phases. Additionally, high-fidelity reservoir models contain a large number of grid blocks thus requiring even more computation time. Developed methods to effectively manage computational resources attempt to reduce the size of the problem by reducing the number of components (i.e. lumping, pseudo-models, etc.) or the number of blocks (i.e. upscaling) while still conserving key features of the full compositional solution. These proxy models are useful for optimization or history matching problems where a large number of model runs is required and simulation with the high-fidelity model is prohibitively expensive. Besides reducing the size of the problem, it is also possible to increase the efficiency of compositional simulation by accelerating the thermodynamic equilibrium calculations.

Proxy modeling with a reduced number of components increases the efficiency of compositional simulation by reducing the number of equations that need to be solved. In the lumping technique, the problem is reformulated by presenting the formulation with a limited number of equivalent components and delumping the components at the production stream (Rastegar and Jessen 2009a, 2009b). Further techniques rely on a numerical representation of the method of characteristics. Tang and Zick (1993) propose a limited compositional reservoir simulator that solves a four-component problem by solving an equivalent pseudo-ternary problem. Ganapathy and Voskov (2018) proposes a physics-based method for proxy modeling, termed multi-scale reconstruction in physics. In their method, a pseudo-binary proxy model is used to solve any transport problem with an arbitrary number of components. Their model is applied to four and eight-component systems with either the equation of state or constant K-value thermodynamics and accurately locates the major features of the compositional displacement profile (Chen and Voskov 2020).

Further developments in the space of compositional simulation focus on accelerating phase-state or thermodynamic equilibrium calculations. Compositional Space Parameterisation (CSP) and the following Compositional Space Adaptive Tabulation (CSAT) techniques parameterize the entire or part of the tie-line space (Voskov and Tchelepi 2008, 2009). CSP and CSAT are further extended to achieve an equation-of-state-free general-purpose compositional simulation where the governing differential equations are projected on the tie-simplex space (Zaydullin et al. 2013). Additional techniques that reduce the cost of phase behavior computation include the reduced variable method and the shadow region method (Pan and Tchelepi 2011; Rasmussen et al. 2003).

As a logical extension of parametrization approaches, the Operator-Based Linearization (OBL) method has been introduced in Voskov (2017). In the OBL approach, the governing equations of the physical problem are linearized and represented as a combination of space- and state-dependent functions. In addition, a uniform mesh is introduced into the physical space of the problem where operators are interpolated. This limited, linear approximation of the physical space provides better nonlinear convergence with controlled error (Lyu et al. 2021; Wang et al. 2020).

Recently, the application of machine learning techniques to solve linear or nonlinear partial differential equations (PDEs) has gained interest for their ability to solve PDEs without any labeled data of the solution. These machine learning techniques such as constrained learning and physics-informed machine learning (PIML) utilize artificial neural networks (NN) to solve complex equations by encoding the known physics of the problem into the loss function of the NN as prior-knowledge (Raissi et al. 2019). For PDEs, this is achieved by including the residual form of the equation as a regularization constant in the loss function in addition to a regressive term that uses initial and boundary conditions. In the loss functions, the regularization term effectively constrains the hypothesis space of the neural network thus allowing the physics-informed neural network (PINN) to learn the solution without any labeled data on the interior of the solution domain (Raissi et al. 2019; Stewart and Ermon 2016).

Raissi et al. (2019) and Fuks and Tchelepi (2020) demonstrate the application of PIML for nonlinear PDEs. Raissi et al. (2019) use PIML to accurately solves the Burgers equation, a quasi-linear PDE, and find that with a small number of initial and boundary points that their nonlinear behavior is accurately captured. They conclude that if the PDE is well posed, PIML can solve the problem provided that the neural network is expressive and a sufficient number of collocation points are used. Furthermore, Fuks and Tchelepi (2020)

expressly tests the limitations of PIML by application to a hyperbolic two-phase displacement problem. They find that encoding the residual in the loss function does not produce an adequate estimation of saturation irrespective of the number of collocation points or neural architecture. In particular, the neural network is unable to resolve the location of the shock in space and time which is a defining feature of the solution.

Stewart and Ermon (2016) apply constrained learning to convolutional neural networks to track the movement of objects without the use of any labeled data and demonstrate the usefulness of PIML techniques for small data regimes where data is sparse. An additional advantage of PIML for solving PDEs is that automatic differentiation is used to supply derivatives for the residual. Consequently, they are presented as an efficient alternative to traditional numerical modeling techniques such as the Galerkin, collocation, or finite volume methods for higher dimensional problems (Blechschmidt an Ernst 2021). However, for application to hyperbolic problems where an exact solution is required, as opposed to a qualitative approximate solution, PIML remains inadequate (Fuks and Tchelepi 2021).

Besides the application of PIML to solve forward problems, PIML can be used to handle the inverse problem for the data-driven discovery of partial differential equations where unknown parameters within the nonlinear operator of the PDE are approximated from observations of the auxiliary quantity. Raissi et al. (2019) use PIML to qualitatively estimate the pressure and accurately determine hidden parameters within Navier-Stokes and Burger's equation.

In this work, PIML is used to train the operator space of the physics-based proxy of Ganapathy and Voskov (2018) and improve performance for a compressible, gas-injection problem at isothermal conditions. Our approach embeds neural networks in the nonlinear operator of the proxy model and uses the Lebesgue integral of the forward solution with respect to the full compositional model in the loss function to improve the performance of the proxy model. Our method is implemented in TensorFlow and utilizes both automatic differentiation and operator-based linearization approaches. Transfer training is employed for increased computational efficiency and solution accuracy.

Simulation framework

Operator based linearization (OBL)

Compositional simulation of a gas injection problem is carried out in the Delft Advanced Research Terra Simulator (DARTS). DARTS utilizes operator-based linearization (OBL) where the mass conservation equation is written in operator form and a mesh is introduced into the physical space. In this approach, operator values are computed for the mesh vertices onto which a multi-linear interpolant is applied for a continuous representation of the physics (Voksov 2017).

The mass conservation equation for n_c components and n_p phases is written in its molar formulation for an isothermal displacement. The mass conservation equation of component i reads:

$$\frac{\partial}{\partial t}\left(\phi\sum_{j=1}^{n_p} x_{ij}\rho_j s_j\right)+\nabla\cdot\sum_{j=1}^{n_p} x_{ij}\rho_j \boldsymbol{u}_j = -\sum_{j=1}^{n_p} x_{ij}\rho_j q_j, \qquad i=1,\ \ldots,\ n_c \tag{1}$$

where x_{ij} is the mole fraction of component i in phase j, ρ_j is the molar phase density and s_j the phase saturation and ϕ the porosity. Furthermore, $\boldsymbol{u}_j$ is the Darcy flow velocity and is a function of absolute permeability k, phase relative permeability $\boldsymbol{k}_{rj}$ and phase viscosityμ_j,

$$\boldsymbol{u}_j = -k\frac{\boldsymbol{k}_{rj}}{\mu_j}\nabla p. \tag{2}$$

The ratio of relative permeability and viscosity is henceforth referred to as phase mobilityλ_j. Finite volume discretization is applied to the general mass conservation equation with two-point flux approximation and backward Euler approximation in time:

$$V\phi\left(\left(\sum_{j=1}^{np} x_{ij}\rho_j s_j\right)^{n+1} - \left(\sum_{j=1}^{np} x_{ij}\rho_j \boldsymbol{u}_j\right)^{n}\right) - \Delta t\sum_{l\in L}\left(\sum_{j=1}^{np} x_{ij}^{l}\rho_j^{l}T_j^{l}\Delta p\right) + V\sum_{j=1}^{np} x_{ij}\rho_j q_j = 0, \qquad i = 1, \ldots, n_c \tag{3}$$

where V is the control volume, ϕ the porosity and T_j^l the phase transmissibility across the grid block interface L. The phase saturation is calculated as a function of phase density and molar phase fractionν_j :

$$s_j = \frac{\nu_j/\rho_j}{\sum_k^{np} \nu_k/\rho_k}. \tag{4}$$

Following the OBL approach, the discretized equations are rewritten in their algebraic, operator form. In this operator form, state and spatial operators are introduced. State variables (ω) are a function of physical unknowns, pressure p and the overall composition z, whereas spatial variables (ξ) are a function of spatial coordinates (e.g., block geometry or permeability). The residual equation reads:

$$r_i(\omega, \xi) = V(\xi)\phi_0(\xi)(\alpha_i(\omega) - \alpha_i(\omega_n)) - \Delta t\sum_l \beta_i(\omega)T^l(\xi)\Lambda(\omega)\Delta p = 0, \qquad i = 1, \ldots, n_c. \tag{5}$$

The residual equations are defined according to the total velocity formulation and contain $2\times n_c+1$ operators. These operators are defined according to Eqs. *6,7* and *8* where α corresponds to the accumulation operator, β the flux operator and Λ the total mobility which in turn is embedded into the flux operator. The equation is solved with the Newton-Raphson method and the fully implicit method for time approximation.

$$\alpha_i(\omega) = \left(1 + c_r\left(p - p_{ref}\right)\right)\sum_{j=1}^{np} x_{ij}\rho_j s_j = \rho_{tot} z_i \tag{6}$$

$$\beta_i(\omega) = \frac{\sum_{j=1}^{np} x_{ij}\lambda_j\rho_j}{\Lambda} \tag{7}$$

$$\Lambda(\omega) = \sum_{j=1}^{np} s_j = 1. \tag{8}$$

To ensure thermodynamic phase equilibrium, the fugacities f of each component in both phases must be equal,

$$f_{i,o}(p, T, x_i) = f_{i,g}(p, T, y_i). \tag{9}$$

Auxiliary relations are that the sum of mole fractions and phase saturations are equal to one,

$$\sum_{i=1}^{n_c} x_{ij}, \qquad j = 1, \ldots, n_p, \tag{10}$$

$$\sum_{j=1}^{np} s_j = 1. \tag{11}$$

To determine the phase behavior of the system, flash calculations are carried out based on constant equilibrium ratios (i.e. K-values) for each component (Orr 2007). In this approach, the phase split is determined by solving the Rachford-Rice equation for the vapor fraction v,

$$h(v) = \sum_{i=1}^{n} \frac{z_i(K_i - 1)}{v(K_i - 1) + 1}, \tag{12}$$

and further deriving the partitioning coefficients for liquid and vapor x_i and y_i from from the tie-line:

$$z_i = x_i(1 - v) + y_i v. \tag{13}$$

Pseudo-binary model

The multi-scale reconstruction of physics (MSRP) method proposed by Ganapathy and Voskov (2018) is utilized for a four-component gas vaporization problem with injection gas and initial oil compositions as described in Table 1 below. The reference solution to this problem is shown in Fig. 1 The injection gas consists of *97%* of methane while the initial oil composition consists of a mixture of all four components. In a quaternary displacement, the solution is defined according to three key tie-lines, namely the initial, injection and cross-over tie-lines. A tie-line is a line in the compositional space where the liquid and gas phases are in thermodynamic equilibrium for a given pressure and temperature and is expressed by the stability condition. For gas injection processes, the compositional path enters and exits the two-phase zone through the extension of the initial and injection tie-lines and is marked by discontinuities in the compositional and saturation profile (i.e. shocks). Consequently, three zones are identified in the compositional profile according to their phase state. These zones are the single-phase vapor, the single-phase liquid and the two-phase zones and are delineated by the leading and trailing shocks. Downstream of the leading shock the mixture is liquid at initial composition conditions while upstream of the trailing shock the composition consists purely of *C1*. Further, the area between these two shocks is occupied by both phases simultaneously.

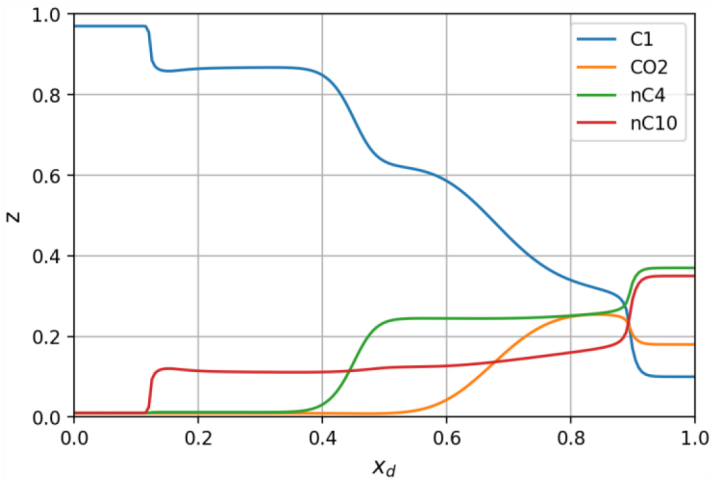

Figure. 1—Compositional (mole fraction) profile for a four-component vaporizing gas drive. The profile is characterized by the leading shock and trailing shocks located at approximately x_d=*0.90* andx_d=*0.12*. From left to right, they delineate three zones where single-phase vapor, both phases and single-phase liquid are present.

Table 1—Vaporizing gas drive.

Component	C1	CO_2	nC4	nC10
Initial	0.10	*0.18*	*0.37*	*0.35*
Injection	*0.97*	*0.01*	*0.01*	*0.01*
K-value	*2.5*	*1.5*	*0.5*	*0.05*

Since the solution of the compositional problem is characterized by the leading and trailing shocks, a lower dimensional proxy model by Ganapathy and Voskov (2018) is used to partially construct part of the solution. This method is able to reduce a compositional problem with an arbitrary number of components to a pseudo-binary problem where the number of components is reduced to two and thus the number of equations is reduced to$2\times n_b$. Their analytic approach relies on the fact that the displacement path can be constructed according to key tie-lines of the system as the leading- and trailing shocks can only enter and leave the two-phase zone along their tie-line extensions (Ganapathy and Voskov 2018; Orr 2007). The first stage of the MSRP reconstruction is able to accurately reconstruct a part of the conservative solution as it manages to locate the leading and trailing shocks in space and time for an incompressible transport problem. Here their model is adapted to include flow as well as transport and used for an isothermal compressible gas vaporization problem.

The location of the shocks for the injection component *C1* is reconstructed. The pseudo-operator space for component i is parameterized according to pseudo-composition z_r and pressure p. The lever rule is applied with fixed partitioning coefficients computed at initial and injection compositions,

$$v = \frac{z_r - x_i}{y_i - x_i}, \tag{14}$$

where z_r is equal toz_i. The resulting vapor fraction, v, and corresponding saturation are used to define the operator space for the initial and injection tie-lines$\beta\hat{}(ini/inj)$. Thereafter, the pseudo-operator space $\beta_r(\omega)$ is defined as the convex hull of the union of each of the key tie-lines,

$$\beta_{ri}(\omega) = conv\left(\beta_i^{ini} \cup \beta_i^{inj}\right). \tag{15}$$

Similarly, $\alpha_r(\omega)$ is defined as the product of the total density and pseudo-compositionz_r,

$$\rho_{tot,r} = conv\left(\rho_{tot}^{ini} \cup \rho_{tot}^{inj}\right), \tag{16}$$

$$\alpha_r(\omega) = \rho_{tot,r} z_r. \tag{17}$$

The analytical construction is illustrated for the flux operator in one dimension (incompressible, transport) in Fig. 2 and in two dimensions (compressible, flow) in Fig. 3a. The operator space switches from the initial to the injection tie-lines at their point of intersection in the nonlinear, two-phase zone. The pseudo-operator $\alpha_r(\omega)$ is shown in Fig. 3b.

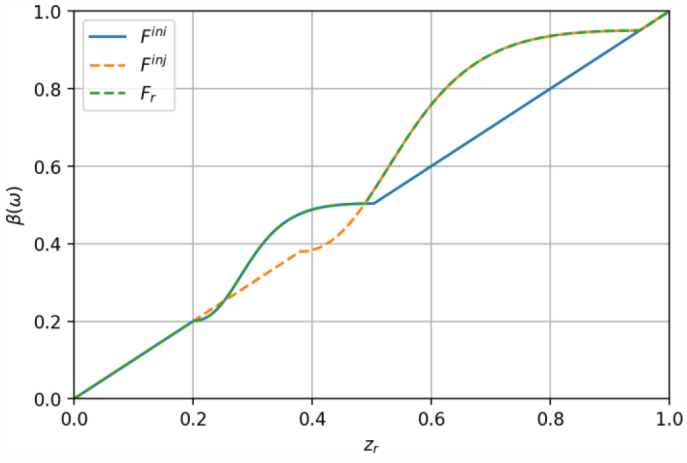

Figure. 2—The convex hull of the union of the initial and injection tie-lines is taken to define the pseudo-operator space. For the incompressible transport problem the β operator corresponds to the fractional flow curve.

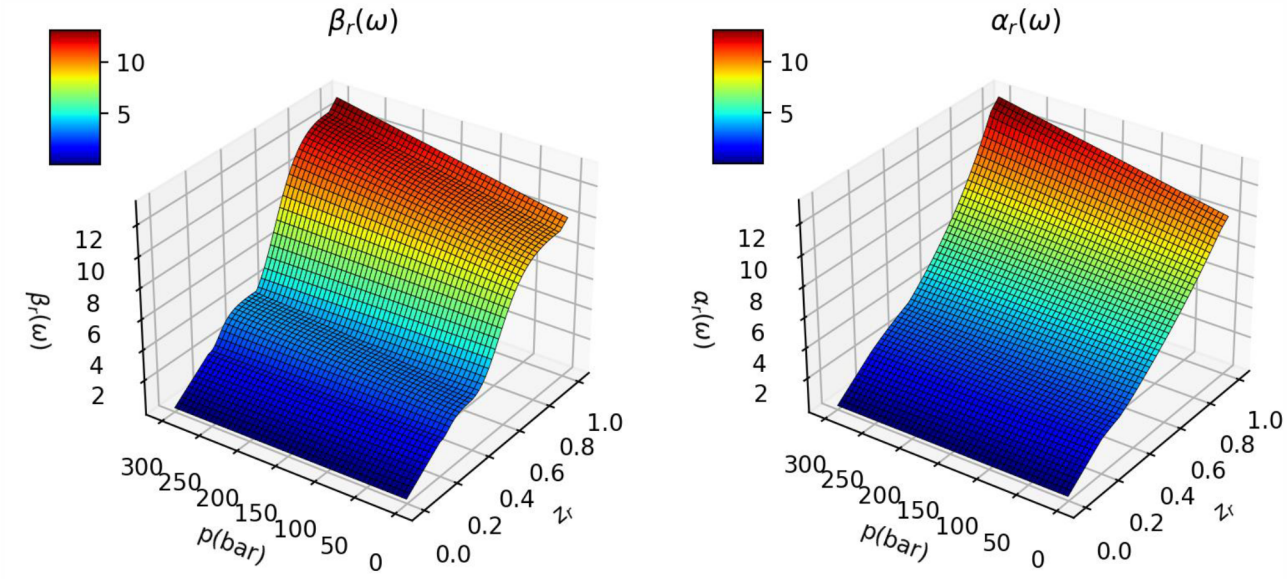

Fig. 3a—Flux operator.

Fig. 3b—Accumulation operator.

Figure. 3—Operator space of the binary proxy model.

The total mobility operator is defined by tabulation of values from the reference solution with respect to the state variables. Care is taken to carry out the reference simulation with variable bottom-hole pressure controls in order to ensure that data points are distributed across the range of the entire parameter space. After that, the scattered data points are used to compute a structured OBL mesh by application of a linear interpolant. Fig. 4 shows the gathered points and the resulting interpolation for the total mobility operator.

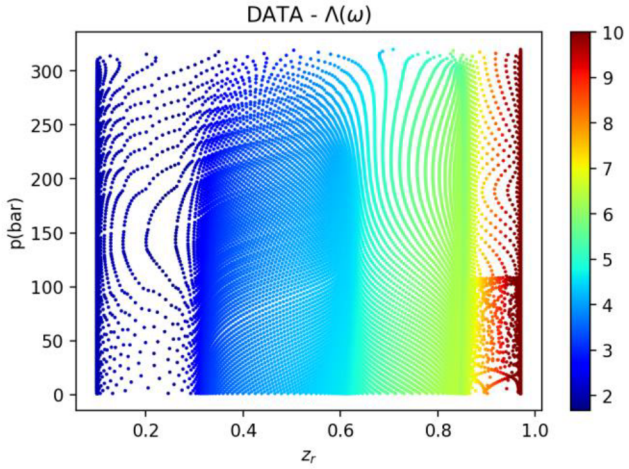

Fig. 4a—Collected dta points from the reference simulation.

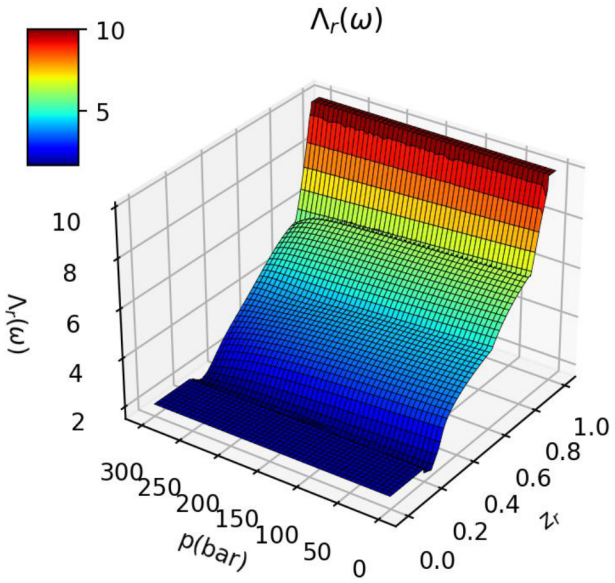

Fig. 4b—Resulting interpolated operator space. In extremum, the total mobility is equal to 2 and 10 in the single-phase liquid and vapor zones.

Figure. 4—Total mobility operator of the binary proxy model.

Machine learning

Deep neural networks' ability to function as universal function approximators are used to approximate and adapt the operator space of the pseudo-binary model. Training is carried out with a specialized regularization constant in the loss function. Deep neural networks are composed of n_l series of functions a,

$$f_\theta(\boldsymbol{X}) = a_{n_l}\left(a_{n_l-1}\ \left(\ldots\left(a_2\ \left(a_1(\boldsymbol{X})\right)\right)\right)\right), \tag{18}$$

where each hidden layer consists of a stack of artificial neurons which process input matrix $\boldsymbol{X}_i$ as the weighted sum of weights $\boldsymbol{W}_i$ and biases $\boldsymbol{b}_i$ before passing through activation function σ,

$$a_i(\boldsymbol{X}) = \sigma\left(\boldsymbol{W}_i\boldsymbol{X}_i + \boldsymbol{b}_i\right), \qquad i = 1,\ \ldots, n_l. \tag{19}$$

Consequently function f is parameterized according to the ensemble of weights and biases θ,

$$\theta = \left\{W_1, W_2,\ \ldots,\ \ W_{n_l}, b_1, b_2,\ \ldots, b_{n_l}\right\}. \tag{20}$$

Often, PIML is applied to a nonlinear PDE of the form,

$$r := z_t + N, \tag{21}$$

where z_t is a scalar quantity and N the nonlinear operator. The residual form of the PDE,

$$r_\theta = \frac{\partial z_\theta}{\partial t} + \frac{u_{tot}}{\phi}\frac{\partial F}{\partial x}, \tag{22}$$

is included in the loss function in addition to a regular regressive loss term. In the first term, the mean squared error (MSE) of a randomly selected set of boundary/initial points (N_z) is computed and in the second term. In the second term, the residual is computed for a random set of collocation points (N_r) and used as a regularization constant. The loss function then reads:

$$L(\theta) = L_z(\theta) + L_r(\theta) = \frac{1}{N_z}\sum_{i=0}^{N_z}\left|z^i_{bc} - z^i_\theta\right|^2 + \frac{1}{N_r}\sum_{i=0}^{N_r}\left|r^i_\theta\right|^2. \tag{23}$$

In our study, a supervised learning scheme applies physics-informed machine learning to adapt the operator space of the aforementioned proxy model and improve its predictive performance. This is achieved by embedding neural networks into the nonlinear operator of the PDE and utilizing the forward solution in a regularization term of the loss function. Our approach integrates the distance between the leading and trailing shocks of the proxy model with respect to the reference model and penalizes the loss function so that the neural network learns the operator space for which the misfit of the shocks is minimal. This can be effectively implemented as an application of the Lebesgue integration in the loss function instead of the Riemann integral.

After solving the proxy and computing the loss function, TensorFlow's reverse-mode automatic differentiation is applied to handle back-propagation and update the ensemble of weights and biases θ (Abadi et al. 2016). In the first step, the analytical approximation of the operator space is effectively loaded into the neural network. Training in this step is carried out for *600* training iterations and the Adam stochastic optimizer with a constant learning rate of *0.001* is used. Training is carried out for a one-dimensional homogeneous reservoir with constant permeability. Furthermore, the training scheme is described according to neural network architecture, optimizer, and activation function. The architecture of the neural network is identical to that of Raissi et al. (2019) and Fuks and Tchelepi (2020). It consists of *8* hidden layers with *20* neurons per layer. Standard gradient descent is used for the optimizer with a constant learning rate of*0.0001*. A small learning rate is expressly chosen since it is necessary for the operator space to maintain the part of its form that defines the shocks in accordance with the velocity constraint of the method of characteristics. The

neural network weights are initialized randomly with Xavier/Glorot initialization. The activation function σ is the hyperbolic tangent function *tanh* and introduces nonlinearity to the system.

Results

This section describes the main results achieved with the transfer learning approach. Three models are compared, the reference model with n_c components and the binary model before (base ML) and after transfer training (transfer ML). Training is carried out with a one-dimensional reservoir model and the trained model is further validated with two-dimensional heterogeneous layers of the SPE10 model.

1D-training

The operator space of the total mobility operator $\Lambda(\omega)$ and the flux operator $\beta(\omega)$ is adapted with our transfer training approach.

In the first step, the neural network is exclusively trained to approximate the OBL mesh with labeled data from the analytical construction.

In the second step, weights are transferred, and the loss function is redefined. the estimated operator space is used within a fully implicit solver embedded in the loss function of the neural network as described previously. The solver utilizes the OBL approach as it interpolates required values from the vertices of the OBL mesh. Training is carried out for a one-dimensional homogeneous reservoir containing *200* grid blocks where the permeability is equal to *100 mD* and the length of the reservoir is*200 m*. During the training period, the bottom-hole pressures are fixed at the limits of the OBL mesh, namely *1 bar* and*300bar*. The Lebesgue integral of the leading and trailing shocks is computed at *10* discrete time steps ranging from t=*1 day* to*10 days*. The range at which the Lebesgue integral of misfit is calculated lies between z_r=*0.90* and *0.96* for the trailing shock and between z_r=*0.11* and *0.30* for the leading shock.

The evolution of the loss function in the second step is plotted in Fig. 5 and the evolution of the misfit of the trailing and leading shocks at every time step, *ts*, and training iteration is given in Fig. 6. A global minimum is reached at training step 7 after which improvements remain negligible. The misfit of the leading and trailing shock decreases and the estimation of the proxy model is improved with respect to the conservative reference solution of n_c components. Changes in the operator space after training with the forward solution are visualized in Fig. 7.

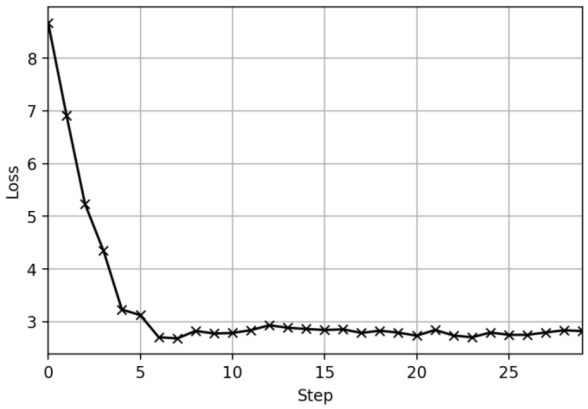

Figure. 5—Evolution of the loss function.

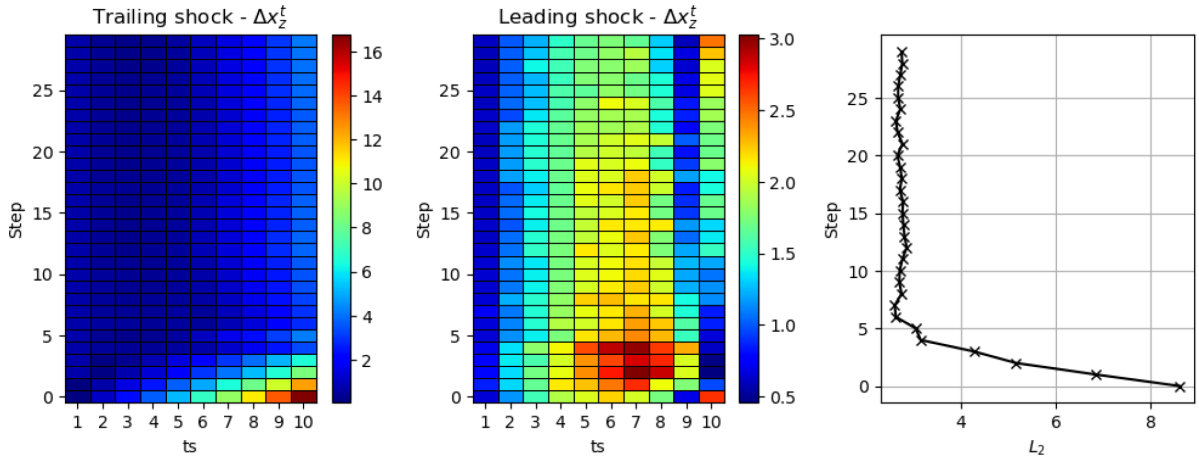

Figure. 6—Evolution of the Lebesgue integral of the trailing and leading shocks at every time step ts, and training step.

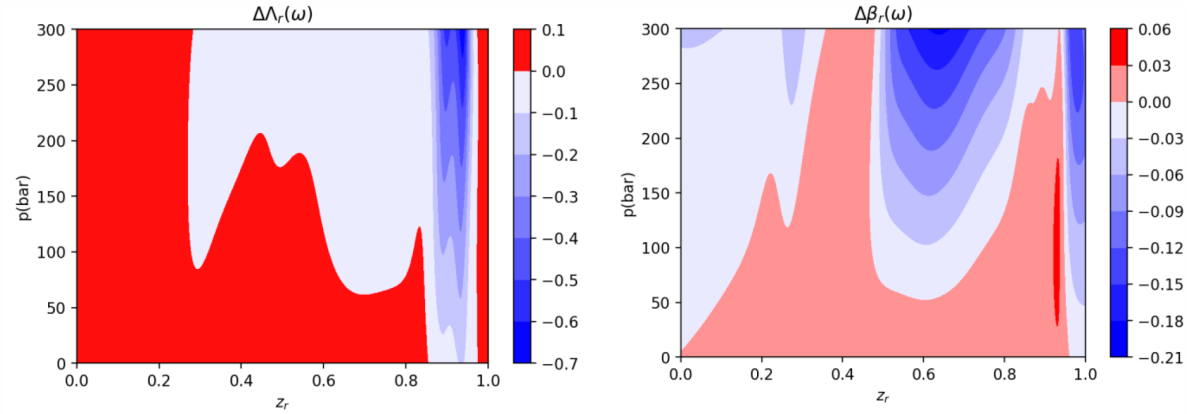

Fig. 7a—Total mobility operator.

Fig. 7b—Flux operator.

Figure. 7—Difference of the operator space after training. The greatest changes in the operator space are concentrated in the area relevant to the trailing shock.

Horizontal layers of SPE10

The trained model is applied to two-dimensional heterogeneous layers from the SPE10 geological model. The two-dimensional models consist of (*220×60*) cells. The SPE10 geological model is divided into two parts, namely the Tarbert formation (top *35* layers) and the Upper Ness formation (bottom *50* layers), which are representations of prograding near shore and fluvial environments (Christie and Blunt 2001). The cells are of size $\Delta x=3m$, $\Delta y=6m$ and $\Delta z=0.6m$. A *5* -spot well pattern is used with an injection well in the middle at fixed bottom hole pressure of *300bars* injecting at *97%* of methane and *4* production wells in the corners at*1bar*. The breakthrough times of the first and last breakthroughs are reported for each of the models. The first and last breakthrough coincides with the breakthrough of the leading and trailing shock at any of the four productionwells. In addition, the quality of the prediction of the locations of the shocks in space and time is quantified by analysis of the phase state of each grid cell.

The performance of each model is evaluated by quantifying the error of the phase-state classification, *S*, of the resulting compositional profile. The phase-state maps identify three zones within the reservoir corresponding to single- and two-phase zones where gas (2), oil and gas (1), or oil (0) are present. The state of each cell is classified according to the tie-line end points of the injection and initial tie-lines,

$$T\left(S_r(x,y,t)-S_{ref}(x,y,t)\right)=1, \qquad for \quad S_r-S_{ref}\neq 1, \tag{24}$$

and the error of the model is quantified as the ratio of the number of misclassified, grid blocks with respect to the reference model and the total number of grid blocks,

$$Error(t) = \frac{\sum_i^{n_b} T\left(S_{r,i} - S_{ref,i}\right)}{n_b}. \tag{25}$$

Figs. 8a and 8b illustrate the distribution of the phases and the error of the proxy model before and after training, for layer 7 of the SPE10 model. The boundary between the zones delineates the location of the leading and trailing shock within the reservoir and thus the error gives an approximation of how well the location of the shock is approximated in space and time. From the error maps, it is observed that the estimation of the location of the trailing shock is visibly improved for the trained model whereas the estimation of the leading shock remains the same or is slightly better. The evolution of the error as a function of time is reported in Fig. 9b and the permeability of layer 7 is shown in Fig. 9a. The error increases as the displacement front grows and the region of displacement increases.

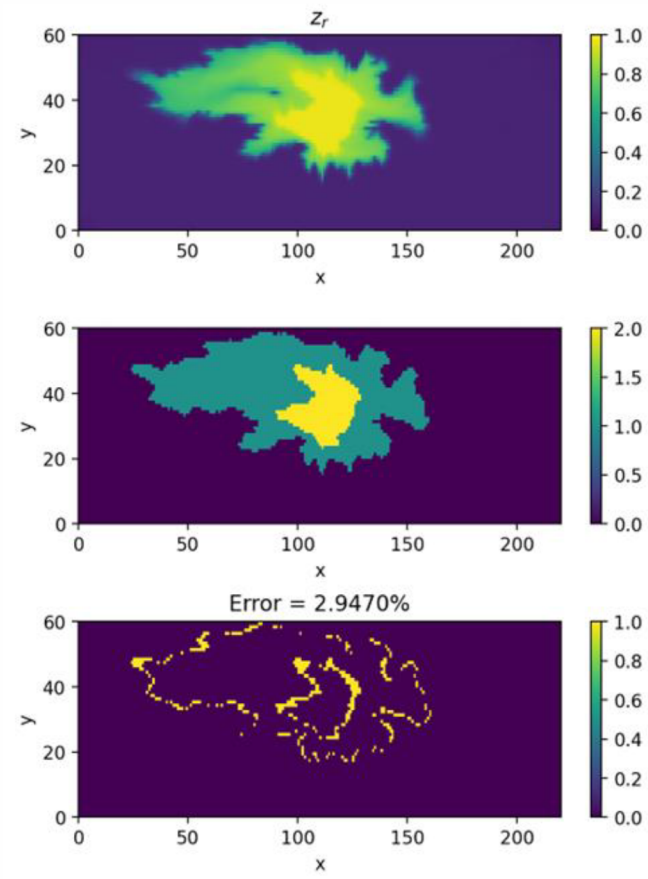

Fig. 8a—Base ML proxy model.

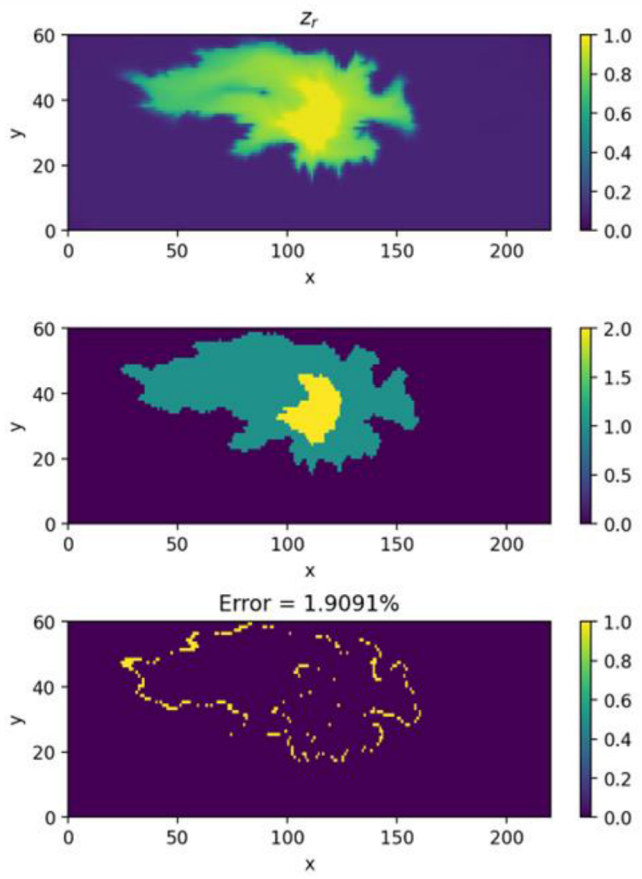

Fig. 8b—Transfer ML proxy model.

Figure. 8—Distribution of the phase-states and the corresponding error at $t = 150$ days for layer 7 of the SPE10 model. The phase-states correspond to single- and two-phase zones where gas (2), oil and gas (1), or oil (0) occupy the grid block. The transfer ML model indicates a marked improvement in the estimation of the trailing shock.

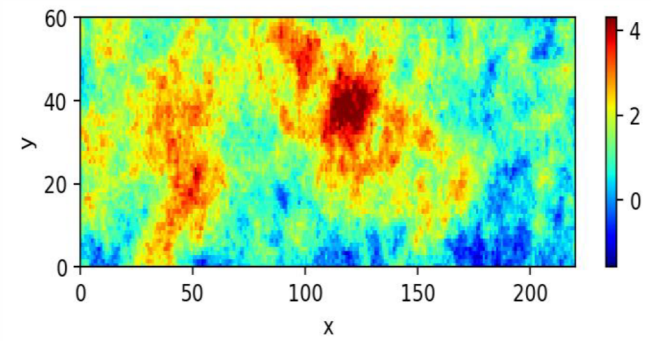

Fig. 9a—Log10 of the permeability distribution of layer 7 of the SPE10 model.

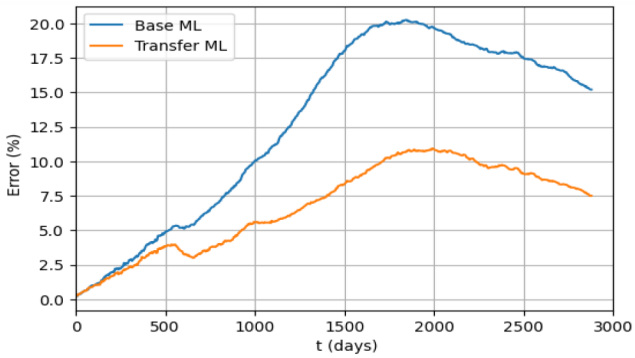

Fig. 9b—Error of the phase-state classification at breakthrough of the leading shock.

Figure. 9—Evolution of the error for layer 7 of the SPE10 model before (base ML) and after transfer training (transfer ML).

The performance before and after training of the proxy model is compared for the top *15* layers of the SPE10 model. The breakthrough times of the leading and trailing shocks are reported per layer in Figs. 10 and 11 together with the corresponding error of the phase-state classification at breakthrough. It is found that the trained model consistently outperforms the model trained at the first stage in terms of error and breakthrough time of the trailing shock while the breakthrough time of the leading shock effectively remains the same as the base ML model already makes a good estimation. The error of the trained model at breakthrough remains below*7.5%*. The average difference in breakthrough time for the trained and untrained models with respect to the reference model is *293days* versus *570days* for the trailing shocks and *15days* versus *16days* for the leading shock.

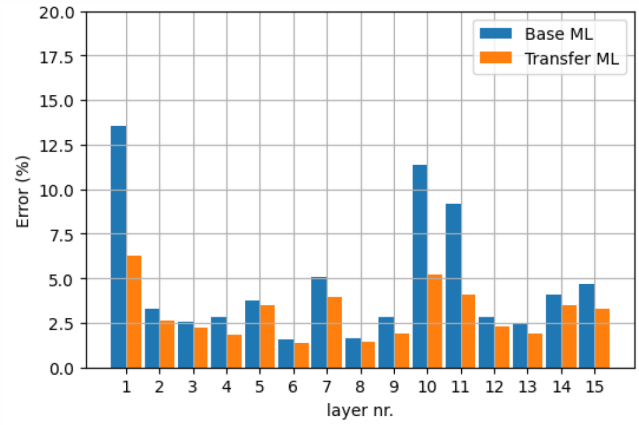

Fig. 10a—Error of the phase-state classification at breakthrough of the leading shock.

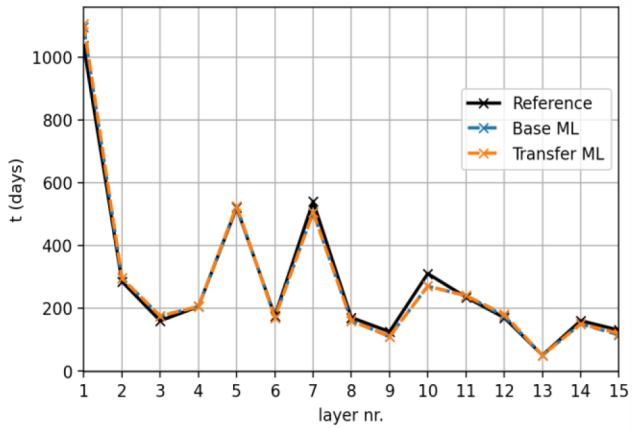

Fig. 10b—Breakthrough times of the leading shock at one of the wells.

Figure. 10—The breakthrough time of the leading shock at one of the wells and the corresponding error of the phase-state maps is assessed for the top 15 layers of the SPE10 model.

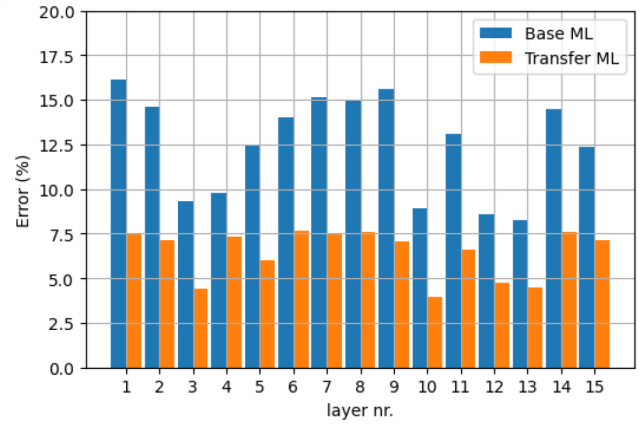

Fig. 11a—Error of the phase-state classification at breakthrough of the trailing shock.

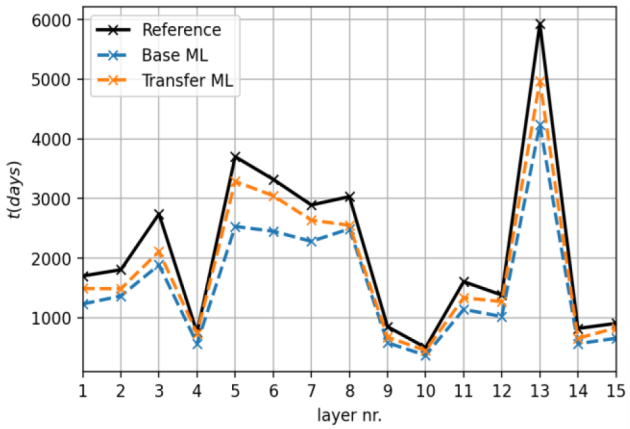

Fig. 11b—Breakthrough times of the trailing shock at one of the wells.

Figure. 11—The breakthrough time of the trailing shock at one of the wells and the corresponding error of the phase-state maps is assessed for the top 15 layers of the SPE10 model.

Conclusion

Compositional simulation can be prohibitively expensive for optimization and history-matching problems where a large number of model runs are required. Therefore, proxy models that use a reduced number of components while maintaining key parts of the high-fidelity solution are an attractive alternative for these applications. In this report, the proxy model based on Multiscale Reconstruction in Physics is used for a compressible gas injection problem. Neural networks' ability to operate as universal function approximators is leveraged to learn the physical space of the full compositional simulation in the reduced dimension (1D) and improve the proxy model for multi-dimensional simulation. Our approach uses an adaptation of PIML by subjecting the operator space of the proxy model to neural networks and including the Lebesgue integration of the forward solution in the loss function. Improvements in the estimation of breakthrough times are realized for the fifteen layers of the SPE10 model. This method can be further extended for developing a generic physical proxy model with lower degrees of freedom trained at data points generated by a high-fidelity physical model or based on real-world observations.

Acknowledgement

We thank Dr. Jeroen Groenenboom and Dr. Ali Fadili of Shell for several helpful and insightful discussions.

Nomenclature: Acronyms

CSP	Compositional Space Parameterisation
CSAT	Composiational Space Adaptive Tabulation
DARTS	Delft Advanced Research Terra Simulator
OBL	Operator-based Linearization
NN	Neural Network
MSRP	Multiscale Reconstruction in Physics
PIML	Physics Informed Machine Learning
PINN	Physics Informed Neural Network
PDE	Partial Differential Equation

Variables

$\boldsymbol{k}$	absolute permeability
k_r	relativie permeability

f_{ij} fugacity of component i in phase j
n_b number of grid blocks
n_c number of components
n_p number of phases
n_l number of hidden layers
p pressure
q source/sink
s_j saturation of phase j
S phase state
v_j molar fraction of phase j
T^l geometric transmissibility at interface l
x_{ij} mole fraction of component i in phase j
x_i liquid fraction of component i
y_i vapor fraction of component i
V bulk volume
u Darcy velocity
z overall composition
X input feature matrix
W_i weight matrix of layer i
b_i biases of layer i

Greek symbols

λ_j mobility of phase j
μ_j viscosity of phase j
ρ_j molar density of phase j
ϕ porosity
α_i accumulation operator of component i
β_i flux operator of component i
Λ total mobility operator
ω state variable
ξ geometric variable
θ ensemble of neural network weights and biases
σ activation function

Superscripts

ini initial
inj injection

Subscripts

o oil phase
g gas phase
i component
j phase
tot total
r pseudo-binary solution/operator space
ref reference solution/operator space

References

Abadi, M., Agarwal, A., Barham, P., et al 2016. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. arXiv. https://doi.org/10.48550/arXiv.1603.04467.

Blechschmidt, J. and Ernst, O. G. 2021. Three ways to solve partial differential equations with neural networks — A review. *GAMM-Mitteilungen*, **44**(2). https://doi.org/10.48550/arXiv.2102.11802.

Chen, Y. and Voskov, D. 2020. Optimization of CO2 injection using multi-scale reconstruction of composition transport. *Computational Geosciences*, **24**(2): 819–835. https://doi.org/10.1007/s10596-019-09841-8.

Christie, M. A. and M. J. Blunt. Tenth SPE Comparative Solution Project: A Comparison of Upscaling Techniques. *SPE Reservoir Evaluation & Engineering*, **4**: 308–317. https://doi.org/10.2118/72469-PA.

Fuks, O. and Tchelepi, H. A. 2020. Limitations of Physics Informed Machine Learning for Nonlinear Two-Phase Transport In Porous Media. *Journal of Machine Learning for Modeling and Computing*, **1**: 19–37. 10.1615/JMachLearnModelComput.2020033905.

Ganapathy, C. and Voskov, D. 2018. Multiscale reconstruction in physics for compositional simulation. *Journal of Computational Physics*, **375** (1), 747–762. https://doi.org/10.1016/J.JCP.2018.08.046.

Lyu, X., Voskov, D., Tang, J., and Rossen, W. R. 2021. Simulation of Foam Enhanced-Oil-Recovery Processes Using Operator-Based Linearization Approach. *SPE J*, **26**(04), 2287–2304. SPE-205399-PA. https://doi.org/10.2118/205399-PA.

Orr, F. 2007. Theory of Gas Injection Processes. Holte, Denmark: Tie-Line Publications.

Raissi, M., Perdikaris, P., and Karniadakis, G. 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, **378**, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045.

Rasmussen, C. P., Krejbjerg, K., Michelsen, M. L., and Bjurstram, K. E. 2003. Increasing Computational Speed of Flash Calculations with Applications for Compositional, Transient Simulations. Presented at the SPE Annual Technical Conference and Exhibition, Denver, Colorado, October 2003. SPE-84181-MS. https://doi.org/10.2118/84181-MS.

Rastegar, R. and Jessen, K. 2009a. A Flow Based Lumping Approach for Compositional Reservoir Simulation. Presented at the SPE Reservoir Simulation Symposium, The Woodlands, Texas, USA, All Days. SPE-119160-MS. https://doi.org/10.2118/119160-MS.

Rastegar, R., and Jessen, K. 2009b. Lumping and Delumping for Integrated Compositional Modeling. Presented at the SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, October 2009. SPE-125017-MS. https://doi.org/10.2118/125017-MS.

Pan, H. and Tchelepi, H. A. 2011. Compositional Flow Simulation Using Reduced-Variables and Stability Analysis Bypassing. Presented at the SPE Reservoir Simulation Symposium, The Woodlands, Texas, USA, February 2011. SPE-142189-MS. https://doi.org/10.2118/142189-MS.

Stewart, R. and Ermon, S. 2016. Label-Free Supervision of Neural Networks with Physics and Domain Knowledge. arXiv. https://doi.org/10.48550/arXiv.1609.05566.

Tang, D. and Zick, A. 1993. A New Limited Compositional Reservoir Simulator. presented at the SPE Symposium on Reservoir Simulation, New Orleans, Louisiana, February 1993. SPE-25255-MS. https://doi.org/10.2118/25255-MS.

Voskov, D. V. 2017. Operator-based linearization approach for modeling of multiphase multi-component flow in porous media. *Journal of Computational Physics*, **337**, 275–288. https://doi.org/10.1016/j.jcp.2017.02.041.

Voskov, D. V. and Tchelepi, H. A. 2009. Compositional Space Parameterization: Theory and Application for Immiscible Displacements. *SPE J*, **14**(3), 431–440. SPE-106029-PA. https://doi.org/10.2118/106029-PA.

Voskov, D. and Tchelepi, H. A. 2008. Compositional Space Parametrization for Miscible Displacement Simulation. *Transport in Porous Media*, **75**(1), 111–128. https://doi.org/10.1007/s11242-008-9212-1.

Wang, Y., Voskov, D., Khait, M., and Bruhn, D. 2020. An efficient numerical simulator for geothermal simulation: A benchmark study. *Applied Energy*, **264**, https://doi.org/10.1016/j.apenergy.2020.114693.

Zaydullin, R., Voskov, D. V., and Tchelepi, H. A. (2013). Nonlinear Formulation Based on an Equation-of-State Free Method for Compositional Flow Simulation. *SPE J*, **18**(02), 264–273. SPE-146989-PA. https://doi.org/10.2118/146989-PA.

Appendix

Table A1—Hydrodynamic parameters.

	Oil	Gas
Porosity	*0.3*	*0.3*
Residual saturation	*0.0*	*0.0*
End points k_r (s_i)	*1.0*	*1.0*
Saturation exponents	*2.0*	*2.0*
Viscosity *(cP)*	*0.5*	*0.1*
Density *(kg/m³)*	600	200
Compressibility *(1/bar)*	*10^(-5)*	*10^(-3)*

Table A2—Molar weights of each component in g/mol

C1	*CO_2*	*nC4*	*nC10*
16.043	*44.010*	*58.123*	*114.520*