Adaptive POD model reduction for solute transport in heterogeneous porous media

Calogero B. Rizzo · Felipe P. J. de Barros · Simona Perotto · Luca Oldani · Alberto Guadagnini

Abstract We study the applicability of a model order reduction technique to the solution of transport of passive scalars in homogeneous and heterogeneous porous media. Transport dynamics are modeled through the advection-dispersion equation (ADE) and we employ Proper Orthogonal Decomposition (POD) as a strategy to reduce the computational burden associated with the numerical solution of the ADE. Our application of POD relies on solving the governing ADE for selected times, termed snapshots. The latter are then employed to achieve the desired model order reduction. We introduce a new technique, termed Snapshot Splitting Technique (SST), which allows enriching the dimension of the POD subspace and damping the temporal increase of the modeling error. Coupling SST with a modeling strategy based on alternating over diverse time scales the solution of the full numerical transport model to its reduced counterpart allows extending the benefit of POD over a prolonged temporal window so that the salient features of the process can be captured at a reduced computational cost. The selection of the time scales across which the solution of the full and reduced model are alternated is linked to the Péclet number (Pe), representing the interplay between advective and dispersive processes taking place in the system. Thus, the method is adaptive in space and time across the heterogenous structure of the domain through the combined use of POD and SST and by way of alternating the solution of the full and reduced models. We find that the width of the time scale within which the POD-based reduced model solution provides accurate results tends to increase with decreasing Pe. This suggests that the effects of local scale dispersive processes facilitate the POD method to capture the salient features of the system dynamics embedded in the selected snapshots. Since the dimension of the reduced model is much lower than that of the full numerical model, the methodology we propose enables one to accurately simulate transport at a markedly reduced computational cost.

Keywords Proper Orthogonal Decomposition · Model Reduction · Heterogeneous Porous Media · Flow and Transport · Computational Efficiency

C. B. Rizzo · F. P. J. de Barros
Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, California, USA

S. Perotto · L. Oldani
MOX, Dipartimento di Matematica, Politecnico di Milano, Milano, Italy

A. Guadagnini
Dipartimento di Ingegneria Civile e Ambientale, Politecnico di Milano, Milano, Italy

1 Introduction

Improving our ability to provide reliable and computationally efficient approaches to quantify transport process in porous media is of major importance to a wide range of applications. Most notably, these include the characterization of the feedback between anthropogenic activities and the subsurface environment, with direct implications on the assessment of water quality which is key to physically based development of modern strategies addressing the water-energy-food nexus.

Analysis of flow and transport phenomena in porous media often involves investigations within domains of large extent [42,43]. Practical difficulties associated with the way one can rigorously include pore-scale modeling in the analysis of typical laboratory and field-scale settings lead to depicting of solute transport in such media

through effective models [14, 27, 38, 39, 45]. Several alternative modeling options are available in this context, a common choice being a continuum-based representation grounded on the standard advection-dispersion equation (ADE) on which we focus here. Evolution of transport scenarios of practical interest to industrial and environmental applications typically occurs over temporal scales associated with the duration of several days, months and sometimes years. High quality numerical approximations are required to capture extreme values of chemical concentrations, i.e., peak values and low concentrations, constituting key environmental performance metrics with severe implications on human health through the proper quantification of the exceedance of toxicity thresholds at sensitive locations in the system [3]. Accurate numerical solution of the ADE typically requires considerable computational time and the use of appropriate discretization techniques [19]. Computational time requirements become a major challenge when several transport scenarios needs to be evaluated, as is typical in modern probabilistic risk assessment approaches [34]. It then becomes relevant to develop methods to decrease the complexity of the discrete model associated with the governing process equations [15, 16], while preserving accuracy of the solution.

Here, we focus on the analysis of the Proper Orthogonal Decomposition (POD) technique [30, 40] and explore its ability to reduce the computational burden associated with the solution of the ADE under a variety of transport conditions characterized by diverse relative strengths of advective and dispersive processes. POD is one of the most widely used model order reduction techniques and has been applied to a wide range of problems in the fields of data analysis, statistics and/or dynamic systems [12, 23, 24, 32]. The approach has been recently employed for the solution of fully saturated flow and transport in subsurface reservoirs [25, 26, 28, 31] and it has been shown to have the potential to considerably decrease the discrete problem size while maintaining a relatively high accuracy with respect to the solution of the original governing equation. It has also been employed in the context of preliminary studies addressing uncertainty quantification of flow in groundwater systems to speed up numerical Monte Carlo simulations of flow in the presence of random forcing and system parameters [35, 36].

The key feature of model reduction techniques is that the solution of a given partial differential equation (PDE) can be well approximated by a linear combination of a limited number of basis functions which can be conveniently selected. The coefficients of this linear combination can then be computed through the

solution of the reduced system obtained via a Galerkin projection of the governing PDE. In the POD framework, the basis functions identifying the finite subspace where the model is projected are the eigenfunctions of an integral operator whose kernel is given by the spatial correlation function of the state variable of interest. A critical feature of the application of the POD technique to dynamically evolving scalar fields of the kind described by an ADE is related to the way this finite subspace is built. This is typically accomplished through (a) solving the governing PDE at a set of predefined time steps termed snapshots, (b) applying Principal Component Analysis (PCA) to the snapshots, and (c) employing the resulting principal components as basis functions to characterize the above mentioned finite subspace [11]. The proper selection of the snapshots is a critical point in the workflow, because it controls the errors associated with the model reduction. Examples of studies providing guidelines for snapshot selection can be found in application related to groundwater [44] and two-phase [10, 18] flows, mass tranfer in heterogeneous environments [21, 29] as well as oceanic flows [9]. The problem is particularly challenging in the presence of diverse competing physical processes driving the evolution of the system, such as in the case of the ADE where the space-time distribution of dissolved chemical concentration is governed by the relative importance of advective and dispersive/diffusive processes. We investigate the manner in which the interplay between these processes impacts the effectiveness of POD-based model reduction techniques for the computationally efficient and accurate numerical solution of the ADE in homogeneous and heterogeneous porous media under typical subsurface environmental conditions.

In general, the snapshots are extracted from one or multiple full simulations. Here we build the POD projection matrix by alternating between the solution of the full numerical transport model and its reduced counterpart. Switching from full to reduced model (or vice versa) is driven by the residual error associated with the reduced problem. This enables us to control the magnitude of the modeling error through a user-defined tolerance. In addition, we introduce a new approach, termed *Snapshot Splitting Technique* (SST), to guide the selection of the snapshots. This technique leads to an increased capability of the reduced model to provide accurate approximations for a sustained period of time.

The structure of this work is as follows. Section 2 includes a brief presentation of the ADE model we consider, the POD method and the mixed Finite Element-based POD method (FE-POD) we employ. In section 3 we illustrate the use of the mixed FE-POD method

to the solution of solute transport in homogeneous and heterogeneous media. Section 4 is devoted to the presentation and discussion of our model reduction strategy and algorithm in the presence of advective and diffusive/dispersive processes of competing strength. Section 5 includes our conclusions and outlines of future developments.

2 Reduction of the ADE Through Proper Orthogonal Decomposition

We introduce here the advection-dispersion equation (ADE), which we employ to describe solute transport in a porous medium at the continuum scale, and illustrate the theoretical basis and workflow for the application of the POD for the model reduction. We perform our analysis by relying on a dimensionless formulation which enables us to discriminate through the Péclet number (Pe) the relative importance of advection and diffusion/dispersion processes on the effectiveness of the POD approach to model reduction.

2.1 Dimensionless problem formulation

The general format of the n-dimensional ADE employed to depict the evolution of chemical concentration following a continuous point injection within a Darcy-scale velocity field $\mathbf{u}$ is:

$$\frac{\partial c^*}{\partial t^*} - \nabla \cdot (\mathsf{D}\nabla c^*) + \mathbf{u} \cdot \nabla c^* = c_{\text{in}}^* \delta\left(\mathbf{0}, t^*\right), \quad (1)$$

where $c^* : \mathbb{R}^n \to \mathbb{R}$ is the solute concentration, D is a tensor embedding the effects of diffusive and local-scale dispersive processes, c_{in}^* is the solute concentration injected per unit time, δ is the Dirac delta function and t^* is the time. We consider a steady-state velocity field $\mathbf{u}$ given by Darcy's law:

$$\mathbf{u} = -\frac{\mathsf{K}}{\phi}\nabla h. \quad (2)$$

Here, ϕ and K are the porosity and the hydraulic conductivity tensor of the porous medium, respectively. The steady-state spatial distribution of hydraulic head h is governed by:

$$\nabla \cdot [\mathsf{K}\nabla h] = 0 \quad (3)$$

equipped with proper boundary conditions.

In a two-dimensional system (i.e., $n = 2$) with x and y denoting principal directions, we have:

$$\mathsf{D} = \begin{bmatrix} D_x & 0 \\ 0 & D_y \end{bmatrix}, \ \mathbf{u} = \begin{bmatrix} u_x \\ u_y \end{bmatrix}, \quad (4)$$

where D_x and D_y are constant. Here, we consider the average flow direction to be aligned along the x-axis and define the Péclet number as:

$$Pe = \frac{\langle u_x \rangle L}{D_x}, \quad (5)$$

where $\langle u_x \rangle$ is the average value of u_x, the component of $\mathbf{u}$ along x, and L is a characteristic length scale of the system. Next, we introduce the following dimensionless quantities:

$$x = \frac{x^*}{L}, \quad y = \frac{y^*}{L}, \ c = \frac{c^* u_x}{c_{\text{in}}^* L},$$

$$t = \frac{t^* u_x}{L}, \mathcal{D} = \frac{D_y}{D_x}, \mathbf{v} = \frac{\mathbf{u}}{\langle u_x \rangle}. \quad (6)$$

The above dimensionless quantities allow recasting (1) as:

$$\frac{\partial c}{\partial t} - \frac{1}{Pe}\left(\frac{\partial^2 c}{\partial x^2} + \mathcal{D}\frac{\partial^2 c}{\partial y^2}\right) + v_x\frac{\partial c}{\partial x} + v_y\frac{\partial c}{\partial y} = \delta\left(\mathbf{0}, t\right). \quad (7)$$

Finally, we present the algebraic formulation of problem (7) using a finite element (FE) method and define the solution by $c_f(\cdot)$ and the corresponding vector by $\mathbf{c}_f$. We can formulate the *full discrete problem* as follows:
For each $k \geq k_0$ find $\mathbf{c}_f^k \in \mathbb{R}^{n_f}$ such that

$$\left(\frac{\mathsf{M}}{\Delta t} + \mathsf{A}\right)\mathbf{c}_f^{k+1} = \mathbf{f}^{k+1} + \frac{\mathsf{M}}{\Delta t}\mathbf{c}_f^k, \quad (8)$$

with $\mathbf{c}_f^k$ the solution at time $t_k = k\Delta t$ and with initial condition $\mathbf{c}_f^{k_0} = \mathbf{c}_{f,0}^{k_0}$. The corresponding discrete problem is characterized by n_f degrees of freedom. The matrix M is the mass matrix, the matrix A is the stiffness matrix which is linked to $\mathcal{D}$, Pe and $\mathbf{v}$, $\mathbf{f}$ is a known vector and Δt is the time step.

2.2 Proper Orthogonal Decomposition method

The basic idea of Proper Orthogonal Decomposition (POD) is to reduce the original problem into a new simpler problem where the solution lies in a subspace of reduced dimensionality, as compared to the space where the full model is defined. The construction of such a subspace is a challenging issue in the model development. A commonly adopted technique relies on solving the full model for a given number of time intervals, usually termed snapshots and then relying on the information embedded in them to build the desired subspace. In this context, our strategy starts from the computation of the snapshots through the solution of the full model (7) from the initial system state up to a

given simulation time t_0. A key point in this approach is then to select the snapshots and t_0 in a way that the salient features of the early system dynamics are adequately captured. These are then employed to advance the solution of the ensuing reduced model beyond t_0, as the concentration further progresses in the domain. In this sense, the construction of the solution subspace can be performed by establishing a suitable relationship between a proper space V_{n_f} in which the solution of the full model lies and where each element can be represented by a n_f-dimension vector, and a reduced subspace V_{n_p} with dimension $n_p \ll n_f$. Given a vector $\mathbf{x} \in \mathbb{R}^{n_f}$ that represents an element of V_{n_f}, one can always find its counterpart $\hat{\mathbf{x}} \in \mathbb{R}^{n_p}$ as:

$$\hat{\mathbf{x}} = \mathsf{P}^T \mathbf{x}, \tag{9}$$

where the columns of the n_f-by-n_p projection matrix P are an orthonormal basis spanning the subspace V_{n_p}.

From an operational standpoint, computation of P requires constructing the n_p basis vectors that span the subspace V_{n_p}. We do so by selecting n_p snapshots $(\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_{n_p})$ where $\mathbf{c}_i \in \mathbb{R}^{n_f}$ and defining a matrix:

$$\mathsf{X} = \begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \cdots & \mathbf{s}_{n_p} \end{bmatrix}, \tag{10}$$

where $\mathbf{s}_i$ is the i-th snapshot $\mathbf{c}_i$ normalized with respect to the Euclidean norm. We define:

$$\mathsf{C}_s = \mathsf{X}^T \mathsf{X}, \tag{11}$$

where C_s is an n_p-by-n_p symmetric matrix. Using the spectral theorem, we can rewrite C_s as:

$$\mathsf{C}_s = \mathsf{G}\Lambda\mathsf{G}^T, \tag{12}$$

where G is the orthogonal matrix of the eigenvectors of C_s and Λ is a diagonal matrix whose entries are the eigenvalues of C_s. Since the dimension of C_s is usually small, the computation of matrices G and Λ requires a small amount of time compared to the requirement for the solution of the full numerical model.

By exploiting the well-known relation between the spectral decomposition in (12) and the singular value decomposition of X [25, 26] we obtain that:

$$\mathsf{P} = \mathsf{X}\mathsf{G}\Lambda^{-1/2}. \tag{13}$$

so that $\mathsf{P}^T\mathsf{P} = \mathsf{I}$ with I the identity matrix of order n_p.

The matrix P contains all information about the relationship between the two spaces V_{n_f} and V_{n_p} and the columns of the projection matrix $\tilde{\mathbf{p}}_i$ are a set of basis vectors of V_{n_p}. Note that in general:

$$\mathsf{P}\mathsf{P}^T \neq \mathsf{I}, \tag{14}$$

where the identity matrix I is here of order n_f. Equation (14) is key to characterize the POD dynamics. A

vector $\mathbf{v} \in V_{n_f}$ can be split as $\mathbf{v} = \tilde{\mathbf{v}} + \mathbf{v}_0$, where $\tilde{\mathbf{v}}$ and $\mathbf{v}_0$ belong to the subspace generated by P and to the associated orthogonal subspace, respectively. Therefore $\mathbf{v}_0 \cdot \boldsymbol{\phi} = 0$ for every $\boldsymbol{\phi}$ linear combination of the POD basis vectors. Using the properties of the projection matrix P, we can write:

$$\tilde{\mathbf{v}} = \mathsf{P}\mathsf{P}^T \mathbf{v}. \tag{15}$$

The matrix $\mathsf{P}\mathsf{P}^T$ acts as a filter extracting the component of a function belonging to the POD subspace. We note that both vectors $\tilde{\mathbf{v}}$ and $\mathbf{v}$ have a dimension equal to n_f. In other words, increasing the process dynamics captured by the POD subspace (e.g., through an appropriate selection of a given number of snapshots or by increasing the number of snapshots) leads to a decreased error induced by the model reduction. Given the dynamics of the system considered, it is then relevant to derive a flexible formulation which enables us to increase the dimension of the POD subspace by relying on a fixed number of snapshots, i.e., on a given computational effort. We do so in the following section and then assess the benefit of such a formulation by a suite of computational examples.

In practice, the POD based reduced model is obtained by projecting equations (8) over the subspace induced by the matrix P, i.e., the subspace containing all the possible linear combinations of the columns of P. We introduce the solution of the reduced problem $c_p(\cdot)$ with its corresponding vector $\mathbf{c}_p$. The *reduced problem* using the POD method reads:
For each $k \geq k_0$ find $\mathbf{c}_p^k \in \mathbb{R}^{n_p}$ such that

$$\mathsf{P}^T \left(\frac{\mathsf{M}}{\Delta t} + \mathsf{A} \right) \mathsf{P}\, \mathbf{c}_p^{k+1} = \mathsf{P}^T \mathbf{f}^{k+1} + \mathsf{P}^T \left(\frac{\mathsf{M}}{\Delta t} \right) \mathsf{P}\, \mathbf{c}_p^k, \tag{16}$$

with $\mathbf{c}_p^{k_0} = \mathsf{P}^T \mathbf{c}_{f,0}^{k_0}$ the projection of the initial value.

2.3 Snapshot Splitting Technique (SST) for increasing the POD subspace dimension

Next, we describe a new approach to improve the classic POD method. Given n_p snapshots $(\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_{n_p})$ where $\mathbf{c}_i \in \mathbb{R}^{n_f}$, for each snapshot $\mathbf{c}_i$ we define n_l new snapshots $\mathbf{k}_{ij}$ such that:

$$\mathbf{k}_{ij} = g_j(\mathbf{c}_i), \tag{17}$$

where $g_j(\cdot) : \mathbb{R}^{n_f} \to \mathbb{R}^{n_f}$ for $i = 1, \cdots, n_p$ and $j = 1, \cdots, n_l$. If the vectors associated with each snapshot are independent, then n_l basis vectors are obtained for each snapshot, i.e., the POD subspace will have a dimension of $n_p \times n_l$. In such a case, the columns of matrix (10) are the vectors $\mathbf{k}_{ij} \in \mathbb{R}^{n_f}$ normalized with respect
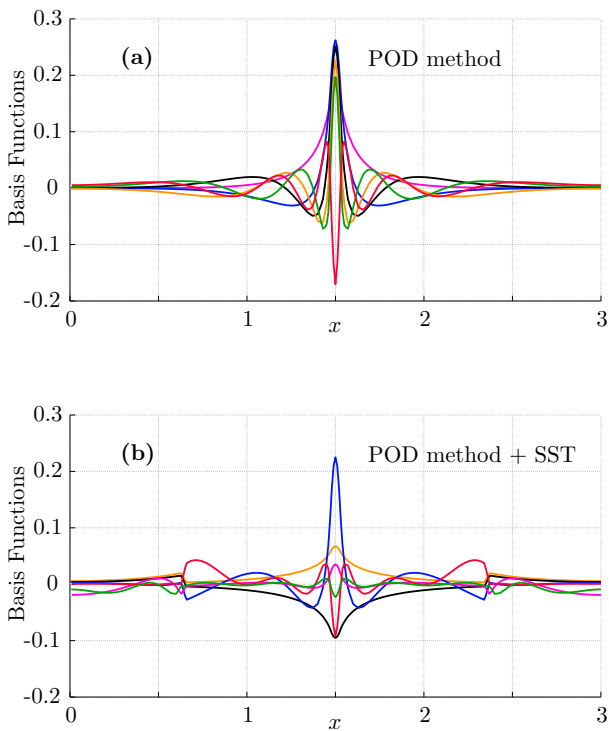
Fig. 1: Distribution along a longitudinal profile (i.e., aligned in the x direction) of the basis functions $\tilde{\mathbf{p}}_i$ ($i = 1, \cdots, 7$) obtained (a) without and (b) with the *Snapshot Splitting Technique* (SST) for a purely diffusive setting. The solute source is located at $x = 1.5$.

to the Euclidean norm. The POD method can then be applied as illustrated in 2.2.

The strategy for the selection of the functions $g_j(\cdot)$ is not unique, as long as the generated vectors are independent. The key concept we employ here for the definition of $g_j(\cdot)$ is that the domain can be partitioned into areas where the solution dynamics are similar. The logarithmic sampling is a natural choice considering that most of the available analytical solutions of the ADE are expressed in terms of either the exponential function or the complementary error function [20]. A logarithmic distribution has the additional benefit of uniformly sampling the range of relevant concentration values, without disregarding significant orders of magnitude relevant to industrial and environmental applications, where the detection of minute chemical concentrations can be critical. We define the maximum coefficient of the snapshots set as:

$$c_{\max} = \max_i \max_j c_{ij}, \qquad (18)$$

where $i = 1, \cdots, n_p$ and c_{ij} is the j-th component of the vector $\mathbf{c}_i$. We partition $[0, c_{\max}]$ into n_l intervals

introducing $n_l + 1$ values ξ_j with $j = 0, \cdots, n_l$ such that:

$$0 = \xi_0 < \xi_1 < \xi_2 < \cdots < \xi_{n_l - 1} < \xi_{n_l} = c_{\max}. \qquad (19)$$

Finally, we define the functions $g_j(\cdot)$ such that:

$$[g_j(\mathbf{v})]_m = \begin{cases} v_m & \text{if } v_m < \xi_j, \\ 0 & \text{otherwise,} \end{cases} \qquad (20)$$

where $j = 1, \cdots, n_l$, $m = 1, \cdots, n_f$ and v_m is the m-th component of the vector $\mathbf{v} \in \mathbb{R}^{n_f}$. Thus, for each j, the modified snapshots $\mathbf{k}_{ij}$ will be different from zero in areas where the concentration is below ξ_j. In our application, we select a logarithmic distribution for ξ_j allows us to split the snapshots so that the area covered by each $\mathbf{k}_{ij}$ is similar for $j = 1, \cdots, n_l$. Since we are studying the case of continuous injection, the value $c_{\max}$ occurs at the source location. A logarithmic distribution of ξ_i takes into account that the fast solution dynamics are close to the source. The same strategy is followed, for example, in the analysis of tracer tests and of pumping tests (e.g., see [7]). As a consequence, our technique allows enriching the POD subspace in these critical zones of the evolving solute concentration. A priori, we may adopt other distributions for ξ_i, e.g. a Gaussian one, by ensuring that it is able to equally split the modified snapshots $\mathbf{k}_{ij}$ using equation (20). We have found the logarithmic distribution to be effective for the cases studied in this paper. However, one may be interested in modifying this choice to adapt the SST method to other scenarios.

To highlight the differences between the way the POD subspace is constructed by employing the typical approach illustrated in Section 2.2 and our proposed technique to correct such construction by relying on snapshots defined in (17), we term our approach as *Snapshot Splitting Technique* (SST) and exemplify its beneficial effect by considering a purely diffusive problem (i.e., $\mathbf{u} = \mathbf{0}$ in (1)) with a continuous point injection located at $x = y = 1.5$ in a 3×3 rectangular domain representing a homogeneous porous medium (see also Section 3 for details about the numerical solution of the full model). Figure 1 depicts the distribution along a longitudinal profile (i.e., aligned in the x direction) of the basis functions obtained without (Figure 1a) and with (Figure 1b) the application of the SST. In this illustrative example we select the snapshots $\mathbf{c}_i$ from the first 10 time steps within which the full model is solved. Using the approach illustrated in Section 2.2, the results clearly show that all basis functions are significantly different from zero solely at locations very close to the source. This implies that (a) employing the POD subspace generated via these basis functions does not lead to an accurate representation of the system dynamics

at locations far away from the source, so that (b) the ensuing reduced model can propagate in time only the information associated with system states close to the source. Otherwise, application of the SST to increase the dimension of the POD subspace (while relying on a fixed number of snapshots) allows constructing a set of basis functions which carry significant information at a variety of locations in the domain.

3 Application of POD to Advective-Dispersive Transport driven by Continuous Solute Injection

In this section we apply the methods discussed in Section 2 to reduce the dimensionality of a transport problem described by equation (7). We start by considering a homogeneous porous medium where a uniform velocity field ($v_x = 1$ and $v_y = 0$) takes place. We employ a Finite Element (FE) approach ($\mathbb{P}1$ elements) for spatial discretization and the Backward Euler scheme for time discretization with discretization step $\Delta t = 0.05$. We recall that while we ground our examples on FE, POD reduction can also be applied in the presence of diverse numerical methods, e.g., Finite Volumes. The computational domain is a 5×3 rectangle with a structured triangular mesh. Each length unit is divided into 35 segments, the computational grid being then composed by 36750 triangles. The number of degrees of freedom of the $\mathbb{P}1$ space with open boundary conditions coincides with the number of nodes of the numerical grid, i.e., 18656. A continuous point injection source is placed at $x = y = 1.5$ in all settings we examine here. When the SST illustrated in Section 2.3 is employed for the construction of the POD subspace, we choose $n_l = 10$ with a logarithmic distribution for the ξ_j as explained in 2.3.

In this study, we set $\mathcal{D} = 1$. We explore the effect of the Péclet number (Pe) on the behavior of the reduced model, as compared against the full numerical solution $\mathbf{c}_f^k$ of the model (8). We use the following procedure for the computation of $\mathbf{c}_p^k$ at time step k:

(**S1**) For the first N_t time steps ($k \leq N_t$), $\mathbf{c}_p^k$ is computed by solving the full numerical model (8), so $\mathbf{c}_p^k = \mathbf{c}_f^k$.

(**S2**) We compute the projection matrix (13) via the POD method upon relying on the solution computed at step **S1**. The snapshots are taken as $\mathbf{c}_f^k$ at time step $k \leq N_t$.

(**S3**) Using the projection matrix and the discrete problem defined in (16), we compute $\mathbf{c}_p^k$ for additional N_a time steps.

In all our computational examples we consider $N_t = 10$ and $N_a = 90$. It is important to notice that N_t and hence the size of the reduced linear system, $n_p = 10$, is

considerably smaller than the size of the system associated with the full model n_f (i.e., 18656 in this case). The computational time associated with each step of the reduced model is negligible when compared to the time demanded by a step of the full model. The influence of N_t and N_a on the optimization of the procedure is case dependent and we leave this analysis to future studies.

We perform computations with the reduced model in the absence or in the presence of the SST introduced in Section 2.3. When the SST is employed during step **S2**, the solution will be denoted as c_{p+SST}^k or $\mathbf{c}_{p+SST}^k$. Note that in case of SST the reduced model dimension is equal to $n_p \times n_l = 100$ which is still much smaller than the full numerical model dimension n_f. All results in this study have been generated using Matlab (Windows 10, Intel(R) Core(TM) i7-6700 CPU @ 340GHz 3.41 GHz). A single step using the full linear system (i.e., Equation (8)) requires a computation time $t_f = 0.068s$ on average, a single step using the reduced linear system (i.e., Equation (16)) only requiring $t_r = 0.00025s$. Two orders of magnitude separate the full from the reduced model at each time step, so that the run time of the reduced problem is virtually negligible, as compared with the one characterizing the full system.

The quality of the ensuing solutions obtained with the reduced model is compared by computing the relative error E_s^k between the full numerical problem solution c_f and the reduced problem solutions, i.e., c_p or c_{p+SST}, which is computed for each time step k as:

$$E_s^k = \frac{\left\| c_f^k - c_s^k \right\|_{L^2(\Omega)}}{\left\| c_f^k \right\|_{L^2(\Omega)}}, \tag{21}$$

where subscript $s = p+SST$ or p depending on whether the SST is employed or not in the construction of the reduced model.

Figure 2 depicts the relative error (21) as a function of the number of time steps k and for three selected Péclet numbers, i.e., $Pe = 30, 60, 90$, identifying mildly to highly advective transport settings. The relative error increases in a nonlinear fashion with Pe and with the time elapsed since the last snapshot is computed from the full numerical model ($k = 10$ in the example). The plateau to which the relative errors tends is due to the open boundary conditions that allow solute mass to exit the domain at the downstream boundary. Figure 2 reveals that employing the proposed SST enables us to increase the effectiveness of the information content embedded in the snapshots extracted from the solution of the full model, this leading to a remarkable reduction in the associated relative error. For instance, we
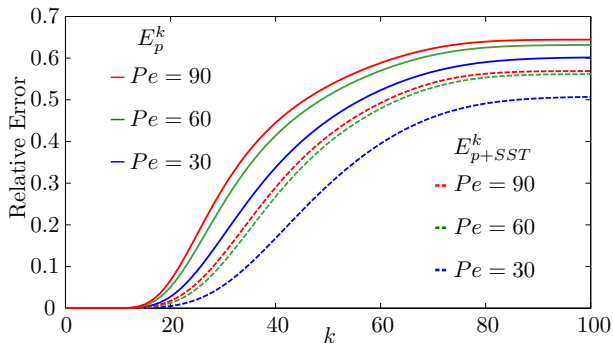
Fig. 2: Effect of the *Snapshot Splitting Technique* (SST) on the evolution of the relative error (21) with the number of time steps, k, after the reduced model is employed (full model is run for $k \leq 10$) for a homogeneous system with diverse values of Pe.
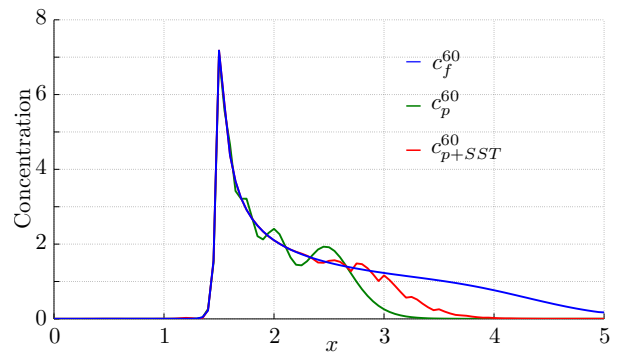

Fig. 3: Spatial distributions of solute concentration computed by the full model and the reduced model with and without activating the *Snapshot Splitting Technique* (SST) at simulation time step $k = 60$. Results are depicted for the homogeneous test setting along the longitudinal cross-section passing through the injection point for $Pe = 30$.

observe that for $Pe = 30$ the relative error is approximately 10% after roughly 35 or 25 simulation steps, resorting or not resorting to SST, respectively. Thus, SST yields an improved accuracy of the reduced model for the same number of full model snapshots.

As a complement to Figure 2, Figure 3 depicts the spatial distributions of solute concentration c_f^k, c_p^k and c_{p+SST}^k along the longitudinal cross-section at $y = 1.5$ (i.e., passing through the injection point) for $Pe = 30$ and time step $k = 60$, i.e., after 50 time steps from the time at which the latest snapshot from the full model has been taken. Notably, the reduced solution associated with the SST, c_{p+SST}^k, displays much less severe spatial fluctuations than its counterpart c_p^k. This result confirms the beneficial effect of the proposed SST to the reduced model even after a considerable time has elapsed since the observation time associated with the last snapshot computed via the full model.

Figures 4a, 4b and 4c respectively depict the spatial distributions of solute concentration obtained by the full model solution and the POD approximation without and with the use of SST. Numerical results are illustrated for the observation time corresponding to $k = 60$ at which the relative error is high and the approximation is not so reliable in both cases. Even as severe spatial fluctuations can be observed, it is clear that employing the SST tends to drive the solution of the reduced model towards the behavior displayed by the true system dynamics, as represented more faithfully by the full model solution. In particular, the application of the SST renders an approximation of improved quality in a larger area around the point source, when compared against the traditional application of the POD method.

As an additional and challenging test case, we then consider the performance of our reduced order modeling strategy to model chemical transport in a heterogeneous porous medium. In this context, modern probabilistic approaches to environmental risk quantification [4,6,22,46] rely on the assessment of the probability that concentration of a given chemical species at a target environmental location exceeds a given threshold value, which is typically linked to human health or to environmental constraints. Other environmental performance metrics of interest in the context of the environmental risk assessment associated with solute migration in the subsurface include, e.g., peak values of concentrations, and/or early/late times of arrival of solute at sensitive locations in the aquifer [3]. The stochastic nature of the groundwater flow and transport problems is typically linked to the spatially heterogeneous distribution of aquifer hydraulic properties, such as hydraulic conductivity K. The latter is typically conceptualized as a stochastic process of space and the flow and transport equations become stochastic differential equations [42,47]. In this sense, evaluation of the probability linked with the above described environmental metrics is typically performed in a numerical Monte Carlo framework. The Monte Carlo method is essentially based on (a) the synthetic generation of multiple realizations of spatial distributions of hydraulic conductivity fields, honoring certain statistical properties (which are typically estimated from available data), and (b) the solution of the flow and transport problem for each of these replicates. Since the accurate assessment of target probability values typically requires
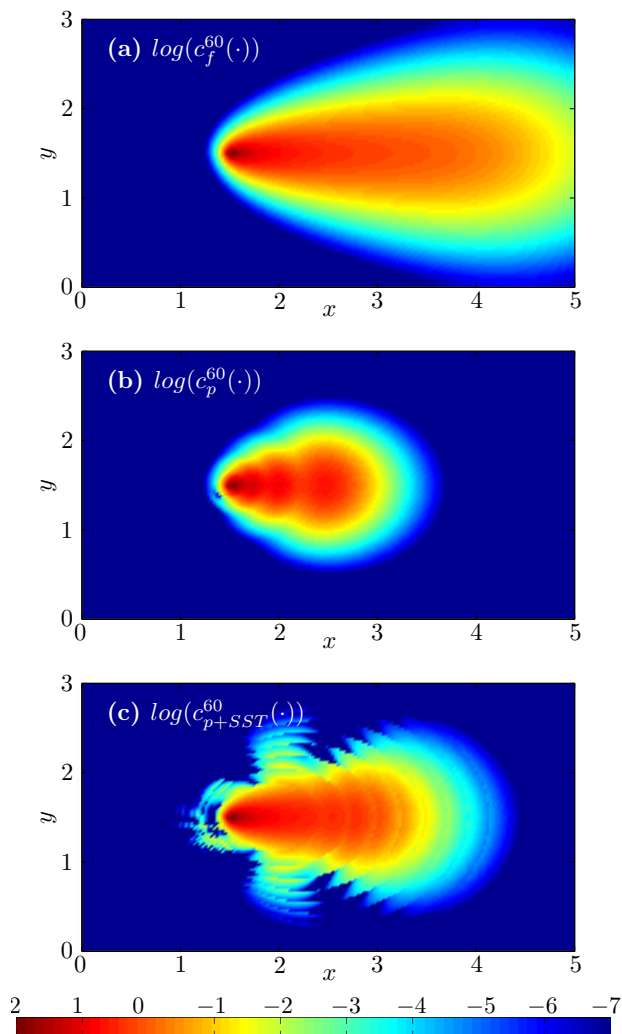
Fig. 4: Spatial distributions of solute concentration obtained by (a) the full model solution and the POD approximation (b) without and (c) with the use of the *Snapshot Splitting Technique* (SST). Results correspond to the longitudinal profiles depicted in Figure 3 and refer to $k = 60$. The relative error is already markedly high and clearly highlights the need of updating the POD subspace. Colors are associated with a logarithmic scale.

performing a large number of Monte Carlo simulations [1, 34, 35], having at our disposal a computationally efficient methodology for the solution of transport is of remarkable interest. Consistent with the classical approach to stochastic groundwater hydrology [13, 42], we model the (natural) logarithm of hydraulic conductivity, $Y = \log K$, as a second-order stationary correlated Gaussian random process of space. We then generate a synthetic random realization of a statistically isotropic field Y. The latter is characterized by a zero

mean and unit variance, and is associated with an exponential isotropic covariance model with correlation scale $\lambda = 2L$. Generation of Y is performed through the widely tested SGeMS software [41].

In practical cases, the Péclet number is computed using the K correlation scale as characteristic length. Therefore, we introduce the effective Péclet number $\widehat{Pe}$. The relation between the effective Péclet number $\widehat{Pe}$ and the Péclet number defined in (5) is:

$$\widehat{Pe} = \frac{\langle \mathbf{u} \rangle \lambda}{D} \simeq 2Pe, \tag{22}$$

where $\langle \mathbf{u} \rangle$ is the average velocity magnitude over the domain and D is the diffusion coefficient of the system (in this case $D = D_x = D_y$).

The steady-state Darcy velocity field $\mathbf{u}$ is obtained by solving the governing equations (2) and (3) upon setting permeameter-like boundary conditions to the system, corresponding to a unit head drop across the domain.

Following the procedure used for the homogeneous case, the transport problem is solved by employing the FE method ($\mathbb{P}1$ elements) for space discretization and the Backward Euler scheme for time discretization with $\Delta t = 0.05$. The spatial distributions of c_f^k, c_p^k are then calculated for the same value of Pe considered in the homogeneous setting, corresponding to $\widehat{Pe} = 60, 120, 180$. Figure 5 depicts the temporal evolution of the relative error (21), suggesting that, while the overall quality of the reduced model solution deteriorates with time, it does so to a lesser extent than the corresponding solution associated with the homogeneous set-up (see also Figure 2). This result might be related to the observation that the solute plume tends to follow the high conductivity paths in the system [17] and the effect of these paths close to the source tend to be influential to the plume behavior over time [5]. As such, capturing these features by the POD techniques enables the reduced solution to mimic the full model for longer time than in the homogeneous setting, where no preferential conductivity path is present. The beneficial effects of the SST for the enrichment of information embedded in the POD subspace is clear also in the presence of heterogeneity (Figure 5). This notwithstanding, it can be noted that one cannot propagate the reduced model solution in time indefinitely, without incurring in significant approximation errors. This is clearly observed in Figure 6, where the spatial distribution of c_f^k, c_p^k and c_{p+SST}^k are depicted at time step $k = 60$ for $Pe = 30$.
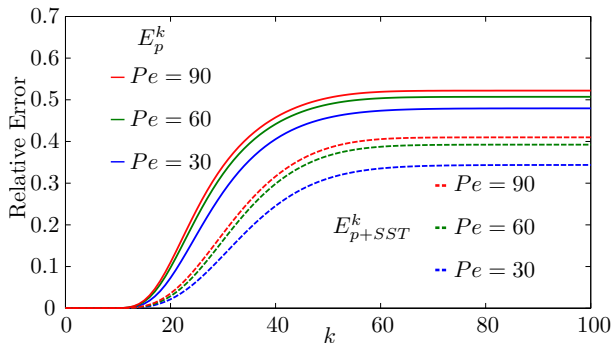
Fig. 5: Effect of the *Snapshot Splitting Technique* (SST) on the evolution of the relative error (21) with the number of time steps, k, after the reduced model is employed (full model is run for $k \leq 10$) for a heterogeneous system with diverse values of Pe.

4 Updating of the Projection Subspace

The results of Section 3 clearly show that the error between the reduced and full model solutions tends to increase with time after the last snapshot has been observed and tends to increase until it becomes not negligible after a certain time has elapsed. This is due to the fact that the information content in the generated POD subspace (even by relying on SST) is not large enough to capture all of the key details of the dynamically evolving concentration field on the basis of a limited number of snapshots. As such, the quality of the approximation tends to deteriorate as time advances since the solution significantly changes compared to the early time states. The strategy we follow here, in order to make optimal use of the POD model reduction strategy in these types of dynamically evolving systems, is based on alternating between the full model FE and POD-FE solution of the transport problem. By doing so, the POD subspace is dynamically updated over discrete temporal intervals where snapshots are extracted from the FE full problem solution and processed to construct the reduced space basis. Examples of applications of model updating strategies can be found in [8, 29].

For convenience, and given the results of Section 3, we illustrate our findings by relying on the projection matrix (13) constructed through the SST introduced in 2.3. Figure 7 depicts the temporal evolution of the error (21) obtained by solving the full model in the heterogeneous setting previously described for a fixed number $N_t = 10$ of time steps and alternating the use of the full model with the reduced POD-FE solution for diverse numbers of iterations, i.e., $n_{POD} = 10, 20, 30$, and for $Pe = 30, 60, 90$. It can be noted that alternating
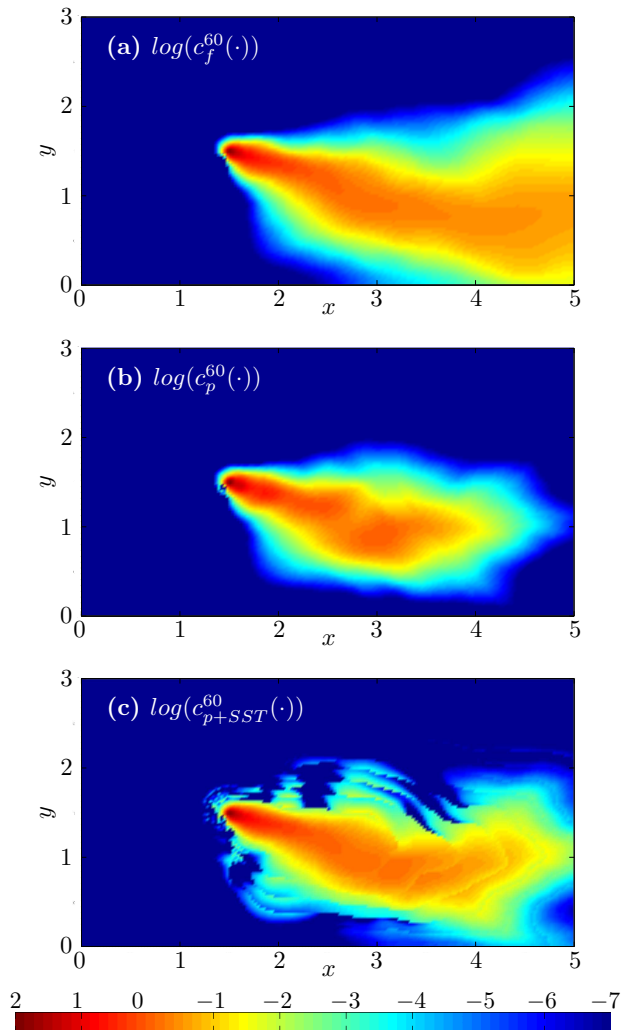

Fig. 6: Spatial distributions of solute concentration obtained by (a) the full model solution and the POD approximation (b) without and (c) with the use of the *Snapshot Splitting Technique* (SST) and refer to $k = 60$. The relative error is already markedly high and clearly highlights the need of updating the POD subspace. Colors are associated with a logarithmic scale.

between the two solutions enables one to update the POD basis in time so that the updated POD subspace can follow closely the system dynamics. The global error is seen to consistently decrease after each full problem iteration cycle. This highlights the need for an update in time of the POD basis to obtain a reduced model yielding a good accuracy and suggests that there is a beneficial cumulative effect to the quality of the reduced model solution of subsequent basis updates.

The results shown in Figure 7 suggest that setting the number of iterations n_{POD} a priori does not allow a good adaptation of the POD algorithm to the system
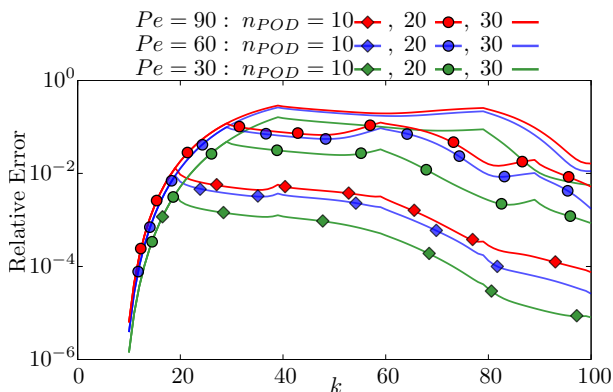
Fig. 7: Temporal evolution of the relative error E^k (21) for diverse values of Pe and employing a subspace update after a fixed interval n_{POD}.
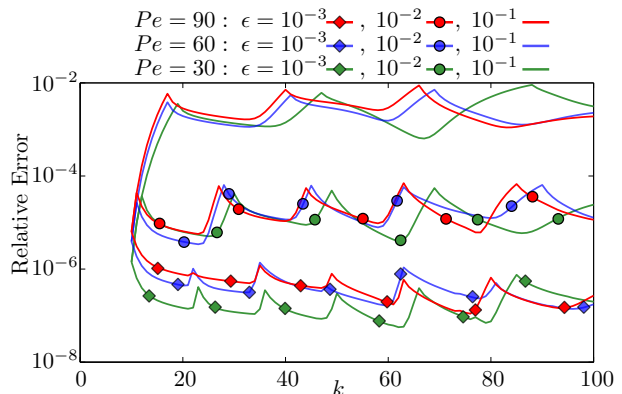

Fig. 8: Temporal evolution of the relative error E^k (21) for diverse values of Pe and imposing a different threshold on the residual $\mathcal{R}_k$ (23). Peaks of the error identify the times when the POD subspace is updated. The total number of reduced steps is 60, 40, 30 for $\epsilon = 10^{-1}, 10^{-2}, 10^{-3}$, respectively.

dynamics. For example, it is seen that the concentration configuration in the domain changes rapidly after a few time steps from the beginning of solute injection so that the relative distance between the full model solution and the POD subspace increases sharply within few time steps and a model update is required. Otherwise, the variability of concentration in space tends to be smooth out after a few time steps so that a reduced model based on this type of information is prone to retain a relatively high global accuracy for a sustained period of time. This behavior can be quantified and embedded in the dynamic POD update procedure through the definition of the following residual for each time step:

$$\mathcal{R}_k = \frac{\left\| (\frac{\mathsf{M}}{\Delta t} + \mathsf{A})\mathsf{P}\mathbf{c}_p^k - (\mathbf{f}^k + \frac{\mathsf{M}}{\Delta t}\mathsf{P}\mathbf{c}_p^{k-1}) \right\|_2}{\left\| \mathbf{f}^k + \frac{\mathsf{M}}{\Delta t}\mathsf{P}\mathbf{c}_p^{k-1} \right\|_2}, \qquad (23)$$

where $||\cdot||_2$ is the Euclidean norm. In general, the residual (23) is different from the residual of the reduced model (16). Note that residual (23) does not depend of the full model solution and it is computed using only the solution at the current reduced step projected onto the full discrete space through the matrix P. We then employ (23) to design the adaptive algorithm for the update of the POD subspace basis which is described in the following:

(**U1**) A solution $\mathbf{c}_p^k$ is computed by solving the full discrete model, i.e., $\mathbf{c}_p^k = \mathbf{c}_f^k$, for the first N_t time steps ($N_t = 10$ in our example, i.e., $k \leq 10$).

(**U2**) The projection matrix is computed via the POD method relying on the full model solution computed in **U1**, $\mathbf{c}_p^k$ representing a snapshot.

(**U3**) The model is projected to the POD subspace yielding the reduced discrete model which is solved until

the residual $\mathcal{R}_k$ in (23) attains a given threshold value, ϵ.

(**U4**) Starting from the last solution obtained in **U3**, we switch to the solution of the full discrete model; in our example, we solve the full model for additional N_t time steps, a full analysis of the feedback between N_t and the transport setting being outside the scope of this work.

(**U5**) The solution computed in **U4** together with the old snapshots are employed to update the projection matrix used in the POD reduction method. Transport simulation progresses from step **U3** with the newly constructed reduced model.

Figure 8 depicts the temporal evolution of the global relative error (21) as a function of k and for different choices of Pe and of the selected threshold ϵ, for the same problem setting associated with Figure 7. It can be noted that the error is consistently low, due to the improved efficiency according to which the subspace updating is performed. As expected, the largest global relative errors are related to the selected threshold. Lowering the latter yields to a frequent update, resulting in an increased computational time.

Note that adding new snapshots to update the projection matrix during the last step (**U5**) increases the size of the reduced problem. However, the snapshot updating takes place only a limited number of times before the longest simulation time is reached. As a consequence, the dimension of the reduced problem is always much smaller than that of the full problem. It is further noted that, depending on the problem, in case of long simulation times one might need to progressively reduce
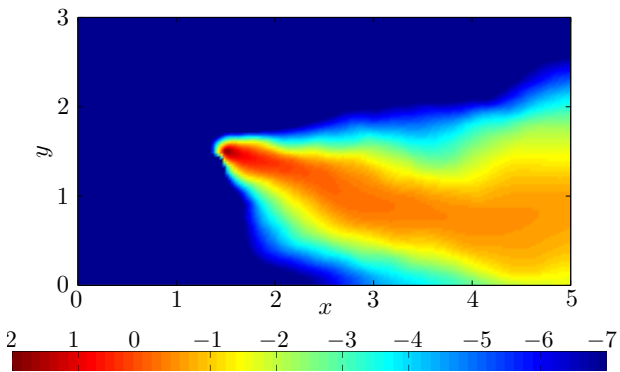
Fig. 9: Spatial distributions of solute concentration obtained by the alternating algorithm solution. Colors are associated with a logarithmic scale.

the number of snapshots included in the set (otherwise the size of the projection matrix P can grow in such a way that the advantage of the method is hampered at some point). This issue can be handled, for example, by progressively removing the oldest (in time) snapshots as time increases, once a fixed size for P is set. The choice of the latter is guided by the environmental setting and goal one selects and by the computational power available.

Finally, using the subspace update strategy presented in this section, we can simulate a dispersion and advection problem for a given time interval using a reduced amount of computational resources and being able to control the error through the value of the threshold ϵ. With reference to this point, Figure 9 clearly shows that concentration values computed at $k = 60$ using the alternating algorithm with $Pe = 30$ and $\epsilon = 10^{-1}$ is almost identical to the solution resulting from the full numerical problem shown in Figure 6a. It is remarkable to note that the results for 60 time steps out of 100 have been obtained by way of the reduced model. Since the linear system associated with the reduced model is much smaller than the full model system, the computational time associated with the reduced stage is essentially negligible, thus resulting in a speed-up of 2.5 (ratio between total number of steps, 100, and the number of full model steps, 40).

5 Summary and Conclusions

Our work leads to the following major conclusions:

1. We present a new algorithm for improving the performance of the numerical solution of transport problems associated with migration of conservative dissolved chemicals in a porous medium under the ac-

tion of advective and diffusive/dispersive processes. The proposed algorithm is based on a Proper Orthogonal Decomposition (POD) model reduction approach and employs a combination of the full and reduced model, the solution of which is alternated in time. The adaptive model reduction strategy we propose relies on (i) employing a *Snapshot Splitting Technique* (SST), which enables us to enrich the information content associated with the basis employed to construct the POD subspace on the basis of a given number of snapshots, and (ii) alternating between the solution of the full and reduced model through the definition of a threshold on the algebraic residual (23). We remark that the optimization of the model alternation sequence as a function of the domain heterogeneity and transport setting is outside the scope of the present contribution, which is keyed to illustrate and demonstrate the viability of the proposed strategy. As such, our work shows that using our subspace updating strategy enables us to simulate an advection-dispersion problem for a given time interval using a reduced amount of computational resources and with the ability to control the computational error through a desired threshold ϵ.

2. We explore the effect of the Péclet number on the quality of the reduced model approximation. We find that the POD method is associated with an improved performance for low Péclet numbers. This is related to the observation that the snapshots employed to build the basis vectors are taken from the full model solution and the concentration dynamics for low Péclet numbers are strongly linked to the past system states. On the contrary, in the presence of high values for the Péclet number (i.e., when advection strongly dominates), the method will automatically switch to the solution of the full model. This causes a minimal computational overhead due to the POD subspace computation. This apparent drawback has a negligible impact from an operative standpoint because the computational time required to compute the POD basis is much smaller than the time required for a full simulation step. As a future development, it would be interesting to extend the range of applicability of our approach by using more sophisticated model reduction methods, e.g., by coupling our strategy with Lagrangian based model reduction [33] or by resorting to more recent techniques such as hierarchical model reduction [2, 37].

3. The improvement associated with the proposed SST enables us to increase the number of steps within which the reduced model can provide a viable solu-

tion of the system behavior in the time-alternating algorithm, yielding a numerical solution associated with increased computational efficiency.

The joint use of the *Snapshot Splitting Technique* and of the time alternating model reduction algorithm has the potential to be extended to problems governing scalar transport in diverse types of flow fields and media. Cycling between the solution of the full and reduced models can provide a marked decrease of the complexity of a discrete problem characterized by a considerably high physical dimension according to which an ad hoc fast solver can be built. Additional advantages of the proposed framework can be thus itemized:

- it is adaptive, automatically switching between reduced and full model according to the value of the residual;
- it is independent of the discretization method, being based on the algebraic formulation;
- it reproduces the solution of the original model when no reduction is allowed (e.g., with a high advective field);
- it is possible to localize the reduction by applying the reduction procedure only in certain regions of the domain by resorting, for instance, to a domain decomposition approach;
- it provides a snapshot of the solution at a fixed time, which can be advantageously exploited when dealing with coupled system.

Since the proposed approach is directly applied to the discretized problem, future investigations might include the treatment of additional complex features associated with flow and transport in heterogeneous porous media, such as nonlinear chemical reactions and multi-phase flows.

References

1. Ballio, F., Guadagnini, A.: Convergence assessment of numerical monte carlo simulations in groundwater hydrology. Water resources research **40**(4) (2004)

2. Baroli, D., Cova, C.M., Perotto, S., Sala, L., Veneziani, A.: Hi-pod solution of parametrized fluid dynamics problems: preliminary results. MS&A series (In press, 2017)

3. de Barros, F.P.J., Ezzedine, S., Rubin, Y.: Impact of hydrogeological data on measures of uncertainty, site characterization and environmental performance metrics. Advances in Water Resources **36**, 51–63 (2012)

4. de Barros, F.P.J., Fiori, A.: First-order based cumulative distribution function for solute concentration in heterogeneous aquifers: Theoretical analysis and implications for human health risk assessment. Water Resources Research **50**(5), 4018–4037 (2014)

5. de Barros, F.P.J., Nowak, W.: On the link between contaminant source release conditions and plume prediction uncertainty. Journal of Contaminant Hydrology **116**(1), 24–34 (2010)

6. de Barros, F.P.J., Rubin, Y.: A risk-driven approach for subsurface site characterization. Water resources research **44**(1) (2008)

7. Bear, J.: Dynamics of fluids in porous media. Courier Corporation (2013)

8. Bergmann, M., Bruneau, C.H., Iollo, A.: Enablers for robust pod models. Journal of Computational Physics **228**(2), 516–538 (2009)

9. Cao, Y., Zhu, J., Luo, Z., Navon, I.: Reduced-order modeling of the upper tropical pacific ocean model using proper orthogonal decomposition. Computers and Mathematics with Applications **52**(8-9), 1373–1386 (2006)

10. Cardoso, M., Durlofsky, L., Sarma, P.: Development and application of reduced-order modeling procedures for subsurface flow simulation. International Journal for Numerical Methods in Engineering **77**(9), 1322–1350 (2009)

11. Chatterjee, A.: An introduction to the proper orthogonal decomposition. Current Science **78**(7), 808–817 (2000)

12. Crommelin, D., Majda, A.: Strategies for model reduction: Comparing different optimal bases. Journal of the Atmospheric Sciences **61**(17) (2004)

13. Dagan, G., Neuman, S.P.: Subsurface flow and transport: a stochastic approach. Cambridge University Press (2005)

14. Dentz, M., Le Borgne, T., Englert, A., Bijeljic, B.: Mixing, spreading and reaction in heterogeneous media: A brief review. Journal of Contaminant Hydrology **120**, 1–17 (2011)

15. van Doren, J., Markovinovic, R., Jansen, J.D.: Reduced-order optimal control of waterflooding using pod. In: 9th European Conference on the Mathematics of Oil Recovery (2004)

16. van Doren, J.F., Markovinović, R., Jansen, J.D.: Reduced-order optimal control of water flooding using proper orthogonal decomposition. Computational Geosciences **10**(1), 137–158 (2006)

17. Edery, Y., Guadagnini, A., Scher, H., Berkowitz, B.: Origins of anomalous transport in heterogeneous media: Structural and dynamic controls. Water Resources Research **50**(2), 1490–1505 (2014)

18. Efendiev, Y., Gildin, E., Yang, Y.: Online adaptive local-global model reduction for flows in heterogeneous porous media. Computation **4**(2), 22 (2016)

19. Esfandiar, B., Porta, G., Perotto, S., Guadagnini, A.: Impact of space-time mesh adaptation on solute transport modeling in porous media. Water Resources Research **51**(2), 1315–1332 (2015)

20. Fetter, C.W., Fetter Jr, C.: Contaminant hydrogeology, vol. 500. Prentice hall New Jersey (1999)

21. Ghommem, M., Presho, M., Calo, V.M., Efendiev, Y.: Mode decomposition methods for flows in high-contrast

porous media. global–local approach. Journal of computational physics **253**, 226–238 (2013)

22. Henri, C.V., Fernàndez-Garcia, D., Barros, F.P.: Probabilistic human health risk assessment of degradation-related chemical mixtures in heterogeneous aquifers: Risk statistics, hot spots, and preferential channels. Water Resources Research **51**(6), 4086–4108 (2015)

23. Jolliffe, I.: Principal component analysis. Wiley Online Library (2005)

24. Kowalski, M.E., Jin, J.M.: Model-order reduction of nonlinear models of electromagnetic phased-array hyperthermia. Biomedical Engineering, IEEE Transactions on **50**(11), 1243–1254 (2003)

25. Kunisch, K., Volkwein, S.: Galerkin proper orthogonal decomposition methods for parabolic problems. Numerische Mathematik **90**(1), 117–148 (2001)

26. Kunisch, K., Volkwein, S.: Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. SIAM Journal on Numerical Analysis **40**(2), 492–515 (2002)

27. Le Borgne, T., Dentz, M., Carrera, J.: Lagrangian statistical model for transport in highly heterogeneous velocity fields. Physical Review Letters **101**(9), 090,601 (2008)

28. Li, H., Luo, Z., Chen, J.: Numerical simulation based on pod for two-dimensional solute transport problems. Applied Mathematical Modelling **35**(5), 2489–2498 (2011)

29. Li, X., Hu, B.X.: Proper orthogonal decomposition reduced model for mass transport in heterogenous media. Stochastic Environmental Research and Risk Assessment **27**(5), 1181–1191 (2013)

30. Lumley, J.L.: The structure of inhomogeneous turbulent flows. Atmospheric Turbulence and Radio Wave Propagation pp. 166–178 (1967)

31. Luo, Z., Li, H., Zhou, Y., Xie, Z.: A reduced finite element formulation based on pod method for two-dimensional solute transport problems. Journal of Mathematical Analysis and Applications **385**(1), 371–383 (2012)

32. Ly, H.V., Tran, H.T.: Proper orthogonal decomposition for flow calculations and optimal control in a horizontal cvd reactor. Tech. rep., DTIC Document (1998)

33. Mojgani, R., Balajewicz, M.: Lagrangian basis method for dimensionality reduction of convection dominated nonlinear flows. arXiv preprint arXiv:1701.04343 (2017)

34. Moslehi, M., Rajagopal, R., de Barros, F.P.J.: Optimal allocation of computational resources in hydrogeological models under uncertainty. Advances in Water Resources **83**, 299–309 (2015)

35. Pasetto, D., Guadagnini, A., Putti, M.: Pod-based monte carlo approach for the solution of regional scale groundwater flow driven by randomly distributed recharge. Advances in Water Resources **34**(11), 1450–1463 (2011)

36. Pasetto, D., Putti, M., Yeh, W.W.G.: A reduced-order model for groundwater flow equation with random hydraulic conductivity: Application to monte carlo methods. Water Resources Research **49**(6), 3215–3228 (2013)

37. Perotto, S.: A survey of hierarchical model (hi-mod) reduction methods for elliptic problems. In: Numerical simulations of coupled problems in engineering, pp. 217–241. Springer (2014)

38. Porta, G., Bijeljic, B., Blunt, M., Guadagnini, A.: Continuum-scale characterization of solute transport based on pore-scale velocity distributions. Geophysical Research Letters (2015)

39. Porta, G., Thovert, J.F., Riva, M., Guadagnini, A., Adler, P.: Microscale simulation and numerical upscaling of a reactive flow in a plane channel. Physical Review E **86**(3), 036,102 (2012)

40. Rapún, M.L., Vega, J.M.: Reduced order models based on local pod plus galerkin projection. Journal of Computational Physics **229**(8), 3046–3063 (2010)

41. Remy, N., Boucher, A., Wu, J.: Applied geostatistics with SGeMS: a user's guide. Cambridge University Press (2009)

42. Rubin, Y.: Applied stochastic hydrology. Oxford Univ. Press, New York (2003)

43. Sahimi, M.: Flow and transport in porous media and fractured rock: from classical methods to modern approaches. John Wiley & Sons (2011)

44. Siade, A.J., Putti, M., Yeh, W.W.G.: Snapshot selection for groundwater model reduction using proper orthogonal decomposition. Water Resources Research **46**(8) (2010)

45. Siena, M., Guadagnini, A., Riva, M., Bijeljic, B., Nunes, J.P., Blunt, M.: Statistical scaling of pore-scale lagrangian velocities in natural porous media. Physical Review E **90**(2), 023,013 (2014)

46. Tartakovsky, D.M.: Assessment and management of risk in subsurface hydrology: A review and perspective. Advances in Water Resources **51**, 247–260 (2013)

47. Zhang, D.: Stochastic methods for flow in porous media: coping with uncertainties. Academic press (2001)