Ranking Viscous Finger Simulations to an Acquired Ground Truth with Topology-aware Matchings

Maxime Soler, Martin Petitfrere, Gilles Darche, Melanie Plainchault, Bruno Conche and Julien Tierny
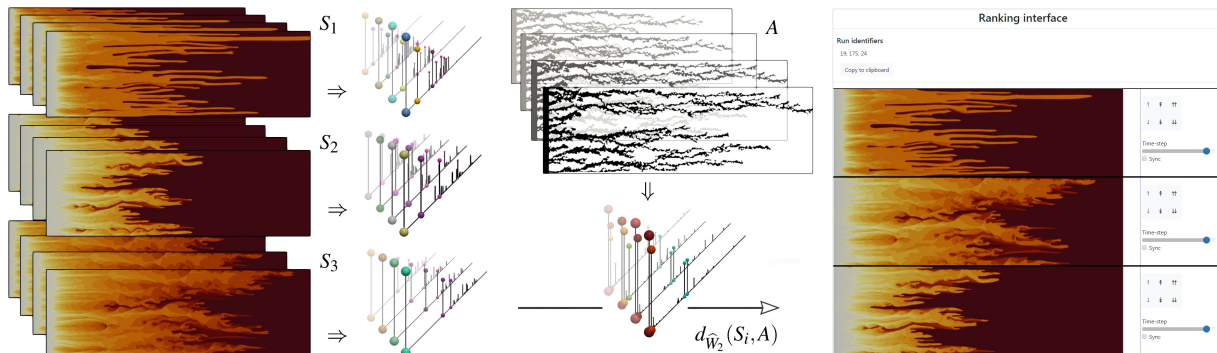
Figure 1: Overview of the ranking framework. An ensemble of viscous fingering simulation runs S_1, S_2, S_3 is launched, and the persistence diagram of newly available time-steps can be computed *in-situ* (left). Only persistence diagrams for which there is a matching ground truth image (center, top) are computed. Diagrams of every simulation are compared with the diagrams of the ground truth (center, bottom) at matching time-steps. This comparison, based on a metric $\widehat{W}_2$ combining the notions of persistence and geometry, outputs a distance measurement, which can be integrated over time to form the metric $d_{\widehat{W}_2}$. This produces a final ranking (right) which characterizes the quality of simulations, allowing experts to select and explore best performing runs automatically.

ABSTRACT

This application paper presents a novel framework based on topological data analysis for the automatic evaluation and ranking of viscous finger simulation runs in an ensemble with respect to a reference acquisition. Individual fingers in a given time-step are associated with critical point pairs in the distance field to the injection point, forming persistence diagrams. Different metrics, based on optimal transport, for comparing time-varying persistence diagrams in this specific applicative case are introduced. We evaluate the relevance of the rankings obtained with these metrics, both qualitatively thanks to a lightweight web visual interface, and quantitatively by studying the deviation from a reference ranking suggested by experts. Extensive experiments show the quantitative superiority of our approach compared to traditional alternatives. Our web interface allows experts to conveniently explore the produced rankings. We show a complete viscous fingering case study demonstrating the utility of our approach in the context of porous media fluid flow, where our framework can be used to automatically discard physically-irrelevant simulation runs from the ensemble and rank the most plausible ones. We document an in-situ implementation to lighten I/O and performance constraints arising in the context of parametric studies.

1 INTRODUCTION

The chaotic nature of fluid flows makes it difficult to account for the propagation of initial uncertainties in numerical models, or uncertainties in model parameters. To predict uncertain phenomena, thanks to the increase in computing power in recent years, Monte Carlo methods have been broadly used, for instance in climate modeling, forecasts, statistical physics, chemistry and astrophysics. The idea is to compute a large number of simulations, called an *ensemble*, while densely sampling the space of input parameters. A *post-mortem* comparison (i.e. performed *after* simulations have been completed) to experimentally acquired data can then determine which simulations produced the most realistic outcomes and how input parameters affect their variability.

Specifically, in reservoir engineering, an area of petroleum engineering concerned with fluid flow through porous media, it is important to quantitatively predict well productions, i.e. the quantity of oil that can be extracted. For example, injector wells are commonly drilled to inject water and flush the oil in place, which is extracted by producer wells; simulation can help well placement optimization and production forecast. Numerical models are subject to parameter uncertainties, and can be tuned by launching randomly sampled ensemble simulations. Usually, reference production rates and well pressures are history-matched with the simulated ensembles, which ideally would allow domain experts to restrict the space of input parameters. This history match procedure is usually applied at the field scale (oil and gas reservoirs), but also at the core scale (a few decimeters) when lab engineers want to match the behavior of experimental corefloods.

In practice, notably in the domain of Darcy-type simulations at the core scale, production and pressure data is not sufficient to infer model parameters. Further measuring tools have been recently integrated in lab experiments in order to constrain the parameter space, by monitoring the *saturation* scalar fields through X-rays, so as to obtain information on phase velocities and residual saturations. Here the saturation measures the volume fraction of a given phase in the geometrical domain. Observing these scalar fields seems relevant when the fluid behaves in a particularly chaotic way, so that simulations which are not physically adequate could be detected. The case of the viscous fingering phenomenon, an instability which occurs at the interface between two fluids of distinct viscosity in porous media, is of particular interest.

In this context, for all simulations quantitatively reproducing production and pressure data, experts have to visually inspect each member of the ensemble to further discard non-physical simulations. This process is currently performed manually and can be time consuming. Moreover, the viscous fingering process involves a notoriously chaotic and unstable geometry. In particular, two different viscous fingering simulations can both be realistic from an expert's point of view (and yield valid physical properties for reservoir exploitation) even though saturation would admit fingers with a drastically different shape and distribution in space. This high geometrical variability makes it particularly challenging to derive a meaningful distance metric to compare saturation scalar fields

between a simulation and a ground truth.

For studying scalar fields, topological data analysis (TDA) has been used in recent years as a robust and reliable setting, allowing one to hierarchically define features of interest in the data [32]. Its applicability to time-varying data [10, 91], ensembles [35] and comparisons [92] makes it a reliable candidate for assessing the likeliness of simulations in an ensemble given a ground truth. Although several approaches have explored the promising potential of TDA for extracting and characterizing the features of interest in viscous fingering simulations [36, 59], no approach has been proposed to estimate the similarity between two time-varying viscous fingerings based on topological representations.

In this application paper, we address the aforementioned issues by proposing a novel framework, based on topological data analysis, for quantitatively ranking simulations from an ensemble with respect to a ground truth in a viscous fingering case study. This framework allows experts to easily separate the most realistic simulations from the most unrealistic ones. It is based on a new approach for comparing temporal sequences of *persistence diagrams*, specifically adapted to the problem of viscous fingering. Extensive experiments quantitatively show the superiority of our approach compared to traditional alternatives. The framework also includes an interactive visual system for exploring the output rankings. Finally, we report a complete case study for which the presented approach has been applied *in-situ* (i.e. during the simulation).

1.1 Related work

Since the work presented in this paper involves multiple domains (viscous finger simulations, topological data analysis, time-varying data), this section briefly presents their respective related work.

Viscous fingering is a well-known instability encountered in soils and porous media [77], arising from the unfavorable mobility ratio between an injected fluid and the fluid in place, for instance when injecting water in a highly viscous oil. These phenomena have been studied in the context of petroleum engineering at multiple scales [38, 90]. Other factors than the viscosity ratio are at play, such as properties inherent to the medium in which the fingering takes place [48, 87, 102, 103]. In practice, performing waterflood in highly viscous oil can lead to physical instabilities resulting in fingering patterns, with water flowing in preferential paths and bypassing large quantities of oils. To prevent this phenomenon, polymer can be injected in order to increase the water viscosity, therefore making the injection front more stable, and leading to increased macroscopic oil recovery [29]. There are multiple numerical models that can describe the evolution of fluids in the context of water floods; some have been qualitatively compared to acquisitions [29, 73, 80], but the literature lacks robust ways to quantitatively compute their difference.

In-situ: as current trends in super-computing indicate an increase of the computing power that evolves faster than memory, IO and network bandwidth, new paradigms for scientific simulation are needed. The simulation of flow in porous media, of key importance for studying viscous fingers, is particularly affected by data movement problematics, as models keep increasing in size, and high-resolution time sampling is required for observing realistic simulations. Over recent years, solutions for limiting data movement were developed in this perspective, such as *in-situ* [2, 67, 71, 75, 105] and *in-transit* [5, 62] models, with a clear ambition to reach toward exascale computing [22, 96] in the forthcoming years. However, to the best of our knowledge, no data analysis method has yet been proposed for the in-situ analysis of viscous fingers.

Topological data analysis (TDA) techniques [28, 32, 47, 69] have been used over the course of recent years because of their ability to hierarchically identify features in scalar data in a generic, robust [19, 33] and efficient manner. They have been applied in various scientific domains, such as computational fluid dynamics [36, 51], turbulent combustion [9], material sciences [45], biological imaging [8, 14],

chemistry [6, 40], astrophysics [81, 97], ensemble clustering [35], compression [93] or feature tracking [92]. One of the reasons for the successful applications of TDA is the possibility for experts to easily translate high-level domain-specific notions in terms of topological data structures, which are abstractions related to geometrical aspects or discrepancies in the data. Among such abstractions are persistence diagrams [31, 33], contour trees [13], Reeb graphs [7, 68, 101], Morse-Smale complexes [43]. For instance, in astrophysics the cosmic web can be extracted by querying the most persistent 1-separatrices of the Morse-Smale complex connected to maxima of matter density [97]. Similar TDA applications can be found in the above examples. Topological data analysis techniques have also been applied *in-situ* [54], which demonstrates their interest and relevance in the context of large-scale simulations. It has also been used to study the viscous fingering phenomenon, for instance in ensembles of particle simulations, to determine how the resolution affects fingers [36, 59] or to provide frameworks for their visual exploration and interpretation [58]; but never, to our knowledge, for the purpose of comparing simulations, in particular with a reference.

Feature-oriented distances: for comparing simple discrete scalar fields such as images, intuitive approaches are point-wise geometric distances such as the Euclidean and chord distances, or distances with a statistic awareness such as the Mahalanobis distance or correlation coefficients [16]. In specific applicative domains, however, the experts' knowledge should be accounted for to gain a more precise insight of what is of actual interest in the data and which patterns or subsets are interesting to compare. Consequently, feature-oriented distance definitions are exposed in the remainder of this section. Associating geometrical loci in scalar data based on a high-level definition of features of interest often relies on computing the overlap of geometrical sub-domains [9, 10, 78, 82–85, 91]. Such methods are used for feature tracking in time-varying data [86]. On another note, Transportation theory offers an important continuous formulation of this problematic, with the notion of a *Wasserstein* or *Earth mover's* distance [50, 57, 61], which has gained interest in recent years [26, 55, 94, 95]. In the discrete setup, when applied to topological structures such as persistence diagrams, transport-based matching methods suffer from instabilities in the geometrical domain [21], for which the underlying metric can be specifically corrected [92] depending on the context. Though this family of approaches for computing distances between features based on transport seems promising for the problem of comparing viscous fingers, there is, to our knowledge, no work studying such an application.

1.2 Contributions

This application paper makes the following new contributions:

1. **Approach:** we present a novel analysis framework allowing to select relevant members in a simulated ensemble given a ground truth. The system yields a ranking that allows experts to visually explore the most likely simulations and discard the most unrealistic ones.

2. **Metrics:** new topological metrics for comparing time-varying viscous fingers are introduced, based on the Wasserstein matching of persistence diagrams, specifically tuned for the viscous fingering phenomenon and integrated over time.

3. **Case study:** a complete case study of a viscous fingering simulation ensemble is documented, along with a proof-of-concept in-situ implementation of our approach.

4. **Evaluation:** the metrics and ranking framework are qualitatively evaluated with feedback from domain experts. The quantitative performance of our approach is also analyzed and its superiority over traditional alternatives is demonstrated.

2 PRELIMINARIES

This section describes the context of reservoir simulation, introduces our formal setting and the metrics which are extended by our work.

2.1 Darcy-type porous media simulation

There are multiple models for simulating flow in porous media. Though our viscous finger analysis framework is not limited to a specific simulation model, we introduce here Darcy-type simulations, for which the physics is governed by quantities averaged over control volumes. We consider diphasic flow with oil and water.

Eq. 1 describes mass conservation, where $i \in \{o, w\}$ is the oil and water phase; ϕ is the porosity of the medium; ρ_i is the mass density of phase i; S_i is the *saturation* of phase i (it stands for the volume fraction of phase i); q_i is the well source term (injection/production) of phase i; and $\mathbf{v}_i$ is the velocity of phase i. If V_{tot} denotes the total volume, then the mass of component i is given by $m_i = V_{tot}\phi\rho_i S_i$.

$$\frac{\partial}{\partial t}(\phi\rho_i S_i) = -\nabla \cdot \rho_i \mathbf{v}_i + q_i \tag{1}$$

Darcy's law is an equation that describes fluid flow in porous media, determined experimentally by H. Darcy in 1856 for one phase [27], and which can be derived from the Stokes equations [104]. Its extension to multiphasic flow is given in Eq. 2, where $\mathbf{v}_i$ is the velocity of phase i; $\mathbf{K}$ is the absolute permeability tensor of the porous medium; μ_i is the viscosity of i; $\mathbf{g}$ is the acceleration of gravity; P_i is the pressure of phase i; kr_i is the relative permeability of phase i. In our model, kr_i is a function of water saturation.

$$\mathbf{v}_i = -\mathbf{K}\frac{kr_i}{\mu_i}(\overrightarrow{\nabla}P_i - \rho_i\mathbf{g}) \tag{2}$$

Furthermore, as shown in Eq. 3, oil saturation can be simply expressed in terms of water saturation, and water pressure can be expressed in terms of oil pressure, with P_c being the *capillary pressure*, a function of water saturation.

$$\begin{cases} S_w = 1 - S_o \\ P_w = P_o - P_c \end{cases} \tag{3}$$

In this model, the unknowns are the saturations S_i and pressures P_i. The system formed by Eq. 1, 2, and 3 can then be solved numerically to yield the evolution of fluid in porous media under Darcy's approximation. Moreover, models exist [11,56,98] for expressing P_c as a function of S_w, which can be obtained experimentally through centrifugal fan experiments. Relative permeabilities kr_o and kr_w, also functions of S_w, are more elusive. Numerous models have been proposed in the literature in various contexts [1,12,18,23,24,37,53], and there is a number of methods for building them from interpretation of lab experiments [30,46,60,66,74]. Their correct definition, however, is key to a realistic description of flow in porous media, and can be quite difficult to obtain depending on the recovery mechanism, especially in processes involving severe viscous fingering patterns (in which case Darcys law can become approximate) or when dealing with an extra fluid phase, like an injected gas phase [4], notably because of the limited availability of experimental measurements. In the remainder of this work, relative permeabilities are considered as an input parameter of simulations.

Most of reservoir simulators are based on finite volumes discretizations of Eq. 1, 2, 3 on a gridded 2D or 3D model, in which independent variables are constant in each grid block. These quantities must be determined at each time-step by solving the sets of non-linear conservation equations. The results shown in the experiments section were obtained in the 2D case with an in-house research reservoir simulator [49,70] using an IMPES scheme (IMplicit Pressure, Explicit Saturation) [17], which separately computes saturation with an explicit time approximation, and pressure with an implicit one. At every time-step, scalar data defined on control volumes is updated. As there are multiple variables, the simulator outputs multiple fields, like phase pressures and saturations. The pressure field is very diffusive, and in the diphasic case the saturation

is constrained by Eq. 3. Thus, a good indicator of the simulation state is the scalar field of water saturation S_w, which we will use as input data in the following.

2.2 Persistence diagrams

This subsection describes our formal setting. It contains definitions adapted from [100]. An introduction to TDA can be found in [32].

Input data: for each time step, the input saturation data is considered as a piecewise linear (PL) scalar field $f : \mathcal{M} \to \mathbb{R}$ defined on a PL d-manifold $\mathcal{M}$ with $d = 2$ in our application. Scalar values are given at the vertices of $\mathcal{M}$ and linearly interpolated elsewhere.

Critical points: if $w \in \mathbb{R}$ is an isovalue, the *sub-level set* of w, noted $f_{-\infty}^{-1}(w)$, is the pre-image of the open interval $(-\infty, w)$ under f: $f_{-\infty}^{-1}(w) = \{q \in \mathcal{M} \mid f(q) < w\}$. Symmetrically, the *sur-level set* is $f_{+\infty}^{-1}(i) = \{q \in \mathcal{M} \mid f(q) > w\}$. These two objects serve as segmentation tools in multiple analysis tasks [8,9,14]. The points $q \in \mathcal{M}$ where the topology of $f_{-\infty}^{-1}(f(q) - \varepsilon)$ differs from that of $f_{-\infty}^{-1}(f(q) + \varepsilon)$ for $\varepsilon \to 0$ are called the *critical points* of f. They can be classified according to their *index* $\mathcal{I}$: 0 for minima, 1 for 1-saddles, $d - 1$ for $(d - 1)$-saddles, d for maxima.

Persistence diagrams: the set of critical points of f can be visually represented by a topological abstraction called the *persistence diagram* [19,33] (Fig. 2). Specifically, the topological Elder Rule [32] states that critical points can be arranged in a set of pairs, such that each critical point appears in only one pair (c_i, c_j) with $f(c_i) < f(c_j)$ and $\mathcal{I}(c_i) = \mathcal{I}(c_j) - 1$. Such a pairing indicates that a topological feature of $f_{-\infty}^{-1}(i)$ (connected component, cycle, void, etc.) created at critical point c_i dies at the critical point c_j. For example, as the value i increases, if two connected components of $f_{-\infty}^{-1}(i)$ meet at a saddle c_j of f, the *youngest* of the two (the one with the highest minimal value, c_i) *dies* at the advantage of the oldest (the one with the lowest minimal value). Critical points c_i and c_j form a *persistence pair*.

A classical representation of the persistence diagram $\mathcal{D}(f)$ embeds each pair (c_i, c_j) as a point in the 2D plane at coordinate $(f(c_i), f(c_j))$. The height of the pair $P(c_i, c_j) = |f(c_j) - f(c_i)|$ is called the *persistence* and denotes the life-span of the topological feature created at c_i and destroyed at c_j. In 3D, the persistence of pairs linking critical points of index $(0, 1)$, $(2, 3)$ and $(1, 2)$ respectively denotes the life-span of connected components, voids and non-collapsible cycles of $f_{-\infty}^{-1}(i)$. In the following, we will focus on $(1, 2)$ persistence pairs (involving maxima).

The interest of this visual representation in practice is that it quickly hints at the distribution and relative importance of critical points. Small oscillations due to noise in the input data are typically represented by pairs with low persistence, in the vicinity of the diagonal. In contrast, the most prominent topological features are associated with large vertical bars (Fig. 2, b).

Persistence diagrams are used in many applications, for instance as a visual help for interactively tuning simplification thresholds in multi-scale segmentation tasks, either based on the Reeb graph [14,41,68,72,99,101] or the Morse-Smale complex [43,44,76].

2.3 Metrics between Persistence diagrams

Metrics have been defined to evaluate the distance between scalar fields $f, g : \mathcal{M} \to \mathbb{R}$. The L^p-norm $||f - g||_p$ is a classical example.

In the context of TDA, multiple metrics [15,19] have been introduced to compare two persistence diagrams $\mathcal{D}(f)$ and $\mathcal{D}(g)$.

Critical point pairs in persistence diagrams can be associated with a point-wise distance, noted d_p inspired by the L^p-norm. Given two persistence pairs $a = (a_x, a_y) \in \mathcal{D}(f)$ and $b = (b_x, b_y) \in \mathcal{D}(g)$, d_p can be defined as:

$$d_p(a, b) = (|a_x - b_x|^p + |a_y - b_y|^p)^{1/p} \tag{4}$$

The *Wasserstein* distance [50,61], noted W_p, between persistence diagrams $\mathcal{D}(f)$ and $\mathcal{D}(g)$ can then be defined as:
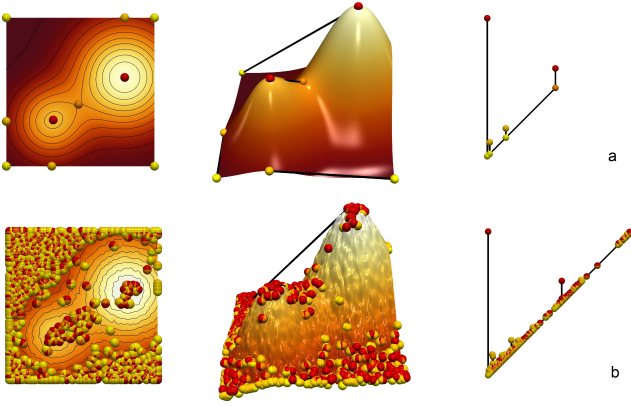
Figure 2: A smooth (top row) and a noisy (bottom row) scalar field, defined on a 2D domain (left), with their 3D terrain representation (middle) and persistence diagrams (right). Critical points are represented as spheres (red: maxima, orange: saddles, yellow: minima). The largest pairs in the diagrams correspond to the two main hills.
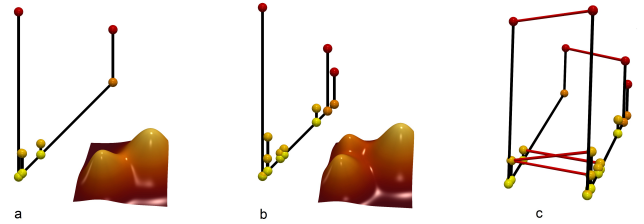

Figure 3: Persistence diagrams of two distinct 2D scalar fields (a, b), matched by the Wasserstein metric (c, matching pairs are linked with red segments). The third hill of (b), captured by the rightmost persistence pair, is discarded by the matching.

$$W_p\big(\mathscr{D}(f),\mathscr{D}(g)\big) = \min_{\phi\in\Phi}\left(\sum_{a\in\mathscr{D}(f)} d_p\big(a,\phi(a)\big)^p\right)^{1/p} \quad (5)$$

where Φ is the set of all possible assignments ϕ mapping each persistence pair $a\in\mathscr{D}(f)$ to a persistence pair $b\in\mathscr{D}(g)$ with identical critical indices $\mathscr{I}$ or to its diagonal projection, noted diag(a) – which corresponds to the removal of the corresponding feature from the assignment, with a cost $d_p(a,\text{diag}(a))$. It is illustrated in Fig. 3. In practice, the Wasserstein distance is computed by solving a variant of the assignment problem [52, 63, 64, 92].

In the applications, this point-wise distance d_p can be fine-tuned to better account for the layout of critical points in the geometrical domain $\mathscr{M}$, as done in applications such as feature tracking [92], resulting in the following *lifted* point-wise distance:

$$\widehat{d_p}(a,b) = (\alpha_x|a_x - b_x|^p + \alpha_y|a_y - b_y|^p + \beta_x\delta_x^p + \beta_y\delta_y^p)^{1/p} \quad (6)$$

with $\delta_x = |x_a - x_b|$ and $\delta_y = |y_a - y_b|$ in the 2D case, and where (x_a, y_a) stands for the coordinates of the extremum of the persistence pair a in the geometrical domain $\mathscr{M}$. Coordinates are taken at extrema rather than saddles or midpoints because extrema usually bear more meaning in the applications, but this can be adapted to the specificity of applicative cases.

In the above equation, the coefficients α_x, α_y, β_x and β_y need to be properly tuned for the target application. A geometrically lifted version of the Wasserstein distance, noted $\widehat{W}_p$, can then be introduced as:

$$\widehat{W}_p\big(\mathscr{D}(f),\mathscr{D}(g)\big) = \min_{\phi\in\Phi}\left(\sum_{a\in\mathscr{D}(f)} \widehat{d_p}\big(a,\phi(a)\big)^p\right)^{1/p} \quad (7)$$

3 ANALYSIS FRAMEWORK

This section describes the problem of representing viscous fingers appearing in time-varying saturation fields, comparing them across simulations, and our approach for addressing this problem. In the following, we will note each time step of the reference ground-truth acquisition A_t and each time step of a simulation run S_t. Then, the goal of our framework is to efficiently compute relevant similarity measures, to rank simulation runs in order of increasing distance to the acquisition, so as to present to the experts the most plausible simulations for further inspection (Fig. 1).

Note that, in the rest of the paper, we will consider only one ground-truth acquisition data set, as the acquisition process (further detailed in Sec. 4.1) is long (several months) and involves expensive instrumentations (the process requires the acquisition machinery to operate in a high pressure environment to replicate the reservoir

conditions). Because of the rarity of such acquired data, the number of *available* acquired time steps A_t is in practice significantly lower than the number of simulated time steps S_t. The simulator is thus set up to output additional time steps corresponding to a specific set of volumes of injected water, which were recorded for each time step A_t. This physical criterion allows us to reliably match in time acquired and simulated time steps. The correspondence between simulation and acquisition timesteps is then given and reliable.

3.1 Feature representation

As discussed in the introduction, trying to reproduce the viscous fingering phenomenon with Darcy-type simulation software is very challenging because the fingering geometry greatly varies when one modifies input parameters, even slightly. In particular, the input parameters considered here are the relative permeabilities kr_i. When comparing a simulation to an acquired ground-truth, this great geometrical variability challenges traditional image based distances, either point-wise based (L_2 norm) or morphing based [26]. Moreover, the raw geometry of the viscous fingers can be insufficient in practice to identify all plausible simulations. Indeed, two geometrically different simulations can be deemed equally plausible by the experts if they share more abstract similarities, involving the number of fingers, their prominence and their progress in the porous medium. Thus, a proper feature representation, capable of abstracting these informations, is required to correctly represent the viscous fingering. Fig. 4 illustrates the extent to which the geometry of fingers may vary across simulations and how clearly distinct simulations can be judged as equally plausible by the experts.

The water saturation scalar field allows one to visually identify fingers, because they form a clear, sharp frontier with the background (as the geometric domain was initially filled with oil). The first step for identifying fingers then consists in extracting a sub-level set $f_{-\infty}^{-1}(w)$ of water saturation, for an isovalue w chosen properly, to extract the geometric domain $\mathscr{M}$ where fingers are effectively present. In our use case, based on discussions with experts, we set in practice this isovalue parameter once for all to 0.12. Let $\mathscr{F} = f_{-\infty}^{-1}(w)$ be that sub-part of $\mathscr{M}$. The same workflow can be
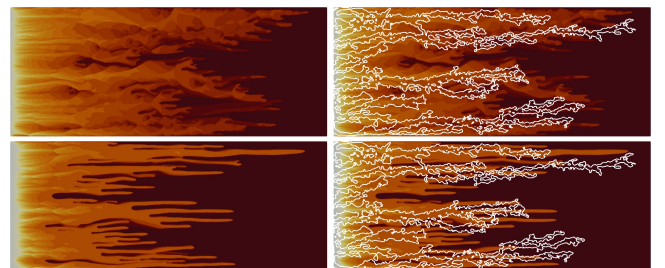

Figure 4: Late time-steps of two Darcy-type simulation runs launched with different model parameters (left column). The ground truth obtained with X-rays is contoured in white (right column, superimposed). Runs exhibit a very chaotic finger geometry.
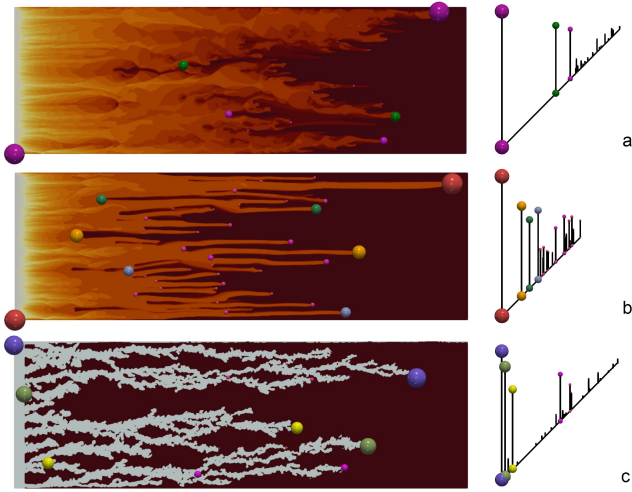
Figure 5: Simulated time-steps (left column, a and b) and the matching ground truth image (left column, c). Critical points are represented with spheres, and the corresponding persistence diagram is shown on the right. As every critical point belongs to only one persistence pair, the color of spheres encodes their persistence pair; and their diameter encodes its height (the larger, the higher the persistence). The most important fingers can clearly be identified by looking at the most persistent pairs in diagrams (right column). For instance, we can see that the three most important fingers in the acquisition are the purple, green and yellow ones.

applied on acquired X-ray images.

To compare a simulation to an acquired ground-truth, a naive strategy consists in estimating overlaps between the sub-level sets of saturation of the simulation and the acquisition, for a given time-step, and use the area of such an overlap as a measure of likeliness. However, this purely geometric approach appears to be inadequate in practice due to the important variability in the number and shape of fingers, which then would not be accounted for (see Fig. 4).

A natural way of characterizing fingers while taking their shape into consideration is to provide $\mathcal{F}$ with a *descriptive* scalar field, for instance a geodesic distance from the injection point. Here, the injection point is the left boundary of the domain, so the scalar field can simply be the x geometrical coordinate. Local maxima of this new scalar field would then correspond to the tips of viscous fingers, and saddles to valleys between fingers. Since they correspond to finger tips, maxima of the x geometrical coordinate provide a useful information to represent the progress of each finger in the porous medium. Moreover, in this setting, the persistence of the pair involving each maximum directly represents the length of the corresponding finger, which can be used as a reliable measure of importance given this application, to distinguish the main fingers from noise. The persistence diagram directly captures this information, in a robust and hierarchical setting. Fig. 5 illustrates the correspondence between fingers in the domain and pairs of critical points in persistence diagrams. In this context, persistence diagrams seem to be a promising feature representation for viscous fingers, since they efficiently describe their number, progress through the porous medium as well as their prominence.

3.2 Metrics between time-varying persistence diagrams

Considering that viscous fingers are captured by persistence diagrams, computing the similarity of a simulation with respect to a ground truth would require, in a first step, to compute distances between persistence diagrams. As outlined in the introduction, metrics have been introduced for this purpose, notably the (2-)Wasserstein distance in the *birth-death* space, noted W_2 (Eq. 5).

A drawback of W_2 is that it does not take into consideration

geometrical information, other than coming from the birth-death space. Fig. 6 illustrates this limitation. To avoid this problem, a *lifted* adaptation of W_2, noted $\widehat{W}_2$, including geometrical components can be considered (Eq. 7). It is subject to input parameters indicating the importance given to each geometrical component. Note that, the *Earth mover's distance* [57], noted *EMD*, which is an alternative of interest too for our application, is a special case of $\widehat{W}_2$, for $\alpha_x = \alpha_y = 0$. It is similar to W_2, but it only operates on the geometrical space instead of the birth-death space. Thus, the lifted Wasserstein distance $\widehat{W}_2$ can be interpreted as a blend between the W_2 distances in the diagram birth-death space and in the geometrical domain.

In practice, an important characteristic of a viscous fingering simulation run is the moment when the longest finger arrives at the right boundary, called *breakthrough time*. Correctly predicting this event is essential because once it is reached, it means a preferential path has been formed, allowing water to easily flow through, impacting production. Thus, the position of local maxima (i.e. fingertips) is more important than the position of saddles (i.e. finger branchings).

Then, given a time-step t, to compare the persistence diagrams coming from a simulation S_t and the acquisition A_t, metrics should be more sensitive to the advancement of fingertips, then to the global extent of fingers, and lastly to their y location in the domain. Thus, at this point, we propose to select the following metrics:

- The Earth mover's distance for local maxima: $EMD(S_t, A_t)$

- The 2-Wasserstein distance: $W_2(S_t, A_t)$

- The 2-Wasserstein distance, lifted to include geometrical information (the position of critical points): $\widehat{W}_2(S_t, A_t)$. As in this application, the advancement of fingertips is much more important than their vertical position in the domain, we only consider the x-coordinate of critical points. Thus, lifting coefficients (cf. Eq. 6) are $\beta_x = 10/\gamma$ (γ being the extent of the geometrical domain), $\beta_y = 0$, and $\alpha_x = \alpha_y = 1/\rho$ (ρ being the range of the scalar function). The values of these lifting parameters have been adjusted empirically based on discussions with experts.

Characterizing the evolution of fingers through time raises the necessity to integrate these metrics, as they are intended to evaluate the proximity between persistence diagrams for a single time-step t. Thus, to measure the distance from a time-varying simulation S to the time-varying acquired ground truth A, we introduce time-integrated versions, based on the L_2 norm, of the above metrics:

- $d_{EMD}(S,A) = \left(\sum_t \left(EMD(S_t, A_t) \right)^2 \right)^{1/2}$

- $d_{W_2}(S,A) = \left(\sum_t \left(W_2(S_t, A_t) \right)^2 \right)^{1/2}$

- $d_{\widehat{W}_2}(S,A) = \left(\sum_t \left(\widehat{W}_2(S_t, A_t) \right)^2 \right)^{1/2}$

As suggested by the experts, the displacement speed of the saturation front is key to predicting breakthrough time. They suggested to match in priority simulations which display compatible fronts in terms of velocity during the experiment. Given fingers are captured by persistence diagrams, a possibility for appreciating their evolution with respect to that suggestion would be to compute the sequence of distances between diagrams in successive time steps. In other words, for each couple of consecutive time steps t and $t+1$, compute a distance between S_t and S_{t+1} (Sect. 2.3), and integrate for all time steps. Here the chaotic behavior displayed by fingers when input simulation parameters change need not be taken into account: we are considering a unique simulation run, which has temporal coherence, therefore it is easier to choose a fitting metric. As shown in Fig. 7, a working solution is the 2-Wasserstein distance, lifted to give more importance to the y coordinate of maxima: $\breve{W}_2(S_t, S_{t+1})$, with lifting coefficients $\beta_x = 0$, $\beta_y = 10/\gamma$, $\alpha_x = \alpha_y = 1/\rho$ (γ is the geometrical extent; ρ is the scalar range). Because of the variability in the number of fingers, however, considering the difference of traveled distances alone could be problematic, for many little fingers going
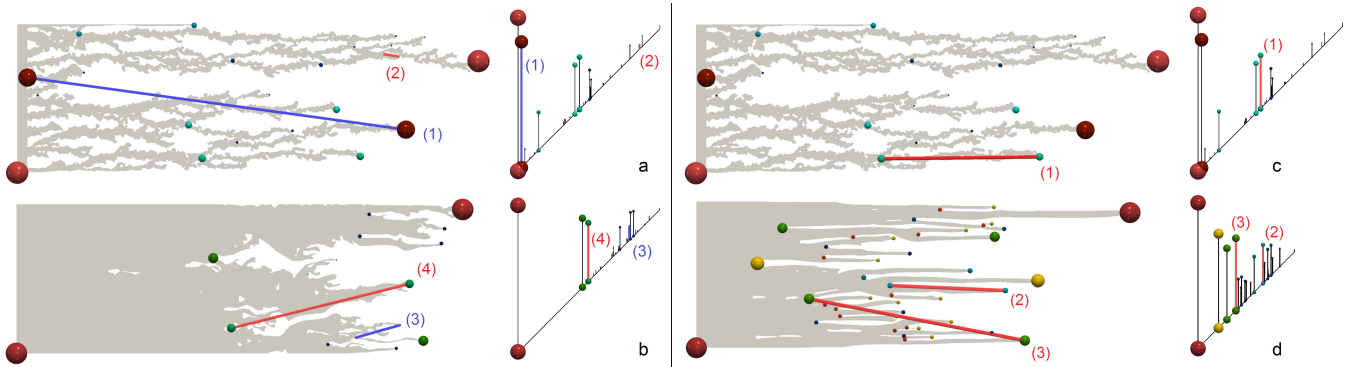
Figure 6: Limitations of matching methods based on geometry only (a, b) and matching methods based on persistence only (c, d). As the Earth mover's distance (left) only considers the geometrical location of extrema, it can incorrectly associate critical points belonging to unrelated fingers. For instance, the large pair in the acquisition, represented with a blue segment (a, (1)), is matched to a pair with low persistence (b, (3)) because their maxima are geometrically close; and the large red finger in the middle of a simulation (b, (4)) is matched to a small protrusion (a, (2)) attached to the largest finger in the acquired image. On the right, the reference metric for matching persistence diagrams, the 2-Wasserstein metric, is shown to associate the bottom finger in the acquisition (c, (1)) to a finger in the middle of a simulation (d, (2)), because their persistence is comparable. Taking both geometry and persistence into account, a *lifted* version of the Wasserstein metric associates (c, (1)) to (d, (3)), which is farther away in terms of persistence, but has the nearest maximum, and is qualitatively the best match.
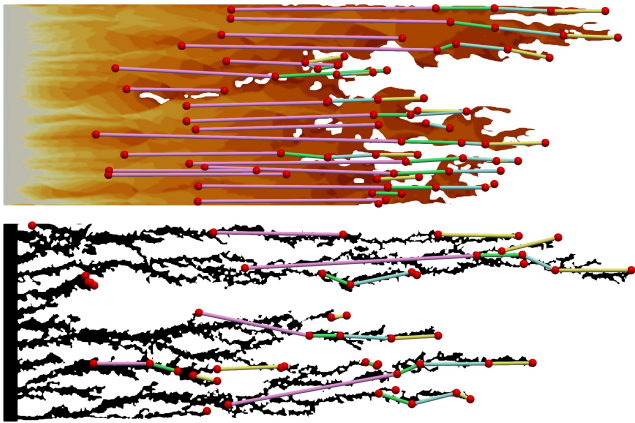

Figure 7: Critical point trajectories based on optimal matchings. Within a given simulation (or the acquisition, bottom), the geometrical coherence of fingers allows us to use a lifted version (that gives importance to the *y*-coordinate of fingers) of the Wasserstein metric to correctly track the evolution of persistence pairs. Comparing the mean distance traveled by fingers between simulations and the acquisition, for each pair of time-steps, is proposed as a *velocity*-aware metric.

slow could compare close to few fast fingers. We then consider the mean traveled distance per finger. Thus, if n_{A_t} (resp. n_{S_t}) denotes the number of fingers in the acquisition (resp. simulation) at time-step t, we propose to evaluate the velocity-oriented difference by:

- $d_{\widecheck{W}_2}(S, A) = \left(\sum_t \left(\frac{1}{n_{S_t}} \widecheck{W}_2(S_t, S_{t+1}) - \frac{1}{n_{A_t}} \widecheck{W}_2(A_t, A_{t+1}) \right)^2 \right)^{1/2}$

Then, given an ensemble of time-varying viscous fingering simulations, each run S can be compared to the reference acquired ground-truth A and runs can be ranked in increasing order of distance to A and presented to the experts for further visual inspection.

3.3 In-situ deployment

Doing feature extraction and comparisons can be problematic for very large numbers of simulations, in terms of data movement. Fortunately, computing the metrics we just presented does not require to have all time-steps available at once, and hence may be done in a progressive fashion. We propose, within our framework, to implement the computation of metrics comparing the acquired reference to the simulation *in-situ*, that is, without storing time-steps to the

disk first. Precomputed persistence diagrams for the acquisition are first loaded in memory. Whenever the simulation attains a time for which there is a corresponding acquisition time step, the saturation scalar field is passed to our analysis pipeline, which applies a threshold, extracts the persistence diagram, and computes the per time step distance to the acquisition diagram (for instance $\widehat{W}_2(S_t, A_t)$). The distance can then be accumulated as the simulation unfolds. The *in-situ* application of our pipeline is optional: time-steps can still be saved to the disk and the pipeline applied *post-mortem* if desired.

3.4 Visual interface

Each metric previously mentioned naturally produces a ranking of simulations, from the most to the less plausible ones. We propose a way to visually inspect those rankings with a lightweight HTML+Javascript application, as illustrated in Fig. 8. Note that we use the same interface for two tasks, to allow experts to *(i)* visually explore the rankings generated by our framework and to *(ii)* produce a ground-truth reference ranking (for the quantitative evaluation, Sec. 4.1). This visual interface offers linked views of the saturation scalar fields, to visually compare simulations runs, for a given time step t which can be interactively selected. If needed (in particular to generate a reference ranking, cf. Sec. 4.1), the experts can interactively modify the suggested ranking by displacing a selected run up or down the ranking, either by unit or long jumps (typically skipping 10 or 50 positions, useful for the a priori reference ranking).

4 CASE STUDY

This section exposes our experimental setting, details a complete viscous fingering use case and summarizes the results of our approach in terms of performance and quality, compared to classical methods.

4.1 Experimental protocol

The behavior of a slab, initially filled with oil and water at connate water saturation, then subject to a water injection in reservoir conditions is captured through X-rays: X-ray images are processed in order to be converted to maps of the fluid saturations within the slab. 2D simulations are then launched with varying input parameters in order to match the simulation results to the experimental measurements and to the fluid saturation maps derived from X-ray images. The resemblance of fingers can be taken into account manually by experts, involving an interpretation of X-rays and an assessment of likeliness according to their expertise. A reference ranking of simulations is then produced by the experts with the help of our visual
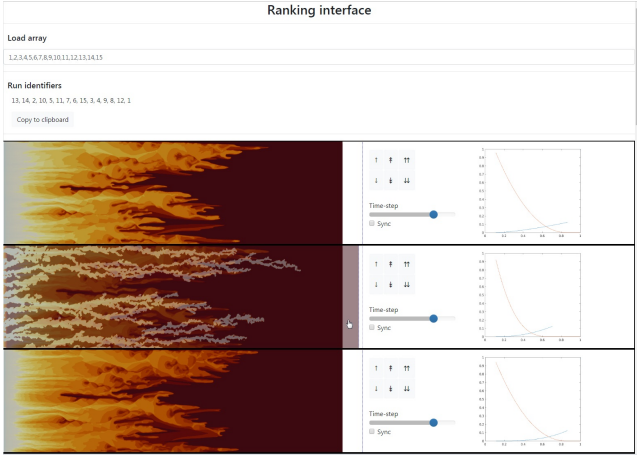
Figure 8: Lightweight web interface for exploring and modifying simulation rankings. An ordered list of runs can be loaded as an input. Time-steps of runs are then displayed on the left pane; they can be hovered with the mouse to be compared with the matching acquired image. A slider allows to select the time-step to compare. Users can edit the ranking with swapping buttons. For each run, kr curves (input parameters of the simulation model) are displayed on the right.
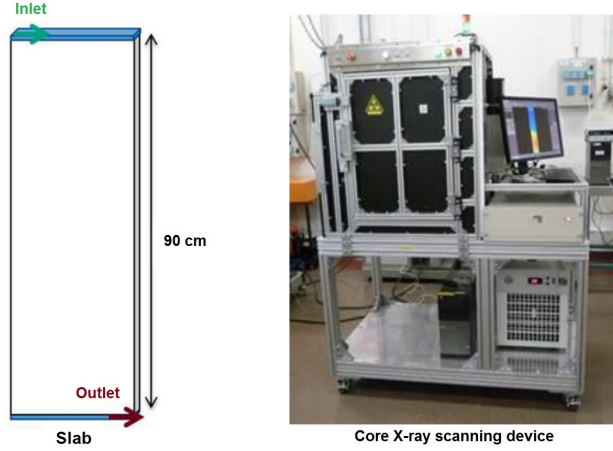

Figure 9: Schematic view of the slab used for the acquisition (left, experimental protocol described in [29]). It is disposed vertically during the capture. On the right, a typical X-ray scanning device for imaging flow in porous media is shown.

interface (Sec. 3.4), and is compared to the rankings generated by the metrics proposed in our framework (Sec. 3.2). The performance and quality of our approach are then evaluated.

Acquisition: The detailed experimental setup is that of [29, 34], further described in [88–90]. We consider slabs of Bentheimer sandstone ($30\times90\times2.45$ cm), with porosity of approximately 23% and absolute permeability of 2.5 Darcy (when $S_w = 1$). The slab is coated with two epoxy layers. Three grooves are cut into the first layer on the extreme faces, and connected to injection and production rails. It is mounted vertically in a 2D X-ray scanning rig (Fig. 9). Slabs first undergo cleaning and calibration processes. A tracer test validates the homogeneous behavior of the slab, then oil is injected to reach initial conditions. The system is then aged at 50°C and ambient pressure for a month, to get closer to field conditions. The water injection rate is kept constant at 3 cm^3/h, which corresponds to the velocity in fields far from wells. One of the fluids is doped with an X-ray absorbing chemical for increasing the contrast. The scanner is equipped with an X-ray source (40 to 60 kV at maximum 0.4 mA) and a camera capturing a slice of 0.5×11.5 cm. The camera moves in horizontal rows along the slab. A scan for a 30×30 cm image takes
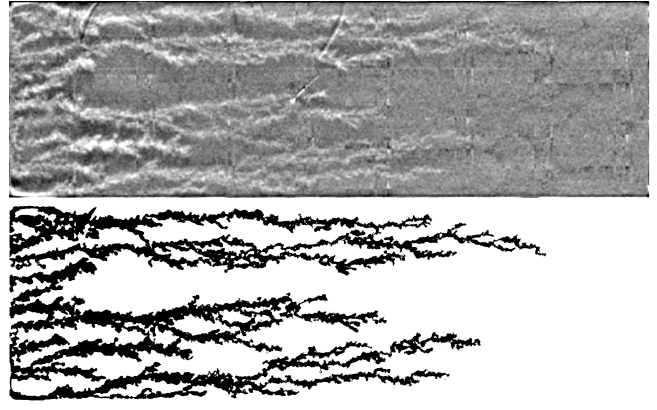

Figure 11: De-noised X-ray capture (top) and segmented fingers (bottom). Fingers were manually detoured by experts.

4-5 min, during which the fluid has moved by about 0.1 to 0.2 mm. The captured images, which are noisy and exhibit severe vertical and horizontal artifacts, are filtered [88], and manually segmented by an expert (Fig. 11) to differentiate fingers from the background, hence forming a reference finger geometry $\mathscr{F}_A$.

Simulations: The input parameters of the 2D simulations are relative permeabilities (kr_w and kr_o). In our model, they are a function of water saturation S_w. We consider relative permeabilities in the form of simple Corey curves (Fig. 10, Eq. 8, [11]), subject to parameters kr_o^0 (oil relative permeability endpoint), S_{or} (residual oil saturation), and power law exponents n_c and n_w. Other quantities like S_{wc} (connate water saturation) and kr_w^0 (water relative permeability endpoint) are determined by measurement.

$$\begin{cases} kr_o(S_w) = kr_o^0 \times \left(\frac{1-S_w-S_{or}}{1-S_{or}-S_{wc}}\right)^{n_c} \\ kr_w(S_w) = kr_w^0 \times \left(\frac{S_w-S_{wc}}{1-S_{or}-S_{wc}}\right)^{n_w} \end{cases} \quad (8)$$

The parameters of these curves, kr_o^0, S_{or}, n_c and n_w, were randomly sampled and selected using the algorithm by Wootton, Sergent, Phan-Tan-Luu [79] to ensure a good initial covering of the space. The geometrical domain is discretized on a regular grid of 290×890 blocks. 200 runs were launched on 400 simulation nodes (2 MPI ranks per run), then time-steps for which there was a corresponding X-ray image were saved (8 available segmented images).
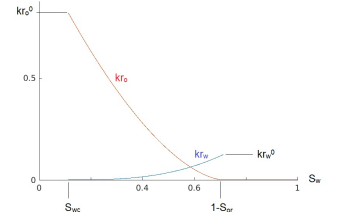

Figure 10: Typical relative permeability curves: for oil (red) and water (blue). In this graph, the X-axis is water saturation S_w. Intuitively, it represents the extent to which the flow of a phase (say the flow of oil) is inhibited by the presence of another (say the presence of water).

Expert ground truth: Images of the water saturation field were captured for each simulation at available X-ray time-steps, for experts to manually form a reference ranking. During this process, experts would quickly discard runs deemed too far from the X-ray image (because fingers are advancing too slow, too fast, or in a too diffusive fashion). Then, they would closely look at the shape and advancement of fingers when comparing two close runs. We use our lightweight web based visual interface (Sec. 3.4, Fig. 8), to alleviate this tedious process. Note that images from all simulations were necessary for the experts to form the reference ranking, so the corresponding simulated time-steps were saved to the disk. Once this reference is formed, later analyses can be done *in-situ*.

Table 1: Time performance comparison (CPU time in seconds), for a single time-step, between the *in-situ* implementation (everything is computed during the simulation using the local CPU's memory) and the *post-mortem* implementation (the scalar fields are analyzed and compared in a post-processing stage).

Step	CPU time (s)		Detailed CPU time (s)	
	in-situ	*post-mortem*		
Simulation iteration	3.096	3.096		
Time step storage		0.076		
Catalyst analysis	1.111		0.063	Persistence diagram
			0.002	Distance
			0.001-	Distance storage
			1.046	Catalyst overhead
Data transfer		0.021		Lustre to workstation
Data conversion		0.246		.unrst to .vtk
Paraview analysis		2.189	0.084	Persistence diagram
			0.002	Distance
			2.085	Paraview overhead
Analysis time	**1.111**	2.532		
Total processing	4.207	5.628		

4.2 In-situ performance

In this section, we evaluate the quantitative gains of using our approach *in-situ* (Sec. 3.3), in terms of time and storage. Tab. 1 provides a CPU time comparison of our analysis pipeline based on the lifted 2-Wasserstein metric (Sec. 3.2), for the two different strategies: *(i) in-situ*, where the analysis is run on the fly during the simulation and without data storage and *(ii) post-mortem*, where selected time steps are stored to disk to be analyzed after the simulation has finished. Persistence diagrams are computed using the algorithm by Gueunet et al. [42], and Wasserstein distances are computed using the exact approach by Soler et al. [92], both available in the Topology ToolKit (TTK, [100]). The *in-situ* implementation is based on Catalyst [3], which is called by the simulation code at selected time steps to run a python script instantiating our analysis pipeline.

The numbers are given for a single simulation time-step, therefore at the finest possible time resolution (about ten thousand time-steps are required to complete a run). Figures are averages on the time-steps of a typical run. *In-situ* computations are done on a super-computer (among the 51st of TOP500 Nov. 2018) with Xeon(R) E5-2680v3 processors, the post-process is done on a local workstation with a Xeon(R) E5-2640v3 processor, so there is a difference in performance obtained for computing persistence diagrams.

Ideally, overheads due to different data layouts and conversions (lines *"Catalyst overhead"* and *"Paraview overhead"*, Tab. 1) in the simulator and VTK/ParaView would not enter into account (if the simulator were to directly output a VTK data array). We are left with two unneeded stages in the *post-mortem* approach: time step writes and data transfer. Selecting 8 time-steps from 200 simulations, this amounts to 155.2 s. of IO time versus approximately 0.6 ms. necessary to write the 1,600 doubles in the *in-situ* case (representing the 1,600 distance estimations), which is 260,000 times faster.

In terms of data storage, the *post-mortem* strategy requires to store and potentially transfer 3.28 GiB (2.1 MiB per time-step) of data, versus 12.5 KiB for 1,600 doubles (representing the 1,600 distance estimations), which is 275,000 times lighter.

Thus, overall, the in-situ instantiation of our framework reduces data movement by 5 orders of magnitude, while dividing by 2.3 the time required to analyze a time-step (line *"Analysis time"*, Tab. 1).

4.3 Ranking quality

In this section, we evaluate quantitatively the relevance of the rankings obtained with each of the metrics discussed in Sec. 3.2, and compare them to rankings obtained with *overlap methods*, traditionally used for associating geometrical sub-domains [9, 10, 78, 82, 91]. Let $\mathscr{F}_{A_t}$ be the acquired finger geometry (Sec. 4.1) and $\mathscr{F}_{S_t}$ be the sub-level set of the simulated water saturation at time t. The overlap $O(A_t, S_t)$ between A_t and S_t is the volume of $\mathscr{F}_{A_t} \cap \mathscr{F}_{S_t}$ divided by the volume of $\mathscr{F}_{A_t} \cup \mathscr{F}_{S_t}$. From this we can define a distance:

Table 2: Quality of rankings. Kendall coefficients between each ranking and the reference ground truth formed by experts are computed (closest to 1 is best). Since the order in which the poorest runs are ordered in the expert's ranking is arbitrary, coefficients are also computed for the (50 and 25) best simulations according to each method. The best coefficient for each case is shown in bold.

Method	O	W_2	$\widehat{W}_2$	EMD	$\check{O}$	$\check{W}_2$
All	0.37	0.25	0.26	0.15	0.12	**0.41**
Top 50	0.22	0.46	**0.66**	0.47	-0.29	0.46
Top 25	0.13	0.29	**0.84**	0.70	-0.13	0.42

- $d_O(A, S) = \left(\sum_t \left(1 - O(A_t, S_t) \right)^2 \right)^{1/2}$

Integrating the overlap $\check{O}_t(S) = 1 - O(S_t, S_{t+1})$ between S_t and S_{t+1} for a single simulation and comparing it to the integrated overlap for the acquisition yields a *velocity*-oriented version:

- $d_{\check{O}}(A, S) = \left(\sum_t \left(\check{O}_t(S) - \check{O}_t(A) \right)^2 \right)^{1/2}$

At this point, we need to compare different rankings to the reference ranking constituted by experts. Let R_1 and R_2 be two rankings of n simulations. One of the most commonly used methods for computing a degree of similarity between R_1 and R_2 is **Kendall's** τ [25, 39]: for all couples $(r_i, r_j) \in R_1^2$ and $(s_i, s_j) \in R_2^2$:

$$\tau = \frac{2}{n(n-1)} \sum_{i<j} \text{sign}(r_i - r_j)\text{sign}(s_i - s_j) \quad (9)$$

It corresponds to the number of pairs (i, j) for which r_i and r_j in R_1 have the same ordering as s_i and s_j in R_2 minus the number of pairs for which the orderings in R_1 and R_2 are different. In other words, it is the difference between the number of concordant pairs and the number of discordant pairs. The closer this number is to 1 in absolute value, the more compatible the rankings, τ being close to -1 indicates that the two rankings are in reverse order.

Over the 200 examined runs, many were quickly discarded by experts during the manual ranking, because they were too far from the acquisition. Thus, as the order of the poorest runs is not important to the experts, we also compute the similarity with the reference ranking for the best (top 50 and top 25) identified runs according to each method. This enables us to focus the evaluation of the performance of our framework in separating plausible from non-plausible runs. Resulting Kendall coefficients are exposed in Tab. 2.

Observing lines 2 and 3 in Tab. 2, we can first note that the overlap method (column "O") does not perform well. This behavior was expected because of the very chaotic geometry of fingers. The Wasserstein method, which is the traditional reference metric for comparing persistence diagrams, is shown in column "W_2". The Earth mover's distance method (column "EMD") only takes the geometrical information of extrema into account, regardless of their persistence. It seems to perform better than W_2, which is unexpected, because EMD can wrongly associate small-scale details to large-scale ones. The lifted Wasserstein method, which includes persistence information and favors a geometrical direction, is shown in column "$\widehat{W}_2$". As it achieves the best overall Kendall coefficients, it seems that $\widehat{W}_2$ manages to combine the advantages of both EMD and W_2, not just being a simple interpolation between the two. Lastly, metrics based on the distances traveled by fingers (columns $\check{O}$ and $\check{W}_2$) do not appear to be able to produce relevant rankings.

4.4 Expert feedback

In this section we expose a qualitative appreciation, collected from experts, and a discussion of ranking results. We show in Tab. 3 the qualitative appreciation of rankings. The poor performance of velocity-based metrics ($\check{O}$ and $\check{W}_2$) was unexpected. Looking at the rankings, we see that aberrant runs are considered close to the ground truth by these two metrics. There are three types of aberrant

Table 3: Appreciation of the top-25 rankings returned by each method. Diffuse runs, slow runs and runs in common with the expert's ranking are counted for each method. Each ranking is shown to an expert using our web interface and their appreciation is noted.

Method	O	W_2	$\widehat{W_2}$	EMD	$\breve{O}$	$\breve{W_2}$
too diffuse	0	4	0	5	0	7
too slow	0	0	0	0	17	0
common	0	10	21	18	0	11
appreciation	poor	good	best	poor	wrong	wrong

Figure 12: Diffuse run example. The tips of fingers grow wider than their base, forming a sort of inverted funnels. The saturation field does not exhibit a very sharp frontier with the background.

runs: too slow, too fast, and too diffusive (i.e. whose finger tips grow large and do not form a very sharp frontier with the background, as illustrated in Fig. 12). $\breve{O}$ gives a good score to runs that are too slow, and $\breve{W_2}$, on the contrary, scores highly runs that are too diffusive.

The number of slow runs is counted in Tab. 3. The $\breve{O}$ approach, based on overlaps in consecutive time-steps, does not discard them. The reason for this is that in simulations, the water saturation front is very smooth though in the acquisition fingers display a quite dendritic structure. Thus, the overlap between successive time-steps of smooth fingers going slow compares close to the overlap between successive time-steps of thin fingers going fast. The number of runs which exhibit a very diffusive behavior is also counted. These diffuse fingers seem to give trouble to the $\breve{W_2}$ metric (and also to *EMD* and W_2). This is because in the set of available simulations, among all which are diffuse some inevitably end up at the exact same advancement as the acquisition when the threshold stage (Sec. 3.1) is applied. Taking into account the number of fingers ($\breve{W_2}$) or considering their branching events (W_2) is apparently insufficient to discard them. Note that the $\widehat{W_2}$ metric (and even W_2) were well appreciated by the experts because the top simulations in their rankings display fingers whose tips are quite close to the acquisition near breakthrough time, though for W_2, there seems to be a higher distance variability. As for the basic overlap method O, it fails to identify the real best simulations, though it does not incorrectly bring out aberrant runs either (be it too diffuse or too slow). Its ranking, though, feels random to the experts. Overall, the best performing metric seems to be $\widehat{W_2}$, as confirmed quantitatively and qualitatively.

Taking a step back, the approach we proposed is appealing to experts because it allows them to include geometrical information into their parametric studies, in an autonomous and systematic way (instead of manually inspecting and checking runs). Using the $\widehat{W_2}$-based ranking results, we present in Fig. 13 all permeability curves and those yielding the best simulation runs. We see no clear pattern arising, either visually or numerically with respect to L_2 or Haussdorff [65] distances between curves. This confirms the well-known difficulty of calibrating *kr* curves.

5 CONCLUSION

In this application paper, we presented a framework for enabling the automatic comparison and ranking of simulation runs to an acquired ground truth. We presented a set of metrics specifically adapted to this task in the case of viscous fingering in porous media. After
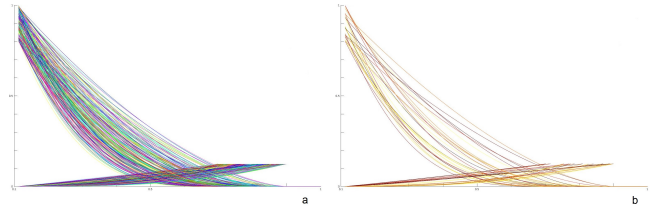

Figure 13: Input relative permeability curves of all 200 simulations (left, with random colors) and for the 25 best selected runs (right, darker is closest to ground truth). No clear pattern was seen that could discriminate relative permeability curves yielding the best fingers.

evaluation, we identified the best fitting approach (the $\widehat{W_2}$ metric, which computes a geometrically tuned Wasserstein distance between simulation and acquisition persistence diagrams, on a per-time-step basis). This quantitative measurement method supplements the expert's, and allows them to automatically form a subjective ranking close to one they would have manually produced. We demonstrated the possibility and showed the advantage of implementing the computation of this metric in-situ, speeding up the analysis pipeline by a factor of 2.3 and reducing data movement by 5 orders of magnitude. We proposed a lightweight web interface to explore automatically generated rankings and manually edit them. As with the best metric $\widehat{W_2}$, there are still some diffuse runs in the ranked best fifty, we believe it could be further enhanced, for instance by considering the sharpness of the water saturation front, or by augmented the persistence diagram with the individual volume of fingers. Besides, though in our experimental setting, simulations took place in a 2D domain, nothing in our approach is restrictive to this case. In future work, our framework could be experimented with simulation models other than Darcy's. Moreover, it could also be applied to 3D cases. Our overall approach would be usable as-is in these scenarios, but its meta-parameters (water saturation threshold and lifting coefficients) would likely need to be adjusted. Automatically optimizing the values of these meta-parameters is also a promising direction for future work. Furthermore, the metrics introduced in this work have been mostly motivated empirically based on interactions with domain experts. In the future, we will consider a theoretical investigation of the stability of these metrics (based on stability results on the Bottleneck [19] and Wasserstein [20] metrics). Further, ways to capture finger merging events, for instance by considering richer topological structures such as Reeb Graphs, could also be considered, as this seems to happen in acquired images.

On another note, on the set of 200 simulations, we were not able to identify a regime of best-matching input parameters. The combination of our metric with production and pressure data (at injectors/producers) in a follow-up study would be interesting in this regard. Trying to understand the influence of the space of input parameters, here *kr* curves, proves quite challenging. In our study, only four parameters (power law exponents and endpoints) were sampled, yielding a four-dimensional space, but the number of sampled parameter may be significantly higher. In particular, this study may be extended to permeability curves other than based on simple Corey power laws. We think it would be insightful to develop a visual interface for exploring such spaces of model parameters. We hope to see, in future studies, how accounting for the geometrical and topological quality of a modeled phenomenon can be used to infer or restrict model parameters.

REFERENCES

[1] F. O. Alpak, L. W. Lake, S. M. Embid, et al. Validation of a modified carman-kozeny equation to model two-phase relative permeabilities. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, 1999.

[2] U. Ayachit, A. Bauer, E. P. Duque, G. Eisenhauer, N. Ferrier, J. Gu, K. E. Jansen, B. Loring, Z. Lukić, S. Menon, et al. Performance analysis, design considerations, and applications of extreme-scale in situ infrastructures. In *SuperComputing*, 2016.

[3] U. Ayachit, A. Bauer, B. Geveci, P. O'Leary, K. Moreland, N. Fabian, and J. Mauldin. Paraview catalyst: Enabling in situ data analysis and visualization. In *Proc. of the First Workshop on In Situ Infrastructures for Enabling Extreme-Scale Analysis and Vis.*, pp. 25–29. ACM, 2015.

[4] L. Baker et al. Three-phase relative permeability correlations. In *SPE Enhanced Oil Recovery Symposium*, 1988.

[5] J. C. Bennett, H. Abbasi, P. T. Bremer, R. Grout, A. Gyulassy, T. Jin, S. Klasky, H. Kolla, M. Parashar, V. Pascucci, P. Pebay, D. Thompson, H. Yu, F. Zhang, and J. Chen. Combining in-situ and in-transit processing to enable extreme-scale scientific analysis. In *SC*, 2012.

[6] H. Bhatia, A. G. Gyulassy, V. Lordi, J. E. Pask, V. Pascucci, and P.-T. Bremer. Topoms: Comprehensive topological exploration for molecular and condensed-matter systems. *J. of Comp. Chem.*, 2018.

[7] S. Biasotti, D. Giorgio, M. Spagnuolo, and B. Falcidieno. Reeb graphs for shape analysis and applications. *TCS*, 2008.

[8] A. Bock, H. Doraiswamy, A. Summers, and C. Silva. Topoangler: Interactive topology-based extraction of fishes. *IEEE TVCG*, 2018.

[9] P. Bremer, G. Weber, J. Tierny, V. Pascucci, M. Day, and J. Bell. Interactive exploration and analysis of large scale simulations using topology-based data segmentation. *IEEE TVCG*, 2011.

[10] P. T. Bremer, G. Weber, V. Pascucci, M. Day, and J. Bell. Analyzing and tracking burning structures in lean premixed hydrogen flames. *IEEE TVCG*, 2010.

[11] R. H. Brooks and A. T. Corey. Hydraulic properties of porous media and their relation to drainage design. *Transactions of the ASAE*, 1964.

[12] F. M. Carlson et al. Simulation of relative permeability hysteresis to the nonwetting phase. In *SPE annual technical conference and exhibition*. Society of Petroleum Engineers, 1981.

[13] H. Carr, J. Snoeyink, and U. Axen. Computing contour trees in all dimensions. *Computational Geometry*, 24(2):75–94, 2003.

[14] H. Carr, J. Snoeyink, and M. van de Panne. Simplifying flexible isosurfaces using local geometric measures. In *IEEE VIS*, 2004.

[15] F. Chazal, D. Cohen-Steiner, M. Glisse, L. J. Guibas, and S. Oudot. Proximity of Persistence Modules and their Diagrams. In *SoCG*, 2009.

[16] C.-C. Chen and H.-T. Chu. Similarity measurement between images. In *29th Annual International Computer Software and Applications Conference (COMPSAC'05)*, vol. 2, pp. 41–42. IEEE, 2005.

[17] Z. Chen, G. Huan, and B. Li. An improved IMPES method for two-phase flow in porous media. *Transport in porous media*, 2004.

[18] G. L. Chierici et al. Novel relations for drainage and imbibition relative permeabilities. *Soc. of Petroleum Engineers Journal*, 1984.

[19] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. In *Symp. on Comp. Geom.*, pp. 263–271, 2005.

[20] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, and Y. Mileyko. Lipschitz functions have L_p-stable persistence. *Foundations of Computational Mathematics*, 2010.

[21] D. Cohen-Steiner, H. Edelsbrunner, and D. Morozov. Vines and vineyards by updating persistence in linear time. In *Symp. on Comp. Geom.*, 2006.

[22] D. A. S. C. A. Committee. Synergistic challenges in data-intensive science and exascale computing. Technical report, DoE Advanced Scientific Computing Advisory Committee, Data Sub-committee, 2013.

[23] A. T. Corey et al. The interrelation between gas and oil relative permeabilities. *Producers monthly*, 1954.

[24] A. T. Corey, C. Rathjens, et al. Effect of stratification on relative permeability. *Journal of Petroleum Technology*, 8(12):69–71, 1956.

[25] C. Croux and C. Dehon. Influence functions of the spearman and kendall correlation measures. *Stat. Methods & Applications*, 2010.

[26] M. Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Proc. of NIPS*, 2013.

[27] H. Darcy. *Les fontaines publiques de la ville de Dijon*. 1856.

[28] L. De Floriani, U. Fugacci, F. Iuricich, and P. Magillo. Morse complexes for shape segmentation and homological analysis: discrete models and algorithms. *Comp. Grap. For.*, 2015.

[29] R. de Loubens, G. Vaillant, M. Regaieg, J. Yang, A. Moncorgé, C. Fabbri, G. Darche, et al. Numerical modeling of unstable waterfloods and tertiary polymer floods into highly viscous oils. *SPE Journal*, 2018.

[30] D. A. DiCarlo, S. Akshay, M. Blunt, et al. Three-phase relative permeability of water-wet, oil-wet, and mixed-wet sandpacks. *SPE Journal*, 5(01):82–91, 2000.

[31] H. Edelsbrunner and J. Harer. Persistent homology – a survey. *American Mathematical Society*, 2008.

[32] H. Edelsbrunner and J. Harer. *Computational Topology: An Introduction*. American Mathematical Society, 2009.

[33] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Disc. Compu. Geom.*, 2002.

[34] C. Fabbri, R. De Loubens, A. Skauge, P. Ormehaug, B. Vik, M. Bourgeois, D. Morel, G. Hamon, et al. Comparison of history-matched water flood, tertiary polymer flood relative permeabilities and evidence of hysteresis during tertiary polymer flood in very viscous oils. In *SPE Asia Pacific Enhanced Oil Recovery Conference*, 2015.

[35] G. Favelier, N. Faraj, B. Summa, and J. Tierny. Persistence atlas for critical point variability in ensembles. *IEEE TVCG*, 2019.

[36] G. Favelier, C. Gueunet, and J. Tierny. Visualizing ensembles of viscous fingers. In *IEEE Scientific Visualization Contest*, 2016.

[37] D. H. Fenwick, M. J. Blunt, et al. Network modeling of three-phase flow in porous media. *SPE Journal*, 3(01):86–96, 1998.

[38] C. H. Gao et al. Advances of polymer flood in heavy oil recovery. In *SPE heavy oil conference and exhibition*, 2011.

[39] A. Göktas and Ö. Isçi. A comparison of the most commonly used measures of association for doubly ordered square contingency tables via simulation. *Metodoloski zvezki*, 8(1):17, 2011.

[40] D. Guenther, R. Alvarez-Boto, J. Contreras-Garcia, J.-P. Piquemal, and J. Tierny. Characterizing molecular interactions in chemical systems. *IEEE TVCG*, 20(12):2476–2485, 2014.

[41] C. Gueunet, P. Fortin, J. Jomier, and J. Tierny. Contour forests: Fast multi-threaded augmented contour trees. In *IEEE LDAV*, 2016.

[42] C. Gueunet, P. Fortin, J. Jomier, and J. Tierny. Task-based augmented merge trees with Fibonacci heaps. In *IEEE LDAV*, 2017.

[43] A. Gyulassy, P. T. Bremer, B. Hamann, and V. Pascucci. A practical approach to morse-smale complex computation: Scalability and generality. *IEEE TVCG*, 14(6):1619–1626, 2008.

[44] A. Gyulassy, D. Guenther, J. A. Levine, J. Tierny, and V. Pascucci. Conforming morse-smale complexes. *IEEE TVCG*, 2014.

[45] A. Gyulassy, A. Knoll, K. Lau, B. Wang, P. Bremer, M. Papka, L. A. Curtiss, and V. Pascucci. Interstitial and interlayer ion diffusion geometry extraction in graphitic nanosphere battery materials. *IEEE TVCG*, 2015.

[46] J. Hagoort et al. Oil recovery by gravity drainage. *Society of Petroleum Engineers Journal*, 20(03):139–150, 1980.

[47] C. Heine, H. Leitte, M. Hlawitschka, F. Iuricich, L. De Floriani, G. Scheuermann, H. Hagen, and C. Garth. A survey of topology-based methods in visualization. *Comp. Grap. For.*, 2016.

[48] G. M. Homsy. Viscous fingering in porous media. *Annual review of fluid mechanics*, 19(1):271–311, 1987.

[49] S. Jaure, A. Moncorge, R. de Loubens, et al. Reservoir simulation prototyping platform for high performance computing. In *SPE Large Scale Computing and Big Data Challenges in Reservoir Simulation Conference and Exhibition*. Society of Petroleum Engineers, 2014.

[50] L. Kantorovich. On the translocation of masses. *AS USSR*, 1942.

[51] J. Kasten, J. Reininghaus, I. Hotz, and H. Hege. Two-dimensional time-dependent vortex regions based on the acceleration magnitude. *IEEE TVCG*, 2011.

[52] M. Kerber, D. Morozov, and A. Nigmetov. Geometry helps to compare persistence diagrams. *J. Exp. Algorithmics*, 2017.

[53] J. Killough et al. Reservoir simulation with history-dependent saturation functions. *Society of Petroleum Engineers Journal*, 1976.

[54] A. Landge, V. Pascucci, A. Gyulassy, J. Bennett, H. Kolla, J. Chen, and T. Bremer. In-situ feature extraction of large scale combustion

simulations using segmented merge trees. In *SuperComputing*, 2014.

[55] H. Lavenant, S. Claici, E. Chien, and J. Solomon. Dynamical optimal transport on discrete surfaces. In *SIGGRAPH Asia 2018 Technical Papers*, p. 250. ACM, 2018.

[56] M. Leverett et al. Capillary behavior in porous solids. *Transactions of the AIME*, 142(01):152–169, 1941.

[57] E. Levina and P. Bickel. The earthmover's distance is the mallows distance: some insights from statistics. In *IEEE ICCV*, 2001.

[58] T. Luciani, A. Burks, C. Sugiyama, J. Komperda, and G. E. Marai. Details-first, show context, overview last: Supporting exploration of viscous fingers in large-scale ensemble simulations. *IEEE TVCG*, 25(1):1225–1235, 2019.

[59] J. Lukasczyk, G. Aldrich, M. Steptoe, G. Favelier, C. Gueunet, J. Tierny, R. Maciejewski, B. Hamann, and H. Leitte. Viscous fingering: A topological visual analytic approach. In *Applied Mechanics and Materials*, vol. 869, pp. 9–19. Trans Tech Publ, 2017.

[60] D. J. MacAllister, K. C. Miller, S. K. Graham, C.-T. Yang, et al. Application of x-ray ct scanning to determine gas/water relative permeabilities. *SPE formation evaluation*, 8(03):184–188, 1993.

[61] G. Monge. Mémoire sur la théorie des déblais et des remblais. *Académie Royale des Sciences de Paris*, 1781.

[62] K. Moreland, R. Oldfield, P. Marion, S. Jourdain, N. Podhorszki, V. Vishwanath, N. Fabian, C. Docan, M. Parashar, M. Hereld, M. E. Papka, and S. Klasky. Examples of in transit visualization. In *Proc. of the 2nd Int. Workshop on Petascale Data Analytics: Challenges and Opportunities*, PDAC '11, pp. 1–6. ACM, 2011.

[63] D. Morozov. Dionysus. http://www.mrzv.org/software/dionysus, 2010. Accessed: 2016-09-15.

[64] J. Munkres. Algorithms for the assignment and transportation problems. *J. of the Society for Industrial and Applied Mathematics*, 1957.

[65] J. Munkres. *Topology*. Pearson Education, 2014.

[66] M. Oak, L. Baker, D. Thomas, et al. Three-phase relative permeability of berea sandstone. *Journal of Petroleum Technology*, 1990.

[67] P. OLeary, J. Ahrens, S. Jourdain, S. Wittenburg, D. H. Rogers, and M. Petersen. Cinema image-based in situ analysis and visualization of mpas-ocean simulations. *Parallel Computing*, 55:43–48, 2016.

[68] V. Pascucci, G. Scorzelli, P. T. Bremer, and A. Mascarenhas. Robust on-line computation of Reeb graphs: simplicity and speed. *ToG*, 2007.

[69] V. Pascucci, X. Tricoche, H. Hagen, and J. Tierny. *Topological Data Analysis and Visualization: Theory, Algorithms and Applications*. Springer, 2010.

[70] L. Patacchini, R. De Loubens, A. Moncorge, A. Trouillaud, et al. Four-fluid-phase, fully implicit simulation of surfactant flooding. *SPE Reservoir Evaluation & Engineering*, 17(02):271–285, 2014.

[71] M. Rasquin, P. Marion, V. Vishwanath, B. Matthews, M. Hereld, K. Jansen, R. Loy, A. Bauer, M. Zhou, O. Sahni, J. Fu, N. Liu, C. Carothers, M. Shephard, M. Papka, K. Kumaran, and B. Geveci. Electronic poster: Co-visualization of full data and in situ data extracts from unstructured grid cfd at 160k cores. In *Proc. of the 2011 Companion on HPC Networking, Storage and Analysis Companion*, SC '11 Companion, pp. 103–104. ACM, 2011.

[72] G. Reeb. Sur les points singuliers d'une forme de Pfaff complètement intégrable ou d'une fonction numérique. *Acad. des Sci.*, 1946.

[73] A. Riaz, G.-Q. Tang, H. A. Tchelepi, and A. R. Kovscek. Forced imbibition in natural porous media: Comparison between experiments and continuum models. *Physical Review E*, 75(3):036305, 2007.

[74] J. Richardson, J. Kerver, J. Hafford, J. Osoba, et al. Laboratory determination of relative permeability. *J. of Petroleum Tech.*, 1952.

[75] M. Rivi, L. Calori, G. Muscianisi, and V. Slavnic. In-situ visualization: State-of-the-art and some use cases. 2011.

[76] V. Robins, P. Wood, and A. Sheppard. Theory and algorithms for constructing discrete morse complexes from grayscale digital images. *IEEE Trans. on Pat. Ana. and Mach. Int.*, 2011.

[77] P. G. Saffman and G. I. Taylor. The penetration of a fluid into a porous medium or hele-shaw cell containing a more viscous liquid. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 245(1242):312–329, 1958.

[78] H. Saikia and T. Weinkauf. Global feature tracking and similarity estimation in time-dependent scalar fields. *Comp.. Graph. For.*, 2017.

[79] J. Santiago, M. Claeys-Bruno, and M. Sergent. Construction of

space-filling designs using wsp algorithm for high dimensional spaces. *Chemometrics and Intelligent Laboratory Systems*, 113:26–31, 2012.

[80] J. Sharma, S. B. Inwood, A. Kovscek, et al. Experiments and analysis of multiscale viscous fingering during forced imbibition. *SPE Journal*, 2012.

[81] N. Shivashankar, P. Pranav, V. Natarajan, R. van de Weygaert, E. P. Bos, and S. Rieder. Felix: A topology based framework for visual exploration of cosmic filaments. *IEEE TVCG*, 2016.

[82] D. Silver. Object-oriented visualization. *IEEE Comput. Graph. Appl.*, 15(3):54–62, May 1995.

[83] D. Silver and X. Wang. Volume tracking. In *Visualization '96. Proceedings.*, pp. 157–164, Oct 1996.

[84] D. Silver and X. Wang. Tracking and visualizing turbulent 3d features. *IEEE TVCG*, 3(2):129–141, Apr 1997.

[85] D. Silver and X. Wang. Tracking scalar features in unstructured data sets. In *Visualization '98. Proceedings*, pp. 79–86, Oct 1998.

[86] D. Silver and X. Wang. Visualizing evolving scalar phenomena. *Future Generation Computer Systems*, 15(1):99 – 108, 1999.

[87] A. Skauge, T. Horgen, B. Noremark, and B. Vik. Experimental studies of unstable displacement in carbonate and sandstone material. In *IOR 2011-16th European Symposium on Improved Oil Recovery*, 2011.

[88] A. Skauge, P. A. Ormehaug, T. Gurholt, B. Vik, I. Bondino, G. Hamon, et al. 2-d visualisation of unstable waterflood and polymer flood for displacement of heavy oil. In *Improved Oil Recovery*, 2012.

[89] A. Skauge, K. Sorbie, P. A. Ormehaug, and T. Skauge. Experimental and numerical modeling studies of viscous unstable displacement. In *European Symposium on Improved Oil Recovery*, 2009.

[90] T. Skauge, B. F. Vik, P. A. Ormehaug, B. K. Jatten, V. Kippe, I. Skjevrak, D. C. Standnes, K. Uleberg, A. Skauge, et al. Polymer flood at adverse mobility ratio in 2d flow by x-ray visualization. In *SPE EOR Conference at Oil and Gas West Asia*, 2014.

[91] B. S. Sohn and C. Bajaj. Time-varying contour topology. *IEEE TVCG*, 12(1):14–25, 2006.

[92] M. Soler, M. Plainchault, B. Conche, and J. Tierny. Lifted wasserstein matcher for fast and robust topology tracking. In *IEEE Symposium on Large Data Analysis and Visualization*, 2018.

[93] M. Soler, M. Plainchault, B. Conche, and J. Tierny. Topologically controlled lossy compression. In *2018 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 46–55. IEEE, 2018.

[94] J. Solomon, F. de Goes, G. Peyré, M. Cuturi, A. Butscher, A. Nguyen, T. Du, and L. J. Guibas. Convolutional wasserstein distances: efficient optimal transportation on geometric domains. *ACM ToG*, 2015.

[95] J. Solomon, G. Peyré, V. G. Kim, and S. Sra. Entropic metric alignment for correspondence problems. *ACM Trans. Graph.*, 2016.

[96] S. W. Son, Z. Chen, W. Hendrix, A. Agrawal, W. k. Liao, and A. Choudhary. Data compression for the exascale computing era - survey. *Supercomputing Frontiers and Innovations*, 1(2), 2014.

[97] T. Sousbie. The persistent cosmic web and its filamentary structure: Theory and implementations. *Royal Astronomical Society*, 2011.

[98] J. Thomeer et al. Introduction of a pore geometrical factor defined by the capillary pressure curve. *Journal of Petroleum Technology*, 1960.

[99] J. Tierny and H. Carr. Jacobi fiber surfaces for bivariate Reeb space computation. *IEEE TVCG*, 23(1):960–969, 2016.

[100] J. Tierny, G. Favelier, J. A. Levine, C. Gueunet, and M. Michaux. The Topology ToolKit. *IEEE TVCG*, 24(1):832–842, 2017.

[101] J. Tierny, A. Gyulassy, E. Simon, and V. Pascucci. Loop surgery for volumetric meshes: Reeb graphs reduced to contour trees. *IEEE TVCG*, 15(6):1177–1184, 2009.

[102] M. Trojer, M. L. Szulczewski, and R. Juanes. Stabilizing fluid-fluid displacements in porous media through wettability alteration. *Physical Review Applied*, 3(5):054008, 2015.

[103] T. Tsuji, F. Jiang, and K. T. Christensen. Characterization of immiscible fluid displacement processes with various capillary numbers and viscosity ratios in 3d natural sandstone. *Adv. in Water Resou.*, 2016.

[104] S. Whitaker. Flow in porous media i: A theoretical derivation of darcy's law. *Transport in porous media*, 1(1):3–25, 1986.

[105] H. Yu, C. Wang, R. W. Grout, J. H. Chen, and K. L. Ma. In situ visualization for large-scale combustion simulations. *IEEE Computer Graphics and Applications*, 30(3):45–57, May 2010.