Appeared in: IEEE Transactions on Visualization and Computer Graphics, 18(12):2041–2050, 2012.

Augmented Topological Descriptors of Pore Networks for Material Science

Daniela M. Ushizima, *Member, IEEE*, Dmitriy Morozov, Gunther H. Weber, *Associate Member, IEEE*, Andrea G.C. Bianchi, James A. Sethian, and E. Wes Bethel, *Member, IEEE*

Abstract—One potential solution to reduce the concentration of carbon dioxide in the atmosphere is the geologic storage of captured CO_2 in underground rock formations, also known as carbon sequestration. There is ongoing research to guarantee that this process is both efficient and safe. We describe tools that provide measurements of media porosity, and permeability estimates, including visualization of pore structures. Existing standard algorithms make limited use of geometric information in calculating permeability of complex microstructures. This quantity is important for the analysis of biomineralization, a subsurface process that can affect physical properties of porous media. This paper introduces geometric and topological descriptors that enhance the estimation of material permeability. Our analysis framework includes the processing of experimental data, segmentation, and feature extraction and making novel use of multiscale topological analysis to quantify maximum flow through porous networks. We illustrate our results using synchrotron-based X-ray computed microtomography of glass beads during biomineralization. We also benchmark the proposed algorithms using simulated data sets modeling jammed packed bead beds of a monodispersive material.

Index Terms—Reeb graph, persistent homology, topological data analysis, geometric algorithms, segmentation, microscopy.

1 Introduction

Over the past centuries, the atmospheric carbon dioxide concentration has steadily increased. Since 1751, nearly 337 billion tons of CO_2 produced by the combustion of fossil fuels and cement production were emitted into the atmosphere [4]. The growing consensus is that these increases in CO_2 are disruptive to the earth's climate, and that a reduction in atmospheric CO_2 is necessary to avoid catastrophic effects for the environment. While research in alternate energy production is vital in finding new ways to generate energy without producing CO_2, a large amount of existing infrastructure and technology are tied to the combustion of fossil fuels. Consequently, there is a need to remove carbon after combustion processes and to store it for long periods of time. One vehicle to this solution is carbon sequestration. By carbon sequestration, we refer to the process of first capturing CO_2 produced by combustion and then storing it in underground rock formations via subterranean injection.

The objective of the Energy Frontier Research Center (EFRC) for Nanoscale Control of Geologic CO_2 (NCGC) is to investigate the processes that play a role in the geologic sequestration of carbon dioxide in fluid-rock systems. Such investigations are necessary to ensure that the technology developed for storing CO_2 in deep subsurface rock formations is safe and effective. To gain a better understanding of these processes, researchers of the NCGC conduct experiments and perform simulations that model the interaction of CO_2 with rock systems, investigating processes in fluid-rock systems from the molecular scale to the pore scale [17]. They generate large amounts of experimental data for different parameters—e.g., a variation of brine mixtures at different temperature and pressure values in porous media—while keeping track of immiscible fluid structures, dissolution, and precipitation [2, 3]. Numerical computer models calculate fluid-rock mechanics at pore scales, infiltration stability and instabilities, and emergent structures. Understanding these processes is key in describing the flow and reactive transport of CO_2-rich fluids in geologic reservoirs and to develop approaches for controlling the flow of CO_2 in reservoirs that effectively and efficiently fill pore space with injected CO_2. The EFRC also utilizes the Advanced Light Source (ALS) at LBNL, particularly its Beamline 8.3.2, for the tomographic imaging of fluid invasions of 3D porous materials, precipitation, and dissolution [2].

Our work, in collaboration with EFRC, includes analysis and visualization of both experimental data and output of numerical simulations. One area of particular interest is characterizing pore morphology, a key material descriptor, which is useful for carbon sequestration because porous media are present both in the subsurface as well as in the prepared composites. While there are algorithms to estimate media porosity and permeability from measured image data [5, 25, 3], these algorithms make limited use of geometric information coming from experiments using microtomography.

In this paper, we introduce a new approach to characterize porous materials using extraction of pore networks and pockets by topological algorithms. Our analysis framework uses a multi-stage pipeline to extract information about pore networks from a sample. The input consists of cross-sectional images resulting from high resolution X-ray computed tomography (microCT). In the first step, we compute a Boolean segmentation into material and void space from these slices. This segmentation serves as the input to our topological analysis of channels through the medium.

Porous media properties, such as permeability, can be measured in different directions. In our application, the depth-direction is assumed: we are interested in the flow through the material from top to bottom. Taking advantage of this explicit orientation, we use a topological construction, a Reeb graph, to describe the structure of the interstitial volume, i.e., void or pore space. To obtain a multiscale descriptor, our algorithm finds pores with some minimum diameter. Specifically, fixing a width threshold α, we restrict our view to the pores wide enough to accommodate a sphere of radius α. Our graph describes the distinct paths through the material such spheres can pervade. As we vary the width threshold α, we obtain a family of graphs providing a multiscale representation of channels through the medium.

Based on this family of graphs, we characterize the flow behavior through the sample. For this purpose, we augment the edges of each graph with their "widths," analogous to approximating their capacity for carrying flow. By computing the maximum flow through this graph, we obtain a single number that describes the flow behavior

- *All authors are with the Computational Research Division, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720. E-Mail: {DUshizima, DMorozov, GHWeber, AGCBianchi, JASethian, EWBethel}@lbl.gov.*
- *G.H. Weber is also with the Department of Computer Science, University of California, Davis, One Shields Avenue, Davis, CA 95616. E-Mail: ghweber@ucdavis.edu.*
- *A.G.C. Bianchi is also with the Department of Physics, University of Ouro Preto, Ouro Preto, MG, Brazil. E-Mail: andrea@iceb.ufop.br.*
- *J.A. Sethian is with the Department of Mathematics, University of California, Berkeley, 970 Evans Hall #3840, Berkeley, CA 94720. E-Mail: sethian@math.berkeley.edu.*

through the pores of prescribed thickness.

The flow graph computation attributes a pervasiveness value to each edge in the graph. Also, an edge with a large capacity can only carry as much of the flow as its inlets and outlets support. In particular, a dead-end edge with no outlets carries no flow, no matter what its capacity. Thus, the graph augmented with this information describes how much an individual channel contributes to the overall aggregate flow through the pore network contained within a given material sample. This structure emphasizes significant properties by determining a simplified pore network, represented by a graph with a subset of the connected edges linking the source (top of the stack) to the sink (bottom of the stack). Our contribution is to propose a descriptor to enable a comparative analysis of materials by deriving measures from these graphs, e.g., the maximum flow curve, with the advantage of being relatively robust to small undulations on the object surface and often incurring a lower computational cost [11] in comparison with graphs obtained from a medial axis.

In order to illustrate the most prominent pore bodies of the material, we detect pockets inside of the pore channels and represent them as spheres that fit inside the channels. Each pocket is characterized by two values: the radius of the sphere of maximum size, enclosed in the pocket, and the maximum radius of a sphere that can pass through the pocket. The pair of these values characterizes the prominence of a pocket. Larger differences in radii signify more pronounced and thus, more significant pockets.

To show that these geometric and topological descriptors enhance our ability to estimate material permeability, we isolate microtomography artifacts from our analysis. For this purpose, we run molecular dynamics algorithms, as described by Skoge et al. [20], to obtain numerical constructions of jammed packed bead beds, with identical and non-overlapping spheres. The synthetic data sets resemble a monodispersive material and exhibit complex porous structure. Since it is possible to calculate porosity and permeability for these synthetic materials both analytically and numerically, these data sets serve as a valuable benchmark and calibration of the proposed topological methods. Subsequently, we test our framework on real experimental data, acquired using high-resolution, synchrotron-based X-ray microtomography of glass beads during biomineralization.

We review previous work in porous media description, particularly focusing on methods to extract pore networks in Sect. 2. In Sect. 3, we describe both simulated materials and an experimental sample, which are processed using software capabilities for segmenting microCT data into dense material and void space, discussed in Sect. 4. After defining the interstitial volume, our algorithms extract topological structures, such as the Reeb graph, persistence diagrams, and maximum flow (max-flow) curves, as detailed in Sect. 5. We illustrate our results in Sect. 6 by visualizing detected structures, calculating porosity and permeability of eleven different samples, and showing graphs displaying the permeability decay for homogeneous precipitation using max-flow curves. We further explain how to use these curves to compare materials at different stages of biomineralization. Finally, we conclude our investigation with future developments of the proposed tools and perspectives for a broader application of these new technologies.

2 Background and Related Work

2.1 Porosity and Permeability of Materials

In the context of earth sciences, permeability is an important measure that characterizes how much fluid can pass through a porous material. Porosity—the percentage of void space in a material—is used in the calculation of permeability as part of the Kozeny–Carman (KC) equation. KC is a permeability formulation stating that the flow is proportional to the pressure drop and inversely proportional to the fluid viscosity. According to the literature [25, 3], the KC relation provides a poor fit for permeability data measured using a differential pressure transducer, and pore geometry extracted from microCT data can improve the accuracy of the permeability calculation.

While the estimation of permeability for porous media remains a challenging issue from both theoretical and experimental points of view [23], the permeability of beds of close packed spheres has been extensively studied as one of the basic models of porous media. Particularly for monodispersed spheres, it is possible to calculate KC, as described by Zaman et al. [25], as

$$\kappa = d^2\theta^3/180(1-\theta)^2, \tag{1}$$

where κ is permeability, θ is porosity and d is the sphere diameter. Other models include the Rumpf and Gupte models [25], whose formulation is valid only for porosities ranging from 0.35 to 0.70. Our synthetic data sets go beyond the validity of these models, therefore we limit our calculations of Eq. 1 to calculate κ for simulated porous materials and we establish a parallel between permeability and max-flow curves through the use of topological networks as discussed in Sect. 5.

2.2 Geometric Pore Space Characterization

Modern approaches to studying permeability focus on incorporating pore-space geometry in derived measures. To this end, they characterize pore geometry as a network of connected pockets (pore bodies). To extract geometry, most existing techniques examine the medial axis of the pores, i.e., the set of points inside the pores that do not have a unique closest point on the material. In three-dimensional space, the medial axis is a two-dimensional stratified space. Typically, models of a material channel network are used to describe aggregate flow characteristics, and thus, it becomes necessary to thin the medial axis into a graph structure [16, 15, 19] describing the connectivity of pore space. In contrast, the immediate output of our algorithm is a graph requiring no such post-processing. Moreover, to distinguish between the channels of different widths, we construct a family of graphs which has an immediate geometric interpretation. Similar to Jones and Ma [14], we propose an exploratory visualization scheme to emphasize relevant flow trajectories with applications to porous media. However, instead of focusing on a locality-based visualization of interactions between the geometry of solid structures and the flow trajectory, our method aims to extract the pore network for estimating the fluid flow using topological descriptors.

Silin and Patzek [19] study the connectivity of pockets—or pore bodies—within the material, which they define, in different terms, as the maxima of the distance function to the material; we use a similar definition in Sect. 5. When studying pockets, the main challenge is in pruning noisy maxima that do not reflect the features of the pore space. This pruning is where our works diverge: they merge pockets when their largest inscribed spheres overlap. On the other hand, we rely on the theory of persistent homology [7] to distinguish between significant and noisy pockets, letting the user make a decision about significance from their distribution.

The representation in [19] relies on a stick-and-ball simplification of the pore network, where spheres represent pockets (pore bodies) and sticks signify pore throats that connect them. Both in their and our representation the spheres have the same meaning. However, instead of connecting pore bodies with straight line segments, we use a Reeb graph as the representation of the network; we draw edges so that they follow the centroids of the connected regions, corresponding to the 3D interstitial space. We further augment edges of the Reeb graph with capacity information (the cross-sectional area measured in pixels). After computing maximum flow through the graph with respect to these capacities, we eliminate edges that do not contribute to the flow, thus eliminating dead ends and highlighting trajectories in the material that are potentially involved in the flow.

Lindquist et al. [16] characterize porous material properties using the medial axis. They compute the latter using a front-based approach that starts "fires" at grain boundaries, propagates into voids and leaves only pore voxels where fire arriving from opposite directions cancels itself out. By labeling "fires" based on connected grains, and requiring fires to have different labels for canceling each other out, their approach eliminates irregularities in medial axis detection. Based on the medial axis, the authors subsequently characterize the geometry of pore space using (i) the "burn number" distribution, i.e., the distance distribution of void voxels to the material; (ii) the volume distribution

of connected components of the medial axis; and (iii) tortuosity computed as the length distribution of paths from one side of the sample to the other. Their approach differs from ours in that they use the medial axis for characterizing the pore space instead of the Reeb graph. Furthermore, when computing properties such as connected components and tortuosity, they treat each voxel of the medial axis as node in a graph, leading to much larger graphs.

Building on this work, Lindquist [15] extracts the "backbone" of the medial axis. His method subsequently uses the skeleton to detect pores and to derive measures, such as a coordination number (degree) of the vertices in the backbone and channel length. This work also introduces three algorithms to grow the backbone into three-dimensional pores and provides measures, such as area and volume.

Gyulassi et al. [9] extract filament structures in porous solids. The goal is to characterize changes in a solid foam during the impact of a projectile. For this purpose, they use the Morse–Smale complex to compute a simplified representation of the distance field that removes noise. They extract filament structure in the porous solid as a subset of this complex. By defining a distance between graphs, they characterize how much the foam changes due to the impact of a projectile. While their method also considers porous materials and structures related to the medial axis, they are concerned with substantially different properties of solids. Our approach also differs from theirs in that we perform a quantitative analysis by associating a cross-sectional filament area with the edges of the Reeb graph computed from the three-dimensional data set.

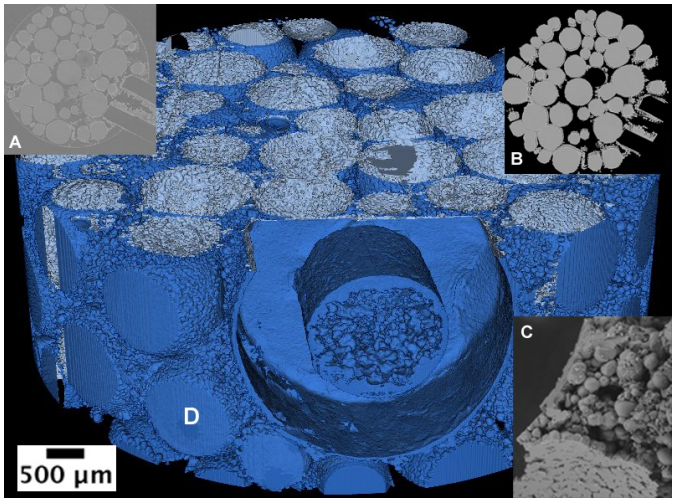


Fig. 1. MicroCT of porous media: (A) cross-section of glass bead column, inoculated with *S. pausterii* that promote calcite precipitation; cross-section is input to our software Quant-CT, which outputs segmented slices as in (B); rendering of the segmentation result for the whole stack in (D) using VisIt; SEM image in (C) emphasizes the result of biomineralization, which clogs the void space, cementing the pore channels.

3 Materials: Pore Networks in Geological Formations

Potential subsurface regions that are candidates for carbon sequestration are retired wells. Reservoir rocks must be porous and permeable to allow extraction of hydrocarbons, properties that are also necessary for carbon injection and storage in the subsurface. Injection processes must guarantee that CO_2 remains in the subsurface, but the technology still needs to be improved. At the EFRC, experiments are performed using a variety of materials that influence carbon sequestration, such as soil and rock samples from natural environments, and prepared composites combined with microorganisms. These samples can be placed in column vessels where temperature and pressure conditions can be controlled to monitor the mechanical properties and the distribution of CO_2, and its viscosity, density, and surface tension with brine.

Monitoring the samples includes the measurement of the media porosity, and permeability, among other characterizations of pore morphology. For example, such descriptors are important to address biomineralization, a natural subsurface process that can profoundly alter the physical properties of porous materials. Biomineralization refers to the synthesis of inorganic mineral-like materials by living organisms through the combination of chemical elements and such organisms in the same system. Bacteria can form inorganic crystals either intracellularly or extracellularly [24], generating by-products, such as the precipitation of calcium carbonate, as shown in Fig. 1. Calcite is a mineral phase that can precipitate during subsurface remediation or geotechnical engineering processes, with possible behavior variations of the system, e.g., flow alteration and soil strengthening. Research with microorganisms [3] showed that microbial precipitation of $CaCO_3$ strikingly reduced the permeability of porous media, consequently affecting fluid flow and transport properties. This paper describes tools to address pore description, illustrating our results with microCT from experiments using porous material under biogenic $CaCO_3$ precipitation induced by *Sporosarcina pasteurii* [3], and synthetic images that simulate porous media.

3.1 Porous media from microCT

MicroCT from porous media is obtained using high-resolution synchrotron-based X-ray spectromicroscopy, which is a nondestructive technique for seeing inside solid objects. The process of imaging the entire 3D rock sample involves acquiring projection views at several equally spaced angular positions, which after transformations [12], turn into cross-sections of the object. Each voxel conveys information about the X-ray attenuation and density of the scanned material as a gray level value. MicroCT image stacks are often isotropic with a pixel resolution at micrometer scale. They can be used to create 3D virtual models that are useful in carbon sequestration research, particularly in probing the structure of porous materials. Here, we describe microCT samples acquired at the ALS Beamline 8.3.2 [1] with energies between 10 and 45keV, with a 1% bandpass, CCD camera Cooke PCO 4000, Kodak chip with 4008×2672 pixels, 14 bit, 9 micron square pixels.

Fig. 1 illustrates results of calcite precipitation induced in a glass bead-packed column in biogenic mixture, using the microbe *S. pasteurii*. The glass beads data set has dimensions $3337 \times 3337 \times 483$ pixels, 4.49 μm resolution and a total of 10GB, originally, and the core region corresponding to the inner part of the experimental vessel is selected, with x-y dimensions of 1393×1398.

3.2 Porous media from simulation

Our synthetic test data consists of jammed packings of hard spheres, generated using a molecular-dynamic code proposed by Skoge et al. [20]. These codes enable a generation of bead beds inside a cylindrical container of a specified radius, here $R = 200$ pixels, with different packing ratios (media volume fraction), similar to the column vessel used to enclose the samples before imaging, as described in Sect. 3.1. We consider monodisperse spherical random packings of n spheres, for n in $[50, 2000]$, and sphere diameters d in [20,94] pixels. These simulations are often used to understand the equilibrium and dynamical properties of materials, such as simple fluids, colloids, glasses, and granular media. Fig. 7a illustrates the result of simulating a bead bed with packing equal to 10% comprised of 1,500 spheres with a diameter equal to 20 voxels. We considered synthetic packs that are both above and below the suspension limit, imitating grains in fluid or rocks, respectively.

4 Image Analysis

Our computational analysis chain starts with synchrotron based X-ray microCT cross-sectional images as input, which are the result of reconstructed data using the filtered back projection algorithm [12]. The slices represent the linear attenuation coefficient map of the scanned object, compressed into different shades of gray, e.g., eight-bit intensity images, often degrading the intensity variation. Other degrada-

tions may arise from the reconstruction, both intrinsic and extrinsic artifacts.

Ushizima et al. [21, 22] describe a computational analysis workflow for carbon sequestration research that includes image filtering, segmentation, and feature detection and analysis. This pipeline includes algorithms that analyze microCT data sets from soil samples, rocks, glass beads and simulated data, representing jammed packing of glass beads. They introduced and designed a new algorithm and software framework, Quant-CT [21, 22], which processes and quantifies structures from cross-sections, and uses VisIt [6] to visualize results as illustrated in Fig. 1 and Fig. 8a. The current algorithms can differentiate porous media (high density material) from void (background, low density media) using a Boolean classifier, and extract features such as volume, surface area, granularity spectrum, porosity, and more recently, permeability. Quant-CT supports quick user interaction, including the ability for the user to train the algorithm via subsamples, and provide its core algorithms with an automated parametrization [21]. As illustrated in Fig. 1a, this method takes microCT slices as input to an efficient greedy algorithm for statistical region merging (SRM) that runs in linear time. This algorithm reduces image artifacts by filtering sets of slices with a bilateral filter (BF). One of the challenges of using BF is finding good parametrizations of the photometric and the geometric kernels for anisotropic smoothing. Ushizima et al. showed that one could calculate the photometric parameter as proportional to the signal-to-noise ratio [21], aiming to minimize the issue of reflection over the large spectrum of X-ray attenuation values of the scanned materials. This enables automated tuning of scale parameters, based on statistics extracted from patches fed into the filtering algorithm by the user. After image enhancement, they perform image segmentation using SRM to identify two-phase volumes, dense material (e.g. rock, beads) and void (pore volume). This improved algorithm leverages the SRM ability to build a segmentation tree dependent on a scalar parameter, a.k.a. predicate, but originally tackles the oversegmentation issue of SRM by using user-selected patch statistics to fuse regions into void and solid phases. These tests illustrate that it can be applied to high-resolution X-ray microtomography stacks up to tens of gigabytes in size. This quantification has been used to support computational modeling of the fluid dynamics into the pore space of the host object.

5 Topology-based Extraction of Pockets and Pore Networks

To analyze the physical properties of the material we summarize the geometry of its pore structure. The binary classification of the CT stack of images gives us access to the necessary information in the experimental data. Our goal is to both visualize the internal structure and compare multiple materials. Since we can also extract the same descriptors for the simulation data, we can compare physical and simulated materials. We study a pair of multiscalar geometric descriptors: the first is an (augmented) graph reflecting the connectivity of the pore network; the second is a diagram of "pockets" within the material. Together they not only serve as a compact visual representation of the material, but they also enable the computation of various derived quantities as well as the identification of pore bodies.

Reeb graphs. We assume a given orientation of the porous medium and use its height function h, similar to the computation proposed in [11]. We consider a function $h : \mathbb{X} \to \mathbb{R}$, defined on the three-dimensional void space $\mathbb{X}$ of the material, that assigns to every point its depth in the stack. Our analysis relies on Reeb graphs, which describe the topology of level sets of real-valued functions. Given the function h, we say that two points $x, y \in \mathbb{X}$ are equivalent, $x \sim y$, if they belong to the same connected component of some level set. The quotient space $\mathbb{X}/\sim$, called a *Reeb graph* of h, identifies all the equivalent points together. Intuitively, it collapses connected components of the level sets of the height function, and preserves their connectivity, as illustrated in Fig. 2b.

As the illustration suggests in two dimensions, the Reeb graph matches our intuition of the pore network in the material. But we can capture even more information. To differentiate between channels of varying thickness we turn to an auxiliary construction. Let M denote the space occupied by the material; let $\mathrm{d}_M : \mathbb{X} \to \mathbb{R}$ be the distance function to this space, $\mathrm{d}_M(x) = \inf_{y \in M} \|x - y\|$. We restrict our attention to the α-superlevel set of d_M, $M^\alpha = \mathrm{d}_M^{-1}[\alpha, \infty)$, which is generically a three-manifold with boundary. This subset of the void space has a simple geometric meaning: a point x belongs to the superlevel set M^α if a ball of radius α, centered at x, does not intersect the material, $\mathrm{B}(x, \alpha) \cap M = \emptyset$. By increasing the threshold α, we restrict our view to wider channels.

Already for small values of the threshold α, the empty space in the slices of the CT scan becomes disconnected, introducing an interesting topology that describes the pathways a particle of size α can take through the material. We examine the Reeb graph by restricting the height function to the superlevel set, $h_{|M^\alpha}$. Its branching structure reflects the connectivity of the "wide pathways" in the material, see Fig. 8c. In fact, we get a 1-parameter family of graphs that we can use to describe the permeability, porosity, and other quantities of interest.

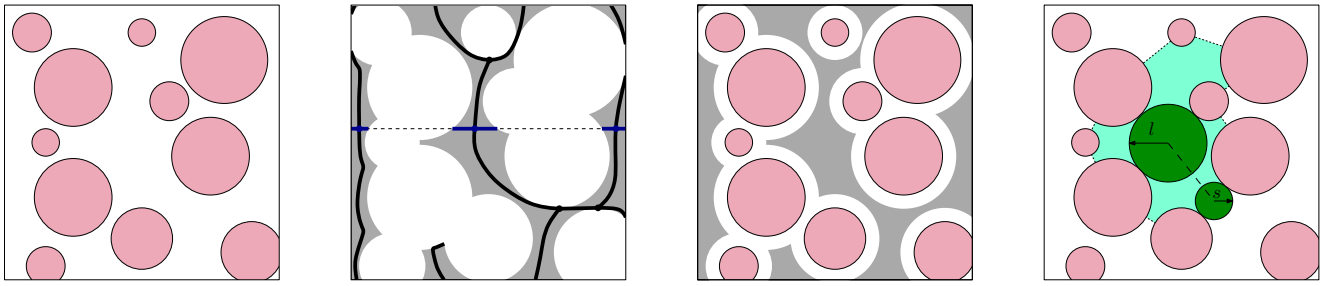
Reeb graph computation. The efficient computation of Reeb graphs is an active research topic. Until recently, the fastest algorithm was a randomized approach due to Harvey et al. [10], running in expected time $\mathrm{O}(m \log m)$. Parsa [18] has recently reported on an optimal deterministic algorithm, which always runs in time $\mathrm{O}(m \log m)$, where m is the number of edges in the domain. In our case, the data has a special property: the depth function h, whose Reeb graph we want to compute, is aligned with one of the coordinate axes—we take advantage of this structure in our computation.

To compute the Reeb graph, it suffices to identify the connected components of the superlevel set M^α of the distance function, restricted to each slice in the stack. A depth-first search in each slice performs this labeling in linear time. Subsequently, we identify which components at level z in the stack are adjacent to which components at level $z+1$, as illustrated in Fig. 3a. A traversal of the points at a given level identifies all such connections—the edges of the Reeb graph—in linear time. This axis-aligned structure presents one more advantage: the computation of the Reeb graph is completely parallel—all slices are treated independently.

Maximum flow. To illustrate the utility of a graph representation of the porous material, we employ the Reeb graph as the input to a max-flow computation. Each vertex in the Reeb graph corresponds to a connected component in the level set of the height function. Every edge connects two such components together. We map the two components into the same plane (from the two planes vertically consecutive in the stack) and find their intersection. We assign the area of this intersection as the capacity to the edge; see Fig. 3a.

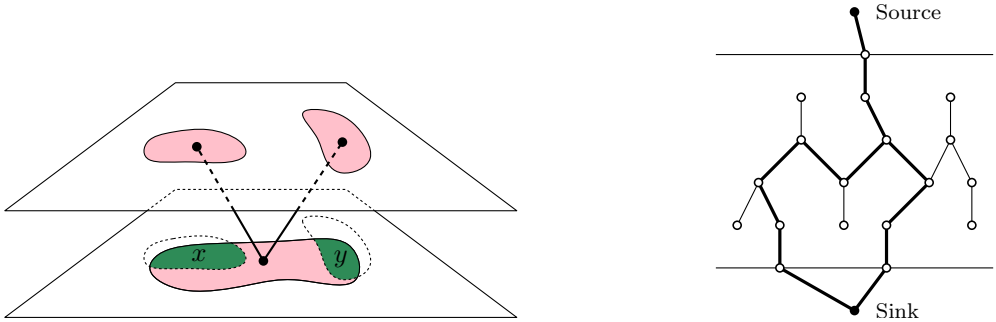
Once we extract the capacities of all the edges, we add two auxiliary vertices: a source vertex connected to all the components of the top slice and a sink vertex connected to all the components of the bottom slice. We use the Ford–Fulkerson algorithm [8] to find the max-flow from the source to the sink. By definition, this quantity expresses how much flow can travel across the graph without exceeding the edge capacities and conserving the flow on the vertices, i.e., the outflow of any vertex is the same as its inflow. We view this measure as a nuanced description of the geometry of the material: unlike porosity, which only measures the percentage of the empty space, max-flow accounts for what space is reachable in the material. Particularly interesting is the assignment of the flow to individual edges in the solution to the max-flow problem; see Fig. 3b. First, it eliminates the dead branches that do not connect the top to the bottom of the material; this pruning is especially helpful when visualizing the graph. Second, it highlights the bottlenecks in the network, the edges that are saturated to their full capacity.

Our ability to examine Reeb graphs at different superlevel sets of the distance function provides us with a multiscalar view of the flow through the material, reflecting its complicated geometry. As we increase the threshold α of the distance function, the material thickens and the pores shrink. Therefore, the maximum flow decreases as illustrated in Fig. 4. One interpretation of the resulting graphs is particularly convenient for our application. The graphs tell us how



(a) Void space (b) Reeb graph (c) Wide pathways (d) Pocket

Fig. 2. A two-dimensional illustration of the constructions of a Reeb graph, wide pathways, and a pocket. We stress that the figure is two-dimensional only for clarity of the illustration; all the constructions involved in our computation happen in three dimensions. (a) A schematic porous material in pink, with the void region in white. (b) Reeb graph of the height function h, restricted to the wide pathways M^α. The dark blue components in a level set of the function collapse to single points in the Reeb graph. (c) Wide pathways, shaded in gray, are a superlevel set M^α of the distance function, d_M, to the material. (d) A pocket of persistence $l-s$ in green. The largest sphere that fits inside has radius l, while the largest sphere that can escape has radius s.



(a) Edge capacities (b) Flow graph

Fig. 3. Schemas of the edge capacities and Reeb graph calculations. (a) Three components (shaded in pink) in two consecutive slices of the stack are represented by three vertices and two edges in the Reeb graph. The capacity assigned to the edges is calculated as the area of the intersection of the respective components, shaded in green in the lower slice. Accordingly, the left edge is assigned capacity x; the right edge is assigned capacity y. (b) Full Reeb graph with dead-ends and a detected subgraph that carries the maximum flow, highlighted in the thicker line.

much flow the material carries after homogeneous precipitation, which evenly clogs the pores.

Persistent pockets. To refine our representation of the porous material, we augment the Reeb graph with significant pockets, where the liquid can accumulate. We use the theory of persistent homology [7] to detect such pockets and quantify their significance.

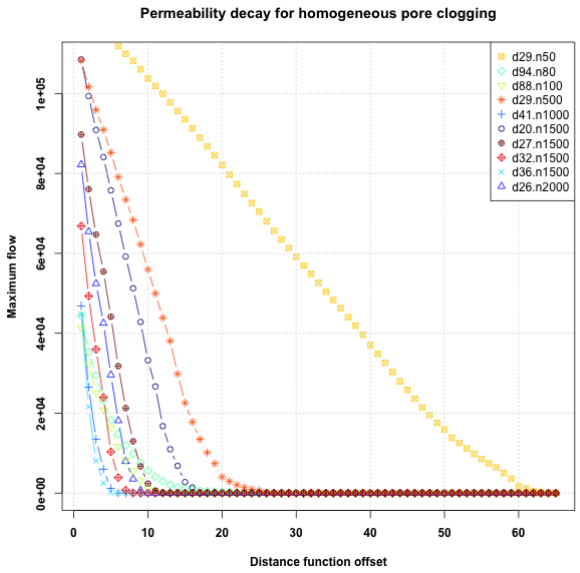
The distance function d_M, once again, supplies the necessary information. We track the connected components in its superlevel sets as we decrease the defining threshold α. As we decrease α from infinity to zero, whenever it equals the value of a local maximum of the function d_M, a new component appears in the superlevel set M^α. We assign this value to the component, and say that it is born at α. Occasionally, as our threshold passes values of certain saddles, such components merge. When two components born at α_1 and α_2 merge at α_s, with $\alpha_s \leq \alpha_1 \leq \alpha_2$, we say that the younger component merged into the older one, and pair α_1 and α_s. Once we sweep α from infinity to zero, all maxima, except for the global one, become paired to saddles. Thus, we obtain a collection of pairs $(s_i, l_i)_i$, each representing a pocket in the material. The difference $l_i - s_i$ is the *persistence* of the pocket, centered at the maximum at distance l_i from M.

A simple geometric intuition justifies our language. The material has a pocket with parameters (s,l) if the largest empty sphere that fits inside has radius l, while the largest sphere that can escape has radius s; see Fig. 2d. The larger the difference $l-s$ between these radii, the more pronounced the pocket. By plotting all the pairs (s,l) in the plane, with s as the abscissa and l as the ordinate of the point, we get a *persistence diagram*, see Fig. 6; points further away from the diagonal represent more persistent pockets. The diagram serves as a summary of material properties: it highlights significant scales, shows all the pockets at once, and enables a comparison of multiple materials.

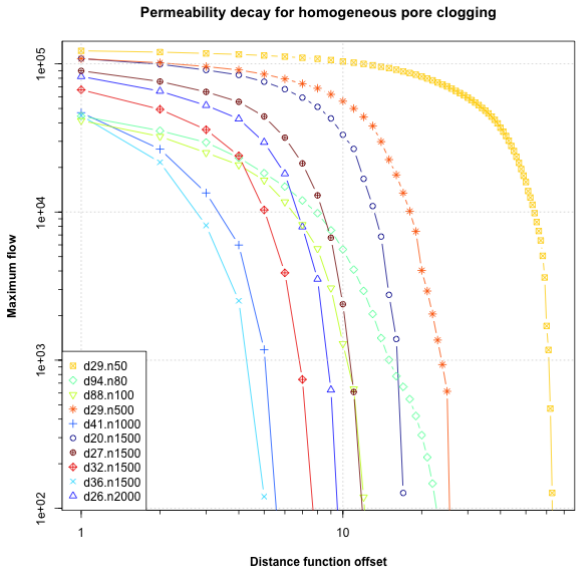
6 Results

We computed porous material descriptors, using Reeb graphs and persistence diagrams to extract pore network, max-flow curves and pockets. We tested these algorithms using different materials. In total, we considered ten synthetic bead beds and one experimental microCT data set of borossilicate beads of approximately 1mm, coated with bioprecipitates. The goals of this analysis were: (i) to demonstrate that max-flow curves, extracted from Reeb graphs support a broader, multiscale description of the flow throughout the material as opposed to single scalars as porosity and permeability (Figs. 4 and 5); (ii) to enable a comparison of different materials regarding interstitial space (Fig. 4); (iii) to model homogeneous precipitation on clean bead packs to better understand biomineralization in real experiments (Table 1); (iv) to compare our findings with standard representations as porosity-permeability curves (Fig. 5); and (v) to visualize the pore network and pockets (pore bodies) from simulated materials (Fig. 7) and microtomography of beads injected with microorganisms that induce bioprecipitation (Fig. 8).

Fig. 4 shows maximum flow curves for the synthetic samples with different granularity and porosity. Both linear and logarithmic representations are shown to emphasize differences between curves and



(a) Linear



(b) Logarithmic

Fig. 4. Maximum flow for synthetic materials, for different sphere diameters and packing, as we increase threshold α of the distance function.

facilitate an interpretation of the slopes. The last column of Table 1 shows the slope calculated for each curve in Fig. 4. On the horizontal axis of Fig. 4, the lower the distance function offset value, the wider the void space of the sample. Higher values of the distance function offset are analogous to higher levels of precipitation in a particular sample. Each curve is drawn using a different color and symbol, and labeled according to the diameter (d) of the bead, followed by number (n) of beads. The curve $d29.n50$ corresponds to beads with $d = 29$, $n = 50$, with a higher permeability due to the low packing (1%), and therefore, a high porosity (99%) and a larger amount of void space in comparison to the other samples. In comparing curves of different materials, the fluid flow of a sample decayed slower for materials with high porosity, which was expected. However, the comparison between materials with similar porosity and adverse permeabilities can be less obvious—e.g., samples $d29.n50$ and $d29.n500$. Particularly in these

Table 1. Porosity, permeability, and slope of tangent to max-flow curves, calculated for synthetic data sets used in Fig.4.

Data set	Porosity	Permeability (pel^2)	Slope
d36.n1500	0.42	1.68	-14,019.40
d41.n1000	0.44	2.44	-13,569.20
d32.n1500	0.6	7.53	-12,677.86
d26.n2000	0.71	16.14	-12,151.36
d27.n1500	0.76	29.46	-10,532.90
d20.n1500	0.9	166.83	-8,039.12
d29.n500	0.9	334.14	-5,848.84
d88.n100	0.45	13.61	-4,608.08
d94.n80	0.46	17.35	-3,539.35
d29.n50	0.99	43,929.40	-2,166.34

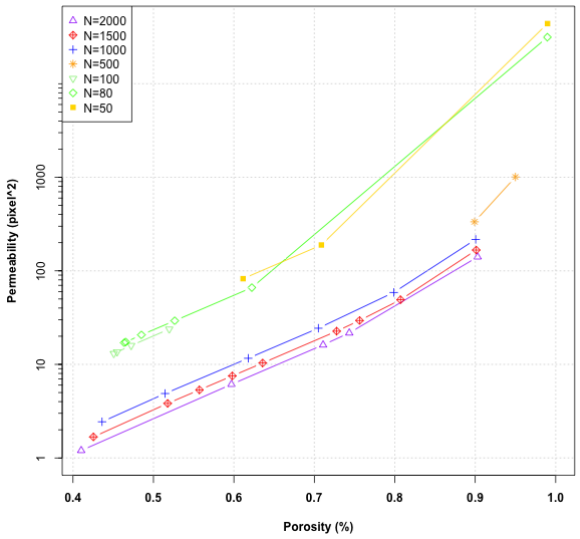


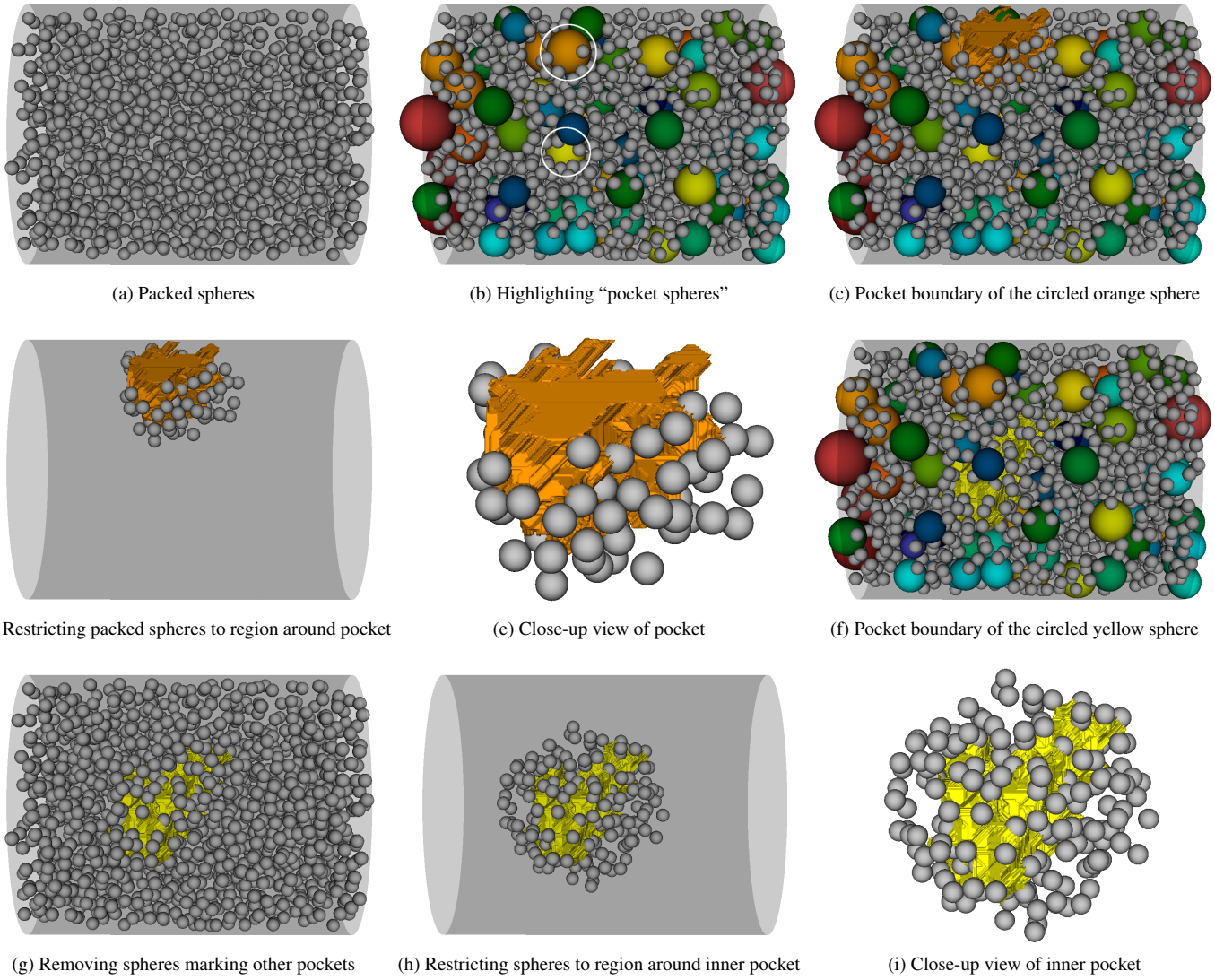
Fig. 5. Porosity-permeability curves for material comparison: each curve corresponds to samples with the same number of beads, with a decreasing sphere diameter, i.e., increasing permeability/porosity.

cases, the max-flow curves can quickly illustrate how the channel network would be affected by narrowing processes as precipitation.

As a rule of thumb, we can calculate linear regressions on these curves, focusing on the values before the max-flow function assymptotically approaches zero. In general, we observed that the slope of the straight line, obtained via linear regression, will be higher for materials with higher permeability. Also, the granularity is reflected on the slope, as illustrated in the differences between curves $d36.n1500$ and $d41.n1000$. When considering samples with similar coarseness, as $d26.n2000$ and $d27.n1500$, the sample with a higher porosity will present a higher slope.

As expected, the number of particles and the sphere diameter influence both porosity and permeability as illustrated in Fig. 5 and Table 1. Fig. 5 includes more data sets (with a total of 35 simulated bead beds) than in Fig. 4 for better visualization of the porosity-permeability relationship, when increasing the spheres diameter, but keeping the same number of spheres. Similar to the experiments performed by Zaman [25], we found that porosity and permeability are directly proportional, according to the considered simulated data. The porosity-permeability curve is a standard representation, but it can be confusing in comparisons of materials with different properties, as illustrated in Table 1, where materials with similar porosity can exhibit a very different permeability.

Fig. 7 shows the visualization of a synthetic material, simulated using 1500 spheres, where each sphere has a radius of 20 pixels. Fig. 7a



(a) Packed spheres (b) Highlighting "pocket spheres" (c) Pocket boundary of the circled orange sphere

(d) Restricting packed spheres to region around pocket (e) Close-up view of pocket (f) Pocket boundary of the circled yellow sphere

(g) Removing spheres marking other pockets (h) Restricting spheres to region around inner pocket (i) Close-up view of inner pocket

Fig. 7. Visualization of pockets in synthetic packed sphere data set comprising a cylinder packed with $n = 1500$ spheres of constant radius $r = 20$ pixels. (a) Showing only the packed spheres: notice how difficult it is to estimate the size of the empty space between particles. (b) Representing pockets as spheres with the largest radius of a sphere fitting into the pocket highlights the empty space between spheres. Sphere color corresponds to size using a modified rainbow color map. (c), (d), (e) Boundary of the pocket corresponding to the circled orange pocket sphere. (f), (g), (h), (i) Boundary of the circled yellow pocket sphere in the interior.

shows only the packed spheres in the cylinder. It is difficult to perceive the size of the voids between them or whether there are any voids at all. To get an overview of their structure, Fig. 6 shows the persistence diagram that summarizes the distribution of the pockets in this data set. Points represent individual pockets. The ordinate l represents the largest sphere that fits inside a pocket. Consequently, the vertical position of the points shows the size of the void space. The further the points are shifted upwards, the larger the voids or pockets in the solid. The abscissa s represents the size of the largest sphere that can escape from the pocket. It characterizes the size of the largest throat leading out of a pocket. The distance of a point to the diagonal, $l - s$, expresses the prominence of a pocket. Both the persistence diagram and Fig. 7 show pockets of persistence at least five pixels to prune topological noise.

To highlight the geometry of the pockets, Fig. 7b shows "pocket spheres," i.e., the largest sphere that fits into an individually detected void between the packed spheres. Visualizing these spheres, colored according to their radius, already gives an impression of the size of the free space between spheres. The main goal in coloring spheres is being able to distinguish between them and make it possible to determine whether two spheres have similar sizes. For this purpose, we use a modified rainbow color map to assign colors. To obtain more information about pockets, we show the detected pocket boundaries for the pockets corresponding to the two spheres circled by white in Fig. 7b. The motivation to show both the orange and yellow pockets is to inspect the differences between pockets adjacent to the cylindrical container vessel (in orange), where the simulation of homogenous precipitation is different from internal pockets (in yellow), which also tend to be larger than the more external pockets.

Fig. 7c shows the boundaries of the pocket corresponding to the circled orange sphere close to the cylinder containing the packed spheres. To better discern its boundaries, we hide the pocket spheres and restrict display to all packed spheres within a certain distance from its centroid in Fig. 7d and provide a close-up view in Fig. 7e. We can clearly see the outlines of the detected void. Fig. 7f shows the boundary of the pocket corresponding to the circled yellow sphere in the interior. While it is significantly larger than the orange pocket, Fig. 7g illustrates that it is obscured by the packed spheres, even if we hide

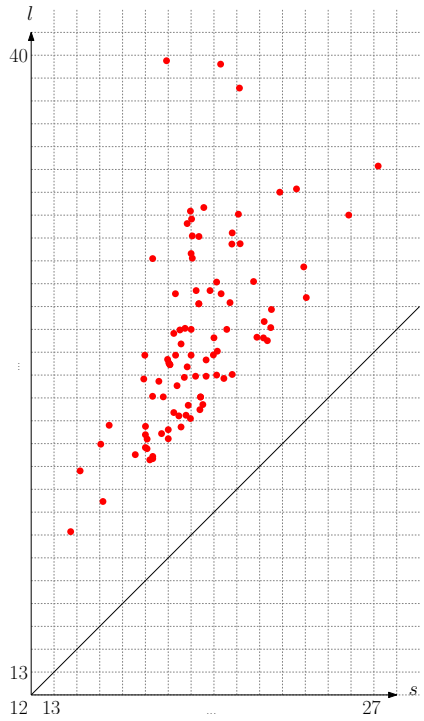


Fig. 6. Persistence diagram showing the distribution of the pockets visualized in Fig. 7. Each pocket is represented by a point, where the ordinate l represents the radius of the largest sphere that fits inside, and the abscissa s, the radius of the largest sphere that can escape. Both are measured in pixels. The line corresponds to zero persistence. The larger the distance of a point to the line, the more pronounced its corresponding pocket.

the pocket spheres. Fig. 7h shows a view where all packed spheres further away than a specified radius from the centroid are removed, and Fig. 7i shows a close-up view of this pocket. It is clearly visible that there are large regions of empty space between the spheres and that pockets correspond to connected regions between the spheres.

Fig. 8 shows the results of our analysis for a data set of calcite precipitation induced by a biogenic mixture. Fig. 8a shows the result of segmenting the imaging data into material (gray) and void space (empty). To analyze the flow through the sample, we show the flow graph along with the sample in Fig. 8b, to illustrate that the maximum flow coincides with voids in the segmentation. To analyze the flow behavior further, we only show a cylindrical cut through the segmentation along with the flow graph (Figs. 8c through 8f). Figs. 8c and 8d show a cylindrical cut through the material that corresponds to a region of large flow through the sample. The first figure also shows pocket spheres around the cylindrical cut to illustrate where pockets in the sample are detected. The flow through the sample coincides with the voids in the material. Figs. 8e and 8f examine why other regions in the sample show no contribution to the flow graph. The cylindrical cut is blocked by small deposits that fill most of the voids. These deposits correspond to calcite precipitation produced by the microbes (*S. pasteurii*) in the experiment.

7 Conclusions and Future Work

In this paper, we have presented topological descriptors, customized to the problem of characterizing both synthetic materials from 3D simulations of packed spheres and 3D high-resolution microtomography. We have proposed a new descriptor of homogeneous deposition on porous materials, and its obstruction effect on interstitial space, given different slopes obtained from linear regression on the maximum flow curves. The main advantages of our approach is the combination of image-processing, topological analysis, multiscale approaches and visualization aimed at gaining a deeper understanding of pore structures, illustrating results using both experimental and synthetic data sets. We introduced the use of max-flow slopes as a descriptor to a enable comparison between materials with different porosity and permeability values, that will exhibit adverse behaviour during precipitation, which could not be predicted by using only porosity-permeability curves.

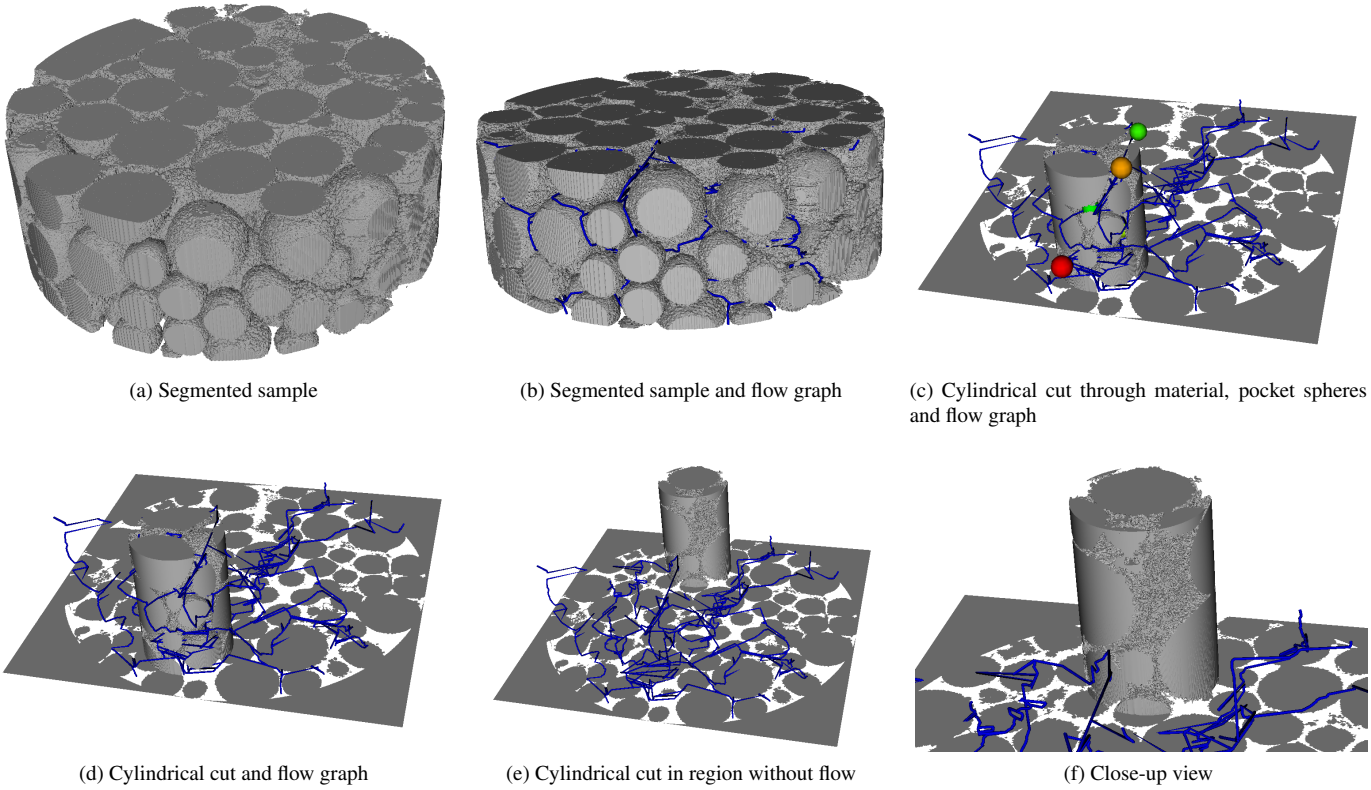
The visualization of pore bodies helps domain scientists developing an intuition regarding fluid storage. It also facilitates the identification of similar pockets, which correspond to spheres with alike colors. On the other hand, the max-flow curves and their slopes support a comparison of materials with a different fluid flow potential. In addition, the max-flow curves can indicate permeability decay when the pore network is subjected to depositing, and consequently, pore clogging. We simulated a homogeneous deposition by considering different offsets of the distance function. These results are relevant to carbon sequestration research because bioprecipitation can accumulate and block the network. Therefore, bioprecipitation could be helpful in trapping CO_2 in the subsurface; this paper lays out initial work in quantifying deposition in porous materials. The current tools for topological descriptor extraction and the visualization of pore networks and pore bodies are prototypes to be deployed in material science research. Our approach to data representation is a key step in extracting estimates to be used as statistical descriptors, e.g., the distributions of pore bodies conditioned to pore throats.

Moving forward, we will compute these topological descriptors for more samples acquired by microCT and study in detail how they illustrate differences between materials. We will also apply these techniques to numerical simulation data, with the aim of comparing numerical simulation results to physical experiments. Finally, we will re-examine related work and perform more detailed comparisons of derived descriptors. In particular, we are interested in deriving properties, such as the pore throats (as in [19]) from our representation, with the advantage of not requiring preprocessing as the removal of redundant "ribs." In our current representation, we only represent the largest throat explicitly, but having information about all pore throats would enhance the expressiveness of our method. Other investigations might include fluid flow numerical methods to solve the Navier Stokes [13] equations considering concise representations of the pore network.

Future efforts will focus on computing up-scaled continuum properties (e.g., single phase permeability, capillary pressure curves, relative permeability for CO_2/brine systems, effective diffusivity, dispersivity) from our microstructural representation, so we can address the impact of geometries to geological carbon sequestration. Such tools will enable a prediction of large scale behavior of plumes from small samples of rocks taken from deep underground—microstructure can provide insight (and hopefully quantitative estimates) of the key parameters (and parameter relationships) needed for those models. Before CO_2 is injected at a site, extensive multiphase flow modeling is performed to model how the plume will evolve with time; usually these modeling exercises are far from truly predictive and one reason is the paucity of data available to describe the transport properties. Since core flood experiments (the way to actually measure some of these values) are time-consuming, expensive, they require large samples, and they are hard to repeat, pore-scale imaging of small samples and subsequent modeling could be one way of obtaining a wide variety of property estimates for input into larger scale simulations.

Acknowledgments

We thank Donald J. DePaolo, Center Director of Nanoscale Control of Geologic CO_2, an Department of Energy (DOE) Energy Frontier Research Center (EFRC). We also thank J. Ajo-Franklin for the microtomography data and valuable feedback regarding porous media characterization. Further thanks to A. MacDowell, D. Parkinson at the Advanced Light Source, Beamline 8.3.2, which is the U.S. DOE Office of Science, Office of Basic Energy Research Sciences, of the U.S. DOE (contract DE-AC02-05CH11231). This work was partially supported by the Director, Office of Science, Office and Advanced Scientific Computing Research, of the U.S. DOE under Contract No. DE-AC02-05CH11231 through the Scientific Discovery through Advanced Computing (SciDAC), the SciDAC Institute of Scalable Data Management,



(a) Segmented sample

(b) Segmented sample and flow graph

(c) Cylindrical cut through material, pocket spheres and flow graph

(d) Cylindrical cut and flow graph

(e) Cylindrical cut in region without flow

(f) Close-up view

Fig. 8. Visualization of pockets and flow graph in the calcite data set. (a) Segmented sample of calcite. (b) Flow graph following the voids in the sample. (c), (d) Cylindrical cut through region of high flow. (e), (f) Cylindrical cut through region of low flow. Pores are "blocked" by calcite precipitation produced by microbes (*S. pasteurii*) in the experiments.

Analysis and Visualization (SDAV), the grant "Topology-based Visualization and Analysis of High-dimensional Data and Time-varying Data at the Extreme Scale," and the National Energy Research Scientific Computing Center (NERSC). It was also supported in part by the Applied Mathematical Science subprogram of the Office of Energy Research, U.S. Department of Energy, under Contract No. DE-AC03-76SF00098. Andrea G. C. Bianchi was also supported by Capes (BEX 0761/11-7), Capes/Nanobiomed, INEO/CNPq, CNPq (472565/2011-7) and Fapemig (APQ-00802-11).

Legal Disclaimer

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

References

[1] Lawrence Berkeley National Laboratory (LBNL) Advanced Light Source (ALS). `http://microct.lbl.gov/`. Access date: June 6, 2012.

[2] J. Ajo-Franklin. Using synchrotron micro tomography for pore-scale monitoring of $CaCO_3$ precipitation and CO_2 flow. In *Symposium of Nanoscale Control of Geologic CO_2 - Energy Frontier Research Center*, Berkeley, CA, USA, 2010.

[3] R. Armstrong and J. Ajo-Franklin. Investigating biomineralization using synchrotron based X-ray computed microtomography. *Geophysical Research Letters*, 38(4), 2011.

[4] T. Boden, G. Marland, and R. Andres. Global, regional, and national fossil-fuel CO_2 emissions. *Carbon Dioxide Information Analysis Center - ORNL*.

[5] W. D. Carrier and F.ASCE. Goodbye, Hazen; Hello, Kozeny-Carman. *Journal of Geotechnical and Geoenvironmental Engineering*, 129(11):1054–1056, Nov 2003.

[6] H. Childs, E. S. Brugger, K. S. Bonnell, J. S. Meredith, M. Miller, B. J. Whitlock, and N. Max. A contract-based system for large data visualization. In *Proceedings of IEEE Visualization 2005*, pages 190–198, 2005.

[7] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete and Computational Geometry*, 28(4):511–533, 2002.

[8] L. R. Ford and D. R. Fulkerson. Maximal flow through a network. *Canadian Journal of Mathematics*, 8:399–404, 1956.

[9] A. Gyulassy, V. Natarajan, M. Duchaineau, V. Pascucci, E. M. Bringa, A. Higginbotham, and B. Hamann. Topologically clean distance fields. *IEEE Trans.on Vis. and Comput. Graph. (Proceedings IEEE Visualization 2007)*, 13(6):1432–1439, November/December 2007.

[10] W. Harvey, Y. Wang, and R. Wenger. A randomized $O(m \log m)$ time algorithm for computing Reeb graphs of arbitrary simplicial complexes. In *Proceedings of the Annual Symposium on Computational Geometry*, pages 267–276, 2010.

[11] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. L. Kunii. Topology matching for fully automatic similarity estimation of 3d shapes. In *Proc. of the 28th Annual Conference on Computer graphics and Interactive Techniques*, SIGGRAPH '01, pages 203–212, New York, NY, USA, 2001. ACM.

[12] J. Hsieh. *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*. SPIE—The International Society for Optical Engineering, Bellingham, WA, USA, 2003. ISBN-13: 978-0819444257.

[13] L. Jing. A review of techniques, advances and outstanding issues in numerical modelling for rock mechanics and rock engineering. *International Journal of Rock Mechanics and Mining Sciences*, 40(3):283 – 353, 2003.

[14] C. Jones and K.-L. Ma. Visualizing flow trajectories using locality-based rendering and warped curve plots. *IEEE Trans. Vis. Comput. Graph.*, pages 1587–1594, 2010.

[15] W. Lindquist. Quantitative analysis of three dimensional x-ray tomographic images. In *Developments in X-ray Tomography III, Proceedings of SPIE*, volume 4503, pages 103–115, Bellingham, WA, USA, 2002.

[16] W. B. Lindquist, S.-M. Lee, D. A. Coker, K. W. Jones, and P. Spanne. Medial axis analysis of three dimensional tomographic images of drill core samples. *Journal of Geophysical Research. Solid Earth*, 101(B4):8297–8310, 1996.

[17] P. Nico, J. B. Ajo-Franklin, A. McDowell, D. B. Silin, L. Tomutsa, S. M. Benson, and Y. Wu. Synchrotron x-ray micro- tomography and geological CO_2 sequestration. *Advances in Computed Tomography for Geomaterials - GeoX 2010*, pages 374–380, 2010.

[18] S. Parsa. A deterministic $O(m \log m)$ time algorithm for the reeb graph. In *Proceedings of the Annual Symposium on Computational Geometry*, 2012.

[19] D. Silin and T. Patzek. Pore space morphology analysis using maximal inscribed spheres. *Physica A: Statistical and Theoretical Physics*, 371(2):336–360, Nov. 2006.

[20] M. Skoge, A. Donev, F. H. Stillinger, and S. Torquato. Packing hyperspheres in high-dimensional euclidean spaces. *Phys. Rev. E*, 74:041127:1–11, 2006.

[21] D. Ushizima, J. Ajo-Franklin, A. Macdowell, P. Nico, D. Parkinson, E. Bethel, and J. Sethian. Statistical segmentation and porosity quantification of 3d x-ray microtomography. *SPIE Optics and Photonics: XXXIV Applications of Digital Image Processing*, 8135-1:1–14, 2011.

[22] D. M. Ushizima, G. H. Weber, J. Ajo-Franklin, Y. Kim, A. Macdowell, D. Morozov, P. Nico, D. Parkinson, D. Trebotich, J. Wan, and E. W. Bethel. Analysis and visualization for multiscale control of geologic CO_2. *Journal of Physics: Conference Series, Proceedings of SciDAC 2011*, July 2011.

[23] F. J. Valdes-Parada, J. A. Ochoa-Tapia, and J. Alvarez-Ramirez. Validity of the permeability Carman-Kozeny equation: A volume averaging approach. *Physica A*, 388:789–798, 2009.

[24] N. Yoshida, E. Higashimura, and Y. Saeki. Catalytic biomineralization of fluorescent calcite by the thermophilic bacterium geobacillus thermoglucosidasius. *Appl. Environ. Microbiol.*, 76(21):7322–7327, 2010.

[25] E. Zaman and P. Jalali. On hydraulic permeability of random packs of monodisperse spheres: direct flow simulations versus correlations. *Physica A*, 389:205–214, 2010.